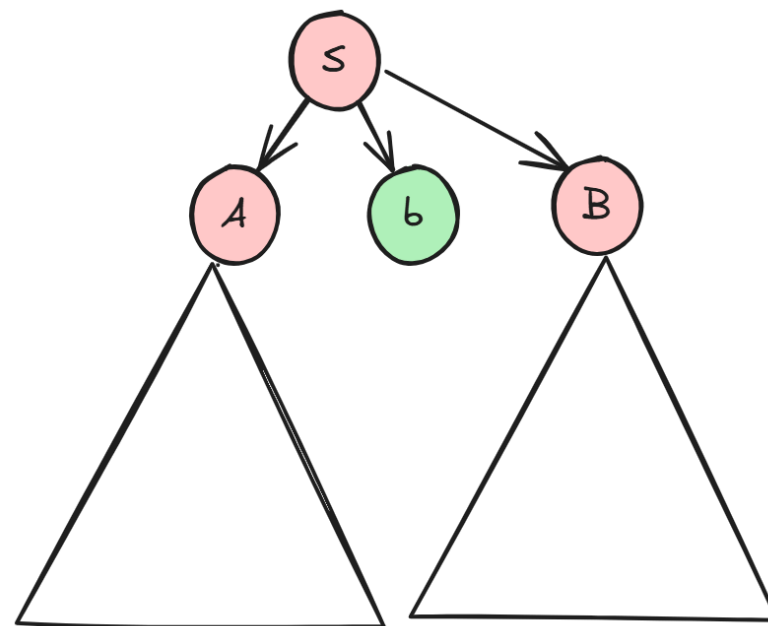


Skladovni avtomati in gramatike

Uroš Čibej



Pregled

- Kontekstno neodvisne gramatike
- Skladovni avtomati
- Primeri
- Povezava med gramatikami in skladovnimi avtomati

Literatura

- Sipser razdelek 2.1 (delno) in 2.2

Prepisovalni sistemi

Formalni sistemi, ki nize preoblikujejo na podlagi prepisovalnih pravil oblike:

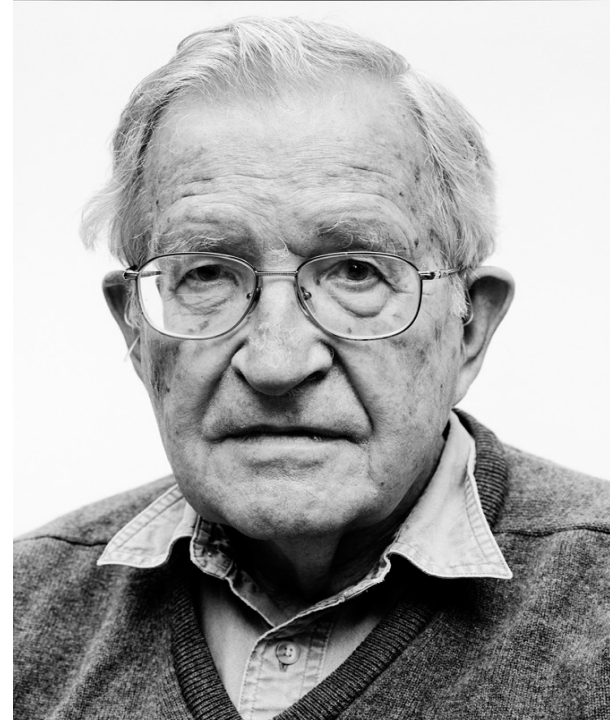
$$u \rightarrow v$$

Prepisovanje je transformacija niza xuy v niz xvy , če obstaja pravilo $u \rightarrow v$

Obstaja veliko različnih prepisovalnih sistemov: λ račun, Thue sistemi, gramatike,...

