# Unifikace

**Petr Štěpánek**

**S využitím materialu Krysztofa R. Apta**

**2006**

**Unifikace** je ten typ substitucí, který dělá rezoluční metodu efektivním prostředkem automatického dokazování. Jsou to substituce, které mají klíčový význam pro logické programování, (tedy dokazování v Hornově logice) a v širším smyslu pro strojové dokazování v celé predikátové logice.

Intuitivně řečeno, jde o proces, který převádí dva nebo více termů na společnou instanci (pokud existuje) pomocí určitých substitucí.

Mějme například termy $f(a,y,z)$ a $f(x,b,z)$, kde $a$, $b$ jsou konstanty a $x, y, z$ proměnné. Jednu z jejich společných instancí

$$f(a,b,a) \tag{1}$$

můžeme získat substitucí

$$\gamma = \{x/a,\ y/b,\ z/a\}$$

Společnou instanci $f(a,b,z)$ (2)

získáme také substitucí $\theta = \{x/a,\ y/b\}$

Takovým substitucím říkáme unifikace. První substituce $\gamma$ je speciál-ním případem druhé substituce $\theta$, která se jeví jako „obecnější" než substituce $\gamma$. Také term (1) je instancí termu (2). Substituce $\gamma$ zbytečně váže proměnnou $z$ konstantou $b$.

V tomto případě je $\theta$ nejobecnější unifikace termů

$$f(a,y,z) \quad \text{a} \quad f(x,b,z)$$

Začneme formální definicí pojmu, který je důležitý pro zavedení pojmu nejobecnější unifikace.

**Definice.**

Říkáme, že substituce $\theta$ je obecnější než substituce $\gamma$, jestliže $\gamma = \theta\eta$ pro nějakou substituci $\eta$. Častokrát se tento vztah vyjadřuje nerovností $\gamma \leq \theta$.

Substituce $\theta$ je obecnější než $\gamma$, jestliže $\gamma$ lze získat ze substituce $\theta$ složením (aplikací) s další substitucí $\eta$. Protože za $\eta$ můžeme vzít také prázdnou substituci $\varepsilon$ platí, že každá substituce je obecnější než ona sama.

Tato relace je tedy reflexivní. Využijeme-li asociativnost skládání substitucí, ukážeme, že je to také tranzitivní relace.

I když vypadá jednoduše, předchozí substituce skrývá mnoho jemností.

## Příklady 1.

•Jak bychom očekávali, substituce $\{x/y\}$ je obecnější než substituce $\{x/a, y/a\}$, protože $\{x/y\}\{y/\mathrm{a}\}$ se rovná $\{x/a, y/a\}$

•Naproti tomu může překvapit, že substituce $\{x/y\}$ není obecnější než substituce $\{x/a\}$.

Jednoduchou úvahou zjistíme, že kdyby pro nějakou substituci $\eta$ platilo $x/a \in \{x/y\}\eta$, potom $y/a \in \eta$ a tedy $y \in Dom(\{x/y\}\eta)$.

(iii) Podobně $\{x/f(y,z)\}$ není obecnější než $\{x/f(a,a)\}$ , kde proměnné $y$ a $z$ jsou různé (i když jedna z nich může být proměnná $x$).

Postupujeme sporem, kdyby první substituce byla obecnější než druhá , potom by rovnost $\{x/f(y,z)\}\gamma = \{x/f(a,a)\}$ platila pro nějakou substituci $\gamma$ .

Tedy substituce $\gamma$ by obsahovala vazby $y/a$ a $z/a$ , z nichž alespoň jedna není vazbou pro $x$ . To by znamenalo, že $Dom(\{x/f(y,z)\}\gamma \neq Dom(\{x/f(a,a)\}$ - spor.

Tyto příklady nebo spíše použité metody stojí za zapamatování, protože hrají svou úlohu ve výkladu teorie logického programování.

Obdobou lemmatu o Variantách pro termy je následující tvrzení pro substituce.

**Lemma .** (Přejmenování)

Pro libovolné dvě substituce $\theta$, $\eta$ platí

$$\eta \leq \theta \text{ a } \theta \leq \eta \text{ právě když } \eta = \theta\gamma$$

pro nějaké přejmenování $\gamma$, takové že $Var(\gamma) \subseteq Var(\theta) \cup Var(\eta)$

Důkaz ($\Rightarrow$) Nechť $t$ je term takový, že $Var(t) = Var(\theta) \cup Var(\eta)$. Tento předpoklad nemusí být vždy splněn. Pokud jazyk neobsahuje žádný funkční symbol četnosti větší než 1 takový term nemusí existovat. V takovém případě jazyk rozšíříme o jeden binární funkční symbol a pokračujeme v důkazu v tomto rozšířeném jazyce.

