R O Z D Z I A Ł   I I I

Gramatyki regularne

Przypomnijmy, że gramatykami regularnymi nazywamy wszystkie te gramatyki generatywne, których wszystkie reguły produkcji mają postać $A \to P$ lub $A \to PB$, gdzie $A, B \in V_N$, a $P \in V_T^*$.

1. Postać normalna gramatyk regularnych

Twierdzenie 3.1 (twierdzenie o postaci normalnej).

Każdy język regularny może być generowany przez gramatykę o regułach produkcji następującego kształtu:

(i) $X \to aY$,

(ii) $X \to \lambda$,

gdzie $X,Y \in V_N$, zaś $a \in V_T$.

Idea twierdzenia.

W dowodzie posłużymy się metodę sprowadzalności dowolnej gramatyki regularnej do gramatyki w postaci normalnej z powyższego twierdzenia. Możliwość jej wykorzystania zdeterminuje nam wyprowadzalność każdego ze słów generowanego przez nią języka regularnego. Operację tę wykonuje się w następujący sposób: najpierw (wychodząc oczywiście na samym początku z symbolu początkowego S) stosujemy (tyle razy, jaka jest długość wyprowadzanego słowa) reguły rozszerzające kształtu (i), a następnie stosujemy do ostatniego symbolu (nieterminalnego) tzw. regułę zerowa kształtu (ii) (jest to reguła „zbijająca” ów symbol nieterminalny). Mamy więc:

$$S \to a_1 A_1 \to a_1 a_2 A_2 \to \ldots \to a_1 a_2 \ldots a_{n-1} A_{n-1} \to a_1 a_2 \ldots a_{n-1} a_n A_n \to a_1 a_2 \ldots a_n.$$

Zauważmy, że za każdym razem stosowaliśmy tu reguły produkcji do jedynego, stojącego zawsze na samym końcu symbolu nieterminalnego, dzięki czemu sekwencyjnie „od lewej do prawej” generowaliśmy dane słowo, a na końcu „decydowaliśmy się” na zakończenie derywacji wygenerowanego już słowa.

Uwaga.

Można podać także analogiczne twierdzenie, zastępując w nim jedynie reguły kształtu „(ii) $X \to \lambda$” regułami kształtu „(ii') $X \to a$”. Każda z derywacji wyglądałaby wówczas w następujący sposób:

$$S \to a_1 A_1 \to a_1 a_2 A_2 \to \dots \to a_1 a_2 \dots a_{n-1} A_{n-1} \to a_1 a_2 \dots a_{n-1} a_n.$$

Byłaby więc ona o jeden krok krótsza. Również tu za każdym razem stosowaliśmy reguły produkcji do jedynego, stojącego zawsze na samym końcu symbolu nieterminalnego, dzięki czemu sekwencyjnie „od lewej do prawej” generowaliśmy dane słowo, derywację tą kończąc jednak już nie regułą zerową, lecz tzw. regułą końcową. Pozwoliło to nam skrócić ją o jeden krok, gdyż w przypadku tym każda reguła dobudowuje tu po jednej, kolejnej literze do generowanego słowa (nie ma już reguł pustych).

Tak sformułowane stwierdzenie jest oczywiście prawdziwe, jak to pokazuje poniższe (będące dowodem konstrukcyjnym) jego **uzasadnienie**:

Po sprowadzeniu gramatyki regularnej do postaci normalnej z twierdzenia 3.1, usuwamy z niej każdą z reguł kształtu $X \to \lambda$, przy czym o ile tylko $Z \to bX$ (dla pewnego $Z \in V_N$ i $b \in V_T$) jest już regułą tej gramatyki, to do gramatyki tej dodajemy jeszcze regułę końcową $Z \to b$. Tak otrzymana gramatyka jest równoważna z gramatyką z twierdzenia 3.1. □ **(**uzasadnienia)

Dowód (tw. 3.1).

Niech $G = <V_N, V_T, S, F>$ będzie gramatyką regularną. Zatem wszystkie jej reguły produkcji są postaci:

a) $A \to P$ lub

b) $A \to PB$,

gdzie $A, B \in V_N$, a $P \in V_T^*$.

Gdy $|P| = 0$ (tj. $P = \lambda$), to reguły te (tj. postaci $A \to \lambda$ i $A \to B$) zostawiamy bez zmian (mimo, że tylko pierwsza z nich czyni zadość tezie twierdzenia).

Każdą regułę produkcji gramatyki G postaci a), w której $|P| = 1$ zastępujemy dwiema regułami postaci $A \to PW$, $W \to \lambda$, gdzie A i P są określone j.w., a W jest nowym symbolem nieterminalnym (tj. $W \notin V_N$). Obie te reguły są w postaci z dowodzonego twierdzenia.

Każdą regułę produkcji gramatyki G postaci b), w której $|P| = 1$ zostawiamy bez zmian (spełnia zadość wymaganemu kształtowi reguł z twierdzenia).

W pozostałych regułach produkcji gramatyki G, $|P| > 1$. Każdą z nich:

(i) gdy jest kształtu $A \to a_1 \dots a_n$ (gdzie $A \in V_N$, zaś $a_1,\dots, a_n \in V_T$) wówczas zastępujemy ją ciągiem reguł $A \to a_1Y_1$, $Y_1 \to a_2Y_2,\dots, Y_{n-1} \to a_nY_n$, $Y_n \to \lambda$, gdzie wszystkie Y_i $(1 \le i \le n)$ są nowymi nieterminałami (tj. $Y_i \notin V_N$, a $Y_i \neq Y_j$ dla wszystkich $i \neq j$),

(ii) gdy jest kształtu $A \to a_1 \ldots a_nB$ (gdzie $A, B \in V_N$, zaś $a_1,\ldots, a_n \in V_T$), to wówczas zastępujemy ją ciągiem reguł $A \to a_1Z_1$, $Z_1 \to a_2Z_2,\ldots, Z_{n-1} \to a_nB$, gdzie wszystkie Z_i $(1 \le i \le n)$ są nowymi nieterminałami (tj. $Z_i \notin V_N$, a $Z_i \neq Z_j$ dla wszystkich $i \neq j$).

W ten sposób otrzymujemy gramatykę $G' = \langle V_N', V_T, S, F'\rangle$, w której:

– $V_N' = V_N \cup \{W\} \cup \{Z_i\}_{i\in I} \cup \{Y_j\}_{j\in J}$, gdzie I i J są dwoma zbiorami indeksów nowych nieterminałów odpowiednio postaci Z_i i Y_j,

– F' opisaliśmy wyżej.

W G' występują jedynie reguły produkcji kształtu:

1) $X \to aY$,

2) $X \to Y$,

3) $X \to \lambda$,

gdzie $X, Y \in V_N'$, zaś $a \in V_T$.

Zauważmy, że gramatyki G i G' są sobie równoważne, a w G' w stosunku do pożądanej gramatyki w postaci z twierdzenia nadmiarowe są jedynie reguły postaci $X \to Y$, gdzie $X,Y \in V_N'$.

Korzystając z faktu, że każda z derywacji w G' jest „łańcuchowa” w stosunku do derywacji w G, tworzymy (analogicznie jak w dowodzie twierdzenia 2.2 o postaci normalnej Chomsky’ego) następującą konstrukcyjną procedurę pozbywania się nadmiarowych reguł produkcji postaci 2).

Najpierw dla każdego $X \in V_N'$ oznaczamy przez U(X) zbiór tych wszystkich nieterminałów, z których można wyprowadzić X jedynie za pomocą reguł tejże postaci, powiększony o zbiór jednoelementowy $\{X\}$.

Zatem $Y \overset{*}{\underset{G}{\to}} X$ witw, gdy $Y \in U(X)$, gdzie $X,Y \in V_N'$.

Mając tak określony zbiór U(X) (dla dowolnego $X \in V_N'$) - z gramatyki G' możemy wyeliminować wszystkie reguły z F' postaci $X \to Y$ (gdzie $X,Y \in V_N'$), otrzymując tym samym gramatykę $G'' = \langle V_N', V_T, S, F''\rangle$ o następujących regułach produkcji F'':

1) $\ulcorner Z \to \lambda \urcorner \in F'' \Leftrightarrow \exists_{X\in V_{N'}}: (Z \in U(X) \wedge \ulcorner X \to \lambda \urcorner \in F')$ (tj. $\forall_{Z\in U(X)} \ulcorner Z \to \lambda \urcorner \in F''$);

$K \to L \to Z \to Y \to X \to \lambda$ - reguły F'
- reguły F''

Rys. 3.1.

2) Podobnie, $\ulcorner Z \to aV \urcorner \in F'' \Leftrightarrow \exists_{X\in V_{N'}}: (Z \in U(X) \wedge \ulcorner X \to aV \urcorner \in F')$.

K→ L → Z →Y→ X → aV - reguły F'
- reguły F''

Rys. 3.2.

Oczywiście G' ≡ G'', a zatem (ponieważ również G ≡ G') mamy, że G ≡ G'', tj. L(G) = L(G''). □

2. Wykresy gramatyk regularnych

W gramatykach regularnych wszystkie reguły produkcji mają postać A → P lub A → PB, gdzie A, B ∈ V_N, a P ∈ V_T. Na mocy twierdzenia o postaci normalnej (tw. 3.1), możemy je zastąpić regułami produkcji jedynie kształtu X → aY i X → λ, gdzie X,Y ∈ V_N, zaś a ∈ V_T.

W wykresach gramatyk regularnych wierzchołki będą oznaczać symbole nieterminalne, a krawędzie reprezentować symbole terminalne. Tak więc zapis

$$A \xrightarrow{a} B$$ Rys. 3.3.

oznacza, że A → aB jest regułą produkcji rozważanej gramatyki.

Regułę A → λ umieszczamy na wykresie podkreślając już istniejące na nim A:

$$\underline{A} \xrightarrow{a} B$$ Rys. 3.4.

Symbol początkowy oznaczać będziemy kropą („•”) u góry:

$$S^{\bullet}$$ Rys. 3.5.

Przykład 3.1.

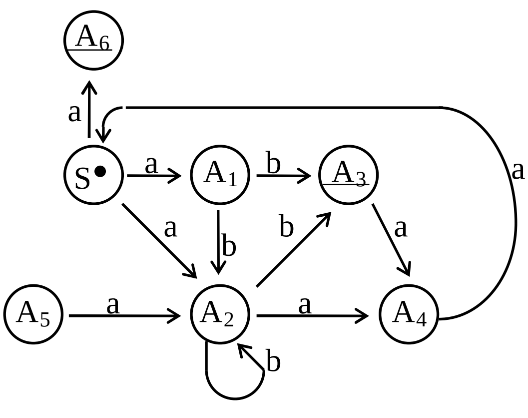
Niech G = <V_N, V_T, S, F>, gdzie V_N = {S, A_1, ..., A_5 }, V_T = {a, b},

a F = {S → a, S → aA_1, S → aA_2, A_1 → bA_2, A_1 → bA_3, A_2 → bA_2,

A_2 → bA_3, A_2 → aA_4, A_3 → aA_4, A_3 → λ, A_4 → aS, A_5 → aA_2}.

