

Počítačová podpora v archeologii

Jiří Macháček

(editor)

Ústav archeologie a muzeologie

Filozofická fakulta Masarykovy univerzity

Brno 1997

*Práce vznikla s podporou grantu Fondu rozvoje VŠ 1996 č. 554, Ústavu archeologie a muzeologie
Filozofické fakulty Masarykovy univerzity, Filozofické fakulty Masarykovy univerzity a grantu Gran-
tové agentury AV ČR č. A 9002603*

© Jiří Macháček 1997

ISBN 80-210-1562-4

1. PŘEDMLUVA (VLADIMÍR PODBORSKÝ)	9
2. ÚVOD (JIŘÍ MACHÁČEK, EDITOR)	11
3. DATABÁZOVÉ SYSTÉMY (JAROSLAV SMUTNÝ)	15
3.1 ÚVOD	15
3.2 DATABÁZOVÉ MODELY	16
3.2.1 <i>Klasické souborové zpracování dat</i>	16
3.2.2 <i>Hierarchické databázové modely</i>	17
3.2.3 <i>Síťové databázové modely</i>	17
3.2.4 <i>Relační databázové modely</i>	18
3.2.5 <i>Objektové databázové modely</i>	19
3.3 DALŠÍ DĚLENÍ DATABÁZOVÝCH SYSTÉMŮ	19
3.4 STRUČNÝ POPIS NEJZNÁMĚJŠÍCH DATABÁZOVÝCH PROGRAMŮ	20
3.4.1 <i>Access</i>	20
3.4.2 <i>Foxpro</i>	20
3.4.3 <i>Visual dBase</i>	21
3.4.4 <i>Oracle</i>	21
3.4.5 <i>SQL server</i>	22
3.5 ZÁKLADNÍ PRAVIDLA PRO TVORBU DATABÁZE	22
3.5.1 <i>Sestavení struktury datové báze</i>	23
3.5.2 <i>Zabezpečení dat</i>	31
3.5.3 <i>Vytvoření spojení s jinými programy</i>	31
3.6 ZÁVĚR	32
4. METODA ZÁKLADNÍHO ZPRACOVÁNÍ ARCHEOLOGICKÝCH VĚDECKÝCH DAT S POMOCÍ POČÍTAČOVÉ PODPORY (JIŘÍ MACHÁČEK)	33
4.1 ÚVOD	33
4.2 DEFINOVÁNÍ PROBLÉMU	33
4.2.1 <i>Příklad 1</i>	34
4.3 VÝBĚR DAT A TVORBA STRUKTURY DATABÁZE	34
4.3.1 <i>Příklad 2</i>	35
4.4 TESTOVÁNÍ RELEVANTNOSTI DAT	36
4.4.1 <i>Příklad 3</i>	37
4.5 PLNĚNÍ DATABÁZE	37
4.5.1 <i>Příklad 4</i>	38
4.6 DATABÁZOVÉ DOTAZY	38
4.6.1 <i>Příklad 5</i>	38
4.7 VIZUALIZACE DAT	39
4.7.1 <i>Příklad 6</i>	39
4.8 TESTOVÁNÍ VÝSLEDKŮ	41
4.8.1 <i>Příklad 7</i>	42
4.9 INTERPRETACE	43
4.9.1 <i>Příklad 8</i>	44
4.10 LITERATURA	44
5. KOMENTÁŘ KE „KÓDU MORAVSKÉ DOMÁCÍ ENEOLITICKÉ KERAMIKY“ (PAVEL KOŠTUŘÍK, JIŘÍ MACHÁČEK)	47
5.1 ÚVOD	47
5.2 TVORBA ARCHEOLOGICKÉHO KÓDU PRO PRAVĚKOU KERAMIKU (OBECNÉ ZÁSADY A PRAKTICKÁ REALIZACE)	47
5.3 DATABÁZE MORAVSKÉ DOMÁCÍ ENEOLITICKÉ KERAMIKY	50

5.3.1	Struktura databázové tabulky	50
5.3.2	Deskripční kód moravské domácí eneolitické keramiky	52
5.4	PŘÍKLADY UŽITÍ DATABÁZE	89
5.5	LITERATURA	91
6.	VÝPOČETNÍ TECHNIKA A ZPRACOVÁNÍ KAMENNÉ ŠTÍPANÉ INDUSTRIE (PETR NERUDA)	93
6.1	ROZDĚLENÍ CELKOVÉ ANALÝZY	93
6.1.1	Analýza jader	94
6.1.2	Analýza odbitých kusů	95
6.2	KONKRÉTNÍ APLIKACE	98
6.3	LITERATURA	104
7.	ARCHEOLOGICKÁ DATABÁZE ČECH (MARTIN KUNA)	105
7.1	ÚVOD	105
7.2	ZDROJE INFORMACÍ ADC	105
7.3	STRUKTURA DAT	106
7.4	SYSTÉM ARCHIV	107
7.5	SBĚR DAT	108
7.6	AKTUÁLNÍ STAV ADC	109
7.7	LITERATURA	109
8.	STÁTNÍ ARCHEOLOGICKÝ SEZNAM ČR – INFORMAČNÍ SYSTÉM ARCHEOLOGICKÝCH NALEZIŠŤ (DARA BAŠTOVÁ, LENKA KRUŠINOVÁ, ZUZANA SKLENÁŘOVÁ, PETR VOLFÍK)	115
8.1	ÚVOD	115
8.2	PRINCIPY A TVORBA DATABÁZE ARCHEOLOGICKÝCH NALEZIŠŤ	115
8.3	STRUKTURA DATABÁZE	117
8.4	APLIKACE SAS	119
8.4.1	Normalizace databáze, referenční integrita	120
8.4.2	Datový model databáze SAS	120
8.4.3	Stav a vývoj aplikace	122
8.5	PROBLEMATIKA ŘEŠENÍ INTRAVILÁNŮ	122
8.6	STÁTNÍ ARCHEOLOGICKÝ SEZNAM JAKO INFORMAČNÍ SYSTÉM	123
8.7	LITERATURA	123
9.	GEOGRAFICKÉ INFORMAČNÍ SYSTÉMY (MILAN KONEČNÝ)	127
9.1	CO JSOU GIS? DEFINICE A TERMINOLOGIE	127
9.1.1	Proč jsou GIS tak žádané?	127
9.1.2	Integrované systémy	129
9.1.3	Co jsou geografická data?	129
9.1.4	Složky GIS	129
9.1.5	Metadata	132
9.1.6	Datový model	133
9.1.7	Ukládání dat	133
9.2	ZDROJE DAT V ČR	133
9.2.1	ZABAGED	133
9.2.2	Archiv grafických dat KN	135
9.2.3	Databáze správních hranic územních celků	135
9.2.4	Vojenský topografický informační systém	135
9.3	NOVÉ TRENDY ROZVOJE GIS	136
9.3.1	GIS v sítích	136
9.3.2	Státní informační systém ČR	138

9.3.3	Co je „information superhighway?“ Úloha GIS.....	139
9.3.4	GIS a informační dálnice.....	140
9.3.5	Evropské aktivity: GI2000.....	141
9.3.6	Česká národní prostorová informační infrastruktura.....	141
9.4	LITERÁRNÍ A ČASOPISECKÉ ZDROJE O GIS A DIGITÁLNÍ KARTOGRAFIÍ.....	141
9.5	LITERATURA.....	142
10.	TEORIE A PRAXE ZPRACOVÁNÍ ARCHEOLOGICKÝCH VÝZKUMŮ S POMOCÍ PROSTŘEDKŮ GIS/LIS (MICHAL KUČERA, JIŘÍ MACHÁČEK).....	145
10.1	ÚVOD.....	145
10.2	GEOGRAFICKÉ INFORMAČNÍ SYSTÉMY – DEFINICE A DĚLENÍ.....	145
10.2.1	Definice GIS.....	145
10.2.2	Dělení GIS.....	146
10.3	FÁZE TVORBY GIS.....	147
10.3.1	Úvodní studie.....	147
10.3.2	Sběr dat.....	148
10.3.3	Správa dat.....	151
10.3.4	Analýzy nad daty.....	151
10.3.5	Prezentace dat.....	153
10.4	PROGRAMOVÉ VYBAVENÍ KA FF MU.....	153
10.4.1	MicroStation.....	153
10.4.2	MGE-PC2.....	154
10.5	PROJEKTY ŘADY POHAN.....	156
10.5.1	Úvod do problematiky projektů (dokumentace archeologického výzkumu).....	156
10.5.2	Úvodní studie.....	156
10.5.3	Sběr dat.....	163
10.5.4	Správa dat.....	164
10.5.5	Analýzy nad daty.....	164
10.5.6	Prezentace.....	165
10.6	LITERATURA.....	165
10.7	UŽIVATELSKÉ PROSTŘEDÍ POHAN.....	166
11.	GEOGRAFICKÝ INFORMAČNÍ SYSTÉM A VÝZKUM PRAVĚKÉ SÍDELNÍ STRUKTURY (MARTIN KUNA).....	173
11.1	ÚVOD.....	173
11.2	VÝZNAM GIS PRO ANALÝZU ARCHEOLOGICKÝCH A GEOGRAFICKÝCH DAT.....	174
11.3	ZÁKLADNÍ FUNKCE GIS IDRISI.....	174
11.4	VYTVOŘENÍ A ÚPRAVY ARCHEOLOGICKÉ MAPY.....	176
11.5	INTERPOLACE DAT.....	179
11.6	PŘÍKLADY.....	180
11.6.1	Příklad 1: Kontinuita obytných areálů.....	180
11.6.2	Příklad 2: Počet a rozsah sídelních areálů.....	181
11.6.3	Příklad 3: Vztah obytných areálů ke krajině.....	185
11.7	DODATEK: POZNÁMKY K VSTUPU DAT DO GIS A TISKU MAP.....	190
11.8	LITERATURA.....	192
12.	STATISTICKÁ ANALÝZA ARCHEOLOGICKÝCH DAT (ZDENĚK WEBER).....	197
12.1	ÚVOD.....	197
12.2	STATISTIKA PRO ARCHEOLOGY.....	197
12.2.1	Prvek, znak, soubor, výběr, základní soubor.....	199
12.2.2	Indikátory a pravidla jejich výběru.....	201
12.2.3	Studium vazeb mezi vlastnostmi (znaky) a jejich hodnotami.....	202
12.3	STATISTICKÉ PROGRAMOVÉ VYBAVENÍ POČÍTAČŮ.....	203

12.4	STATISTICKÉ PAKETY	204
12.4.1	<i>Praktická demonstrace funkce programu WINKS</i>	205
12.4.2	<i>Doporučený postup při statistické analýze archeologických dat</i>	205
12.5	TABULKOVÉ KALKULÁTORY	207
12.5.1	<i>Statistické funkce EXCELLU</i>	211
12.5.2	<i>Funkce pro správu seznamů a databází</i>	213
12.6	GRAFY DAT	214
12.7	VÍCEROZMĚRNÁ STATISTICKÁ ANALÝZA	214
12.8	ZÁVĚRY	214
12.9	LITERATURA	214
13.	K VYUŽITÍ SERIACE PŘI DATOVÁNÍ SÍDLIŠTNÍ KERAMIKY (VLADIMÍR SALAČ)	215
13.1	ÚVOD	215
13.2	PROBLEMATIKA DATOVÉ ZÁKLADNY	215
13.2.1	<i>Data ze sídlišť a pohřebišť</i>	216
13.2.2	<i>Příklad – data z laténských sídlišť v SZ Čechách</i>	216
13.3	PROGRAM – KOMBINAČNÍ ANALÝZA ARCHEOLOGICKÝCH NÁLEZŮ (KAAN)	220
13.3.1	<i>Tvorba matic</i>	220
13.3.2	<i>Práce s maticemi</i>	221
13.3.3	<i>Seriace</i>	223
13.4	PŘÍKLAD. SERIACE LATÉNSKÝCH SÍDLIŠTNÍCH KERAMICKÝCH SOUBORŮ ZE SZ ČECH	223
13.4.1	<i>Břešťany</i>	223
13.4.2	<i>Bílinsko</i>	229
13.4.3	<i>SZ Čechy</i>	230
13.5	POZNÁMKY K UŽITÍ SERIACE PŘI DATOVÁNÍ SÍDLIŠŤ	231
13.6	ZÁVĚR	234
13.7	LITERATURA	235
14.	SYNTÉZA STRUKTUR FORMALIZOVANÝMI METODAMI – VEKTOROVÁ SYNTÉZA (EVŽEN NEUSTUPNÝ)	237
14.1	FORMÁLNÍ ASPEKT ARCHEOLOGICKÝCH STRUKTUR	238
14.2	POSTUP FORMALIZOVANÉHO ŘEŠENÍ	239
14.2.1	<i>Krok 0 (sestavení deskriptivního systému)</i>	239
14.2.2	<i>Krok 1 (výpočet korelační matice)</i>	240
14.2.3	<i>Krok 2 (výpočet faktorů)</i>	240
14.2.4	<i>Krok 3 (rotace faktorů)</i>	241
14.2.5	<i>Krok 4 (faktorová skóre)</i>	242
14.2.6	<i>Krok 5 (validace)</i>	243
14.3	PŘEDNOSTI FORMALIZOVANÉHO HLEDÁNÍ STRUKTUR	243
14.4	PROSTOROVÝ ASPEKT ARCHEOLOGICKÝCH STRUKTUR	244
14.5	PŘÍKLADY	244
14.5.1	<i>Velikost zlomků laténské keramiky</i>	244
14.5.2	<i>Loděnický potok</i>	248
14.5.3	<i>Vikletice (šňůrové pohřebiště)</i>	254
14.6	PRAKTICKÉ PROVEDENÍ	257
14.7	LITERATURA	257

1. Předmluva (Vladimír Podborský)

Koncem 60. let se k nám dostaly první informace o možnostech zpracování hromadných dat pomocí výpočetní techniky. Obtížně se tehdy opatrovala základní literatura k této problematice a zkušeností s novou nadějnou metodikou práce s objemnými archeologickými soubory bylo pramálo. Nebyly ani výkonné kalkulačky a o počítačích a kybernetice jsme slyšeli jen z dálky.

O to více lákaly nové možnosti, zvláště když z rozvíjejících se velkých terénních výzkumů narůstal materiálový mobiliár do rozměrů, které hrozily zahlcením a o jehož analyticko-syntetickém využití jsme měli jen velmi nejasné představy. Východisko jsme tehdy spatřovali v matematicko-statistických postupech a přípravou pro jejich aplikaci, nejprve prostou „kupeckou“, ve výhledu snad mechanickou, bylo vytváření souborů děrných štítků se základními formalizovanými informacemi o archeologických pramenech, zejména o keramice; tehdy se ještě nevědělo nic o počítačových databázích.

Jak tomu bývá při zavádění nových technologií, řada vyznavačů nové „počítačové“ archeologie podlela iluzím o jedinečnosti a „samospasitelnosti“ matematicko-statistických metod a funkci lidského faktoru při závěrečných výstupech archeologických výzkumů byla ochotna přenést na počítače, k nimž si ale přístup zjednávala jen s největšími obtížemi; počítač sám měl řešit obtížné vědecké problémy za němé asistence badatele, který pouze shrne a zveřejní jeho výpočty. Na druhé straně se zformovala konzervativní opozice archeologů, kteří v nových metodách „ohne Seele“ viděli jen únikovou hru od skutečné vědy, iluzorní až fantaskní oportunistus, v kybernetice pak v krajních případech dokonce ideologicky nebezpečnou buržoasní metodologii. Každopádně ještě dlouho v 70. letech u nás platilo, že „dnešní badatelé používají počítačů se stejnou samozřejmostí jako antičtí služeb delfské věštinny“...

Za takových poměrů a při denních problémech s chodem pracovišť a udržením pracně sestavených týmů bylo účelné především připravit vhodné numerické deskripční systémy zpracovávaných archeologických pramenů a soubory výchozích dat, které by se později – až se výpočetní technika dostane do pracoven – řádně vytěžily a vyhodnotily. Na Katedře archeologie brněnské Filozofické fakulty několik takových „numerických kódů“ vzniklo a řada mladých archeologů se pro nové nekonvenční postupy práce s prameny zanítila. Myslím, že jim to bylo ku prospěchu zvláště poté, co se počítače i u nás staly běžnou pracovní pomůckou.

Bylo ovšem třeba, aby se radikálně změnila poměry. Přístrojové vybavení a nové technologie se mohly rozšířit teprve po revoluční proměně naší společnosti. Pionýři „nové archeologie“ mezitím sestárli, většinou však neztratili nadšení a s elánem se chopili nových možností, nastupujícím generacím se otevřely široké, dříve netušené perspektivy: nové metodologie, výjezdů do zahraničí, kontaktů s badatelskými centry pomocí Internetu atd.

Je potěšitelné, že na současném Ústavu archeologie a muzeologie Filozofické fakulty MU v Brně se tradice hledání nových metod udržela a našla schopné, pracovité a nadšené pokračovatele. Že se daří soustřeďovat kolem tohoto učiliště řadu předních českých specialistů v „počítačové“ archeologii (ale nejen v ní), je neméně významné; pro výchovu nových badatelů je to nesmírně užitečné. Předkládaná skripta by to měla dokumentovat.

Brno 2. ledna 1997

2. Úvod (Jiří Macháček, editor)

Hromadné zavádění počítačů do všech oblastí lidské činnosti pracujících s informacemi je jeden z nejradykálnějších projevů technického pokroku posledních desetiletí. Nevyhýbá se ani společenským disciplínám. Jak ukazují mnohé příklady (např. ze sociologie, psychologie, muzikologie, lingvistiky, pomocných věd historických), je i zde práce s výpočetní technikou velkým přínosem. K oborům, které mají pro využití počítačů ideální předpoklady, patří také archeologie.

Možnosti aplikace počítačů v archeologii jsou, jak lze soudit podle zkušeností mnohých zahraničních a některých našich institucí, nesmírně široké. Slouží ke zpracování obrovského množství různých typů, nashromážděného generacemi archeologů. Zásadní roli hraje počítač nejen v inventarizaci desítek až stotisícových položek nálezového inventáře, sumarizaci informací z terénních výzkumů či prospekční činnosti, ale i při jejich zpracování a prezentaci. Je při tom využíváno např. databázových systémů, statistických programů a nejnověji i tzv. GIS (Geografických informačních systémů).

Skripta, která připravil početný kolektiv autorů, jsou jedním z prvních pokusů o vysokoškolský studijní materiál, zaměřený speciálně na výuku počítačové podpory v archeologii. Svoji formou jsou proto poněkud odlišná. Nemají za úkol poskytnout jednoduchý návod k práci či sumarizovat v kompaktní formě naše dosavadní znalosti. Záměrem předloženého materiálu je naopak ukázat šíři a různorodost možností, které nové informační technologie v archeologii nabízejí. Kromě teoretických statí a metodických postupů zde proto naleznete i konkrétní aplikace a zkušenosti z praxe. Skripta by měla především probudit zájem o využití výpočetní techniky a inspirovat archeology k hledání nových metodologických cest v jejich oboru.

Skripta jsou rozdělena do tří tematických částí věnovaných databázím, geografickým informačním systémům a statistice.

První část seznamuje s nejrozšířenější počítačovou technologií, sloužící k zpracování velkého objemu dat. V úvodu se J. Smutný věnuje databázím z nejobecnějšího hlediska. Popisuje funkci databáze, představuje některé z nejznámějších databázových programů a definuje základní obecná pravidla pro její tvorbu. V další statí předkládá J. Macháček návrh postupu pro základní zpracování archeologických vědeckých dat s pomocí databází a tabulkových kalkulátorů. Zcela konkrétní aplikací je popis kódu a databáze moravské eneolitické keramiky od P. Košťuríka a J. Macháčka. Možnosti využití výpočetní techniky při zpracování kamenné industrie jsou tématem příspěvku P. Nerudy. Poslední dvě statí jsou věnovány nejrozsáhlejším projektům archeologických databází v České republice – Archeologické databázi Čech (M. Kuna) a Státnímu archeologickému seznamu ČR (L. Krušinová). Jedná se do jisté míry o konkurenční a paralelní projekty, které však hrají v naší archeologii významnou a jedinečnou roli. V příspěvcích se čtenář seznámí s metodologickým pozadím obou projektů a může srovnat různé přístupy při tvorbě archeologických databází velkého rozsahu.

Druhá část je tematicky věnována technologii pro archeology velice přitažlivé. Geografické informační systémy logicky navazují na databáze, k jejímž negrafickým datům dodávají důležité prostorové údaje. Zachycení reálného světa v počítači se tak stává mnohem komplexnější, což v konečných důsledcích umožňuje jeho hlubší pochopení. V úvodní statí této části se M. Konečný zamýšlí z obecného hlediska nad významem GIS pro dnešní společnost. Seznamuje čtenáře se zdroji digitálních geografických dat v České republice a snaží se přiblížit nové trendy rozvoje GIS. Součástí článku je i pojednání o metadatech, které by zasluhovalo pozornost všech archeologů, vytvářejících rozsáhlejší soubory nejen geografických dat. Stať M. Kučery a J. Macháčka je věnována problematice vektorového GIS. Po kapitole o teorii tvorby GIS/LIS následuje popis konkrétní aplikace systému při zpracování výsledků rozsáhlých systematických výzkumů na Pohansku u Břeclavi. Práce je doplněna instrukcemi k pracovnímu prostředí vyvinutému speciálně pro potřeby digitalizace archeologických grafických dat v programu MicroStation. Ve třetím z článků zaměřených na problematiku GIS předvádí M. Kuna funkce rastrového GIS a jeho možnosti ilustruje na příkladu pozoruhodného projektu výzkumu pravěké krajiny ve středních Čechách.

Závěrečné pasáže skript jsou věnovány statistickému vyhodnocení archeologických dat. Z. Weber zde seznamuje čtenáře s počítačovými programy zaměřenými na základní statistiku vědeckých

dat. Práce V. Salače je výsledkem jeho dlouholetých zkušeností se seriací archeologického materiálu. Metodiku tohoto statistického postupu dokumentuje na příkladu zpracování laténské sídlištní keramiky. E. Neustupný představuje možnosti jednoho z mocných statistických nástrojů – faktorové analýzy. Autor se problémem nezabývá pouze v teoretické rovině, ale i v bohatě dokumentovaných praktických příkladech, což je velice důležité pro pochopení složité metody. Pozoruhodné je propojení výsledků statistických analýz s GIS.

Je jisté, že tato skripta nemohla v plné šíři obsáhnout možnosti počítačových technologií při práci archeologa. Nedostalo se v nich místo např. na fotorealistickou vizualizaci archeologických monumentů, prezentaci archeologických dat prostřednictvím Internetu a mnoho dalších témat. Přesto lze doufat, že se stanou zdrojem inspirace pro novou nastupující generaci archeologů, jejichž úkolem bude hromadné zavádění nových metod v praxi.

V Brně 21.1. 1997

Databáze v archeologii

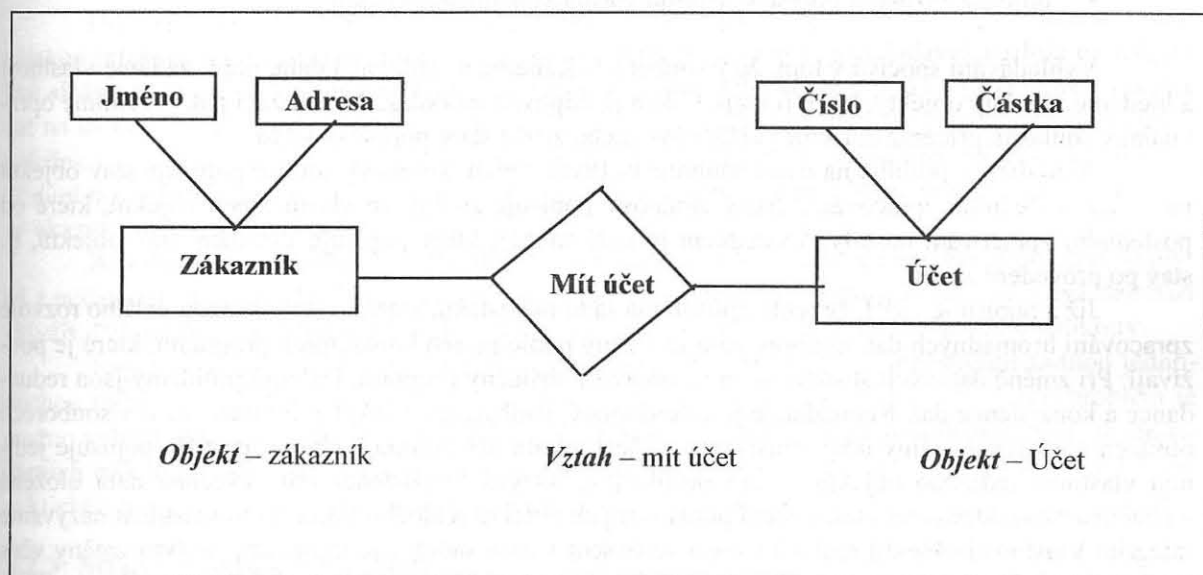
3. Databázové systémy (Jaroslav Smutný)

3.1 Úvod

Potřeba databází vzešla z problémů při zpracování různých ekonomických a evidenčních agend. Definovat a sestavit datový model, který by dobře reprezentoval příslušnou agendu je velmi náročné. Z mnohotvárného a složitého reálného světa musíme vyčlenit člověkem pochopitelnou předmětovou oblast. Z popisu předmětové oblasti musíme abstraktním odvozením vytvořit sémantický (významový) popis projektované báze dat, to znamená její konceptuální schéma. Popis datové báze je v každém případě velmi značným zjednodušením reality. Jediným kritériem jeho správnosti je schopnost zodpovědět všechny požadované dotazy. A dotazy musí být možno formulovat přirozeně, jejich správnost musí být intuitivně zřejmá a formální aparát použitý pro zápis konceptuálního schématu musí být převoditelný do prostředků přímo realizovatelných na počítači. Od počátku 70 let se objevují pokusy o zachycení významu – sémantiky dat ukládaných do informačních systémů. Vznik prakticky použitelného sémantického modelu, který umožňuje popsat konceptuální schéma dat, je spjat se s modelem „Entity-vztah“ (Entity-Relationship, E-R). Tento model umožňuje zachytit ve svých schématech objekty vnějšího světa, které jsou předmětem našeho zájmu v databázovém modelu. Jde tedy o objektově orientovaný model. V něm rozlišujeme dva základní typy objektů:

- entity – objekty v počítači nereprezentovatelné
- popisy (deskripce) – s jejich pomocí popisujeme vlastnosti entit

Popisy pak lze rozdělit do různých typů, přesně určujících operace, které lze s příslušnými popisy provádět. Vztahy mezi popisy a entitami pak nazýváme atributy, vztahy mezi entitami – relace. Pojem relace nemá žádnou souvislost s relacemi v relačním modelu. Příklad modelu E-R je na obr. 1. Jde o klasický případ evidence v bance. Zákazník je reprezentovaný jménem a adresou, konto pak číslem a částkou. Vztah mezi zákazníkem a účtem je „mít účet“. V praxi jsou možné další varianty. Zákazník má více účtů, ale jeden účet náleží právě jednomu zákazníkovi.



Obr. 1.

Z výše popsaného je zřejmé, že pro popis datové báze na konceptuální úrovni se používá obecný model dat. Je to souhrn prostředků pro:

- popis dat
- popis vazeb
- popis omezení pro jednotlivá data

Pomocí modelu E-R se dají snadno popsat objekty, jejich vlastnosti a vztahy, které chceme v databázi zobrazovat. Přičemž model E-R má pro popis struktury databáze podobnou funkci jako má vývojový diagram pro popis algoritmu. Pro realizaci musíme místo abstraktního objektu použít jeho implementaci, kterou je většinou záznamový model. Různé záznamově orientované modely se pak liší implementací vztahů. V další části pak budou popsány modely záznamově orientované (hierarchické, síťové a relační).

V zásadě se dá zpracování agend rozdělít na následovné úlohy:

- Tvorba datové báze
- aktualizace dat
- vyhledávání dat

Pro práci s daty na počítači bylo vyvinuto mnoho nejrůznějších metod, efektivních pro široké spektrum úloh. V následujících odstavcích budou popsány jednotlivé metody z hlediska historického vzniku.

3.2 Databázové modely

3.2.1 Klasické souborové zpracování dat.

Data jsou zpravidla organizována tak, že každý objekt zobrazované reality je popsán záznamem souboru, vlastností objektu položkami záznamu. Popis všech objektů tvoří soubor, což je základní objekt zpracování. Tedy stručně řečeno :

- záznam popisuje všechny vlastnosti jediného objektu
- soubor popisuje všechny vlastnosti všech objektů
- položka záznamu popisuje jedinou vlastnost jediného objektu

Vyhledávání spočívá v tom, že v souboru nalezneme požadovaná data, např. zadáme vlastnost a hledáme všechny objekty, které ji mají. Cílem je odpověď na dotaz. Aktualizaci pak rozumíme opravu dat v souboru, přičemž můžeme přidat nové, nebo zrušit staré popisy objektů.

Aktualizace probíhá na dvou souborech. První z nich, kmenový soubor popisuje stav objektů na konci posledního zpracování, druhý změnový popisuje změny ve vlastnostech objektů, které od posledního zpracování nastaly. Výsledkem je třetí soubor, který popisuje výsledný stav objektů, tj. stav po provedení změn.

Již z popisu je vidět, že tento způsob má řadu nedostaků, které se staly brzdou dalšího rozvoje zpracování hromadných dat. Soubory jsou navrženy podle potřeb konkrétních programů, které je používají. Při změně datových struktur se musí opravit i obslužný program. Dalšími problémy jsou redundance a konzistence dat. Redundance je vícenásobný, nadbytečný výskyt informace. Je-li v souborech obsažen opakovaně stejný údaj, musí mít za všech okolností stejnou hodnotu, protože popisuje jedinou vlastnost jediného objektu. Tento problém se nazývá konzistence dat. Všechna data uložená v souboru musí odpovídat vlastnostem popisovaných objektů reálného světa. Tuto vlastnost nazýváme integrita. Vlastnosti objektů reálného světa se ovšem v čase mění a je třeba, aby se tyto změny včas a při zachování konzistence odrazily v datech, které vlastnosti měnících se objektů popisují.

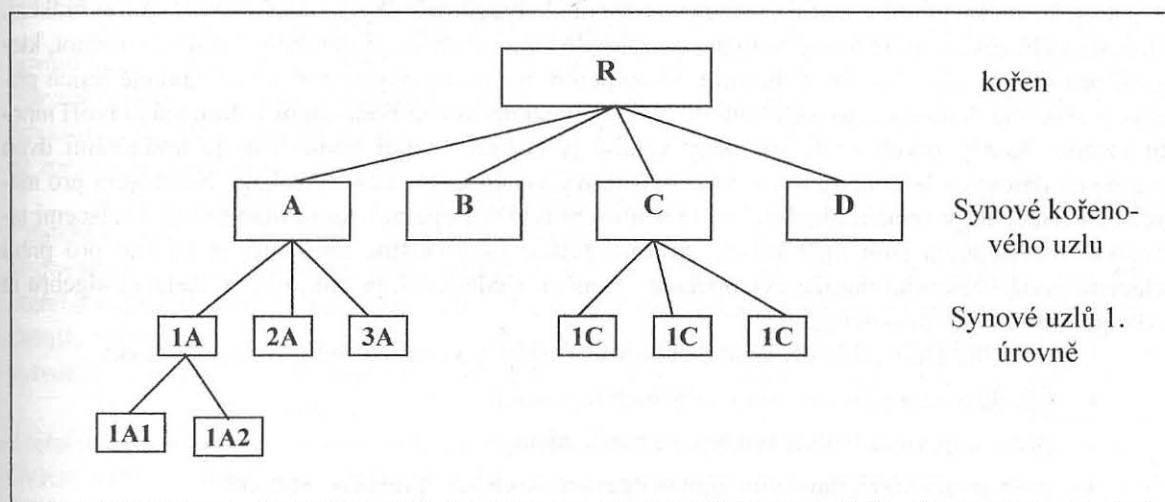
Význačným a velmi pokrokovým krokem bylo oddělení datových souborů od aplikačních programů. Vznikající databázové systémy pak můžeme charakterizovat následujícími vlastnostmi:

- struktury aplikačních programů a datových souborů jsou odděleny

- přístup k datům je možný jen prostřednictvím programů databázového systému nikoli přímo
- dotazy nejsou pevné
- je umožněn přístup více uživatelů současně a vyřešena ochrana dat proti zneužití

3.2.2 Hierarchické databázové modely.

Dalším logickým stupněm ve vývoji databázových modelů je hierarchický databázový systém. V tomto modelu jsou data organizována na základě stromové struktury vycházející z kořene. Jednotlivé datové struktury jsou umístěny na různých úrovních ležících podél větví, které vycházejí z kořene. Datovým strukturám na jednotlivých úrovních se říká uzly. Pokud z uzlu nevychází další větev, říká se mu list. Stromová struktura definuje rodičovské a sourozenecké vztahy mezi různými prvky v databázi. Velmi připomíná organizační schema, které všichni známe pod názvem „pavouk“.



Obr. 2.

Dle obr. 2 je velmi dobře vidět, oč je tento systém lepší než předchozí. Chce-li uživatel najít nějakou informaci, není potřeba prohledávat celý soubor. Stačí, když se požadavek rozloží na jednotlivé složky a jde se od kořene na příslušnou větev a z ní na další větev a pod. . Na fyzické struktuře dat na disku v tomto modelu prakticky nezáleží. Většinou jsou data uložena jako zřetězený seznam položek s ukazateli od otce k synům a od sourozence k sourozenci. Poslední list má ukazatel nulový. Přidání nových položek na jakékoliv úrovni je rovněž poměrně snadné. Musí se pouze změnit koncový ukazatel na ukazatel na další sourozenecký uzel.

Ačkoli je tento model velmi flexibilní, má i některé problematické rysy. První problém vyplývá z počáteční struktury databáze, která je pevně dána. Programátor ji musí definovat, když databázi vytváří. Pak již nemohou být rodičovské vztahy změněny bez úplného přepracování celé struktury.

Nejzávažnějším problémem hierarchických databází však je to, že neposkytují žádnou jednoduchou metodu pro definování křížových vztahů neboli vztahů „mnohý k mnoha“. Tento problém je možno vyřešit ukládáním dat ve více kopiích na různých úrovních. Jiným způsobem je přidání sekundárních rodičovských a sourozeneckých ukazatelů do hierarchické struktury. Pak ale vzniknou četné cyklické vztahy a situace se značně zkomplikuje.

3.2.3 Síťové databázové modely.

Síťový model popisuje databáze, v nichž existují vztahy „mnohý k mnoha“ (tj. vícenásobné rodičovství). Vztahy mezi položkami se obvykle označují jako „množiny“ z důvodu odlišení od striktně rodičovských vztahů v hierarchickém modelu. Síťový databázový model je založen na ukazatelích

(lineárních nebo cyklických, které vyjadřují vztahy mezi jednotlivými položkami. Flexibilita při vytváření vztahu „mnohý k mnoha“ je nejsilnější zbraní síťového modelu. Avšak vzájemné vztahy mezi různými množinami se mohou stát velmi komplikovanými a obtížně mapovanými.

Stejně jako hierarchický může i síťový model poskytovat velkou rychlost zpracování, zvláště použijí-li se indexové ukazatele směřující přímo k první položce v prohledávané množině.

Síťový model má však stejné problémy se strukturou, stejně jako hierarchický model. Jakmile je databáze vytvořena, každá další změna na jiný systém množin vyžaduje, aby byla vytvořena nová struktura. Přidávání nových položek je však oproti předchozímu modelu jednodušší. Stačí pouze definovat novou množinu a změnit potřebné ukazatel, aby se nová množina správně začlenila do zbývajících dat.

3.2.4 Relační databázové modely.

Model relační databáze definoval v roce 1970 Dr. E. F. Codd. Tento model je založen na teorii množin, relacích a relačním kalkulu. Opouští myšlenku rodičovských vztahů mezi různými položkami. Místo toho organizuje data do uspořádaných n-tic. Prvky takových množin se obvykle znázorňují jako řádky tabulky. Každá položka se stává sloupcem tabulky a každý záznam jejím řádkem. Matematickým základem je teorie množin, na základě které se definují množiny datových hodnot, které jsou pro systém atomické (nedělitelné). Seskupením těchto datových hodnot na základě jejich příbuznosti získáme doménu datových hodnot. Vztahy mezi datovými hodnotami v doménách tvoří množinu vztahů. Každý prvek z této množiny vztahů je tvořen dvojicí hodnot, to je hodnotami dvou atomických datových hodnot ze dvou domén. Takový vztah se pak nazývá relace. Nástrojem pro manipulaci s relacemi je relační algebra. To je souhrn pravidel a operací, které manipulují s relacemi jako celky a výsledkem jsou opět relace. Protože relace jsou vlastně množiny, je možné pro práci s relacemi použít základní množinové operace – součin, sjednocení, průnik, rozdíl. Relační algebru je možné použít jako dotazovací jazyk.

V roce 1982 Dr. Codd definoval databázový systém pomocí následujících podmínek:

- všechna data jsou uložena v relačních tabulkách
- neexistují vícenásobné přístupové mechanismy
- dotazovací jazyk musí umožňovat operace – selekce, projekce, spojení

V roce 1985 tyto podmínky pravidla byla upřesněna a doplněna:

- práce s daty je uskutečňována pouze relačními prostředky
- databáze na logické úrovni je dána pouze relacemi
- musí existovat přístupový mechanismus k uloženým datům (název tabulky, název sloupce, hodnota přístupového klíče)
- jsou definovány prázdné hodnoty
- veškerý popis databázových objektů je uchován v relačních tabulkách, manipulace s daty je založena na relační algebře
- existuje jazyk pro definici dat včetně vytváření pohledů a podpory vstupů a výstupů
- existují příkazy pro manipulaci s daty (příkazy – insert, delete, update)
- je možné definovat přístupová práva
- logické relace jsou nezávislé na jejich fyzické reprezentaci (fyzická datová nezávislost)
- funguje logická datová nezávislost

Na obr. 3 je pak praktická ukázka relačního databázového modelu. Jde o zjednodušené přepracování modelu E-R z obr 1.

Jméno	Adresa	Číslo účtu	Částka
Čapek	Veveří 95	00001	10.000,-
Strauss	Slatinská 5	00002	25.000,-
Němcová	Česká 5	00003	15.230,-
Veselý	Grohova 1	00004	78.800,-
Fišer	Veveří 91	00005	125.560,-
Čapek	Veveří 95	00006	16.802,-

Obr. 3.

Relační model má oproti hierarchickému a síťovému modelu řadu výhod. Snad nejdůležitější z nich je naprostá flexibilita v popisu relací mezi různými položkami. Změna struktury databáze spočívá v pouhém přidání nebo zrušení sloupce v tabulce. A to nijak neovlivní ostatní sloupce ani ostatní tabulky. Kdykoliv lze vytvořit novou tabulku jako projekci (podmnožinu) existujících tabulek. Staré tabulky lze podle uvážení zrušit. To, že k provedení není nutno znovu vytvářet celou strukturu datové báze, se pochopitelně odrazí v zachování integrity dat.

Nejdůležitější rozhodnutí, před které jsme při vytváření datových struktur postaveni, je definování tabulek. Proces dekompozice dat do podmnožin pro jednotlivé tabulky se nazývá normalizace. Hlavním cílem relačního databázového modelu je zachování maximální integrity dat. Aby mohl být považován skutečně za relační, musí zcela znemožnit přístup k datům jinak než prostřednictvím dotazů zpracovávaných jím samým.

Relační model vyžaduje, aby programy pracovaly s daty, aniž by znaly jejich umístění v databázi. Zde je velký rozdíl oproti dříve popsaným modelům, kde programy musí vyhledávat data sledováním řady ukazatelů. Program vznášející dotaz na relační databázi prostě požádá o data, které je potřeba vyhledat, systém řízení báze dat provede potřebná vyhledávání a poskytne žádanou informaci. Detaily vyhledávání jsou specifické pro každý programový produkt. Prohledávání může být urychleno vytvořením indexového souboru pro jeden nebo více sloupců tabulky (bude vysvětleno později).

Jedinou nevýhodou relačního databázového modelu je, že tabulka reprezentující vztah v relačním modelu zabírá více místa než reprezentace vztahu v hierarchickém a síťovém modelu. Ovšem naprosto převažující výhodou relačního modelu, jak již bylo popsáno dříve, je fakt, že obsahuje jen dvourozměrné tabulky, které se navíc snadno implementují jako soubory.

Navíc se relační databáze obtížně vyrovnávají s rozměrnými daty proměnné velikosti (texty, obrázky, zvuk, video-sekvence). Děje se tak pomocí datových typů BLOB (Binary Large Objects).

Na závěr tohoto oddílu je třeba podotknout, že naprostá většina databázových programů dnes provozovatelných na nejrůznějších platformách je založena na relačním modelu.

3.2.5 Objektové databázové modely

U objektově orientované databáze jsou určité položky shrnuty do jednoho objektu. Objekt obsahuje data těchto položek a navíc má definovány funkce a způsoby jak s daty nakládat.

Velkou výhodou objektových databází je možnost práce s různými typy dat, tedy i s nestrukturovanými. Tento typ databází je však v současné době ve fázi vývoje, nicméně do budoucna se jeví velmi perspektivní, obzvláště s nástupem a rozvojem multi-mediálních aplikací.

3.3 Další dělení databázových systémů

Dosud jsme se zabývali popisem databázových systémů z hlediska typu datového modelu. Podle způsobu spolupráce aplikace s databázovým systémem lze databázové systémy na :

- souborové databázové systémy (např. typu xBase)
- systémy typu klient -server

Rozdíl mezi oběma typy se projeví především při práci v počítačové síti. Při použití souborového databázového systému veškeré zpracování probíhá na lokálních počítačích. Na počítači typu server jsou uložena pouze data, která se při zpracování přesouvají na lokální počítače. Tento způsob zpracování dat je vhodnější pro nasazení na samostatných počítačích. Při provozu v síti se tento systém projeví zvýšenou síťovou komunikací.

Systémy klient-server pracují na principu dataz-odpověď. Aplikace na lokálním počítači zformuluje dotaz a pošle jej databázovému systému běžícímu na serveru (poskytovateli služeb). Ten jej zpracuje a zpět pošle výsledky operace. Velmi významným prostředkem pro formulování dotazů a příkazů směřovaných z klienta na server je SQL (Structured Query Language).

3.4 Stručný popis nejznámějších databázových programů

3.4.1 Access

Jde o relační databázový produkt od firmy Microsoft. Dnes existuje ve verzi 7.0. Je určen pro operační systémy Windows 3.11, Windows 95, Windows NT 3.51, 4.0. Tohoto produktu lze využít pro vytváření nejrůznějších databázových aplikací – od nejjednodušších až po komplexní systémy. Access umožňuje stejně jako ostatní databázové programy též kategorizovat strukturu tabulek, ve kterých jsou uloženy informace, vytvářet uživatelské vstupní formuláře, zadávat, editovat a mazat data, provádět výpočty a prohlížet nebo tisknout výstupní sestavy, obsahující vybrané informace z jedné nebo více tabulek. Uživatelé všech stupňů pokročilosti mohou pracovat s Accessem interaktivně. Pomocí klávesnice a myši je možno manipulovat s objekty na obrazovce, pracovat s příkazovými tlačítky, vybírat položky ze seznamu a volby z rolovacích menu, připojených k hlavnímu menu. Pro zkušenější uživatele a pro použití v stálých ukončených aplikacích je Access vybaven nástroji pro práci s makry. Makro lze použít pro automatizaci operací tak, že zaznamenáme posloupnost prováděných akcí. Tyto akce lze pak později zopakovat spuštěním makra. Spojením makra s formuláři a sestavami lze vytvořit plně automatizovanou aplikaci s uživatelským menu, která může být použita pracovníky, kteří o Accessu nic nevědí. Pro velké komplexní aplikace Access je vybaven programovacím jazykem, nazvaným Access Basic. Tento jazyk je velmi příbuzný jazyku Visual Basic. Součástí Access Basicu jsou knihovny zahrnující příkazy pro práci s tabulkami, dotazy, formuláři a sestavami Accessu. Access obsahuje rovněž speciální nástroje, které mají rychle a jednoduše pomoci uživatelům. Jde o pomocníky (Wizards), kteří uživatele vedou při tvorbě datové struktury, formuláře, sestavy a dotazu. Tyto vytvořené objekty je potom možné dále upravovat. Access samozřejmě umožňuje vyměňovat informace s ostatními programy (jiné databázové programy, tabulkové kalkulátory, textové editory a pod.), které pracují rovněž pod Windows. Velmi výhodné je také použití Access při správě dat uložených ve větším databázovém systému typu klient-server.

3.4.2 Foxpro

Jde opět o relační databázový produkt od firmy Microsoft. Dnes existuje ve verzi 3.0. Je určen pro operační systémy, Windows 95, Windows NT 3.51, 4.0.

Microsoft FoxPro je již dlouho rychlým a výkonným nástrojem pro programátory aplikací Xbase. Nejnovější verze rozšiřuje jeho možnosti objektovou orientací. Výrazná zdokonalení této verze v rychlém vývoji aplikací, přizpůsobivosti a přístupu k datům udělala z FoxPro plnoprávného člena rodiny vizuálních nástrojů Microsoftu. Microsoft Visual FoxPro 3.0, jak se nyní produkt jmenuje, v sobě kombinuje vizuální vývojové prostředí s novými nástroji pro tvorbu řešení, která propojují aplikace Microsoft Office a Microsoft BackOffice. Jeho přizpůsobivá architektura umožňuje, aby již dříve vytvořené programy Xbase pracovaly v operačních systémech Windows NT a Windows 95. Uživatel si může zvolit takový operační systém, který mu nejlépe vyhovuje.

Ke zdokonalení uživatelských aplikací, napsaných ve Visual FoxPro, je možné nyní používat objekty a ovládací prvky OLE (o co jde, bude vysvětleno později). Pomocí OLE technologie lze ovlá-

dat aplikace Microsoft Office a editovat je přímo na místě z Visual FoxPro. Stovky ovladačů OLE, nabízených dalšími výrobci, tyto možnosti dále rozšiřují.

Data lze propojit s Visual FoxPro bez ohledu na to, zda jsou uložena v aplikacích, databázích klient-server nebo na lokálních paměťových médiích. Vestavěné 32-bitové ovladače ODBC 2.0 (o co jde, bude vysvětleno později) dávají přístup ke všem populárním databázovým serverovým aplikacím.

3.4.3 Visual dBase

Jde o relační databázový produkt od firmy Borland. Dnes existuje ve verzi 5.5. Je určen pro operační systémy, Windows 95, Windows NT 3.51, 4.0. Jde o databázové prostředí velmi podobných vlastností jako Access. Opět i zde se můžeme setkat s maximální snahou tvůrců pomoci uživatelům v návrhu jednotlivých částí aplikace. Program obsahuje tkz. experty, což jsou programové moduly, které v maximální možné míře vedou uživatele při návrhu.

Visual dBase však je více vývojové programátorské prostředí. Jazyk, kterým Visual dBase disponuje, vychází ze standardu Xbase, ale kromě toho velmi vyhovuje všem požadavkům na objektové orientovaný jazyk.

3.4.4 Oracle

Jde o výkonné, relační SQL databázové prostředí typu klient/server. Programy tohoto typu byly donedávna doménou velkých sálových počítačů. Ovšem s nárůstem výkonnosti systémů PC se začaly používat i na téhle platformě. Výše popsané databázové systémy v nejnovějších verzích mohou pracovat v režimu klient/server. Ovšem princip práce je jiný. Data se ze serveru přenášejí ke klientu k dalšímu zpracování. U skutečného systému klient/server je klient zbaven této zátěže. Posílá pouze dotazy SQL a server mu po jejich zpracování zasílá pouze výsledky. Tím je možné zvládnout mnohem složitější úlohy.

Velkou výhodou databází, které přicházejí z prostředí velkých sálových počítačů je integrita dat. Dále tyto systémy podporují transakční zpracování dat. To znamená, datové změny jsou zapisovány na disk až po potvrzení příkazem. Když se zhroutl program, operační systém a pod. dá se vše vrátit do původního stavu. Jsou-li dvě transakce spojeny, systém neumožní provést pouze jednu z nich. Dalšími zabezpečujícími funkcemi jsou transakční deníky zaznamenávající všechny změny databáze, a zrcadlení, zapisující data na dvě oddělená místa. V relačním databázovém modelu je také velmi důležitá referenční integrita, která hlídá propojení (relace) mezi jednotlivými tabulkami modelu. Nesmí zůstat žádní sirotci (odkazy nikam). Systém Oracle rovněž podporuje distribuované databáze. To znamená, že data jsou uložena na různých počítačích. Uživatelé pak mohou pracovat s daty bez ohledu na jejich umístění, nebo-li vypadá to jako když data jsou na jediném počítači. S touto technologií přichází doba veřejných datových skladů, z kterých uživatel je schopen používat jemu přístupná data, která jsou umístěna kdekoli ve světě.

V současné době existuje systém Oracle na cca 90 platformách od laptopů až po paralelní superpočítače. V současné době je na trhu Oracle Workgroup 2000. Toto seskupení zahrnuje pět produktů:

- Oracle 7.2 Workgroup Server
- Personal Oracle 7.2
- Oracle Power Objects
- Oracle Objects for OLE
- Oracle Radio Agents

Workgroup Server je určený pro pracovní skupiny spolupracující hlavně na sítích LAN (ale i WAN). Předností tohoto systému je také velmi intuitivní grafické uživatelské prostředí. Pomocí myši lze velmi jednoduše nastavovat všechny parametry, nastavovat uživatelská práva, zřizovat konta a provádět ostatní administrativní úkony.

Personal Oracle je přizpůsoben pro lokální provoz. Jinak má všechny vlastnosti jako Oracle Workgroup Server. Značnou výhodou jsou poměrně nízké nároky na hardware počítače a kapacitu paměti.

Power Objects je vývojové vizuální prostředí, které je určeno pro vývojáře aplikací.

Objects zajišťuje vývojářům aplikací velmi efektivní přístup k datům v databázích Oracle 7.

Radio Agents je prostředek pro připojení mobilních aplikací k podnikovým databázím. Umožňuje práci i bez stálého spojení se sítí a připojení je možné provést i běžnými komunikačními prostředky jako je telefonní linka.

Velmi zajímavý je rovněž produkt Oracle Media Server. Tento produkt podporuje práci s multimediálními daty.

3.4.5 SQL server

Jde o výkonný databázový server od firmy Microsoft, dnes ve verzi 6.5. Původně byl vyvíjen s ve spolupráci s firmou Sybase pro platformu OS/2, ale časem přešel jeho vývoj zcela do režie Microsoftu. Dnes je výhradně určen pro operační systém Windows NT. Díky velké integraci s tímto operačním systémem využívá maximálně možné míře jeho služeb, funkcí a zabezpečení.

Zároveň je s Windows NT přenositelný na všechny platformy, kde Windows NT pracují. Jde o platformy Intel, Alpha, MIPS a PowerPC. Stejně přenositelné jsou i aplikace, takže není nutno pro každou platformu připravovat speciální moduly. Úzká vazba na Windows NT tak zaručuje uživatelům velkou úsporu nákladů, protože celý systém může růst s výkonem použitého serveru – od jednoprocessorových k víceprocesorovým bez nutnosti složitých zásahů.

Microsoft SQL 6.5 disponuje všemi vlastnostmi, které přísluší spíčkovému serveru.

V dalším textu bude provedeno shrnutí a popis důležitých vlastností datových serverů:

- transakce
- replikace
- triggers
- stored procedures

U distribuovaného zpracování, kdy je zpracování rozloženo přes více serverů, musí být kladen velký důraz na zabezpečení správného provádění transakcí. Transakce je proces, při kterém dochází k práci s daty. Data mohou být měněna, rušena, doplňována. Každý takový proces je uzavřený úsek operací, který začne, provádí operace s daty a pak skončí. A právě konec operace je velmi důležitý. Mohou totiž nastat dva konce operace. Za prvé – vše proběhlo dobře a transakci ukončuje „commit“ – potvrzení o správnosti. Teprve v tomto okamžiku jsou všechny operace, které byly součástí transakce, potvrzeny a nová data jsou platná. Pokud během transakce došlo ke konfliktu (např. výpadek systému a pod.), nastane „rollback“ a data zůstanou beze změny, v podobě jako před transakcí.

Pro distribuované zpracování disponuje SQL 6.5 prostředkem zvaným „dvoufázový commit“. Transakce je zplatněna až tehdy, když jsou zplatněny jednotlivé změny na všech jednotlivých serverech a poté je transakce zplatněna jako celek.

Replikace jsou také operace pro podporu distribuovaného zpracování dat. V podstatě se jedná o automatické zasílání změn dat na místa tzv. předplatitelů. Rozdíl je v tom, že úspěšné přijetí nemusí být potvrzeno. Příkladem může být automatická aktualizace např. účetních dokladů z poboček velkého podniku na centrální server a naopak.

Triggers (spouštěče) jsou procedury (programy), které jsou po vytvoření spojeny s databází a díky platnosti pro všechny aplikace přistupují ke sledovaným datům a reagují podle toho, co mají naprogramováno. Čili mohou uživatele varovat, že něco není v pořádku, mohou spustit program a pod.

Stored procedures jsou programy, které zabezpečují vlastní výkonnou činnost databázového serveru (např. provádění výpočtů).

3.5 Základní pravidla pro tvorbu databáze

Každý návrh databáze by měl projít následujícími fázemi :

- sestavení struktury datové báze
- tvorba formulářů
- sestavení dotazů
- tvorba sestav
- zabezpečení dat
- vytvoření spojení s jinými programy (např. prezentačními)

V další části textu budou vysvětleny jednotlivé kroky při sestavení databáze. Protože není cílem tohoto textu věnovat se konkrétnímu programovému produktu, vše bude vysvětleno v co nejobecnější rovině. Ve výkladu budou použity základní SQL příkazy. Před tím je nutno alespoň velmi stručně vysvětlit, co že to vlastně SQL jazyk je.

SQL (Structured Query Language) vznikl v polovině 70. let jako potřeba standardizovat dotazy tak, aby byla umožněna spolupráce aplikací s různými databázovými systémy. V současné době SQL je standardním prostředkem pro přístup a manipulaci s daty nezávisle na použitém relačním databázovém systému.

3.5.1 Sestavení struktury datové báze

Ještě dříve než začneme používat některý z konkrétních databázových systémů, musíme vytvořit logický návrh databáze, který popisuje objekty, jejich vlastnosti a vztahy mezi nimi. Co budeme považovat za objekty, které jejich vlastnosti popíšeme a které vztahy mezi objekty budou do návrhu zahrnuty, záleží na tom, jaké výsledky od zpracování očekáváme. Jinými slovy na dotazech, na které chceme získat odpověď z báze dat.

Jak bylo uvedeno již dříve, relační datový model reprezentuje množina dvourozměrných tabulek (viz. obr. 3). Na každou tabulku se odkazujeme jejím jménem. Každý řádek reprezentuje datovou větu (record) a každý sloupec pak datovou proměnnou. Množina datových proměnných (položek, fields) pak definuje datovou strukturu tabulky (structure).

Při sestavování tabulek musíme přiřadit každé proměnné (dále položce) jméno a datový typ. U některých položek pak ještě určit maximální délku, implicitní hodnotu a formát zobrazení. Jméno položky může obsahovat běžné alfanumerické znaky. Není vhodné používat znaky typu „&, *, ^, #, @, (,), %“ a pod.. Na pořadí položek v tabulce v podstatě nezáleží. Nicméně je dobré si pořadí položek v tabulce promyslet, protože tato posloupnost položek ve struktuře tabulek slouží jako implicitní pořadí zobrazení. Jméno položky je jedinečné pro danou tabulku. V různých tabulkách mohou být sloupce se stejným jménem. Počet položek tabulky je většinou omezen na 255, počet řádků je omezen pouze kapacitou ukládacího media (pevného disku). Může být i nulový. Pak je tabulka prázdná.

Většina databázových programů umožňuje při sestavování tabulek použít tři režimy:

- režim „návrh tabulky“ (Table Design) v němž ručně přiřazujeme jména a datové typy
- režim „průvodce tabulkou“ v rámci kterého vybíráme z již předem výrobcem software definovaných tabulek a podle uvážení modifikujeme jména položek a datové typy
- použitím SQL příkazů

Postup 2. je vhodný pro méně zkušené tvůrce, 3. je nejvíce používán správci velkých datových serverů.

V dalších řádcích budou popsány alespoň velmi obecně datové typy položek používaných při sestavování tabulek. Při použití konkrétních produktů musíme počítat samozřejmě s odlišnostmi.

3.5.1.1 Nejběžnější datové typy

Text – textové položky mohou obsahovat všechny znaky, které lze zadávat z klávesnice. To znamená číslice, písmena, interpunkční znaménka a grafické symboly. Délka textové položky může být až 255 znaků, ale zabírá pouze prostor nutný pro uchování vložených údajů.

Memo – tato položka může obsahovat všechny znaky, které lze zadávat z klávesnice. Délka této položky může být až 32 000, ale zabírá pouze prostor nutný pro uchování vložených údajů.

Number – položka tohoto typu může obsahovat pouze číslice, desetinnou tečku nebo znaménko minus. Položek typu „number“ může být více druhů:

byte – obsahuje pouze kladné hodnoty typu integer (celá čísla) od 0-255

integer – celá čísla od -32 768 do 32 767

long integer – celá čísla od -2 147 483 648 do 2 147 483 647

single – číselné hodnoty v jednoduché přesnosti od $-3.4 \cdot 10^{38}$ do $3.4 \cdot 10^{38}$

double – číselné hodnoty v dvojitě přesnosti od $-1.797 \cdot 10^{308}$ do $1.797 \cdot 10^{308}$

Counter – položka tohoto typu uchovává hodnoty „long integer“, které jsou automaticky o jedna zvyšovány přidáním nového záznamu.

Date/Time – tato položka reprezentuje dny v kalendáři nebo hodiny.

Ano/Ne – položky tohoto typu obsahují logické hodnoty, které mohou být vyjádřeny jako 0/1, ano/ne, nebo true/false.

OLE object – tato položka uchovává data, která nejsou reprezentována znaky. Mohou to být například grafické objekty, zvukové záznamy, fotografie, video sekvence a pod. Databázové prostředky obvykle nabízí dva způsoby práce s tímto typem proměnné. První možnost je, když do proměnné vložíme úplnou kopii objektu (fotka, obrázek, schema). Současně objekt může (ale také nemusí) existovat mimo databázi. Další možností je možnost definovat spojení mezi OLE objektem a dokumentem vytvořeným externím programem. V takovém případě do datové položky je vložena pouze informace, která je potřeba pro nalezení uloženého souboru s daným objektem.

Nutno podotknout, že vypsání typů položek reprezentují datové typy databázového programu „Access“. Jiné databázové programy používají ekvivalentní datové typy, i když s jiným označením.

Jak již bylo dříve uvedeno, lze tabulku vytvořit SQL příkazem. Jeho syntaxe je uvedena níže. Velké a malé písmena SQL jazyk nerozlišuje.

```
CREATE TABLE <jmeno tabulky>  
    (<polozka> <typ> (<delka>),  
    (<polozka> <typ> (<delka>)] ...);
```

Data do tabulky pak ukládáme po řádkách příkazem:

```
INSERT INTO <jmeno tabulky> VALUES (<seznam konstant>);
```

Jako malý příklad uvedeme vytvoření struktury následující datové tabulky a naplníme ji dvěma řádky.

```
CREATE TABLE vcelari ( jmeno      TEXT(20),  
                      adresa      TEXT(25),  
                      narozen     DATE,  
                      kraj        TEXT(15),  
                      poc_vcelstev INTEGER,  
                      litru_medu   SINGLE);
```

```
INSERT INTO vcelari
```

```
VALUE ('Jaroslav Smutný', 'Veveří 95', '27.01.1970', 'Jihomoravský', 5, 20);
```

```
INSERT INTO vcelari
```

```
VALUE ('Jan Poláček', 'Grohova 78', '25.11.1950', 'Jihočeský', 24, 125);
```

3.5.1.2 Indexy

Indexy hrají v každé databázi důležitou roli. Je to vlastně pomocná složka databáze, která uchovává informace jedné nebo několika položek v tabulce. Ty jsou nazývány „indexy, indexové klíče“. Indexy můžeme definovat už při vytváření tabulek, nebo později. Indexy slouží k urychlení operací jako je seřazení a hledání záznamů podle klíčových položek. Takže komplexně řečeno, indexové klíče jsou vhodné pro následující účely:

- k urychlení procesu zobrazení záznamu v jiném pořadí, než v jakém byly zadány
- k urychlení vyhledávání jednotlivých záznamů podle jednoznačných hodnot v klíčových položkách
- k urychlení dotazů, které vybírají záznamy podle obsahu klíčových položek
- k zamezení zadávání záznamů se stejnými indexovými klíči
- podporu vazeb mezi tabulkami
- k určení implicitního pořadí zobrazování záznamů

Většina tabulek by měla mít minimálně jeden zvláštní klíč, který se nazývá „primární klíč“. Tento určuje implicitní pořadí zobrazení při prohlížení záznamů a navíc slouží jako jednoznačný identifikátor. Následující obrázek velmi zjednodušeně ukazuje princip použití indexových klíčů. V tomto případě jde o vzestupné seřazení podle položky „příjmení“.

Příjmení	Jméno
Smutný	Jaroslav
Adam	Jan
Kadaňka	Pavel
Zvěřina	Pavel
Bílek	Miroslav

Číslo věty	Příjmení
2	Adam
5	Bílek
3	Kadaňka
1	Smutný
4	Zvěřina

V jazyce SQL vypadá příkaz na indexování následovně:

```
ORDER BY <seznam klíčů>
```

```
SELECT *
```

```
FROM vcelari
```

```
ORDER BY litru_medu DESC;
```

DESC – znamená sestupně

3.5.1.3 Relace

Před dalším postupem musí být provedena optimalizace struktur datové báze. Většina produktů založených na relačním databázovém modelu umožňuje pracovat s mnoha tabulkami současně. Takže lze se vyhnout zbytečnému ukládání stejné informace vícekrát. Z jedné tabulky lze odstranit informace, které obsahuje jiná tabulka. Tento způsob šetří místo na disku a brání nesrovnalostem. Další možnost, jak se vyhnout nadbytečnosti je zlepšit konzistenci dat a snažit se mít hlavní tabulky co nejmenší.

V relačním databázových modelech je možné tabulky různým způsobem provázat. Aby se mohly tabulky provázat, musí existovat tzv. propojovací pole, s jehož pomocí můžeme realizovat nejčastěji užívané propojení „otcovské tabulky“ se „synovskou tabulkou“. Přitom způsob vazby v aplikaci není statický, neměnný, ale naopak ho můžeme dynamicky měnit. Tabulka, která byla v jednom relačním uspořádání „otcem“, klidně může být v jiném uspořádání „synem“, nebo jiným potomkem. To je také hlavní rozdíl, kterým se relační model liší od dřívějších modelů dat.

Nejjednodušší typ vazby je „1:1“. To znamená, že pro každý záznam v jedné tabulce existuje jeden záznam v tabulce jiné. Složitější je vazba „1:n“, kde jednomu záznamu v jedné tabulce odpovídá několik záznamů v tabulce jiné. Nejúplnějším typem je vazba „n:n“. Pro kterýkoli záznam z jedné tabulky může být několik záznamů v tabulce druhé. Pokud je mezi tabulkami vazba n ku n, je nutné většinou spojit je třetí tabulkou, která má vazbu jedna ku n s ostatními dvěma.

Pro každou dvojici příbuzných tabulek musí být spojující položky stejného datového typu a většinou by měly mít stejnou délku. U většiny programových produktů není nutné, aby měly stejné jméno, ale je to vhodné z důvodu přehlednosti návrhu.

V jazyce SQL vypadá příkaz na sestavení relace a dotazu následovně:

```
SELECT <seznam sloupců>
FROM <seznam tabulek>
WHERE <podmínka spojení>
```

Operace spojení je provedena příkazem „SELECT“ tak, že v klauzuli „FROM“ uvedeme seznam jmen spojovaných tabulek a v klauzuli „WHERE“ uvedeme podmínku spojení.

Na závěr této kapitoly se pokusme vytvořit databázový systém, který by našel uplatnění při evidenci v půjčovně aut. Na tomto příkladu je vidět proces normalizace. Normalizace, jak bylo posáno již dříve, nám pomůže k odstranění opakujících se skupin a nadbytečných (redundantních) informací. Normalizací se z celkového výběru atributů vytvářejí postupně jednodušší relace.

Zároveň je nutné, aby si čtenář uvědomil některá fakta. Například, že někteří zákazníci si mohou půjčit auto vícekrát, nebo, že půjčovna může mít více stanovišť, kde jsou auta zaparkována.

po logické analýze problému
prvotní struktura dat

z kód	příjmení	adresa	a kód	model	s kód	stanoviště	cena	půjčeno	vráceno
-------	----------	--------	-------	-------	-------	------------	------	---------	---------

po normalizaci
zákazník

z kód	příjmení	adresa
-------	----------	--------

pronájem

z kód	a kód	půjčeno	vráceno
-------	-------	---------	---------

auto

a kód	model	cena	s kód
-------	-------	------	-------

stanoviště

s kód	stanoviště
-------	------------

Obr. 5.

Po logické analýze problému jsme tedy dospěli ke čtyřem datovým tabulkám. Přes šedě zvýrazněná políčka je provedena relační vazba.

3.5.1.4 Formuláře

Pokud jsme vytvořili datové struktury a jsme si jisti, že jsou správné a úplné, je možné přistoupit k vytváření vstupních formulářů. Ve většině databázových aplikacích hrají formuláře rozhodující roli při vkládání a editaci dat. Vstupní formulář navrhujeme většinou tak, že vzhled se podobá papírovým formulářům, které běžně známe. To samozřejmě urychlí vkládání dat, usnadní kontrolu a orientaci. Velkou výhodou je rovněž to, že tam, kde operátor může při vkládání dat udělat chyby možností nejednoznačného zadání, zavedeme ve formuláři objekty, pomocí kterých operátor pouze vybírá data z předem připravených grafických objektů. Mezi tyto objekty patří např. seznamy, překryvné seznamy, číselníky, výběrová logická pole, přepínače. Na následujících obrázcích 6, 7 budou ukázány některé typy formulářů.

Zaměstnanci

Margaret Peacock

Osobní údaje:

Adresa: 4110 Old Redmond Rd.

Město: Redmond Region: WA

PSČ: 98052 Země: USA

Telefon domů: (206) 555-8122

Formální oslovení: **Paní**
Dr.
Slečna
Paní

Datum narození:

Poznámky: Margaret odmaturovala z anglické literatury na Concordia College (1958), v roce 1966 absolvovala Kulinářský institut. V době od července do listopadu 1992 dočasně

Informace o firmě

Záznam: 4 z 9

Obr. 6.

Objednávky

Plátce: Alfred's Futterkiste Příjemce: Alfred's Futterkiste

Obere Str. 57 Obere Str. 57

Berlin 12209 Berlin 12209

Německo Německo

Podejice: Peacock, Margaret

Přeprava: Speedy United Federal

Číslo obj.: 10702 Datum objednávky: 10.10.1994 Dodací dne: 21.11.1994 Datum odeslání: 18.10.1994

Výrobek:	Cena:	Množství:	Sleva:	Výsledná cena:
Lakkalkööri	450,00 Kč	15	0%	6 750,00 Kč
Aniseed Syrup	250,00 Kč	6	0%	1 500,00 Kč
*				

Tisk faktury

Mezisoučet:	8 250,00 Kč
Dopravné:	598,50 Kč
Celkem:	8 848,50 Kč

Záznam: 1 z 830

Obr. 7.

3.5.1.5 Dotazy

Jedním z nástrojů pro přetváření surových dat na užitečnou informaci je schopnost vybrat zobrazovaná data, určit které informace chceme v daném okamžiku vidět a jak tyto informace vhodně uspořádat.

Téměř každý programový systém umožňuje opět několik možností, jak sestavit a provádět dotazování :

- pomocí SQL příkazů
- pomocí návrháře dotazů

Dotazy představují hlavní účel, k němuž byl jazyk SQL vybudován. Sestavují se pomocí syntakticky bohatého příkazu **SELECT**, který může mít tvar velmi jednoduchý, ale také velmi komplikovaný. Obecný popis příkazu vypadá takto:

```
SELECT <seznam jmen sloupců definujících projekci> / *  
FROM <jméno tabulky>  
[WHERE <podmínka výběru>];
```

Tímto příkazem tedy provedeme selekci a projekci nad uvedenou tabulkou. Je-li místo seznamu sloupců uvedena hvězdička, má význam výběru všech sloupců. Podmínku selekce uvádíme v klauzuli **WHERE**. Pro zápis podmínky můžeme používat :

- relační operátory (=, !=, <>, >, >=, <=, **BETWEEN**, **IN**, **LIKE**)
- logické operátory (**AND**, **OR**, **NOT**)
- množinové operátory (**INTERSECTION**, **UNION**, **MINUS**)
- konstanty a jména sloupců tabulky

Operátor „**BETWEEN**“, který testuje příslušnost do intervalu, se používá ve tvaru :

```
BETWEEN <dolní mez intervalu> AND <horní mez intervalu>
```

Pro test příslušnosti do množiny se dá použít operátor „**IN**“ :

```
IN (<seznam prvků množiny>)
```

Množinový operátor „**INTERSECTION**“ je možné samozřejmě nahradit logickým operátorem „**AND**“ a množinový operátor „**UNION**“ je možné nahradit logickým operátorem „**OR**“. Operátor „**MINUS**“, který označuje množinový rozdíl, se používá k vyhledávání řádek, které mají vlastnost „**A**“ a současně nemají vlastnost „**B**“. Je možno jej nahradit použitím logických operátorů „**AND**“ a „**NOT**“. Proto řada implementací jazyka SQL tyto operátory nemusí používat.

Jako příklad uveďme následující dotaz na tabulku občanů, kdy chceme zobrazit jména a data nástupu do zaměstnání občanů, kteří navíc bydlí na Praze 6 (PSČ – 160 00). Konstrukce dotazu vypadá následovně:

```
SELECT prijmeni, jmeno, nastup  
FROM obcane  
WHERE psc='160 00';
```

Další složitější dotaz by mohl vypadat takto. Zajímají nás pracovníci, kteří nastoupili po Silvestru 1996, mají plat vyšší než 8000,- Kč. Údaje chceme seřadit podle data nástupu vzestupně, při stejném datu nástupu podle platu sestupně.

```
SELECT prijmeni, jmeno, nastup, plat
```

FROM obcane
WHERE nastup > {31.12.96} AND plat > 8000
ORDER BY nastup ASC, plat DESC;

Tabulku výsledků je možné zobrazit na obrazovku, vytisknout na tiskárnu, uložit do další tabulky nebo do textového souboru. Potom je nutné předešlý dotaz doplnit o následující fráze :

TO PRINTER – tisk na tiskárně

TO CONSOLE – na obrazovku

INTO TABLE <nazev tabulky> – do databázového souboru, který může být přechodný, nebo trvalý.

TO FILE <nazev souboru> – do textového souboru

Nutno ještě podotknout, že součástí klauzule „SELECT“ mohou být i jiné vestavěné funkce (např. Trim() – ořezání všech mezer, Left() – ořezání všech mezer zleva a pod.). Rovněž je možná kombinace vypočítaného sloupce a konstantního sloupce. Dotaz by mohl vypadat následovně:

```
SELECT typ, cena_ks, pocet_ks, cena_ks * pocet_ks, cena_ks * pocet_ks * 0,22
FROM zboží
WHERE typ= 'vysavač';
```

Velmi důležité jsou také „agregační dotazy“. Jejich princip spočívá v tom, že se zpracovávají hodnoty z celých sloupců tabulky, jejichž výsledkem je jediná hodnota. Každou chvíli je potřeba zjišťovat sloupcové statistiky, jako jsou součty, průměry nebo výskyty (četnosti). V SQL se agregační dotazy realizují většinou s pomocí „sloupcových funkcí“.

SQL nabízí následující sloupcové funkce:

SUM(jméno sloupce) – součet hodnot ve sloupci

AVG(jméno sloupce) – aritmetický průměr hodnot ve sloupci

MIN(jméno sloupce) – minimální hodnota ve sloupci

MAX(jméno sloupce) – maximální hodnota ve sloupci

COUNT(jméno sloupce) – počet hodnot ve sloupci

3.5.1.6 Vytváření podotázek

V jazyce SQL můžeme dotazy řetězit v tom smyslu, že pro formulaci dotazu na vyšší úrovni můžeme použít výsledky dotazu na nižší, podřízené úrovni. Nižší úroveň dosáhneme vnořením příkazu „SELECT – FROM – WHERE“ do příkazu „SELECT – FROM – WHERE“. Podotázky můžeme rozdělit na podotázky dodávající právě jednu hodnotu a na podotázky dodávající n-tici hodnot. Podle tohoto kritéria můžeme podotázky dosazovat do formulací otázek vyššího řádu. Z důvodu vysvětlení uvedeme následující příklad.

Chceme nalézt všechny včelaře, kteří patří do stejné místní organizace jako včelař Novák Jaroslav. Poznáme to podle toho, že mají stejné jméno předsedy. Můžeme postupovat ve dvou krocích. Nejprve nalezneme jméno předsedy včelaře Nováka Jaroslava a potom dalším „SELECT“, kam toto jméno zapíšeme, nalezneme všechny včelaře, jejichž předseda se jmenuje stejně. Místo toho však lze zkráceně napsat:

```
SELECT jmeno, adresa
FROM vcelari
WHERE predseda=
(SELECT predseda
FROM vcelari
WHERE jmeno='Novák, Jaroslav');
```

V uvedeném příkladu jsme získali jedinou hodnotu, ale obecně můžeme získat celou množinu hodnot.

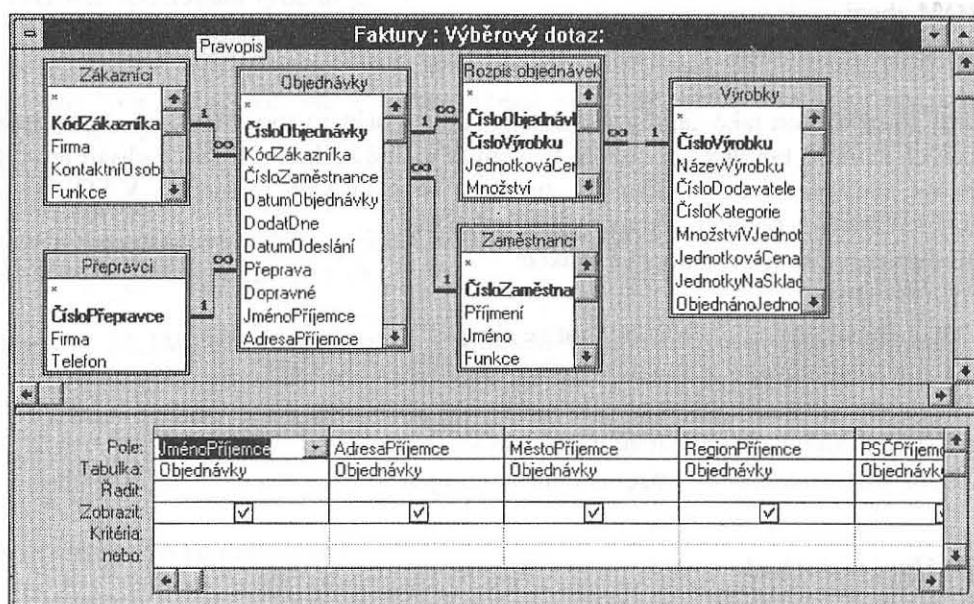
Další zadání by mohlo vypadat následovně. Naleznete včelaře z kraje Jihomoravského, který má více včelstev než kterýkoli včelař u kraje Jihočeského.

```

SELECT jmeno, adresa
FROM vcelari
WHERE medu > ALL
      (SELECT vcelstev
       FROM vcelari
       WHERE kraj='Jihočeský')
AND kraj='Jihomoravský'
ORDER BY vcelstev DESC;

```

V závěru pasáže „dotazy“ si ukážeme na obr. 8 konstrukci dotazu pomocí návrháře.



Obr. 8.

Jde o návrh dotazu pomocí návrháře databázového programu Access. Dle obr. 8 je vidět, že veškeré zadávání parametrů dotazu se děje výběrem myši nebo z klávesnice. Prvním krokem při definici nového dotazu je určení zdroje dat, kterým může být jedna nebo více existujících tabulek nebo dotazů. Ve vícetabulkových dotazech se musí dále nadefinovat vazby mezi tabulkami. Ve spodní části okna dotazu je výstupní mřížka. Do ní se zadává popis dotazu, položek a počítaných hodnot, které chceme vidět, pořadí třídění a výběrová kritéria. Každý sloupec výstupní mřížky představuje jednu datovou jednotku, položku nebo počítanou položku.

3.5.1.7 Sestavy

Sestavy slouží ke komplexnímu zobrazení dat podle různých kritérií. Data tištěná v sestavě mohou být odvozena z jedné nebo více tabulek či dotazů. Každá sestava se skládá z oddílů, které jsou definovány svou pozicí v tiskovém pořádku. Základních oddílů v sestavě bývá sedm.

záhlaví sestavy – tištěno pouze na začátku sestavy

záhlaví stránky – tištěno jednou na začátku každé strany, jde většinou o název sestavy, datum, záhlaví sloupců a pod.

záhlaví skupiny – tištěno jednou na začátku každé skupiny, obvykle zahrnuje obsah položek nebo výrazů, které definují skupinu

detailní oddíl – tištěno pro každý záznam odpovídající výběrové množiny, může obsahovat položka, vypočtené hodnoty a vysvětlující text

paty skupiny – tištěno jednou na konci každé skupiny, obvykle obsahuje součet nebo jiné statistické přehledy pro danou skupinu

paty stránky – tištěna na konci každé stránky, obsahuje většinou číslo stránky, datum a pod.

paty sestavy – je tištěna jednou na konci sestavy, může být použita k tisku shrnujícího přehledu, celkových součtů nebo statistických přehledů

Asi nejlepší metoda jak vytvořit vyhovující sestavu, je pracovat zesponu nahoru. Tedy začít detailními oddíly, které obsahují data odvozená z jednotlivých položek zdrojové tabulky nebo dotazu. Teprve až jsme spokojeni s velikostí a rozmístěním objektů v detailním oddílu, můžeme stanovit správnou pozici odpovídajících položek v dalších oddílech jako jsou nadpisy sloupců v oddílu záhlaví stránky tak, aby odpovídaly sloupcům pod nimi. Dále můžeme umístit celkové statistiky do do dílu paty skupiny a do oddílu paty sestavy pod odpovídající sloupce.

3.5.2 Zabezpečení dat

Obsahuje-li databáze citlivé informace, které by mohly být předmětem zneužití (personální, účetní data), neměl by k těmto informacím mít přístup každý, kdo s daným počítačem pracuje.

Zabezpečení dat je prováděno většinou na několika úrovních. Každý uživatel je vždy příslušníkem určité skupiny uživatelů. Jak celé skupině, tak i jednotlivým uživatelům je možné přidělit určité oprávnění k manipulaci s daty. Každá skupina a každý uživatel je definován určitou identifikační posloupností. Každý uživatelský účet může mít také heslo. V heslech i ve jménech skupin i uživatelů se většinou rozlišují velká a malá písmena. Standardně jsou vždy předdefinovány dva uživatelské účty – a to účet administrátorů a běžných uživatelů. Administrátor je zodpovědný za správu celé databáze, tedy i za přidělování uživatelských účtů a hesel.

Na závěr této pasáže je nutno podotknout, že účinnost zabezpečení dat velmi závisí na netechnických podmínkách. Všechna opatření jsou bezcenná, sdílejí-li uživatelé hesla, což bývá praxe, která může začít velmi nevinně, musí-li například zaměstnanec splnit práci někoho, kdo je nepřítomen nebo zaneprázdněn.

3.5.3 Vytvoření spojení s jinými programy

Někdy je potřeba z nejrůznějších důvodů importovat i exportovat data z námi používaného databázového programu do jiné aplikace. Tou může být například tabulkový kalkulátor, který je obzvláště vhodný pro složité matematické výpočty, nebo pro různé druhy dalších analýz. Dále to může být specializovaný statistický program nebo účetní program, či některý ze skupiny prezentačních programů.

V podstatě existují dvě základní metody, jak data využít v různých jiných aplikačních programech :

- export a import dat v nejrůznějších formátech
- datový most mezi aplikacemi

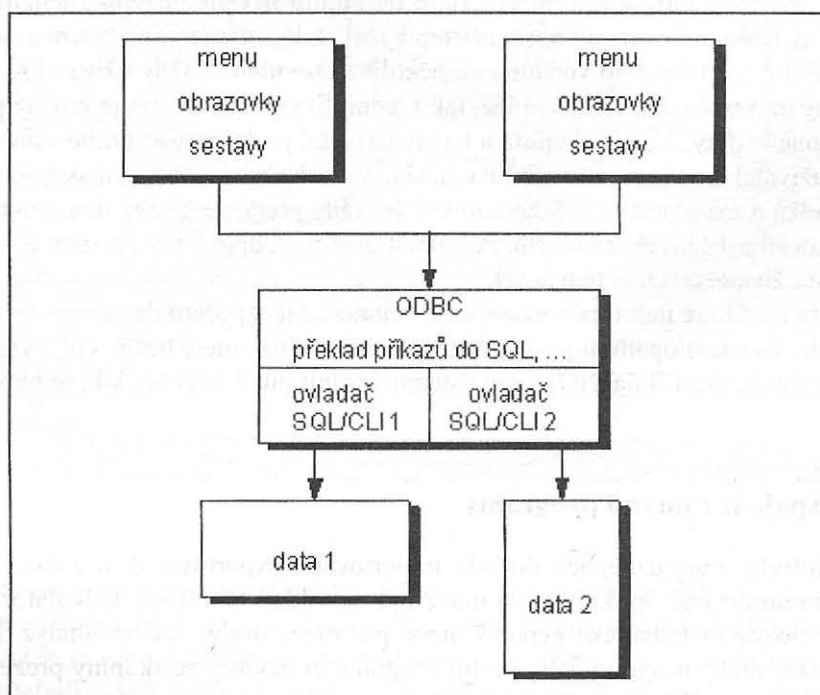
V prvním případě (import/export) se vytvoří kopie daného datového souboru. S tou se pak pracuje zcela nezávisle na originálu. Některé programové produkty umožňují zpřístupnění některých podporovaných typů souborů. V tomto případě se kopie zdrojového souboru nevytváří.

Nutno podotknout, že mechanismus pro import, export a zpřístupnění souborů je pro všechny podporované externí formáty více méně stejné. Přesto je detailní přístup, pro nevyhnutelné rozdíly

mezi těmito formáty a díky programům, které je vytvořily, mírně odlišný. Sdílení dat vyžaduje vždy jisté znalosti o práci externích programů. Pokud je nemáme, může se stát výmena dat obtížnou.

Druhým způsobem zpřístupnění „cizích datových souborů“ je použití datového mostu. K nejznámějším patří ODBC (Open Database Connectivity). ODBC je rozhraní MS-Windows pro připojení se k databázi, definuje API (Aplikation Programming Interface) pro přístup k databázím. Jeho záměrem je soustředit v sobě funkční oblasti společné pro mnoho databázových produktů. Rozhraní ODBC je fyzicky realizováno tak, že pod společnými funkčními nástroji je sada samostatných ovladačů pro přístup ke konkrétním systémům. Aplikace pracuje s databází pomocí ODBC funkcí, které jsou uloženy v dynamické knihovně (DLL). Tyto funkce jsou určeny pro klienta databáze. Mezi DLL funkcemi a vlastní databází se nachází databázový ovladač, který převádí obecné požadavky na jakoukoli databázi do tvaru pro databázi konkrétní. Podmínkou je mít ovladač dané databáze. Podstatným rysem ODBC je fakt, že s hostitelským systémem lze komunikovat výhradně pomocí funkcí, které jsou v ODBC zabudovány. Omezením ODBC je pouze to, že standardně neřeší navigační přístup k databázím. Může být realizován pouze v případě, že ho podporují ODBC-ovladače jednotlivých systémů. Zároveň není uzpůsobeno pro zpracování dotazů z více datových zdrojů zároveň.

Na obr. 9 je zobrazeno více databázových systémů přístupných pomocí jednotného rozhraní ODBC – vylepšení předchozího způsobu. Kromě jednotného přístupu k ODBC, které tak zajišťuje přístup ke všem systémům, které ODBC podporují, nabízí aplikacím již některé připravené funkce, které mohou být využity oběma systémy a nemusí být tak v obou samostatně řešeny (např. převod různých tvarů dotazů na SQL, síťový provoz hostitelského systému). Nenabízí heterogenní dotazy a standardně neřeší navigační přístup.



Obr. 9.

3.6 Závěr

Na závěr pojednání o databázích by bylo dobré říci, že snahou autora bylo seznámit čitatele s velmi rozsáhlou, ale také zajímavou oblastí softwarového inženýrství. Bohužel omezení počtu stránek, které byly vyčleněny na tuto problematiku, zabránilo možnosti věnovat se některým pasážím důkladněji. Přesto se domnívám, že i čtenář, který se nikdy nesetkal s problematikou databází, se po přečtení tohoto příspěvku náležitě zorientuje v dané problematice. Samozřejmě, že při práci s konkrétním produktem bude nutné dále rozšiřovat své znalosti.

4. Metoda základního zpracování archeologických vědeckých dat s pomocí počítačové podpory (Jiří Macháček)

(Práce vznikla s podporou grantu reg.č. 404/94/0410 Grantové agentury ČR)

4.1 Úvod

Využití počítačů při zpracování vědeckých dat je v současné době samozřejmostí. Ani archeologie není v tomto směru výjimkou. Počítačové technologie se již staly trvalou součástí metodologické výbavy archeologa. Jako každá metoda se však i použití počítačů musí řídit určitými pravidly. V následujících kapitolách si ukážeme, jak postupovat při základním zpracování archeologických dat za pomoci databázových a tabulkových programů.

Předkládám metodu komplexní, ne však tak složitou, aby ji nemohl zvládnout průměrně obeznámený uživatel běžně dostupných a rozšířených kancelářských softwarů. Metoda nebazíruje na znalosti speciálních matematických či statistických postupů, vyžaduje však částečné obeznámení s databázovými a tabulkovými programy a některými jednoduchými statistickými metodami.

Základní zpracování archeologických dat s počítačovou podporou lze rozčlenit do následujících kroků:

- 1) Definování problému
- 2) Výběr dat a tvorba struktury databáze
- 3) Testování relevantnosti dat
- 5) Plnění databáze
- 6) Databázové dotazy
- 7) Vizualizace dat
- 8) Testování výsledků
- 9) Interpretace

Pro získání důvěryhodného výsledku je nutno při práci projít všemi navrženými kroky. Jednotlivé kroky se pokusím rozvést a v maximální míře přiblížit za pomoci příkladů z praxe.

4.2 Definování problému

První bod metody hraje v celém sledu výše vyjmenovaných kroků nejdůležitější roli, neboť dává našemu snažení náplň a smysl. Měla by mu proto být věnována zvláště zvýšená pozornost. Definování otázky, na kterou se budeme snažit odpovědět, souvisí ve velké míře s filozoficko-metodologickou výbavou a postoji badatele. Není účelem této statě rozebírat teoretické problémy vědy, dovolil bych si však krátce upozornit na některé skutečnosti.

Již samo využití počítačů predikuje náš pohled na svět a vnáší novou dynamiku do vývoje teoretické základny archeologie. Hromadné používání počítačů umocňuje proces nastartovaný v 60. letech v anglosaském světě, kdy zde ve střetu o základní teoretické otázky zvítězilo pojetí Nová archeologie nad tradicionalistickým přístupem. Jedna ze změn spočívá právě ve formulaci otázky. Tázací zájmena CO či JAK jsou zde nahrazována slovem PROČ. Odpověď na otázku, tážící se po smyslu jevů, je mnohem složitější a předpokládá komplexní přístup k řešení problému, což právě výpočetní technika velice usnadňuje. Přeneseno do praktické roviny lze říci, že záměrem moderní archeologie již není ptát se např. na to, jaký artefakt se vyskytoval na určitém sídlišti v určité době, ale PROČ se tam vyskytoval.

Formulace otázky je velice důležitá pro další postup práce, neboť otázka předurčuje odpověď. Špatně položená otázka může „zhatit“ výsledek dlouhé a úmorné práce již v začátku. Lze proto jen

doporučit, aby před započítím rozsáhlejšího projektu bylo definování problému diskutováno v širším kolektivu, v případě studentských prací se zkušeným pedagogem.

4.2.1 Příklad 1

JEDNOTLIVÉ KROKY NAVRHOVANÉ METODIKY BUDOU DOKUMENTOVÁNY NA PŘÍKLADU PRÁCE J. MACHÁČKA: PODUNAJSKÝ TYP ANEB KERAMIKA STŘEDODUNAJSKÉ KULTURNÍ TRADICE, BRNO 1994.

V průběhu 6.– 8. stol. dochází ve střední Evropě k hlubokým zásahům do společenských struktur a k radikální přeměně materiální kultury. Tento proces se odráží i v archeologických nálezech. Kromě výrazných složek inventáře jakým jsou např. zbraně či šperky, můžeme změny zaregistrovat i u nenápadnější, o to však početnější skupiny nálezů – keramiky.

K zásadnímu zlomu ve vývoji keramiky dochází s nástupem charakteristického zdobení a obtáčení. Je to kombinace vlastností, která se stala pro slovanskou keramiku typickou a lze ji sledovat až do vrcholného středověku. Na území ohraničeném na západě horním Mohanem a středním Labem, na severu povodím horní Odry a horní Vislou, na východě obloukem Karpat a na jihu karpatskou kotlinou je tato keramika v průběhu 7. a 8. stol. stylově poměrně jednotná.

Jedna ze zásadních otázek důležitých pro pochopení celého mechanismu vývoje slovanské keramiky souvisí s tím, zda je zdobená a obtáčená keramika na celém výše zmíněném území opravdu homogenní či heterogenní a PROČ tomu tak bylo.

4.3 Výběr dat a tvorba struktury databáze

Data a metody, které vybereme pro zpracování našeho úkolu, úzce souvisejí s otázkou, kterou řešíme. V této stati se zaměříme na problémy, které umožňují využití databází a tabulkových procesorů. Jedná se především o hromadné jevy, čítající řádově stovky až tisíce případů.

Před započítím vlastní práce je nutno vyřešit, jaké informace může a má databáze poskytovat. Abychom mohli takové rozhodnutí učinit je nutné nejdříve vytvořit předběžný model studovaného problému (Neustupný 1986, 532), který bude vycházet z našich předchozích zkušeností a znalostí. Z něho pak odvodíme data důležitá pro řešení našeho úkolu. Jako příklad nám může posloužit např. studium sociální struktury starobronzové společnosti. Na základě našich dosavadních znalostí je zřejmé, že relevantní informace důležité pro řešení našeho problému lze získat z pohřebišť. Víme např., že hrobové jámy význačných jedinců s bohatou výbavou zde byly hlubší a širší. Na základě tohoto tvrzení zahrneme do naší databáze informace o rozměrech hrobů. Z modelu dále např. vyplývá, že velký význam z hlediska identifikace společenského postavení pohřbených jedinců sehrával počet a typ určitých milodarů např. dýček, šperků, nárameníků, honosnějších zbraní, možná i kamenný zával, některá antropologická zjištění atd. atd. Všechny tyto informace by bylo dobré v databázi zachytit. Je nutno si však uvědomit, že každý jev či předmět má obrovské množství vlastností – znaků, které lze zaznamenat. Naším zájmem ale není generování velkého objemu data, nýbrž získání relevantních informací, jejichž rozbor povede k efektivnímu řešení. Musíme srovnávat časovou náročnost, která je ovlivněna kvantitou dat, jež budeme o jednotlivých jevech či předmětech ukládat, se snahou o získání co nejúplnější informace, vedoucí ke komplexnímu pojetí problému. S ohledem na obě hlediska, tedy čas a úplnost informací (informace, které zaznamenáme, nebudou nikdy úplné!) je nutno před započítím práce zvolit kompromisní – efektivní – řešení.

Pokud jsme si dobře ujasnili údaje, se kterými budeme pracovat, je namístě navrhnout databázi, do které je budeme ukládat. Dle složitosti dat je nutno rozhodnout, zda použijeme více tabulek (dle terminologie programů MS Access či Paradox), navzájem propojených relačními spojeními, či tabulku jednu. Práci s více tabulkami lze doporučit všude tam, kde vaše data tvoří více logických celků,

např. keramika z hrobů (jedna tabulka), rozměry hrobů (druhá tabulka), antropologické údaje (třetí tabulka) (relačně propojeno číslem hrobu). Takový přístup je zvláště vhodný pro týmovou práci, při které plní jednotliví členové skupiny různé tabulky, které jsou při finálním vyhodnocení spojeny do jednoho databázového systému, ve kterém lze volně kombinovat data z různých relačně propojených tabulek.

Po stanovení potřebného počtu tabulek je nutno vytvořit jejich vnitřní strukturu. Databázová tabulka je složena z řádků a sloupců. Řádky tvoří jednotlivé záznamy (records) (Jedním záznamem může být např. jeden hrob, jedna nádoba, jedno hradiště). Sloupce (pole-fields) jsou kategoriemi informací. To jsou vlastnosti (znaky) jevů či předmětů, které jste se po předchozí úvaze rozhodli zaznamenávat (např. hloubka, šířka, tvar hrobu; barva, výzdoba, výška nádoby; plocha, datace, nadmořská výška hradiska). V polích je nutno rozložit informace na co nejmenší logické části, což později využijeme při vytváření dotazů a hledání souvztažností mezi nimi (např. pro jméno a příjmení v databázové tabulce zaměstnanců vytvořte dvě různá pole). Rozdělená data lze v databázi různě kombinovat, spojená data však lze jen těžce oddělit.

Nikdy nevkládejte do tabulek vypočítávaná data, např. délko/šířkové indexy. Počítačové programy nabízejí nástroje, které to udělají za vás a bez chyb.

Nezapomeňte umístit do tabulky údaj, který jednoznačně identifikuje záznam (tzv. identifikátor) (např. číslo hrobu, inv. číslo nálezu). V jedné tabulce se nesmí vyskytnout dva stejné identifikátory. Identifikátor může zároveň sloužit jako index (viz výše: Smutný, J.: Databázové systémy).

V databázi není možné zapsat jednu vlastnost (znak) různými slovy či symboly. Kdyby k tomu došlo, způsobilo by to velké komplikace při vyhledávání a filtrování databáze. Pro plnohodnotnou a kvalitní práci s archeologickou databází je proto nutno použít při deskripci vlastností (znaků) předmětů či jevů určitý jednotný systém, v archeologii obecně nazývaný kód. Jeho potřeba souvisí se snahou o objektivizaci a formalizaci informací, které zaznamenáváme. Kódy mohou být, v závislosti na počtu a charakteru znaků, které sledujeme, jednoduché, obsahující pouze několik termínů, či složité (např. Podborský, Kazdová, Koštuřík, Weber 1977 – Numerický kód moravské malované keramiky) (viz níže: Koštuřík P., Macháček J.: Komentář ke „Kódu moravské domácí eneolitické keramiky“ a Salač V.: K využití seriace při datování sídlištní keramiky).

Sledované znaky dělíme na kvantitativní a kvalitativní. Hodnoty znaků kvantitativních zjišťujeme měřením. Znaky kvalitativní vyjadřují kvalitu (barva, tvar, pohlaví atd.). V kódu musí být ošetřeny jak znaky kvantitativní (přesné určení rozměrů, které měříme; jednotky, ve kterých měříme atd.), tak i kvalitativní (přesná terminologie jednotlivých vlastností – výzdobné prvky, tvar nádob atd.). Dobře vytvořený kód snižuje ztrátu hodnoty dat, způsobenou subjektivním přístupem jednotlivých pracovníků.

Jestliže jsme s úspěchem prošli touto přípravnou fází zpracování archeologických dat, můžeme začít se sběrem dat, nejdříve pro testovací sérii.

4.3.1 Příklad 2

Vzhledem k velkému množství materiálu a obtížnosti celé problematiky, související s přeměnou keramiky lepené v ruce v keramiku obtáčenou a zdobenou a jejím dalším regionálním rozrůzněním, je nezbytné zvolit takovou metodu a výběr materiálu, který by umožnil efektivně dojít k určitým závěrům. Rozhodl jsem se pro jednoduchý rozbor, ve kterém budeme sledovat především tvar nádob, jeho závislost na prostoru, čase a funkci. Jako pramennou základnu jsem zvolil nejlépe a nejkompaktněji probádaná pohřebiště ze Slovenska doplněná některými lokalitami z Moravy a Rakouska. Ve své analýze se omezují na keramiku z hrobů, ze které lze (v porovnání s keramikou z objektů na sídlištních) získat kompletnější informaci o tvaru nádob.

Pokusil jsem se pracovat s maximálním množstvím nálezů, proto jsem využíval lehce dostupné informace z literatury. Nastíněný postup však měl velkou nevýhodu v tom, že ve většině prací byly vy publikovány pouze nejzákladnější metrické údaje (tj. výška, průměr ústí, průměr dna, někdy i průměr výdutě), které umožňují pouze velmi hrubou tvarovou charakteristiku.

I tak však máme k dispozici hodnoty, které klasik bádání o staroslovanské keramice J. Eisner považoval za důležité pro popis tvaru hradištních nádob. Jeho definici lze považovat za jistý druh předběžného modelu, ve které na základě svých zkušeností uvádí, jaké informace jsou pro studium keramiky významné:

„Měříme výšku nádoby a srovnáváme ji s její největší šířkou i s šířkou jejího ústí a dna. Jsou hrnce široké i štíhlé. Mluvíme o hrncích s vejčítým, baňatě vejčítým až soudkovitým břichem, o hrncích šálkovitých (okřínovitých) a mísovitých, mají-li široké ústí a nevalnou výšku, o hrncích lahvovitých, jsou-li v hrdle značně zúženy, a štíhlých hrncích pohárovitých neboli kalichovitých, mají-li úzké dno a ústí do široka rozevřené. U soudkovitého hrnce je rozdíl mezi šířkou hrdla a dna nevalný..... Malé rozdíly v tvaru nemusily vzniknout úmyslně..... Ovšem v některých končinách může určitý tvar převládnout...“ (EISNER 1966, 155)

Na základě těchto úvah byla pro každou z lokalit vytvořena speciální tabulka, obsahující následující pole:

- název lokality
- č. hrobu
- typ hrobu (kostrový, žárový)
- identifikátor nádoby
- technologie výroby (pouze v ruce, obláčená)
- výzdoba (ano, ne)
- typ nádoby (pražský typ, podunajský typ, potiský typ – uváděno pouze některými autory)
- výška nádoby
- průměr ústí nádoby
- průměr dna nádoby
- průměr maximální výdutě (u některých lokalit chybí)

4.4 Testování relevantnosti dat

V předcházející kapitole byla zmíněna snaha o maximální efektivitu při sběru dat. Jednou ze základních podmínek tohoto přístupu je výběr vhodných informací, umožňujících odpovědět na položenou otázku. Jsou však jevy či předměty a jejich znaky/vlastnosti, které jsme se na základě našeho předběžného modelu rozhodli sledovat, pro řešení studované problematiky opravdu relevantní? Odpověď na tuto otázku je velice důležitá, protože naše práce je ve fázi těsně předcházející vlastnímu plnění databáze – tj. časově, příp. i finančně nejnáročnější část celého zpracování dat. Máme poslední možnost modifikovat databázovou strukturu. Její změny v průběhu plnění databáze jsou totiž nesmírně náročné, leckdy prakticky nemožné.

Abychom co nejvíce minimalizovali riziko špatného výběru, je nejdříve nutné zpracovat tzv. pilotní projekt. V tomto projektu budeme na vzorku dat testovat, zda vlastnosti, které hodláme sledovat a zaznamenávat, mohou dát odpověď na naše otázky a zda opravdu poskytují žádané informace.

Nejlepším postupem je vybrat menší uzavřený soubor dat, který přednostně zpracujeme v dalších navržených krocích, počínaje plněním databáze, přes databázové dotazy, vizualizaci dat k pokusu o stanovení hypotézy. Je vhodné vybrat taková data, která byla již dříve zpracována pomocí jiné metody a obě zjištění srovnat. Pokud budou výsledky pilotního projektu uspokojivé, tzn. takové, které ukáží, že typ dat zahrnutých do našeho výběru může poskytnout dostatečné podklady pro interpretaci, můžeme přistoupit k vlastnímu plnění databáze.

V této fázi je vhodné testovat i kvalitu vstupních dat (hlavně z hlediska jejich statistické věrohodnosti), a to především v případě, že je přebíráme z méně věrohodných zdrojů (starší výzkumy, katalogy atd.) .

4.4.1 Příklad 3

O tom, že vlastnosti keramiky, které budeme v naší práci sledovat, mají vypovídací hodnotu, nás přesvědčí srovnání výsledku analýzy menšího souboru získaného pomocí naší metodiky s tradičním archeologickým zpracováním.

Jeden z mála těch, kteří se pokusili o hlubší rozbor keramiky ze slovansko-avarského pohřebiště, byl A. Točík (1963). Část materiálu z pohřebiště v Prši rozdělil na základě tvarových podobností do čtyř keramických skupin (1-4).

Použijeme-li naše data a především tvarové indexy d/v a u/v (viz níže v kap. Databázové dotazy), zjistíme, že Točíkova seskupení nádob (1-3) nachází svůj odraz i v našem vyjádření. Na rozdíl od prvních třech skupin však nejsme schopni identifikovat skupinu 4 (sk. 4 je z hlediska námi sledovaných znaků heterogenní), pro kterou jsou pravděpodobně charakteristické jiné znaky než ty, které můžeme naší metodou postihnout např. nesledovaný index u/v_{yd} či d/u (viz níže v kap. Databázové dotazy).

Dalším příkladem, při jehož řešení byly použity tradiční archeologické metody, a na který můžeme zároveň aplikovat naše postupy, je třídění keramiky z pohřebiště v Devínské Nové Vsi.

V klasické práci vydělil J. Eisner (1952) tři hlavní skupiny keramiky – podunajský, potiský a pražský typ. Jednotlivé skupiny se neodlišovaly pouze technologií, materiálem, výzdobou, ale i tvarem. Poslední charakteristika je obsažena i v našich datech.

Na základě zpracování těchto dat pomocí počítačových technologií (hl. index d/v a u/v) se ukázalo, že tvarové rozdíly mezi podunajským, potiským a pražským typem z Devínské Nové Vsi lze poměrně dobře vyjádřit i pomocí našich dat a ověřit statistickým testováním.

Z obou pilotních projektů je zřejmé, že výsledky intuitivní práce erudovaných odborníků s malými soubory z větší části odpovídají výsledkům, které jsme schopni získat z našich dat. Zásadní rozdíl mezi oběma postupy spočívá ve skutečnosti, že zatímco tradičními metodami lze zpracovat maximálně desítky kusů, můžeme pomocí počítačů pracovat s tisíci a posunout tak naše poznání na novou kvalitativní úroveň.

4.5 Plnění databáze

Sběr dat je časově nejnáročnější část zpracování archeologických informací pomocí databázových programů. Abychom zbytečně neplývali svými silami či svěřenými hodnotami, lze k této fázi přistoupit teprve po uspokojivém absolvování předcházejících kroků.

Než začneme s vlastním sběrem dat, je nutno vytvořit takové pracovní prostředky, které maximálně zefektivní a usnadní naši práci. Tento postup je nutný především tehdy, pracujeme-li ve větším týmu lidí, z nichž někteří jsou zaangažováni pouze pro sběr dat.

Vhodnou pomůckou jsou formuláře, a to ať již papírové či počítačové.

Papírové formuláře použijeme všude tam, kde není z rozličných důvodů možné vkládat data přímo do počítače (nedostatečný počet počítačů, nekvalifikovaný personál, špatné klimatické podmínky atd.). Jejich výhoda spočívá v tom, že je lze použít prakticky kdykoliv a kdekoliv, nevýhodný je narůst práce spojené s převodem dat z papírového formuláře do počítače.

Pokud lze při sběru vkládat data přímo do počítače, je dobré použít databázové formuláře. Ty nabízejí jednoduchý způsob zobrazení dat, a podobně jako papírový formulář určují data, která chceme shromažďovat. „Izolují“ uživatele od základní databázové tabulky, která je poměrně nepřehledná a pro vyplňování neefektivní. Formuláře tak šetří čas a brání vzniku chyb v textu. Při jejich tvorbě lze použít různé instruktivní popisky, grafické efekty aj. Formuláře nabízejí nejvhodnější rozvržení pro zadávání, změny a zobrazení záznamů v databázi. V jednom formuláři lze spojit více tabulek dohromady, a tak zefektivnit jejich vyplňování. Data můžeme vkládat pomocí seznamu, na kterém jsou uvedeny položky, které lze do příslušného pole vložit. Tato funkce zabraňuje chybnému vyplnění či použití termínů, které se nevyskytují v našem deskripčním systému, což je z hlediska kvality databáze velice důležité.

V procesu sběru dat je nutná vzájemná kontrola, sloužící k včasné identifikaci chyb různého typu a rozsahu (nebezpečné jsou zvláště chyby systémové). Mějte na paměti, že svou prací vytváříte důležité informace, jejichž hodnota je přímo závislá na jejich kvalitě.

V závěru této fáze by mělo dojít k „pročištění“ databází, tzn. vyhledání a odmazání duplicitních záznamů, odstraňování překlepů atd.

4.5.1 Příklad 4

Pro vlastní analýzu keramiky 7.-8. stol. byla sestavena databáze z více než 2000 nádob pocházejících z 11 lokalit.