Podle předpokladu lemmatu je $t\eta$ instance $t\theta$ a $t\theta$ je instance $t\eta$ , podle Lemmatu o Variantách lze sestrojit přejmenování $\gamma$ , takové že

$$t\eta = t\theta\gamma$$

Podle stejného lemmatu a podle volby termu $t$ také

$$Var(\gamma) \subseteq Var(\theta) \cup Var(\eta)$$

Odtud

$$Var(\theta\gamma) \subseteq Var(t) \quad \text{a} \quad Var(\eta) \subseteq Var(t)$$

.

a proto $\eta = \theta\gamma$ .

($\Leftarrow$) Stačí na obě strany předpokládané rovnosti $\eta = \theta\gamma$ použít inverzní přejmenování $\gamma^{-1}$ a dostaneme $\eta\gamma^{-1} = \theta\gamma\gamma^{-1} = \theta$ .

Následující pojmy hrají klíčovou roli v dalším výkladu.

**Definice** (Unifikace termů)

(i) Říkáme, že substituce $\theta$ je *unifikací termů* $s$ a $t$ jestliže $s\theta = t\theta$ . V takovém případě říkáme, že termy $s$ *a* $t$ jsou *unifikovatelné* .

(ii) říkáme, že $\theta$ je *nejobecnejší unifikace* (zkratka *mgu*) $s$ a $t$ je-li to unifikace termů $s$ a $t$ a je obecnější než všechny unifikace $s$ a $t$ .

(iii) Nejobecnější unifikace (mgu) $\theta$ termů $s$ a $t$ je *silná* , jestliže pro libovolnou unifikaci $\eta$ těchto termů platí $\eta = \theta\eta$ .

Definované pojmy vycházejí z přirozené intuice: nejobecnější unfikace dvou termů vytvoří společnou instanci obou termů “nejobecnějším způsobem” , tedy bez zbytečných vazeb.

To znamená, že $\theta$ je nejobecnější unifikace nějakých dvou termů, jestliže pro každou jejich unifikaci $\eta$ lze sestrojit substituci $\gamma$, takovou že $\eta = \theta\gamma$.

Pojem silné mgu není tak důležitý, ale Robinsonův algoritmus unifikace algoritmus Martelliho a Montanariho i jiné algoritmy unifikace produkují unfikace (pokud existují) s touto vlastností.

Tento jev je známý nejenom z Informatiky, ale také z Matematiky: když sestrojujeme nějaký objekt (zobrazení, možinu atd.) výsledek má často silnější vlastnosti než bylo původně požadováno.

V našem výkladu popíšeme a budeme analyzovat jen algoritmus Martelliho a Montanariho.

## Příklady.

(i) Mějme termy $f(g(x,a), z)$ a $f(y,b)$ . Potom

$$\eta = \{x/c,\ y/g(c,a),\ z/b\}$$
$$\theta = \{y/g(x,a),\ z/b\}$$

jsou dvě možné unifikace těchto termů a $\theta$ je obecnější než unifikace $\eta$ , protože

$$\eta = \{x/c,\ y/g(c,a),\ z/b\} = \{y/g(c,a),\ z/b\}\{x/c\} = \theta\{x/c\}$$

přitom není těžké ukázat, že $\theta$ je mgu pro termy $f(g(x,a), z)$ a $f(y,b)$ dokonce je to silná unifikace, protože platí

$$\eta = \{x/c,\ y/g(c,a),\ z/b\} = \{\ y/g(x,a),\ z/b\}\{x/c,\ y/g(c,x),\ z/b\} = \theta\eta$$

## Varovné příklady

(ii) Uvažujme termy $f(g(x,a), z)$ a $f(g(x,b), b)$ . Nejsou unfikovatelné, protože pro žádnou substituci $\theta$ nemůže platit $a\theta = b\theta$.

(iii) Také termy $g(x,a)$ a $g(f(x),a)$ nejsou unfikovatelné protože pro každou substituci je term $x\theta$ vlastním podtermem termu $f(x)\theta$ .

Takzvaný *unifikační problém* formulovaný tak, že pro libovolné dva termy je třeba rozhodnout zda jsou unifikovatelné, má positivní řešení.

Lze sestrojit algoritmus, který se zastaví s neúspěchem, pokud dané termy nejsou unifikovatelné a v opačném případě sestrojí nejobecnější unifikaci obou termů, dokonce takovou, která je silná.

V obecném případě dva termy nejsou unifikovatelné ze dvou důvodů:

a) Dva složené termy, které začínají různými funkčními symboly nelze unifikovat.

Tento případ je ilustrován příkladem (ii), nejjednodušším v této kategorii, dvě různé konstanty nelze unfikovat,

b) libovolnou proměnnou $x$, která je vlastním podtermem nějakého termu $t$, nelze s termem unfikovat (i když se nabízí substituce $\{x/t\}$). Tento případ jeznázorněn v příkladu (iii) na nejednodušším možném příkladu, kde $t = f(x)$.

Oba případy neúspěchu unifikace jsou součástí každého unifikačního algoritmu. Druhý z nich je třeba řešit testem výskytu proměnné $x$ v termu $t$. Proto si v anglické literatuře vysloužil jméno "Occur-check failure".