Kontekstno neodvisne gramatike

- jezikoslovec, matematik, politolog, ... Noam Chomsky
- zajem strukture naravnih jezikov
- hierarhija gramatik:
 - linearne
 - **kontekstno neodvisne**
 - kontekstno odvisne
 - neomejene gramatike



Primer

Stavek → Osebek Povedek Predmet

Osebek → *Miha* | *Janez* | *Jure*

Povedek → *igra* | *pije* | *nosi*

Predmet → *knjigo* | *pivo* | *kolo* | *klobuk*

Primer

$Besedilo \rightarrow Stavek \text{ Ločilo } Besedilo \mid \varepsilon$

$Ločilo \rightarrow . \mid ! \mid ?$

Formalna definicija KNG

$$G = \langle V, T, P, S \rangle$$

- V - končna množica spremenljivk
- T - končna množica končnih simbolov
- P - končna množica pravil (produkcij)
 - $A \rightarrow (V \cup T)^*$
- $S \in V$

Izpeljava v KNG \Longrightarrow

Korak izpeljave je ena uporaba prepisovalnega pravila

$$\beta A \gamma \Longrightarrow \beta \alpha \gamma$$

Če obstaja produkcija $A \rightarrow \alpha$

Če je uporabljenih več pravil, zapisujemo \Longrightarrow^*

Jezik KNG

Jezik gramatike G :

$$L(G) = \{w \in T^* \mid S \xRightarrow{*} w\}$$

Primer

Gramatika za jezik:

$$L = \{a^n b^n\}$$

$$S \rightarrow aSb$$

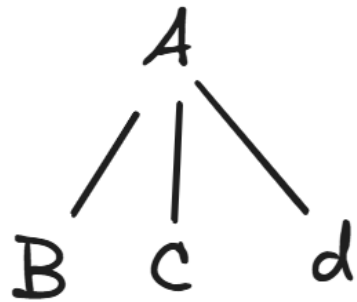
$$S \rightarrow \varepsilon$$

Izpeljava za niz $aabb$

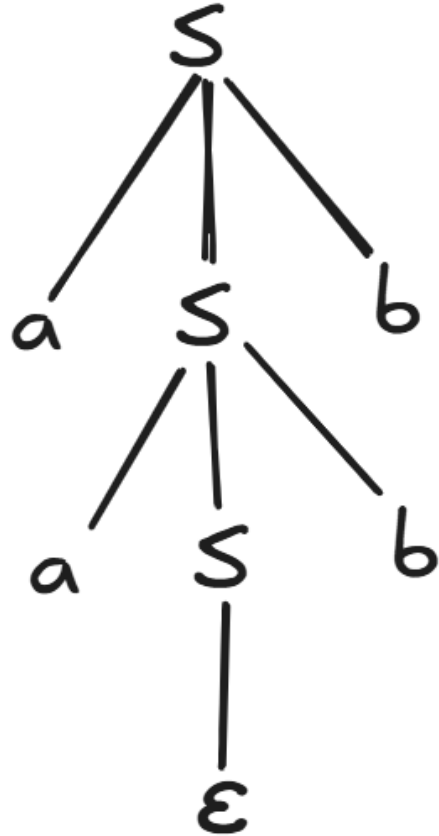
$$S \implies aSb \implies aaSbb \implies aabb$$

