Rozstrzygalne hierarchie topologiczne regularnych języków drzew

Autoreferat rozprawy doktorskiej

Filip Murlak

16 stycznia 2008

1 Mierzenie trudności

Podczas 40 lat badań nad regularnymi językami słów dwie miary trudności okazały się przydatne i inspirujące: hierarchia indeksu, odzwierciedlająca kombinatoryczną złożoność automatu rozpoznającego język oraz hierarchia Wadge'a, która jest bardzo precyzyjnym uszczegółowieniem hierarchii borelowskiej. Zaskakujące związki między tymi hierarchiami zostały w pełni ukazane w pracach Klausa Wagnera o hierarchii zwanej dziś hierarchią Wagnera.

Biorąc hierarchię Wagnera za wzór, chcielibyśmy uzyskać równie pełny i nie mniej elegancki opis dla języków drzew. Uda się nam to jedynie częściowo.

1.1 Automaty w weryfikacji

Automaty na drzewach zaprojektowane przez Michaela O. Rabina jako narzędzie do wykazania rozstrzygalności monadycznej logiki drugiego rzędu są dziś, obok rachunku μ, podstawowym narzędziem teoretycznym w modelowaniu i weryfikacji systemów współbieżnych. W jednym ze stosowanych podejść, drzewo reprezentuje możliwe zachowanie systemu, a automat reprezentuje warunek poprawności. Efektywność znanych metod weryfikacji zależy istotnie od prostoty sposobu, w jaki wyrażony jest warunek poprawności. Tymczasem, przy modelowaniu rzeczywistych systemów, warunki poprawności są często prezentowane w sposób daleki od optymalności. Stąd wypływa potrzeba algorytmów upraszczających reprezentację warunku.

Najciekawszą miarą złożoności warunku poprawności jest głębokość wzajemnego zagnieżdżenia pozytywnych i negatywnych ograniczeń na wydarzenia mające miejsce nieskończenie wiele razy podczas pracy systemu. Formalizacja tego kryterium w teorii automatów prowadzi do pojęcia indeksu. Zwykle, możliwe zachowania systemu są również prezentowane jako automat, a wtedy weryfikacja sprowadza się do problemu zawierania języków, co redukuje się do problemu niepustości. Złożoność problem niepustości zależy przede wszystkim od indeksu automatu: znane algorytmy są wykładnicze ze względu na indeks, ale wielomianowe ze względu na ilość stanów [4, 8, 30].

Stąd bierze się motywacja badania indeksu automatów. Kluczowym zagadnieniem jest tu tzw. problem indeksu, czyli skonstruowanie automatu o minimalnym indeksie rozpoznającego zadany język. Dotychczas odkryto algorytmy rozwiązujące problem indeksu dla języków ścieżkowych [12, 23] i, ogólniej, dla języków rozpoznawanych przez automaty deterministyczne [25, 35]. Podejście poprzez rachunek μ zaowocowało procedurą rozstrzygająca, czy dana formuła modalnego rachunku μ jest równoważna formule logiki modalnej [26].

1.2 Trudność topologiczna

Klasyczna deskryptywna teoria mnogości klasyfikuje podzbiory przestrzeni polskich ze względu na najprostszą definicję za pomocą rzutu, negacji i przeliczalnej sumy bazowych zbiorów otwartych. Wynikająca z tej klasyfikacji hierarchia borelowska i rzutowa może być alternatywną miara złożoności regularnych języków. Przykłady Skurczyńskiego [32] i twierdzenie o luce Niwińskiego i Walukiewicza [24] sugerują, że indeks automatu i ranga borelowska rozpoznawanego języka są ze sobą blisko związane. Porównanie tych hierarchii może być ciekawe same w sobie, ale staje się szczególnie ważne w świetle potencjalnych zastosowań w teorii weryfikacji.

Innym celem przyświecającym badaniom nad topologiczną złożonością języków jest porównywanie różnych modeli obliczeń. Który model jest silniejszy dla języków drzew: automaty deterministyczne, czy słabe alternujące? Wiadomo, że istnieją języki rozpoznawane przez automaty deterministyczne, ale nie przez słabe automaty alternujące i odwrotnie. Jak porównywać, jeśli nie przez inkluzję? Jeszcze bardziej nietypowe porównanie: języki drzew rozpoznawane przez automaty deterministyczne a języki słów rozpoznawane przez deterministyczne automaty ze stosem. Jak porównywać drzewa ze słowami?

Obiecującym narzędziem jest porządek Wadge'a oparty na istnieniu ciągłych redukcji między podzbiorami przestrzeni polskich. Porządek ten definiuje hierarchię, która jest uszczegółowieniem hierarchii borelowskiej i dostarcza bardzo subtelnej miary trudności zbioru. Klasy języków drzew i słów można porównywać poprzez porównanie wysokości hierarchii Wadge'a obciętej do tych klas.