4.6 Databázové dotazy

Jedním z hlavních důvodů použití databází při zpracování archeologických dat je schopnost těchto počítačových programů rychle a efektivně řadit, třídít, kombinovat, vyhledávat a sumarizovat naše informace. V databázových dotazech lze za pomoci relací spojovat i data z více databázových tabulek. Moderní databázové systémy umožňují provádět výpočty (délko/šířkové indexy, procentuální zastoupení aj.) a vytvářet nová pole, která obsahují výsledky těchto výpočtů. Při velkém počtu záznamů (řádově tisíce) je to prakticky jediný způsob jak zpracovat velký objem dat.

Pomocí databázových výběrových dotazů a relačního propojení více tabulek zjistíme během několika sekund např. nejen, který z velkomoravských hrobů na Moravě je nejhlubší (pokud ovšem máme tyto informace v databázi), ale i jaké předměty se v něm vyskytly. Jestliže jsem dostali informaci, že zemřelý z tohoto hrobu měl při sobě meč, téměř okamžitě po modifikaci naší otázky zjistíme rozměry všech hrobů na Moravě, v nichž se vyskytla tato zbraň. Hned nato se dozvíme, v kolika případech byl meč v kombinaci s keramikou, jaký měla tato keramika ornament, jak často je tento ornament na nádobě se zaobleným okrajem, ve kterých sídlištních objektech se vyskytují zaoblené okraje či zda se tyto objekty objevují v superpozici s jinými objekty atd. atd. atd.

Kromě výběrových dotazů se nabízejí i další databázové technologie. Pro hromadné zpracování dat jsou důležité především tzv. křížové dotazy (dle terminologie MS Access) neboli kontingenční tabulky (dle terminologie programu MS Excel). Pomocí těchto výkonných nástrojů získáme shrnutí značného objemu dat v kompaktní, dobře čitelné a srozumitelné formě. Takto zpracovaná data umožňují snadné sledování jejich tendence k seskupování, ev. k vytváření určitých trendů a závislostí, tzn. umožňují identifikaci projevů archeologických struktur. Jako příklad bychom mohli uvést výstup z databázové tabulky fragmentů keramiky (jeden fragment = jeden záznam) s informacemi např. o výzdobě či tvaru keramiky a objektech, z nichž pocházejí. Pomocí křížového dotazu téměř okamžitě zjistíme např. počty všech možných kombinací mezi typy okrajů a typy výzdoby v našem souboru či zastoupení jednotlivých keramických typů v objektech. V křížových dotazech programu MS Access lze kombinovat až čtyři znaky, v případě přípravy dat ve výběrovém dotazu i více. Jedná se nejspíše o jedinou technologii mimo oblast složitých statistických a matematických algoritmů, jež nabízí tak rozsáhlé možnosti kombinace a sumarizace velkoobjemových dat.

Výstupy z křížových dotazů, ev. kontingenčních tabulek jsou ideálním východiskem k vizualizaci dat pomocí grafů (viz níže) i k tvorbě deskriptivních matic, které jsou východiskem pro další statisticko-matematické zpracování (viz níže: Neustupný E.: Syntéza struktur formalizovanými metodami -vektorová syntéza, Salač V.: K využití seriace při datování sídlištní keramiky).

Možnosti vytváření databázových dotazů jsou obrovské. Závisí především na množství našich dat a struktuře databáze, která musí umožnit maximum kombinací mezi různými daty.

4.6.1 Příklad 5

V průběhu databázového zpracování keramiky z hrobů 7.-8. stol. bylo vytvořeno několik dotazů, ve kterých byla keramika tříděna podle svých vlastností jak v rámci jedné tabulky (např. nádoby lepené

v ruce, obtáčené, zdobené či nezdobené), tak v rámci více tabulek. V dotazech byly totiž spojovány záznamy splňující určitou podmínku (např. zdobená keramiky ze žárových hrobů) z několika tabulek (lokalit) do tabulky jedné. Takto nově vytvořené soubory dat umožnily srovnání keramiky z různých hledisek, tedy nejen mezi jednotlivými lokalitami, ale i regiony (skupinami lokalit), příp. mezi veškerou keramikou určitých vlastností vzhledem ke keramice jiných vlastností bez ohledu na lokalitu, z níž pocházely.

U všech nádob byl v dotazech automaticky vypočítán tvarový index (d/v): dno dělené výškou a index (u/v): ústí dělené výškou. Pomocí nich pak zjednodušíme tvar nádoby do lichoběžníku o jednotné výšce = 1, jehož rovnoběžné základny tvoří ústí a dno.

Tam, kde to bylo možné, byl vypočítán i index (u/vyd) (ústí děleno výdutí), který zjednodušeně charakterizuje utváření horní části nádoby, a index (d/u) vyjadřující poměr mezi dnem a ústím.

Databáze tak byla doplněna o následující pole:

- index d/v
- index u/v
- index u/vyd
- index d/u

4.7 Vizualizace dat

Obrovské množství výsledků, které lze získat pomocí databázových dotazů, není pouze velkým přínosem databázových programů, ale i určitým problémem. Pokud nebudeme pracovat s více-rozměrovými statistickými technologiemi jako např. clusterové či faktorové analýzy (viz níže: Neustupný E.: Syntéza struktur formalizovanými metodami – vektorová syntéza), které dokáží vyhodnotit mnoho znaků najednou, stojíme před problémem, jak se v získaných výsledcích orientovat. Prostředkem, který poskytne rychlou a názornou informaci o povaze zkoumaného jevu je graf. Graf je zobrazením vzájemných vztahů dvou nebo více proměnných veličin pomocí symbolů. Usnadňuje vyhodnocování dat a porovnání různých hodnot z databáze.

Používáme několik typů grafů. Jedním ze základních grafů je sloupcový graf, který umožňuje porovnání jednotlivých záznamů (např. počet střepů v jednotlivých objektech), dále graf výsečový, jenž zobrazuje poměr jednotlivých částí k celku (např. poměr jednotlivých typů okrajů v sídlištním objektu) a především XY (bodový) graf. Ten zobrazuje vztah nebo stupeň vztahu mezi několika hodnotami. Výsledek je vykreslen jako body dané řadou souřadnic XY. Tímto grafem je možno především vyjádřit tendenci dat ke shlukování, příp. sledovat vývoj a změnu dat v nestejných časových intervalech. XY (bodové) grafy jsou nejčastěji používány pro vědecké účely.

Vytváření grafů umožňují i některé databázové programy. Obecně je však vhodnější použít pro tvorbu grafů specializované softwary zvané tabulkové kalkulátory (např. Excel, Lotus 123). Tyto programy nabízejí kromě grafů i silné nástroje k analýze dat pomocí vzorců, umístěných v buňkách. Buňka, která je tvořena průsečíkem řádku a sloupce, má jednoznačnou adresu (např. B4 – písmenem jsou označovány sloupce a číslem řádky v tabulce), a lze tak na ni odkazovat.

V tabulkových kalkulátorech můžeme provádět rychle a poměrně jednoduše výpočty z velkého množství dat. Programy nabízejí nejen matematické, ale i základní statistické funkce, např. tvorba histogramů, výpočet průměrů, mediánů, směrodatných odchylek, regresí atd (viz níže: Weber Z.: Statistická analýza archeologických dat).

4.7.1 Příklad 6

Indexy d/v a u/v byly vypočítány u všech nádob. Použijeme-li tyto indexy do XY (bodového) grafu, kde na vodorovné ose leží index d/v a na svislé ose index u/v , zjednodušíme tvar nádoby do lichoběžníku o jednotné výšce = 1, jehož rovnoběžné základny tvoří ústí a dno, a můžeme tak zidealizované tvary nádob porovnávat mezi sebou bez ohledu na jejich absolutní rozměry (graf 1).

Použitím indexu u/v a jeho kombinací s indexem d/v v XY (bodovém) grafu získáme další zajímavý pohled na celý soubor keramiky, v němž jsou vhodně kombinovány tři hodnoty – ústí, maximální výduť a dno.

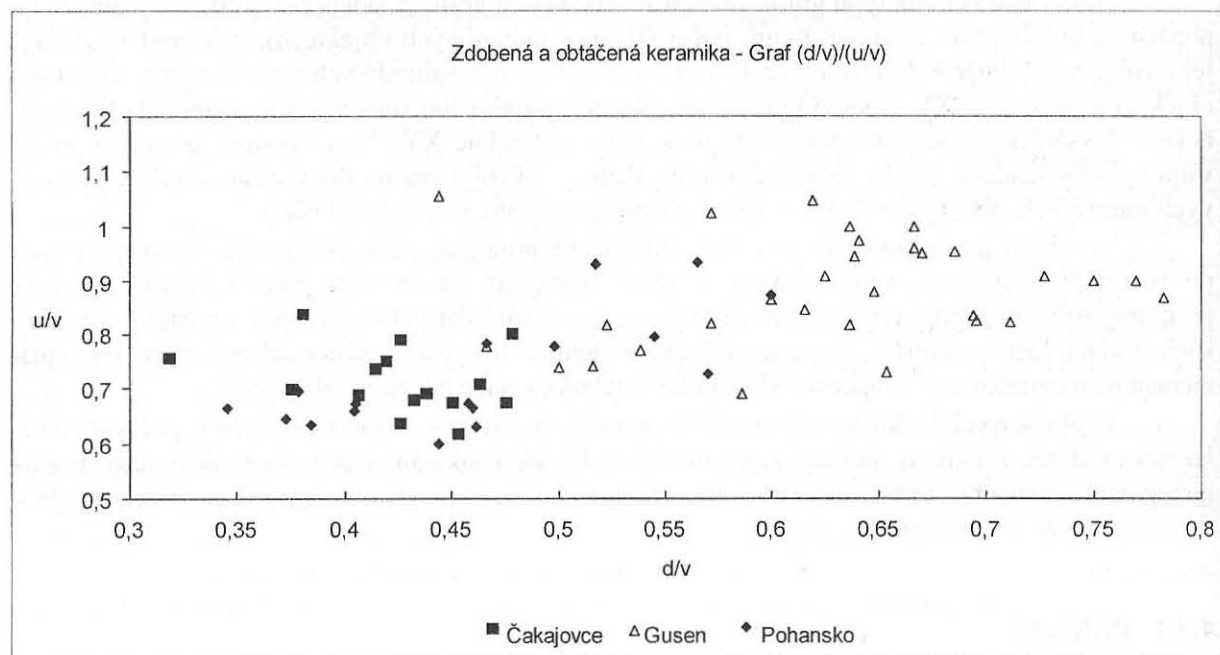
Vytvořil jsem sérii grafů, kde můžeme dobře sledovat poměry mezi tvarovým spektrem nádob z různých pohřebišť, z nichž některé zde budu prezentovat.

Pro větší názornost a přesnost jsem použil i grafického vyjádření rozdělení relativní četnosti dat z jednotlivých pohřebišť. Podle rozdělení četnosti lze totiž posoudit, jak se základní data z hlediska své velikosti rozpadají na jednotlivé četnosti, tedy počty případů připadající příslušným intervalům o stejné velikosti, které jsou definovány zvláštním statistickým postupem, vycházejícím z variačního rozpětí souboru. Rozdělení četnosti bylo automaticky vypočítáno v tabulkovém kalkulátoru MS Excel. Na vodorovné ose jsou vždy vyneseny hodnoty tvarového indexu (při této metodě můžeme sledovat vždy jen jeden index!), rozdělené do stejně velkých intervalů, a na ose svislé hodnoty relativní četnosti, které těmto kategoriím odpovídají. Relativní četnost se vypočítá jako poměr jednotlivých absolutních četností k rozsahu souboru. Křivka pak spojuje takto vynesené body. Vzniklý spojnicový graf nazýváme polygonem (graf 2).

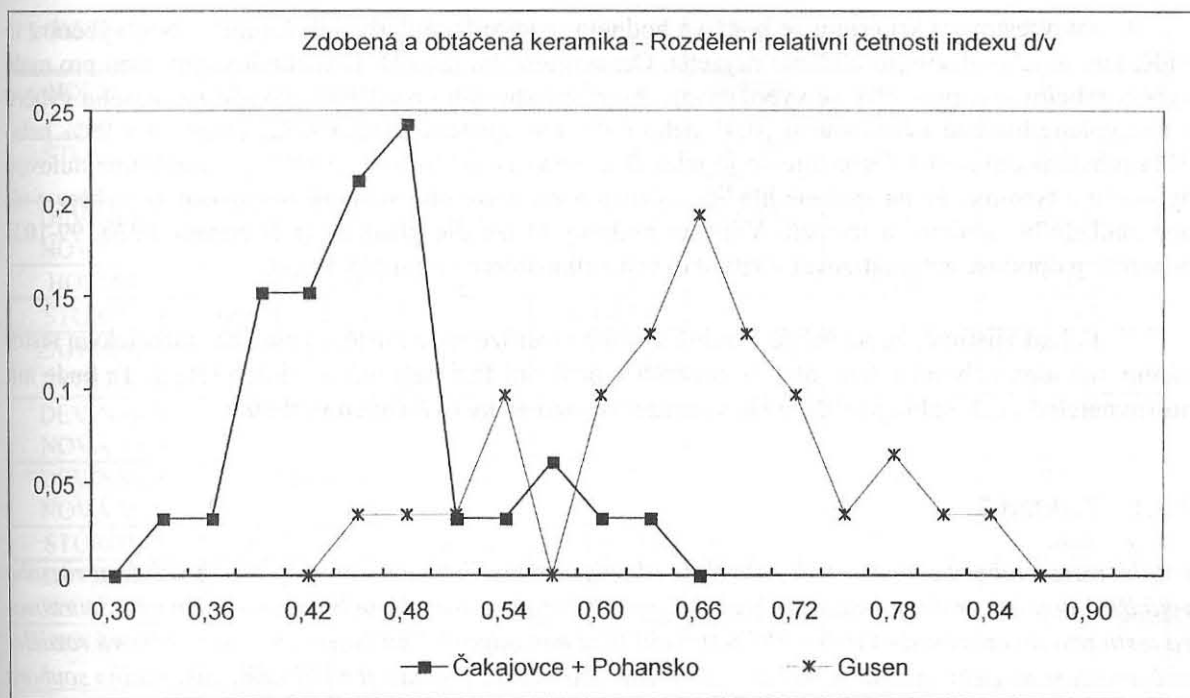
Podobně jako při předchozí analýze jsem vytvořil sérii grafů, ve kterých lze srovnávat polygony jednotlivých lokalit a vytvořit si tak představu o jejich vzájemném vztahu.

V tečkových grafech i polygonech relativní četnosti se zřetelně rýsují tři okruhy keramických tvarů. V oblasti nízkých hodnot se koncentruje keramika ze žárových pohřebišť, ve střední oblasti keramika z kostrových hrobů jihozápadního Slovenska a vysoké indexy se objevují u nádob z hornorakouského Gusenu.

Rozdílily zaznamenáváme i mezi pohřebišti z jihozápadního Slovenska, kde se vydělují dvě skupiny: Devínska Nová Ves s Záhorskou Bystricí a Štúrovo s Holiarema a Novými Zámky.



Graf 1. Zdobená a obtáčená keramika – graf (D/V)/(U/v).



Graf 2. Zdobená a obtáčená keramika – Rozdělení relativní četnosti indexu d/v.

4.8 Testování výsledků

Lze předpokládat, že pokud projdeme všemi navrženými kroky, dospějeme k určitým zjištěním, které naleznou svůj odraz jednak ve výsledku databázových dotazů, jednak v grafech. Naše data se budou různě seskupovat nebo budou vyjadřovat určité tendence a závislosti. Zjištěná strukturovanost zkoumaných dat je projevem zákonitosti v archeologických pramenech (Neustupný 1986, 542). Tyto struktury je nutno interpretovat.

Než však přejdeme k vlastní interpretaci, musíme ověřit spolehlivost našich zjištění. Jestliže to data svým typem a rozsahem umožňují, lze k dosažení tohoto cíle ve velké míře použít metod matematické statistiky.

Statistika nám dává k dispozici různé postupy, kterými vyjadřujeme statistickou věrohodnost našich výsledků. K elementárním z nich patří např. zjišťování významnosti shody či rozdílu mezi dvěma výběry, tj. vzorky dat, které zkoumáme (např. tvar nádob ze dvou lokalit), či statistické závislosti mezi dvěma proměnnými tzv. korelaci (např. mezi časem a tvarem nádob) (viz níže: Weber Z.: Statistická analýza archeologických dat).

Většina metod je popsána ve statistických příručkách. Přiblížíme si proto jako příklad pouze jedinou z nich. Jedná se o Kolmogorovův-Smirnovův test pro dva nezávislé výběry, který je v archeologii obecně používán (Smetánka 1971). Je to metoda, kterou lze jednoduše a přitom poměrně spolehlivě zhodnotit statistickou významnost rozdílu mezi dvěma soubory nálezů.

Patří do velké skupiny testů shody. Pomocí zvoleného statistického postupu testujeme, zda dva výběry pocházejí z jediného základního souboru. Jedná se o neparametrický test, který neklade velké nároky na předem dané podmínky, které jsou u jiných metod vyžadovány. Je založen na hodnocení rozdílu kumulativních četností dvou výběrů.

Můžeme ho využít ve dvou případech,

- a) když jsou výběry menší nebo rovny 40, v tomto případě musí mít stejné rozsahy
- b) když jsou výběry větší než 40, pak mohou být různě velké.

V prvním případě používáme absolutní kumulativní četnost, v druhém relativní kumulativní četnost.

Jako testovací kritérium se používá hodnota rozdílu kumulativních četností obou výběrů v té třídě, kde je jeho absolutní hodnota největší. Označujeme ho jako D . Kritické hodnoty jsou pro malé výběry tabelovány, pro velké se vypočítávají. Mají své speciální rozdělení, závislé na rozsahu výběrů a na zvolené hladině významnosti (0,01 nebo 0,05, tzn. zjištěný výsledek lze přijmout s 99% nebo 95% pravděpodobností). Označujeme je jako $D_{0,01}$ nebo $D_{0,05}$. Jestliže je $D > D_{0,05}$, zamítáme nulovou hypotézu a tvrdíme, že na zvolené hladině významnosti nelze oba soubory považovat za výběry z téhož základního souboru a naopak. Výpočet hodnoty D lze dle příruček (Reisenauer 1970, 99-103) poměrně jednoduše automatizovat v tabulkových kalkulátorech typu MS Excel.

Pokud zjistíme, že na určité hladině významnosti lze naše zjištění podepřít statistickým testováním, nic nám nebrání v tom, abychom přešli k poslední fázi naší práce – interpretaci. Ta bude mít nesrovnatelně vyšší váhu, jestliže bude vycházet ze statisticky ověřených výsledků.

4.8.1 Příklad 7

Vztahy mezi soubory z kostrových pohřebišť jihozápadního Slovenska lze díky jejich velkému rozsahu vyjádřit i pomocí exaktních statistických postupů. Využijeme osvědčeného Kolmogorovova-Smirnova testu pro dva nezávislé výběry. Při testování hledáme odpověď na otázku, zda se výběrová rozdělení četností statisticky významně neliší, neboli zda oba výběry pocházejí z jediného základního souboru (Reisenauer 1970, 99).

Postupně budeme testovat jednotlivé indexy a vždy dvě lokality mezi sebou. Výsledky statistických testů umístíme do přehledné tabulky (tab 1).

Statistická pozorování potvrzují na vysoké hladině významnosti naše závěry o tvarových skupinách keramiky na jihozápadním Slovensku, ke kterým jsme dospěly na základě předchozího databázového zpracování a vizualizace dat v grafech. Naše závěry se tak z hlediska statistiky jeví jako spolehlivé.

LOKALITA 1	LOKALITA 2	INDEX	D	$D_{0,05}$		VÝSLEDEK
DEVÍNSKA NOVÁ VES	HOLIARE	d/v	0,055814	0,1808	$D < D_{0,05}$	Shoda
DEVÍNSKA NOVÁ VES	ŠTÚROVO	d/v	0,084453	0,17695	$D < D_{0,05}$	Shoda
HOLIARE	ŠTÚROVO	d/v	0,059656	0,203988	$D < D_{0,05}$	Shoda
DEVÍNSKA NOVÁ VES	ZÁHORSKÁ BYSTRICA	d/v	0,194296	0,19598	$D < D_{0,05}$	Shoda
HOLIARE	ZÁHORSKÁ BYSTRICA	d/v	0,216142	0,2207	$D < D_{0,05}$	Shoda
ŠTÚROVO	ZÁHORSKÁ BYSTRICA	d/v	0,228261	0,2175	$D > D_{0,05}$	Rozdíl
DEVÍNSKA NOVÁ VES	NOVÉ ZÁMKY	d/v	0,0693152	0,1359833	$D < D_{0,05}$	Shoda
HOLIARE	NOVÉ ZÁMKY	d/v	0,040011	0,1696746	$D < D_{0,05}$	Shoda
ŠTÚROVO	NOVÉ ZÁMKY	d/v	0,0436494	0,1654861	$D < D_{0,05}$	Shoda
ZÁHORSKÁ BYSTRICA	NOVÉ ZÁMKY	d/v	0,2027559	0,18569	$D > D_{0,05}$	Rozdíl
DEVÍNSKA NOVÁ VES	ŠTÚROVO	u/v	0,207839	0,1807095	$D > D_{0,05}$	Rozdíl
DEVÍNSKA NOVÁ VES	HOLIARE	u/v	0,111163	0,18395	$D < D_{0,05}$	Shoda
ŠTÚROVO	HOLIARE	u/v	0,145668	0,204529	$D < D_{0,05}$	Shoda
DEVÍNSKA NOVÁ VES	ZÁHORSKÁ BYSTRICA	u/v	0,177853	0,208997	$D < D_{0,05}$	Shoda

LOKALITA 1	LOKALITA 2	INDEX	D	D _{0,05}		VÝSLEDEK
HOLIARE	ZÁHORSKÁ BYSTRICA	u/v	0,244383	0,2299	D>D _{0,05}	Rozdíl
ŠTÚROVO	ZÁHORSKÁ BYSTRICA	u/v	0,329484	0,2273	D>D _{0,05}	Rozdíl
DEVÍNSKA NOVÁ VES	NOVÉ ZÁMKY	u/v	0,16609	0,13984	D>D _{0,05}	Rozdíl
HOLIARE	NOVÉ ZÁMKY	u/v	0,11819	0,1695	D<D _{0,05}	Shoda
ŠTÚROVO	NOVÉ ZÁMKY	u/v	0,04876	0,16598	D<D _{0,05}	Shoda
ZÁHORSKÁ BYSTRICA	NOVÉ ZÁMKY	u/v	0,28072	0,19304	D>D _{0,05}	Rozdíl
DEVÍNSKA NOVÁ VES	HOLIARE	u/vyd	0,235732	0,18471	D>D _{0,05}	Rozdíl
DEVÍNSKA NOVÁ VES	ŠTÚROVO	u/vyd	0,355407	0,20256	D>D _{0,05}	Rozdíl
ŠTÚROVO	HOLIARE	u/vyd	0,200094	0,21609	D<D _{0,05}	Shoda
DEVÍNSKA NOVÁ VES	ZÁHORSKÁ BYSTRICA	u/vyd	0,21565	0,257383 (D _{0,01})	D<D _{0,01}	Shoda
HOLIARE	ZÁHORSKÁ BYSTRICA	u/vyd	0,353912	0,227547	D>D _{0,05}	Rozdíl
ŠTÚROVO	ZÁHORSKÁ BYSTRICA	u/vyd	0,483258	0,2422714 43	D>D _{0,05}	Rozdíl
DEVÍNSKA NOVÁ VES	ŠTÚROVO	d/u	0,214286	0,20162	D>D _{0,05}	Rozdíl
DEVÍNSKA NOVÁ VES	HOLIARE	d/u	0,09232	0,18470	D<D _{0,05}	Shoda
ŠTÚROVO	HOLIARE	d/u	0,17235	0,2152	D<D _{0,05}	Shoda
DEVÍNSKA NOVÁ VES	ZÁHORSKÁ BYSTRICA	d/u	0,117862	0,21349	D<D _{0,05}	Shoda
HOLIARE	ZÁHORSKÁ BYSTRICA	d/u	0,079825	0,22636	D<D _{0,05}	Shoda
ŠTÚROVO	ZÁHORSKÁ BYSTRICA	d/u	0,117676	0,24036	D<D _{0,05}	Shoda

Tab. 1.

4.9 Interpretace

Cílem našeho snažení je zjištění struktury archeologických pramenů a jejím prostřednictvím snad i zákonitostí v chování lidí dávných věků (Neustupný 1986, 537).

Postup, který jsem zde navrhl, je velmi jednoduchým příkladem toho, že i pomocí základního softwarového vybavení lze některé archeologické struktury identifikovat. Můžeme zjistit, zda se naše archeologická data seskupují podle určitých principů, korelují spolu nebo se naopak chovají zcela nahodile. Je jisté, že pomocí speciálních matematických algoritmů, kterými jsou např. faktorové či clusterové analýzy, lze získat mnohem komplexnější a podrobnější pohled na data, je však také jisté, že navržený postup umožňuje identifikovat struktury, které by nebylo možno tradičním způsobem vůbec postihnout (např. z důvodu příliš velkého množství dat).

Zjištěním zákonitostí v archeologických datech však naše práce vůbec nekončí, naopak spíše začíná. Abstraktní struktury musíme spojit s projevem živé kultury dávných lidí. To je vesměs možné pouze s pomocí archeologických modelů. Modely jsou pro pravěká společenstva odvozována především z etnoarcheologických pozorování, pro historická období i z písemných zpráv. Zjistíme-li shodu mezi našimi archeologickými strukturami a strukturami modelu, lze usuzovat na jejich shodný význam či funkci (Neustupný 1986, 544-548), a tak je interpretovat.

4.9.1 Příklad 8

Pomocí databázových, grafických a statistických nástrojů jsem se pokusil o zjištění struktury zdobené a obtáčené keramiky 7.-8. stol. především z hlediska její regionální podmíněnosti.

Z rozboru jednoznačně vyplynulo, že zkoumaný materiál lze podle tvarů rozdělit do tří velkých skupin:

a) ze slovanských žárových pohřebišť

b) z kostrového pohřebiště v Horním Rakousku

c) z jihoslovenských kostrových pohřebišť doby avarského kaganátu

Dále bylo exaktně prokázáno, že se v rámci jihoslovenských kostrových pohřebišť liší nádoby z lokalit v nejbližším okolí Dunaje (Holiare, Štúrovo, Nové Zámky) od keramiky z prostoru Bratislavské brány (Devínska Nová Ves, Záhorská Bystrica)

Na základě našich znalostí historických i archeologických jsem vytvořil model etnokulturního obrazu střední Evropy v 6.-8. stol., který jsem následně srovnal se strukturami zjištěnými zpracováním keramiky.

Dospěl jsem k názoru, že rozhodující pro vznik tvarové škály zdobené a obtáčené keramiky 7.-8. stol. byl místní substrát, na který zdobená a obtáčená keramika navazovala, příp. vlivy, které na ni bezprostředně působily. V případě žárových pohřebišť se jednalo o starší horizont keramiky pražského typu, na jižním Slovensku o substrát tvořený v různých poměrech z potiského a pražského typu, a v Horním Rakousku o vliv merovejské/raně německé keramiky či zboží produkovaného pozdně antickým obyvatelstvem. Zjištěné skutečnosti jsou důležité i pro diskusi o etnicitě lidu zdobícího své nádoby vlnovkou. Pokud totiž určujeme u substrátní keramiky pražského a potiského typu původ jejich výrobců (pražský typ – slovanský (Borkovský 1940), potiský typ – avarský (např. Bóna 1973)), není důvodu, proč bychom to stejně nemohli činit i u tvarů keramiky zdobené a obtáčené, odvozených z těchto typů.

Pro diskusi o etnicitě jsou důležité i výsledky, které se podařilo získat analýzou keramiky z kostrových pohřebišť jižního Slovenska. Nádoby z lokalit v nejbližším okolí Dunaje, tedy ze sféry bezprostředního avarského vlivu, se liší od keramiky z nejsevernější periferie kaganátu, kde jsou i podle různých příznaků v pohřebním ritu předpokládána určitá specifika. Tvary z Devínské Nové Vsi a Záhorské Bystrice, které tuto oblast reprezentují, se podle našeho pozorování více přibližují keramice z čistě slovanských žárových pohřebišť (v některých parametrech jsou prakticky identické).

I když jsem takto nemohl zpracovat zdobenou a obtáčenou keramiky 7.-8. stol. z celého širokého území, na kterém se vyskytovala, je podle zpracovaného vzorku jasné, že přes blízkou příbuznost nebyla všude homogenní, ale již od počátku diferencovaná. Jak vyplývá i z dalších zjištění, které dokládám na jiných místech, byla na různorodou keramiku, která se do střední Evropy dostala spolu se skupinami nového obyvatelstva, aplikována pozdně římská hrnčířská tradice. Vznikl horizont, který je i přes základní jednotu poměrně heterogenní. Vystává tak před námi pozoruhodný obraz regionální diferenciacie, utvářené kontaktem více raně středověkých etnokulturních komplexů a odraz tohoto procesu v archeologických strukturách.

4.10 Literatura

Bóna, I. 1973: VII. száradí Avar telep_ülések és Árpád – kori magyar falu Donaujvárosban, Fontes Archaeologici Hungariae, Budapest.

Borkovský, I. 1940: Staroslovanská keramika ve střední Evropě. Praha.

Eisner, J. 1952: Devínska Nová Ves. Bratislava.

Macháček, J. 1994: Podunajský typ aneb keramiky středodunajské kulturní tradice. Brno. (rpk dipl. práce)

Neustupný, E. 1986: Nástin archeologické metody, Archeologické rozhledy 38, 525-549.

Podborský, V. – Kazdová, E. – Košťurík, P. – Weber, Z. 1977: Numerický kód moravské malované keramiky. Problémy deskripce v archeologii, Brno.

Reisenauer, R. 1970: Metody matematické statistiky a jejich aplikace. Praha.

Smetánska, Z. 1971: Kolmogorov-Smirnovův test pro dva výběry. In: *Nové archeologické metody I*, Praha, 140-146.

Točík, A. 1963: Pohřebisko a sídlisko z doby avarské říše v Prši, SIA XI, 121-198.

5. Komentář ke „Kódu moravské domácí eneolitické keramiky“

(Pavel Koštuřík, Jiří Macháček)

5.1 Úvod

Rozsáhlé systematické archeologické výzkumy a rovněž předstihové výzkumy vyvolávají potřebu efektivního zpracovávání nebývalého množství archeologického materiálu, které si dnes bez využití počítačové techniky lze jen těžko představit. Na Katedře archeologie byl učiněn první pokus o hromadné vyhodnocení archeologických nálezů z neolitického sídliště v Těšeticích-Kyjovicích, okr. Znojmo již v sedmdesátých letech, kdy na základě analýzy malované keramiky vzniklo kolektivní dílo zabývající se touto problematikou (Podborský, Kazdová, Koštuřík, Weber 1977). Bohužel tehdejší stav výpočetní techniky nebyl u nás na náležitě výši a úrovni, a proto teprve dnes dochází k opětovnému sestavování kódů pro pravěkou keramiku. Při vyhodnocování archeologických nálezů z hradiska u Kramolína, okr. Třebíč opět vyvstal problém jak se vyrovnat se zpracováním velkého množství keramiky. Jako nejvhodnější řešení se jeví vyhodnocení získané kolekce prostřednictvím databázových systémů.

Kolekci eneolitické keramiky, která představuje přes 1 800 kusů (kromě nálezů keramiky typu Retz-Křepice, keramiky jordanovské a zvoncovitých pohárů), zahrnující v to keramiku nálevkovitých pohárů, bolerázskou keramickou produkci a nevýraznou složku jevišovické keramiky, není únosné podrobně vyhodnocovat pouze verbálním způsobem. Pomocí kódu a vhodného programu dosáhneme mnohem preciznějších výsledků – po stránce typologické a výzdobné – a též se nám podaří zachytit určité vazby a vztahy k dalším archeologickým kategoriím.

5.2 Tvorba archeologického kódu pro pravěkou keramiku (obecné zásady a praktická realizace)

Při vytváření archeologického kódu pro pravěkou keramiku je nutno vycházet z následujících obecných zásad:

- 1) Zevrubné seznámení s archeologickým materiálem v rámci studovaného období, příp. kultury.
- 2) Vydělení zájmové oblasti z hlediska regionálního i časového. S ohledem na variabilitu archeologických pramenů je nutno zvolit takové území příp. období, ve kterém je studovaný materiál relativně homogenní.
- 3) Na základě řešené problematiky a předběžného studia materiálu je nutné vydělení kategorií, které budou strukturovat deskripční systém. Mezi základní patří:
 - část obecná, do níž je zahrnuta např. lokalizace nálezů, muzeologické údaje, údaje z archeologického terénního výzkumu aj.
 - část typologická (tvarová analýza archeologického materiálu, např. tvary okrajů)
 - část metrická (např. výška nádoby)
 - část fyzikálně-technologická (údaje o keramické hmotě)
 - část výzdobná (umístění a druh výzdoby)
- 4) Na základě předcházejícího rozdělení vytvořit strukturu databáze, stanovit její jednotlivá pole (sloupce). Je dobré databázi budovat tak, aby umožnila doplňování nově zjištěných znaků na hodnocené keramice. Pokud již existuje deskripční systém (kód) popisující podobnou problematiku, je vhodné vyjít z jeho struktury, a tak si zabezpečit možnost případného srovnání dat.
- 5) Na základě předběžného studia materiálu stanovit a definovat kvality odpovídající jednotlivým znakům (např. odstíny barev, tvary okrajů, vzory výzdoby) a tyto kvality příp.

zakódovat. Při tvorbě kódu je nutno postupovat od jevů obecnějších k detailním (např. lišta → lišta vodorovná s rohlíčkovitými jamkami). Definice znaků a jejich kvalit musí být před začátkem plnění databáze konečná a není jí již možno v průběhu práce měnit (lze ji pouze doplňovat nově zjištěnými, doposud nedefinovanými znaky a kvalitami). Při publikování výsledků vědecké práce je nutno kompletní deskripční systém zveřejnit.

- 6) Aplikovat výsledky přípravné práce v konkrétním počítačovém programu.

Na výše uvedených zásadách byla založena i příprava kódu moravské eneolitické keramiky a databáze, aplikované pro vyhodnocení eneolitické domácí keramiky z hradiska u Kramolína. Na archeologické části se kromě P. Koštuříka podílel i M. Salaš z Moravského zemského muzea v Brně a část programátorskou realizoval Z. Lenhart z téže instituce. Pro vytvoření kódu byl využit především analogický archeologický materiál z jihozápadomoravských eneolitických výšinných sídlišť, především jeho katalogové zpracování, provedené A. Medunovou-Benešovou (1972; 1973; 1977; 1981; 1986). Na základě podrobné typologické, technologické a výzdobné analýzy eneolitické keramiky (kultura nálevkovitých pohárů, kultura s kanelovanou keramikou a jevišovická kultura) byla vytvořena systematika rozboru eneolitické keramiky a zápis se znaky jednotlivých keramických fragmentů či celých tvarů do navržené databáze. Při sestavování jednotlivých druhů výzdoby či dalších důležitých projevů této keramiky byl vzat za základ prostor jihozápadní Moravy, kde v rámci eneolitu předpokládáme určitý specifický vývoj. Do jaké míry bude možné uplatnit využití takto sestaveného kódu pro celou Moravu, prokáže až další zpracování větších kolekcí z rozmanitých částí Moravy. V současné době provádí testování tohoto kódu P. Procházková, která v rámci diplomové práce analyzuje eneolitické památky hmotné kultury z Olomouce a okolí. Při vytváření kódu pro moravskou domácí eneolitickou keramikou bylo počítáno rovněž s možností vytvořený systém vhodně doplňovat a rozšiřovat, což umožní běžné korektury stávajícího systému.

Vytvoření kódu, jak jsme se již zmínili, bylo inspirováno předcházejícím kódem, pomocí něhož se vyhodnocovaly kolekce kultury s moravskou malovanou keramikou (Podborský, Kazdová, Koštuřík, Weber 1977). Od vzniku tohoto kódu uplynulo téměř dvacet let a výpočetní technika doznala výrazných změn. Proto se také pojetí kódu pro zpracovávání domácí eneolitické keramiky poněkud liší od kódu pro vyhodnocování moravské malované keramiky. Především se upustilo od zakódování některých údajů v obecné části, neboť současná úroveň počítačů umožňuje jiné zapsání obecných informací o nálezech, než bylo běžné v 70. letech. Také není zakódován sektor (SK) a čtverec lokality (CTV), neboť v dnešních databázových systémech, umožňujících vkládat a vyhodnocovat různé kombinace písmen a číslic, lze zapisovat přímo původní označení sektorů a čtverců a ne pouze jejich kódů. Podobná úprava byla provedena v přímém zapisování objektů (COBJ). Rovněž zachycení hloubky jednotlivých keramických fragmentů či celých nádob je možné zapisovat přímo (VRS), nejen pouze kódy. K dalším změnám došlo v části metrické, kde na rozdíl od předchozího kódu lze tloušťku stěny nádoby (TS) uvádět přímo v daném rozměru.

Výrazných změn v kódu doznala část výzdobná. Na rozdíl od moravské malované keramiky je domácí eneolitická keramika nepoměrně chudší a malování se nevyskytuje vůbec. Rozšiřuje se především škála vhloubené a plastické výzdoby, která dosahuje velké rozmanitosti a některé projevy této výzdoby jsou pro určité úseky vývoje domácí eneolitické keramiky zvláště charakteristické.

U vhloubené výzdoby lze rozlišit několik technik:

1. Vpichy (01-16). Jsou to spíše menší zásahy do povrchu keramiky v různých částech nádoby. Objevují se buď samostatně, nebo v kombinaci s jinými technikami výzdoby. Mají rozmanité tvary a uspořádání na povrchu nádob. Nejčastěji se vyskytují na vnějším povrchu keramických tříd (obr. 6).
2. Vrypy (21-29). Jsou zpravidla větší a protáhlejší než vpichy a nedosahují takové rozmanitosti jako vpichy. Přítomnost na jednotlivých částech keramických tvarů není jednotná. Objevují se nejčastěji v horizontálním uspořádání, ale další struktura není výjimkou. Rovněž jako vpichy mohou být přítomny s další formou výzdoby (obr. 7).
3. Jamky (31-48). Na rozdíl od vpichů a vrypů jsou větší a tvarově dosahují též početnější rozmanitosti. Převažuje u nich opět horizontální orientace. Na některých fragmentech vystupují v různém uspořádání a někdy též pokrývají téměř celé tělo nádoby. Stejně jako

předchozí tvary vhloubené výzdoby se mohou na nádobách objevovat společně s dalšími technikami výzdobného projevu (obr. 8).

4. Rýhy (51-69). Představují výraznou složku vhloubené výzdoby. Příznačná je pro ně tvarová rozmanitost a uspořádání na povrchu keramických tříd. Jednak vystupují na povrchu nádob samostatně, ale v mnoha případech v kombinaci nejenom s dalšími druhy vhloubené výzdoby, nýbrž i s výzdobou plastickou, kterou výrazně doplňují (obr. 9).
5. Žlábký (71-93). Podobně jako rýhy dosahují velké tvarové rozmanitosti a svérázného uspořádání na nádobách. Kromě žlábků horizontálních, svislých a šikmé orientace se objevují další tvary, které vytvářejí různé obrazce. Mnohé z nich jsou charakteristické pouze pro určité období, což umožňuje blíže datovat keramiku s příslušnou výzdobou do určitého eneolitického období. Vrcholu dosáhla výzdoba pomocí žlábků v časovém horizontu C1 jevišovické stratigrafie (obr. 10).
6. Brázděný vpich (101-115). Tato zvláštní technika vhloubené výzdoby navazuje na chronologicky starší výzdobnou techniku, která se objevila především ve starším eneolitu a na jihozápadní Moravě má své zastoupení v mnoha kolekcích. Mladší brázděný vpich, který lze sledovat v rámci analýzy domácí eneolitické keramiky, má zastoupení až v souvislosti s jevišovickou kulturou a je to do jisté míry určitá retardace předchozího staroeneolitického vývoje. V rámci jevišovické kultury se objevuje na keramice v rozmanitém provedení a uspořádání. Často vytváří linie horizontální, vertikální či šikmé orientace, ale rovněž složité geometrické obrazce (obr. 11).
7. Prstování (120-121) není hojnou výzdobou eneolitické keramiky ve srovnání s předchozími technikami. Je to zvláštní úprava povrchu keramiky, která byla prováděna ještě před vypálením (obr. 12).
8. „Bernburgská výzdoba“ (130). Objevuje se zcela ojediněle a je důkazem kontaktů jihozápadomoravského prostředí se západními oblastmi Evropy. Představuje typickou výzdobu pro oblast Německa, odkud se dostala do moravského prostředí. Její přítomnost souvisí nepochybně až s jevišovickou kulturou (obr. 12).
9. Slámování (140) je zvláštní úprava povrchu keramiky před vypálením. V moravském prostředí vystupuje nejvíce v jevišovické kultuře, ale první objevení této úpravy povrchu keramiky může souviset již i s KK. Slámování pokrývá často větší část těla keramických tříd a vystupuje běžně v kombinaci s jinými tvary výzdoby (obr. 12).
10. Voštinování (141). Někdy bývá zaměňováno se slámováním. Je to zvláštní technika úpravy povrchu keramiky, kde na rozdíl od slámování se na povrch keramického tvaru nanáší zvláštní vrstvička hlíny v podobě jakéhosi „nastříknutí“, což způsobuje zhrubnutí povrchu (podobá se určitému druhu omítky). Obliba této techniky úpravy povrchu nesouvisí pouze s jevišovickou kulturou, nýbrž její počátky sahají do časného eneolitu. V evropském prostředí ji dokládají nálezy již z jordanovské kultury na území Wroclavska. Není vyloučeno, že se rovněž objeví i v našem časném eneolitu (obr. 12).

Početně a tvarově pestřejší je plastická výzdoba. U domácí moravské eneolitické keramiky rozlišujeme následující tvary:

1. Výčnělky (1-38). Některé bezprostředně navazují na předchozí neolitické tvary či jsou s nimi totožné (například tvary 02 a 04). Jiné se od neolitických výrazně odlišují a tvoří typické projevy plastické výzdoby. Jejich zastoupení na určitých částech keramických tříd je různé a často se vyskytují s další výzdobou jak plastickou, tak i vhloubenou (obr. 13-14).
2. Zvláštní plastické tvary (41-54). Mezi ně patří různě tvarované lišty, které nesou často i další výzdobu. Převahu mají plastické obloučky, často však i složitější tvary. Jejich přítomnost na povrchu keramických tříd není většinou samostatná, nýbrž se objevují ve spojitosti s dalšími výzdobnými projevy (obr. 15).
3. Provrtané výčnělky (61-93). Podobně jako výčnělky bez provrtu vytvářejí na keramických třídách charakteristickou výzdobu. Některé z nich se podobají neolitickým, jiné představují zcela nové provedení. Podobně jako u výčnělek bez provrtu vystupují na různých částech keramických forem, často ve spojitosti s další výzdobou. Některé z nich budou

mít s největší pravděpodobností chronologický význam, poněvadž se vyskytly pouze v omezeném časovém období (obr. 16-17).

4. Ucha (101-129). Zatímco v době neolitické představovaly jen ojedinělý tvar plastické výzdoby, je jejich výskyt v pozdní době kamenné velmi četný a tvary jsou dosti rozmanité. Podle technologie rozlišujeme ucha pásková, která mají v tomto období naprostou převahu, a ucha tyčinkovitá. Dále jsou ucha většinou ještě zdobena jak plastickou, tak i vhloubenou výzdobou a vytvářejí mnohdy zvláštní tvary. Rovněž některé tvary mají význam pro chronologické postavení, neboť jejich výskyt se váže na určité časové období (obr. 18-19).
5. Lišty (131-209). Představují nejtypičtější eneolitickou výzdobu keramických tvarů. Objevují se jednak v hladkém provedení, jednak vytvářejí rozmanité obrazce často výzdobně velmi náročné. Jejich uspořádání na nádobách se různí, stejně jako kombinace, které často vytvářejí na určitých částech nádob složité tvary skládající se s plastické i vhloubené výzdoby. Některé lišty budou mít nesporně chronologický význam, především krátké hladké formy se objevují pouze v jevišovické kultuře (obr. 20-23).
6. Okrajové lišty (211-258). Jsou příbuzné s normálními lištami zdobícími různé části keramických tříd. Okrajové lišty vystupují přímo na okrajích nebo těsně pod okrajem v jednoduchém či několikanásobném provedení. Ojediněle jsou hladké, většinou však nesou vhloubenou výzdobu. Některé z nich mají zastoupení v průběhu celého eneolitu, jiné se vážou pouze na určité období. Okrajové lišty vytvářejí specifický ráz eneolitické keramiky a většina z nich je umístěna na hrncích nebo hrncovitých tvarech (obr. 24-26).

Jeden ze zásadních přínosů kódu moravské domácí eneolitické keramiky lze spatřovat ve sjednocování terminologie jak v oblasti typologické, tak i výzdobné. Skladba plastické a vhloubené výzdoby umožnila vytvoření jednotného popisu, a tím i předpoklad porovnávat mezi sebou soubory moravské domácí eneolitické keramiky z rozličných oblastí. Prostřednictvím zpracovaných kolekcí bude snad možné upřesnit či částečně korigovat naše dosavadní znalosti o vývoji sledovaných eneolitických kultur na Moravě. Snad v souvislosti s dalšími novými poznatky z terénu či díky novým a kvalitním souborům dojde k vytvoření detailnějších pohledů na vývoj alespoň některých domácích eneolitických kultur. Již tento poznatek by přinesl v rámci poznávání našich eneolitických kultur mnoho nového a postupně může docházet k doplňování znalostí o době eneolitické na Moravě.

Vytváření různých grafů z analyzovaných souborů a jejich vzájemné porovnávání bude zajištěno vhodně doplňovat mozaiku našich poznatků. Prostřednictvím různých kombinací, které je umožněno využitím počítačové techniky, získáme poznatky o vzájemném postavení a doplňování jednotlivých druhů výzdoby na určitých částech nádob. Takto zjištěné kombinace prvků a druhů výzdoby bychom bez využití počítačové techniky nikdy nezískali. Do budoucna lze předpokládat, že vyhodnocením kvalitních souborů dostaneme přesnější představu o vývoji některých moravských domácích eneolitických kultur a snad též prostřednictvím vhodně zvolených kombinací i zcela nové informace o vztazích mezi určitými kulturami či kulturními skupinami.

5.3 Databáze moravské domácí eneolitické keramiky

5.3.1 Struktura databázové tabulky

V následující tabulce je zobrazena struktura databáze moravské eneolitické keramiky. V navržené databázi odpovídá jeden řádek jednomu jedinci (keramickému fragmentu či celé nádobě) a sloupce jeho znakům. Podle přiloženého návrhu je možno vytvořit databázi, která bude plně kompatibilní s našimi daty.

Struktura databázové tabulky moravské domácí eneolitické keramiky				
Část databáze	Sloupec (Pole)	Význam zkratky	Datový typ pole	Velikost pole
A. Část obecná	Obec	název obce (například Kramolín)	Text	30
	INV	inventurní číslo, barva (h – hnědá, š – šedá, č – černá, čš – černošedá, hš – hnědošedá)	Text	6
	Poloha	individuální název tratě (polohy)	Text	30
	DAT	datování	Text	3
	SK	sektor lokality	Text	1
	CTV	čtverec	Text	3
	DROB	druh objektu	Číslo	8
	COBJ	číslo objektu	Číslo	8
	CAST	část objektu	Text	2
	VRS	vrstva objektu (píše se přímo údaje bez cm)	Text	7
B. Část typologická	DOCH	stupeň dochovalosti exempláře	Číslo	8
	OKR	tvár okraje	Číslo	8
	HR	tvár hrdla	Číslo	8
	PL	tvár plecí	Číslo	8
	SP	tvár spodku nádoby	Číslo	8
	DN	profil dna	Číslo	8
	NZ	tvár nožky	Číslo	8
	PRF	profilace nádoby	Číslo	8
	KT	keramická třída (typ a varianta)	Číslo	8
C. Část metrická (v cm)	PRO	průměr okraje	Číslo	8
	PNH	průměr nasazení hrdla	Číslo	8
	PRV	průměr výdutě (šířka střepu)	Číslo	8
	PD	průměr dna	Číslo	8
	PRP	průměr podstavy	Číslo	8
	CV	celková výška (výška střepu)	Číslo	8
	VH	výška hrdla	Číslo	8
	VP	výška plecí	Číslo	8
	VSP	výška spodku nádoby	Číslo	8
	VN	výška nožky	Číslo	8
D. Část fyzikálně technologická	TS	tloušťka stěny nádoby (fragmentu)	Číslo	8
	TVS	tvrdost stěny nádoby (fragmentu)	Číslo	8
	KH	druh keramické hmoty	Číslo	8
	PKH	příměs v keramické hmotě	Číslo	8
	PV	úprava povrchu vně	Číslo	8
	PU	úprava povrchu uvnitř	Číslo	8
E. Část výzdobná	DV	druh výzdoby vně a uvnitř	Číslo	8
	RVV	rozložení vhloubené výzdoby vně a uvnitř	Číslo	8
	RPV	rozložení plastické výzdoby vně a uvnitř	Číslo	8
	VV	vhloubená výzdoba	Číslo	8
	PVX	plastická výzdoba	Číslo	8

5.3.2 Deskripční kód moravské domácí eneolitické keramiky

V této kapitole předkládáme deskripční systém moravské eneolitické domácí keramiky. Pro jednotlivá pole (sloupce) databáze jsou zde přehlednou formou shrnuty kódy a jejich význam. Pole, jejichž charakter nevyžaduje použití kódu (např. Obec, CTV, SK), zde již nejsou uváděna.

Pro zvýšení produktivity práce je dobré vytvořit ve vaší databázi formulář, ve kterém lze s výhodou využít kombinovaných polí (dle terminologie programu MS Access). Ty jsou připojeny ke každému poli a je v nich umístěna kompletní nabídka kódů popisujících danou charakteristiku (např. OKR, RVV, VV). Požadovanou hodnotu si nemusíte pamatovat a psát ji, stačí ji pouze vybrat v seznamu. Kombinované pole tak zajistí minimalizaci chyb překlapy a vyloučí použití termínů, které se nevyskytují v kódu. Budete-li chtít použít kombinované pole v programu MS Access, vytvořte dva sloupce; do prvního, svázaného, sloupce zadejte kódová označení (př. 141), která se budou ukládat do příslušného pole (např. VV), s nímž je seznam svázan, do druhého, nesvázaného, sloupce zadejte významy kódů, které se nebudou v databázi ukládat, a které budou sloužit pouze jako pomoc při výběru správného řádku v seznamu (např. voštinování). Přesný návrh kombinovaných polí k jednotlivým položkám je umístěn v následujících tabulkách. V levém sloupci, který svážete s patřičným polem, jsou vypsány kódy, pravém sloupci jejich význam. Do kombinovaného pole s kódy vhloubené a plastické výzdoby umístěte ještě třetí nesvázaný sloupec s popisem skupin výzdoby. Tato informace vám pomůže v snadnější orientaci v rozsáhlých seznamech.

DAT (datování)	
Kód	Význam kódu
0	eneolit obecně
1	kultura nálevkovitých pohárů
2	starší stupeň kultury nálevkovitých pohárů
3	baalberský stupeň kultury nálevkovitých pohárů
4	mladší stupeň kultury nálevkovitých pohárů
5	bolerázský stupeň kultury s kanelovanou keramikou
6	kultura s kanelovanou keramikou
7	jevišovická kultura
8	bošácká skupina
9	skupina Hunyadihalom

DROB (druh objektu)	
Kód	Význam kódu
01	ojedinelý nález
02	povrchový sběr
03	kulturní vrstva
10	jáma (obecně)
11	stavební jáma
13	odpadní jáma
14	výrobní jáma
15	kulová jáma
17	silo
18	hliník
20	dům (obecně)
21	kulový dům
22	hliněný dům
23	cihlový dům
24	kamenný dům
25	sруб

DROB (druh objektu)	
Kód	Význam kódu
26	zemnice
27	polozemnice
28	dílna
30	konstrukční a fortifikační prvky
31	základový žlab
32	násyp (val)
33	příkop
34	kamenná zídka
35	palisáda
40	topeniště
41	ohniště
42	hliněná pec
43	kamenná pec
48	žároviště
50	komunikace
51	mazaníková kra
52	studna
53	vodojem
56	ulice
57	stezka
58	most
60	hrob (obecně)
61	rituální kostrový hrob
62	nerituální (neúplný) kostrový hrob
63	rituální žárový kremační hrob
64	kremační pozůstatky
67	hromadný hrob
68	zvířecí pohřeb
70	hromadný nález surovinový obecně
71	depot
72	keramický depot
73	výrobní depot
74	obchodní depot
75	poklad
78	votivní depot
80	kultovní objekt (obecně)

VRS (vrstva objektu)	
Kód	Význam kódu
01	povrch
0-20	0-20 cm
20-40	20-40 cm
40-60	40-60 cm
60-80	60-80 cm
220-240	220-240 cm

DOCH (stupeň dochovalosti)	
Kód	Význam kódu
1	okraj
2	hrdlo
3	plece
4	výduť
5	spodek nádoby
6	dno
7	nožka
8	rekonstruovatelný exemplář
9	zcela dochovaný exemplář

OKR (tvar okraje) – obr. 1	
Kód	Význam kódu
1	kyjovitý
2	zaoblený
3	zúžený
5	límcovitý
6	zaostřený
8	rovný
9	dovnitř seříznutý

HR (tvar hrdla) – obr. 2	
Kód	Význam kódu
1	prohnuté
2	kónické
3	přímé
4	přímé s vybočeným okrajem

PL (tvar plecí) – obr. 2	
Kód	Význam kódu
1	klenutý
2	kónický
3	přímý
4	prohnutý

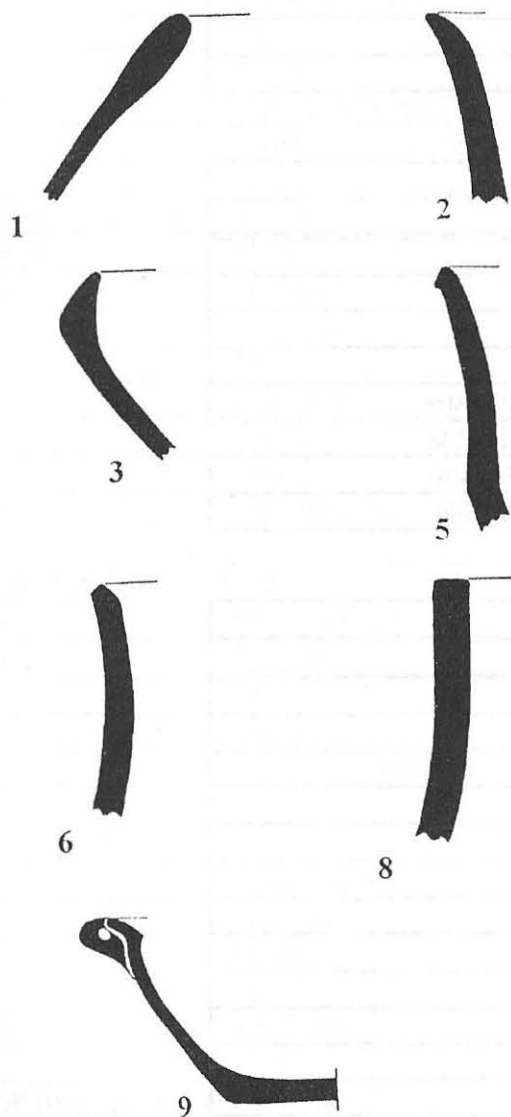
SP (tvar spodku nádoby) – obr. 2	
Kód	Význam kódu
1	klenutý
2	kónický
3	přímý
4	prohnutý

DN (profil dna) – obr. 3	
Kód	Význam kódu
1	ostrý
2	oblý
3	nožkovitý
8	dno nevyznačeno

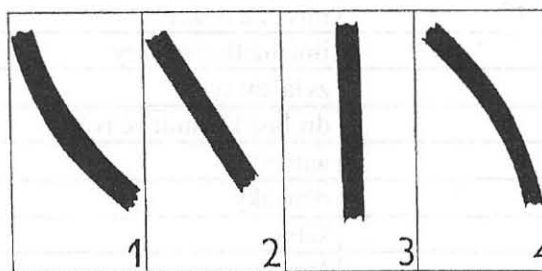
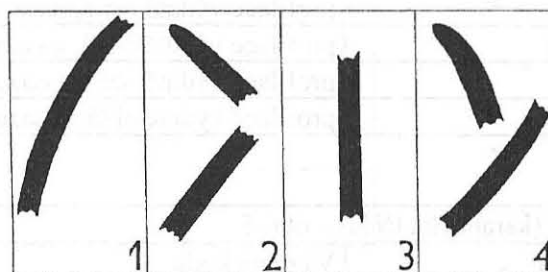
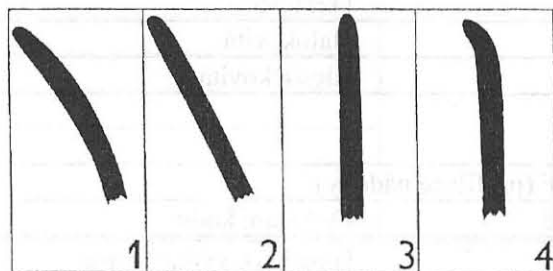
NZ (tvar nožky) – obr. 4	
Kód	Význam kódu
1	válcovitá
3	kónická
5	mírně silně prohnutá
6	prohnutá
7	křížová
8	lalokovitá
9	destičkovitá

PRF (profilace nádoby)	
Kód	Význam kódu
1	profilace výdutě ostrá
2	profilace výdutě oblá
3	nasazení hrdla ostré
4	nasazení hrdla oblé
5	profilace výdutě ostrá, nasazení hrdla ostré
6	profilace výdutě ostrá, nasazení hrdla oblé
7	profilace výdutě oblá, nasazení hrdla ostré
8	profilace výdutě oblá, nasazení hrdla oblé

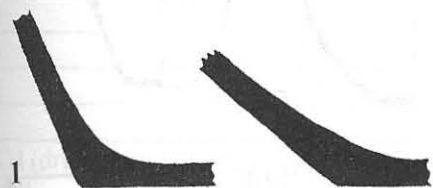
KT (keramická třída) – obr. 5	
Kód	Význam kódu
01	hrnce
02	jiné hrncovité typy (poháry)
03	mísy
04	mísy na nožce
06	jiné užitkové tvary
07	zvláštní typy
09	drobné keramické tvary
10	amfory
11	džbánky
12	šálky
13	čerpáky
14	hmoždíře
15	rendlíky
16	láhev s límcem
17	plastika
18	buben



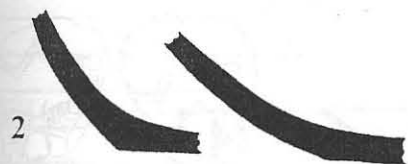
Obr. 1. Tvary okrajů moravské domácí eneolitické keramiky.



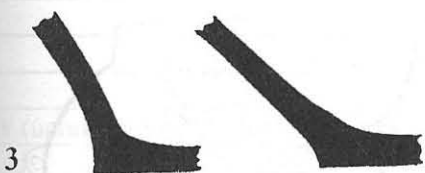
Obr. 2. Tvary základních částí nádob (hrdlo, plece, spodek).



1



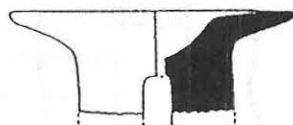
2



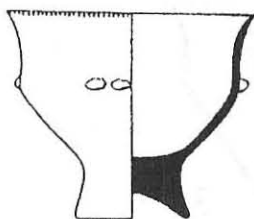
3



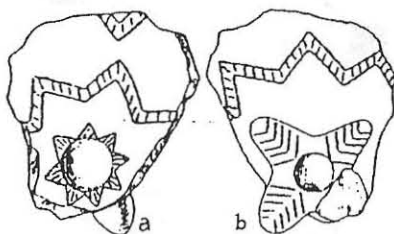
8



1

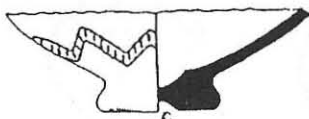


3



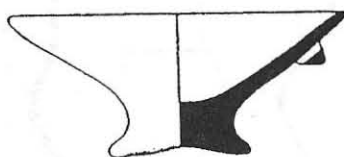
a

b



c

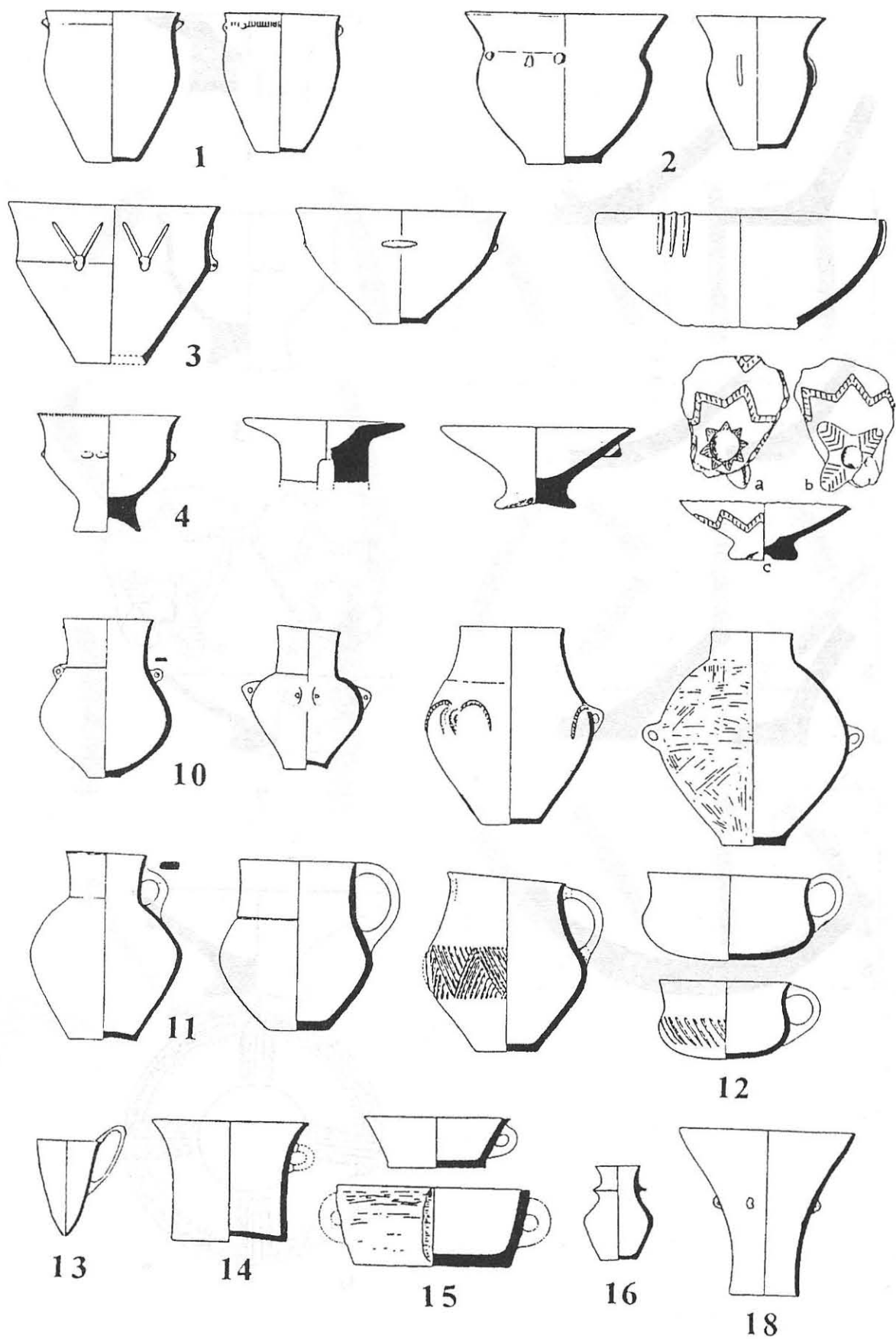
7



9

Obr. 3. Profily den moravské domácí eneolitické keramiky.

Obr. 4. Tvary nožek moravské domácí eneolitické keramiky.



Obr. 5. Keramické třídy moravské domácí eneolitické keramiky.

TVS (tvrdost stěny nádoby)	
Kód	Význam kódu
1	velmi měkká (oděr prsty)
2	měkká (lze rýpat nehtem)
3	normální (lze rýpat měděným plechem)
4	tvrdá (lze rýpat ocelovým nožem)

KH (druh keramické hmoty jedince)	
Kód	Význam kódu
1	plavená
2	jemnozrnná (do 1 mm)
4	středozrnná (do 2 mm)
6	zrnitá (do 3 mm)
7	hrubozrnná (nad 3 mm)

PKH (příměs v keramické hmotě)	
Kód	Význam kódu
1	příměs barviva
2	příměs šamotu
3	příměs kostí
4	příměs mušlí
5	příměs tuhy
8	příměs slídy
9	příměs organická

PV (úprava povrchu vně)	
Kód	Význam kódu
1	zrnitý, hrubý
2	hrubě modelovaný
4	hlazený
5	leštěný
6	potažený zvláštní vrstvou
7	potažený zvláštní vrstvou a hlazený
8	potažený zvláštní vrstvou a leštěný

PU (úprava povrchu uvnitř)	
Kód	Význam kódu
1	zrnitý, hrubý
2	hrubě modelovaný
4	hlazený
5	leštěný
6	potažený zvláštní vrstvou
7	potažený zvláštní vrstvou a hlazený
8	potažený zvláštní vrstvou a leštěný

DV (druh výzdoby jedince vně a uvnitř)	
Kód	Význam kódu
03	vhloubená
04	plastická

DV (druh výzdoby jedince vně a uvnitř)	
Kód	Význam kódu
34	vhlobená a plastická
05	inkrustace
35	vhlobená a inkrustace

RVV (rozložení vhlobené výzdoby vně a uvnitř)	
Kód	Význam kódu
1	okraj
2	hrdlo
3	plece
4	výduť
5	spodek
6	dno
7	nožka
8	ucho
12	okraj a hrdlo
23	hrdlo a plece
24	hrdlo a výduť
34	plece a výduť
35	plece, výduť a spodek
45	výduť a spodek
50	celé tělo
	Pokud je výzdoba uvnitř přidává se 0, například 03, 050, 035

RPV (rozložení plastické výzdoby vně a uvnitř)	
Kód	Význam kódu
01	okraj
02	hrdlo
03	plece
04	výduť
05	spodek
06	dno
07	nožka
12	okraj a hrdlo
14	okraj a výduť
23	hrdlo a plece
24	hrdlo a výduť
34	plece a výduť
35	plece a spodek
45	výduť a spodek

Výzdoba na moravské domácí eneolitické keramice se rozpadá na dvě hlavní složky:

1. Vhlobenou výzdobu, která představuje takové výzdobné projevy, které se zahlubují do povrchu keramických tříd (vpichy, vrypy, jamky, rýhy, žlábký, brázděný vpich, prstování, bernburgská výzdoba, slámování, voštinování).

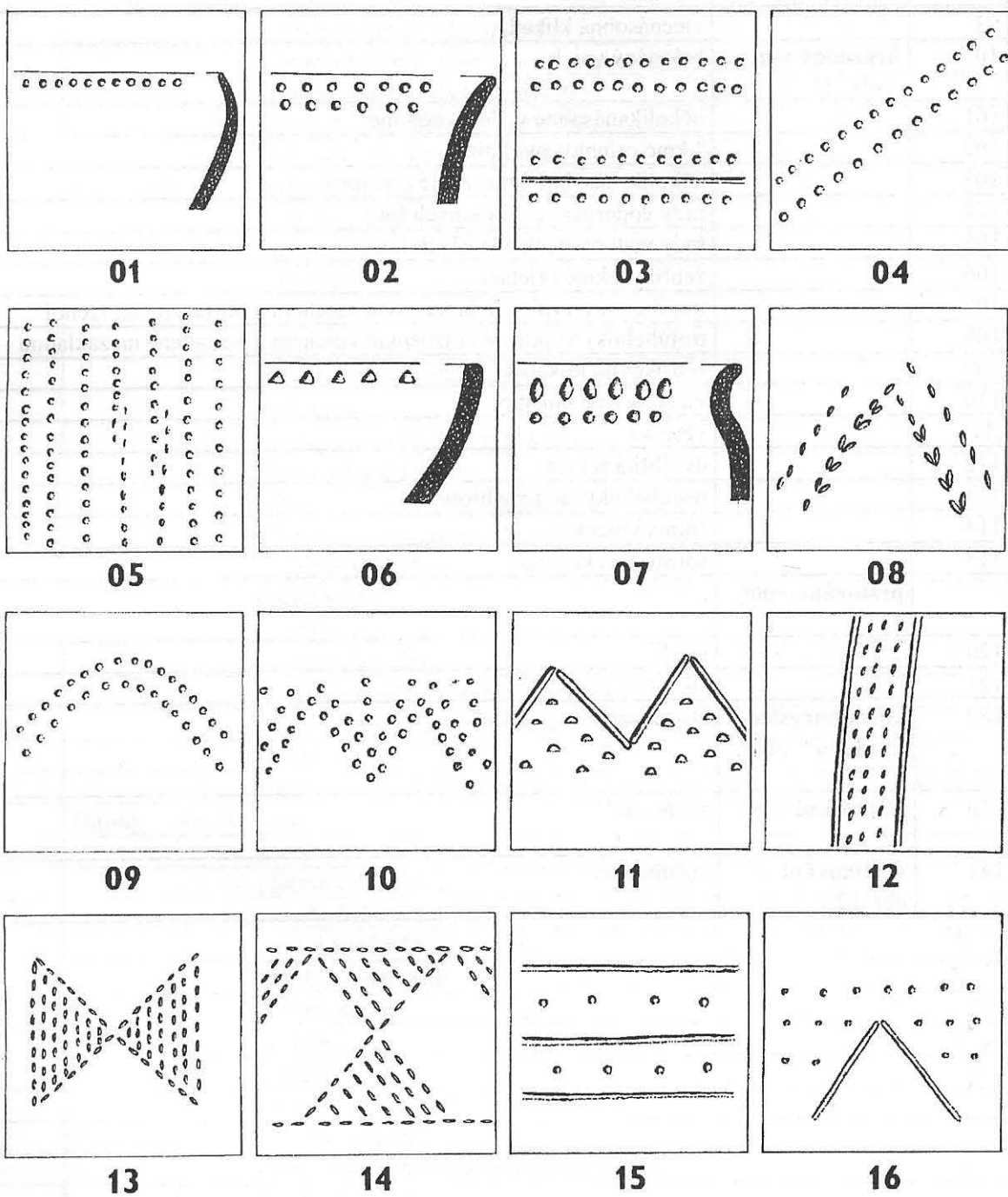
2. Plastickou výzdobou rozumíme vše, co vystupuje na povrchu keramických tříd, kromě voštinování. Patří sem výčnělky, zvláštní plastické tvary, provrtané výčnělky, ucha, lišty a okrajové lišty.

VV (vhloubená výzdoba)		
Kód	Skupiny výzdoby	Význam kódu
01	vpichy – obr. 6	vpichy
02		vpich v horizontálních řadách
03		dvě horizontální řady vpichů
04		šikmé řady vpichů
05		svislé řady vpichů
06		trojúhelníkový vpich v horizontální řadě
07		nepřavidelný vpich v horizontálních řadách
08		hrotitý vpich v klikatce
09		dvojitý oblouk
10		trojúhelníky z vpichů
11		trojúhelníky vyplněné vpichy
12		páska vyplněná vpichy
13		dvojbřítá sekera z vpichů
14		přesypací hodiny z vpichů
15		vpichy ohraničené liniemi
16		vpichy s klikatkou
20	vrypy (na rozdíl od vpichů jsou protáhlejší) – obr. 7	vrypy
21		svislé vrypy
22		obdélníkové vrypy v řadách
23		vrypy v horizontálních dvojřadách
24		vrypy ve více horizontálních řadách
25		vrypy ve tvaru šipky
26		horizontální řady segmentových vrypů
27		kapkovitý vryp v horizontální řadě
28		pásky vyplněné vrypy
29		zvlněný tělískový vryp
30	jamky – obr. 8	jamky
31		horizontální řada jamek
32		jamky ve svislých řadách
33		oválné jamky v horizontálním uspořádání
34		oválné jamky střídavě položené ve dvou řadách
35		oválné jamky střídavě položené ve třech řadách
36		zahrocené jamky na těle nádoby
37		šikmo orientované oválné jamky v horizontální řadě
38		jamky nepřavidelného tvaru v horizontální řadě
39		jamky trojúhelníkového tvaru postavené na hrot v horizontální řadě
40		jamky trojúhelníkového tvaru postavené na základně v horizontální řadě
41		jamky trojúhelníkového tvaru postavené na základně ve dvou horizontálních řadách
42		jamky půlměsíkového tvaru v horizontálním uspořádání
43		jamky tvaru segmentu v horizontálním uspořádání

VV (vhloubená výzdoba)		
Kód	Skupiny výzdoby	Význam kódu
44		„kávová zrnka“ v horizontální řadě
45		jamky obdélníkového tvaru v horizontálním uspořádání
46		okrouhlé jamky na těle nádoby
47		plastické trojúhelníky vyplněné jamkami
48		girlandy vytvořené jamkami
50	rýhy – obr. 9	rýhy
51		vodorovné rýhy přerušované svislými
52		šikmé rýhy
53		svislé rýhy
54		„i“ motiv
55		šikmé klikatky
56		několikanásobné klikatky
57		mřížování
58		motiv nepravidelné „rybí kosti“
59		hroty na těle nádoby
60		žebříkovité klikatky – jednoduchá
61		žebříkovité klikatky – dvojité
62		spojené kosočtverce
63		ptačí stopa
64		křížky
65		obrvené kolečko
66		rýhovaná pole v kombinaci s jamkami
67		rýhy ve tvaru „nepravidelných vlčích zubů“
68		„rybí kost“
69		trojúhelníky zavěšené na rýze
70	žlábký – obr. 10	žlábký
71		vodorovné žlábkování
72		šikmé žlábkování
73		svislé žlábkování
74		několikanásobný půloblouk
75		žebřík vodorovný
76		žebřík šikmý
77		žebřík svislý
78		„hrábě“
79		žlábkovaná větévka
80		žlábkované „x“
81		několikanásobné žlábkování do klikatky
82		vlčí zuby
83		nepravidelné vlčí zuby
84		žlábkování nepravidelného uspořádání
85		nepravidelné žlábkování v polích
86		žlábkování do čtvrtin kruhu
87		„rybí kost“
88		„rybí kost“ svislé orientace
89		žlábkovaná klikatka
90		trojúhelník vyplněný horizontálními žlábký
91		šikmá vlnice
92		nepravidelné žlábkování kombinované jamkami

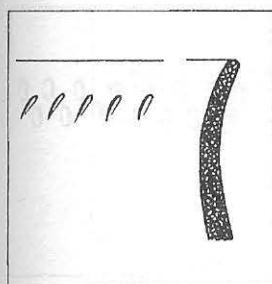
VV (vhloubená výzdoba)		
Kód	Skupiny výzdoby	Význam kódu
93		vícenásobná klikatka
100	brázděný vpich – obr. 11	brázděný vpich
101		několikanásobné vodorovné linie
102		šikmo orientované linie
103		několikanásobné linie svislé orientace
104		řady vodorovných a šikmých linií
105		řady vodorovných a svislých linií
106		žebříky šikmé orientace
107		trojúhelníky vyplněné brázděným vpichem a postavené na vrchol
108		trojúhelníky vyplněné brázděným vpichem a postavené na základnu
109		žebříkovitá klikatka
110		vícenásobné klikatky
111		větévka
112		dvojbřítá sekera
113		trojúhelníky spojené hroty
114		šikmá klikatka
115		soustředná kolečka
	prstování – obr. 12	
120		svislé
121		šikmé
130	„bernburská výzdoba“ -obr. 12	„bernburská výzdoba“
140	slámování -obr. 12	slámování
141	voštinování – obr. 12	voštinování

Vpichy

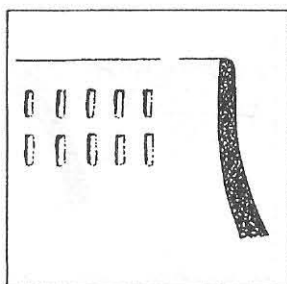


Obr. 6. Vhloubená výzdoba – vpichy (1-16).

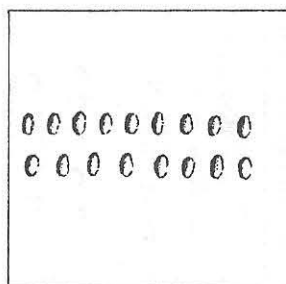
Vrypy



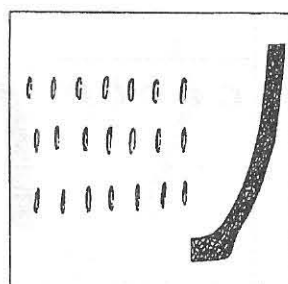
21



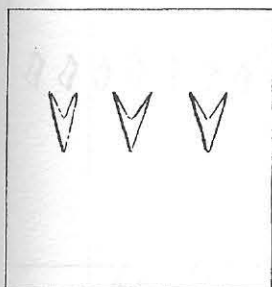
22



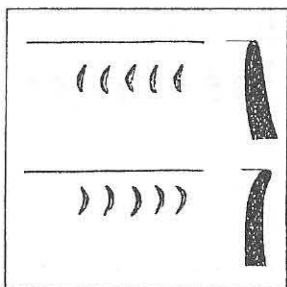
23



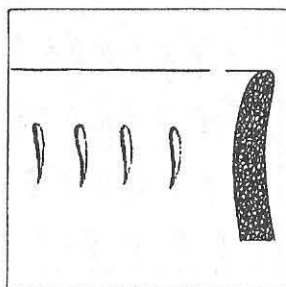
24



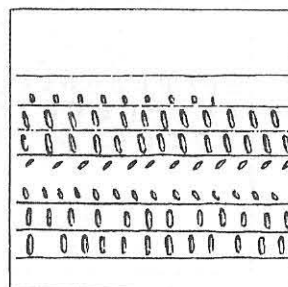
25



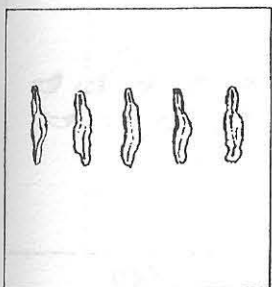
26



27



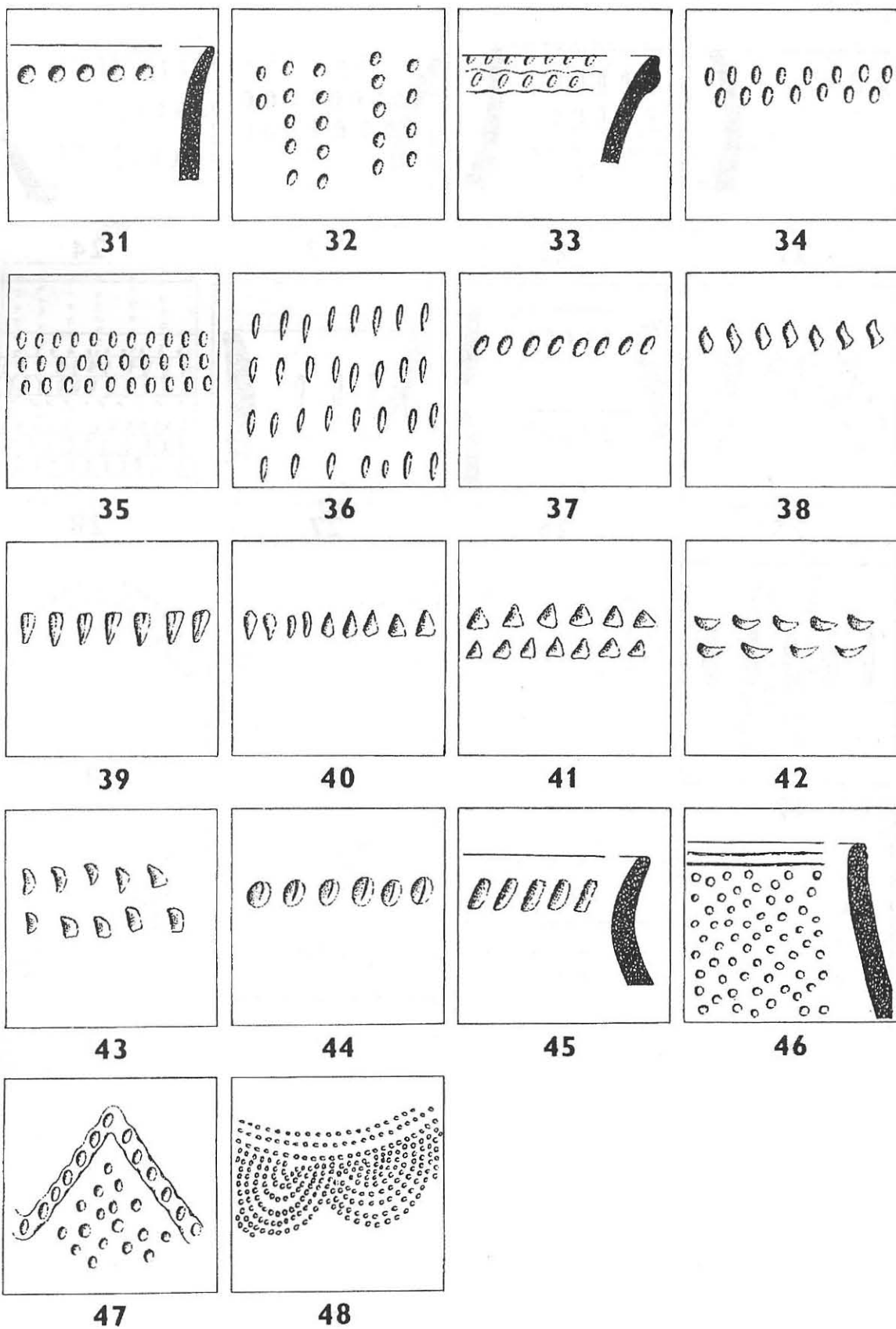
28



29

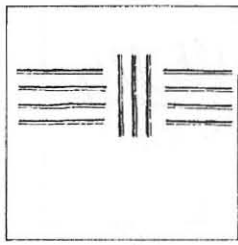
Obr. 7. Vhloubená výzdoba – vrypy (21-29).

Jamky

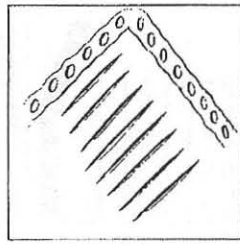


Obr. 8. Vhloubená výzdoba – jamky (31-48).

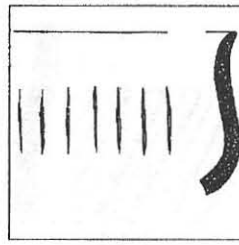
Rýhy



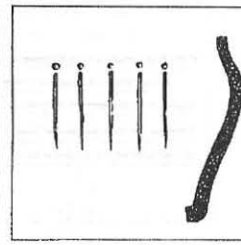
51



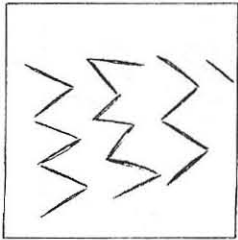
53



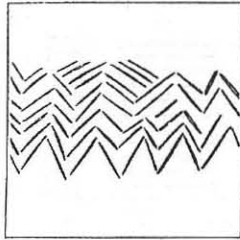
54



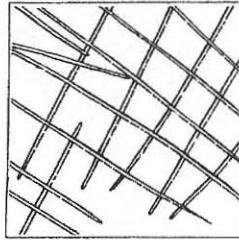
55



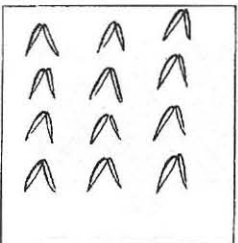
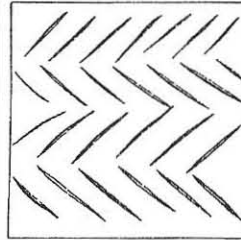
56



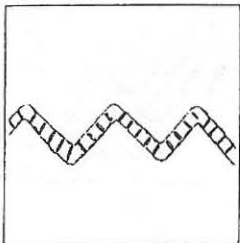
57



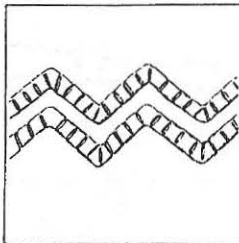
58



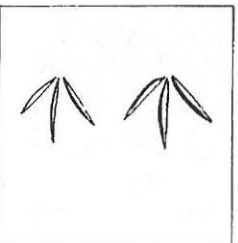
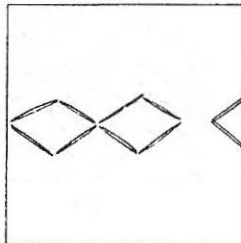
60



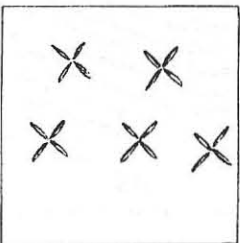
61



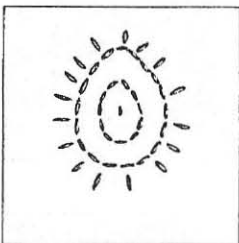
62



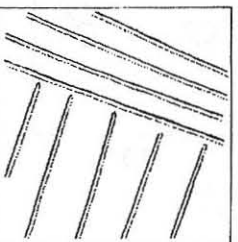
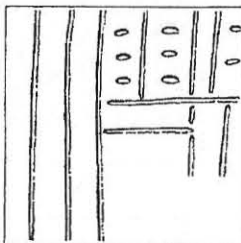
64



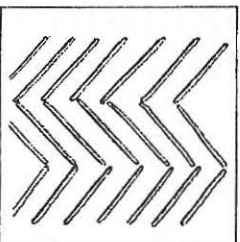
65



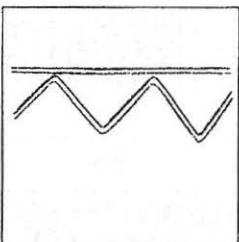
66



68

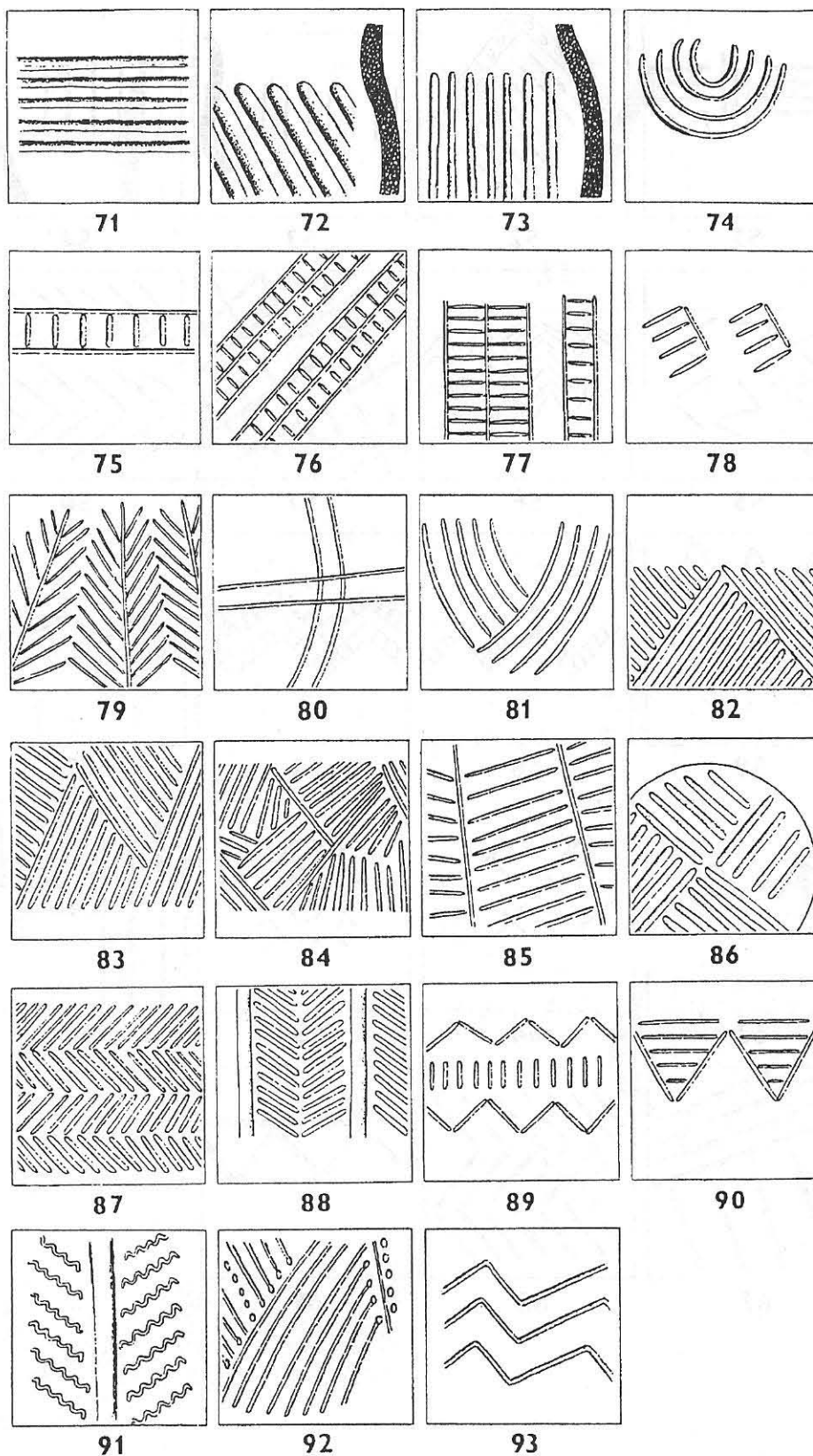


69



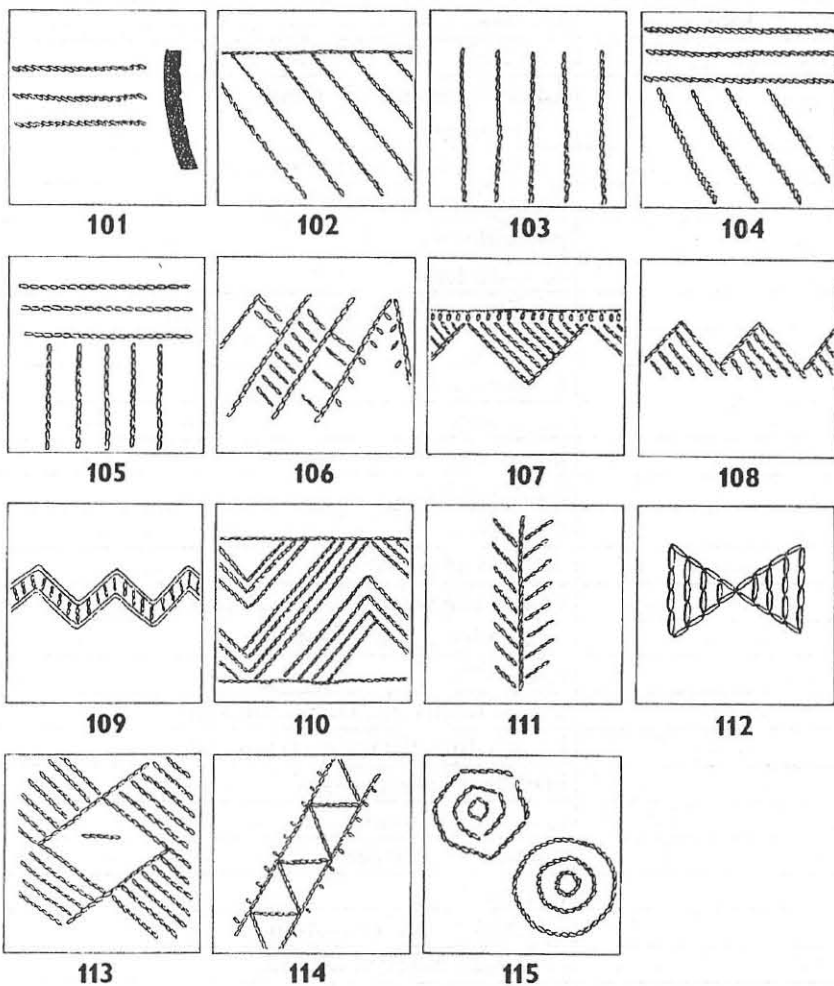
Obr. 9. Vhloubená výzdoba – rýhy (51-69).

Žlábký



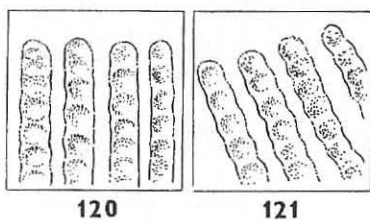
Obr. 10. Vhloubená výzdoba – žlábký (71-93).

Brázděný vpich

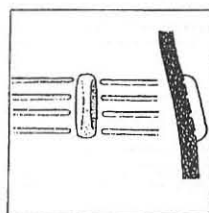


Obr. 11. Vhloubená výzdoba – brázděný vpich (101-115).

Prstování



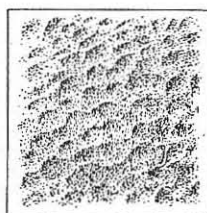
Bernburgská výzdoba



Slámování



Voštinování



Obr. 12. Vhloubená výzdoba – prstování (120-121), bernburgská výzdoba (130), slámování (140), voštinování (141).

PVX (plastická výzdoba)		
Kód	Skupiny výzdoby	Význam kódu
00	výčnělky – obr. 13-14	výčnělky
01		rohatý s jamkou na vrcholu
02		svise protáhlý
03		svise protáhlý s jamkou na vrcholu
04		drobný nevýrazný
05		polokulovitý
06		ve tvaru luku
07		s jamkou na vrcholu
08		sedlovitý
09		fazolovitý
10		jazykovitý
11		kuželovitý
12		knoflíkovitý
13		vysedlý
14		vodorovně protáhlý
15		vodorovně protáhlý s jamkou na vrcholu
16		fazolovitý šikmé orientace
17		kapkovitý
18		svislá dvojice lišt nad výčnělkem
19		jazykovitý s lištami ve tvaru „U“
20		jazykovitý přesekávaný
21		jazykovitý rozeklaný
22		obloukovitě hráněný
23		„pavouk“
24		„pavouk“ s přesekáváním
25		se dvěma svislými lištami
26		se svislými a šikmými lištami
27		s obráceným „W“
28		s obráceným přesekávaným „W“
29		sedlovitý s d'ubkovanými prohnutými lištami ve tvaru obráceného „W“
30		fazolovitý s paličkami
31		čočkovitý
32		výčnělek s obloukovitou lištou
33		výčnělek s přesekávanými lištami
34		výčnělek s přímou či dovnitř prohnutou lištou
35		výčnělek s vousy
36		držadlo
37		výčnělek se svislými lištami
38		výčnělek s paprskovitými lištami
40	zvláštní plastické tvary – obr. 15	zvláštní plastické tvary
41		obrácená „V“ lišta
42		obrácená „U“ lišta s přeseky
43		„U“ lišta s jamkami
44		obrácená dvojitá „U“ lišta
45		obrácená „W“ lišta
46		trojitá obrácená „V“ lišta s jamkami

PVX (plastická výzdoba)		
Kód	Skupiny výzdoby	Význam kódu
47		dvojitá obrácená „V“ lišta s jamkami
48		plastická „U“ lišta
49		„vous“
50		přesekávaný „vous“
51		„vous“ s háčkovitým ukončením
52		dvě trojice prohnutých přesekávaných lišt
53		krátká hladká lišta šikmé orientace
54		krátká hladká lišta svislé orientace
60	provrtané výčnělky – obr. 16-17	provrtané výčnělky
61		hráněné ostré s horizontálním provrtem
62		hráněné oblé s horizontálním provrtem
63		horizontálně protáhlé s horizontálním provrtem
64		svisle protáhlé s horizontálním provrtem
65		hráněné s horizontálním provrtem
66		prožlábnuté s horizontálním provrtem
67		tunelovité
68		tunelovité prožlábnuté
69		širší prožlábnuté s horizontálním otvorem
70		s žebrem uprostřed a horizontálním otvorem
71		trojnásobně prožlábnutý s horizontálním otvorem
72		horizontálně protáhlé s vertikálním otvorem
73		svisle protáhlé s vertikálním otvorem
74		svislé tunelovité
75		svislé válcovité s vertikálním otvorem
76		s d'ubkovanými hranami a vertikálním otvorem
77		svisle protáhlé se třemi žebry a vertikálním otvorem
78		subkutánní ouško ve stěně
79		brylovité
80		„cívka“
81		horizontálně provrtané s lištou ve tvaru obráceného „V“
82		horizontálně provrtané s lištou ve tvaru obráceného „W“
83		horizontálně provrtané s přesekávanou lištou ve tvaru obráceného „V“
84		horizontálně provrtané s dvojicí šikmých přesekávaných lišt
85		horizontálně provrtané s přesekávanou lištou ve tvaru obráceného „W“
86		horizontálně provrtané s šikmými d'ubkovanými lištami
87		tunelovité s lištou svislou nad
88		tunelovité s „V“ lištou nad
89		svisle provrtané s „V“ lištou nad
90		horizontálně provrtané ve tvaru „hlavy krávy“
91		horizontálně provrtané s lištami ve tvaru beraních rohů
92		horizontálně provrtané ve tvaru rohů
93		horizontálně provrtané s vybíhajícími lištami
100	ucha – obr. 18-19	ucha
101		pásková úzká
102		pásková úzká vytažená nad okraj

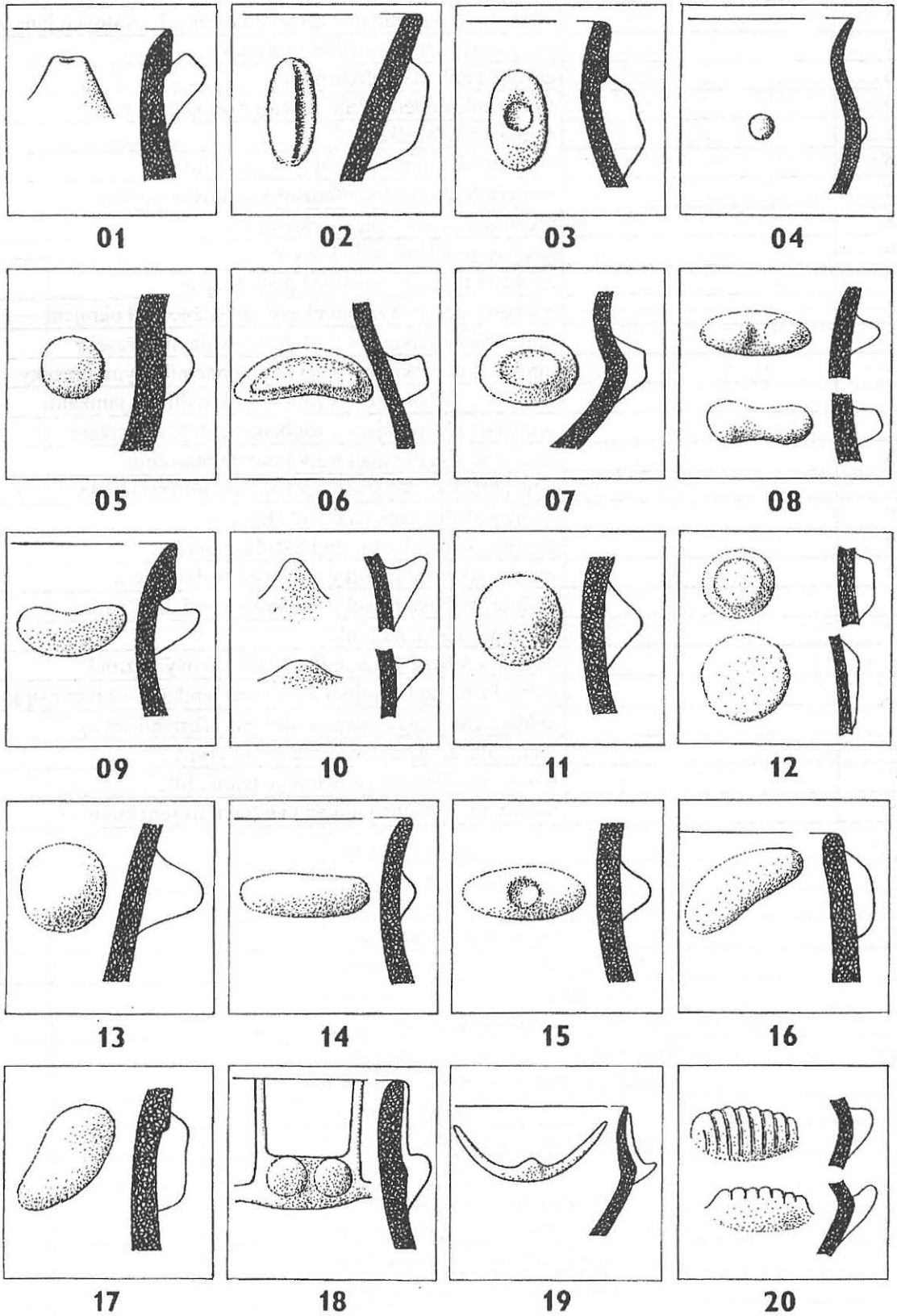
PVX (plastická výzdoba)		
Kód	Skupiny výzdoby	Význam kódu
103		pásková úzká prožlábnutá
104		pásková úzká prožlábnutá s jamkami na hraně
105		ansa lunata
106		ansa cornuta
107		pásková úzká vytažená nad okraj a rozeklaná
108		pásková široká
109		pásková široká prožlábnutá
110		pásková s výčnělkem
111		prožlábnutá hráněná
112		zdůrazněná
113		pásková prožlábnutá vytažené nad okraj
114		s výčnělkem nahoře
115		se čtyřmi výčnělky nahoře
116		pásková vytažená nad okraj, nahoře rozeklaná
117		pásková s vousem u kořenu
118		pásková s přesekávaným vousem u kořenu
119		pásková se svislými „U“ lištami u kořenu
120		s „roh“
121		typu „hunyadihalom“
122		se dvěma výčnělky u kořene
123		s přesekávaným či hladkým vousem a s výčnělkem uprostřed
124		páskové ucho s výčnělkem
125		vertikálně provrtané ucho
126		vícenásobně prožlábnuté ucho
127		tyčinkovité ucho
128		„zaječí ucho“
129		„baalberské ucho“
130	lišty – obr. 20-23	lišty
131		vodorovné s okrouhlými jamkami
132		vodorovné s vpichy ve tvaru hrotu šípů
133		vodorovné s vpichy ve dně jamky
134		vodorovné přesekávané
135		vodorovné s rohlíčkovitými jamkami
136		vodorovné s oválnými jamkami
137		vodorovné d'ubkované
138		vodorovné dvojité s přeseky
139		vodorovné dvojité se zahrocenými jamkami
140		vodorovné dvojité s oválnými jamkami
141		vodorovné dvojité s d'ubkováním
142		vodorovné trojité s přeseky
143		vodorovné trojité s oválnými jamkami
144		vodorovné krátké s přeseky
145		svislé krátké
146		svislé krátké dvojnásobné
147		svislé krátké trojnásobné
148		svislé krátké čtyřnásobné
149		svislé krátké vícenásobné

PVX (plastická výzdoba)		
Kód	Skupiny výzdoby	Význam kódu
150		svislé krátké přesekávané
151		svislé krátké dvojnásobné přesekávané
152		svislé krátké trojnásobné přesekávané
153		svislé krátké čtyřnásobné přesekávané
154		svislé krátké d'ubkované
155		svislé krátké dvojnásobné d'ubkované
156		svislé krátké trojnásobné d'ubkované
157		šikmé s přesečky
158		šikmé s jamkami
159		šikmé s d'ubkami
160		šikmé dvojité s přesečky
161		šikmé dvojité s jamkami
162		šikmé dvojité s d'ubkami
163		přesekávané ve tvaru obráceného „V“
164		paprščité s přesečky
165		paprščité s jamkami
166		paprščité s d'ubkami
167		ve tvaru kříže zdobené jamkami
168		několikanásobné svislé
169		obloukovité trojnásobné
170		obloukovité čtyřnásobné s přesečky
171		svislé paličky
172		ve tvaru obráceného nepravidelného „V“
173		různě uspořádané
174		plastická vlnice
175		plastická klikatka
176		plastická klikatka s přesečky
177		vodorovné se zavěšenou klikatkou
178		vodorovné přesekávané se zavěšenou klikatkou
179		vodorovné s jamkami a se zavěšenou klikatkou
180		vodorovné d'ubkované se zavěšenou klikatkou
181		vodorovné přesekávané se zavěšenou přesekávanou klikatkou
182		vodorovné s jamkami se zavěšenou klikatkou zdobenou jamkami
183		vodorovné s jamkami a zavěšeným „U“
184		vodorovné d'ubkované se zavěšenou d'ubkovanou klikatkou
185		vodorovné dvojité zdobené jamkami a se zavěšenou klikatkou
186		vodorovné dvojité d'ubkované s klikatkou
187		vodorovné dvojité d'ubkované se zavěšeným d'ubkovaným „U“
188		vodorovné dvojité zdobené přesečky se zavěšenou klikatku zdobenou přesečky
189		vodorovné dvojité zdobené jamkami a zavěšenou klikatku zdobenou jamkami
190		vodorovné dvojité zdobené d'ubkami a zavěšenou klikatku zdobenou d'ubkami
191		vodorovné trojité s přesečky se zavěšenou klikatku zdobenou přesečky

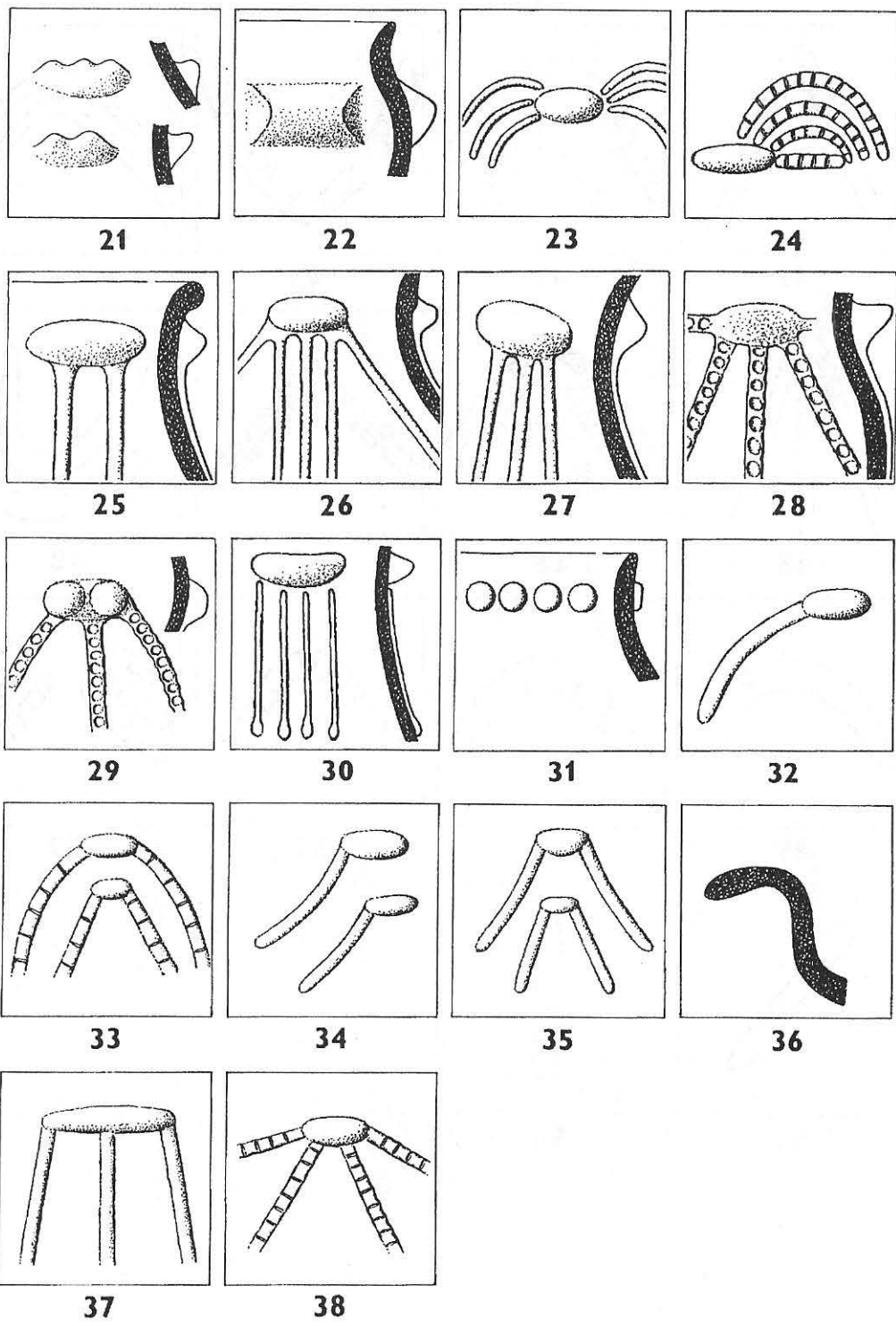
PVX (plastická výzdoba)		
Kód	Skupiny výzdoby	Význam kódu
192		vodorovné trojité s d'ubkami se zavěšenou klikatkou zdobenou d'ubkami
193		vodorovné trojité s dvojitou klikatkou
194		vodorovné dvojité se svislými trojnásobnými krátkými lištami
195		svislá lišta nad d'ubkovanou horizontální lištou
196		vodorovné dvojité s několikanásobnými zavěšenými svislými lištami
197		vodorovné dvojité s přesečky se zavěšenou dvojicí svislých krátkých lišt
198		vodorovné dvojité d'ubkované se zavěšenou dvojicí svislých krátkých lišt
199		vodorovné dvojité s přesečky s trojicí zavěšených krátkých svislých lišt
200		vodorovné dvojité s jamkami s krátkými zavěšenými lištami zdobenými jamkami
201		vodorovné dvojité s d'ubkovaním a se zavěšenými dvojicemi krátkých svislých d'ubkovaných lišt
202		vodorovné dvojité s jamkami a s trojicí zavěšených svislých lišt zdobených jamkami
203		plastická klikatka mezi dvojnásobnými či trojnásobnými vodorovnými lištami
204		dvojitá klikatka mezi vodorovnými lištami
205		vodorovná lišta sevřená klikatkami
206		klikatka s kuličkami na vrcholech
207		plastický strom
208		složité plastické kombinace
209		s d'ubkami ve tvaru kávových zrněk
210	okrajové lišty – obr. 24-26	okrajové lišty
211		límcovité hladké
212		límcovité vykrajované
213		límcovité přesekávané
214		límcovité zdobené segmenty
215		límcovité zdobené oválnými jamkami
216		límcovité zdobené okrouhlými jamkami
217		límcovité vykrajované zdobené oválnými jamkami
218		límcovité vykrajované zdobené šikmými okrouhlými jamkami
219		límcovité zdobené zvlněnými jamkami
220		zdobené vpichy
221		zdobené trojúhelníkovými přesečky
222		zdobené oválnými přesečky
223		zdobené kapkovitými jamkami
224		zdobené hrotitými jamkami
225		zdobené obdélníkovými jamkami
226		zdobené segmentovými jamkami
227		vykrajované d'ubkované
228		zdobené dvěma řadami vrypů
229		límcovité vykrajované s dvojitou řadou okrouhlých jamek
230		zdobené dvojitou řadou okrouhlých jamek

PVX (plastická výzdoba)		
Kód	Skupiny výzdoby	Význam kódu
231		zdobené dvěma řadami střídavě uložených oválných jamek
232		dvojitě zdobené oválnými jamkami
233		dvojitě zdobené d'ubkováním
234		límcovité zdobené třemi řadami oválných jamek
235		trojnásobné hladké
236		trojnásobné zdobené oválnými jamkami
237		trojnásobné zdobené přerušově oválnými vpichy
238		vícenásobné zdobené d'ubkováním
239		hladké, umístěné pod okrajem
240		zdobené přesečky, umístěné pod okrajem
241		zdobené kapkovitými jamkami, umístěné pod okrajem
242		umístěné pod okrajem a zdobené oválnými přesečky
243		umístěné pod okrajem a zdobené obdélníkovými přesečky
244		umístěné pod okrajem a zdobené okrouhlými jamkami
245		umístěné pod okrajem a zdobené oválnými jamkami
246		umístěné pod okrajem a zdobené d'ubkováním
247		umístěné pod okrajem a zdobené zvlněnými tělísky
248		dvojitě hladké umístěné pod okrajem
249		dvojitě, jedna hladká, druhá d'ubkovaná
250		dvojitě zdobené přesečky umístěné pod okrajem
251		dvojitě zdobené přesečky na okraji a pod okrajem
252		věnc v'ýčnělků na hrdle
253		dvojitě zdobené jamkami ve tvaru kávových zrněk
254		jednoduchá pod okrajem a zdobená jamkami ve tvaru šipky
255		jednoduchá pod okrajem a zdobená šikmými vrypy
256		límcovitá, zdobená obdélníkovými vrypy
257		límcovitá, zdobená jamkami ve tvaru „M“
258		límcovitá, zdobená jamkami ve tvaru trojúhelníku

Výčnělky

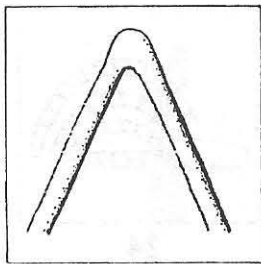


Obr. 13. Plastická výzdoba – výčnělky (1-20).

