Fogalomtár a Formális nyelvek és automaták tárgyhoz

(A törzsanyaghoz tartozó definíciókat és tételeket * jelöli.)

Definíciók

Univerzális ábécé: Szimbólumokok egy megszámlálhatóan végtelen halmazát univerzális ábécének nevezzük.

* **Ábécé:** Ábécének nevezzük az univerzális ábécé egy tetszőleges véges részhalmazát.

* **Betű:** Az ábécé elemeit betűknek hívjuk.

* **Szó:** Az X ábécé elemeinek egy tetszőleges véges sorozatát az X ábécé feletti szónak nevezzük. Ha X nem lényeges vagy egyértelmű, akkor szóról beszélünk.

* **Nyelv:** X^* valamely részhalmazát (azaz 2^{X^*} valamely elemét) az X ábécé feletti nyelvnek nevezzük.

* **Nyelvosztály (nyelvcsalád):** Nyelvek valamely összességét nyelvosztálynak, nyelvcsaládnak hívjuk.

* **Két szó konkatenációja:** Az $u = t_1 \cdots t_k$ és $v = t'_1 \cdots t'_\ell$ szavak konkatenációja alatt az $uv := t_1 \cdots t_k t'_1 \cdots t'_\ell$ szót értjük. (A két szó egymás utáni leírásával kapott szó.)

* **Szó hatványa:** Legyen u egy szó, nemnegatív egész hatványai $u^0 := \varepsilon$, $u^1 := u$, $u^n := u^{n-1}u$. *(rekurzív definíció)*

* **Szó megfordítása:** Legyen $u = t_1 \cdots t_k$ egy szó, ekkor u megfordítása $u^{-1} := t_k \cdots t_1$.

Homomorfizmus: A $h : X^* \longmapsto Y^*$ konkatenációtartó leképezéseket homomorfizmusnak nevezzük. h konkatenációtartó leképezés, ha tetszőleges $u, v \in X^*$ szó esetén $h(uv) = h(u)h(v)$.

Homomorfizmus nyelvekre való kiterjesztése: $h(L) := \bigcup_{u \in L} \{h(u)\}$.

* **Két nyelv metszete, uniója, különbsége, szimmetrikus differenciája:** A nyelv is egy halmaz (szavak halmaza), ezeket mint halmazokon vett műveleteket értelmezzük.

* **Nyelv komplementere:** Egy X ábécé feletti nyelv komplementerén a halmazelméleti értelemben vett komplementert értjük X^*-ra nézve.

* **Két nyelv konkatenációja:** Legyenek L_1, L_2 nyelvek. Ekkor az L_1 és L_2 nyelvek konkatenációján az $L_1 L_2 := \{uv \mid u \in L_1 \wedge v \in L_2\}$ nyelvet értjük.

* **Nyelv hatványa:** Legyen L egy nyelv, nemnegatív egész hatványai $L^0 := \{\varepsilon\}$, $L^1 := L$, $L^n := L^{n-1}L$. *(rekurzív definíció)*

* **Nyelv lezártja (iteráltja):** Legyen L egy nyelv. $L^* := \bigcup_{i=0}^{\infty} L^i$ az L nyelv lezártja.

Nyelv pozitív lezártja (iteráltja): Legyen L egy nyelv. $L^+ := \bigcup_{i=1}^{\infty} L^i$ az L nyelv pozitív lezártja.

* **Nyelv megfordítása:** Legyen L egy nyelv. $L^{-1} := \{u^{-1} \mid u \in L\}$ az L nyelv megfordítása.

* **Részszó:** v részszava u-nak, ha léteznek olyan w_1, w_2 szavak, melyre $u = w_1 v w_2$.

Szó egy prefixe: v az u szó prefixe, ha van olyan w szó, hogy $u = vw$. v valódi prefix, ha $v \neq \varepsilon, u$.

Szó prefixhalmaza: Legyen u egy szó. $\mathrm{Pre}(u) := \{\, v \mid v \text{ prefixe } u\text{-nak}\}$ az u szó prefixhalmaza.

Szó legfeljebb i hosszúságú prefixhalmaza: $\mathrm{Pre}(u,i) := \mathrm{Pre}(u) \cap X(u)^{\leq i}$.

* **Szó i hosszúságú prefixe:** $\mathrm{pre}(u,i) := \begin{cases} u & \ell(u) \leq i \\ v & v \in \mathrm{Pre}(u) \wedge \ell(v) = i \end{cases}$.

Szó egy suffixe: v az u szó suffixe, ha van olyan w szó, hogy $u = wv$. v valódi suffix, ha $v \neq \varepsilon, u$.

Szó suffixhalmaza: Legyen u egy szó. $\mathrm{Suf}(u) := \{\, v \mid v \text{ suffixe } u\text{-nak}\}$ az u szó suffixhalmaza.

Szó legfeljebb i hosszúságú suffixhalmaza: $\mathrm{Suf}(u,i) := \mathrm{Suf}(u) \cap X(u)^{\leq i}$.

* **Szó i hosszúságú suffixe:** $\mathrm{suf}(u,i) := \begin{cases} u & \ell(u) \leq i \\ v & v \in \mathrm{Suf}(u) \wedge \ell(v) = i \end{cases}$.

Nyelv prefixhalmaza: Legyen L egy nyelv. $\mathrm{Pre}(L) := \bigcup_{u \in L} \mathrm{Pre}(u)$ az L nyelv prefixhalmaza.

Nyelv suffixhalmaza: Legyen L egy nyelv. $\mathrm{Suf}(L) := \bigcup_{u \in L} \mathrm{Suf}(u)$ az L nyelv suffixhalmaza.

* **Reguláris nyelvek:** *(rekurzív definíció)*

- az elemi nyelvek, azaz $\emptyset$, $\{\varepsilon\}$, $\{a\}$ $(a \in U)$
- azon nyelvek, melyek az elemi nyelvekből az unió, konkatenáció és lezárás műveletek véges számú alkalmazásával állnak elő

* **Reguláris kifejezések:** *(rekurzív definíció)*

- az elemi reguláris kifejezések, azaz $\emptyset$, ε, a $(a \in U)$
- ha R_1 és R_2 reguláris kifejezések akkor $(R_1 \cup R_2), (R_1 R_2), R_1^*$ is reguláris kifejezések
- a reguláris kifejezések halmaza a legszűkebb halmaz, melyre a fenti két pont teljesül

* **X ábécé feletti reguláris kifejezések:** *(rekurzív definíció)*

- az elemi reguláris kifejezések, azaz $\emptyset$, ε, a $(a \in X)$
- ha R_1 és R_2 X ábécé feletti reguláris kifejezések akkor $(R_1 \cup R_2), (R_1 R_2), R_1^*$ is X ábécé feletti reguláris kifejezések
- ax X ábécé feletti reguláris kifejezések halmaza a legszűkebb halmaz, melyre a fenti két pont teljesül

* **X ábécé feletti általánosított reguláris kifejezések:** *(rekurzív definíció)*

- az elemi reguláris kifejezések, azaz $\emptyset$, ε, a $(a \in X)$
- ha R_1 és R_2 X ábécé feletti általánosított reguláris kifejezések akkor $(R_1 \cup R_2)$, $(R_1 R_2)$, R_1^*, $(R_1 \cap R_2)$, $\overline{R}_1$ is X ábécé feletti általánosított reguláris kifejezések
- az X ábécé feletti általánosított reguláris kifejezések halmaza a legszűkebb halmaz, melyre a fenti két pont teljesül

* **Reguláris kifejezések szemantikája:** *(rekurzív definíció)*

- az $\emptyset$, ε, a reguláris kifejezések rendre az $\emptyset$, $\{\varepsilon\}$, $\{a\}$ nyelveket reprezentálják
- ha R_1 az L_1 illetve R_2 az L_2 nyelvet reprezentálja, akkor $(R_1 \cup R_2), (R_1 R_2), R_1^*$ rendre az $L_1 \cup L_2, L_1 L_2, L_1^*$ nyelveket reprezentálja.

* X **ábécé feletti általánosított reguláris kifejezések szemantikája:** *(rekurzív definíció)*

- az $\emptyset$, ε, a elemi reguláris kifejezések rendre az $\emptyset$, $\{\varepsilon\}$, $\{a\}$ nyelveket reprezentálják
- ha R_1 az L_1 illetve R_2 az L_2 nyelvet reprezentálja, akkor $(R_1 \cup R_2), (R_1 R_2), R_1^*, (R_1 \cap R_2), \overline{R}_1$ rendre az $L_1 \cup L_2, L_1 L_2, L_1^*, L_1 \cap L_2, \overline{L}_1$ nyelveket reprezentálja.

* **Rekurzívan felsorolható nyelv:** Az L nyelv rekurzívan felsorolható $\Longleftrightarrow$ ha létezik A algoritmus, mely az elemeit felsorolja. Felsoroló algoritmus: Az A algoritmus outputjára szavakat állít elő, s így a nyelv összes szavát (és csak azokat) felsorolja.

* **Parciálisan rekurzív nyelv:** Az L nyelv parciálisan rekurzív $\Longleftrightarrow$ létezik olyan A parciálisan eldöntő algoritmus, melynek inputjára tetszőleges szót helyezve eldönti, benne van-e a nyelvben ($u \in L$ szó esetén *igen* válasszal áll le, míg $u \notin L$ esetén nem terminál, vagy ha terminál, akkor *nem* választ ad).

* **Rekurzív nyelv:** Az L nyelv rekurzív $\Longleftrightarrow$ létezik olyan A eldöntő algoritmus, melynek inputjára egy tetszőleges u szót helyezve eldönti, benne van-e az L nyelvben (mindig terminál, *igen* a válasz, ha u eleme az L nyelvnek, és *nem* a válasz ellenkező esetben).

* **Generatív nyelvtan (grammatika):** A $G = \langle T, N, \mathcal{P}, S \rangle$ négyest nyelvtannak nevezzük, ahol T a terminális, N a nyelvtani (nemterminális) jelek egymástól diszjunkt ábécéje, $\mathcal{P}$ (produkciós) szabályoknak egy véges halmaza, ahol minden $P \in \mathcal{P}$ szabály $p \to q$ alakú, $p, q \in (T \cup N)^*$ és p tartalmaz legalább egy nyelvtani jelet, továbbá $S \in N$, melyet kezdőszimbólumnak nevezünk.

* **Mondatforma:** $(T \cup N)^*$ elemeit mondatformáknak nevezzük.

* **Közvetlen levezetés (nyelvtanban):** Az α mondatformából közvetlenül levezethető a β mondatforma, ha léteznek γ_1, γ_2 mondatformák és $p \to q \in \mathcal{P}$, hogy $\alpha = \gamma_1 p \gamma_2$ és $\beta = \gamma_1 q \gamma_2$. Jelölése: $\alpha \underset{G}{\to} \beta$.

