Szymon Toruńczyk

1 Automaty niedeterministyczne

Automat niedeterministyczny $\mathcal{A}$ jest wyznaczony przez następujące składniki:

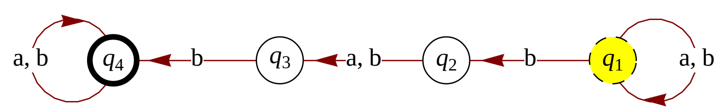
- Alfabet skończony A
- Zbiór stanów Q
- Zbiór stanów początkowych Q_I
- Zbiór stanów końcowych Q_F.
- Relacja przejścia
$$\delta \subset Q \times A \times \ Q.$$

A zatem, relacja przejścia jest pewnym zbiorem trójek (p, σ, q), które będziemy zapisywać tak: $p \xrightarrow{\sigma} q$. Jeżeli dla każdego stanu p i dla każdej litery σ istnieje najwyżej jeden stan q taki, że $p \xrightarrow{\sigma} q$, to automat $\mathcal{A}$ jest deterministyczny.

Mówimy, że słowo $w = w_1 w_2 \dots w_k$ jest akceptowane przez automat $\mathcal{A}$ jeżeli *istnieje* ciąg stanów $q_0, \dots q_k$, spełniające następujące warunki:

- $q_0 \in Q_I$
- $q_k \in Q_F$.
- $q_0 \xrightarrow{w_1} q_1 \xrightarrow{w_2} q_2 \cdots \xrightarrow{w_k} q_k$

Przykład 1 Automat, rozpoznający język $\{a+b\}^* \cdot b \cdot \{a+b\} \cdot b \cdot \{a+b\}^*$ przedstawiamy graficznie tak:



Możemy rozszerzyć pojęcie relacji przejścia w następujący sposób:

> Dla słowa $w = w_1 w_2 \dots w_k$ napiszemy $p \xrightarrow{w} q$, jeżeli istnieje ciąg $q_0, q_1, \dots q_k$ taki, że

- $p = q_0$,
- $q = q_k$,
- $q_0 \xrightarrow{w_1} q_1 \xrightarrow{w_2} q_2 \cdots \xrightarrow{w_k} q_k$

Naturalne jest przyjęcie konwencji, że $p \xrightarrow{\varepsilon} q$ wtedy i tylko wtedy, gdy $p = q$.

Możemy teraz przeformułować definicję akceptacji:

> Słowo w jest akceptowane przez automat $\mathcal{A}$ wtedy i tylko wtedy, gdy $p \xrightarrow{w} q$ dla pewnego $p \in Q_I$ oraz $q \in Q_F$.

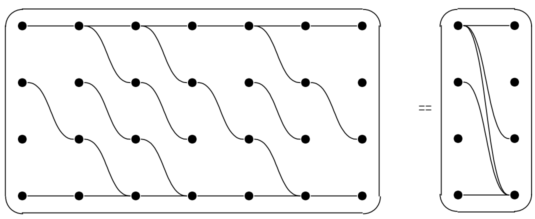
Według powyższych oznaczeń, każde słowo w wyznacza nam relację binarną $\xrightarrow{w}$ na zbiorze Q, i z samej definicji wynika, że relacja wyznaczona przez konkatenację słów v oraz w jest złożeniem odpowiednich relacji

$$\xrightarrow{v \cdot w} = \xrightarrow{v} \diamond \xrightarrow{w},$$

gdzie złożenie relacji $\xrightarrow{1}$ oraz $\xrightarrow{2}$ jest relacją $\xrightarrow{1} \diamond \xrightarrow{2}$ zdefiniowaną tak:

$$p \xrightarrow{1} \diamond \xrightarrow{2} r \iff \text{istnieje } q \text{ takie, że } p \xrightarrow{1} q \text{ oraz } q \xrightarrow{2} r.$$

Przykład 2 Poniżej przedstawiony jest ciąg relacji wyznaczonych przez kolejne litery słowa $w = abbaba$ w automacie z poprzedniego przykładu oraz ich złożenie. W tym ciągu relacji istnieje ścieżka od stanu początkowego 1 po lewej stronie do stanu akceptującego 4 po prawej stronie. A zatem, słowo w jest akceptowane przez automat. Równoważnie, wynikowa relacja zawiera krawędź od stanu początkowego 1 do stanu akceptującego 4, więc słowo jest akceptowane.



