2 Unifikacja

Rozważamy język z symbolami funkcyjnymi i termy. Przypuśćmy, że mamy termy

$$t = f(x, g(a, b)) \text{ i } s = f(g(y, b), x).$$

W tych termach x i y są zmiennymi, a i b – stałymi, a f i g – dwuargumentowymi symbolami funkcyjnymi.

Klasyczne zadanie unifikacji to pytanie, czy dla danych termów t i s można podstawić w nich za zmienne takie termy, aby w wyniku podstawiania otrzymać termy identyczne?

Dla zdefiniowanych wyżej termów odpowiedź jest pozytywna; tak jest, jeżeli np. za x podstawiamy $g(a, b)$, za y – stałą a.

Załóżmy dodatkowo, że symbol f ma własność $f(x, y) = f(y, x)$. Teraz interesuje nas, czy jest podstawienie, które spowoduje, że termy t i s po podstawieniu staną się równe lub pozostaną różne, ale z założonej równosci wywnioskujemy ich równość. Na to pytanie dla podanych termów t i s damy odpowiedź pozytywną. W tym przypadku świadczy o tym podstawienie za y stałej a.

3 Zadanie unifikacji łącznej

Rozważamy zbiór zmiennych i pewien zbiór stałych i symboli funkcyjnych. Niech T oznacza zbiór termów zbudowanych z danych zmiennych, stałych i symboli funkcyjnych. Niech $\approx$ będzie binarną relacją w zbiorze T. Zadaniem unifikacji nazywamy następujący problem: dla danych termów t_1, t_2 ze zbioru T należy odpowiedzieć na pytanie, czy istnieje podstawienie (termów z T) takie, że po wykonaniu tego podstawienia w termach t_1 i t_2 otrzymujemy termy t'_1 i t'_2 takie, że $t'_1 \approx t'_2$.

Przyjmując, że $\approx$ jest relacją równości otrzymujemy klasyczne zadanie unifikacji ze znanym algorytmem rozstrzygającym to zadanie.

Często rozważa się zadania unifikacji, w których relacja $\approx$ jest definiowana za pomocą pewnego układu równości. Dokładniej, bierzemy zbiór R par termów, pary te interpretujemy jako równości i zbiór R domykamy ze względu na następujące reguły wnioskowania:

1. $\dfrac{}{t = s}$ dla wszystkich równości $t = s \in R$,
2. $\dfrac{}{t = t}$,
3. $\dfrac{t_1 = t_2}{t_2 = t_1}$,
4. $\dfrac{t_1 = t_2, t_2 = t_3}{t_1 = t_3}$,
5. $\dfrac{t_i = s_i \text{ dla } i = 1, \ldots, n}{f(t_1, \ldots, t_n) = f(s_1, \ldots, s_n)}$,

6. $\dfrac{t_1 = t_2}{\sigma t_1 = \sigma t_2}$ dla dowolnego podstawienia σ.

Otrzymany w ten sposób zbiór oznaczamy symbolem $E(R)$. Zbiór ten wyznacza relację

$$t_1 \approx t_2 \Leftrightarrow t_1 = t_2 \in E(R).$$

Niech teraz język zawiera stałe i jeden dwuargumentowy symbol f. Zadanie unifikacji łącznej polega na rozwiązaniu zadania unifikacji dla $R = \{f(x, f(y, z)) = f(f(x, y), z)\}$. i relacji $\approx$ zdefiniowanej jak wyżej.

Zadanie unifikacji łącznej jest równoważne problemowi niesprzeczności dla równań w słowach.

Dobierając odpowiednio zbiór R można otrzymać zadanie unifikacji równoważne dziesiątemu problemowi Hilberta. Zadanie to jest nierozstrzygalne. Nierozstrzygalne jest również zadanie unifikacji z relacją $\approx$ wyznaczoną przez zbiór

$$\begin{aligned} R = \{&f(x, f(y, z)) = f(f(x, y), z), g(x, g(y, z)) = g(g(x, y), z), \\ &g(x, f(y, z)) = f(g(x, y), g(x, z)), g(f(x, y), z) = f(g(x, z), g(y, z))\}. \end{aligned}$$

4 Zbiór słów

Alfabetem nazywamy dowolny, niepusty zbiór A, elementy alfabetu nazywamy literami. Symbol A^* oznacza zbiór słów nad alfabetem A. Słowo nad alfabetem A można definiować jako ciąg liter z A. Jeżeli mówimy o słowach, to ciąg jednoelementowy utożsamiamy z jego elementem. Tak więc zachodzi inkluzja $A \subseteq A^*$. Ciąg zeroelementowy nazywamy słowem pustym i oznaczamy symbolem ε.