* **Közvetett levezetés (nyelvtanban):** Az α mondatformából közvetetten levezethető a β mondatforma, ha létezik $k \in \mathbb{N}$ és $\gamma_0, \gamma_1, \ldots, \gamma_k$ mondatformák, hogy $\alpha = \gamma_0, \beta = \gamma_k$ és minden $i \in [0, k-1]$ esetén $\gamma_i \underset{G}{\to} \gamma_{i+1}$. Jelölése: $\alpha \overset{*}{\underset{G}{\to}} \beta$. Ha fontos, hogy éppen k lépésben: $\alpha \overset{k}{\underset{G}{\to}} \beta$. Ekkor k a levezetés hossza.

* **Nyelvtan által generált nyelv:** A $G = \langle T, N, \mathcal{P}, S \rangle$ nyelvtan által generált nyelv $L(G) := \{u \in T^* \mid S \overset{*}{\underset{G}{\to}} u\}$.

Ekvivalens nyelvtanok: A G_1 és G_2 nyelvtanok ekvivalensek ($G_1 \sim G_2$), ha $L(G_1) = L(G_2)$.

Kváziekvivalens nyelvtanok: A G_1 és G_2 nyelvtanok kváziekvivalensek ($G_1 \underset{\text{kv}}{\sim} G_2$), ha $L(G_1) \setminus \{\varepsilon\} = L(G_2) \setminus \{\varepsilon\}$.

* **Nyelvtanok típusai:** Egy G nyelvtan i. típusú ($i \in \{0, 1, 2, 3\}$), ha a szabályai a táblázatban megadott alakúak *(Alaptípus szabályai oszlop)*.

* **Nyelvtanok megszorított típusai:** Egy G nyelvtan megszorított i. típusú ($i \in \{1, 2, 3\}$), ha a szabályai a táblázatban megadott alakúak *(Megszorított típus szabályai oszlop)*.

* **Környezetfüggetlen nyelvtan:** 2. típusú nyelvtan.

* **Környezetfüggő nyelvtan:** Megszorított 1. típusú nyelvtan.

* **Normálformák:** Egy G nyelvtan i-es normálformájú ($i = 1, 2, 3$), ha a szabályai a táblázatban megadott alakúak *(Normálforma szabályai oszlop)*. Az táblázatban megadott 1. típusú normálforma elnevezése Kuroda normálforma, a 2. típusúé Chomsky normálforma.

Típus	Alaptípus szabályai	Megszorított típus szabályai	Normálforma szabályai
0.	nincs további megkötés	$p \to q$, ahol $q \in (T \cup N)^+$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	$AB \to A$ $(A, B \in N)$; $BA \to A$ $(A, B \in N)$; +Kuroda NF szabálysémái
1.	$p \to q$, ahol $\ell(p) \leq \ell(q)$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	$\gamma_1 A \gamma_2 \to \gamma_1 q \gamma_2$, ahol $\gamma_1, \gamma_2 \in (T \cup N)^*$, $A \in N, q \in (T \cup N)^+$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	(Kuroda) $A \to a$ $(A \in N, a \in T)$; $A \to BC$ $(A, B, C \in N)$; $AB \to AC$ $(A, B, C \in N)$; $BA \to CA$ $(A, B, C \in N)$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán
2.	$A \to q$, ahol $A \in N, q \in (T \cup N)^*$	$A \to q$, ahol $A \in N, q \in (T \cup N)^+$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	(Chomsky) $A \to a$ $(A \in N, a \in T)$; $A \to BC$ $(A, B, C \in N)$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán
3.	$A \to uB$ vagy $A \to u$, ahol $A, B \in N$, és $u \in T^*$	$A \to aB$ vagy $A \to a$ ahol $A, B \in N, a \in T$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	$A \to \varepsilon$ vagy $A \to aB$, ahol $A, B \in N$ és $a \in T$

(A táblázatban konvencionálisan S a kezdőszimbólum)

Greibach normálforma Legyen $G = \langle T, N, \mathcal{P}, S \rangle$ egy nyelvtan. A G nyelvtan Greibach-féle normálformájú, ha szabályai a következő alakúak: $S \to \varepsilon$, ahol S kezdőszimbólum és ha van ilyen szabály, akkor S nem fordul elő szabály jobboldalán; $A \to aQ$, ahol $A \in N, a \in T$, és $Q \in N^*$.

Zsákutca: Olyan nyelvtani jel, melyből az adott 2. típusú nyelvtanban nem vezethető le terminális szó.

Nem elérhető nyelvtani jel: Olyan nyelvtani jel, mely az adott 2. típusú nyelvtanban semmilyen, a kezdőszimbólumból történő levezetésben nem szerepel.

Zsákutcamentes nyelvtan: 2. típusú nyelvtan, melynek semmelyik nyelvtani jele se zsákutca.

Összefüggő nyelvtan: 2. típusú nyelvtan, mely nem tartalmaz nem elérhető nyelvtani jelet.

Redukált nyelvtan: Zsákutcamentes és összefüggő 2. típusú nyelvtan.

* **Láncszabály:** Egy $G = \langle T, N, \mathcal{P}, S \rangle$ nyelvtanban $p \to q \in \mathcal{P}$ láncszabály, ha $p, q \in N$.

* **Láncszabálymentes nyelvtan:** Egy nyelvtan láncszabálymentes, ha nincs láncszabálya.

* **Epszilonszabály:** Egy $G = \langle T, N, \mathcal{P}, S \rangle$ nyelvtanban $p \to q \in \mathcal{P}$ epszilonszabály, ha $q = \varepsilon$.

* **Korlátozott epszilonszabály (KeS):** Egy $G = \langle T, N, \mathcal{P}, S \rangle$ nyelvtanra teljesül a korlátozott epszilonszabály, ha nincsenek epszilonszabályai az $S \to \varepsilon$ szabály esetleges kivételével, de

ebben az esetben minden $p \to q \in \mathcal{P}$ szabályra $q \in (T \cup N \setminus \{S\})^*$ (azaz semelyik szabály jobboldala sem tartalmazza a kezdőszimbólumot).

* **Epszilonmentes nyelvtan:** Egy nyelvtan epszilonmentes, ha teljesül rá a korlátozott epszilon-szabály.

Nyelvtani transzformáció: A Φ nyelvtani transzformáció olyan eljárás, mely egy G nyelvtanból, egy másik, $\Phi(G)$ nyelvtant készít.

Ekvivalens nyelvtani transzformáció: Egy Φ nyelvtani transzformáció ekvivalens nyelvtani transzformáció, ha minden G nyelvtanra $G \sim \Phi(G)$.

Kváziekvivalens nyelvtani transzformáció: Egy Φ nyelvtani transzformáció kváziekvivalens nyelvtani transzformáció, ha minden G nyelvtanra $G \underset{\text{kv}}{\sim} \Phi(G)$.

Típusmegőrző nyelvtani transzformáció: Legyen $\mathcal{G}_\pi$ a π típusú nyelvtanok osztálya. Egy Φ nyelvtani transzformáció megőrzi a π típust, amennyiben minden $G \in \mathcal{G}_\pi$ esetén $\Phi(G) \in \mathcal{G}_\pi$.

* **Nyelvek típusai:** Egy L nyelv (megszorított) i. típusú, ha létezik olyan (megszorított) i. típusú nyelvtan, mely L-et generálja. *(l. megszorítási tétel).*

* **Reguláris nyelv:** 3. típusú nyelv. *(l. Kleene tétel)*

* **Környezetfüggetlen nyelv:** 2. típusú nyelv.

* **Környezetfüggő nyelv:** 1. típusú nyelv.

Nyelvi operátorra zárt nyelvcsalád: Legyen Ψ n-változós nyelvi operátor, azaz ha $L_1, \ldots, L_n$ nyelvek, akkor $\Psi(L_1, \ldots, L_n)$ is legyen nyelv. Az $\mathcal{L}$ nyelvcsalád zárt a Ψ nyelvi operátorra, ha $L_1, \ldots, L_n \in \mathcal{L}$ esetén $\Psi(L_1, \ldots, L_n) \in \mathcal{L}$.

BNF: Egy Backus-Naur forma (BNF), a következő építőkövekből áll: ⟨szöveg⟩,::=,{,},|, egyéb karakterek. ⟨szöveg⟩ a fogalmak, az egyéb karakterek a terminálisok. Egy BNF szabály baloldalán pontosan 1 fogalom áll, jobboldalán fogalmak, terminálisok, {,},| jelek sorozata áll, úgy hogy a sorozat helyesen zárójelezett a {,} zárójelekkel. A két oldalt a ::= szimbólum választja el egymástól.

Zárójelek és alternatívák jelentése: Az $F ::= \gamma_1\{\alpha_1|\cdots|\alpha_n\}\gamma_2$ formula az $F ::= \gamma_1\alpha_1\gamma_2$, $\ldots$, $F ::= \gamma_1\alpha_n\gamma_2$ formulákat reprezentálja.

Szemantikája: Adott BNF szabályok egy halmaza (zárójelek és alternatívák nélkül). Az α mondatformából (terminálisok és fogalmak sorozata) közvetlenül levezethető a β mondatforma, ha léteznek γ_1, γ_2, ξ mondatformák és $F ::= \xi$ BNF szabály, hogy $\alpha = \gamma_1 F \gamma_2$ és $\beta = \gamma_1\xi\gamma_2$. A közvetett levezetést az eddigiekhez hasonlóan definiáljuk. Egy adott fogalom azon terminális sorozatok halmazát reprezentálja, mely belőle, mint 1 hosszúságú mondatformából közvetetten levezethető.

EBNF: A BNF-hez képest két további jelölést használunk. @γ, ahol γ egy fogalom, terminális, vagy csoport a 0, 1, 2, stb. hosszuságú, csak γ-ból álló sorozatokat reprezentálja. γ_k^n, ahol $0 \leq k \leq n$ természetes egész számok és γ egy fogalom, terminális, vagy csoport a $k, k+1, \ldots, n$ hosszúságú γ sorozatokat reprezentálja. *Szemantika:* mint a BNF-nél.

Szintaxisgráf: Szintaxisgráf alatt olyan irányított gráfot értünk, mely a következő tulajdonságokkal rendelkezik. Egyetlen forrása és egyetlen nyelője van. Az élek címkézetlenek, a szögpontok lehetnek címkézettek és címkézetlenek. A címkéknek két fajtája van, az egyik téglalap, a másik ellipszis alakú. Ezen felül a gráfnak van egy neve. A téglalap alakú címkéket és a gráf nevét fogalmaknak, az ellipszis alakú címkéket terminálisoknak nevezzük.