Można wreszcie poszukiwać absolutnej miary złożoności zbioru. Do tego również posłużyć może hierarchia Wadge'a, jako że ustawia języki w porządek bardzo zbliżony do liniowego. W standardowej teorii mnogości, hierarchia Wadge'a ma tę własność jedynie dla zbiorów borelowskich, ale przyjmując dobrze umotywowane dodatkowe aksjomaty determinacji można zapewnić tę własność dla wszystkich regularnych języków słów i drzew. W ten sposób można pytać jak wysoko sięga dany model obliczeń. Wiemy dużo o językach słów [6, 9, 31, 39], ale prawie nic na temat języków drzew. Na których poziomach są języki drzew rozpoznawane przez automaty deterministyczne? A słabe alternujące?

2 Automaty

Przez Σ^* oznaczamy zbiór skończonych słów nad skończonym alfabetem Σ, a przez Σ^ω zbiór nieskończonych słów nad alfabetem Σ. Symbol T_Σ oznacza zbiór *pełnych drzew binarnych nad alfabetem* Σ, tzn. funkcji $t : \{0,1\}^* \mapsto \Sigma$. Jeśli nie stwierdzimy inaczej, "drzewo" oznacza pełne drzewo binarne nad jakimś alfabetem.

Będziemy pracowali głównie z automatami deterministycznymi i słabymi, ale dla zapewnienia ogólnego formalizmu, zdefiniujemy automaty w ich najogólniejszej, alternującej wersji.

Gra parzystości to gra z pełną informacją, rozgrywana przez dwóch graczy, Adama i Ewę. Prezentujemy ją jako krotkę $(V_\exists, V_\forall, E, v_0, \mathrm{rank})$, gdzie $V_\exists$ i $V_\forall$ są (rozłącznymi) zbiorami pozycji Adama i Ewy, $E \subseteq V \times V$ dla $V = V_\exists \cup V$ jest relacją opisująca możliwe ruchy, $v_0 \in V$ jest ustalona pozycja początkową, a $\mathrm{rank} : V \to \{0, 1, \ldots, n\}$ jest funkcją priorytetu.

Gracze rozpoczynają rozgrywkę kładąc żeton w v_0, następnie przesuwają żeton zgodnie z relacją E (zawsze do następnika bieżącej pozycji), konstruując ścieżkę w grafie (V, E). Ruch jest wybierany przez Adama lub Ewę, w zależności od tego, do kogo należy bieżąca pozycja. Jeśli gracz nie może sie ruszyć, przegrywa. W przeciwnym wypadku, rezultatem rozgrywki jest nieskończona ścieżka. Ewa wygrywa rozgrywkę, jeśli największy priorytet pojawiający sie nieskończenie często na tej ścieżce jest parzysty. W przeciwnym wypadku wygrywa Adam.

Automat alternujący na drzewach $A = \langle \Sigma, Q_\exists, Q_\forall, q_0, \delta, \mathrm{rank}\rangle$ składa się ze skończonego alfabetu Σ, skończonego zbioru stanów Q podzielonego na stany egzystencjalne $Q_\exists$ i uniwersalne $Q_\forall$ z ustalonym stanem początkowym q_0, relacji przejścia $\delta \subseteq Q \times \Sigma \times \{0, 1, \varepsilon\} \times Q$ oraz funkcji priorytetu $\mathrm{rank} : Q \to \omega$. Drzewo wejściowe t jest akceptowane przez A, o ile Ewa ma strategię wygrywającą w grze parzystości $\langle Q_\exists \times \{0, 1\}^*, Q_\forall \times \{0, 1\}^*, (q_0, \varepsilon), E, \mathrm{rank}'\rangle$, gdzie $E = \{((p, v), (q, vd)) \colon v \in \mathrm{dom}(t),\ (p, t(v), d, q) \in \delta\}$ i $\mathrm{rank}'(q, v) = \mathrm{rank}(q)$.

Mówimy, że automat jest *deterministyczny*, o ile Ewa nie ma wyboru, a Adam może wybierać jedynie kierunek ruchu: w lewo lub w prawo (bez ε-przejść). Formalnie, żądamy $Q = Q_\forall$ oraz $\delta : Q \times \Sigma \times \{0, 1\} \to Q$. Jeśli zrezygnujemy z warunku $Q_\exists = \emptyset$, ale zachowamy ograniczenie dla Adama, otrzymamy definicję automatów *niedeterministycznych*. Automaty *słabe* to automaty alternujące spełniające warunek $(p, \sigma, d, q) \in \delta \implies \mathrm{rank}\, p \leq \mathrm{rank}\, q$.

Automaty na słowach definiujemy analogicznie, jedynie $\delta \subseteq Q \times \Sigma \times Q$. W przypadku automatów alternujących należy dopuścić ε-przejścia.

3 Indeks i ranga borelowska

Indeksem Mostowskiego–Rabina automatu nazywamy parę

$$(\min \mathrm{rank}\, Q, \max \mathrm{rank}\, Q)\,.$$

Przeskalowując funkcję rank możemy założyć, że $\min \mathrm{rank}\, Q$ jest równe 0 lub 1. Indeksy są zatem elementami $\{0, 1\} \times \omega \setminus \{(1, 0)\}$. Dla indeksu (ι, κ), symbolem $\overline{(\iota, \kappa)}$ oznaczamy *indeks dualny*: $\overline{(0, \kappa)} = (1, \kappa + 1)$, $\overline{(1, \kappa)} = (0, \kappa - 1)$.