Test proměnných se ve většině implementací vynechává pro svou časovou náročnost. Korektností algoritmu unifikace v takovém případě se budeme podrobně zabývat později.

V důkazech korektnosti se užívá následující jednoduché lemma.

**Lemma.** (Vazba proměnné a termu)

Nechť $\theta$ je libovolná substituce , $x$ je libovolná proměnná a $t$ je libovolný term. Potom platí

$$x\theta = t\theta \quad \text{právě když} \quad \theta = \{x/t\}\theta \qquad (1)$$

Důkaz. Snadno se nahlédne, že pro libovolnou substituci platí

$$x\{x/t\}\theta = t\theta \tag{2}$$

odkud dostáváme

$$x\theta = t\theta \quad \text{právě když} \quad x\theta = x\{x/t\}\theta. \tag{3}$$

Zbývá ověřit, že ve tvrzení (3) lze za předpokladu uvedeného na levé straně pravou stranu ekvivalence nahradit rovností substitucí $\theta$ a $\{x/t\}\theta$.

Je-li $y$ libovolná proměnná různá od $x$ platí

$$y\{x/t\} = y$$

$$y\{x/t\}\theta = y\theta \tag{4}$$

Potom (2) a (4) a definice rovnosti ukazují, že pravou stranu tvrzení (3) lze nahradit rovností substitucí $\theta = \{x/t\}\theta$.

## Unifikační algoritmy.

První unifikační algoritmus sestrojil J. A. Robinson při zavedení Resoluční metody v roce 1965. Idea tohoto algoritmu je přirozená, oba termy $s$ a $t$ čte současně zleva a identifikuje takzvané dvojice neshody. Nicméně není tak jednoduchý jak by se mohlo zdát.

Měřením experimentů měl Robinsonův algoritmus přibližně kvadratickou složitost v počtu symbolů termů $s$ a $t$. Tento algoritmus má řadu variant, které umožňují zrychlení výpočtu. Těmito algoritmy se nebudeme zabývat.

Podrobně rozebereme unifikační algoritmus Martelliho a Montanariho (dále jen MM) z roku 1982 a dokážeme jeho korektnost.

Tento algoritmus místo unifikace dvou termů $s$ a $t$ řeší zdánlivě obecnější problém unifikování konečných množin dvojic termů, které jsou zapsány jako množiny rovností tvaru

$$\{s_1 = t_1, s_2 = t_2, \dots, s_n = t_n\} \tag{1}$$

Rovnost zde znamená jen výraz $s_i = t_i$ kde na obou stranách rovnosti mohou být různé výrazy, nejde tedy o rovnici v algebraickém smyslu .

MM algoritmus pracuje s takovými množinami rovností. K popisu algoritmu je nutné zavést potřebné značení a terminologii.

**Definice.** (aplikace substituce na množinu rovností)

(i) Výsledek aplikace substituce $\theta$ na množinu rovností $E$ tvaru (1) značíme $E\theta$ a definujeme rovností

$$\{s_1 = t_1, s_2 = t_2, \dots, s_n = t_n\}\theta = \{s_1\theta = t_1\theta, s_2\theta = t_2\theta, \dots, s_n\theta = t_n\theta\}$$

(ii) Říkáme, že $\theta$ je unifikace množiny rovností $E$ tvaru

$$\{s_1 = t_1 , s_2 = t_2 , \dots , s_n = t_n \}$$

jestliže platí $s_1\theta = t_1\theta , s_2\theta = t_2\theta , \dots , s_n\theta = t_n\theta$ .

Jsou-li $s$ a $t$ termy, jednoprvková množina $\{s = t\}$ má stejnou množinu unifikací jako termy $s$ a $t$ .

(iii) Říkáme, že unifikace $\theta$ množiny rovností $E$ je nejobecnější unifikace množiny $E$ (krátce *mgu*) , je-li obecnější než všechny unifikace množiny $E$ ,

(iv) říkáme, že nejobecnější unifikace $\theta$ množiny $E$ je silná, jestliže pro libovolnou unifikaci $\eta$ množiny $E$ platí $\eta = \theta\eta$

**Definice.** (Vlastnosti množin rovností)

(i) Říkáme, že *dvě množiny rovností jsou ekvivalentní* jestliže mají stejnou množinu unifikací,

(ii) říkáme, že *množina rovností je vyřešená* je-li tvaru

$$\{x_1 = t_1, x_2 = t_2, \dots, x_n = t_n\}$$

kde proměnné $x_i$ , $1 \leq i \leq n$ jsou navzájem různé a žádná z nich se nevysktuje v termech $t_j$ , $1 \leq j \leq n$ .

**Lemma.** (Vyřešené množiny rovností a unifikace)

Je-li $E = \{x_1 = t_1, x_2 = t_2, \dots, x_n = t_n\}$
vyřešená množina rovností, potom substituce

$$\theta = \{x_1/t_1, x_2/t_2, \dots, x_n/t_n\}$$

je silná nejobecnější unifikace množiny $E$ . Říkáme, že substituce $\theta$ je unifikace určená množinou $E$ .

