



Geoinformatika

V – Transformace dat

jaro 2020

Petr Kubíček

kubicek@geogr.muni.cz

**Laboratory on Geoinformatics and Cartography (LGC)
Institute of Geography
Masaryk University
Czech Republic**

Chyby v datech

- Při vkládání dat do systému **není možné zabezpečit správnost 100% zadání dat.**
- **Identifikace chyb** je velice obtížná. Obvykle se data kontrolují **vizuálně**. Dalším způsobem kontroly chyb prostorových dat je proces vytváření topologie neboli **topologické čištění dat.**
- GIS mají většinou schopnosti **procházet místa s potenciální chybou** a umožní uživateli interaktivně odstranit případné chyby.



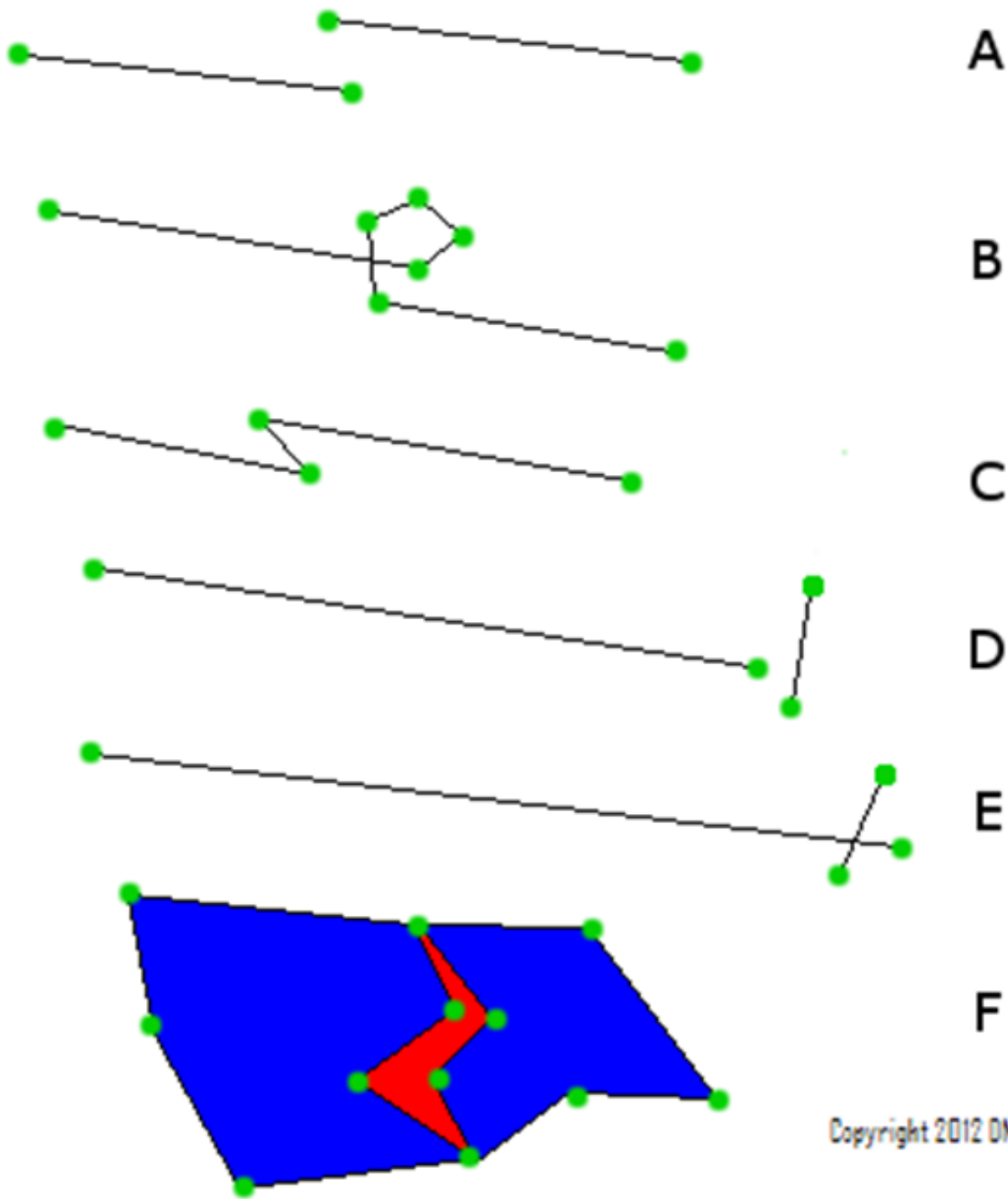
Možné chyby při zadávání

- **Nekompletnost dat** - scházejí body, linie, polygony.
- **Chybné umístění** prostorových dat - chyby vycházející ze špatné kvality vstupních dat nebo z nedostatečné přesnosti při digitalizování.
- **Zkreslení prostorových dat** - chyby z nepřesností vstupních dat (deformace podkladových dat, zkreslení již existující analogové kresby).
- **Špatná vazba** mezi prostorovými a atributovými daty.
- **Atributy jsou chybné** nebo nejsou kompletní – velice častá chyba zvláště pokud jsou atributy pořizovány z různých zdrojů v různých časech.

Chyby při vytváření topologie

- **Třísky a mezery (Sliver and gaps)** - jev nastává, když jsou dvě hranice digitalizovány z různých zdrojů, ačkoli v terénu představují jednu a tu samou. V takovém případě jsou linie představující tutéž hranici neidentické (nepřerývají se)
- **Mrtvé konce (dead ends)** - nedotahy a přetahy.
- **Duplikátní linie** (hlavně v CAD, ale i u některých GIS, které z toho vytváří regulární polygon) reprezentující stejný objekt.
- Pokud se používá pro reprezentaci polygonů metoda hranic a centroidů, tak i **přiřazení více centroidů jednomu polygonu**.

Obvyklé chyby v GIS geometrii





Topologické čištění dat

- **Jednotlivé úlohy**
 - **Eliminace duplikátních linií** (stejných i podobných).
 - **Odstraňování nedotahů** a přetahů.
 - **Nalezení průsečíků** dvou nebo více liniových prvků s následující segmentací.
 - **Odstranění mezer** (souvisí s nedotahy).
- **Topologicky čistá data** jsou taková data, nad kterými **je možné vytvořit topologii**, aniž by se jakkoli změnila jejich poloha.
- Pro tvorbu topologicky čistých dat se používají **topologické koncepty** (konektivita, definice plochy, sousednost).



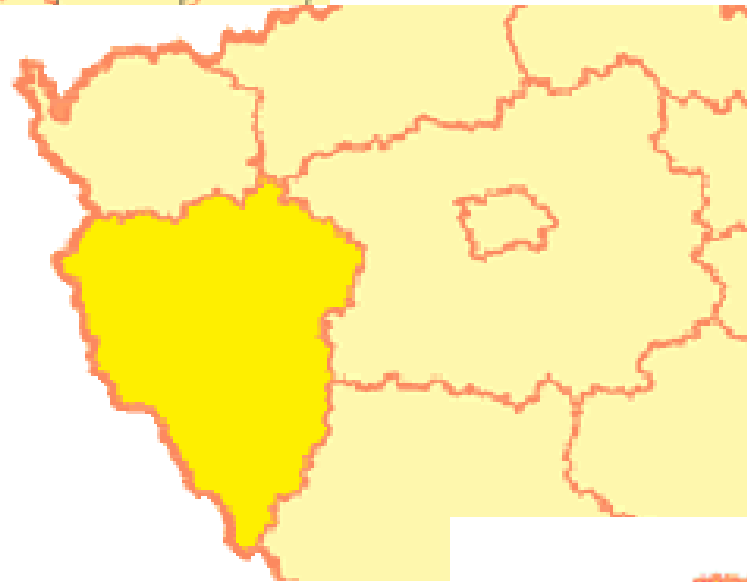
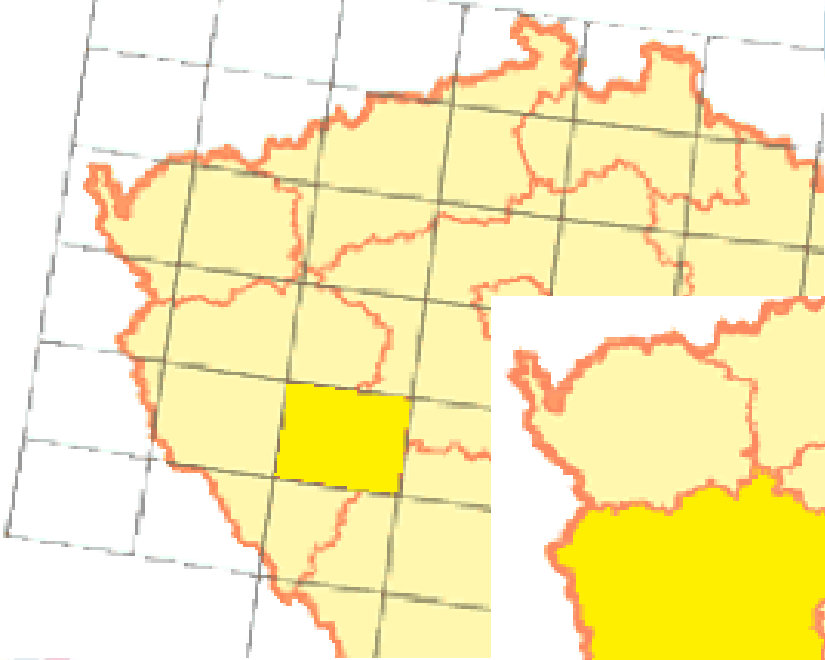
Chyby právního charakteru