Na indeksach definiujemy porządek

$$(\iota, \kappa) < (\iota', \kappa') \text{ o ile } \kappa - \iota < \kappa' - \iota'\,.$$

W szczególności, indeksy wzajemnie dualne są nieporównywalne. *Hierarchia indeksu* dla pewnej ustalonej klasy automatów składa się ze zbiorów języków rozpoznawanych przez automaty o coraz większych indeksach (patrz Rys. 1).

$$
\begin{tikzcd}[column sep=large]
(1,1) \arrow[dr, dash] & (1,2) \arrow[dr, dash] & (1,3) \arrow[dr, dash] & (1,4) & \cdots \\
(0,0) \arrow[ur, dash] & (0,1) \arrow[ur, dash] & (0,2) \arrow[ur, dash] & (0,3) & \cdots
\end{tikzcd}
$$

Rysunek 1: Hierarchia indeksu Mostowskiego–Rabina.

Fundamentalną własnością hierarchii indeksu jest jej ścisłość, tzn. istnienie dla każdego indeksu (ι, κ) języków rozpoznawanych przez automat o indeksie (ι, κ), ale nie przez automat o indeksie $\overline{(\iota, \kappa)}$. Ścisłość hierarchii dla deterministycznych automatów na drzewach wynika wprost ze ścisłości dla deterministycznych automatów na słowach [39]: jeśli język L wymaga co najmniej indeksu (ι, κ), to takiego samego indeksu wymaga język drzew, które na skrajnie lewej gałęzi mają słowo z L. Dla automatów niedeterministycznych i alternujących hierarchie są również ścisłe [3, 22].

Drugim kluczowym zagadnieniem jest *problem indeksu*, czyli kwestia wyznaczania minimalnego indeksu potrzebnego do rozpoznania zadanego języka. Jeśli mamy dany język deterministyczny, możemy pytać o minimalny indeks automatu deterministycznego rozpoznającego ten język. Dla słów eleganckie rozwiązanie podali Niwiński i Walukiewicz [23]. Wprost z tego rezultatu wynika rozstrzygalność tego problemu dla drzew [19].

Dla języka deterministycznego można pytać również o indeks niedeterministyczny, czyli minimalny indeks automatu niedeterministycznego rozpoznającego ten język (może on być niższy, niż minimalny indeks automatu deterministycznego). Problem ten okazał się trudny i został rozwiązany dopiero niedawno w [25]. Rozstrzygalność problemu indeksu dla automatów niedeterministycznych pozostaje jednym z najważniejszych problemów otwartych dotyczących automatów na nieskończonych strukturach.

Tutaj interesuje nas głównie *słaba hierarchia indeksu*, tzn. hierarchia języków rozpoznawanych przez słabe automaty o indeksie (ι, κ). Ścisłość tej hierarchii została wykazana przez Mostowskiego [16] poprzez równoważność z problemem głębokości alternacji kwantyfikatorów w słabej monadycznej logice drugiego rzędu, której ścisłość udowodnił Thomas [33]. Rozstrzygalność *słabego problemu indeksu*, czyli obliczania minimalnego indeksu słabego automatu rozpoznającego zadany język, jest również kwestią otwartą. W [19] podaliśmy metodę obliczania minimalnego indeksu słabego automatu deterministycznego rozpoznającego zadany język deterministyczny. Tutaj rozwiązujemy słaby problem indeksu z wejściem ograniczonym do automatów deterministycznych

Słaba hierarchia indeksu jest ściśle związana z hierarchią borelowską. Pracujemy ze standardową cantorowską topologią na Σ^ω i T_Σ. Obie te przestrzenie są przestrzeniami polskimi, tzn. zupełnymi ośrodkowymi przestrzeniami metrycznymi. Co więcej, obie są homeomorficzne ze zbiorem Cantora.

Klasa *zbiorów borelowskich* w przestrzeni topologicznej X to domknięcie klasy zbiorów otwartych na dopełnienie i przeliczalną sumę. Wewnątrz tej klasy budujemy tzw. hierarchię borelowską. Początkowe skończone poziomy tej hierarchii zdefiniowane są jak następuje:

$\Sigma^0_1(X)$ – otwarte podzbiory X,

$\Pi^0_k(X)$ – dopełnienia zbiorów z $\Sigma^0_k(X)$,

$\Sigma^0_{k+1}(X)$ – przeliczalne sumy zbiorów z $\Pi^0_k(X)$.

Na przykład, $\Pi^0_1(X)$ to zbiory domknięte, $\Sigma^0_2(X)$ to zbiory F_σ, a $\Pi^0_2(X)$ to zbiory G_δ. Przyjmujemy konwencję, że $\Pi^0_0(X) = \{X\}$ i $\Sigma^0_0(X) = \{\emptyset\}$.