Důkaz. Nejprve si všimneme, že $\theta$ je unifikace $E$: pro $i = 1, 2, \dots, n$ platí $x_i\theta = t_i$ a také $t_i\theta = t_i$ protože podle předpokladu žádné $x_i \varepsilon Dom(\theta)$ se nevyskytuje v $t_j$.

Dále předpokládejme, že $\eta$ je libovolná unifikace množiny $E$. Potom pro $i = 1, 2, \dots, n$ platí $x_i\eta = t_i\eta = x_i\theta\eta$ protože $t_i = x_i\theta$. Dále pro libovolnou proměnnou $x$, která není prvkem $Dom(\theta) = \{x_1, x_2, \dots, x_n\}$ dostáváme $x\eta = x\theta\eta$, protože $x = x\theta$. Celkem tedy $\eta = \theta\eta$ a $\theta$ je silná nejobecnější unifikace množiny $E$.

Z důkazu samého bezprostředně plyne, že k sestrojení nejobecnější unifikace pro nějakou množinu rovností $E$ stačí transformovat $E$ na ekvivalentní množinu rovností ve vyřešeném tvaru.

Martelliho a Montanariho algoritmus takovou transformaci provede, pokud je to možné, a skončí neúspěchem jestliže to nejde.

## Martelliho a Montanariho unifikační algoritmus

Je-li dána množina rovností $E$ nedeterministicky vyber z množiny $E$ rovnost $e$ v některém z následujících tvarů a transformuj množinu $E$ akcí, která je přiřazena k $e$ podle následujících pravidel.

(1) rovnost $f(s_1, s_2, \ldots, s_n) = f(t_1, t_2, \ldots, t_n)$ *nahraď rovnostmi* $s_1 = t_1, s_2 = t_2, \ldots, s_n = t_n$

(2) $f(s_1, s_2, \ldots, s_n) = g(t_1, t_2, \ldots, t_m)$ kde $f \neq g$, *znamená* STOP *s neúspěchem ,*

(3) $x = x$ *tuto rovnost vynechej,*

(4) rovnost $t = x$, kde $t$ není proměnná, *nahraď rovností* $x = t$ ,

(5) je-li $x = t$ rovnost kde $x$ není prvkem $Var(t)$ ale vyskytuje se v některé jiné rovnosti z $E$ — *aplikuj substituci* $\{x/t\}$ *na všechny ostatní rovnosti z* $E$

(6) je-li $x = t$ rovnost, kde $x$ je prvkem $Var(t)$ a $x \neq t$ — STOP *s neúspěchem*

MM algoritmus končí výpočet, jestliže je zastaven příkazem STOP v akci (2) nebo (6), nebo v případě, že nelze žádnou jinou akci provést.

Chceme-li použít MM algoritmus k unifikaci termů $s$ , $t$ aktivujeme ho vstupem $E = \{s = t\}$ .

Povšimněme si, že akce (1) zahrnuje i případ $c = c$ pro libovolnou konstantu $c$ , který vede k vynechání této rovnosti (konstanta je funkční symbol, který nevyžaduje žádné argumenty) .

Podobně akce (2) zahrnuje i případ dvou různých konstant.

**Věta.** (Korektnost Montanariho a Martelliho algoritmu)

Algoritmus Martelliho a Montanariho vždy ukončí svůj výpočet. Pokud je vstupní množina rovností $E$ unifikovatelná, algoritmus úspěšně skončí výpočet a vydá vyřešenou množinu rovností, která určuje silnou nejobecnější unifikaci možiny $E$. Jinak výpočet končí neúspěchem.

Důkaz je důsledkem následujících čtyř lemmat.

**Lemma 1.** MM algoritmus vždy končí svůj výpočet.

Důkaz lemmatu. Nejprve zavedeme potřebné pojmy:

a) je-li dána množina rovností $E$, říkáme, že proměnná $x$ je vyřešená v $E$, jestliže $E$ obsahuje rovnost $x = t$ pro nějaký term $t$ a jediný výskyt proměnné $x$ v množině $E$ je na levé straně této rovnosti. Jinak říkáme, že proměnná $x$ není vyřešená v $E$.

b) Ke každé množině rovností $E$ přiřadíme následující tři přirozená čísla.

$uns\,(E)$ - počet nevyřešených proměnných v $E$,

$lfun\,(E)$ - celkový počet výskytů funkčních symbolů na levých stranách rovností v $E$,

$card\,(E)$ - počet rovností v $E$.

Uvažujeme uspořádané trojice přirozených čísel

$$(uns\,(E)\,,\, lfun\,(E)\,,\, card\,(E))$$

a ukážeme, že každá úspěšná akce algoritmu provedená na množinu $E$ s výsledkem $E'$ sníží pořadí trojice

$$(uns\,(E')\,,\, lfun\,(E')\,,\, card\,(E'))$$

v lexikografickém uspořádání trojic nejméně o jedno místo.

