

# INSTITUT TEORETICKÉ INFORMATIKY NA MU

## INSTITUTE FOR THEORETICAL COMPUTER SCIENCE AT MU

Jan Obdržálek, Vojtěch Řehák

V tomto příspěvku prezentujeme zkušenosti z probíhající spolupráce mezi Fakultou informatiky Masarykovy university a průmyslovými partnery (zejména společností ANF DATA spol. s r.o.) v rámci výzkumného centra Institut teoretické informatiky. Kromě obecného popisu možností a forem spolupráce se detailně věnujeme třem konkrétním projektům, na jejichž příkladě demonstrujeme funkčnost a realizovatelnost popsaných principů.

Institut teoretické informatiky, ANF DATA, informatika, formální verifikace

We present our experience of the on-going collaboration between the Faculty of Informatics at Masaryk University and its industrial partners (especially ANF DATA spol. s r.o.) within the framework of the Institute for Theoretical Computer Science research centre. After laying out the general framework and forms of collaboration we describe in detail three different projects, on which we demonstrate the real-life functioning and realizability of the described cooperation framework.

Institute for Theoretical Computer Science, ANF DATA, computer science, formal verification

### O INSTITUTU TEORETICKÉ INFORMATIKY

Institut teoretické informatiky (ITI) byl založen v roce 2005 jako společné výzkumné centrum následujících institucí:

- Matematicko-fyzikální fakulta Univerzity Karlovy v Praze
- Fakulta informatiky Masarykovy univerzity v Brně
- Fakulta aplikovaných věd Západočeské univerzity v Plzni
- Matematický ústav Akademie věd České republiky
- Ústav informatiky Akademie věd České republiky

Předmětem výzkumných aktivity institutu jsou metody, algoritmy a struktury teoretické informatiky a jejich aplikace v oboru informačních technologií (IT). Hlavní cíle mohou být shrnuty následovně:

*Provádět výzkum nejvyšší kvality.* Mezi pracovníky institutu patří mj. některé klíčové postavy české matematiky a informatiky. ITI integruje několik zavedených výzkumných skupin do jednoho celku. Institut se účastní organizace vědeckých seminářů, konferencí a workshopů. Velký důraz je kladen na zavedení a udržení spolupráce s významnými výzkumníky z celého světa.

*Vzdělávat novou generaci výzkumníků.* Institut aktivně vyhledává nové mladé talenty. Již během svých magisterských studií jsou studenti zapojováni do výzkumných aktivit centra. Vel-

ký důraz je kladen na udržení kontaktu s doktorskými studenty a začínajícími vědci, kteří jsou momentálně v zahraničí.

*Podporovat spolupráci s průmyslem.* Institut spolupracuje s významnými průmyslovými partnery, kteří mohou ovlivnit stávající výzkumné priority. Tito partneři také pomáhají vyhodnotit praktický dopad nových výsledků.

V tomto příspěvku se zaměříme na poslední bod, tedy spolupráci s průmyslovými partnery. V současnosti ITI na Fakultě informatiky spolupracuje se dvěma firmami, Aponia Software s.r.o. a ANF DATA s.r.o. (součást koncernu Siemens). Firma Aponia Software je výrobcem navigačního softwaru a integrovaných navigačních řešení – především GPS navigace. Jejich nejznámějším produktem je software Be-on-road. Tématem spolupráce je především využití nejnovějších matematických poznatků v teorii grafů při plánování trasy jízdy. Mezi řešené problémy patří plánování trasy na základě velkého množství dat (např. automapa Evropy), které však musí být realizováno v zařízeních s velmi omezeným výpočetním potenciálem jako jsou PDA. Dále se v rámci spolupráce řeší dynamická změna trasy například v důsledku dopravní uzávěry nebo minutí odbočky, a nejsou opomíjeny ani aspekty týkající se vhodné prezentace výsledné trasy řidiči. Protože spolupráce s Aponia Software začala teprve nedávno, nezmiňujeme zde žádné výsledky. Na druhou stranu spolupráce s ANF DATA již běží delší dobu, a proto se budeme věnovat především jí. Předtím však zmíníme, jak obecně vypadá spolupráce mezi ITI a partnerskou firmou.

## FORMY SPOLUPRÁCE

Spolupráce se uskutečňuje formou projektů. Každý projekt má nejméně dva koordinátory – jednoho z univerzity a jednoho z partnerské firmy. Při výběru tématu projektu se vychází ze dvou základních požadavků: 1) téma by mělo odrážet potřeby firmy v horizontu několika příštích let a 2) téma by mělo být natolik zajímavé, že by v jeho rámci mohla být formulována i doktorská disertace. V rámci každého projektu jsou v první fázi zformulována zadání pro bakalářské a diplomové práce. Tím dojde k zapojení studentů do práce na tématu. Samozřejmě práce na tématu není podmíněna zapsanou bakalářskou či diplomovou prací.

