Topologia a logika — języki definiowalne w MSO + U

Michał Skrzypczak

Uniwersytet Warszawski

26 lipca 2010

http://www.mimuw.edu.pl/~mskrzypczak/dokumenty/

Motywacja

- Schemat in $\rightarrow$ out nie jest adekwatny do opisu serwera.
- Naturalny model obliczenia: nieskończony ciąg stanów maszyny.
- Problem: jak definiować własności serwera, by można je było weryfikować?

Przykład

Serwer pocztowy ma trzy stany: prace wewnętrzne, odbieranie i wysyłanie. Chcemy zagwarantować, że zawsze, prędzej czy później serwer będzie w stanie odbieranie.
Formalnie, możemy ten warunek zapisać tak:

$$\forall_{n\in\mathbb{N}}\exists_{k\geq n} S(k) = \mathrm{odbieranie}.$$

Problem

Logika pierwszego rzędu nad $(\mathbb{N}, +, \cdot)$ *jest* ***nierozstrzygalna***.

Definicja

Logika MSO *(monadyczna drugiego rzędu) to logika drugiego rzędu, w której kwantyfikatory drugiego rzędu wiążą tylko zbiory obiektów (a nie np. relacje).*

Przykład

W logice MSO *nad* ω *można wyrazić, że liczba* n *jest parzysta:*

$$\exists_{P \subseteq \omega}\, 0 \in P \ \wedge\ n \in P \ \wedge\ \forall_{p \in P}\, s(p) \notin P \ \wedge\ s(s(p)) \in P$$

Twierdzenie (Büchy '68)

Logika MSO *nad* $(\mathbb{N}, <) = \omega$ *jest rozstrzygalna.*

Automaty

Definicja

Biegiem automatu $\mathcal{A}$ na danym słowie $\alpha \in \Sigma^\omega$ nazywamy dowolny ciąg $(q_0, q_1, \ldots) \in Q^\omega$ zgodny z relacją przejścia w automacie.

Definicja

Warunek Büchiego jest określony przez podzbiór $F \subseteq Q$. Bieg $q_0, q_1, \ldots$ jest akceptujący, jeśli istnieje $f \in F$ taki, że nieskończenie wiele razy $q_i = f$.

Definicja

Warunek Mullera jest określony przez $\mathcal{F} \subseteq 2^Q$. Bieg $q_0, q_1, \ldots$ jest akceptujący, jeśli istnieje $F \in \mathcal{F}$ takie, że F to dokładnie zbiór stanów występujących nieskończenie często.

Rozstrzygalność MSO

Twierdzenie (Safra, MacNaughton)

Niedeterministyczne automaty z warunkiem Büchiego akceptują te same języki co deterministyczne automaty z warunkiem Mullera.

- Konstruujemy automaty dla prostych formuł (np. $x \in X$).
- Automaty Mullera są zamknięte na sumę, przecięcie i dopełnienie.
- Automaty Büchiego są zamknięte na rzutowanie (niedeterminizm).
- Więc też automaty Mullera są zamknięte na niedeterminizm.
- Bierzemy formułę φ. Indukcyjnie ze względu na jej strukturę konstruujemy automat $\mathcal{A}$ taki, że $\alpha \models \varphi \Leftrightarrow \alpha \in L(\mathcal{A})$.
- Spełnialność formuły przekłada się na własność grafową $\mathcal{A}$.

Definicja

Logika $WMSO$ (słaba monadyczna drugiego rzędu) posiada tę samą składnię co MSO, ale kwantyfikatory wiążą tam wyłącznie skończone zbiory.

Uwaga

Dla drzew $WMSO$ ma mniejszą siłę wyrazu, niż MSO (prosty dowód topologiczny).
Dla słów $MSO = WMSO$ bo mamy automaty deterministyczne.

Przypomnienie

$$L(\varphi) = \{\alpha \in \Sigma^\omega : \alpha \models \varphi\}$$

Na przykład dla $\varphi = \forall_n \exists_{k \geq n} Z(k)$ język $L(\varphi)$ to „istnieje nieskończenie wiele liter Z".