Je zřejmé, že žádná akce algoritmu nezmění vyřešenou proměnnou na nevyřešenou, takže *uns* ($E$) nikdy nevzroste. Dále platí:

- akce (1) sníží hodnotu *lfun* ($E$) o jednotku,
- akce (3) nezmění hodnotu *lfun* ($E$) , ale sníží hodnotu *card* ($E$) o jednotku,
- akce (4) sníží hodnotu *lfun* ($E$) nejméně o jednotku a
- akce (5) sníží hodnotu *uns* ($E$) o jednotku .

Přitom akce (2) a (6) zakončují výpočet neúspěchem. Protože lexikografické uspořádání trojic přirozených čísel nemá nekonečný klesající řetězec, algoritmus při každém vstupu zakončí svůj výpočet.

**Lemma 2.** Každá úspěšná akce nahradí množinu rovností ekvivalentní množinou rovností.

Důkaz. Lemma platí pro akci (1) protože pro libovolnou substituci $\theta$ platí

.
$$f(s_1, s_2, \ldots, s_n)\theta = f(t_1, t_2, \ldots, t_n)\theta$$

právě když

pro každé $i = 1, 2, \ldots, n$ platí $s_i\theta = t_i\theta$ .

Pro akce (3) a (4) je platnost tvrzení zřejmá. Pro akci (5) uvažujme dvě množiny rovností

.
$$E \cup \{x = t\} \quad \text{a} \quad E\{x/t\} \cup \{x = t\} ,$$

kde $E\{x/t\}$ označuje množinu rovností, která vznikne z $E$ provedením substituce $\{x/t\}$ na každou rovnost z $E$ .

Potom platí

$\theta$ je unifikace $E \cup \{x = t\}$ <=> $\theta$ je unifikace $E$ a platí $x\theta = t\theta$

(vazba proměnné a termu) <=> $\{x/t\}\theta$ unifikuje $E$ a platí $x\theta = t\theta$

<=> $\theta$ unifikuje $E\{x/t\}$ a platí $x\theta = t\theta$

<=> $\theta$ unifikuje $E\{x/t\} \cup \{x = t\}$

**Lemma 3.** Jestliže výpočet algoritmu skončí úspěšně, výsledkem je vyřešená množina rovností.

Důkaz. Jestliže algoritmus úspěšně skončí, pak k výsledné množině rovností není použitelná žádná akce. Speciálně nejsou použitelné akce (1) , (2) a (4) , takže levé strany všech rovností jsou proměnné.

Navíc nejsou použitelné ani akce (3) , (5) a (6) takže tyto proměnné jsou navzájem různé a žádná z nich se nvyskytuje na pravé straně nějaké rovnosti.

**Lemma 4.** Skončí-li algoritmus neúspěchem, potom množina rovností v okamžiku neúspěchu není unifikovatelná.

Důkaz. Je-li neúspěch způsoben akcí (2), potom v odpovídajícím kroku je vybrána rovnost

$$f(s_1, s_2, \dots, s_n) = g(t_1, t_2, \dots, t_m)$$

a pro žádnou substituci $\theta$ neplatí

$$f(s_1, s_2, \dots, s_n)\theta = g(t_1, t_2, \dots, t_m)\theta .$$

Je-li neúspěch způsoben akcí (6) , v posledním kroku je prvkem odpovídající množiny rovností nějaká rovnost $x = t$ , taková, že $x \in Var(t)$ a pro žádnou substituci $\theta$ neplatí $x\theta = t\theta$ , protože $x\theta$ je vlastním podtermem termu $t\theta$ .

Tím je věta dokázána. (QED jak říkali staří latiníci.)

## Příklady.

a) Mějme množinu rovností

$$\{k\,(z, f\,(x, b, z)) = k\,(h(x), f\,(g(a)\,, y, z))\} \qquad (1)$$

Akce (1) je aplikovatelná a dostaneme

$$\{z = h(x)\,, f\,(x\,,b\,,z) = f\,(g(a), y, z))\}$$

Vybereme-li druhou rovnost a použijeme opět Akci (1) dostáváme

$$\{z = h(x)\,, x = g(a)\,, b = y\,, z = z\}$$

Po použití Akce (4) a Akce (3) dostaneme

$$\{z = h(x)\,, x = g(a)\,, y = b\}$$

Nyní můžeme použít Akci (5) a dostaneme vyřešenou množinu rovností

$$\{z = h(g(a)\,, x = g(a)\,, y = b\}$$

Teď už není žádná akce aplikovatelna a podle věty o vyřešené množině rovností, dostáváme nejobecnější unifikaci množiny (1)