Szemantikája: Adott szintaxisgráfok egy halmaza. Az α mondatformából (terminálisok és fogalmak sorozata) közvetlenül levezethető a β mondatforma, ha léteznek γ_1, γ_2, ξ mondatformák és F fogalom, hogy $\alpha = \gamma_1 F \gamma_2$, $\beta = \gamma_1 \xi \gamma_2$, továbbá létezik olyan szintaxisgráf, melynek címkéje F és a gráfban létezik irányított út[1] a forrásból a nyelőbe, mely mentén a címkézett szögpontok címkéi az irányított út[1] által meghatározott sorrendben éppen ξ-t adják. A közvetett levezetést az eddigiekhez hasonlóan definiáljuk,. Egy adott fogalom azon terminális sorozatok halmazát reprezentálja, mely belőle, mint 1 hosszúságú mondatformából közvetetten levezethető.

* **n-verem:** n-verem alatt a következő $(2n+5)$-öst értjük:

$\mathcal{V} = \langle Q, T, \Sigma_1, \ldots, \Sigma_n, \delta, q_0, \sigma_1, \ldots, \sigma_n, F \rangle$, ahol

Q	az állapotok halmaza (ez legyen véges halmaz),
T	egy ábécé, a bemenő ábécé,
Σ_i	az i-edik verem ábécéje,
δ	az állapotátmeneti függvény,
$q_0 \in Q$	kezdőállapot,
σ_i	az i. verem kezdőszimbóluma, ahol $\sigma_i \in \Sigma_i$,
$F \subseteq A$	a végállapotok halmaza.

Az állapotátmeneti függvény $\delta : Q \times (T \cup \{\varepsilon\}) \times \Sigma_1 \times \cdots \times \Sigma_n \rightarrow 2^{Q \times \Sigma_1^* \times \cdots \times \Sigma_n^*}$ alakú függvény, melyre megköveteljük, hogy értékkészlete véges halmazokból álljon.

* **Veremautomata:** 1-verem.

* **Konfiguráció:** Konfigurációnak nevezzük azoknak az adatoknak az összességét, melyektől a gép elkövetkezendő működése függ.

A konfigurációk a következő alakú $(n+2)$-esek: $[q, v, \alpha_1, \ldots, \alpha_n]$, ahol:
- q az aktuális állapot,
- v az input szó még elolvasatlan része,
- α_i az i-edik verem tartalma.

* **Közvetlen konfigurációátmenet:** Közvetlen konfigurációátmenetről beszélünk, ha $\mathcal{V}$ egy lépésben vált át egyik konfigurációból a másikba, azaz $[q, u, \alpha_1, \ldots, \alpha_n] \underset{\mathcal{V}}{\dashv} [q', v, \beta_1, \ldots, \beta_n]$ akkor és csak akkor, ha van olyan $t \in T \cup \{\varepsilon\}$, hogy $u = tv$, továbbá minden $i \in [1, n]$ esetén van olyan $\sigma_i \in \Sigma_i$ és $\gamma_i, \tau_i \in \Sigma_i^*$, amelyekre $\alpha_i = \sigma_i \gamma_i, \beta_i = \tau_i \gamma_i$, valamint $(q', \tau_1, \ldots, \tau_n) \in \delta(q, t, \sigma_1, \ldots, \sigma_n)$. Ha itt $t = \varepsilon$, akkor ε-mozgásról beszélünk.

* **Közvetett konfigurációátmenet:** A közvetett konfigurációátmenet a közvetlen átmenet reflexív, tranzitív lezártja. Jelölése: $\overset{*}{\underset{\mathcal{V}}{\dashv}}$.

* **Kezdőkonfiguráció:** Az u szóhoz tartozó kezdőkonfiguráció: $[q_0, u, \sigma_1, \ldots, \sigma_n]$.

* **Termináló konfiguráció:** Egy K konfiguráció termináló konfiguráció, ha nincs rákövetkezője, azaz $\nexists\, K'$ konfiguráció, hogy $K \underset{\mathcal{V}}{\dashv} K'$.

* **Végállapottal elfogadó konfiguráció:** Végállapottalelfogadó egy $[q, \varepsilon, \beta_1, \ldots, \beta_n]$ konfiguráció,

[1] Itt az út egy ponton többször is áthaladhat. Ez valójában a gráfelméleti *séta* fogalomnak felel meg.

ha $q \in F$.

* **Üres veremmel elfogadó konfiguráció:** Üres veremmel elfogadó egy $[q, \varepsilon, \beta_1, \ldots, \beta_n]$ konfiguráció, ha $\beta_1 = \varepsilon$.

* **Szó végállapottal elfogadása:** A $\mathcal{V}$ n-verem végállapottal elfogadja az u szót, ha létezik átmenet az u szóhoz tartozó kezdőkonfigurációból végállapottal elfogadó konfigurációba.

* **Szó üres veremmel elfogadása:** A $\mathcal{V}$ n-verem üres veremmel elfogadja az u szót, ha létezik átmenet az u szóhoz tartozó kezdőkonfigurációból üres veremmel elfogadó konfigurációba.

* **Végállapottal elfogadott nyelv:** A $\mathcal{V}$ által végállapottal elfogadott nyelv: $L^{\mathrm{F}}(\mathcal{V}) = \big\{ u \in T^* \;\big|\; [q_0, u, \sigma_1, \ldots, \sigma_n] \underset{\mathcal{V}}{\overset{*}{\dashv}} [q, \varepsilon, \beta_1, \beta_2, \ldots, \beta_n]$ valamely $q \in F$-re$\big\}$.

* **Üres veremmel elfogadott nyelv:** A $\mathcal{V}$ által üres veremmel elfogadott nyelv: $L^{\varepsilon}(\mathcal{V}) = \big\{ u \in T^* \;\big|\; [q_0, u, \sigma_1, \ldots, \sigma_n] \underset{\mathcal{V}}{\overset{*}{\dashv}} [q, \varepsilon, \varepsilon, \beta_2, \ldots, \beta_n] \;\big\}$.

* **Determinisztikus n-verem:** Egy adott $\mathcal{V}$ n-verem determinisztikus, ha minden konfigurációnak legfeljebb 1 rákövetkezője van.

  Ekvivalens definíció:

  Egy adott $\mathcal{V}$ n-verem determinisztikus, ha a következő két feltétel teljesül:

  - Minden $(q, t, \sigma_1, \ldots, \sigma_n) \in D_\delta$ esetén $\big| \delta(q, t, \sigma_1, \ldots, \sigma_n) \big| \leq 1$,
  - $\delta(q, \varepsilon, \sigma_1, \ldots, \sigma_n) \neq \emptyset$ esetén minden $t \in T$-re $\big| \delta(q, t, \sigma_1, \ldots, \sigma_n) \big| = 0$.

* **Determinisztikus veremautomata:** determinisztikus 1-verem.

* **Véges, ε-átmenetes nemdeterminisztikus automata (εNDA):** 0-verem

* **Véges, nemdeterminisztikus automata (NDA):** 0-verem, melyre $D_\delta \subseteq \{(q, t) \,|\, q \in Q,\, t \in T\}$.

Véges, parciális determinisztikus automata (PDA): Determinisztikus 0-verem, melyre $D_\delta \subseteq \{(q, t) \,|\, q \in Q,\, t \in T\}$.

* **Véges determinisztikus automata (VDA):** Determinisztikus 0-verem, melyre $D_\delta = \{(q, t) \,|\, q \in Q,\, t \in T\}$.

* **Állapotátmeneti függvény általánosítása:** Legyen $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ egy véges determinisztikus automata. Rekurzívan definiáljuk $\delta(q, u)$ értékét ($q \in Q$, $u \in T^*$). $\delta(q, \varepsilon) := q$, $\delta(q, t)$ már definált ($t \in T$). $\delta(q, ut) := \delta(\delta(q, u), t)$ ($u \in T^*$, $t \in T$).

* **VDA által felismert nyelv általánosított állapotátmeneti függvénnyel megadva:** Legyen $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ egy véges determinisztikus automata. Az $\mathcal{A}$ által felismert (elfogadott) nyelv a következő: $L(\mathcal{A}) := \{ u \in T^* \;\big|\; \delta(q_0, u) \in F \}$.

Automata megadása táblázattal: A sorok megfelelnek Q elemeinek, az oszlopok T (általánosabb automata esetén $X = T \cup \{\varepsilon\}$ vagy $X = \mathcal{R}(T)$) elemeinek. A $q \in Q$ sornak és $t \in T$ (vagy általánosabban $t \in X$) oszlopnak megfelelő cella tartalma $\delta(q, t)$. A kezdőállapotot a $\rightarrow$, a végállapotokat $\leftarrow$ szimbólummal megjelöljük.

* **Automata megadása átmenetdiagrammal:** Irányított gráf, ahol a szögpontok és az élek is címkézettek. A szögpontok megfelelnek Q elemeinek (a kezdőállapotot a $\rightarrow$ szimbólummal megjelöljük, F elemeit bekarikázzuk), míg az élek T (általánosabb automata esetén $X = T \cup \{\varepsilon\}$ vagy $X = \mathcal{R}(T)$) elemeinek. $q \in Q$-ból vezet $t \in T$ (vagy általánosabban $t \in X$) címkéjű irányított él $q' \in Q$-ba, akkor és csak akkor, ha $q' = (\in)\delta(q, t)$.

Általánosított szekvenciális automata: $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ általánosított szekvenciális automata, ha Q egy véges halmaz, az állapotok halmaza, T egy véges halmaz, az inputszavak ábécéje, $q_0 \in Q$ a kezdőállapot, $F \subseteq Q$ a végállapotok halmaza, továbbá az állapotátmeneti függvény $\delta : Q \times \mathcal{R}(T) \rightarrow 2^Q$ alakú függvény, melyre megköveteljük, hogy véges tartójú legyen, azaz minden $q \in Q$-hoz csak véges sok $R \in \mathcal{R}(T)$ esetén teljesül, hogy $\delta(q, R) \neq \emptyset$.

Általánosított szekvenciális automata által elfogadott szó: Az $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ általánosított szekvenciális automata elfogadja az $u \in T^*$ szót $\Leftrightarrow$ $\exists n \in \mathbb{N}$, $u_1, \ldots, u_n \in T^*$, $R_1, \ldots R_n \in \mathcal{R}(T)$ és $q_1, \ldots, q_n \in Q$, melyre $u = u_1 \cdots u_n$, továbbá minden $1 \leq i \leq n$ esetén $u_i \in L(R_i)$, $q_i \in \delta(q_{i-1}, R_i)$ és $q_n \in F$.

Átmenetdiagrammos megadás esetén: $\mathcal{A}$ elfogadja az $u \in T^*$ szót, ha van irányított út[1] q_0-ból valamely F-beli állapotba, mely út[1] mentén az élek címkéje az adott sorrendben $R_1, \ldots, R_n$ és $u \in L(R_1 R_2 \cdots R_n)$.