- **Při pořizování dat je nutné brát v potaz i právní souvislosti problematiky, kdo má na data obchodní práva, zda je možné data využívat pro akademické, soukromé, či obchodní účely.**
- **Zdroje obvykle přesně popisují možnosti využití a omezují zejména komerční či veřejné použití dat (i jako podkladu).**
- **Ochrana dat (vodotisk, záměrné chyby).**

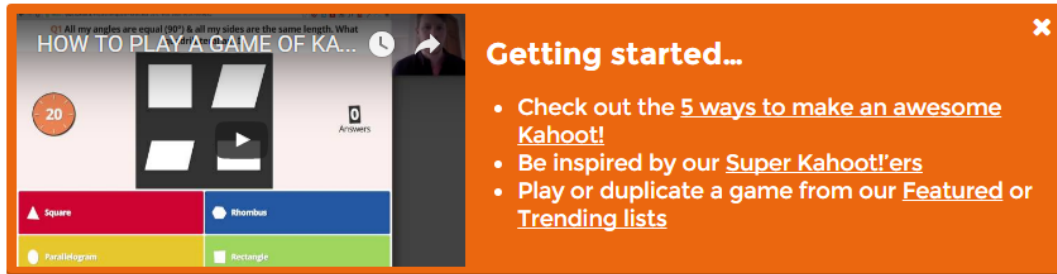


Uchovávání a zpracování dat

- **Pravidelné (např. mapové listy).**
 - Na disku je každý mapový list v jednom souboru (resp. ve více souborech se stejným jménem, lišících se pouze příponou) či adresáři.
- **Nepravidelné (mapové listy, zájmové území - katastrální území, území národního parku, okresu, kraje ...).**
 - Na disku je každé zájmové území v jednom souboru (resp. ve více souborech se stejným jménem, lišících se pouze příponou) či adresáři.
- **Bezešvé (Seamless)**
 - Celé zájmové území je uloženo v jednom souboru, adresáři či databázi).



Geoinformatika



Getting started...

- Check out the [5 ways to make an awesome Kahoot!](#)
- Be inspired by our [Super Kahoot!ers](#)
- Play or duplicate a game from our [Featured](#) or [Trending lists](#)

Create a new kahoot



Quiz

Introduce, review and reward



Jumble

Brand NEW game



Discussion

Initiate and facilitate debate



Survey

Gather opinion and insight



Komplexní GIS schéma

Transformace dat

- modelu
- polohy
- formátu

Sběr dat

- editace
- import

Uložení dat

- | | |
|-------------|------------------------------|
| Návrh | - struktura
- datové typy |
| Manipulace | - dotazování
- indexování |
| Dokumentace | - metadata |

Analýza dat

- průzkum
- modelování

Prezentace dat

Vizuální

- kartografická
 - statické mapy
 - dynamické mapy
 - uživatelské rozhraní
- nekartografická
 - grafická
 - textová

Nevizuální

- export
- řídicí povely

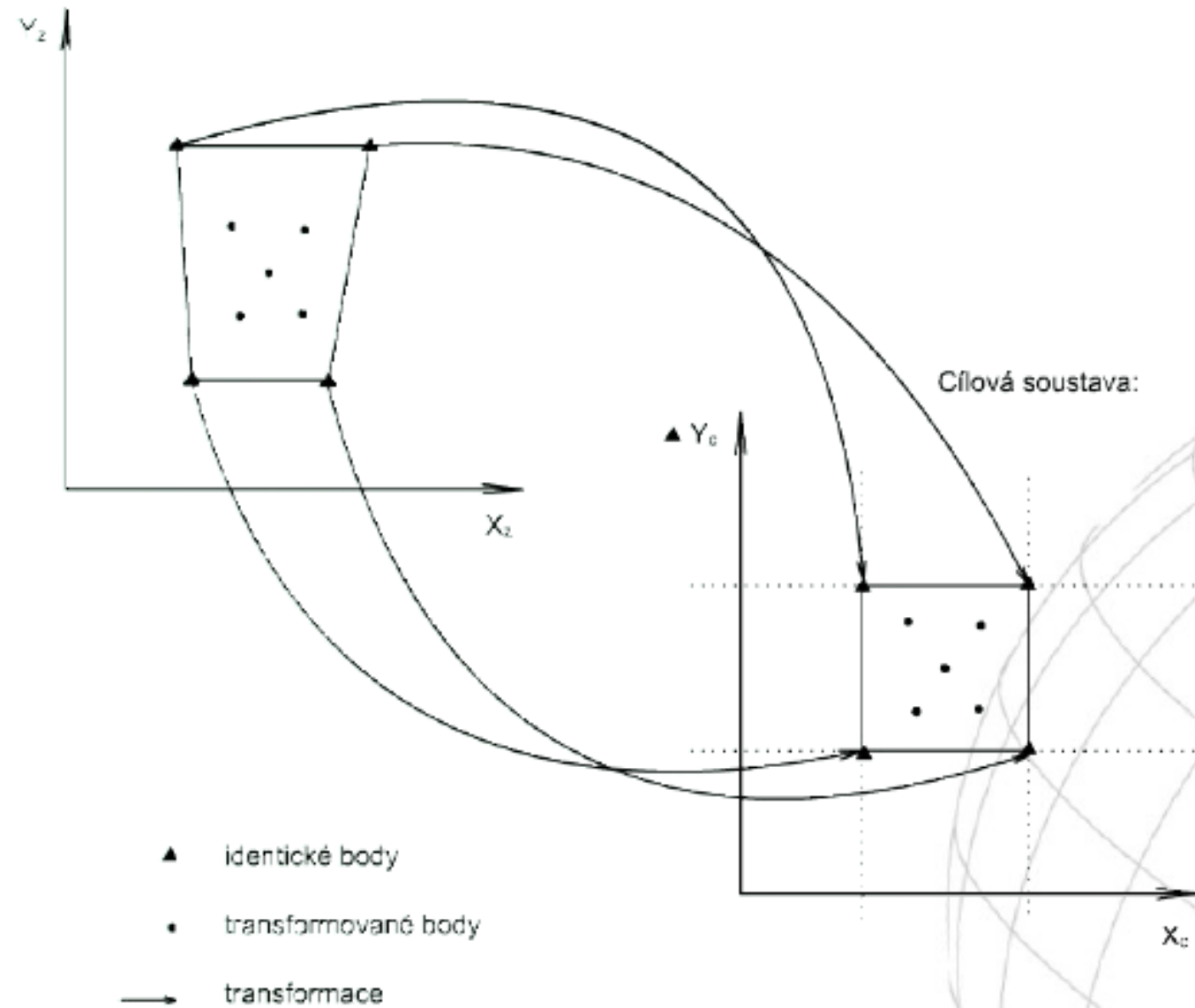


Transformace

- **Polohová – geometrická transformace**
 - Lineární
 - Afinní
 - Projektivní
- **Datového modelu**
 - RAVE – rastr to vector
 - VERA – vector to rastr
- **Formátu**

Geometrické transformace

Zdrojová soustava:

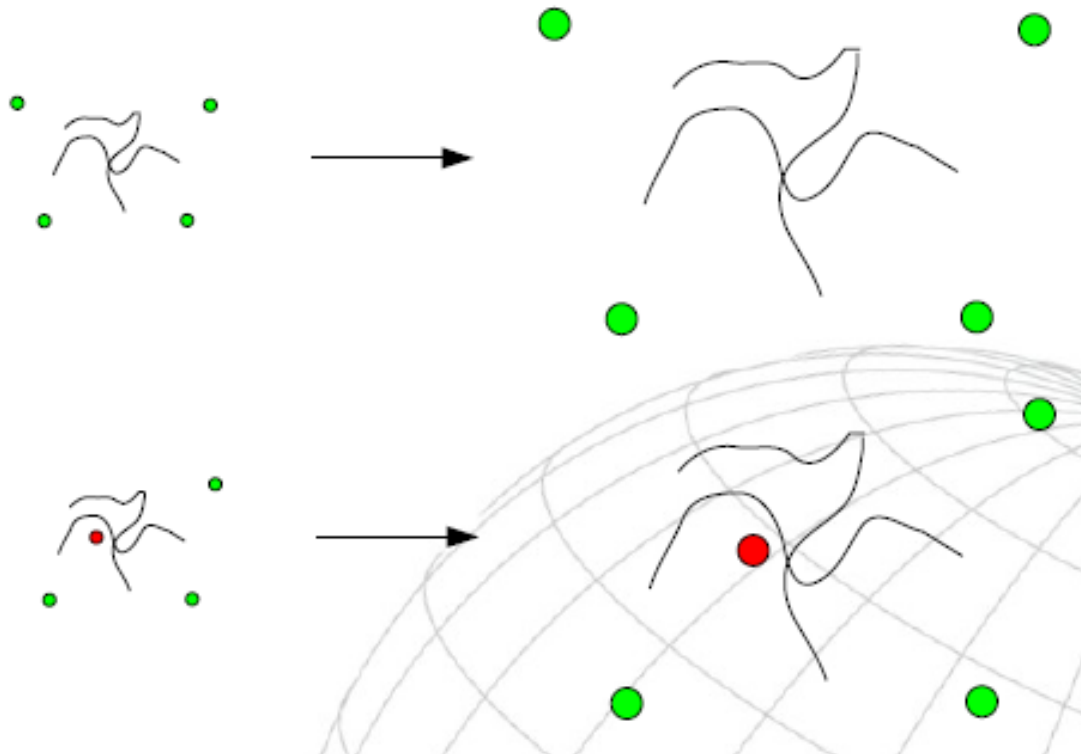


Transformace mezi rovinnými pravoúhlými souřadnicemi jsou založeny na poznání přesné polohy vybraných identických bodů.

Geometrické transformace - volba identických bodů

- U výběru dvojic identických bodů je také vhodné mít na paměti, že je nutné je vybírat co nejbližže okrajům transformovaného území, aby nebyly způsobeny nežádoucí deformace na okrajích.

» dobře





Geometrické transformace - identické body a transformační koeficienty

- **Transformační koeficienty** jsou hodnoty, vypočtené z dvojic identických bodů, kterými se vyjadřuje přechod od zdrojové souřadnicové soustavy do cílové.
- U transformace se ale obvykle **používá více identických bodů, než je nutné** pro výpočet transformačních koeficientů.
- Hodnoty transformačních koeficientů se pak vypočtou **metodou nejmenších čtverců**, kde se minimalizuje suma rozdílů v poloze mezi souřadnicemi transformovaných bodů (více – Matematická kartografie).
- **Transformace** je například **posun a změna měřítka**.



Geometrické transformace – typy transformací

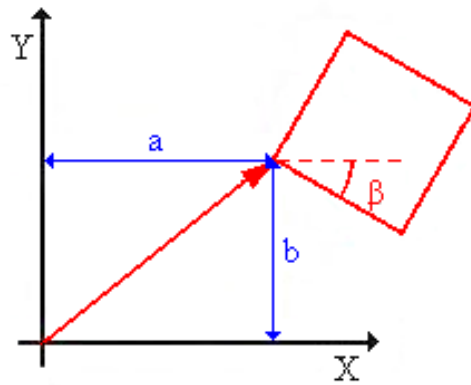
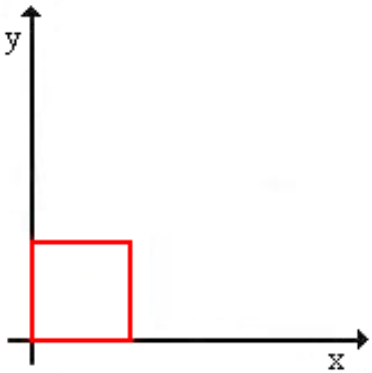
Transformace souřadnicového systému mezi rovinnými pravoúhlými souřadnicemi:

- Lineární konformní transformace (LKT)
- Afinní transformace (polynomická prvního řádu a polynomické transformace vyšších řádů)
- Projektivní transformace



Lineární konformní transformace

- $X(X,Y)$ - nové souřadnice
- $x(x,y)$ - staré souřadnice
- B - úhel otočení
- m - změna měřítka
- $p(a,b)$ - posun
- **Transformační koeficienty** (m, B, a, b) lze vypočítat již ze dvou dvojic identických bodů $(X1,Y1), (X2,Y2)$ a původní $(x1,y1), (x2, y2)$.



Zápis rovnicí

$$X = m \cdot \cos(B) \cdot x - m \cdot \sin(B) \cdot y + a$$

$$Y = m \cdot \sin(B) \cdot x + m \cdot \cos(B) \cdot y + b$$

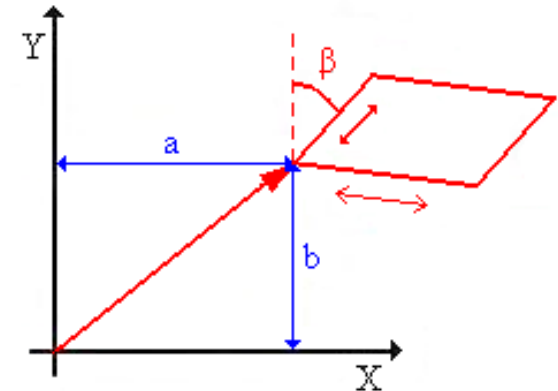
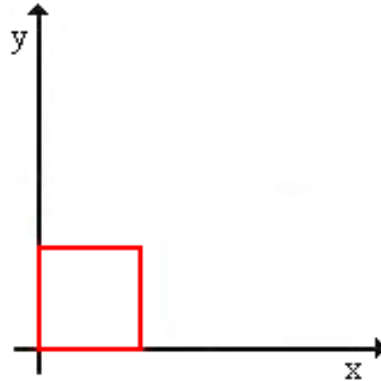
Helmertova transformace – speciální případ LKT; $m = 1$

Lineární konformní transformace

- Posun
- Rotace
- Uniformní změna měřítka (v obou osách stejná)
- **Zachovává tvar objektu! (konformní)**
- Je potřeba dvou dvojic identických bodů

Afinní transformace

- **Jednotlivé souřadnice nejsou na sobě závislé – změna měřítka v různých směrech.**
- **$X(X, Y)$** - nové souřadnice
- **$x(x, y)$** - staré souřadnice
- **A** - regulární matice
- $p(c, f)$ - posun
- Transformační koeficienty (a, b, c, d, e, f) lze spočítat ze tří dvojic identických bodů.



Zápis rovnicí

$$X = a \cdot x + b \cdot y + c$$

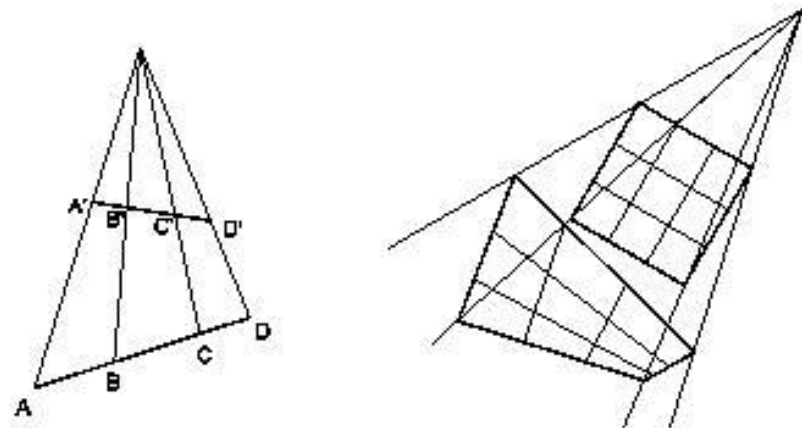
$$Y = d \cdot x + e \cdot y + f$$

Afinní transformace

- Posun
- Rotace
- Neuniformní změna měřítka (v každé ose jinak – zkosení)
- „Z obdélníka kosodélník“
- Je potřeba **tří dvojic identických bodů**