Jeszcze ogólniejsze klasy języków tworzą *hierarchię rzutową*. Będziemy potrzebowali wyłącznie jej najniższego poziomu:

$\Sigma^1_1(X)$ – *analityczne* podzbiory X, czyli rzuty podzbiorów borelowskich X^2 z topologią produktową,

$\Pi^1_1(X)$ – dopełnienia zbiorów z $\Sigma^1_1(X)$.

Kiedy tylko przestrzeń X jasno wynika z kontekstu będziemy ją pomijali w zapisie i używali notacji Σ^0_1, Π^0_1, itd.

Niech $\varphi : X \to Y$ będzie ciągłym przekształceniem przestrzeni topologicznych. Mówimy, że φ jest *redukcją* $A \subseteq X$ do $B \subseteq Y$, o ile $\forall_{x\in X}\ x \in A \leftrightarrow \varphi(x) \in B$. Zauważmy, że jeśli B jest w pewnej klasie z opisanych wyżej hierarchii, to A również. Dla dowolnej klasy $\mathcal{C}$ zbiór B jest *$\mathcal{C}$-trudny*, o ile dla każdego $A \in \mathcal{C}$ istnieje redukcja A do B. Hierarchia topologiczna jest ścisła dla przestrzeni polskich, więc jeśli zbiór jest $\mathcal{C}$-trudny, to nie może być w żadnej niższej klasie. Jeśli $\mathcal{C}$-trudny zbiór B jest elementem klasy $\mathcal{C}$, to mówimy, że jest *$\mathcal{C}$-zupełny*.

Skurczyński zaobserwował, że hierarchia borelowska jest nieskończona nawet po obcięciu do języków rozpoznawalnych przez słabe automaty.

Twierdzenie 1 (Skurczyński [32])**.** *Dla każdego $n \geq 1$,*

- *istnieje* Π^0_n*-zupełny język rozpoznawalny przez słaby automat o indeksie* $(0, n)$,
- *istnieje* Σ^0_n*-zupełny język rozpoznawalny przez słaby automat o indeksie* $(1, n + 1)$.

Następne twierdzenie pokazuje, że wynik Skurczyńskiego jest dokładny.

Twierdzenie 2 (Duparc, Murlak [7]; Mostowski [16]). *Dla każdego słabego automatu* A

- *jeśli* A *ma indeks* $(0, n)$*, to* $L(A) \in \Pi^0_n$,
- *jeśli* A *ma indeks* $(1, n + 1)$*, to* $L(A) \in \Sigma^0_n$.

Sądzimy, że przeciwna implikacja jest również prawdziwa: język Π^0_n rozpoznawalny przez słaby automat może być rozpoznany przez słaby automat o indeksie $(0, n)$ (i dualnie dla Σ^0_n). W tym studium pokażemy, że hipoteza ta zachodzi, jeśli ograniczymy się do języków deterministycznych.

W 2002 roku Niwiński i Walukiewicz odkryli zaskakującą dychotomię: każdy deterministyczny język jest albo bardzo prosty, albo bardzo trudny.

Twierdzenie 3 (Niwiński, Walukiewicz [24]). *Dla deterministycznego automatu* A*,* $L(A)$ *jest albo rozpoznawalny przez słaby automat o indeksie* $(0, 3)$ *(a więc jest* Π^0_3*) albo jest nieborelowski (więc również nie jest rozpoznawalny przez słabe automaty). Równoważny automat może być skonstruowany w czasie rozwiązywania problemu niepustości.*

W [19] rozszerzyliśmy ten wynik do rozstrzygalności hierarchii borelowskiej dla automatów deterministycznych.

Twierdzenie 4 (Niwiński, Walukiewicz [24]; Murlak [19]). *Niech* A *będzie automatem deterministycznym. Ranga borelowska* A *może być wyznaczona w czasie znajdowania stanów produktywnych* A.

Wykazaliśmy także, że charakteryzacja wykorzystana w Twierdzeniu 4 działa również w przypadku słabego indeksu i w ten sposób rozwiązaliśmy słaby problem indeksu dla automatów deterministycznych.

Twierdzenie 5. *Dla deterministycznego języka hierarchia borelowska i słaba hierarchia indeksu są równe (Fig. 2) i rozstrzygalne w czasie znajdowania produktywnych stanów automatu.*

$$
\begin{array}{ccccccccccccc}
\Sigma^0_0=(1,1) & & & & \Sigma^0_1=(1,2) & & & & \Sigma^0_2=(1,3) & & & & \\
& \diagdown & & \diagup & & \diagdown & & \diagup & & \diagdown & & & \\
& & \begin{array}{c}\Sigma^0_1\cap\Pi^0_1\\ \|\\ (1,2)\cap(0,1)\end{array} & & & & \begin{array}{c}\Sigma^0_2\cap\Pi^0_2\\ \|\\ (1,3)\cap(0,2)\end{array} & & & & \begin{array}{c}\Sigma^0_3\cap\Pi^0_3\\ \|\\ (1,4)\cap(0,3)\end{array} & & \\
& \diagup & & \diagdown & & \diagup & & \diagdown & & \diagup & & \diagdown & \\
\Pi^0_0=(0,0) & & & & \Pi^0_1=(0,1) & & & & \Pi^0_2=(0,2) & & & & \Pi^0_3=(0,3)
\end{array}
$$

Rysunek 2: Dla automatów deterministycznych hierarchie są równe.