Drevesa izpeljav

$$A \implies BCd$$



$$S \implies aSb \implies aaSbb \implies aabb$$

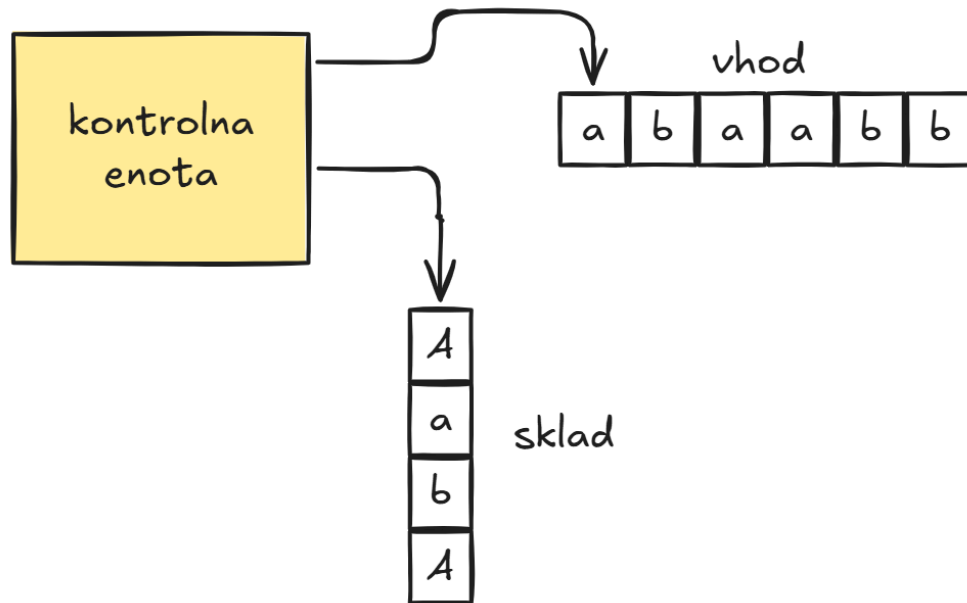


Uporaba KNG v računalništvu

- obdelava naravnega jezika
- definicija formata podatkov
- kompresija podatkov
- **programski jeziki (prevajalniki, tolmači)**
 - Demo: <https://tree-sitter.github.io/tree-sitter/7-playground.html>

Razširitev končnih avtomatov

- vpeljemo koncept neomejenega pomnilnika
- dostopamo ga na specifičen način (kot sklad, LIFO)



Kako bi lahko sklad uporabili za $L = \{a^n b^n\}$?

Formalna definicija SA

$$M = \langle Q, \Sigma, \Gamma, \delta, q_0, F \rangle$$

- Γ - abeceda sklada
- $\delta : Q \times \Sigma_\varepsilon \times \Gamma_\varepsilon \rightarrow 2^{Q \times \Gamma_\varepsilon}$

$$\Sigma_\varepsilon = \Sigma \cup \{\varepsilon\}$$

$$\Gamma_\varepsilon = \Gamma \cup \{\varepsilon\}$$

Podajanje SA

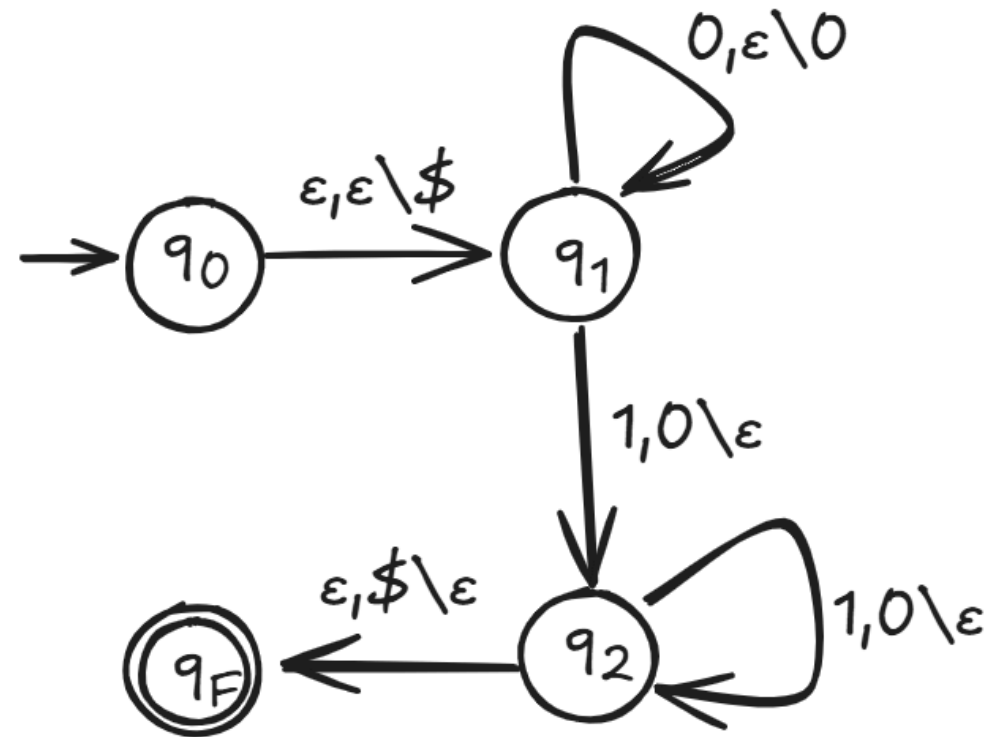
$$\delta(q_0, \varepsilon, \varepsilon) = (q_1, \$)$$