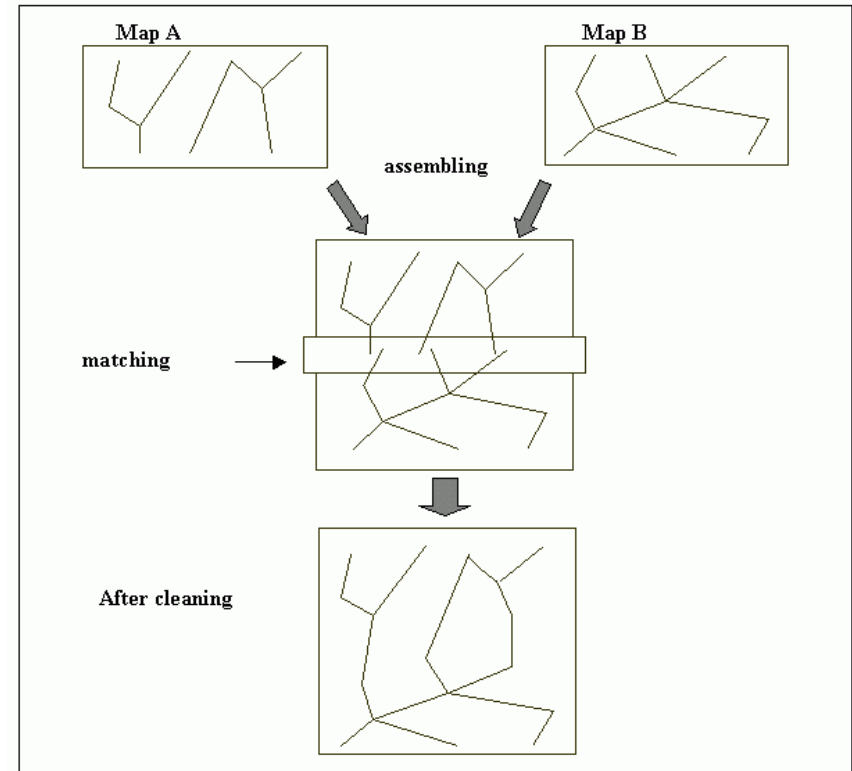
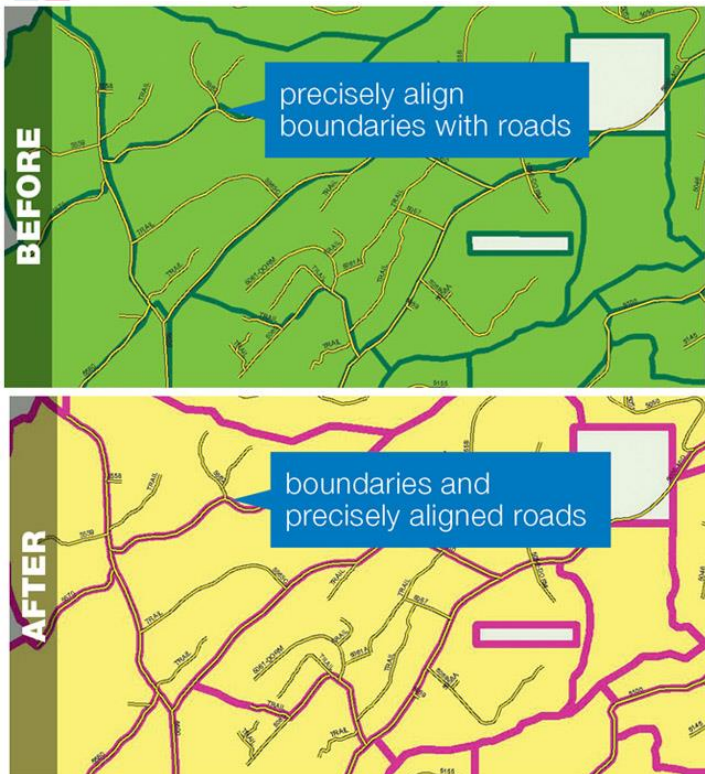
Projektivní transformace

- Transformace jednoho rovinného prostoru do druhého – vhodné pro data s menšími deformacemi.
- Posun
- Rotace
- „Z obdélníka lichoběžník“
- Je potřeba **čtyř dvojic identických bodů**

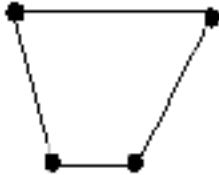
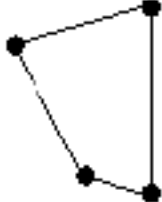





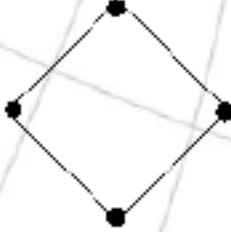


Dalsi typy geometrických transformací - Rubber sheeting, edge matching

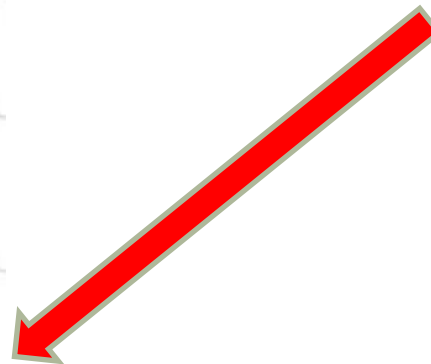
- **Rubber sheeting** - pro zdeformované mapy - lineární transformace po částech.
- **Edge Matching** - sjednocení okrajů mapy. V důsledku dělení plochy na mapové listy, odpovídá rubber sheetingu, ale platí pouze pro okraje mapových listů.



Projevy geometrických transformací

Transformace	Před	Po
Projekce		
Afinita		
Změna měřítka		
Rotace		

- **Helmertova?**
- **LKT, $m=1$**





Transformace datového modelu

- Jelikož pro některé analýzy jsou vhodnější vektorové reprezentace dat a pro jiné zase rastrové, GIS systémy pracující s oběma typy nabízejí nejrůznější nástroje umožňující a usnadňující převod mezi oběma reprezentacemi.
- Převod z rastrové do vektorové podoby se nazývá **vektORIZACE (RAVE)**, opačný proces z vektorové do rastrové podoby je **rasterizace (VERA)**.

Vektorizace

Ruční

- Vše dělá operátor (případně za asistence počítače při přichytávání vektorových prvků na existující rastrovou kresbu - tzv. „čtvrtaautomatická“).

Poloautomatická

- Operátor zvolí počátek rastrové linie, systém se pokusí identifikovat rastrový objekt, ukáže operátorovi směr, kterým se vektorizace bude ubírat, a při potvrzení ze strany operátora, se vydá vektorizovat, dokud nenarazí na nějakou překážku (mezera, křižovatka) či sporný bod, kde se zastaví a čeká na operátorovu odezvu (jestli má pokračovat, v jakém směru má pokračovat, ...).
- Existují dva módy poloautomatické vektorizace, podle způsobu přichytávání:
 - na **střed rastru** (používaný pro vektorizaci linií),
 - na **okraj rastru** (používaný pro vektorizaci polygonů).

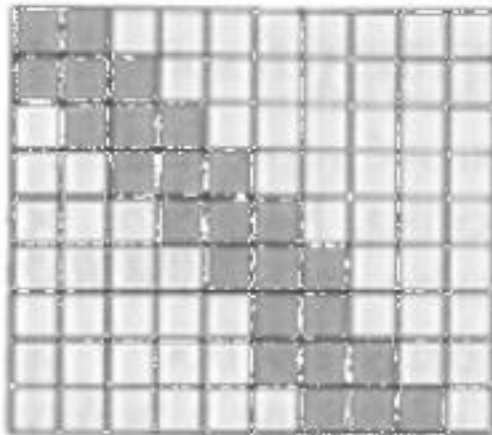
Poloautomatická

- **Přichytávání na okraj** je pro počítač výrazně jednodušší, jelikož vektorizační software pouze hledá hranu v rastrovém obrazu, které se drží.
- **Přichytávání na střed** je složitější a pro identifikaci středu vektorizovaného objektu se využívá principu nazývaného „**skeletizace**“, který vychází z principů používaných v automatické vektorizaci.

1)



2)



3)



Poloautomatická

- Pro poloautomatickou vektorizaci lze obvykle **nastavit několik důležitých parametrů** pro zautomatizování činnosti.
- Mezi tyto parametry patří:
- kvalita rastrového podkladu (jestli jsou objekty homogenní oblasti či ne),
- maximální přípustná šířka linie,
- akceptovatelná mezera v rastrové linii (při digitalizaci čerchovaných a jiných čar),
- akceptovatelný úhel mezi částmi linie a variabilita (jak reaguje systém na změny šířky pouze v jednom směru).

Automatická vektorizace

- Při automatické vektorizaci probíhá převod rastr->vektor automatizovaně, **bez aktivní účasti operátora**.
- Algoritmy automatické vektorizace vycházejí z algoritmů zpracování digitalizovaného obrazu a umělé inteligence.
- Tuto metodu však většinou **nelze použít pro převod běžných analogových podkladů**, ale pouze pro již tištěné mapy z digitálních podkladů (podobně jako OCR).