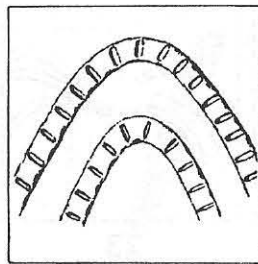


Obr. 14. Plastická výzdoba – výčnělky (21-38).

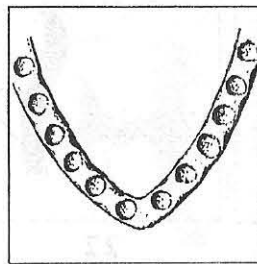
Zvláštní plastické tvary



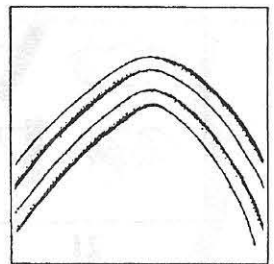
41



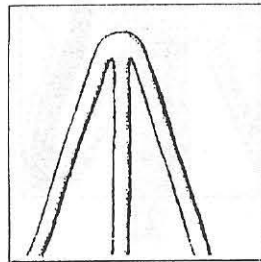
42



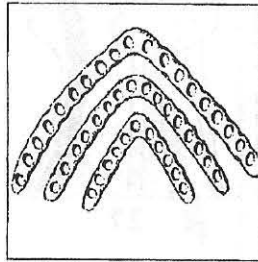
43



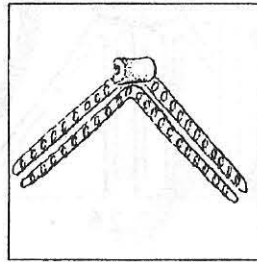
44



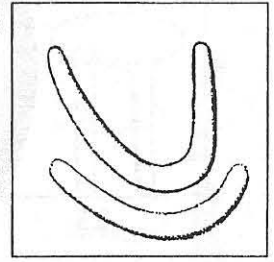
45



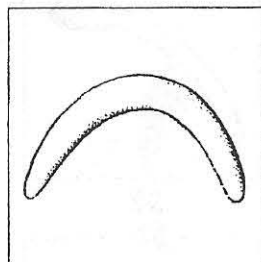
46



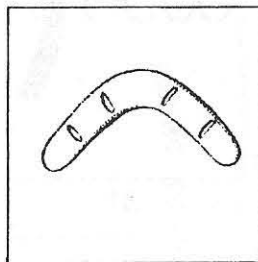
47



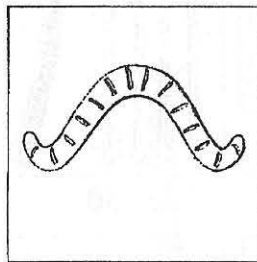
48



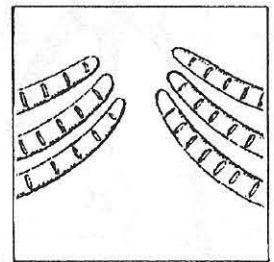
49



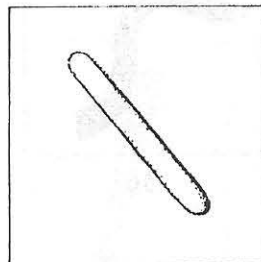
50



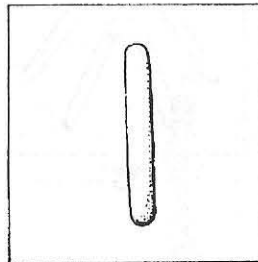
51



52



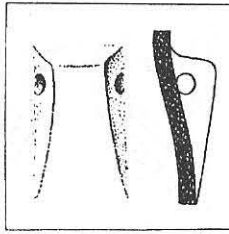
53



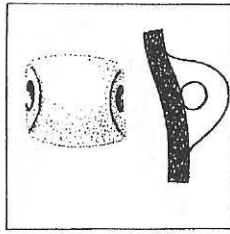
54

Obr. 15. Plastická výzdoba – zvláštní plastické tvary (41-54).

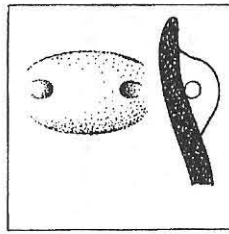
Provrtané výčnělky



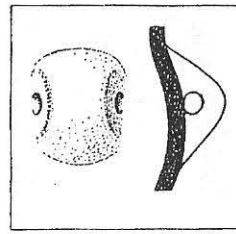
61



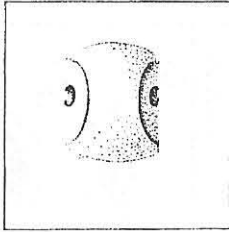
62



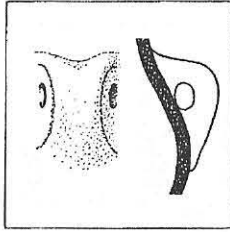
63



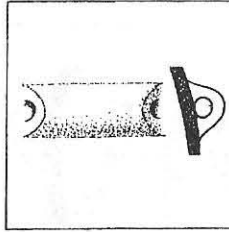
64



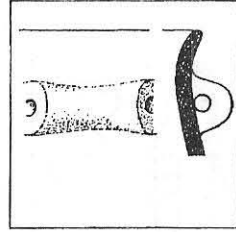
65



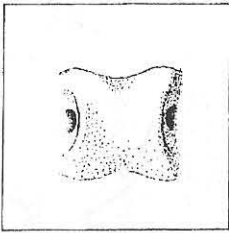
66



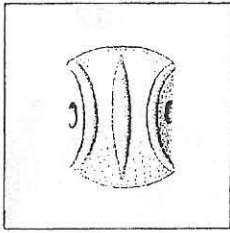
67



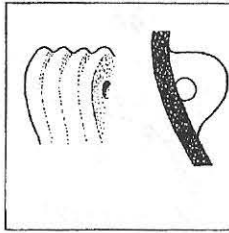
68



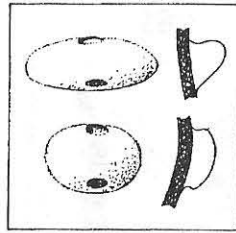
69



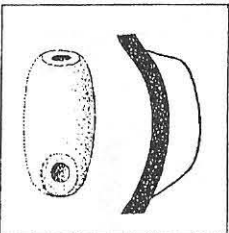
70



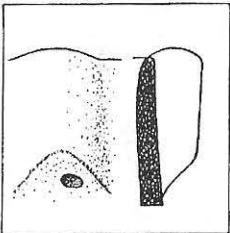
71



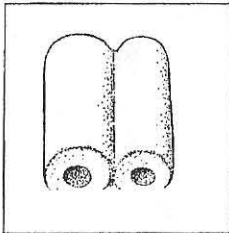
72



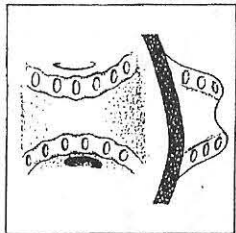
73



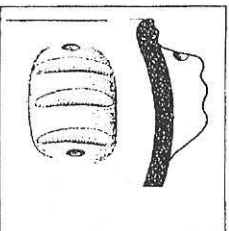
74



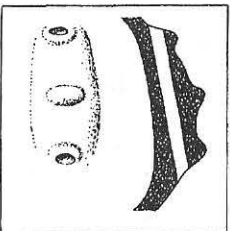
75



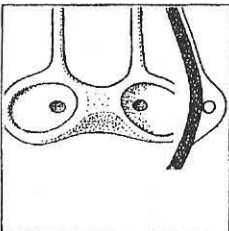
76



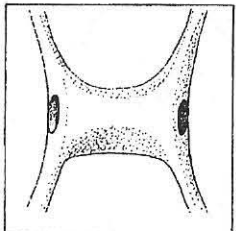
77



78

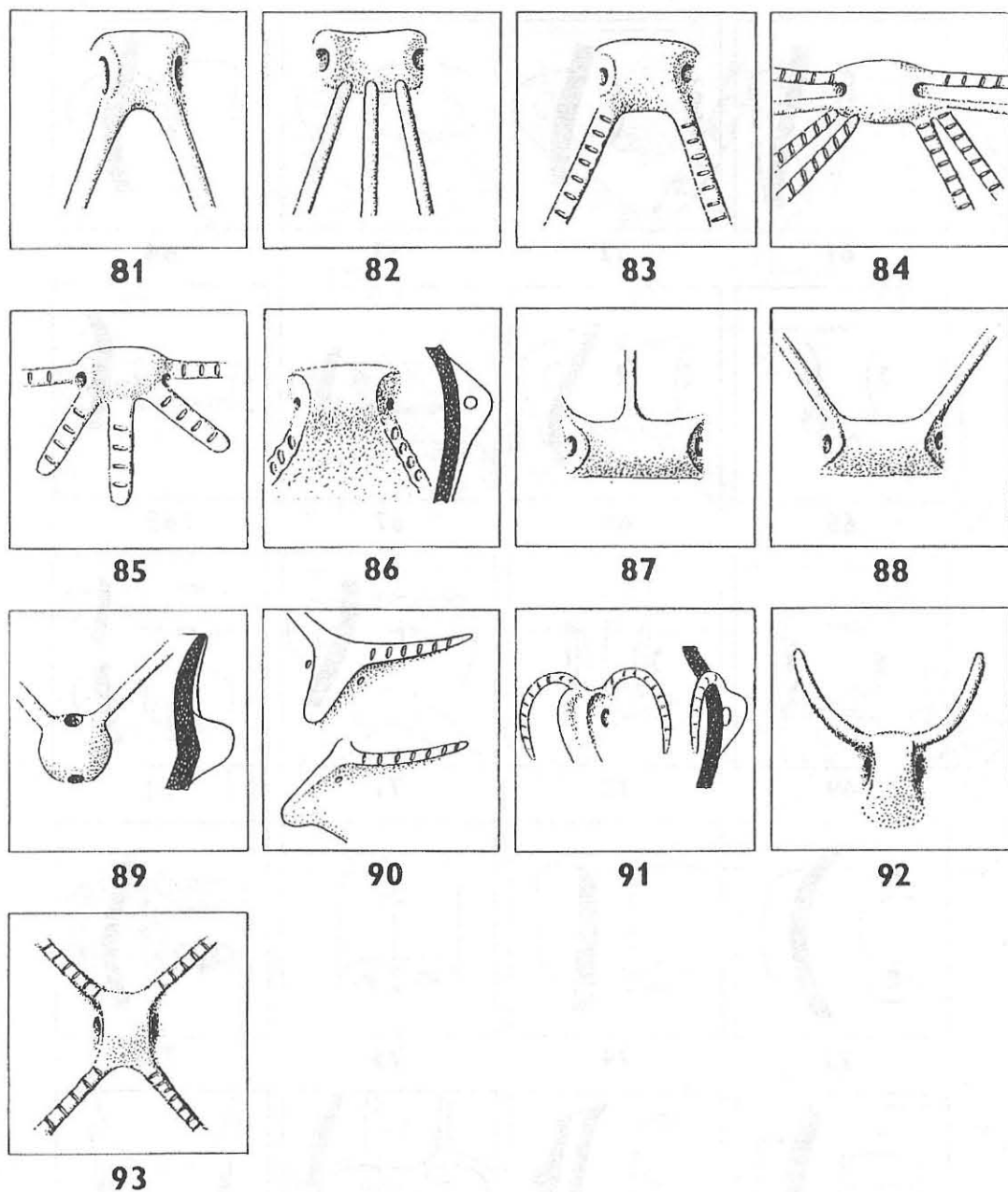


79



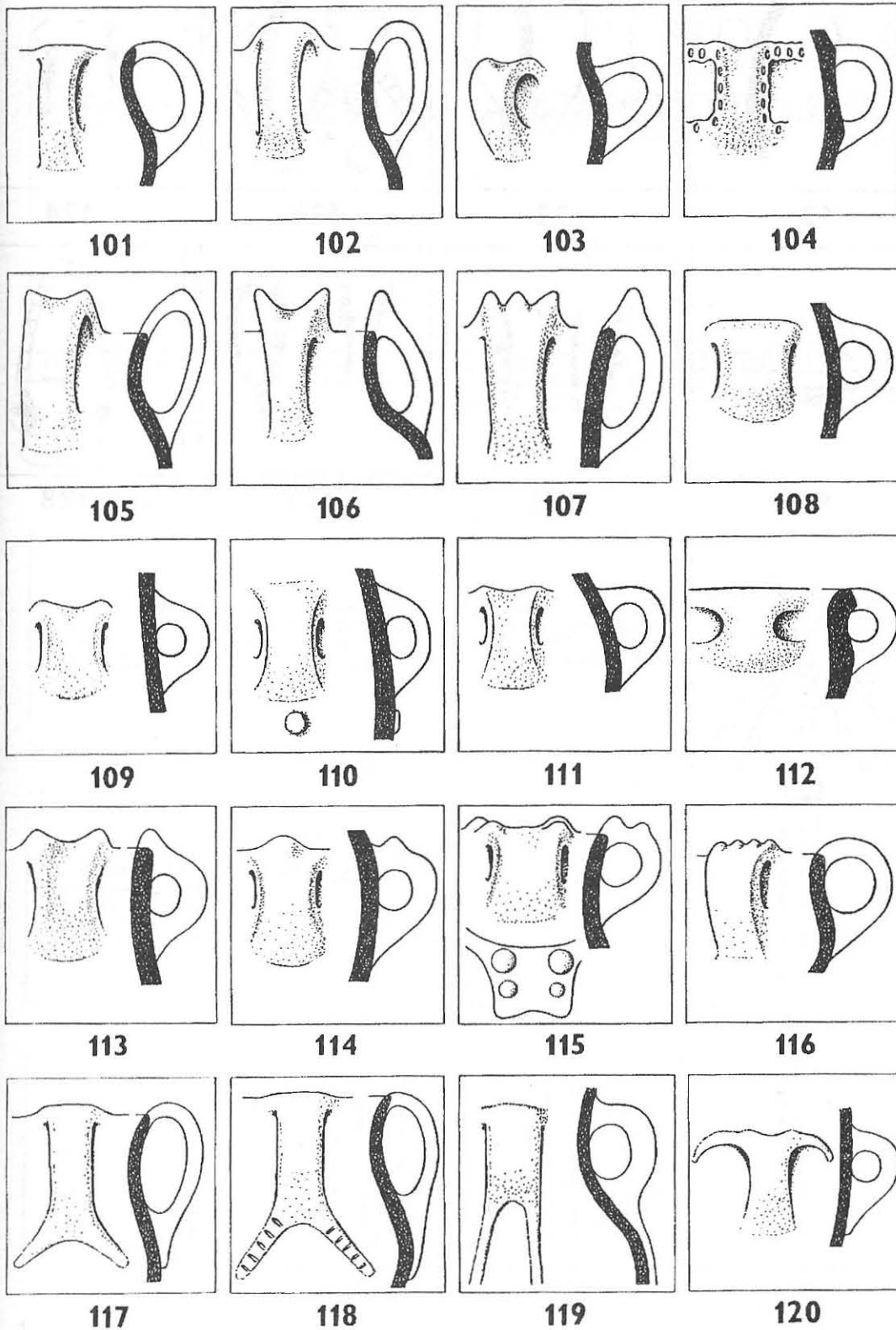
80

Obr. 16. Plastická výzdoba – provrtané výčnělky (61-80).

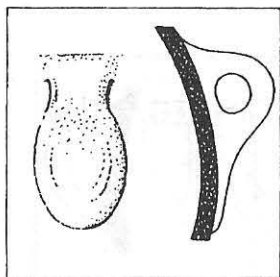


Obr. 17. Plastická výzdoba – provrtané výčnělky (81-93).

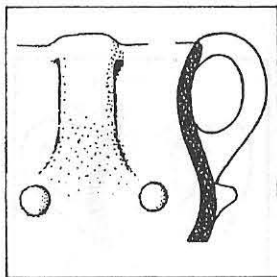
Ucha



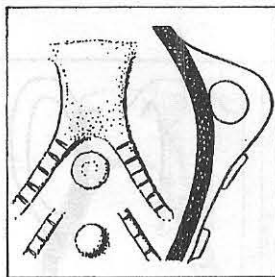
Obr. 18. Plastická výzdoba – ucha (101-120).



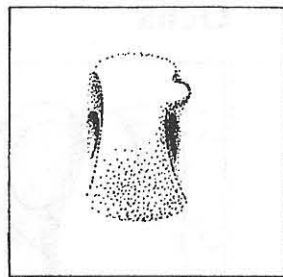
121



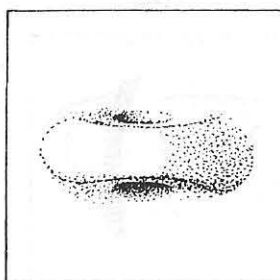
122



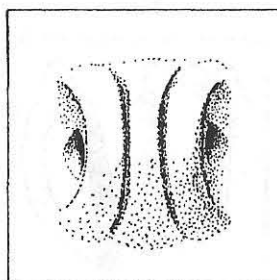
123



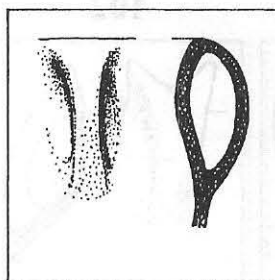
124



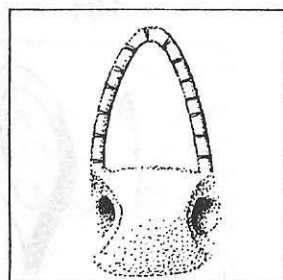
125



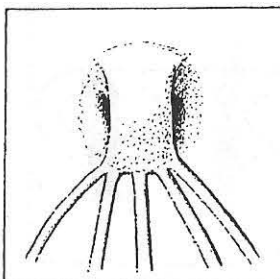
126



127



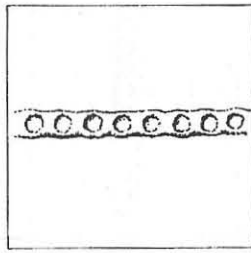
128



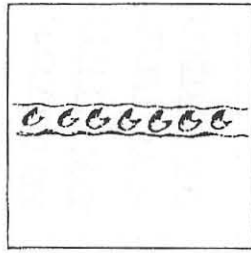
129

Obr. 19. Plastická výzdoba – ucha (121-129).

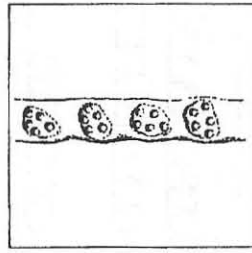
Lišty



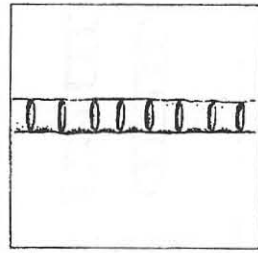
131



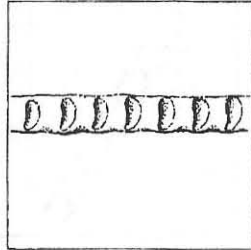
132



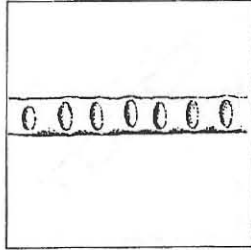
133



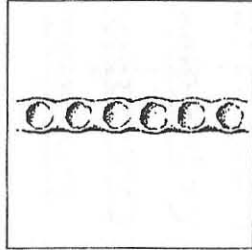
134



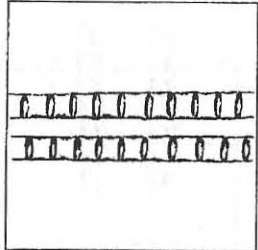
135



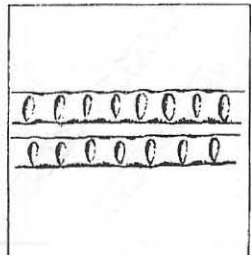
136



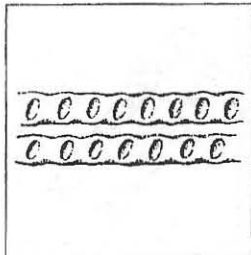
137



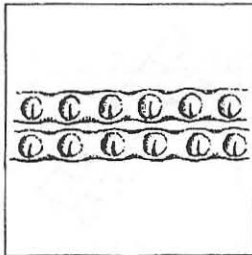
138



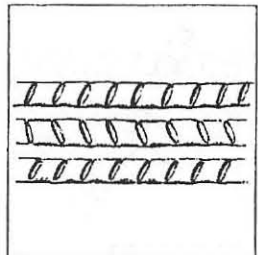
139



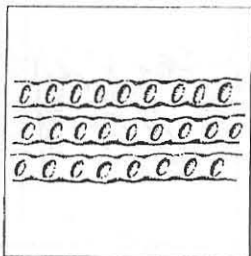
140



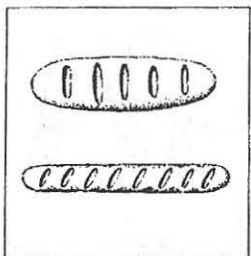
141



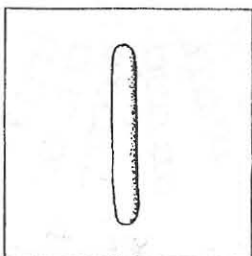
142



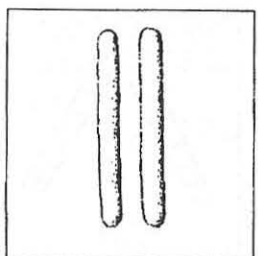
143



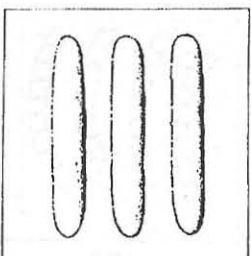
144



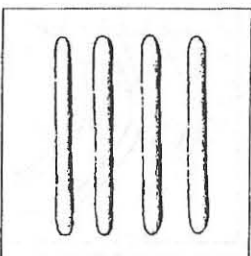
145



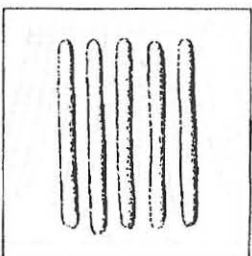
146



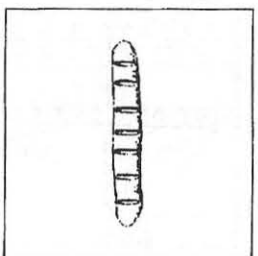
147



148

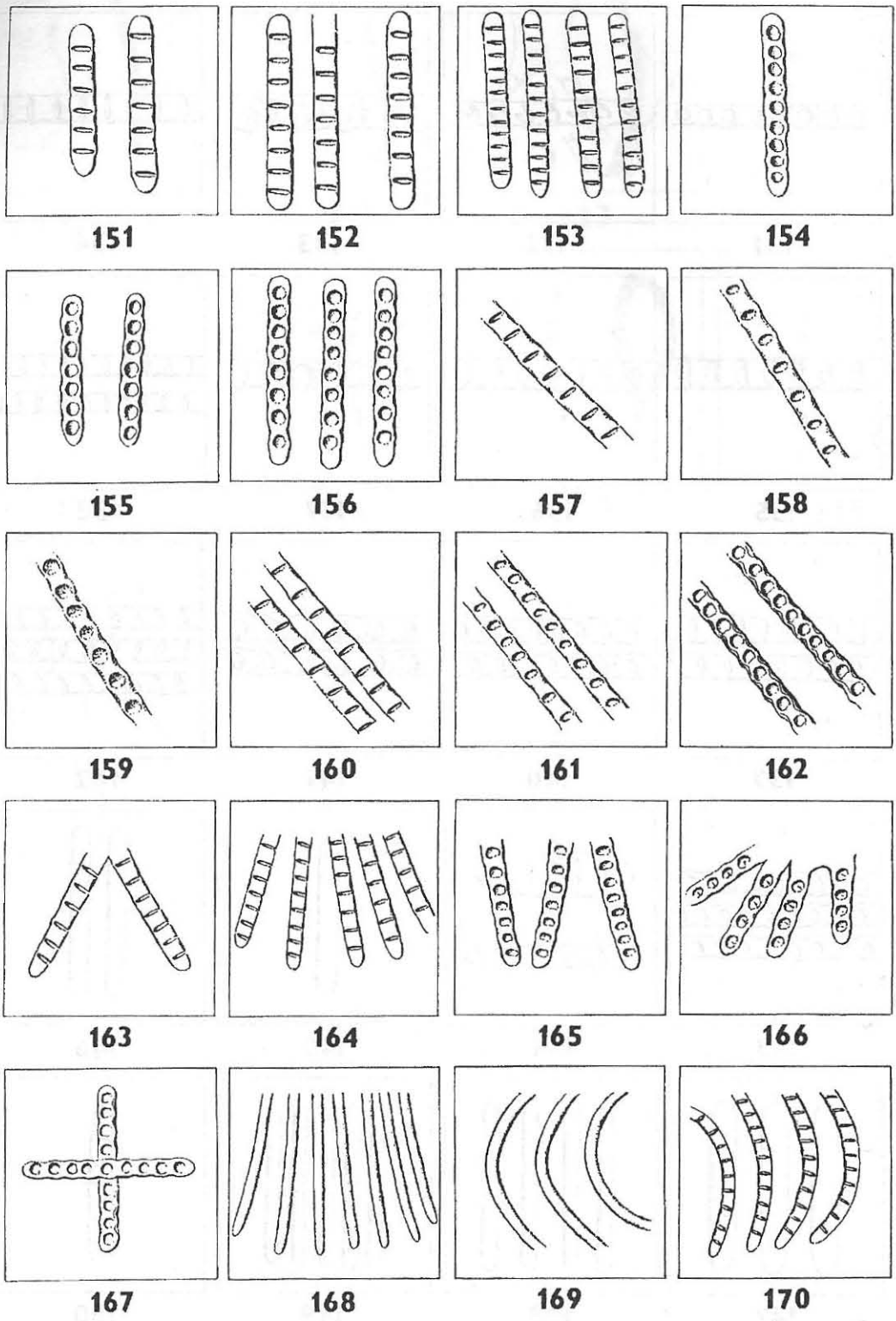


149

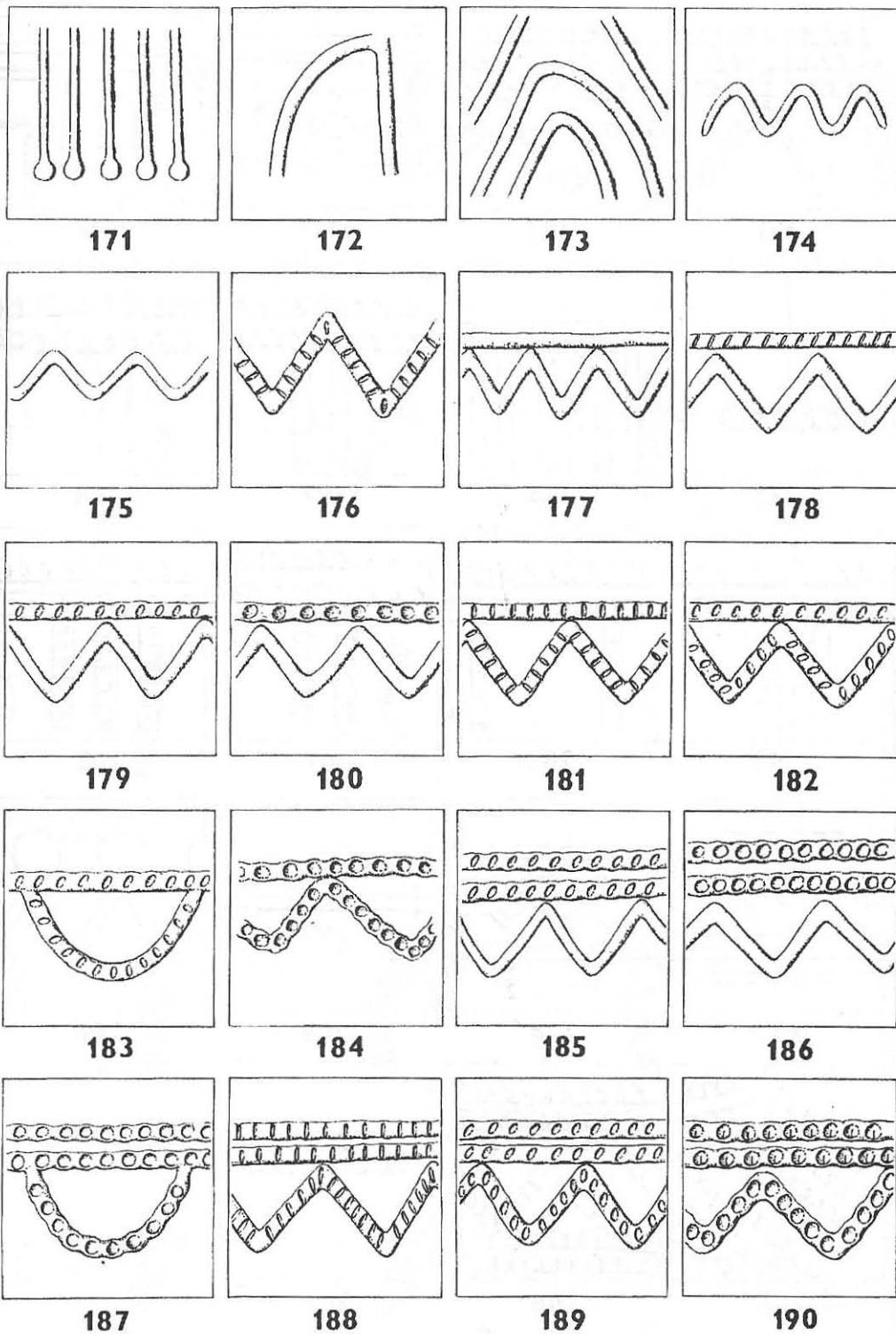


150

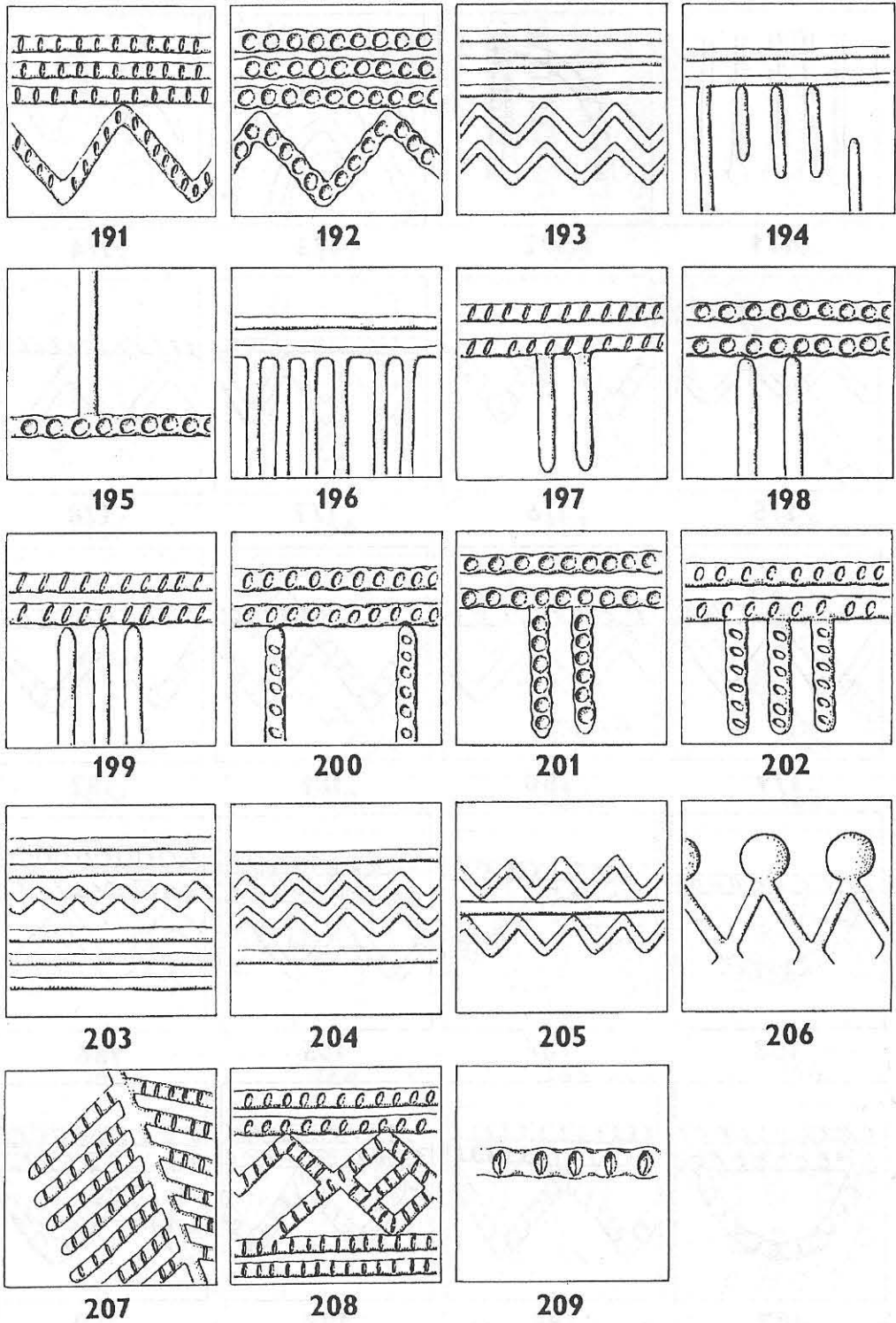
Obr. 20. Plastická výzdoba – lišty (131-150).



Obr. 21. Plastická výzdoba – lišty (151-170).

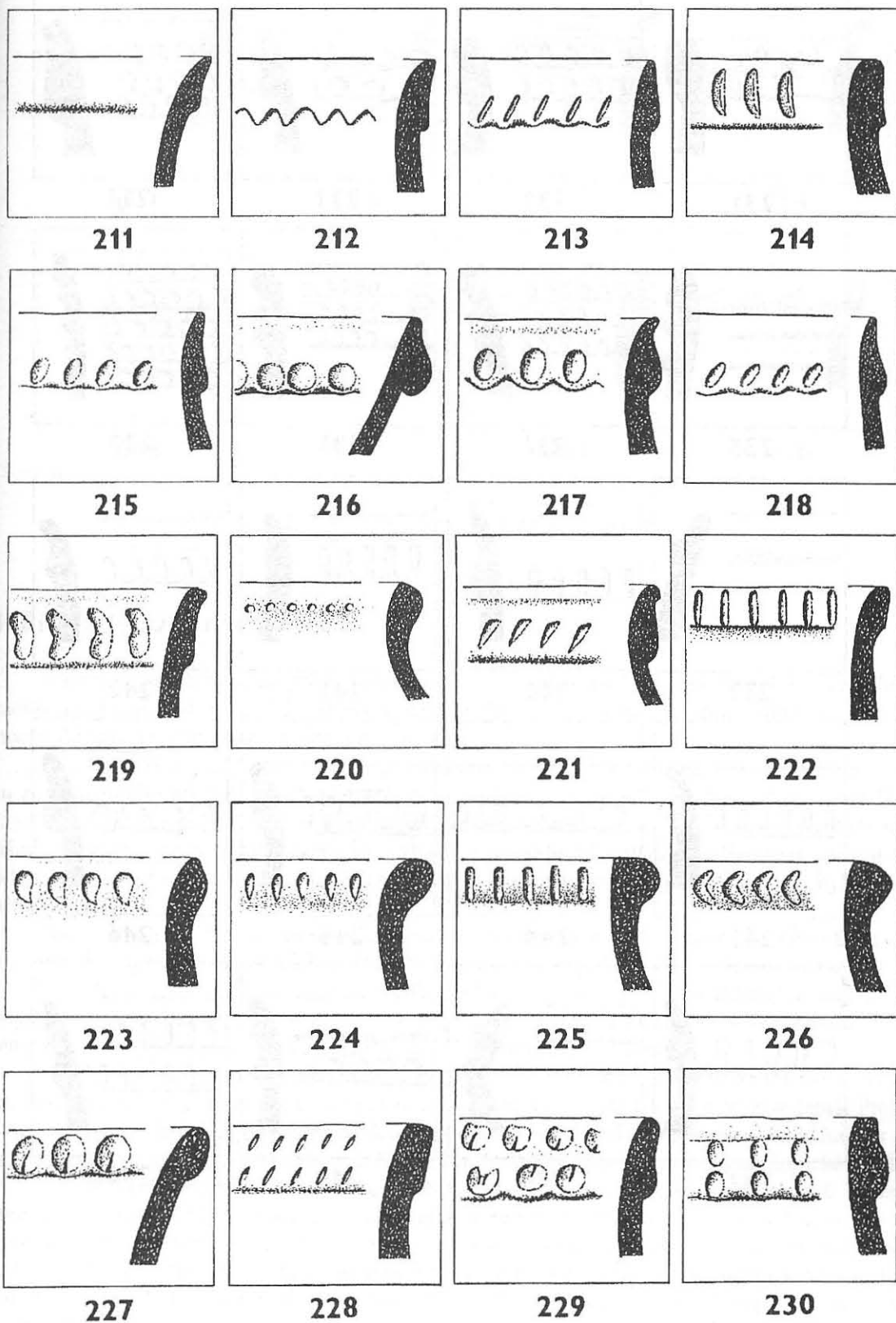


Obr. 22. Plastická výzdoba – lišty (171-190).

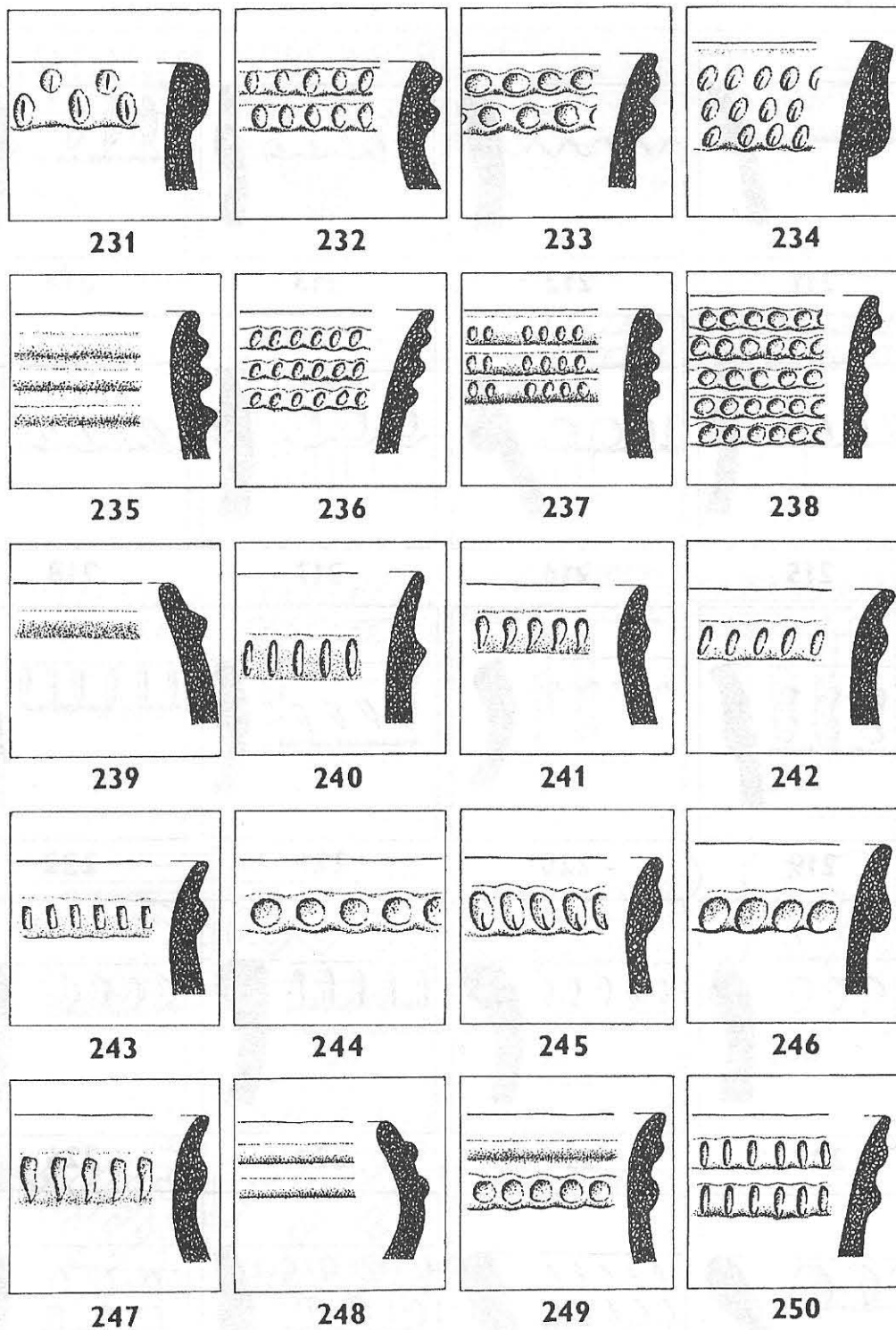


Obr. 23. Plastická výzdoba – lišty (191-209).

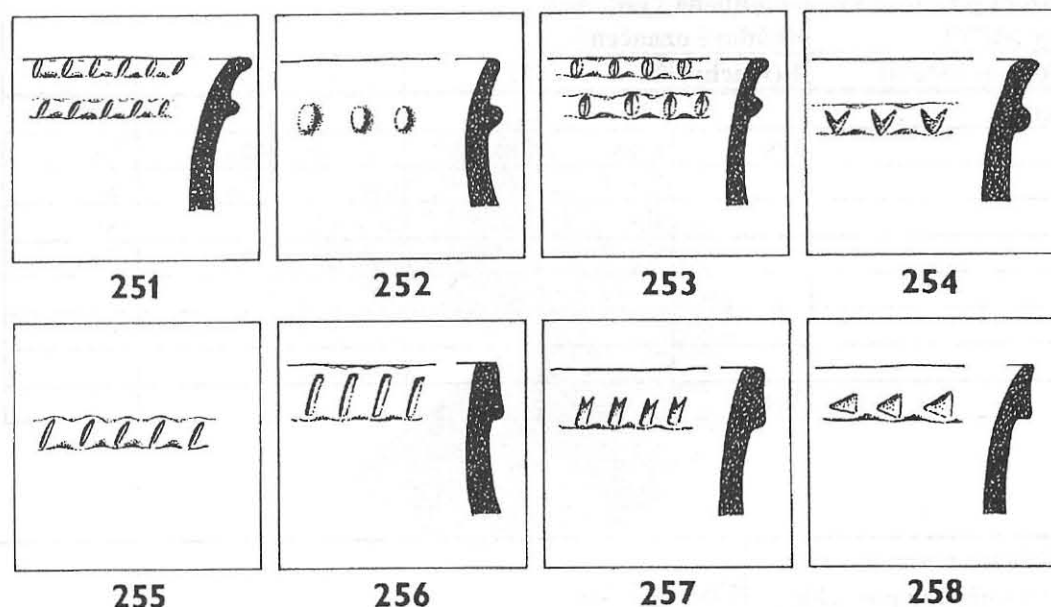
Okrajové lišty



Obr. 24. Plastická výzdoba – okrajové lišty (211-230).



Obr. 25. Plastická výzdoba – okrajové lišty (231-250).



Obr. 26. Plastická výzdoba – okrajové lišty (251-258).

5.4 Příklady užití databáze

V poslední kapitole si na příkladu programu MS Access ukážeme, jak lze informace z databáze moravské domácí eneolitické keramiky využít. V databázových programech vyhodnocujeme data pomocí vyhledávání, filtrování a pomocí dotazů.

K jednodušším patří filtrování podle jednoho kritéria. Chceme-li např. zjistit veškerou keramiku ze sídlištních jam stačí zadat do filtru, pracujícího v režimu tabulky, kód 10 pro pole DROB a ihned se nám zobrazí údaje o všech keramických fragmentech a celých nádobách, pocházejících ze sídlištních jam (pro tento i další úkoly lze vytvořit i jednoduchý dotaz). Použijeme-li operant NOT a kód 10 nahradíme kódem 03 (kulturní vrstva), zobrazí se nám všichni jedinci, kteří nebyli nalezeni v kulturní vrstvě.

Použijeme-li při filtraci operantu Or (nebo), můžeme zjistit údaje např. o všech okrajích, rekonstruovatelných nádobách a celých nádobách ze souboru (pro pole DOCH kritérium: 1 Or 8 Or 9).

Při filtrování lze kombinovat i různá pole. Zadáme-li např. v poli DOCH (stupeň dochovanosti exempláře) hodnotu 9 (tzn. zcela dochovaný exemplář) a v poli KT (keramická třída) hodnotu 18 (tzn. buben), získáme informace o všech zcela dochovaných bubnech z lokality.

K velmi mocným databázovým nástrojům patří tzv. křížové dotazy (dle terminologie MS Access). Používají se pro shrnutí dat v kompaktní formě, která je dobře čitelná a srozumitelná. Program MS Access nabízí pro jejich tvorbu průvodce, s jehož pomocí se jedná o relativně jednoduchou záležitost. V databázi eneolitické keramiky z Kramolína můžeme identifikovat archeologické struktury např. v podobě keramických tříd, které lze vydělit pomocí kombinací jednotlivých znaků keramiky. Výstupem z programu MS Access je např. tabulka, v jejímž levém sloupci jsou zachyceny všechny varianty umístění plastické výzdoby na nádobě a v prvním řádku jednotlivé druhy vhloubené výzdoby (tab. 1). V buňkách tabulky pak můžeme nalézt počty kombinací mezi jednotlivými kvalitami těchto znaků. Pokud nás zajímá dále např. zastoupení typů okrajů, ev. kombinovaných např. s úpravou vnějšího povrchu, a jejich výskyt v jednotlivých objektech, není problémem takový křížový dotaz vytvořit (tab. 2). V dotazu je možno kombinovat až čtyři znaky. Tyto výstupy lze vizualizovat pomocí různých typů grafů, a tak usnadnit jejich interpretaci (graf 1, 2).

Identifikací archeologických struktur ve zkoumaném materiálu naše práce nekončí, ale je to pouze prostředek pro další etapu poznávání pravěkých společností.

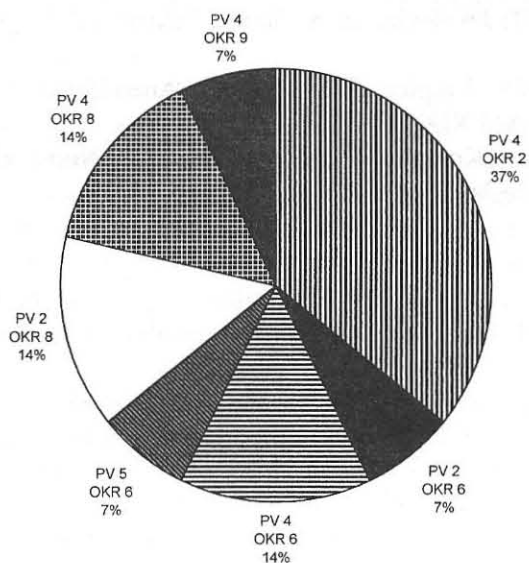
rozložení plastické výzdoby (RPV) (kódové označení)	vhloubená výzdoba (VV) (kódové označení)					
	1 (vpichy)	21	22	140	141
1 (okraj)					1ks
2	1ks	17ks	5ks	6ks	7ks
3		1ks		3ks	3ks
4		4ks			1ks
12				4ks	13ks
20					
23					1ks
24					
34	2ks				

Tab. 1.

Kombinace tvaru okraje a úpravy vnějšího povrchu		Číslo objektu						
OKR	PV	20	35	36	37	38
2 (zaoblený)	2 (hrubě m.)		5ks	2ks	1ks	1ks
2	4	5ks	8ks		8ks	1ks
3	4					1ks
4	4			1ks		
4	5		1ks			
5	2		1ks			
5	4			1ks	1ks	
6	2	1ks	2ks			
6	4	2ks	6ks			
6	5	1ks				
8	1		1ks			
8	2	2ks	1ks	2ks	1ks	4ks
8	4	2ks	6ks		1ks	4ks
8	5		1ks			
9	2		5ks	1ks	1ks	
9	4	1ks	9ks		2ks	

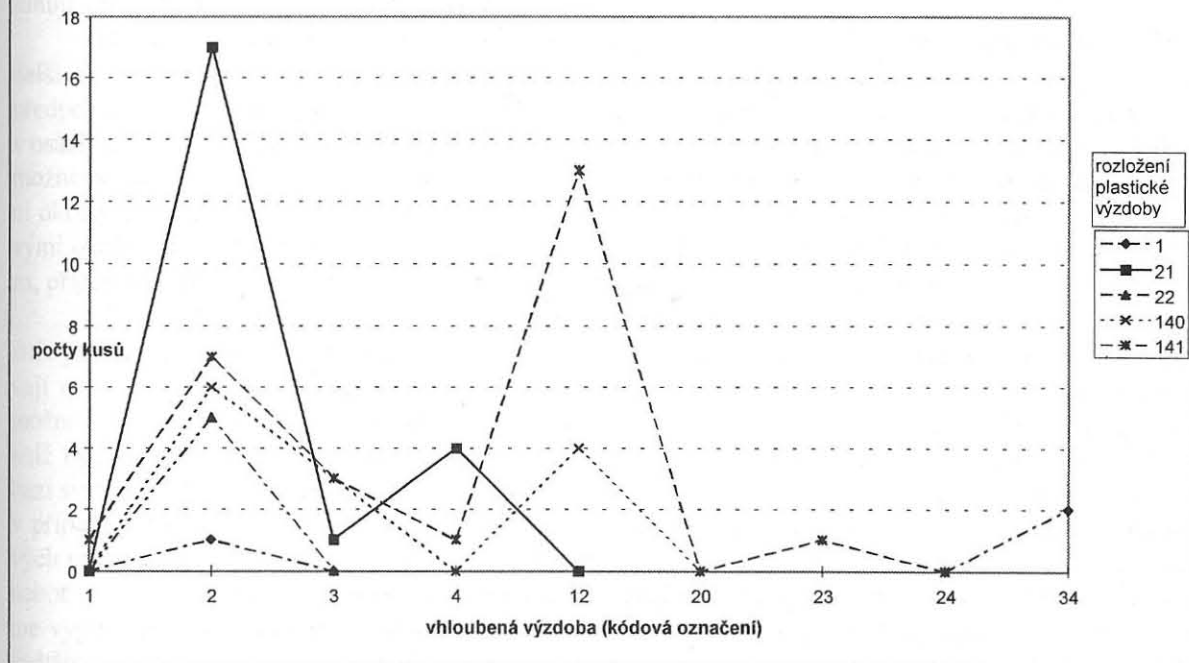
Tab. 2.

Objekt 20 - Kombinace tvaru okraje a úpravy vnějšího povrchu v %



Graf 1.

Závislost některých variant vhloubené výzdoby na rozložení plastické výzdoby



Graf 2.

5.5 Literatura

- Medunová-Benešová, A. 1972: Jevišovice-Starý Zámek. Schicht B – Katalog der Funde, FAM VI, Brno.
- Medunová-Benešová, A. 1973: Grešlové Mýto. Äneolithische Höhensiedlung „Nad Mírovcem“. Katalog der Funde, FAM VII, Brno.

- Medunová-Benešová, A. 1977: „Pallardiho Hradisko“. Eine Äneolithische Höhensiedlung bei Vysočany, Bez. Znojmo. Katalog der Funde. FAM IX, Brno.
- Medunová-Benešová, A. 1981: Jevišovice-starý Zámek. Schicht C2, C1, C. Katalog der Funde. FAM XIII, Brno.
- Medunová-Benešová, A. 1986: Křepice, Bez. Znojmo. Äneolithische Höhensiedlung „Hradisko“. Katalog der Funde. FAM XIX, Brno.
- Podborský, V. – Kazdová, E. – Košťuřík, P. – Weber, Z. 1977: Numerický kód moravské malované keramiky. Problémy deskripce v archeologii, Brno.



6. Výpočetní technika a zpracování kamenné štípané industrie (Petr Neruda)

Využití výpočetní techniky pro analýzu kamenné štípané industrie je v podstatě stejné jako u ostatních druhů archeologického materiálu. Snahou je získat porovnatelná data podle jednotného systému a za použití databázových a tabulkových programů provést jejich analýzu směřující k vytvoření modelu výroby (operačního schématu), distribučního modelu apod. Nejvíce změn, kterým podlehnou kamenné štípané objekty během různých stádií své existence je zapsáno ve formě několika atributů. Jejich řádné poznání a interpretace je prostředkem sloužícím ke konstrukci technologických rekonstrukcí a záznamů objektů, které jsou založeny na statistické nebo distribuční analýze kamenných souborů (Schild 1980, 57).

Jisté specifické rysy oproti ostatnímu archeologickému materiálu zde však existují a projevují se už v celkovém charakteru popisovaného souboru a často v jeho nálezových okolnostech. Kusy jsou klasifikovány v souvislosti s jejich objevením se během procesu jádrové a nástrojové preparace, exploatace a remodelace. Při vytváření klasifikačního systému je třeba si uvědomit, že většina morfologických údajů, které jsme schopni popsat je aspoň do určité míry náhodná. Je sice možné vyrobit více méně stejný kamenný broušený, keramický či bronzový předmět, ale je téměř nemožné vyrobit sérii metricky a do značné míry i morfologicky stejných štípaných předmětů. U nástrojů je sice tvarová podobnost větší a lépe srovnatelná, ale i tak zde existuje určitá variabilita pro jeden typ nástroje, se kterou je třeba počítat. Je tedy nutné vytvořit si nejen určitý vůdčí tvar (typ), ale i jeho tvarový rozptyl. I s touto definiční variabilitou však musíme počítat s určitou chybou (nejednoznačné zařazení), která se však většinou v celkovém zpracování a porovnání ztrácí, neboť kolekce štípané industrie obsahují většinou tisíce inventárních čísel. 1\

Druhým a často ne nepodstatným problémem jsou nálezové okolnosti, které predeterminují další možnosti zpracování směřující k vytvoření speciálních sídlištních modelů na nalezišti. To však předpokládá solidní archeologický výzkum s přesnou dokumentací jednotlivých nalezených předmětů v osách x/y, pokud možno s přesnou výškovou nivelací korelovanou se stratigrafickým sledem. Pak je možné po provedení remontáží a analyzování industrie rozčleňovat plochu sídliště na jednotlivé vnitřní okrsky charakterizované určitou činností. Na Moravě však většinou disponujeme takovými nálezovými okolnostmi, že většinou přichází v úvahu pouze technologická, typologická a surovinová analýza, přičemž vnitřní členění plochy podle zájmových aktivit nám v detailech uniká.

Z hlediska statistické metody zpracování je výhodná ta skutečnost, že kolekce štípané industrie jsou většinou dosti rozsáhlé, takže případné chyby v determinaci se v celkovém zpracování stávají méně výrazné. I to je jeden z důvodů, proč využít výpočetní techniku a rozsáhlé programové možnosti, které se nám dnes naskýtají, zvláště můžeme-li přenášet data mezi jednotlivými programy, aniž bychom je museli znova přepisovat (databáze, tabulkový procesor, grafické programy např. na bázi systému GIS apod.). Je však účelné rozložit celkovou analýzu na jednotlivé komponenty, které je v případě nutnosti možné relačně spojovat (v databázích dBase IV, Paradox, Acces apod., tabulkových procesorů: Excel, Lotus 1-2-3 apod.). Spojením všech znaků do jednoho formuláře není žádoucí, neboť ne na všech předmětech lze sledovat všechny znaky (u databáze nepoužité na nástroje nemůžeme vyplnit položky týkající se nástroje) a množství nevyplněných položek by pak soubor neúměrně zvětšovalo. Navíc rozdělením analýzy do více složek můžeme vyplnit jen ty formuláře, které potřebujeme pro danou hladinu zhodnocení industrie. Zde máme na mysli zejména speciální nástrojové analýzy, které můžeme provádět dodatečně a pro celkovou analýzu kolekce je relačně propojíme s vyššími stupni analýz (nástrojová a technologická).

6.1 Rozdělení celkové analýzy

Základní zpracování kolekce štípané industrie je založeno na analýze souboru jader, technologické a nástrojové analýze databáze. Jsou sice hodnoceny odděleně, ale celkový obraz popisované

industriie je skládán právě z těchto separátních částí, které samy o sobě bez vzájemné vazby nevytvářejí dostatečně objektivní poznání. Může se totiž stát, že úštěp s vícesměrnými jizvami na dorsální straně, který patří ke skupině úštěpů oddělených z jádra se změněnou orientací, může být totálně mylný v souboru, který náleží taxonomické jednotce bez takových jáder. Na druhou stranu, některé z technik jádrové preparace produkují jádra s více směrovými jizvami na jejich dorsální straně. Proto je nutné, po dílčím vyhodnocení jader a debitáže, navzájem korelovat získané výsledky a analyzovat případné chyby. Celý proces pak připomíná cyklické střídání vyhodnocování a korelování výsledků zpětně k datovým podkladům. Jádrová a debitážová determinace by nám měla dostatečně charakterizovat sledovanou industrii tak, aby bylo možné vytvořit rekonstrukci operačních schémat a distribučních modelů. Pak mohou následovat speciální analýzy dílčích znaků (analýzy drasadel, rydel, hrotů apod.). Všechny soubory, které zpracovávají data o debitáži jsou pak navzájem propojitelné v relační databázi a tím dochází k opakování pouze jedné položky, která slouží k relačnímu spojení (např. inventární číslo).

Důležitým faktorem prokřížení jsou pak použité suroviny, které mohou či nemusí ovlivnit celkovou skladbu industriie. Důležitější než druh výchozího materiálu je často jeho forma, daná tvarem, zrnitostí materiálu a obsahem inhomogenit a kazů. Je to právě často surovina, která si vyžádá změnu či přizpůsobení technologie jako formu řešení individuálních surovinových kvalitativních nedostatků.

Při vytváření popisného systému se bohužel asi nevyhneme některým subjektivním faktorům (daným stavem poznání), které jsou jistým protikladem k objektivním (neutrálním) typům dat, které by měly aspoň převažovat. Typickým příkladem je zařazení produktu do systému výrobního procesu a typologická klasifikace nástrojů. Oba systémy se totiž mění podle stáří klasifikované industriie. Markantní je tento problém mezi středem a mladopaleolitickými kolekcemi. Například pro taubachien, ale i ostatní středopaleolitické kultury, je obtížné použít mladopaleolitický systém produkce. Stejně tak seznamy typů jsou vzhledem ke skladbě industriie odlišné. Aspoň pro technologickou analýzu máme dvě možnosti řešení: vytvořit nový determinační kód anebo upustit od zařazování produktů do systému produkce a popisovat polotovary pouze morfologicky. Druhý přístup by byl jistě neutrálnější a použitelný pro všechny typy kolekcí, ale bylo by nezbytné, zavést pro dostatečný popis morfologie ještě další atributy: počet negativů na dorsální ploše, jejich vzájemná orientace a procentuální zastoupení kůry. Naskytá se však otázka, zda tím nevytvoříme přehřel nejrůznějších kombinací znaků, ale jednoznačnější determinaci stádií výrobních procesů bude obtížné použít. Tento systém bude v budoucnu aplikován na střední paleolit a budou testovány jeho výpovědní hodnoty.

Typologické zařazení náleží je problém, který je možné řešit podobně. V poslední době se od přesné typologické klasifikace upouští a zpracování nástrojů probíhá pouze v intencích několika nástrojových tříd. Vlastní morfologie nástroje, svědčící o způsobu výroby a jeho případných reutilizacích, je pak pojata jako kombinování dílčích objektivních znaků (opět problém vyhodnocení) a typologické číslo slouží pouze k rychlé orientaci či k jednoduché prezentaci nástrojové skladby. Není nutné ho však jednoznačně zavrhnout, neboť své výsledky v minulosti již přinesl, ale je asi nezbytné provést jeho korelaci s novým přístupem.

Jednou z nejobtížnějších částí je vlastní syntéza nasbíraných dat. Vzhledem k jejich možnostem je žádoucí provádět rekonstrukce sledovaných procesů pokud možno multiaspektuálně. Jako příklad můžeme uvést určení operačních schémat, které pak můžeme rozdělit i podle surovin, které jsou aspoň v kolekcích se štípanou industrií jedním ze základních prvků vícevrstevné analýzy. Je pak na badateli, aby určil, do jaké míry jsou tyto úrovně podobné a do jaké míry jsou samostatné nebo slučitelné. Samozřejmě, že takové úrovně je možné vytvořit i z jiných dat, ale čím více jich určíme, tím obtížnější je celkové vyhodnocení a prezentace. Někdy je snaha po maximální multiaspektualitě spíše nevhodná, protože osvětlit takový komplikovaný obraz je nesmírně obtížné a někdy zřejmě i nemožné.

6.1.1 Analýza jader

Pro rekonstrukci operačního schématu je zpracování jader jednou z nejdůležitějších složek hodnocení industriie. Hlavním rysem, který tvoří jeho základ je způsob redukce masы výchozí suroviny, tj. způsob její exploatace střídáním preparačních, těžebních a repreparačních výrobních postupů. Prvním krokem je zařazení posuzovaného kusu do systému produkce. Narážíme však na problém, ve kterém stádiu výrobního procesu se nalezená jádra nacházejí a proč byla vlastně opuštěna či zahozena.

Zanechání prejadra před úpravou a exploatací, stejně jako odhození těžce vytěženého jádra je zdrojem informací o procedurách a technikách preparace jader, jejich reparacích a těžbě. Exploatace a často i násilné přetvarování jádra zpravidla zachová některé stopy jejich předcházejícího tvaru, které nadále existují mezi stopami z posledních stadií exploatace (Schild 1980, 58). Tvar jádra nebo způsob exploatace se v průběhu těžby mění, tzn. že jádra se změnou orientací jsou výsledkem změny orientace po té, co původní těžní plocha a podstava již neodpovídaly žádaným technickým parametrům. Tato skutečnost je dokazována tím, že tento typ jader se vyskytuje v mladopaleolitických kolekcích u kusů, které jsou poměrně malé, často s nereparovatelnými vadami v jedné z ploch, tj. ve stádiu vytěženosti. Z toho také vyplývá, že soubor jader by měl být posuzován spíše jako celek, než jako heterogenní soubor individuálních technologických jevů. Je nutné pokusit se ho rozčlenit na jednotlivá stadia výrobního procesu (mohla být opuštěna v různých stádiích) a pak posoudit, kdy se objevují určité jevy a s čím souvisí. Pak lze snáze rozlišit, které znaky mají obecnější platnost a které jsou řešením individuálních problémů. Při pouhém typologickém rozčlenění a klasifikaci jednotlivých znaků bychom mohli mechanicky prohlásit, že soubor obsahuje např. 40 jednopodstavových, 50 dvoupodstavových jader a 36 kusů se změnou orientací. To samo o sobě je pouze pozitivistické konstatování stavu věci bez určení vztahů a příčin jevů.

Je tedy nezbytné položit si soubor otázek, jejichž zodpovězení nás nutí přistoupit ke klasifikaci z obecnějšího hlediska s určením modelu a zákonitostí. Ukazuje se, že mezi nejdůležitější z nich patří: byl typ jádra determinován jakostí suroviny, jejími rozměry nebo vyplýval z tradice zpracování bez ohledu na jiné faktory? Nejedná se o kombinaci obou? Kolik a které technologické postupy byly využívány a čím se řídila jejich aplikace? Srovnání souboru jader z tohoto hlediska je jistě více signifikantní, než pouhé kvantitativní srovnání jednotlivých typů jader a některých jejich úprav.

Pro analýzu jader byl použit pouze jeden formulář, neboť téměř u všech kusů je možné vyplnit dané položky a tudíž je daný formulář schopen ekonomicky a dostatečně popsat analyzovaný soubor. Návrh atributů, které lze na jádrech sledovat, je vyobrazen na formuláři (formulář 1). 2\

6.1.2 Analýza odbitých kusů

U debitáže (ústěpů, čepelí a na nich zhotovených nástrojů) jsme postaveni před poněkud odlišný problém, než je tomu u jader. Datové možnosti jsou u jednotlivých kusů rozdílné. Nejmarkantnější tento problém vynikne v nesourodosti dat, které jsme schopni sledovat u debitáže nevyužitých nástrojů a u nástrojů. V případě, že byl polotovár použit pro nástroj, stírají se technologické znaky na předmětu (retuši), ale přibývají kvalitativní data o vlastním nástroji. Dále je možné provádět hodnocení některých speciálních znaků (např. u rydel), které se vyskytují pouze v některých nástrojových třídách a v jiných ne. Při použití jednoho formuláře by docházelo nejen ke zbytečným ztrátám paměti počítače a záznamových médií, ale i ke ztrátám časovým při vyplňování údajů. Soubor několika formulářů je pak postaven na systému analýzy od nejobecnějšího k nejspeciálnějšímu. U všech kusů je možné určit technologická data (zařazení do systému produkce), klasifikovat suroviny, stav zachování a určit jejich metriku, tzn. všechna data společná všem kusům debitáže. Následuje nástrojová analýza, která již neobsahuje výše uvedená data a obsahuje informace o vlastním nástroji, tj. rozměry nástrojových částí, jejich umístění, charakter, metriku apod. Jestliže chceme dále analyzovat určité nástrojové třídy, pak je možné vytvořit formuláře na ně specializované (dále bude uveden případ pro zpracování rydel).

6.1.2.1 Technologická analýza

Jak již bylo výše uvedeno, představuje formulář technologické analýzy nejzákladnější zpracování debitáže (formulář 2). Její význam spočívá ve skutečnosti, že zde mohou být hodnoceny i ty kusy industrie, které nejsou zpracovány jako nástroje, tedy údaje, které byly při obvyklém typologickém přístupu ztraceny nebo nedostatečně zhodnoceny. Přitom pro hodnocení operačního schématu je nenástrojová debitáž vhodnější než nástroje právě z hlediska zachovaných znaků z exploatace jader. Tato část kolekce kamenné industrie je důležitá mimo jiné i pro určení případné funkce sídliště nebo jeho části. Nástroje jsou v této úrovni hodnoceny jen z hlediska použitých polotovarů (technologie)

a jejich typologické aspekty jsou ve formuláři zhodnoceny pouze číselným označením typu nástroje (tato položka však může být vypuštěna), a v podrobnější variantě jsou pak relačně dostupné ve formuláři nástrojové analýzy. V této části zpracování industrie narážíme asi na nejvíc problémů při její klasifikaci. Oproti typologické klasifikaci, která je poměrně jednoznačná, je zařazení polotovaru k určitému výsledku výrobního procesu častěji problematické. Někdy se totiž na jednom předmětu objeví i více znaků patřící různým produktům výroby a je pak zřejmě nejlepší, pokusit se určit ten, který těsně předcházela vlastnímu odbití. Jako příklad může sloužit remontáž z lokality Hošťálkovice-Hladový vrch, kde se tento jev projevil hned u několika kusů (Neruda 1995). I tak je nutné počítat s určitou chybou, ale vzájemné poměry mezi produkty z preparačního, těžebního a reparačního stádia výrobního procesu budou při velkých objemech dat dostatečně signifikantní.

Ne zcela jednoznačné je i určení suroviny. Zde je třeba počítat předem s určitými predeterminanty. Asi nejvýznamnějším je u paleolitických kolekcí přítomnost patiny, která někdy zcela překrývá povrch předmětu a tím znesnadňuje až znemožňuje použití makroskopického i mikroskopického určení bez destrukce předmětu. Navíc se zdá, že makroskopické určování surovin je dnes již dosti nedostatečné, zvláště vezmeme-li velkou variabilitu, kterou některé druhy suroviny mají (rohovec typu Stránská skála, moravské jurské rohovce). Přínosné by bylo i podrobnější petrografické rozdělení některých variabilnějších rohovců (např. moravských jurských).

S poměrně velkou chybou lze prozatím počítat při určování techniky odbití, tj. určení použitého otloukače. Existuje sice základní definice použití měkkého a tvrdého otloukače, ale po uplatnění zůstává dosti nejednoznačně klasifikovatelných kusů. Osvědčilo se vykazovat tyto předměty jako neurčené a pracovat pouze s těmi, které se z hlediska určení zdály jednoznačnější. V kapitole o konkrétní aplikaci si ukážeme, že výsledky získané z této části analýzy zcela korespondovaly s výsledky získanými z jiných znaků, a že tudíž mají svou vypovídací hodnotu.

6.1.2.2 Nástrojová analýza

Tato analýza je zaměřena na získání základních údajů o nástrojových znacích popisovaného předmětu, které lze rozdělit na složku typologickou, metrickou a morfologickou, popisující tvar a následně i způsob výroby nástroje. Popis těchto znaků je normalizován do 8 sektorů a na ventrální a dorzální plochu. Tím je zabezpečena souměřitelnost popisu znaků a jejich lokalizace (formulář 3).

6.1.2.3 Speciální nástrojová analýza

Zpracovaný formulář (formulář 4 a 5) je určen k detailnímu rozboru rydel a rydlových odštěpů, ale podobné by bylo možné vytvořit i pro ostatní nástrojové třídy. Výsledky takových zpracování nástrojů by měly zodpovídat otázky způsobů výroby s výrobními finesami a reutilizace nástrojů. V případě potřeby informací o použitých polotovarech je možné použít pomoci relačního spojení i technologické a nástrojové formuláře.

Formulář pro zpracování jader	
Inventární číslo <input type="text"/>	Poloha <input type="text"/>
Výrobní proces	
Stádium <input type="checkbox"/>	Výsledek <input type="checkbox"/>
Forma suroviny <input type="checkbox"/>	Typ jádra <input type="checkbox"/>
Úprava podstavy <input type="checkbox"/>	Preparace <input type="checkbox"/>
Tvar příčného průřezu <input type="checkbox"/>	
Tvar podélného průřezu <input type="checkbox"/>	
Tvar jádra <input type="checkbox"/>	
Metrika	
Délka <input type="text"/>	Poškození <input type="checkbox"/>
Šířka <input type="text"/>	
Tloušťka <input type="text"/>	
Surovina <input type="checkbox"/>	Kvalita <input type="checkbox"/>
Poznámka	

Formulář 1.

Technologický formulář	
Inventární číslo <input type="text"/>	Poloha <input type="text"/>
Výrobní proces	
Stádium <input type="checkbox"/>	Výsledek <input type="checkbox"/>
Počet negativů <input type="checkbox"/>	Orientace negativů <input type="checkbox"/>
Technika výroby	
Talon <input type="checkbox"/>	Abráze <input type="checkbox"/>
Průřez <input type="checkbox"/>	Otloukač <input type="checkbox"/>
Lateralita <input type="checkbox"/>	Osost <input type="checkbox"/>
Metrika	
Délka <input type="text"/>	Poškození <input type="checkbox"/>
Šířka <input type="text"/>	Zachovalá část <input type="checkbox"/>
Tloušťka <input type="text"/>	
Surovina <input type="checkbox"/>	Kvalita <input type="checkbox"/>
Poznámka	

Formulář 2.

Nástrojový formulář									
Inventární číslo <input type="text"/>	Poloha <input type="text"/>								
Typ <input type="checkbox"/>	Podtyp <input type="checkbox"/>								
Retuš									
Dorsální	1 2 3 4 5 6 7 8								
Tvarování	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Rozsah	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Sklon	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Morfologie	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Ventrální	1 2 3 4 5 6 7 8								
Tvarování	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Rozsah	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Sklon	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Morfologie	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
D1 <input type="text"/>	Š1 <input type="text"/>								
D2 <input type="text"/>	Š2 <input type="text"/>								
T1 <input type="text"/>	T2 <input type="text"/>								
Poznámka									

Formulář 3.

Specializovaný formulář - rydla					
Inventární číslo <input type="text"/>					
Počet rydel <input type="checkbox"/>					
Morfologie rydel					
	1 2 3 4				
Podstava rydla	<table border="1"><tr><td></td><td></td><td></td><td></td></tr></table>				
Sklon rydlové plochy	<table border="1"><tr><td></td><td></td><td></td><td></td></tr></table>				
Obnovení rydla	<table border="1"><tr><td></td><td></td><td></td><td></td></tr></table>				
Preparace	<table border="1"><tr><td></td><td></td><td></td><td></td></tr></table>				
Pořadí	<table border="1"><tr><td></td><td></td><td></td><td></td></tr></table>				
Poznámka					

Formulář 4.

Specializovaný formulář - rydlivé odštěpy	
Inventární číslo	<input type="text"/>
Pořadí vzniku	<input type="checkbox"/>
Preparace hrany	<input type="checkbox"/>
Způsob vyběhnutí	<input type="checkbox"/>
Podstava	<input type="checkbox"/>
Poznámka	

Formulář 5.

6.2 Konkrétní aplikace

Výsledky analýz mohou sledovat různé směry zájmů, tzn. že můžeme s nasbíranými daty nakládat podle různých zájmových sfér. Sledovanému záměru je pak podřízen výběr nasbíraných údajů a výběr metod analýzy. Většinou jsou archeologické situace a nálezy aspoň do určité míry jedinečné, a tak je obtížné vytvořit zcela pevný způsob zpracování. Je však velice výhodné zhodnotit datové možnosti nálezů a určit si předem, které analýzy a způsoby prezentace budeme volit pro sledovaný soubor, čímž si ulehčíme práci a vyhneme se tak zbytečným analýzám, které stanovený problém neřeší.

A) Jednou z nejčastějších výsledků, ke kterým zpracováním kamenné industrie směřujeme je rekonstrukce operačního schématu analýzou jader, debitáže a pokud možno i skládáním jader a odbitých polotovarů do prvotních tvarů. Jako poměrně dobrý příklad můžeme uvést určení operačních schémat pro micoquienské souvrství v jeskyni Kůlna (Boěda 1995). Odlišil dva hlavní přístupy, kterými byly vytvářeny polotovary a následně nástroje na této lokalitě. Prvním z nich je metoda přímého tvarování (facetage), kde výchozí blok suroviny je přímo opracován na polotovar pomocí plošného obíjení povrchu a další, teď již nástrojovou retuší, upraven na vlastní nástroj bifaciálního charakteru s různými technologickými a morfologickými parametry, které se mění podle funkce i na jednom předmětu. Úštěpy vzniklé tímto odbíjením jsou v tomto případě odpadem, který však byl také dále využíván, ale nepředstavoval cílový produkt. Druhou metodou je pak odbíjení úštěpů z jader, kdy tyto úštěpy představují žádaný polotovar pro výrobu nástrojů. Žádoucími by byly i další statistické podklady týkající se jednotlivých metod výroby, které by např. charakterizovaly surovinovou skladbu ve vztahu k jednotlivým způsobům produkce nástrojů. Zdá se, že jednou z dosti signifikantních analýz je rozbor technologie odbíjení polotovarů, která aspoň u mladopaleolitických kolekcí dosti koresponduje s procesem produkce polotovarů a nástrojů.

Poměrně názorným příkladem prokřížení několika atributů je analýza způsobu odbití ve vztahu polotovar – talon – otloukač. Z databáze vytvoříme filtrováním menší soubor, který obsahuje celé a bazální kusy (předměty, na nichž je zachován talon), pak můžeme vhodně zadanou dotazovou větou seřadit informace tak, aby bylo možné vytvořit tabulku, ze které jsou výsledky snadněji zhodnotitelné (TAB.1). V dolní části tabulky jsou pak prováděny sumarizační výpočty těch závislostí, které nás aktuálně zajímají. Na tento druh úloh je výhodné použít tabulkové procesory, které mohou celý proces

zautomatizovat a celá práce se díky přívětivému uživatelskému prostředí zpříjemní a hlavně zrychlí. graf 1 pak znázorňuje použití otloukačů v jednotlivých stádiích výrobního procesu v mladopaleolitické kolekci. Z výsledků analýz můžeme konstatovat závislosti o použití tvrdého otloukače v preparačním (zejména u úštěpů) a reparačním stádiu výrobního procesu (u všech polotovarů), zatímco pro těžbu čepelí a zejména jejich cílových variant byl využíván měkký otloukač. Abrazí těžní hrany pak sleduje toto schéma a je vázána spíše na měkký otloukač u cílových polotovarů. Podobně bychom mohli dále pokračovat s určováním závislostí mezi polotovarem a typem talonu s následným vztahem k použité technice odbití. Vezmeme-li konkrétní polotovary, na které jsou tyto atributy navázány, můžeme konstatovat, že výrazné rozlišení polotovarů podle procesu výroby se promítá i do výrazného střídání technik použitých pro odbíjení polotovarů.

V případě, že kolekce obsahuje kusy, které jsme schopni zpětně poskládat, získáváme nový kvalitativní znak celé industrie zejména pro tvorbu operačního schématu. Příkladem může být operační schéma z lokality Hošťálkovice – Hladový vrch, kde se podařilo určit i některé individuální postupy štípače a zrekonstruovat celý postup exploatace jádra (Neruda 1995).

Zajímavým problémem, se kterým se můžeme setkat, je determinace čepel/úštěp ve středopaleolitických kolekcích, tj. zda určit čepel pouze morfologicky nebo metricky. Řešením by mohlo být užití morfologického kritéria (paralelní hrany a vnitřní negativy) s tím, že se pak za použití tabulkového procesoru vynesou rozměry do metrického X/Y grafu, ve kterém mohou vyniknout 2 základní varianty. Buď získáme poměrně jednotně rozptýlenou jednu skupinu dat, ze které vyplývá, že metrický aspekt nehrál výraznou roli při snaze o získání paralelního polotovaru a nebo získáme více méně 2 oddělené skupiny dat, které ukazují na tu skutečnost, že spolu s morfologickým charakterem odštěpu byla brána v úvahu i metrická kvalita polotovaru.

B\ Rekonstrukce využití suroviny a debitáže může být korelována s ekonomikou využití nástrojů, která pak vyplývá ze specializovaných nástrojových analýz. Surovina představuje jeden ze základních technologických determinantů, které nutí rozložit kolekci na surovinově rozdílné části a srovnat jejich operační schémata, čímž se mohou projevit rozdíly ve využívání jednotlivých surovin. Tento systém determinace jsem v poněkud jednodušší formě uplatnil pouze jednou, a to u kolekcí ze severní Moravy, které nejevily surovinovou pestrost a tudíž ve sledovaném materiálu byl hlavní a téměř jedinou surovinou eratický silicit. Proto byly technologické a typologické údaje prokříženy pouze v jedné hladině, takže výhoda využití výpočetní techniky pro takové zpracování není tak výrazná, jak by tomu mohlo být. Řada nových prací by ale měla tuto problematiku jasně demonstrovat. Jednou ze základních informací, které se snažíme získat pro určení ekonomického modelu, je technologická skladba kolekce, tzn. kvantitativní nebo procentuální zastoupení jednotlivých stádií a výsledků výrobního procesu a jejich využití na nástroje. Je tedy nutné vyčíslit technologickou skladbu debitáže nepoužité na nástroje a u nástrojů. Při optickém vyjádření závislosti je výhodné použít např. sloupcový sumarizační graf (graf 2), ze kterého vyplývá, že cílové polotovary jsou preferovány, ale že jsou využívány i ostatní stádía výrobního procesu. Tuto analýzu můžeme dále konkretizovat srovnáním zastoupení výsledků výrobního procesu a jejich využití na nástroje například pro těžební stádium výrobního procesu (graf 3). Tyto výsledky bylo prozatím možné získat z technologického formuláře. Předpokládáme však, že nás zajímá, které polotovary byly využívány pro určité nástrojové třídy.4\ Relace propojíme formulář pro technologické zpracování a zpracování nástrojů a z vyfiltrovaných údajů můžeme např. vytvořit graf 4, který nám sledované údaje graficky porovná. Může se vyskytnout případ, typický spíše pro středopaleolitické kolekce (např. taubachien), kdy nejsme schopni předem jednoznačně vytvořit systém zařazení polotovarů do stádií a výsledků výrobního procesu, protože není jasné, které kusy měly být cílové. Pak je zřejmě výhodné rozdělit soubor debitáže pouze podle morfologických znaků – počet negativů na dorsální ploše, směr jejich orientace a procento kůry na tomto povrchu – a z výše uvedených analýz pro graf 2 až 4 určit, které z nich jsou preferovány pro výrobu nástrojů, případně pro kterou nástrojovou třídu a v jaké rozměrové skupině.

Z jedné kolekce se například podařilo určit, že všechny použité polotovary pro škrabadla měly poměrně jednotnou tloušťku, byly fragmetární s původní lomovou plochou a jejich délka měla rovněž minimální rozptyl. Zdá se tedy, že byly vybírány silnější čepel (bez specializace na určité stádium výrobního procesu) a že byly vybírány i určité rozměry nebo tyto rozměry byly záměrně redukovány tak, aby nástroj odpovídal určité nástrojové „normě“. Vezmeme-li u téže kolekce nástrojovou třídu rydel, musíme konstatovat silnou specializaci při výběru polotovaru (čepel bez kůry), pokud možno delších rozměrů a konstantní tloušťky okolo 8 mm. Při použití specializovaného formuláře pro rydla

se můžeme také zabírat problematikou přetvarování – reutilizace nástrojů (vhodná jsou např. i drasidla). Nabídnuté výsledky jsou pouze tím základním, co bychom měli získat. V případě potřeby můžeme sledovat využití polotovarů i v rámci jedné nástrojové třídy, navíc ve vztahu k rozměrům, pravidelnosti polotovaru apod.

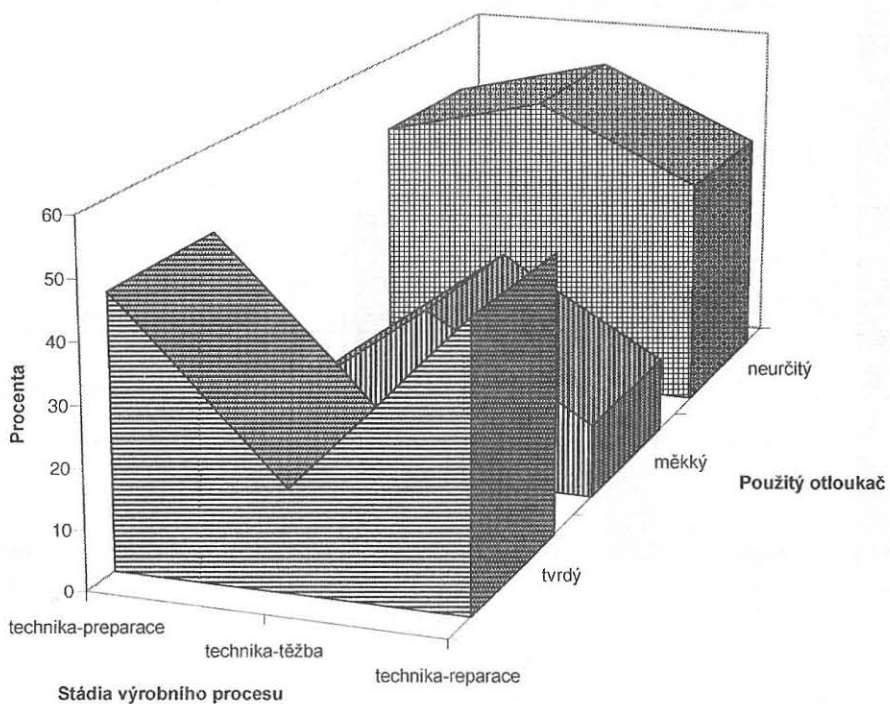
C\ Určení funkce sídliště je žádoucí zejména u regionálně orientovaných prací, při kterých se pokoušíme porovnávat několik lokalit mezi sebou. Různé funkce sídliště ovlivňují technologickou i typologickou skladbu souboru a je pak úkolem základní technologicko-typologické analýzy určit, které znaky souvisejí s jiným kulturním nebo časovým substrátem a které mohou vyplývat z odlišných funkcí porovnávaných nalezišť. V základních rysech lze říci, že ateliérové lokality mají výraznou převahu neretušovaných polotovarů a jader nad nástroji, které by se měly vyskytovat spíše výjimečně. Na nalezištích s loveckým zaměřením je zastoupení neretušovaných polotovarů nižší a skladba nástrojů ukazuje spíše na lovecké praktiky (např. hroty) a upřesnit sledovanou situaci mohou i nálezné okolnosti (viz. bod D\). Na stálých sídlištích je pak skladba nástrojů pestrá a skladba neretušované debitaže a jader se liší podle dostupnosti surovinových zdrojů.

D\ Spaciální modely na sídlišti jsou jistě tím nejzajímavějším, co můžeme získat. Pomáhají nám určit ekonomické, ale i společenské poměry na sídlišti, osvětlit nám vlastní život tehdejších lidí. Velmi názorným příkladem je situace, kterou analyzovali na lokalitě Etiolles ve Francii. Remontážemi a technologickým rozбором industrie byl určen následující model využití bloku suroviny. Nejprve byl blok suroviny mimo vlastní naleziště oštípán do formy počátkového jádra, které pak bylo dále upraveno na sídlišti do tvaru vhodného pro těžbu polotovarů. Tato těžba se odehrála na dvou místech. U jednoho z ohnišť bylo toto jádro velice kvalitně exploatováno a byly z něho vytěženy dlouhé čepele. V určitém stádiu vytěženosti se jádro přesunulo k druhému ohništi, kde již nebyly výsledky snažení tak dokonalé a mohlo jít o činnost, připomínající experimentování. Vědci vysvětlují tento jev jako dětské učení se procesům štípání. Z toho můžeme vyvodit nejen ekonomii využití bloku suroviny, ale i vnitřní členění prostoru sídliště podle zájmových oblastí (Olive-Pigeot-Taborin 1991, 38 a 41).

Jiným příkladem, který souvisí s loveckými stanicemi, je spojení technologické analýzy s náleznými okolnostmi a přírodovědnými rozbory. Na lokalitě Combe Sauniere (Chadelle-Geneste-Plisson 1991) se našly místa s bazálními zlomky solutrénských hrotů a místa s jejich terminálními ekvivalenty. Místa, kde se nalézaly hlavně terminální části byla podrobena přírodovědným analýzám, zaměřeným na výbrusy půdy a ty ukázaly, že tyto prostory obsahují krev zvířat. Celkový výsledek lze stručně vyhodnotit takto: 1. na lokalitě existovaly místa, kam byla přinášena ulovená zvířata a následně porcována. Terminální části hrotů byly zabodnuty v jejich tělech a při čtvrcení byly uvolňovány ven. Nálezy bazálních částí hrotů na jiném místě souvisely s tím, že ratiště oštěpů donesli zpět na lokalitu, aby je použili pro nové oštěpy.

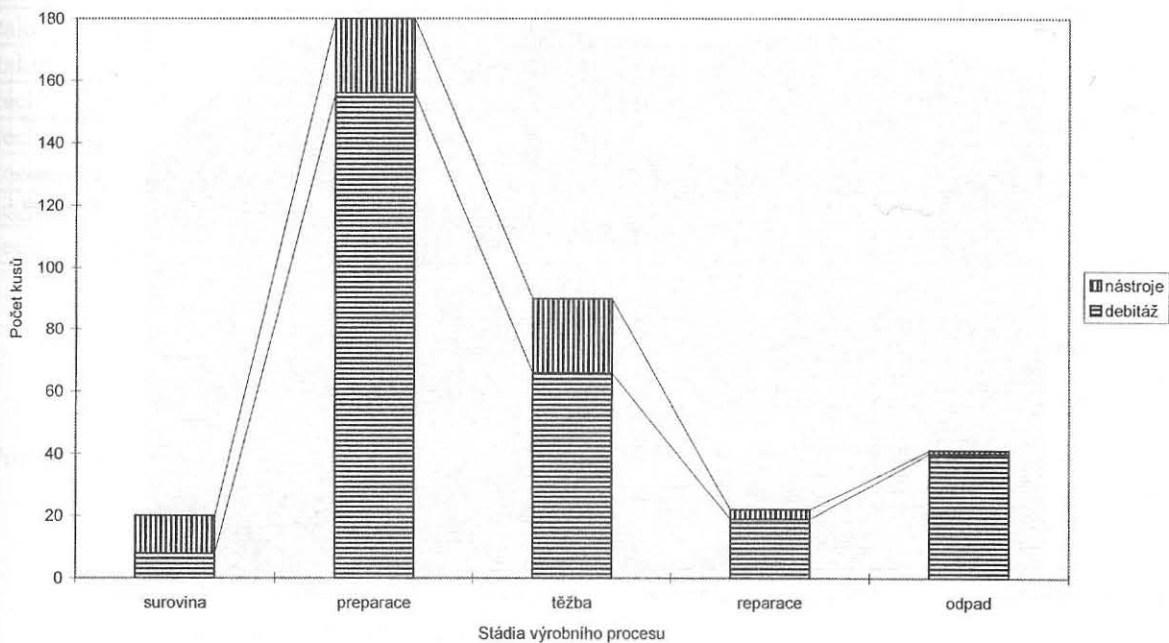
Je zřejmé, že vlastní analýza kamenné industrie je složena ze dvou základních procesů: determinace a vyhodnocení. Zatímco popis atributů je více méně mechanická práce, postavená na solidní znalosti kamenné technologie, je vyhodnocení poměrně problematičtější úsekem práce.^{5\} Je totiž nutné rozhodnout, co se získanými informacemi, jakým metodologicko-statistickým metodám je podrobit, aby jsme se co nejvíce přiblížili ke skutečné rekonstrukci zkoumaných jevů. Je tedy zřejmé, že je velmi obtížné vytvořit pevnou kuchačku, podle níž bychom mechanicky postupovali, neboť by se tím vytratila osobní invence badatele. Je to snad nemožné i z toho důvodu, že každý materiál má svá specifika, které musíme respektovat a přizpůsobit se jim.

Technika odbíjení ve stádiích výrobního procesu



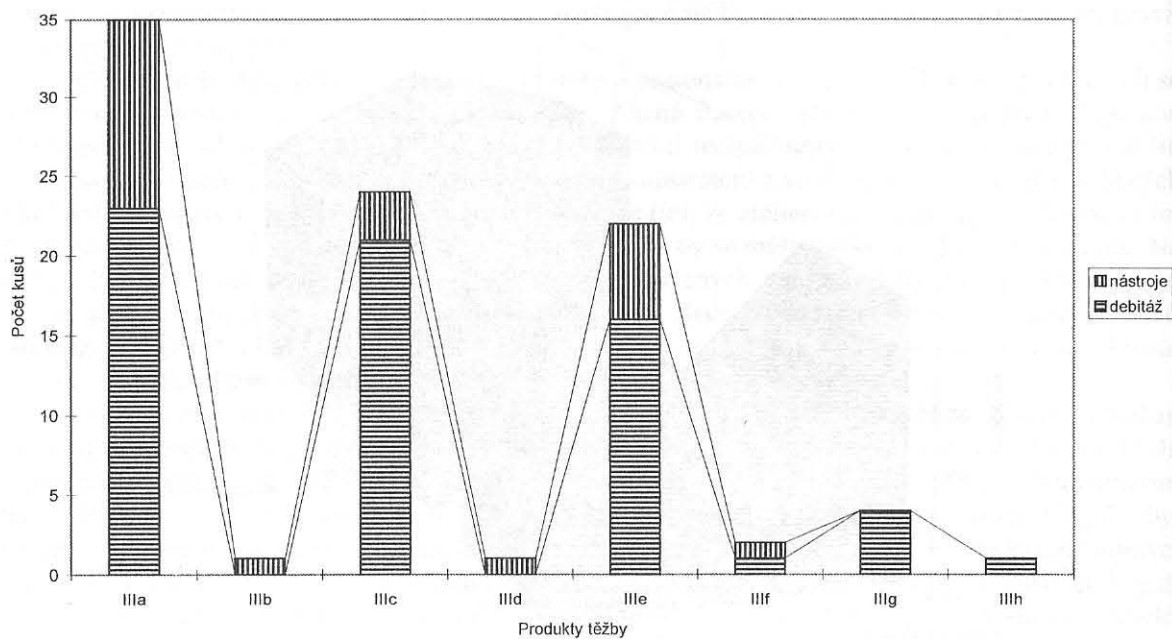
Graf 1.

Využití stádií výrobního procesu na nástroje



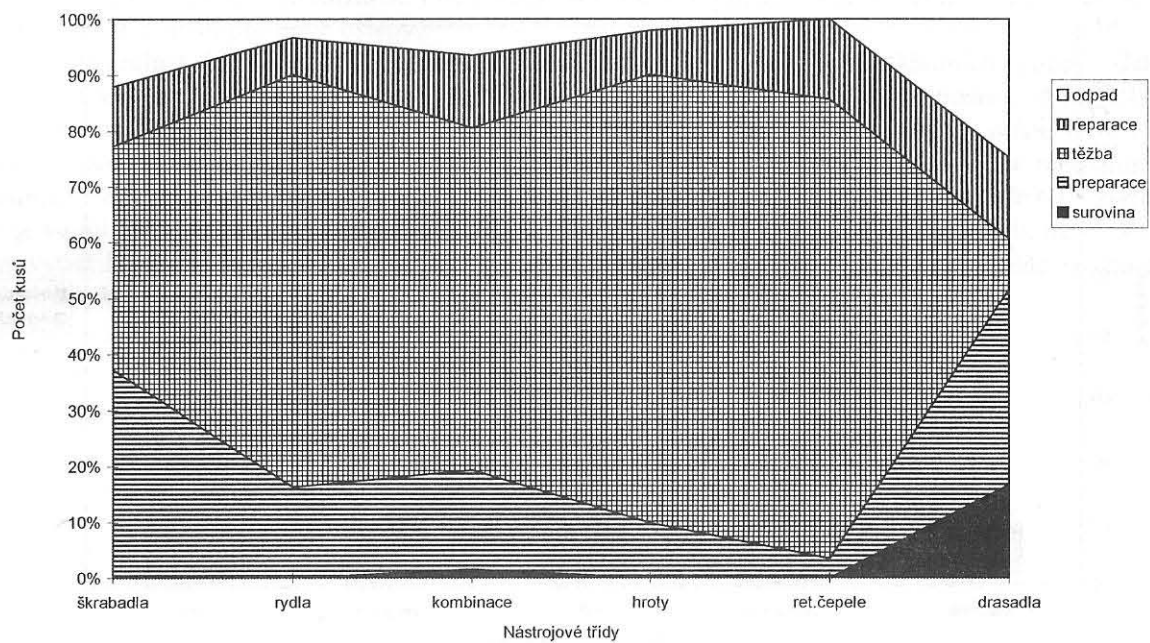
Graf 2.

Využití produktů z těžby



Graf 3.

Využití stádií výrobního procesu pro nástrojové třídy



Graf 4.

Tab. 1		otloukač													
otloukač		měkký					tvrdý					neurčitý			
	k	p	f	l	b	k	p	f	l	b	k	p	f	l	b
talón		2	1			38	32	1	2	5	16	10		1	4
ústěp s 100 % kůry	1	3				19	32	1	2	11	20	24	2	4	7
ústěp z hrany		0				9	2	1	1	1	9			3	2
čepel s kůrou		1				2	6		2		1	8		4	4
čepel z hrany		18		2	2	1	6	1	1	3	1	18	3	8	8
preparační čepel		6	1			2	13	2		8	1	23	1	14	10
ústěp bez kůry	0	23	0	7	4	2	19	3	3	2	1	36	11	18	18
ústěp s bokem jádra	1	8	1			2	10		3	1	2	33		10	3
ústěp s lat.negativy	1	18	4	3	3	1	10	1	1	5	4	16	2	20	13
ústěp s II neg.po lat.neg.		3	2				2					2		4	
čepel bez kůry		66	11	23	3		19	1	4	16	2	36	6	50	29
čepelka bez kůry		2	1		1		1		1	5		5	1	7	13
čepel s bokem jádra		10	2	2						1	1	3	2	3	6
čepelka s bokem jádra		1					1			1			1		4
čepel s lat.neg.po II neg.				1								1			
tableta						1	9	5	1	2		7	7	1	1
odražená těžní plocha		1				4	9	1		3	2	2	2	1	4
sekundární hrana z jádra		7	1		1		4		1	1		9	3		2
talón u prepar. ústěpů	1	5	1	0	0	57	73	4	5	17	37	43	2	8	13
talón u prepar.čepelí	0	25	1	2	2	5	25	3	3	11	3	49	4	26	22
talón u těžebních ústěpů	1	31	1	7	4	4	29	3	6	3	3	69	11	28	21
talón u těžebních čepelí	1	100	20	28	7	1	33	2	6	28	7	62	12	84	65
talón u repar. ústěpů	0	1	0	1	0	5	18	6	1	5	2	10	9	2	5
talón u repar. čepelí	0	7	1	0	1	0	4	0	1	1	0	9	3	0	2
talony pro ústěpy	2	37	2	8	4	66	120	13	12	25	42	122	22	38	39
talony pro čepele	1	132	22	30	10	6	62	5	10	40	10	120	19	110	89
talony u preparací	1	30	2	2	2	62	98	7	8	28	40	92	6	34	35
talony u těžby	2	131	21	35	11	5	62	5	12	31	10	131	23	112	86
talony u reparací	0	8	1	1	1	5	22	6	2	6	2	19	12	2	7
technika pro ústěpy			53					236					263		
technika pro čepele			195					123					348		
technika pro preparací			37					203					207		
technika pro těžbu			200					115					362		
technika pro reparace			11					41					35		

Tab. 1.