4 Porządek Wadge'a

Pojęcie ciągłej redukcji zdefiniowane w poprzedniej sekcji indukuje prepo-rządek na zbiorach. Dla przestrzeni topologicznych X,Y i zbiorów $A \subseteq X$, $B \subseteq Y$, definiujemy $A \leq_W B$ („A jest redukowalne do B w sensie Wadge'a"), o ile istnieje ciągła redukcja z A do B, tzn. ciągła funkcja $\varphi\colon X \to Y$, taka że $A = \varphi^{-1}(B)$. Mówimy, że A i B są *równoważne w sensie Wadge'a*, $A \equiv_W B$, o ile $A \leq_W B$ i $A \leq_W B$. Definiujemy również ostry preporządek: $A <_W B$, o ile $A \leq_W B$ i $B \not\leq_W A$. *Porządek Wadge'a* to porządek indukowany przez $\leq_W$ na klasach abstrakcji podzbiorów przestrzeni polskich. Porządek Wadge'a ograniczony do zbiorów borelowskich nazywamy *hierarchią Wadge'a*.

Naszym celem jest wykazanie, że porządek Wadge'a jest rozstrzygalny dla deterministycznych języków drzew. Moglibyśmy próbować podejść do problemu bezpośrednio, porównując strukturę danych automatów. Wybraliśmy jednak inny sposób.

Definiujemy cztery sposoby składania automatów: złożenie sekwencyjne $\oplus$, złożenie równoległe $\wedge$, alternatywę $\wedge$ oraz (ι,κ)-replikację $\xrightarrow{(\iota,\kappa)}$. Operacje $\oplus$, $\vee$ i $\xrightarrow{(1,1)}$ wykorzystujemy do zdefiniowania rodziny kanonicznych automatów, które posłużą nam jako wzorzec dla klas abstrakcji deterministycznych języków względem równoważności Wadge'a. Definiujemy automaty C_α dla każdego $0 < \alpha < \omega^{\omega\cdot 3}$ oraz D_α, E_α dla $0 < \alpha < \omega$ lub $\alpha = \omega^{\omega\cdot 2}\alpha_2 + \omega^\omega\alpha_1 + n$ gdzie $\alpha_2 < \omega^\omega$, $0 < \alpha_1 < \omega^\omega$ i $n < \omega$. Dla automatów kanonicznych dowodzimy, że tworzą ścisłą hierarchię względem porządku $\leq_W$ na rozpoznawanych językach.

Twierdzenie 6. *Niech* $0 < \alpha \leq \beta < \omega^{\omega\cdot 3}$. *Jeśli tylko odpowiednie automaty są zdefiniowane, zachodzi* $D_\alpha \not\leq C_\alpha$, $D_\alpha \not\geq C_\alpha$, $D_\alpha < E_\beta$, $C_\alpha < E_\beta$, *a dla* $\alpha < \beta$, $C_\alpha < C_\beta$, $E_\alpha < D_\beta$, $E_\alpha < C_\beta$.

Następnie musimy pokazać, że powyższa hierarchia zawiera wszystkie deterministyczne języki drzew (z dokładnością do równoważności Wadge'a).

$$
\begin{array}{ccccccccccccc}
C_1 & & C_2 & & & C_\omega - C_{\omega+1}\ \cdots & C_{\omega^\omega} & & C_{\omega^\omega+1} & & & C_{\omega^\omega+\omega} - C_{\omega^\omega+\omega+1}\ \cdots & \\
& \diagdown\ \diagup & & \diagdown & & & & \diagdown\ \diagup & & \diagdown & & & \\
& E_1 & & E_2\ \cdots & & & & E_{\omega^\omega} & & E_{\omega^\omega+1}\ \cdots & & & \cdots \\
& \diagup\ \diagdown & & \diagup & & & & \diagup\ \diagdown & & \diagup & & & \\
D_1 & & D_2 & & & & D_{\omega^\omega} & & D_{\omega^\omega+1} & & & &
\end{array}
$$

Rysunek 3: Porządek Wadge' dla automatów kanonicznych.

Okazuje się, że musimy rozszerzyć rodzinę automatów kanonicznych o trzy dodatkowe elementy $C_{\omega^{\omega\cdot 3}}$, $C_{\omega^{\omega\cdot 3}+1}$, $C_{\omega^{\omega\cdot 3}+2}$, stanowiące szczyt hierarchii. Rozszerzona rodzina automatów kanonicznych jest zamknięta na wszystkie operacje składania. Potem wykazujemy, że każdy automat deterministyczny (z dokładnością do równoważności Wadge'a) może być otrzymany poprzez składanie automatów C_1 i D_1 za pomocą tych operacji. W ten sposób otrzymujemy następujące twierdzenie.

Twierdzenie 7. *Dla każdego automatu deterministycznego na drzewach A istnieje automat kanoniczny rozpoznający język równoważny w sensie Wadge'a. Automat ten może być wyznaczony w czasie obliczania stanów akceptujących A.*

To prowadzi nas wprost do celu: dla zadanych automatów A i B, rozstrzygamy, czy $L(A) \leq_W L(B)$, poprzez porównanie równoważnych automatów kanonicznych.