Automatická vektorizace

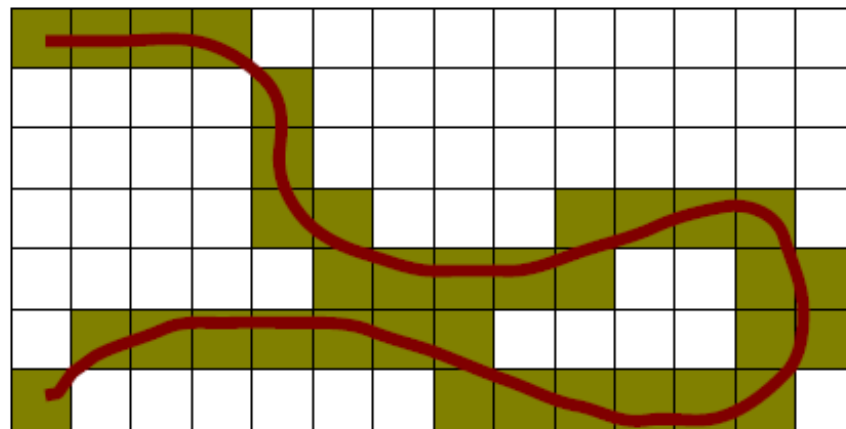
Princip automatické vektorizace pro jednotlivé typy základních objektů:

- **Body** - zpracovávající program vyhledá střed buňky reprezentující bod a zjistí jeho souřadnice a zaznamená je spolu s identifikátorem bodu v rastru (obvykle barva, či nějaká skalární hodnota).
- **Linie** - automatická vektorizace linií funguje na principu hledání kostry (skeletu, odtud skeletizace) objektů, což je metoda velice často používaná pro ztenčování objektů. Po nalezení skeletu jsou pak pouze vyhledány na sebe napojené pixely (v rámci 4 nebo 8 okolí) a ty tvoří požadovanou linii.
- **Polygony** - podobně jako u poloautomatické vektorizace jsou hledány hrany objektů a ty pak převáděny do linií. Poté se ze všech uzavřených liniiových objektů vytvoří polygony.

Rasterizace

Princip

- Provádí se jako překryt vektorové vrstvy na rastrovou mřížku (o určené velikosti buňky) a přiřazení hodnoty této buňky z vybraného atributu.
- Při rasterizaci je nejdůležitější určit správnou velikost buňky výsledného rastru (která bude dostatečně velká pro požadované účely, ale přitom nebude příliš velká pro možnosti hardware, které zpracovává rastr).





Generalizace

Proč vůbec je generalizace v GIS potřebná:

- **Ekonomické požadavky** - svět nelze nikdy modelovat úplně přesně, vždy je to kompromis přesnost/cena.
- **Požadavky redukce objemu dat**
 - čím více je dat, tím je větší možnost udělat chybu a čím je přesnější (intenzivnější) měření, tím je větší šance ovlivnění dílčích měření individuální chybou.
 - generalizace slouží k odfiltrování těchto chyb a konsolidaci.
- **Víceúčelovost požadavků pro údaje** - z jedné digitální reprezentace dat je nutné vytvářet mapy s různými informacemi i v různých měřítkách, často velice rozdílných.
- **Požadavky zobrazování a komunikace (percepce-vnímání) dat**
 - vychází z kartografických doporučení některých limitů, při jejichž překročení se mapy stávají nečitelnými (př. Max 10 gr. znaků na cm²).

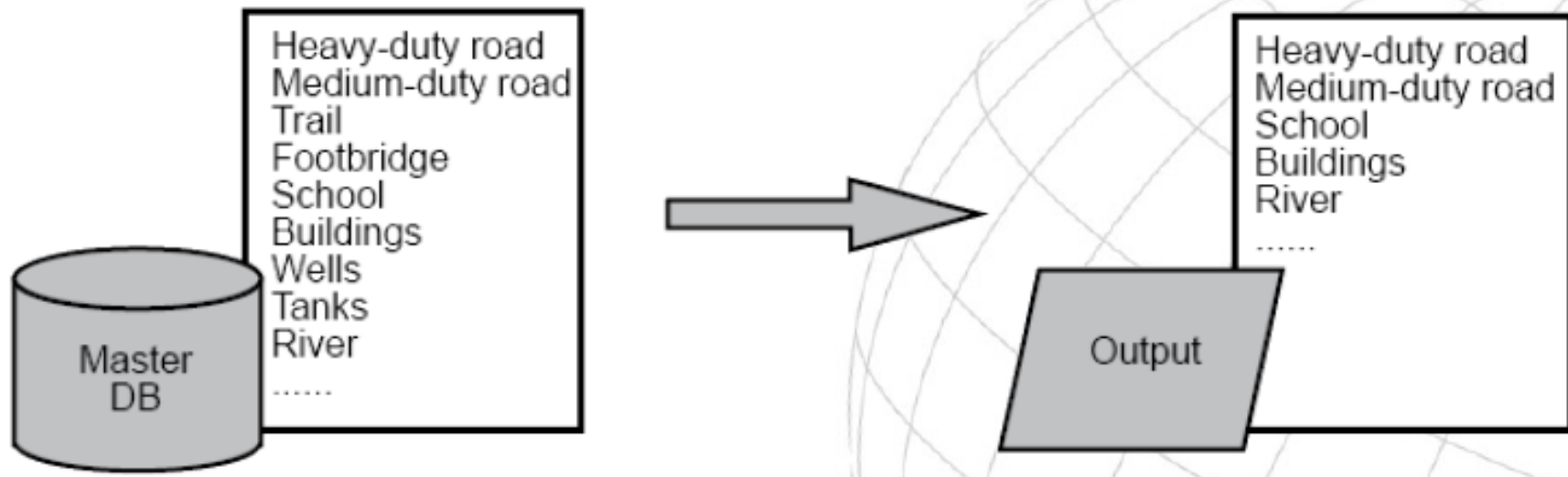
Přehled metod

Vybrané generalizační metody užívané v kartografii a GIS

- **Selekce** (výběr prvků)
- **Eliminace** (eliminace prvků)
- **Zjednodušení** (zjednodušení prvků)
- **Agregace** (kombinování malých prvků do větších)
- **Prostorová redukce** (collapse)
- **Typifikace** (redukce hustoty prvků)
- **Zvýraznění** (přehnání, exaggeration)
- **Reklasifikace** a spojení (spojení prvků se stejnými vlastnostmi)
- **Řešení konfliktů** (posunutí méně důležitých prvků)
- **Redukce vrcholů** (Coordinate Thinning)

- Selekce – výběr tématických vrstev

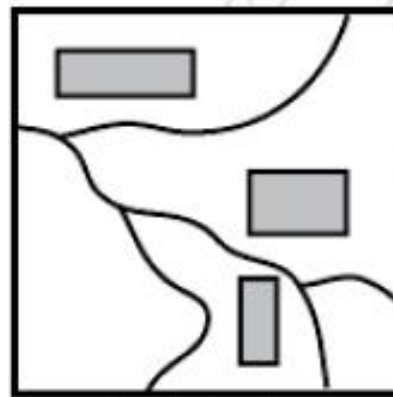
Preselection—Selecting certain feature classes from a master database for the inclusion in the final map. What to be selected depends on the target map scale and purpose. The preselected features will participate further generalization operations.



Eliminace

- Eliminace – odstranění prostorově nedůležitých prvků

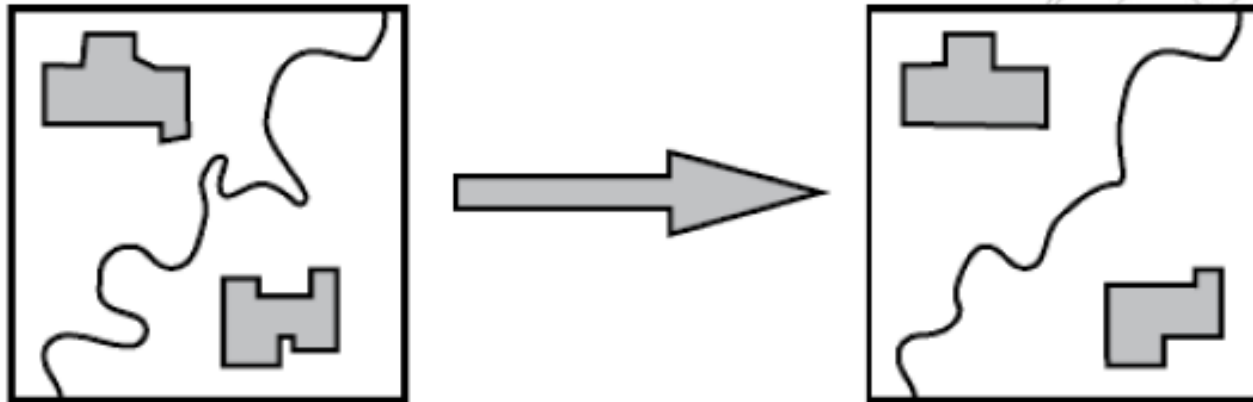
Elimination—Selectively eliminating features that are too small, too short, and too insignificant to be presented in the final map; for example, small islands, short roads, little villages, and so on.