Általánosított szekvenciális automata által felismert nyelv: $L(\mathcal{A}) = \{u \in T^* \,|\, \mathcal{A}$ elfogadja u-t$\}$.

* **Minimális automata:** Valamely $L \in \mathcal{L}_3$-hoz adott minimális állapotszámú véges, determinisztikus automatát L minimális automatájának nevezzük.

* **Összefüggő automata:** Egy $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ véges, determinisztikus automata összefüggő, ha minden $q \in Q$ esetén létezik $u \in T^*$ szó, hogy $\delta(q_0, u) = a$.

* **Automata állapotra vonatkozó maradéknyelve:** Legyen $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ egy véges determinisztikus automata. Az $\mathcal{A}$ automata $q \in Q$-ra vonatkozó maradéknyelve $L(\mathcal{A}, q) := \{v \in T^* \,|\, \delta(q, v) \in F\}$.

* **Automata ekvivalens állapotai:** Legyen $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ egy véges determinisztikus automata. $q \sim q' \;\Leftrightarrow\; L(\mathcal{A}, q) = L(\mathcal{A}, q')$ $\big(\Leftrightarrow\; \forall u \in T^* : (\delta(q, u) \in F \;\Leftrightarrow\; \delta(q', u) \in F)\big)$.

Automaták ekvivalens állapotai: Legyenek $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ és $\mathcal{A}' = \langle Q', T, \delta, q_0', F' \rangle$ véges determinisztikus automaták. $q \sim q' \;\Leftrightarrow\; L(\mathcal{A}, q) = L(\mathcal{A}', q')$ $(q \in \mathcal{A},\ q' \in \mathcal{A}')$.

Automaták ekvivalenciája: Legyenek $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ és $\mathcal{A}' = \langle Q', T, \delta, q_0', F' \rangle$ véges determinisztikus automaták. $\mathcal{A} \sim \mathcal{A}' \;\Leftrightarrow\; q_0 \sim q_0'$.

* **Automata faktorautomatája:** Legyen $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ egy véges determinisztikus automata. $\mathcal{A}$ faktorautomatája $\mathcal{A}/_\sim := \langle \{C_q\}_{q \in Q}, T, \delta', C_{q_0}, \mathcal{F} \rangle$, ahol

- C_q az q-val ekvivalens állapotok osztálya, melynek reprezentánsa q,
- $\mathcal{F} = \{C_q \;\big|\; q \in F\}$,
- $\delta'(C_q, t) = C_{\delta(q,t)}$, azaz δ'-t egy tetszőleges reprezentánssal definiáljuk.

* **Redukált automata:** Egy adott $\mathcal{A}$ véges, determinisztikus automata redukált automata, ha minden $q, q' \in A$ esetén $q \sim q' \;\Leftrightarrow\; q = q'$, azaz nincsenek különböző ekvivalens állapotai.

* **Automata i-ekvivalens állapotai:** Legyen $\mathcal{A} = \langle Q, T, \delta, q_0, F\rangle$ egy véges determinisztikus automata. Azt mondjuk, hogy $q \overset{i}{\sim} q'$ (q i-ekvivalens q'-vel), ha minden $u \in T^{\leq i}$ esetén $\big(\delta(q,u) \in F \;\Leftrightarrow\; \delta(q',u) \in F\big)$ $(i \geq 0)$.

Automaták izomorfiája: Legyenek $\mathcal{A}_i = \langle Q_i, T, \delta_i, q_0^{(i)}, F_i\rangle$ $(i = 1, 2)$ véges, determinisztikus automaták. Ekkor $\mathcal{A}_1$ és $\mathcal{A}_2$ izomorfak, ha létezik $\varphi : Q_1 \to Q_2$ kölcsönösen egyértelmű ráképezés, melyre a következők teljesülnek:
- $\varphi(q_0^{(1)}) = q_0^{(2)}$,
- $\varphi(F_1) = F_2$,
- δ-t megőrzi, azaz minden $q_1 \in Q_1$ és minden $t \in T$ esetén $\varphi(\delta_1(q_1,t)) = \delta_2(\varphi(q_1), t)$.

Direkt szorzat automata: Legyenek $\mathcal{A}_i = \big\langle Q_i, T, \delta_i, q_0^{(i)}, F_i\big\rangle$, véges, determinisztikus automaták, $i = 1, 2$. Az $\mathcal{A}_1 \times \mathcal{A}_2 = \langle Q_1 \times Q_2, T, \delta_1 \times \delta_2, (q_0^{(1)}, q_0^{(2)}), F_\times\rangle$ automatát direkt szorzat automatának hívjuk. $\mathcal{A}_1 \times \mathcal{A}_2$ működése komponensenként párhuzamosan történik, amit a $\delta_1 \times \delta_2$ jelöléssel fejezünk ki. Formálisan: $(\delta_1 \times \delta_2)\big((q_1, q_2), t\big) = \big(\delta_1(q_1,t), \delta_2(q_2,t)\big)$. A végállapotok halmaza feladatonként változhat.

Például legyen $\odot$ a $\cap, \setminus, \triangle$ műveletek közül az egyik. Feladat: konstruálni egy $\mathcal{A}_\odot$ automatát, melyre fennáll, hogy $L(\mathcal{A}_\odot) = L(\mathcal{A}_1) \odot L(\mathcal{A}_2)$. Ekkor
- $F_\cap := F_1 \times F_2$,
- $F_\setminus := F_1 \times (Q_2 \setminus F_2)$,
- $F_\triangle := \big(F_1 \times (Q_2 \setminus F_2)\big) \cup \big((Q_1 \setminus F_1) \times F_2\big)$.

Knuth-Morris-Pratt (KMP) automata: Legyen $m \in T^*$ egy szó. Az $m = m_1 m_2 \cdots m_{\ell(m)}$ mintához tartozó $\mathcal{A}^m$ Knuth-Morris-Pratt automata (vagy röviden KMP automata) a következő. $\mathcal{A}^m = \langle \{q_i\}_{0 \leq i \leq \ell(m)}, T, \delta^m, q_0, \{q_{\ell(m)}\}\rangle$, ahol

$$\delta^m(q_i, x) = q_j \iff j = \begin{cases} \ell(m) & i = \ell(m) \\ \max\{\ell(w) \mid w \in \mathrm{Pre}(m) \cap \mathrm{Suf}(m_1 \cdots m_i x)\} & i < \ell(m) \end{cases}.$$

* **Nyelv szóra vonatkozó maradéknyelve:** Egy L nyelv $p \in T(L)^*$-ra vonatkozó maradéknyelve $L_p := \{v \in T(L)^* \;\big|\; pv \in L\}$.

Összes levezetések gráfja: Legyen $G = \big\langle T, N, \mathcal{P} = \{p_1 \to q_1, \ldots, p_n \to q_n\}, S\big\rangle$ tetszőleges nyelvtan. G összes levezetéseinek gráfja olyan végtelen, irányított gráf, melynek ponjtai $(T \cup N)^*$ elemeinek felelnek meg. α-ból β-ba van i, j címkéjű $(i \geq 1, j \geq 1)$ él, ha $\alpha \underset{G}{\to} \beta$ és $\alpha = \gamma_1 p_j \gamma_2$, $\beta = \gamma_1 q_j \gamma_2$, $i = \ell(\gamma_1) + 1$.

* **Szintaxisfa:** Legyen $G = \big\langle T, N, \mathcal{P}, S\big\rangle$ tetszőleges 2-es típusú nyelvtan. A t nemüres fát G feletti szintaxisfának nevezzük, ha megfelel a következő tulajdonságoknak:
- pontjai $T \cup N \cup \{\varepsilon\}$ elemeivel vannak címkézve.
- belső pontjai N elemeivel vannak címkézve.
- ha egy belső pont címkéje X, a közvetlen leszármazottjainak címkéi pedig balról jobbra olvasva $X_1, X_2, \ldots, X_k$, akkor $X \to X_1 X_2 \ldots X_k \in \mathcal{P}$.
- az ε-nal címkézett pontoknak nincs testvére.

Legbal levezetés: A legbal levezetés olyan levezetés, hogy ha a levezetés folyamán a mondatforma i. betűjén helyettesítés történik, akkor a korábbi pozíciókat $(1., \ldots, i{-}1.)$ a levezetés már a további lépésekben nem érinti, azok változatlanul maradnak.

[1]Itt az út egy ponton többször is áthaladhat. Ez valójában a gráfelméleti *séta* fogalomnak felel meg.

Legjobb levezetés: A legjobb levezetés olyan levezetés, hogy ha a levezetés folyamán a mondatforma hátulról i. betűjén helyettesítés történik, akkor a későbbi pozíciókat (hátulról $1., \dots, i-1.$) a levezetés már a további lépésekben nem érinti, azok változatlanul maradnak.

Legbal mondatforma: Valamely $L(G)$-beli szó legbal levezetése során előforduló mondatforma.

Legjobb mondatforma: Valamely $L(G)$-beli szó legjobb levezetése során előforduló mondatforma.

Egyértelmű nyelvtan: $G \in \mathcal{G}_2$ egyértelmű nyelvtan, ha minden $u \in L(G)$-nek pontosan egy szintaxisfája létezik.

Egyértelmű nyelv: Létezik 2. típusú egyértelmű nyelvtan, ami generálja.

Lényegesen nem egyértelmű nyelv: Nem létezik 2. típusú egyértelmű nyelvtan, ami generálja.

* **Szintaktikus elemzések alapfeladata:** Legyen adva egy $G = \langle T, N, \mathcal{P}, S \rangle \in \mathcal{G}_2$ második típusú nyelvtan és egy $u \in T^*$ szó. Az elemzési algoritmusok feladata azt eldönteni, hogy u szó eleme-e $L(G)$-nek és ha igen, akkor felépíteni u egy G feletti szintaxisfáját.

* **Felülről lefelé elemzés:** A szintaxisfát a gyökértől, azaz a kezdőszimbólumtól próbálja felépíteni.

* **Alulról felfelé elemzés:** A szintaxisfát a levelektől, azaz az elemzendő szótól próbálja felépíteni.

LL(k)nyelvtan: A G 2-es típusú nyelvtan $LL(k)$ nyelvtan, ha tetszőleges

$S \xrightarrow[G,\mathrm{lb}]{*} vA\alpha_1 \xrightarrow[G,\mathrm{lb}]{} v\gamma_1\alpha_1 \xrightarrow[G]{*} vw_1$ és $S \xrightarrow[G,\mathrm{lb}]{*} vA\alpha_2 \xrightarrow[G,\mathrm{lb}]{} v\gamma_2\alpha_2 \xrightarrow[G]{*} vw_2$ levezetések esetén abból, hogy $\mathrm{pre}(w_1, k) = \mathrm{pre}(w_2, k)$ következik, hogy $\gamma_1 = \gamma_2$.