Języki ω-regularne

Definicja

Języki ω-regularne, to języki definiowane (równoważnie) przez:

- *logikę MSO,*
- *niedeterministyczne automaty Büchiego,*
- *deterministyczne automaty Mullera,*
- *wyrażenia regularne z operacją L^ω,*
- *logikę $WMSO$,*
- *homomorfizmy w monoidy skończone.*

Własności ilościowe

Przykład

Serwer jest nieskończenie wiele razy w stanie 1 a odcinki czasu pomiędzy kolejnymi wejściami w ten stan są wspólnie ograniczone.

$$L_B = \{0^{n_0}10^{n_1}10^{n_2}1\ldots : \exists_{K\in\mathbb{N}}\forall_{i\in\mathbb{N}} n_i \leq K\} \subset \{0,1\}^\omega$$

Problem

Powyższy język nie jest regularny.

Pomysł: dodać do logiki nowy kwantyfikator.

Logika $\mathrm{WMSO} + \mathrm{U}$

Definicja

Formuła $\mathrm{U}_X \varphi(X)$ jest prawdziwa wtw. istnieją dowolnie duże ***skończone*** *zbiory X spełniające $\varphi(X)$.*

Twierdzenie (Bojańczyk '09)

Słaba logika monadyczna drugiego rzędu z kwantyfikatorem U (ozn. $\mathrm{WMSO} + \mathrm{U}$) jest rozstrzygalna.

Dowód.

Definiujemy deterministyczny model automatu z licznikami, które nie są czytane podczas biegu. Pokazujemy że automaty te są zamknięte na *słabe* rzutowanie (czyli słaby kwantyfikator $\exists$). $\square$

Zalety logiki $\mathrm{WMSO} + \mathrm{U}$:

- jest rozstrzygalna,
- stanowi ścisłe rozszerzenie MSO,
- ma adekwatny model automatowy,
- łapie język L_B i wiele podobnych.

Problemy otwarte

A co z pełną logiką monadyczną z kwantyfikatorem U (ozn. $\mathrm{MSO} + \mathrm{U}$)?

1. Czy jest rozstrzygalna?
2. Czy posiada adekwatny model automatu?
3. Jak *złożone* języki łapie?

Pomysł

Zbadajmy topologiczną złożoność języków definiowalnych w różnych logikach i przez różne modele automatów. Porównamy je i zobaczymy co wyjdzie.

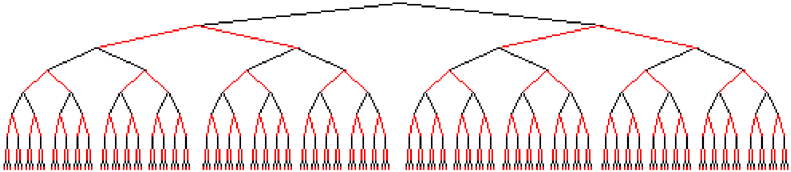
Zbiór Cantora

Definicja

Zbiór Cantora to Σ^ω z metryką

$$d(\alpha, \beta) = 2^{-n}$$

gdzie n to pierwsza pozycja gdzie α, β się różnią.



Hierarchia borelowska

Definicja indukcyjna dla $\eta < \omega_1$.

Definicja

- $\mathbf{\Sigma}^0_1$ — *zbiory otwarte (sumy kulek),*
- $\mathbf{\Pi}^0_1$ — *zbiory domknięte (dopełnienia otwartych),*
- $\mathbf{\Sigma}^0_\eta$ — *przeliczalne sumy zbiorów z* $\bigcup_{\tau<\eta} \mathbf{\Pi}^0_\tau$,
- $\mathbf{\Pi}^0_\eta$ — *dopełnienia zbiorów z* $\mathbf{\Sigma}^0_\eta$,
- $\mathbf{\Delta}^0_\eta = \mathbf{\Sigma}^0_\eta \cap \mathbf{\Pi}^0_\eta$.