Zjednodušení

- Zjednodušení **tvaru** prvků, např. redukce počtu vrcholů

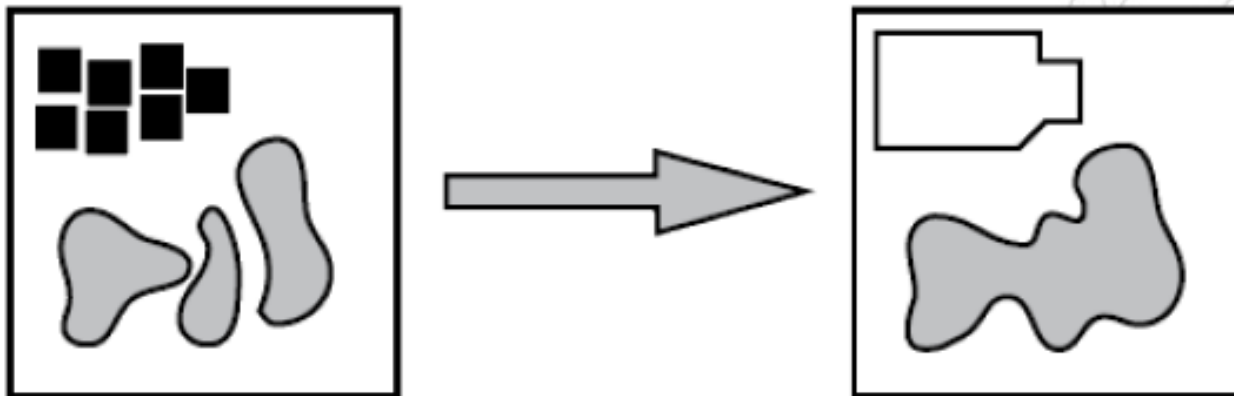
Simplification—Removing unnecessary detail, such as extraneous bends and fluctuations, from a line or an area boundary without destroying its essential shape.



Agregace

- Agregace – kombinování malých prvků do větších

Aggregation—Combining features in close proximity or adjacent features into a new area feature; for example, forming a built-up area from a cluster of buildings or joining patches of crop fields into a large agricultural area.

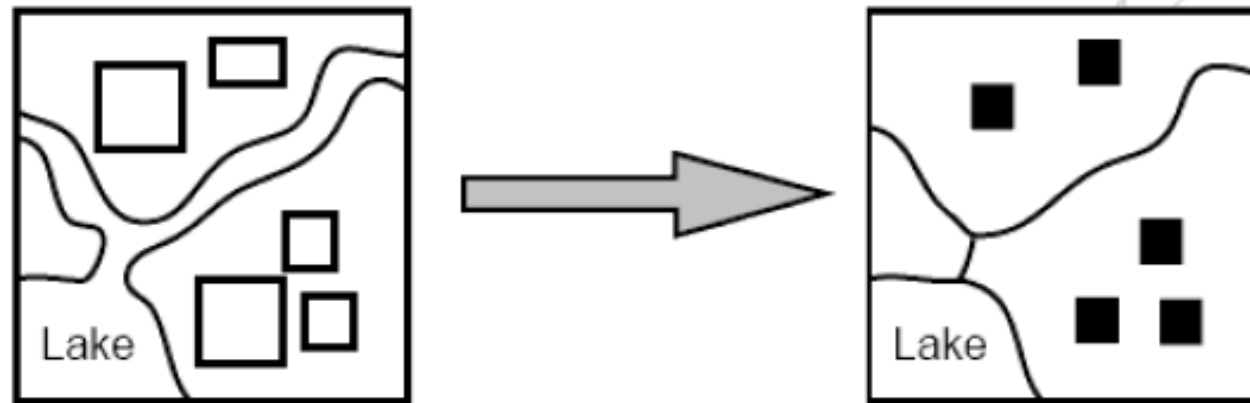




Prostorová redukce (collapse)

- Prostorová redukce – redukce dimenze prvku nebo jeho prostorového rozměru (například polygon na linii)

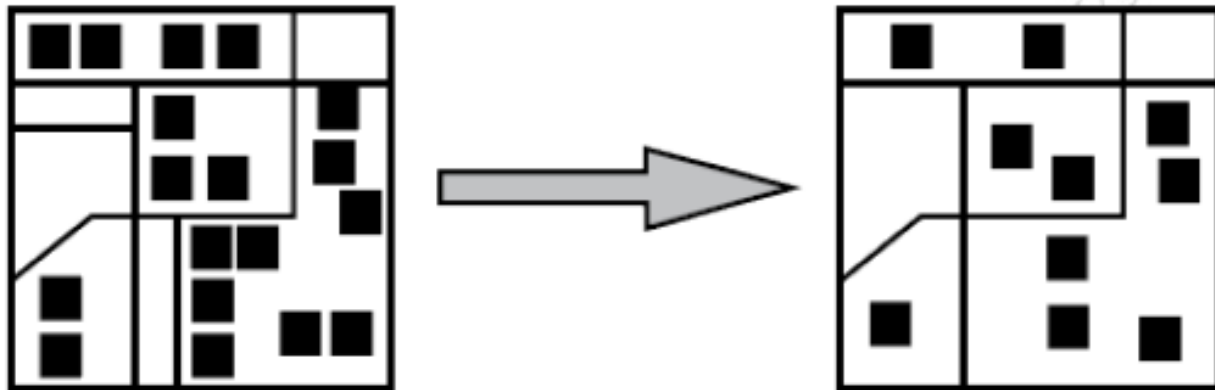
Collapse—Reducing a feature dimension or the representation of its spatial extent; for instance, changing an area feature to a linear or point feature, changing a multiple-line feature to a single-line feature, and so on.



Typifikace

- Typifikace – redukce hustoty prvků a LoD, zachování prostorového vzoru.

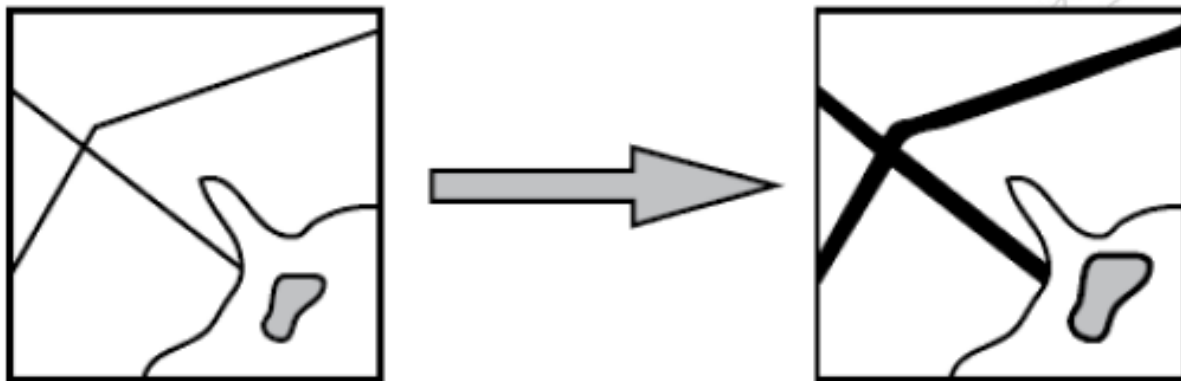
Typification—Reducing feature density and the level of detail while maintaining the representative distribution pattern and visual impression of the original feature group; for example, reducing the amount of detail in a drainage network without losing the impression of its structure.



Zvýraznění

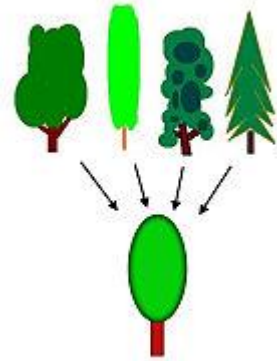
- Zvýraznění – opak prostorové redukce, prostorové zvýraznění (zvětšení) prvku.

Exaggeration—Increasing the spatial extent of a feature representation for the purpose of emphasis and legibility; for example, enlarging the size of an island, which is otherwise small enough to be removed, to include it for its significance as a navigational point of reference.

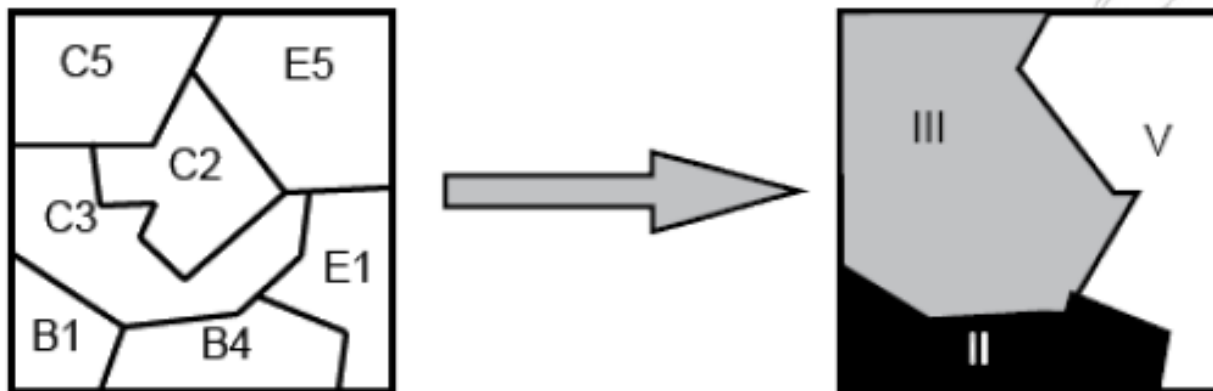


Reklasifikace a spojení

- Reklasifikace a spojení – spojení prvků se stejnými vlastnostmi do jednoho, například vrstvy listnatých a jehličnatých lesů spojit do vrstvy lesů – při změně měřítka.