Słowa łączymy (konkatenujemy). Konkatenacja jest operacją łączną. Zbiór A^* jako algebra jest rozważany z konkatenacją i słowem pustym. Jest więc algebrą z działaniem łącznym, które z kolei ma element neutralny. Taką algebrę nazywamy półgrupą z jednością lub monoidem. Będziemy rozważać najczęściej półgrupy z jednością. Wobec tego słowo półgrupa będzie oznaczać półgrupę z jednością, chyba, że zostanie napisane coś innego.

W algebrze A^* zachodzą prawa skracania: jeżeli $wx = wy$ (albo $xw = yw$), to także $x = y$. W naturalny sposób definiujemy pojęcie długości słowa. Zbiór słów jest uporządkowany relacją

$$x \leqslant y \Leftrightarrow \exists u \in A^* \; xu = y.$$

Jeżeli $x \leqslant y$, to x nazywamy prefiksem y. Zdefiniowana relacja jest porządkiem częściowym. Dodatkowo, dla tego porządku zachodzi następujące prawo: jeżeli $x \leqslant u$ i $y \leqslant u$, to $x \leqslant y$ lub $y \leqslant x$.

5 Homomorfizmy, półgrupy wolne

Jeżeli M_1 i M_2 są półgrupami (z jednością) i funkcja $f : M_1 \to M_2$, to f jest homomorfizmem wtedy i tylko wtedy, gdy $f(\varepsilon) = \varepsilon$ oraz $f(xy) = f(x)f(y)$ dla wszystkich $x, y \in M_1$. Jak zwykle definiujemy podobne pojęcia takie, jak izomorfizm.

Półgrupa M (z jednością) jest wolna wtedy i tylko wtedy, gdy jest izomorficzna z A^* dla pewnego zbioru A.

Przypuśćmy, że rozważamy półgrupę M, zbiór X i włożenie $i : X \to M$. W tej sytuacji mówimy, że półgrupa M ma własność uniwersalności (dla zbioru X) wtedy i tylko wtedy, gdy dla każdej półgrupy z jednością P, dla każdej funkcji $\alpha : X \to P$ istnieje dokładnie jeden homomorfizm $\varphi : M \to P$ taki, że $\alpha = \varphi \circ i$.

Lemat 5.1 *Półgrupa A^* ma własność uniwesalności dla zbioru A i naturalnego włożenia A w A^*.* $\square$

6 Generatory półgrupy

Zdefiniujmy dwie operacje w zbiorze podzbiorów półgrupy M. Jeżeli $X, Y \subseteq M$, to

$$XY = \{xy : x \in X \wedge y \in Y\}.$$

Zamiast XX będziemy też pisać X^2. Analogicznie definiujemy X^n. Symbol X^* oznacza również $\bigcup_{n \in N} X^n$. Zauważmy, że symbol X^* ma już kilka znaczeń.

Zbiór X generuje półgrupę M, jeżeli $M = X^*$.

Lemat 6.1 *Jeżeli M jest podpółgrupą A^* i $X = (M \setminus \{\varepsilon\}) \setminus (M \setminus \{\varepsilon\})^2$, to X jest minimalnym (lepiej: najmniejszym) zbiorem generatorów M.*

Dowód. Oczywiście, $X \subseteq M$, a więc $X^* \subseteq M^* = M$. Przeciwne zawieranie dowiedziemy pokazując przez indukcję ze względu na długość a, że jeżeli $a \in M$, to $a \in X^*$. Przypuśćmy, że $a \in M$. Są możliwe trzy przypadki.

Przypadek 1: Istnieją słowa $x, y \neq \varepsilon$ takie, że $a = xy$. W tym przypadku korzystamy z założenia indukcyjnego.

Przypadek 2: $a = \varepsilon$. Jest to przypadek oczywisty.