$$
\begin{array}{ccccccccc}
\mathbf{\Sigma}^0_1 & & \mathbf{\Sigma}^0_2 & & \mathbf{\Sigma}^0_3 & & \mathbf{\Sigma}^0_4 & & \\
 & \diagdown \quad \diagup & & \diagdown \quad \diagup & & \diagdown \quad \diagup & & \diagdown & \\
 & \mathbf{\Delta}^0_2 & & \mathbf{\Delta}^0_3 & & \mathbf{\Delta}^0_4 & & & \dots \\
 & \diagup \quad \diagdown & & \diagup \quad \diagdown & & \diagup \quad \diagdown & & \diagup & \\
\mathbf{\Pi}^0_1 & & \mathbf{\Pi}^0_2 & & \mathbf{\Pi}^0_3 & & \mathbf{\Pi}^0_4 & &
\end{array}
$$

Zbiory borelowskie

Definicja

Przez $\mathcal{B}$ oznaczamy rodzinę

$$\bigcup_{\eta<\omega_1} \mathbf{\Sigma}^0_\eta \cup \mathbf{\Pi}^0_\eta.$$

Zbiory te nazywamy borelowskimi podzbiorami zbioru Cantora.

Definicja

Każdy zbiór borelowski $B \in \mathcal{B}$ ma dobrze określoną pozycję w hierarchii (swoją złożoność), jako najmniejszą liczbę porządkową η, że $B \in \mathbf{\Sigma}^0_\eta \cup \mathbf{\Pi}^0_\eta$.

Zbiory borelowskie — własności

Fakt

Rodzina $\mathcal{B}$ to σ-ciało: jest zamknięta na dopełnienia i przeliczalne sumy i przecięcia.

Fakt

Hierarchia borelowska jest ścisła — żadna inkluzja na poniższym obrazku nie jest równością.

$$
\begin{tikzcd}
\mathbf{\Sigma}^0_1 \arrow[dr, dash] & & \mathbf{\Sigma}^0_2 \arrow[dr, dash] & & \mathbf{\Sigma}^0_3 \arrow[dr, dash] & & \mathbf{\Sigma}^0_4 \arrow[dr, dash] & \\
& \mathbf{\Delta}^0_2 \arrow[ur, dash] \arrow[dr, dash] & & \mathbf{\Delta}^0_3 \arrow[ur, dash] \arrow[dr, dash] & & \mathbf{\Delta}^0_4 \arrow[ur, dash] \arrow[dr, dash] & & \cdots \\
\mathbf{\Pi}^0_1 \arrow[ur, dash] & & \mathbf{\Pi}^0_2 \arrow[ur, dash] & & \mathbf{\Pi}^0_3 \arrow[ur, dash] & & \mathbf{\Pi}^0_4 \arrow[ur, dash] &
\end{tikzcd}
$$

Hierarchia rzutowa

Definicja

Zbiór $A \subseteq \Sigma^\omega$ *jest analityczny (ozn.* $\mathbf{\Sigma}^1_1$*), wtw. gdy istnieje borelowski zbiór* $B \subseteq \Sigma^\omega \times Y$ *taki, że*

$$A = \{\alpha \in \Sigma^\omega : \exists_{y \in Y} (\alpha, y) \in B\}.$$

- $\mathbf{\Pi}^1_n$ to dopełnienia zbiorów $\mathbf{\Sigma}^1_n$,
- $\mathbf{\Sigma}^1_{n+1}$ to rzuty zbiorów $\mathbf{\Pi}^1_n$,
- $\mathbf{\Delta}^1_n = \mathbf{\Sigma}^1_n \cap \mathbf{\Pi}^1_n$.

$$\begin{tikzcd}
 & \mathbf{\Sigma}^1_1 \arrow[dr, dash] & & \mathbf{\Sigma}^1_2 \arrow[dr, dash] & & \mathbf{\Sigma}^1_3 \arrow[dr, dash] & \\
\mathbf{\Delta}^1_1 \arrow[ur, dash] \arrow[dr, dash] & & \mathbf{\Delta}^1_2 \arrow[ur, dash] \arrow[dr, dash] & & \mathbf{\Delta}^1_3 \arrow[ur, dash] \arrow[dr, dash] & & \cdots \\
 & \mathbf{\Pi}^1_1 \arrow[ur, dash] & & \mathbf{\Pi}^1_2 \arrow[ur, dash] & & \mathbf{\Pi}^1_3 \arrow[ur, dash] &
\end{tikzcd}$$

Własności hierarchii rzutowej

Fakt

Hierarchia rzutowa jest ścisła.