Z kształtu reguł produkcji F widzimy, że jest to gramatyka regularna. Dla celów wykresu „zła” jest tutaj jedynie reguła S → a. Dla A_6 ∉ V_N mamy, że jest ona równoważna parze reguł S → aA_6 i A_6 → λ.

Tak powstały zbiór F' = (F \ {S → a}) ∪ {S → aA_6, A_6 → λ} jest już odpowiedni dla celów sporządzenia następującego wykresu powyższej gramatyki:



Rys. 3.6.

□

Ciąg wierzchołków $A_1, \ldots, A_k$ nazywamy *drogą* w wykresie gramatyki regularnej G witw, gdy istnieją $a_1, \ldots, a_{k-1} \in V_T$ takie, że reguły $\ulcorner A_1 \to a_1\, A_2 \urcorner$, $\ulcorner A_2 \to a_2\, A_3 \urcorner, \ldots, \ulcorner A_{k-1} \to a_{k-1}\, A_k \urcorner \in F$. Droga $A_1, \ldots, A_k$ wyznaczona jest przez wartości $a_1, a_2, \ldots, a_{k-1}$. Nie jest to jednak jednoznaczne.

Przykład 3.1 - c.d.

Wyrażenie „aba" możemy wyznaczyć m.in. na jednej z następujących trzech dróg: S, A_2, A_3, A_4; S, A_1, A_3, A_4; S, A_1, A_2, A_4. □

Drogę $A_1, \ldots, A_k$, w której $A_1 = S$, a A_k jest symbolem podkreślonym (tj. zastępującym regułę $A_k \to \lambda$) nazywamy *drogą końcową*.

Lemat.

Dla dowolnej gramatyki regularnej G, $P \in L(G)$ witw, gdy istnieje co najmniej jedna droga końcowa wyznaczona przez P. □

Pojęcia symboli pasywnych, aktywnych, osiągalnych i nieosiągalnych w danej gramatyce zostały określone dla gramatyk bezkontekstowych w paragrafie 4. poprzedniego rozdziału. Ponieważ jednak każda gramatyka regularna jest też bezkontekstowa, zatem możemy przyjąć, że tym samym są one określone również dla gramatyk regularnych.

Przykład 3.1 - c.d.

A jest tu symbolem nieosiągalnym i jednocześnie brak jest symboli pasywnych. □

Ponieważ do nieosiągalnego symbolu nieterminalnego nie prowadzi żadna droga od symbolu początkowego S, zatem z wykresu możemy usunąć reprezentujący go wierzchołek.

Podobnie, możemy usunąć z wykresu gramatyki regularnej wszystkie symbole pasywne, jako nie umożliwiające wyprowadzenie żadnego słowa w powyższej gramatyce (nie prowadzi z nich bowiem żadna droga do jakiegokolwiek z symboli końcowych).

Możemy rozstrzygnąć (na mocy twierdzeń z 4. paragrafu poprzedniego rozdziału, bądź też bezpośrednio z wykresu gramatyki regularnej), które symbole nieterminalne są pasywne i które symbole są nieosiągalne.

Przez konsekwentne usuwanie z danej gramatyki regularnej wszystkich symboli nieosiągalnych i pasywnych otrzymujemy prostszą (równoważną jej) gramatykę regularną, której wykres jest spójny, a każdy jego wierzchołek jest albo końcowy albo leży na przynajmniej jednej drodze końcowej (żaden z nich nie jest więc zbędny).

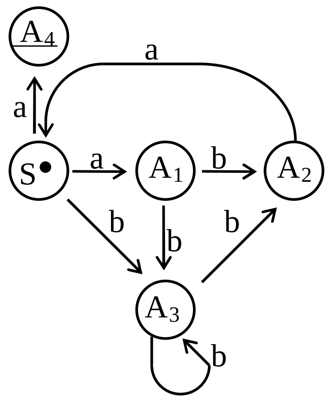
Podobnie, jak to miało miejsce w przypadku gramatyk bezkontekstowych - o gramatyce takiej (tj. takiej, w której żaden symbol nieterminalny nie jest ani pasywny ani nieosiągalny) będziemy mówić, że jest ona w p o s t a c i z r e d u k o w a n e j.

Oprócz operacji wyzbywania się symboli nieterminalnych i pasywnych, na gramatykach regularnych (a więc i na ich wykresach) wprowadza się jeszcze jedną operację, a mianowicie operację zastąpienia tej gramatyki równoważną jej gramatyką, w której jednak symbol początkowy występuje jedynie po lewej stronie reguł produkcji. Polega ona na:

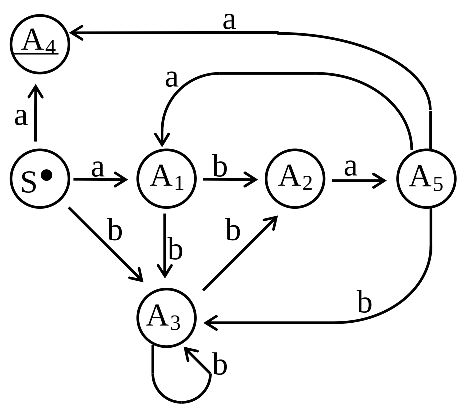
1) wyznaczeniu wszystkich produkcji kształtu $A \to aS$, gdzie S jest symbolem początkowym, $A \in V_N$, zaś $a \in V_T$,
2) zastąpieniu symbolu S w tych produkcjach przez nowy symbol nieterminalny $K \notin V_N$ (a więc tym samym dołączeniu K do V_N),
3) dołączeniu do zbioru reguł produkcji F produkcji $K \to aA$ w tych i tylko w tych przypadkach, w których w F występuje już produkcja $S \to aA$.

Rozpatrzmy to na poniższym przykładzie.

Przykład 3.2.



Rys. 3.7.



Rys. 3.8.

Dla gramatyki z rys. 3.7, otrzymujemy gramatykę z rys. 3.8 w następujący sposób:

1) jedyną regułą produkcji kształtu $A \to aS$ (gdzie S jest symbolem początkowym, $A \in V_N$, zaś $a \in V_T$), jest tu reguła $A_2 \to aS$;
2) regułę tę zastępujemy nową regułą produkcji $A_2 \to aA_5$, gdzie A_5 jest nowym symbolem nieterminalnym;
3) ze względu na występowanie produkcji $S \to aA_1$, $S \to bA_3$ i $S \to aA_4$, dopisujemy produkcje $A_5 \to aA_1$, $A_5 \to bA_3$ i $A_5 \to aA_4$. □

Zadanie 3.1: Wykonaj ponownie zadanie 1.9 tak jednak, by gramatyka ta była regularna. Narysuj jej wykres. □

Zadanie 3.2: Nad alfabetem polskim skonstruuj gramatykę regularną generującą język, w którym cząstka „nie” występuje:

a) w każdym wyrazie (na początku, w środku, czy na końcu),

b) w każdym wyrazie na początku (j. np. w wyrazie „nieco”),

c) w każdym wyrazie na końcu (j. np. w wyrazie „dranie”),

d) w każdym wyrazie w środku (j. np. w wyrazie „koniec”),

e) jeżeli już, to na pewno nie na początku. □

3. Wyrażenia regularne

Na początku odnotujmy następujący **fakt**:

Każdy język skończony jest regularny.

Dowód.

Niech $L = \{P_1, P_2, ..., P_n\}$, gdzie $\forall_{1 \le i \le n}\ P_i \in V_T^*$. Tak określony język L jest skończony (bo składa się ze skończonej liczby słów, bo $card(L) = n < \infty$). Określmy gramatykę $G = <V_N, V_T, S, F>$, gdzie $V_N = \{S\}$, a $F = \{S \to P_1, S \to P_2, ..., S \to P_n\}$. Ze względu na kształt reguł produkcji, gramatyka G jest regularna. □

Wniosek.

Każdy język skończony jest również bezkontekstowy, kontekstowy i struktur frazowych, jako że każdy język regularny jest również językiem bezkontekstowym, kontekstowym i struktur frazowych. □

Obecnie możemy przejść już do właściwego zagadnienia tego paragrafu.

Niech V będzie skończonym alfabetem. Wówczas przez REG(V) oznaczamy z b i ó r w s z y s t k i c h w y r a ż e ń r e g u l a r n y c h n a d a l f a b e t e m V, określony w następujący sposób:

(i) $V \subseteq REG(V)$ (tj. $\forall_{a \in V}\ a \in REG(V)$),

$\lambda \in REG(V)$,

$\emptyset \in REG(V)$ (tj., tzw. wyrażenie stałe $\emptyset$ jest wyrażeniem regularnym);

(ii) $\forall_{\alpha,\beta \in REG(V)}\ (\alpha)^*, (\alpha\beta), (\alpha \cup \beta) \in REG(V)$,

gdzie przez $(\alpha)^*$ rozumiemy i-tą krotność wyrażenia α (i=0,1,2,...).

Przykład 3.3.

Gdy $V = \{0, 1\}$, to wyrażeniami regularnymi nad tym alfabetem V będą m.in.:

$((01)\cup(001))$, $((01)\cup(001))^*$, $((110)^*(0)\cup(11)\cup\emptyset)$. □

Widzimy więc, że wyrażenia regularne są pewnymi, syntaktycznie poprawnymi, napisami, zbudowanymi z symboli: $\varnothing$, λ, *, $\cup$, (i) oraz z liter alfabetu V. Znając moce wiązań (najmocniej wiąże „*”, słabiej konkatenacja, a najsłabiej suma) - w zapisie wyrażenia regularnego możemy opuszczać zbyteczne nawiasy. Dzięki temu, wyrażenia z powyższego przykładu, możemy prościej zapisać następująco: $01\cup001$, $(01\cup001)^*$, $(110)^*0\cup11\cup\varnothing$.