$\xrightarrow{a} \diamond \xrightarrow{b} \diamond \xrightarrow{b} \diamond \xrightarrow{a} \diamond \xrightarrow{b} \diamond \xrightarrow{a} \quad = \quad \xrightarrow{abbaba}$

Tak więc, czytając słowo w, możemy obliczać „on-line” relację $\xrightarrow{w}$ w następujący sposób:

$$\xrightarrow{w_{[i]}} = \xrightarrow{w_{[i-1]}} \diamond \xrightarrow{w_i}$$

gdzie $w_{[i]} = w_1 w_2 \dots w_i$.

Robi to automat *deterministyczny* $\mathcal{B}$ opisany poniżej.

- Alfabetem jest alfabet automatu $\mathcal{A}$
- Zbiorem stanów jest zbiór R_Q wszystkich relacji binarnych na zbiorze Q. Jest ich dokładnie $2^{|Q| \times |Q|}$.
- Stanem początkowym jest $\xrightarrow{\varepsilon}$ (tzn. relacja identycznościowa).
- Stany akceptujące to takie relacje $\rightsquigarrow$, że $p \rightsquigarrow q$ dla pewnych $p \in Q_I$, $q \in Q_F$.
- Relacja przejścia określona jest następująco. Automat $\mathcal{B}$ będąc w stanie $\rightsquigarrow$ i czytając literę σ przechodzi do stanu $\rightsquigarrow \ \diamond \xrightarrow{\sigma}$.

W ten sposób, po wczytaniu słowa w, automat $\mathcal{B}$ oblicza relację $\xrightarrow{w}$, czyli relację przejścia wyznaczoną przez to słowo w automacie $\mathcal{A}$. Jeżeli ta relacja umożliwia przejście ze stanu z Q_I do stanu z Q_F, to automat $\mathcal{B}$ akceptuje. Jasne jest więc, że automat $\mathcal{B}$ akceptuje dokładnie te słowa, co automat $\mathcal{A}$. Także widać, że $\mathcal{B}$ faktycznie jest deterministyczny. Automaty podobne do $\mathcal{B}$ określimy mianem *automatów półgrupowych*. W tym celu wprowadzimy parę definicji.

Definicja 1.1 Półgrupą nazywamy zbiór $(S, \ \cdot \ )$ wyposażony w operację mnożenia $\cdot$ spełniającą warunek

$$(s \cdot t) \cdot u = s \cdot (t \cdot u). \qquad (*)$$

Dodatkowo będziemy wymagać, by półgrupa S zawierała element neutralny, tzn. taki element e, że

$$s \cdot e = e \cdot s = s$$

dla każdego $s \in S$ (zazwyczaj takie półgrupy nazywa się *monoidami* lub *półgrupą z jednością*).

Warunek $(*)$ nazywa się *łącznością*.

Przykład 3 Półgrupa $(\mathbb{Z}_6, \times)$ składająca się z liczb $0, 1, 2, 3, 4, 5$ wraz z operacją mnożenia modulo 6.

Przykład 4 Półgrupa funkcji na zbiorze n-elementowym, wraz z operacją składania funkcji ($\circ$).

Przykład 5 Półgrupa macierzy $k \times k$ o elementach z pierścienia $\mathbb{Z}_n$, z operacją mnożenia macierzy.

Przykład 6 Dowolna grupa.

Przykład 7 Półgrupa relacji binarnych na zbiorze n-elementowym, wraz z operacją składania relacji ($\diamond$).

Definicja 1.2 *Automat półgrupowy* $\mathcal{S}$ jest opisany przez następujące składniki:

- Alfabet Σ
- Półgrupa $(S, \cdot)$
- Funkcja przejścia $\delta \colon \Sigma \to S$
- Zbiór akceptujący $T \subset S$.