Twierdzenie (Souslin)

Zbiory borelowskie to dokładnie $\mathbf{\Delta}^1_1$.

$$\mathcal{B} = \mathbf{\Delta}^1_1 = \mathbf{\Sigma}^1_1 \cap \mathbf{\Pi}^1_1.$$

Fakt

Każda z klas $\mathbf{\Sigma}^1_n, \mathbf{\Pi}^1_n$ jest zamknięta na przeliczalne sumy i przecięcia. Każda z klas $\mathbf{\Delta}^1_n$ to σ-ciało.

Funkcje ciągłe

Definicja

Funkcja $f\colon \Sigma^\omega \to \Gamma^\omega$ *jest ciągła, jeśli (równoważnie):*

- *zachowuje granice ciągów zbieżnych,*
- *dla dowolnego otwartego* U, *zbiór* $f^{-1}(U)$ *jest otwarty,*
- f *jest (niekoniecznie synchronicznym) przepisywaczem.*

Przepisywacz czyta kolejne litery słowa $\alpha \in \Sigma^\omega$ i raz na jakiś czas (niekoniecznie synchronicznie) decyduje się napisać kolejną literę tworzonego słowa $\beta \in \Gamma^\omega$. Po przeczytaniu całego α musi powstać nieskończone słowo β.

Ciągłe redukcje

Definicja

Funkcja $f\colon \Sigma^\omega \to \Gamma^\omega$ *jest ciągłą redukcją* $X \subseteq \Sigma^\omega$ *do* $Y \subseteq \Gamma^\omega$*, jeśli*

$$X = f^{-1}(Y).$$

Fakt

Jeśli Y *leży w klasie borelowskiej lub analitycznej* $\mathcal{C}$*, a* f *jest ciągłą redukcją* X *do* Y*, to* X *też leży w* $\mathcal{C}$*.*

Fakt

Istnieją zbiory zupełne w sensie ciągłych redukcji we wszystkich klasach borelowskich i analitycznych typu Σ, Π*.*

Fakt

Automat deterministyczny jest synchronicznym przepisywaczem słowa wejściowego $\alpha \in \Sigma^\omega$ w ciąg stanów — element Q^ω.
Czyli funkcja $\alpha \mapsto (q_0, q_1, \ldots)$ jest ciągła.

Fakt

Dla danego automatu deterministycznego Mullera, zbiór jego biegów akceptujących $A \subseteq Q^\omega$ leży w klasie $\mathbf{\Delta}^0_3$.
Analogicznie w przypadku deterministycznych automatów z licznikami.

Wnioski

Wniosek

Złożoność topologiczna języka definiowalnego przez automat deterministyczny jest niewiększa niż złożoność warunku akceptacji.

Wniosek

Zarówno języki ω-regularne, jak też definiowane w $\mathrm{WMSO} + \mathrm{U}$ *leżą w klasie* $\mathbf{\Delta}^0_3 = \mathbf{\Sigma}^0_3 \cap \mathbf{\Pi}^0_3$.

Wniosek

Języki rozpoznawane przez niedeterministyczne automaty z borelowskim warunkiem akceptacji są analityczne:

$$\left\{\alpha \in \Sigma^\omega : \exists_{(q_0, q_1, \ldots) \in Q^\omega} (\alpha, (q_i)) \in A\right\}$$

Wyniki (Hummel, Toruńczyk, S.)

Twierdzenie

W logice $\mathrm{MSO} + \mathrm{U}$ można, dla każdego $i \in \mathbb{N}$, zdefiniować język zupełny w klasie $\mathbf{\Pi}^0_{2i}$.
Ponadto języki te są również definiowane przez alternujące automaty z licznikami.

Twierdzenie

W logice $\mathrm{MSO} + \mathrm{U}$ można zdefiniować język nieborelowski.