$$\{z / h(g(a)) , x / g(a) , y / b\}$$

b) Uvažujme množinu

$$\{k (z,\ f(x, b, z)) \ = \ k (h (x),\ f( g (z), y, z))\} \qquad (2)$$

která je přiřazena ke dvojici termů

$$k (z,\ f(x, b, z)) \quad \text{a} \quad k (h (x),\ f( g (z), y, z))$$

Provedením Akce (1) získáme množinu

$$\{z = h (x),\ f(x, b, z) = f( g (z), y, z))\}$$

Opětovné použití Akce (1) na druhou rovnost dává množinu

$$\{z = h (x), x = g(z), b = y, z = z\}$$

Nyní Akce (4) a (3) použité na třetí a čtvrtou rovnost dávají množinu

$$\{z = h\,(x),\ \ x = g(z),\ \ y = b\}$$

Zvolíme-li druhou rovnost, Akce (5) je použitelná a dává výsledek

$$\{z = h\,(g\,(z)),\ \ x = \ g\,(z),\ \ y = b\}$$

to ještě není vyřešená soustava. První rovnost ještě není vyřešená, protože obsahuje proměnnou $z$ na obou stranách rovnosti.

Zde je použitelná jen Akce (6), která zakončí výpočet neúspěchem.

Cvičení. Sestrojte množinu rovností, pro kterou Martelliho a Montanariho algoritmus dává dvě různé mgu.

## *Occur-check.*

Test konfliktu proměnných, tedy test

.                    " *proměnná x se nevyskytuje v termu  t*"

se v nějaké podobě najde v každém unifikačním algoritmu. Ve většině implementací Prologu se pro zvýšení efektivnosti unifikačního algoritmu tento test vynechává.

V případě Martelliho a Montanariho algoritmu to znamená, že v Akci (5) vynecháme test proměnných (tedy occur-check) a vypustíme celou poslední Akci (6).

V této situaci máme dvě možnosti, které závisí na tom, zda nabízející se substituci $\{x/t\}$ provedeme jen na samotný term $t$ .

a) jestliže to uděláme, nekorektnost výpočtu může nastat tím, že $x$ se vyskytuje v $t$ a tedy také v $t\{x/t\}$ .

b) jestliže to neuděláme, potom nesprávný výsledekm může vzniknout jako v případě jediné rovnosti $x = f(x)$ , která vede na substituci $\{x/f(x)\}$.

V praxi se zpravidla implementuje možnost b) . Zjednodušený algoritmus unifikace přesto v mnoha případech generuje korektní výpočty.

Zatím nemáme k dispozici klíčový pojem výpočtů, které provádějí logické programy. Pojem korektntosti výpočtů teprve budeme definovat. Pak teprve budeme moci analyzovat důsledky zjednodušení unifikačního algoritmu.

Této problematice se budeme věnovat podrobněji později. Zatím ukážeme vlastnosti nejobecnějších unifikací, které jsou generovány známými unifikačními algoritmy. Je to nutná příprava k řešení takových problémů, jako je právě "occur chceck".

Termín 'unifikace' používáme volně, bez uvádění termů nebo množin rovností, které unifikují.

## Vlastnosti substitucí, unifikací a mgu.

**Definice.** (Ekvivalentní substituce)

Říkáme, že dvě unifikace $\theta$ a $\eta$ *jsou ekvivalentní* , jesliže pro nějaké přejmenování $\gamma$ platí $\eta = \theta\gamma$ .

Ukážeme, že pro libovolné dva termy jsou všechny nejobecnejší unifikace (mgu) ekvivalentní. To je pro naše účely vitální fakt. Dříve se však budeme zabývat vlastnostmi, které nám dovolí rozlišovat mezi různými typy substitucí a mgu.

**Definice.** (Idempotentní substituce)

Říkáme, že substituce $\theta$ je *idempotentní* . jestliže $\theta\theta = \theta$ .

**Věta.** (Idempotence)

Nejobecnější unifikace (mgu) je silná, právě když je to idempotentní mgu.

Důkaz. (=>) Nechť $\theta$ je silná mgu. $\theta$ je ovšem také unifikace, proto z definice silné mgu dostáváme $\theta = \theta\theta$ . Tedy $\theta$ je idempotentní mgu.

(<=) Předpokládejme naopak, že $\theta$ je idempotentní mgu a $\eta$ je unifikace. Potom pro nějakou substituci $\gamma$ platí $\eta = \theta\gamma$ . Využijeme idempotence $\theta$ a dostaneme $\eta = \theta\gamma = \theta\theta\gamma = \theta\eta$. Mgu $\theta$ je tedy silná.

**Lemma.** (Kriterium idempotence)

Substituce $\theta$ je idempotentní, právě když $Dom(\theta) \cap Ran(\theta) = 0$ .

Cvičení.

- Dokažte předchozí lemma,
- dokažte, že množina idempotentních substitucí není uzvřená na skládání,
- jsou-li však $\theta$ a $\eta$ idempotentní substituce takové, že

$$Dom(\theta) \cap Ran(\eta) = 0$$

dokažte, že potom $\theta\eta$ je také idempotentní substituce,

- dokažte, že pro libovolný term $s$ a idempotentní substituci $\theta$ platí

$$Var(s\theta) \cap Dom(\theta) = 0 .$$

**Příklad.** (mgu nemusí být idempotentní substituce .)

Pro libovolný term $t$ je substituce $\{x/y, y/x\}$ mgu $t$ a $t$ . Není ale idempotentní, protože $\{x/y, y/x\}\{x/y, y/x\} = \varepsilon$ .