Poznámky:

- 1) Tato nejednoznačnost se projevuje nejčastěji u klasifikace použitého otloukače k odbití polotovaru nebo při technologickém určení polotovaru. Celá situace se dá přirovnat např. ke klasifikaci charakteru keramického těsta v pozdějších obdobích pravěku.
- 2) Databázové pole je výhodné definovat úsporně s přihlédnutím k co možná neekonomičtějšímu využití operační paměti.
- 3) Pro rozlišení tvrdého a měkkého otloukače je používáno následujících kritérií: tvrdý otloukač: výrazně vystouplý bulbus s bodovým výstupkem na talonu z ventrální strany polotovaru vzniklý při bodovém kontaktu s otloukačem. Na bulbu jsou patrné dostředné

části kružnic (aspoň jedna) s středem v místě úderu. měkký otloukač: na ventrální straně talonu není výstupek, tudíž nelze určit bodové místo úderu (neexistuje). Je to dáno tím, že se měkký otloukač při úderu rozkládá v kontaktu se surovinou na větší ploše. Výsledkem je římsička na hraně ventrální plochy polotovaru a podstavy.

- 4) Bylo by samozřejmě možné zpřesnit tuto analýzu na jednotlivá typologická čísla, ale výsledek i v grafické formě by byl značně nepřehledný.
- 5) Pro klasifikaci atributů je, kromě znalosti skutečného autentického štípaného materiálu, téměř nezbytná i znalost experimentálních výsledků, ať již zprostředkovaně, nebo i z autopsie. Nejsou sice samospasitelné, ale umožní nám představit si nesmírnou morfologickou variabilitu, která se projevuje v konkrétních archeologických kolekcích, ale umožní nám i určit některé jevy, které by, bez testování byly těžce analyzovatelné (funkční partie nástrojů). Na druhou stranu si můžeme zpětně vyzkoušet, zda technologické charakteristiky, které jsme analýzou určili, přinesou konkrétní výsledky i při napodobení určeného operačního schématu.

6.3 Literatura

- Boěda, E. 1995: Caracteristiques techniques des chaînes opératoires lithiques des niveaux Micoquiens de Kůlna (Tchecoslovaquie), Actes du Colloque de Miskolc. PALEO-Supplément No 1, 57-72.
- Chadelle, J.-P. – Geneste, J.-M. – Plisson, H. 1991: Processus fonctionnelle de formation des assemblages technologiques dans les sites du Paléolithique supérieur. Les pointes de projectiles lithiques du Solutréen de la grotte de Combe Sauniere (Dordogne, France), in: 25 ans D'ÉTUDES TECHNOLOGIQUES EN PRÉHISTOIRE ET D'HISTOIRE D'ANTIBES, Éditions APD-CA, Juan-les-Pins, 275-287.
- Neruda, P. 1995: Technologická analýza remontáže gravettienské industrie z lokality Hošťálkovice-Hladový vrch, AMM, Sci.soc. LXXX, 1/2 29-44.
- Olive, M. – Pigeot, N. – Taborin, Y. 1991: Il y a 13 000 ans à Étioilles. Paris-Argenton-sur-Creuse.
- Schild, R. 1980: Introduction to Dynamic Technological Analysis of Chipped Stone Assemblages. In: Unconventional Archaeology, New Approaches and Goals in polish Archaeology. Wroclaw-Warszawa-Krakow-Gdańsk 1980, 57-85.

Doporučená literatura:

- Bordes, F. 1967: Considérations sur la Typologie et les techniques dans le Paléolithique, Quartär 18, 25-56, Pl. I-VIII.
- Cahen, D. – Karlin, C. – Keeley, L.H. – Van Noten, F. 1980: Méthodes d'analyse technique, spatiale et fonctionnelle d'ensembles, Helinium 20, 3, 209-259.
- Inizan, M.-L. et al. 1995: Technologie de la pierre taillée, Préhistoire de la pierre taillée. Tome 4. Meudon: CREP.
- Tailler! Pour quio faire: préhistoire et technologie lithique II. Recent progress in microwear studies, Studia Praehistorica Belgica 2, 65-76. Tervuren: Musée Royal d'Afrique Centrale.
- Terminologie et Technologie, Préhistoire de la pierre taillée 1, C.R.E.P. Paris 1980.
- Économie du débitage laminaire, Préhistoire de la pierre taillée 2, C.R.E.P. Paris 1984.

7. Archeologická databáze Čech (Martin Kuna)

7.1 Úvod

Počítačové databáze jsou v dnešní době již nepostradatelnou součástí archeologie, a to jak při zpracování jednotlivých terénních výzkumů nebo kulturně historických témat, tak při evidenci archeologických památek v rámci regionu a státu. Teprve databáze činí rychle narůstající objem dat všeho druhu plně přístupným, tříditelným a přenosným.

Nutnost průběžně vytvářet souborný přehled archeologického nálezového fondu v měřítku celého státu si česká archeologie uvědomovala dlouho. Řada kroků tímto směrem byla uskutečněna již v minulých desetiletích, a to např. vydáváním periodika *Výzkumy v Čechách*, zveřejněním soupisů nálezových zpráv a zavedením jednotné prostorové identifikace PIAN v Archeologickém ústavu v Praze, vytvářením regionálních soupisů nalezišť apod. Operativní využití nashromážděných informací bylo však až do konce 80. let podstatně omezeno nedostatkem výpočetní techniky. Tento stav se v posledních letech rychle mění. Vzniká tak především reálná možnost převodu rozsáhlých souborů dat do podoby počítačových databází jako informačního zdroje nové kvality.

Práce na vytvoření databázového systému evidence archeologických nálezů, nalezišť a výzkumů z území Čech byly v Archeologickém ústavu AV ČR v Praze zahájeny již na počátku 90. let. Od této doby zde postupně vzniká rozsáhlý databázový soubor, *Archeologická databáze Čech (ADC)*, poskytující základní údaje o rozmístění a charakteru archeologických pramenů v Čechách a postupu jejich dosavadního výzkumu. Problematika vytváření, správy a využití této databáze zahrnuje několik hlavních tematických okruhů, kterým odpovídají následující kapitoly.

Přehledné informace o podobných „národních archeologických databázích“ v Evropě obsahuje sborník C.U. Larsena (1992). Další podobné projekty již delší dobu probíhají v Polsku (Konopka 1980) a Maďarsku (Bakkay, Kalicz a Sági 1966); v poslední době byl zahájen nový projekt centrální evidence archeologických nalezišť i na Slovensku (Bujna a kol. 1993). Kritikou stavu souborného utřídění pramenné základny v Čechách se naposledy zabývali M. Kuna a J. Klápště (1990), obecnou problematiku databází a konkrétní příklad zpracování okresní databáze shrnul E. Neustupný (1994, 1996). Podrobnější informace o ADC obsahuje uživatelská příručka systému ARCHIV (Kuna, Křivánková a Krušinová 1995).

7.2 Zdroje informací ADC

Archeologická databáze Čech vzniká z několika zdrojů:

- (a) Tzv. *roční zprávy* o nových archeologických výzkumech a údaje o výzkumech nově zpracovaných do podoby nálezové zprávy.
- (b) *Novější fondy* archivu nálezových zpráv ARÚ Praha. Od r. 1984 měla hlášení a nálezové zprávy již standardní podobu a vesměs prošla evidencí PIAN (lokalizace koordináty). Revize těchto dat a jejich převod do databáze jsou již ukončeny.
- (c) *Starší fondy* archivu nálezových zpráv ARÚ Praha. Tato část databáze vzniká převodem publikovaných soupisů nálezových zpráv z let 1919-1952 (Charvátová, Spurný a Venclová 1992), 1955-1964 (Justová 1968) a záznamů v BZO z let 1963-1987. Převod těchto dat je již před dokončením; na rozdíl od předchozích datových bloků zde bude v budoucnu ještě zapotřebí provést revizi a doplnit chybějící údaje (např. lokalizaci koordináty).
- (d) Další *databázové soubory různého původu*, které jsou pro území Čech k dispozici. Jde např. o kartotéku MONUMIS, vypracovanou v ARÚ Praha v 80. letech, kartotéku archeologických výzkumů na území bývalého Východočeského kraje, zpracovanou muzeem v Hradci Králové, nebo databázi mostecké expozitury ARÚ (dnes ÚAPPSZČ). Stejným způsobem lze do ADC připojovat jakékoli další regionální nebo tematické okruhy dat.

Shromáždění takového souboru dat představuje *první etapu* tvorby informačního systému, která se v současné době již uzavírá. Získaná data jsou ovšem informacemi s rozdílnou mírou úplnosti, spolehlivosti i vypovídací hodnoty. Předpokládáme, že tato data jsou pouze východiskem pro další zpracování a ve své současné podobě zůstanou v databázi pouze do té doby, než budou nahrazena přesnějšími údaji.

Obsahem *druhé etapy* tvorby databáze by mělo být, kromě průběžného začleňování údajů z nových výzkumů, především doplnění informací o výzkumech, které nejsou evidovány v ARÚ Praha, a dále celková revize záznamů, včetně např. jejich lokalizace koordináty. Tento úkol však není možné provést bez účasti dalších pracovišť, zejména regionálních ústavů a muzeí, a to jak z důvodu kapacitních, tak proto, že často pouze v těchto institucích jsou k dispozici potřebné informace. Na zájmu a možnostech regionálních archeologických pracovišť spolupracovat na vytváření kvalitnější podoby centrální databáze bude proto do značné míry záviset další vývoj ADČ.

7.3 Struktura dat

Evidenční jednotkou ADČ je „archeologická akce“, čili archeologický výzkum, průzkum, nález či jiné zjištění, které má vlastní prostorové vymezení, nálezové okolnosti, autora a dobu výzkumu. *Archeologická akce* je jednotkou *primární dokumentace*, tj. je nezávislá na případných budoucích změnách hodnocení nálezů, jako je jejich předatování, změny ve funkční interpretaci atd. Archeologickou akci dále členíme na *komponenty* (přesněji na komponenty akce), kterými rozumíme nálezové jednotky, související s určitou specifickou funkcí (např. sídliště) a obdobím (zpravidla archeologickou kulturou). Jednotlivé komponenty mohou mít v rámci akce vlastní, přesnější prostorové určení (nezbytné je to zejména rozsáhlejších akcí). Při sběru údajů o nových výzkumech chápeme jako samostatnou akci i každou sezónu dlouhodobého výzkumu.

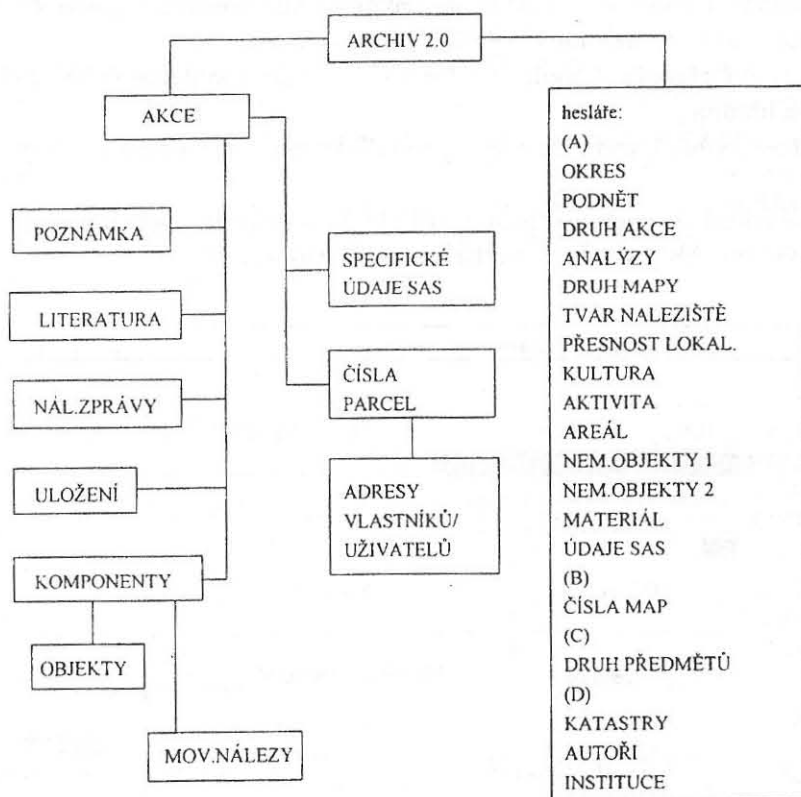
Členění databáze podle archeologických akcí vychází z tradičního modelu evidence terénních výzkumů v ARÚ Praha. Má své nesporné výhody, neboť organizace databáze podle jiných kritérií (např. podle nalezišť nebo komponent) je nesrovnatelně náročnější na přípravu dat, je závislá na správnosti odhadu prostorového rozsahu naleziště a chronologického určení nálezů; navíc při ní vždy dochází k určité úpravě či úniku vstupních informací.

V některých případech je ovšem nutné pojem archeologické akce chápat poněkud volněji. Např. při zpracování *starších nálezů* jsou často jednotlivé epizody výzkumu určité lokality zpětně nerekonstruovatelné, jako akci (tj. záznam) zde proto musíme vymezit celou sérii následných archeologických zásahů v určitém prostoru (např. v cihelně, vymezené stejnými koordináty). Naopak, např. v případě systematických *povrchových sběrů* musíme jako „akci“ (tj. samostatným záznamem) popisovat každé naleziště, polygon sběru nebo jinou referenční jednotku s vlastním prostorovým určením. Jako specifický druh archeologické akce chápeme také jednorázovou *identifikaci určitého celku* archeologických pramenů, který je v terénu vymezenitelný pomocí nemovitých pozůstatků (jako je tomu často např. u mohylníků, hradišť, feudálních sídel, zaniklých středověkých vsí atd.).

Každá archeologická akce (záznam databáze) obdrží během zpracování individuální *identifikační kód*. Tento kód je neopakovatelný a umožňuje akci jednoznačně rozlišit (kód se skládá ze SPZ okresu, prvních deseti znaků jména katastru, roku provedení akce a počítačem náhodně generovaného trojmístného čísla). Výhodou tohoto kódu je to, že je sám o sobě srozumitelný a umožňuje vyhledat v databázi příslušný záznam i tehdy, pokud známe jen některé z klíčových údajů (např. jen jméno katastru a přibližný rok výzkumu). Každé nové akci je přiděleno i *číslo PIAN*. S tímto číslem je akce zanesena do map ZM 1:10000 (v případě historického jádra Prahy do Map evidence nemovitostí 1:1000), které slouží ke kontrole prostorového určení a k snadnější prostorové orientaci v přechodném období než se stane běžným používání GIS. S vedením evidence PIAN v tištěných mapách počítáme přinejmenším do té doby, než je bude možné nahradit digitálními mapovými podklady.

Přidělením nezaměnitelného identifikačního kódu každému záznamu (akci) vytváří ADČ možnost jednotné identifikace terénních výzkumů na území Čech. Na tyto identifikátory bude v budoucnu možné připojit záznamy v dalších, specializovaných databázích (např. inventáře muzeí, bibliografie, archivy fotografií a plánů apod.).

ADC je relačním databázovým systémem. Záznamy o jednotlivých akcích obsahují kromě základního popisu a lokalizace také odkazy na literaturu, uložení nálezů a především výčet jednotlivých komponent, k nimž se dále připojují výčty kategorií nemovitých a movitých nálezů. Schéma těchto vztahů ukazuje obr. 1.



Obr. 1. Schéma struktury databázového systému Archeologické databáze Čech, program ARCHIV 2.0

7.4 Systém ARCHIV

Systém ARCHIV je počítačový program, vytvořený v oddělení prostorové archeologie pro manipulaci s ADC (autoři M. Kuna a D. Křivánková). Skládá se z databázových souborů, určených pro ukládání záznamů, dále z databází-heslářů, které slouží ke kontrole správnosti zápisu hlavních údajů, a vlastního programu ARCHIV. Tento program umožňuje prostřednictvím uživatelských menu, tj. bez nároků na zvláštní znalosti uživatele v oblasti počítačů, provádět s databází všechny základní operace. Systém ARCHIV 2.0 je vytvořen v programovacím jazyku dBASE IV pro DOS a jeho distribuční verze obsahuje i tzv. run-time modul, s jehož pomocí lze program spustit na libovolném počítači bez ohledu na to, zda je na něm instalován software dBASE. Základní operace v systému ARCHIV 2.0 jsou:

- ukládání (zápis) dat (obr. 2);
- editace záznamů (opravy, rušení a doplňování záznamů);
- výběr dat a jejich uložení do nových souborů (databázových i textových);
- tisk záznamů v podobě různých výstupních sestav.

Prostřednictvím menu lze provádět i další operace, potřebné při správě větších databází a výměně dat mezi různými pracovišti, jako např.:

- vytváření záložních kopií v komprimované podobě;
- obnova datových souborů ze záložních kopií;
- připojování souborů (např. datových bloků nezávisle vytvořených na jiném pracovišti).

Pro další verzi systému počítáme s rozšířením programu o menu, umožňující např.:

- automatické odhalování duplicitních záznamů;
- grafické zobrazení vybraných dat na obrazovce v rozsahu jednoho listu ZM 1:10000 a jeho výstup na tiskárnu v měřítku 1:25000;
- automatickou transformaci ADC do databázových souborů s jinou strukturou (např. do struktury databáze užívané ÚAPPSZČ v Mostě);
- automatický přepočítání koordinátů PIAN do souřadnic systému JTSK, S-42 nebo zeměpisných souřadnic;
- vytváření souborů, které lze načíst geografickými informačními systémy.

Všechny uvedené programy již jsou v ARÚ Praha k dispozici, jejich postupné začleňování do budoucích verzí systému ARCHIV bude vycházet z potřeb uživatelů.

AKCE		počet záznamů :	
BLOK	_____	ČJ.	_____/____
HLÁŠENÍ	_BZO_	NÁL.ZPR.	_____
LITER.	_____	MATER.	_____
JINÝ ZD.	_____	KAT.2.	_____
KATASTR	_____	KAT.3.	_____
4.	_____	5.	_____
ADMIN.O.	_____		
OKRES	_____	PŘ.Č.	_____/____
PODNĚT	_____	DRUH A.	____/____/____
ANALÝZY	____/____/____		
AUTOR A.	_____		
INSTIT.	_____		
ROK	_____	PŘESNÉ DATUM	_____
LOKAL.	_____		
D.MAPY	_____	Č.MAPY/NÁZEV	_____/_____
TVAR NAL.	_____	Č.PIAN	_____
PŘESNOST	_____		
SOUŘADNICE			
1.	____:____	2.	____:____
3.	____:____	4.	____:____
5.	____:____	6.	____:____
7.	____:____	8.	____:____
9.	____:____	10.	____:____
11.	____:____	12.	____:____
AUT.ZÁP.	_____	ROK ZÁP.	_____
PŘÍLOHY : STR	_____	OBR	_____
PLÁNY	_____	FOTO	_____
přechod na další části formuláře CTRL+END			

Obr. 2. Obrazovka pro ukládání a editaci záznamů v systému ARCHIV 2.0.

Dosavadní zkušenosti ukázaly určité problémy v oblasti hardwaru i softwaru. Původní verze systému ARCHIV byly koncipovány pro počítače typu PC 386 nebo 486 a distribuovány s run-time modulem jazyka dBASE IV. Tento software ovšem nevyhovuje např. počítačům typu Pentium, které vyžadují run-time modul jazyka dBASE V. V současné době jsou již k dispozici instalační diskety systému ARCHIV 2.0 v obou verzích. Jelikož rychlý vývoj hardwaru a softwaru (např. prostředí Windows95) může přinášet další problémy s instalací systému ARCHIV, je nutné případné obtíže řešit s autory programu. Převod systému ARCHIV do jiného softwarového prostředí (např. Windows) však bude smysluplný až v okamžiku, kdy se výkonnější hardware a software stane základním vybavením všech archeologických pracovišť a kdy budou sjednoceny některé z jeho technických parametrů (např. zápis českých znaků); učinit tento krok předčasně by znamenalo ohrozit vzájemnou kompatibilitu ukládaných souborů a možnosti efektivní výměny dat.

7.5 Sběr dat

Sběr dat pro ADC je zajišťován jednak činností oddělení prostorové archeologie (zpracování starších fondů, transformace jiných databází), jednak autory terénních akcí a projektů (podávání zpráv

o nových výzkumech, zpracování regionálních datových souborů). Data jsou buď nejprve zpracována na formulářích, nebo jsou do počítače ukládána přímo (systém ARCHIV počítá s výměnou dat na disketě a umožňuje připojování souborů). K popisu archeologických akcí jsou používány standardní *formuláře ZAA* (Zpráva o archeologické akci; obr. 3-6), případně formuláře modifikované rozšířením o rubriky památkové ochrany (ZAA-SAS). Při popisu akcí se využívají povolené, případně doporučené okruhy výrazů, tzv. *hesláře* (třídníky).

Informace o nových výzkumech přicházejí do ARÚ Praha na základě dohod mezi organizacemi oprávněnými k provádění archeologických výzkumů a Archeologickým ústavem v Praze. Zpráva je podávána o každé archeologické akci (sezóně výzkumu) v témže kalendářním roce, kdy akce proběhla, a má charakter tzv. *roční zprávy*, tj. souhrnné informace o výzkumu (roční etapě). Zprávy o archeologických výzkumech je možné podávat i v digitální podobě v systému ARCHIV. Jelikož odborné zpracování rozsáhlejších akcí může vyžadovat delší dobu, stejný formulář ZAA je také přílohou každé dodatečně vypracované *nálezové zprávy*.

7.6 Aktuální stav ADČ

Ke konci r. 1996 obsahovala ADČ tyto datové bloky:

Blok	Počet záznamů	Lokalizace koordináty
roční zprávy a NZ od r. 1990 (BZO 1988n.)	4525	+
roční zprávy a NZ 1984-1990	4724	+
starší roční zprávy (BZO 1963-1987)	3157	-
starší NZ (1919-1952, 1955-1964)	3822	-
činnost expozitury Most (1953-1992)	1798	+
kartotéka MVČ Hradec Králové	14572	-
projekt MONUMIS	11281	-
regionální výzkumné projekty ARÚ	2155	+
celkem	46034	

Data ADČ jsou, spolu se systémem ARCHIV 2.0 a uživatelskou příručkou, *bezplatně k dispozici všem archeologickým pracovištím v ČR*. K podmínkám předání dat patří pouze dodržování všeobecných zásad odborné práce, tj. zejména (a) další nešíření předaných dat bez souhlasu institucí, které se podílely na jejich sběru, (b) nevyužívání dat ke komerčním účelům bez dohody s těmito institucemi a (c) podávání zpráv o vlastních výzkumech.

Jedním z vedlejších výstupů ADČ je také periodická publikace *Výzkumy v Čechách (BZO)*. V současné podobě tato publikace shrnuje veškeré roční zprávy a nálezové zprávy, došlé do ARÚ Praha v určitém časovém úseku. Obsah publikace je k dispozici také na disketě, a to jak v podobě datové tabáze, tak textového souboru.

7.7 Literatura

- Bakkay K., Kalicz N., Sági K. 1966: *Veszprém megye régészeti topográfiája* (Magyarország régészeti topográfiája), Budapest.
- Bujna J., Kuzma I., Doliak D. a Jenis J. 1993: *Centrálna evidencia archeologických nálezísk na Slovensku*, Slovenská archeológia 41, 367-386.
- Charvátová K., Spurný V., Venclová N. 1992: *Nálezové zprávy StAÚ 1919-1952*, Praha (ARÚ).
- Justová J. 1968: *Nálezové zprávy Archeologického ústavu ČSAV 1955-1964*, Archeologické studijní materiály 6, Praha (ARÚ).
- Konopka M. 1980: *Carte des sites archéologiques en Pologne: méthodes et organisation*, Archeologia Polona 21-22 (1983), 187-216.

- Kuna M., Klápště J. 1990: Poznámky ke koncepci terénní archeologické práce, *Archeologické rozhledy* 42, 435-445.
- Kuna M., Křivánková D., Krušinová L. 1995: ARCHIV 2.0. Systém Archeologické databáze Čech. Uživatelská příručka, Praha (ARÚ a SÚPP).
- Larsen C.U. (ed.) 1992: Sites & monuments. National archeological records, København (Nationalmuseet).
- Neustupný E. 1994: Role databází v archeologii, *Archeologické rozhledy* 46, 123-130.
- 1996: Databáze nalezišť okresu Chrudim, *Archeologické rozhledy* 48, 126-134.

Číslo	Název	Adresa	Telefon	Fax	E-mail
1	Archeologické ústředí	Prague			
2	Archeologické ústředí	Prague			
3	Archeologické ústředí	Prague			
4	Archeologické ústředí	Prague			
5	Archeologické ústředí	Prague			
6	Archeologické ústředí	Prague			
7	Archeologické ústředí	Prague			
8	Archeologické ústředí	Prague			
9	Archeologické ústředí	Prague			
10	Archeologické ústředí	Prague			

ZPRÁVA O ARCHEOLOGICKÉ AKCI

Systém ARCHIV - ARÚ AVČR Praha, odd. prostorové archeologie

AKCE

Zaškrtněte některou z možností, nebo vyplňte podle heslářů. V rubrice DRUH AKCE můžete uvést max. tři hesla (čísla), oddělená lomítkem. V rubrice ANALÝZY uveďte libovolný počet hesel, oddělených lomítkem.

Soubor:

ZPRÁVA <input type="checkbox"/> (Hlášení)	KATASTR (hlavni)	DALŠÍ KAT.Ú.	
<input type="checkbox"/> RZ - BZO	ADHIN.O.	OKRES	PR.Č.AKCE (pro ARÚ)
<input type="checkbox"/> Nál.zpr. <input type="checkbox"/> (Liter.) <input type="checkbox"/> (Mater.) <input type="checkbox"/> (Jin.zd.)	PODNET	DRUH AKCE	ANALÝZY
VEDOUcí VYZKUMŮ	INSTITUCE	ROK (od-do)	PŘESNÉ DATUM

LOKALIZACE/OKOLNOSTI/JINÉ OZNAČENÍ (max. 110 znaků)	DRUH MAPY <input type="checkbox"/> ZH25 <input type="checkbox"/> ZH10 <input type="checkbox"/> 1:1000 jiný:	Č.MAPY Pozn.: Souřadnice všech bodů jedné akce je nutno udat od Z:J s.č.jednoho listu!
Základní údaj, blíže označující naleziště/akci. Podle možnosti uveďte traťové jméno, ppč., jméno majitele pozemku, údaj o původu nálezů atd. Rozveďte v POZNÁMCE.	NALEZIŠTĚ JE <input type="checkbox"/> v daném Bodě (bodech) <input type="checkbox"/> na dané Línii <input type="checkbox"/> v Ploše mezi body <input type="checkbox"/> v Okruhu daného bodu	PŘESNOST LOKALIZACE <input type="checkbox"/> 1 přesně (geodeticky) <input type="checkbox"/> 2 s rozptylem cca 25 m <input type="checkbox"/> 3 s rozptylem 200-300 <input type="checkbox"/> 4 pouze kat.území
SOUŘADNICE 1. : 2. : 3. : 4. : 5. : 6. : (mm od Z:J s čáry)	7. : 8. : 9. : 10. : 11. : 12. :	Více než 6 párů souřadnic uvádějte jen u dálk.výkopů (31,42). V těchto případech uveďte přesnou lokalizaci jednotliv. komponent!
ZAPSAL:	ROK ZÁP.	PODPIS:

POZNÁMKA (Doplňující popis naleziště či nálezů; pokračujte na volném vloženém listě.)

LITERATURA (Výběr, zejména původní publikace a novější materiálové práce s dalšími odkazy.)

Příjmení, jm. autora(ů), rok publ., str., zkratka citace (název časopisu, názvu monografie)

NÁLEZOVÉ ZPRÁVY (U čj. nález. zpráv připojte "NZ", u čj. mimo ARÚ do závorky zkratku instituce.)

KOMPONENTY

Rubriky vyplňujte podle heslářů. PŘESNOU DATACI uveďte, jen je-li to možné, a to libovolným srozumitelným výrazem v délce do 25 znaků. V rubrice AKTIVITA můžete použít až 5 hesel, oddělených lomítkem, avšak pouze u dále nedělitelných komponent. PŘESNOU LOKALIZACI uvádějte pouze tehdy, není-li totožná s lokalizací celé akce. V rubrice NADM.VÝŠKA uveďte nadmořskou výšku v metrech (od-do). URČIL vyplňte, nejde-li o určení autora zápisu, a připojte, zda jde o určení podle materiálu, nebo druhotného zdroje (literatura, inventární kniha).

A	KULTURA	PŘESNÁ DATACE				AKTIVITA	AREÁL	NADM.V.
	PRES.1. LOK. :	2. :	3. :	4. :	5. :	6. :	URČIL autor záp. ?:	mater. [] dr.zd. []

B	KULTURA	PŘESNÁ DATACE				AKTIVITA	AREÁL	NADM.V.
	PRES.1. LOK. :	2. :	3. :	4. :	5. :	6. :	URČIL autor záp. ?:	mater. [] dr.zd. []

C	KULTURA	PŘESNÁ DATACE				AKTIVITA	AREÁL	NADM.V.
	PRES.1. LOK. :	2. :	3. :	4. :	5. :	6. :	URČIL autor záp. ?:	mater. [] dr.zd. []

D	KULTURA	PŘESNÁ DATACE				AKTIVITA	AREÁL	NADM.V.
	PRES.1. LOK. :	2. :	3. :	4. :	5. :	6. :	URČIL autor záp. ?:	mater. [] dr.zd. []

E	KULTURA	PŘESNÁ DATACE				AKTIVITA	AREÁL	NADM.V.
	PRES.1. LOK. :	2. :	3. :	4. :	5. :	6. :	URČIL autor záp. ?:	mater. [] dr.zd. []

F	KULTURA	PŘESNÁ DATACE				AKTIVITA	AREÁL	NADM.V.
	PRES.1. LOK. :	2. :	3. :	4. :	5. :	6. :	URČIL autor záp. ?:	mater. [] dr.zd. []

G	KULTURA	PŘESNÁ DATACE				AKTIVITA	AREÁL	NADM.V.
	PRES.1. LOK. :	2. :	3. :	4. :	5. :	6. :	URČIL autor záp. ?:	mater. [] dr.zd. []

H	KULTURA	PŘESNÁ DATACE				AKTIVITA	AREÁL	NADM.V.
	PRES.1. LOK. :	2. :	3. :	4. :	5. :	6. :	URČIL autor záp. ?:	mater. [] dr.zd. []

I	KULTURA	PŘESNÁ DATACE				AKTIVITA	AREÁL	NADM.V.
	PRES.1. LOK. :	2. :	3. :	4. :	5. :	6. :	URČIL autor záp. ?:	mater. [] dr.zd. []

Obr. 4. Formulář Zpráva o archeologické akci – strana 2.

8. Státní archeologický seznam ČR – informační systém archeologických nalezišť

(Dara Baštová, Lenka Krušinová, Zuzana Sklenářová, Petr Volfík)

8.1 Úvod

Změněná společenská situace po roce 1989, z níž vyplynula složitější struktura majetkoprávních poměrů, postavila archeologickou odbornou veřejnost před požadavek jasného stanovení objektu svého zájmu jak na úrovni daného stavu poznání, tak kontinuálním zachycováním jeho dalšího vývoje. Pracovníci orgánů státní správy (Okresních úřadů, Ministerstva kultury ČR) a archeologové, kteří připravují odborná vyjádření pro rozhodování v územním a stavebním řízení, nejvíce postrádají soupis archeologických nalezišť na území celého státu. Soupis je základním předpokladem efektivní ochrany archeologického kulturního dědictví.

Od druhé poloviny 80. let se věnuje zvýšená pozornost přípravě nadregionálních soupisů archeologických nalezišť. Dosavadní práce byly totiž zaměřeny jiným směrem: nekladly si nárok na celostátní měřítko, zpracovávaly buď menší region, vymezený správně nebo geograficky, případně řešily problematiku prostorového rozšíření nalezišť nebo předmětů určitého druhu, resp. období. Takové soupisy mohou sloužit a také slouží jako podklad pro sběr informací na celostátní úrovni. V rámci úkolu generální aktualizace organizované tehdejší Státním ústavem památkové péče a ochrany přírody (SÚPPOP) iniciovali jeho pracovníci částečnou revizi a doplnění informací o archeologických památkách, kterou provedl Archeologický ústav ČSAV v Praze. Pro soupis se mezi archeology vžil název „kartotéka MONUMIS“; v ARÚ byl převeden do podoby jednoduché databáze, používané pro účely archeologické památkové péče. Jeho značnou nevýhodou bylo a je to, že obsahoval většinou pouze archeologické nemovité kulturní památky, tj. archeologická naleziště zapsaná v dané době do státních seznamů nemovitých kulturních památek dle zákona č. 22/1958 Sb., o kulturních památkách, a zahrnoval pouze území Čech. Chybějící soupis archeologických nalezišť do určité míry suploval archiv nálezových zpráv Archeologických ústavů AV ČR v Praze a v Brně, který však obsahuje pouze záznamy o archeologických akcích, pokud ovšem hlášení o akci bylo zasláno do ARÚ, což před začátkem platnosti zákona č. 20/1987 Sb., o státní památkové péči, zpravidla prováděli pouze jejich pracovníci; hlášení regionálních archeologů se často ukládala spolu s další dokumentací v archivech jejich institucí. To vše je navíc dostupné pouze v sídlech ústavů. Další informace o archeologických nálezech jsou uloženy v muzeích jako součást evidence sbírkových fondů a jejich dokumentace.

Tento stav rozložení informací o archeologických nalezištích dává tušit konkrétní problémy, které nastávají při snaze o standardizaci struktury archeologických informací a organizaci jejich sběru. Jejich vyřešení je nezbytným předpokladem tvorby archeologických databází.

8.2 Principy a tvorba databáze archeologických nalezišť

Konkrétní ochranu archeologických nalezišť lze zajistit jen tehdy, podaří-li se vyhotovit soupis archeologických nalezišť v co nejkratším čase (max. 10 let) a zároveň pro co nejrozsáhlejší území – nejlépe celého státu – v souvislosti s formulací zákona č. 20/1987 Sb., o státní památkové péči ve znění pozdějších předpisů, v němž pojem „území s archeologickými nálezy“ zahrnuje veškerá archeologická naleziště. Nezasupitelnost tohoto pojmu a jeho výkladu je dána zásadním rozdílem, který pro ochranu archeologických nalezišť znamená znění zákona č. 20/1987 Sb., o státní památkové péči oproti zákonu č. 22/1958 Sb., o kulturních památkách. Předcházející předpis totiž definoval pojem „kulturní památka“ jako kulturní statek, který je dokladem historického vývoje společnosti (§ 2). Kulturní památky se z evidenčních důvodů zapisovaly do státních seznamů a chráněny byly i kulturní památky do těchto seznamů nezapsané (§ 7). Archeologické kulturní památky jsou často rozpoznatelné

až narušením jejich podstaty, to jest určením archeologických pramenů skrytých zpravidla pod povrchem terénu (výjimečně i nad povrchem, např. v zásepech kleneb a stropů). Jejich získání archeologickým výzkumem je podmíněno současným zásahem a zničením části informačního potenciálu naleziště. Přestože archeologické nálezy jsou movité a nemovité věci, příp. jejich soubory, které jsou významnými doklady historického vývoje, životního způsobu a prostředí společnosti od nejstarších dob do současnosti (§ 2 zákona č. 20/1987 Sb.), od účinnosti tohoto zákona je prohlášeno MK ČR a do Ústředního seznamu kulturních památek zapsáno zatím minimální množství z celkového počtu dosud známých archeologických nalezišť.

Skutečnost, že v naší republice dosud neexistuje celostátní soupis archeologických nalezišť, je mnohaletým dluhem české archeologické obce její pramenné základně. Po r. 1989 se praxe zásadním způsobem změnila, investoři pod tlakem finančních nároků usilují o co nejrychlejší průběh akcí a archeologické památky pro ně představují jak nevídané zdržení prací, tak růst nákladů.

Každý pokus o vytvoření celostátní databáze území s archeologickými nálezy je v první fázi vždy především kompromisem mezi obrovským množstvím známých, ale nechráněných nalezišť a akutní potřebou mít co nejdříve k dispozici údaje o jednom konkrétním nalezišti. Tato potřeba vedla ke zrodu myšlenky Státního archeologického seznamu ČR. Archeologický ústav AV ČR v Praze přistoupil k problému archeologické databáze Čech z jiného směru: za základní evidenční jednotku považuje archeologickou akci (Kuna-Křivánková-Krušinová 1995, 10). Takováto struktura databáze však nevyhovuje potřebám archeologické památkové péče (jedině jako podklad pro další zpracování), protože několik záznamů o akci se může týkat jediného území s archeologickými nálezy. Snaha o naplnění databáze dostatečným množstvím dat vede k přednostnímu využívání údajů s přesnou prostorovou identifikací prozkoumaných ploch, které však tvoří pouze část celkového objemu dostupných údajů. Starší, méně přesné údaje se přitom nerevidují a nemohou tak sloužit ochraně a záchraně archeologických nálezů, kde se nelze vyhnout vymezení rozsahu naleziště.

Specifikum Státního archeologického seznamu ČR je dáno zaměřením na archeologická naleziště, v dikci památkového zákona „území s archeologickými nálezy“. Můžeme je rozdělit na čtyři kategorie:

- I. území s pozitivně prokázaným výskytem archeologických nálezů,
- II. území, na němž dosud nebyl pozitivně prokázán výskyt archeologických nálezů, ale určité indicie mu nasvědčují; pravděpodobnost výskytu archeologických nálezů 51-100% (svědectví písemných pramenů, výsledky geofyzikálního průzkumu, letecké prospekce a dalších nedestruktivních metod, těsná blízkost UAN kategorie I.),
- III. území, na němž nebyl dosud rozpoznán a pozitivně prokázán výskyt archeologických nálezů a ani mu nenasvědčují žádné indicie, ale jelikož předmětné území mohlo být osídleno či jinak člověkem využito, existuje 50% pravděpodobnost výskytu archeologických nálezů (veškeré území státu kromě kategorie IV),
- IV. území, na němž není reálná pravděpodobnost výskytu archeologických nálezů (veškerá vytěžená území – doly, lomy, cihelny, pískovny apod., kde byly odtěženy vrstvy a uloženy nad geologickým podložím).

Z tohoto pojetí jednoznačně vyplývá, že databáze údajů o archeologických akcích vzhledem k výše uvedenému nedostatku nemohou soupis archeologických nalezišť nahradit; zachycují pouze výběr z jejich celkového počtu určený archeologickou aktivitou v minulosti, a navíc plochy, které dokumentují, byly mnohdy již vytěženy archeologickým výzkumem a jde tedy vlastně o území s archeologickými nálezy kat. IV.

Je zřejmé, že i pojetí databáze Státního archeologického seznamu má své slabé místo, které se ukrývá v riziku nepřesnosti vymezení rozsahu konkrétních území s archeologickými nálezy: nutí totiž zpracovatele, aby na základě vlastních znalostí a zkušeností interpretoval jednotlivé indicie k jeho určení. Výsledkem je pracovní zjednodušení, které nicméně umožňuje znázornění území s archeologickými nálezy v dostatečně konkrétním přehledném měřítku a je pro účely archeologické památkové péče dostačující. Uvedenou nevýhodu vyvažuje také skutečnost, že proces tvorby databáze nelze nikdy považovat za ukončený – údaje o známých nalezištích mohou být kdykoli revidovány a kdykoli může být doplněn záznam týkající se naleziště nově objeveného. Tato možnost však musí být zároveň předpokladem a požadavkem na kvalitu databáze: není-li její struktura dostatečně flexibilní a schopná průběžné aktualizace, její obsah v krátké době zastarává.

Využití databáze území s archeologickými nálezy k řešení zásadních otázek vědy je zcela zřejmé: soubor údajů o jednotlivých nalezištích propojený s podrobnou mapou, na níž je vyznačena jejich poloha a rozsah, poskytuje – můžeme-li jej v dané chvíli považovat za relativně úplný – jedinečné východisko pro studium vývoje a struktury osídlení regionu nebo celého území státu. Tyto možnosti značně rozšiřují analytické nástroje rychle se rozvíjejícího oboru geografických informačních systémů (GIS).

Využití pro potřeby ochrany archeologického kulturního dědictví je dalším cílem, k němuž by mělo vytváření jakékoli databáze archeologických nalezišť směřovat od samého počátku. Databáze spolu s vyznačením území s archeologickými nálezy do mapy, převedená do vrstvy informací v geografickém informačním systému, slouží k vyhledávání možných střetů zájmu archeologické památkové péče s hospodářskými zájmy v území již ve stadiu záměru (projektu) porovnáním „archeologické“ mapy s předpokládaným umístěním stavby nebo jiného provozu spjatého se zemními pracemi. V praxi může tento nástroj účinně používat k ochraně a záchraně archeologických nalezišť pouze výkonný orgán státní správy, tj. v současné době okresní úřad, v úzké spolupráci s archeologickým pracovištěm.

Databáze by měla obsahovat základní charakteristiky uspořádané v okruzích vycházejících ze zájmů a konkrétních potřeb uživatelů. Tvorba archeologické databáze je založena na spolupráci archeologů a inženýrů; struktura základní informační věty se odvíjí od sladění potřeb uživatelů a možností softwarových nástrojů. Intenzita těchto přípravných kroků by neměla být podceňována zvláště tehdy, je-li databáze nedílnou součástí informačního systému nadregionálního rozsahu, který vyžaduje vyřešení správy dat, tj. jejich údržbu, aktualizaci a doplňování jak na regionální, tak na celostátní úrovni. Teprve po dosažení konsensu zúčastněných stran (archeologů, inženýrů, uživatelů) na struktuře základní informační věty a na volbě SW nástrojů lze informační systém zároveň koncipovat pro využití v náročném prostředí GIS.

8.3 Struktura databáze

S ohledem na výše uvedenou nutnost sladit zájmy zúčastněných stran vycházela struktura databáze archeologických nalezišť vyvíjené pro účely ochrany v co největší možné šíři z položek, které již zavedl v databázi ARCHIV Archeologický ústav AV ČR v Praze. Zásadní odlišnost pojetí evidenční jednotky, úpravy a doplňky struktury základní informační věty si vynutilo celostátní zaměření projektu SAS ČR (zejména doplnění hesel kultur rozšířených na území Moravy a Slezska), zohlednění celostátních projektů zaměřených na ochranu movitých archeologických nálezů, resp. sbírkových předmětů, ohled na potřeby uživatelů na úrovni okresních úřadů a regionálních archeologických pracovišť a zejména záměr vytvořit databázi jako nedílnou součást tvorby uživatelské aplikace v prostředí GIS.

Základní evidenční jednotkou databáze vyvíjené v projektu SAS ČR je archeologické naleziště. Území s archeologickými nálezy je povinně definováno číslem mapového listu ZM 1:10 000 a pořadovým číslem archeologického naleziště vyznačeného na mapovém listě. Složením těchto dvou povinných údajů program automaticky vytvoří pořadové číslo SAS. Nejde o klasické pořadové číslo, tento postup nicméně umožňuje velmi rychle podle něj identifikovat příslušné území s archeologickými nálezy a navíc se výrazně snižuje nebezpečí vzniku duplicitních čísel. Pro registraci zpracovávaných území a řešitelů a pozdější vyhledávání bloků informací správce celostátních dat je vyplňována položka, která označuje ucelený soubor záznamů a v níž je ve zkratce určeno území nebo zpracovatel a rok, kdy byl uvedený soubor zpracován.

Pro potřeby státní správy, má-li být ochrana nalezišť účinná, lze pracovat jediné se stávající sítí katastrů (vodítkem je územní identifikační registr) a obecně se stávajícím správním uspořádáním.

Databáze musí řešit i případy, kdy území s archeologickými nálezy leží na rozhraní dvou nebo více katastrálních území; je třeba rozhodnout, které se bude považovat za hlavní a naleziště pod ním bude vedeno, aby se záznamy zbytečně nemnožily. Vzhledem k tomu, že v jednom okrese může existovat více katastrálních území stejného jména, osvědčilo se zejména pro manipulaci s kartami uvádět také příslušnou administrativní obec (počítač se orientuje podle kódů).

Pro podchycení tradičního jména naleziště, obvykle odlišného od výše uvedených administrativně správních údajů, je žádoucí zřídit položku, ve které lze uvést jméno trati, hradu, zaniklé středověké

vesnice (např. Týřov, Svídna, Holubí hlava, Leskoun apod.). Velmi stručný popis polohy lze uplatnit v samostatné položce. Zaznamenání stávajícího stavu povrchu naleziště (porost, způsob obdělávání, zastavba) je důležitým východiskem pro další strategii ochrany – pokud je např. na významném nalezišti pole, je možno při přípravě územního plánu uplatnit požadavek vynětí ze zemědělského půdního fondu, zatravnění, stavební uzávěry apod. Jednou z charakteristik umístění naleziště v území je nadmořská výška, je vhodné uvést výškové rozpětí v prostoru naleziště nebo jeden údaj u bodového nálezu.

Podle zavedeného standardu určují prostorové vymezení naleziště položky „druh mapy“, „tvar naleziště“ definovaný souřadnicemi odměřovanými od západních a jižních sekčních čar mapového listu, „přesnost lokalizace“ vyjadřující, s jakou přesností je poloha naleziště určena. Souřadnice naleziště je nutno uvádět za sebou tak, aby jejich sled vytvářel linii obvodu naleziště. Nemohou být zapsány nahodile, osvědčuje se maximálně 12 párů koordinátů v mm. Jestliže se naleziště rozkládá na více mapových listech, je zvykem koordináty odměřovat od sekčních čar západnějšího a jižnějšího mapového listu, takže výsledné hodnoty souřadnic jsou pak součtem hodnot naměřených na obou listech. Tím je dáno jednoznačné určení naleziště na hlavním mapovém listě, k němuž tedy nemusí existovat více záznamů.

V databázi by se mělo pro kontrolu u každého záznamu evidovat, kdo jej vypracoval a odpovídá za správnost určení naleziště a všech ostatních údajů (vždy archeolog), dále kdo převedl záznam do aplikace, vhodné je však dokumentovat i jména těch, kdo naleziště objevili a kdo na něm prováděli výzkumy nebo průzkumy.

Vyplňování databází naráží na rozdílné vyjadřovací zvyklosti a jazykové znalosti zpracovatelů jednotlivých záznamů. Zejména v archeologii, kde se silně projevuje subjektivita jednotlivců, může terminologickým nedorozuměním zabránit pouze užívání jednotného hesláře. Protože však předmět této vědy obsahuje mnoho proměnných, nelze se obejít bez „Poznámky“, do níž zpracovatel může psát libovolný text nevázaný na úzký okruh povolených výrazů.

Databáze vytvářená pro účely památkové ochrany by neměla postrádat údaj, zda v době zápisu je naleziště nějakým způsobem chráněno na základě nějaké právní normy, ať již samostatně nebo jako součást většího celku (např. památkové rezervace, památkové zóny), a to včetně území ochrany přírody. Jestliže zpracovatel doporučuje zajistit ochranu (u lokality zatím nechráněné) nebo změnit stávající míru ochrany, měl by mít možnost svůj názor uvést. Podle tohoto záznamu je pak možno se orientovat při přípravě návrhů na prohlášení nalezišť za kulturní památky. U chráněných je zapotřebí uvést také rejstříkové číslo kulturní památky, zda bylo kolem této památky vyhlášeno ochranné pásmo nebo zda archeologické naleziště leží v ochranném pásmu památky jiné; pokud ano, kdy a kým bylo vyhlášeno.

Určitou část informací zpracovatel čerpá z literatury, kterou charakterizuje běžná skladba bibliografických údajů, případně pro účely databáze redukováná – autor (vyplatí se uvádět iniciály křestního jména), rok vydání, název monografie, u článku název a ročník časopisu, v němž byl otištěn (není nezbytné uvádět název článku a často stačí i zkrácená podoba názvu monografie) a stránkový rozsah.

Pocházejí-li údaje z náleзовých zpráv nebo hlášení, uvádějí se jejich čísla jednací a název instituce, ve které je náleзовá zpráva pod uvedeným číslem jednacím uložena.

Pro archeologickou odbornou veřejnost je samozřejmé při tvorbě databáze k jakémukoli účelu uvést časové určení; pak je třeba specifikovat, ke kterým položkám informační věty se datace vztahuje.

Koncepce každé databáze se musí vyrovnat se základním principem vztahu časového určení nálezů k prostorovému vymezení rozsahu jejich výskytu. Buď bude zachycovat striktně pozitivní archeologická zjištění, nebo bude archeologické nálezy interpretovat, tj. dávat je do souvislostí, jsou-li zřejmé nebo odhadnutelné na základě současného stavu poznání. Výsledků prvního přístupu nelze bez následné interpretace využít pro účely archeologické památkové péče. Zároveň při plnění takové databáze mohou unikat neúplné vstupní informace, které indikují mnohdy významná archeologická naleziště, rozsáhlá a zasluhující ochranu, ale přitom nepřesně časově zařazená nebo dosud nedatovaná.

Pro určení doby a trvání osídlení lokality, jeho druhu a charakteru osídlení se stal obvyklým standard systému Archiv, který zahrnuje datování a kulturní určení podle možnosti, uvedení aktivit, které jsou nálezy doloženy, určení, jaké areály v době osídlení na lokalitě existovaly, jméno odborníka, který stanovil přítomnost dané komponenty na lokalitě a údaj, podle jakého druhu zdroje ji stanovil (protože svědectví různých informačních zdrojů má různou hodnotu).

Je-li pro danou komponentu určena nadmořská výška přesněji, než je nadmořská výška celého naleziště, je dobré ji uvést zvlášť.

Oddíl komponent je možno rozšířit o rubriku, do níž lze ke každé z nich upřesnit její prostorové umístění na vícekrát osídlené nebo využívané lokalitě, pakliže se územní rozsah jednotlivých složek liší. To pak umožňuje specialistům podchytit a řešit určité tematické okruhy a otázky na základě provázání komponenty s nemovitými objekty a movitými nálezy.

Z důvodů evidenčních, odborných i památkářských je na místě zaznamenávat zjištěné nemovité objekty, ať jde o viditelné nadzemní relikty staveb nebo o podzemní, zničené stavbou nebo archeologickým výzkumem. Nemovité nálezy jsou určeny svým druhem, specifickou charakteristikou a počtem. Nemovité a movité nálezy se sledují ve vazbě na danou komponentu.

Vzhledem k možnosti následného napojení databází movitých archeologických nálezů a archeologických sbírkových předmětů, plněných jednotlivými archeologickými institucemi, byl Státní archeologický seznam ČR koncipován tak, aby mohl registrovat i movité nálezy: charakterizuje je jejich druh, materiál, počet, popř. další vlastnosti, a uložení předmětů sledované až na úroveň evidenčních čísel, přičemž se rozlišuje, zda jde o osobu právnickou, tj. instituci, nebo fyzickou. Tu identifikuje jméno, název, adresa, PSČ, sídlo nebo bydliště. Pokud se místo uložení nálezů liší od adresy uvedené osoby, je vhodné je uvést zvlášť, rovněž jsou-li depozitáře organizovány dle časového zařazení nálezů. Pro základní přehled o stavu zpracování a objemu movitých nálezů lze uvést, zda předměty jsou umyty, laboratorně zpracovány, množství může být uvedeno pouze orientačně – např. počet beden, sáčků apod. Tyto údaje jsou důležité u nalezišť, která byla předmětem dlouholetých systematických nebo záchranných archeologických výzkumů, jejichž výsledky nejsou dodnes zveřejněny.

Rubriky určující majetkové právní vztahy, nezbytné pro efektivní uplatnění zájmů archeologické památkové péče, odpovídají způsobu definice těchto rubrik v informačním systému památkové péče vyvíjeném v odboru evidence a dokumentace SÚPP. Jde o parcely a adresy vlastníků nebo uživatelů, katastrální území a typ parcely, který souvisí s rozdělením parcel v některých katastrálních územích na stavební a pozemkové. V návrzích na prohlášení archeologických nalezišť za kulturní památky je třeba uvádět aktuální údaje. Počítá se i s navázáním dat jejich prostorového vymezení v katastrálních mapách na záznamy v databázi. Na tento krok pak může bezprostředně navazovat vypracovávání pasportů nemovitých kulturních památek stávajících i navrhovaných.

8.4 Aplikace SAS

Pro automatizovaný sběr dat v rámci projektu SAS ČR byla vytvořena aplikace, při jejíž tvorbě jsme vycházeli z těchto základních požadavků :

- vzhled a funkční podobnost s papírovým formulářem
- jednoduchost obsluhy
- integrace používaných heslářů
- datová kompatibilita s aplikací „Ústřední seznam kulturních památek“
- normalizovaná struktura databáze
- široká nabídka kritérií pro výběr dat z databáze
- připravenost dat pro přímé využití v prostředí GIS

Jako prostředí pro běh aplikace byl zvolen operační systém Microsoft Windows. Tím byl zajištěn základní uživatelský komfort ve standardním prostředí známém převážně většině uživatelů osobních počítačů.

K vlastnímu vývoji aplikace byl zvolen nástroj pro tvorbu databází firmy Borland Paradox pro Windows, a to v aktuální verzi 5.0. Tento produkt byl vybrán především z následujících důvodů :

- osvědčil se při tvorbě jiných databázových aplikací ve Státním ústavu památkové péče, především v aplikaci Ústřední seznam kulturních památek – tím je také zajištěna přímá kompatibilita databáze SAS s touto aplikací
- je to nástroj na obsluhu jednoduchý, ale přitom velmi výkonný – obsahuje také tzv. Runtime modul, který umožňuje spouštění aplikací vytvořených v Paradoxu i na počítačích nevybavených tímto komplexním programem.

8.4.1 Normalizace databáze, referenční integrita

Při tvorbě aplikace byl hlavní důraz kladen na vytvoření kvalitního datového modelu struktury databáze, který by se co nejvíce přibližoval tzv. normalizovanému stavu.

– normalizace databáze je vlastně proces rozdělení rozsáhlé databáze do několika tabulek, které jsou pak vzájemně relačně provázány. Databáze se normalizuje především ze dvou důvodů:

Normalizace databáze zabraňuje vzniku duplicitních dat a tím šetří čas a místo na disku a v některých případech zamezuje vzniku chyb. Sledujeme-li např. v takové databázi parcely, na kterých se nachází určité naleziště, nebudeme muset do každého záznamu s parcelou vkládat jméno a adresu jejího vlastníka. Zabrало by to příliš času a snadno bychom udělali chybu např. tím, že bychom použili pro stejnou osobu různě aktuální adresu. Lépe je tedy udržovat takovou tabulku vlastníků zvlášť, pak i případná aktualizace adresy se provádí pouze jednou.

– normalizace rovněž usnadňuje výběr potřebných informací z databáze. Do tabulky se základními údaji o nalezištích nebudeme zřejmě ukládat jednotlivá katastrální území, do kterých dané naleziště zasahuje, protože v takovém případě by potom bylo z této tabulky obtížné vygenerovat seznam všech zatím archeologicky dotčených katastrů, nebo naopak najít všechna naleziště, která zasahují do námi zvoleného katastru. Jelikož je i počet dotčených katastrálních území u jednotlivých nalezišť různý, nevěděli bychom v případě umístění všech údajů do jedné tabulky, jaký počet polí pro katastry rezervovat. U naleziště s jedním katastrem bychom zbytečně plýtvali místem rezervovaným na disku, v případě více dotčených katastrů by se nám zase jejich výčet nemusel do tabulky vejít.

Neméně důležitou částí návrhu struktury databáze je volba tzv. primárního klíče v každé definované tabulce. Pod pojmem primární klíč rozumíme pole nebo kombinaci polí, které se používá k jednoznačné identifikaci záznamů v tabulce. Tabulka je pak automaticky uspořádána podle hodnot v primárním klíči a do tabulky nelze vložit více záznamů s toutéž hodnotou primárního klíče.

V neposlední řadě je třeba nadefinovat vazby-relace mezi jednotlivými tabulkami. S tím úzce souvisí problém se zajištěním tzv. referenční integrity. Je to proces, který zabraňuje vzniku tzv. „osiřelých“ záznamů, kdy se např. v databázi vyskytuje údaj o literatuře, ale chybí záznam o nalezišti, ke kterému se vztahuje. Paradox pro tento účel nabízí nástroje, které celý proces automatizují.

Definujeme-li vazbu referenční integrity, deklarujeme, že každému záznamu v jedné tabulce (tzv. dětská tabulka) musí být přiřazen právě jeden záznam v jiné tabulce (rodičovské tabulce). Jakmile v Paradoxu takovou vazbu nadefinujeme, automaticky zajistí, aby nedošlo k událostem, jež by vedly ke vzniku „osiřelých“ záznamů, tj. záznamů v dětské tabulce, které nemají relační záznamy v rodičovské tabulce. Tyto „zakázané“ události jsou následující:

– do dětské tabulky není možno vložit záznam, ke kterému neexistuje relační záznam v rodičovské tabulce

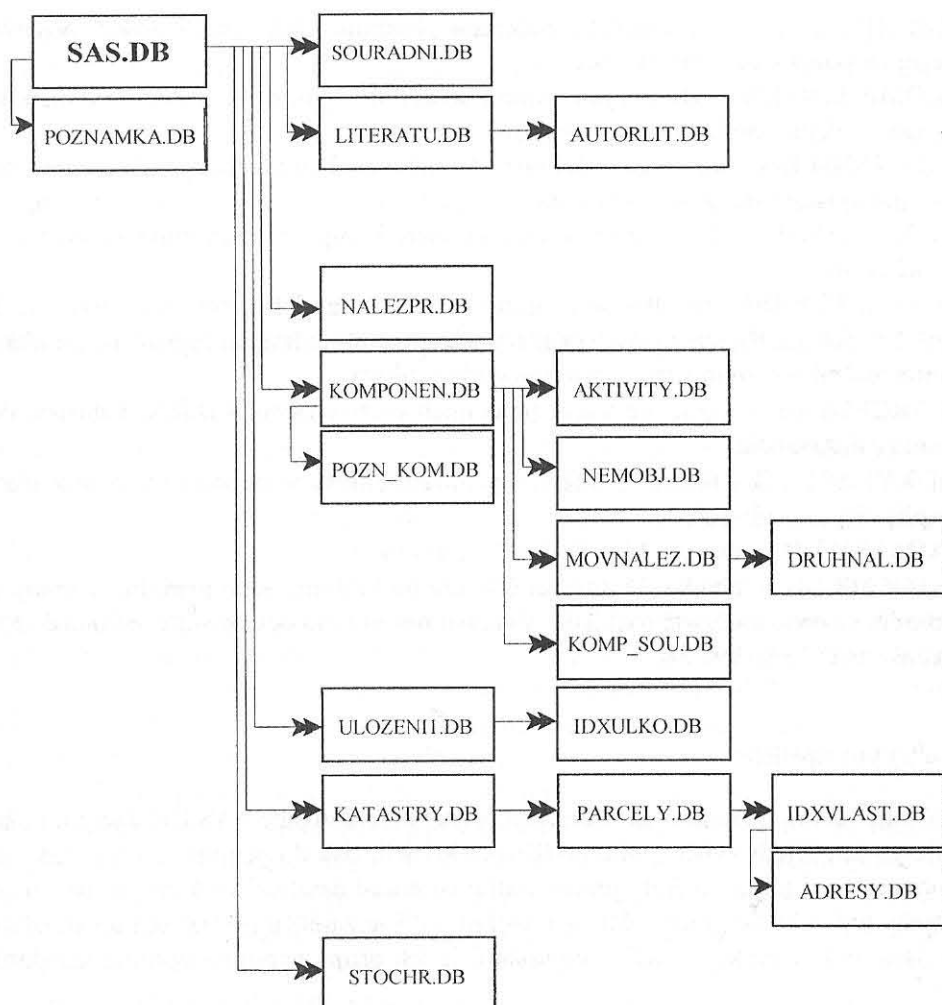
– z rodičovské tabulky nelze smazat záznam, ke kterému existují relační záznamy v dětské tabulce

– v rodičovské tabulce nelze provést změny v polích klíče, aniž by byly provedeny příslušné změny v relační dětské tabulce. Paradox nabízí v tomto případě dvě možnosti : buď může takovýmto změnám zabránit, nebo je může automaticky provést ve všech relačních záznamech.

Je zřejmé, že důsledné využití principu referenční integrity vede k zajištění vysoké bezpečnosti a konzistence dat.

8.4.2 Datový model databáze SAS

Po zhodnocení všech výše uvedených požadavků byl pro databázi SAS navržen datový model uvedený na obr.1. V tomto modelu jsou zobrazeny jednotlivé tabulky databáze SAS a jejich vzájemné vazby, přičemž jednoduchá šipka znázorňuje vazbu 1:1 (k jednomu záznamu v rodičovské tabulce existuje maximálně jeden záznam v tabulce dětské), šipka dvojitá vazbu 1:N (k jednomu záznamu v rodičovské tabulce může existovat několik záznamů v dětské tabulce). Probereme si tedy jednotlivé tabulky a jejich vzájemné vazby:



Obr.1.

- * SAS.DB – hlavní tabulka celé databáze. Obsahuje základní údaje o nalezišti, především identifikačního charakteru. Primární klíč tvoří tzv. pořadové číslo SAS – jednoznačný identifikátor naleziště, který je vytvořen spojením čísla mapového listu ZM10 a pořadového čísla naleziště na tomto listu. Pomocí tohoto čísla jsou také realizovány všechny přímé vazby na další tabulky. Tabulka využívá hesláře okresů a hesláře institucí oprávněných k archeologickým výzkumům.
- * POZNAMKA.DB – tabulka obsahující ke každému nalezišti text neomezené délky uložený v tzv. memo položce, určená pro vkládání libovolných doplňujících informací, které nelze zařadit do jiných polí databáze.
- * SOURADNI.DB – obsahuje výčet všech souřadnic potřebných pro vymezení polygonu naleziště v rámci mapového listu ZM10.
- * LITERATU.DB – tato tabulka obsahuje ke každému nalezišti přehled literatury.
- * AUTORLIT.DB – obsahuje ke každé literatuře výčet jejích autorů.
- * NALEZPR.DB – seznam nálezových zpráv ke každému nalezišti.
- * KOMPONEN.DB – tabulka obsahující výčet archeologických komponent rozlišených na daném nalezišti. Využívá hesláře archeologických kultur a areálů.
- * POZN_KOM.DB – zde lze vložit ke každé komponentě doplňující text neomezené délky.
- * AKTIVITY.DB – obsahuje ke každé komponentě výčet aktivit s využitím hesláře.
- * NEMOBJ.DB – obsahuje ke každé komponentě výčet nemovitých objektů vyskytujících se na daném nalezišti, používá hesláře pro druh a charakteristiky objektů.
- * MOVNALEZ.DB – obsahuje ke každé komponentě výčet materiálů movitých nálezů použitých k identifikaci komponenty s využitím hesláře.

- * DRUHNAL.DB – ke každému materiálu obsahuje výčet všech druhů movitých nálezů a jejich počet s využitím hesláře.
- * KOMP_SOU.DB – tabulka pro uložení souřadnic vymežujících rozsah dané komponenty v rámci plochy celého areálu naleziště.
- * ULOZENI1.DB – obsahuje výčet institucí nebo osob, u kterých je uložen movitý materiál z daného naleziště, jeho počet a stav.
- * IDXULKO.DB – obsahuje informaci, ke které komponentě (komponentám) se dané uložení vztahuje.
- * KATASTRY.DB – tabulka obsahující výčet katastrálních území, na kterých se dané naleziště rozkládá. Využívá jako heslář tzv. Územně identifikační registr, tj. standard státního informačního systému pro prostorovou identifikaci.
- * PARCELY.DB – obsahuje výčet parcelních čísel v rámci každého katastru, do kterých naleziště zasahuje.
- * IDXVLAST.DB – tabulka sloužící k přiřazení adresy vlastníka či uživatele dané parcely z připojeného adresáře.
- * ADRESY.DB – adresář vlastníků či uživatelů parcel.
- * STOCHR.DB – tabulka již dosažených stupňů kulturní nebo přírodní ochrany území, na kterém se dané naleziště rozkládá. Využívá heslář typů ochrany dle zákona č. 20/1987 Sb. a zákona č. 114/1992 Sb.

8.4.3 Stav a vývoj aplikace

Na základě tohoto modelu byla vytvořena první verze aplikace SAS ČR, která umožňuje vkládání a aktualizaci dat, jejich výběr podle rozličných kritérií, tisk do připravených sestav, zálohování dat a jejich přenos mezi jednotlivými zpracovateli a centrální databází ve Státním ústavu památkové péče. Pro potřeby jednoduché práce s daty v prostředí GIS je zajištěn převod dat do standardního výměnného formátu .dbf, v praxi se však více uplatní jejich přímé napojení pomocí standardů ODBC nebo SQL.

Aplikace byla v průběhu roku 1996 zdarma distribuována na jednotlivá archeologická pracoviště spolupracující na projektu Státní archeologický seznam. V roce 1997 předpokládáme vývoj druhé verze aplikace, která bude řešit především následující požadavky:

- přechod do 32-bitového prostředí Windows95/NT
- možnost ukládání a prohlížení fotodokumentace
- doplnění o funkce pro výběr a přenos dat pro potřeby okresních úřadů v rámci vývoje uživatelské aplikace v prostředí GIS ArcView
- další zvýšení uživatelského komfortu

Specifickým úkolem při vývoji druhé verze bude sladit používané datové struktury s novými standardy státního informačního systému. Jde především o oblast územní identifikace (katastrální území, obec, okres, parcela) a oblast identifikace osob a organizací (adresy vlastníků a uživatelů parcel, uložení movitých nálezů). Problémem je, že většina těchto standardů ještě nemá konečnou podobu a prochází poměrně bouřlivým vývojem. Jejich zavádění do praxe ve státní správě je také zatím velmi nerovnoměrné.

8.5 Problematika řešení intravilánů

Samostatným problémem je registrace území s archeologickými nálezy v intravilánech historických sídel. Nejvýznamnější historická jádra měst byla prohlášena za městské památkové rezervace a památkové zóny, ovšem bez ohledu na rozsah území s archeologickými nálezy, který zhusta přesahuje jejich hranice. Sledování výskytu archeologických nálezů v těchto územích vyžaduje speciální strukturu informační věty pro předem definovaný soubor území. Tyto soubory lze volit podle daného stavu plošné ochrany (MPR, MPZ), z hlediska definice historických sídel, z hlediska stavu a počtu

těch, která z nich zatím byla archeologicky zkoumána, z hlediska potřeby registrace prozkoumaných ploch na území plošně chráněných území apod. Vzhledem k nutnosti přizpůsobení se měřítku mapových pokladů intravilánů, které jsou v současné době často k dispozici v digitálním tvaru (rastrová i vektorová data), si řešení tohoto úkolu vyžádá náročnou přípravnou fázi zaměřenou na porovnání výhodnosti volby jednotlivých hledisek, z nichž lze ke tvorbě databáze zaměřené na historická sídla přistoupit. Jelikož na toto téma existuje několik projektů, je žádoucí využít všech dosavadních zkušeností a dílčích výsledků.

8.6 Státní archeologický seznam jako informační systém

Specifičnost archeologického fondu, která spočívá v nemožnosti jednorázově definovat konečnou množinu archeologických nalezišť, je důvodem k vytváření samostatného informačního systému. V projektu Státní archeologický seznam ČR (SAS ČR) bylo zvoleno zvláštní řešení, které využívá analytických operací k modelování určitých závislostí. Analytickými nástroji disponují složité SW produkty v prostředí geografických informačních systémů (GIS): umožňují předpokládat výskyt dosud nezjištěných archeologických nalezišť a tím pomáhají snížit nebezpečí jejich zničení. Řešení nejmodernějšími prostředky v GIS bylo zvoleno po prověření přístupu k obdobným úkolům v zahraničí. Z důvodu rozmístění uživatelské verze PC ArcView na okresních úřadech v ČR a také proto, že disponuje požadovanými analytickými nástroji, vyhovoval konkrétně SW PC ArcInfo a ArcView. Základní evidenční jednotka databáze, tj. území s archeologickými nálezy, představuje jednak záznam v databázi, jednak má grafickou podobu v GIS, oba jsou vzájemně provázány v prostředí ArcView. Koncepce sběru dat vychází ze shody většiny archeologických pracovišť na nutnosti vytvoření celostátní archeologické databáze. Státní ústav památkové péče vyzývá archeologická pracoviště ke spolupráci a sjednává obsah a termíny smluv. Jejich předmětem je shromažďování odborných podkladů podle metodiky SAS ČR z území definovaného rozsahem mapových listů ZM 1:10 000. Cílem snažení je dosáhnout ochoty regionálních pracovišť k dalšímu podílu na aktualizaci a doplňování odborných podkladů jako odborného partnera výkonného orgánu státní správy, tj. okresního úřadu. S okresními úřady se uzavírají smlouvy o spolupráci a vzájemném spolupůsobení při plnění projektu SAS ČR, který sleduje zvýšení ochrany archeologických nalezišť. Součástí obsahu smluv je zajištění a sdílení digitalizovaných dat v rastrové nebo vektorové podobě a údajů o majetkoprávních poměrech na území s archeologickými nálezy. Všechna dostupná grafická a popisná data z daného regionu, průběžně aktualizovaná, budou využita v uživatelské aplikaci v prostředí PC ArcView instalované na okresních úřadech a partnerských archeologických pracovištích.

8.7 Literatura

- Bařtová, D. – Krušinová, L. – Sklenářová, Z. – Volfík, P. (1996) Státní archeologický seznam ČR. Uživatelská příručka. Praha.
- Biow, L. (1994) HELP ! Paradox for Windows, Brno.
- Kuna, M. – Křivánková, D. – Krušinová, L. (1995) ARCHIV 2.0. Systém archeologické databáze Čech. Uživatelská příručka. Praha.
- Ministerstvo hospodářství (1996) Standardy státního informačního systému České republiky. Praha.

Geografické informační systémy v archeologii

9. Geografické informační systémy (Milan Konečný)

9.1 Co jsou GIS? Definice a terminologie

Kolébku GIS je Kanada, kde takto byly nazvány v 60. letech tohoto století. Téměř současně se GIS objevily i v USA. Zkratka GIS znamená Geografické Informační Systémy, někteří autoři dávají přednost termínu geoinformační systémy. Slovo geografický je používán ve smyslu prostorový, odtud široce používaný termín prostorové informační systémy. Jako LIS (Land Information Systems) bývají označovány informační systémy o území. Termín geografické informační systémy (GIS) je jak v naší, tak i ve světové literatuře používán s naprostou převahou.

Geografické informační systémy zaznamenávají v teorii i praxi doslovný „boom“. Jejich rozvoji jsou věnovány obrovské finanční částky, ale na druhé straně jejich využití, a to buď samostatně nebo v poslední době stále častěji v kombinaci s jinými technologiemi také nové finanční prostředky vytváří. Na projekty GIS jsou zaměřeny četné aktivity Evropského společenství, i vyspělých zaoceánských zemí, zejména USA, Kanady, Austrálie či Nového Zélandu.

9.1.1 Proč jsou GIS tak žádané?

Prostorový pohled na data může často vyvolat nové pohledy a nová vysvětlení. Může ukázat souvislosti a závislosti, které s použitím pouze klasických výzkumných metod nebyly jednak patrné, jednak tradiční metody nebyly vůbec schopné nebo dostatečně účinné pro pochopení, správu, řízení a využití rozmanitých aktivit nebo přírodních a ekonomických zdrojů. Například pro zvolení vhodného místa pro stavbu hotelu, zdravotního střediska nebo prodejny automobilů v rámci nějakého regionu (okresu) může být vyžadováno porovnání až 200 tematických map. Tradičními postupy je to téměř nepředstavitelné, ale s pomocí GIS se taková operace může stát běžnou záležitostí. To vše za předpokladu, že v databázi máme uloženy kvalitní tematické informace.

GIS také umožňuje rychlý přístup k administrativním záznamům, např. vlastnictví realit, údajům o daních, komunálních sítích a potrubích, která jsou doplněna údaji o geografické poloze.

GIS vyvolal velký zájem o využití moderní výpočetní techniky v těch oblastech, které byly dříve považovány z hlediska využití nových technologií spíše za konzervativní. Dnes mohou, např. v archeologii, geografii, geologii, ale také kriminalistice či marketingu, poskytnout precizní nástroj pro zpracování a interpretaci prostorových informací.

Pro řadu lidí jsou fascinující mapové a jiné grafické výstupy. Po prvních letech využívání počítačové grafiky, kdy mapy byly transformovány do rozmanitých nepěkných čtvercových, trojúhelníkových nebo šestiúhelníkových sítí, je dnes možné vytvořit takové mapy, které laik a často i odborník nerozezná od klasických. Navíc, díky možnostem „desk top mapping“, lze vytvářet mapy v širokém rozsahu i v domácích podmínkách, za využití dostupných technických prostředků.

Grafické výstupy jsou též významně rozvíjeny a doplňovány různými vizualizačními prostředky a multimédií. I ty jsou stále více využívány spolu s GIS, jakožto významným nástrojem řízení, rozhodování a přijímání řešení.

GIS můžeme rozumět systém počítačového hardware, software a postupů vytvořených pro podporu sběru, řízení, manipulace, analýzy, modelování a zobrazování prostorově vztahovaných dat pro řešení složitých plánovacích a řídicích problémů.

GIS je systém obsahující prostorově vztahovaná data, která mohou být analyzována a konvertována do podoby informace podle plánovaného specifického využití nebo aplikací. Klíčovým rysem GIS je analýza dat za účelem produkce nové informace.

GIS také může být jakýkoliv informační systém pro řízení, který musí:

- sbírat, ukládat a vyhledávat informace založené na prostorovém umístění,
- identifikovat místa v rámci určité oblasti zájmu, za použití specifických kritérií,

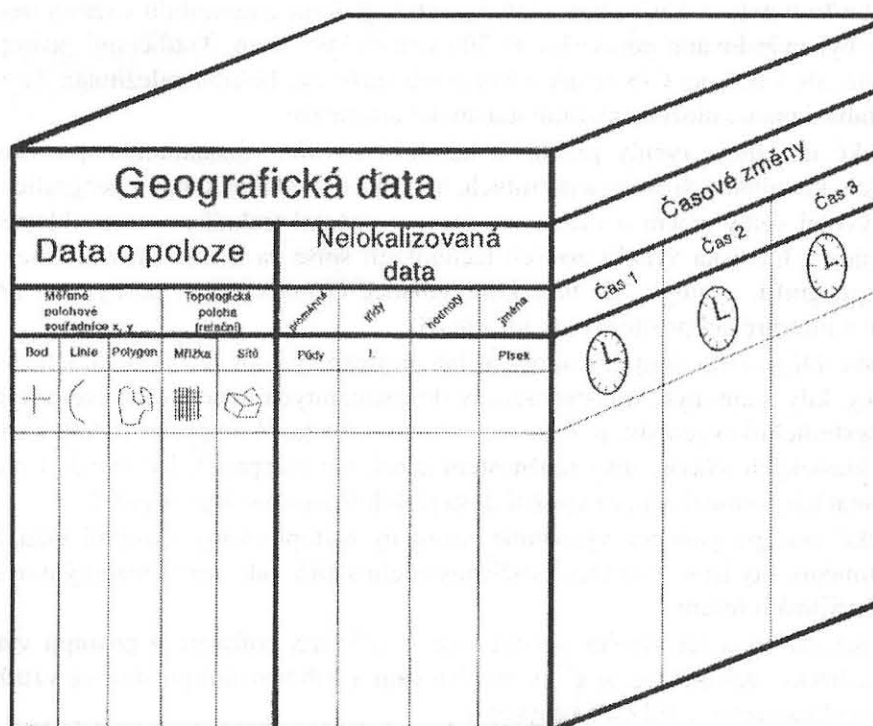
- zkoumat vzájemné vztahy mezi datovými soubory v určité oblasti zájmu (např. v geografii, archeologii, geologii, marketingu, aj.),
- prostorově analyzovat data vztažená k oblasti zájmu za účelem přijímání rozhodnutí v této oblasti,
- usnadňovat výběr a dodávání dat do specifických aplikačních analytických modelů, pomocí nichž je možné odhadnout důsledky zvolených alternativ na vybrané prostředí,
- zobrazovat vybrané prostředí jak graficky, tak i numericky, a to jak před nebo po provedení analýzy.

Nejčastější chybou při nasazování GIS je, že není akceptován jako celistvý systém. Často je snaha vytvářet pouze jednotlivé jeho části. Někdy je GIS omezen pouze na systémy používané pro digitální kartografii resp. počítačovou grafiku. Tyto systémy ale většinou nejsou použitelné pro zpracování velkých objemů dat a zejména nemají analytické funkce, které jim umožňují zpracovávat data rozmanitých tematik, jež jsou uložena v různých tematických vrstvách databáze.

Klíčovým prvkem a současně velmi často největším problémem každého GIS jsou data a informace. Na obr. 1 uvádíme jeden z prvních příkladů charakteristik dat v GIS.

Sběr dat, jejich údržba a aktualizace tvoří až 80% nákladů na celý GIS. Kvalita dat a způsob jejich sběru je prozatím příliš rozdílný jak v rámci jednotlivých států a jejich institucí, tak i z hlediska kontinentálního resp. globálního. Proto je otázka kvality dat, jejich možným chybám či tvorbě metadata, standardů, legislativním aspektům nakládání s daty a dalším problémům, věnována v poslední době velká pozornost.

Tři konceptuální složky geograficky založených informačních systémů

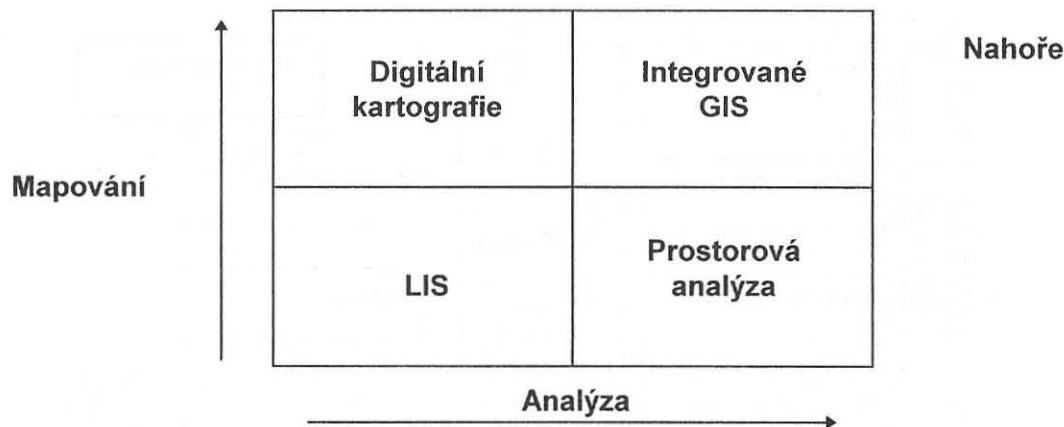


podle Dangermond J., 1983

Obr. 1.

9.1.2 Integrované systémy

Za integrovaný GIS považujeme takový systém, který má vysokou úroveň schopností pro mapování a prostorovou analýzu. Obr. 2 znázorňuje, jak v současnosti čtyři užívané termíny pro GIS: digitální (počítačová resp. automatizovaná kartografie), informační systémy o území, prostorové analýzy, mohou být uspořádány podle dvou proměnných – schopností mapování a analýzy.



Obr. 2. Klasifikace GIS. (podle Grimshaw D.J. 1994).

9.1.3 Co jsou geografická data?

Jádrem GIS jsou geografická data. Od jiných druhů dat se odlišují tím, že popis objektů reálného světa zahrnuje jejich **polohu** s ohledem na známý souřadnicový systém, jejich **atributy**, které nemají polohový charakter (jméno majitele pozemku, cena pozemku, barva, typ zemědělského využití, apod.) a konečně zaznamenávají **prostorové (topologické) vztahy**, jež popisují, jak jsou jednotlivé jevy propojeny.

Dalšími charakteristikami geografických dat (kromě lokalizace), je jejich zpravidla velký objem, rozměrovost a kontinuita.

9.1.4 Složky GIS

Složky GIS dokumentuje obr. 3. Atributy i prostorová data jsou sbírána pokud možno tak, aby umožňovala co nejlepší vyjádření reálného světa. Podívejme se na jednotlivé složky uvedeného modelu podrobněji.

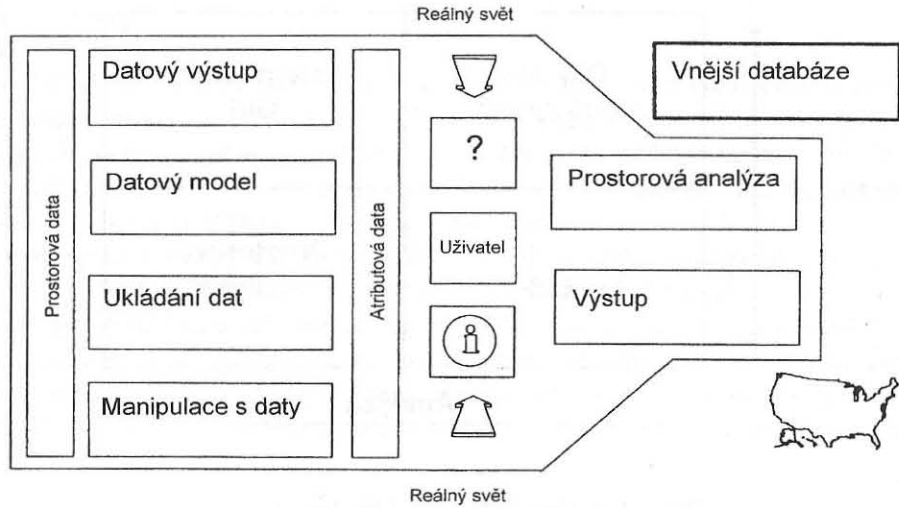
9.1.4.1 Vstup dat

V současnosti jsou dostupné dále uvedené metody vstupu dat. Poněvadž rozsah této publikace je omezen, doporučujeme P.T. čtenářům, aby si v případě zájmu vyhledali konkrétní literaturu o jednotlivých metodách a blíže se s nimi seznámili.

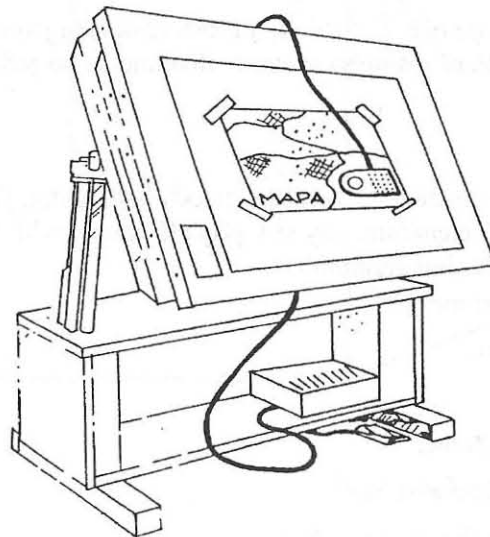
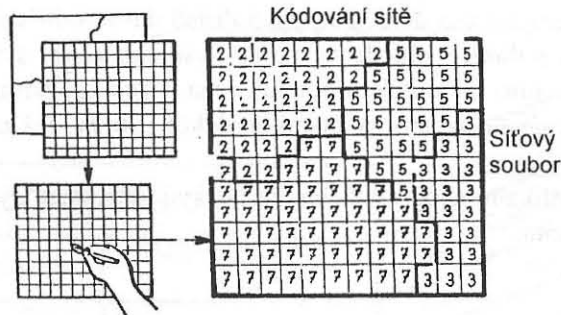
Jde o následujících šest metod:

- digitalizace (obr.4),
- skenování
- dálkový průzkum Země,
- GPS-globální poziční systémy
- přímé vstupy z geodetických měření
- fotogrammetrie

Pozn.: Poslední jmenovaná metoda je u nás v současnosti v útlumu a není ani dostatečně vyučována na univerzitách. Přitom s pomocí tzv. digitální ortofotografie lze dosáhnout až třetí třídy geodetické přesnosti. Tato dynamicky se rozvíjející část fotogrammetrie je využívána při pozemkových úpravách, zjišťování nových objektů v krajině a může být velmi prospěšná i pro archeologický výzkum. V současnosti tuto metodu v ČR aplikuje s dostatečnou přesností jen několik firem, v Brně např. Intermap CR s.r.o.



Obr. 3.



Obr. 4.

Data jsou pro využití v počítači upravována, proces je označován jako **geokódování**.

Geokódování je konverze prostorové informace do tvaru čitelného pro počítač. Geokódování zahrnuje jak vlastní proces, metody převodu, tak i koncepce, podle nichž jsou postupy používány. Týká se jak určení typu mapy a jejího měřítka, tak i hodnocení přesnosti.

Přesnost v tomto kontextu rozeznáváme dvojí, a to:

- **přesnost (accuracy)**, lze charakterizovat jako blízkost pozorování k hodnotám skutečným nebo těm hodnotám, jež jsou za skutečné pokládány. V případě mapové přesnosti půjde o přesnost grafických charakteristik map k jejich skutečné hodnotě nebo pozici. Můžeme říci, že např. mapa 1:50 000 má přesnost 1 m. Přesnost geografické informace má polohovou, tematickou a časovou složku.
- **relativní přesnost** obecně může být míra vnitřní konzistence polohových měření v daném souboru dat. Může jít o měření přesnosti jednotlivých rysů na mapě ve srovnání s jinými rysy na téže mapě. Např. jde o přesnost mapy vůči místní polohové síti (jež nemusí být vztažena k zemskému geoidu). Pro mnohé lokální účely je relativní přesnost významnější než přesnost absolutní.
- **absolutní přesnost** je měření polohy rysů na mapě ve srovnání s jejich polohou na povrchu Země. Jinými slovy jde o přesnost mapy ve vztahu k zemskému geoidu.
- **preciznost (precision)** je mírou schopnosti rozlišovat mezi velmi blízkými hodnotami. Jinými slovy, jde o přesnost s jakou pracuji. Např. mapu můžeme vektorizovat s velkou přesností.

Mapovací standardy přesnosti jsou obecně chápány jako akceptovatelná chyba, která nesmí být překročena. V případě kreslicích nebo zobrazovacích zařízení se přesnost vztahuje k toleranci na zobrazovaných nebo grafických charakteristikách, jež jsou relativní k originálnímu souřadnicovému souboru. Přesnost geografické databáze je dána také přesností jejího obsahu, např. chyby v zadání nebo vynechávky charakteristik, topologická integrita grafické a negrafické informace, apod.

Důležitým, efektivitu ovlivňujícím faktorem geokódování je pochopení geografických vlastností a souvislostí, které lze z geografických dat odvodit resp. určit. Převod dat do počítačem čitelného tvaru není tak velkým problémem. Problémem je, aby se z co největšího množství těchto dat daly vytvořit potřebné informace.

Geokódování se provádí ze dvou důvodů:

1. se speciálním mapovacím záměrem, např. zachycování linií, jež vytvářejí polygony. Kupř. hranice jednotlivých archeologických nalezišť v kontextu nalezišť v celém regionu či celého kontinentu.
2. obecně účelové geokódování, při němž sbíráme každý kousek informace o určitém území a do-
sazujeme si jej podle možných kartografických souvislostí. Tento postup má velkou potenciální důležitost, je obdobou významu existence topografických map pro dané území.

Geokóduje se zpravidla na dvou úrovních:

1. grafické prvky mapy jsou převáděny do čísel, mapa může být reprodukována pomocí metod počítačové kartografie. Mnoho současných databází existuje v této podobě;
2. kódujeme důležitou topologickou informaci spjatou s našimi daty. Pro potřeby GIS je zakódování topologie podstatné pro využití a uchování kartografických datových souborů.

Pro pochopení této odlišnosti uvádíme příklad symbolizace cest a řek. Jestliže digitální mapa obsahuje silnici a řeku, které se navzájem protínají resp. vedou souběžně, pak netopologickým přístupem bude, že obě nakreslíme přes sebe. Při zakódování topologie bychom měli poznat, že buď silnice kříží řeku mostem, kdy zakreslíme pouze kartografický znak silnice a přerušíme kartografický znak řeky resp. přidáme kartografický znak mostu.

9.1.5 Metadata

Data jsou tradičně úzkým místem při využití geoinformačních technologií. Jsou téměř vždy drahá, typově různorodá, zčásti nedostupná, neúplná nebo neaktuální či nedůvěryhodná. Pro odstranění uvedených nedostatků je potřeba vytvořit mechanismus, který zabezpečuje ochranu, správu a efektivní využití geografických dat. Nástrojem jeho realizace je vytvoření báze metadat.

Metadata jsou data o datech. Jsou jedním z prvků obecně zaváděné informační kultury a úspěšným předpokladem informační či geoinformační činnosti.

Báze metadat, jakožto nejefektivnější prostředek organizace metadat, je předurčená pro uspořádání a popis strukturních prvků jednotek uchování informace v její číselné i nečíselné podobě. Cílem je zajištění získávání a výměny informací mezi jejími držiteli (producenty, tvůrci) a těmi, kdo ji využívají (spotřebiteli, uživateli).

K problematice metadat lze přistupovat ze tří aspektů:

1. **Metadata jako prostředek inventarizace informačních zdrojů.** Jde o systematicky uspořádané přehledy, např. regionální geologické, geografické, archeologické, aj. tematiky, katalogy nebo databáze se specializovaným popisem jednotek uchování dat, jejich částí, prvků a skupin prvků. Zavedení metadat je předpokladem cíleného a úspěšného nalezení zdrojů dat, jejich hodnocení z pohledu potřeb uživatelů a zpracovatele a také problémové orientace GIS jako celku v procesu projektování nového GIS. Velké a dlouhodobé projekty GIS by měly zahrnovat postupy sběru metadat a považovat je za jednu z povinných etap jejich projektování. Víceúčelový charakter využití GIS při řešení rozmanitých, ne vždy prostorových úloh, předpokládá dostatečně univerzální a podle možností co nejúplnější přehled charakteristik popisovaných dat.
2. **Metadata jako součást resp. prvek geoinformačního systému.** Inventarizace a ocenění informačních zdrojů o území je obvykle jednou z prvních etap projektování GIS, která předurčuje (společně s analýzou typů úloh, které budou řešeny potenciálními uživateli) obsah databáze GIS. Přitom je možný formalizovaný popis uchovávané informace o zdrojových materiálech na počítačových nosičích a manipulace s nimi v rámci standardních systémů řízení DB (tj. organizace v podobě databáze metadat). Konečně, metadata mohou být zahrnuta v podobě samostatného bloku do využívaného GIS, kde báze metadat společně s „objektovou“ databází bude zajišťovat specifické úkoly spojené s administrativou databáze GIS tím, že shromažďuje a zabezpečuje přístup k datům o zdrojích, o vznikajících a výsledných datech, jež jsou systémem vytvářeny v procesu jejich využití.
3. **Metadata zapojená do procesu výměny a předávání dat mezi jejich držiteli (správci) a uživateli v lokálních nebo globálních sítích.** Obě strany přirozeně potřebují podrobného „meta-průvodce“.

Vytvoření takového průvodce vyžaduje podrobnou specifikaci druhu dat, jejich původu, kvality, aktuálnosti, plnosti a dostupnosti. Sémanticky se utváří vyčerpávající popis vlastních dat o objektech. To zabezpečí těsné *spojení dat a metadat*, které vylučuje potřebu nějakého doplňujícího komentáře o předávaných datech. Je to důležité i pro proces výměny nebo předávání dat, který je obvykle doprovázen konverzí do jiných formátů. Jednoznačnost a úplnost konverze musí vycházet z odpovídajících částí metadatové databáze.

V mechanismu výměny dat je nutné vyčlenit část, která se bude týkat problematiky bází metadat. Přitom vytvořené standardy výměny dat musí obsahovat jejich vyčerpávající popis.

Báze metadat mají velký význam při projektování distribuovaných prostorových datových bází, systémů, které pracují v režimu kolektivního využívání a také systémů, kde je nutná efektivní a rychlá výměna prostorově-souřadnicových dat v národních a regionálních měřítkách (informační dálnice).

Jako příklad si uvedeme popis národních informačních zdrojů Australského byra minerálních zdrojů, geologie a geofyziky, který zahrnuje 253 databází (z toho 57 bibliografických a 196 faktogra-

fických), jež sestavilo 37 australských organizací podle následující tabulky (upraveno podle Koskarev A.V., 1995):

Jméno databáze

Akronym

Typ/statut databáze: bibliografická/faktografická, „uzavřená“ (dále neobnovovaná)/aktivní (s pravidelnou aktualizací dat)

Tematika

Vztah dat k území

Časové období, z něhož data pocházejí

Klíčová slova-deskriptory, které se vztahují k tematice a místu

Databázový systém-typ počítače, využívaného systému řízení databáze, magnetický nosič

Objem dat v databázi (počet uchovávaných jednotek)

Výstupy: typ výstupní dokumentace (text, tabulky grafika, mapy, aj.)

Dokumentace: odvolání na návod uživateli aj. instruktivní materiály, popis z literárních zdrojů

Komentář a doplňková informace

Kontaktní osoba : adresa pro kontakt

Datum sestavení popisu

9.1.6 Datový model

Existují dva hlavní modely:

1. vektorový
2. rastrový

Oba je u většiny systémů možné spojit v model hybridní.

Podrobnější informaci o modelech lze nalézt v dalším příspěvku těchto skript (viz Kučera M., Macháček, J.: Teorie a praxe zpracování archeologických výzkumů s pomocí prostředků GIS/LIS).

Uvedme zde jen pro srovnání stručné charakteristiky analytických schopností rastrových a vektorových GIS.

Typické analytické schopnosti rastrových a vektorových GIS:

Rastrové GIS:

lokální operace

překódování

skládání vrstev

operace na lokálních sousedech

operace na vzdálenějších sousedech

operace v zónách

Vektorové GIS:

reklasifikace, rušení hran, spojování

topologická analýza

vytváření zón

měření

operace nad povrchy (2.5 D)

síťová analýza

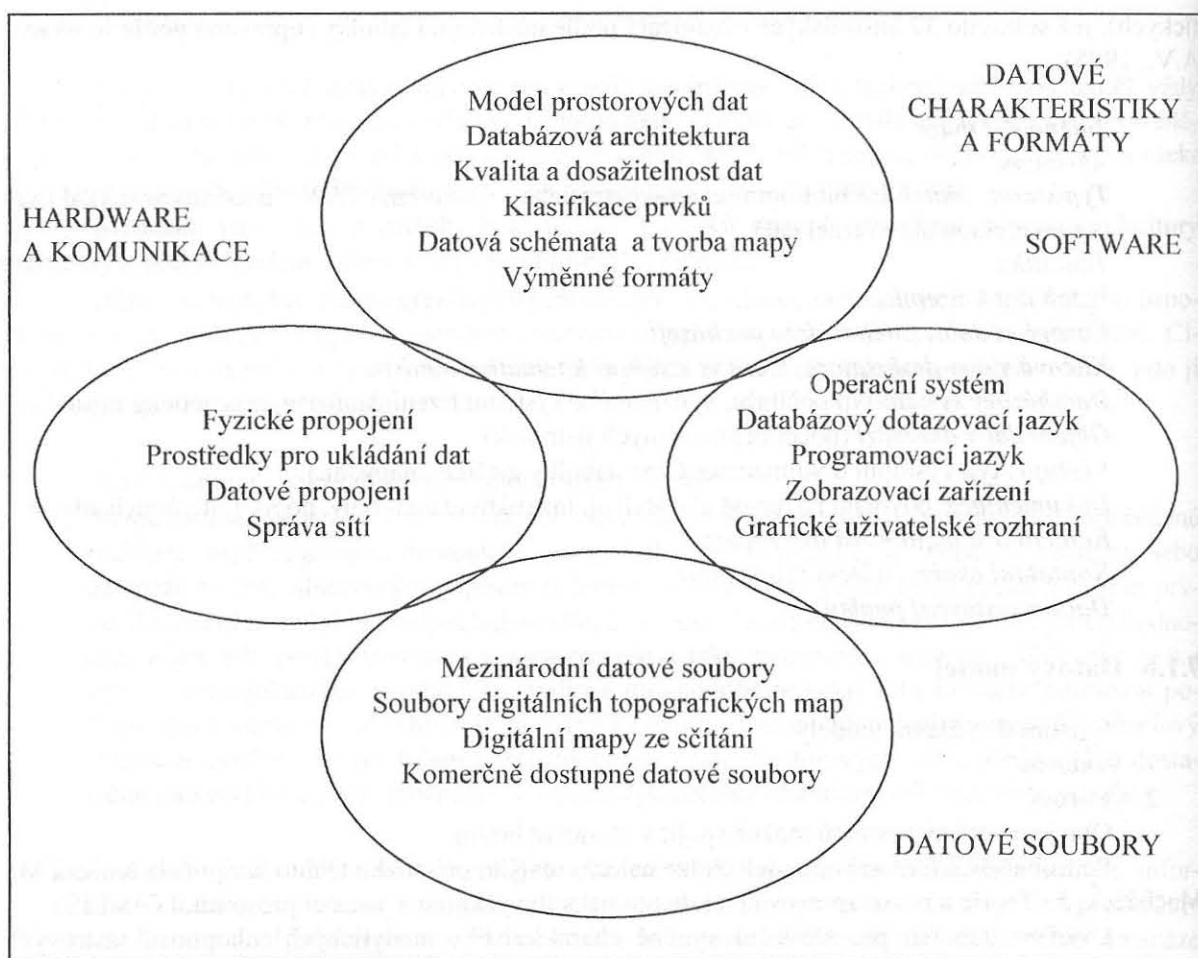
9.1.7 Ukládání dat

Uskutečňuje se na pevných discích, magnetických páskách, CD, magnetooptických discích a dalších nosičích. K uvedené problematice doporučuji prostudovat příslušnou specializovanou literaturu.

9.2 Zdroje dat v ČR

9.2.1 ZABAGED

V České republice je postupně dokončována Základní báze geografických dat (ZABAGED 1 a 2). Jsou vytvářeny a harmonizovány standardy (obr. 5) pro datovou výměnu, jež mohou být využity i v mezinárodním měřítku.



Obr. 5. Kategorie GIS standardů (Zdroj: McLaughin, 1991).

Vznik ZABAGED byl vyvolán nízkou úrovní organizace formování GIS u nás, které se dosud projevuje mimo jiné v nedostatku standardů pro tvorbu geografických dat, v nejasném vymezení práv a povinností zúčastněných subjektů, podmínek pro všestranné využití jednou vytvořených dat a dalších pravidel a úmluv týkajících se této oblasti. ZABAGED je součástí státního informačního systému, kde snad sehraje integrační úlohu vytvořením společné geometrické osnovy všech prostorově orientovaných informačních systémů s odpovídající úrovní podrobnosti. Garantem vzniku a postupné realizace projektu a vytvoření konečného produktu je Český úřad zeměměřičský a katastrální.

Data do projektu ZABAGED byla získána rastrovou digitalizací tiskových podkladů základní mapy ČR v měřítku 1:10 000 a konverzí do:

- rastrového modelu ZABAGED 2, jenž je tvořen počítačově využitelným mapovým pozadím a byl realizován pro celé státní území v roce 1994. Je považován za dočasnou náhradu ZABAGED 1 a po dobu jejího budování jako její doplněk.
- topologicko-vektorového modelu ZABAGED 1, s dočasným rastrovým zobrazením polohopisu intravilánu a trvalou reprezentací reliéfu ve tvaru síťového digitálního modelu, který bude ukončen až v roce 2003. První výstupy jsou k dispozici od roku 1995.

ZABAGED bude harmonizován se vznikajícími evropskými normami pro geografickou informaci. Tím budou vytvářeny předpoklady pro bezkonfliktní uplatnění ve vnitrostátní a mezinárodní výměně geografických dat a zabezpečena její efektivní využitelnost k informačnímu servisu pro:

- tuzemské GIS určené k plnění cílených poznávacích a rozhodovacích funkcí ve správě, službách, apod.,

- b) tvorbu státního mapového díla tematických map středních měřítek při přechodu na jejich automatizovanou realizaci,
- c) zahraniční GIS k realizaci různých mezinárodních programů na podporu evropské poznávací a rozhodovací součinnosti (Koncepce 1994).

Souhrnně – k začátku roku 1997 jsou na Českém úřadu zeměměřičském a katastrálním (ČÚZK) resp. Zeměměřičském úřadu (ZÚ) k dispozici:

- rastrová podoba ZABAGED 2 1: 10 000
- rozpracován ZABAGED 1, 1:10 000 (28% území ČR)
- rastrová bežešvá podoba topografických map 1:50 000, ČR, barevná i černobílá verze
- rastrová i vektorová bežešvá mapa 1:500 000 pro celou ČR, barevná i černobílá verze

9.2.2 Archiv grafických dat KN

Týká se skenování katastrálních map. Skenovány jsou především mapy bývalých pozemkových evidencí, jejichž vedení bylo na střediscích geodézie (dnešních katastrálních úřadech) ukončeno v 50. letech, v menším rozsahu i vedené (evidenční) katastrální mapy. Rastrový obsah těchto map je afinní transformací upraven na původní rozměr mapového listu a ve formátu CIT uložen na archivační media. V této podobě nebo po případné transformaci do systému JTSK je distribuován (po dohodě s místním katastrálním úřadem) uživatelům. Pro přehlednější evidenci digitalizovaného mapového fondu byl v ZÚ vyvinut grafický prostředek, systém ARCHIV.

9.2.3 Databáze správních hranic územních celků.

Vznikla digitalizací zákresu správních hranic v ZM 50. Obsahuje celorepublikový soubor okresních hranic a dále (strukturovaně po okresech) soubory vrstev obecních hranic a hranic katastrálních území, včetně jejich definičních bodů, které jsou situovány v centru osídlení. prostřednictvím institucí, které spravují Územně identifikační registr byly doplněny vrstvy kódů (čísel Českého statistického úřadu) a názvů správních jednotek. Databáze správních hranic je vedena v grafickém prostředí MicroStation. Data jsou distribuována ve formátech .DGN nebo databázích a přehledů, založených na územně správním principu. Slouží rovněž jako podklad pro vedení databáze bežešvých administrativních hranic v Evropě (CERCO, Megrin Group).

9.2.4 Vojenský topografický informační systém.

Vojenský topografický informační systém (dále VTIS) představuje soubor technicko – technologických (výpočetních, automatizačních a komunikačních) prostředků, programového vybavení, technologií a metodik zpracování informací o území, zvláště pak topografických dat a dalších grafických nebo alfanumerických souvisejících podkladů.

VTIS je součástí Státního informačního systému (SIS).

Hlavními produkty VTIS jsou:

- databáze topografických informací o území v digitálních formách (vektorové, rastrové a maticové),
- soubory vojenských topografických map v analogové formě měřítek 1:25 000, 1:50 000, 1:100 000 a 1:200 000,
- odborné publikace, normy a standardy vydávané k zabezpečení standardizace při aplikaci dat a mapových podkladů v AČR.

Východiskem pro výstavbu VTIS jsou vojenské topografické mapy měřítkové řady 1:25 000 až 1:200 000 a technologický systém k jejich trvalé aktualizaci a obnově.

K dispozici, a to i pro potřeby civilního sektoru (např. regionálních GIS) jsou:

- digitální model území v měřítku 1:200 000, označovaný jako DMU 200. Obsahuje informace o vodstvu, sídlech, komunikacích, rostlinném a půdním krytu, energetických vedeních, produktovodech a hranicích. V současné době je prováděna aktualizace se zaměřením na doplnění základních informací o komunikacích, sídlech a hranicích.
- digitální model reliéfu 2. generace, tzv. DMR -2. Jedná se o soubor nadmořských výšek ve čtvercové síti 100m x100 m.

V současné době jsou obě databáze převáděny do ekvivalentních datových formátů západoevropských standardů:

- digitální model reliéfu 1. generace (DMR-1) , což je rozsáhlá DB výškopisných údajů, zobrazených ve čtvercové síti 1 km x 1 km, přičemž výškopisný údaj uvádí maximální výšku území v příslušném kilometrovém čtverci,
- digitální ekvivalenty topografických map měřítek 1:50 000, 1:100 000, 1:200 000 a 1:1 mil. jako databázi souborů zobrazujících obraz vojenských topografických map uvedených měřítek v rastrové formě. Z dat je možno v GIS vytvořit bezešvou elektronickou mapu. Tato data budou rozšířena o digitální ekvivalenty topografických map měřítek 1:25 000 a 1:50 000,
- digitální model území , DMU je jedním z nejvýznamnějších produktů. Jde o vektorový topografický model, který přesností, rozlišovací úrovní a mírou generalizace odpovídá rozlišovací úrovni vojenských topografických map měřítko 1:25 000. Je konstruován objektové podle sémantických typů objektů definovaných katalogem topografických objektů, který vychází z obsahu mezinárodního katalogu DIGEST 1-2 a značkového klíče Topo-4-3. Jsou registrovány údaje o vodstvu, sídlech, komunikacích, rostlinném a půdním krytu, energetických vedeních, produktovodech, hranicích a objektech mikroreliéfu. Hlavním zdrojem dat jsou vojenské topografické mapy měřítko 1:25 000 a dříve vytvořené účelové databáze vybraných topografických informací v digitálních formách. Údaje jsou z hlediska přesnosti registrovány v závislosti na významu objektu ve čtyřech třídách přesnosti v rozmezí od 0,10 m do 25 m polohové odchylky.

Pozn.: Termínem DIGEST (Digital Geographic Information Exchange Standard) rozumíme digitální geografický informační výměnný standard, který byl původně vyvinut zeměmi NATO. V posledních letech je stále více uplatňován i v civilní oblasti a začíná sehrávat integrační roli při propojování vojenských a civilních digitálních kartografických děl.

9.3 Nové trendy rozvoje GIS

9.3.1 GIS v sítích

Zapojení GIS do sítí souvisí s obecnou problematikou zapojení všech informačních systémů, a to i prostorových, do sítí. Cílem je učinit informace, jež je v různých GIS, LIS či CAD k dispozici, dostupné široké veřejnosti, přičemž z tohoto propojení by měli užitek jak uživatelé, tak i správci jednotlivých systémů.

Prostorové informační sítě, které vznikaly v průběhu 90. let, se především zaměřovaly na převod textových dat z hostitelských počítačů na terminály uživatelů. Typickým příkladem je oblast GIS či LIS. V současnosti sítě stále více převádějí velké objemy vektorových a obrazových dat.