Z Twierdzenia 6 i 7 wynika również, że hierarchia Wadge'a dla deterministycznych języków drzew ma wysokość $\omega^{\omega\cdot 3} + 3$. To dużo więcej niż ω^ω dla regularnych języków słów [39], ale dużo mniej niż $(\omega^\omega)^\omega$ dla deterministycznych bezkontekstowych języków słów [6], $(\omega_1^{CK})^\omega$ dla języków słów rozpoznawanych przez deterministyczne maszyny Turinga [31], czy – nieznana dotychczas dokładnie – olbrzymia liczba porządkowa $\xi > \varepsilon_0$ dla niedeterministycznych bezkontekstowych języków słów [9].

5 Stopnie Wadge'a

Słynne twierdzenie Martina o determinacji [15] pozwala bardzo precyzyjnie opisać kształt hierarchii Wadge'a.

Twierdzenie 8 (Lemat Wadge'a)**.** *Dla borelowskich języków L, M zachodzi*

$$L \leq_W M \quad lub \quad L^{\complement} \geq_W M\,.$$

Innymi słowy, szerokość hierarchii Wadge'a jest nie większa niż 2 i jeśli dwa zbiory L i M są nieporównywalne, to $L \equiv_W M^{\complement}$. Oznacza to, że porządek Wadge'a jest prawie liniowy. Jeśli $L \equiv_W L^{\complement}$, to mówimy, że L jest *samodualny.*

Drugą kluczową własność hierarchii Wadge'a zapewnia następujące twierdzenie.

Twierdzenie 9 (Wadge, Martin, Monk)**.** *Hierarchia Wadge'a jest dobrze ufundowana.*

Podsumowując, pozycja języka w hierarchii Wadge'a jest (z dokładnością do dopełnienia) zadana przez jej wysokość. Stąd następująca definicja. *Stopniem Wadge'a* języka, który nie jest samodualny nazywamy liczbę porządkową zdefiniowaną indukcyjnie:

- $d_W(\emptyset) = d_W(\emptyset^{\complement}) = 1$,
- $d_W(L) = \sup\{d_W(M) + 1 : M \text{ nie jest samodualny}, M <_W L\}$ dla $L >_W \emptyset$.

Dla zbiorów samodualnych definiujemy

$$d_W(L) = \sup\{d_W(M) : M \text{ nie jest samodualny}, M <_W L\} .$$

Powyższa definicja jest prawie identyczna jak definicja wprowadzona przez J.Duparca. Oryginalna definicja Wadge'a liczyła również zbiory samodualne (patrz [5]).

Jednym z najbardziej zaskakujących odkryć na temat hierarchii Wadge'a jest fakt, że arytmetyczna struktura na stopniach Wadge'a jest odzwierciedlona w prostych operacjach teoriozbiorowych, które na dodatek mają naturalną interpretację w języku gier Wadge'a. Już w swojej pracy doktorskiej [37] Wadge zdefiniował operacje spełniające

$$\begin{aligned} d_W(L_1 + L_2) &= d_W(L_1) + d_W(L_2) , \\ d_W(\sup_n L_n) &= \sup_n d_W(L_n) , \\ d_W(L_1 \cdot \omega_1) &= d_W(L_1) \cdot \omega_1 , \end{aligned}$$

dla dowolnych języków $L_1, L_2, \ldots \subseteq \Sigma^\omega$. My adaptujemy te operacje do przypadku drzew w taki sposób, żeby powyższe równości nadal zachodziły i używamy ich do policzenia stopni Wadge'a języków deterministycznych.

Dla każdej liczby porządkowej $\alpha < \omega^{\omega\cdot 3}$ istnieje jedyne przedstawienie

$$\alpha = \omega^{\omega\cdot 2}(\omega^k l_k + \ldots + \omega^0 l_0) + \omega^\omega(\omega^k m_k + \ldots + \omega^0 m_0) + \omega^k n_k + \ldots + \omega^0 n_0 \,.$$

Zdefiniujmy $j_{\mathrm{Det}}(\alpha)$ wzorem

$$j_{\mathrm{Det}}(\alpha) = \omega_1^\omega(\omega^k l_k + \ldots + \omega^0 l_0) + \omega_1(\omega_1^k m_k + \ldots + \omega_1^0 m_0) + \omega^k n_k + \ldots + \omega^0 n_0$$

dla $\alpha < \omega^{\omega\cdot 3}$ oraz $j_{\mathrm{Det}}(\omega^{\omega\cdot 3}) = \omega_1^{\omega+1}$, $j_{\mathrm{Det}}(\omega^{\omega\cdot 3}+1) = \omega_1^{\omega_1}$.