Jeżeli a $\in$ REG(V), to przez L(a) oznaczać będziemy j ę z y k d e n o t o w a n y (oznaczany) p r z e z to w y r a ż e n i e r e g u l a r n e, określony (indukcyjnie po długości a) w następujący sposób:

1) gdy $\alpha = \varnothing$, to $L(\alpha) = L(\varnothing) = \varnothing$ (tj. wyrażenie stałe $\varnothing$ generuje język pusty),
2) gdy $\alpha = \lambda$, to $L(\alpha) = L(\lambda) = \{\lambda\}$,
3) gdy $|\alpha| = 1$ (tj. $\alpha = a \in V$), to $L(\alpha) = \{\alpha\} = \{a\}$,
4) gdy $\alpha \in REG(V)$, to $L(\alpha^*) = (L(\alpha))^*$ (tj. jest domknięciem Kleene’go języka $L(\alpha)$),
5) gdy nadto $\beta \in REG(V)$, to $L(\alpha\beta) = L(\alpha)L(\beta)$, a $L(\alpha\cup\beta) = L(\alpha) \cup L(\beta)$.

Przykład 3.4.

Zobaczmy, jakie języki są denotowane przez przykładowe wyrażenia regularne.

Wyrażenie regularne:	Język przez nie denotowany:
$(a)^* = a^*$	$\{a\}^* = \{\lambda, a, aa, aaa, \ldots\}$
$((a) \cup (b))^* = (a \cup b)^*$	$\{a, b\}^* = \{\lambda, a, b, aa, ab, ba, bb, aaa, \ldots\}$
$((a)^*(b)) = a^*b$	$\{a\}^*\{b\} = \{b, ab, aab, \ldots\}$
$b \cup ab^*$	$\{b\} \cup \{a\}\,\{b\}^* = \{b, a, ab, abb, \ldots\}$

Tab. 3.1.

Zauważmy, że w ostatnim przypadku, słowo „babbb” nie jest elementem tego języka. Znak „$\cup$” oznacza bowiem sumę języków, a nie konkatenację słów. □

Zauważmy, że $\{a,b\}\{a\}^* = \{a\}\{a\}^* \cup \{b\}\{a\}^*$ (=\{a,aa,aaa,...,b,ba,baaa,...\}), a więc tym samym, że $(a \cup b)a^* = aa^* \cup ba^*$ $(= a^+ \cup ba^*)$. Prawo o tym schemacie jest prawdziwe dla wszystkich wyrażeń.

W ogóle, dla dowolnych wyrażeń regularnych P, Q, R, zachodzi:

1) $P \cup (Q \cup R) = (P \cup Q) \cup R$,
2) $P(QR) = (PQ)R$,
3) $P \cup Q = Q \cup P$,
4) $P(Q \cup R) = PQ \cup PR$,
5) $(P \cup Q)R = PR \cup QR$,
6) $\lambda P = P\lambda = P$,
7) $P^* = \lambda \cup PP^*$,
8) $P^* = (\lambda \cup P)^*$,
9) $\varnothing P = P\varnothing = \varnothing$,
10) gdy $P = R \cup PQ$, a $Q \neq \lambda$, to $P = RQ$.

Wspomnijmy jeszcze, że reguły 1) - 10) wraz z regułą podstawiania i przechodniością dają nam zupełny system aksjomatów. Oznacza to, że w przypadku dowolnej formuły języka wyrażeń regularnych (tj. dowolnej równości wyrażeń regularnych), o ile nie jest ona zdaniem (tj. posiada zmienne wolne) - wówczas możemy zamienić ją na zdanie poprzez tzw. domknięcie, tj. poprzedzenie jej kwantyfikatorami ogólnymi wiążącymi wszystkie jej zmienne wolne. Wówczas z powyższych aksjomatów można wywieść bądź powyższe zdanie, bądź jego negacje.

Zależność między językami regularnymi, a językami denotowanymi przez wyrażenia regularne określa następujące twierdzenie:

Twierdzenie 3.2.

Każde wyrażenie regularne denotuje pewien język regularny, a każdy język regularny jest denotowany przez pewne wyrażenie regularne.

Idea twierdzenia:

Twierdzenie to mówi, że klasa języków denotowanych przez wyrażenia regularne pokrywa się z klasą języków regularnych tj., że istnieje wzajemna odpowiedniość między wyrażeniami regularnymi, a gramatykami regularnymi.

Dowód.

1) Pokażemy najpierw, że każde wyrażenie regularne denotuje pewien język regularny, tj. że $\forall_{\alpha \in REG(V)}\ L(\alpha) \in \mathcal{L}_3$.

Otrzymujemy to z definicji języka denotowanego przez wyrażenie regularne.

Proste wyrażenia regularne $\varnothing$, λ, a_1, ..., a_k opisują oczywiście języki $\varnothing$, $\{\lambda\}$, $\{a_1\}$, ..., $\{a_k\}$, które, chociażby ze względu na swą skończoność (patrz: fakt na początku tego paragrafu), są regularne. Ich regularność można także wykazać podając reguły produkcji ich gramatyk:

- dla języka pustego $\varnothing$: $S \to S$,
- dla języka $\{\lambda\}$: $S \to \lambda$,
- dla języka $\{a_i\}$ $(1 \leq i \leq k)$: $S \to a_i$.

Z twierdzenia 1.1 otrzymujemy, że języki regularne są zamknięte na operacje regularne (sumę, iloczyn i domknięcie Kleene'go), podobnie, jak to ma miejsce w przypadku wyrażeń regularnych. Zatem również każdemu złożonemu wyrażeniu regularnemu, odpowiada pewien język regularny, rozbudowywany w identyczny sposób, jak to wyrażenie regularne. □ (1)

2) Pozostaje nam jeszcze pokazać, że każdy język regularny jest denotowany przez pewne wyrażenie regularne, tj. że dla każdego języka regularnego istnieje wyrażenie regularne, które go denotuje.

Niech $L = L(G)$ będzie językiem regularnym wyznaczonym przez gramatykę regularną $G = <V_N, V_T, S, F>$. Z twierdzenia 3.1, gramatykę tę możemy sprowadzić do postaci normalnej o regułach produkcji jedynie kształtu $X \rightarrow aY$ i $X \rightarrow \lambda$, gdzie $X,Y \in V_N$, zaś $a \in V_T$. Dalej sprowadzamy ją do postaci zredukowanej (tj. bez symboli pasywnych i nieosiągalnych). Nadto, jak już to wyżej zauważyliśmy (w przedostatnim akapicie na str. 48), zachodzi następujący **fakt**:

Wykres gramatyki regularnej w postaci zredukowanej jest spójny, a każdy jego wierzchołek jest albo końcowy, albo leży na przynajmniej jednej drodze końcowej. □ (faktu)

Nasz dowód przeprowadzimy indukcyjnie, że względu na liczbę łuków wykresu gramatyki G (tj. względem liczby reguł produkcji gramatyki G).

a) Jeżeli wykres gramatyki G ma $n = 0$ łuków, to (ze względu na powyższy fakt) składa się on albo z jednego wierzchołka zarazem początkowego i końcowego (wtedy język $L = \{\lambda\}$) albo nie zawiera żadnych wierzchołków (wtedy język $L = \emptyset$). Wprost z definicji wynika, że każdy z tych języków jest denotowany przez pewne wyrażenie regularne (odpowiednio: $\emptyset$ i λ).

b) Załóżmy obecnie, że dla wykresu każdej gramatyki G w postaci normalnej zredukowanej, mającego $s < n$ łuków, język $L' = L(G')$ jest generowany przez pewne wyrażenie regularne.

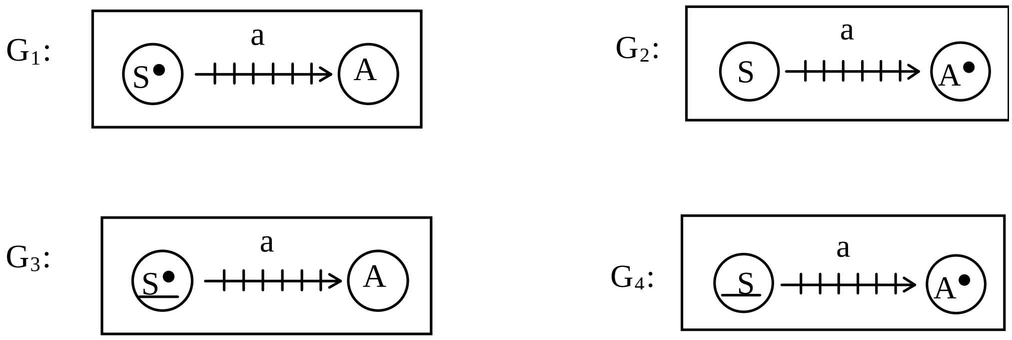
Z wykresu gramatyki G usuwamy jeden łuk wychodzący z wierzchołka początkowego np. łuk odpowiadający produkcji $S \rightarrow aA$, a tak otrzymaną gramatykę oznaczamy przez $\mathbf{G_1}$.

Przez $\mathbf{G_2}$ oznaczamy gramatykę o tym samym wykresie co G_1, z tą różnicą, że nie wierzchołek S, ale wierzchołek A jest początkowy.

Przez $\mathbf{G_3}$ oznaczamy gramatykę o tym samym wykresie co G_1, z tą różnicą, że jedynym wierzchołkiem początkowym i końcowym jest wierzchołek S.

Przez $\mathbf{G_4}$ oznaczamy gramatykę o tym samym wykresie co G_1, z tą różnicą, że wierzchołek A jest początkowy, zaś S jest jedynym wierzchołkiem końcowym.

Wykresy tych gramatyk pokazane są schematycznie na poniższym rysunku:



Rys. 3.9.

Każda z tych gramatyk ma n - 1 łuków, ale mogą one nie spełniać założenia naszego faktu, a zatem - mogą również nie spełniać jego tezy. Sprowadzając je do postaci zredukowanej:

- nie zwiększamy liczby ich krawędzi (będziemy więc mogli stosować do nich założenie indukcyjne),
- otrzymujemy wykresy gramatyk spełniające tezę faktu.

Na podstawie założenia indukcyjnego, języki $L(G_1)$, $L(G_2)$, $L(G_3)$ i $L(G_4)$ są generowane przez wyrażenia regularne. Pokażemy, że zachodzi równość:

$$L(G) = L(G_1) \cup L(G_3) \cdot [\{a\}L(G_4)]^* \cdot \{a\}L(G_2). \qquad (1)$$

Rzeczywiście:

- gdy droga końcowa w G nie wykorzystuje krawędzi a z S do A, to jest to również droga końcowa w G_1; podobnie - każda droga końcowa w G_1 jest droga końcowa w G,
- gdy droga końcowa w G wykorzystuje co najmniej raz krawędź a z S do A, to jest ona postaci:

$$\underbrace{x}_{\text{I blok}} \quad \underbrace{ay_1\ ay_2\ \dots\ ay_m}_{\text{II blok}} \quad \underbrace{az}_{\text{III blok}}, \qquad (2)$$

gdzie:

- x jest drogą końcową w G_3,
- $y_1, y_2, \dots, y_m$ - to drogi końcowe w G_4,
- z jest drogą końcową w G_2.