Nyél: Alulról felfelé elemzés esetén az olyan visszahelyettesíthető részre, mely egy legjobb mondatforma valamely legjobb levezetésének utolsó lépésében áll elő, a nyél elnevezést használjuk.

LR(k) nyelvtan: A G 2-es típusú nyelvtan $LR(k)$ nyelvtan, ha tetszőleges

$S \xrightarrow[G,\mathrm{lj}]{*} \alpha_1 A v_1 \xrightarrow[G,\mathrm{lj}]{} \alpha_1\gamma_1 v_1 \xrightarrow[G]{} w_1 v_1$ és $S \xrightarrow[G,\mathrm{lj}]{*} \alpha_2 B v_2 \xrightarrow[G,\mathrm{lj}]{} \alpha_2\gamma_2 v_2 \xrightarrow[G]{} w_2 v_2$ legjobb levezetések esetén abból, hogy $\alpha_1\gamma_1\,\mathrm{pre}(v_1, k)$ és $\alpha_2\gamma_2\,\mathrm{pre}(v_2, k)$ valamelyike kezdőszelete a másiknak, következik, hogy $\alpha_1 = \alpha_2$, $A = B$ és $\gamma_1 = \gamma_2$.

Tételek

* **Tétel:** Nem minden nyelv írható le nyelvtannal.

* *(Church-tézis)* Minden valamilyen konstruktív módon megadható nyelv leírható nyelvtannal.

* **Tétel:** *(Megszorítási tétel)* $\mathcal{L}_{\text{megszi}} = \mathcal{L}_i$ $(i = 1, 2, 3)$.

* **Tétel:** *(Normálforma tétel)* $\mathcal{L}_{\text{nfi}} = \mathcal{L}_i$ $(i = 0, 1, 2, 3)$.

  $i = 1$ esetén Kuroda normálforma tétel, $i = 2$ esetén Chomsky normálforma tétel a neve.

* **Tétel:** *(Greibach NF tétel)* Minden $G \in \mathcal{G}_2$ nyelvtanhoz létezik G' Greibach normálformájú nyelvtan, melyre $G' \sim G$.

* **Tétel:** *(Zártsági tétel)* Az $\mathcal{L}_i$ $(i = 0, 1, 2, 3)$ nyelvosztályok zártak az unió, konkatenáció és a lezárás műveletekre.

* **Tétel:** $\mathcal{L}_0 \subseteq \mathcal{L}_{\text{RekFel}}$, $\mathcal{L}_0 \subseteq \mathcal{L}_{\text{ParcRek}}$, $\mathcal{L}_1 \subseteq \mathcal{L}_{\text{Rek}}$,

* **Tétel:** $\mathcal{L}_{\text{VDA}} = \mathcal{L}_{\text{PDA}} = \mathcal{L}_{\text{NDA}} = \mathcal{L}_{\varepsilon\text{NDA}} (= \mathcal{L}_{0\text{V}}) = \mathcal{L}_3$

Tétel: Legyen $\mathcal{A}$ egy véges, determinisztikus automata, ekkor $\mathcal{A}/_\sim$ redukált és $L(\mathcal{A}/_\sim) = L(\mathcal{A})$.

Tétel: *(Izomorfia tétel)* Legyenek $\mathcal{A}_1$ és $\mathcal{A}_2$ összefüggő, redukált és egymással ekvivalens automaták. Ekkor $\mathcal{A}_1 \cong \mathcal{A}_2$.

* **Tétel:** *(Kleene tétel)* $\mathcal{L}_{\text{REG}} = \mathcal{L}_3$.

* **Tétel:** Az $\mathcal{L}_3$ nyelvosztály zárt a komplementer, a metszet, a különbség és a szimmetrikus differencia műveletekre.

* **Tétel:** *(Chomsky nyelvhierarchia)* $\mathcal{L}_0 \supset \mathcal{L}_1 \supset \mathcal{L}_2 \supset \mathcal{L}_3$.

Tétel: *(Myhill-Nerode tétel)* $L \in \mathcal{L}_3$ akkor és csak akkor, ha $\left|\{L_p\}_{p \in T^*}\right| < \infty$, ahol $T = T(L)$ az L nyelv ábécéje.

* **Tétel:** *(Kis Bar-Hillel lemma)* Minden $L \in \mathcal{L}_3$ nyelvhez van olyan $n = n(L) \in \mathbb{N}$ nyelvfüggő konstans, hogy minden $u \in L$, $\ell(u) \geq n$ szó esetén van u-nak olyan $u = xyz$ felbontása $(x, y, z \in T(L)^*)$, melyre
  - $\ell(xy) \leq n$,
  - $\ell(y) > 0$,
  - minden $i \geq 0$ egész esetén $xy^iz \in L$.

* **Tétel:** *(Nagy Bar-Hillel-lemma)* Minden $L \in \mathcal{L}_2$ nyelvhez vannak olyan $p = p(L)$, $q = q(L) \in \mathbb{N}$ nyelvfüggő konstansok , hogy minden $u \in L$, $\ell(u) \geq p$ szó esetén van u-nak olyan $u = xyzvw$ felbontása $(x, y, z, v, w \in T(L)^*)$, melyre
  - $\ell(yzv) \leq q$,
  - $\ell(yv) > 0$,
  - minden $i \geq 0$ egész esetén $xy^izv^iw \in L$.

* **Tétel:** A determinisztikus veremautomaták által elfogadott nyelvek osztálya valódi részhalmaza a veremautomaták által elfogadott nyelvek osztályának.

* **Tétel:** A végállapottal és az üres veremmel elfogadó veremautomaták által elfogadható nyelvek nyelvosztálya megegyezik.

* **Tétel:** $\mathcal{L}_{1\mathrm{V}} = \mathcal{L}_2$.

* **Tétel:** $\mathcal{L}_{n\mathrm{V}} = \mathcal{L}_0$, $(n \geq 2)$.

* **Lemma:** Tetszőleges $G \in \mathcal{G}_2$ nyelvtan, $Z \in T \cup N \cup \{\varepsilon\}$ és $\alpha \in (T \cup N)^*$ esetén $Z \xrightarrow[G]{*} \alpha$ akkor és csak akkor, ha létezik t G feletti szintaxisfa, melyre $\mathrm{gy}(t) = Z$ és $\mathrm{front}(t) = \alpha$.

Tétel: *(Lináris nyelvi egyenletek megoldóképlete)* Ha R_1 és R_2 reguláris kifejezések és $\varepsilon \notin L(R_1)$, akkor az $R_1X \cup R_2 = X$ egyenlet egyértelmű megoldása $X = R_1^* R_2$.

Tétel: *(Lineáris nyelvi egyenletrendszerek megoldhatósága)* Legyen $\mathbf{Mx} \cup \mathbf{v} = \mathbf{x}$ nyelvi egyenletrendszer, ahol az $\mathbf{M} = \begin{bmatrix} L_{11} & \cdots & L_{1n} \\ \vdots & \ddots & \vdots \\ L_{n1} & \cdots & L_{nn} \end{bmatrix}$ nyelvmátrix és a $\mathbf{v} = \begin{bmatrix} L_1 \\ \vdots \\ L_n \end{bmatrix}$ nyelvvektor adottak és $\mathbf{x} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ ismeretlen nyelvekből álló vektor. Ha $\mathcal{L} = \bigcup_{j=1}^{n} \bigcup_{k=1}^{n} \{L_{jk}\} \cup \bigcup_{j=1}^{n} \{L_j\} \subseteq \mathcal{L}_i$ $(i \in \{0, 1, 2, 3\})$ és $\varepsilon \notin L_{jk}$ $(1 \leq j \leq n, 1 \leq k \leq n)$, akkor az egyenletrendszernek egyértelműen létezik megoldása, melynek elemei $\mathcal{L}$ elemeiből reguláris műveletekkel megkaphatók.

Tétel: $\mathcal{L}_{\mathrm{LL}(k)} \subset \mathcal{L}_{\mathrm{LR}(k)} = \mathcal{L}_{1\mathrm{DV}}$.

Algoritmusok

Álterminálisok bevezetése ($\Phi_{\textbf{Ált}}$)

Input: $G = \langle T, N, \mathcal{P}, S \rangle$ nyelvtan.
Output: G' nyelvtan, melynek csak $A \to a$ $(A \in N, a \in T)$ sémájú szabályai tartalmaznak terminálist és $G' \sim G$.
Minden $t \in T$ terminálisra t valamennyi előfordulását $\mathcal{P}$-beli szabályokban egy új (nem N-beli), terminálisonként egyedi Q_t nyelvtani jelre cseréljük.
Minden $t \in T$ terminálisra hozzáadjuk a szabályrendszerhez a $Q_t \to t$ szabályt.

0. típusú ε-mentesítés ($\Phi_{\textbf{0epsz}}$)

Input: $G = \langle T, N, \mathcal{P}, S \rangle$ nyelvtan.
Output: G' epszilonmentes nyelvtan, melyre $G' \underset{\text{kv}}{\sim} G$.
Minden $Z \in T \cup N$ és $p \to \varepsilon \in \mathcal{P}$ $(p \in (T \cup N)^* N (T \cup N)^*)$ esetén hozzáadjuk a szabályrendszerhez a $Zp \to Z$ és $pZ \to Z$ szabályokat, majd az epszilonszabályokat elhagyjuk.

2. típusú ε-mentesítés ($\Phi_{\textbf{2epsz}}$)

Input: $G = \langle T, N, \mathcal{P}, S \rangle$ 2. típusú nyelvtan.
Output: G' megszorított 2. típusú nyelvtan, melyre $G \sim G'$.
Az első lépésben meghatározzuk a $H := \{A \in N \mid A \xrightarrow[G]{*} \varepsilon\}$ halmazt. Ehhez rekurzívan definiáljuk a H_i $(i \geq 1)$ halmazokat:
$H_1 := \{A \in N \mid A \to \varepsilon \in \mathcal{P}\}$,
$H_{i+1} := H_i \cup \{A \in N \mid \exists A \to Q \in \mathcal{P} : Q \in H_i^*\}$.
$H_1 \subseteq H_2 \subseteq \dots \subseteq H_i \subseteq \dots$, és mivel a H_i halmaz elemszáma felülről korlátos ezért stabilizálódik a sorozat, azaz egy i_0 indextől kezdődően biztosan azonosak lesznek ezek a halmazok, ez a H_{i_0} lesz a H halmaz.
A második lépésben a H halmaz ismeretében átalakítjuk a nyelvtant a kellő alakúra.
$S \notin H$ esetén:
$G' := \langle T, N, \bar{\mathcal{P}}, S \rangle$, ahol $A \to \bar{q} \in \bar{\mathcal{P}}$ akkor és csak akkor, ha $\bar{q} \neq \varepsilon \ \wedge\ \exists \mathrm{A} \to q \in \mathcal{P}$, hogy $\bar{q}$-t q-ból néhány (esetleg nulla) H-beli jel elhagyásával kapjuk.
$S \in H$ esetén:
$\bar{\mathcal{P}}$-hez vegyük hozzá még az $S' \to S \,|\, \varepsilon$ szabályokat és S' legyen az új kezdőszimbólum.
Megjegyzés: Φ_{2epsz} megőrzi a 2. és 3. típust.