Přenos dat je jedním z výrazných znaků globální informatizace. Dříve jsme byli zvyklí, že při jednoduchých přenosech, telefonních či televizních, se hlas přenášel beze změn. Podobně tištěné slovo se přenášelo pomocí knih, novin a časopisů. Moderní komunikační sítě vnášejí přidanou hodnotu tím, že přenosové schopnosti spojují s ostatními informačními funkcemi. Tím se objevují základní rysy komunikační sítě s mnoha propojenými počítači, od velkých sálových přes připojená stolní pra-

coviště až po osobní počítače a malé terminály. Právě toto telekomunikační propojení zprostředkuje funkci přenosu. Řešena může být nejrůznějšími způsoby a probíhat může uvnitř budovy, do sousední ulice či města nebo na jiný kontinent.

Rozeznáváme místní, lokální síť (**LAN**, local area networks) nebo dálkové síť (**WAN**, wide area networks)

Tam, kde síť nejen přenáší různé formy informací, ale vykonávají i funkce tvorby, zpracování a uchovávání informací, obohacují probíhající vědeckou nebo hospodářskou činnost. Z toho důvodu se jim říká síť s přidanou hodnotou (**VAN**, value-added networks).

Současný rozvoj využití Internetu resp. jeho lokální varianty, Intranetu, a vůbec myšlenka informační dálnice (viz. dále), tento vývoj potvrzuje. Vícenásobní uživatelé, multimediální databáze jsou a budou postupně interaktivně dostupné prostřednictvím regionálních, národních a globálních sítí. Velké textové a grafické soubory budou elektronicky rozesílány a přijímány. Videotelefony a videokonference budou stále prohlubovat síťové propojení prostorových informačních systémů. Organizace budou mít technologickou svobodu jako nikdy předtím.

Splnění požadavku lepšího propojení vyplývá z rostoucího množství rozmanitých aplikací, ale také vyžaduje integraci dat a sdílení dat. Zkušenosti z vyspělých států (Austrálie, Severní Amerika, Velká Británie) dokazují, že jakmile se zpřístupní soubory základních dat, vzápětí vzrostou mnohonásobně aplikace prováděné nad těmito daty. Hnací síla tohoto vývoje čerpá z několika zdrojů:

1. kapitalizace vyplývající z možnosti odvozování produktů a služeb s přidanou hodnotou (value-added). Platí to zejména v interaktivním prostředí, v němž může být informace dosažena po každé, když je to potřeba. Zvýší se úloha tzv. příležitostného uživatele.
2. jakmile objem dat roste, zaměří se pozornost na náklady za zdvojená data, vzroste zájem o koordinaci a integraci, a to z ekonomických a výkonnostních důvodů. Např. práce s jednou databází zajišťuje, že každý operuje se stejnou informací, přesností, zdvojené vyhledávání resp. kreslení se redukuje. Vznikne možnost provádět široké aplikace vyžadující data z různých, fyzicky často vzdálených a obsahově tematicky rozdílných zdrojů.

Realizace těchto záměrů vede přes integraci existujících heterogenních datových systémů a vývoj aplikačních software pro co největší využití informační báze, jež je často geograficky rozptýlená. Pokud je to dosaženo, vzniká pro data rozmístěná původně v různých informačních systémech, přidaná hodnota (added utility).

Dosažení lepšího propojení závisí jak na technických předpokladech, jimiž jsou vysoce rychlostní digitální telekomunikace, tak i problematice spojené s distribuovanými informačními systémy. Přitom jde o integraci dat z různých zdrojů, bez dříve nutné centralizace, více logickou, než fyzickou integraci. Dosažení této úrovně integrace vyžaduje rozvinout a kontrolovat následující aspekty.

1. **Bezpečnost.** Jedním z nejproblematictějších bloků elektronicky propojených databází je bezpečnost hostitelských dat, tj. zajištění přístupu pouze pro legitimní uživatele.
2. **Vlastnictví.** Obecně nedořešený problém, mezinárodně diskutovaný.
3. **Kvalita dat.** Existuje několik položek vztahených ke standardům:
 - a) **všeobecnost datového modelu**, tj. konceptu použitého pro popis dat v databázi. Koncept resp. významy spojené s datovými vrstvami by měly být konzistentní mezi různými uživatelskými skupinami. Obvyklejší je doposud opačná situace, že data a informace zaznamenaná a uložená na různých místech mají nekompatibilní datové struktury;
 - b) **přesnost dat a informací uložených v rozmanitých databázích**, a to v čase, prostoru a tematické. Týká se to měřítek map, kartografických projekcí a transformací, prostorové preciznosti a hustoty, definice rysů (charakteristik), klasifikačních systémů, atd. Přitom data mohou být na různých místech a v různých databázích, např. souřadnice vybraných bodů v jednom systému, adresy v jiném, telefonní čísla účastníků v dalším, atd.);
 - c) **kompletnost dat** se může měnit podle míst uložení i jednotlivých uživatelů. Některé obory nemusí být závazné, jiné nemusí být ověřené. Některé databáze byly vytvořeny pro účelové využití (např. účetnictví, popis naleziště), které nepodporuje širší funkce (marketing, strategické plánování, plánování výzkumu);
 - d) **platnost**; tj. vyjádření aktuálnosti dat a pravidelnost jejich aktualizace.

4. **Podpora sítě a expertizy.** Zpravidla není dostupná ve všech místech a pro všechny specialisty napojené na síť. Správci nebo klíčová centra budou poskytovat servis týkající se diagnózy chyb v celé síti, síť udržovat a obnovovat a také využívat experty v různých místech sítě.

Vývoj směřuje k distribuovaným informačním systémům, přes výše zmíněné problémy, z následujících důvodů:

- výhody datové integrace jsou již diskutovány,
- kvůli kladným (a rovněž záporným) prvkům spojeným s vlastnictvím lokálních dat a jejich kontrolou,
- operační výkon může být v distribuovaných systémech zvyšován podle rostoucích požadavků, na rozdíl od vysokých nákladů na střediskové počítače,
- distribuované systémy jsou více odolné, kupř. při haváriích telekomunikačního spojení,
- nové technologie mohou být postupně propojovány s multimediálními daty v prostorových informačních systémech.

9.3.2 Státní informační systém ČR

Vědoma si výše uvedeného stavu, který se zákonitě začal projevovat i v nedostatečném využití výpočetní techniky a již existujících dat pro výrazně lepší činnost státní správy, zejména z hlediska jednotlivého občana, přijala česká vláda 20. září 1995 „**Usnesení Vlády ČR: Záměr a postup výstavby Státního informačního systému České republiky**“. Konstatuje se v něm, že:

„Státní informační systém (SIS) umožňuje zajistit efektivní shromažďování, předávání a využívání informací ve shodě se zákony, a to nejen státní správou, samosprávou, mezinárodními institucemi, ale i občanem“.

Návrh vychází z názoru že pokud systém nakládání s informacemi v rámci státní správy považujeme za prostředek k racionalitě sběru, využívání a ukládání informací, pak je třeba prioritně respektovat následující zásady:

1. nevyžadovat od poskytovatele informací takovou informaci, která není potřeba,
2. minimalizovat styk poskytovatele informací se státní správou,
3. zajistit přítomnost informace vždy a pouze tam, kde je využívána,
4. zajistit bezpečnost informace před zneužitím,
5. využít veškerých dostupných technických prostředků pro maximálně efektivní využití informace a minimalizaci ekonomických nákladů.

Základní principy SIS vycházejí z toho, že existují informace, které se týkají občana, ekonomického subjektu nebo území a které neslouží jen omezenému počtu resortů.

Existují úkony styku občana a státní správy, při kterých není nezbytná přítomnost občana, a které lze řešit zprostředkovaně. Je možné zlepšit organizaci práce při zajišťování komunikace mezi státní správou, samosprávou a občanem.

Dosažením přímého (on-line) přístupu k informacím lze zabránit vytváření duplicit a zrychlit správní úkony.

V rámci SIS bude vytvořena distribuovaná relační databáze, která umožní získávat a využívat data o občanech, právnické osobě, území, případně dalších organizacích a zároveň statistické údaje, potřebné pro činnost státní správy a samosprávy, zajistit odpovídající datovou komunikaci mezi centry státní správy.

Databáze bude zpočátku obsahovat dvě základní tabulky, z nichž se odvozují všechny ostatní údaje potřebné pro styk občana a státní správy. Jde o tabulku informací o občanech a o tabulku informací o ekonomických subjektech.

Působnost SIS bude celostátní a počítá se jak s jeho působností ve státních orgánech, tak i v přenesených působnostech na jiné orgány, např. obce s přenesenou působností státní správy. Ve vztahu k vyžadování a poskytování informací státem má jako partnera občana, právnickou nebo fyzickou

osobu. Tyto kroky si vyžádají analýzu a tvorbu nové legislativy ve sféře poskytování, ochrany a nakládání s informacemi a povede k návrhu systémových změn v oblasti kompetencí, způsobu ochrany a nakládání s daty. Legislativa musí vycházet z premisy jednoznačnosti a kontrolovatelnosti systémů.

Jak ukazují zkušenosti vyspělých evropských i zámořských států (USA, Austrálie, Nového Zélandu, aj.), dříve nebo později bude potřeba vyřešit vztah i k soukromému sektoru. Národní informační infrastruktury těchto států jsou budovány mj. pro podporu soukromého podnikání a dalších aktivit vedoucích mimo jiné ke vzniku nových firem a tvorbě pracovních míst. Zde bude třeba hledat vztah mezi oběma dosud v informační oblasti dost vzdálenými sektory.

Dalším problémem je stupeň utajení informací. Ochrana osobních a dalších pro jednotlivce i stát významných dat je nutná, ale v současnosti jsou některá data a informace chráněna příliš přísně. Mám na mysli např. přístup k datům v oblasti zdravotního stavu obyvatelstva a určování zdravotních rizik. Nesmí jít o uveřejňování jmen jednotlivců s určitou chorobou, ale o údaje o výskytu jednotlivých chorob na co rozměrově nejmenší prostorovou jednotku, což následně umožní analýzu možných příčin výskytu určitých chorob, které může být závislé na stavu životního prostředí, ale i jiných faktorech.

S tím souvisí také stanovení pravidel pro obchodování s daty a informacemi a s tím spojené problémy s ochranou dat a autorskými právy na ně. Dosavadní přístup např. k mapovým dílům pouze z hlediska materiálního produktu zcela opomíjí skutečnost, že v mapě je potřeba chránit i jednotlivé informační vrstvy do ní vložené, způsob jejich spojení, apod.

Při využití SIS by měl požadavek na informaci vycházet z potřeby rozhodnutí, které není bez této informace možné přijmout a provést. To znamená přehodnotit velký objem informací a navrhnout jejich redukci nebo obsahové změny, které budou velice složité a časově náročné.

U SIS jde zejména o záměr uspořádat informace ve státní sféře. Dát jim efektivní strukturu pro státní správu z pohledu využití a ekonomiky zpracování. V současné době jsou informace uspořádány tak, že každý, kdo vykonává nějaké rozhodnutí, má příslušné informace ve svém databázovém systému a také je sám vyžaduje od poskytovatele. Pokud dojde k přehodnocení objemu a obsahu informací podle výše uvedených kritérií, dostaneme se do stavu minimalizace nároků na poskytovatele, při racionálním způsobu využití této informace pro výkon státní působnosti. V současnosti informace vyžaduje každý orgán samostatně. Prvním krokem bude redukce požadavku informací z hlediska identifikace poskytovatele.

Z toho důvodu je SIS postaven na identifikaci žadatele o rozhodnutí na základě minimálních informací, u občana např. jméno, příjmení a rodné číslo. Přitom je potřeba vyloučit duplicity a chyby. Ukládat informaci na jednom místě (okresní úřady) a odsud ji v případě potřeby získávat. Na ně pak budou navazovat veškeré informace, které má státní správa k dispozici.

Geografická informace bude v budoucnu zcela jistě významnou součástí i státních informačních systémů, jejichž název se pravděpodobně přizpůsobí jejich novému širšímu obsahu a asi bude znít „**národní prostorová resp. geografická informační infrastruktura**“, která bude **součástí národní informační infrastruktury**.

GIS a digitální kartografie musí napomáhat ke vzniku nových koncepčních systémů pro řízení společnosti a prostředí. Všechny druhy systémů, včetně negeografických, budou navzájem propojeny novými informačními technologiemi. Úloha GIS a digitální kartografie bude definována určitým systémem zpracování informací. Geografický pohled přitom bude mít kritickou důležitost. Velké a bez nadsázky téměř neuvěřitelné perspektivy se ukazují v souvislosti s budováním informačních dálnic (information superhighway).

9.3.3 Co je „information superhighway?“ Úloha GIS.

Informační dálnice je komplikovaný a technicky náročný projekt, který zabezpečí nejen rychlý přenos velkého množství dat, ale zejména umožní uživatelům (např. GIS) přístup do řady databází se standardizovanými, a tedy obecně přístupnými daty.

Za informační superdálnici je považován transparentní soubor odpovídajících si, vzájemně propojených a vysoce rychlých sítí, doručujících aktuální interaktivní data, texty, zvuk, hlas, video, film všem jejich tvůrcům a zákazníkům.

Informační dálnice bude zahrnovat existující a budoucí sítě a technologie: fixované sítě, mobilní sítě, „chytré karty“, GPS, které umožní řadu operací na informační dálnici.

Kombinace informační dálnice a GIS technologií umožní vznik mnoha nových aplikací, a to zejména vzájemné propojení a pružné sdílení velkoobjemových dat lokalizovaných na různých místech, a to v krátkém čase. Přitom budou k dispozici rozmanitá a aktuální data z různých zdrojů (GIS, DPZ, digitální fotogrammetrie, GPS, statistiky, aj.), přičemž data budou dostupná pro všechny uživatele.

9.3.4 GIS a informační dálnice

V současnosti existují dvě představy pro vytvoření GIS aplikací s ohledem na informační dálnici:

- 1) Extrapolace současné situace vyvozuje nové způsoby realizace toho, co děláme nyní. Technologické trendy, kombinace technologických a sociálních trendů jsou využívány k předpovědím vývoje s ohledem na aplikace GIS. Jsou přitom očekávány:
 - dynamičtější procesy,
 - méně složité procesy,
 - nižší ceny.
- 2) Představa budoucnosti vyvozuje nalezení způsobů, jak dělat nové věci novými způsoby.

Schopnosti informační dálnice jsou aplikovány v kombinaci s vlastnostmi GIS, včetně využití multimédií, vizualizace, aj. prostředků

Kombinace informační dálnice a GIS technologií umožní vznik mnoha nových aplikací:

- Vzájemné propojení: lokalizace dat se stane méně důležitá, než vzájemné propojení mezi systémy umožňující pružné sdílení dat mezi různými místy,
- Vysoká rychlost: technické prostředky umožní výměnu obrovských objemů GIS dat v relativně krátkém čase,
- Aktuální data z mnoha zdrojů: pro použití v GIS budou k dispozici rozmanité zdroje dat,
- Dostupnost pro všechny tvůrce a zákazníky: vlastnictví systémů a dat nepodmiňují budování a využití GIS.

Komunikační prostředky umožní tvorbu a využití datových center GIS:

- elektronická datová centra nabídnou přístup k souborům jak komerčních, tak nekomerčních dat pro GIS pomocí informační dálnice,
- přístup do elektronických datových center GIS zlepší funkčnost a pružnost GIS v jednotlivých podnicích a organizacích,
- v jednotlivých případech: pro odpověď na otázky pomocí vlastního GIS potřebujeme doplňkovou informaci,
- GIS upgrade: přidání nových dat do DB a rozšíření jeho funkcí,
- GIS aktualizace: datové soubory odvozované z externích zdrojů by měly být periodicky aktualizovány.

Předpokládá se, že elektronická datová centra s komerčními i nekomerčními GIS daty budou přístupná po informační dálnici.

Informační dálnice povede k decentralizaci úkolů spojených s operacemi GIS:

- kromě zachování vlastnictví datových souborů, musí tzv. „clearing houses“ pečovat o sběr dat, konverzi formátů a aktualizaci,
- GIS servis může nabízet zpracovatelské kapacity pro operace s GIS v jednotlivých organizacích nebo dálkový přístup pro uživatele po informační dálnici,
- některé organizace namísto vlastních GIS specialistů a systémových operátorů dají přednost „vnějším zdrojům“ a zaměří se na efektivní využití systému při vlastním podnikání (Lauterbach, 1995).

9.3.5 Evropské aktivity: GI2000

Nejvýznamnější připravovaná evropská aktivita je charakterizována v materiálu z 30. prosince 1995 s názvem „K evropské geografické informační infrastruktuře (EGII)“ *The EGII Discussion Document*. Aktivita je označována akronymem **GI2000**.

Geografická informace je definována jako složitá, rychle rostoucí a důležitá část informační společnosti s mnoha aplikacemi v nadnárodních orgánech a na úrovni vlád států, podnikání a výzkumu. Geografická informace je důležitá kvůli svému obsahu a významu elektronické a prostorové informace vzrůstá pro plánování, správu území, marketingové studie, životní prostředí, obnovu energetických zdrojů, záchranné služby a stavy nouze, péči o zdraví, politické analýzy a mnoho dalších aktivit. Ačkoliv nové technologie a aplikace pro geografické informace jsou vyvíjeny rychle, příští vývoj v Evropě je omezován řadou rozdílů ve způsobech, jak je geografická informace sbírána, ukládána a distribuována v různých zemích a různých oblastech státní správy a obchodu. Poněvadž nedigitální geografická informace byla specializovanou aktivitou organizovanou jednotlivými státy a profesemi rozmanitými způsoby, neexistuje ani „evropská politika“ týkající se digitální geografické informace, ani snadno dostupné základní datové soubory a podpůrné technologie a znalostní infrastruktury.

Současný rozvoj průmyslu s geografickou informací je odhadován na 18-30% ročně. Je způsoben mezinárodními a celoevropskými aplikacemi a pokroky počítačové technologie. Růst zmíněných aplikací a příhraničních projektů je omezován absencí dat pro celou Evropu, která by měla příslušnou kvalitu a přiměřené ceny, rozdílnými mezinárodními přístupy a standardy a politickými a legislativními opatřeními. Pokud neexistují standardy, projevuje se to na až dvojnásobné prodejní ceně dat, což velmi nepříznivě působí na jejich uživatele.

Klíčovým cílem projektu je poskytnout širokou, snadno dostupnou platformu s vysoce kvalitními základními daty, v rámci jednotné evropské infrastruktury.

9.3.6 Česká národní prostorová informační infrastruktura.

Česká národní prostorová informační infrastruktura by měla být tvořena stabilním celostátně rozšířeným a dohodnutým souborem pravidel, standardů a postupů pro tvorbu, sběr, výměnu a využití geografické informace jako součásti SIS. Databázové soubory by měly být snadno dostupné, měl by existovat metadatový servis, který by napomáhal data snadno lokalizovat potenciálními uživateli.

Takto pojatá koncepce počítá s napojením na budoucí Evropskou geografickou informační infrastrukturu (EGII).

9.4 Literární a časopisecké zdroje o GIS a digitální kartografii.

V česky psané literatuře neexistuje aktuální publikace, která by vyčerpávajícím způsobem popisovala oblast GIS. Poslední ucelené dílo, které bylo ve své době prvním v tehdejší střední a východní Evropě, bylo vydáno v r. 1985 M. Konečným a Karlem Raisem (viz. Seznam literatury) a bylo provázeno stejnojmenným učebním textem. Práce vznikala v období využívání tzv. velkých počítačů, nezahrnuje tedy vývoj zaspoočatý příchodem PC a další nejmodernější výpočetní techniky. V zahraničí existuje celá řada velmi dobrých učebnic a publikací, jejichž úplný seznam zde nelze uvést (navíc jsou většinou finančně nákladné). Citujeme jen vybrané tituly (viz též část Literatura):

- Antenucci J.C. et al.: *Geographic Information Systems. A Guide to the Technology*. Chapman & Hall, New York, London. 1991.
- Aronoff S.: *Geographic Information Systems: A Management Perspective*. WDL Publications. 1989.
- Burrough P.A.: *Principles of Geographical Information Systems*. Oxford Science Publications. 1986.
- Digitální data v informačních systémech. Ed.: V. Voženílek. Antrim. 1996
- Elements of Spatial Data Quality. Ed. By S.C. Guptill & J. L. Morrison. ICA. Pergamon. 1995.
- Handling Geographical Information. Methodology and Potential Applications. Edited by I. Masser & M. Blakemore. Longman Scientific & Technical. J. Wiley & Sons. N.Y. 1991.

- Innovations in GIS 1. Ed. By M.F. Worboys. Taylor & Francis. 1994.
- Introductory Readings in Geographic Information Systems. Edited by D.J. Peuquet and Duane F. Marble. Taylor & Francis. 1990.
- Koškarev A. V., Tikunov V.S.: Geoinformatika. Kartgeocentr-Geoizdat. Moskva. 1993
- Krcho J.: Morfometrická analýza a digitálne modely georeliéfu. Veda. Bratislava 1990.
- Langram G.: Time in Geographic Information Systems. Taylor & Francis. 1993.
- Maguire D.J., Goodchild M.J., Rhind D.W.: Geographical Information Systems. Vol. 1 Principles, Vol.2 Applications. Longmann Scientific & Technical. 1991.
- Mitášova I., Pezlar Z., Veverka B.: Základy teorie systemov a kybernetiky s aplikaciami v geodezii a kartografii. Alfa. Bratislava. 1990.
- Networking Spatial Information Systems. Edited by Newton P.W., Zwart P.R. and Cavill M.E. John Wiley & Sons. 1995 .
- Robinson A.H., Morrison J.L. et al.: Elements of Cartography. Sixth Edition. John Wiley & Sons. 1995.

Dalším cenným zdrojom informácií jsou odborné geografické a kartografické časopisy, více zahraniční, než domácí. Jmenujme nejprestižnější a nejoblíbenější z nich:

- International Journal of Geographical Information Systems (IJGIS), editovaný P. Fisherem, K.C. Clarke a D.J. Abelem, vychází osmkrát ročně a je vydáván vydavatelstvem Taylor&Francis Ltd v Londýně. Vychází od r. 1986.
- Geoinformatika, editovaná P. Bergougnouxem a vydávaná Kluwer Academic Publishers v Norwell, MA, USA, začal vycházet v r. 1996.
- Oba uvedené časopisy jsou zaměřeny na teoretické a vědecké otázky rozvoje GIS a disciplín s nimi souvisejících.
- GIS Europe, Europe's Geographic Information Systems Magazine, vydávaný Geoinformation International v Cambridge, Velká Británie. Vychází od r. 1991.
- GIS World, vycházející od r. 1987 ve Fort Collins, USA. Časopis je předchůdcem GIS Europe.
- Business Geographics. The Geographic Technology Magazine for the Business Community. Vychází od r. 1992 ve Fort Collins, USA.
- Všechny tři časopisy mají dobrou odbornou úroveň a navíc barevnou, atraktivní úpravu. Redaktoři GIS Europe často komentují konference a vývoj v oblasti GIS ve středoevropských zemích.

Další skupinu tvoří časopisy tematicky odborně zaměřené. Příkladem mohou být časopisy: GIM-International Journal for Geomatics, vycházející od r. 1986 v Lemmer, Nizozemí, Photogrammetric Engineering & Remote Sensing, vydávaný od r. 1979 v Bethesda, MA, USA

Dále vycházejí firemní časopisy předních firem působících v oblasti GIS, např.:

- Global Link, Intergraph's International Mapping/GIS and Civil Newsletter, vycházející od r. 1994,
- ARC News, Environmental Systems Research Institute, Inc., vycházející od r. 1978 v Redlands, CA, USA

Řadu informací lze najít i na Internetu (např. konference GIS-L) nebo na WWW stránkách komisi Mezinárodní kartografické asociace (ICA) resp. dalších organizací i soukromníků.

9.5 Literatura

- Albrecht J.: Universal GIS-Operations. A Task-Oriented Systematization of Data Structure-Independent GIS Functionality Leading Towards a Geographic Modeling Language. ISPA, Vechta, 1996
- Davis S., Davidson B.: Vize roku 2020. NIS ČR. Praha 1995.
- Europe and the Global Information Society: Recommendations to the European Council (Bangemann report), 1994

- Grimshaw D.J.: Bringing Geographical Information Systems into Business. Longman Scientific & Technical, 1994.
- Koncepce Základní báze geografických dat (ZABAGED), ČÚZK, Praha, 1994.
- Konečný M.: Environmental GIS-Research in Eastern Europe. 7-15. In: Proceedings on International Symposium on Regional Environmental Change and GIS. Keio University Shonan Fujisawa Campus and Meiji University Izumi Campus, Japan, 1994.
- Konečný M.: The GIS Superhighway for Government Decision Support. In: Space and Time in Environmental Information Systems. 618-623. 9th International Symposium on Computer Science for Environmental Protection CSEP'95. Metropolis-Verlag, Marburg. 1995
- Konečný M.: Kartografie a geografické informační systémy pro 21. století. Cartography and Geographic Information Systems for 21st century. 205-220. In: Univerzity na prahu nového tisíciletí. Sborník mezinárodní konference 2.-3.11. 1994, sv. 1, Universitas Masarykiana Brunensis, Brno, Česká republika, 1995.
- Konečný M., Rais K.: Geografické informační systémy, Folia Fac. Scient. Nat. Univ. Purk., Brno, 196, 1985.
- Konečný M., Mikšovský M.: Světová kartografie a její výhled. Geodetický a kartografický obzor 8/92, 1992.
- Konečný M., Mikšovský M.: Nové trendy v kartografii. Geodetický a kartografický obzor 2/96, 37-40, 1996.
- Koškarev A.V., : Metadannyje. 52-55. In: GIS – obozrenije, Leto 95. Žurnal po peredovym tehnologijam v geoinformatike. 1995
- Koškarev A.V., Tikunov V.S.: Geoinformatika, 212, Kartgeocentr-Geodezizdat, Moskva, 1993
- Kremers, H.: Meta-Data: Essentials and Quality of Information Systems, Workshop 2 Materials, Brno-GIS-1994 Conference, 1994.
- Kubo S.: GIS for the Twenty-One Century, K 2-5, In: Proceedings Brno-GIS-1994: Europe in Transition – The Context of GIS. Laboratory on Geoinformatics and Cartography, Brno, 1994.
- Lauterbach F.A.: The Information Highway to Performance. The Impact on GIS. GIS in Business Proceedings, Madrid, 1995.
- Morrison J.L.: What is the Status of Worldwide Topographic Mapping in the Twentyfirst century? Keynote paper Cambridge Conference for National Mapping Organisations, Cambridge, UK, 25 July – 1 August, 1995.
- Neumann J.: Geografická informace. Český výkladový a anglicko-český a česko-anglický překladový slovník. MH ČR. Praha 1996.
- Newton P.W., Zwart P.R. and Cavill M.E., eds.: Networking Spatial Information Systems. John Wiley & Sons, 1995.
- Němec I.: Informace není bič, ale nástroj k odpovědnému rozhodování. Hospodářské noviny, 15. března 1996, s. 7.
- The EGII Discussion Document, GI2000-Towards European Geographic Information Infrastructure (EGII), 31 december 1995, 22 p.
- Usnesení Vlády ČR: Záměr a postup výstavby Státního informačního systému České republiky. 20. září 1995. Praha.
- Veverka B.: Kartografie a GISy v ČR – vývoj, zkušenosti a perspektiva. In: Sborník referátů kartografické konference, Plzeň, 1994.
- Voženílek V.: Czech GIS Data: In Context of European Data Availability. In: Proceedings of Brno GIS 1994, Brno

10. Teorie a praxe zpracování archeologických výzkumů s pomocí prostředků GIS/LIS (Michal Kučera, Jiří Macháček)

(Práce vznikla s podporou grantu reg.č. 404/94/0410 Grantové agentury ČR a grantu AKTION Österreich – Česká republika 10p1)

10.1 Úvod

Zpracování dat formou informačního systému se stává jedním z nejefektivnějších způsobů sběru, organizace a využití velkých objemů informací, tolik typických pro konec 20. století. Informační systémy jsou využívány ve státních institucích, armádě, bankách i soukromých firmách, informační systém lze nalézt v automobilu nebo pomocí něho vyhledat spoj na nádraží. Speciální odnoží informačních systémů jsou GIS (**Geografické Informační Systémy**), jejichž základní charakteristikou je využití prostorových dat.

Za pojmem GIS se skrývá široké pole vědních odvětví. Jen namátkou: informatika, geodézie, fotogrammetrie, dálkový průzkum Země, kartometrie, statistika a jiné. Pro tento průnik disciplín existuje speciální termín – **geomatika** (*geomatics*).

Předložená stať si klade za úkol stručně přiblížit problematiku GIS po stránce teoretické, dále informuje, formou jakési technické zprávy, o nasazení prostředků GIS/LIS na Katedře archeologie Masarykovy univerzity v Brně. Podává také podrobný popis tvorby projektů v prostředí programu MGE-PC2, zpracovávajících části archeologické lokality Břeclav – Pohansko. Jednotlivé fáze budování systému, obecně popsané v úvodní kapitole věnované teorii GIS, jsou vysvětleny a rozebrány na zmíněných konkrétních aplikacích.

10.2 Geografické informační systémy – definice a dělení

Teorie geografických informačních systémů (*Geographical Information System*) je poměrně mladá, a proto nemá jednotné třídění ani terminologii.

Označení GIS zavedl v roce 1963 R.F. Tomlinsen. Dodnes ale neexistuje jednoznačná standardní definice pojmu GIS, ačkoliv na stránkách odborné literatury a tisku probíhá již několik let debata jak o samotné definici, tak i o termínech problematiku provázejících. Situaci komplikuje skutečnost, že kolébkou GIS jsou anglicky mluvící země (první aplikace vytvořeny v polovině 70-ých let v Kanadě). Překlady anglických termínů se mnohdy velmi odlišují, často dochází k tvorbě neologismů. Následující kapitoly by měly populárně naučnou formou problematiku GIS přiblížit.

10.2.1 Definice GIS

Zkratka GIS vznikla ze tří výše uvedených slov. Zamysleme se nad významem jednotlivých slov pojmu GIS (Klimeš 1996).

- **geografický** – adjektivum od slova geografie. Slovo geografie pochází z řečtiny a je složeno ze dvou slov:
geo- (řecky gé = Země) je první částí složených slov, znamená Země, týkající se Země
-grafie (řecky grafein = psát) druhá část složených slov. Vyjadřuje, že jde o zapisování, popisování, kreslení, rýsování či tisk toho, co uvádí první část složených slov
Geografie je potom věda popisující Zemi, věda zabývající se geosférou.

- **informační** – adjektivum od slova informace. Informace je údaj, směrovaná zpráva o charakteristikách, stavech systémů, procesů a jejich částí, případně množství údajů, přenesených sdělovacím kanálem za jednotku času.
- **systém** – účelově definovaná množina (entita, soubor) prvků (částí, jednotlivých objektů) a vazeb mezi nimi, které spolu určují vlastnosti, chování a funkce systému jako celku. Prvek systému je na dané úrovni nedělitelná část. Vazba systému realizuje spojení mezi prvky systému. Vazby a prvky tvoří strukturu systému.

Informační systém je potom systém, jehož vazby tvoří informace. Mezi prvky tohoto systému probíhá výměna informací. Výměna probíhá i mezi systémem a jeho okolím. Okolí je taktéž tvořeno množinou prvků, které sice nejsou součástí systému, ale vykazují k některým jeho prvkům vazby. Tyto realizuje vstup a výstup systému. Takto definovaný systém nazýváme otevřený (Konečný – Rais 1985).

Na základě výše uvedeného je možno definovat GIS:

„Geografický informační systém je systém lidí a technických a organizačních prostředků, který provádí sběr, přenos, uložení a zpracování údajů za účelem tvorby informací vhodných pro další využití v geografickém výzkumu a jeho praktických aplikacích“ (Konečný – Rais 1985).

„Organizovaný souhrn počítačové techniky, programového vybavení, geografických dat a zaměstnanců navržený tak, aby mohl efektivně získávat, ukládat, aktualizovat, manipulovat, analyzovat a zobrazovat všechny druhy geograficky vztažených informací. Pomocí GIS lze provádět komplex prostorových operací, které by bylo obtížné, časově náročné nebo prakticky nemožné provádět jinými prostředky“ (ARCDATA 1993).

Velmi podobný je následující výklad pojmu GIS (Vojtička 1995): „GIS je organizovaný soubor počítačového hardveru, softveru, geografických údajů a personálu určený na efektivní zber, uchovávanie, obnovovanie, manipuláciu, analýzu a zobrazovanie všetkých geografických vzťahov informácie.“

Chmelík (1995) ve své práci cituje R. Billa a D. Fritsche: „Geoinformační systém je počítačově podporovaný systém, který se skládá z technického vybavení, programového vybavení, dat a uživatele. Pomocí něho je možno prostorová data v digitální formě pořizovat a redigovat, ukládat a reorganizovat, modelovat a analyzovat, stejně jako alfanumericky či graficky prezentovat.“

Netradičně GIS definuje Babický (1994): „GIS je systém, jehož cílem je snižovat entropii informací o území.“

Zmíněná vymezení pojmu GIS musíme upřesnit. Vycházejí z výše uvedených významů slov geografický, informační a systém, je nutno dodat:

- prvky GISu jsou jednoznačně definovány a je jich konečný počet
- na rozdíl od tabulkových procesorů a databázových produktů patřících také do kategorie informačních systémů, GIS navíc obsahuje prostorové údaje, informace o poloze, geometrii, topologii objektu.

10.2.2 Dělení GIS

Jak bylo zmíněno v úvodu, definovat GIS je problematické, neboť zkratka GIS značí technologii, teorii, metodologii i vlastní aplikace. S přihlédnutím ke skutečnosti, že hranice mezi jednotlivými skupinami je velmi nezřetelná, se množina GIS dělí následovně (Orlík 1993, Kaláb 1993):

Systémy pro digitální mapování (CAM - *Computer Aided Mapping*) – pořizování a správa digitálních mapy. Začínají nahrazovat klasické metody tvorby a údržby mapy. Mají rozsáhlé editační možnosti a jsou součástí či nadstavbou CAD (*Computer Aided Design, Drafting*) systémů. Do skupiny CAM patří také systémy pro poloautomatickou vektorizaci.

Informační systémy o území (LIS *Land Information System*). Tyto systémy umožňují vedení a správu digitálních map. Mapy je možné doplnit o textová data, vesměs ve formě databázových tabulek, s možností jednodušších analýz. Většina těchto systémů pracuje nad daty ve vektorové formě. Spojení databáze s prvky výkresu je skrze jednoznačný identifikátor. Systémy neumí řešit topologické vztahy mezi objekty. **Do této kategorie**

patří MGE-PC2. Speciální variantou LIS systémů jsou AM/FM (*Automatic Mapping/Facility Management*) – systémy pro správu rozsáhlých sítí (komunikačních, inženýrských) především pro města a průmyslové podniky. Bývají problémově orientované.

Grafické informační systémy (GIS *Geographical Information System*) Teprve zde hovoříme o opravdových GISech. Jedná se zpravidla o složité, obrovské systémy, zhusta fungující na počítačové síti. Rozlišujeme systémy rastrové a vektorové (srov.obr 2). Vektorově orientované systémy vycházejí z topologického datového modelu. Samozřejmě umožňují i některé operace a analýzy nad rastrem. Naopak rastrově orientované GISy zajišťují import dat ve vektorovém formátu, což rozšiřuje možnosti analýz. Nejmodernější systémy dovolují provádět i modelování: „Co se stane když...?“ (*What if...*).

Systémy pro manažerské mapování (DMS – *Desktop Mapping System*) – prohlížečky map umožňující i analýzy. Slouží především k prezentaci a k rozhodování. Nedovolují sběr dat a zpravidla ani jejich editaci. Lze však připojit fotografie, zvuky a videosekvence. Dále pak systémy řeší výstup do tabulkových a textových procesorů.

10.3 Fáze tvorby GIS

Při pohledu na GIS z hlediska časového a logického postupu tvorby se rozlišují tyto fáze tvorby GISu:

- úvodní studie
- sběr dat
- správa dat
- analýza nad daty
- prezentace dat

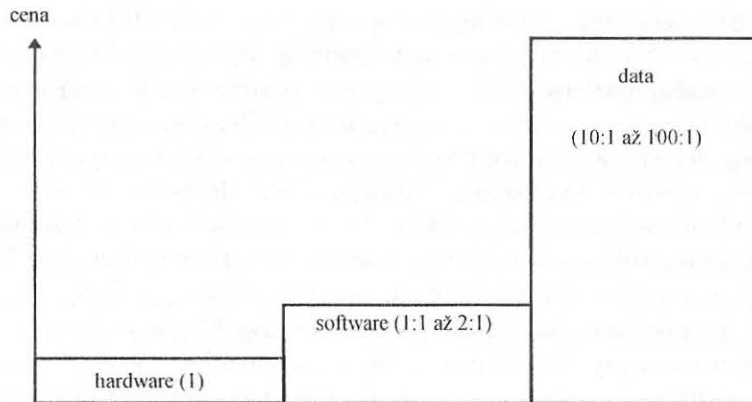
Jednotlivé etapy tvorby GISu následují v uvedeném pořadí. Jakmile je základní struktura a datová základna GISu vytvořena dochází k periodické obnově systému (aktualizace dat, modernizace hardware a software). V tomto případě se jednotlivé etapy mohou prolínat.

10.3.1 Úvodní studie

Vybudování funkčního systému, který v dostatečné míře uspokojí požadavky uživatele, je finančně a časově velmi náročnou záležitostí. Prvním krokem při vytváření GISu je proto sestavení týmu pracovníků, kteří provedou strategickou studii. Základním kritériem je poměr nákladů na systém a zisku, který tento systém přinese. Náklady netvoří pouze prvotní nákup hardwaru a softwaru, ale především obrovské investice do sběru, aktualizace a správy dat, výškolení personálu, modernizace počítačového a programového vybavení, jeho údržba (obr.1). Návratnost vložených prostředků se přitom nepočítá na týdny či měsíce, nýbrž na roky, případně desetiletí.

Úvodní studie musí být proto velmi pečlivá. Výběru definitivního programového i hardwarového vybavení by mělo předcházet navrzení struktury a vytvoření **pilotního projektu** pomocí 2–3 předem zvolených produktů.

Teprve na základě odzkoušení, zhodnocení a diskuse nad pilotními projekty je možno se definitivně rozhodnout pro jeden produkt. Následuje jeho nákup, instalace, výchova pracovníků. Je navržena definitivní struktura projektu (způsob organizace dat) vzhledem k jejich budoucímu využití (správa, analýzy, prezentace). Návrh si zaslouží velkou pozornost, neboť kompletnost a přesnost databáze určuje kvalitu analýz a výsledných produktů.



Obr.1. Poměr cen jednotlivých komponentů GISu (zpracováno podle sborníku EGIS 1990)

10.3.2 Sběr dat

Sběr dat je základní a zároveň nejdražší etapou budování GISu (obr.1). Cílem je vytvoření digitální databáze (mapové, atributů), která je základem budoucího systému. Předpokladem úspěšného, efektivního sběru dat je kvalitně připravená metodika a technologie sběru a volba vhodných zdrojů dat.

Provádění úprav v databázi v průběhu analýz nebo dokonce až při tvorbě tematických map je velmi nákladné, časově náročné a mnohdy i téměř nemožné. Navíc dobře navržená databáze je použitelná i pro více projektů.

Základním atributem geografických dat je prostorovost. Prostorovost dat znamená jejich jednoznačnou polohovou lokalizaci pomocí souřadnic ve zvoleném jednotném souřadnicovém systému (třírozměrném nebo rovinném).

Dále je nutné mít na zřeteli kvalitu dat. Rozhodují tato kritéria hodnocení kvality dat (Chmelík 1995):

1. obsahová úplnost vzhledem k požadavkům uživatele
2. geometrická přesnost dat a jejich věcná správnost
3. aktuálnost dat

10.3.2.1 Zdroje dat

Zdrojem dat jsou všechny dostupné a využitelné soubory informací o zájmovém území, především zájmové území (část geosféry) samotné. Jako zdroje dat slouží :

- mapy (základní, státní, topografické, tematické, účelové, geografické, aj.)
- plány, náčrty, výkresy
- výstupy z CAD (mapy a jiné výkresy v digitální formě, digitální modely terénu, aj.)
- informace z jiných informačních systémů
- statistické ročenky, zprávy, slovníky, rejstříky, přehledy, archivní údaje, aj.

10.3.2.2 Metody sběru dat

Metody sběru dat pro GIS lze rozdělit do dvou základních skupin:

10.3.2.2.1 Získávání dat přímo ze zájmového území

Sem patří především **geodetické metody**. Tyto zahrnují jak klasické měření úhlů a délek, v současnosti realizovaná především pomocí elektronických dálkoměrů a totálních stanic, tak metody GPS (*Global Position System* – určení polohy bodu v prostoru využitím družic), případně kombinace terestrických a družicových metod. Data jsou získávána v přímém kontaktu s objektem.

Ke geodetickým metodám mají blízko **fotogrammetrické metody**. Data o rozměru, tvaru a poloze objektu v prostoru se nezískávají přímým měřením v terénu, nýbrž vyhodnocením měřických snímků. Především pro environmentální aplikace GIS se používají **metody dálkového průzkumu Země (DPZ)**.

Další metody přímého získávání dat (především negrafických) jsou místní šetření, rozhovory, ankety, průzkumy, aj.

10.3.2.2 Přebírání dat z již dříve vytvořených souborů dat

Jedná se o využití částí, případně celých existujících souborů informací o zájmovém území. Velmi výhodné je získávání dat z digitálních médií (výkresy CAD programů, digitální výstupy z jiných informačních systémů, aj.), kdy se soubory relativně snadno importují do příslušného GIS produktu.

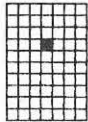

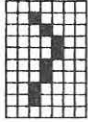

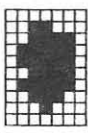
Převod analogových podkladů (map, plánů, archívních údajů, aj.) je mnohem náročnější, neboť veškerá data je třeba, většinou ručně, vložit do počítače.

10.3.2.3 Vstup dat do počítače

Výsledkem sběru dat, popsaného v předchozí kapitole jsou soubory informací v digitální (souřadnice, předpisy kresby, databázové tabulky, aj.) nebo analogové formě (snímky, mapy, zápisníky, atd.). Těmito daty je nutno podle předem vytvořených kritérií a metodiky naplnit databáze vytvářeného informačního systému.

Vstup dat v digitální formě se provádí formou importu a následné úpravy souborů pomocí příslušných softwarových prostředků.

Vstup **analogových negrafických informací** do počítače se děje obvykle prostřednictvím klávesnice, někdy je možné text scannovat s následným použitím programu pro rozpoznání písma OCR (*Optical Character Recognition*). Takto získané informace jsou v počítači uloženy v podobě databází, textových nebo tabulkových souborů.

prvek	vektor forma		rastr forma	
	digitální	analogová	digitální	analogová
bod	souřadnice x, y	•	pixel	
čára	posloupnost souřadnic x, y		pixel	
plocha	uzavřená posloupnost souřadnic x, y		pixel	

Obr. 2. Rastrová a vektorová data (zpracováno podle sborníku EGIS 1990).

Grafické předlohy se do počítače převádějí buď do vektorového nebo rastrového formátu viz obr.2. Převod do rastrového formátu se děje pomocí scannerů (skenerů). Výsledkem scannování (skenováním) je sejmutý rastrový obrázek (bitmapa).

Převod předlohy do vektorového formátu lze uskutečnit vektorizací přímo z grafické předlohy pomocí digitizéru (tabletu), případně vektorizací naskenované předlohy na obrazovce počítače. Vektorová struktura dat rozlišuje tyto útvary:

Bodový prvek je vyjádřen diskrétní polohou definující mapovaný prvek, jehož hranice a tvar jsou příliš malé, než aby mohl být vyjádřen jako linie či plocha.

Liniový prvek je uspořádaný soubor souřadnic, které po svém zřetězení vyjadřují tvar linie takového prvku v mapě, jenž je příliš úzký, než aby mohl být vyjádřen jako plocha. Případně to je prvek, nemající šířku např. vrstevnice.

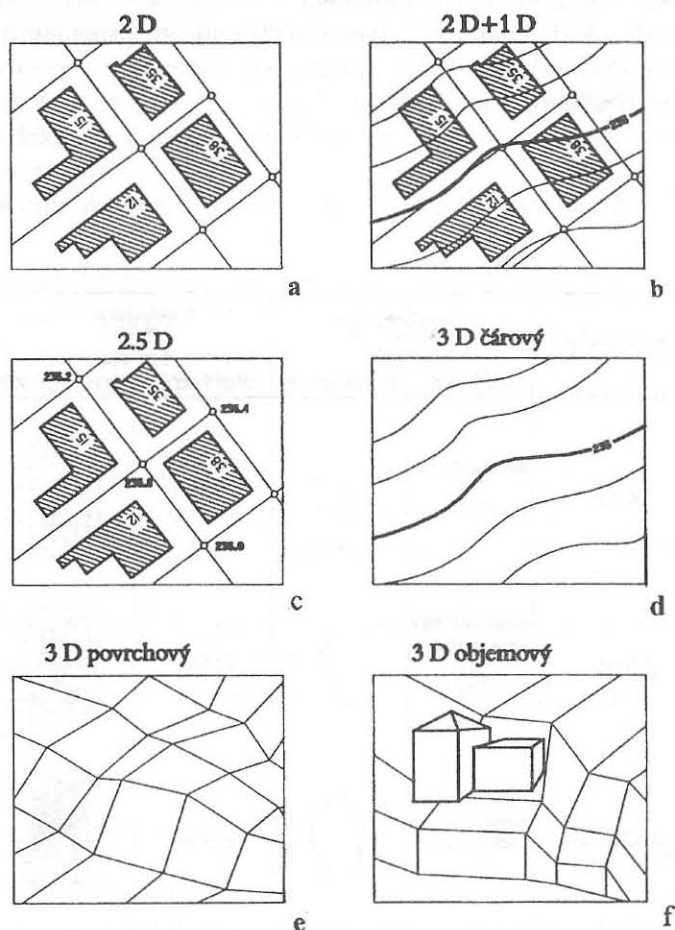
Plošný prvek je uzavřený obrazec, jehož hranice vymezují homogenní oblast.

Jak jsme již uvedli, základním atributem dat geografického informačního systému je prostorovost. Z hlediska geometrie je možné data rozdělit následovně (sborník EGIS 1990):

2D (dvourozměrná) – jsou dány dvě souřadnice (X,Y) (obr.3-a).

2D+1D (dvou a jednorozměrná data) – k datům určených v rovině je přiřazena vrstva s informací o třetím rozměru (vrstevnice), neexistuje ale vztah mezi polohou a výškou (obr.3-b).

2.5D (dvou a půlrozměrná) – k polohovým datům je třetí rozměr (kóta) přiřazen jako atribut (obr.3-c).



Obr. 3. Dělení dat z hlediska geometrie (zpracováno podle sborníku EGIS 1990).

3D (trojrozměrná) – existuje více možností struktury dat

- planimetrie a připojený digitální model terénu (obr.3-d).
- povrchový model (*surface model*) Jedná se o povrch generovaný ze čtvercové nebo trojúhelníkové sítě bodů (obr.3-e).
- objemový model (*volume model*), kde je terén modelován pomocí krychlí, kuželů (obr.3-f).

4D (čtyřrozměrná data) – jako čtvrtý rozměr funguje čas (časoprostorová informace)

10.3.2.4 Objekt, topologie

Základním principem GISu je, zjednodušeně řešeno, spojení grafické a negrafické databáze do smysluplného celku. Grafické elementy jsou umístěny ve výkresovém souboru, tzv. **mapě**, negrafické údaje jsou uloženy ve formě databázových tabulek, v **tabulkách atributů**. Jejich svázání s prvky výkresu (mapy) zajišťuje jednoznačný identifikátor (podobným způsobem je realizováno relační spojení tabulek v databázi). Z grafického elementu se stává **objekt**.

Objektem se rozumí konkrétní kvalitativní a kvantitativní jednotka. Objekt má vlastní identitu, je fyzicky, geometricky a polohově určen, je tematicky jasně identifikovatelný. Objekt je součástí hierarchie (např. objekt **DUB**, znázorněný v mapě obrázkem dubu a nesoucí databázový záznam o stáří, rozměrech, data vysazení, atd., je součástí skupiny objektů **LES LISTNATÝ**).

Objekty mohou být určeny nejen svojí polohou v prostoru, ale i prostorovými vztahy mezi nimi (vzájemné spojení, sousednost, návaznost). Hovoříme o **topologii**. Topologie je matematický postup založený na explicitním definování prostorových vztahů mezi objekty.

10.3.3 Správa dat

Správou dat se rozumí (Chmelík 1990) soubor technických, technologických a personálních opatření s cílem, co nejlépe organizovat data uložená v počítači, včetně jejich ochrany, zajištění přístupu k nim. Systém musí být otevřený s možností aktualizace dat.

10.3.4 Analýzy nad daty

Jedním ze základních cílů budování GISu je zkvalitnění a zrychlení rozhodování. GIS má umožnit řešení problémů, které by bez jeho použití bylo velmi obtížné, ne-li nemožné. Po sběru a vybudování funkční struktury dat přichází fáze jejich využívání, fáze analýz.

10.3.4.1 Základní typy otázek, které GIS umí vyřešit

V úvodní kapitolách vymezujících pojem a dělení GIS jsme načrtli klasifikaci produktů GIS s poznámkou, že hlavním parametrem třídění GIS produktů je právě jejich schopnost analýz.

Systémy GIS i LIS dokáží zodpovědět dotazy, týkající se (ARCDATA 1993):

- 1) polohy – dotaz zjišťuje, co se nachází na konkrétním místě.
- 2) podmínky – je třeba vyhledat místo, které splňuje jisté podmínky.

Pouze velké GISy řeší i dotazy týkající se

- 3) trendů – zahrnuje oba předchozí dotazy a zjišťuje změny v analyzované oblasti v průběhu času.
- 4) prostorového uspořádání – zjišťuje se pravidelné prostorové uspořádání určitého jevu v závislosti na zadaných podmínkách.
- 5) modelování – analýzy typu „Co se stane, když...“ („What if...“)

10.3.4.2 Nástroje pro analýzu

Kromě dotazování na tabulku atributů, např. pomocí databázového jazyku SQL (*Structured Query Language*) s možností lokalizace vybraných objektů a jevů na mapě dovolují GISy i složitější analýzy. K nim patří především:

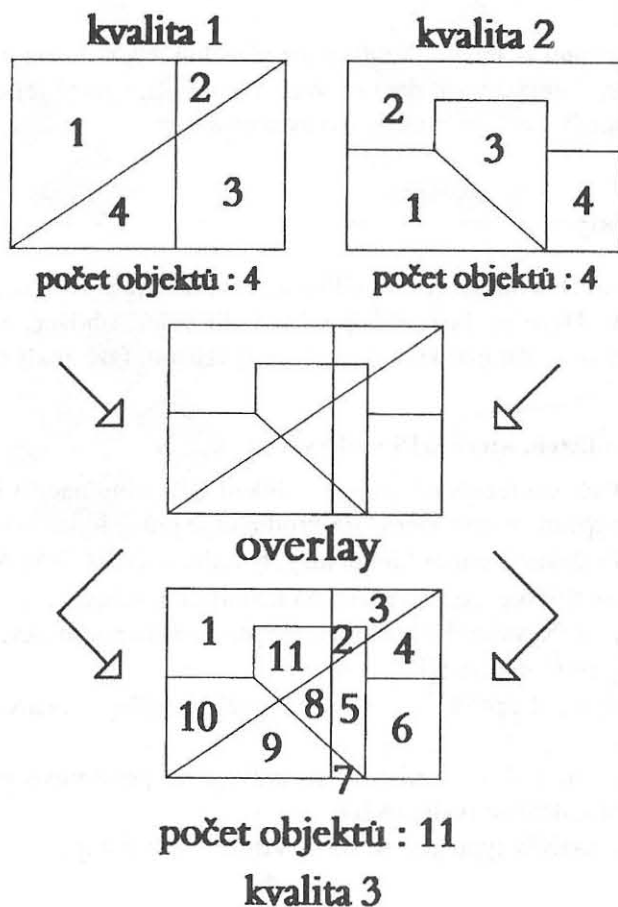
Generování obalových zón (buffers). Tato funkce slouží k prostorovému vymezení dotazu. Obalová zóna požadovaných parametrů se vytvoří kolem jednoho či více vybraných objektů. Např. *Vyber a zobraz jehličnaté stromy se stupněm poškození 3 vzdálené více než 6 a méně než 17 km od elektrárny.* V archeologické aplikaci se můžeme dotázat například: *Vyber a zobraz neopevňená rovinná sídliště kultury horákovské, nacházející ve vzdálenosti 15 km od hradisek téže kultury.*

Analýzy s využitím topologie, kdy je dotaz na tabulku atributů a případně geometrická data navíc determinován podmínkou týkající se topologie. Klasickým příkladem je například dotaz: *Vyber mi všechny parcely o rozloze větší než 4000 m², vzdálené 800 m severně od kostela.*

Příklad z archeologie: *Vyber a zobraz pohřebiště únětické kultury ve vzdálenosti 30 km východně od řeky Moravy.*

Velmi silným nástrojem GIS je **topologické overlay** (výraz se nepřekládá). Jedná se o překrývání polygonových vrstev, přičemž vznikají nové vrstvy s požadovanými vlastnostmi (obr.4). Příklad: *Ve vrstvě KULTURY vyber zemědělské plochy oseté pšenicí. Dále vyber zemědělské plochy s výměrou nad 50 ha ve vrstvě VÝMĚRA PLOCH. např. ZEMĚDĚLSKÁ PŮDA O VÝMĚŘE VĚTŠÍ NEŽ 50 HA OSETÁ PŠENICÍ.*

Příklad z archeologie: *Ve vrstvě PŮDY vyber spraše. Dále vyber sídliště s lineární keramikou v nadmořské výšce větší než 250 m n.m. S použitím overlay je vytvořena nová vrstva SÍDLIŠTĚ S LINEÁRNÍ KERAMIKOU V NADMOŘSKÉ VÝŠCE VĚTŠÍ NEŽ 250 M N.M. NA SPRÁŠI.*



Obr.4. Topologické overlay (zpracováno podle sborníku EGIS 1990).

Modelování jevů, které mají svoji analogii v reálném světě. Geografický model je souhrn funkcí a procedur využívajících prostorové analýzy, jenž umožňuje získání nových kvalitativních a kvantitativních informací. Modely jsou aplikovány především u environmentálních GISů, tedy při řešení problémů životního prostředí, deformací reliéfu, aj.

K dalším analytickým nástrojům patří funkce dovolující měření, výpočty, statistické funkce. GIS poskytuje výsledky ve formě reportů, tabulek, grafů. Toto je už ale otázka, týkající se prezentace dat.

10.3.5 Prezentace dat

Poslední fází budování GISu je prezentace dat. Prezentace se děje buď pouhým výběrem požadovaných dat, nebo demonstrováním výsledků analýz.

Data lze prezentovat v digitální formě (interaktivně) na obrazovce počítače (terminálu) nebo ve formě grafické. Ke zpřístupnění digitálních dat uživateli GISu slouží tzv. prohlížečky, jednoduše ovladatelné programové produkty, pomocí nichž „surová“ data z počítače efektně zobrazí na obrazovce počítače. Pomocí prohlížečky se dají provádět také jednodušší analýzy. K vybraným objektům lze zobrazit rastrové obrázky (nascannované fotografie), záznamy videa a animace, případně i zvuky. V současné době je možné implantovat prohlížečky do sítě Internet.

Grafické výstupy představují především tematické mapy, plány, výkresy, tabulky, atlasy, statistické přehledy, databázové reporty, realizované na plotru (vektorovém či rastrovém) a tiskárně.

10.4 Programové vybavení KA FF MU

10.4.1 MicroStation

Programový produkt MicroStation fy. Bentley patří k největším a zároveň nejlepším grafickým programům CAD (*Computer Aided Design*).

10.4.1.1 Obecná charakteristika programu

Prostředí programu MicroStation připomíná a jeho ovládání připomíná systém Windows. Jednotlivé příkazy lze buď vypsat na povelový řádek, nebo použít příslušnou ikonu. Základní pohyby v programu a jeho ovládnutí na nižší uživatelské úrovni je intuitivní a tudíž poměrně snadné.

10.4.1.2 Práce s programem

Vektorová kresba (výkres), kterou vytváříme na obrazovce počítače je soubor *.dgn. Základním prvkem výkresu je entita. Entitou se rozumí úsečka, kružnice, oblouk, aj.

Každá entita nese svoje atributy: barvu, tloušťku čáry, typ čáry (plná, čerchovaná, tečkovaná, aj). Dále má entita určité vlastnosti. Je to především informace, zda představuje desku nebo otvor, v případě desky je-li vyplněná barvou či ne.

Jednotlivé entity se sdružují do vrstev – množin spolu souvisejících entit. Vrstvy si lze představit jako průsvítky. Například ve vrstvě 1 jsou zachyceny komunikace, ve vrstvě 2 zástavba. Zapnou-li se obě vrstvy najednou, vidíme zástavbu včetně komunikací. Program rozlišuje celkem 63 vrstev.

MicroStation je silným nástrojem nejen pro vlastní kreslení, ale také pro editaci a modifikaci nakreslených prvků (entit). Systém umožňuje úpravy pracovního prostředí a to jak na uživatelské úrovni (viz. appendix Uživatelské prostředí **POHAN**), tak na úrovni programátorské (implantován programovací jazyk MDL). Těmito prostředky je možné upravit poměrně obecný CAD do podoby výkonné aplikace pro sběr dat.

MicroStation dovoluje propojení elementů výkresu se záznamy z databázových tabulek, a tím vytvořit jednoduchý LIS. Tento proces je ale uživatelsky poměrně náročný, výhodnější je použít některou z GIS/LIS aplikací nad MicroStation pracujících. K těmto programům patří i produkty MGE a MGE-PC2.

10.4.2 MGE-PC2

MGE-PC2 je nástroj LIS pro sběr, modifikaci, dotazování a analýzu geografických dat. Je nadstavbou programu MicroStation, pro MS DOS či WINDOWS. V této kapitole jsou uvedeny základní informace o software MGE – PC2. Podrobněji budou některé funkce rozebrány v kapitole týkající se projektů řady POHAN.

10.4.2.1 Obecná charakteristika systému

MGE-PC2 je PC – variantou softwarového GIS produktu MGE (*Modular GIS Environment*) fy. Intergraph určená pro systém DOS, WINDOWS a WINDOWS NT. Podobně jako „velké“ MGE pracuje i jeho verze pro osobní počítače na platformě CADu MicroStation fy. Intergraph (dnes již Bentley), konkrétně pod verzí 5.0. Podstatou programu je spojení výkresů, souborů *.dgn s externí databází. Ideální je relační databáze Oracle, systém si rozumí i s formátem Xbase. Výhodou je, že při tvorbě tabulek formátu Xbase přímo v projektu v MGE-PC2 není nutné mít k dispozici samostatný databázový program, protože software MGE-PC2 dokáže požadované tabulky sám vytvořit. Je možno volit mezi formáty: dBase III+, dBase IV, a FoxPro.

10.4.2.2 Projekt

Projekt (*project*) je nejvyšší hierarchickou jednotkou. Projekt zahrnuje veškerá geometrická i negrafická data, dotazy, prostorová vymezení a výsledky analýz, které se týkají daného zájmového území. Projekt se skládá ze dvou základních částí – dvou základních adresářů umístěných v podstatě libovolně v počítači. Jejich spojení zajišťuje prostřednictvím nastavení cest schéma (*scheme*) projektu.

V prvním adresáři jsou uchovávány tabulky atributů, tabulky databázového serveru MicroStation a ovládací tabulky pro MGE.

Druhý adresář (nesoucí jméno projektu) obsahuje následující podadresáře:

..\data	parametry pro převod projektu do systému MapInfo
..\dgn	výkresy (<i>maps</i>)
..\fence	uložené výběrové ohrady (<i>fence</i>)
..\idx	geografické indexy (<i>geographic index</i>) a mapu sousedství (<i>vicinity map</i>)
..\qry	uložené dotazy (<i>query</i>)
..\rpt	zprávy (<i>reports</i>)
..\seed	zdrojové výkresy (<i>seed files</i>)
..\setup	soubory nastavení a konfigurací
..\ulf	uložené výběrové množiny (<i>selection sets</i>), fronty (<i>queue</i>)
..\zone	uložené obalové zóny (<i>buffer zones</i>)

Pomocí funkce **Export** se oba adresáře automaticky zkomprimují do souboru *.mpd. Tímto způsobem lze projekt přenášet na odlišné platformy (UNIX), do „velkého“ MGE.

10.4.2.2.1 Struktura projektu

Základní jednotkou projektu je objekt (*feature*). Ten je definován typem, kódem, barvou, umístěním do vrstvy, stylem a tloušťkou čáry. Informace o objektech sdružuje databázová tabulka *feature.dbf*. K objektu je pomocí jednoznačného identifikátoru (*mblink*) navázána negrafická informace – jeden záznam z tabulky atributů. Tabulky atributů jsou uživatelem vytvořené a vyplněné databázové tabulky. Pro jeden druh objektu lze definovat pouze jednu tabulku. Databázové tabulky je možno vytvářet přímo v MGE-PC2. Druhou možností je připojení externě vytvořených tabulek.

MGE-PC2 dovoluje vytvářet objekty dvěma způsoby:

- přímou digitalizací definovaných objektů, kdy se digitalizovaným elementům rovnou přiřadí vlastnosti objektu s tím, že je možno ihned vypsát záznam atributů v tabulce

- přisouzením funkce objektu předem nakresleným elementům výkresu pomocí procedury **Feature maker**. Také zde je možné ihned vyplnit prázdný záznam, případně vybrat požadový záznam z externě připojené tabulky pomocí dotazu (*query*)

Objekty jsou fyzicky umístěny ve výkresech formátu *.dgn, které v prostředí MGE dostaly pojmenování mapy (*maps*). Na logické úrovni jsou objekty seskupeny do kategorií (*categories*).

Je jasné, že kategorie může sdružovat více map. Tyto jsou s kategorií svázány umístěnými objekty a pomocí geografického indexu (*geographical index*) – speciálního výkresu, ve kterém jsou nakresleny hranice jednotlivých map příslušejících do určité kategorie. Geografický index ulehčuje lokalizaci objektů při dotazu (*query*) a logickém výběru (*feature display*).

10.4.2.3 Dotazy

10.4.2.3.1 Jednoduché dotazování

Nejatraktivnější částí systému je tvorba dotazů. MGE -PC2 obsahuje pro práci s dotazy poměrně silný aparát (tvorba dotazů, jejich kopírování, provádění, ukládání a opětné vyvolání), založený na jazyku SQL (*Structure Query Language*). Je možné dotazovat se pouze na tabulku atributů nebo pouze na určitý prvek, dotaz lze navíc prostorově vymežit.

Dotazování **pouze na tabulku atributů** se využívá především k výběru záznamů z externě vytvořené tabulky pro připojení (*link*) k objektu.

Cílem většiny dotazů bývá především lokalizace otázkou vymezeného prvku nebo souboru prvků, spojená se současným zobrazením příslušného záznamu.

10.4.2.3.2 Prostorové vymezení dotazu (*spatial constraint*)

Dotaz (*query*) je možné prostorově determinovat následujícími způsoby:

- mapou (*map*)
- ohradou (*fence*) – funkce shodná s MicroStation
- obalovou zónou (*buffer zone*) – obalové zóny je možno tvořit kolem bodů, čar i areálových objektů. Zóny můžeme ukládat, znovu vyvolávat a slučovat (*merge*). Buferové zóny se mohou stát součástí mapy, dokonce i samostatným objektem.
- výběrovou množinou (*selection set*)

Jednotlivá prostorová vymezení je možné uložit do souboru a kdykoli je znovu vyvolat a použít pro další dotazy.

Po vytvoření dotazu (vypsáním SQL podmínky, nebo interaktivně) a jeho specifikování se výsledek barevně zvýrazní (okrajem či výplní) na obrazovce počítače. Při listování vybranými záznamy se lokalizuje nalistovaný záznam v mapě. Uživatel tak vidí objekt, jehož atributy si právě prohlíží či modifikuje. Z dotazem vymezených záznamů lze sestavit zprávu (*report*). Absenci relačních dotazů nahrazuje možnost uložit výsledek databázové otázky do výběrové množiny (*selection set*) a touto posléze vymežit další dotaz.

10.4.2.4 Prezentace

Samotné MGE-PC2 neobsahuje žádný modul určený k prezentaci. Program pouze využívá možnosti MicroStation (plotování, sejmutí obrazovky do bitmapy, export do výměnného formátu *.dxf). Výsledek databázového dotazu (*report*) lze importovat např. do tabulkového procesoru, kde je možné jej upravit do formy tabulky a grafu.

Pro interaktivní prezentaci je možné importovat zkomprimovaný projekt *.mpd do prohlížečky VistaMap.

10.5 Projekty řady POHAN

10.5.1 Úvod do problematiky projektů (dokumentace archeologického výzkumu)

Základním zdrojem informací pro archeologa je dokumentace, vytvářená přímo v terénu, na archeologické lokalitě. V terénu se vyhotovuje **dokumentace kresebná, písemná a fotografická**. Kresebná část se provádí především v měřítku 1:20, detaily potom 1:10 na milimetrový papír. Kreslí se tužkou a pastelkami. Prvky se zaměřují ortogonální metodou s použitím pásma, olovnice, několika skládacích metrů, případně pentagonu. Každá odkrytá úroveň se výškově určí metodou plošné nivelace, přičemž se do měřického náčrtu zaznamenávají buď absolutní výšky v systému Balt po vyrovnání, nebo relativní kóty vztažené ke vhodně stabilizovanému výškovému bodu. Při nivelaci je nutné volit body co nejpělivěji, aby při jejich minimálním počtu byl co nejdříve postižen reliéf terénu.

Vzniklé měřické náčrtky (plány) slouží nejen jako podklad pro interpretaci a analýzy, ale také se z nich sestavuje přehledný plán části či celé plochy výzkumu. Mají-li být použity pro publikační účely, překreslují se perem.

Ze stručného nástinu způsobu kresebné dokumentace archeologického výzkumu – dodnes používané – je cítit archaičnost. Absence novějších měřících i zpracovatelských technologií je způsobena nedostatkem financí. Moderní technika se v archeologii používá zpravidla pouze pro geodetické práce velkého rozsahu (zaměření lokality jako celku do mapy), případně pro prvotní prospekci – použití geofyzikálních metod.

Je zřejmé, že práce při dokumentování výzkumu by ulehčilo použití totální stanice se automatickým záznamem naměřených hodnot a následným zpracováním zápisníku automatizovaně v počítači. Výsledkem by byl soubor souřadnic (rovinných x, y , případně prostorových x, y, z). Vypočtené souřadnice by sloužily jako výchozí data pro další zpracování v PC – vytvoření kresby pomocí CAM/CAD systému, případně náročnější modelování situace pomocí GIS/LIS produktů.

Výhoda tohoto způsobu sběru dat je nasnadě. Zcela by odpadlo udržování čtvercové sítě a nivelace plochy, kresebná dokumentace by se zredukovala na stylizovaný náčrt „od ruky“.

V případě **písemné dokumentace** se použití výpočetní techniky přímo nezbytné. Počítač vybavený vhodným textovým, databázovým a tabulkovým procesorem se stává standardním nástrojem pro práci archeologa. Nálezová situace se již nezpracovává formou ručně psaných deníků výzkumu a nálezových zpráv, ale je, nejlépe přímo v terénu, vkládána do počítače.

Také metodika **fotografické dokumentace** by mohla být zefektivněna. Zatímco užití metod analogové i analytické fotogrammetrie je pro dokumentování výzkumu, vzhledem k potřebnému technickému i personálnímu vybavení, téměř nemožné, nabízí se možnost tyto metody nahradit aplikováním softwarových produktů pro zpracování rastru. V rámci diplomové práce (Kučera 1996) bylo provedeno srovnání přesnosti zpracování hrobové jámy metodou dvousmímkové fotogrammetrie (vyhodnocení na Stecometru a následné analytické zpracování v programu ORIENT) a jednoduchá transformace nescanovaného snímku pomocí programového modulu IRAS/C, pracujícího pod programem MicroStation. Transformace dopadla při srovnání s velmi přesnou fotogrammetrickou metodou velmi dobře, nyní je třeba ji otestovat na různých nálezových situacích. Každopádně je vhodné především u složitých, cenných nálezových situací provádět fotografování tak, aby bylo možné snímky označit, byť s výhradami, za měřické. Znamená to vytyčit v rámci zájmového prostoru před fotografováním vřícovací body.

10.5.2 Úvodní studie

10.5.2.1 Formulace problému

Předchozí kapitola stručně pojednává o způsobech archeologické dokumentace a zmiňuje se o nutnosti použití výpočetní techniky při sběru dat v terénu. Výsledkem práce archeologa v terénu je kvantum informací, dat. Tyto informace jsou pouze výchozím materiálem pro další zpracování, analýzy, interpretace a publikování. Vezmeme-li v úvahu definici GIS (kap.10.2.1), zjistíme, že technologie GIS se k aplikacím pro archeologie přímo nabízí.

10.5.2.2 Výběr programového vybavení

Rozhodnutí o koupi akademických verzí programů MicroStation a MGE-PC2 pro Katedru archeologie FF MU padlo na jaře roku 1995. Předcházela mu několikátýdenní diskuse. Základní požadavky byly jasně formulovány (viz předchozí kapitola), bylo nutné zvolit vhodné programové vybavení, v závislosti na dostupném hardware. Jako alternativa k MGE-PC2 se uvažovalo o systému ARC/INFO. Díky velmi výhodným finančním podmínkám zvítězilo MGE-PC2. Dalším nezanedbatelným faktorem hovořícím pro MGE-PC2 byla skutečnost, že výstupy programu MicroStation (výkresy *.dgn) se stávají standardním formátem přenosu dat mezi státními institucemi i soukromými (geodetickými) firmami. Popis programů přinesla kapitola 10.4.

10.5.2.3 Návrh struktury projektu. Stručně o projektech řady POHAN

Po důkladném prostudování manuálů programu MGE-PC2 byla navržena struktura dvou pilotních projektů, zpracovávajících dílčí části archeologické lokality Pohansko u Břeclavi. Projekty byly nazvány podle lokality, v názvech dále figuruje letopočet sezony výzkumu.

Projekt **POH 95** zahrnuje data záchranného výzkumu, prováděného v roce 1995 v části lokality u loveckého zámečku. Metodika dokumentace tohoto výzkumu byla plně podřízena budoucímu zpracování dat formou LIS (*Land Information System* – informační systém o území).

Sběr dat pro tento projekt probíhal přímo na lokalitě, jednotlivé plány byly okamžitě po zaměření a nakreslení v terénu převáděny do počítače, nálezová situace se popisovala v terénu pomocí formuláře (dle principů kontextuální formulářové archeologie). Formulář byl ihned vkládán do počítače. Pouze technické problémy (krátká životnost akumulátorů notebooku) nedovolily realizovat původní záměr vkládat data do počítače přímo na ploše výzkumu.

Nutná kvalita geometrických dat kladla zvýšené nároky na přesnost zakreslování nálezové situace do měřických náčrtů. Pro srovnání byla při sběru dat použita i fotogrammetrie.

10.5.2.4 Projekt POH 86-88

Jiného rázu je projekt **POH 86-88**. Jeho náplní jsou data, získaná z dokumentace výzkumu řemeslnického areálu v bývalé lesní školce. Zpracováním výsledků výzkumů z let 1986 až 1988 byl vytvořen komplexní postup převodu dat archivovaných v analogové formě do prostředí LIS.

Cílem tohoto pilotního projektu je ověřit vhodnost použití technologie LIS pro zpracování terénní dokumentace z let 1961-1990. Zastavme se nyní u návrhu struktury tohoto projektu.

10.5.2.4.1 Požadavky na projekt:

- Archivace dat v digitální formě (Převod grafické a písemné dokumentace do digitální formy bez ztráty informační hodnoty. Další podmínkou je, aby data nebyla při digitalizaci interpretována.)
- Možnost analýz a prezentace (Maximální nahrazení klasického způsobu vědeckého zpracování dokumentace a materiálu zpracováním s využitím výpočetní a reprodukční techniky. Systém má za úkol postihnout celý proces zpracování archeologických dat od jejich sběru až po analýzy, prezentaci a publikování.)

10.5.2.5 Struktura projektu POH 86-88

U CAD systémů platí zkušenost, že je snadnější spojit jednotlivé prvky ve větší celky, než dělit větší celek na jednotlivé prvky. Vzhledem k této skutečnosti byla struktura projektu navržena co nejpodrobnější (viz. tabulky 1,2,3).

10.5.2.5.1 Členění projektu

Ve struktuře projektu je možné vysledovat trojí členění
– vertikální, horizontální a logické.

Vertikální členění projektu vychází z existující dokumentace výzkumů v lesní školce. Tato totiž zachycuje nálezovou situaci ve třech úrovních:

- vyrýsování na povrchu
- stav po částečném vykopání objektů – (tedy včetně nálezů a kontrolních profilů a bloků)
- stav po kompletním vykopání plochy

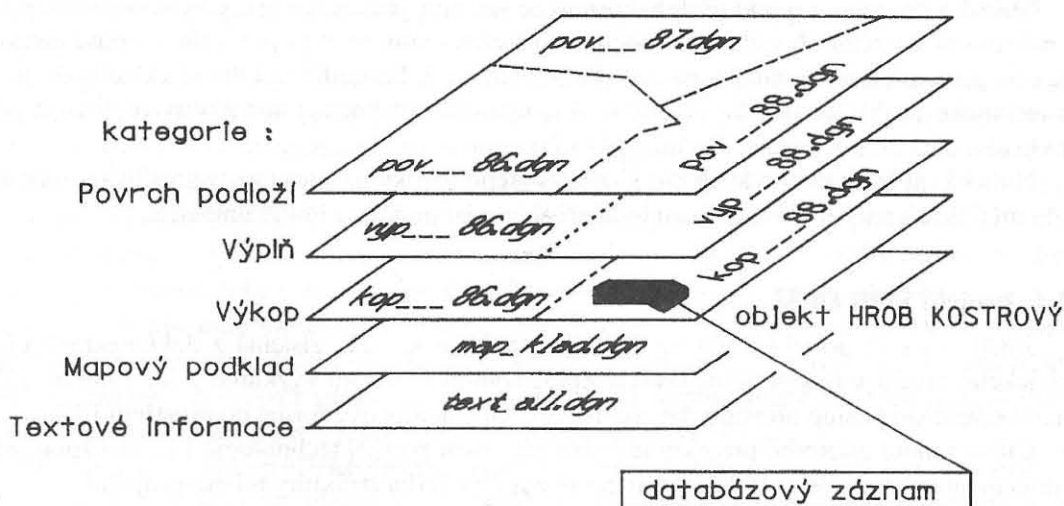
Z těchto tří základních úrovní vychází návrh kategorií projektu. Jsou to kategorie **Povrch podloží, Výplň a Výkop**.

Další dvě kategorie **Mapový podklad** a **Textové informace** **nečlení projekt vertikálně, nýbrž logicky**.

Horizontální členění projektu představuje rozdělení jednotlivostí (objektů, hrobů, materiálů) do výkresů (*maps*). Mapa představuje vždy část úrovně, vykopané během určité sezony, je tedy podmnožinou kategorie.

Grafickou část pilotního projektu, jenž zpracovává sezony 1986, 1987 a 1988, si je tedy možné představit jako skupinu fólií, z kterých skládáme jednotlivé obrázky. Viz obr.5.

Do jednotlivých výkresů (*map*) jsou kresleny jednotlivé objekty (*features*). Každý objekt má grafické atributy (barvu, umístění do vrstvy, druh entity, aj.), případně nese i záznam z tabulky atributů.



Obr. 5. Vztah kategorie, mapy, objektu a databáze (výsek z pilotního projektu POH 86-88).

10.5.2.5.2 Tabulky atributů

Návrh databázových tabulek (tabulek atributů) je dalším velmi důležitým krokem. Do projektu POH 86-88 vstoupily tabulky **objekty** a **hrob_kostrový**. Tabulky co nejdříveji postihují obsah písemné dokumentace výzkumu (deníky výzkumu a nálezové zprávy). Návrh databází přináší tab. 4 a 5.

10.5.2.5.3 Názvy kategorií a objektů

Pojmenování kategorií (množin logicky souvisejících objektů) názorně podává tab. 1, kde se nacházejí také názvy příslušných geografických indexů. Souvislost kategorií s mapami vystihuje obr.5.

Názvy objektů obsahuje tab.2 Každý objekt v tabulce je kromě svého jména, jednoznačného identifikátoru (*mblink*), názvu kategorie, do níž náleží a grafických atributů, charakterizován připojenou databázovou tabulkou atributů a kódem.

Název objektu se skládá z vlastního jména objektu a zkratky názvu kategorie, např. hrob kostrový (KOP).

Kód objektu je šestimístné číslo. První tři cifry kódu informují o kategorii, do které objekt náleží:

- 001 ... Výkop
- 002 ... Výplň
- 003 ... Povrch podloží
- 004 ... Textové informace
- 005 ... Mapový podklad

Jak již bylo zmíněno první tři uvedené kategorie vymezují fyzickou polohu objektu ve vertikálním směru, zbylé dvě určují polohu logickou (opět ve vertikálním směru dle zásady postupu od obecného ke konkrétnímu, od celku k detailu).

Čtvrtá číslice v kódu označuje skupinu objektů.

- 1 ... Objekty (základní tvar)
- 2 ... Materiály (keramika, kámen, kosti...)
- 3 ... Vnitřní tvar objektů
- 4 ... Profil v půdorysu a jeho náležitosti
- 5 ... Texty (čísla objektů, poznámky...)
- 6 ... Prvky mapového podkladu

Poslední dvě cifry informují o umístění objektu do vrstvy výkresu programu MicroStation. Například kód 002239 označuje objekt kategorie Výplň patřící do skupiny Materiály. Objekt je umístěn ve vrstvě 39 (viz tab. 1 a 2).

10.5.2.5.4 Pojmenování map

Mapami jsou v systému MGE-PC2 míněny výkresy * .dgn. Digitalizované výkresy se dělí do dvou základních skupin:

- a) **výkresy představující jednu rozlišenou a definovanou vrstvu výkopu** – tento typ výkresů vzniká digitalizací polní dokumentace (měřických náčrtů), případným geodetickým zaměřením terénní situace. Patří sem i výkresy s čísly jednotlivých objektů (komplexů, kontextů).
- b) **výkresy mapových podkladů a textových informací** – jedná se o vektorizované mapové podklady, výkresy s výsledky analýz, apod. Tato skupina výkresů nevzniká přímou digitalizací archeologické dokumentace.

skupina a)

Název výkresu vypadá následovně:

A A A _ _ X X .dgn

A A Azkratka souvisí se jménem kategorie (obr.5)

_ _ _volné pozice pro poznámku (mezivrstvy, nestandardní situace)

X Xrok výkopu (zkráceně 1988⇒ 88)

Význam zkratk je tento:

pov = (**povrch**) – mapa zachycuje stav na povrchu podloží, je patrné vyrýsování objektů (různá kvalita zeminy)

vyp = (**výplň**) – mapa zachycuje stav po odebrání povrchové vrstvy, v objektech jsou ponechány profily, kontrolní bloky a nálezová situace uvnitř objektů (např. ohniště, kamenné destrukce, koncentrace mazanice, aj.)

kop = (**výkop**) – mapa zachycuje konečný stav po vybrání objektů na podloží (rostlý terén)

txt = (**výkres s textovými informacemi**) – mapa zachycuje čísla objektů(komplexů, kontextů), poznámky

skupina B)

Názvy výkresů jsou tvořeny individuálně dle druhu výkresu (výkresy obsahující výsledky analýz, uložené obalové zóny, statistické informace, grafy, texty, schémata, aj.).

Výkresy, obsahující vektorizované mapové podklady mají ve svém názvu na prvních třech pozicích zkratku : **map**, například:

map_klad.dgn ... výkres zachycuje klad mapových listů Katastrální mapy (ZMVM) měřítko 1:1000 a mapy měřítko 1:10 000

map_old.dgn ... výkres zachycuje vektorizovanou část mapového listu účelové mapy lokality měřítko 1:1000.

MSLINK	CNAME (název kategorie)	INDEXNAME (název indexu)
1	Mapový podklad	mp
2	Povrch podloží	pov
3	Výpín	vyp
4	Výkop	kop
5	Textové informace	ti

Tab. 1. Tabulka kategorií MGE-PC2 POH 86-88 (category).

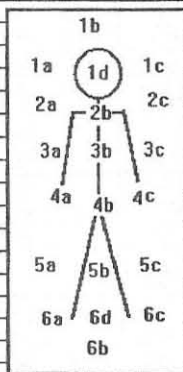
MSLINK	FOODE (kód)	FNAME (název objektu)	CATEGORY (kategorie)	TYPE (typ)	LEVEL (sazka)	COLOR (barva)
2	003126	hranice vrstvy (POV)	2	2	26	11
3	003236	kámen (POV)	2	3	36	12
4	003243	kost (POV)	2	3	43	4
5	002121	sídlíštní objekt (VYP)	3	3	21	7
7	002122	hrob kostrový (VYP)	3	3	22	92
8	002123	kulová jamka (VYP)	3	3	23	21
9	002135	recentní výkop (VYP)	3	0	35	19
10	002124	žlab (VYP)	3	3	24	9
11	002328	vnitřní tvar objektu (VYP)	3	0	28	135
12	002456	profil, popis (VYP)	3	0	56	0
14	002243	kost (VYP)	3	3	43	4
15	002239	mazanice (VYP)	3	3	39	6
16	002238	keramika (VYP)	3	3	38	3
17	002240	uhlík (VYP)	3	3	40	0
18	002237	kámen opracovaný (VYP)	3	3	37	2
19	002241	struska (VYP)	3	3	41	15
20	002242	železný nástroj (VYP)	3	3	42	14
21	003130	zásah v podloží (POV)	2	0	30	0
22	002236	kámen (VYP)	3	3	36	12
23	001121	sídlíštní objekt (KOP)	4	3	21	7
24	001122	hrob kostrový (KOP)	4	3	22	92
25	001123	kulová jamka (KOP)	4	3	23	21
26	001135	recentní výkop (KOP)	4	0	35	19
27	001124	žlab (KOP)	4	3	24	9
28	001328	vnitřní tvar objektu (KOP)	4	0	28	135
29	001243	kost (KOP)	4	3	43	4
30	001239	mazanice (KOP)	4	3	39	6
31	00138	keramika (KOP)	4	3	38	3
32	001240	uhlík (KOP)	4	3	40	0
34	001241	struska (KOP)	4	3	41	15
35	001242	železný nástroj (KOP)	4	3	42	14
38	003239	mazanice (POV)	2	3	39	6
39	003238	keramika (POV)	2	3	38	3
40	002126	hranice vrstvy (VYP)	3	2	26	11
41	001126	hranice vrstvy (KOP)	4	2	26	11
42	001236	kámen (KOP)	4	3	36	12
43	005608	vrstevnice 0.5 m (MAP)	1	0	8	86
44	005607	vrstevnice 1.0 m (MAP)	1	0	7	84
45	005603	budova bez rozlišení (MAP)	1	3	3	9
46	005605	vodní plochy (MAP)	1	0	5	23
47	005604	cesta (MAP)	1	0	4	0
48	005610	okraj mapy (MAP)	1	0	10	35
49	005613	čtvercová síť (MAP)	1	0	13	106
50	005606	polohopis bez rozlišení (MAP)	1	0	6	0
51	005601	hektarová síť (MAP)	1	0	1	35
52	005615	čísla sloupců a řad (MAP)	1	0	15	12
53	005619	mapový list 1 :10 000 (MAP)	1	0	19	0
54	005620	mapový list 1 :1 000 (MAP)	1	0	20	0
55	005614	hranice výkopu (MAP)	1	0	14	1
56	004558	číslo kulové jamky (TXT)	5	0	58	0
57	004557	číslo objektu (TXT)	5	0	57	0

Tab. 2. Tabulka objektů MGE-PC2 POH 86-88 (feature).

MSLINK	MAPNAME (název mapy)	CATEGORY (kategorie)
1	pov 86.dgn	2
2	vyp 88.dgn	3
3	vyp 86.dgn	3
4	pov 87.dgn	2
5	pov 88.dgn	2
6	vyp 87.dgn	3
7	kop 88.dgn	4
8	kop 86.dgn	4
9	kop 87.dgn	4
10	mapa old.dgn	3
11	map klad.dgn	1
12	txt_all.dgn	5

Tab. 3. Tabulka map MGE-PC2 POH 86-88 (maps).

Sloupec	Typ	Délka	Vyvěštitelka
HROB	Číslo	8	číslo objektu
HROB ID	Text	254	pomocný identifikátor objektu (a,b,....)
ROK	Text	254	rok, v němž proběhl výzkum
CTVEREC	Text	254	číslo čtverce
TVAR	Text	254	tvar objektu
STENY 1	Text	254	šířkost stěn příčných
STENY 2	Text	254	šířkost stěn podélných
PROFIL 1	Text	254	profil příčný
PROFIL 2	Text	254	profil podélný
DNO 1	Text	254	příčný tvar dna
DNO 2	Text	254	podélný tvar dna
ORIENTACE	Text	254	orientace vzhledem k světovým stranám
DEL MAX	Číslo	8	maximální délka objektu
SIR MAX	Číslo	8	maximální šířka objektu
HL STUPNE	Číslo	8	hloubka stupně
HL MAX	Číslo	8	maximální hloubka objektu
STRAT NAD	Text	254	č. kontextu stratigraficky pod popisovaným kontextem
STRAT POD	Text	254	č. kontextu stratigraficky nad popisovaným kontextem
JEDINCI	Číslo	8	počet jedinců v hrobě
POLOHA	Text	254	poloha těla
L RUKA	Text	254	poloha levé ruky
P RUKA	Text	254	poloha pravé ruky
L NOHA	Text	254	poloha levé nohy
P NOHA	Text	254	poloha pravé nohy
P RUK DIS	Ano/Ne	1	dislokace pravé ruky
L RUK DIS	Ano/Ne	1	dislokace levé ruky
P NOH DIS	Ano/Ne	1	dislokace pravé nohy
L NOH DIS	Ano/Ne	1	dislokace levé nohy
LEBKA DIS	Ano/Ne	1	dislokace lebky
ZEBRA DIS	Ano/Ne	1	dislokace žeber
DEL KOSTRY	Číslo	8	délka kostry in situ
ZACHOVAL	Text	254	zachovalost kostry
RAKEV	Ano/Ne	1	přítomnost rakve
RAK DEL	Číslo	8	délka rakve
RAK SIR	Číslo	8	šířka rakve
RAK TVAR	Text	254	tvar rakve
NALEZY	Ano/Ne	1	přítomnost milodarů v hrobě
1A	Text	254	lokalizace milodarů
1B	Text	254	
1C	Text	254	
1D	Text	254	
2A	Text	254	
2B	Text	254	
2C	Text	254	
3A	Text	254	
3B	Text	254	
3C	Text	254	
4A	Text	254	
4B	Text	254	
4C	Text	254	
5A	Text	254	
5B	Text	254	
5C	Text	254	
6A	Text	254	
6B	Text	254	
6C	Text	254	



Tab. 4. Návrh tabulky hrob_kostroy (část první).

Sloupec	Typ	Délka	Vysvětlivka
6D	Text	254	
KER	Ano/Ne	1	přítomnost keramiky (nádoby)
KER JINE	Ano/Ne	1	přítomnost keramiky (jiné)
ZV KOST	Ano/Ne	1	přítomnost zvířecích kostí
ZV K KS	Číslo	8	počet zvířecích kostí
LID KOST	Ano/Ne	1	přítomnost lidských kostí
KOST ART	Ano/Ne	1	přítomnost kostěných artefaktů
KAMENY	Ano/Ne	1	přítomnost kamenů
KAM OBJ	Číslo	8	objem kamenů
KAM ART	Ano/Ne	1	přítomnost kamenných artefaktů
MAZ	Ano/Ne	1	přítomnost mazanice
MAZ KS	Číslo	8	počet kusů mazanice
MAZ HM	Číslo	8	hmotnost mazanice
ZELEZO	Ano/Ne	1	přítomnost železa (nespecifikováno)
ZEL ART	Ano/Ne	1	přítomnost železných artefaktů
STRUSKA	Ano/Ne	1	přítomnost strusky
STR KS	Číslo	8	počet kusů strusky
STR HM	Číslo	8	hmotnost strusky
UHLIKY	Ano/Ne	1	přítomnost uhlíků
DREVO	Ano/Ne	1	přítomnost dřeva
JINE	Ano/Ne	1	jiné
NAL POPIS	Text	254	popis nálezu
POC VRST	Číslo	8	počet vrstev
CHARAKTER	Text	254	popis vrstev
KUL JAM	Ano/Ne	1	přítomnost kúlových jamek
KUL FOC	Číslo	8	počet kúlových jamek
KUL MISTO	Text	254	umístění kúlových jamek
KUL POPIS	Text	254	popis kúlových jamek
KONSTRUKCE	Ano/Ne	1	přítomnost konstrukcí
KON POPIS	Text	254	popis konstrukcí
SOUVISI S	Text	254	souvisí s.....
POZNAMKA	Text	254	poznámka
ANALYZA1	Číslo	8	výsledky databázových analýz
ANALYZA2	Číslo	8	výsledky databázových analýz
ANALYZA3	Číslo	8	výsledky databázových analýz
MSLINK	Číslo	10	identifikátor pro MGE-PC2
MAPID	Číslo	10	identifikátor pro MGE-PC2

Tab. 4. Návrh tabulky hrob_kostrový (část druhá).

Sloupec	Typ	Délka	Vysvětlivka
OBJEKT	Číslo	8	číslo objektu
OBJEKT ID	Text	150	pomocný identifikátor objektu (a,b,.....)
ROK	Text	250	rok, v němž proběhl výzkum
CTVEREC	Text	250	číslo čtverce
TVAR	Text	250	tvár objektu
STENY 1	Text	250	šikmost stěn příčných
STENY 2	Text	250	šikmost stěn podélných
PROFIL 1	Text	250	profil příčný
PROFIL 2	Text	250	profil podélný
DNO 1	Text	250	příčný tvar dna
DNO 2	Text	250	podélný tvar dna
ORIENTACE	Text	250	orientace vzhledem k světovým stranám
DEL MAX	Číslo	8	maximální délka objektu
SIR MAX	Číslo	8	maximální šířka objektu
HL STUPNE	Číslo	8	hloubka stupně
HL MAX	Číslo	8	maximální hloubka objektu
STRAT NAD	Text	250	č. kontextu stratigraficky pod popisovaným kontextem
STRAT POD	Text	250	č. kontextu stratigraficky nad popisovaným kontextem
KER	Text	1	přítomnost keramiky (nádoby)
KER JINE	Text	1	přítomnost keramiky (jiné)
ZV KOST	Text	1	přítomnost zvířecích kostí
ZV K KS	Číslo	8	počet zvířecích kostí
LID KOST	Text	1	přítomnost lidských kostí
KOST ART	Text	1	přítomnost kostěných artefaktů
KAMENY	Text	1	přítomnost kamenů
KAM OBJ	Číslo	8	objem kamenů
KAM ART	Text	1	přítomnost kamenných artefaktů
MAZ	Text	1	přítomnost mazanice
MAZ KS	Číslo	8	počet kusů mazanice
MAZ HM	Číslo	8	hmotnost mazanice
ZELEZO	Text	1	přítomnost železa (nespecifikováno)

Tab. 5. Návrh tabulky objekty (část první).

ZEL ART	Text	1	přítomnost železných artefaktů
STRUSKA	Text	1	přítomnost strusky
STR KS	Číslo	8	počet kusů strusky
STR HM	Číslo	8	hmotnost strusky
UHLIKY	Text	1	přítomnost uhlíků
DREVO	Text	1	přítomnost dřeva
JINE	Text	1	jiné
NAL POPIS	Text	254	popis nálezu
POC VRST	Číslo	8	počet vrstev
CHARAKTER	Text	254	popis vrstev
KUL JAM	Text	1	přítomnost kúlových jamek
KUL POC	Číslo	8	počet kúlových jamek
KUL MISTO	Text	250	umístění kúlových jamek
KUL POPIS	Text	254	popis kúlových jamek
PEC	Text	1	přítomnost otopných zařízení
PEC MISTO	Text	250	popis a umístění otopných zařízení
KONSTRUKCE	Text	1	přítomnost konstrukcí
KON POPIS	Text	250	popis konstrukcí
INTERPRETA	Text	250	předběžná interpretace
SOUVISI S	Text	250	souvisí s.....
POZNAMKA	Text	254	poznámka
MSLINK	Číslo	10	identifikátor pro MGE-PC2
MAPID	Číslo	10	identifikátor pro MGE-PC2

Tab. 5. Návrh tabulky objekty (část druhá).

10.5.3 Sběr dat

10.5.3.1 Vstup dat do systému

10.5.3.1.1 Negrafická data

Negrafická data vstupují do systému prostřednictvím klávesnice. Obsah písemné dokumentace je vkládám do databáze Access. Vkládání je pro pracovníky uživatelsky příjemné, neboť vstup do databáze je řešen pomocí kvalitně vytvořeného formuláře. Ten zajišťuje databázi také jednoznačnost (vkládání dat pomocí combo boxů, zaškrtování alternativ, apod.)

Takto vytvořené databáze se formou dotazu exportují do formátu Xbase. Vzniklé tabulky objekty a hrob_kostrovy vstupují do MGE-PC2.

10.5.3.1.2 Grafická data

Grafické podklady se do počítače přenášejí pomocí digitizéru (tabletu) Kurta XGT. Digitalizace není objektová, data se ukládají v surovém stavu do výkresů formátu *.dgn. Tyto výkresy potom vstupují do MGE-PC2, kde se z elementů CADu MicroStation stávají objekty (*features*).

Postup digitalizace v prostředí MicroStation

Grafická předloha (plán) se vsune pod upevňovací fólii tabletu. Pod právě digitalizovaný výkres v počítači se „podsvítí“ jako referenční výkres (*reference file*) čtvercová síť, pomocí které se na obrazovce počítače lokalizuje digitalizovaná oblast -čtverec. Rohy tohoto čtverce slouží jako vřícovací body pro transformaci plánu do počítače. Jednotlivé rohy čtverce v plánu se přiřadí k rohům identického čtverce na obrazovce. K transformaci grafické předlohy do počítače stačí tři jednoznačně identifikovatelné body, neležící v přímce. Při větším počtu bodu provádí program MicroStation vyrovnání. Poté se začnou digitalizovat jednotlivosti plánu. Detailní popis digitalizace přináší kapitola PRACOVNÍ PROSTŘEDÍ POHAN.

10.5.3.2 Pracovní prostředí POHAN

Program MicroStation umožňuje uživateli rozsáhlé úpravy, týkající se jak vzhledu programu, tak i jeho funkcí.

- 1) **jednoznačnost** – v prostředí lze kreslit pouze pomocí vytvořených ikon. Jednotlivé ikony v sobě skrývají nejen příkaz pro vykreslení určitého typu elementu, ale i definici jeho atributů (tloušťka a typ čáry, barva, umístění do vrstvy, aj.). Nastavení odpovídá definici

geometrických atributů objektů v MGE-PC2, při převodu elementů na objekty není nutno tyto atributy měnit. Všechny výkresy mají shodný zdrojový výkres (*seed file*).

- 2) **zjednodušení editace a plotování** – díky jednoznačné definici atributů elementů je velmi snadné provádět jejich výběr, úpravy, lze si vytvořit knihovnu souborů plotru pro jednotlivé způsoby prezentace (plotr, tiskárna)
- 3) **počestění a podstatné zjednodušení programu MicroStation** – v současné době nevlastní Katedra archeologie FF MU českou licenci programu MicroStation. Částečné počestění anglické verze společně s podstatným okleštěním funkcí programu přispívá ke zjednodušení práce v systému. Program na uživatele působí přehledněji, což se projevuje na úspoře času při zaškolování pracovníků pro digitalizaci.

Součástí prostředí je i nastavení názvů vrstev, nastavení uživatelských typů čar a vlastní knihovna značek (viz apendix POHAN. Stručný popis uživatelského prostředí).

10.5.4 Správa dat

10.5.4.1 Uchovávání dat v počítači

Data se na pevném disku počítače uchovávají v surovém stavu (výkresy, databáze) jako digitální forma dokumentace výzkumu (v nejbližší době se začnou vypalovat na CD disk). Vybraná data potom vstupují do LIS projektu.

10.5.5 Analýzy nad daty

Provádění analýz je věcí uživatele nikoli tvůrce systému. Přesto jsme již při vytváření projektů některých z možností dotazování (srov. kap. 3.) využili:

- **dotaz pouze na tabulku atributů** (slouží k prohledávání negrafické databáze (listování) bez lokalizace objektů v mapě. Systém tak slouží jako jednoduchý databázový program. Této formy dotazování se využívá především k vyhledání databázového záznamu pro připojení ke grafickému elementu)
- **dotaz na mapu i tabulku atributů** (představuje klasické vyhledání objektu v mapě a jeho zvýraznění podle potřeb uživatele. Na obrazovce lze připojenými záznamy listovat s možností lokalizace aktuálního objektu)
např. *Vyber sídlištní objekty delší než 6.5 m a užší než 2.1 m (objekty specifického tvaru – žlabovité), obsahující železo a strusku a vybarvi je červeně.*
- **dotaz s využitím obalové zóny (buffer)** (velmi účinný nástroj pro analýzu, kombinuje předchozí možnosti s prostorovým vymezením)
např. *Vyber sídlištní objekty delší než 6.5 m a užší než 2.1 m, obsahující železo a strusku vzdálené méně než 2m od studní (Takto můžeme sledovat prostorovou závislost těchto specifických objektů na zdrojích vody, což může napomoci k jejich ev. interpretaci).*

Výsledek dotazu je možné uložit jako zprávu (*report*) a dále zpracovat např. v programu Excel. Obalové zóny lze ukládat nejen do souboru, ale také přímo do výkresu a vytvářet z nich nové objekty (nesoucí například statistické informace). Výsledek analýzy může být následně využit také jako jedna z podmínek dalšího dotazu.

Díky velmi podrobnému členění databázových tabulek projektu a možnostem MGE-PC2 strukturovat dotazy pomocí logických operátorů a prostorového vymezení je možné provádět i velmi komplikované analýzy.

10.5.6 Prezentace

10.5.6.1 Způsoby prezentace

Výsledky analýz i vlastní projekt je možné prezentovat buď interaktivně v počítači (prohlížečka VistaMap, použití programu pro převod projektu a následná prezentace v síti Internet), nebo v rámci grafických a textových výstupů, tabulek, fólií, aj.

Pro publikační účely je vhodné využívat funkci uložení vzhladu obrazovky do formy bitmapy – Capture screen a následný import takto vzniklého obrázku do textového nebo tabulkového procesoru (Word, Excel, aj.)

Základním způsobem výstupu je v případě CAD systémů (vektorové grafiky) vykreslení požadovaného výkresu (mapy) nebo jeho části na kreslicím zařízení (plotru). Následující kapitola představuje návrh nomenklatury výkresů projektu POH 86-88.

10.5.6.2 Názvosloví výkresů

Názvosloví a číslování grafických výstupů je navrženo podobně jako nomenklatura prvků projektu. Číslo (kód) výkresu se skládá ze dvou částí:

A / B

Část **A** představuje číslo (označení) výkresu.

Část **B** je zkrácený kód kategorie.

Grafické výstupy ve formě výkresů dělíme do čtyř skupin:

1) **mapy a přehledky** – vyznačují se měřítkem 1:100 a menším. **Mapy** informují o umístění výzkumu vzhledem k okolí. Jejich kód začíná písmenem **M**.

Přehledky lokalizují jednotlivé objekty v rámci plochy výzkumu. Kód začíná písmenem **P**.

2) **nákresy jednotlivých objektů** – měřítko zpravidla 1:20 a větší. Část kódu před lomítkem je číslo objektu (komplexu).

3) **speciální výkresy** – zachycují vybranou část plochy výzkumu. Měřítko je větší než 1:100 a menší než 1:20. Kód začíná písmenem **S**.

4) **ostatní** – nestandardní obrázky (schémata pro tisk, aj.)

Pro grafické výstupy je vhodný pauzovací papír, který umožňuje skládání a prohlížení úrovní nad sebou. Výstup na papír ulehčuje a jednotný ráz výkresů zajišťuje použití knihovny buněk s buňkami legendy, popisové tabulky, vyplnění tabulky, směrové růžice, měřítka a rámu výkresů formátu A4 až A1. Součástí projektu jsou také předem definované soubory ovládání plotru Mutoh 500, obsahující přiřazení barev k jednotlivým perům plotru.

10.6 Literatura

Babický, T. 1994: K definici GIS, Geodetický a kartografický obzor, 40(82), č.9, s.196.

Chmelík, M.1996: GIS help. (v tisku)

Kaláb, P. 1993: Zkušenosti s budováním grafických informačních systémů, CAD, ročník 3, , č.4, s.76-77.

Klimesš, L. 1986: Slovník cizích slov. SPN, Praha.

Konečný, M. – Rais, K. 1985: Úvod do geografických informačních systémů. UJEP, Brno.

Kučera, M. 1996: Aplikace digitálních metod sběru a analýzy dat v archeologii za použití prostředků GIS pracujících na platformě systému MicroStation. Diplomová práce, FAST VUT Brno.

Orlík, T. 1993: (Geo)grafické informační systémy – zelená větev Computer Aided Designu, CAD, ročník 3, č.4, s.58-62.

Vojtička, O. 1995: Využití údajov katastra nehnutelnosti na tvorbu GIS/LIS, Geodetický a kartografický obzor, 41(83), č.4, s.77-80.

Firemní literatura, sborníky přednášek:

ARCDATA s.r.o. 1993: Seznamte se s GIS, systémem ARC/info. (překlad knihy ESRI Inc.:Understanding GIS.), Praha

10.7 Uživatelské prostředí POHAN

Charakteristika

Pracovní prostředí **POHAN** je nástroj sloužící ke sběru archeologických dat. Zvládnutí práce v prostředí **POHAN** předpokládá krátké zaškolení v programu MicroStation. Uživatel by měl dokázat pracovat s myší a kurzorem tabletu. Je nutné znát základní ovládání oken (podobné WINDOWS).

Instalace

Pracovní prostředí **POHAN** se instaluje z diskety. Její obsah je následující:

install.bat instalační batch soubor
SYSTEM adresář obsahující komponenty uživatelského prostředí **POHAN**

Postup instalace:

Disketu vložte do mechaniky počítače, přepněte se na ni a spusťte program install.bat. Během instalace se řiďte instrukcemi na obrazovce. Instalace komponentů prostředí **POHAN** do příslušných adresářů programu MicroStation proběhne automaticky.. Adresáře budou potom vypadat následovně:

V adresáři ...*adresář MicroStationu*\WSMOD přibude adresář **POHAN** obsahující podadresáře DATA, SEED, SYMB, CELL.

adresář DATA zahrnuje soubory:

arch_poh.stg skupina nastavení (*group settings*) – předvolby pro kreslení archeologických objektů
pohan.lvl soubor s názvy vrstev
pohan.mnu soubor s definicí funkčních kláves

adresář SEED zahrnuje soubory:

seed_poh.dgn zdrojový výkres (*seed file*) pro zakreslení půdorysu
pro_seed.dgn zdrojový výkres (*seed file*) pro zakreslení profilu

adresář SYMB zahrnuje soubory:

fonty.rsc soubor s vybranými českými a anglickými vektorovými fonty
cara_poh.rsc soubor s nastavenými typy čar

adresář CELL zahrnuje soubory:

cell_poh.cel knihovna značek využívaných v prostředí **POHAN**

V adresáři ...*adresář MicroStationu* \WSUI přibude adresář **POHAN** obsahující 19 souborů ustn.m01 až ustn.m19.

V adresáři ...*adresář MicroStationu* \CONFIG\USER přibudou soubory:

pohan.ucf soubor s nastavením cest prostředí
pohan.upf soubor s nastavením preferencí
pohan.pcf soubor s nastavením projektu

Spouštění programu s prostředím POHAN

Po spuštění programu MicroStation vyberte v menu **USER** v položce **SELECT DEFAULT WORKSPACE** prostředí **POHAN** a potvrďte volbu. Změna se projeví při dalším startu MicroStation.

Pozn.: MicroStation s POHANem lze spustit také přímo z DOSu z adresáře USTATION příkazem USTATION -wupohan. Je vhodné si pro spouštění programu vytvořit jednoduchý dávkový (batch) soubor.

Proveďte nastavení systémových proměnných [viz. dokumentace MicroStation User's Guide. Bentley Systems, Inc. 1994, s. 14-8 až 14-11.].

Práce v prostředí POHAN

Po spuštění programu se objeví anglický **MICROSTATION MANAGER**. Zde se vybere a otevře výkres.

Po otevření výkresu se na obrazovce uvidíte základní okno MicroStation.