Rzeczywiście:

– najpierw może bowiem wychodząc z S wcale nie wykorzystywać krawędzi a z S do A i znaleźć się w końcu w S (za to odpowiedzialny jest I blok; odpowiada on językowi $L(G_3)$),
– następnie może m razy (gdzie $m \in N_0$) po przejściu krawędzią a z S do A powrócić z powrotem do S (za to odpowiedzialny jest II blok; odpowiada on językowi $[\{a\}L(G_4)]^*$),
– i na koniec, musi wykorzystać krawędź a przechodząc z S do A, by dalej już, nie wykorzystując tej krawędzi, przejść do jednego ze stanów końcowych gramatyki G (za to odpowiedzialny jest III blok; odpowiada on językowi $\{a\}L(G_2)$).

Z powyższego wywodu widać także, że każda droga postaci (1) jest końcowa w G. Jak już to zauważyliśmy, języki występujące po prawej stronie równości są generowane przez wyrażenia regularne, zatem (z indukcyjnej definicji wyrażeń regularnych) - również język L(G) jest generowany przez pewne wyrażenie regularne. Kończy to dowód całego twierdzenia. □

Mając udowodnione powyższe twierdzenie, możemy pokusić się o poszukanie metody:

1) znajdywania gramatyki regularnej odpowiadającej danemu wyrażeniu regularnemu,
2) znajdywania wyrażenia regularnego odpowiadającego danej gramatyce regularnej.

Ad 1.

Najprościej jest tego dokonać posługując się wykresem gramatyk regularnych, a więc (tym samym) gramatykami w postaci normalnej (z twierdzenia 3.1). Gdybyśmy mieli udowodnione twierdzenie 1.1, moglibyśmy robić to następująco: dla gramatyki regularnej w postaci normalnej (będącej szczególną postacią gramatyki regularnej), stosujemy procedurę przedstawioną w dowodzie twierdzenia 1.1, a następnie tak otrzymane gramatyki regularne sprowadzamy do postaci normalnej. My jednak postąpimy inaczej. Mając mianowicie wykresy gramatyk regularnych, będziemy wykonywać na nich operacje regularne w taki sposób, by od razu otrzymać wykresy gramatyk regularnych (a więc tym samym gramatyki regularne w postaci normalnej).

Twierdzenie 3.3.

Jeżeli języki L_1 i L_2 są generowane przez gramatyki regularne w postaci normalnej (odpowiednio G_1 i G_2), to języki otrzymane przez stosowanie na nich operacji regularnych:

a) $L_1 \cup L_2$,

b) $L_1 L_2$,

c) L_1^*

są również generowane przez pewne gramatyki regularne w postaci normalnej.

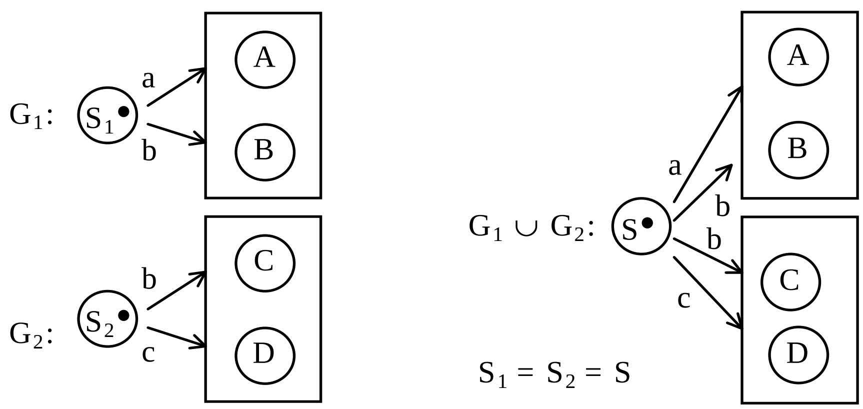
Dowód.

Dowód przeprowadzimy konstrukcyjnie, że względu na stosowane operacje regularne. Najpierw jednak, bez straty ogólności rozważań, możemy przyjąć że:

– obie gramatyki G_1 i G_2 mają rozłączne zbiory symboli nieterminalnych (w przeciwnym przypadku możemy je rozłączyć - pokrywającym się symbolom nadać w gramatyce G_2 nowe, nie występujące dotychczas nazwy),

– za pomocą ostatniej operacji przedstawionej na koniec drugiego paragrafu tego rozdziału, symbole początkowe S_1 i S_2 tych gramatyk zostały odseparowane (tj. gramatyki te zastąpiliśmy równoważnymi im gramatykami, w których jednak symbol początkowy występuje jedynie po lewej stronie reguł produkcji).

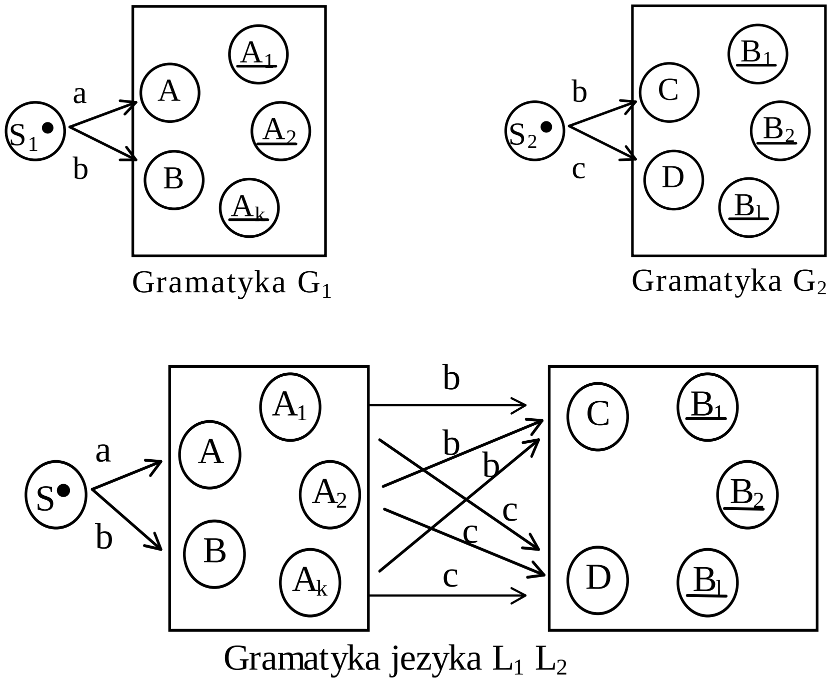
a) By otrzymać gramatykę języka $L_1 \cup L_2$, kładziemy po prostu $S_1 = S_2 = S$, jak to pokazano na rysunku 3.10.



Rys. 3.10.

b) Konstrukcja gramatyki generującej język $L_1 L_2$.

Niech gramatyka G_1 ma symbole końcowe $A_1, A_2, \ldots, A_k$. Usuńmy z gramatyki G_2 symbol S_2, a następnie połączmy strzałkami wychodzącymi z symboli końcowych $A_1, A_2, \ldots, A_k$ te symbole, do których dochodziły strzałki z S_2 (patrz rys. 3.11). Tak otrzymana gramatyka generuje język $L_1 L_2$.



Rys. 3.11.

Zauważmy jeszcze, że symbole końcowe $A_1, A_2, \ldots, A_k$ gramatyki G_1, w nowej gramatyce nie są już symbolami końcowymi. Końcowe są w niej tylko symbole końcowe gramatyki G_2 (tj. $B_1, B_2, \ldots, B_l$).

Zauważmy również, że postępowanie to polega tu w gruncie rzeczy na:

- usunięciu symbolu S_2 i zastąpieniu reguł $S_2 \to bC$, $S_2 \to cD$ regułami:

$$A_1 \to bC,\ A_2 \to bC, \ldots, A_k \to bC,$$

$$A_1 \to cD,\ A_2 \to cD, \ldots, A_k \to cD,$$

- oraz usunięciu reguł $A_1 \to \lambda$, $A_2 \to \lambda$, ..., $A_k \to \lambda$ gramatyki G_1.

c) Konstrukcja gramatyki języka $L_1{}^*$ polega na sklejeniu w G_1 wszystkich <u>osiągalnych</u> wierzchołków końcowych z wierzchołkiem początkowym, tj. na położeniu:

$$S_1 = A_1,\ S_1 = A_2, \ldots, S_1 = A_k.$$

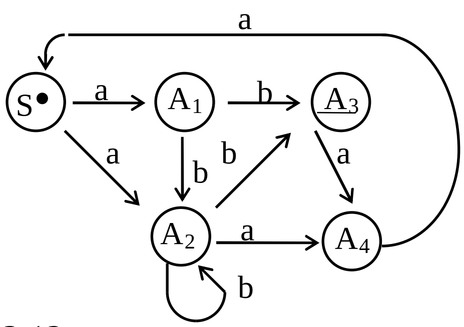
Operacja ta polega w istocie na:

– zastąpieniu reguł postaci $A_i \to aB$ i $B \to bA_i$ regułami $S_i \to aB$ i $B \to bS_i$ dla każdego $i = 1, 2, \ldots, k$,

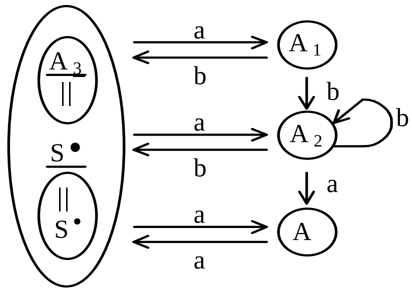
– zastąpieniu reguł $A_i \to \lambda$ $(i = 1, 2, \ldots, k)$ regułą $S_1 \to \lambda$.

Do zakończenia dowodu pozostaje nam jedynie formalne udowodnienie, ze konstrukcje a), b) i c) dają odpowiednio gramatyki generujące języki $L_1 \cup L_2$, L_1L_2 i L_1, co pozostawiam już jednak zainteresowanemu czytelnikowi.

Przykład 3.5.



Rys. 3.12.



Rys. 3.13.

Dla gramatyki z rys.3.12 (jest to gramatyka z przykładu 3.1 bez wierzchołków A_5 i A_6), po zastosowaniu powyższej konstrukcji, otrzymujemy gramatykę z rys. 3.13. □

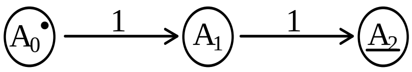
Jak wyglądają wykresy gramatyk regularnych prostych wyrażeń regularnych wiemy już z pierwszej części dowodu twierdzenia 3.2. Z metodą konstrukcji wykresów gramatyk regularnych dla złożonych wyrażeń regularnych zapoznaliśmy się z kolei w dowodzie twierdzenia 3.3. Obecnie możemy więc już zobrazować na konkretnym przykładzie działanie naszej metody.