$$\delta(q_1, 0, \varepsilon) = (q_1, 0)$$

$$\delta(q_1, 1, 0) = (q_2, \varepsilon)$$

$$\delta(q_2, 1, 0) = (q_2, \varepsilon)$$

$$\delta(q_2, \varepsilon, \$) = (q_F, \varepsilon)$$



Trenutni opis SA

Stanje, ostanek besede na vhodu, sklad

$$(q, w, s)$$

prehod med trenutnimi opisi

$$(q, aw, z_1\alpha) \vdash (r, w, z_2\alpha)$$

če obstaja $\delta(q, a, z_1) = (r, z_2)$

Jezik skladovnega avtomata

$$L = \{w \mid (q_0, w, \varepsilon) \stackrel{*}{\vdash} (q, \varepsilon, \alpha), q \in F\}$$

Moč nedeterminizma

$$L = \{a^i b^j c^k \mid i = j \vee i = k\}$$

za ta jezik nujno potrebujemo nedeterminizem

A lahko za isti jezik napišemo KNG?

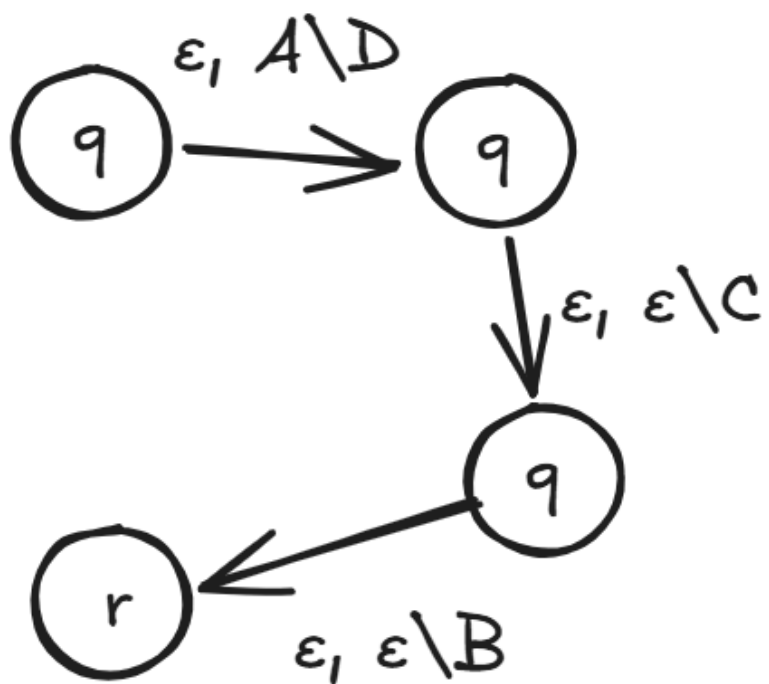
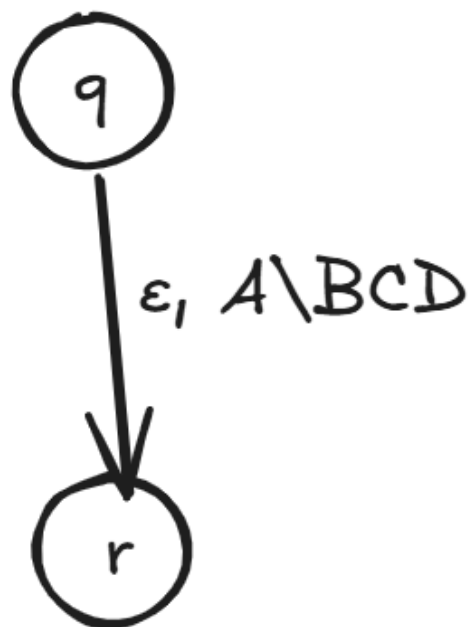
$$L = \{a^i b^j c^k \mid i = j \vee i = k\}$$