Funkcja przejścia określa, jaki element półgrupy S wyznacza każda z liter $\sigma \in \Sigma$.

Powiemy, że automat $\mathcal{S}$ *akceptuje* słowo $w = w_1 w_2 \ldots w_k$ jeżeli

$$\delta(w_1) \cdot \delta(w_2) \cdots \delta(w_{k-1}) \cdot \delta(w_k) \in T.$$

Zauważmy, że bez wpływu na język rozpoznawany przez automat $\mathcal{S}$, możemy założyć, że półgrupa S jest generowana przez zbiór elementów $\{\delta(\sigma) : \sigma \in \Sigma\}$, tzn. że każdy element S jest ilocznynem pewnego ciągu takich elementów.

Każdy automat półgrupowy $\mathcal{S}$ wyznacza w naturalny sposób automat deterministyczny, rozpoznający ten sam język. Jego stany to elementy półgrupy S, stanem początkowym jest jej element neutralny, stany akceptujące to elementy zbioru T i wreszcie, automat ten, będąc w stanie s i wczytując literę σ przechodzi do stanu $s \cdot \delta(\sigma)$.

Z kolei dowolny automat niedeterministyczny $\mathcal{A}$ wyznacza automat półgrupowy określony następująco:

- Alfabet to alfabet automatu $\mathcal{A}$
- Półgrupa automatu to półgrupa relacji R_Q na zbiorze Q stanów automatu $\mathcal{A}$,

- Funkcja przejścia zadana jest wzorem $\delta(\sigma) = \xrightarrow{\sigma}$, gdzie $\xrightarrow{\sigma}$ jest relacją przejścia wyznaczoną przez literę σ w automacie $\mathcal{A}$,
- Zbiór akceptujący to $\{\rightsquigarrow \in R_Q : \exists p \in Q_I, q \in Q_F : p \rightsquigarrow q\}$.

Widzimy więc, że wszystkie trzy klasy automatów – półgrupowe, deterministyczne i niedeterministyczne – mają tę samą siłę wyrazu (i rozpoznają języki regularne).

2 Automaty dwukierunkowe

Deterministyczny automat dwukierunkowy $\mathcal{D}$ działa w następujący sposób. W kolejnych krokach, $\mathcal{D}$ wczytuje literę z taśmy wejściowej i, w zależności od tej litery, kierunku z którego przyszedł oraz stanu w jakim się znajduje, decyduje się, zgodnie z ustaloną tablicą reguł, przejść w prawo lub w lewo oraz zmienić stan na inny (w bardziej powszechnym wariancie zakłada się, że decyzja nie zależy od kierunku przyjścia). Przyjmujemy, że przed pierwszą i za ostatnią literą słowa na taśmie wejściowej stoją specjalne markery końca słowa. Automat zaczyna przed markerem początku słowa, w stanie początkowym. Gdy przekroczy marker końca słowa, kończy swój bieg. W zależności od stanu, w którym się znajdzie, słowo będzie zaakceptowane lub nie.

A zatem, formalnie, dwukierunkowy automat deterministyczny ma następujące składniki:

- Alfabet skończony A oraz dwa markery specjalne $begin, end \notin A$
- Zbiór stanów Q
- Stan początkowy $q_I \in Q$
- Zbiór stanów końcowych Q_F
- Funkcja przejścia

$$\delta\colon (A \cup \{begin, end\}) \times \{d_{\leftarrow}, d_{\rightarrow}\} \times Q \quad \rightarrow \quad \{d_{\leftarrow}, d_{\rightarrow}\} \times Q$$

  ($d_{\leftarrow}$ oraz $d_{\rightarrow}$ oznaczają tutaj kierunki ruchu automatu)