Przykład 3.6.

Znajdź gramatykę regularną generującą język dany wyrażeniem regularnym:

$$[01 \cup (11)^{*}]01.$$

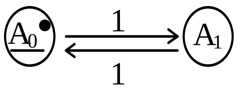
Język {11} jest generowany przez gramatykę:



(1)

Rys. 3.14.

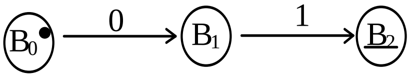
Sklejając symbol A_2 z symbolem A_0 otrzymujemy wykres gramatyki języka $\{11\}^{*}$:



(2)

Rys. 3.15.

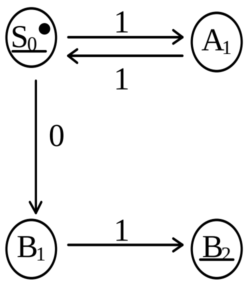
Z kolei gramatykę języka {01} przedstawia wykres:



(3)

Rys. 3.16.

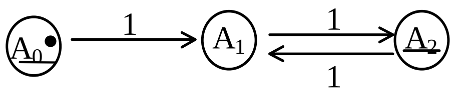
Sklejając wykresy (2) i (3) (poprzez utożsamienie $A_0 = B_0 = S$) - powinniśmy otrzymać wykres gramatyki języka $\{01\} \cup \{11\}^{*}$, jednak (jak się okazuje po poniższym rys. 3.17) - otrzymujemy wykres gramatyki języka $(\{11\}^{*} \cup \lambda)\{01\}$:



(4)

Rys. 3.17.

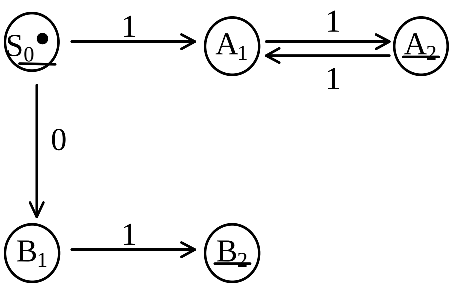
Spowodowane jest to faktem, że symbol S jest tu tak początkowy jak i końcowy. Aby usunąć ową niedogodność powodującą przekłamanie mocy generatywnej tworzonej gramatyki - odpowiednio przekształcamy gramatykę (2) do poniższej postaci (2'):



(2')

Rys. 3.18.

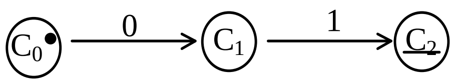
a następnie łączymy ją z gramatyką (3) otrzymując tym samym w miejsce gramatyki (4) właściwą już gramatykę (4'):



(4')

Rys. 3.19.

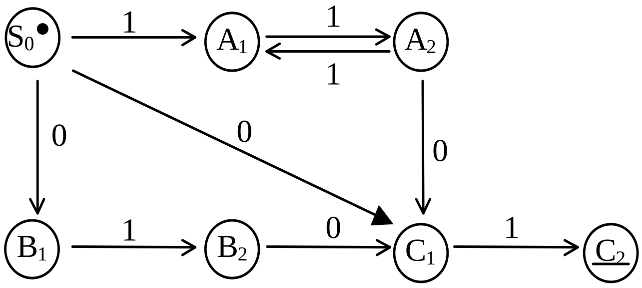
Pozostaje nam jeszcze dołączyć gramatykę języka {01}, ale ponieważ już ona wystąpiła, bierzemy ją z nowymi symbolami nieterminalnymi C_0, C_1 i C_2 :



(3')

Rys. 3.20.

Usuwając z gramatyki (3') symbol początkowy C_0 i łącząc strzałkami z 0 wszystkie symbole końcowe (S_0, A_2 i B_2) gramatyki (4') z wierzchołkiem C_1 (bo w (3') mieliśmy: $C_0 \rightarrow 0C_1$) - otrzymujemy szukany wykres gramatyki języka $(\{01\} \cup \{11\}^*)\{01\}$:



(5)

Rys. 3.21.

Oczywiście, symbole końcowe gramatyki (4') przestają być symbolami końcowymi w gramatyce (5). □

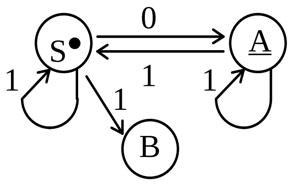
Opisane tu zagadnienie znajdywania gramatyki regularnej odpowiadającej danemu wyrażeniu regularnemu, nazywamy z a g a d n i e n i e m s y n t e z y g r a m a t y k i .

Ad 2.

Twierdzenie 3.2 pozwala również opisywać dowolny język regularny poprzez wyrażenie regularne, bezpośrednio przedstawiające kształt wszystkich generowanych przez niego słów. Metoda ta została przedstawiona w drugiej części dowodu tego twierdzenia. Prześledźmy ją na poniższym przykładzie:

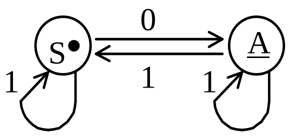
Przykład 3.7.

Określić (poprzez podanie wyrażenia regularnego) język generowany przez gramatykę daną następującym wykresem:



Rys. 3.22.

Zgodnie z procedurą przedstawioną w dowodzie - tworzymy najpierw wykres jej postaci zredukowanej

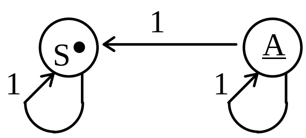


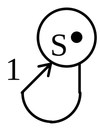
Rys. 3.23.

a następnie (usuwając strzałkę z 0 „idącą” z S do A) wykresy gramatyk G_1, G_2, G_3 i G_4 oraz ich postacie zredukowane i na ich podstawie określamy generowane przez nie języki:

wykresy gramatyk   ich postacie zredukowane $L(G_i)$

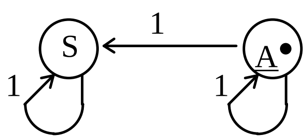
G_1:

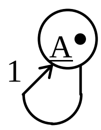




$\varnothing$

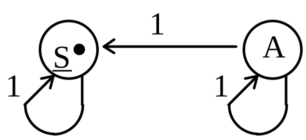
G_2:

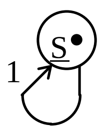




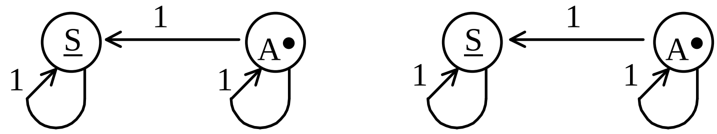
$\{1^*\}$

G_3:





$\{1^*\}$

G_4: 

$\{11^*\}$

Tab. 3.2.

Stąd

$$L(G) = L(G_1) \cup L(G_3) \cdot [\{a\}L(G_4)]^* \cdot \{a\}L(G_2) =$$
$$= \varnothing \cup \{1^*\} \cdot [\{0\}\{11^*\}]^* \cdot \{0\}\{1^*\} =$$
$$= \{1^*\} \cdot \{011^*\}^* \cdot \{01^*\}.$$

Gramatyka ta generuje więc język regularny opisany wyrażeniem regularnym $1^* (011^*)^* 01^*$, czyli wyrażeniem regularnym $1^* (01^+)^* 01^*$. □

Wadą tej metody jest to, że w przypadku gramatyki bardziej rozbudowanej (o większej liczbie krawędzi), otrzymuje się nadal bardzo skomplikowane wyrażenia regularne wymagające żmudnego skracania.

Opisane tu zagadnienie znajdywania wyrażenia regularnego odpowiadającego danej gramatyce regularnej, nazywamy z a g a d n i e n i e m a n a l i z y g r a m a t y k i .

Zadanie 3.3: Niech A, B i C będą wyrażeniami regularnymi. Niech $A \cap B$ będzie wyrażeniem regularnym opisującym język $\{A\} \cap \{B\}$, a A' niech będzie wyrażeniem regularnym opisującym język $V^* \setminus \{A\}$. Podobnie, niech A^T będzie operacją transpozycji taką, że $\{A^T\} = \{A\}^T$. Udowodnij, że:

a) $A(BC) = (AB)C$,
b) $AB \cup AC = A (B \cup C)$,
c) $BA \cup CA = (B \cup C) A$,
d) $\lambda A = A\lambda = A$,
e) $\varnothing \cup A = A$,
f) $A\varnothing = \varnothing = \varnothing A$,
g) $(A^*)^* = A^*$,
h) $ABB^* \cup AB^* = AB^*$,
i) $\lambda^* = \lambda$,
j) $\varnothing^* = \lambda$,
k) $(A \cup B)^* = (A^* \cup B^*)^* = (A^*B^*)^*$,
l) $A \cup B = B \cup A$,
m) $(A \cup B) \cup C = A \cup (B \cup C)$,
n) $(A \cup B)' = A' \cap B'$,
o) $(A \cap B)' = A' \cup B'$,
p) $(AB)^T = B^T A^T$,
q) $(A \cup B)^T = A^T \cup B^T$,
r) $(A^T)^* = (A^*)^T$,
s) $A^* = (\lambda \cup A)^*$,
t) $\lambda^T = \lambda$,
u) $\varnothing^T = \varnothing$,
v) $A'' = A$. □

4. Języki regularne, a języki liniowe i bezkontekstowe. Gramatyki samorozszerzające

Gramatykę bezkontekstową G nazywamy l i n i o w ą , gdy wszystkie jej reguły produkcji są postaci $A \to P$ lub $A \to Q_1BQ_2$, gdzie $A, B \in V_N$, zaś $P, Q_1, Q_2 \in V_T^*$.
Z kolei gramatykę liniową nazywamy p r a w o s t r o n n i e l i n i o w ą , gdy $Q_2 = \lambda$, a l e w o s t r o n n i e l i n i o w ą, gdy $Q_1 = \lambda$.
Oznaczmy:

– przez Lin - klasę wszystkich języków liniowych, generowanych przez gramatyki liniowe,
– przez PL - klasę wszystkich języków prawostronnie liniowych, generowanych przez gramatyki prawostronnie liniowe,
– przez LL - klasę wszystkich języków lewostronnie liniowych, generowanych przez gramatyki lewostronnie liniowe,
– $\mathcal{L}_3$ przez Reg,
– $\mathcal{L}_2$ przez CF (skrót od „contex free").