Obsah jednotlivých roletových nabídek

*Pozn.: Jednotlivé funkce se vyvolávají pomocí horkých kláves **CTRL+písmeno** přímo ze základního **MICROSTATION MANAGERU**, případně stisknutím příslušného písmena, pohybujete-li se již v roletovém men. Základní nabídky (**Soubor, Palety...**) vyvoláte pomocí akcelerátoru **ALT+podtržené písmeno**.*

Soubor

- Založ nový výkres <CTRL> <N>
- Otevři výkres <CTRL> <O>
- Zavři výkres <CTRL> <Z>
- Komprese výkresu <CTRL> <K>
- Ulož nastavení <CTRL> <U>
- Konec

Editace

- Undo
- Redo

Nastavení

- Aktivní úhel
- Formát souřadnic
- Vlastnosti vrstev <CTRL> <S>
- Přesný vstup

Pohled

- Zapnout
- Kaskáda
- Vlastnosti pohledu <CTRL> <P>
- Vrstvy pohledu <CTRL> <V>

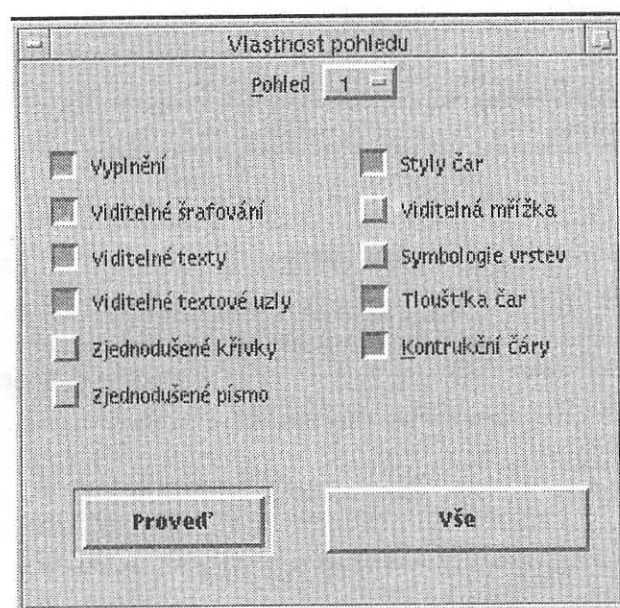
Palety

- Základní palety
- Hlavní paleta

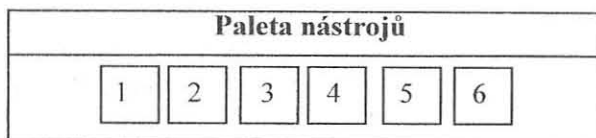
- Ovládání pohledu
- Modifikace elementu
- Výběrová ohrada
- Měření
- Nastavení tabletu
- Digitalizace
 - Objekty – základní obrys
 - Materiály
 - Vnitřní kresba objektu
 - Pomocná kresba
 - Texty – čísla
- Zavři subpalety
- Digitalizace
 - Nastavení tabletu
 - Předvolby
- Zachycení
 - Tlačítka
- User
 - Prostředí
 - Vyber komponenty uživatelského prostředí
 - Vyber uživatelské prostředí
- Preference
- Help

Roletové nabídky **Soubor**, **Editace**, **Nastavení**, **Pohled**, **Zachycení**, **User** a **Help** skrývají do češtiny přeložené a zjednodušené nabídky standardního prostředí programu MicroStation. Jejich ovládání lze zvládnout intuitivně, případně odkazují na literaturu [MicroStation – User's Guide. Bentley Systems, Inc. 1994.]

vzhled upraveného dialogového okna
(původně *View Attributes*)



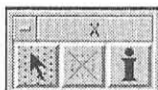
Zcela přebudované nebo nové jsou nabídky **Palety** a **Digitalizace**. Při jejich popisu se používá číslování jednotlivých ikon podle následujícího schématu:



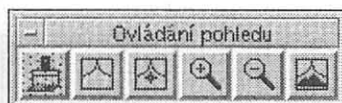
Popis nabídky Palety

Nabídka Palety obsahuje :

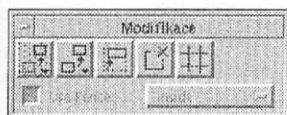
Základní palety



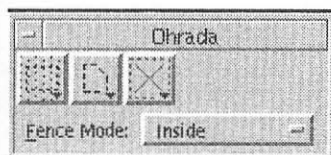
Hlavní paleta



Ovládání pohledu



Modifikace elementu



Výběrová ohrada

Měření - paleta shodná s původní paletou programu MicroStation

Všechny výše zmíněné palety vznikly počestěním a úpravou standardních palet MicroStation. Nová je pouze ikona 3 na Hlavní paletě, která je grafickou obdobou příkazu **analyze**. Slouží k získání informací o elementu ve výkresu.

Nastavení tabletu



Tablet

- 1.....zapnutí přichytávání na mřížku
- 2.....vypnutí přichytávání na mřížku
- 3.....lokalizace vektorizované oblasti výkresu. Pomocí minimálně tří bodů neležících na přímce se grafická předloha umístí (transformuje) do výkresu v počítači.

Postup:

- 1) Kurzorem tabletu vyberte a potvrďte bod v grafické předloze.
- 2) Na obrazovce lokalizujte identický bod, potvrďte volbu.
- 3) Opakujte postup 1) a 2) pro zvolený počet bodů.

Digitalizace

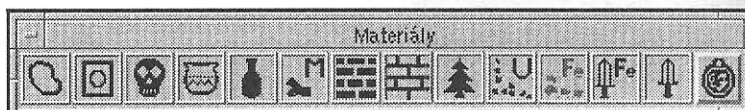
- 1.....sídlíšní objekt
- 2.....hrob kostrový
- 3.....kúlová jamka
- 4.....žlab
- 5.....recentní výkop
- 6.....hranice vrstev
- 7.....zásah v podloží

Objekty – základní obrys



Materiály

- 1.....kámen
- 2.....kámen opracovaný
- 3.....kost
- 4.....keramika
- 5.....sklo
- 6.....mazanice
- 7.....cihly, krytina
- 8.....malty
- 9.....organické látky
- 10...uhlík
- 11...struska
- 12...železný nástroj
- 13...barevné kovy
- 14...vzácné kovy



Vnitřní kresba objektu

- 1.....vnitřní tvar objektu
- 2.....neviditelné hrany
- 3.....spádovky (kolmé čáry)



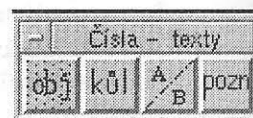
Pomocná kresba

- 1.....neidentifikovatelné objekty (pomocné čáry)
- 2.....profil
- 3.....pomocný bod



Číslo - texty

- 1.....číslo objektu
- 2.....číslo kúlové jamky
- 3.....popis profilu
- 4.....poznámka



Popis nabídky Digitalizace

Nastavení tabletu

Ovládání nabídky viz [MicroStation - User's Guide. Bentley Systems, Inc. 1994.]

Předvolby

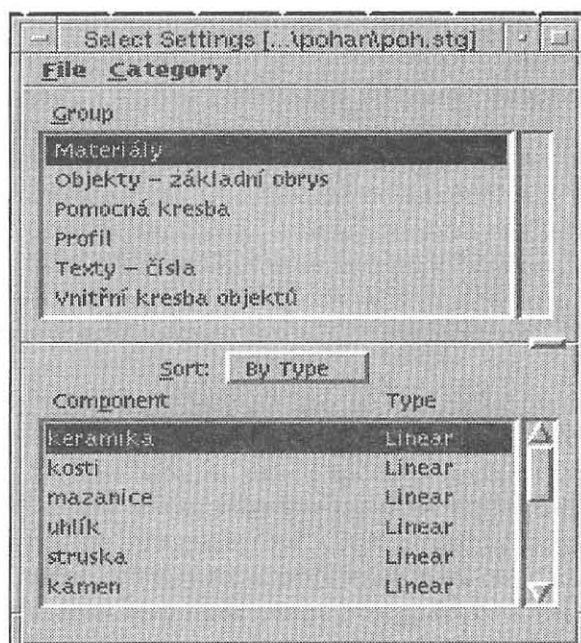
Nabídka zahrnuje soubor Skupin nastavení (*Settings group*), který je textovou analogií výše popsaných ikon.

poznámka: *Obsah ikon Číslo - texty a Pomocná kresba je zde sloučen ve skupinu Pomocná kresba, texty.*

Nabídka **Předvolby** navíc obsahuje skupinu **Profil**, sloužící k zakreslování příčných řezů (profilů) objektu. Ve skupině se nacházejí následující komponenty pro kreslení: šrafování, obrys kontextu, pomocný obrys, popis profilu, číslo kontextu.

Práce s Předvolbami je totožná s prací s ikonami. Na uživateli záleží pro kterou alternativu se rozhodne. Prostor umožňuje používat obě zároveň.

Předvolby - ukázka nabídek



Definované funkční klávesy (je možné přizpůsobit uživateli)

- < F1 >zapnutí chytání na mřížku
- < F2 >informace o elementu
- < F3 >zoom + (zvětšení)
- < F4 >zoom - (zmenšení)
- < F5 >zoom okno
- < F6 >zapnutí všech vrstev
- < F7 >vypnutí všech vrstev (mimo aktivní vrstvy)
- < F8 >uzavření elementu
- < F9 >připojení referenčního výkresu
- < F11 > ...záložní kopie (uložení výkresu do souboru * .bak)
- < CTRL > < F9 > ...odpojení referenčního výkresu
- < CTRL > < F10 > ...DOS Prompt („odskok“ do DOSu)
- < CTRL > < F12 > ... ukončení programu (s uložením nastavení a kompresí výkresu)

Další informace o pracovním prostředí POHAN a možnostech jeho získání najdete na internetové adrese http://www.phil.muni.cz/archoe/index_cz.html.

11. Geografický informační systém a výzkum pravěké sídelní struktury (Martin Kuna)

11.1 Úvod

Výpočetní technika, a to především databázové systémy a statistické programy, přinesla v posledních desetiletích přelom v efektivním využití archeologických dat. Software typu geografických informačních systémů (GIS) je dalším krokem v tomto směru. GIS jsou softwarovými produkty, zaměřenými na zpracování prostorových informací. Jejich přínosem je zejména možnost přímé vizualizace archeologických situací v reálném (geografickém) prostoru a možnost analýzy jejich uspořádání a vzájemných souvislostí s prvky geografického prostředí.

GIS se v archeologii uplatňují zejména ve dvou oblastech. Za prvé, GIS se stávají nástrojem uchování a zpřístupnění archeologických informací v jejich prostorovém kontextu. V této roli jsou GIS používány při dokumentaci terénních výzkumů a jako součást informačních systémů na úrovni lokality, regionu nebo celého státu (pojem GIS je zde, podobně jako v případě pojmu „databáze“, užíván jak k označení samotného softwaru, tak výsledného souboru informací, který je tímto prostředkem zpřístupněn uživateli). GIS jsou proto významným doplňkem databází, využívaných pro potřeby péče o archeologické dědictví.

Druhou oblastí využití GIS v archeologii je studium prostorové struktury archeologických pramenů a jejich geografického prostředí. Obě zmíněné oblasti použití, tj. dokumentace archeologických dat, motivovaná spíše praktickými ohledy, a na teoretických motivech založené studium, spolu úzce souvisejí, nicméně implikují dílčí rozdíly v prioritách při výběru dat, ve způsobech dokumentace a v neposlední řadě i při výběru konkrétního softwaru. V současné době již totiž existuje celá řada produktů typu GIS, lišících se podstatně nejen cenou, ale i složitostí obsluhy a využitelností pro různé úkoly archeologické práce.

Tzv. *vektorové* GIS (pojem vektorový zde souvisí s technikou počítačového záznamu prostorových objektů jakožto bodů-vektorů v prostoru a pomocí body vymezených linií a polygonů) obecně umožňují velmi přesnou a z hlediska kapacity paměti počítače úspornou evidenci dat. Proto jsou využívány např. ve státní správě, kde se stávají základem rozsáhlých informačních systémů, občas již zahrnujících i vrstvy archeologických informací (např. Okresní úřad Chrudim, připravují se další; srov. Neustupný 1996a). K nevýhodám těchto systémů patří jejich vyšší cena, uživatelská náročnost a v některých případech i absence analytických funkcí, které jsou potřebné z hlediska odborných cílů archeologie (např. GIS PC ARC/INFO). Naproti tomu tzv. *rastrové* GIS (uchovávající každou z mapových vrstev jako „rastr“ čtverců o určité velikosti) jsou orientovány spíše na odborné úkoly. Vzhledem k tomu jsou zpravidla vybaveny větším množstvím analytických funkcí, bývají levnější a uživatelsky přístupnější. Jejich jistou nevýhodou je komplikované propojení s doplňujícími databázovými informacemi, velké nároky na paměť počítače (omezuující v praxi rozsah studovaného území) a nižší kvalita reprodukce prostorové informace, včetně tištěného výstupu. Je však pravděpodobné, že rychlý vývoj softwaru povede brzy ke stírání rozdílů mezi těmito dvěma hlavními druhy GIS. Použitelnost různých typů GIS k odborným úkolům se bude rozšiřovat i s rozvojem technických parametrů dostupného hardwaru.

Užití GIS v archeologii má dosud v celosvětovém měřítku poměrně krátkou historii. Jeho počátky sahají do 80. let a souvisejí s rozvojem počítačového zpracování map, zejména ve Spojených státech. Za mezník je považován rok 1985, kdy byla této problematice věnována vlastní sekce na konferenci Společnosti pro americkou archeologii (Kvamme 1995). Sám pojem GIS se v archeologii objevuje teprve v druhé polovině 80. let s nástupem softwarových výrobků ARC/INFO, MOSS, GRASS nebo IDRISI, původně vyvinutých pro potřeby armády, státní správy, v případě IDRISI pro potřeby univerzitního výzkumu v oblasti přírodních věd (zejména pro analýzu satelitních snímků). Současné možnosti GIS jsou, kromě řady dílčích publikací, nejlépe ilustrovány učebnicovou prací V. Gaffneyho a Z. Stančiče *GIS approaches to regional analysis: A case study of the island of Hvar* (1991; rec.

Gojda 1993) a dvěma základními sborníky příspěvků specializovaných konferencí, shrnujícími přístupy k GIS v kontextu americké a evropské archeologie (Allen, Green a Zubrow eds. 1990; Lock a Stančič eds. 1995).

V České republice bylo využití GIS pro potřeby evidence archeologických dat a dokumentace výzkumů v současné době již zahájeno na více archeologických pracovištích (např. SÚPP Praha, ÚAPPSZČ Most, Archaia Praha aj.). Naproti tomu možnosti aplikace v analýze prostorové struktury archeologických dat jsou prozatím rozvíjeny jen několika výzkumnými projekty oddělení prostorové archeologie ARÚ Praha; první výsledky již byly publikovány (Kuna a Adelsbergerová 1995; Neustupný 1995a; Neustupný 1994, 1996b; Neustupný a Vencel 1995, Neustupný a Venclová 1996; Kuna 1996). Na zkušenosti z těchto projektů navazuje i tento příspěvek. Po technické stránce se tyto projekty opírají o vektorový GIS PC ARC/INFO, sloužící zejména k digitalizaci mapových vrstev, a rastrový GIS IDRISI v. 4.0 a v. 1.0 pro Windows (produkt Clark University, Worcester, Mass., USA), používaný k vlastní analýze archeologických a geografických dat.

11.2 Význam GIS pro analýzu archeologických a geografických dat

Nesporným přínosem archeologie posledních tří desetiletí je důraz na poznání strukturálních otázek vývoje společnosti, k nimž patří např. vztah společnosti k přírodnímu prostředí, sledování demografických trendů, organizace sídelních areálů atd. Řešení těchto otázek předpokládá uchopit archeologické prameny jako pozůstatky strukturovaného systému, který má formální a prostorovou dimenzi, včetně specifického vztahu ke konkrétnímu geografickému prostředí. V GIS lze spatřovat metodologický nástroj převratného významu, který umožňuje nejen novým způsobem řešit řadu existujících teoretických otázek, ale který nutně povede k formulování otázek nových. Jelikož veškerá archeologická data jsou informacemi prostorovými, GIS v budoucnu nezbytně ovlivní všechny roviny archeologické práce: od sběru dat v terénu, přes metodiku analýzy dat až po formulaci nových teoretických pojmů. Z hlediska odborných cílů archeologie je mimořádně významné, že GIS umožňují:

- (a) Operativní promítání formálních vlastností archeologických pramenů do reálného (geografického) prostoru a naopak, tj. převádění geografických (prostorových) vlastností do prostoru formálního. Příkladem může být např. promítnutí archeologických typů do topografické nebo speciální mapy a zpětné načtení určitých geografických údajů (vzdálenost od vodního zdroje, typ geologického podloží apod.) do databáze.
- (b) Analýzu prostorového uspořádání (např. plošný rozsah, hustota, vzájemná vzdálenost) celků vzniklých terénním výzkumem nebo formální analýzou a na tomto základě definici nových prostorových celků archeologických pramenů.
- (c) Zjišťování takových vlastností geografického prostoru, které jsou významné z hlediska chování minulých populací, avšak které nejde zjistit v běžně dostupných mapách (např. mapa sklonu a orientace svahu, areálů dostupnosti); vytváření libovolných kombinací mapových prvků a vrstev.
- (d) Přístup k prostorovým datům jako k souvislému povrchu plynule se měnících vlastností, tedy nikoliv jen jako k množině nesouvislých pozorování.
- (e) Vizualizaci prostorových dat způsoby, které neumožňuje běžná mapa (např. profily terénem, trojrozměrné modely) a její využití pro archeologickou analýzu a interpretaci.

11.3 Základní funkce GIS IDRISI

Základní princip rastrových GIS je velmi jednoduchý. Určitý prostor (zpravidla jde o reálný prostor definovaný geografickými souřadnicemi, ale může jít o jakýkoli prostor s libovolnými souřadnicemi) je rozdělen do sítě buněk (tzv. rastru) o zvolené velikosti. Každé buňce jsou přiřazeny tři numerické proměnné: souřadnice X a souřadnice Y (popisující její polohu) a třetí proměnná, která označuje stav (hodnotu) dané buňky, někdy též označovaná jako souřadnice Z. Souřadnice Z jednotlivých buněk rastru mohou vyjadřovat buď určitou měřitelnou hodnotu, příslušející dané buňce mapy (např.

nadmořskou výškou), nebo stav (vlastnost) dané buňky (např. mohou být pořadovým číslem archeologického naleziště, označením druhu geologického podloží apod.). V prvním případě jde zpravidla o reálná čísla, v druhém případě o čísla celá (označující tzv. nominální proměnnou), fungující jako identifikátory určité třídy objektů mapy.

Uspořádání buněk s určitou hodnotou Z vytváří v počítači digitální analogii tištěné dvourozměrné mapy, přičemž přesnost zobrazení zvolených objektů závisí především na velikosti buněk, čili „hrubosti“ užitého rastru. Velikost buněk lze definovat libovolně, je však přitom nutné brát ohled na kapacitu pevného disku a rychlost počítače. Je-li souřadnicí Z celé číslo (*integer*), zabírá každá buňka 2 byty paměti, u čísel reálných dvojnásobek. V některých aplikacích, např. při studiu geomorfologie terénu, je často nezbytné pracovat s rastrem poměrně malým (10x10m a menším); z toho vyplývá, že každá mapová vrstva regionu o rozloze 10x10 km se bude skládat z 1,000,000 buněk a zabírat zhruba 2 až 4 MBy paměti. Tento rozsah mapových vrstev představuje zřejmě rozumnou hranici velikosti z hlediska jejich zpracovatelnosti dnes dostupnou výpočetní technikou: běžné operace s vrstvami tohoto rozsahu na počítači typu PC 486 nebo Pentium trvají kolem 10 vteřin, složitější pak několik minut; některé výpočty však mohou trvat i řadu hodin.

Podstata GIS spočívá v rozkladu topografické, speciální nebo archeologické mapy na jednotlivé „vrstvy“ (coverage, image). Takovou vrstvou může být např. vodní síť, výškopis, poloha současných sídel, typy geologického podloží nebo výskyt archeologických nálezů. GIS v první řadě umožňuje zobrazit mapovou vrstvu na obrazovce v libovolném výřezu, odečítat z obrazovky přesnou polohu (X, Y) vybraných bodů a hodnotu (Z) příslušné buňky, překládat přes rastrový obraz libovolný počet obrazů vektorových a tímto způsobem vizualizovat vzájemné vztahy mezi vrstvami. Možné je i konstruovat na obrazovém podkladě jedné vrstvy vrstvu novou v podobě bodů, linií nebo polygonů. Při zobrazení mapové vrstvy je každé hodnotě Z , případně určitému intervalu hodnot, přiřazena jedna z barev (IDRISI 4.0 i IDRISI pod Windows pracuje se 16 nebo 256 barevnými odstíny, paletou šedých odstínů a dalšími paletami; uživatel si také může definovat palety vlastní, viz Dodatek).

Povaha rastrové mapy jako množiny prostorových jednotek, charakterizovaných numerickou proměnnou, však umožňuje provádět s mapovými vrstvami řadu dalších operací, ve kterých teprve spočívá hlavní význam rastrových GIS. Tyto operace lze sdužit do několika tématických okruhů (Eastman 1992):

(a) *Základní okruh*

Operace související s vytvářením rastrových map z vektorových souborů a jejich úpravami. Z věcného hlediska jsou nejvýznamnější tyto operace (v závorce je vždy uvedeno jméno operace):

- reklasifikace (*reclass*) vrstvy, umožňující vybrat z mapy jen určitou třídu jevů, sdužit určité jevy nebo rozřadit plynulé hodnoty do nových, uživatelsky smysluplných tříd (jako příklad uveďme výběr a označení všech ploch o určité nadmořské výšce, svažitosti nebo typu půdy) a vytvoření nové vrstvy s vybranými prvky;
- přiřazení (*assign*) nových hodnot, uložených v databázi, objektům mapové vrstvy;
- interpolace (*interpol*, *intercon*) prostorově nesouvislých dat do souvislého povrchu plynule se měnících hodnot. Jednoduchou metodou interpolace je i tzv. filtrace (*filter*) mapové vrstvy, během níž je každé buňce přiřazena hodnota průměru hodnoty buňky a osmi bezprostředně sousedících buněk.

(b) *Okruh geografických analýz*

Operace, umožňující vizualizaci prostorových vztahů uvnitř vrstvy a vytváření vrstev nových dle požadavků uživatele. Pomocí těchto operací lze zjistit a uložit jako novou vrstvu např.:

- celkový počet buněk s určitou hodnotou, tedy plochu různých objektů mapy (*area*);
- profil hodnot Z podél definované linie (*profile*);
- součet, rozdíl a výsledky dalších aritmetických operací mezi dvěma mapovými vrstvami (*overlay*);
- vzdálenost od vybraných objektů mapy, např. od nejbližšího vodního toku, významných nalezišť apod. (*distance*);
- dostupnost různých míst prostoru na základě jejich eukleidovské vzdálenosti od výchozích bodů a různých průchodností terénu, případně různých terénních překážek (*cost*);

- optimální spojení bodů v prostoru s ohledem na průchodnost terénu (*pathway*);
- areály typu Thiessenových polygonů (*thisessen, allocate*);
- sklon a orientaci svahu na základě digitálního výškopisného modelu (*surface*);
- rozsah povodí (spádového území vodního toku) a areál dohledu z vybraných míst mapy (*watershed, viewshed*);
- „zázemí“ vybraných areálů na základě „poptávky“, definované hodnotou polygonů jedné mapy a „nabídky“, stanovené hodnotami geografického prostředí v mapě druhé (*hinterland*).

(c) *Okruh grafického vyhodnocení vrstev*

Operace specificky zaměřené na klasifikaci satelitních snímků.

(d) *Okruh statistických analýz*

Soubor operací, spojených s celkovým vyhodnocením mapové vrstvy. K základním operacím patří:

- vytvoření histogramu četností hodnot buněk ve zkoumaném obraze (*histo*);
- extrahování (*extract*) celkového součtu hodnot, průměru, směrodatné odchylky atd. buněk jedné mapy, spadajících do různých polygonů mapy jiné (např. počet nalezišť na jednotlivých druhých geologického podloží);
- výpočet trendu v prostorovém uspořádání hodnot v mapě (*trend*), generování náhodného vzorku z buněk mapy (*random, sample*), stanovení některých parametrů uspořádání hodnot mapy (*pattern, autocorr, regress*);
- výpočet korelací a map faktorových skóre ze série mapových podkladů pomocí analýzy hlavních komponent (*principal component analysis, time series analysis*) a další analýzy.

11.4 Vytvoření a úpravy archeologické mapy

Rastrový GIS převádí prostorové objekty, v běžné mapě zobrazené pomocí bodů, linií a ploch do pravidelné sítě buněk, tzv. rastru. Převodem do rastrové mapy se v závislosti na velikosti rastru mění tvar zobrazovaných objektů (bodové objekty se mění na čtverec o ploše jedné buňky, linie se mění na řady buněk a plochy na skupiny buněk). Do rastrového obrazu jsou objekty archeologické mapy (právě tak jako jakékoli jiné objekty) převedeny z vektorového tvaru, který je možno vytvořit buď úpravou běžné databáze nebo digitalizací mapy na digitizéru (podrobnější údaje viz Dodatek). Z věcného hlediska je možné mapu archeologických objektů pojmout několika způsoby:

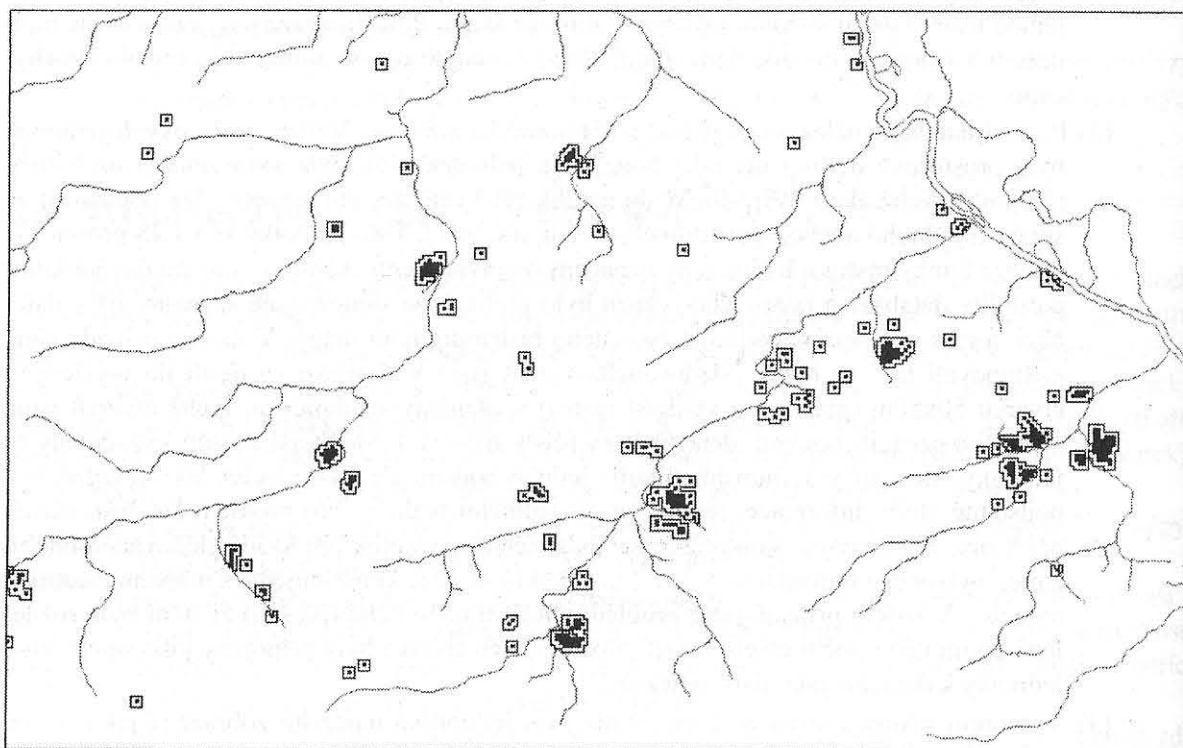
- (a) Hodnoty buněk rastrové mapy reprezentují výskyt (přítomnost) určitých objektů mapy; všem „obsazeným“ buňkám je přiřazena stejná hodnota.
- (b) Hodnoty buněk reprezentují označení jednotlivých objektů (např. pořadová čísla nalezišť). Tímto způsobem lze i v rastrovém tvaru mapy podržet informaci o vzájemné souvislosti buněk patřících týmž prostorovým objektům (stejně číslo označuje všechny buňky, patřící k polygonu určitého archeologického naleziště).
- (c) Hodnoty buněk mapy reprezentují nějakou vlastnost mapovaných objektů. Objektem zde může být jak bod nebo polygon, *vymezující* určité naleziště, tak prostorový celek, který *zahrnuje* nějaké archeologické fakty (např. čtverec povrchového sběru; rozlišení pojmů *vymezující* a *zahrnující* polygon viz Neustupný 1996b; Neustupný a Venclová 1996). Hodnota buňky může vyjadřovat např. počet nálezů ve čtvrcích povrchového sběru apod.

Obsahově různé typy rastrové mapy jsou vzájemně transformovatelné. Připojením hodnot určitých proměnných, uložených v databázi, k identifikátorům mapy typu (b) můžeme získat digitální mapu typu (a) nebo (c). Naopak z těchto map můžeme procedurami GIS automaticky generovat mapu typu (b), v níž každé prostorově souvislé skupině buněk o stejné hodnotě (označující např. přítomnost nálezů určitého stáří, určitý počet nálezů atd.) je přiřazen vlastní identifikátor a tím je tato skupina označena jako samostatný objekt mapy.

Tyto procedury představují jeden z významných přínosů GIS v prostorové analýze archeologických dat, neboť umožňují podle potřeby analyzovat uspořádání různých proměnných v prostoru a na tomto základě vytvářet nové prostorové celky.

Ilustraci některých možností GIS provedeme na nálezech z oblasti Brandýska, která byla v letech 1986-1990 intenzivně sledována z pracoviště OM v Brandýse n.L. Rozsah studovaného území je 16x10 km a tvoří jej povodí Vinořského a Mratínského potoka a několika dalších menších potoků, levobřežních přítoků Labe. Nálezový fond tvoří informace o starších nálezech na daném území, výsledky záchranných výzkumů 80. let a souběžně prováděných povrchových sběrů. Rozšíření nálezového fondu systematickými povrchovými sběry v projektu ALRNB v letech 1991-1995 nebylo podstatné; i tyto nálezy však byly do databáze začleněny. Zejména povrchové sběry 80. let pokryly velkou část daného území (obr. 1 nahoře), takže pro ta archeologická období, která se vyznačují relativně větší archeologickou výrazností, můžeme získaný materiál považovat za poměrně reprezentativní vzorek někdejšího osídlení v ohledu celkového počtu a rozmístění pravěkých areálů. Archeologické mapy jsme konstruovali pro pět pravěkých období (mladší doba bronzová, pozdní doba bronzová, doba halštatská, doba laténská a doba římská). Vytvoření a zpracování rastrových map proběhlo v těchto krocích:

- (1) Sestavení *databáze nálezových jednotek*, tj. chronologických a prostorových celků, zjištěných jednou archeologickou akcí. Záznamy databáze odpovídaly jednotlivým nálezovým jednotkám, sloupce (pole) databáze pak jednotlivým archeologickým obdobím. V úvahu byly brány všechny nálezy sídlištního charakteru a nálezy z povrchových sběrů, u nichž předpokládáme souvislost se sídlištními aktivitami; pohřebiště byla v této fázi ponechána stranou. Při popisu nálezových jednotek jsme se rozhodli rozlišovat pouze přítomnost a nepřítomnost nálezů určitého období, neboť heterogenní charakter dat neumožňoval vzájemné kvantitativní srovnání nálezů v jednotné škále. Pole databáze tedy obsahovala hodnoty 0 a 1 podle toho, zda dané období bylo či nebylo danou archeologickou akcí zachyceno.
- (2) Převod databáze nálezových jednotek na *databázi nalezišť*. Většina nálezových jednotek byla prostorově definována jako bod. Řada jednotek však byla vymezena tímž bodem (jednotky týchž akcí), případně body natolik blízkými, že tyto jednotky lze považovat za součásti jednoho a téhož prostorového celku, naleziště. Tyto jednotky se v GIS promítnou do téže buňky rastru a budou tedy označeny stejným identifikátorem; pro zachování kompatibility databáze a rastrového obrazu bylo proto třeba sloučit dané záznamy již v databázi, a to s ohledem na velikost zvoleného rastru digitální mapy. V daném případě jsme postupovali tak, že rozsah sledovaného území jsme v databázi rozdělili do myšlených čtverců 50x50m (plánovaná velikost rastru) a záznamy spadající do týchž čtverců jsme sloučili a označili novými identifikátory (čísla nalezišť). Složitější postup vyžadovaly ty jednotky, které byly definovány nikoliv jedním bodem, ale jako plocha, kterou nelze, bez podstatné ztráty informace, redukovat do jediného bodu. Tento problém lze řešit různě, např. tím, že sestavíme samostatné databáze těchto jednotek pro každé sledované období, z nich vytvoříme samostatné mapové vrstvy a ty posléze kombinujeme s mapami ostatních nalezišť. V našem případě jsme problém obešli tím, že větší (plošná) zjištění byla rozdělena do určitého počtu čtverců v síti 50x50m, tyto čtverce byly připojeny jako samostatné jednotky k databázi jako další naleziště.
- (3) Vytvoření rastrové *mapy nalezišť*, v níž jsou jednotlivá naleziště zobrazena jako buňky s vlastními identifikátory, odpovídajícími pořadovým číslům nalezišť v databázi. Naleziště jsou zde chápána jako diskrétní prostorové celky, sdružující na empirickém základě nálezy z různých terénních akcí podle jejich prostorové souvislosti (v daném případě odpovídala naleziště jednotlivým buňkám mapy, tj. čtvercům 50x50m).
- (4) Výběr (*dbidris*) proměnných, tj. jednotlivých období, z databáze a přiřazení (*assign*) jejich hodnot příslušným bodům mapy nalezišť. Výsledkem jsou *mapy výskytu nálezů* jednotlivých období. Hodnoty buněk (0 nebo 1) v našem případě rozlišovaly pouze přítomnost a nepřítomnost (sídlíštních) nálezů určitého stáří v daném bodě (buňce). Rastr o velikosti 50x50m jsme zvolili na základě úvahy, že zhruba takováto plocha může odpovídat rozsahu obytného areálu jedné pravěké usedlosti a odrážet tedy jednu pravěkou *sídelní událost*.



Obr. 1 nahoře: Mapa současného stavu výzkumu Brandýska (mikroregion Vinořského a Mratínského potoka). Rozměry území 16 x 10km; orientace S-J. Světle šedé plochy: polygony povrchového sběru projektu ALRNB v letech 1991-1995; tmavší šedé plochy: polygony povrchového sběru 1986-1990; černé body: starší archeologické výzkumy a výzkumy OM Brandýs n.L. v letech 1986-1990.

Dole: Mapa výskytu nálezů a komponent mladší doby bronzové. Černé body odpovídají čtvercům 50x50m (jednotlivým sídelním událostem). Linie vymezují jednotlivé komponenty tím, že k ploše s nálezem přiřazují jednu řadu sousedních prázdných čtverců.

Jsou-li archeologické nálezy určitého období vzájemně vzdáleny tak, že spadají do různých buněk rastru, s velkou pravděpodobností nejde o jednu a tutéž sídelní událost; naopak, pokud jsou nálezy téhož období vzájemně vzdáleny méně, nelze to vyloučit. Je-li tento předpoklad správný, pak počet obsazených buněk rastru vyjadřuje minimální počet sídelních událostí určitého období. (Tento postup by šlo ještě dále zpřesňovat, neboť je zřejmé, že do různých čtverců mohou v našem případě někdy padnout nálezy jednotky bližší než 50m a naopak v rámci téhož čtverce se mohou vyskytnout jednotky vzdálené teoreticky až 70m). Digitální rastrovou mapu můžeme analyzovat dvojím způsobem: jako plochu, pokrytou nalezišti s výskytem nálezů určitého druhu nebo jako množinu nalezišť určité kategorie. (V našem případě bude výsledek totožný, neboť každé naleziště odpovídá jedné buňce rastru, jindy však mohou naleziště zobrazena jako polygony s různým počtem buněk rastru). Tento postup ovšem nemusí být vždy dostačující, neboť počet a rozmístění nalezišť jsou jednak vždy poplatné charakteru výzkumu, jednak se z takového obrazu vytrácí informace o vzájemné souvislosti prostorově blízkých nálezů (nalezišť) stejného stáří. Prostorové celky chronologicky současných nálezů obdobné funkce nazýváme *komponentami* (viz Neustupný 1993); pomocí GIS lze takové celky poměrně snadno generovat:

- (5) Mapy výskytu nálezů rozdělíme procedurou *group* do nových prostorových celků, a to tak, že každé souvislé skupině buněk se stejnou hodnotou je přiřazen nový identifikátor. Výsledek představuje *mapu komponent*, tj. prostorově souvislých celků nálezů určitého druhu a stáří. Před touto operací také můžeme mapy výskytu různým způsobem upravit. Takovou úpravou může být např. spojení buněk, které jsou vzájemně vzdáleny méně než je mez, kterou stanovíme. Procedurou *filter*, *distance*, *cost* nebo *buffer* můžeme totiž ke každé buňce s výskytem příslušného nálezu přiřadit buňky sousední do libovolné vzdálenosti, doplnit buňky, které leží mezi dvěma buňkami s výskytem nálezů apod. Při těchto procedurách můžeme také použít masku, odvozenou z map přírodního prostředí a omezit „rozdávání“ komponent geografickými faktory. Tím se lze vyhnout např. tomu, aby automaticky generované komponenty zasahovaly např. přes vodní toky, přes plochy prudkých svahů apod. (obr. 1 dole).

11.5 Interpolace dat

Vytvořením mapy komponent převádíme archeologické prameny z empirických celků, vzniklých terénním výzkumem, do prostorových celků, definovaných na základě analýzy jejich formálních a prostorových vztahů. Mapa komponent však stále představuje zobrazení sídelního prostoru v podobě nespojitých jednotek (hodnot). Tím je ovšem jak do určité míry zkreslen jeho původní charakter geografického kontinua, tak jsou omezeny možnosti jeho analýzy vůči souvislým, plynule se měnícím vlastnostem krajiny. Určitá část mapy buď patří nebo nepatří komponentě určitého stáří, ale hodnoty jednotlivých buněk nemohou vyjádřit, jak významný z hlediska celku je daný bod prostoru a s jakou pravděpodobností můžeme na tom či onom místě mimo vymezené komponenty očekávat nálezy další. Vytvoření mapy, v níž určité vlastnosti prostoru jsou vyjádřeny souvislými, plynule se měnícími hodnotami je možné různými procedurami interpolace dat. Výsledky interpolace mohou být velmi rozdílné, a proto volba konkrétního postupu musí odrážet posouzení charakteru dat, jejich reprezentativnosti a úkolů, k jejichž řešení má upravená mapa sloužit.

V našem případě jsme předpokládali, že význam určité části prostoru z hlediska jeho někdejšího osídlení se odráží především v množství sídelních událostí, které proběhly v jeho rámci. Významné areály se proto budou projevovat jako větší koncentrace buněk mapy, zatímco areály přechodně či náhodně osídlené budou reprezentovány ojedinělými buňkami mapy, vzdálenými od hlavních areálů. Dále můžeme předpokládat, že pravděpodobnost výskytu dalších areálů klesá se vzdáleností od hlavních center sídelních aktivit, které byly situovány v místech, které daná komunita považovala za optimální. Každému místu mapy tedy můžeme přiřadit určitý význam podle toho, kolik doložených sídelních událostí je situováno v jeho okolí a na základě vzdálenosti, které daný bod mapy od těchto míst odděluje.

Naším předpokladům nejlépe vyhovovala jednoduchá metoda filtrace (*filter*) mapy. Procedura filtrace přisuzuje každé buňce mapy průměrnou hodnotu, vypočtenou z hodnoty dané buňky a hodnot osmi bezprostředně sousedních buněk. Jelikož jsme pracovali v rastru 50x50m, dvojí filtrací map výskytu nálezů byly interpolovány původní hodnoty buněk v okruhu 100m. Tímto postupem (a) se podstatně snížily hodnoty prostorově osamocených buněk, představujících ojedinělé sídelní události mimo hlavní těžiště obytných areálů; (b) vytvořil se 100m široký pás plynulého přechodu mezi polygony s hodnotou 1 (tj. polygony s výskytem nálezů) a okolním prázdným prostorem a (c) došlo k propojení buněk vzdálených až 200m do větších polygonů s nenulovými hodnotami. Po každé filtraci byly automaticky odstraňovány ty nově obsazené buňky mapy, které překračovaly hranice, které pokládáme z hlediska archeologických areálů za omezující (bylo zabráněno, aby se polygony nenulových hodnot rozrůstaly a spojovaly přes vodní tok; obr. 2 nahoře).

Takto upravené mapy poskytují odlišný obraz osídlení než pouhá mapa výskytu nálezů. Interpolovaná mapa je sice stále závislá na reprezentativnosti vstupních dat, jsou v ní však potlačeny náhodné faktory v uspořádání nálezů a je odhadován význam každého bodu v krajině na základě jeho prostorového vztahu k známým areálům v plynulé škále souvislého povrchu hodnot.

11.6 Příklady

11.6.1 Příklad 1: Kontinuita obytných areálů

Jedním z cílů výzkumu bylo sledovat vzájemný vztah mezi areály jednotlivých pravěkých období a vysvětlení tohoto vztahu jako určité zákonitosti vývoje pravěkého osídlení. Interpolované (filtrované) mapy osídlení umožňují provést korelaci mezi jednotlivými mapovými vrstvami. Mapy sledovaných období byly korelovány procedurou *pca* (principal component analysis), v níž je stanovení korelačních koeficientů jedním z vedlejších výsledků (spolehlivost této procedury byla zkontrolována i převodem buněk mapy do databáze a výpočtem korelačních koeficientů pomocí statistického balíku SPSS). Procedura *pca* určuje korelační koeficient mezi všemi dvojicemi mapových vrstev na základě porovnání hodnot v polohově korespondujících buňkách celé mapy; výběr buněk zahrnutých do analýzy však lze omezit maskou). Získané korelační koeficienty jsou obsaženy v tab. 1.

	BR.ML	BR.PO	HALSTA	LATEN	RIM
BR.ML	1.00	0.66	0.25	0.26	0.13
BR.PO	0.66	1.00	0.35	0.43	0.13
HALSTA	0.25	0.35	1.00	0.53	0.13
LATEN	0.26	0.43	0.53	1.00	0.18
RIM	0.13	0.13	0.13	0.18	1.00

Tab. 1. Prostorová korelace osídlení v obdobích mladší doby bronzové až doby římské.

Na základě korelačních koeficientů můžeme formulovat několik závěrů:

- (a) Osídlení všech sledovaných období vykazuje silnou prostorovou korelaci, neboť všechny korelační koeficienty jsou statisticky významné, a to na hladině významnosti 99% (hladinu významnosti zde však nelze brát jako pevný statistický údaj, neboť vstupní data byla modifikována interpolací). Silnou prostorovou korelaci map lze vysvětlit obdobnou polohou pravěkých obytných areálů v dlouhodobém časovém měřítku, tj. jako dlouhodobou stabilitu vyšších sídelních celků, sídelních areálů.
- (b) Hodnoty korelačních koeficientů ukazují určitou zákonitost. Patrný je zejména fakt, že nejsilnější korelace existují vždy mezi kulturami chronologicky nejbližšími, přičemž s chronologickou vzdáleností hodnota koeficientů klesá. Tuto skutečnost lze nejlépe vysvětlit specifickým charakterem mobility obytných areálů. Obytné areály byly zřejmě periodicky přemisťovány na krátké vzdálenosti, zřejmě do bezprostředního sousedství areálů opouštěných. Tento proces mohl mít podobu víceméně náhodného (nesměrovaného)

pohybu, při kterém se postupně přemísťovalo i těžiště, kolem kterého posuny jednotlivých pravěkých usedlostí oscilovaly. Příčinu, proč vzájemně korelují, byť mnohem slaběji, i chronologicky vzdálenější období, lze hledat v relativně malé ploše území, které bylo k dispozici jedné pravěké komunitě (sídelní areál). Plocha, po které se obytné areály jedné komunity mohly pohybovat, byla omezená, a proto muselo docházet i k prostorovým průnikům obytných areálů vzdálenějších období; z hlediska čistě pravděpodobnostního však k těmto průnikům muselo docházet méně často, což je patrné z korelačních koeficientů.

- (c) Hodnoty korelačních koeficientů naznačují, že podle míry vzájemné korelace lze sledovaných pět období rozdělit do tří skupin. Hodnoty korelačního koeficientu nad 0.5 ukazuje dvojice mladší d. bronzová – pozdní doba bronzová a dvojice halštat – latén. Kultury každé z těchto dvojic mají vůči kulturám z druhé dvojice korelaci vždy poněkud slabší (0.25 – 0.43), přičemž korelační koeficienty všech těchto kultur ve vztahu k době římské jsou shodně ještě nižší. Prokázat tuto strukturu statisticky by bylo obtížné, neboť všechny korelační koeficienty jsou ještě v rozmezí statisticky významných hodnot, nicméně však na tomto základě můžeme formulovat hypotézu, že prostorové uspořádání pravěkých areálů odráží tři širší fáze vývoje: (a) mladší a pozdní doba bronzová, (b) doba halštatská a laténská a (c) doba římská), jejichž vnitřní prostorové vazby jsou silnější než vazby vůči kulturám ostatním.

11.6.2 Příklad 2: Počet a rozsah sídelních areálů

Proceduru *pca* (principal component analysis) lze použít nejen k výpočtu korelačních koeficientů, ale i k provedení *vektorové syntézy* (faktorové analýzy) mapových vrstev. Tímto postupem lze zjistit hlavní osy variability prostorového uspořádání pravěkého osídlení a vytvořit *mapy faktorových skóre*, které lze interpretovat jako mapy významnosti prostoru z hlediska dlouhodobého vývoje sídelní struktury (základními pojmy, týkajícími se této metody, se nebudeme zabývat, neboť jsou vysvětleny na jiném místě tohoto svazku; Neustupný 1996c). Jelikož procedura *pca* v GIS IDRISI má některá vážná omezení (neumožňuje např. omezit počet vypočítávaných faktorů a rotaci faktorů), doporučujeme k takovým úkolům použít GIS v kombinaci s jiným statistickým programem (v našem případě šlo o SPSS 6.1; tuto metodu viz též Neustupný 1996b; Neustupný a Venclová 1996). Postup probíhal v následujících krocích:

- (1) Sjednocením všech sledovaných mapových vrstev (*overlay*) byla vytvořena souborná mapa, v níž byl každé buňce s nenulovou hodnotou přiřazen vlastní identifikátor (celkem 6218 buněk).
- (2) V databázovém systému dBASE IV byla vytvořena databáze, v níž počet záznamů odpovídal počtu nenulových buněk souborné mapy a počet polí počtu sledovaných map (plus jedno pole, obsahující identifikátor buňky).
- (3) Procedurou *dbidris* byly do příslušných polí databáze přeneseny hodnoty z jednotlivých (interpolovaných) archeologických map.
- (4) Databáze byla podrobena analýze hlavních komponent v balíku SPSS 6.1, přičemž pro každý z významných faktorů a každou buňku mapy bylo vypočteno faktorové skóre a přiřazeno k databázi jako další pole.

Na základě zvážení vlastních čísel faktorů byly nakonec extrahovány tři faktory, které celkově vysvětlují přes 80% variability daného datového souboru. Faktorové koeficienty těchto tří faktorů ukazují zřetelně základní strukturu prostorových vztahů mezi jednotlivými kulturami a rozdělují je do tří hlavních fází, naznačených již korelačními koeficienty v předchozím příkladě: (a) mladší a pozdní doba bronzová, (b) doba halštatská a laténská a (c) doba římská. Tento výsledek souhlasí s výsledkem získaným pouhou korelací map, na rozdíl od korelace však vektorová syntéza více zdůraznila rozdíly v prostorovém uspořádání (tab. 2).

Faktor	vlastní číslo	% variability	kumul.% var.
1	1.99	39.8	39.8
2	1.12	22.4	62.2
3	0.98	19.7	81.9
4	0.54	10.8	92.7
5	0.36	7.3	100.0

Rotovaná matice faktorů:

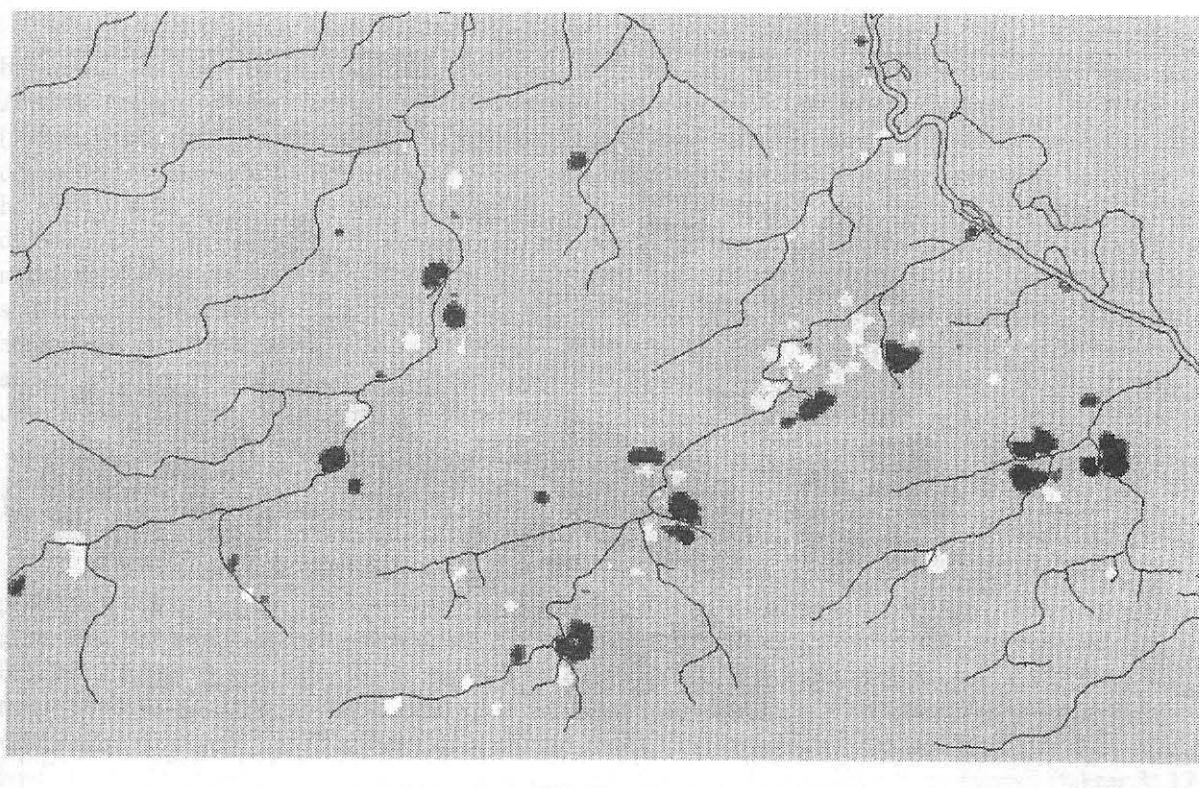
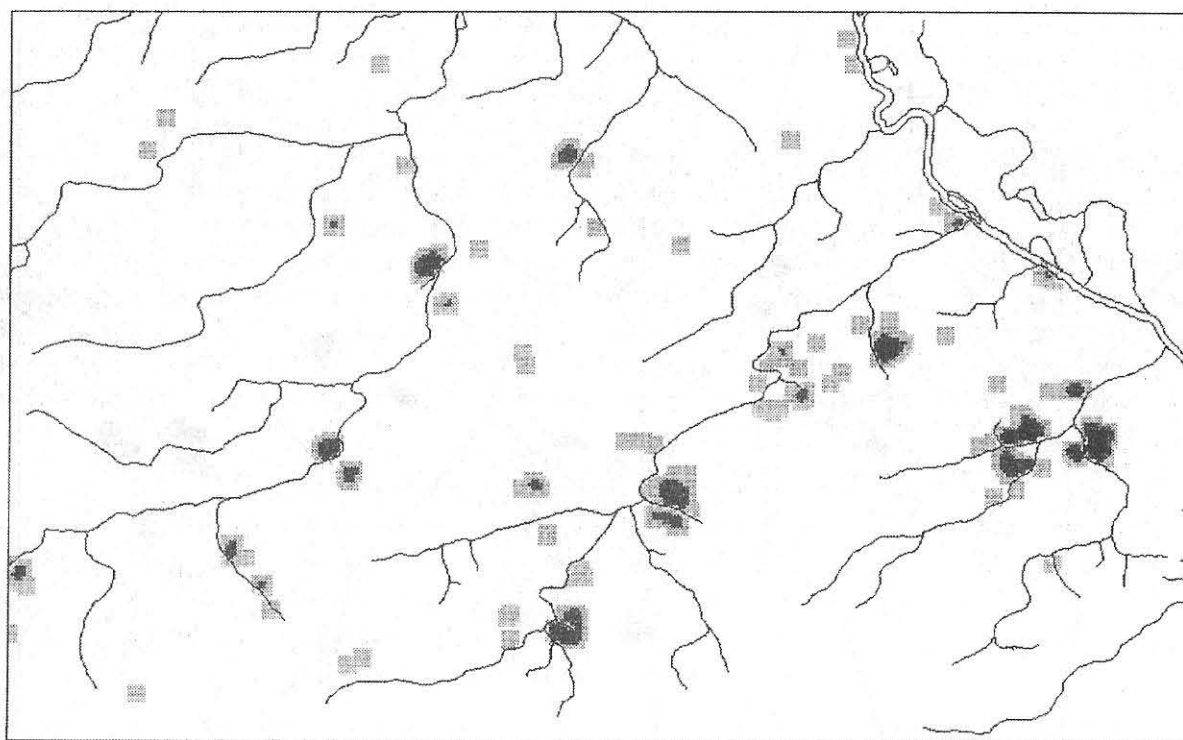
mapa	faktor 1	faktor 2	faktor 3
BR.ML	0.92	0.01	0.02
BR.PO	0.85	0.28	0.00
HALSTA	0.12	0.84	0.00
LATEN	0.11	0.85	0.05
RIM	0.01	0.04	0.99

Tab. 2. Popis faktorů, zjištěných vektorovou syntézou.

- (5) Hodnoty faktorových skóre byly pomocí *dbidris* a *assign* přeneseny zpět do GIS, kde z nich vznikly samostatné mapové vrstvy. Faktorová skóre přiřazují každé buňce mapy statistickou významnost z hlediska jednoho ze tří faktorů. Ze statistického hlediska jsou za signifikantní považována faktorová skóre větší než 1.00; tyto plochy mohou být z map faktorových skóre vyjmuty jejich reklasifikací (*reclass*).

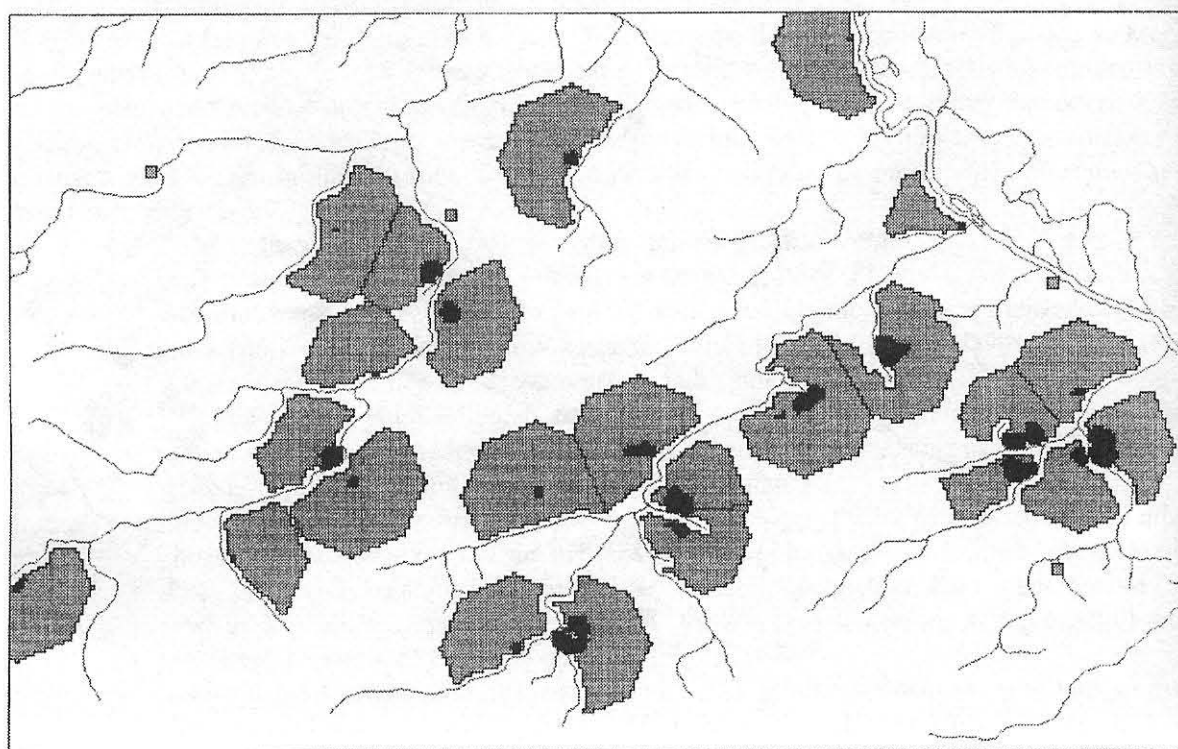
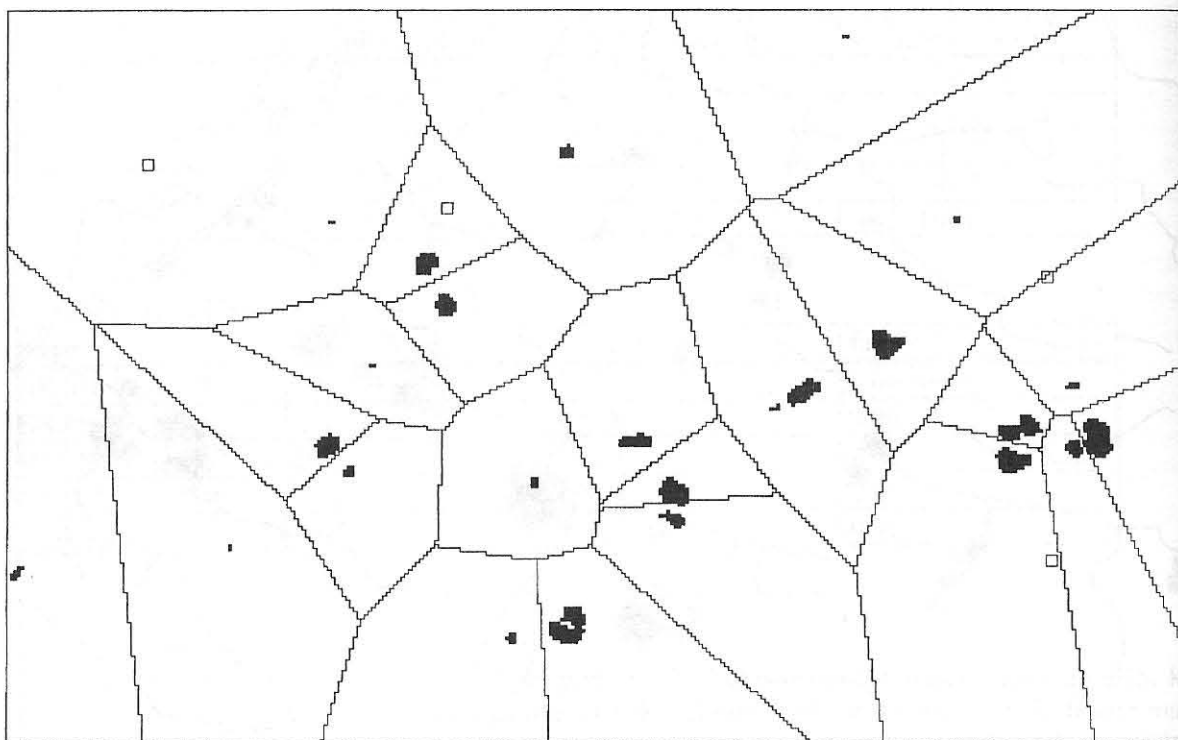
Místa s faktorovým skóre větším než 1.00 můžeme považovat za prostorové celky, ve kterých je objektivně prokázána vysoká intenzita osídlení některého z období, charakteristického pro určitý faktor, případně souběh osídlení několika (v případě našich faktorů nejvýše dvou) charakteristických období. Nejpravidelnější strukturu ukazuje uspořádání těchto ploch v prvním faktoru, typickém pro mladší a pozdní dobu bronzovou (obr. 2 dole); na jeho základě můžeme proto nejlépe formulovat několik obecných závěrů:

- Zjištěné faktory sdružují vždy následná archeologická období, což do značné míry nezávisle potvrzuje smysluplnost výsledků vektorové syntézy. Plochy s vysokým faktorovým skóre můžeme považovat za dlouhodobě využívané obytné areály, v nichž došlo k nadprůměrné kumulaci sídlištních pozůstatků. Tato místa mohla představovat jakási těžiště sídlištních aktivit v delším časovém úseku, jádra sídelních areálů.
- Všechna zjištěná jádra sídelních areálů jsou vzájemně výrazně oddělena, a to buď prázdným prostorem nejméně 850m, nebo vodním tokem. Z toho můžeme usuzovat, že každé ze zjištěných jader odpovídá samostatnému sídelnímu areálu.
- Jelikož většina jader zahrnuje nálezy obou období, typických pro daný faktor (tj. mladší i pozdní doby bronzové), můžeme je považovat za jádra sídelních areálů souběžně existujících pravěkých komunit (nikoliv výsledek posunu sídelních areálů během daného časového úseku). Zjištěný počet jader proto s vysokou pravděpodobností reprezentuje počet pravěkých komunit, obývajících dané území.
- Jádra sídelních areálů vykazují jasnou tendenci k určitému shlukování, a to buď ve dvojicích na protilehlých březích vodních toků, nebo ve větších shlucích na soutocích. Vodní toky tedy tvořily přirozené hranice sídelních areálů, avšak umístění obytných areálů sledovalo snadnější komunikaci mezi sousedními komunitami. Zajímavé rovněž je, že podobné uspořádání mají v některých částech území i současná sídla, navazující na síť středověkých vesnic.
- Umístění jader sídelních areálů naznačuje, že obytný prostor byl v rámci teritoria komunity (sídelní areál jako celek) situován excentricky a výrobní areály (pole, pastviny, využívaná část lesa) byly převážně umístěny vždy jen na jedné straně vodního toku.



Obr. 2 nahoře: Interpolovaná mapa nálezů mladší doby bronzové. Světle šedé polygony: hustota 0.01-0.2; tmavošedé polygony: 0.2-0.6; černé polygony: 0.6-1.0 sídelní události na čtverec (buňku) mapy.

Dole: Mapa faktorových skóre faktoru 1. Bílá: skóre menší než -0.5 (tj. plochy pro faktor 1 spíše atypické a velmi atypické); světle šedá (pozadí): skóre -0.5 až 0.5; tmavošedá: skóre 0.5 až 1.0; černá: skóre větší než 1.0 (plochy typické pro faktor 1).



Obr. 3 nahoře: Mapa Thiessenových polygonů kolem sídelních jader faktoru I (mladší a pozdní doba bronzová). Prázdné čtverce představují pravděpodobná jádra dalších sídelních areálů.

Dole: Rekonstruované sídelní areály mladší a pozdní doby bronzové. Šedé čtverce představují pravděpodobná jádra dalších sídelních areálů.

Pomocí některých procedur GIS můžeme sídelní areály modelovat:

- (1) Procedurou *thiessen* nebo *allocate* vytvoříme kolem každého z jader plochu tzv. Thiessenových polygonů. Thiessenovy polygony přiřazují každý bod mapy tomu výchozímu prostorovému objektu, k němuž jsou nejbližší (obr. 3 nahoře).
- (2) Takto získané Thiessenovy polygony byly omezeny předpokládanými areály dostupnosti, vypočtenými pro jádra sídelních areálů. Areály dostupnosti byly zkonstruovány jako plochy o poloměru 1000m, které nesměly překročit vodní toky.

Předpokládané sídelní areály (celkem 25) jsme definovali průnikem obou výše zmíněných map (obr. 3 dole). Každému sídelnímu jádru byla tedy přiřazena plocha buněk, které (a) jsou situovány danému jádru blíže než kterémukoli z jiných jader; (b) jsou v okruhu 1000m od daného jádra; (c) leží na stejném břehu vodního toku jako dané sídelní jádro. U takto definovaných sídelních areálů byly změřeny jejich plochy (*area*), shrnuté v tab. 3 (minimum bylo 45 ha, maximum 202 ha).

Plocha sídelních areálů (ha)	Počet areálů
< 50	1
51-100	4
101-150	12
151-200	6
> 200	1

Tab. 3. Plocha sídelních areálů mladší a pozdní doby bronzové v mikriregionu Vinořského a Mratinského potoka.

Z tab. 3 vyplývá především pravidelnost v rozmístění sídelních jader mladší a pozdní doby bronzové, bez které by jednotlivé sídelní areály nemohly nabývat plochy o tak standardní velikosti. Zároveň velikost předpokládaných areálů odpovídá modelu, nedávno formulovanému D. Dreslerovou (1995). Dreslerová předpokládá, že se průměrná komunita v mladším pravěku skládala zhruba ze čtyř domácností. Taková komunita potřebovala zhruba 2 ha obytné plochy, 4 hektary polí a 20 ha úhoru, 1 ha pastvin a 60 ha lesa jako zdroje palivového a stavebního dřeva a zdroje krmiva pro zimní krmení dobytka. Dále Dreslerová uvádí, že při tomto způsobu čerpání přírodních zdrojů nedochází k trvalému narušení ekologické rovnováhy krajiny; takto uspořádané sídelní areály proto mohou být víceméně stabilní během dlouhého období. Naše výpočty předložený model jasně potvrzují. Pouze tři sídelní areály jsou menší než uváděná kritická mez (87 ha); i v těchto případech však existuje kolem sídelních areálů ještě určitý prázdný prostor, který komunita mohla využít za předpokladu mírného zvětšení okruhu svých aktivit nad odhadnutou mez 1000m.

11.6.3 Příklad 3: Vztah obytných areálů ke krajině

Přestože prostor sídelních areálů byl poměrně omezený, poskytoval pravěkým komunitám příležitost využít krajinné prostředí různými způsoby. K aspektům variabilního využití krajiny patřilo i umístění vlastního obytného areálu, které odráželo jak faktory ekonomické, tak společenské a symbolické. Základem zkoumání těchto aspektů byla poloha sídelních jader, definovaných v mapách faktorových skóre jako plochy s hodnotou větší než 1.00. Tyto mapy odpovídají třem faktorům, získaným vektorovou syntézou, a představují hlavní obytné areály tří širších období: mladší a pozdní doby bronzové (faktor 1; 25 areálů), doby halštatské a laténské (faktor 2; 31 areálů) a doby římské (faktor 3; 17 areálů). Na tomto místě se budeme zabývat rozbořem dvou aspektů jejich polohy: vzdálenosti od vodního toku a vztahem k reliéfu krajiny.

Závislost polohy pravěkých obytných areálů na vodním zdroji není třeba zvlášť prokazovat; v našem příkladě proto šlo pouze o stanovení přesnějších kvantitativních parametrů. Vodní síť byla digitalizována podle současného stavu, opraveného na základě map 1. vojenského mapování z konce 18. stol. Procedurou *distance* byla vytvořena mapa vzdáleností od nejbližšího vodního toku (obr. 5 – barevná příloha) a z této mapy byly extrahovány (*extract*) průměrné hodnoty pro každé ze sídelních jader jednotlivých období (faktorů). Souhrnný výsledek ukazuje, že více než 90% sídelních jader bylo

situováno ve vzdálenosti do 300m od vodního zdroje. Podobným způsobem byly zpracovány i mapy náleží jednotlivých období; celkové součty sídelních událostí zahrnuje poslední sloupec tab. 4. Tyto hodnoty ukazují, že jednotlivá pravěká sídliště se mohla vyskytnout i ve vzdálenostech větších; tento výskyt má však charakter spíše náhodný. Rozdíly mezi jednotlivými chronologickými obdobími nebyly významné.

Průměrná vzdálenost (m)	Počet sídelních jader			celkem událostí
	faktor 1	faktor 2	faktor 3	
0-100	3	9	4	337
100-200	17	14	9	654
200-300	4	3	3	245
300-400	–	2	1	92
400-500	1	3	–	49
500-600	–	–	–	6
600-700	–	–	–	2
700-800	–	–	–	1
800-900	–	–	–	–
>900	–	–	–	5
celkem	25	31	17	1391

Tab. 4. Vzdálenost obytných areálů od vodního zdroje.

Jako jeden ze zajímavých geografických parametrů umístění areálů se jeví jejich vztah k vertikálnímu členění krajiny. Je známo, že některé kultury pravěku preferovaly místa vyvýšená, např. ostrožny, okraje teras apod., jiné naopak údolní dno nebo spodní partii svahů na samém okraji nivy. Při sledování tohoto aspektu jsme vyšli z *digitálního výškopisného modelu krajiny (DEM)*. DEM je digitální mapou, která vzniká interpolací vrstevnicové mapy nebo bodových měření do souvislého povrchu, ve kterém je každé buňce přiřazena vypočtená hodnota nadmořské výšky (obr. 6 – barevná příloha). DEM je základem řady dalších mapových vrstev, jejichž vytváření patří ke standardním funkcím GIS, např. mapy sklonu svahu, orientace svahu, areálu dohledu, povodí atd., a slouží i k trojrozměrné ortogonální vizualizaci vybraných úseků krajiny (*ortho*).

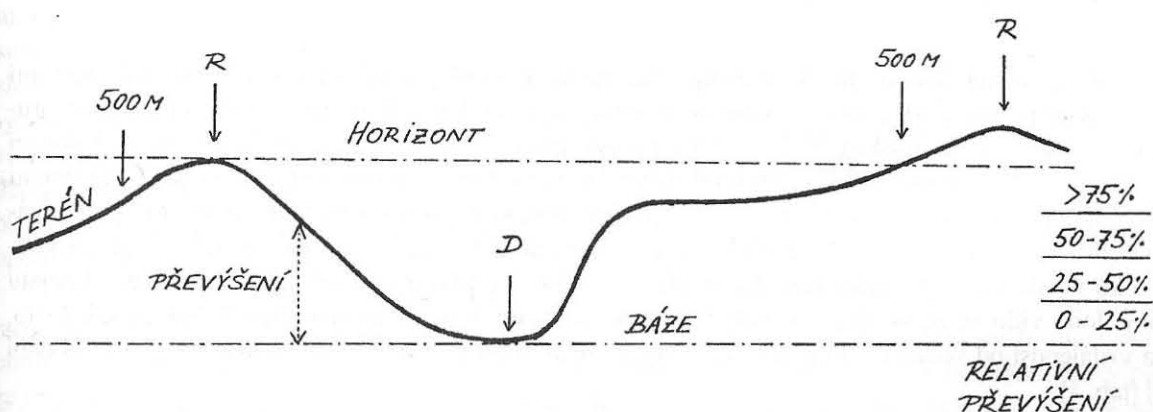
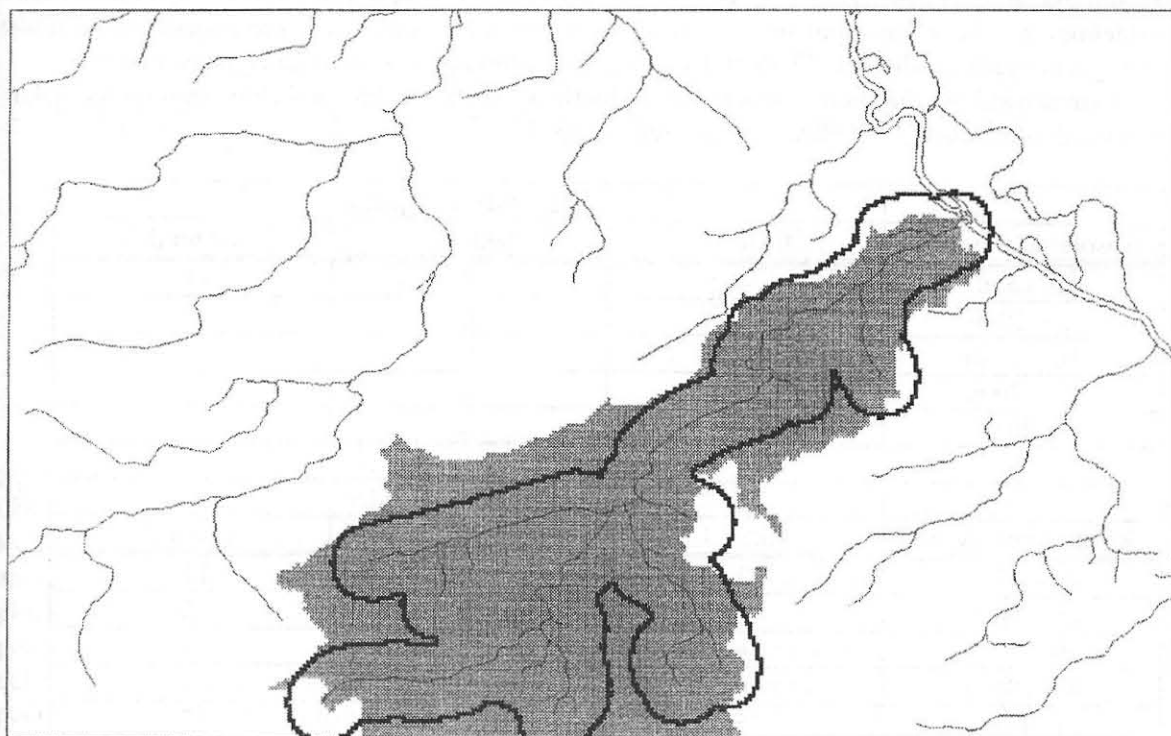
Z digitálního výškopisného modelu v rastru 10x10m jsme vytvořili mapy převýšení, které stanovují výškový rozdíl mezi každou z buněk mapy a úrovní údolního dna nejbližšího vodního toku. Mapy převýšení byly získány následujícím způsobem:

- (1) Z mapy nadmořské výšky terénu (DEM) byly vyříznuty linie, polohově odpovídající vodním tokům, tj. nejhlubším místům údolních den. Tyto linie byly interpolovány (*intercon*) do souvislého povrchu stejným způsobem jako vrstevnicová mapa, čímž byla celou krajinou proložena rovina, spojující místa lokálně nejmenší nadmořské výšky (vrstva BÁZE; obr. 4 dole).
- (2) Kolem vodních toků byla konstruována linie, spojující nejvyšší hodnoty nadmořské výšky v oblasti užívané pro obytné areály, tj. v prostoru do vzdálenosti 500m od vodního toku. Tato linie byla definována jako vnější okraj plochy průniku povodí (*watershed*) a vzdálenosti 500m od vodního toku. Linie byla totožná s rozvodím potoků, pokud probíhalo ve vzdálenosti menší než 500m od potoka, nebo s izolinií 500metrové vzdálenosti od potoka, pokud rozvodí bylo vzdálenější (obr. 4 nahoře).
- (3) Touto linií byla vyříznuta mapa nadmořské výšky (DEM) a získané hodnoty byly interpolovány do souvislého povrchu, představující rovinu nejvyšších nadmořských výšek v prostoru vhodném pro umístění obytných areálů (vrstva HORIZONT; obr. 4 dole).
- (4) Aritmetickými operacemi (*overlay*) s těmito vrstvami byla získána *mapa převýšení* a *mapa relativního převýšení* nad nejbližším vodním tokem. Mapa převýšení zobrazuje rozdíl

v nadmořské výšce mezi každou buňkou mapy a úrovní nejbližšího údolního dna (v metrech); mapa relativního převýšení udává převýšení každé buňky jako poměrnou část největšího možného převýšení v dané části mapy. Provedené operace jsou objasněny v následujících vzorcích:

$$\text{PŘEVÝŠENÍ} = \text{DEM} - \text{BÁZE}$$

$$\text{RELATIVNÍ PŘEVÝŠENÍ} = \frac{\text{PŘEVÝŠENÍ}}{(\text{HORIZONT} - \text{BÁZE})} * 100$$



Obr. 4 nahoře: Povodí Vinořského potoka, vypočtené z digitálního výškopisného modelu (šedá plocha). Černou linií vymezena vzdálenost 500m od potoka.

Dole: Schéma konstrukce mapy převýšení a mapy relativního převýšení. „500 M“: vzdálenost 500m od vodního toku; R: rozvodí; D: nejhlubší místo údolního dna.

Při rozlišení (absolutního) převýšení a relativního převýšení jsme vycházeli z úvahy, že každý z obou prvků může být potenciálně měřítkem jiných aspektů sídelního chování. Absolutní převýšení areálu, měřené od údolního dna, může zachytit např. vztah areálů k režimu vodního toku a k údolní nivě, kterou mohou některé kultury vyhledávat, jiné nikoliv. Z tohoto hlediska mají význam zejména nižší hodnoty převýšení. Čím výše nad vodním tokem však je určitý areál situován, tím více ztrácí tento (absolutní) údaj smysl, neboť každá část krajiny nabízí k osídlení reliéf s jiným převýšením, a tudíž výsledné hodnoty jsou obtížně srovnatelné. Relativní převýšení tento faktor odstraňuje, neboť vychází z výšky horního okraje osídleného terénu, z „horizontu“, a vůči němu poměřuje polohu ostatních míst krajiny. Z tohoto pohledu nabývá areál na horním okraji údolí vždy stejnou (maximální) hodnotu, bez ohledu na absolutní hloubku údolí. Předpokládáme-li např., že záměrem pravěké komunity mohlo být situovat obytný areál do polohy tak „dominantní“, jak to jen umožňoval charakter daného sídelního areálu, získává pojem relativního převýšení smysl, neboť v tomto případě by se relativní převýšení všech areálů mělo blížit 100%, třebaže absolutní převýšení se může podstatně lišit.

Extrahování příslušných hodnot pro jednotlivá sídelní jádra proběhlo analogicky jako v předchozích příkladech. Výsledky jsou obsaženy v tab. 5:

Absolutní převýšení	Počet sídelních jader		
	faktor 1	faktor 2	faktor 3
0 – 5 m	4	10	11
5 – 10 m	11	6	5
10 – 15 m	7	9	–
15 – 20 m	3	2	1
> 20 m	–	4	–

Relativní převýšení	Počet sídelních jader		
	faktor 1	faktor 2	faktor 3
0 – 25 %	3	10	11
25 – 50 %	6	5	5
50 – 75 %	6	7	–
75 – 100 %	10	9	1
celkem	25	31	17

Tab. 5. Absolutní a relativní převýšení sídelních jader nad nejbližším vodním tokem.

Mezi oběma částmi tabulky existuje sice základní shoda, avšak i rozdíly, které dokumentují výše uvedené předpoklady. Jednoznačné preference vůči nejnižše položeným částem krajiny pozorujeme u faktoru 3 (doba římská). Malé absolutní převýšení je v tomto případě spolehlivým indikátorem vztahu areálů tohoto období vůči údolním dnům a spodním částem svahů. Tuto závislost se dosavadní badání často pokoušelo vysvětlit např. sušším charakterem klimatu nebo zájmem situovat obytné areály co nejbližše vodního toku. Klimatické vysvětlení je zde však nepřesvědčivé, neboť u malých vodních toků ve sledovaném území nemohlo docházet k větším záplavám při jakýchkoli výkyvech klimatu a kromě toho vidíme, že ve stejných polohách jsou situovány i areály jiných období, byť ne tak často. Sama vzdálenost od vodního zdroje není také argumentem, neboť ta se v zásadě u areálů různých období neliší (tab. 4).

Rozptyl areálů dalších dvou faktorů ukazuje na užitečnost měření převýšení oběma uvedenými způsoby. Zatímco na základě absolutního převýšení by se mohlo zdát, že tendence zaujímat nejvýše položené polohy je nejsilnější u faktoru 2 (doba halštatská a laténská), vztah k relativnímu převýšení připisuje tuto tendenci nejvýrazněji pro faktor 1 (mladší a pozdní doba bronzová). Zejména v mladší a pozdní době bronzové tedy můžeme předpokládat, že obytné areály byly záměrně situovány do míst v rámci dostupné plochy nejvýše položených; pro dobu halštatskou a laténskou je tento záměr méně průkazný. Halštatské a laténské období (faktor 2) ovšem charakterizuje určitá dvojnásobnost, neboť jejich areály jsou nejpočetněji zastoupeny v obou krajních polohách. Tento fakt odráží zřejmě

vztah halštatského období k ostrožnám (na kterých se často vyskytují i nálezy laténské), ale zároveň i k polohám na okraji údolní nivy, kde často předjíhá areály laténského a římského období.

Přestože rozdíly mezi jednotlivými obdobími ve vztahu k převýšení jsou nesporné, závažným problémem je prokázání intencionálnosti umístění areálů v každém ze sledovaných období. Záměrnost v poloze sídelních jader můžeme prokázat, porovnáme-li jejich pozorované rozdělení v jednotlivých kategoriích převýšení s rozdělením „očekávaným“. Za „očekávané“ můžeme považovat takové rozdělení, které by vzniklo při náhodném umístění sídelních jader do krajiny: četnosti výskytu sídelních jader v jednotlivých třídách převýšení by pak odpovídaly plošnému rozsahu těchto tříd. Určitým problémem je zde ovšem vymezení celku, definujícího rozsah a vzájemný poměr jednotlivých tříd. Za potenciálně osidlitelný prostor můžeme považovat pás podél vodních toků do vzdálenosti 300m nebo 500m (tab. 4). Plochy, odpovídající jednotlivým úrovním relativního převýšení v takto definovaných celcích, ukazuje tab. 6:

Kategorie	Pás 300m (km ²)	Pás 300m (%)	Pás 500m (km ²)	Pás 500m(%)
0 – 25 %	28.61	40.96	32.61	31.48
25 – 50 %	15.06	21.56	18.45	17.81
50 – 75 %	14.06	20.13	23.56	22.74
75 -100 %	12.12	17.35	29.02	28.01
celkem	69.85	100.00	103.64	99.99

Tab. 6. Rozsah ploch s různým relativním převýšením nad nejbližším vodním tokem.

Na základě celkového počtu sídelních jader v každém období a procentuálního zastoupení tříd převýšení můžeme stanovit „očekávané“ četnosti výskytu areálů (např. u faktoru 1 je očekávané rozdělení areálů cca 10, 5, 5, 5, zatímco pozorované rozdělení je 3, 6, 6, 10). Statistickým testem chí-kvadrát můžeme vypočítat, s jakou pravděpodobností je tento rozdíl výsledkem náhody, resp. určitého záměru v sídelním chování. Rozdělení areálů faktoru 1 je jasně nenáhodné, vezmeme-li jako základ plochu pásu do 300m (pravděpodobnost náhody je menší než 0.5%); v rámci širšího pásu, v němž podstatně narůstá plocha s největším převýšením, je pravděpodobnost náhody větší (10-25%). U sídelních jader faktoru 2 je záměrnost uspořádání neprokazatelná, u faktoru 3 je naopak velmi pravděpodobná, a to v obou případech.

Z uvedeného příkladu můžeme uzavřít, že během sledovaného období dochází k významné změně sídelního chování. Třebaže rozsah osídleného území se během celého období víceméně nemění, dochází v rámci sídelních areálů k nenáhodným posunům obytných jader, případně i k dílčím posunům sídelních areálů (např. na druhý břeh vodního toku apod.). Na konci doby bronzové můžeme sledovat výraznou preferenci relativně nejvyšších poloh v dostupné krajině. Tato tendence je patrná i ve starších obdobích pravěku (eneolit, starší doba bronzová) a může mít několik příčin: od potřeby vizuální kontroly co největší části sídelního areálu (polí apod.) po důvody symbolické. Nejméně pravděpodobné se zdá vysvětlení strategickými důvody, neboť pak by bylo nelogické, proč právě v obdobích, kdy mobilita populací a možnost vzájemných střetů zřejmě rostou (halštat, latén, řím), dochází k opouštění takto situovaných areálů.

Ke změně sídelního chování dochází postupně, v průběhu doby halštatské a laténské, přičemž jasně odlišné principy výběru obytného areálu zjišťujeme až v době římské. Zdá se, že ani v tomto období nejde o pouhé vymizení jakéhokoli (ekonomického nebo symbolického) významu, který poloze obytného areálu připisovaly předchozí pravěké kultury (což byla původní hypotéza, srov. Kuna a Adelsbergerová 1995; Neustupný 1995b, 51), ale rovněž o určitý záměr, jehož přesný smysl nám dosud uniká. Za zmínku ovšem stojí i to, že na areály doby římské často navazují areály následných kultur, přinejmenším až do střední doby hradištní.

11.7 Dodatek: Poznámky k vstupu dat do GIS a tisku map

Do podoby digitálního obrazu lze geografická či archeologická data převést různým způsobem. Geografická data jsou reprezentována body (např. výškopisné kóty), liniemi (např. vodní síť, vrstevnice) nebo polygony (sídla, lesy, typy půdy). Digitalizaci těchto dat je zpravidla nutno provést pomocí digitizéru, na němž je každý objekt mapy obtažen a uložen do paměti počítače i s přiděleným identifikátorem (hodnotou). Tento postup vyžaduje příslušné vybavení (digitizér, software: v ARÚ Praha je k digitalizaci map používán PC ARC/INFO) a určité zkušenosti; pro archeologa je proto většinou výhodnější, nechat si takové mapy digitalizovat na objednávku.

Nejvýznamnější, ale i nejnákladnější geografickou mapovou vrstvou je digitální výškopisný model. Podklad pro vytvoření tohoto modelu lze nejlépe vytvořit vektorizací tiskových podkladů Základních map (vrstevnic), které jsou k dispozici v Kartografickém úřadu v Praze (pro měřítko 1:200000 až 1:10000), a to buď v tištěné podobě nebo v podobě digitálního obrazu sejmutého skenerem. Vektorizaci (tj. převod tištěného či naskenovaného obrazu do podoby vektorově zapsaných linií s příslušnými identifikátory) může provést některá ze specializovaných firem. Vektorizace jednoho mapového listu ZM 1:10000 (s vrstevnicemi v intervalech 2m) stojí, i s poplatkem za použití mapového díla, cca 2000-4000.– Kč. Jelikož Kartografický úřad sám také provádí podobnou vektorizaci Základních map, bude zde možné v příštích letech zakoupit již hotový DEM (nebo vektorizované vrstevnice jako podklad k němu). Vektorizaci vrstevnic z map 1:25000 (vrstevnice v intervalu 5m) pro území celé republiky paralelně provádí i Vojenský topografický ústav v Dobrušce; i zde bude v budoucnu DEM k dispozici.

Z týchž institucí lze získat i tiskové podklady, případně již hotové digitální mapy jiných geografických prvků (vodopis, komunikace, sídla, lesy apod.). Pořízení těchto mapových vrstev je méně nákladné než získání DEMu.

Digitalizace archeologických dat je většinou snazší a může být bez větších problémů provedena archeologem. Výchozím tvarem dat je databáze archeologických nalezišť (akcí), ve které je každému nalezišti (akci) přidělen jeden nebo několik párů souřadnic, definujících jeho polohu a tvar. Podobné databáze na archeologických pracovištích již běžně existují a mohou být tedy, po provedení obsahové revize, víceméně rovnou převáděny do GIS. Poloha archeologických nalezišť je zpravidla vyjádřena různými typy souřadnic, např. v systému JTSK, S-42, ve stupních zeměpisné šířky a délky, nebo způsobem zavedeným v ARÚ Praha, udávajícím vzdálenost v milimetrech od západní a jižní sekční čáry mapového listu Základních map. V současné době existují konverzní programy, které umožňují vzájemně převádět všechna z těchto prostorových určení. Uživatel tedy může zvolit kterýkoli ze souřadnicových systémů a do něj převést prostorová určení všech nalezišť; pracovat však samozřejmě lze i v libovolně jinak definované souřadnicové síti (např. přiřadíme-li rohu zkoumaného území nulové souřadnice a vzhledem k nim určíme polohu každého naleziště). Z hlediska napojení archeologických dat na geografické údaje je ovšem nejvýhodnější použít některou z kilometrových sítí, jejichž užívání je v ČR běžné (S-42, JTSK).

Z databáze, která obsahuje jednotné prostorové určení nalezišť, automaticky vytvoříme textový soubor, který lze procedurou *arcidris* převést do vektorového souboru v GIS IDRISI a z něj pak vytvořit rastrovou mapu (*polyras*). Textový soubor tohoto typu musí obsahovat identifikátor naleziště, počet bodů, které (v případě nalezišť definovaných jako polygony) dané naleziště definují, a jejich souřadnice (tab. 7). V případě malých souborů lze takový textový soubor vytvořit i ručně.

Tištěné výstupy mapových vrstev zajišťuje v GIS IDRISI procedura *paint*, která komunikuje s barevnými inkoustovými tiskárnami (např. řady Hewlett-Packard). Protože tištěné barvy neodpovídají přesně barvám obrazovky, je zpravidla nutné (zejména při záměru reprodukovat tištěnou mapu v publikaci) barevné odstíny pro tisk pečlivě zvolit a upravit. Barvy mapy jsou definovány jako kombinace tří základních barev (červené, zelené a modré), z nichž každá může mít intenzitu 0 až 63 bodů. Skladba základních barev v jednotlivých barevných třídách 16stupňové škály palety IDRISI je ukázána v tab. 8. Stanovením jiného vzájemného poměru základních barev (IDRISI pracuje buď se 16 nebo 256 barevnými třídami) modifikujeme odstíny barev mapy. Uživatel může definovat řadu vlastních barevných „palet“, které lze uložit v počítači jako samostatné soubory a které mohou (na obrazovce i při tisku) podle potřeby nahradit standardní palety GIS IDRISI.

NAL_BODY.TXT	NAL_POLY.TXT	
10 3469000 5560000	10 5	11 5
11 3469000 5559000	3469000 5560000	3469000 5559000
12 3468000 5559000	3469100 5560000	3469100 5559000
end	3469100 5559900	3469100 5558900
	3469000 5559900	3469000 5558900
	3469000 5560000	3469000 5559000
	end	end
		end

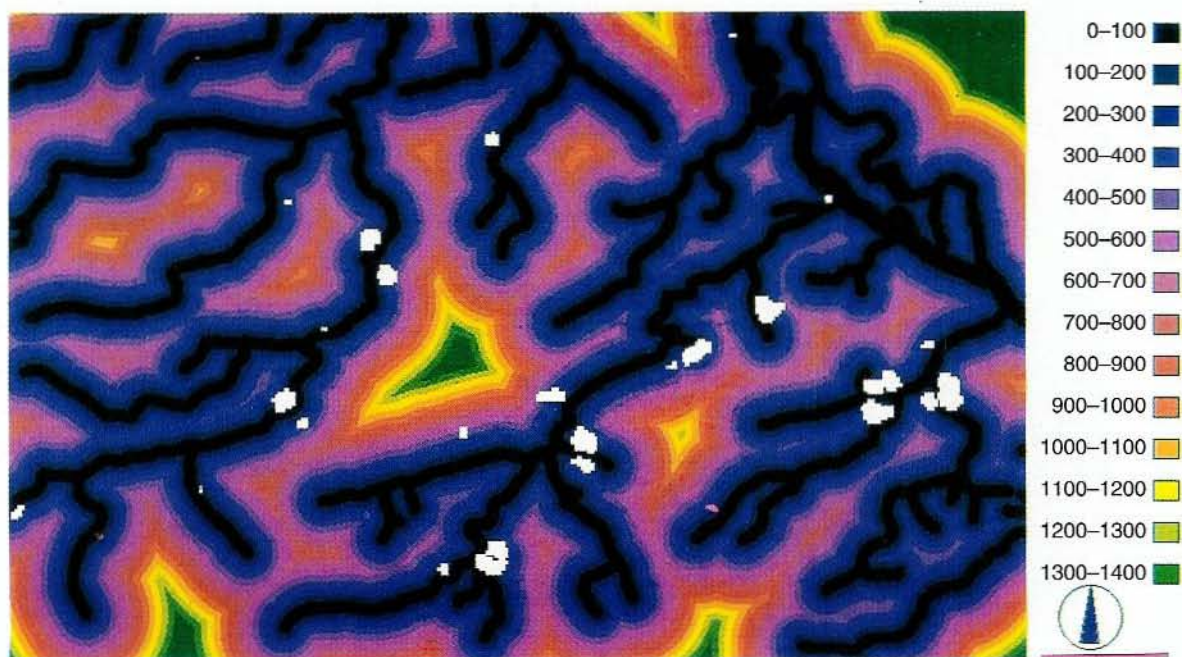
Tab. 7. Formát textových souborů pro převod databáze do vektorového souboru v GIS IDRISI. Soubor NAL_BODY.TXT vymezuje naleziště jako body. Prvním číslem v každém řádku je identifikátor (číslo) naleziště, následují souřadnice X a Y. Souřadnice jsou udány v metrech, v síti S-42. Soubor je uzavřen řetězcem „end“ na posledním řádku. Soubor NAL_POLY.TXT vymezuje každé naleziště jako polygon (v daném případě čtverec 100x100m). Čísla na prvním řádku znamenají identifikátor naleziště a počet bodů, kterými je naleziště vymezeno. Následují řádky se souřadnicemi jednotlivých bodů obvodu polygonu (v daném případě rohů čtverce); první a poslední bod jsou vždy totožné. Každý polygon je ukončen řetězcem „end“; celý soubor je uzavřen řetězcem „end“ na dvou řádcích.

Hodnota	červená	zelená	modrá	popis barvy
0	0	0	0	černá
1	0	0	23	modročerná
2	0	0	47	tmavomodrá
3	0	0	63	modrá
4	23	0	63	světle modrá
5	47	0	63	tmavofialová
6	63	0	63	fialová/růžová
7	63	0	47	světle fialová
8	63	0	23	světle červená
9	63	0	0	červená
10	63	20	0	hnědá
11	63	46	0	okrová
12	63	63	0	žlutá
13	47	63	0	světle zelená
14	23	63	0	zelená
15	0	47	0	tmavozelená
	63	63	63	bílá

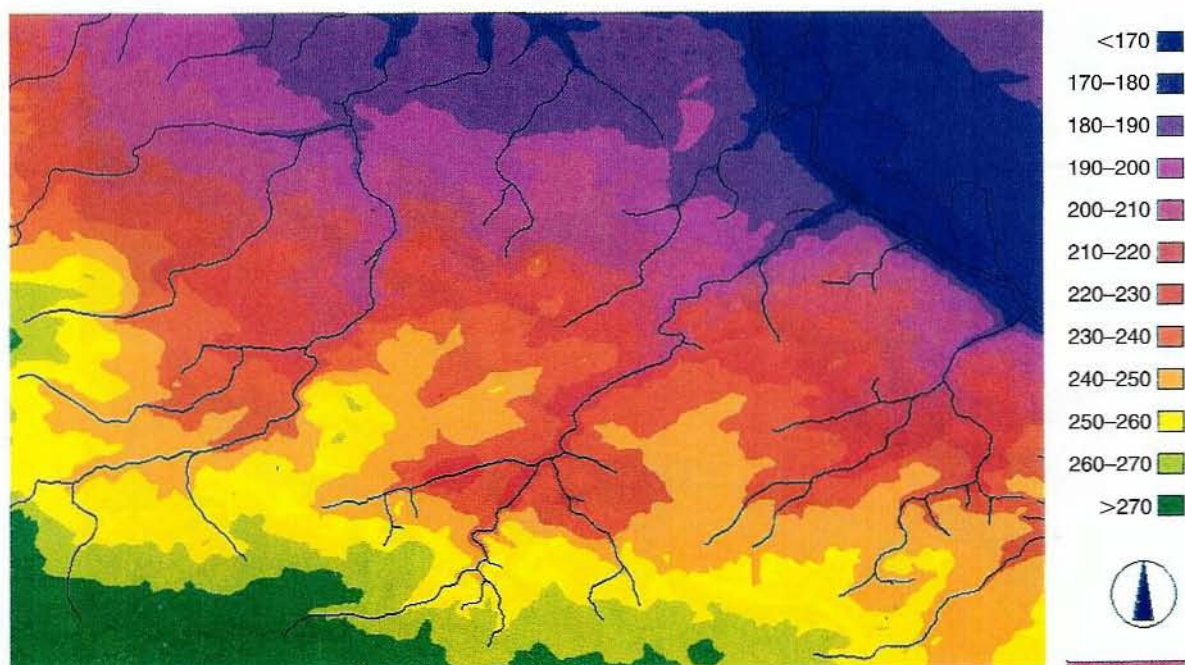
Tab. 8. Skladba barev 16stupňové škály v GIS IDRISI.

11.8 Literatura

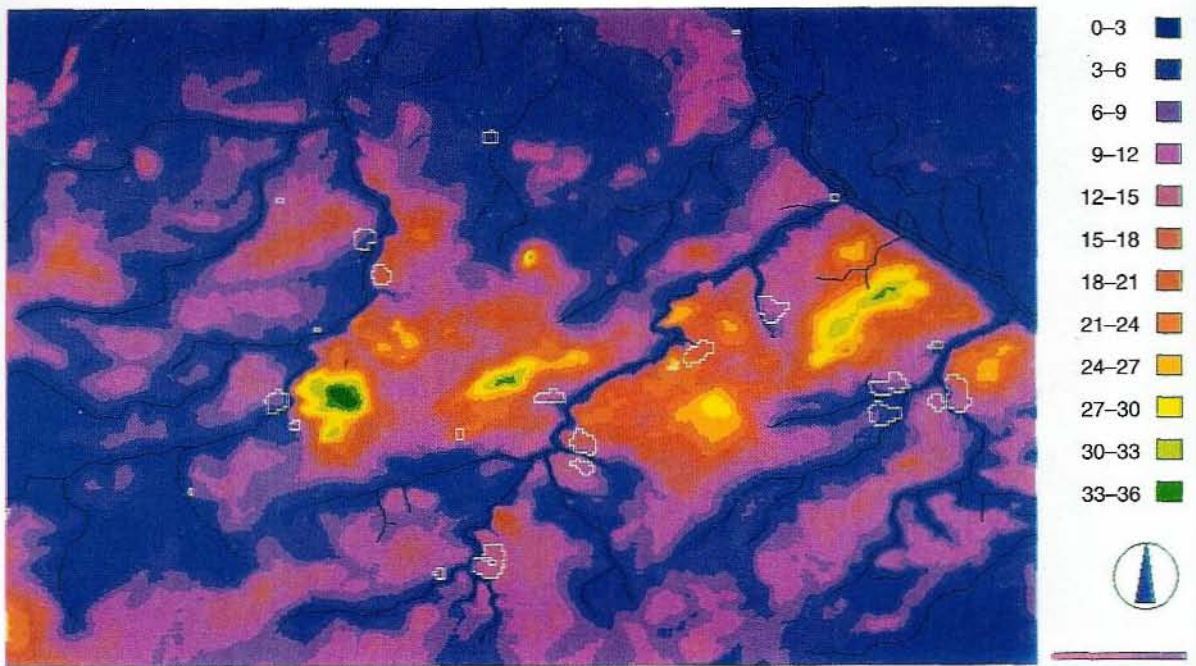
- Allen K.M.S., Green S.W., Zubrow E.B.W. (eds) 1990: *Interpreting space. GIS and archaeology*, New York – London – Philadelphia (Taylor & Francis).
- Dreslerová D. 1995: A socio-economic model of a prehistoric micro-region, in: Kuna M., Venclová N. (eds), *Whither archaeology? Papers in honour of E. Neustupný*, Praha (ARÚ), 145-160.
- Eastman J.R. 1992: *IDRISI*, version 4.0, Worcester, Mass. (Clark University).
- Gaffney V.L., Stančič Z. 1991: *GIS approaches to regional analysis: a case study of the island of Hvar*, Ljubljana.
- Gojda M. 1993: Gaffney V. L., Stančič Z., *GIS approaches to regional analysis: a case study of the island of Hvar*, Ljubljana 1991, *Památky archeologické* 84, 164-165 (recenze).
- Kuna M. 1996: *GIS v archeologickém výzkumu regionu: vývoj pravěké sídelní oblasti středních Čech*, *Archeologické rozhledy* 48, v tisku.
- Kuna M. a Adelsbergerová D. 1995: *Prehistoric location preferences: an application of GIS to the Vinořský potok project, Bohemia*. In: Lock, G. a Stančič, Z. (eds), *Archeology and Geographical information systems*, London (Taylor & Francis), 117-131.
- Kvamme K.L. 1995: *A view across the water: the North American experience in archaeological GIS*. In: Lock, G. and Z. Stančič (eds), *Archeology and Geographical Information Systems: a European Perspective*. London (Taylor & Francis), 1-14.
- Lock G. and Z. Stančič (eds), *Archeology and Geographical Information Systems: a European Perspective*. London (Taylor & Francis).
- Neustupný E. 1993: *Archeological method*. Cambridge (CUP).
- 1994: *Předvídaní minulosti* *Technický magazín* 11, 58-60.
- 1995a: *Beyond GIS*. In: Lock, G. and Z. Stančič (eds.), *Archeology and Geographical Information Systems: a European Perspective*. London (Taylor & Francis), 133-139.
- 1995b: *Pravěk. Lidé v dějinách* (učebnice pro 2. stupeň ZŠ), Praha (Fortuna).
- 1996a: *Databáze nalezišť okresu Chrudim*, *Archeologické rozhledy* 48, 126-134.
- 1996b: *Polygons in archaeology*, *Památky archeologické* 87 (v tisku).
- 1996c: *Syntéza struktur formalizovanými metodami (vektorová syntéza)*, v tisku.
- Neustupný, E. a Vencl, S. 1995: *Formal methods at Hostim*. In: S. Vencl, *Hostim. Magdalenian in Bohemia*, *Památky archeologické – Supplementum* 4, 205-224.
- Neustupný E. a Venclová N. 1996: *Využití prostoru v latěnu: region Loděnice*, *Archeologické rozhledy* 48, v tisku.



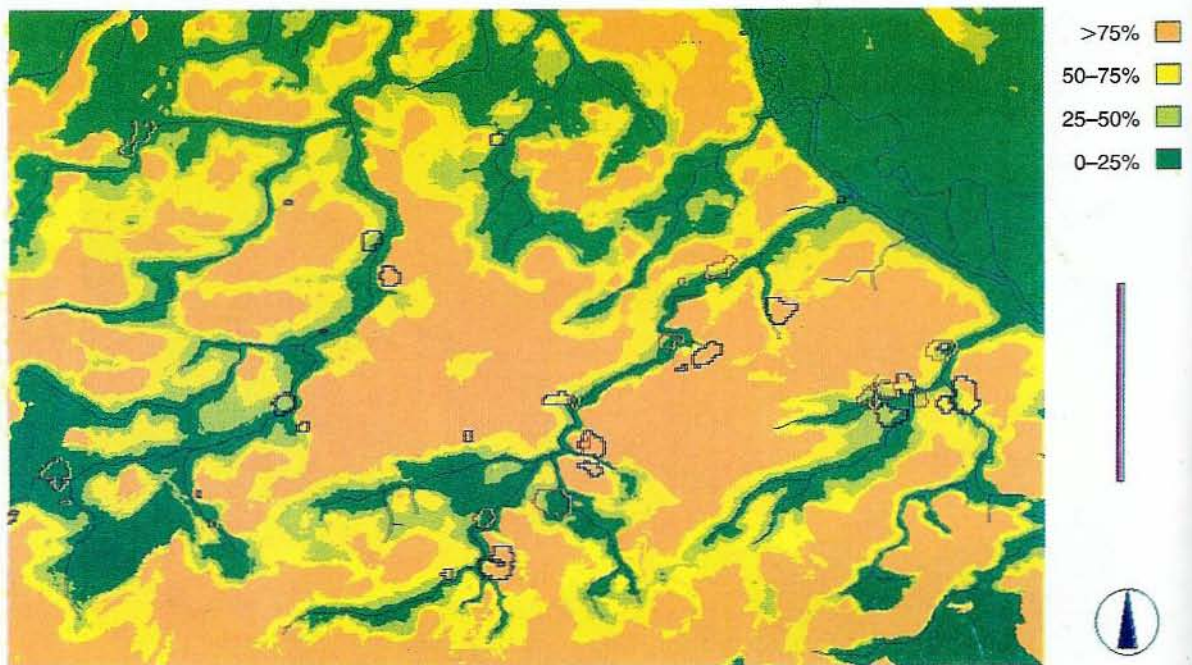
Obr. 5.: Vzdálenost od vodního toku (v metrech). Bíle vyznačena sídelní jádra mladší a pozdní doby bronzové.
Délka měřítka: 2 km.



Obr. 6. Digitální výškopisný model území (nadmořská výška v metrech). Délka měřítka: 2 km.



Obr. 7. Mapa převýšení (v metrech). Bílými liniemi vyznačena jádra sídelních areálů mladší a pozdní doby bronzové. Délka měřítka: 2 km.



Obr. 8. Mapa relativního převýšení (v procentech). Černými liniemi vyznačena jádra sídelních areálů mladší a pozdní doby bronzové (faktor 1); červenými liniemi jádra areálů doby římské (faktor 3). Délka měřítka: 3 km.

Statistika a archeologie

12. Statistická analýza archeologických dat (Zdeněk Weber)

12.1 Úvod

Archeologická data mají značnou cenu danou jednak obvykle nákladným způsobem jejich získávání, jednak povahou potenciální informace v datech obsažené. Přirozeně tak vzniká značný tlak na kvalitu jejich analýzy. Pojem statistika říká něco jiného teoretickému matematikovi a úplně něco jiného pracovníkovi statistického úřadu. Minula doba velkých sálových počítačů, kde nad stohy skládaného papíru, potištěného výpisy programů a nepřehlednými výstupními sestavami se skláněli nejméně tři lidé. Matematik, který nerozuměl počítači, programátor, který nerozuměl statistice a odborník – zákazník, pro kterého se to všechno dělalo a který obvykle příliš nerozuměl ani jednomu ani druhému. Ten pak rozhodoval jako Werichův „pekařův císař“ – ... my mu nerozumíme, ale my mu věříme... – a samozřejmě za to platil. Vznikly tak rozsáhlé knihovny matematických a statistických programů, pomocí kterých se buď ve zdrojovém nebo předkompilovaném tvaru přímo na místě sestavovaly konkrétní aplikace.