**Definition.** (Relevantní unifikace)

Říkáme, že unifikace $\theta$ termů $s$ a $t$ je relevantní (vzhledem k daným termům) jestliže platí

$$Var(\theta) \subseteq Var(s) \cup Var(t)$$

.

- Relevantní unifikace dvou termů používá jen proměnné obsažené v těchto termech (i když ne nutně všechny), ale jiné proměnné do unifikace nezavádí.
- Prohlédnutím Martelliho a Montanariho unfikačního algoritmu se lze přesvědčit, že tento algoritmus produkuje jen relevantní unifikace.
- Toto pozorování vede k následujícímu tvrzení, které ukazuje proč jsou idempotentní mgu zajímavé.

**Věta.** (Relevance)

Každá idempotentní mgu $\theta$ nějakých termů $s$, $t$ je relevantní.

Důkaz. Nejprve dokážeme následující

Pozorování. Pro libovolné substituce $\theta$ , $\eta$ platí

$$Var(\theta) \subseteq Var(\eta) \cup Var(\theta\eta) \qquad (1)$$

.

Je-li $x \in Var(\theta)$ , potom $x \in Dom(\theta)$ nebo $x \in Ran(\theta)$ , k důkazu (1) stačí ověřit

$$Dom(\theta) - Dom(\theta\eta) \subseteq Ran(\eta) \qquad (2)$$

.

a

$$Ran(\theta) - Dom(\eta) \subseteq Ran(\theta\eta) \qquad (3)$$

Důkaz (2) , nechť $x \in Dom(\theta) - Dom(\theta\eta)$ . Pak pro nějaký term musí platit $x/t \in \theta$ (tedy nutně $t \neq x$) a $t\eta = x$ . Tedy $t$ je proměnná a $x \in Ran(\eta)$ odkud plyne (2).

Důkaz (3) , předpokládejme, že

$$x \varepsilon\ Ran(\theta)\ - Dom(\eta) \qquad (4)$$

.
potom pro nějakou dvojici $y/t\ \varepsilon\ \theta$ platí $y \neq t$, $x\ \varepsilon$ Var($t$) a $x \varepsilon\ Var\ (t\eta)$ .

Nyní, uvažujme dva případy:

a) kdyby $y = t\eta$, potom $t$ je proměnná a $t / y\ \varepsilon\ \eta$ . Ale protože jsme již ukázali, že $x\ \varepsilon\ Var(t)$ , pak musí platit $x = t$ . To znamená, že $x\ \varepsilon\ Dom(\eta)$ , a to je ve sporu s předpokladem (4) . Tedy,

b) $y \neq t\eta$ , to znamená, že $y / t\eta\ \varepsilon\ \theta\eta$ a $x\ \varepsilon\ Ran\ (\theta\eta)$ , protože podle předpokladu (4) $x$ je prvkem $Ran(\theta)$ a není prvkem $Dom(\eta)$ . Tím je (3) a také (1) dokázáno .

[konec důkazu pozorování]

Důkaz Věty. Podle předpokladu je $\theta$ idempotentní mgu termů $s$ , $t$ . Podle Věty o idempotenci je $\theta$ silná unifikace . Je-li $\eta$ li bovolná unifikace daných termů, dostáváme $\eta = \theta\eta$ . Z (1) dostáváme

$$Var(\theta) \subseteq Var(\eta) \tag{5}$$

.

Můžeme předpokládat, že $\eta$ je mgu $s$ a t sestrojená Martelliho a Montanariho algoritmem . Víme, že $\eta$ je relevantní a podle (5) je také unifikace $\theta$ relevantní. [konec důkazu věty]

**Věta.** (Ekvivalence mgu a substituce)

Nechť $\theta_1$ je mgu množiny rovností $E$ . Potom pro libovolnou substituci $\theta_2$ platí

$$\theta_2 \text{ je mgu množiny } E \text{ právě když } \theta_2 = \theta_1\gamma$$

.

pro nějaké přejmenování $\gamma$ takové, že

$$Var(\gamma) \subseteq Var(\theta_1) \cup Var(\theta_2)$$

.

Důkaz. (=>) Je-li $\theta_2$ také mgu množiny $E$, potom $\theta_2$ je obecnější než $\theta_1$ a naopak. Z věty o přejmenování pak existuje přejmenování $\gamma$, které splňuje pravou stranu ekvivalence.

(<=) Je-li $\gamma$ přejmenování, které splňuje pravou stranu ekvivalence, potom z věty o přejmenování je $\theta_2$ obecnější než $\theta_1$. Přitom $\theta_2 = \theta_1\gamma$ je unifikace množiny $E$, a to znamená, že $\theta_2$ je mgu množiny $E$.