Przyjmijmy również skrótowe zapisy g.Lin, g.PL, g.LL, g.Reg i g.CF, oznaczające odpowiednio dowolną gramatykę liniową, prawostronnie liniową, lewostronnie liniową, regularną i bezkontekstową, a przez j.Lin, j.PL, j.LL, j.Reg i j.CF - odpowiednio generowane przez nie języki (liniowy, prawostronnie liniowy, lewostronnie liniowy, regularny i bezkontekstowy).

Zauważmy, że wprost z powyżej wprowadzonych definicji mamy:

1) Lin $\subseteq$ CF,
2) Reg = PL,
3) PL $\subseteq$ Lin,
4) LL $\subseteq$ Lin.

Twierdzenie 3.4.

Każda gramatyka lewostronnie liniowa generuje język typu 3 (tj. LL $\subseteq \mathcal{L}_3$).

Dowód.

Niech $G = \langle V_N, V_T, S, F \rangle$ będzie gramatyką lewostronnie liniową, w której $V_N = \{S, A_1, ..., A_z\}$. Zakładamy, że symbol początkowy S nie występuje po prawej stronie reguł produkcji (jeśli by występował, to zgodnie z procedurą przedstawioną na końcu drugiego paragrafu tego rozdziału, pozbywamy się go). Niech $P \in V_T^*$.

Konstruujemy gramatykę regularną $G' = \langle V_N, V_T, S, F'\rangle$ z następująco określonymi regułami produkcji F':

1) jeśli $\ulcorner S \to P \urcorner \in F$, to $\ulcorner S \to P \urcorner \in F'$,

2) jeśli $\ulcorner A_k \to P \urcorner \in F$, to $\ulcorner S \to PA_k \urcorner \in F'$,

3) jeśli $\ulcorner A_k \to A_jP \urcorner \in F$, to $\ulcorner A_j \to PA_k \urcorner \in F'$,

4) jeśli $\ulcorner S \to A_jP \urcorner \in F$, to $\ulcorner A_j \to P \urcorner \in F'$.

Pokażemy, że $L(G) \subseteq L(G')$. Niech więc $P \in L(G)$; pokażemy, że również $P \in L(G')$.

1) Jeśli $\ulcorner S \to P \urcorner \in F$, to P można wyprowadzić w G bezpośrednio z symbolu początkowego S, a zatem można je wyprowadzić również tą samą drogą w gramatyce G' (bo $\ulcorner S \to P \urcorner \in F'$), a więc $P \in L(G')$.

2) Jeśli zaś $\ulcorner S \to P \urcorner \notin F$, to słowo P można tak rozbić na części $P = P_nP_{n-1}...P_2P_1$, że jego wywód w G będzie wyglądał następująco (nad strzałkami zaznaczono numery stosowanych reguł produkcji):

$$S \xrightarrow{4} A_{i_1}P_1 \xrightarrow{3} A_{i_2}P_2P_1 \xrightarrow{3} A_{i_3}P_3P_2P_1 \xrightarrow{3} A_{i_{m-1}}P_{m-1}P_{m-2}K\ P_3P_2P_1 \xrightarrow{2}$$
$$\xrightarrow{2} P_mP_{m-1}K\ P_3P_2P_1. \qquad (*)$$

Pokażemy, że w gramatyce G' można je otrzymać na drodze następującego wywodu:

$$S \xrightarrow{2} P_mA_{i_{m-1}} \xrightarrow{3} P_mP_{m-1}A_{i_{m-2}} \xrightarrow{3} P_mP_{m-1}P_{m-2}A_{i_{m-3}} \xrightarrow{3} K \xrightarrow{3} P_mP_{m-1}K\ P_3P_2A_{i_1} \xrightarrow{4}$$
$$\xrightarrow{4} P_mP_{m-1}K\ P_3P_2P_1. \qquad (**)$$

Rzeczywiście:

– ponieważ w $(*)$ stosowaliśmy $\ulcorner A_{i_{m-1}} \xrightarrow{2} P_m \urcorner \in F$, więc w $(**)$ mogliśmy stosować $\ulcorner S \xrightarrow{2} P_mA_{i_{m-1}} \urcorner \in F'$;

– ponieważ w $(*)$ stosowaliśmy

$\ulcorner A_{i_1} \xrightarrow{3} A_{i_2}P_2 \urcorner$, $\ulcorner A_{i_2} \xrightarrow{3} A_{i_3}P_3 \urcorner$, ..., $\ulcorner A_{i_{m-2}} \xrightarrow{3} A_{i_{m-1}}P_{m-1} \urcorner \in F$,

więc w $(**)$ mogliśmy stosować odpowiednio

$\ulcorner A_{i_2} \xrightarrow{3} P_2A_{i_1} \urcorner$, $\ulcorner A_{i_3} \xrightarrow{3} P_3A_{i_2} \urcorner$, ..., $\ulcorner A_{i_{m-1}} \xrightarrow{3} P_{m-1}A_{i_{m-2}} \urcorner \in F'$;

– i na koniec, ponieważ w $(*)$ stosowaliśmy "$S \xrightarrow{4} A_{i_1}P_1$" $\in F$, więc w $(**)$ mogliśmy stosować $\ulcorner A_{i_1} \xrightarrow{4} P_1 \urcorner \in F'$.

Zatem dla dowolnego słowa P, jeśli $P \in L(G)$, to $P \in L(G')$, czyli w sumie $L(G) \subseteq L(G')$, co właściwie kończy dowód. □

Dodajmy jeszcze, że z symetrii wynika inkluzja w drugą stronę, co w sumie daje nam, że L(G) = L(G').

Wnioski:

1) $\mathcal{L}_3$ jest zamknięta ze względu na operację odbicia zwierciadlanego;

2) Każdy język typu 3 (tj. prawostronnie liniowy) jest generowany przez pewną gramatykę lewostronnie liniową (tj. LL = Reg $\subseteq$ PL).

Uzasadnienie (wniosków).

1) Niech $a_1a_2...a_n$ będzie dowolnym słowem generowanym przez g.PL. Wówczas $a_n...a_2a_1$ jest słowem generowanym przez pewną g.LL, w której reguły produkcji otrzymujemy przez konsekwentną zamianę kolejności (na odwrotną) wszystkich symboli w następnikach wszystkich reguł produkcji (tak więc np. regułę $A \rightarrow P_1 \dots P_n B$ zastępujemy regułą $A \rightarrow BP_n \dots P_1$).

Niech $L \in \mathcal{L}_3$, tj. L jest językiem regularnym. Ponieważ klasa gramatyk regularnych pokrywa się z klasą gramatyk prawostronnie liniowych (Reg = PL), zatem L jest j.PL. Wnosimy stąd, że L^{-1} jest j.LL, tj. $L^{-1} \in$ L (z poprzedniego twierdzenia). □ (1)

2) Z rozważań powyższych wynika, że gdy L jest j.PL, to jego inwersja L^{-1} jest również j.PL, czyli $L = (L^{-1})^{-1}$ jako jego inwersja jest j.LL. Zatem dowolny j.PL (tj. regularny) jest też lewostronnie liniowy. □ (2)

Z twierdzenia i z drugiego wniosku z niego otrzymujemy następujący wniosek:

Wniosek.

Siła generacyjna gramatyk prawostronnie liniowych jest taka sama jak gramatyk lewostronnie liniowych. □

W sumie mamy więc, że: Reg = PL = LL $\subseteq$ Lin $\subseteq$ CF. Pokażemy, że Reg $\subset$ Lin, tj. że nie wszystkie gramatyki liniowe generują języki regularne. W tym celu rozpatrzmy na początek poniższy przykład.

Przykład 3.8.

Niech $L_1 = \{a^n b^n c^k: n \geq 1, k \geq 1\}$, a $L_2 = \{a^k b^n c^k: n \geq 1, k \geq 1\}$. Oba te języki są generowane przez gramatyki liniowe G = <{S, A}, {a, b, c}, S, F>, gdzie:

- w gramatyce generującej język L_1, F = {$S \rightarrow Sc$, $S \rightarrow Ac$, $A \rightarrow aAb$, $A \rightarrow ab$},
- w gramatyce generującej język L_2, F = {$S \rightarrow aS$, $S \rightarrow aA$, $A \rightarrow bAc$, $A \rightarrow bc$}.

Generowanie dowolnego słowa w pierwszej z tych gramatyk wygląda następująco:

$$S \rightarrow Sc \rightarrow Scc \rightarrow \dots \rightarrow Sc^{k-1} \rightarrow Ac^k \rightarrow aAbc^k \rightarrow \dots \rightarrow a^{n-1}Ab^{n-1}c^k \rightarrow a^n b^n c^k.$$

Podobnie wygląda generowanie dowolnego słowa w drugiej z tych gramatyk.

$L_1 \cap L_2 = \{a^n b^n c^n: n \geq 1\}$, a to nie jest język bezkontekstowy (z ostatniego wniosku z 4 paragrafu 2 rozdziału). □

Widać stąd, że przecięcie dwóch j.Lin nie koniecznie musi dawać j.CF, a tym bardziej (ze względu na "Lin $\subseteq$ CF"), że przecięcia j.Lin $\cap$ j.CF, j.CF $\cap$ j.Lin i j.CF $\cap$ j.CF nie muszą dawać j.CF. O tym, że j.CF daje nam zawsze przecięcie j.CF z j.Reg, orzeka poniższe twierdzenie.

Twierdzenie 3.5.

Jeśli $L \in \mathcal{L}_2$, a $L' \in \mathsf{L}_3$, to $L \cap L' \in \mathcal{L}_2$.

Dowód.

Niech $L \in \mathcal{L}_2$, a $L' \in \mathsf{L}_3$. Rozważymy dwa przypadki ze względu na należenie λ do L.

<u>I przypadek</u> - gdy $\lambda \notin L$.

Należy pokazać, że $L \cap L' \in \mathcal{L}_2$.

Niech L (jako język bezkontekstowy) będzie generowany przez gramatykę $G = \langle V_N, V_T, S, F\rangle$ w postaci normalnej Chomsky'ego (tj. mającą reguły produkcji jedynie postaci $X \rightarrow YZ$ i $X \rightarrow a$).

Niech L' (jako język regularny) będzie generowany przez gramatykę $G' = \langle V_N', V_T', S', F'\rangle$ w postaci normalnej z twierdzenia 3.1 (tj. mającą reguły produkcji jedynie postaci $X \rightarrow aY$ i $X \rightarrow \lambda$).

Oznaczmy przez $W = \{X_1, X_2, ..., X_k\}$ zbiór tych wszystkich symboli nieterminalnych gramatyki G', dla których $\ulcorner X_i \rightarrow \lambda \urcorner \in F'$, gdzie $1 \leq i \leq k$ (dla uproszczenia, jednak bez straty ogólności, przyjęliśmy tu, że jest to właśnie k pierwszych elementów nieterminalnych).