S bouřlivým rozvojem malých osobních výpočetních prostředků se postupně ztrácelo výsadní postavení statistiků a programátorů při analýze dat. Odborníci si uvědomili, jak je pro ně samé výhodné, když si mohou zpracovat výsledky svých měření a výzkumů sami a také je graficky prezentovat. Vznikají firmy z vědeckých a výzkumných ústavů a univerzit, které vyvíjejí a prodávají statistický software nikoliv již univerzální, ale specificky oborově zaměřený. Na přelomu 70. a 80. let spolu s růstem kvality hardware se projevuje výrazná diferenciací vlastností statistických programů, vyžadovaná trhem. Tehdy ale vzniká vážné nebezpečí, se kterým se u běžného software (např. textové editory) nesetkáme. Tímto nebezpečím je ve statistice *nepoučený uživatel*. Ten může snadno založit závažné rozhodnutí na základě:

- volby nesprávné metody statistické analýzy, která poskytne nesmyslné výsledky,
- nesprávné interpretace správných výsledků.

Protože uživatelské manuály nebývají dostačující, je třeba věnovat pozornost základnímu statistickému vzdělání uživatelů.

12.2 Statistika pro archeology

Pod pojmem *statistika* obvykle rozumíme vědu o zjišťování, zpracování a rozboru (hodnocení a výkladu) číselných údajů, shromažďovaných buď k popisu rozsáhlých souborů nebo k redukci rušivých odchylek, vyvolávaných náhodnými činiteli. Vypracovává metody založené na předpokladu, že zjišťované údaje jsou realizací náhodných veličin a že účelem je bližší určení neznámého parametru v jejich rozdělovacím zákonu. Je založena na *počtu pravděpodobnosti*.

Předkládaná skripta jsou především určena posluchačům archeologie, kteří nemají dostatečnou matematickou průpravu, ale přesto budou potřebovat statisticky zpracovat a vyhodnotit velký objem kvalitativních i kvantitativních dat získaných excerpcí odborné literatury, z vlastních měření artefaktů, nebo z databází terénních výzkumů. Cílem vedle pochopení podstaty statistických metod a smyslu jimi získaných výsledků, je také orientace ve výběru nejvhodnějších postupů a poznání počítačové podpory statistické analýzy dat pomocí standardního i speciálního software (viz navazující stat').

V současné době citelně chybí moderní počítačově orientovaná učebnice statistiky pro archeology. Ve většině dosavadních učebnic jde buď o

- *obecnou teorii statistiky* aplikovanou převážně na sociálně-ekonomické oblasti, nebo o
- *matematickou statistiku*, tedy část matematiky, zabývající se úlohami ocenění charakteristik pravděpodobnostních rozdělení na základě modelů náhodných pozorování.

Archeologům je dnes v době širokého využívání výpočetní techniky zapotřebí spíše znalostí z *teorie analýzy dat* – tedy speciálních částí aplikované statistiky, zabývající se

- *způsoby výběru, uspořádání a formalizace obvykle vícerozměrných dat do datových souborů a databází,*
- *metodami jejich optimálního zpracování,*
- *správnou interpretací a prezentací výsledků.*

Sylaby minimálních znalostí ze statistických metod, plynoucích z těchto požadavků:

1. Základní statistická terminologie. Hlavní prostředky statistiky. Zdroje archeologických dat. Měření, stupnice, data.
2. Jednorozměrné archeologické soubory. Základní operace s nimi. Statistické charakteristiky. Teoretická rozdělení.
3. Některá speciální rozdělení. Odhady parametrů a intervaly spolehlivosti.
4. Testování dat: testy významnosti, shody, extrémních odchylek, náhodnosti; testy parametrické a neparametrické. Vyrovnání dat.
5. Zobrazování a prezentace výsledků statistické analýzy.
6. Vícerozměrná archeologická data a jejich analýza.
7. Náhodné veličiny. Závislost funkční, statistická. Korelace. Kvalitativní data. Transformace kvantitativních znaků na kvalitativní.
8. Prostor znaků. Archeologický prostor. Matice dat. Ověření normality.
9. Klasifikace, taxonomie, uspořádání. Transformace dat.
10. Míry vzdálenosti, podobnosti a rozptýlenosti archeologických objektů.
11. Seskupování objektů, clustery. Seriace.
12. Korespondenční analýza.
13. Shlukovací analýza. Dendrogramy.
14. Faktorová analýza.
15. Prezentace, věrohodnost a interpretace výsledků analýzy obsáhlých vícerozměrných archeologických datových souborů.

Potřebné vědomosti lze získat z následujícího stručného *výběru studijní literatury*:

Reisenauer, R.: *Metody matematické statistiky a jejich aplikace*, SNTL – Práce, 1970

Hanousek, J., Charamza, P.: *Moderní metody zpracování dat – matematická statistika pro každého* (disketa), Grada, Praha 1992

Ajvazjan, S. et al.: *Metody vícerozměrné analýzy*, SNTL, Praha 1981

Řehák, J., Řeháková, B.: *Analýza kategorizovaných dat v sociologii*, Academia, Praha, 1986

Hebák, P., Hustopecký, J.: *Průvodce moderními statistickými metodami*, SNTL, Praha, 1990

Fedorov-Davodiv, G.A.: *Statističeskije metody v archeologii*, Vysšaja škola, Moskva, 1987

Runyon, R.P.: *Nonparametric statistics. A cotemporary Approach.* (ruský překlad: *Spravočnik po neparametričeskoj statistike. Sovremennyj podchod*), Finansy i statistika, Moskva, 1982.

Ihm, P. et al.: *Statistik in der Archäologie*, *Archaeo-Physika*, 9, Rheinland-Verlag, Koeln, 1978.

Fletcher, M., Lock, G.R.: *Digging for Numbers: Elementary Statistics for Archeologists*, Oxford's University Committee for Archaeology Monograph, 33, Oxford, 1991.

Baxter, M.J.: *Exploratory Multivariate Analysis in Archaeology*, Edinburgh University Press, Edinburgh 1994.

ČSN 01 0225 : *Aplikovaná statistika. Testy shody empirických rozdělení s teoretickými* (1981)

ČSN 01 0222 : *Aplikovaná statistika. Testy odlehlosti výsledků pozorování* (1980)

ČSN 01 0223 : *Pravidla stanovení odhadů a konfidenčních mezí pro parametry normálního rozdělení* (1985)

ČSN 01 0104 : *Teorie pravděpodobnosti a aplikovaná statistika. Termíny, definice, označení* (1990)

ČSN 01 0230 : *Aplikovaná statistika. Analýza rozptylu* (1988)

ČSN 01 0250 :Statistické metody v průmyslové praxi (1973)

ČSN 01 0252 :Statistické metody v průmyslové praxi II. Základní neparametrické metody (1975)

ČSN 01 0253 :Statistické metody v průmyslové praxi III.Základní parametrické metody (1976)

Některé poznámky k vybraným problémům statistické terminologie a některých postupů ve statistice používaných.

12.2.1 Prvek, znak, soubor, výběr, základní soubor

Předmětem statistického zkoumání jsou hromadné jevy a to i takové, u nichž není hned patrna numerická podstata. Cílem je pak vytvoření spolehlivých závěrů o povaze sledovaných jevů na základě informací, obsažených v numerických údajích. Zkoumaný objekt v případě **jednorozměrné analýzy** je zastoupen jedinou hodnotou, které obecně říkáme *prvek*. Skupina prvků tvoří *soubor* (statistický soubor). Ten musí mít některé vlastnosti, vyjadřující *hledisko věcné, časové a prostorové*, společné všem prvkům. Každý prvek, kterému chybí třeba jediná vlastnost, znehodnocuje statistický soubor, do kterého byl zařazen natolik, že již není vhodný pro sledování zkoumaného jevu. Vytýčení společných vlastností je základním požadavkem při vytváření statistických souborů. Čím více je společných vlastností, tím více je *soubor homogenní*. Další společnou vlastností všech prvků určitého souboru jsou vyšetřované vlastnosti. Všechny prvky musí mít aspoň jednu sledovanou vlastnost, kterou nazýváme měřený *znak*. Vedle toho mohou (ale nemusí) mít prvky ještě jiné vlastnosti, tzv. *ostatní vlastnosti*, jimiž se blíže nezabýváme, protože nejsou pro řešení našeho problému podstatné.

Sledované znaky běžně dělíme na

- kvantitativní, tj. znaky vyjádřené určitou hodnotou fyzikální veličiny, zjišťované měřením (tíha, rozměry, pevnost, počet střepů jedné nádoby atd.) Předem nutno zvolit jednotky a přesnost měření (0,01 mm; 0,1 kg). Všechny hodnoty je třeba uvádět jednotně, aby nikdo nemusel provádět další početní úpravy.
- kvalitativní (např. barva nádob, druh hrobu, tvar nářadí), které seskupujeme do přirozených skupin (malované střepy, žárové hroby, kulové jámy) a hodnotíme pomocí různých stupnic (škál): barevné spektrum, stupnice tvrdosti.
- alternativní tj. znaky, které mají jen dvě možné obměny: celistvá x rozbitá nádoba, zachovalý x porušený skelet.

Hodnota či kvalita sledovaného znaku se u jednotlivých prvků souboru mění. Znak je tedy veličinou *proměnnou*. Nelze-li na základě určité zákonitosti předem stanovit konkrétní hodnotu proměnné veličiny u daného prvku, je výskyt každé jednotlivé hodnoty proměnné veličiny důsledkem náhodného působení známých i neznámých vlivů. Takovou proměnnou označujeme jako *náhodnou veličinu* (např. hloubka hrobu, jeho rozměry, tloušťka střepu).

Podle druhu informace, které nám znaky o zkoumaných skutečnostech poskytují a jaké úkony či operace umožňují, rozlišujeme:

- nominální znaky – umožňují nám klasifikaci objektů
- ordinální znaky – umožňují porovnání míry intenzity variant objektu
- kardinální znaky – umožňují číselné vyjádření hodnoty existující varianty.

Adekvátně pak dělíme i příslušné pojmy na *klasifikační, komparační a metrické*.

Někdy se rozčleňují znaky do tří kategorií podle toho, jak jejich varianty umožňují používat matematické operace. Rozlišujeme pak znaky *intenzivní* (umožňující jejich klasifikaci) a znaky *extenzivní* (umožňující realizaci numerických operací). Mezi těmito krajnostmi jsou znaky *inextenzivní*. Ty neumožňují numerické operace v pravém slova smyslu (např. vysoký x nízký), ale lze je vhodně transformovat (např. vysoký 10 cm x vysoký 5 cm) a pak realizovat vhodnou numerickou operaci.

V praxi se setkáváme se smíšenými tříděními znaků na

- **kvalitativní čili nominální:** Pomocí nich lze roztrždit soubory (výběry, populace) do jednotlivých kategorií podle variant daného znaku. Varianty těchto znaků však nevyjadřují míru přítomnosti (kvantity) znaku u jeho nositele. Neumožňují posoudit, zda nositel má daný znak ve vyšší míře či stupni než nositel jiné varianty. Uspořádání variant nominálních znaků je libovolné, neboť neznáme kritérium, dle kterého by se dal stanovit kvantitativní vztah mezi variantami, jejich vzdálenost, odstup atd. Z matematických vztahů lze použít pouze identitu (=) a odlišnost (\neq). Pro charakteristiku nominálních dat se nejčastěji používají deskriptivní statistiky (např. modus – tj. nejpočetnější varianta; poměry početního zastoupení variant v souboru buď v absolutních hodnotách, v procentech nebo v indexových číslech). Při *vícerozměrných* znacích však existují postupy, kterými lze kvalitativní znaky promítnout do kvantitativní roviny. Používá se určitého hodnotícího hlediska, speciálních stupnic a pod. (Např. barvy se dají kvantitativně hodnotit pomocí frekvenčních spekter nebo vlnových délek). Nominální znaky podrobněji dělíme na:

- alternativní čili dichotomické – třídí znaky na dvě vzájemně se vylučující varianty;
 - množné čili polytomické – znaky s větším počtem variant. Nutnou podmínkou je vyhraněnost a jednoznačnost variant (nesmí se překrývat).
- Každý prvek daného souboru může patřit jen do *jediné* varianty.

- **kvantitativní** – dělené dále na :

- 1) **pořadové čili ordinální znaky**, které se od nominálních znaků odlišují tím, že jejich varianty jsou uspořádané a to podle intenzity v nich obsaženého daného znaku. Od kardinálních znaků se odlišují tím, že u jejich variant se nesleduje vzájemná vzdálenost, že velikost intervalů mezi nimi není stejná, ani známá resp. číselně vyjádřená.

Ordinální znaky se dále rozdělují podle způsobu, jakým se určuje pořadí jejich variant na:

- *porovnávací* (není žádná předem daná pořadová stupnice, varianty se setřídí dle míry zastoupení (intenzity) v nich sledovaného znaku),
- *zařadovací* (předem se vymezí pořadí variant, tj. zadá se jejich „stupnice“).

Toto rozdělení ordinálních znaků je důležité pro volbu postupu při výpočtu míry statistické závislosti:

- pro porovnávací ordinální znaky se používá Kendallův koeficient τ ,
- pro zařadovací ordinální znaky se použije Spearmanův koeficient pořadové korelace ρ .

- 2) **číselné čili kardinální znaky** – jsou měřitelné znaky, jejichž varianty lze uspořádat dle pořadí, uvádět do číselně vyjádřitelných vztahů – např. číselným vyjádřením velikosti intervalů nebo přímo číselnou hodnotou variant. Pro měření je zapotřebí měřicí soustava, která se obvykle přebírá hotová nebo se někdy musí vypracovat.

Kardinální znaky se dělí na : *intervalové* a *podílové*.

- **INTERVALOVÉ ZNAKY** jsou takové kardinální znaky, které neoznačují jen celkové pořadí variant dle intenzity (stupně) znaku – což je typické pro pořadové znaky – ale přímo velikost intervalu mezi jednotlivými stupni znaku. Rozsah intervalů přitom nemusí být stejný. Umožňují určit o kolik (o jakou číselnou hodnotu) je jedna varianta znaku větší než jiná varianta. Neumožňují však určit, kolikrát je tato varianta větší či menší než varianta druhá (což je specifické právě pro podílové znaky!) Podle míry informace se tedy intervalové znaky řadí mezi znaky pořadové a podílové. Varianty intervalových znaků dovolují počítat z nich aritmetické průměry (což není přípustné u znaků pořadových!) a další charakteristiky: rozložení četností, standardní

odchylku, variační koeficient, korelační koeficient mezi nárůstem jednoho a druhého intervalového znaku.

- *PODÍLOVÉ ZNAKY* se také někdy označují jako *ZNAKY POMĚROVÉ*. Od intervalových znaků se odlišují tím, že jejich varianty se nevyjadřují v libovolně stanovených stupních intervalových stupnic, ale jsou dané v číselných hodnotách soustavy, ve které je známa přirozená *nulová* hodnota. Odtud název těchto znaků, protože jimi lze vyjádřit hodnotu jedné varianty jako *podíl* resp. násobek varianty druhé. Charakter podílových znaků mají i tzv. *SYNTETICKÉ* znaky, jejichž hodnoty získáváme sumací znaků parciálních (např. jako součet pozitivních odpovědí na otázku přítomnosti či nepřítomnosti (presens – absens) jiných parciálních znaků. Zjištění druhu znaku v konkrétních případech je důležité. Např. matematické vyjádření funkční závislosti $y = x^2$ je nepřípustné, nejde-li o podílové znaky!

Vztahy podílových znaků k intervalovým jsou charakterizovány možností *redukce* podílových znaků na intervalové, tyto lze redukovat na pořadové a pořadové znaky pak redukovat na nominální. Tyto redukce bývají někdy účelné, jestliže umožňují jednodušší analytické postupy, úsporu času a práce. Opačný postup není možný. Nominální znaky nelze technicky transformovat na znaky pořadové, pořadové pak na číselné.

Druhy znaků nutno respektovat při **výběru metod** zjišťování statistických závislostí.

- mezi nominálními znaky se statistická závislost určuje pomocí *koeficientu kontingence*,
- mezi pořadovými znaky pomocí *neparametrických metod výpočtu koeficientu pořadové korelace*,
- mezi kardinálními znaky pomocí *parametrických metod výpočtu koeficientu korelace*.

12.2.2 Indikátory a pravidla jejich výběru.

Indikátorem rozumíme takový znak, jež je vybraný z jiných znaků proto, že nejlépe odpovídá studovanému jevu, resp. jeho aspektu či jedné složce, resp. vyjadřuje základní kvalitu zkoumaného jevu. Výběr znaků a jejich volba za indikátory nesmí být jen intuitivní, ale spíše výsledkem uplatnění určitých zásad:

- je nutné soubor znaků přesně vymezit, zvážit jejich důležitost a stupeň adekvátnosti, vyřadit ty znaky, jež jsou již obsaženy v jiných výrazněji vymezených znacích. Z nich volíme indikátory.
- Indikátory mají mít rozlišovací schopnost, t.j. jejich přítomnost či nepřítomnost má vyjádřit mimo jiné i to, zda daný jev patří do kategorie zkoumaných či nikoliv.
- u složitějších jevů, které nelze výrazně vymezit a sledovat pomocí jediného znaku jako indikátoru, volíme více tzv. parciálních indikátorů, ze kterých se potom skládá tzv. *syntetický indikátor*.
- při volbě indikátorů se dává přednost znakům s empirickým významem, t.j. znakům jež vyjadřují „nejjednodušší určení“, tedy v hierarchii „všeobecné – zvláštní – jedinečné“ tomu poslednímu pojmu – jedinečnému.
- dáváme přednost indikátorům, jejichž varianty vyhovují požadavku kvantitativního zjišťování změn a projevů – např. formou měření.
- dáváme přednost znakům ordinálním před nominálními, dále znakům kardinálním před ordinálními resp. nominálními, dále potom znakům podílovým před intervalovými.

V analýze jednorozměrných dat jsou důležité *statistické charakteristiky*, podávající informaci o studovaném datovém souboru. Jsou to :

- charakteristiky polohy – aritmetický průměr (bývá někdy ovlivňován krajními naměřenými polohami) a proto se používá median,
- charakteristiky rozptýlení – rozpětí, rozptyl, směr.odchylka, variační koeficient.

Nezanedbatelné jsou také *testy významnosti*, kterými řešíme problémy ověření určitého předpokladu např. zda zkoumaný výběr pochází ze základního souboru s určitým rozdělením, zda odlehlá hodnota je důsledek hrubé chyby měření atd. Obecný postup vyžaduje:

- volbu hladiny významnosti,
- formulaci nulové a alternativní hypotézy,
- volbu testovacího kritéria a
- interpretaci výsledků.

Druhy testů: jednostranné x oboustranné, parametrické x neparametrické. Pokud nepotřebujeme detailní závěry, stačí pracovat s testem oboustranným. Test neparametrický je výhodný, neznámeli přesně typ rozdělení základního souboru.

12.2.3 Studium vazeb mezi vlastnostmi (znaky) a jejich hodnotami

V archeologii jsou vazby mezi jevy obvykle statistické. Statistika však pouze prokazuje existenci vazeb, určuje její sílu (těsnost). Interpretace smyslu a příčin vazeb je plně záležitost archeologická.

Obvykle se rozlišují případy vazeb:

1. mezi kvantitativními znaky,
2. mezi kvalitativními znaky, jež se dají uspořádat,
3. mezi kvalitativními znaky, které se uspořádat nedají,
4. mezi kvantitativními a kvalitativními znaky. Pro tento případ je nutno buď přiřadit kvalitativním znakům vid (tvar) kvantitativních (tedy je uspořádat), nebo kvantitativním znakům cestou rozdělení jich do intervalů jim přiřadit vid (tvar) uspořádaných znaků kvalitativních.

Mimo to je nutné rozlišovat sledování vazeb :

- a) mezi libovolnou hodnotou jednoho příznaku a libovolnou hodnotou dalšího příznaku.,
- b) mezi všemi hodnotami jednoho příznaku a všemi hodnotami dalšího příznaku,
- c) jediného příznaku s několika dalšími.

Studium vazeb již souvisí s tzv. **vícerozměrnou statistikou, vícerozměrnými soubory, vícerozměrnou analýzou.**

Základním podkladem jsou *datové matice* typu $n \times p$, kde řádky odpovídají jednotlivým studovaným objektům, sloupce potom jednotlivým zjišťovaným znakům. O úspěchu analýzy takové matice rozhoduje, jak se podaří vyjádřit měřitelnými statistickými znaky jednotlivé vlastnosti, o které se zajímáme. Právě přechod od víceméně neurčitého komplexu vlastností k seznamu přesně vymezených měřitelných znaků je složitý proces, pro který lze těžko předem poskytnout obecně platné návody.

Pokud se v datové matici vyskytuje časový prvek, musíme uvažovat o *datovém komplexu*, který navíc nemusí být úplný. Dostáváme potom serii různých typů analýz včetně časových řad a stochastických procesů.

Mnohdy se množina objektů člení na k podmnožin, tedy datová matice se rozpadá **horizontálně**. Objekty ve skupinách jsou si z určitého hlediska podobné – např. tím, že mají shodnou hodnotu některého uvažovaného faktoru. Pokud je takové rozčlenění do skupin dáno, je cílem porovnání skupin mezi sebou – **analýza rozptylu**. Někdy je nalezení vhodného rozčlenění souboru naopak cílem analýzy – **shluková (cluster) analýza**.

V nesymetrických úlohách zkoumání závislostí mezi proměnnými nemají všechny proměnné stejný význam. Často uvažujeme p_1 faktorů, které ovlivňují nebo mohou ovlivňovat p_2 výsledných

proměnných. Datová matice se pak rozpadá **vertikálně** na submatici faktorů **X** typu $n \times p_1$, a submatici výsledných proměnných **Y** typu $n \times p_2$. Velmi často se vyskytují případy :

- $p_1 = 1, p_2 = 1$ – jednofaktorová jednorozměrná analýza rozptylu, jednoduchá regrese,
- $p_1 > 1, p_2 = 1$ – vícefaktorová jednorozměrná analýza rozptylu, vícenásobná regrese, běžná diskriminační analýza .

Datové matice typu $m \times p$, kde $p = 2$ jsou tak zvané **kontingenční** tabulky se dvěma vstupy. Pokud se jedná o kvantitativní znaky označujeme tabulku jako **korelační**.

Závěrem se stručně zmíníme o některých **typech tzv. multivariační analýzy**:

- *faktorová analýza* vyjadřuje vztah mezi sledovanými proměnnými jako výsledek působení jiných proměnných – faktorů. Z určení lineárních korelací mezi proměnnými se matematicky konstruují faktory, kterými lze vysvětlit korelace mezi proměnnými. Předpokládá se obvykle, že tyto korelace jsou lineárními kombinacemi faktorů, jež jsou společné všem proměnným. Kromě toho se mohou určovat i faktory specifické, jež jsou určeny pouze jedinou proměnnou.
- *analýza komponentní* (hlavních komponent) – je druhem faktorové analýzy, kde pojem faktoru je nahrazen komponentou. Odlišuje se od ní tím, že extrahované komponenty mají vysvětlit nikoliv korelace, ale rozptyl původních proměnných.
- *analýza kánonická* – má za cíl určit maximální lineární korelace mezi dvěma skupinami objektů.
- *analýza Q* – typ fakt.analýzy, kdy vstupem je korelační matice typu $n \times n$, jejímiž prvky jsou korelační koeficienty mezi vzorky.
- *analýza R* – opět je typem fakt.analýzy, kdy vstupem je korelační matice typu $p \times p$, jejímiž prvky jsou korelační koeficienty mezi proměnnými.

Všechny typy těchto analýz jsou početně náročné operace a jsou rozumně realizovatelné jen pomocí vhodných statistických software pro PC. O tomto více v následující kapitole.

12.3 Statistické programové vybavení počítačů

Do statistického programového vybavení počítačů zahrnujeme:

- statistické pakety (balíky)
- tabulkové kalkulátory (spreadsheets)
- speciální grafické programy
- databáze.

V dalším se zaměříme přednostně na ta programová vybavení pod operačním systémem Windows (jsou oproti systému DOS více propracované, lépe podporované a uživatelsky příjemné), která mají reálnou použitelnost ve specifickém vědním oboru, jakým je archeologie a jsou obecně dostupná. Podrobněji rozebereme prvé dva body přehledu s ohledem na maximální dostupnost tohoto vybavení. Speciální grafické programy se statistickými operacemi jsou používány hlavně pro analýzu elektrických signálů a v archeologii nemají zatím žádné široké uplatnění. Pomineme také podrobný rozbor programů výhradně určených pro tvorbu, správu a exploataci databází. Ty lze podstatně jednodušeji a efektivněji vytvářet a vyhodnocovat pomocí tabulkových kalkulátorů.

12.4 Statistické pakety

Zatím nejdále je **WinBASP** ve verzi 5.36 (prosinec 1996). Je to **Bonnský archeologický statistický paket** pro **Windows**, který vznikl přepracováním praxí prověřeného programu **BASP** pro **DOS** (poslední verze 4.5.) Jsou to nekomerční specializované programy, nabízené formou shareware na Internetu, které je možné po vyzkoušení legalizovat [1]. Program je průběžně testován velkou skupinou odborníků a zjištěné nedostatky jsou operativně v nejbližší následující verzi odstraněny. Je zaměřen výhradně na *vícerozměrnou statistickou analýzu archeologických dat*, jako je korespondenční analýza, shlukovací analýza, seriace, faktorová analýza, sociální a prostorová analýza, práce s mapami (geografické shlukování, geografická seriace, GIS aj.) a další. Podrobněji o **WINBASP** pojednáme jinde.

Ve grafickém vybavení paketů je zřejmý trend návratu statistiky ke grafické analýze zpracovávaných dat. Rozhodnutí, kterou statistickou metodu nutno zvolit, je stále na odborníkovi, i když pakety se snaží toto rozhodování co nejvíce usnadnit. Proto mu nabízejí přípravu dobrých a přehledných podkladů. Např. tabulka o velikosti 5 x 10 polí ve své původní podobě přehledná není. Statistik musí znát charakter zpracovávaných údajů – tedy jedná-li se o např. časovou řadu, o 5 časových řad, o 10 objektů s 5 měřenými vlastnostmi, o 5 krát deset měřených hodnot nebo o data k vícerozměrné regresi. Nechá si proto nejdříve údaje adekvátně zobrazit a teprve na základě získaného grafu se kvalifikovaně rozhodne o dalším postupu.. Vedle možnosti vykreslování grafů, demonstrujících zpracování dat určitou metodou (regresní křivky, korelogramy, periodogramy aj.), nabízí pakety také trojrozměrnou grafiku (3 D).

Pro jednorozměrná a dvojrozměrná statistická data z oblasti archeologie je vhodný statistický software **WINKS** verze 4.2 (TexaSoft, Texas, USA). Je to nová windows verze jeho předchůdce **KWIKSTATU**, mnoha cenami oceněného programu pro statistickou analýzu dat a generování grafů, pracujícího pod operačním systémem **DOS**. **WINKS** implementuje různé standardní statistické testy a vytváří jejich vhodné grafické prezentace. Oba statistické programy lze nalézt na Internetu [2] jako shareware a po vyzkoušení legalizovat. Popíšeme dále stručně možnosti **WINKSu**.

Hlavní přednosti tohoto paketu statistických programů pro uživatele:

- možnost výběru vhodné procedury na analýzu dat
- metodicky dobře zvládnuté popisy a postupy analýzy dat
- vkládání dat nebo jejich používání v běžných formátech např. dBase, Lotus 1-2-3, ASCII
- poskytnutí kompletní analýzy na jednom místě, takže uživatel nemusí spouštět další programy na realizaci analytických mezikroků
- interpretace výsledků analýzy v tiskové i grafické formě, aby měl uživatel usnadněno rozhodování na jejich základě. Interpretace je např. formou upozornění, že nejsou splněny předpoklady, za kterých lze zvolený výstup vykonat.

Mezi klady tohoto programu patří Průvodce analýzami a velmi podrobný asi 60 stránkový **Manuál** a **Help**. Menu **WINKS** se v ničem neodlišuje od struktury jiných programů pod **Windows**. Mezi položkami souborového menu (**File**) jsou submenu pro založení nové databáze dat (**New database**), otevření již existujících databází (**Open database**), smazání databáze (**Delete database**), dále submenu s různými utilitami (**import**, sestavy, třídění v databázi, výstup dat) a pro správu poznámkových sešitů (**Journal**), nastavení tiskárny (**Printer setup**). Menu „Analyse“ umožňuje výběr aktivního analytického modulu. V oblasti statistiky to jsou moduly popisné statistiky (**Descriptive statistics**), grafy a diagramy (**Graph/Charts**), t-testy (**t-tests**), neparametrické porovnávání (**Non-parametric comparisons**), lineární a nelineární regrese, korelace, frekvenční analýza, chi-kvadrát, jednocestná i dvojecestná **ANOVA**, analýza časových řad (**Time serie analysis**), vícenásobné porovnávací procedury atd.

Mezi různými typy grafů najdeme histogramy, četnost výskytu jevu v dvourozměrné souřadné soustavě, 3 D grafy, koláče a další. Pod Menu je na obrazovce umístěna nástrojová lišta, kde se nachází mimo obvyklé i tři netypické ikony: Repeat (opakuje poslední analytickou proceduru) a Journal (sešit). Ten se obvykle využívá na vytvoření kolekce výstupů analýz z WINKS procedury do jednoho souboru, který se později dá vytisknout nebo importovat do textového editoru. Třetí ikonou X se aktivuje aktuální přepočítávání dat a zobrazení histogramu.

12.4.1 Praktická demonstrace funkce programu WINKS

Máme otevřít databázi, zobrazit popisné statistiky pro jednu proměnnou a nakonec analýzy zobrazit histogram. Podrobné statistické analýzy reprezentuje v Nástrojích (Tools) ikona X. Po kliknutí na ni se zobrazí dotaz, zda chceme databázi vytvořit, otevřít nebo tuto volbu zrušit. Otevřeme příkladovou databázi Example.dbf. Otevření databáze aktivuje dialog, kde si vybereme jméno analyzované položky – např. informaci o věku(Age). Po chvíli se objeví popisné statistiky z dat v položce Age. Okno textového výstupu (View Results) předloží výsledky v textové podobě, které lze uložit do souboru přes menu Save. Vhodný graf dané analýzy se objeví po stisku tlačítka Graph. Obdržíme histogram dat, která se nachází v otevřené databázi pod položkou Age. Typ grafu si můžeme změnit volbou některého tlačítka na horním okraji výstupního okna (Counts – počet výskytů, NormDist – přepínání mezi frekvenční a distribuční funkcí).

Kdybychom neměli v uvedeném případě vytvořenu žádnou databázi, volili bychom položku menu File-New database. WINKS nám tehdy nabídne výběr z několika druhů různě strukturovaných databází.

Pro analýzy dat se obvykle používá volba „Independence Group t-test or ANOVA. Pak se zobrazí tabulkový kalkulátor se dvěma položkami Group (informace o číselném kódu testu) a Obs(obsahuje položku, jež se bude zkoumat). Po napsání hodnot vybereme z menu Analyse položku t-test/ANOVA a pokračujeme jako v příkladě s existující databází.

Při rozpacích si může vždy zavolat nápovědu (HelpMenu) kliknutím na ikonu otazníku (?). Ta obsahuje velký počet různých instruktivních příkladů aplikací a také podrobný seznam hesel v Obsahu. Kliknutím vyvoláme pomocné podrobné popisy vhodných postupů či významů voleb.

12.4.2 Doporučený postup při statistické analýze archeologických dat.

(Má obecnou platnost pro vyhodnocování dat a není omezen pouze na popisovaný statistický paket Winks)

a) Vyberete si potřebnou úlohu v následující nabídce:

1. Chcete použít popisnou statistiku (Descriptive Statistics) jedné proměnné
2. Chcete použít popisnou statistiku pro dvě závisle proměnné
3. Chcete porovnávat dvě proměnné – nezávislé nebo párové
4. Chcete porovnávat více než dvě proměnné – nezávislé nebo závislé
5. Chcete zkusit asociaci mezi dvěma proměnnými
6. Chcete zkusit asociaci mezi více než dvěma proměnnými
7. Definice některých v tabulkách použitých pojmů.

b) Podle příslušného čísla vaší volby přejděte na dále uvedenou tabulku. Ta vám nabídne potřebný modul pro správnou realizaci zvoleného záměru:

Ad 1) Popisná statistika a grafy. Použitelné procedury.

PROMĚNNÉ	DRUH DAT	MODUL	PROCEDURY
jedna	normální	statistický	průměr, odchylka, histogram, konfidenční interval, rámeček, 5ti ciferná data
jedna	nenormální	statistický	medián, rámeček, histogram, 5ti ciferná data
jedna	kategorická	křížových tabulek	frekvence, piktogram
jedna	časová		grafy časových serií

Ad 2) Popisná statistika a grafy. Použitelné procedury.

PROMĚNNÉ	DRUH DAT	MODUL	PROCEDURY
dvě závislé	normální	statistický nebo regresní	Pearsonův korelační koeficient, 2 D graf
dvě závislé	nenormální	statistický nebo regresní	Spearmanův korelační koeficient, 2 D graf
dvě závislé	kvalitativní	křížových tabulek	Křížové tabulky, 3 D graf

Ad 3) Srovnávací testy pro dva soubory

PROMĚNNÉ	DRUH DAT	MODUL	PROCEDURY
dvě závislé	normální	t-test/ANOVA	párový t-test
dvě závislé	nenormální	neparametrický	Friedmanův test
dvě závislé	dichotomní	křížových tabulek	Mc Nemarův test
dvě nezávislé	normální	t-test/ANOVA	ANOVA
dvě nezávislé	nenormální	neparametrický	Mann - Whitneyův U-test
dvě nezávislé	kvalitativní	křížových tabulek	Chí-kvadrát (homogenita)

Ad 4) Srovnávací testy pro více souborů

PROMĚNNÉ	DRUH DAT	MODUL	PROCEDURY
více závislých	normální	t-test/ANOVA	opakovaná ANOVA
více závislých	nenormální	neparametrický	Friedmanova ANOVA
více závislých	dichotomní	neparametrický	Cochranův Q-test
více nezávislých	normální	t-test/ANOVA	ANOVA
více nezávislých	nenormální	neparametrický	Kruskal-Wallis
více nezávislých	kvalitativní	křížových tabulek	Chí-kvadrát

Ad 5) Asociační test mezi dvěma soubory

PROMĚNNÉ	DRUH DAT	MODUL	PROCEDURY
závislé	normální	regresní	Pearsonova korelace, jednoduchá lineární regrese
závislé	nenormální	regresní	Spearmanova korelace
závislé	kvalitativní	křížových tabulek	Chí-kvadrát (nezávislost)
závislé	smíšená	regresní	Spearmanova korelace

Ad 6) Asociační test pro více souborů

PROMĚNNÉ	DRUH DAT	MODUL	PROCEDURY
více závislých	normální	regresní	vícenásobná regrese
více závislých	nenormální	nelineární analýzy	Kendalova částečná pořadová korelace
více závislých	kvalitativní	nelineární analýzy	diskriminační analýza

Ad 7) Definice některých pojmů.

Normální data jsou data dobře aproximovatelná Gaussovou distribucí.

Nenormální data jsou data, která nejsou normálně rozdělena.

Kategorická data jsou nominální data např. žena-muž, hnědý-modrý-zelený.

Kvantitativní data jsou numerická data, označovaná jako kardinální proměnné.

Dichotomická data jsou kategorická data, jež nabývají pouze dva stavy (ano/ne, jev existuje/neexistuje).

Kvalitativní data popisují obvykle přívlastky (barva předmětu, socioekonomickou třídu, sex).

12.5 Tabulkové kalkulátory

označované někdy jako tabulkové procesory (např. velmi rozšířený MS Excell) jsou také vybaveny řadou statistických nástrojů a velmi bohatou grafikou. V mnoha případech mohou sloužit jako „víceúčelový nástroj“ pro tvorbu databází dat a jejich následnou analýzu a grafickou prezentaci.

Tak např. **MS EXCELL v. 5.0 i 7.0** obsahuje celkem 19 komplexních procedur pro analýzu dat, z nichž 13 je vyloženě statistických a zbývající se mohou použít pro úpravy a zkoumání převážně numerických dat. Aktivujeme je z menu **Nástroje/Analýza dat**. Pokud v menu **Nástroje** položku **Analýza dat** nenaleznete, je možné si je dodatečně doinstalovat. Spustíte si instalační program a instalujete doplněk **Analytické nástroje**. Pro naše účely si uvedeme stručný popis všech procedur (podrobnosti o každé z nich naleznete v **Nápovědě**):

ANOVA – JEDEN FAKTOR.

Jednoduchá analýza rozptylu s testem hypotézy, že střední hodnoty několika náhodných výběrů jsou si rovny. Hladina významnosti se uvažuje implicitně 0,05 – je možné ji upravit. Vstupní oblast musí mít minimálně dvě souvislé oblasti. Výsledky lze uložit na nový list nebo do nového sešitu. Analýza představuje rozšíření testů pro střední hodnoty dvou výběrů, jako je t-test. Lze ji použít i pro náhodné výběry s rozdílným počtem hodnot.

ANOVA -DVA FAKTORY S OPAKOVÁNÍM.

Provádí jednoduchou analýzu rozptylu pro více skupin dat a pro každou skupinu je více náhodných výběrů. Hladina významnosti je implicitně 0,05 – je ji možné upravit. Vstupní oblast musí mít minimálně dvě souvislé oblasti. Výsledky lze uložit na nový list či do nového sešitu.

ANOVA – DVA FAKTORY BEZ OPAKOVÁNÍ.

Provede analýzu rozptylu pro dvojné třídění, kdy se v jednom vzorku nepoužívá více než jeden výběr. Testuje hypotézu, že střední hodnota dvou nebo více náhodných výběrů jsou shodné. Hladina významnosti, podmínka vstupní oblasti a uložení výsledků stejně jako v předchozím heslu.

KORELACE.

Měří vztah mezi dvěma množinami dat a to tak, že se neprojevuje závislost na použitých jednotkách měření. Korelační koeficient se počítá jako kovariance dvou množin dat, dělená součinem jejich směrodatných odchylek. Vstupní oblast musí mít minimálně dvě souvislé oblasti. Výsledky se uloží na nový list nebo sešit. Korelace umožňuje zjištění těsnosti statistické závislosti mezi dvěma soubory dat. Korelační koeficient je kladný, je-li závislost přímá a záporný, je-li závislost nepřímá. Pokud jsou hodnoty v obou množinách nezávislé, bude hodnota korelace blízká nule. Je-li vzájemná závislost souborů lineární, je korelační koeficient 1 nebo -1. Dva shodné soubory mají korelační koeficient roven jedné.

KOVARIANCE.

Slouží ke zjištění, jak na sobě závisejí dva náhodné výběry. Je citlivá na jednotky měření použité ve výběrech. Vstupní oblast musí mít minimálně dvě souvislé oblasti. Výsledky lze uložit na nový list nebo do nového sešitu. Kovariance je kladná při přímé závislosti a záporná, je-li závislost nepřímá. Jsou-li hodnoty v obou výběrech na sobě nezávislé, bude hodnota kovariance blízká nule.

POPISNÁ STATISTIKA.

Vygeneruje soustavu univariantních statistických hodnot. Hladina významnosti se uvažuje implicitně 0,05 a lze ji upravit. Vstupní oblast může mít více souvislých oblastí dat ve sloupcích. Výsledky se uloží na nový list nebo do nového sešitu. Po označení položky „Celkový přehled“ se vygeneruje tabulka s těmito statistickými charakteristikami pro každý sloupec zvlášť: *střední hodnota, standardní chyba střední hodnoty, medián, modus, směrodatná odchylka, rozptyl výběru, špičatost, šikmost, rozdíl hodnot max-min, nejmenší hodnota, největší hodnota, součet hodnot v souboru, počet prvků, k-tá největší hodnota, k-tá nejmenší hodnota, věrohodnost při zadané hladině významnosti (implicitně 95 %).*

Tato analýza slouží jako výchozí pro další navazující procedury.

F-TEST.

Dvouvýběrový test pro porovnání rozptylu dvou datových souborů. Hladina významnosti se uvažuje implicitně 0,05 a je možné ji upravit. Na vstupu jsou dva sloupce nebo řádky pro dva soubory dat. Výsledky lze uložit na nový list nebo do nového sešitu. Výsledkem analýzy je tabulka obsahující: *střední hodnotu, rozptyl, počet pozorování, rozdíl, F, P ($F <= 1$) a Fkrit* pro každý soubor zvlášť.

HISTOGRAM.

Je v podstatě frekvenční analýzou dat. Rozdělí soubor do tříd (skupin). Umožňuje generovat normální histogram, tříděný histogram (tzv. pareto) a kumulativní procentní podílový histogram. Na výstupu může být mimo histogramů též graf.. Výsledky lze uložit na nový list nebo do nového sešitu. Číslo se započte do skupiny, pokud je menší nebo rovno horní hranici třídy a nepatří do předchozí skupiny.

POŘADOVÁ STATISTIKA.

Vytvoří tabulku, která bude obsahovat pořadové číslo a procentní pořadí pro každou hodnotu v množině dat. Výsledky lze uložit na nový list nebo do nového sešitu a to volitelně do řádků nebo sloupců. Pomocí tohoto nástroje je možné realizovat relativní zařazení hodnot v souboru dat.

VZORKOVÁNÍ.

Vytvoření reprezentativního vzorku dat z rozsáhlého souboru podle zvolené vzorkovací metody (periodická nebo náhodná), vzorkovací periody a počtu vzorků. Výstupní tabulku se vzorkem dat můžeme uložit na nový list nebo do nového sešitu. Tato analýza se použije, když vstupní soubor je příliš rozsáhlý a nelze jej celý zpracovat nebo vykreslit do grafu. Pro další práci se použije vytvořený vzorek, který má vlastnosti celého souboru. Máme-li podezření na periodicitu dat, pak můžeme vytvořit vzorek, který bude obsahovat údaje jen z určité části cyklu. Např. obsahují-li vstupní data hodnoty pozorování za čtvrtletí, pak vzorkováním s periodou 4 ve výstupní oblasti údaje za stejná čtvrtletí.

PÁROVANÝ t-TEST NA STŘEDNÍ HODNOTU.

Dvouvýběrový párový Studentův t-test pro zjištění, zda střední hodnoty dvou souborů jsou různé (shodnost rozptylů obou souborů se nepředpokládá). Na vstupu jsou dvě oblasti analyzovaných dat a obě musí být buď ve sloupcích nebo v řádcích. Hladina významnosti se uvažuje implicitně 0,05 a lze ji upravit. Dále se zadává předpokládaný (hypotetický) rozdíl středních hodnot (0 znamená, že se předpokládají shodné střední hodnoty). Na výstupu je tabulka, která obsahuje 11 statistických charakteristik pro oba soubory. Výsledky lze uložit na nový list nebo do nového sešitu. Tabulka obsahuje tyto charakteristiky: *střední hodnota, rozptyl pozorování, Pearsonova korelace, hypotetický rozdíl středních hodnot, rozdíl, t-stat, $P(T \leq t)$ (1), t-krit (1), $P(T \leq t)$ (2), t-krit (2)*.

Párový test se používá tehdy, jsou-li pozorování ve výběrech přirozeným způsobem spárována, např. při dvojím testování skupiny – před a po experimentu.

DVOUVÝBĚROVÝ t-TEST S ROVNOSTÍ ROZPTYLŮ

Dvouvýběrový Studentův t-test pro testování shodnosti středních hodnot dvou souborů dat za předpokladu shodnosti jejich rozptylů. Na vstupu jsou dvě oblasti analyzovaných dat a obě musí být buď ve sloupcích nebo v řádcích. Hladina významnosti se uvažuje implicitně 0,05 a je ji možné upravit. Dále se zadává předpokládaný (hypotetický) rozdíl středních hodnot (0 znamená, že se předpokládají shodné střední hodnoty). Na výstupu je tabulka, která obsahuje 11 statistických charakteristik pro oba soubory. Výsledky lze uložit na nový list nebo do nového sešitu.

DVOUVÝBĚROVÝ t-TEST S NEROVNOSTÍ ROZPTYLŮ

Dvouvýběrový Studentův t-test pro testování shodnosti středních hodnot dvou souborů dat za předpokladu, že jejich rozptyly nejsou shodné. Na vstupu jsou dvě oblasti analyzovaných dat a obě musí být buď ve sloupcích nebo v řádcích. Hladina významnosti se uvažuje implicitně 0,05 a je ji možné upravit. Dále se zadává předpokládaný (hypotetický) rozdíl středních hodnot (0 znamená, že se předpokládají shodné střední hodnoty). Na výstupu je tabulka, která obsahuje 10 statistických charakteristik pro oba soubory. Výsledky lze uložit na nový list nebo do nového sešitu.

Test se používá v případě, kdy studované soubory dat jsou různé. Pokud se testuje soubor vzniklý před nějakou akcí se souborem vzniklým po akci, doporučuje se použít párový test.

DVOUVÝBĚROVÝ z-TEST NA STŘEDNÍ HODNOTU

Dvouvýběrový z-test testuje hypotézu o rozdílnosti středních hodnot dvou souborů při známých rozptylech. Na vstupu jsou dvě oblasti analyzovaných dat a obě musí být buď ve sloupcích nebo v řádcích. Dále se musí zadat rozptyl pro oba soubory (ten se zjistí např. z popisné statistiky). Hladina významnosti se uvažuje implicitně 0,05 a je možné ji upravit. Na výstupu je tabulka, která obsahuje 9 statistických charakteristik pro oba soubory. Výsledky lze uložit na nový list nebo do nového sešitu.

EXPONENCIÁLNÍ VYROVNÁVÁNÍ (VYHLAZOVÁNÍ)

Umožňuje vyhlazení průběhu dat s možností předpovědi. Předvídá hodnotu, na základě předpovědi pro předchozí úsek, opravenou o chybu předchozí předpovědi. Vstupní oblast musí tvořit jeden sloupec nebo řádek s minimálně 4 buňkami. Výstupní oblast musí být na stejném listu jako oblast vstupní. Implicitně je nastaven koeficient útlumu 0,3 (prázdný rámeček). Lze zadat požadavek na vygenerování grafu a požadavek na výpočet standardní chyby.

FOURIÉROVA ANALÝZA

Jde o rychlou Fouriérovu transformaci (FFT) pro analýzu periodických dat a lineárních systémů. Umožňuje též inverzní transformaci (pro získání původních dat ze souboru dat již transformovaných). Počet vstupních hodnot musí být mocninou dvou (např. 16, 32, 64, 128 ... max 1024). Menší počet dat, chybějících do nutného počtu, lze nahradit nulami, nebo přebývající data ze souboru vynechat. Výsledky analýzy lze uložit na nový list nebo do nového sešitu.

POHYBLIVÝ (KLOUZAVÝ) PRŮMĚR

Je průměr počítaný ze zadaného pohyblivého intervalu (implicitně 3 hodnoty, lze upravit). Na výstupu může být tabulka a graf (histogram). Tabulku lze doplnit o vypočtené standardní chyby. Vstupní dat musí být uspořádána do jednoho sloupce nebo řádku s minimálně 4 buňkami (políčky). Výstupní oblast musí být na stejném listu se vstupními daty. Klouzavý průměr poskytuje informaci o trendu s vyhlazením výkyvů. Při předpovědi se dává větší váha posledním hodnotám. Předpovědi odrážejí trend, ale opožďují se. Jak narůstá délka pohyblivého intervalu, stávají se předpovědi méně citlivé ke krátkodobým výkyvům a tím pomaleji odrážejí poslední trendy. Pro výpočet váženého pohyblivého průměru (jednoho intervalu) můžeme použít funkci SOUČIN.SKALÁRNÍ.

GENERÁTOR PSEUDONÁHODNÝCH ČÍSEL

Vyplní oblast jednoho nebo více sloupců nezávislými náhodnými čísly, která pocházejí z některého rozdělení. Náhodné proměnné mají řadu použití ve statistických simulacích. Použitím této procedury získáme soubor hodnot charakterizovaný daným pravděpodobnostním rozdělením. (Např. pro hodnocení souboru výšek jednotlivců z kosterních pozůstatků můžeme použít normální rozdělení, soupis pohřebišť určité kultury na lokalitách dané oblasti bude mít Bernoulliho rozdělení). Generátor umožňuje umělé vytvoření souborů pseudonáhodných čísel pro jeden ze následujících 7 typů rozdělení:

Rovnoměrné Je charakterizováno dolní a horní hranicí. Čísla jsou vybírána se stejnou pravděpodobností z celého zadaného intervalu. Obvykle se používá rovnoměrné rozdělení v intervalu 0...1.

Normální Je charakterizováno střední hodnotou a směrodatnou odchylkou. Obvykle se používá střední hodnota rovna 0 a směrodatná odchylka rovna 1.

Bernoulliho Je charakterizováno pravděpodobností výskytu v daném výběru. Výsledná hodnota je 0 nebo 1.

Stejného rozdělení by bylo možné dosáhnout například následujícím postupem: použijeme náhodné rovnoměrné rozdělení na intervalu 0...1. Když hodnota náhodné proměnné bude menší nebo rovna pravděpodobnosti výskytu, přiřadíme Bernoulliho náhodné proměnné hodnotu 1; jinak jí přiřadíme hodnotu 0.

Binomické Je charakterizováno pravděpodobností výskytu v jednom výběru a celkovým počtem výběrů.

Stejného rozdělení by bylo možné dosáhnout například tak, že bychom náhodné proměnné přiřadili hodnotu získanou generováním daného počtu Bernoulliho náhodných hodnot a jejich sečtením.

Poissonovo Je charakterizováno hodnotou lambda rovnou převrácené hodnotě střední hodnoty. Poissonovo rozdělení se často používá pro charakterizaci počtu nějakých událostí, které se odehrají za časový úsek.

Podle vzorků Je charakterizováno dolní a horní hranicí, krokem, počtem opakování každého čísla a počtem sekvencí (neboli kolikrát se má opakovat celá posloupnost).

Diskrétní Je charakterizováno vstupní oblastí se dvěma sloupci. V levém z nich jsou uvedeny hodnoty, v pravém jsou uvedeny odpovídající pravděpodobnosti výskytu. Součet pravděpodobností musí být roven 1.

Obecně pro generátor pseudonáhodných čísel platí, že pokud na vstupu zadáme stejný tzv. „základ (jádro)“, vygeneruje po opětovném spuštění stejnou posloupnost pseudonáhodných čísel. Výsledky lze uložit na nový list nebo do nového sešitu. Generátor umožňuje efektivní ověřování různých modelů na základě vytvořené množiny dat s pravděpodobnostním rozdělením. Pro generování pseudonáhodných čísel uvnitř intervalu 0 a 1 (tj. bez těchto hraničních hodnot) lze také použít funkci NÁHČÍSLO (.). Pro získání náhodných čísel rozložených mezi libovolně zadanými hodnotami užijeme funkci RANDBETWEEN (.).

REGRESE

Provádí lineární *prostou* nebo *vícenásobnou* regresi proložením přímkou množinou pozorování s použitím metody nejmenších čtverců. Na vstupu jsou dvě množiny dat. První množina obsahuje závislá data (Y) a musí být pouze v jednom sloupci. Druhá množina obsahuje data nezávislá (X) a musí mít stejný počet řádků jako množina závislá, počet sloupců X_i je limitován ($i = 1 \dots \max 16$). Lze si vybrat, zda regresní přímka má procházet počátkem (implicitně neprochází) a nastavit úroveň věrohodnosti (implicitně je 95 %, což odpovídá hladině významnosti 0,05). Výstupní tabulka má nejméně 7 sloupců a obsahuje tabulku analýzy rozptylu (ANOVA), koeficienty, standardní chybu odhadu y , hodnoty koeficientu determinance R^2 , počet pozorování a standardní chybu koeficientů. Na výstupu mohou být 3 grafy: *graf s rezidui*, *graf regresní přímky* a *graf pravděpodobnosti*. Výsledky lze uložit na nový list nebo do nového sešitu. Regrese se používá k analýze ovlivnění jedné závislé proměnné hodnotami jedné nebo několika nezávisle proměnných.

Procedura umožňuje velmi jednoduše spočítat *nelineární regresi exponenciální*. Podmínkou je, že vstupní množiny dat obsahují přirozené logaritmy (\ln) závislých i nezávislých dat. Rovněž také výsledky jsou udány jako přirozené logaritmy vypočtených hodnot. Blíže lze nalézt v Nápovědi k funkcím LOGLINREGRESE a LINREGRESE.

Pro úplnost je užitečné zde uvést úplný soupis statistických funkcí, které jsou ve MS EXCEL v.5 a 7 obsaženy. Tyto statistické funkce jsou obvykle používány na množinu čísel a vrací příslušné statistické hodnoty. Nejsou jimi pouze součty a odchylky, ale také některé jednoduché lineární a exponenciální aproximace křivek. Několik z těchto funkcí vrací jako výsledek pole.

Pozor na některé výrazné rozdíly v pojmenování funkcí u verzí programů *lokalizovaných* (převedených do češtiny) a originálních (anglických). Např. PRŮMĚR = AVERAGE, PRŮM.ODCHYLKA = AVEDEV, SMODCH = STDEVP, SMODCH.VÝBĚR = STDEV, POČET = COUNT atd. Tyto rozdíly se také vyskytují i v menu, submenu a nápovědě lokalizovaných programů. V této stati jsou přednostně uvažovány lokalizované verze.

12.5.1 Statistické funkce EXCELLU

D BETADIST	Vrátí hodnotu distribuční funkce rozdělení beta
D BETAINV	Vrátí inverzní hodnotu distribuční funkce rozdělení beta
D BINOMDIST	Vrátí distribuční funkci pro binomické rozdělení
P CORREL	Vrátí korelační koeficient mezi údaji ze dvou polí
P COVAR	Vrátí hodnotu kovariance
T CRITBINOM	Vrátí nejmenší přirozené číslo, pro které má pravděpodobnostní funkce binomického rozdělení hodnotu větší nebo rovnu kritické hodnotě alfa
P ČETNOSTI	Vrátí vertikální pole, které popisuje rozdělení dat do kategorií podle velikosti
D CHIDIST	Vrátí jednostrannou distribuční funkci chí-kvadrát
D CHIINV	Vrátí inverzní jednostrannou distribuční funkci chí-kvadrát
T CHITEST	Provede test nezávislosti
P CONFIDENCE	Vrátí interval věrohodnosti pro střední hodnotu základního souboru
Z DEVSQ	Vrátí součet čtverců odchylek
D EXPONDIST	Vrátí distribuční funkci exponenciálního rozdělení
D FDIST	Vrátí hodnotu distribuční funkce F rozdělení
D FINV	Vrátí hodnotu inverzní funkce k distribuční funkci F rozdělení
P FISHERINV	Vrátí hodnotu inverzní funkce k Fisherově transformaci
P FISHER	Vrátí Fisherovu transformaci čísla x
A FORECAST	Vrátí odhad hodnoty y v bodě x na základě lineární regresní přímky
T FTEST	Vrátí výsledek F-testu
D GAMMADIST	Vrátí hodnotu distribuční funkce rozdělení gama
D GAMMAINV	Vrátí hodnotu inverzní funkce k distribuční funkci rozdělení gama

<i>D</i> GAMMALN	Vrátí přirozený logaritmus gama funkce
<i>Z</i> GEOMEAN	Vrátí geometrickou střední hodnotu
<i>Z</i> HARMEAN	Vrátí harmonický průměr
<i>D</i> HYPGEOMDIST	Vrátí hodnotu distribuční funkce hypergeometrického rozdělení
<i>A</i> INTERCEPT	Vrátí průsečík lineární regresní přímky s osou x
<i>P</i> KURT	Vrátí hodnotu špičatosti rozdělení dané množiny
<i>P</i> LARGE	Vrátí k-tou největší hodnotu množiny
<i>A</i> LINTREND	Vrátí předpověď hodnot při lineární aproximaci (trend)
<i>A</i> LINREGRESE	Vrátí nové hodnoty y po proložení původních hodnot přímkou pomocí metody nejmenších čtverců
<i>D</i> LOGINV	Vrátí inverzní funkci k distribuční funkci logaritmicko normálního rozdělení
<i>A</i> LOGLINREGRESE	Vrátí parametry exponenciální aproximace
<i>A</i> LOGLINTREND	Vrátí předpověď hodnot funkce při exponenciální aproximaci
<i>D</i> LOGNORMDIST	Vrátí distribuční funkci logaritmicko normálního rozdělení
<i>Z</i> MAX	Vrátí maximální hodnotu v seznamu argumentů
<i>Z</i> MEDIAN	Vrátí medián zadaných čísel
<i>Z</i> MIN	Vrátí minimální hodnotu v seznamu argumentů
<i>P</i> MODE	Vrátí modus – nejčastěji se vyskytující hodnotu dané posloupnosti hodnot
<i>D</i> NEGBINOMDIST	Vrátí distribuční funkci negativního binomického rozdělení
<i>D</i> NORMDIST	Vrátí distribuční funkci normálního rozdělení
<i>D</i> NORMINV	Vrátí inverzní funkci k distribuční funkci normálního rozdělení
<i>D</i> NORMSDIST	Vrátí distribuční funkci standardního normálního rozdělení
<i>D</i> NORMSINV	Vrátí inverzní funkci k distribuční funkci standardního normálního rozdělení
<i>P</i> PEARSON	Vrátí Pearsonův korelační koeficient
<i>P</i> PERCENTIL	Vrátí hodnotu, která odpovídá k-tému percentilu v poli hodnot
<i>P</i> PERCENTRANK	Vrátí procentní pořadí čísla x v poli hodnot
<i>P</i> PERMUTACE	Vrátí počet permutací pro zadaný počet objektů
<i>Z</i> POČET2	Vrátí počet neprázdných hodnot v seznamu argumentů
<i>Z</i> POČET	Vrátí počet čísel v seznamu argumentů
<i>D</i> POISSON	Vrátí distribuční funkci Poissonova rozdělení
<i>P</i> PROB	Vrátí pravděpodobnost toho, že hodnoty v oblasti x leží v daném intervalu
<i>Z</i> PRŮMĚR	Vrátí průměrnou hodnotu argumentů
<i>P</i> PRŮMODCHYLKA	Vrátí průměrnou absolutní odchylku bodů od střední hodnoty
<i>P</i> QUARTIL	Vrátí hodnotu kvartilů ze zadaného pole dat
<i>D</i> RANDBETWEEN	Vrátí náhodné číslo ze zadaného intervalu
<i>P</i> RANK	Vrátí pořadí argumentu (podle velikosti) v seznamu čísel
<i>P</i> RKQ	Vrátí druhou mocninu Pearsonova korelačního koeficientu pro lineární regresi
<i>D</i> SKEW	Vrátí šikmost rozdělení náhodné veličiny
<i>P</i> SLOPE	Vrátí směrnici lineární regresní přímky proložené zadanými body
<i>P</i> SMALL	Vrátí k-tou nejmenší hodnotu v poli
<i>Z</i> SMODCH.VÝBĚR	Vrátí odhad směrodatné odchylky základního souboru určený z náhodného výběru
<i>Z</i> SMODCH	Vrátí směrodatnou odchylku základního souboru určený z náhodného výběru
<i>P</i> STANDARDIZE	Vrátí normalizovanou náhodnou veličinu s normálním rozdělením
<i>A</i> STEYX	Vrátí standardní chybu při výpočtu lineární regrese
<i>D</i> TDIST	Vrátí hodnotu distribuční funkce Studentova t-rozdělení
<i>D</i> TINV	Vrátí inverzní funkci k TDIST
<i>P</i> TRIMMEAN	Vrátí střední hodnotu datové posloupnosti
<i>T</i> TTEST	Vrátí pravděpodobnost spojenou se Studentovým t-testem
<i>Z</i> VAR.VÝBĚR	Vrátí odhad rozptylu základního souboru
<i>Z</i> VAR	Vrátí rozptyl základního souboru

D WEIBULL	Vrátí hodnotu distribuční funkce Weibullova rozdělení
T ZTEST	Vrátí dvoustrannou hodnotu P, která je výsledkem z-testu

Uvedené statistické funkce se dají rozdělit do 5 skupin (viz první sloupec soupisu): **Z** Základní statistiky, **P** Pokročilé statistické funkce, **A** Aproximační funkce, **D** Distribuční funkce, **T** Testovací funkce. Statistické funkce se v EXCELU používají obdobně jako ve většině programů. Napíšeme jméno funkce (malými anebo velkými písmeny), levou závorku, argumenty funkce oddělené čárkou a uzavírací pravou závorku. Argumenty funkcí mohou být čísla, řetězce znaků, odkazy na buňky nebo jiné funkce (tzv. vnořování funkcí je možné až do sedmé úrovně). Argumenty – typ a rozsah – zjistíme v Návodě. Jestliže argument funkce není ve správném rozsahu, pak funkce vrátí chybu #NUM!; jestliže argument funkce není správného typu, vrací se chyba #VALUE!. Pro vkládání funkcí do buněk lze výhodně použít ikony Průvodce funkcí na liště nástrojů.

Poznámka k DATABÁZÍM v tabulkovém kalkulátoru.

Pod **DATABÁZÍ** rozumíme pravoúhlu spojitou oblast buněk v listu tabulky. Každý řádek tvoří záznam, každý sloupec pole. První řádek obsahuje názvy polí. Statistické zpracování takových dat v EXCELLU, se provádí pomocí *databázových funkcí*, které se používají na hodnoty vybírané z databáze s použitím výběrové podmínky.

12.5.2 Funkce pro správu seznamů a databází

DMAX	Vrátí maximální číslo z vybraných položek v databázi
DMIN	Vrátí minimální číslo z vybraných položek v databázi
DPOČET2	Vrátí počet neprázdných buněk, které splňují daná kritéria
DPOČET	Vrátí počet neprázdných buněk, které splňují daná kritéria
DPRŮMĚR	Vrátí průměrnou hodnotu z vybraných položek v databázi
DSMODCH.VÝBĚR	Vrátí výběrovou směrodatnou odchylku z těch záznamů v databázi, které odpovídají zadaným kritériím
DSMODCH	Vrátí směrodatnou odchylku základního souboru z těch záznamů v databázi, které odpovídají zadaným kritériím
DSOUČIN	Vrátí součin určených položek z těch záznamů v databázi, které odpovídají zadaným kritériím
DSUMA	Vrátí součet položek z těch záznamů v databázi, které odpovídají zadaným kritériím
DVAR	Vrátí výběrový rozptyl položek z těch záznamů v databázi, které odpovídají zadaným kritériím
DVAR.VÝBĚR	Vrátí rozptyl základního souboru položek z těch záznamů v databázi, které odpovídají zadaným kritériím
DZÍSKAT	Vrátí tu položku databáze, která splňuje daná kritéria
SQLREQUEST	Připojí vnější zdroj dat a spustí dotaz ze sešitu
SUBTOTAL	Vrátí souhrn dat v seznamu nebo v databázi

Všechny tyto funkce mají 3 stejné argumenty: jméno databáze, jméno pole a podmínku. Databázové funkce zpracují pouze ty záznamy, které splňují zvolenou výběrovou podmínku. Pomocí příkazů DATA/FILTR/ROZŠÍŘENÝ FILTR lze vybraná data překopírovat do jiné oblasti a tam pak je případně dále samostatně statisticky zpracovat podle výše podrobně uvedených postupů.

12.6 Grafy dat

Jestliže vygenerujeme určitá data je nutné si je vykreslit a dále je analyzovat, třeba zjišťováním, jaký typ křivky představují. MS EXCEL nabízí dva typy grafů: grafy vložené do listu tabulky a grafy na zvláštním listu. Prvé jsou vhodné pro bezprostřední analýzu konkrétních tabulek, druhé slouží pro prezentaci výsledků a jejich dokumentaci. V obou případech je proces tvorby grafu zjednodušen pomocí ikony Průvodce grafem na 4 – 5 kroků. Důležitá je volba typu grafu (nabídka obsahuje 15 typů). Nejužitečnější se jeví typy *graf XY* a *3D povrchový graf*. Graf XY je jediný, který vykresluje skutečný rozsah dat na ose x; ostatní umísťují hodnoty y rovnoměrně rozložené ve směru osy x a proto neposkytují reálnou představu o rozmístění, zvláště jsou-li data ve skutečnosti nerovnoměrná. Graf typu 3D povrchový vykresluje pravouhloú oblast dat jako povrch v trojrozměrném prostoru. Dat pro osu x a y jsou nad a vpravo mřížky pro data na ose z. Graf vykresluje pouze hodnoty na ose z rovnoměrně rozložené podél os x a y. Data x a y jsou využita pouze pro popisy. Všechny grafy lze dodatečně upravovat, doplňovat, transformovat (formátovat). Pokud vadí rovnoměrné rozdělení dat podél os x a y, lze data rozdělit posunutím a vytvořit postačující *drátový* model použitím dvojrozměrného XY grafu.

12.7 Vícerozměrná statistická analýza

Speciálním doplňkem k MS EXCELL do Nástrojů je program **xlSTAT**, nabízený komerčně [3]. Vedle jednoduché a vícenásobné regrese, analýzy variační, kovarianční a diskriminační, nabízí dva programy na shlukovací (cluster) analýzu (metoda centroidní a Wardova), korelační analýzu (pořadovou neparametrickou korelaci podle Kendalla a Spearmana, analýzu metodou hlavních komponent), dále jednoduchou a vícenásobnou korespondenční analýzu, třídění kategorických dat a multidimensionální škálování dle Kruskala. Simuluje také X/Y ploter.

12.8 Závěry

Stat' na základě dostupného software, provozovaného pod operačním systémem Windows podává přehled o využití jejich statistického vybavení pro analýzu archeologických dat.

12.9 Literatura

[1] Internet : <http://www.uni-tuebingen.de/uni/afj/basp.html>

[2] Internet: <http://www.dfw.net/homepage>

[3] Internet: <http://www.kovcomp.co.uk/xlsbroc.html>

13. K využití seriace při datování sídlištní keramiky (Vladimír Salač)

13.1 Úvod

Pojem seriace se v české archeologii zhruba v posledním čtvrtstoletí objevoval a objevuje poměrně často. Vlastních aplikací této metody při zpracování konkrétního materiálu však příliš nemáme.

Východiskem matematického postupu seriace je matice dat. V archeologii jde zpravidla o tabulku, v níž řádek (sloupec) představuje archeologický objekt (hrob, sídlištní objekt, vrstvu apod.) a ve sloupcích (řádcích) jsou obsaženy jeho charakteristiky (např. druhy nálezů z objektů, hrobů apod.). Takováto tabulka se nazývá tabulkou kontingenční. Vlastní vzhled resp. obsah tabulky může být dvojitý, buď tabulka registruje pouze výskyt (1) resp. absenci (0) sledovaných jevů, nebo zaznamenává konkrétní četnosti výskytu (např. kolikrát se v objektu vyskytne střep s rytou vlnicí). V prvním případě jde o zvláštní typ kontingenční tabulky, který v tomto příspěvku budeme nazývat tabulkou absence x presence (dále APT), kdežto název kontingenční tabulka (dále KT) budeme používat u tabulek obsahujících skutečné četnosti.

Úkolem seriace je přerovnávat pořadí řádků a sloupců matice takovým způsobem, aby korelace mezi řádky a sloupci byla maximální. Jestliže tabulka má odpovídající strukturu, srovnají se obsazená pole (popř. pole s nejvyššími četnostmi) na diagonále, popř. v její blízkosti. Úroveň uspořádání obsazených polí v tabulce vzhledem k diagonále vyjadřuje tzv. korelační koeficient, který získáme výpočtem ze všech relativních četností v jednotlivých sloupcích a řádcích. Korelační koeficient může mít maximální hodnotu 1 (jestliže je matice čtvercová a obsazená /nenulová/ pole se nacházejí výhradně na diagonále. Čím více se korelační koeficient blíží k této hodnotě, tím více spolu údaje v řádcích a sloupcích korelují, tj. obsazená pole se nacházejí blíže k diagonále, či jsou přímo na ní.

Seriace lze provádět jak „ručně“, tj. např. tabulku rozstříhat a fyzicky měnit pořadí řádků i sloupců, tak aby obsazená pole ležela poblíž diagonály, nebo početně, tj. měnit pořadí sloupců a řádků pouze teoreticky, vypočítávat hodnoty korelačních koeficientů a hledat nejvyšší z nich. Takovéto postupy lze však použít pouze u nevelkých tabulek, jsou velmi pracné a mohou být i nepřesné. U „ručního“ řazení dokážeme registrovat pouze výskyt či absenci sledovaného jevu, vlastní četnost výskytu zohlednit nedokážeme. U ryze početního přístupu je zase problémem prokázat, že jsme skutečně dosáhli nejvyššího možného korelačního koeficientu, tj. zda by jiné seřazení tabulky nepřineslo koeficient ještě vyšší.

Dnes lze celý proces automatizovat a pomocí příslušného softwaru docílit seřazení KT tak, aby korelační koeficient měl maximální možnou hodnotu, tj. aby obsazená pole byla v tabulce co nejvíce diagonalizována.

Proces seriace i s užívaným počítačovým algoritmem a se základní literaturou podrobně popsal např. P. Ihm (1983; 1990), proto není nutné se mu zde dále věnovat. Podíváme se spíše na některé praktické aspekty spojené s užíváním této metody při řešení otázky chronologie sídlištního materiálu.

Proč se seriace u archeologů těší velké oblibě je zřejmé. Badatelé jsou přesvědčeni, že zadají-li do sloupců správné vlastnosti archeologických objektů, lze pořadí objektů v tabulce získané seriací interpretovat chronologicky, tj. předpokládat, že objekty nacházející se v sousedních řádcích jsou si chronologicky blízké, kdežto objekty na jejich opačných koncích jsou si časové vzdálené.

13.2 Problematika datové základny

Má-li být pořadí objektů (např. řádků) v tabulce chronologicky interpretováno, je nutné, aby byly vybrány takové vlastnosti, které jsou skutečně chronologicky relevantní. Zde se ocitáme v procesu vzniku archeologických dat. Tento proces představuje samostatnou rozsáhlou problematiku, počínaje

kritikou samotných pramenů, přes výběr archeologických situací a nálezů relevantních pro řešení daného problému, až po způsoby zaznamenávání a ukládání provedených pozorování. Posledně se u nás touto problematikou zabýval J. Rulf (1993), na jehož práci zde odkazujeme a problémem se dále nezabýváme.

Je zřejmé, že formálně dokonalý a objektivní matematický postup prováděný automatizovaně pomocí počítače nepřinese správné výsledky, pokud pracujeme s vadnými daty. Podstatnou část chybných dat může a musí eliminovat sám archeolog při získávání a zpracovávání pramenů, tj. při tvorbě dat. Máme na mysli nejen klasickou kritiku pramenů, ale i teoretické úvahy, umožňující správné vytváření dat a jejich zpracování, resp. pomáhající zjistit limity, za kterými práce s daty nemusí přinést objektivní výsledky (na příkladu sídlištní keramiky pojednal takovéto situace E. Neustupný 1996). Tím ovšem proces hodnocení a výběru pramenů (dat) nekončí, neboť v průběhu práce se zpravidla ukáží jejich další nedostatky, které archeolog běžnými pracovními postupy (předem) neodhalí. *Kritika pramenů resp. dat tedy není jednorázovou počáteční pracovní etapou, nýbrž prakticky nekončícím průběžným procesem.*

13.2.1 Data ze sídlišť a pohřebišť

Seriace se dosud nejméně úspěšně aplikuje při chronologii pohřebišť (u nás např. E. Neustupný 1978), při zpracovávání sídlišť bývá užívána zřídka. Důvod je zřejmě způsoben základními rozdíly v samotné prameně resp. datové základně. Předně nálezy z hrobů lze zpravidla považovat za uzavřené nálezné celky: všechny nálezy jsou současné. U sídlištních nálezů toto základní východisko postrádáme. Podstatný rozdíl je i v množství materiálu a sledovaných vlastností. Rozdíl je zpravidla i ve stavu bádání. Jestliže u pohřebišť je již předem nastin chronologie více či méně znám, u sídlišť tomu tak vždy není, tj. jestliže u nálezů z pohřebišť je značná část chronologicky významných prvků známa, u sídlišť bývá nutné je teprve hledat. Významnou odlišností je i množství pramenů resp. nálezů. Zatímco u hrobových celků lze zpravidla poměrně snadno zanést do sloupců tabulky všechny nálezy (např. Mirón 1986), u sídlištního materiálu to bývá velmi obtížné (cf. zpracování sídlišť lengyelské kultury dle vypracovaného kódu; Podborský – Kazdová – Košťuřík – Weber 1977). Ne nepodstatnou odlišností je značně vyšší torzovitost nálezů v případě sídlištního materiálu.

U nálezů ze sídlišť je tedy kritika pramenů, tvorba dat, jejich ukládání a manipulace s nimi podstatně obtížnější, než u nálezů z pohřebišť. Základní obtíž při přípravě dat pro seriaci sídlištního materiálu představuje nutnost výběru pramenů a jejich charakteristik, které hodláme sledovat.

13.2.2 Příklad – data z laténských sídlišť v SZ Čechách

Chronologie laténské sídlištní keramiky v SZ Čechách byla řešena z podstatné části pomocí seriace. Podívejme se tedy, jak v tomto konkrétním případě probíhala práce s daty.

Údaje, které jsou nezbytné pro řešení chronologie sídlištní keramiky se pochopitelně netýkají pouze tohoto materiálu. Je nutné registrovat údaje o regionu odkud soubory pocházejí, stejně jako údaje o sídlišti a objektu, kde byly nalezeny. Proto vznikla soustava popisných systémů, umožňujících jednotný popis a zaznamenávání potřebných pozorování. Tyto údaje byly zpravidla převedeny do numerického kódu a uloženy do systému databází.

Databáze regiony (REG.DBF)

- registruje číslo a název regionu a jeho kód
rozsah: 5 regionů

Databáze lokality (LOK.DBF)

- registruje příslušnost k regionu, číslo lokality, název lokality a typ lokality (dvorec, oppidum, neopevněné sídliště apod.)
rozsah: 22 lokalit

Databáze objektů (OBJ.DBF)

- registruje příslušnost objektu k lokalitě (číslo lokality), číslo objektu (reálné číslo přidělené při výzkumu), typ objektu převedený do numerického kódu (typář cf. obr. 1), slovní pojmenování

typu objektu, jeho délku, šířku a hloubku, objem zahloubeného objektu, počet nalezených střepů, hustotu střepů v objektu (ks/m^3), výskyt nekeramických nálezů, předběžné datování, datování nekeramických nálezů a datování po analýze počítačem

rozsah: 250 objektů

Databáze nálezů/keramiky (NAL.DBF)

- registruje obecné údaje o nálezu (střepu): příslušnost k lokalitě a objektu, vrstvu, identifikační číslo, sektor objektu; dále vlastnosti keramického zlomku – tvar nádoby (typář obr. 2), tvar okraje (typář obr. 3), tvar dna, materiál, typ povrchu, umístění výzdoby a její druh (obojí je možné registrovat dvakrát; typář obr. 4), průměr ústí, max. délku, max. sílu, plochu, váhu, počet totožných jedinců a odkaz na vyobrazení.

rozsah: ca 50 000 střepů, uložených v 15 000 záznamech

V těchto čtyřech databázích propojitelných pomocí společných položek, jsou uloženy informace, které byly považovány za relevantní pro řešení chronologie sídlištní keramiky (schéma cf. obr. 5; cf. též Salač 1993). Nyní je naší logickou snahou vytvořit z dat matici (KT), která se stane předmětem seriace. Vzhledem k množství dat musíme pochopitelně usilovat o automatizované vytváření kontingenčních tabulek, kde řádek bude představovat jeden objekt a ve sloupcích bude zaznamenáván výskyt určitých vlastností keramiky nalezené v objektu.

100	110	120	130	140	150	
POLOŽENICE GRUBENHAUSER						
NUMERICKÝ KÓD SÍDLIŠTNÍCH OBJEKTU V SZ ČECHÁCH NUMERISCHER KODE DER SIEDLUNGSOBJEKTE IN NW BÖHMEN	200	300	400	500	510	
	ZÁHL-STAVBY EINGETIEFTE BAUT.	ZÁSÖBNÍ JÁMY VORRATSGRUBE	HLINKY LEHMGRUBEN	JAMKY-ŽLABY PFÖSTENLÖCHER		
600	610	620	630	640	650	660
PECE ÖFEN					HRNČIŘSKÉ PECE TÖPFERÖFEN	TAVÍCÍ PECE SCHMELZÖFEN
700	710	720	730	740	750	760
JAMY GRUBEN						
800	995		999		TOPENIŠTĚ FEUERSTELLE	
KULTURNÍ VRSTVA KULTURSCHICHT	TORZA OBJEKTŮ OBJEKTRESTE		OSTATNÍ (VÝJIMKY) ANDERE (AUSNAHMEN)			

Obr. 1. Numerický kód sídlištních objektů doby laténské v SZ Čechách.

Již jen ze stručného popisu databází (cf. též obr. 1-4) je zřejmé, že množství charakteristik zachycených v datech vztahujících se k jednomu sídlištnímu objektu je tak veliké, že nutně musíme pracovat s výběrem. Jinak řečeno, ve sloupcích zamýšlené matice není možné registrovat výskyt všech typů a variant sledovaných charakteristik, neboť práce s KT o několika stech sloupců by byla velmi obtížná. Mnohem důležitější ovšem je, že naším cílem není sledovat všechny zjištěné charakteristiky, neboť je zřejmé, že všechny nemají stejný chronologický význam (cf. např. chronologický význam položek tvar okraje a síla střepu). Výběr jednotlivých charakteristik (např. tvar nádoby, tvar okraje

apod.) které se mají objevit ve sloupcích KT, lze v dnešních databázových systémech snadno zajistit příslušnými příkazy. Zásadní potíží však způsobuje archeologův požadavek, nesledovat pouze jednotlivé *charakteristiky* keramiky, ale především jejich kombinace, které nadále nazýváme *vlastnostmi*.

Např. mísa se zataženým okrajem (typy 221, 222, 223; cf. obr. 1) sama o sobě nemá chronologický význam. Je-li však tato mísa vyrobena na kruhu (typ povrchu 100, 110) a je-li zároveň zevnitř (umístění výzdoby 200) zdobená vhlazovanou vlnicí (typ výzdoby 621 atd., viz typář na obr. 4), představuje již chronologicky citlivý nále. Snahou archeologa tedy bude, aby se výskyt mísy se zataženým okrajem vyrobené na kruhu a s vhlazovanou vlnicí na vnitřním povrchu registroval v jediném sloupci.

Problém je však složitější. Jak jsme konstatovali, u sídlištního materiálu ne vždy předem odhadneme chronologickou citlivost určité vlastnosti. Zůstaneme-li u příkladu mísy s vhlazovanou výzdobou, existují hypotézy, že vhlazovaná výzdoba se objevuje nejprve na vnitřku nádob a teprve později na vnějším povrchu. Existuje i předpoklad, že vhlazovaná výzdoba se vyskytuje dříve na keramice točené na kruhu než na nádobách vyráběných v ruce. Vysloven byl však i přesně opačný názor, tj. že hrubé mísy s vhlazováním jsou starší a představují vzory pro mladší keramiku točenou na kruhu. Jen u jediného tvaru, jediné výzdoby, dvou typů umístění a dvou typů povrchu, tedy potřebujeme otestovat všechny vzájemné kombinace, přičemž výsledek, tj. chronologický význam předem neznáme. Např. je možné, že chronologicky citlivá je pouze základní vlastnost, tj. mísa se zataženým okrajem a vhlazování, přičemž technologie výroby ani umístění nejsou vůbec důležité.

Z predešlého odstavce vyplývá, že při tvorbě chronologie sídlištní keramiky je nutné definovat značné množství vlastností. Proto se ukázalo užitečné vytvořit samostatnou databázi vlastností, jejichž chronologický význam chceme testovat (cf. obr. 6). V této databázi jsou zadány pomocí intervalů čísel požadované vlastnosti, tj. kombinace různých jednotlivých charakteristik. Tím odpadá námáhavé opakované zadávání definic jednotlivých vlastností.

Zadání našeho příkladu (mísa se zataženým okrajem vyrobená na kruhu s vhlazováním vně) tedy bude vypadat následovně: kód tvaru nádoby (cf. obr. 1) se musí pohybovat v intervalu <221,223> a zároveň kód povrchu musí být v intervalu <100,199> a zároveň kód umístění výzdoby musí být v intervalu <100,199> a zároveň kód výzdoby (cf. obr. 4) musí být v intervalu <620,639>. Pomocí takovéto definice získáváme jeden sloupec KT. Výpis z jedné dílčí databáze vlastností přináší obr. 6, náš příklad je na posledním řádku. Takovýmto způsobem lze nadefinovat vlastnosti, které chceme mít zastoupené v seriaci, tj. vlastnosti, které mají být obsaženy ve sloupcích matice (KT), která bude vstupovat do seriace.

DO STĚNY LZE VĚPSAT PŘÍMKU IN DIE WAND KANN MAN EINE GERADE HINEINSCHREIBEN		1 OBLOUK V KŘÍVCE STĚNY 1 BOGEN IN DER WANDLINIE		2 OBLOUKY V KŘÍVCE STĚNY 2 BOGEN IN DER WANDLINIE		3 OBLOUKY V KŘÍVCE STĚNY 3 BOGEN IN DER WANDLINIE	
ZÁKLADNÍ DEFINICE Ø ÚSTÍ > V Ø DER MÜN- DUNG > H		Ø ÚSTÍ ≤ V Ø DER MÜN- DUNG ≤ H		Ø ÚSTÍ < V Ø DER MÜN- DUNG < H		Ø ÚSTÍ < V Ø DER MÜN- DUNG < H	
MÍSY SCHÜSSELN	SOUDKY FÄSSCHEN	SOUDKY FÄSSCHEN	MÍSY SCHÜSSELN	MÍSY SCHÜSSELN	HRNCE TÖPFE	POHÁRY POKALE	
200	100	100	200	200	300	600	
211	110	121	221	230	311	321	331
212		122	222		312	322	332
			223				
			224				
			225				
CEDNÍKY 800 - TYPY SEIHER 800 - TYPEN NICHT UNTERSCHIEDEN				SITULY		SITULEN 400	
				410		420	
				LÁHVE		FLASCHE 500	
						510	
				ZÁSOBNICE		VORRATSGEFÄSSE 700	
						710	
						720	
VARIANTY MÍSY ESDVITÉ PROFILACE VARIANTEN DER S-PROFIL SCHÜSSELN		231	232	233	234	NÁDOBY SE ZATA- ZENÝM OKRAJEM GEFÄSSE MIT EIN- BIEGENDEM RAND MĚŘENÍ ÚHLU α MESSUNG DES WINKELS α	

Obr. 2. Numerický kód latéské sídlištní keramiky v SZ Čechách – tvary nádob.

100 SOUDKY - FÄSSCHEN									
	1	2	3	4	5	6	7	8	9
110									
120									
130									
140									

200 MÍSY - SCHÜSSELN									
	1	2	3	4	5	6	7	8	9
210									
220									
230									
240									
250									
260									
270									
280									

300 HRNCE - TÖPFE									
	1	2	3	4	5	6	7	8	9
310									
320									
330									
340									
350									

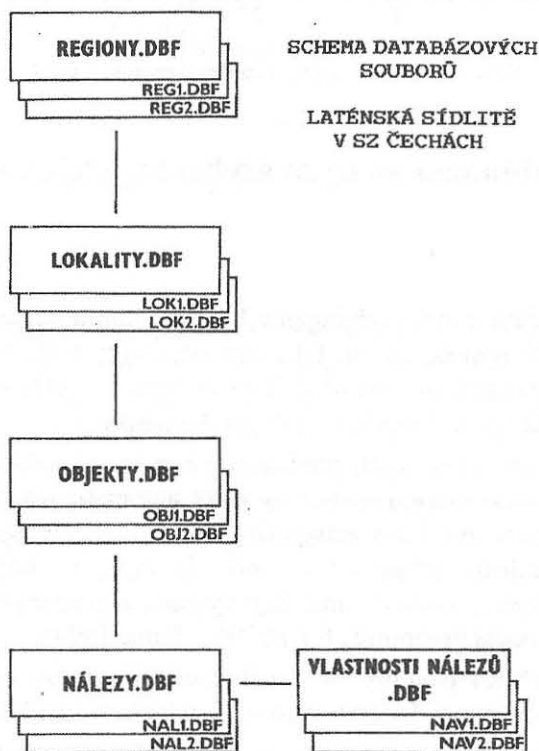
Obr. 3. Numerický kód latéské sídlištní keramiky v SZ Čechách – okraje nádob (část).

400 JEMNÁ VÝZDOBA FEINE VERZIERUNG									
	1	2	3	4	5	6	7	8	9
410									
420									
430									
440									
450									

500 KOLKOVANÁ VÝZDOBA STEMPERVERZIERUNG									
	1	2	3	4	5	6	7	8	9
510									
520									
530									
540									
550									

600 VHLAZOVANÁ VÝZDOBA									
	1	2	3	4	5	6	7	8	9
610									
620									
630									
640									
650									

Obr. 4. Numerický kód latéské sídlištní keramiky v SZ Čechách – výzdoba (část).



Obr. 5. Systém databázových souborů registrujících data vybraná k řešení chronologie latéské sídlištní keramiky.

CIS. VL.	OZNA- GENI	VLASTNOSTI NALEZU															
		TVA NAD OD	TVA NAD DD	TVA OKR OD	TVA OKR DD	TVA DNA OD	TVA DNA DD	POV RCH OD	POV RCH DD	UNI VYZ OD	UNI VYZ DD	DRU VUZ OD	DRU VUZ DD	PRU LST OD	PRU LST DD		
1	T01	100	199	0	0	0	0	0	0	0	0	0	0	0	0		
7	T07	110	219	0	0	0	0	0	0	0	0	0	0	0	0		
11	T11	121	221	0	0	0	0	0	0	0	0	0	0	0	0		
12	T12	122	222	0	0	0	0	0	0	0	0	0	0	0	0		
13	T13	133	233	0	0	0	0	0	0	0	0	0	0	0	0		
14	T14	144	244	0	0	0	0	0	0	0	0	0	0	0	0		
15	T15	155	255	0	0	0	0	0	0	0	0	0	0	0	0		
17	T17	177	277	0	0	0	0	0	0	0	0	0	0	0	0		
20	T20	200	300	0	0	0	0	0	0	0	0	0	0	0	0		
26	T26	260	360	0	0	0	0	0	0	0	0	0	0	0	0		
27	T27	270	370	0	0	0	0	0	0	0	0	0	0	0	0		
30	T30	300	400	0	0	0	0	0	0	0	0	0	0	0	0		
33	T33	330	430	0	0	0	0	0	0	0	0	0	0	0	0		
41	T41	700	799	0	0	0	0	0	0	0	0	0	0	0	0		
43	T43	410	419	0	0	0	0	0	0	0	0	0	0	0	0		
44	T44	420	429	0	0	0	0	0	0	0	0	0	0	0	0		
48	T48	910	910	0	0	0	0	0	0	0	0	0	0	0	0		
49	T01-02	231	231	241	244	0	0	0	0	0	0	0	0	0	0		
50	T02-02	232	232	242	247	0	0	0	0	0	0	0	0	0	0		
51	T03-03	233	233	251	251	0	0	0	0	0	0	0	0	0	0		
52	T04-03	234	234	253	253	0	0	0	0	0	0	0	0	0	0		
54	T05	100	199	130	137	0	0	0	0	0	0	0	0	0	0		
55	T06	100	199	131	137	0	0	0	0	0	0	0	0	0	0		
56	T07	100	199	132	132	0	0	0	0	0	0	0	0	0	0		
57	T08	100	199	134	134	0	0	0	0	0	0	0	0	0	0		
58	T09	100	199	135	135	0	0	0	0	0	0	0	0	0	0		
59	T010	100	199	136	136	0	0	0	0	0	0	0	0	0	0		
60	T011	100	199	138	138	0	0	0	0	0	0	0	0	0	0		
62	T013	700	799	710	719	0	0	0	0	0	0	0	0	0	0		
63	T014	700	799	720	729	0	0	0	0	0	0	0	0	0	0		
65	T019	700	799	730	739	0	0	0	0	0	0	0	0	0	0		
68	TP1	120	129	0	0	0	0	200	249	0	0	0	0	0	0		
69	TP2	220	229	0	0	0	0	100	199	0	0	0	0	0	0		
70	TV1-04	231	232	0	0	0	0	0	0	0	0	410	419	0	0		
71	TV2-04	233	234	0	0	0	0	0	0	0	0	420	429	0	0		
72	TV3	400	499	0	0	0	0	0	0	0	0	230	239	0	0		
73	O1-03	0	280	289	0	0	0	0	0	0	0	0	0	0	0		
74	O2-03	0	340	349	0	0	0	0	0	0	0	0	0	0	0		
83	P03	0	0	0	0	0	0	200	249	0	0	0	0	0	0		
88	P08	0	0	0	0	0	0	250	249	0	0	0	0	0	0		
93	PUV1	0	0	0	0	0	0	100	199	200	299	620	639	0	0		
94	PUV2	0	0	0	0	0	0	100	199	100	199	610	614	0	0		
95	PUV3	0	0	0	0	0	0	100	199	100	199	610	614	0	0		
96	PUV4	0	0	0	0	0	0	100	199	200	299	610	614	0	0		
97	PUV5	0	0	0	0	0	0	200	499	100	199	620	639	0	0		
98	PUV6	0	0	0	0	0	0	200	499	100	199	620	639	0	0		
99	UV1	0	0	0	0	0	0	0	0	100	199	620	639	0	0		
100	UV2	0	0	0	0	0	0	0	0	100	199	610	614	0	0		
101	UV3	0	0	0	0	0	0	0	0	200	299	600	609	0	0		
102	UV4	0	0	0	0	0	0	0	0	200	299	620	639	0	0		
103	UV5	0	0	0	0	0	0	0	0	200	299	620	639	0	0		
112	V09	0	0	0	0	0	0	0	0	0	0	131	131	0	0		
113	V10	0	0	0	0	0	0	0	0	0	0	140	149	0	0		
115	V12	0	0	0	0	0	0	0	0	0	0	160	169	0	0		
116	V13	0	0	0	0	0	0	0	0	0	0	170	179	0	0		
117	V14	0	0	0	0	0	0	0	0	0	0	180	189	0	0		
121	V18	0	0	0	0	0	0	0	0	0	0	210	219	0	0		
124	V21	0	0	0	0	0	0	0	0	0	0	230	239	0	0		
125	V22	0	0	0	0	0	0	0	0	0	0	234	234	0	0		
126	V25	0	0	0	0	0	0	0	0	0	0	310	319	0	0		
129	V24	0	0	0	0	0	0	0	0	0	0	320	329	0	0		
130	V27	0	0	0	0	0	0	0	0	0	0	330	339	0	0		
131	V28	0	0	0	0	0	0	0	0	0	0	340	349	0	0		
134	V31	0	0	0	0	0	0	0	0	0	0	410	429	0	0		
138	V35	0	0	0	0	0	0	0	0	0	0	410	429	0	0		
140	V37	0	0	0	0	0	0	0	0	0	0	430	439	0	0		
142	V39	0	0	0	0	0	0	0	0	0	0	500	599	0	0		
145	V45	0	0	0	0	0	0	0	0	0	0	610	619	0	0		
150	V47	0	0	0	0	0	0	0	0	0	0	620	629	0	0		
151	V48	0	0	0	0	0	0	0	0	0	0	620	629	0	0		
152	V49	0	0	0	0	0	0	0	0	0	0	630	639	0	0		
153	V50	0	0	0	0	0	0	0	0	0	0	710	719	0	0		
159	V53	0	0	0	0	0	0	0	0	0	0	720	729	0	0		
160	PRIKLAD	221	223	0	0	0	0	100	199	100	199	620	639	0	0		

Obr. 6. Výpis z dílčí databáze (NAVBI.L.DBF) vlastností nálezů (keramiky), která registruje vlastnosti z Bilinska.

13.3 Program – Kombinační analýza archeologických nálezů (KAAN)

13.3.1 Tvorba matic

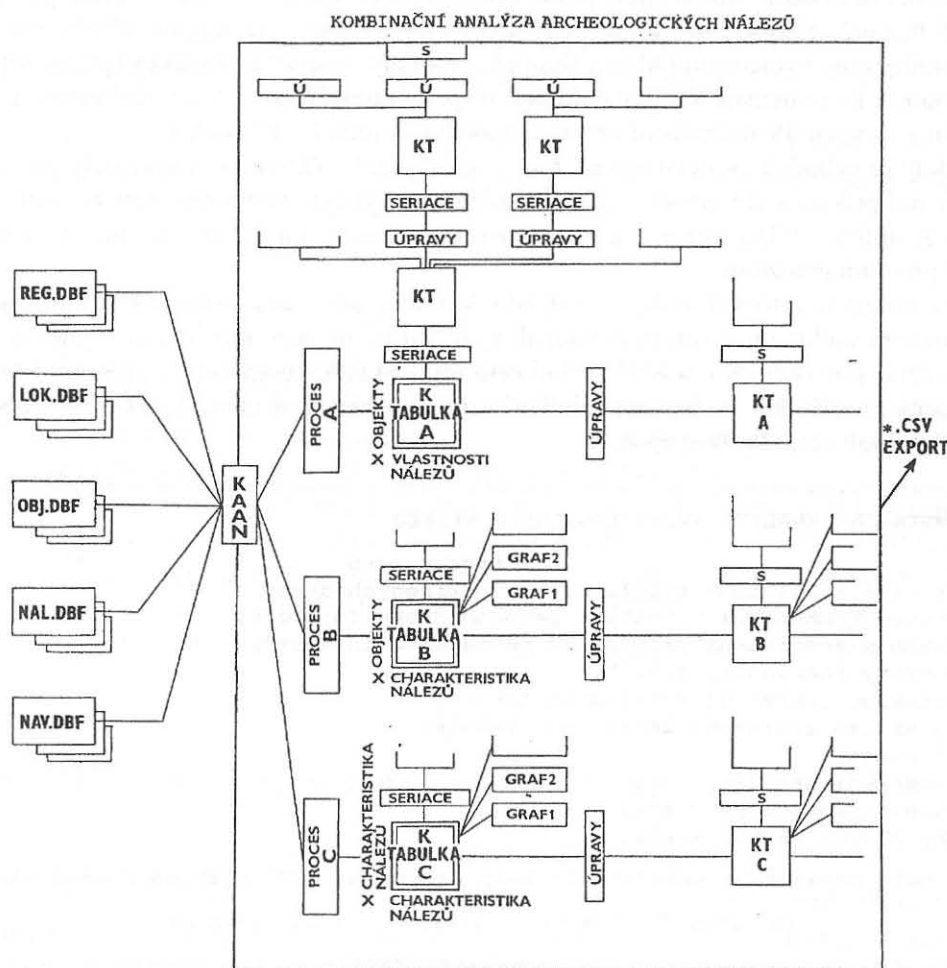
Máme-li vytvořena data, která potřebujeme k řešení chronologie, tj. data zachycující údaje o místech nálezů a o nálezech samotných, ale také charakteristiky a vlastnosti nálezů (tzv. deskriptory), které považujeme za relevantní, stojíme před klíčovou úlohou – vytvořit matici, kde v řádcích budou (v našem příkladě) objekty a ve sloupcích vlastnosti keramiky.

Protože vlastnost (např. ve sloupci) představuje kombinaci jednotlivých charakteristik nelze v databázových systémech požadovanou matici vytvářet automatizovaně pomocí běžných příkazů, proto byl vytvořen zvláštní program, který automatizovaně vytváří z výše uvedených databází matici (KT), kterou požadujeme (schéma programu viz. obr. 7). Program, který se nazývá „Kombinační analýza archeologických nálezů – KAAN“ umožňuje vytvářet z uvedených databází tři druhy kontingenčních tabulek (podrobný popis programu cf. Kubálek – Salač 1994).

Typ A – v řádcích KT jsou objekty, ve sloupcích vlastnosti keramiky (nálezy), které jsou nadefinované a uloženy v databázi vlastností nálezů (NAV.DBF).

Typ B – v řádcích KT jsou objekty, ve sloupcích jediná vybraná charakteristika (např. pouze tvar, pouze výzdoba atd.). Program v tomto případě nepracuje s databází NAV.DBF.

Typ C – v řádcích KT je vybrána jedna charakteristika keramiky (nálezu) a ve sloupcích jiná vybraná charakteristika. Např. v řádcích je tvar nádoby (co řádek, to jedna varianta tvaru) a ve sloupcích průměr ústí (co sloupec, to jeden rozměr uvedený v cm). Program pracuje pouze s databází NAL.DBF.



Obr. 7. Schéma práce programu KAAAN.

13.3.2 Práce s maticemi

Pro řešení chronologických otázek mají největší význam matice typu A, proto všechny příklady budeme uvádět na tomto typu; způsob práce s typy B a C je totožný.

Program KAAAN umožňuje automatizovanou tvorbu KT, což je základní předpoklad pro zpracovávání sídlištního materiálu, neboť oproti práci s pohřebišti je třeba vyvážení matic zpravidla mnohokrát opakovat. Vytvořením dílčích účelových databází, které velké databáze omezí např. dle regionu (např. pouze Bílinsko), nebo dle typu objektů (např. pouze polozemnice) umožní zaměřit se detailněji na dílčí problém (např. chronologii jediného regionu apod.).

Vytvořené matice je však vždy nutné dále upravit, neboť zpravidla obsahují řádky (objekty) s malým počtem obsazených polí nebo mající dokonce všech polí nulu (sledované vlastnosti se v objektech vyskytly zřídka nebo vůbec ne). Takovéto řádky nejen, že nemají pro seriaci žádný smysl, ale nulový řádek celý matematický postup zablokuje, takže jej nelze provést. Tyto řádky (sloupce) je nutné odstranit. Program KAAAN umožňuje automatické odstranění nulových řádků a sloupců.

V matici se mohou dále vyskytnout objekty, o kterých víme, že jsou buď superpozicemi, nebo jejich soubory byly skartovány, mají problematické nalezové okolnosti apod. I tyto objekty je nutné z matice odstranit, neboť nejen, že jejich chronologická hodnota je sporná, ale mohou ovlivnit uspořádání

celé matice, tj. i pořadí (datování) kvalitních souborů. Dále může KT obsahovat řádky (sloupce) pouze s jedním obsazeným polem. I tyto řádky resp. sloupce mají pro chronologii minimální význam. Podobně je tomu i u sloupců (vlastností), které mají obsazena všechna pole, neboť vlastnost vyskytující se ve všech objektech k řešení chronologie zpravidla také nepřispěje. Za určitých okolností ani vlastnost vyskytující se dvakrát, třikrát apod. nemá chronologický význam. Někdy chceme pozorovat matici v seriaci bez určité konkrétní vlastnosti či objektu. V souladu s těmito požadavky umožňuje program automatizované vyřazování řádků a sloupců, nesplňují-li určité požadavky (příkaz typu „odstraň všechny sloupce, které nemají obsazena alespoň tři pole“) umožňuje i vyřazování libovolných jednotlivých řádků a sloupců dle požadavku uživatele („odstraň řádek č. 38“ apod.).

Někdy je výhodné určité sloupce i řádky sečíst, např. některé vlastnosti byly příliš úzce definovány (viz náš příklad s vhlazovanou mísou – zde může být výhodné sečíst sloupce sledující vhlazování zvlášť zevnitř a a zvlášť vně a sledovat obecně kombinaci – mísa a vhlazování). I takovouto práci s maticemi program umožňuje.

Aby program splňoval určitý uživatelský komfort, umožňuje ukládat KT v libovolném okamžiku (v různém stadiu zpracování), či naopak vyvolání již uložené tabulky, dále grafické zobrazení řádku či sloupce, převedení KT se skutečnými četnostmi na tabulku absence x presence (APT), možné je i převedení absolutních hodnot na relativní (100% může představovat obsah jednoho řádku či sloupce nebo obsah celé tabulky) apod.

```
KOMBINACNI ANALYZA ARCHEOLOGICKYCH NALEZU

                                uvodni menu
1 tvorba kombinacni tabulky z dat porizenych dBase
2 vstup kombinacni tabulky z pevneho disku (diskety)
3 zaznam kombinacni tabulky na pevny disk (disketu)
4 upravu kombinacni tabulky
5 ciselne zobrazeni kombinacni tabulky
6 graficke zobrazeni kombinacni tabulky
7 seriace
8 obrácení tabulky
9 konec prace s programem
LOK, OBJCH, NAL01, NAV31,

<F1> help obrazovky, <Shift><F1> help programu, <F3> prehled funkci klaves.

(c) Ing. T. Kubalek - PhDr. V. Salac, unor1992
```

Obr. 8. Základní menu programu KAAAN.

Na tomto místě nelze zacházet do detailů, základní možnosti programu naznačuje ukázka hlavního menu programu. Zůstaňme pouze u konstatování, že chceme-li se zabývat seriací sídlištního materiálu, musíme se vypořádat s vytvořením relativně značného počtu dat, dále s jeho smysluplným uložením v určitém systému umožňujícím následně automatizované vytváření KT. Vytvořené tabulky je zpravidla nutné automatizovaně upravovat. Zhruba v těchto momentech se setkáváme (narozdíl od zpracování pohřebišť, kde většinou pracujeme s jednou či několika málo základními maticemi) při zpracování sídlištního materiálu s největšími potížemi. Máme-li je však vyřešeny, můžeme přistoupit k seriaci.

13.3.2.1 Poznámky k terminologii

Abychom si v následujících příkladech usnadnili popisování matic, budeme užívat následující terminologii.

Dlouhý sloupec (řádek) – sloupec v němž je obsazeno velké množství polí, tj. vlastnost (charakteristika) je zastoupena u velkého množství objektů (řádků) – zpravidla jde o výskyt v nadpolovičním množství polí (údaj v posledním řádku matice, vytvořené programem KAAAN)

Krátký sloupec (řádek) – sloupec, v němž je obsazeno pouze několik polí

Souvislý sloupec (řádek) – sloupec, který má po seriaci většinu obsazených řádků v pořadí za sebou

Nesouvislý sloupec (řádek) – sloupec, který má po seriaci velké mezery mezi obsazenými řádky

Těžký sloupec (řádek) – sloupec s nadprůměrně vysokými četnostmi, cf. údaj v předposledním řádku tabulky.

Lehký sloupec (řádek) – sloupec s podprůměrnými četnostmi

Těžké (lehké) pole – pole v matici s nad- resp. podprůměrnou četností v rámci sloupce, či celé tabulky.

Tyto termíny jsou analogicky užívány pro objekty či vlastnosti. Pojmenování jsou v konkrétních případech spojována – např. souvislý dlouhý těžký sloupec (řádek, objekt apod.), krátký lehký sloupec (objekt, řádek) atd.

13.3.3 Seriace

Máme-li již KT připravenou natolik, že chceme, aby byla provedena seriace, můžeme ji pomocí programu KAAN provést. Lze volit, zda má být provedena seriace registrující pouze výskyt či absenci vlastnosti (seriace absence x presence, dále SAP), nebo seriace pracující se skutečnými četnostmi, ať již absolutními či relativními (kvantitativní seriace, dále SK).

Po provedení seriace se zpravidla ukazuje, že uspořádání KT nemůže být chápáno jako konečné. KT je nutné dále upravovat a seriaci provádět opakovaně. Program umožňuje po každém kroku provedení výše zmíněných úprav KT, zároveň je možné pozorovat, jak se po případných úpravách mění korelační koeficient (dále jen Kf), tj. zda provedené úpravy vedly k jeho zvýšení či naopak. Některé možnosti práce s programem KAAN, tj. některé aspekty využití seriace, si ukážeme na následujících příkladech.

13.4 Příklad. Seriace laténských sídlištních keramických souborů ze SZ Čech

Vzhledem k nedostatku zkušeností s nasazením seriace ve větším měřítku se zdálo nevhodné pracovat od počátku se všemi daty, tj. s rozsáhlými maticemi. Jako nejvýhodnější se ukázalo zpracovávat samostatně nejprve jednotlivá sídliště, následně regiony a teprve na závěr provádět celkovou seriaci. To umožnilo nejen lépe kontrolovat a interpretovat pořadí objektů v tabulce po seriaci, ale postupně i poznávat význam jednotlivých vlastností keramiky pro řešení chronologických otázek.

Na počátku vznikla databáze vlastností keramiky (NAV.DBF), která obsahovala 160 definic vlastností. Již jen prvotní vytvoření celkové matice ukázalo, že některé vlastnosti materiál vůbec nevykazuje, jiné se vyskytly zcela ojediněle. Z matice rovněž vyplývalo, které vlastnosti jsou vázány regionálně. Na základě tohoto předběžného pozorování byly vytvořeny dílčí databáze vlastností keramiky typických pro jeden region – např. Bílinsko (NAVBIL.DBF, výpis obr. 6). Nejprve byla prováděna seriace sídliště v Břešťanech, analogicky tedy vznikly dílčí databáze obsahující objekty pouze z Břešťan (OBJBRE.DBF) a nálezy pouze z tohoto sídliště (NALBRE.DBF). První seriace byly prováděny na břešťanském materiálu především proto, že toto sídliště bylo před lety zpracováno klasickým postupem (včetně „ruční seriace“). Jeho chronologický vývoj byl tedy v hrubých rysech znám, což umožnilo snáze porovnávat výsledky počítačového zpracování a běžného archeologického postupu při řešení otázek datování.

13.4.1 Břešťany

Vyhodnocení laténského sídliště v Břešťanech bylo publikováno včetně výsledku „ruční seriace“ a jeho interpretace (Salač 1984, 266; zde obr. 9).