Někteří ze studentů bakalářského studia pak mohou v práci na projektu pokračovat i v navazujícím magisterském studiu, a podobně úspěšní magisterští studenti mohou pokračovat v rámci postgraduálního studia. Postupně se tak vytváří studentské týmy pracující na daném problému (cílem může být například vývoj softwarových nástrojů). Absolventi pak mohou nastoupit do firmy jako zaměstnanci.

Výše uvedený přístup má přínos jak pro univerzitu, tak pro partnerskou firmu. Z hlediska univerzity je to zdroj nových atraktivních témat pro výzkum, zpětná vazba a v neposlední řadě je to možnost vyzkoušet si nové koncepty na reálných datech. U nových výzkumných projektů je možno již při jejich formulování uvažovat o jejich realizaci ve spolupráci s průmyslem. Spolupráce s průmyslovými partnery také zvyšuje atraktivitu studia z pohledu studentů.

Partnerská firma pak profituje ze spolupráce hned několika způsoby. Zaprvé je to možnost získat kvalifikované pracovníky a definovat jejich odborné zaměření. Zvláště za současného stavu, kdy poptávka po kvalifikovaných zaměstnancích v oblasti IT převyšuje nabídku, to poskytuje nezanedbatelnou konkurenční výhodu. Zadruhé je to přímý přístup k výsledkům výzkumu a lidem, kteří ho prováděli – mj. kontakt na nejlepší české výzkumníky.

Konečně pro studenty je tu možnost podílet se na výzkumu, který produkuje prakticky použitelné výsledky. Během studia se student může zúčastnit stáží v partnerské firmě, kde může jednak poznat její prostředí a zároveň i prezentovat své schopnosti. Výsledkem je pak možnost získat zajímavé (a dobře placené) zaměstnání ihned po absolvování studia.

Pro navázání spolupráce je tedy nutné nejprve identifikovat okruhy spolupráce a oslovit garanta budoucí spolupráce na FI MU. Přesné vymezení obsahu spolupráce je pak otázkou dvoustranné dohody mezi budoucími partnery. Centrum ITI může v rámci svého růstu tyto projekty aktivně podpořit (pokud jsou konzistentní s výše uvedenými principy)

## SPOLUPRÁCE S ANF DATA

V následujícím textu budou představeny tři různé projekty ve spolupráci s firmou ANF DATA s.r.o. Kromě tématu se liší hlavně svým průběhem. V dalším textu se proto zaměříme převážně na průběh spolupráce a omezíme technické detaily.

Prvním z projektů je „*Statická analýza programů v jazyce C*“. Zadání projektu vychází z potřeby automaticky detekovat chyby v programech během jejich vývoje (a před tím, než budou nasazeny v praxi). Jisté typy chyb (memory leaks, chyby při práci s vlákny) je velmi obtížné objevit kontrolou kódu a objevují se nepravidelně. Proto bývá časově i finančně velmi náročné takové chyby najít a odstranit.

Zadáním projektu bylo vytvořit nástroj, který je schopen alespoň některé typy chyb automaticky objevit. V první fázi jsme se zaměřili na nabídku již existujících nástrojů. Zjistili jsme, že většinu těchto nástrojů není možné použít přímo ve vývojářské firmě, ale pouze na základě kontraktu (je nutno zaslat kód k analýze a vyčkat na výsledek, platí se za každý analyzovaný řádek). Nevýhod je tu hned několik: opakovaná analýza stojí další peníze, je tu prodloužení vzniklé zasláním kódu a čekáním na výsledek analýzy, a v neposlední řadě nemožnost změnit nástroj podle potřeb vývojářské firmy. Proto bylo rozhodnuto pokusit se na FI v rámci ITI vyvinout vhodný nástroj odstraňující výše uvedené nevýhody.

Projekt začal v roce 2006, kdy byl vytvořen tým skládající se z vedoucího, který je zaměstnancem ITI, tří studentů pracujících na projektu v rámci své diplomové práce, jednoho studenta bakalářského studia a dvou studentů pracujících na izolovaných částech nástroje po dobu jednoho semestru v rámci projektu do relevantního předmětu (např. projekt z překladačů).

V první fázi se sledovaly možnosti existujících nástrojů a relevantní teoretické výsledky – tato fáze měla podobu pravidelných seminářů, kde studenti prezentovali přidělené články a posléze se o nich vedla diskuze. Na základě této fáze bylo rozhodnuto o iniciální struktuře nástroje a byla formulována konkrétní témata pro diplomové práce a s nimi související úkoly. Vlastní vývoj včetně základního testování probíhal do června 2007, kdy byly odevzdány a úspěšně obhájeny všechny tři diplomové práce (dvě z nich s nejlepším možným hodnocením).

V létě 2007 pak tři studenti (dva z nich již vlastně absolventi) nastoupili do firmy ANF DATA na měsíční stáž, během které měli vyzkoušet vyvinutý nástroj (v jeho tehdejší podobě) na reálných kódech, a kromě toho identifikovat nejčastěji se vyskytující typy chyb a možné další směry vývoje nástroje. Stáž byla velmi úspěšná – nástroj (stále ve vývoji) odhalil dosud neobjevené chyby v hojně používaném a testovaném softwaru. Také došlo ke shodě, jakým směrem bude vývoj pokračovat.