Láncmentesítés ($\Phi_{\textbf{Lánc}}$)

Input: $G = \langle T, N, \mathcal{P}, S \rangle$ 1. típusú nyelvtan.
Output: G' 1-es típusú láncszabálymentes nyelvtan, melyre $G' \sim G$.
Az első lépésben meghatározzuk minden $A \in N$ esetén a $H(A) := \{B \in N \mid A \xrightarrow[G]{*} B\}$ halmazt.
Ehhez rekurzívan definiáljuk a $H_i(A)$ $(i \geq 0)$ halmazokat:
$H_0(A) := \{A\}$,
$H_{i+1}(A) := H_i(A) \cup \{B \mid \exists C \in H_i(A) \wedge C \underset{G}{\to} B\}$.
$H_0(A) \subseteq H_1(A) \subseteq \dots \subseteq H_i(A) \subseteq \dots$, és mivel a $H_i(A)$ halmaz elemszáma felülről korlátos ezért stabilizálódik a sorozat, azaz egy i_0 indextől kezdődően biztosan azonosak lesznek ezek a halmazok, ez a $H_{i_0}(A)$ lesz a $H(A)$ halmaz.

A második lépésben a $H(A)$ halmazok ($A \in N$) ismeretében átalakítjuk a nyelvtant a kellő alakúra: $G' = \langle T, N, \mathcal{P}', S \rangle$ lesz az új nyelvtan, ahol

$$\mathcal{P}' = \{u_1 A_1 u_2 A_2 \cdots u_n A_n u_{n+1} \to \beta \mid u_1, \ldots, u_{n+1} \in T^* \wedge A_1, \ldots, A_n \in N \wedge \beta \in (T \cup N)^* \wedge \exists B_1 \in H(A_1), \ldots, B_n \in H(A_n) : u_1 B_1 u_2 B_2 \cdots u_n B_n u_{n+1} \to \beta \in \mathcal{P}\}.$$

Megjegyzések: **1.** $\Phi_{\text{Lánc}}$ megőrzi az 1., 2., és 3. típust. **2.** $\Phi_{\text{Lánc}}$ alkalmazható nem feltétlen ε-mentes 3. típusú nyelvtanokra is. Ilyenkor is $G' \sim G$ és $\Phi_{\text{Lánc}}$ megőrzi a 3. típust.

Hosszredukció (Φ_{Hossz})

Input: $G = \langle T, N, \mathcal{P}, S \rangle$ 1-es típusú nyelvtan.
Output: G' 1-es típusú nyelvtan olyan szabályokkal, melyeknek baloldala és jobboldala legfeljebb 2 hosszúságú, továbbá $G' \sim G$.
Legyen $X_1 X_2 \ldots X_m \to Y_1 Y_2 \ldots Y_n$ ($m \geq 2, n \geq m$) hosszúságot nem csökkentő szabály. ($X_1, X_2, \ldots X_m, Y_1, Y_2 \ldots Y_n \in N$)
A szabály szimulációja a $Z_1, Z_2, \ldots, Z_{n-2}$ új nyelvtani jelek bevezetésével:

$$\begin{aligned}
&X_1 X_2 \to Y_1 Z_1,\\
&Z_1 X_3 \to Y_2 Z_2,\\
&\qquad\vdots\\
&Z_{m-3} X_{m-1} \to Y_{m-2} Z_{m-2},
\end{aligned}$$

Továbbá ha $n = m$, akkor

$$Z_{m-2} X_m \to Y_{m-1} Y_m,$$

egyébként ($n > m$ esetén):

$$\begin{aligned}
&Z_{m-2} X_m \to Y_{m-1} Z_{m-1},\\
&Z_{m-1} \to Y_m Z_m,\\
&\qquad\vdots\\
&Z_{n-3} \to Y_{n-2} Z_{n-2},\\
&Z_{n-2} \to Y_{n-1} Y_n.
\end{aligned}$$

$m = 1$ esetén:

$$\begin{aligned}
&X_1 \to Y_1 Z_1,\\
&Z_1 \to Y_2 Z_2,\\
&\qquad\vdots\\
&Z_{n-3} \to Y_{n-2} Z_{n-2},\\
&Z_{n-2} \to Y_{n-1} Y_n.
\end{aligned}$$

Megjegyzések: **1.** A fenti algoritmust 3. típusú G nyelvtan esetén is definiáljuk. Ekkor $m = 1$ és $Y_1, \ldots, Y_{n-1} \in T$, $Y_n \in N \cup \{\varepsilon\}$. Ebben az esetben az algoritmus outputja egy olyan G' nyelvtan, melynek $p' \to q'$ szabályaira $q' \in TN \cup T \cup N \cup \{\varepsilon\}$. **2.** Φ_{Hossz} megőrzi az 1., 2., és az **1.** megjegyzéssel definiált algoritmus esetén a 3. típust továbbá a láncmentességet és az epszilonmentességet.

1-es típusú nyelvtanok normálformára hozása (Φ_{1NF})

Kuroda normálforma

Input: G 1-es típusú nyelvtan.
Output: G' Kuroda normálformájú nyelvtan, melyre $G' \sim G$.
Lépései:

1. Álterminálisok bevezetése

2. Láncmentesítés

3. Hosszredukció

4. Az $AB \to CD$ $(A \neq C, B \neq D)$ sémájú szabályok eliminálása

(Az $AB \to CD$ $(A \neq C, B \neq D)$ sémájú szabályokat az $AB \to AW$, $AW \to CW$, $CW \to CD$ szabályokkal helyettesítjük, ahol W új, egyedi nyelvtani jel.)

Zsákutcamentesítés ($\Phi_{\mathbf{Zsák}}$)

Input: G 2. típusú nyelvtan.

Output: G' zsákutcamentes 2. típusú nyelvtan, melyre $G' \sim G$.

Az első lépésben meghatározzuk a $J := \{A \in N \mid \exists u \in T^*,\ A \xrightarrow[G]{*} u\}$ halmazt. Ehhez rekurzívan definiáljuk a J_i $(i \geq 1)$ halmazokat:

$J_1 := \{A \in N \mid \exists u \in T^*, A \to u \in \mathcal{P}\}$,

$J_{i+1} := J_i \cup \{A \in N \mid \exists A \to Q \in \mathcal{P} : Q \in (J_i \cup T)^*\}$.

$J_1 \subseteq J_2 \subseteq \cdots \subseteq J_i \subseteq \ldots$, és mivel a J_i halmaz elemszáma felülről korlátos ezért stabilizálódik a sorozat, azaz egy i_0 indextől kezdődően biztosan azonosak lesznek ezek a halmazok, ez a J_{i_0} lesz a J halmaz.

Ezek után a G nyelvtant úgy alakítjuk át, hogy elhagyunk minden olyan szabályt, mely tartalmaz $(N \setminus J)$-beli nyelvtani jelt.

Összefüggővé alakítás ($\Phi_{\mathbf{Öf}}$)

Input: G 2. típusú nyelvtan.

Output: G' összefüggő 2. típusú nyelvtan, melyre $G' \sim G$.

Az első lépésben meghatározzuk a $K := \{A \in N \mid \exists \alpha \in (T \cup N)^* A (T \cup N)^*,\ S \xrightarrow[G]{*} \alpha\}$ halmazt. Ehhez rekurzívan definiáljuk a K_i $(i \geq 0)$ halmazokat:

$K_0 := \{S\}$; $K_1 := K_0 \cup \{A \in N \mid \exists \alpha \in (T \cup N)^* A (T \cup N)^*,\ S \to \alpha \in \mathcal{P}\}$,

$K_{i+1} := K_i \cup \{A \in N \mid \exists B \in K_i,\ \alpha \in (T \cup N)^* A (T \cup N)^*,\ B \to \alpha \in \mathcal{P}\}$.

$K_1 \subseteq K_2 \subseteq \cdots \subseteq K_i \subseteq \ldots$, és mivel a K_i halmaz elemszáma felülről korlátos ezért stabilizálódik a sorozat, azaz egy i_0 indextől kezdődően biztosan azonosak lesznek ezek a halmazok, ez a K_{i_0} lesz a K halmaz.

Ezek után a G nyelvtant úgy alakítjuk át, hogy elhagyunk minden olyan szabályt, melynek baloldala $(N \setminus K)$-beli.

2-es típusú nyelvtanok redukciója ($\Phi_{\mathbf{Red}}$)

Input: G 2-es típusú nyelvtan.

Output: G' redukált (zsákutcamentes és összefüggő) 2-es típusú nyelvtan, melyre $G' \sim G$.

Lépései:

1. Zsákutcamentesítés

2. Összefüggővé tétel.

2-es típusú nyelvtanok normálformára hozása ($\Phi_{\mathbf{2NF}}$)

Chomsky normálforma

Input: G 2-es típusú nyelvtan.

Output: G' Chomsky normálformájú nyelvtan, melyre $G' \sim G$.

Lépései:

1. Álterminálisok bevezetése

2. ε-mentesítés

3. Láncmentesítés

4. Hosszredukció

3-as típusú nyelvtanok normálformára hozása ($\Phi_{\mathbf{3NF}}$)

Input: G 3-as típusú nyelvtan.
Output: G' 3-as típusú normálformájú nyelvtan, melyre $G' \sim G$.
Lépései:

1. Láncmentesítés

2. Hosszredukció

3. Az $A \to a$ sémájú szabályok eliminálása

(Minden $A \to a$ sémájú szabályt az $A \to aF$ szabállyal helyettesítünk, ahol F új, egyedi nyelvtani jel, és hozzáadjuk még a szabályrendszerhez az $F \to \varepsilon$ szabályt.)