Tworzymy k gramatyk bezkontekstowych:

$$G_i = \langle V_N' \times (V_N \cup V_T) \times V_N', V_T \cup V_T', (S', S, X_i), F''\rangle \quad (1 \leq i \leq k),$$

gdzie przynależność reguł do F'' określamy następująco:

1) $\ulcorner A \rightarrow BC \urcorner \in F$ witw, gdy $\forall_{X,Y,Z \in V_{N'}} \ulcorner (X,A,Y) \rightarrow (X,B,Z)\ (Z,C,Y) \urcorner \in F''$,

2) $\ulcorner A \rightarrow a \urcorner \in F$ witw, gdy $\forall_{X,Y \in V_{N'}} \ulcorner (X,A,Y) \rightarrow (X,a,Y) \urcorner \in F''$,

3) $\ulcorner X \rightarrow aY \urcorner \in F'$ witw, gdy $\ulcorner (X,A,Y) \rightarrow a \urcorner \in F''$.

Jak widać (ze względu na kształt reguł produkcji) - tak określona gramatyka jest bezkontekstowa. Pokażemy, że $L \cap L' = \bigcup_{i=1}^{k} L(G_i)$, a więc tym samym, że $L \cap L'$ jest językiem bezkontekstowym (jako suma języków bezkontekstowych).

Słowo postaci $P = a_1 \ldots a_n$ ($\in V_T^*$) $\in L(G)$ witw, gdy

$$\exists_i \forall_{Z_1,...,Z_{n-1}} \text{ istnieje wyprowadzenie w } G_i :$$

$$(S', S, X_i) \xrightarrow[G]{*} (S', a_1, Z_1)\,(Z_1, a_2, Z_2)\,\ldots\,(Z_{n-1}, a_n, X_i) \qquad (1)$$

(stosowaliśmy tu reguły 1) i 2)).

Z drugiej strony, $P \in L(G')$ witw, gdy

$\exists\, Z_1,\ldots,Z_{n-1}\ \exists\, X_i \in W$: w G' istnieje derywacja postaci:

$$S' \to a_1 Z_1 \to a_1 a_2 Z_2 \to \ldots \to a_1 \ldots a_n X_i \to a_1 \ldots a_n. \qquad (2)$$

Zatem (z 3) punktu definicji F''): $P \in L(G')$ witw, gdy w G_i istnieje derywacja:

$$(S', a_1, Z_1)\,(Z_1, a_2, Z_2)\,\ldots\,(Z_{n-1}, a_n, X_i) \xrightarrow[G]{*} a_1 \ldots a_n. \qquad (3)$$

Koniec derywacji (1) jest początkiem derywacji (3), więc w sumie:

$$P \in L \cap L' \text{ witw, gdy } \exists\, i\colon P \in L(G_i).$$

Tak więc, mamy że: $L \cap L' = \bigcup_{i=1}^{k} L(G_i)$.

<u>II przypadek</u> - gdy $\lambda \in L$.

Konstruujemy gramatykę generującą język $L \setminus \{\lambda\}$. Następnie tworzymy gramatykę generującą język $L' \cap (L \setminus \{\lambda\})$, który (jak już wiemy z dowodu I przypadku) jest bezkontekstowy.

Z definicji, $\lambda \in L \cap L'$ witw, gdy $\lambda \in L$ i $\lambda \in L'$.

Rozważmy dwie możliwości:

a) gdy $\lambda \in L$ i $\lambda \notin L'$ - wówczas $\lambda \notin L \cap L'$, a więc $L' \cap (L \setminus \{\lambda\}) = L' \cap L$, a ponieważ $L' \cap (L \setminus \{\lambda\})$ jest językiem bezkontekstowym, więc i równy mu język $L' \cap L$ też jest bezkontekstowy;

b) jeżeli jednak przy naszym założeniu, że $\lambda \in L$, mamy również że $\lambda \in L'$ (tj. gdy $\lambda \in L \cap L'$), to do określonych już reguł produkcji gramatyki generującej język bezkontekstowy $L' \cap (L \setminus \{\lambda\})$, wprowadzamy dodatkowo reguły $S_0 \to \lambda$ i $S_0 \to S$, a S_0 czynimy symbolem początkowym tej gramatyki na miejsce dotychczasowego symbolu początkowego S. Otrzymany język będzie nadal bezkontekstowy, a jako że jest on identyczny z językiem $L \cap L'$ - kończy to dowód dowodzonego twierdzenia. □

Wniosek 1.

1) j.Lin $\cap$ j.Reg = j.CF,

2) j.Reg $\cap$ j.Reg = j.CF. □

Wniosek 2.

$\mathcal{L}_3 \subset \mathrm{Lin}$.

Dowód.

O tym, że $\mathcal{L}_3 \subseteq \mathrm{Lin}$ już wiemy (patrz początek tego paragrafu). Pozostaje nam udowodnić, że zawieranie to jest właściwe. Dla dowodu nie wprost załóżmy, że $\mathcal{L}_3 = \mathrm{Lin}$. Wtedy jednak języki $L_1 = \{a^n b^n c^k : n\geq1, k\geq1\}$ i $L_2 = \{a^k b^n c^n : n\geq1, k\geq1\}$ byłyby regularne, więc w przecięciu dałyby język bezkontekstowy (z wniosku 1), a język $L_1 \cap L_2 = \{a^n b^n c^n : n\geq1\}$ taki nie jest. □

Wniosek 3.

Klasa $\mathcal{L}_3$ jest właściwym podzbiorem klasy $\mathcal{L}_2$ (tj. $\mathcal{L}_3 \subset \mathcal{L}_2$).

Dowód.

$\mathcal{L}_3 \subset \mathrm{Lin} \subseteq \mathcal{L}_2, \Rightarrow \mathcal{L}_3 \subset \mathcal{L}_2$.

$\nwarrow$ z wn. 2 □

Wniosek 4.

Jeżeli $L_1 \in \mathcal{L}_2$, a $L_2 \in \mathcal{L}_3$, to $L_1 \setminus L_2 \in \mathcal{L}_2$ (tzn. jest językiem bezkontekstowym).

Dowód.

Niech V_T będzie wspólnym alfabetem terminalnym gramatyk generujących te języki. Wtedy $L_1 \setminus L_2 = L_1 \cap (V_T^* \setminus L_2)$ (bo jest to zbiór wszystkich słów nad V_T^* należących do L_1 i zarazem nie należących do L_2). Ponieważ (w myśl powyższego twierdzenia) $L_1 \cap (V_T^* \setminus L_2)$ jest oczywiście językiem regularnym (bo $V_T^* \setminus L_2$ jest językiem regularnym - patrz tw. 6.4), więc $L_1 \setminus L_2$ (jako mu identyczny) jest również językiem regularnym. □

Zastanówmy się jeszcze, co powoduje, że niektóre gramatyki nie dają gramatyk regularnych. W tym celu wprowadźmy następującą definicję:

G r a m a t y k ę $G = \langle V_N, V_T, S, F\rangle$ n a z y w a m y s a m o r o z s z e r z a j ą c ą witw, gdy

$$\exists\, A\in V_N\ \exists\, X,Y\in (V_N\cup V_T)^*:\ A \xrightarrow[G]{*} XAY.$$

Twierdzenie 3.6.

Jeżeli gramatyka bezkontekstowa nie jest samorozszerzająca, to język generowany przez tę gramatykę jest regularny. □

Uwaga.

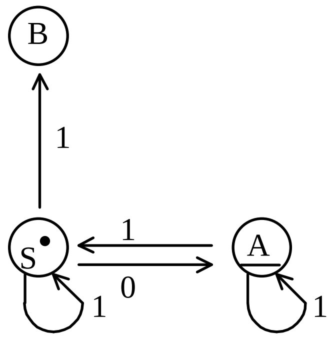
W myśl powyższego twierdzenia, spośród gramatyk bezkontekstowych, tylko gramatyki samorozszerzające mogą nie generować języków regularnych.

5. Gramatyki deterministyczne

W paragrafie tym, mówiąc „gramatyka regularna”, będziemy zawsze mieli na myśli gramatykę regularną w postaci normalnej (z twierdzenia 3.1).
Na ogół dla gramatyki regularnej mogliśmy mieć kilka produkcji o tym samym poprzedniku, a następnikach różniących się jedynie symbolami nieterminalnymi.

Przykład 3. 7 - c. d.

Z sytuacją tą spotykamy się np. w gramatyce z przykładu 3.7, gdzie mieliśmy produkcje $S \to 1S$, $S \to 1B$, $S \to 0A$ oraz $A \to 1A$, $A \to 1S$ i $A \to \lambda$ (patrz rys. 3.24).



Rys. 3.24.

Jest oczywiste, że gramatyki takie są niewygodne, głównie ze względu na niejednoznaczność wywodu słów (słowa te mogą mieć różne drogi końcowe). Dlatego też wprowadzamy pojęcie gramatyki deterministycznej generującej język L, w której każde słowo ma już dokładnie jedną drogę końcową:
Gramatykę regularną $G = \langle V_N, V_T, S, F \rangle$ nazywamy *deterministyczną*, jeśli z faktu, że $\ulcorner A \to aB \urcorner \in F$ i $\ulcorner A \to aC \urcorner \in F$ (dla pewnych $A, B, C \in V_N$ i $a \in V_T$) wynika, że $B = C$. W przeciwnym przypadku gramatykę nazywamy *niedeterministyczną*.
Definicja ta mówi, że gdyby w deterministycznej gramatyce regularnej trafiły się nam przypadkiem dwie reguły produkcji o identycznych poprzednikach i symbolach terminalnych, to właściwie jest to tylko jedna reguła produkcji, bo ma również identyczne symbole nieterminalne po prawej stronie reguł produkcji. Dzięki takiej ich jednoznaczności mamy zdeterminowany wywód dowolnego słowa języka generowanego przez tę gramatykę.

Z kolei, gramatykę regularną $G = \langle V_N, V_T, S, F \rangle$ nazywamy *zupełną*, jeśli:

$$\forall\ A \in V_N\ \forall\ a \in V_T\ \exists\ B \in V_N: \ulcorner A \rightarrow aB \urcorner \in F.$$

Zupełność gramatyki regularnej na jej wykresie poznaje się więc po tym, że z każdego wierzchołka wychodzą strzałki opatrzone wszystkimi symbolami alfabetu terminalnego.

Przykład 3.7 - c.d.

Gramatyka ta jest oczywiście niezupełna. □

Na szczęście jednak, zachodzi następujące twierdzenie.

Twierdzenie 3.7.

Każdą gramatykę regularną można rozszerzyć do gramatyki regularnej i zupełnej generującej ten sam język.

Dowód.