Definicja

Dla $i \geq 0$, *niech* $L_i \subseteq (\mathbb{N}^i)^\omega$ *określone następująco*

$$L_i = \{(v_k)_{k\in\mathbb{N}} : \exists^\infty_{n_1} \exists^\infty_{n_2} \ldots \exists^\infty_{n_i} \exists^\infty_{k\in\mathbb{N}} v_k = (n_1, n_2, \ldots, n_i)\}$$

Fakt

Kolejność wektorów w słowie nie ma znaczenia dla L_i.

Przykład

Ciąg $(0,0),(1,1),\ldots$ *nie należy do* L_2.

Przykład

Słowo zawierające wszystkie wektory postaci (p_i, p_i^j) *po* $i, j \in \mathbb{N}$, *każdy nieskończenie wiele razy, należy do* L_2.
Ale jak zamienimy kolejność współrzędnych, to nie należy do L_2.

Definicja

Niech $W_i\colon (\mathbb{N}^i)^\omega \to \{a,b,c,d\}^\omega$ *będzie kodowaniem..*

Fakt

Język $W_i(L_i)$ *jest* $\mathbf{\Pi}^0_{2i+2}$*-zupełny.*

Intuicyjnie każdy kwantyfikator $\exists^\infty$ podbija poziom o 2, bo dotyczą niezależnych współrzędnych.

Cel

Napiszemy formuły $\mathrm{MSO}+\mathrm{U}$ *definiujące dokładnie języki* $W_i(L_i)$*.*

Lemat

Jeśli $X, Y \subseteq \omega$ oraz $\alpha|_{X\cup Y} \in L_i$, to $\alpha|_X \in L_i$, lub $\alpha|_Y \in L_i$.

Lemat

$\alpha \in L_i$ wtw. istnieje nieskończenie wiele m takich, że $\alpha|_{n_i=m} \in L_{i-1}$.

Lemat

Jeśli $X \subseteq \omega$ i $\alpha|_X$ jest ograniczone na i'tej współrzędnej i

$$\pi_{1,2,\ldots,i-1}(\alpha|_X) \in L_{i-1},$$

to dla pewnego m mamy

$$\pi_{1,2,\ldots,i-1}\alpha|_{n_i=m} \in L_{i-1}.$$

$$L_i = \{(v_k)_{k\in\mathbb{N}} : \exists^\infty_{n_1} \exists^\infty_{n_2} \ldots \exists^\infty_{n_i} \exists^\infty_{k\in\mathbb{N}} v_k = (n_1, n_2, \ldots, n_i)\}$$

Definicja

Niech formuła $\psi_i(X)$ mówi, że i'ta współrzędna w wektorach leżących w zbiorze X jest ograniczona.

Definicja

Niech φ_0 mówi, że słowo jest nieskończone. Dla $i > 0$ niech

$$\varphi_i = \forall_X \psi_i(X) \Rightarrow \exists_Y \psi_i(Y) \wedge X \cap Y = \emptyset \wedge \varphi_{i-1}|_Y.$$

Intuicyjnie φ mówi, że jakikolwiek zbiór ograniczony na i'tej współrzędnej weźmiemy, istnieje rozłączny z nim zbiór, też ograniczony, na którym spełniona jest prostsza formuła.

Fakt

Dla każdego $i \geq 0$

$$L(\varphi_i) = W_i(L_i).$$

Wniosek

W logice $\mathrm{MSO} + \mathrm{U}$ *można definiować zbiory położone dowolnie wysoko w hierarchii borelowskiej.*

Wniosek

Jeśli jakiś deterministyczny model automatu chwyta $\mathrm{MSO} + \mathrm{U}$*, to ma skomplikowany warunek akceptacji — przynajmniej na poziomie* ω *w hierarchii borelowskiej.*

Na marginesie

Fakt (Hummel)

Alternujące automaty z licznikami chwytają języki $W_i(L_i)$.

Wniosek

Alternujące automaty z licznikami mają większą siłę wyrazu niż $\mathrm{WMSO}+\mathrm{U}$.

Problem

Nie wiadomo, czy zachodzi którakolwiek z inkluzji pomiędzy językami definiowalnymi w $\mathrm{MSO}+\mathrm{U}$, a rozpoznawanymi przez alternujące automaty z licznikami.