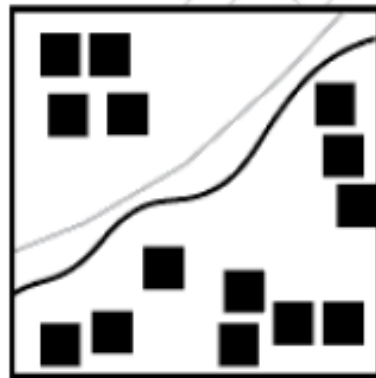
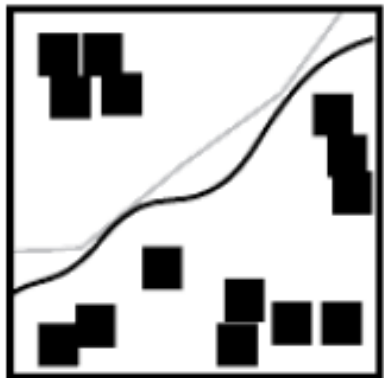
Classification and Symbolization—Grouping features sharing similar geographic attributes into a new, higher-level feature class and representing it with a new symbol.



Řešení konfliktů

- Řešení konfliktů – posunutí některých prvků nacházejících se na jednom místě, **přehlednost mapy je zde kladena nad její absolutní prostorovou správnost.**
- Na původním místě zůstává nejdůležitější prvek (např. silnice na mapě silnic) a posouvají se ostatní (v našem příkladě to může být např. železnice, vodstvo, elektrické vedení,...)

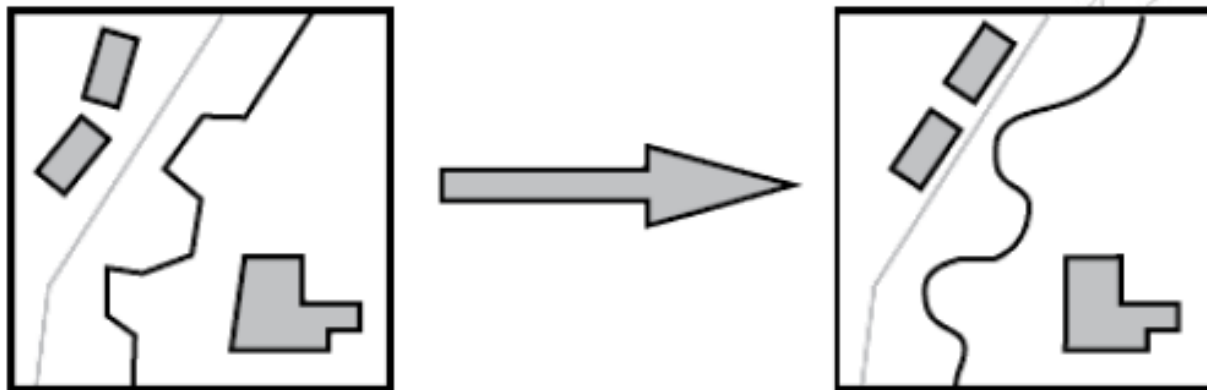
Conflict Resolution (Displacement)—Detecting feature conflicts and then repositioning the less important conflicting features or adjusting feature extents to satisfy the threshold of separation and other cartographic specifications.



Zjemnění

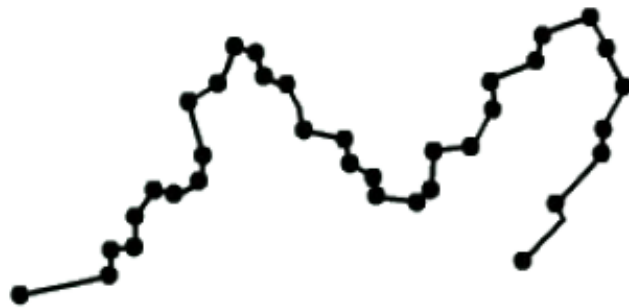
- Zjemnění – úprava vzhledu objektu ke zvýšení **estetičnosti**, například vyhlazení linie řeky.

Refinement—Altering and adjusting a feature's geometry or appearance to improve its aesthetic (visual) impression and to ensure its agreement with reality. Some examples are smoothing a line, squaring a corner, changing the orientation and alignment of a point symbol, correcting the intersecting angles of a contour and a river, and so on.



Redukce vrcholů

- Redukce vrcholů – coordinate thinnig



redukce počtu vrcholů v liniovém prvku - obvykle pomocí definování nejkratší vzdálenosti mezi body (funkce by měla vždy vybrat ty nejvýznamnější) - zjednodušování



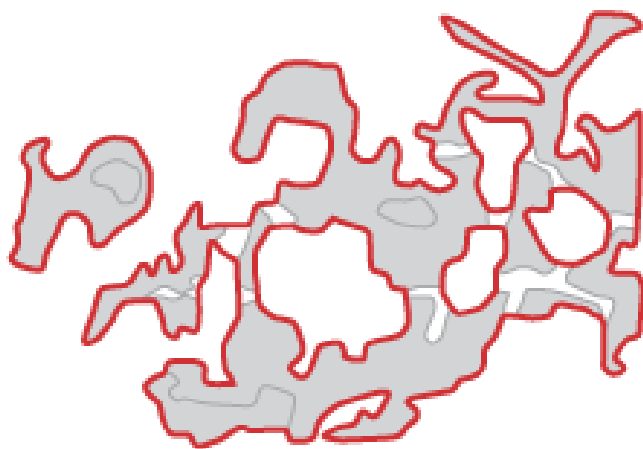


Nástroje generalizace v ArcGIS – příklady a užití

- **Aggregate** Points, Polygons
- **Collapse** Dual Lines To Centerline
- **Merge** Divided Roads
- **Simplify** Building, Line, Polygon
- **Smooth** Line, Polygon

Aggregate Points, Polygons

- **Kombinování menších prvků do větších – nahrazení shluku bodů či objektů (polygonů) jedním velkým objektem.**



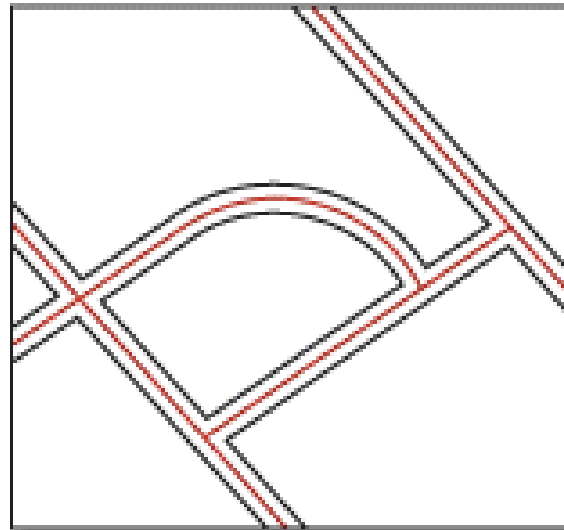
■ Input Feature
■ Aggregated Feature

A) Nonorthogonal features

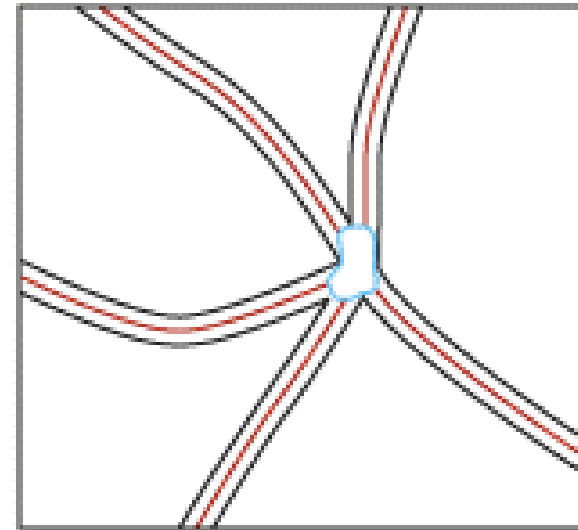
B) Orthogonal features

Collapse Dual Lines To Centerline

- **Prostorová redukce – obrysové linie nahrazeny centrální linií.**



Simple Case



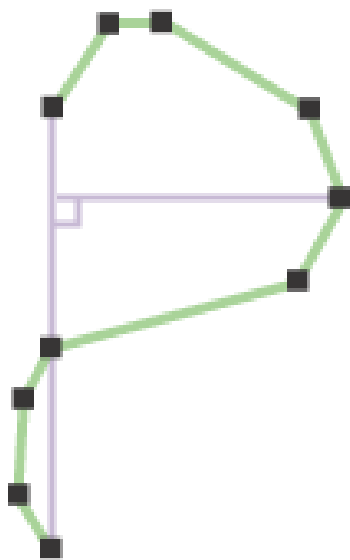
Complicated Intersection

- Casings
- Centerlines (LTYPE 1)
- Unresolved (LTYPE 2)