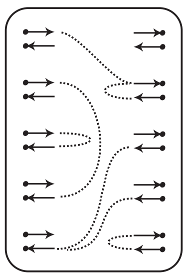
Automat niedeterministyczny tym się różni od deterministycznego, że reguły nie są wyznaczone jednoznacznie w zależności od wczytanej litery, kierunku ruchu i stanu automatu. Automat może mieć kilka możliwości i w każdym ruchu decyduje się na jedną z nich. Formalnie, znaczy to tyle, że δ nie jest już funkcją, ale relacją. A zatem, dla danego słowa, automat niedeterministyczny może mieć wiele *biegów* (czasem nawet nieskończenie wiele, przy zapętleniach), a bieg automatu deterministycznego jest wyznaczony jednoznacznie przez słowo wejściowe. Gdy któryś z możliwych biegów automatu jest zakończony w stanie początkowym, powiemy że automat *akceptuje* dane słowo wejściowe. Zbiór słów akceptowanych przez automat jest *językiem automatu.*

Pokażemy, że deterministyczny automat dwukierunkowy wyznacza w naturalny sposób pewien automat półgrupowy. W tym celu, zdefiniujemy półgrupę T_n „modelującą” działanie automatów dwukierunkowych o n stanach.

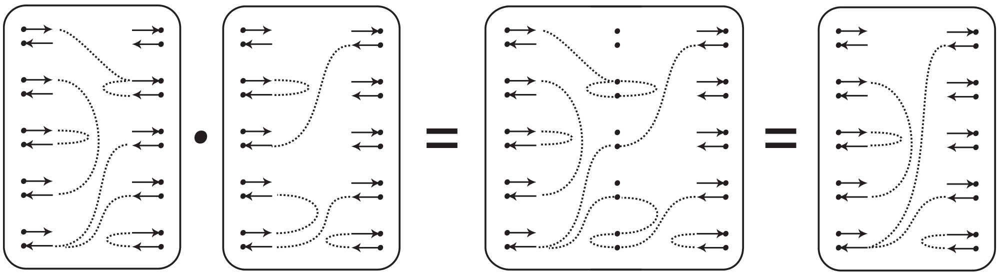
Zdefiniujemy półgrupę T_n *transformacji dwukierunkowych* zbioru $\{1, \dots, n\}$. Elementami tej półgrupy są grafy nieskierowane o $4n$-elementowym zbiorze wierzchołków o następujących etykietach:

- $1_{\bullet\leftarrow}, 2_{\bullet\leftarrow}, \dots, n_{\bullet\leftarrow}$– wierzchołki wyjścia z lewej strony
- $1_{\bullet\rightarrow}, 2_{\bullet\rightarrow}, \dots, n_{\bullet\rightarrow}$– wierzchołki wejścia z lewej strony
- $1_{\rightarrow\bullet}, 2_{\rightarrow\bullet}, \dots, n_{\rightarrow\bullet}$– wierzchołki wyjścia z prawej strony
- $1_{\leftarrow\bullet}, 2_{\leftarrow\bullet}, \dots, n_{\leftarrow\bullet}$– wierzchołki wejścia z prawej strony

Krawędzie mogą łączyć jedynie wierzchołki wejścia z wierzchołkami wyjścia. Przykładowo, lewy wierzchołek wejścia może być połączony z prawym wierzchołkiem wyjścia albo z lewym wierzchołkiem wyjścia (odpowiada to dwukierunkowości automatu). Każdy wierzchołek wejścia może być połączony z najwyżej jednym wierzchołkiem (warunek determinizmu). Oto graficzna reprezentacja przykładowego elementu półgrupy T_5:



Poniższy rysunek, przedstawia, jak będziemy mnożyć dwa elementów półgrupy T_5:



Formalna definicja mnożenia przedstawiona jest poniżej. Aby otrzymać złożenie $s \cdot t$ dwóch takich grafów s, t Sklejamy ze sobą, usuwając ich etykiety, pary wierzchołków o odpowiadających sobie numerach ($k = 1, \ldots n$):

- wierzchołek wyjścia z prawej strony $k_{\to\bullet}$ grafu s z wierzchołkiem wejścia z lewej strony $k_{\bullet\to}$ grafu t
- wierzchołek wyjścia z lewej strony $k_{\bullet\leftarrow}$ grafu t z wierzchołkami wejścia z prawej strony $k_{\leftarrow\bullet}$ grafu s.