Twierdzenie 10. *Dla każdej liczby porządkowej $\alpha \leq \omega^{\omega\cdot 3}+1$ (o ile automaty są zdefiniowane) zachodzi*

$$d_W(L(C_\alpha)) = d_W(L(D_\alpha)) = d_W(L(E_\alpha)) = j_{\mathrm{Det}}(\alpha)\,.$$

Funkcja j_{Det} jest czasami nazywana funkcją Wagnera dla klasy języków – w tym przypadku, dla języków deterministycznych. Pierwsze rezultaty tego typu były związane właśnie z hierarchią Wagnera. Funkcją Wagnera dla regularnych języków słów jest

$$j_{\mathrm{Reg}}(\omega^k n_k + \ldots + \omega^1 n_1 + n_0) = \omega_1^k n_k + \ldots + \omega_1^1 n_1 + n_0\,.$$

Dla deterministycznych bezkontekstowych języków słów funkcję Wagnera podaje J. Duparc [6]:

$$j_{\mathrm{DetCFree}}((\omega^\omega)^k \alpha_k + \ldots + (\omega^\omega)^1 \alpha_1 + \alpha_0) = (\omega_1)^k \alpha_k + \ldots + (\omega_1)^1 \alpha_1 + \alpha_0\,,$$

gdzie $\alpha_i < \omega^\omega$.

Poza wspomnianymi operacjami wprowadzamy również analogi mnożenia przez przeliczalne liczby porządkowe oraz funkcji wykładniczej $\alpha \mapsto \omega_1^{\alpha+\varepsilon}$, gdzie

$$\varepsilon = \begin{cases} -1 & \text{dla } d_C(L) < \omega \\ 0 & \text{dla } d_C(L) = \beta + n \text{ oraz } \mathrm{cof}\beta = \omega_1 \\ +1 & \text{dla } d_C(L) = \beta + n \text{ oraz } \mathrm{cof}\beta = \omega \end{cases}\,.$$

Ta ostatnia operacja została zdefiniowana dla słów w [5]. Okazuje się, że klasa języków rozpoznawalnych przez słabe automaty jest zamknięta na operacje odpowiadające dodawaniu, mnożeniu przez ω oraz funkcji $\alpha \mapsto \omega_1^{\alpha+\varepsilon}$. Stąd wynika dolne ograniczenie na wysokość hierarchii Wadge dla tej klasy języków. Sądzimy, że to ograniczenie jest dokładne.

Twierdzenie 11. *Hierarchia Wadge'a ograniczona do języków drzew rozpoznawalnych przez słabe automaty ma wysokość co najmniej ε_0, gdzie ε_0 jest najmniejszym punktem stałym funkcji $\alpha \mapsto \omega^\alpha$.*

6 Problemy otwarte

Podobnie jak w klasycznej teorii złożoności, prawdziwym wyzwaniem w teorii automatów jest niedeterminizm. Siła jaką daje automatom na drzewach sprawia, że stają się one niezwykle trudne w analizie. Rozstrzygalność hierarchii borelowskiej i hierarchii Wadge'a pozostaje ciągle problemem otwartym, podobnie jak rozstrzygalność hierarchii indeksu.

W tym studium ograniczamy się do bardzo ubogiej wersji niedeterminizmu – obecnej w słabych automatach alternujących. Automaty te rozpoznają dokładnie takie języki L, że L i $L^{\complement}$ są rozpoznawane przez automat niedeterministyczny o indeksie $(1, 2)$ [17, 28], a więc faktycznie stanowią niewielką podklasę języków regularnych. Z drugiej strony, mechanizm ten niewątpliwie ujmuje istotnie wiele niedeterminizmu, jako że zawiera liczne języki, które nie mogą być rozpoznane przez automaty deterministyczne: przykłady Skurczyńskiego pokazują, że języki rozpoznawalne przez słabe automaty mogą mieć dowolnie dużą skończoną rangę borelowską [32], podczas gdy języki deterministyczne są albo Π^1_1-zupełne, albo są w Π^0_3 [24].

Jednak nawet w tym prostym przypadku trzy najważniejsze hierarchie – indeksu, borelowska i Wadge'a – nie są dostatecznie zrozumiane. Wśród najistotniejszych problemów są dwie hipotezy sformułowane w tej pracy, dotyczące wysokości hierarchii Wadge'a oraz równości hierarchii borelowskiej i słabej hierarchii indeksu.

Literatura

[1] A. Arnold. The μ-calculus alternation-depth hierarchy is strict on binary trees. RAIRO-Theoretical Informatics and Applications **33** (1999) 329–339.

[2] A. Arnold, D. Niwiński. Continuous separation of game languages. Manuscript, submitted, 2006.

[3] J. C. Bradfield. The modal mu-calculus alternation hierarchy is strict. Theoret. Comput. Sci. **195** (1998) 133–153.

[4] A. Browne, E. M. Clarke, S. Jha, D. E. Long, W. Marrero. An improved algorithm for the evaluation of fixpoint expressions. Theoret. Comput. Sci. **178** (1997) 237–255.

[5] J. Duparc. Wadge hierarchy and Veblen hierarchy. Part I: Borel sets of finite rank. The Journal of Symbolic Logic **66** (2001).

[6] J. Duparc. A hierarchy of deterministic context-free ω-languages. Theoret. Comput. Sci. **290** (2003) 1253–1300.