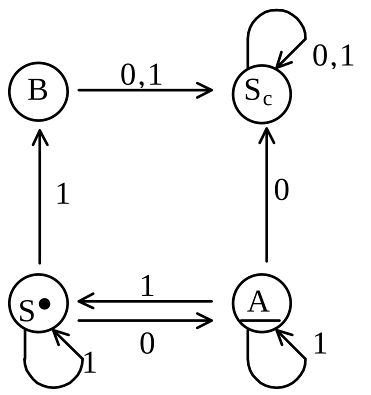
Niech G = $<V_N, V_T, S, F>$ będzie gramatyką regularną. Tworzymy nową gramatykę regularną G' = $<V_N', V_T, S, F'>$, gdzie:

– $V_N' = V_N \cup \{S_e\}$, gdzie S_e jest nowym symbolem nieterminalnym zwanym symbolem ucieczki (jako indeks występuje „e” od angielskiego „escape” = uciekać);
– F' tworzą reguły F powiększone o dwa rodzaje reguł:
  – jeśli dla pewnego nieterminalnego A i terminalnego a w gramatyce G nie ma reguł postaci $A \rightarrow aB$, to dopisujemy regułę $A \rightarrow aS_e$,
  – dopisujemy nadto wszystkie reguły postaci $S_e \rightarrow xS_e$ dla wszystkich symboli $x \in V_T$.

Gwarantuje nam to otrzymanie regularnej gramatyki zupełnej. Widać, że jeśli jakąś drogą dostaniemy się do (nie końcowego!) wierzchołka S_e, to już się z niego nie wydostaniemy. Jasne jest więc, że tak otrzymana gramatyka G' generuje ten sam język, co gramatyka G. □

Przykład 3.7 - c.d.

Uzupełnienie danej gramatyki regularnej G do gramatyki regularnej i zupełnej przedstawia następujący wykres:



Rys. 3.25

Powróćmy do gramatyk deterministycznych. Zachodzi dla nich następujące, główne w tym paragrafie, twierdzenie:

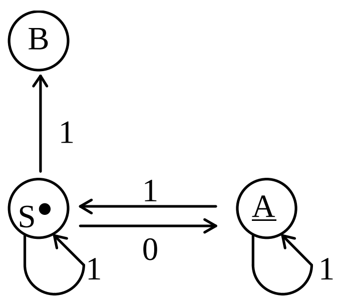
Twierdzenie 3.8 (Myhilla).

Dla każdej gramatyki regularnej (być może niedeterministycznej) istnieje równoważna jej (tzn. generująca ten sam język) gramatyka deterministyczna i zupełna. □

Zamiast dowodu, prześledzimy na konkretnym przykładzie metodę konstrukcji takiej gramatyki.

Przykład 3.9.

Niech G = <V_N, V_T, S, F> będzie gramatyką regularną z przykładu 3.7 (podaną tu jeszcze raz na rys. 3.27). W myśl twierdzenia, w przykładzie tym, skon struujemy równoważną jej deterministyczną i zupełną gramatykę regularną G' = <V_N', V_T', S', F'>.



Rys. 3.26.

Obierzmy S' = {S} $\overset{\text{ozn.}}{=}$ S_0 oraz V_T' = V_T (zawsze tak robimy). Pozostaje nam jeszcze określić zbiory V_N' i F'. Zrobimy to jednocześnie:

– określając bowiem nowe reguły produkcji F' $\overset{\text{ozn.}}{=}$ M (gdzie „M” wzięte jest od „Myhilla”, przy czym fakt, że $\ulcorner A \to aB \urcorner \in$ F' /dla pewnych A, B $\in$ V_N' i a $\in$ V_T'/ oznaczać będziemy M(A,a) = F'(A,a) = B);

– przy okazji będziemy otrzymywać nowe symbole nieterminalne tworząc tym samym zbiór V_N'.

Najpierw otrzymujemy $M(S_0,0) = \{\underline{A}\} = \underline{S_1}$.

Pierwszą równość otrzymujemy tu rozumując następująco: „w gramatyce G będąc w jednym z symboli nieterminalnych zbioru S_0, tj. jedynie w symbolu S, idąc po strzałce 0 dochodzimy jedynie do symbolu nieterminalnego A”; drugą zaś równość otrzymujemy następująco: „ponieważ stan {A} gramatyki G' nie był dotychczas zdefiniowany, więc nadajemy mu nową nazwę S_1”. Tym samym otrzymaliśmy w gramatyce G' regułę produkcji $S_0 \rightarrow 0S_1$ oraz symbol nieterminalny S_1. Ponieważ konstytuujący go symbol A jest końcowy, więc symbol S_1 dodatkowo podkreśliliśmy, co oznacza że w gramatyce G' otrzymaliśmy również regułę $S_1 \rightarrow \lambda$ (czynimy tak zawsze, gdy choć jeden z nieterminałów gramatyki G tworzących nowy nieterminał gramatyki G' jest końcowy).

W identyczny sposób w odniesieniu do symbolu S_0 postępujemy z pozostałymi terminałami (tu jedynie z „1”):

$M(S_0,1) = \{B, S\} = S_2$ (otrzymaliśmy regułę $S_0 \rightarrow 1S_2$ oraz symbol S_2).

Wyżej opisaną procedurę powtarzamy dla wszystkich nowo otrzymanych symboli nieterminalnych gramatyki G', przy czym wykonujemy ją tak długo, dopóki będzie to tylko możliwe. Jeżeli za którymś razem zdarzy się, że jako nowy nieterminał otrzymamy zbiór pusty, to zbiorowi temu nadajemy nazwę S_e (symbolu ucieczki).

Zauważmy, że w następnych krokach konstrukcji gramatyki G' dołączone będą reguły $S_e \rightarrow 0S_e$, $S_e \rightarrow 1S_e$ (a więc z wszystkimi nieterminałami). Symbol S_e rzeczywiście będzie więc spełniał rolę symbolu ucieczki. Jego istnienie spowodowane jest faktem, że (w myśl twierdzenia) przekształcana gramatyka G nie koniecznie musi być zupełna. Dalej otrzymujemy więc po kolei:

$M(S_1, 0) = \emptyset = S_e$ $\qquad$ $M(S_1, 1) = \{\underline{A}, S\} = \underline{S_3}$

(doszły symbole S_3 i S_e oraz produkcje $S_1 \rightarrow 0S_e$, $S_1 \rightarrow 1S_3$ i $S_3 \rightarrow \lambda$);

$M(S_2, 0) = \{\underline{A}\} = \underline{S_1}$ $\qquad$ $M(S_2, 1) = \{B, S\} = S_2$

(doszły produkcje $S_2 \rightarrow 0S_1$ i $S_2 \rightarrow 1S_2$);

$M(S_3, 0) = \{\underline{A}\} = \underline{S_1}$ $\qquad$ $M(S_3, 1) = \{\underline{A}, B, S\} = \underline{S_4}$

(doszedł symbol S_4 oraz produkcje $S_3 \rightarrow 0S_1$, $S_3 \rightarrow 1S_4$ i $S_4 \rightarrow \lambda$);

$M(S_4, 0) = \{\underline{A}\} = \underline{S_1}$ $\qquad$ $M(S_4, 1) = \{\underline{A}, B, S\} = \underline{S_4}$

(doszły produkcje $S_4 \rightarrow 0S_1$ i $S_4 \rightarrow 1S_4$);

$M(S_e, 0) = S_e$ $M(S_e, 1) = S_e$

(dodaliśmy - dające zupełność gramatyki G' - produkcje $S_e \rightarrow 0S_e$ i $S_e \rightarrow 1S_e$).

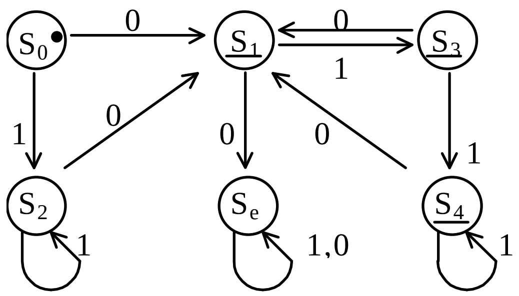
W ten sposób otrzymaliśmy:

- alfabet nieterminalny $V_N' = \{S_0, S_1, S_2, S_3, S_4, S_e\}$
- oraz zbiór reguł produkcji $F' = \{S_1 \rightarrow \lambda, S_3 \rightarrow \lambda, S_4 \rightarrow \lambda,$

$S_0 \rightarrow 0S_1, S_1 \rightarrow 0S_e, S_2 \rightarrow 0S_1, S_3 \rightarrow 0S_1, S_4 \rightarrow 0S_1, S_e \rightarrow 0S_e,$

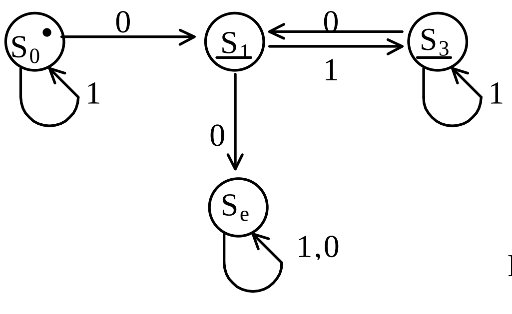
$S_0 \rightarrow 1S_2, S_1 \rightarrow 1S_3, S_2 \rightarrow 1S_2, S_3 \rightarrow 1S_4, S_4 \rightarrow 1S_4, S_e \rightarrow 1S_e\}.$

Jasno widać, że skonstruowana w ten sposób gramatyka jest zarówno deterministyczna jak i zupełna (patrz rys. 3.27).



Rys. 3.27.

Oczywiście, gdybyśmy wyszli nie od gramatyki G, lecz od jej postaci zredukowanej, to od razu otrzymalibyśmy w prostszej formie jej postać deterministyczną i zupełną. Możemy jednak dojść do niej i z bieżącej pozycji, utożsamiając ze sobą wierzchołki S_0 i S_2 oraz S_3 i S_4 (nie zmieni to bowiem ich siły generatywnej).Sytuację te przedstawia rys. 3.28. □



Rys. 3.28. □

Zastosowany tu sposób tworzenia nowej gramatyki nosi nazwę m e t o d y p o t ę g o w e j . Bierze to się z faktu, że nowe wierzchołki tworzą pewien podzbiór zbioru potęgowego starych wierzchołków (stąd nowych wierzchołków może być co najwyżej 2 do potęgi ilość starych wierzchołków).

Twierdzenie Myhilla daje nam w sumie równoważność klasy gramatyk regularnych deterministycznych i zupełnych z klasą gramatyk regularnych. Wynika z tego, że niedeterminizm gramatyki regularnej wcale nie poszerza jej siły generatywnej.