Język nieborelowski

Definicja

Drzewo nad $\mathbb{N}$ to podzbiór $T \subseteq \mathbb{N}^*$ zamknięty na prefiksy. Nieskończoną gałęzią w drzewie T nazywamy $g \in \mathbb{N}^\omega$, takie że

$$\forall_n g|_n \in T.$$

Definicja

Niech $\mathcal{T} \subseteq 2^{\mathbb{N}^*}$ oznacza zbiór wszystkich drzew posiadających nieskończoną gałąź.

Twierdzenie

Zbiór $\mathcal{T}$ jest $\mathbf{\Sigma}^1_1$-zupełny.

Cel

*Zdefiniujemy kodowanie K z przestrzeni $2^{\mathbb{N}^*}$ w $\{a, b, c, d\}^\omega$ i formułę $\mathrm{MSO} + \mathrm{U}$ φ takie, że $T \in \mathcal{T} \Leftrightarrow K(T) \models \varphi$.
Wtedy K stanowić będzie ciągłą redukcję $\mathcal{T}$ do $L(\varphi)$, co pokaże że $L(\varphi)$ jest Σ^1_1-trudne.*

Definicja

Niech $K : 2^{\mathbb{N}^*} \to \{a, b, c, d\}^\omega$ wypisuje wszystkie wierzchołki danego drzewa w pewnym ustalonym porządku, każdy z nich w postaci $c(a^*b)^*a^*c$.

Fakt

K jest funkcją różnowartościową i ciągłą.

Lemat

$T \subseteq \mathbb{N}^*$ ma nieskończoną gałąź wtw. gdy istnieje nieskończony zbiór $G \subseteq T$ taki, że dla każdego $r \in \mathbb{N}$, zbiór

$$V = \{v|_r : v \in G\},$$

jest skończony.

Dowód.

Prosta konsekwencja lematu Königa. □

Lemat

Zbiór $V \subseteq \mathbb{N}^*$ jak wyżej jest skończony wtedy i tylko wtedy, gdy współrzędne występujące w jego elementach są wspólnie ograniczone.

Definicja

Niech $\psi(S)$ mówi, że S zawiera spójne prefiksy bloków postaci $c(a^*b)^*a^*c$, od literki c, do którejś z liter b. Dodatkowo, niech $\psi(S)$ mówi, że liczba liter b w S, we wszystkich blokach jest ograniczona.

Definicja

Niech $\gamma(S)$ mówi, że długości ciągów a^* w S są ograniczone.

Definicja

Niech

$$\varphi = \exists_G \forall_S S \subseteq G \wedge \psi(S) \Rightarrow \gamma(S).$$

Twierdzenie

$T \in \mathcal{T}$ wtedy i tylko wtedy, gdy $K(T) \models \varphi$.

Uwaga

Zbiór $K(\mathcal{T})$ nie jest definiowalny w $\mathrm{MSO}+\mathrm{U}$. Ale $K^{-1}(L(\varphi))=\mathcal{T}$.

Twierdzenie

Nie ma żadnego modelu niedeterministycznego automatu z borelowskim warunkiem akceptacji, chwytającego $\mathrm{MSO}+\mathrm{U}$.

Dowód.

Gdyby jakiś model automatu jak wyżej chwytał $\mathrm{MSO}+\mathrm{U}$, to każdy język definiowalny w $\mathrm{MSO}+\mathrm{U}$ byłby analityczny. Ale logika ta posiada negację, więc byłby też koanalityczny. Więc z twierdzenia Souslina $\mathrm{MSO}+\mathrm{U}$ definiowałoby jedynie borelowskie języki. Pokazany powyżej język $L(\varphi)$ jest Σ^1_1-zupełny, więc nieborelowski. Sprzeczność. $\square$

Co dalej?

- Czy Σ^1_1 to również górne oszacowanie na złożoność języków $\mathrm{MSO}+\mathrm{U}$?
- Czy alternujące automaty z licznikami mogą rozpoznać zbiór $L(\varphi)$?
- Czy jakiś rodzaj *zagnieżdżonego* automatu niedeterministycznego ma szansę uchwycić $\mathrm{MSO}+\mathrm{U}$?

Dziękuję za uwagę, czy są jakieś pytania?