[7] J. Duparc, F. Murlak. On the topological complexity of weakly recognizable tree languages. Proc. FCT 2007, LNCS 4639 (2007) 261-273.

[8] E. A. Emerson, C. S. Jutla The complexity of tree automata and logics of programs. *Proc. FoCS '88*, IEEE Computer Society Press 1988, 328–337.

[9] O. Fi nkel. Wadge Hierarchy of Omega Context Free Languages. Theoret. Comput. Sci. **269** (2001) 283–315.

[10] M. Jurdziński, J. Vöge. A discrete strategy improvement algorithm for solving parity games. *Proc. CAV '00*, LNCS 1855 (2000) 202–215.

[11] A. S. Kechris. *Classical Descriptive Set Theory.* Graduate Texts in Mathematics, Vol. 156, Springer-Verlag 1995.

[12] O. Kupferman, S. Safra, M. Vardi. Relating Word and Tree Automata. *Proc. LICS '96*, 322–332.

[13] L. H. Landweber. Decision problems for ω-automata. Math. Systems Theory **3** (1969) 376–384.

[14] G. Lenzi. A hierarchy theorem for the mu-calculus. *Proc. ICALP '96*, LNCS 1099 (1996) 87–109.

[15] D. A. Martin. Borel determinacy. Ann. Math. **102** (1975) 363–371.

[16] A. W. Mostowski. Hierarchies of weak automata and weak monadic formulas. Theoret. Comput. Sci. **83** (1991) 323-335.

[17] D. E. Muller, A. Saoudi, P. E. Schupp. Alternating automata. The weak monadic theory of the tree, and its complexity. *Proc. ICALP '86*, LNCS 226 (1986) 275–283.

[18] D. E. Muller, P. E. Schupp. Alternating automata on infinite objects, determinacy and Rabin's theorem. *Proc. Automata on Infinite Words 1984*, LNCS 192 (1985) 100–107.

[19] F. Murlak. On deciding topological classes of deterministic tree languages. *Proc. CSL '05*, LNCS 3634 (2005) 428–441.

[20] F. Murlak. The Wadge hierarchy of deterministic tree languages. *Proc. ICALP '06*, Part II, LNCS 4052 (2006) 408-419.

[21] D. Niwiński. An example of non-Borel set of infinite trees recognizable by a Rabin automaton. Manuscript, Warsaw Univeristy, 1985 (in polish).

[22] D. Niwiński. On fixed point clones. *Proc. ICALP '86*, LNCS 226 (1986) 464–473.

[23] D. Niwiński, I. Walukiewicz. Relating hierarchies of word and tree automata. *Proc. STACS '98*, LNCS 1373 (1998) 320–331.

[24] D. Niwiński, I. Walukiewicz. A gap property of deterministic tree languages. Theoret. Comput. Sci. **303** (2003) 215–231.

[25] D. Niwiński, I. Walukiewicz. Deciding nondeterministic hierarchy of deterministic tree automata. *Proc. WoLLiC '04*, Electronic Notes in Theoret. Comp. Sci. 2005, 195–208.

[26] M. Otto. Eliminating recursion in μ-calculus. *Proc. STACS'99*, LNCS 1563 (1999) 531–540.

[27] D. Perrin, J.-E. Pin. *Infinite Words. Automata, Semigroups, Logic and Games.* Pure and Applied Mathematics Vol. 141, Elsevier 2004.

[28] M. O. Rabin. Decidability of second-order theories and automata on infinite trees. Trans. Amer. Soc. **141** (1969) 1–35.

[29] M. O. Rabin. Weakly definable relations and special automata. *Mathematical Logic and Foundations of Set Theory*, North-Holland 1970, 1–70.

[30] H. Seidl. Fast and simple nested fixpoints. Information Processing Letters **59** (1996) 303–308.

[31] V. Selivanov. Wadge Degrees of ω-languages of deterministic Turing machines. Theoret. Informatics Appl. **37** (2003) 67-83.

[32] J. Skurczyński. The Borel hierarchy is infinite in the class of regular sets of trees. Theoret. Comput. Sci. **112** (1993) 413–418.

[33] W. Thomas. A hierarchy of sets of infinite trees. *Proc. Theoretical Computer Science*, LNCS 145 (1982) 335–342.

[34] W. Thomas. Languages, automata, and logic. Handbook of Formal Languages, Vol. 3, Springer-Verlag 1997, 389–455.

[35] T. F. Urbański. On deciding if deterministic Rabin language is in Büchi class. *Proc. ICALP '00*, LNCS 1853 (2000) 663–674.

[36] W. W. Wadge. Degrees of complexity of subsets of the Baire space. Notice Amer. Math. Soc. (1972) A-714.

[37] W. W. Wadge. Reducibility and determinateness on the Baire space. Ph. D. Thesis, Berkeley, 1984.

[38] K. Wagner. Eine topologische Charakterisierung einiger Klassen regulärer Folgenmengen. J. Inf. Process. Cybern. EIK **13** (1977) 473–487.

[39] K. Wagner. On ω-regular sets. Inform. and Control **43** (1979) 123–177.