Pretvorba KNG v SA

Izrek Če je jezik kontekstno neodvisen (definiran z gramatiko G) \implies
 $\exists M \in SA, L(M) = L(G)$

Ideja dokaza: Iz gramatike G zgradimo skladovni avtomat, ki bo nedeterministično generiral (na skladu) izpeljavo vhodnega niza v G .

Substitucija spr. na skladu



Dokaz

$$G = \langle V, T, P, S \rangle \rightarrow M = \langle Q, T, T \cup V, q_0, \{q_F\} \rangle$$

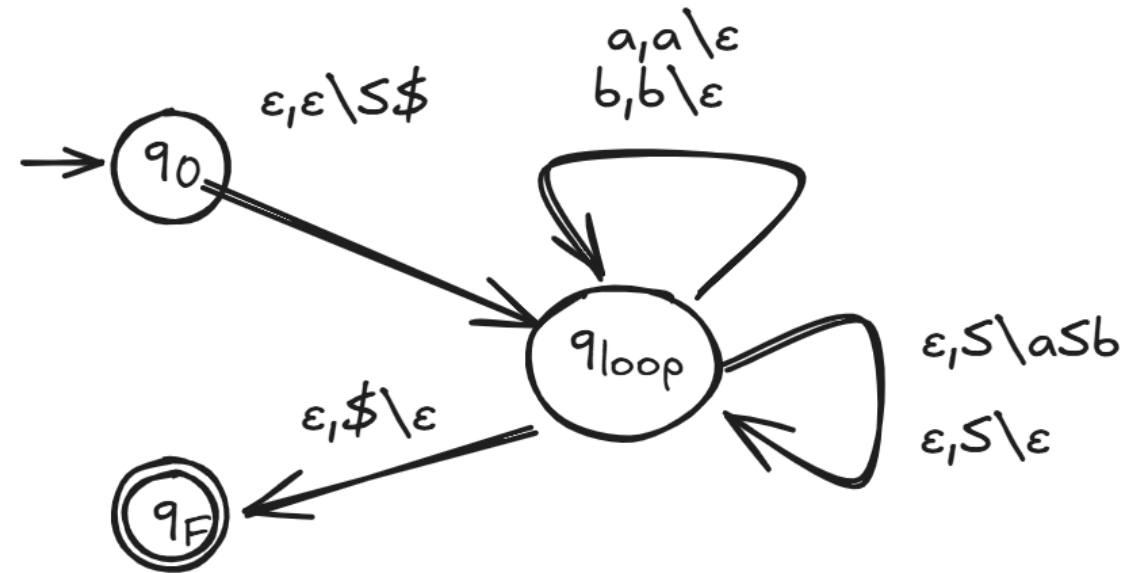
- $Q = \{q_0, q_{loop}, q_F\} \cup Extra$
- **Začetek:** $\delta(q_0, \varepsilon, \varepsilon) = (q_{loop}, S\$)$
- **Konec:** $\delta(q_{loop}, \varepsilon) = (q_F, \backslash \text{varepsilon})\$$
- **Ponavljanje:**
 - i. $\delta(q_{loop}, a, a) = (q_{loop}, \varepsilon)$ za vse $(a \in T)$
 - ii. $\delta(q_{loop}, \varepsilon, A) = \delta(q_{loop}, \alpha)$ za vse $(A \rightarrow \alpha \in P)$

Primer

$$S \rightarrow aSb$$

$$S \rightarrow \varepsilon$$

Domača naloga: odsimulirajte z drevesom trenutnih opisov na besedi $w = aaabbb$



Česa KNG (SA) ne zmorejo

- Obstaja podobna lema o napihovanju kot za RJ
- $L = \{a^n b^n c^n\}$ ni kontekstno neodvisen
- potrebujemo še močnejši model