Polinomiális algoritmus a szóprobléma eldöntésére 2. típus esetén

Cocke-Younger-Kasami (CYK) algoritmus

Input: $G = \langle T, N, \mathcal{P}, S \rangle$ Chomsky normálformájú nyelvtan és egy $u = t_1 \cdots t_n \in T^*$ szó.
Output: `IGEN`, ha $u \in L(G)$. `NEM`, ha $u \notin L(G)$.
Ha $u = \varepsilon$, akkor $u \in L(G) \iff S \to \varepsilon \in \mathcal{P}$.
Legyen A_i a $P_i \in \mathcal{P}$ szabály bal-, q_i pedig a jobboldala. ($A_i \in N$, $q_i \in T \cup N^2$.)
A CYK algoritmus rekurzíven definiál $H_{i,j}$, $1 \le i \le j \le n$ halmazokat $(j-i)$ szerint növekvő sorrendben.

$$H_{i,i} := \{A_k \mid q_k = t_i\},$$

$$H_{i,j} := \{A_k \mid q_k \in \bigcup_{r=i}^{j-1} H_{i,r} H_{r+1,j}\} \quad (i < j).$$

Ha $S \in H_{1,n}$, akkor $u \in L(G)$, különben $u \notin L(G)$.

Lineáris algoritmus a szóprobléma eldöntésére 3. típus esetén

Input: $G = \langle T, N, \mathcal{P}, S \rangle$ 3. típusú normálformájú nyelvtan és egy $u = t_1 \cdots t_n \in T^*$ szó.
Output: `IGEN`, ha $u \in L(G)$. `NEM` ha $u \notin L(G)$.
Az algoritmus rekurzívan kiszámol egy a nyelvtani jelek halmazának részhalmazaiból álló sorozatot.

$$H_0 = \{S\},$$

$$H_{i+1} = \{A \in N \mid \exists B \in H_i \wedge B \to t_{i+1} A \in \mathcal{P}\}.$$

Legyen továbbá $F = \{A \in N \mid A \to \varepsilon \in \mathcal{P}\}$.
$u \in L(G) \Leftrightarrow H_n \cap F \neq \emptyset$.

Minimális automata előállítása

Input: $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ véges, determinisztikus automata.
Output: $L(\mathcal{A})$ minimális automatája.
Lépései:

1. Összefüggővé alakítás

Meghatározzuk a q_0-ból elérhető állapotok H halmazát.

$$H_0 := \{q_0\},$$

$H_{i+1} := H_i \cup \{q \mid \exists\, q' \in H_i \wedge \exists\, t \in T : \delta(q', t) = q\}$,
$H_0 \subseteq H_1 \subseteq \dots \subseteq H_i \subseteq \dots$, és mivel a H_i halmaz elemszáma felülről korlátos ezért stabilizálódik a sorozat, azaz egy i_0 indextől kezdődően biztosan azonosak lesznek ezek a halmazok, ez a H_{i_0} lesz a H halmaz. Az összefüggő automata:
$\mathcal{A}_{\text{össz}} = \langle H, T, \delta\big|_{H\times T}, q_0, F \cap H\rangle$.

2. Redukció

Rekurzívan meghatározzuk az $\mathcal{A}_{\text{össz}}$ automata $\overset{0}{\sim}, \overset{1}{\sim}, \dots$ ekvivalenciáit:

- $q \overset{0}{\sim} q'$, ha $(q \in F \Leftrightarrow q' \in F)$,
- $q \overset{i+1}{\sim} q' \;\Leftrightarrow\; q \overset{i}{\sim} q' \wedge (\forall t \in T : \delta(q,t) \overset{i}{\sim} \delta(q',t))$.

$\overset{0}{\sim} \prec \overset{1}{\sim} \prec \overset{2}{\sim} \prec \dots \prec \sim$, ($\varrho_1 \prec \varrho_2$, ha minden $q, q' \in Q$ esetén $q\varrho_2 q' \Rightarrow q\varrho_1 q'$.)
így az $\overset{i}{\sim}$ az állapotok halmazának egyre finomodó felosztását adja, mely véges sok lépésben stabilizálódik. $i_0 := \min\{i \mid \overset{i}{\sim} = \overset{i+1}{\sim}\}$.
$\mathcal{A}_{\text{össz}}/_{\overset{i_0}{\sim}}$ a minimális automata.

NDA-hoz vele ekvivalens VDA készítése

Input: $\mathcal{A} = \langle Q, T, \delta, q_0, F\rangle$ véges, nemdeterminisztikus automata.
Output: $\mathcal{A}' = \langle Q', T, \delta', q_0', F'\rangle$ véges, determinisztikus automata, melyre $L(\mathcal{A}') = L(\mathcal{A})$.

$Q' := 2^Q$,

$\delta'(\{q_1, \dots q_s\}, t) := \bigcup_{i=1}^{s} \delta(q_i, t) \qquad (q_1, \dots, q_s \in Q, t \in T)$.

$q_0' := \{q_0\}$

$F' := \{A \in 2^Q \mid A \cap F \neq \emptyset\}$.

A q_0'-t tartalmazó $\mathcal{A}'_{\text{össz}}$ komponens meghatározása:
Amikor az állapotokra sorra határozzuk meg az állapotátmeneteket készítünk egy sort δ' értékkészletéről.
Minden lépésben a sor elején levő, még nem vizsgált állapotra meghatározzuk az átmeneteket. Az eljárás akkor ér véget, ha a sor kiürül. Kezdetben a sor egyedül q_0' -t tartalmazza.

3-as normálformájú nyelvtan készítése VDA-hoz

Input: $\mathcal{A} = \langle Q, T, \delta, q_0, F\rangle$ véges, determinisztikus automata.
Output: $G = \langle T, N, \mathcal{P}, S\rangle$ 3-as típusú normálformájú nyelvtan, melyre $L(G) = L(\mathcal{A})$.

$N := Q$,

$S := q_0$,

$q_1 \to tq_2 \in \mathcal{P} \Leftrightarrow \delta(q_1.t) = q_2 \qquad (q_1, q_2 \in Q, t \in T)$,

$q \to \varepsilon \in \mathcal{P} \Leftrightarrow q \in F \qquad (q \in Q)$.

VDA készítése 3-as normálformájú nyelvtanhoz

Input: $G = \langle T, N, \mathcal{P}, S\rangle$ 3-as normálformájú nyelvtan.
Output: $\mathcal{A} = \langle Q, T, \delta, q_0, F\rangle$ véges, determinisztikus automata, melyre $L(\mathcal{A}) = L(G)$.
Lépései:

1. NDA készítése 3NF nyelvtanhoz.

$Q := N$,

$q_0 := S$,

$B \in \delta(A, t) \Leftrightarrow A \to tB \in \mathcal{P} \qquad (A, B \in N, t \in T)$,

$A \in F \Leftrightarrow A \to \varepsilon \in \mathcal{P} \qquad (A \in N)$.

2. NDA-hoz vele ekvivalens VDA készítése.

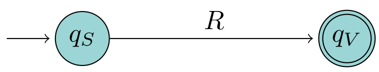
εNDA-hoz vele ekvivalens NDA készítése

Input: $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ εNDA.
Output: $\mathcal{A}' = \langle Q', T, \delta', q_0', F' \rangle$ NDA, melyre $L(\mathcal{A}') = L(\mathcal{A})$.
$Q' := Q,\ q_0' := q_0$.
Egy $q \in Q$ állapot ε-lezártja azon állapotokból áll, ahova q-ból ε-átmenetekkel eljuthatunk. Halmazsorozattal történő rekurzív megadása:
$H_0(q) := \{q\}$.
$H_{i+1}(q) := H_i \cup \bigcup_{q' \in H_i(q)} \delta(q', \varepsilon)$.
$H_0(q) \subseteq H_1(q) \subseteq \cdots \subseteq Q$. A $H_i(q)$ halmazsorozat legfeljebb $|Q|$ lépésben stabizálódik, legyen i_0 a legkisebb index, melyre $H_{i_0}(q) = H_{i_0+1}(q)$. Ekkor $H(q) := H_{i_0}(q)$.
$q' \in \delta'(q, t) \Leftrightarrow \exists q'' \in H(q), q' \in \delta(q'', t)$.
$q \in F' \Leftrightarrow H(q) \cap F \neq \emptyset$.

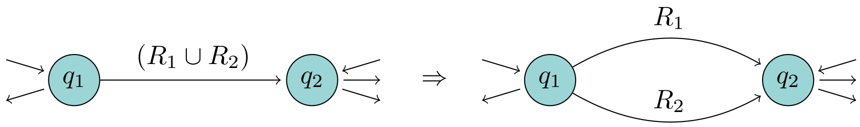
Reguláris kifejezés által leírt nyelvet felismerő VDA készítése

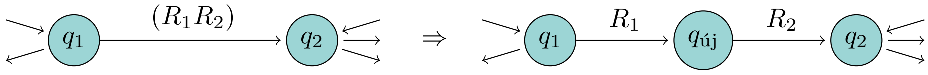
(Automataszintézis)

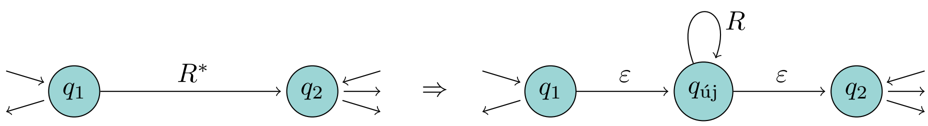
Input: R reguláris kifejezés.
Output: $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ véges, determinisztikus automata, melyre $L(\mathcal{A}) = L(R)$.
Lépései:
0. Általánosított szekvenciális automata készítése reguláris kifejezéshez.
Adott R reguláris kifejezéshez kiindulunk egy $\mathcal{A} = \langle \{q_S, q_V\}, T, \delta, q_S, \{q_V\} \rangle$ általánosított szekvenciális automatából, ahol $\delta(q_S, R) = \{q_V\}$ az egyetlen átmenet. Erre nyílván $L(\mathcal{A}) = L(R)$.



1. Általánosított szekvenciális automata lebontása εNDA-vá
Az alábbi lebontási lépések nem változtatják az elfogadott nyelvet.







Addig bontjuk a reguláris kifejezéseket amíg εNDA-t nem kapunk. (Az $\emptyset$-zal címkézett éleket elhagyjuk.)
2. εNDA-hoz vele ekvivalens NDA készítése
3. NDA-hoz vele ekvivalens VDA készítése

VDA által elfogadott nyelv leírása reguláris kifejezéssel

(Automataanalízis)

Input: $\mathcal{A} = \langle Q, T, \delta, q_0, F \rangle$ VDA.
Output: R reguláris kifejezés, melyre $L(R) = L(\mathcal{A})$.

Lépései:

1. Nyelvi egyenletrendszer felírása az állapotok maradéknyelveire.

- Ha $q \notin F$, akkor a q maradéknyelvére vonatkozó egyenlet: $L(\mathcal{A}, q) = \bigcup_{t \in T} tL(\mathcal{A}, \delta(q,t))$.
- Ha $q \in F$, akkor a q maradéknyelvére vonatkozó egyenlet: $L(\mathcal{A}, q) = \varepsilon \cup \bigcup_{t \in T} tL(\mathcal{A}, \delta(q,t))$.