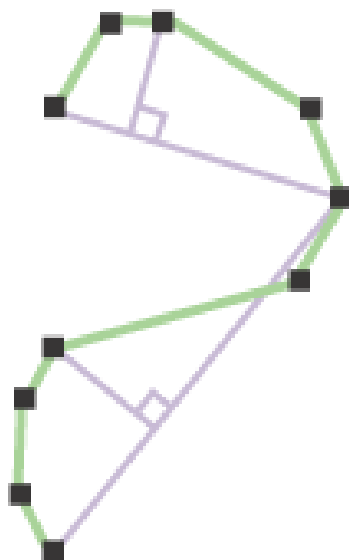


Simplify Building, Line, Polygon

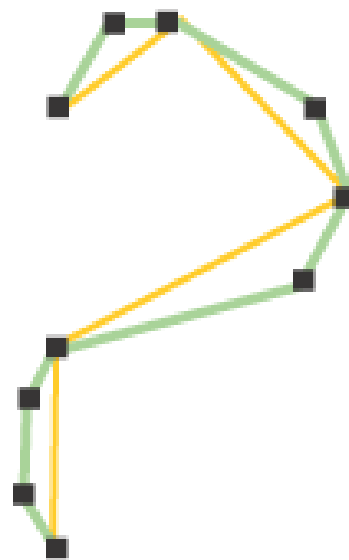
- Douglas –Peucker algoritmus



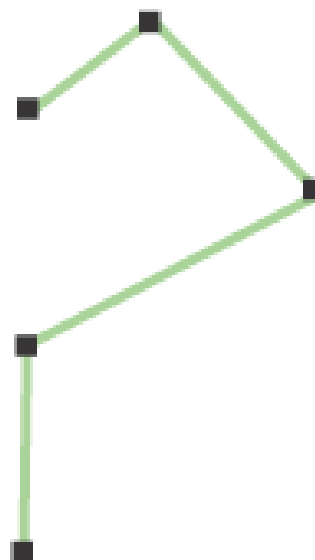
1ST TREND LINE



2ND TREND LINE

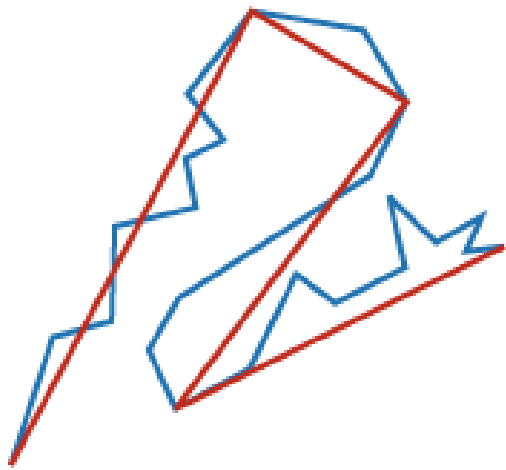


3RD TREND LINE

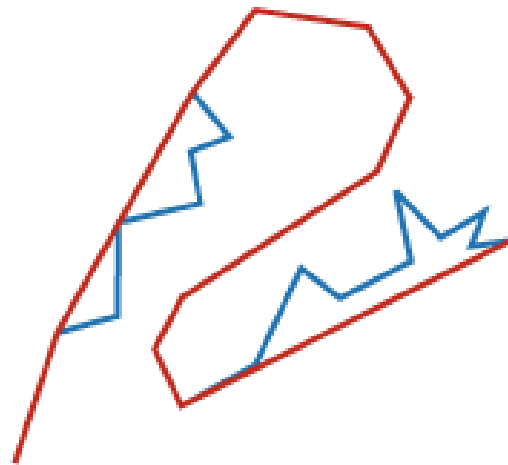


RESULTING ARC

 SIMPLIFICATION TOLERANCE



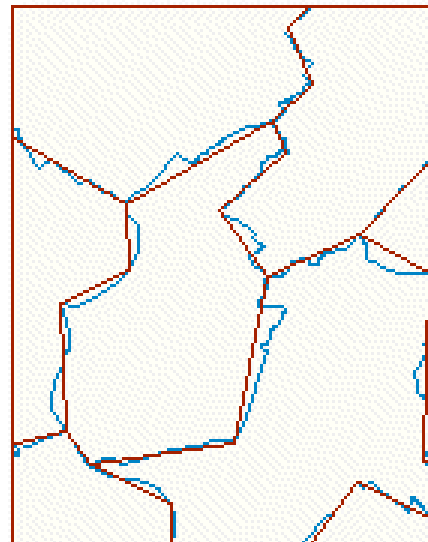
POINT REMOVE



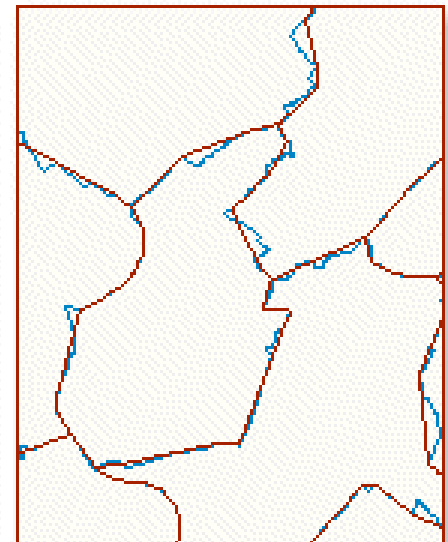
BEND SIMPLIFY

— ORIGINAL
— SIMPLIFIED

Point x
pásmo
(tvar!)



Point
Remove



Bend
Simplify

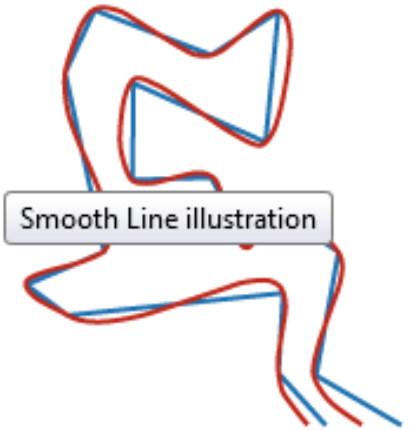


Smooth Line, Polygon

- **Shlazení (polynomální aproximace)**
- **Bézierovy křivky**

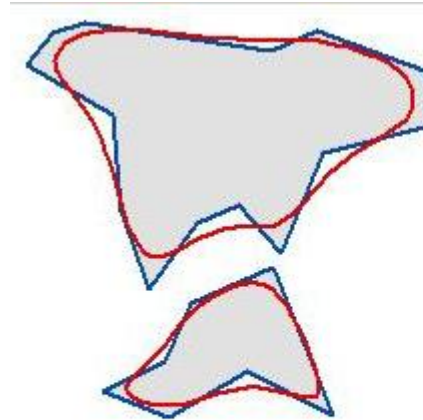
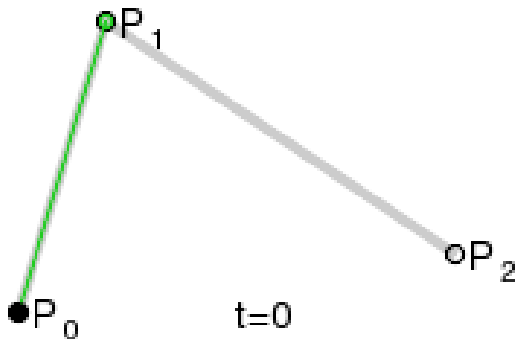


PAEK

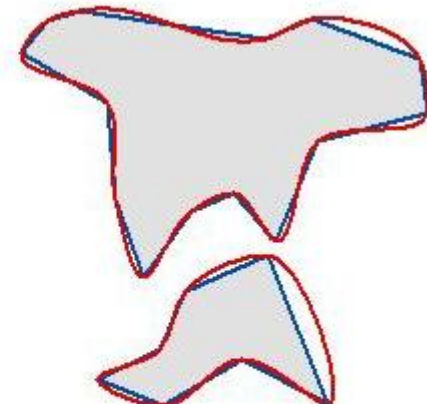


Bezier Interpolation

— ORIGINAL
— SMOOTHED



PAEK



Bezier Interpolation

— ORIGINAL
— SMOOTHED



Automatizace generalizace

- GIS obsahují jen omezené nástroje pro automatizovanou generalizaci, jelikož se jedná o poměrně složitou problematiku, než aby mohla být plně automatizována.
 - Automatizovaně lze provádět pouze **dílčí, specializované kroky** z celého procesu (viz některé výše),
 - Celý proces které musí s ohledem na aplikaci řídit uživatel – **kartograf!**
 - Měřítkové řady a přechody.

Vliv generalizace na kvalitu údajů

- Snižuje se **polohová** (prostorová) přesnost.
- Při snížení polohové přesnosti se může snížit i **atributová** přesnost!
- např. reklasifikace a spojení.