Druhým projektem, který bychom zde chtěli prezentovat, je „*Dynamický BNF parser*“. Cílem tohoto projektu je vytvoření specializovaného programu na základě konkrétních požadavků za strany ANF DATA. Zadání projektu je postaveno na samé hranici zvládnutelnosti a musí využít nejnovějších poznatků v oboru.

Stejně jako v předchozím projektu studenti nejdříve zjišťovali, jestli požadovaný nástroj již neexistuje. Když žádný takový nástroj nebyl nalezen, studenti důkladně prozkoumali odbornou literaturu a v jednom z teoretických článků našli vhodný velmi sofistikovaný algoritmus, který lze použít k řešení zadaného problému. Následovala fáze dokonalého obeznámení se s algoritmem a jeho implementace v programovacím jazyku C++. Vzniklý program obsahuje netriviální začlenění nalezeného algoritmu do oblasti požadované v zadání tohoto projektu.

Na základě tohoto projektu právě vzniká velmi kvalitní bakalářská práce a v dalších fázích vylepšování programu je potenciál i na pokračování, které může být dostatečně silné pro vytvoření navazující magisterské práce. Celý projekt bude zveřejněn a celému světu zpřístupněn pod otevřenou licencí LGPL na webu [sourceforge.net](http://sourceforge.net).

Třetím projektem je "*Automatická detekce chyb v aplikacích na bázi protokolu SIP*". V tomto případě nebylo na začátku projektu žádné konkrétní zadání, ale pouze chuť spolupracovat. Úvodní úsilí bylo tedy věnováno na zkoumání možností budoucí spolupráce. Probíhaly vzájemné konzultace, studium společných oblastí zájmu a v neposlední řadě stáže studentů přímo na pracovišti v ANF DATA. Na základě této nelehké práce, postavené především na otevřeném přístupu k protistraně, byla pro další spolupráci vybrána problematika formální specifikace rozsáhlých projektů postavených na síťových protokolech. Cílem projektu je nalézt vhodný formalismus pro popsání návrhu síťových protokolů a vytvořit softwarový nástroj, který je schopen automaticky detekovat případné chyby, či podezřelá místa v tomto návrhu.

Jde o značně komplexní a ambiciózní projekt, který vyžaduje hlubší znalost jak komunikačních protokolů, tak formalizmů pro popis síťových protokolů i soudobých metod pro automatizovanou verifikaci. Dochází zde tedy k propojení oblasti teoretické informatiky s problematikou velmi úzce svázanou s praktickými problémy.

Projekt je ve fázi specifikace požadavků na plánovaný nástroj. V uplynulém období byly zkoumány existující nástroje a procházeny související teoretické články. Jako příklad uveďme zapůjčení komerčního nástroje, jehož nákup by komerční firmu stál značné finanční prostředky. V prostředí fakulty byl tento program zdarma vyzkoušen. Nebyl však shledán dostatečně dobrým. Firmě, která nástroj zapůjčila, bylo zpětně zasláno oznámení o nalezených chybách a funkčních nedostatcích.

Výsledky celého průzkumu byly shrnuty v technické zprávě (33 stran), které vyšla na Fakultě Informatiky MU. Dále byly poznatky prezentovány na dvouhodinové přednášce určené zaměstnancům ANF DATA. Z této přednášky byl pořízen videozáznam. Stručně lze shrnout, že jako vhodný modelovací formalismus byl vybrán MSC (Message Sequence Charts, ITU-T recommendation Z.120). Nebyl však nalezen žádný existující nástroj splňující cíle projektu. Na druhou stranu zkoumání v oblasti teoreticky stavěných článků potvrzuje realizovatelnost takového nástroje.

V následujícím roce plánujeme vytvoření specifikace požadavků na nově vytvořený nástroj, dále chceme přehledně shrnout známé a publikovat nové verifikační algoritmy pro MSC. No-

vě získané poznatky jsou od podzimního semestru 2007 vyučovány v rámci seminární skupiny jednoho předmětu na Fakultě informatiky. Vyzdvihněme, že je do budoucna je ze strany fakulty zájem o vznik nového takto koncipovaného předmětu. Za zmínku též stojí, že jsme vzhledem rozsáhlosti a potenciálu tohoto projektu začali jednat o možnosti externí financování ze samostatného grantu.

Mgr. Jan Obdržálek, Ph.D., Mgr. Vojtěch Řehák, Ph.D.  
Fakulta informatiky MU  
Botanická 68a  
602 00 Brno  
+420 549 49 4225, +420 549 49 4687  
[obdrzalek@fi.muni.cz](mailto:obdrzalek@fi.muni.cz), [rehak@fi.muni.cz](mailto:rehak@fi.muni.cz)