2. Az egyenletrendszer Gauss-eliminációval történő megoldása $L(\mathcal{A}, q_0)$-ra.

Legyen $Q = \{q_0, q_1 \ldots, q_{n-1}\}$. n egyenletünk van n ismeretlennel. A q_{n-1} állapot maradéknyelvére vonatkozó egyenletből kifejezzük $L(\mathcal{A}, q_{n-1})$-t a többi maradéknyelv függvényében a lineáris nyelvi egyenlet megoldóképlete segítségével. Ezt behelyettesítjük a többi $n-1$ egyenletbe, így marad $n-1$ egyenlet $n-1$ ismeretlennel. Folytatjuk, amíg egy egyenletünk marad, melynek egyetlen ismeretlene $L(\mathcal{A}, q_0)$. Az egyenletet a lineáris nyelvi egyenlet megoldóképlete alapján megoldjuk.
Megjegyzés: Az összes többi maradéknyelvet visszahelyettesítéssel kaphatjuk meg.

VDA előállítása maradéknyelvekből

Input: L nyelv.
Output: Ha $L \in \mathcal{L}_3$, akkor $\mathcal{A}$ VDA, melyre $L(\mathcal{A}) = L$. Ha $L \notin \mathcal{L}_3$, akkor `NINCS`.

Határozzuk meg L szavakra vonatkozó maradéknyelveit, és ha véges sok különböző van, akkor legyenek $p_1, \ldots, p_n \in T(L)^*$ olyan szavak, melyre $L_{p_1}, \ldots, L_{p_n}$ kiadja a maradéknyelvek rendszerét. Az $L_{p_i t}$ $1 \leq i \leq n$, $t \in T(L)$ maradnyelvekről meghatározzuk, mely p_j-re egyezik meg az L_{p_j} maradéknyelvvel. Tehát a VDA:

$\mathcal{A} = \langle \{L_p\}_{p \in T^*}, T, \delta, L_\varepsilon, \{L_p \mid \varepsilon \in L_p\} \rangle$, ahol $\delta(L_p, t) = L_{pt}$.

Jelölések

$A \subseteq B$; $C \subset D$	A részhalmaza B-nek; C valódi részhalmaza D-nek
2^H	H hatványhalmaza
D_f	az f függvény értelmezési tartománya
ε	üres szó
X^*	az összes X feletti szó halmaza
X^+	$X^* \setminus \{\varepsilon\}$, az összes X feletti pozitív hosszúságú szó halmaza
X^i	az összes X feletti i hosszúságú szó halmaza
$X^{\leq i}$	az összes X feletti legfeljebb i hosszúságú szó halmaza
$X^{\geq i}$	az összes X feletti legalább i hosszúságú szó halmaza
$X(u)$; $T(u)$	a legszűkebb ábécé, mely fölött u szó
$X(L)$; $T(L)$	a legszűkebb ábécé, mely felett L nyelv
$\ell(u)$	az u szó hossza
$\ell_t(u)$	az u szóban szereplő t betűk száma
$\ell_H(u)$	az u szóban szereplő H-beli betűk száma
L^*	L lezártja
L^+	L pozitív lezártja
u^{-1}; L^{-1}	u illetve L megfordítása
$\mathrm{Pre}(u)$; $\mathrm{Suf}(u)$	u prefix- illetve suffixhalmaza
$\mathrm{Pre}(L)$; $\mathrm{Suf}(L)$	L prefix- illetve suffixhalmaza
$\mathrm{Pre}(u,i)$; $\mathrm{Suf}(u,i)$	az u szó legfeljebb i hosszúságú prefix- illetve suffixhalmaza
$\mathrm{pre}(u,i)$; $\mathrm{suf}(u,i)$	az u szó i hosszúságú prefixe illetve suffixe
$u \subseteq v$	u részszava v-nek
$\mathcal{R}(X)$	X ábécé feletti reguláris kifejezések halmaza
$\mathcal{R}$	az összes reguláris kifejezés halmaza
$\mathcal{R}_{\text{Ált}}(X)$	X ábécé feletti általánosított reguláris kifejezések halmaza
$L(R)$	az R reguláris kifejezés által reprezentált nyelv
$\alpha \underset{G}{\rightarrow} \beta$	α-ból közvetlenül levezethető β
$\alpha \overset{k}{\underset{G}{\rightarrow}} \beta$	α-ból k lépésben levezethető β
$\alpha \overset{*}{\underset{G}{\rightarrow}} \beta$	α-ból közvetetten levezethető β
$L(G)$	a G nyelvtan által generált nyelv
$\mathcal{G}_i$	i. típusú nyelvtanok osztálya ($i \in \{0,1,2,3\}$)
$\mathcal{G}_{\text{megsz}i}$	megszorított i. típusú nyelvtanok osztálya ($i \in \{0,1,2,3\}$)
$G_1 \sim G_2$	G_1 és G_2 nyelvtan ekvivalensek
$G_1 \underset{\text{kv}}{\sim} G_2$	G_1 és G_2 nyelvtan kváziekvivalensek
$\mathcal{L}_i$	i. típusú nyelvek nyelvosztálya ($i \in \{0,1,2,3\}$)
$\mathcal{L}_{\text{megsz}i}$	megszorított i. típusú nyelvek nyelvosztálya ($i \in \{0,1,2,3\}$) *(l. megszorítási tétel)*
$\mathcal{L}_{\text{nf}i}$	i-es normálformájú nyelvtanok által generált nyelvek nyelvosztálya ($i \in \{0,1,2,3\}$) *(l. normálforma tétel)*
$\mathcal{L}_{\text{RekFel}}$	a rekurzíve felsorolható nyelvek nyelvosztálya
$\mathcal{L}_{\text{ParcRek}}$	a parciálisan rekurzív nyelvek nyelvosztálya
$\mathcal{L}_{\text{Rek}}$	a rekurzív nyelvek nyelvosztálya
$\mathcal{L}_{n\text{V}}$	az n-vermek által elfogadott nyelvek nyelvosztálya
$\mathcal{L}_{\varepsilon\text{NDA}}$	az εNDA-k által elfogadott nyelvek nyelvosztálya

$\mathcal{L}_{\mathrm{NDA}}$	az NDA-k által elfogadott nyelvek nyelvosztálya
$\mathcal{L}_{\mathrm{PDA}}$	a PDA-k által elfogadott nyelvek nyelvosztálya
$\mathcal{L}_{\mathrm{VDA}}$	a VDA-k által elfogadott nyelvek nyelvosztálya
$\mathcal{L}_{\mathrm{REG}}$	a reguláris nyelvek nyelvosztálya
$\Phi_{\text{Ált}}$	Az álterminálisok bevezetésének nyelvtani transzformációja
Φ_{0epsz}	A 0. típusú ε-mentesítés nyelvtani transzformációja
Φ_{2epsz}	A 2. típusú ε-mentesítés nyelvtani transzformációja
$\Phi_{\text{Lánc}}$	A láncmentesítés nyelvtani transzformációja
Φ_{Hossz}	A hosszredukció nyelvtani transzformációja
Φ_{1NF}	Az 1. típusú (Kuroda) normálformára hozás nyelvtani transzformációja
$\Phi_{\text{Zsák}}$	A zsákutcamentesítés nyelvtani transzformációja
$\Phi_{\text{Össz}}$	Az összefüggővé tétel nyelvtani transzformációja
Φ_{Red}	A 2. típusú nyelvtanok redukciójának nyelvtani transzformációja
Φ_{2NF}	A 2. típusú (Chomsky) normálformára hozás nyelvtani transzformációja
Φ_{3NF}	A 3-as normálformára hozás nyelvtani transzformációja
$\mathrm{front}(t)$	a t szintaxisfa leveleinek balról jobbra való összeolvasása
$\mathrm{gy}(t)$	a t szintaxisfa gyökere
$\alpha \xrightarrow[G,\mathrm{lb}]{} \beta$	α-ból legbal levezetéssel közvetlenül levezethető β
$\alpha \xrightarrow[G,\mathrm{lb}]{k} \beta$	α-ból legbal levezetéssel k lépésben levezethető β
$\alpha \xrightarrow[G,\mathrm{lb}]{*} \beta$	α-ból legbal levezetéssel közvetetten levezethető β
$\alpha \xrightarrow[G,\mathrm{lj}]{} \beta$	α-ból legjobb levezetéssel közvetlenül levezethető β
$\alpha \xrightarrow[G,\mathrm{lj}]{k} \beta$	α-ból legjobb levezetéssel k lépésben levezethető β
$\alpha \xrightarrow[G,\mathrm{lj}]{*} \beta$	α-ból legjobb levezetéssel közvetlenül közvetetten levezethető β
$\mathcal{L}_{\mathrm{LL}(k)}$	$\mathrm{LL}(k)$ nyelvtanok által generált nyelvek nyelvosztálya
$\mathcal{L}_{\mathrm{LR}(k)}$	$\mathrm{LR}(k)$ nyelvtanok által generált nyelvek nyelvosztálya
$\mathcal{L}_{\mathrm{1DV}}$	Determinisztikus 1-vermek által elfogadott nyelvek nyelvosztálya

Konvencionális szimbólumhasználatok

$X, Y, \ldots$	ábécék
$t, x, y, z, a, b, c, \ldots$	betűk
$u, v, w, \ldots$	szavak
ε	üres szó
$L, L_1, \ldots$	nyelvek
$\mathcal{L}, \ldots$	nyelvosztályok
$R, R_1, \ldots$	reguláris kifejezések
$G, G_1, \ldots$	nyelvtanok
$\mathcal{G}, \ldots$	nyelvtanosztályok
T	terminális ábécé
N	nyelvtani jelek (nemterminálisok) ábécéje
$a, b, c, \ldots$	terminálisok
$A, B, C, \ldots$	nemterminálisok
S	kezdőszimbólum (csak akkor, ha nincs más k.sz. külön megadva)
$\alpha, \beta, \gamma, \xi, \varrho, \sigma, \tau, \ldots$	mondatformák
$\mathcal{A}, \mathcal{A}_1, \ldots$	véges determinisztikus vagy nemdeterminisztikus automaták
Q, A	automata állapothalmaza
$q, q_1, \ldots a, a_1, \ldots$	automata állapotai
F	automata végállapotainak halmaza
δ	automata állapotátmenet függvénye
$\mathcal{V}, \mathcal{V}_1, \ldots$	veremautomaták
$\Sigma, \Sigma_1, \ldots$	veremábécék
$\sigma, \sigma_1, \ldots$	veremábécé elemei
ϱ	reláció
Φ	nyelvtani transzformáció