Nejprve proto byla provedena za pomoci programu KAAN seriace již publikované tabulky. Můžeme tak nejen ověřit tehdejší výběr 18 sledovaných vlastností, ale i porovnat možnosti „ručního“

Při bližším zkoumání kvantitativního obsahu však je zřejmé, že hranice mezi výskytem „běžným“ a „ojedinělým“ byla značně subjektivní. Ostatně je jistě sporné, zda je zdůvodnitelné chápat výskyt 1 za „ojedinělý“ a výskyt 2 za „běžný“. Proto byla provedena kvantitativní seriace (SK), která bere při diagonalizaci matice v úvahu skutečné hodnoty v jednotlivých polích.

Výsledek SK přináší tabulka S1/B (obr. 10), jejíž korelační koeficient dosahuje vyšší hodnoty než u tabulky předchozí – $K_f = 0,865$. Při řazení objektů vznikla odchylka – na řádcích 3. a 4 a 5. a 7. řádku si objekty zaměnily místa, přičemž obj. 7 řazený do posledního (nejmladšího) břešťanského horizontu se ocitá výše než objekty z 3. horizontu. Toto zjištění vzbuzuje určité pochybnosti o vyčlenění dvou nejmladších horizontů na břešťanském sídlišti.

Další krok představuje práce s podstatně vyšším počtem vlastností keramiky, tj. vytvoření základní matice, kde je ve sloupcích 73 vlastností z databáze NAVBILI.DBF (cf. obr.6). Tato dílčí databáze již reaguje na poznatky získané testováním vlastností a neobsahuje ty, které se na Bílinsku nevykytují. Na druhé straně byl při jejím vytváření kladen důraz na to, aby obsahovala co nejvíce vlastností, tj. aby byla co nejobjektivnější a zároveň, aby zohledňovala kvantitativní výskyt vlastností, které samy o sobě nemají přílišný chronologický význam. Za pomoci této databáze vznikla základní matice (obr. 11), jež byla následně podrobena seriacím procesům. Přitom se ukázalo, že ze 73 vlastností očekávaných na Bílinsku se jich v Břešťanech samotných vyskytuje pouze 57.

Výsledkem SAP je tabulka S2/A (obr. 12) s nepříliš vysokým korelačním koeficientem ($K_f = 0,589$). Ukazuje se, že zařazení velkého množství vlastností prakticky bez výběru, podstatně zhoršuje možnosti SAP. V řazení objektů dochází k nemalým změnám. Objekty nejstarších dvou horizontů si navzájem prohodily místa, podobně jako v nejmladších horizontech. Zajímavé a důležité je, že nedošlo k výměně mezi staršími (LT A-B2) a mladšími horizonty (LT C1-C2), a to ani v takto minimálně uspořádané matici.

Tabulka S2/B (obr. 12) ukazuje výsledek SK; korelační koeficient má hodnotu 0.661, tj. je vyšší než u předchozí tabulky. V tomto případě již řazení objektů spíše odpovídá původní představě, ovšem u mladších objektů opět dochází k záměně pořadí, které se odlišuje od všech předchozích seriacních výsledků.

Vzhledem k dosavadním poznatkům o chronologii laténské keramiky se však zdá být evidentní, že se v objektech vyskytují druhotné intruze. Je možné, že právě tyto intruze zabraňují vyčlenění dvou nejmladších horizontů, či jiným způsobem podstatně ovlivňují výsledky SAP i SK. Pokusíme se tento předpoklad testovat. Vyřadíme z databáze vlastností zahrnuté na tab. S2/B (obr. 12), které budeme považovat za intruze. SAP a SK takto upravené matice ukazují tab. S3/A-B (obr. 13). Výsledek SAP vykazuje nepatrně vyšší korelační koeficient $K_f = 0,608$, přičemž pořadí objektů se prakticky nezměnilo. Matice po SK má rovněž poněkud vyšší korelační koeficient $K_f = 0,679$, přičemž pořadí objektů v horní polovině tabulky se nemění. Naopak u mladších horizontů dochází opět ke změnám.

konkrétní poloha nálezů v objektu apod. V případě Břešťan však nálezové okolnosti neumožňují tento problém dále řešit.

Výsledné pořadí objektů bylo zařazeno do sloupce „datování 2“ záhlaví tabulky S4/B. Toto pořadí bude dále testováno v analýze bílinského regionu (cf. zkratka BR v „datování 2“ u tab. S6-8). Pořadí tabulky S4/B bylo upřednostněno proto, že tabulka nebyla poznamenána lidským faktorem (odstraňování intruzí) a zdá se být tedy objektivnější, byť má nižší korelační koeficient, než tab. S5/B.

13.4.2 Bílinsko

Podobným způsobem jako u nálezů z Břešťan byla provedena seriace souborů z Radovesic. U matice s nejvyšším korelačním koeficientem z této lokality bylo, podobně jako u Břešťan, očíslováno pořadí objektů (zkratka RA v „datování 2“). Nakonec jsme přistoupili k celkové seriaci souborů z Břešťan, Radovesic a ostatních lokalit na Bílinsku.

Výchozí seriaci představuje tab. S6 ($K_f = 0.634$), která, zdá se, poskytuje očekávaný výsledek, tj. pořadí objektů z Břešťan a Radovesic v celkové seriaci v podstatě odpovídá pořadí, které bylo získáno samostatnou seriací těchto sídlišť.

Výsledek takovéto seriace přináší tabulka S6 (obr. 16) ($K_f = 0.642$). Zde uvedené pořadí objektů z Břešťan a Radovesic v podstatě odpovídá pořadí z dílčích seriací těchto sídlišť. Pořadí objektů v KT lze mít z hlediska chronologického již za velmi relevantní, neboť tabulka je obohacena kvalitními soubory z menších lokalit tzv. dvorců v Bílině (č.obj. 6000 /Waldhauser – Holodňák 1984/ a Radovesicích 2 č.obj. 3000), které jsou poměrně dobře datovány nekeramickým materiálem a navíc jsou považovány za jednofázové. Nemohou se zde tedy vyskytovat nesoučasné intruze, vyloučit můžeme i možnost, že by datující (nekeramické) nálezy ke keramickým souborům časově nepatřily. Proto poněkud překvapuje, že objekty z těchto lokalit jsou v některých případech od sebe v tabulce relativně vzdáleny (např. 3001 a 3004, 3006; 6002 a 6004, 6008). Tato skutečnost může být způsobena dlouhými vlastnostmi. Tabulku však zřejmě ovlivňují i objekty, které mají málo typických vlastností (krátké objekty) se zastoupením pouze ve dlouhých vlastnostech.

Následně byla tabulka zbavena dlouhých vlastností a krátkých objektů (v objektu musí být zastoupeny alespoň 4 vlastnosti). Výsledek přináší tabulka S 7 (obr. 16) ($K_f = 0.843$). Na takto uspořádané KT můžeme pozorovat, že objekty z tzv. jednofázových dvorců jsou nyní blíže u sebe (6000 a 3000), ovšem u lokality Radovesice 2 (3000) je mezi objekty přeci jen určitá vzdálenost. Zde je několik možných vysvětlení: lokalita nepředstavuje jednofázové sídliště (jde o torzo většího a déletrvající sídliště), vstupní data (tj. typáře a následně vytvořené vlastnosti) neumožňují jemné rozlišení mladší laténské keramiky, výběr vlastností nebyl proveden správně – nestejná délka vlastností (sloupců) i nestejná váha jednotlivých sloupců i polí způsobuje uspořádání tabulky seriací, které není zcela chronologické, naopak tabulka může náhodně obsahovat velký počet přibližně stejně starých objektů, takže vzdálenost mezi objekty nemá ve všech partiích KT stejný chronologický význam atd. atd. Možností interpretace je velké množství, v neposlední řadě se ukazuje potřeba zkoumat uspořádání seriačních tabulek i jinými statistickými resp. matematickými metodami.

V každém případě je velmi užitečné, jsou-li při řešení chronologických otázek k dispozici krátkodobá či dokonce jednofázová sídliště. Ocitnou-li se totiž objekty z těchto lokalit v tabulce po seriaci v sousedství, znamenají nezávislou kontrolu výsledku – z terénní situace je totiž zřejmé, že musí být současné.

Tabulka S7 (obr. 16) byla pokusně rozdělena do čtyř částí, které jsou v souladu s dosavadními znalostmi datování laténské keramiky interpretovány chronologicky. Druhá a třetí část tabulky byly ještě rozčleněny na podčásti. Ze samotné tabulky však vyplývá, že toto další dělení je spíše spekulativní.

38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	sou-
TP1	TP2	TP3	UV3	V12	V37	V48	V16	V15	T30	PUV	V47	UV2	P00	T07	V53	P00	T01	V48	UV5	V49	UV1	T01	T41	P00	T01	V69	T01	P06	V28	V12	T05	T05	T48	est	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100:	12:	41:	30:	30:	47:	20:	24:	43:	3:	77:	7:	39:	8:	17:	4:	81:	19:	2:	30:	22:	15:	28:	19:	45:	11:	14:	143:	9:	22:	4:	21:	17:	5:	181:	3780:
41:	7:	22:	12:	17:	10:	8:	15:	10:	2:	30:	4:	15:	5:	4:	3:	22:	9:	2:	15:	11:	3:	10:	4:	15:	7:	5:	21:	5:	4:	3:	4:	3:	2:	2:	54:

39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	sou-	
TP1	TP2	TP3	UV3	V12	V37	V48	V16	V15	T30	PUV	V47	UV2	P00	T07	V53	P00	T01	V48	UV5	V49	UV1	T01	T41	P00	T01	V69	T01	P06	V28	V12	T05	T05	T48	est	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4:	25:	34:	3:	11:	3:	8:	30:	41:	74:	7:	28:	8:	17:	39:	44:	3:	19:	28:	142:	10:	14:	13:	44:	19:	22:	11:	8:	22:	4:	21:	17:	3:	18:	1704:	
3:	8:	19:	2:	7:	3:	5:	17:	9:	20:	4:	13:	3:	4:	13:	9:	2:	6:	10:	20:	13:	5:	3:	12:	9:	11:	7:	4:	4:	3:	6:	3:	2:	2:	40:	

Velmi důležité jsou i zde polohy objektů z jednofázových „dvorců“ z Bíliny a Radovesic 2, které zřejmě chronologickou interpretaci uspořádání tabulky nezávisle potvrzují. Na druhou stranu je nutné přiznat, že pořadí objektů v tabulce odpovídá původnímu datování pouze v horní části. Naopak dolní část tabulky dosavadním představám o keramice stupňů LT C2 a LT D1 příliš neodpovídá. Příčin, proč se pomocí seriace nedaří vytvořit jemnější chronologii laténské keramiky, může být celá řada, počínaje příliš hrubými metodami při zkoumání výplní objektů, přes chybný výběr sledovaných vlastností, až po možnost, že objekty na déle trvajících sídlištích prostě blíže datovat pomocí této metody nelze (cf. např. Dreslerová 1995; Neustupný 1996). V každém případě můžeme konstatovat, že výsledek, který představuje např. tabulka S8 je nutné dále zkoumat jak dalšími matematickými metodami (shluková, korespondenční či faktorová analýza atd.), tak i pomocí externí evidence jakou může představovat terénní situace – jednofázová sídliště, naopak vzájemně se porušující objekty apod.

13.5 Poznámky k užití seriace při datování sídlišť

Při řešení chronologie laténské sídlištní keramiky bylo u nás prvně užito při zpracování většího množství sídlištních nálezů automatizovaného vzniku matic z databázi a provádění jejich seriací. Proto je snad vhodné zde shrnout některé zkušenosti s aplikací této metody.

1. Při datování nestačí vytvořit pouze jednu tabulku zachycující výskyt či absenci určitých vlastností (SAP), což je doposud v literatuře běžný jev – její vznik a tudíž i výsledky mo-

pak výhodou patrně je, že metoda klade důraz na celek, zdůrazňuje návaznost a nakonec ukazuje i na problematičnost (archeologem chtěného) členění vývoje. Přesto však je, zdá se, metoda schopna základní změny ve vývoji registrovat.

	4	3	2	7	1	14	6	5	10	9	15	8	11	13	12	
Vikletice I, 26	+	+	+													1
Vikletice I, 28	+	+	+	+												
Vikletice I, 24	+	+	+	+	+	+										
Vikletice I, 102	+	+	+	+	+	+	+	+	+	+	+					2
Vikletice II, 12				+	+	+	+		+	+	+	+				
Vikletice I, 23				+		+	+			+	+	+				
V. Ves 3/61				+		+	+	+	+	+	+	+	+	+		
V. Ves 1/61					+	+		+	+	+	+	+	+	+		
Vikletice I, 88										+	+				+	3
V. Ves 18/62												+	+	+		

Kombinační tabulka prvků keramiky.

Tvary: 1. Kónická mísa se svisle nasazeným okrajem (obr. 4 : 5, 14; 9 : 1, 12); 2. Kónická mísa nečleněná (obr. 4 : 4; 5 : 16, 17; 6 : 6). Plastické prvky nádob: 3. Ploché žebro na kónickém hrdle vyšších tvarů (obr. 4 : 3; 5 : 1; 8 : 4); 4. Úzké žebírko na hrdle (podhrdlí) jemné keramiky (obr. 4 : 6; 5 : 2; 6 : 1; 8 : 1); 5. Ploše vytažená nožka (obr. 8 : 3); 6. Oble hráněný masivní okraj zásobnice (obr. 3 : 6; 9 : 13). Úprava povrchu: 7. Chuchvalecovité zdrsnění (obr. 4 : 8; 6 : 11); 8. Jemné obloukovité nebo svislé, až vlasové rytí povrchu (obr. 3 : 5; *Drda 1974*, obr. 5 : 5); 9. Výrazné husté obloukovité rytí povrchu (obr. 3 : 11); 10. Černý smolný nátěr. Prvky výzdoby: 11. Jednotlivé ryté obloukovité pásy (*Koutecký 1970*, obr. 11 : 11; 12 : 25, 29; 15 : 4; 28 : 2); 12. Ryté mnohonásobné oblouky nebo vlnice v kombinaci se svislým hřebenováním (obr. 7 : 1, 8; *Koutecký 1970*, obr. 28 : 1); 13. Jiná rytá ornamentální výzdoba (*Koutecký 1970*, obr. 12 : 24; 16 : 16, 25; 28 : 12). 14. Nepravidelná vhlazovaná výzdoba na užitkové keramice (obr. 3 : 8; 9 : 4, 5, 13); 15. Pravidelná ornamentální vhlazování výzdoba (obr. 3 : 7; 7 : 4; *Drda 1974*, obr. 5 : 8).

Obr. 18. Vikletice, kombinační tabulka sídlištních objektů (dle *Drda 1977*, 383, doplněno).

7. Je-li metoda aplikována u malých sídlišť (Dreslerová 1995) je nutné obezřetně postupovat při interpretaci výsledků. Můžeme sice zjišťovat matematickými metodami „odstupy“ mezi sousedními objekty, avšak musíme si uvědomit, že mohl zcela uniknout chybějící horizont, tj. časový hiát mezi objekty. Jsou-li pro určitý časový horizont typické určité vlastnosti a chybí-li tento horizont na sídlišti (častěji na jeho prozkoumané části), pak v připravených sloupcích budou nulové hodnoty a tyto sloupce budou odstraněny – lhostejno zda automatizovaně v počítači či v našich hlavách při tzv. ruční seriaci (podobně jako nálezy vyskytující se pouze stopově). Na toto nebezpečí ukazují publikované chronologické závěry z Břešťan (Salač 1984) a Vikletic (Drda 1977). U Vikletic byly všechny objekty na základě kombinační tabulky (obr. 18) datovány do bezprostředně navazujících horizontů v rámci stupňů LT C a D. Seriace prováděná v širším kontextu SZ Čech však zřejmě prokázala značný časový odstup mezi objekty (horizonty; cf. pozici objektů s označením VIK I,II,III v „datování 2“ u tab. S8 – obr. 17). Ostatně příklady z Břešťan a Vikletic ukazují, že chronologické horizonty vytvořené na základě jedné malé lokality, nadto neúplně prozkoumané, mají problematickou hodnotu.
8. Objekty z okrajových horizontů (prvního resp. posledního) by měly být v matici dostatečně zastoupeny. Vyskytnou-li se pouze 1x, pak budou vlastnosti typické pouze pro ně buď zcela vyřazeny nebo budou natolik krátké a lehké, že se objekty nemusí objevit na okrajích tabulky.
9. Z algoritmu seriace vyplývá, že dlouhé a těžké řádky (sloupce) se budou po SK zpravidla vyskytovat uprostřed tabulky (cf. polohu obj. 7 z Břešťan na tab. S1B – obr. 10), podobně jako těžká pole na diagonále. Vždy je tedy nutné zvážit, zda je jejich zařazení do seriace správné. Řešení je obtížné, ale rozhodujícím kritériem by měla být velikost tabulky, čím je KT menší, tím snáze bude těmito jevy deformována.
10. Závažný limit při použití seriace u sídlištní keramiky představuje fakt, že sídlištní objekty mívají zpravidla podstatně odlišná množství stěpů, tj. následně i sledovaných vlastností (jiná situace bývá u hrobů). Tato skutečnost může nemálo ovlivnit pořadí objektů v tabulce. Při výběru objektů a vlastností je nutné mít na paměti, že ideální výsledky by přinesla matice se stejně dlouhými a těžkými řádky a sloupci.
11. Domnívám se, že na zde uvedené matice lze pohlížet i z jiného hlediska – tj. jako na relativně značnou znalostní bázi, která je schopna dávat poměrně stabilní a objektivní výsledky, byť ne natolik jemné, jak bychom snad žádali (např. ve srovnání s chronologií pohřebišť). Je zřejmě opodstatněné, využívat typologické tabulky a databáze (především vlastností) ve spojení s programem KAAAN jako tzv. expertní systém. Tj. do databáze či přímo do matice lze doplnit nový objekt, který bude seriací automatizovaně datován.

13.6 Závěr

V tomto příspěvku jsme si ukázali jeden z mnoha možných přístupů k využití seriace při zpracovávání archeologického materiálu. Snahou bylo ukázat některé momenty práce, které se zřejmě budou opakovat, lhostejno zda bude použit zde uvedený program KAAAN (lze obdržet od autora), tzv. „Bonnská seriace“, či některý z komerčních statistických paketů (např. BMDP). Příspěvek se proto zabýval pouze elementárními možnostmi této metody. Rozsáhlejší programové pakety skýtají další možnosti při práci s maticemi či při provádění seriace. Velice lákavá je např. možnost, že řádky či sloupce nemají stejnou váhu, např. dlouhá vlastnost vyskytující se téměř u všech objektů bude mít v seriaci nižší váhu (méně ovlivní uspořádání tabulky). Naopak krátká vlastnost, kterou považujeme za typickou pro určitý horizont, bude mít váhu vyšší, tj. ovlivní uspořádání tabulky více (cf. např. Herzog – Scollar 1987). Je možné, ba pravděpodobné, že podobným způsobem se podaří uspořádat objekty (hroby) způsobem, který bude lépe odpovídat archeologovým představám. Na druhou stranu stojíme před zásadním zásahem do jinak objektivní pracovní metody. Vzhledem ke složitosti sídlištní problematiky bude nutné stanovit taková pravidla, která by zajistila, aby podobné zásahy nevnesly do metody značný stupeň subjektivismu, který by ji podstatně devalvoval.

Příspěvek ukázal, že seriace je jen jednou z pracovních metod a její aplikace ještě zdaleka nemusí vést k vyřešení problému. V žádném případě není řešením seriace jediné matice vytvořené výběrem z velkého množství dat.

Na druhou stranu automatizovaná tvorba matic, jejich následná seriace a stanovování korelačního koeficientu představují objektivní, bezchybný a opakovatelný proces umožňující kontrolu. Významné je, že tento postup umožňuje čtenáři (uživateli) nejen kontrolovat výsledky svých kolegů, ale dále je rozpracovávat (cf. diskusi o datování pohřebiště v Horathu; Miron 1986; Brugmann 1993).

Seriace je jako každá matematická metoda zcela objektivní, její výsledky jsou však zásadně ovlivněné vstupy (daty) dodanými archeologem. Také interpretace matice uspořádané seriací je již opět ryze archeologickou záležitostí (zdaleka ne každé uspořádání matice lze interpretovat jen chronologicky!). Naší snahou musí být, aby se objektivnost a bezchybnost vstupů a interpretací alespoň blížila objektivnosti a bezchybnosti algoritmu seriace.

13.7 Literatura

- Brugmann, B. 1993: Bemerkungen zur relativen Chronologie mittel- und spätlatènezeitlicher Gräberfelder im Hunsrück-Nahe-Raum, *Archäologisches Korrespondenzblatt* 23, 77-86.
- Drda, P. 1977: Laténská osada Vikletice I, *Archeologické rozhledy* 29, 363-393.
- Dreslerová, D. 1995: A late Hallstatt settlement in Bohemia. Excavation at Jenštějn, 1984. Praha.
- Herzog, I. – Scollar, I. 1987: Ein „Werkzeugkasten“ für Seriation und Clusteranalyse, *Archäologisches Korrespondenzblatt* 17, 273-279.
- Kubálek, T. – Salač, V. 1994: Kombinační analýza archeologických nálezů (KAAN), příručka k programu. Praha.
- Ihm, P. 1983: Korrespondenzanalyse und Seriation, *Arch. Informationen* 6, 18-21.
- 1990: Archäologische Datenverarbeitung mit SAS, *Wiss. Zeitschrift der Humboldt-Univ. zu Berlin, Reihe Gesellschaftswiss.* 39, 284-289.
- Miron, A. 1986: Das Gräberfeld von Horath, *Trierer Zeitschrift* 49, 7-198.
- Neustupný, E. 1978: Mathematics at Jenišův Újezd, in: Waldhauser, J. ed., *Das keltische Gräberfeld bei Jenišův Újezd in Böhmen II*, 40-66.
- 1996: Poznámky k pravěké sídlištní keramice, *Archeologické rozhledy* 48, 490-509.
- Podborský, V. – Kazdová, E. – Košťufík, P. – Weber, Z. 1977: Numerický kód moravské malované keramiky. Brno.
- Rulf, J. 1993: Archeologie, archeologická data a archeologové, *Archeologické rozhledy* 45, 165-172.
- Rulf, J. – Salač, V. 1995: Zpráva o laténské keramice v severozápadních Čechách, *Archeologické rozhledy* 47, 373-417.
- Salač, V. 1984: Laténské sídliště u Břešťan, okr. Teplice, *Archeologické rozhledy* 36, 261-277.
- 1993: Seriation of Iron Age settlement sites in NW Bohemia – system of databases, CAAF program, in: Andresen, J. – Madsen, T. – Scollar, I. eds., *Computing the Past (CAA92)*, 343-348. Aarhus.
- 1995: The density of archaeological finds in settlement features of the La Tène period, in: Kuna, M. – Venclová, N.: *Whither archaeology?*, *Papers of honour Ev•en Neustupný*, 263-276.

Práce představuje jeden z výstupů grantu GA ČR č. 404/94/0606
Za pročetí rukopisu a připomínky děkuji E. Neustupnému a J. Rulfovi

14. Syntéza struktur formalizovanými metodami – vektorová syntéza (Evžen Neustupný)

Následující text pojednává o tématu, které se často označuje jako faktorová analýza. Toto označení je nepřesné, neboť nejde o žádnou analýzu, jestliže se snažíme struktury získat, objevit a prozkoumat (explorovat), a tedy syntetizovat, nikoliv jen ověřit (konfirmovat). Metody z rodiny „faktorové analýzy“ jsou založeny na teorii vektorových prostorů a pojem vektoru je v nich klíčový; proto jsem navrhl je souhrnně označovat jako vektorovou syntézu (Neustupný 1979).

V dalším výkladu předpokládám určitou obeznámenost s archeologickou analýzou a základním rozvržením archeologické metody, ale v podstatě v rozsahu menším než v článku o archeologické metodě (Neustupný 1986). Také terminologicky navazují na své předchozí práce o archeologické metodě (zejména Neustupný 1973, 1986 a 1993).

Při studiu tohoto textu doporučuji nejdříve zběžně přečíst první, teoretickou část („Formální aspekt archeologických struktur“) a pak se obrátit přímo k příkladům; během studia příkladů je vhodné se k teoretické části opakovaně vracet.

Cílem následujících odstavců je usnadnit pochopení výsledků, které dnes automaticky generuje počítačový software. Pokud k takovému pochopení nedojde, hrozí mylné výklady počítačových výstupů.

Formulace úlohy

Po *metodologické stránce* se archeolog setkává s problémy trojího typu:

1. *problémy analytickými* spočívajícími v tom, jak fyzicky (v terénu) rozložit archeologické prameny, které chce studovat, a jak je popsat ve fázi deskripce,
2. *problémy z oblasti syntézy struktur* spočívajícími v nalezení nenáhodnosti (pravidelnosti, zákonitosti) v archeologických pramenech, tj. v mrtvé kultuře
3. *problémy z oblasti interpretace* spočívajícími ve vysvětlení zjištěných pravidelností mrtvé archeologické kultury v termínech někdejší živé skutečnosti.

Empirická metodologie vědy, pokud tyto problémy vůbec vyčleňovala, předpokládala, že jde o tři postupné kroky, které se aplikují jednorázově a vedou tak k víceméně konečnému, definitivnímu poznání. Domnívám se, že tyto problémy ve skutečnosti tvoří *nerozetnutelný kruh*, jehož jednotlivé segmenty se aplikují iterativně (v řadě kroků) a vedou tak ke stále dokonalejšímu poznání (Neustupný 1986, 1993).

V následujícím textu se budeme zabývat problémem syntézy struktur, tj. tím segmentem archeologické metody, který jsem označil číslem 2. Je důležité si uvědomit, že metody, které popíšeme, neřeší nijak problémy ani typu 1 ani typu 3, tedy například samy o sobě nijak nepřispívají k interpretaci. Jestliže v příkladech budeme zmiňovat také interpretaci, je to jen proto, abychom alespoň částečně vysvětlili smysl popisovaných metod v celkovém procesu archeologického poznání.

Archeologické prameny mají dva (a jen dva) druhy pozorovatelných vlastností: formální a prostorové. Jim odpovídají dva aspekty archeologických struktur, které jsou ve skutečnosti projevem těchto pravidelností. Hlavní část tohoto textu pojednává o formálním aspektu struktur, který je z metodologického hlediska dobře prozkoumán; v závěru se zmíníme o aspektu prostorovém a jeho spojení s formálním.

Faktorová analýza, první z aplikovaných metod vektorové syntézy, byla vyvinuta pro potřeby psychologie na počátku 20. století; ovšem před zavedením počítačů nebyla rozsáhleji využitelná. Pronikla do řady disciplín, ale H. Harman v r. 1970 ještě neuváděl žádné použití v archeologii (Harman 1970, 7). Pokud je mi známo, první opravdu stimulující použití v archeologii pochází od L. Binforda a S. Binfordové (1966). Na základě inspirace touto prací realizoval autor tohoto příspěvku první faktorové analýzy v r. 1969 a v r. 1970 už o nich podal zprávu na konferenci v Southamptonu (Neustupný 1973).

14.1 Formální aspekt archeologických struktur

Představme si některé nejběžnější úlohy tradiční archeologie. Stojíme před pohřebištěm, jehož jednotlivé hroby obsahují různé milodary, nebo před sídlištěm, jehož zahloubené objekty obsahují keramické zlomky. Definice toho, co je hrobem, zahloubeným objektem, rozlišení jednotlivých typů milodarů a typů keramiky je záležitostí analýzy a nebudeme je proto v této souvislosti diskutovat. Stejně tak záležitostí analýzy je deskripce vyjádřená tabulkou, v jejíchž řádcích jsou například jednotlivé hroby (zahloubené objekty) a ve sloupcích jsou jednotlivé typy milodarů (typy keramiky), které těmto hrobům (zahloubeným objektům) odpovídají. Jinak řečeno, máme před sebou obvyklou popisnou tabulku (deskripce), ať už ve formě matice nebo relační databáze.

Základní otázky nyní zní:

1. Jsou *typy* milodarů rozděleny v hrobech náhodně, nebo některé z nich *tvorí skupiny*, tj. jejich společný výskyt v hrobech je typičtější než u jiných (v případě laténských hrobů mohou takovou skupinu tvořit např. meče, štíty a hroty kopí) nebo je naopak jejich společný výskyt netypický (případně se vylučují, jako třeba sekeromlaty a náhrdelníky ve šňůrových hrobech)?
2. Lze říci, *kolik* takových skupin je?
3. Lze podat *seznam typů* milodarů, které se spolu vyskytují v uvedených skupinách, případně lze uvést nějaká *čísla* charakterizující příslušnost typů do uvedených skupin?
4. Lze zpětně určit, které ze skupin jsou typické pro jednotlivé hroby, případně *nakolik* jsou pro ně typické?

Uvedené skupiny (typů) nazveme *faktory*; obecně jimi budeme rozumět abstraktní množiny deskriptorů seřazené pomocí určitých číselných charakteristik (později tato čísla nazveme faktorové koeficienty). Takové uspořádání (strukturování) deskriptorů je výsledkem nenáhodné variability, která je skrytě (latentně) obsažena v deskriptivní tabulce.

Jak uvidíme, otázky 2 a 3 lze zodpovědět formalizovanými metodami vektorové syntézy, tedy jinak řečeno metodami z rodiny tzv. faktorové analýzy. Otázku 1 lze těmito metodami zodpovědět pouze tehdy, jestliže formalizované metody pojmem jako metody statistické, otázku 4 pouze některými z uvedených metod.

Všechny čtyři otázky ovšem v určité formě řešila také *tradiční archeologie neformalizovanými metodami*, přičemž číselné charakteristiky faktorů zůstávaly vesměs neurčeny (počet faktorů, míry typičnosti druhů milodarů pro jednotlivé faktory, míry typičnosti hrobů pro jednotlivé faktory, a pochopitelně i pravděpodobnost celkového řešení). Chybění těchto číselných charakteristik sice nijak neimplikuje nesprávnost tradičního řešení, ale současně nám nedává žádné vodítko k důvěře v jeho správnost.

V určité (a to ve velmi omezené) formě lze uvedené otázky řešit *prostředky jednoduché tradiční statistiky* nepřesahujícími analýzu rozptylu. Charakteristikou tohoto způsobu řešení je, že jednotlivé faktory, které ovlivňují variabilitu archeologického kontextu, *musíme znát předem*. Tak např. můžeme předpokládat, že variabilitu milodarů na pohřebišti určuje pohlaví zemřelých. Vezmeme tudíž zvlášť hroby mužské a zvlášť hroby ženské a hledáme, které typy milodarů jsou statisticky významně spojeny buď s jedním nebo s druhým pohlavím. Pomocí analýzy rozptylu (jde o středně náročnou statistickou metodu) můžeme takto posuzovat i celou skupinu typů nebo můžeme současně zkoumat vliv pohlaví a věku. Takové postupy jsou produktivní, ale v podstatě jenom potvrzují existenci pravidelností (struktur), které předpokládáme na základě jiných úvah: samy o sobě nikdy neodhalí nové, neznámé faktory.

Jinak řečeno, tyto metody jsou *konfirmatorní* (tj. potvrzující), nikoliv *exploratorní* (tj. vyhledávající). Stejně tak nevedou k důležitému konstatování, že získané řešení je vzhledem k určitému deskriptivnímu systému optimálním řešením, které *vyčerpává informaci* obsaženou v tomto deskriptivním systému (neobsahuje tedy další informaci, kterou je možno z tohoto systému extrahovat).

14.2 Postup formalizovaného řešení

V následujících odstavcích popíšeme základní kroky (algoritmus) při formalizované syntéze struktur tzv. metodou hlavních komponent. Zaměříme se také na základní terminologii, kterou budeme potřebovat v dalších částech této práce. Logické důvody, proč je takový algoritmus schopen řešit archeologické problémy, jsou uvedeny jinde (Neustupný 1993).

Hlavní komponenty jsou pouze jednou z víceméně ekvivalentních metod vektorové syntézy. Vybrali jsme je proto, že jsou nejbližší matematickému modelu a jsou proto relativně jednoduché. Někteří archeologové dávají přednost jiným podobným metodám (jako je vlastní faktorová analýza, korespondenční analýza, multidimenzionální škálování, částečně tzv. diskriminační a kanonická analýza apod.), protože se domnívají, že lépe odpovídají jejich datům. Většinou jde o nedorozumění nebo nedostatečné pochopení algebraické nebo statistické podstaty řešení: pro běžné archeologické úlohy je většina metod z rodiny faktorové analýzy ekvivalentní, ale všechny nedávají stejné možnosti. Tak např. u korespondenční analýzy nedostáváme rotované faktorové koeficienty, a orientace ve výsledcích je závislá na grafickém znázornění. Na druhé straně korespondenční analýza dává možnost znázornit strukturu objektů a deskriptorů (viz dále) ve stejném prostoru.

14.2.1 Krok 0 (sestavení deskriptivního systému)

Je to krok, který de facto předchází vektorovou syntézu a patří do analýzy. Zde si především všimneme základní terminologie a podmínek, které další kroky algoritmu kladou na konkrétní podobu deskriptivní matice.

Řádky deskriptivní matice nazveme *objekty*, její sloupce označíme jako *deskriptory*. Deskriptory můžeme chápat jako sloupcové vektory, tj. sloupce nějakých hodnot (odtud název *vektorová* syntéza). K tomu, abychom mohli provést další výpočty, musí mít deskriptivní matice následující vlastnosti:

1. musí obsahovat vesměs *reálná čísla*. Reálnými čísly se rozumí počty nebo (fyzikální) rozměry (hodnoty tzv. kardinálních proměnných); lze jimi rozumět i hodnoty tzv. dichotomických proměnných (prezence=1, absence=0). Stavů tzv. nominálních proměnných (např.: proměnná=barva, stvavy= šedá, černá, hnědá) nejsou reálnými čísly, a to ani v případě, že jsou vyjádřeny číslicemi. Tak např. pokud jsou objekty jednotlivé střepy, nelze jako deskriptory zvolit hodnoty „převažující barva hnědá“ a „převažující barva šedočerná“, protože každý střep je převážně buď hnědavý nebo čedočerný. Nuly, kterými bychom zapsali, že nějaké střepy nejsou hnědé (protože jsou šedočerné a nemohou být obojí) jsou tzv. nepravé nuly (Neustupný 1973) a v dalších krocích s nimi nelze provádět smysluplné výpočty (poučení o kardinálních a nominálních proměnných bývá součástí každé příručky statistiky).
2. Deskriptivní matice nesmí obsahovat tzv. *chybějící data*, tj. prázdná políčka. Jestliže objekty jsou hroby, pak deskriptor „délka hrobu“ může být zahrnut do popisného systému pouze tehdy, jestliže je délka známa u všech hrobů. Pokud není, je nutno příslušný hrob s neznámou délkou buď vynechat nebo délku hrobu odhadnout; všechny takto odhadnuté hodnoty jsou ovšem nějakým způsobem neuspokojivé a při jejich větším množství mohou nekontrolovatelným způsobem ovlivnit výsledky.

Jinak řečeno, sloupce deskriptivní matice musí tvořit tzv. lineární vektorový prostor, tj. musí umožňovat základní operace s vektory (v matematice je lineární vektorový prostor definován přesněji). Nepravé nuly, chybějící hodnoty a nominální proměnné jsou nejčastějšími příčinami nelinearity archeologických deskriptivních systémů: do deskriptivních systémů zanášejí prvky, které nejsou čísly a nelze s nimi provádět smysluplné výpočty.

V případě, že výchozí deskriptivní matice není lineárním vektorovým prostorem, je nutno v následujícím kroku 1 použít zvláštní koeficienty; to je vždy možné jen za cenu ztráty informace a za cenu vytvoření nestandardního východiska pro aplikaci dalších metod lineární algebry, nemluvě o tom, že se obvykle zcela ztrácí možnost použití statistických kritérií. Nevýhod je tolik, že takový postup nelze doporučit.

14.2.2 Krok 1 (výpočet korelační matice)

Prvním krokem vektorové syntézy je výpočet koeficientů, které (obecně) vyjadřují *závislost mezi jednotlivými deskriptory*. Teoreticky (nebo zkusmo) lze odvodit velké množství různých koeficientů, které při dalším zpracování dávají různé výsledky. Dříve archeologové napodobovali přírodovědce v hledání „nejlepších“ koeficientů, postupně se však ukázalo, že lineární korelační koeficient (který je v jednoduchém vztahu k eukleidovské vzdálenosti mezi vektory – srov. Neustupný 1973) dává nejlepší výsledky. Je to proto, že je to pojem, který je velmi jednoduše odvoditelný z teorie lineárních vektorových prostorů.

Korelační koeficienty jsou čísla od -1 do +1. Koeficient -1 znamená maximální nekorelovanost deskriptorů (např. situaci, že sekeromlat je ve všech hrobech, v nichž není náhrdelník a naopak), +1 jejich maximální korelovanost. Korelační koeficient 0 znamená zcela „nerozhodný“ stav, nezávislost (statistickou nevýznamnost). V praxi se ovšem takový nerozhodný stav projevuje kolísáním hodnot v určitém rozmezí kolem 0. Pokud korelační koeficienty chápeme jako statistické pojmy, můžeme určit, který interval hodnot kolem 0 je statisticky nevýznamný na určité hladině významnosti. Například lze určit, že korelační koeficienty v intervalu od -0.326 do +0.326 jsou statisticky nevýznamné na hladině významnosti 5% (srov. diskusi k Tabulce 2 a 7).

Korelační koeficienty lze uspořádat do čtvercové matice, jejíž políčko (3,5) ležící na průsečíku 3. řádky a 5. sloupce obsahuje korelační koeficient mezi deskriptorem číslo 3 a 5. Na hlavní diagonále této matice (tj. na spojnici levého horního a pravého dolního rohu matice) budou samé jednotky, protože každý deskriptor maximálně koreluje sám se sebou. Matice je symetrická, tj. pole (3,5) je totožné s polem (5,3); v důsledku toho se každý korelační koeficient v matici (kromě diagonálních) objevuje dvakrát (srov. Tabulku 2 a 7).

Pozn. Některým odborníkům se zdá, že korelační koeficient je výhradně statistický pojem a jeho výpočet (plus všechno, co následuje) patří proto do statistiky. Kdyby to byla pravda, dostali bychom se do nepříjemné situace, protože prakticky na celou vektorovou syntézu by se vztahovala omezení daná různými předpoklady o normálním rozdělení dat apod. Na štěstí lze ale korelační koeficient chápat čistě algebraicky: jako skalární součin odchylek jednotlivých hodnot od průměrných hodnot, přičemž vektory odchylek jsou normalizovány svými eukleidovskými normami (pojem průměru lze rovněž algebraizovat).

14.2.3 Krok 2 (výpočet faktorů)

Korelační matici z kroku 1 tzv. ortogonalizujeme. Jde o vyhledání takových vektorů, které jsou vzájemně (tj. každý jejich pár) nezávislé, tj. jejich skalární součin a tudíž i korelační koeficient je pro každou dvojici ortogonálních vektorů roven 0 (Neustupný 1979, 1993 aj.; detailní výklad pojmu ortogonalizace zde neopakuji). Pokud je korelační matice řádu n (má n řádek a n sloupců), bude takových ortogonálních vektorů maximálně n (proto ani faktorů nemůže být více než n).

Ortogonalizace se provádí matematickým algoritmem pro výpočet tzv. vlastních vektorů a vlastních čísel matice; výpočet je značně komplikovaný, takže i pro malé korelační matice je nutno použít počítač. Většina algoritmů (výpočetních postupů) pro výpočet vlastních vektorů a vlastních čísel řeší oba problémy současně a nelze je oddělit.

Každému vlastnímu vektoru odpovídá jedno tzv. *vlastní číslo*, které souvisí s variabilitou, kterou daný vlastní vektor z celkové variability korelační matice vyjadřuje. Vlastní čísla korelační matice řádu n nabývají hodnot maximálně n , ale zpravidla hodnot menších, protože jejich celkový součet je právě n .

Vlastní vektory jsou určitou formou ortogonálních vektorů a jsou v jednoduchém vztahu k faktorům: z vlastních vektorů dostaneme *faktory* jednoduše tak, že prvky vlastního vektoru postupně násobíme druhou odmocninou z příslušného vlastního čísla (moderní software to samozřejmě dělá automaticky). Takovým násobením vznikne *matice faktorových koeficientů* (neboli tzv. faktorových zátěží), která má n řádek (odpovídajících n deskriptorům) a k sloupců (odpovídajících k faktorům). Faktorové koeficienty nabývají hodnot v rozmezí od -1 do +1, tedy stejných hodnot jako korelační koeficienty; de facto je můžeme považovat za koeficienty, které udávají, jak je daný deskriptor typický

pro daný faktor jako celek. Faktory tedy reprezentují určité „hromadné závislosti“, které se typicky opakují v nějaké množině objektů deskriptivního systému.

Vlastní vektory jsou množiny čísel v absolutní hodnotě menších než jedna; pokud tato čísla násobíme druhou odmocninou vlastního čísla menšího než jedna, dostaneme ještě menší absolutní hodnoty, které mohou být výsledkem „šumu“ v korelační matici. Pokud se tedy nechceme dostat do závislosti na náhodě, musíme zpravidla zvolit počet faktorů menší než n , tj. vyloučit ty faktory, jimž odpovídají velmi malá vlastní čísla.

Neexistuje žádná *jednoduchá* metoda pro rozhodnutí, kolik faktorů máme zvolit jako základ pro další řešení. V praxi se zpravidla aplikuje jedno z následujících čtyř *kritérií* pro počet faktorů:

1. podrží se faktory s vlastními čísly většími než 1;
2. podrží se faktory, jejichž vlastní čísla vysvětlují více než 5% celkové variability obsažené v korelační matici (srov. sloupec 3 v tabulce 3);
3. podrží se faktory, jejichž vlastní čísla vyjadřují dohromady více než $p\%$ celkové variability (kde p je číslo, které si předem stanovíme); srov. sloupec 4 v tabulce 3;
4. sleduje se pokles hodnoty vlastních čísel od největšího k nejmenšímu a počet faktorů se zvolí tam, kde v hodnotách dvou následných vlastních čísel dojde ke „skoku“.

Tato pravidla jsou vysloveně orientační a často neuspokojivá. Jsou případy, kdy i malé vlastní číslo (menší než 1) odpovídá věcně významnému faktoru (Neustupný 1995). Vzhledem k tomu, že výpočet vlastních čísel a vlastních vektorů je na současných počítačích velmi rychlý, je možno experimentovat s řešeními pro různý počet faktorů.

Tyto úvahy by neměly vzbuzovat dojem, že volba počtu faktorů je subjektivní záležitost. Především jsou zde určité meze:

1. jestliže korelační matice má n řádek a sloupců, bude i počet faktorů *maximálně* n ;
2. pokud řešení obsahuje s velmi malých vlastních čísel, bude počet faktorů *maximálně* $n-s$;
3. pokud by $n-s$ faktorů vysvětlovalo příliš velké procento variability korelační matice (např. 96%, srov. sloupec 4 v tabulce 3), nemůžeme ponechat poslední (tj. $n-s$ -tý) faktor, protože musíme počítat s přítomností náhodné variability v matici. Procento náhodné variability musíme nějak odhadnout, nejčastěji podle kvality (významnosti) korelačních koeficientů. Náhodná variabilita ve výši 5% je v mnoha případech nejmenší odhad, ale často dosahuje náhodná variabilita až 30%;
4. ve většině praktických případů podržíme faktory, jejichž vlastní čísla jsou větší než 1. Jestliže takových faktorů je r , bude minimální počet faktorů r .

Odtud vyplývá, že *zpravidla* se počet faktorů bude pohybovat mezi čísly r a $n-s$.

„Správné“ stanovení počtu faktorů se považuje za důležité: volba příliš malého počtu faktorů může způsobit, že nějaký důležitý faktor bude z řešení vypuštěn, volba příliš velkého počtu faktorů může naopak být příčinou zahrnutí „šumu“ do výsledného faktorového řešení. Praktickým prováděním vektorové syntézy ovšem často vzniká dojem, že je víceméně stejně oprávněno několik řešení (např. se 4, 5 nebo 6 faktory), a každé z nich lze často nějak interpretovat.

Pozn. *Začátečníkům v matematických metodách doporučuji, aby termíny jako je vlastní číslo, vlastní vektor, ortogonalizace apod. vzali jen na vědomí do té míry, jak jsou potřebné k ovládní softwaru. Detailnější poučení, zpravidla nepřiliš srozumitelné, lze nalézt v učebnicích lineární algebry, pokračující statistiky a v pracích o faktorové analýze.*

14.2.4 Krok 3 (rotace faktorů)

Jestliže jsme z n vlastních vektorů vybrali k faktorů pro další řešení, předpokládáme, že právě těchto k faktorů dostatečně vyjadřuje variabilitu obsaženou ve výchozí korelační matici. Jednotlivé faktory jsou teď lineárně nezávislé abstraktní objekty. Tyto faktory však mají jednu velkou nevýhodu, která znemožňuje jejich použití v numerické formě: na rozdíl od hodnot vlastních čísel nejsou jejich hodnoty dány jednoznačně, nýbrž závisejí na pořadí sloupců a řádek ve výchozí korelační matici. Naštěstí je možno tzv. *rotací faktorů* dosáhnout jejich transformace do jednoznačného výsledku.

Metod pro rotaci „hrubých“ hodnot faktorů je celá řada, nejpoužívanější mezi nimi je metoda zvaná Varimax. Je založena na myšlence postupných transformací matice faktorových koeficientů do

tzv. *jednoduché struktury*. Jednoduchou strukturou se rozumí taková forma faktorové matice, kde každý deskriptor má velmi vysoký koeficient pouze vzhledem k jednomu faktoru a vzhledem k ostatním faktorům má koeficienty blízké nule. Je samozřejmé, že tento ideální požadavek je zřídka splněn, avšak algoritmus transformuje matici faktorových koeficientů v tomto směru jak dalece je to možné (srov. tabulku 4 a 9). Jak už jsem uvedl, rotace metodou Varimax dává jednoznačný výsledek.

Ačkoliv rotace se snaží o „jednoduchou strukturu“, přece vždy zůstanou v matici faktorových koeficientů případy, kdy jeden deskriptor má významné faktorové koeficienty ke dvěma nebo více faktorům. To je důsledek polytetičnosti deskriptorů, tedy jejich objektivní vlastnosti, kterou nelze žádnými manipulacemi odstranit (někdy se dá tato polytetičnost ovlivnit tzv. šikmými rotacemi).

Faktorové koeficienty, seřazené podle velikosti od +1 do -1, uspořádávají deskriptory podle jejich *typičnosti pro daný faktor* (každý faktor vytváří jiné uspořádání). Je zřejmé, že deskriptory s vysokými kladnými faktorovými koeficienty jsou pro daný faktor vysoce typické, zatímco deskriptory, k nimž patří záporné faktorové koeficienty s vysokými absolutními hodnotami (např. -0.986) jsou pro faktor vysoce *typické pro opak* deskriptorů s vysokými kladnými koeficienty. Deskriptory s koeficienty kolem nuly jsou pro faktor *nevýznamné*, irrelevantní. I v tomto jsou faktorové koeficienty podobné korelačním koeficientům.

Zatímco u korelačních koeficientů lze za určitých předpokladů stanovit statistickou významnost na určité hladině významnosti (viz elementární učebnice statistiky), u faktorových koeficientů je stanovení takové významnosti velmi obtížné. Významnost faktorového koeficientu je ve vztahu k

1. významnosti korelačního koeficientu korelační matice, z níž jsou faktory odvozeny (a tedy k počtu objektů v deskriptivním systému) a
2. k počtu odvozených faktorů a velikosti k nim příslušných vlastních čísel.

V té variantě faktorové analýzy, která je předmětem této příručky, je obtížné udat třeba jen orientačně nějaký vztah. Pokud ovšem faktorové koeficienty seřadíme podle velikosti, často v tomto uspořádání pozorujeme náhlý skok v hodnotách koeficientů (např. skok v řadě 0.986, 0.901, 0.890, 0.854, 0.218, 0.180, 0.130....); do místa takového skoku (mezi 0.854 a 0.218) se pak často klade hranice významnosti. Transformace faktorů do jednoduché struktury prostřednictvím rotací (viz výše) způsobuje, že takové skoky jsou u faktorů velmi častým jevem, protože rotace mají tendenci transformovat absolutní hodnoty faktorových koeficientů buď na nulu nebo na jednotku.

Vodítkem pro rozhodování o významnosti je fakt, že vysoké koeficienty (s absolutními hodnotami 0.9 a 0.8) bývají prakticky vždycky významné a koeficienty v absolutní hodnotě menší než 0.1 bývají obvykle nevýznamné. Často se používají různá věcná kritéria a postupy obdobné validaci faktorů (viz krok 5). Faktor jako celek je tedy objektivním produktem algoritmu vektorové syntézy, ale určení deskriptorů, které jsou pro tento faktor významně typické, je přes uvedenou „skokovitost“ rotovaných faktorů jednou z málo objektivních voleb vektorové syntézy.

Některé faktory jsou tzv. *monopolární* (tj. jejich významné faktorové koeficienty jsou buď jen kladné nebo jen záporné; některé jsou *bipolární* (mají významné koeficienty jak s kladným tak i se záporným znaménkem). Deskriptory, které mají významné koeficienty opačných znamének, tvoří strukturální opozice.

Někdy dochází z numerických důvodů ke změně znamének na opačná u všech faktorových koeficientů některého faktoru; v takovém případě se ale na interpretaci faktoru nic nemění. Už z tohoto důvodu nelze vyvozovat žádné závěry z toho, že „významné“ koeficienty nějakého faktoru jsou kladné nebo záporné. Tato skutečnost se přenáší i na faktorová skóre (srov. krok 4).

14.2.5 Krok 4 (faktorová skóre)

Faktorové koeficienty (faktorové zátěže) udávají, jak je který deskriptor původního deskriptivního systému z kroku 1 typický pro daný faktor jako celek. Můžeme si tudíž představit každý faktor jako nějaký abstraktní objekt (ve skutečnosti je to ovšem struktura) a faktorovou zátěž jako určitou korelaci tohoto abstraktního objektu s vybraným deskriptorem. To je jeden aspekt problému.

Druhý aspekt je typičnost každého z faktorů pro každý z objektů původního deskriptivního systému. Existuje číslo, tzv. faktorové skóre, které tuto typičnost „měří“. Jestliže objektů je m a faktorů k , pak faktorových skóre bude $m \times k$, tedy často velký počet.

Teoreticky mohou faktorová skóre nabývat jakýchkoliv reálných hodnot (záporných i kladných), ale za určitých podmínek jich bude cca 62% v intervalu od -1 do +1, 95% v intervalu od -2 do +2 atd. Průměr hodnot faktorových skóre pro každý faktor je 0 a směrodatná odchylka je 1.

14.2.6 Krok 5 (validace)

Pokud faktorovou analýzu chápeme jako konfirmatorní statistickou metodu testující nějaké struktury (např. struktury získané předtím exploratorní faktorovou analýzou), lze za určitých podmínek získat zamítnutí nebo potvrzení statistické významnosti výsledků. Tento postup je v případě archeologických dat často velmi problematický.

Proto jsem navrhl pro testování výsledků sledovat nikoliv statistickou významnost, nýbrž *validitu* (platnost). Validita se zjišťuje pomocí tzv. externí evidence, tj. na základě objektů nebo deskriptorů, které jsou nezávislé na těch, na nichž je postavena faktorová analýza, které ale přesto vytvářejí stejné struktury. Jestliže je totiž náhodné vytvoření nějaké struktury málo pravděpodobné, pak pravděpodobnost, že *náhodou* bude tatáž struktura vytvořena dvakrát na základě dvou nezávislých souborů evidence, je již zanedbatelně malá.

Pokud např. máme faktorovou analýzu založenou na objektech, jimiž jsou jednotlivé hroby a deskriptory jsou druhy hrobových přídavků (keramiky, ozdoby, nástrojů a zbraní), pak nezávislou externí evidencí může být antropologické (nebo jiné) určení pohlaví a věku osob pohřbených v jednotlivých hrobech. Pokud se např. ukáže, že všechny hroby dospělých mužů mají vysoká kladná faktorová skóre k faktoru č. 1, a že hroby dospělých žen mají vysoká záporná skóre k faktoru č.3, pak je velmi nepravděpodobné, že faktor 1 a 3 by vznikl v důsledku náhody.

Velmi významným druhem externí evidence je poloha objektů v prostoru. Pokud např. dostaneme faktorové řešení na základě sběrových čtverců (objekty) a střepů pravěkých kultur (deskriptory), můžeme za externí evidenci zvolit polohu sběrových čtverců v terénu. Vyneseme na mapu faktorová skóre jednotlivých čtverců a sledujeme, jak se prostorově shlukují (srov. obr. 1-3). Pokud se vysoká kladná skóre shlukují ve skupině sousedních čtverců, a vysoká záporná skóre v jiné skupině, pak je velmi nepravděpodobné, že by faktorové řešení vzniklo náhodou.

14.3 Přednosti formalizovaného hledání struktur

Metody vektorové syntézy nejsou zázračným řešením, které by mohlo nahradit archeology a odstranit jejich pochybnosti, které jsou vždy nutně spojeny s poznáním na základě empirického materiálu. Jsou však velmi efektivní metodou k hledání archeologických struktur, a často také metodou efektivní. Je ovšem důležité pochopit, že se jedná o prostředky k syntéze formálních struktur, nikoliv o neformální interpretaci; ta je samostatným postupem, který *následuje* až po nalezení formálních struktur (Neustupný 1986, 1993 aj.).

Někteří archeologové se domnívají, že formalizované metody odstraňují subjektivitu. To je pravda jen částečně: tyto metody *předpokládají* určitý deskriptivní systém (v kroce 1), který nemůže být z principiálních důvodů plně „objektivní“ v tom slova smyslu, jak objektivitu chápou empiristé. Je ovšem pravda, že v okamžiku, kdy deskripce je dokončena, probíhá algoritmus vektorové syntézy zcela automaticky a tudíž „objektivně“. Jediná možnost, jak archeolog může ovlivnit výsledky, je volbou počtu faktorů. Při diskusi kroku 2 jsem však ukázal, že možnost subjektivního zásahu je omezená a v určitém slova smyslu lze přijmout několik řešení konkrétních úloh.

Pokud není počet faktorů hrubě podceňen, faktorové řešení dává možnost posoudit, zda *jsme* v daném deskriptivním systému nepřehlédlí nějaký zdroj variability (nějaký faktor) nebo zda naopak nepředpokládáme nějaký náhodný zdroj variability (neexistující faktor). Praxe ukazuje, že archeologové mají výraznou tendenci počet faktorů, ovlivňující variabilitu jejich kontextů, jak podceňovat, tak i přeceňovat.

Nelze nezmínit, že metody vektorové syntézy jsou velmi efektivní: jakmile je hotova deskripce v podobě databáze, je další řešení otázkou vteřin nebo minut a je v rámci téže metody vždy plně

reprodukovatelné. Syntézu struktur může provádět i nespécialista a během několika týdnů dosáhnout výsledků, kterých jinak dosahovaly generace tradičních archeologů (srov. Neustupný 1978). Znovu zdůrazňuji, že nespécialista nemůže efektivně provést deskripci a zejména ne následnou interpretaci výsledných faktorů.

14.4 Prostorový aspekt archeologických struktur

Prostorové struktury nebyly doposud systematicky popsány, avšak zjevně k nim patří především různé druhy shlukování. Pokud nějaká formální struktura, například hroby s vysokými faktoro-
vými skóre na faktor 1 určitého faktorového řešení, se na pohřebišti vyskytují v sousedních hrobech (tvoří shluk hrobů), pak prostorová struktura (shluk) validuje formální strukturu a naopak. Je totiž vy-
soce nepravděpodobné, že formální a prostorová struktura by mohla korelovat náhodou. To ovšem
platí, pokud formální a prostorové deskriptory jsou vzájemně nezávislé.

Na těchto skutečnostech je založeno simultánní hledání struktur ve formálním a geografickém
prostoru. To dovoluje posoudit formální a prostorový aspekt archeologických struktur současně
a představuje proto možnost *vyčerpání celé pozorovatelné variability pramenů*. Takové komplexní
hledání struktur využívá současně dva druhy softwaru, který je v archeologických aplikacích nejpro-
duktivnější: software pro multivariatní analýzy a pro GIS (geografické informační systémy). Určité
uplatnění myšlenky komplexního hledání struktur obsahují příklady 3 a 4. Celá metoda, využívající
pojmu polygonu, je rozvinutím jak vektorové syntézy, tak GIS; je podrobněji popsána jinde
(Neustupný 1996, částečně Neustupný a Venclová 1996).

14.5 Příklady

14.5.1 Velikost zlomků laténské keramiky

Jde o velmi jednoduchý příklad, u něhož je možné určité pochopení „zdravým rozumem“.

14.5.1.1 Krok 0

Vladimír Salač detailně změřil, zvážil a popsal 3741 zlomků laténské keramiky z Března
u Chomutova. Z jeho detailního popisu (viz Salač 1996 v tomto svazku) zde vybíráme jen 7 proměn-
ných, použitých jako deskriptory (v následujících výpisech z počítačového softwaru jsou označeny
anglickými zkratkami):

LENGTH – délka střepu, průměr v souboru 5.2 cm

THICKN – síla střepu, v průměru 9.4 mm

WEIGHT – váha střepu, průměr v souboru 22.6 g (standardní odchylka ukazuje na velkou va-
riabilitu!)

DECOR – zdobenost střepu; zdobeno je 7.8% střepů

RIM – okrajovost zlomků; z okraje nádoby pochází 13.3% střepů

BOTTOM – přítomnost zlomků dna; dna je rozpoznatelné u 8.2% střepů

VESS_DET – možnost určit druh nádoby; druh nádoby bylo možno určit u 13.4% střepů

Přehled podává tab. 1.

	průměr	směrodatná odchylka
LENGTH	5.15531	2.09652
THICKN	9.37610	2.52293
WEIGHT	22.58514	39.94085
DECOR	0.07752	0.26745

	průměr	směrodatná odchylka
RIM	0.13285	0.33946
BOTTOM	0.08233	0.27490
VESS_DET	0.13419	0.34090
Počet objektů = 3741		

Tab. 1. Přehled deskriptorů použitých ve faktorové analýze.

Vstupem do následných kroků byla databáze s 3741 záznamem (objekty), z nichž každý byl popsán všemi sedmi deskriptory z Tabulky 1. U všech deskriptorů byla známa hodnota pro všechny objekty, tj. v deskriptivním systému nebyly žádné chybějící hodnoty. Všechny hodnoty byly reálná čísla: v případě výzdoby, okraje, dna a druhu nádoby bylo dosazeno číslo 1, pokud střeš byl zdobený, okrajový, ode dna a bylo možno rozeznat druh nádoby. Číslo 0 znamená, že tomu bylo opačně. V důsledku tohoto pojetí dat databáze reprezentovala lineární vektorový prostor, takže použití metod založených na ortogonalizaci je oprávněné.

Určitý problém tvoří dvojice RIM a BOTTOM. Teoreticky může každý střeš zasahovat od okraje až ke dnu, ale vzhledem k délkám střešů je to málo pravděpodobné. Může proto vzniknout názor, že jde o dva stavy téže nominální proměnné. Z instruktivních důvodů jsme přesto oba členy této dvojice v deskriptivním systému ponechali.

14.5.1.2 Krok 1

Na základě deskriptivního systému z kroku 1 vypočítáme korelační matici (tab. 2). Nejnižší koeficient je -0.109 (mezi RIM a BOTTOM), nejvyšší 0.992 (mezi RIM a VESS_DET). Velká část koeficientů je statisticky významná na hladině 5% (a kromě jedné z korelací dokonce na hladině významnosti 1%).

	LENGTH	THICKN	WEIGHT	DECOR	RIM	BOTTOM	VESS_DET
LENGTH	1.00000						
THICKN	0.27992	1.00000					
WEIGHT	0.54203	0.23639	1.00000				
DECOR	0.11300	-0.01548	0.04011	1.00000			
RIM	0.09385	0.00596	0.01909	-0.01628	1.00000		
BOTTOM	0.20977	0.28843	0.18485	-0.03228	-0.10864	1.00000	
VESS_DET	0.09803	0.00939	0.02212	-0.01441	0.99193	-0.10365	1.00000

Tab. 2. Korelační matice mezi deskriptory příkladu 1. Je tištěna jen spodní trojúhelník matice, horní trojúhelník je s ním symetrický. Kurzivou jsou vyznačeny koeficienty, které nejsou statisticky významné na hladině významnosti 5%.

Vzhledem k tomu, že jde o jednoduchý příklad malého rozsahu, je možné poměrně mnoho vyčíst už prohlížením korelační matice; taková metoda je však velmi problematická u větších matic, jejichž variabilita je určena větším počtem faktorů.

14.5.1.3 Krok 2

Řešením problému vlastních čísel získáme informace na tabulce 3. Je zřejmé, že jen první tři faktory budou mít vlastní čísla větší než 1, ale celých 6 faktorů se na celkové variabilitě obsažené v matici korelačních koeficientů podílí více než 5 procenty; na druhé straně těchto 6 faktorů vyčerpává dohromady 99.9% variability, což je určitě příliš mnoho (nezbylo by téměř nic na variabilitu náhodnou). Oproti tomu první tři faktory s vlastními čísly >1 vyčerpávají 71.2% celkové variability (srov. poslední sloupec v tabulce 3), což může představovat podíl nenáhodné variability; pro další řešení proto zvolíme 3 faktory.

Faktor	vlastní číslo	% variace	kumulativní %
1	2.03649	29.1	29.1
2	1.88509	26.9	56.0
3	1.06155	15.2	71.2
4	0.86629	12.4	83.6
5	0.69743	10.0	93.5
6	0.44510	6.4	99.9
7	0.00805	0.1	100.0

Tab. 3. Vlastní čísla korelační matice z Tabulky 2.

Matici vlastních vektorů zde nereprodukuje, ale lze si o ní učinit určitý obraz z tabulky 5, která obsahuje všech 7 faktorů (rotovaných).

14.5.1.4 Krok 3

Rotací prvních 3 faktorů dostaneme matici faktorových koeficientů na Tabulce 4.

Faktor 1 má dominantní koeficienty k deskriptorům RIM a VESS_DET, mnohem nižší koeficient k LENGTH. Určitým opakem je BOTTOM.

Faktor 2 má relativně menší (ale stále ještě velmi vysoké) koeficienty vzhledem k LENGTH, WEIGHT, THICKN a BOTTOM.

Faktor 3 má vysoký koeficient k DECOR a slabší k LENGTH a WEIGHT. V opozici k němu (se zápornými koeficienty) stojí BOTTOM a THICKN.

	faktor 1	faktor 2	faktor 3
LENGTH	0.11572	0.78931	0.24407
THICKN	-0.00590	0.62500	-0.30587
WEIGHT	0.02808	0.75709	0.18910
DECOR	-0.05356	0.09498	0.84905
RIM	0.99365	0.01372	-0.00627
BOTTOM	-0.17945	0.54717	-0.39266
VESS_DET	0.99332	0.01936	-0.00615

Tab. 4. První tři faktory rotované metodou Varimax. Sloupce obsahují faktorové koeficienty (neboli faktorové zátěže).

Pro srovnání na tabulce 5 uvádíme řešení pro všech 7 faktorů. Při srovnávání s řešením ve třech faktorech budeme za číslem faktoru psát lomítko a počet extrahovaných faktorů (tedy faktor 1/3 je první faktor z tabulky 4 a faktor 1/7 je první faktor z tabulky 5). Především vidíme, že *faktor 7/7*, který odpovídá velmi malému vlastnímu číslu z tabulky 2, je skutečně zanedbatelný. Faktor 6/7 do značné míry opakuje faktor 3/7, i když s jiným důrazem na jednotlivé deskriptory. Faktor 5/7 víceméně osamostatňuje deskriptor DECOR.

Faktor 1/3 odpovídá svou strukturou faktoru 1/7,

faktor 2/3 faktoru 2/7 a 3/7,

kladný pól faktoru 3/3 (kladné koeficienty) zhruba odpovídá faktoru 5/7, a

záporný pól faktoru 3/3 zhruba faktoru 4/7.

Tyto korepondence jsou jen přibližné a z kvantitativního hlediska problematické. Nicméně ukazují, že řešení ve třech faktorech pravděpodobně v dostačující míře reprodukuje variabilitu souboru a to zejména v situaci, kdy nevíme, jak velké faktorové koeficienty máme ještě brát za směrodatné.

Obě řešení vyzvedávají dvojice RIM-VESS_DET, BOTTOM- THICKN, a DECOR-LENGTH, a čtveřici LENGTH-THICKN-WEIGHT-BOTTOM. Z tohoto hlediska musíme také posuzovat význam „správné“ volby počtu faktorů pro celkové řešení. Pokud se spokojíme s kvalitativním řešením, není

chybná volba počtu zvolených faktorů příliš rušivá; pokud nám však záleží na kvantitativních charakteristikách faktorových koeficientů a faktorových skóre, je počet extrahovaných faktorů důležitý.

	faktor 1	faktor 2	faktor 3	faktor 4	faktor 5	faktor 6	faktor 7
LENGTH	0.06413	0.27981	0.13069	0.09648	0.06282	0.94194	0.00004
THICKN	0.00534	0.09761	0.97824	0.13896	-0.01076	0.11857	0.00003
WEIGHT	0.00781	0.95451	0.10284	0.07884	0.01485	0.26802	0.00002
DECOR	-0.01116	0.01424	-0.01005	-0.01837	0.99820	0.05320	0.00002
RIM	0.99630	0.00565	0.00218	-0.04520	-0.00792	0.03499	-0.06345
BOTTOM	-0.06845	0.07421	0.13766	0.98128	-0.01936	0.08700	0.00005
VESS_DET	0.99637	0.00732	0.00454	-0.04081	-0.00614	0.03802	0.06343

Tab. 5. Matice faktorů rotovaných metodou Varimax pro případ přijetí 7 faktorů.

Krok 4 a 5 v souvislosti s tímto příkladem nediskutujeme.

14.5.1.5 Interpretace

V následujících odstavcích velmi stručně a zjednodušeně naznačím vysvětlení získaných faktorů; takové vysvětlení není ovšem součástí syntézy struktur, nýbrž až jejich interpretace v termínech živé kultury. Výpověď faktorů tohoto příkladu o živé kultuře je omezená, protože je to *především* výpověď o zlomcích; informace svědčí proto spíše o transformačních procesech, jimiž se živá kultura mění na mrtvou.

Ve svém kroku 3 vyčlenila faktorová analýza tři faktory, které jsou zcela formálními strukturami: tato metoda nemá vůbec žádné informace o tom, co znamenají jednotlivé deskriptory a pracuje výhradně s výchozí korelační maticí, u níž archeologický význam jednotlivých řádek a sloupců do metody nijak nevchází. Skutečnost, že např. ve faktoru 3/3 je spojena výzdoba s délkou, vyplývá pouze z velikosti korelačních koeficientů; zda 4. řádek matice znamená výzdobu (DECOR) nebo např. plochu střepu, je faktorové analýze zcela lhostejné.

Faktor 1/3 je nepochybně důsledkem skutečnosti, že druh nádoby (VESS_DET) bylo možno určit téměř výlučně u okrajových zlomků (RIM). Současně tento faktor napovídá, že okrajové zlomy jsou delší než ostatní.

Pro živou kulturu to znamená, že druhy nádob se v rámci použitého deskriptivního systému „strukturálně“ liší především svými okraji. Tento poznatek je triviálně známý, ale nemusí *ve stejné míře* platit pro jiné kultury než je laténská.

Faktor 2/3 je „hmotnostní“; říká, že délka, síla a váha střepů spolu úzce souvisí a všechny nabývají vysokých hodnot u zlomků den. Odtud vyplývá např. to, že tenké střepy se rozpadají na menší kusy než střepy silné, a zlomky den do značné míry sledují toto pravidlo.

V oblasti živé kultury (při výrobě laténské keramiky) to může naznačovat určitou ustálenou technologii, např. to, že síla střepu odpovídala velikosti nádoby (pokud přijmeme hypotézu, že velikost střepu koreluje s jeho zakřiveností a tudíž i s velikostí nádoby).

Faktor 3/3 vyčleňuje zdobené zlomky (relativně delší než ostatní) a staví je do protikladu se střepy ode dna a silnými zlomky.

Znamená to, že zdobená keramika se pravděpodobně technologicky odlišovala od průměru keramiky (viz faktor 2/3), že byla relativně tenkostěnná.

Chtěl bych upozornit, že tyto poznatky platí pouze ve spojení s použitým deskriptivním systémem; jeho rozšířením nebo zúžením bychom mohli dostat poněkud jinou podobu faktorů. Pokud ovšem zůstaneme u použitého deskriptivního systému, můžeme mít relativní jistotu, že soubor našich zlomků neobsahuje už další „hromadné“ závislosti mezi deskriptory než ty, které jsme popsali. To je velmi důležitý závěr z každé faktorové analýzy.

Mohlo by se zdát, že vyjmenované výsledky jsou triviální a proto postradatelné. To je nepochybně způsobeno elementarností našeho příkladu. Musíme vzít v úvahu, že faktorová analýza je

schopna správně „vyhmátnout“ právě ty vztahy, které určují variabilitu kontextu. Závěr, že variabilitu uspokojivým způsobem vyjadřují tři faktory, jejichž konkrétní podoby lze určit, je myslím netriviálním řešením.

Chtěl bych upozornit, že faktory lze zpětně porovnat se strukturou matice korelačních koeficientů (tab. 2). Odvodit tyto faktory z matice korelací prohlížením nebo jednoduchými manipulacemi s maticí ovšem nelze. Neexistuje totiž jednoduchá metoda, kterou by bylo možno „rozdělit“ korelační koeficient do několika faktorů (srov. např. přítomnost evidentně významných faktorových koeficientů deskriptorů LENGTH nebo BOTTOM ve všech třech faktorech; pokud neznáme výsledky vektorové analýzy, není důvodu, proč bychom takto nerozdělovali např. deskriptor RIM mezi všechny tři faktory).

14.5.2 Loděnický potok

Projekt Loděnický potok spočíval ve sledování pravěkého až novověkého osídlení malého regionu západně od Prahy (asi 8 x 8 km) pomocí povrchových sběrů. Sběry a následné zpracování nálezů provedla v letech 1993-1995 N. Venclová, která také pořídila databázi nálezů včetně nálezů ze starších výzkumů (Venclová 1994, 1995). Následující odstavce navazují na počítačové zpracování nálezů z doby laténské pomocí faktorové analýzy a GIS (Neustupný a Venclová 1996).

14.5.2.1 Krok 0

Analýza v terénu byla provedena metodikou vypracovanou M.Kunou (1994) s modifikacemi. Pás o ploše asi 43 km² byl rozdělen náhodným vzorkováním do velkých polygonů o rozměrech několika až mnoha hektarů, a v těch pak byly vytyčeny základní polygony (sběrové čtverce) o ploše přibližně 1 ha. Sběry byly provedeny na ploše asi 12 km²; počet základních polygonů byl 1361. V každém polygonu (sběrovém čtverci) byly zaznamenány všechny archeologické nálezy a jejich popis byl uložen v databázi.

Pro studium laténského osídlení pak byly z uvedené databáze vybrány následující deskriptory:

1. laténská keramika
2. SAPCO – středová kolečka vyřiznutá ze švartny (sapropelitu) při výrobě švartnových náramků
3. SAPW – kusy švartny se stopami opracování
4. SAPN – kusy švartny beze stop opracování
5. S_SLAG – kusy hutnické strusky

Přehled o četnosti těchto deskriptorů ve 1361 objektech (sběrových čtvercích) podává tab. 6. Deskriptivní systém ve formě databáze se stal základem dalších kroků. Problémy se zařazením kategorií SAPW, SAPN a S_SLAG do doby laténské diskutuje uvedený článek E.Neustupného a N.Venclové (1996).

Deskriptor	počet kusů	max. v 1 čtverci	průměr v 1 čtverci	STD	neprázdných čtverců
keramika	183	10	0.13	0.62	101
SAPCO	318	23	0.23	1.33	112
SAPW	1152	40	0.85	2.58	357
SAPN	1438	27	1.06	2.42	477
S_SLAG	3358	397	2.47	20.03	282

Tab. 6. Četnosti deskriptorů laténského období v projektu Loděnice.

14.5.2.2 Krok 1

Na základě deskriptivního systému byly vypočítány korelace mezi deskriptory. Čistě formálně, z hlediska tradiční statistiky, jsou tyto korelace velmi hodnotné, tj. všechny až na jednu jsou sta-

tisticky významné na hladině významnosti menší než 1% (jedna na hladině nepatrně vyšší než 1%). Přehled podává tab. 7.

	LATEN	SAPCO	SAPW	SAPN	S_SLAG
LATEN	1.00000				
SAPCO	0.33241	1.00000			
SAPW	0.38608	0.50736	1.00000		
SAPN	0.07351	0.11255	0.39802	1.00000	
S_SLAG	0.24380	0.09584	0.18066	0.06253	1.00000

Tab. 7. Projekt Loděnice, matice korelačních koeficientů. Kurzivou je vyznačen koeficient, který je statisticky významný mírně nad hladinou významnosti 1%; ostatní pod 1%.

14.5.2.3 Krok 2

Přehled vlastních čísel korelační matice z tabulky 7 podává tab. 8. Dvě vlastní čísla jsou větší než 1, ale třetí je jednotce blízké; první tři faktory přitom obsahují téměř 80% variability matice. Při zahrnutí čtvrtého faktoru (který ještě obsahuje 12.6% variace) by faktorové řešení obsahovalo 92.2% variability, což se nám zdálo být příliš mnoho (zbývalo by jen poměrně málo na náhodnou variabilitu). Proto jsme pro další řešení zvolili 3 faktory.

Faktor	vlastní číslo	% variace	kumulativní %
1	2.04034	40.8	40.8
2	1.02865	20.6	61.4
3	0.90799	18.2	79.5
4	0.63074	12.6	92.2
5	0.39228	7.8	100.0

Tab. 8. Projekt Loděnice, vlastní čísla korelační matice.

14.5.2.4 Krok 3

Rotací prvních tří faktorů jsme dostali matici faktorových koeficientů (tab. 9). Tyto koeficienty udávají, jak typické jsou jednotlivé deskriptory pro abstraktní struktury, které se jmenují faktor 1, 2 a 3. Všechny tři faktory jsou v podstatě monopolární.

Faktor 1 charakterizují deskriptory SAPCO, SAPW a LATEN. *Faktor 2* je spojen s deskriptory SAPN a SAPW, zatímco u *faktoru 3* má největší koeficienty struska (S_SLAG), laténská keramika (LATEN) a možná je významná i opracovaná švartna (SAPW).

Některé deskriptory jsou vzhledem k faktorům monotetické (tj. v podstatě definují jen jeden faktor: S_SLAG, SAPN, SAPCO ?), jiné jsou polytetické (tj. jsou charakteristické pro dva nebo tři faktory: LATEN, SAPW).

	Faktor 1	Faktor 2	Faktor 3
LATEN	0.68737	-0.07438	0.38479
SAPCO	0.85588	0.06340	-0.09082
SAPW	0.69719	0.51518	0.10173
SAPN	0.04151	0.95904	0.03106
S_SLAG	0.06551	0.06173	0.95589

Tab. 9. Projekt Loděnice, faktorové koeficienty tří faktorů (rotace Varimax).

14.5.2.5 Krok 4

V návaznosti na faktorové koeficienty byla vypočítána i faktorová skóre udávající, jak je který výzkumný polygon (čtverec) charakteristický pro každý z popsaných faktorů. Matice faktorových skóre má v tomto případě 1361 řádek a 3 sloupce; na Tabulce 10 podávám ukázkou její části. Je zřejmé, že snažit se o rozpoznání nějakých struktur v tak velké matici jejím prohlížením není možné.

polygon	faktor 1/3	faktor 2/3	faktor 3/3
13201	0.61955	2.95895	-0.05281
13202	-0.60569	2.20463	-0.08886
13204	-0.02024	1.40076	-0.27040
13205	-0.37242	1.62095	0.22093
13206	-0.94027	5.29977	-0.12733
13207	-1.26731	7.41805	0.27752
13208	-0.06373	6.05745	-0.28020
13209	-0.33033	1.61593	-0.05580
13210	0.03730	1.87977	-0.28655
13211	-0.39125	1.97044	-0.14907
20101	4.63314	3.06919	1.36573
20102	6.69455	3.91359	9.34419
20103	0.02088	-0.28708	19.17913
20104	0.86845	1.43207	0.00991
20105	4.86920	0.75457	1.46807
20106	6.49341	7.01030	-0.98888
20107	0.57147	0.18383	0.26888
20108	1.38506	-0.47819	-0.08216
20111	1.80938	1.61841	-0.83555
20113	2.84496	-1.80842	2.45091
20120	0.29631	1.70891	2.42236
20121	0.99193	0.05949	0.92742

Tab. 10. Projekt Loděnice, ukáзка části matice faktorových skóre.

14.5.2.6 Krok 5

Faktorová skóre jsou čísla, která charakterizují každý ze základních (sběrových) polygonů ve formálním prostoru tří laténských deskriptorů. Pokud zjištěné faktory (faktorová skóre) mají v geografickém prostoru nenáhodné rodění, je velmi málo pravděpodobné, že faktorové řešení (založené v našem případě na pěti družích nálezů) by vzniklo náhodou: poloha sběrových polygonů totiž představuje vzhledem k nálezům nezávislou externí evidenci. Pokud by ovšem nastal náhodný rozptyl faktorových skóre v geografickém prostoru, nemohli bychom touto externí evidencí validitu (platnost) faktorového řešení prokázat. Museli bychom tedy předpokládat, že faktorové řešení, založené na formálních deskriptorech (tj. na družích nálezů), je důsledkem náhody a není tudíž validní (platné). Jiné možné vysvětlení je, že toto faktorové řešení je výrazem něčeho jiného než prostorového chování lidí.

Tento problém lze testovat tak, že každému sběrovému polygonu přiřadíme postupně jeho faktorové skóre vzhledem k faktoru 1/3, pak 2/3 a nakonec 3/3. Získáme tak tři mapy (obr. 1-3), na nichž lze náhodnost či nenáhodnost prostorového rozdělení faktorových skóre sledovat. Na těchto mapách vidíme vysoký stupeň shlukování sběrových polygonů, což je projev nenáhodného rozdělení faktorových skóre. Můžeme proto uzavřít, že zjištěné faktorové řešení je validní, platné.

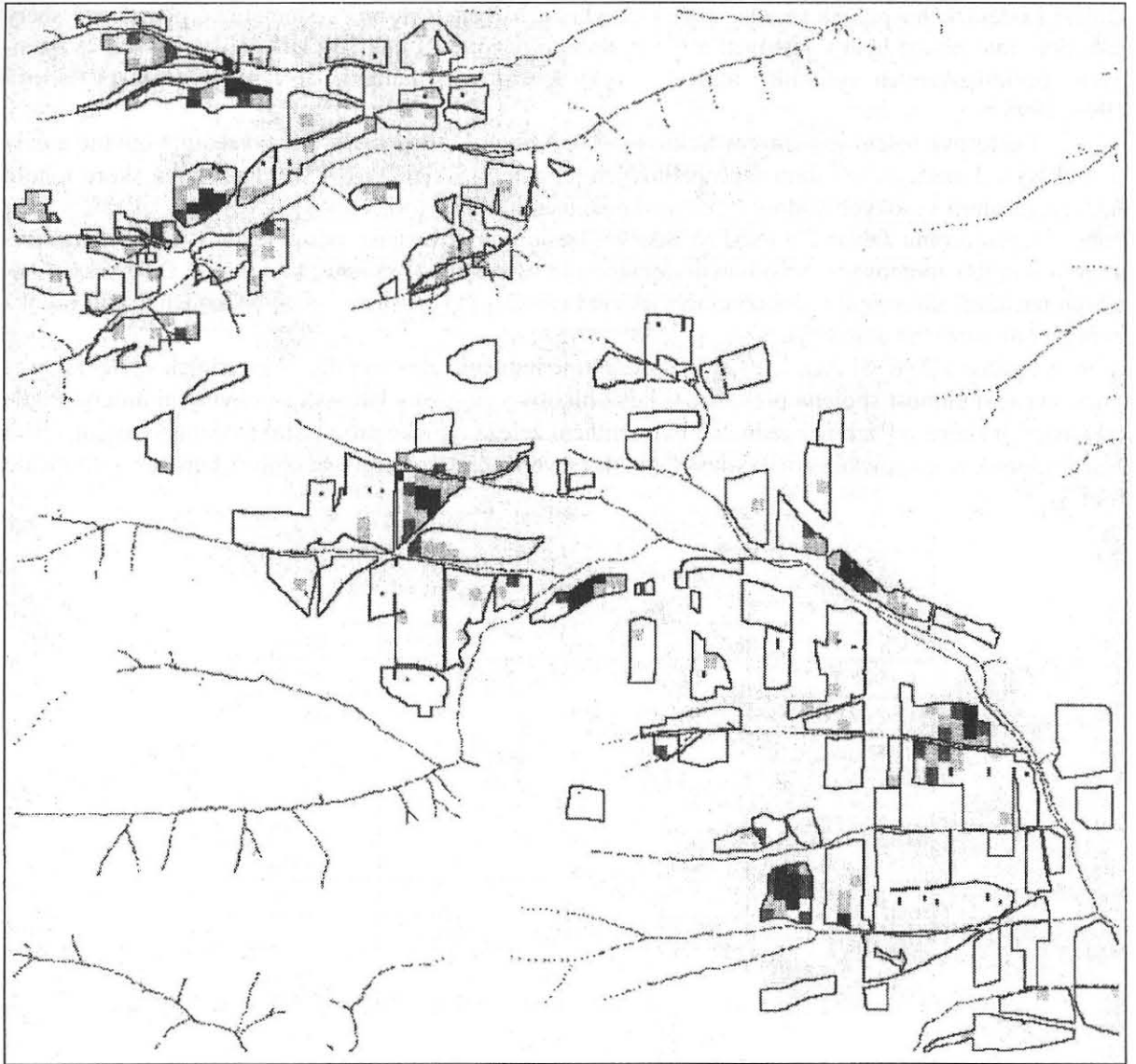
14.5.2.7 Interpretace

Interpretaci zjištěných struktur (formálních i prostorových) jsem společně s N.Venclovou detailně rozvedl na jiném místě (Neustupný a Venclová 1996); zde proto uvádím jen některé výsledky. Oblast Loděnického potoka (a jeho okolí) nevybrala N.Venclová pro systematické povrchové sběry náhodou: tato oblast byla v latěnu B a C (po dobu o něco delší než 100 let) dějištěm dvou významných specializovaných výrobních aktivit: výroby švartnových náramků a výroby železa (Venclová 1994, 1995 aj.).

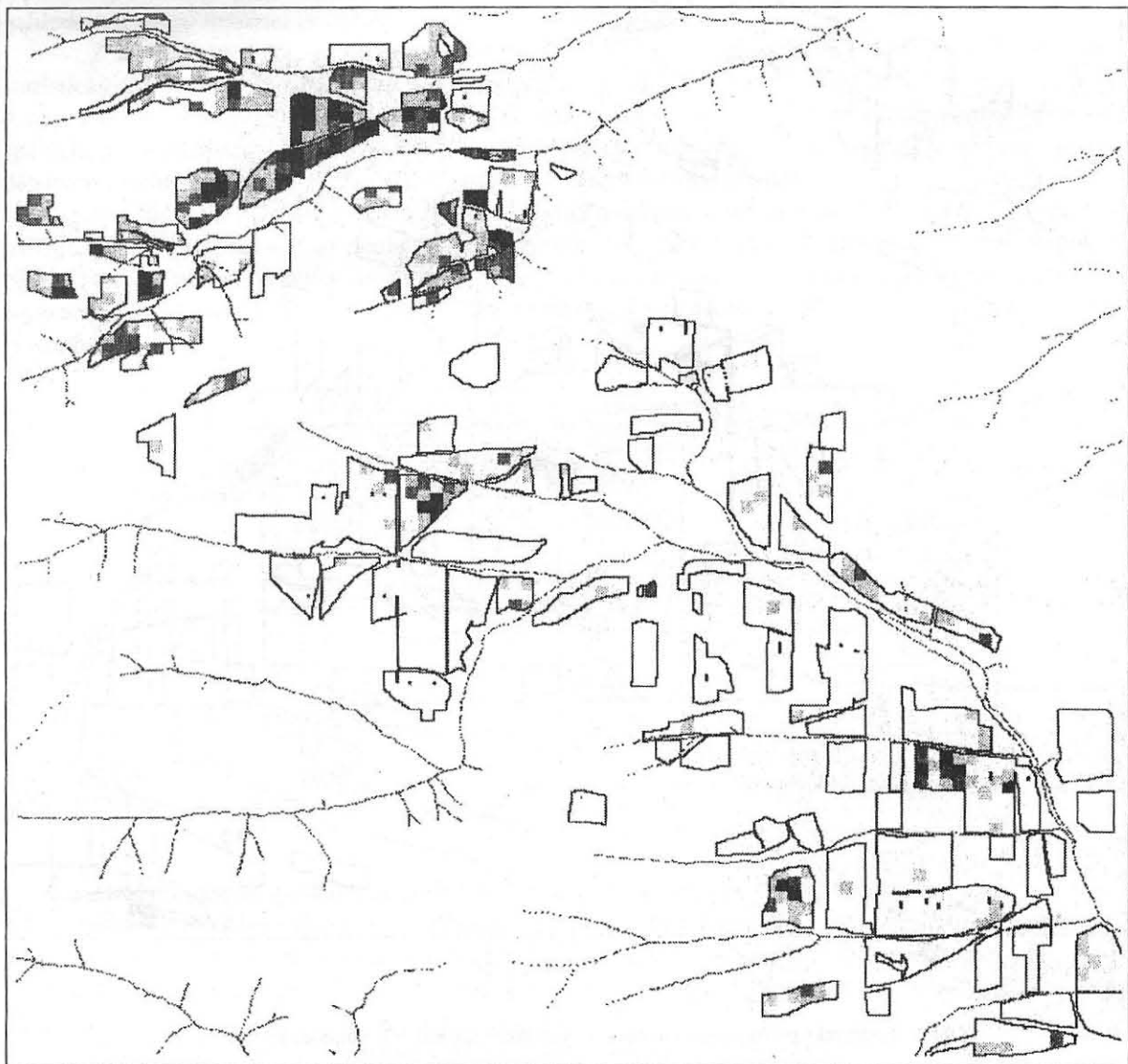
Faktorové řešení je výrazem těchto aktivit. *Faktor 1/3* nepochybně reprezentuje obytné areály (LATEN) s domácí výrobou sapropelitových náramků (SAPCO, SAPW). Faktorová skóre tohoto faktoru dosahují vysokých hodnot v celém zkoumaném regionu (obr.1).

Oproti tomu *faktor 2/3* (SAPN, SAPW) lze interpretovat jako odraz těžby švartny a počátečních fází jejího zpracování. Jeho hlavní doménou je severní část regionu, kde se nad Bakovským potokem nacházel sapropelit v dosažitelném uložení (obr.2). Do prostoru vlastního Loděnického potoka byla zřejmě surovina donášena.

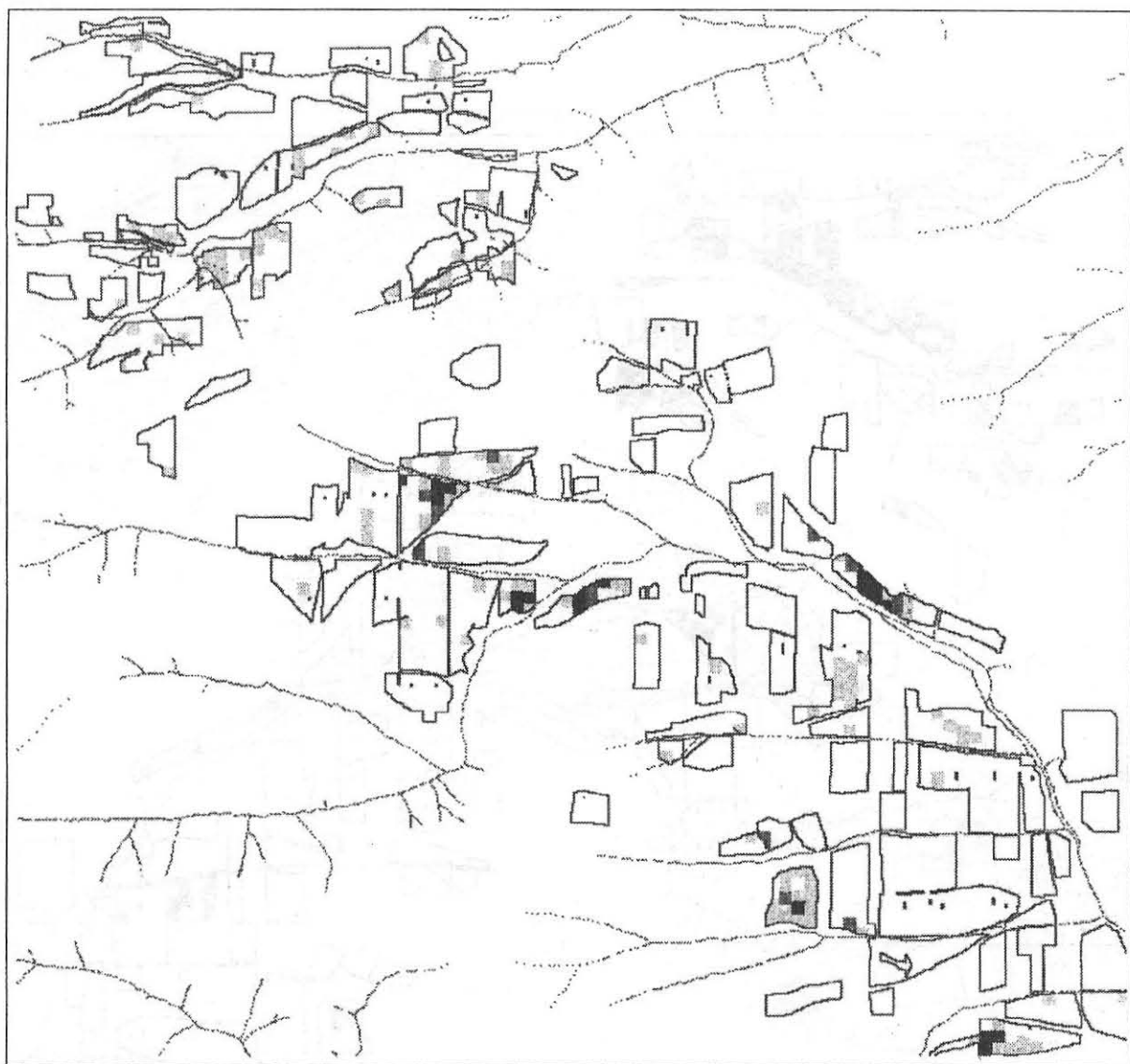
Faktor 3/3 (S_SLAG, LATEN) reprezentuje hutnění železa. Podle dosavadních výsledků byla i tato výrobní činnost spojena převážně (i když nikoliv výlučně) s latěnskými obytnými areály. Podle faktorových skóre větších než jednotka bylo hutnění železa typické pro vlastní Loděnický potok (jižní část regionu), a netypické pro Bakovský potok (severní část regionu), ačkoliv i tam se vyskytovalo (obr.3).



Obr. 1. Loděnický potok, rozložení faktorových skóre faktoru 1/3. Jsou vyznačeny potoky (na jihu Loděnický, na severu Bakovský) a sběrové polygony. Černé plochy: skóre > 2, tmavošedé: skóre od 1 do 2, světlešedé: skóre od 0 do 1, bílé: skóre < 0.



Obr.2. Loděnický potok, rozložení faktorových skóre faktoru 2/3. Stupnice jako u obr. 1.



Obr. 3. Loděnický potok, rozložení faktorových skóre faktoru 3/3. Stupnice jako u obr. 1 .

14.5.3 Vikletice (šňůrové pohřebiště)

Šňůrové pohřebiště ve Vikleticích (Buchvaldek a Koutecký 1970) sestávalo z více než 150 hrobů. Za předpokladu, že hlavní variabilita je určena časem (chronologií), pohlavím a věkem zemřelých, jsem pro popis hrobů vybral 31 proměnných, které jsem pak vzhledem k malým četnostem některých z nich zredukoval na počet 28. Objekty deskriptivního systému se tudíž staly jednotlivé hroby, deskriptory byly většinou druhy hrobové výbavy. Jestliže hrob obsahoval např. (alespoň jednu) amforu, v příslušném poli deskriptivní matice bylo zapsáno číslo 1, pokud amforu neobsahoval, číslo 0. V takovém případě standardní program pro výpočet korelací produkuje korelační koeficient, který je shodný s tzv. Cramerovým V . K řešení bylo vybráno 9 faktorů.

K testování validity faktorů byly vybrány dvě externí evidence. Jednak to byla informace o pohlaví a věku zemřelých: pohlaví je v kultuře se šňůrovou keramikou jednoznačně určeno polohou kostry na pravém nebo levém boku; pro dožití stáří bylo použito mínění fyzických antropologů. Pro každý faktor byly vybrány hroby s faktorovým skóre přesahujícím jednotku; počet takových hrobů pro jednotlivé faktory je udán ve druhém sloupci tabulky 11. V dalších sloupcích následuje počet hrobů

mužských a počet hrobů ženských a jednoduchý statistický test χ^2 (chikvadrát) pro ověření, zda rozdíl v těchto počtech je náhodný (rozdíl je statisticky významný na hladině významnosti 5%, jestliže testovací charakteristika je větší než 3.841). Totéž pak následuje pro pohřby dospělých a pohřby dětí. Upozorňuji, že v případě faktoru 9/9 je při poměru dospělí:děti (12:6) významná převaha dětí, a to vzhledem k velmi malému počtu dětských hrobů na pohřebišti.

Z tabulky je vidět, že soubory hrobových přídavek charakterizované faktory 2/9 a 6/9 jsou statisticky významné v hrobech žen a faktory 3/9 a 8/9 v hrobech mužů. Faktor 3/9 znamená dospělé osoby a faktor 9/9 spíše děti (všimněme si, že tento faktor je častěji charakteristický pro hroby dospělých, ale vzhledem k jejich naprostému převládnutí v souboru znamená i absolutně menší četnost dětských hrobů s výbavou faktoru 9/9 významnost tohoto faktoru pro děti).

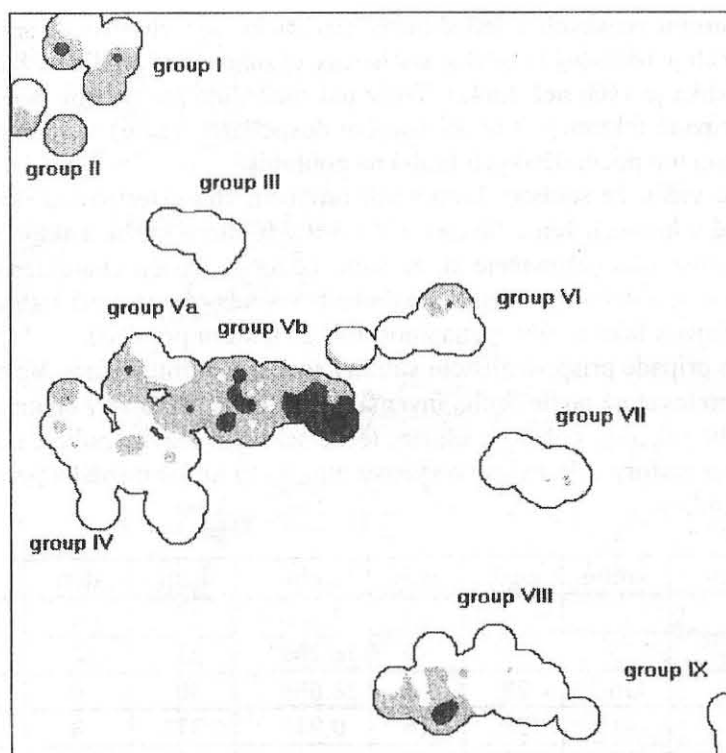
V popsaném případě přispívá zjištění validity současně k interpretaci. Velkou část faktorů bylo ovšem snadné interpretovat už podle druhů inventáře. Tak např. faktor 3/9 charakterizovaly následující deskriptory: ploché sekerky, poháry s úškem, těžké sekery, kulovité palice, sekeromlaty, džbánky, pazourkové nástroje a amfory. Jde zřejmě o výbavu mužských hrobů mladší a pozdní fáze české kultury se šňůrovou keramikou.

Faktor	hrobů	muži	ženy	χ^2	dosp.	děti	χ^2
FA 1/9	33	8	17	3.240	22	4	0.150
FA 2/9	25	1	19	<u>16.200</u>	21	1	1.354
FA 3/9	36	28	0	<u>28.000</u>	30	0	<u>4.422</u>
FA 4/9	41	22	16	0.947	31	8	2.047
FA 5/9-	35	21	12	2.455	30	2	1.244
FA 6/9	38	2	34	<u>28.444</u>	34	3	0.742
FA 7/9-	27	13	11	0.167	19	6	2.777
FA 8/9-	39	30	4	<u>19.882</u>	30	3	0.416
FA 9/9	22	5	13	3.556	12	6	<u>6.744</u>

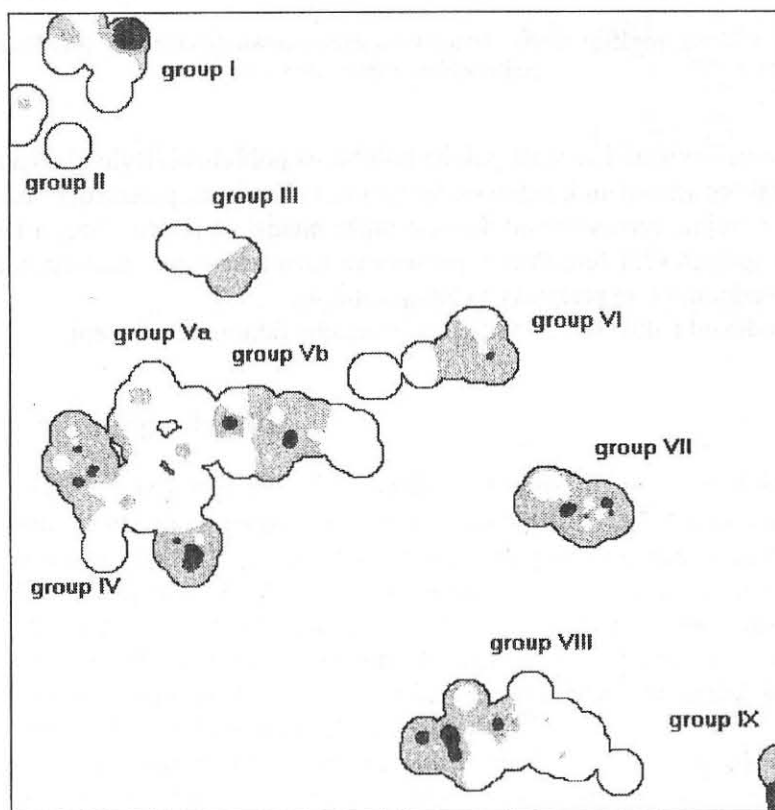
Tab. 11. Vikletice, rozdělení hrobů s faktorovými skóre charakteristickými pro jednotlivé faktory podle pohlaví a věku zemřelých.

Druhou externí evidencí se stala poloha pohřbů na pohřebišti. Bylo zjištěno, že hroby s vysokými faktorovými skóre vzhledem k jednomu faktoru se velice často prostorově shlukují. Tak např. už zmíněný faktor 3/9 zřejmě representoval dospělé muže mladší až pozdní fáze, a faktor 6/9 víceméně současné hroby dospělých (??) žen. Přitom prostorová rozložení těchto faktorů, měřeno faktorovými skóre s vysokými hodnotami, se prakticky vylučují (obr. 4).

To je další důvod k důvěře ve validitu studovaného faktorového řešení.



Obr. 4a. Vikletice, šňůrové pohřebiště, faktor 3/9. Šedě a černě plochy s předpokládaným faktorovým skóre > 0. Stupnice jako u obr. 1



Obr. 4b. šňůrové pohřebiště, faktor 6/9. Šedě a černě plochy s předpokládaným faktorovým skóre > 0. Stupnice jako u obr. 1.

14.6 Praktické provedení

Dnes existuje řada „balíků“ programů, které provádějí faktorovou analýzu. Všechny nejsou stejně vhodné, starší z nich mají komplikované vstupy dat, některé z nich nepočítají faktorová skóre, u jiných je nejasný algoritmus.

Autor tohoto textu dříve používal svůj vlastní program, napsaný v 60. letech (kdy prakticky neexistoval software) v jazyku Fortran IV. Adaptace tohoto programu pro přímý vstup z databáze, pro velké matice dat atd. by si však byla vyžádala nepřiměřeného programátorského úsilí; to vedlo k přechodu na komerční software.

Příklady v této práci jsou vypočítány pomocí softwaru SPSS firmy SPSS, Inc., který je ve svých nových verzích velmi přátelský, není však levný.

V případě SPSS je výběr mezi několika modely „faktorové analýzy“. Vstup do analýzy metodou hlavních komponent je přímo z relační databáze, např. z dBASE IV, ale jsou možné i další varianty vstupu. Podle manuálu nebo „helpu“ se předem zadají požadavky na postup výpočtu (který pak probíhá plně automaticky a prakticky okamžitě) a výstupy. Obvyklý výstup korelačních koeficientů, vlastních čísel a faktorových koeficientů je do textového souboru, odkud se může načíst do rukopisu; faktorová skóre se automaticky zaznamenávají do databáze, z níž byl vybrán deskriptivní systém (pokud je ovšem tato volba předem vybrána). Takovou databázi pak lze použít např. jako vstup do GIS.

Poděkování. Tato práce byla částečně napsána s podporou grantu 404/95/0523 Grantové agentury České republiky. Za posouzení srozumitelnosti textu děkuji D. Dreslerové, M. Kunovi, J. Macháčkovi, V. Salačovi a N. Venclové.

14.7 Literatura

K seznámení s matematickými pojmy je vhodná kterákoliv moderní učebnice lineární algebry, numerické matematiky aj.

Dále uvádím jen práce citované v textu, některé velmi základní příručky a archeologické aplikace vektorové syntézy na český materiál.

Binford, L.R. and S.R. Binford 1966: A preliminary analysis of functional variability in the Mousterian of Levallois facies, *American Anthropologist* 68, 238-295.

Buchvaldek, M. and D. Koutecký 1970: *Vikletice, ein schnurkeramisches Gräberfeld*. Praha

Harman, H.H. 1970: *Modern Factor Analysis*. Chicago and London. The University of Chicago Press.

Kuna, M. 1994: Archeologický průzkum povrchovými sběry – *Archaeological survey by surface collection* (Zprávy České archeologické společnosti, Supplément 23). Praha

Kuna, M. 1996: GIS v archeologickém výzkumu regionu: vývoj pravěké sídelní oblasti středních Čech. *Archeologické rozhledy* 48, 580-604.

Neustupný, E. 1973: Factors determining the variability of the Corded Ware Culture. In: C. Renfrew (ed.): *The explanation of culture change*. London. 725-730.

Neustupný, E. 1973: *Mathematische Untersuchungen zur böhmischen Schnurkeramik – Matematický výskum českej šnúrovej keramiky, Musaica* (Zborník filozofickej fakulty univerzity Komenského), 13-67.

Neustupný, E. 1973: *Jednoduchá metoda archeologické analýzy – A simple method of archaeological analysis*, *Památky archeologické* 64, 169-234.

Neustupný, E. 1978: *Mathematical analysis of an Aeneolithic cemetery*, *Studia praehistorica* 1-2, 238-243.

Neustupný, E. 1978: *Mathematics at Jenišův Újezd*. In: *Das keltische Gräberfeld bei Jenišův Újezd in Böhmen 2*. Teplice. 40-66.

Neustupný, E. 1979: *Vektorová syntéza sídlištní keramiky – Vector synthesis of finds from settlement sites*, *Archeologické rozhledy* 31, 55-74.

- Neustupný, E. 1981: K matematické analýze pravěkých pohřebišť – Zur mathematischen Analyse prähistorischer Gräberfelder. In: Současné úkoly československé archeologie (Valtice 1978). Praha. 190-193.
- Neustupný, E. 1986: Nástin archeologické metody – An outline of the archaeological method, Archeologické rozhledy 38, 525-549.
- Neustupný, E. 1993: Archaeological Method. Cambridge. CUP
- Neustupný, E. 1994: Role databází v archeologii – The role of databases in archaeology, Archeologické rozhledy 46, 121-128.
- Neustupný, E. 1995: Beyond GIS. In: Lock, G. and Z. Stančič (eds.), Archaeology and Geographical Information Systems a European Perspective. London. Taylor&Francis. 133-139.
- Neustupný, E. 1996: Polygons in archaeology, Památky archeologické LXXXVII, 112-136.
- Neustupný, E. and S. Vencl 1995: Formal methods at Hostim. In: S. Vencl: Hostim – Magdalenian in Bohemia, Památky archeologické – Supplementum 4, 205-224.
- Neustupný, E. a N. Venclová 1996: Využití prostoru v latěnu: region Loděnice, Archeologické rozhledy 48-4, 615-642.
- Ueberla, K. 1974: Faktorová analýza. Bratislava. Alfa.
- Venclová, N. 1994: The field survey of a prehistoric industrial region, Památky archeologické – Supplementa I, 239-247.
- Venclová, N. 1995: Specializovaná výroba: teorie a modely – Specialised production: theories and models, Archeologické rozhledy 47, 541-564.

Počítačová podpora v archeologii

Jiří Macháček (editor)

Vydala Masarykova univerzita v Brně roku 1997

1. vydání, 1997

Náklad 500 výtisků

AA – 15,88 VA – 16,21

Sazba Dan Šlosar, LVT FF

Tisk Gloria, Rosice u Brna

Pořadové číslo 2701/FF - 1/97 - 17/99

ISBN 80-210-1562-4