Przypadek 3: Pozostałe możliwości. Jeżeli nie zachodzi przypadek 1, ani 2, to $a \in X$ i tym bardziej $a \in X^*$.

Tak więc X generuje M.

Przypuśćmy, że Y też generuje M i nie zawiera X. Istnieje więc $a \in X \setminus Y$. Z definicji X otrzymujemy, że $a \neq \varepsilon$ oraz $a \notin (M \setminus \{\varepsilon\})^2$. Oczywiście, $(Y \setminus \{\varepsilon\})^n \subseteq (M \setminus \{\varepsilon\})^2$ dla wszystkich $n \geqslant 2$. Wobec tego, $a \notin (Y \setminus \{\varepsilon\})^n$ dla $n \geqslant 2$. Ponieważ $a \in M = Y^* = (Y \setminus \{\varepsilon\})^*$, więc $a \in Y \setminus \{\varepsilon\}$. $\square$

7 Własność jednoznaczności

Przypuśćmy, że X jest podzbiorem półgrupy M. X ma własność jednoznaczności, jeżeli dla wszystkich $x_1, \ldots, x_n, y_1, \ldots, y_m \in X$, jeżeli

$$x_1 \ldots x_n = y_1 \ldots y_m,$$

to $n = m$ i $x_i = y_i$ dla $i = 1, \ldots, n$.

Twierdzenie 7.1 *Jeżeli X generuje półgrupę $M \subseteq A^*$, to X ma własność jednoznaczności wtedy i tylko wtedy, gdy X jest minimalnym zbiorem generatorów i półgrupa M jest wolna.*

8 Twierdzenie o defekcie

Lemat 8.1 *Jeżeli M jest półgrupą zawartą w A^*, to M jest wolna wtedy i tylko wtedy, gdy dla wszystkich $w \in A^*$, $p, q \in M$ warunek $pw, wq \in M$ implikuje, że $w \in M$.*

Wniosek 8.2 *Przekrój wolnych półgrup zawartych w A^* jest półgrupą wolną.*

Wniosek 8.3 *Dla dowolnego $X \subseteq A^*$ istnieje najmniejsza podpółgrupa zawierająca X.*

Twierdzenie 8.4 (o defekcie) *Załóżmy, że $X \subseteq A^*$ jest skończony i półgrupa X^* nie jest wolna. Wtedy jeżeli Y^* jest najmniejszą półgrupą wolną zawierającą X i Y jest jej minimalnym zbiorem generatorów, to Y ma mniej elementów niż X.*

Dowód. Przypuśćmy, że półgrupa X^* nie jest wolna i Y^* jest najmniejszą półgrupą wolną zawierającą X. Zdefiniujemy funkcję $\alpha : X \to Y$ przyjmując, że $x \in \alpha(x)Y^*$ dla dowolnego $x \in X$ ($\alpha(x)$ jest pierwszym czynnikiem w przedstawieniu x w postaci iloczynu elementów Y). Z jednoznaczności przedstawienia dla półgrup wolnych jest to poprawna definicja funkcji.

Ponieważ X^* nie jest wolna i nie ma własności jednoznaczności, więc α nie jest różnowartościowa.

Aby zakończyć dowód wystarczy pokazać, że α przekształca X na Y. Przyjmijmy, że tak nie jest i $z \in Y \setminus \alpha(X)$. Będziemy zajmować się półgrupą $Z^* \subseteq Y^*$ generowaną przez elementy zbioru $Z = (Y \setminus \{z\})\{z\}^*$. Dla tej pólgrupy zachodzi twierdzenie o jednoznaczności rozkładu, więc jest to półgrupa wolna. Dalej, z nie należy do tej półgrupy gdyż w przeciwnym razie z mogłoby być przedstawione w postaci iloczynu elementów na dwa sposoby. W końcu, $X \subseteq Z^*$. Tak jest, ponieważ przedstawienie $x \in X$ w postaci iloczynu elementów zbioru Y nie zaczyna się z i dlatego daje się przedstawić w postaci iloczynu elementów Z. To wszystko razem przeczy minimalności Y^*. □

Wniosek 8.5 *Każde rozwiązanie nietrywialnego równania z dwoma zmiennymi x i y jest postaci $x = s^n$ i $y = s^m$ dla pewnego słowa s.*