Následující dvě lemmata o možnosti rozdělení unifikované množiny rovností $E$ na dvě množiny $E_1$ a $E_2$, o iterativním výpočtu nejobecnější unifikace množiny $E$ a o možné záměně pořadí obou množin při takovém výpočtu jsou technického charakteru a budou využity v dalším výkladu o výpočtech logických programů.

**Věta .** (Iterace)

(i) Nechť $E_1$ a $E_2$ jsou dvě množiny rovností. Předpokládejme, že $\theta_1$ je mgu množiny $E_1$ a $\theta_2$ je mgu množiny $E_2\theta_1$. Potom $\theta_1\theta_2$ je mgu množiny $E_1 \cup E_2$.

(ii) Naopak, je-li množina $E_1 \cup E_2$ unifikovatelná, pak existuje mgu $\theta_1$ množiny $E_1$ a pro libovolnou mgu $\theta_1$ množiny $E_1$ existuje také mgu $\theta_2$ množiny $E_2\theta_1$.

Důkaz. (i) Je-li $e$ rovnost z množiny $E_1$, je unifikována substitucí $\theta_1$, a také substitucí $\theta_1\theta_2$. Dále, je-li $e$ rovnost z množiny $E_2$, potom $e\theta_1$ je rovnost z množiny $E_2\theta_1$. To znamená, že $e\theta_1$ je unifikována substitucí $\theta_2$ a tedy $e$ je unifikována substitucí $\theta_1\theta_2$. Tím je dokázáno, že $\theta_1\theta_2$ je mgu množiny $E_1 \cup E_2$.

(ii) Předpokládejme naopak, že $\eta$ je unifikace $E_1 \cup E_2$. Protože $\eta$ unifikuje tuto množinu, existuje také nějaká mgu $\theta_1$ množiny $E_1 \cup E_2$. Z definice mgu plyne, že $\eta = \theta_1\lambda_1$ pro nějakou substituci $\lambda_1$ .

To znamená, že $\lambda_1$ je unfikace $(E_1 \cup E_2)\theta_1$ a také $E_2\theta_1$. Nechť $\theta_2$ je mgu množiny $E_2\theta_1$ , protože $\lambda_1$ unifikuje stejnou množinu, existuje substituce $\lambda_2$ , taková, že $\lambda_1 = \theta_2\lambda_2$ .

Celkem dostáváme $\eta = \theta_1\lambda_1 = \theta_1\theta_2\lambda_2$ . Tím je dokázáno, že $\theta_1\theta_2$ je mgu množiny $E_1 \cup E_2$.

Poznámka. Prohlédnutím důkazu Iterační věty lze ověřit, že ve znění lemmatu lze každý termín ‘mgu’ nahradit termínem ‘relevantní mgu’ .

**Důsledek.** (Přepínání - Switching)

Nechť $E_1$ , $E_2$ jsou dvě množiny rovností. Předpokládejme, že $\theta_1$ je mgu množiny $E_1$ a $\theta_2$ je mgu množiny $E_2\theta_1$. Potom množina $E_2$ je unifikovatelná a pro každou mgu $\theta_1'$ množiny $E_2$ existuje mgu $\theta_2'$ množiny $E_1\theta_1'$ taková, že $\theta_1'\theta_2' = \theta_1\theta_2$ .

Navíc, $\theta_2'$ lze zvolit tak, že platí

$$Var(\theta_2') \subseteq Var(E_1) \cup Var(\theta_1') \cup Var(\theta_1\theta_2) .$$

Důkaz. Podle Iteračního lemmatu (i) je $\theta_1\theta_2$ mgu $E_1 \cup E_2$ . Tedy $E_2$ je unifikovatelná množina. Nechť $\theta_1'$ je mgu množiny $E_2$ . Opět podle Iteračního lemmatu (ii) existuje mgu $\gamma$ množiny $E_1\theta_1'$ .

Můžeme zvolit $\gamma$ relevantní , tak že

$$Var(\gamma) \subseteq Var(E_1) \cup Var(\theta_1')$$

Nyní podle Iteračního lemmatu (i) je $\theta_1'\gamma$ mgu množiny $E_1 \cup E_2$ .
Podle Věty o ekvivalenci existuje přejmenování $\eta$ takové, že

$$\theta_1'\gamma\eta = \theta_1\theta_2 \qquad (1)$$

a pro $\eta$ navíc platí

$$Var(\eta) \subseteq Var(\theta_1'\gamma) \cup Var(\theta_1\theta_2)$$

Položíme-li $\theta_2' = \gamma\eta$ , podle (1) dostáváme

$$\theta_1'\theta_2' = \theta_1\theta_2$$

a podle Věty o ekvivalenci je $\theta_2'$ mgu množiny $E_1\theta_1'$ .

Přitom platí

$$Var(\theta_2') \subseteq Var(\gamma) \cup Var(\eta) \subseteq Var(E_1) \cup Var(\theta_1') \cup Var(\theta_1\theta_2)$$

Poznámka. Tento důsledek ukazuje, že při iterativním výpočtu mgu může být pořadí kroků zaměněno.