Etykiety pozostałych wierzchołków dwóch grafów pozostają niezmienione, w efekcie czego znów otrzymujemy graf który ma wierzchołki zaetykietowane jako wierzchołki wejścia i wyjścia z lewej i prawej strony. Powiemy, że wierzchołek p jest połączony z wierzchołkiem q w grafie $s \cdot t$, jeżeli w powstałym grafie istnieje ścieżka łącząca odpowiednie wierzchołki. Tak określamy działanie w półgrupie (deterministycznych) transformacji dwukierunkowych T_n.

Niech $\mathcal{D}$ będzie deterministycznym automatem dwukierunkowym, o stanach $1, \ldots, n$, o stanie początkowym 1 oraz zbiorze stanów końcowych Q_F. Wówczas, każda litera $\sigma \in \Sigma \cup \{begin, end\}$ wyznacza pewną transformację dwukierunkową $\hat{\delta}(\sigma)$ zbioru stanów automatu $\mathcal{D}$, mianowicie $\hat{\delta}(\sigma)$ jest grafem opisanym następująco:

- jeżeli $\delta(\sigma, i, d_{\to}) = (j, d_{\to})$, to wierzchołki $i_{\bullet\to}$ oraz $j_{\to\bullet}$ połączone są krawędzią
- jeżeli $\delta(\sigma, i, d_{\to}) = (j, d_{\leftarrow})$, to wierzchołki $i_{\bullet\to}$ oraz $j_{\bullet\leftarrow}$ połączone są krawędzią
- jeżeli $\delta(\sigma, i, d_{\leftarrow}) = (j, d_{\leftarrow})$, to wierzchołki $i_{\leftarrow\bullet}$ oraz $j_{\bullet\leftarrow}$ połączone są krawędzią
- jeżeli $\delta(\sigma, i, d_{\leftarrow}) = (j, d_{\to})$, to wierzchołki $i_{\leftarrow\bullet}$ oraz $j_{\to\bullet}$ połączone są krawędzią

Rozpatrzmy teraz następujący automat półgrupowy $\mathcal{T}$ nad alfabetem Σ.

- Półgrupa automatu $\mathcal{T}$ to T_n.
- Funkcja przejścia automatu $\mathcal{A}$ to $\hat{\delta}$.
- Zbiór akceptujący składa się z tych transformacji dwukierunkowych $s \in T_n$, które spełniają następujący warunek:

transformacja dwukierunkowa $\hat{\delta}(begin) \cdot s \cdot \hat{\delta}(end)$ ma krawędź z wierzchołka $1_{\bullet\rightarrow}$ do wierzchołka $q_{\rightarrow\bullet}$ dla pewnego stanu $q \in Q_F$.

Jest jasne, że automat półgrupowy $\mathcal{T}$ rozpoznaje dokładnie ten język, co automat $\mathcal{D}$. Ponieważ każdy automat półgrupowy wyznacza natychmiast automat deterministyczny (jednokierunkowy), to pokazuje to, że $\mathcal{D}$ rozpoznaje język regularny.

Dla niedeterministycznych automatów dwukierunkowych, można zdefiniować analogiczną do T_n półgrupę niedeterministycznych transformacji dwukierunkowych DT_n, której elementy składają się z podobnych grafów, z tą różnicą, że nie muszą spełniać warunku determinizmu. Mnożenie w tej półgrupie definiuje się w identyczny sposób.

Dla danego dwukierunkowego automatu niedeterministycznego $\mathcal{N}$ można wówczas skonstruować automat półgrupowy, bardzo podobny do tego skonstruowanego powyżej, który rozpoznaje dokładnie ten sam język, co wyjściowy automat. To pokazuje, że nawet niedeterministyczne, dwukierunkowe, automaty rozpoznają jedynie języki regularne.