Kontekstno-neodvisne gramatike za kodiranje in stiskanje podatkov

Janez Podlogar

Univerza v Ljubljani, Fakulteta za matematiko in fiziko

5. 5. 2023

Spomnimo se

S kontekstno-neodvisnih gramatik želimo stisniti poljuben niz.

Spomnimo se

S kontekstno-neodvisnih gramatik želimo stisniti poljuben niz.

Definicija

- *Abeceda* je končna neprazna množica Σ,
- *Množico vseh končnih nizov abecede* Σ označimo s Σ^*,
- *Dolžina niza* w je število znakov abecede v w, označimo z $|w|$.

Spomnimo se

S kontekstno-neodvisnih gramatik želimo stisniti poljuben niz.

Definicija

- *Abeceda* je končna neprazna množica Σ,
- *Množico vseh končnih nizov abecede* Σ označimo s Σ^*,
- *Dolžina niza* w je število znakov abecede v w, označimo z $|w|$.

Primer nizov abecede

Naj bo $\Sigma = \{a, b, c\}$ abeceda, potem sta niza

$$ab \in \Sigma^*, \quad cababcccababcccab \in \Sigma^*$$

Kontekstno-neodvisne gramatike

Definicija

Kontekstno-neodvisna gramatika je četverica $G = (V, \Sigma, P, S)$, kjer je

Kontekstno-neodvisne gramatike

Definicija

Kontekstno-neodvisna gramatika je četverica $G = (V, \Sigma, P, S)$, kjer je

- V končna množica *nekončnih simbolov*,

Kontekstno-neodvisne gramatike

Definicija

Kontekstno-neodvisna gramatika je četverica $G = (V, \Sigma, P, S)$, kjer je

- V končna množica *nekončnih simbolov*,
- abeceda Σ množica *končnih simbolov*,

Kontekstno-neodvisne gramatike

Definicija

Kontekstno-neodvisna gramatika je četverica $G = (V, \Sigma, P, S)$, kjer je

- V končna množica *nekončnih simbolov*,
- abeceda Σ množica *končnih simbolov*,
- $P \subseteq V \times (V \cup \Sigma)^*$ celovita relacija,

Kontekstno-neodvisne gramatike

Definicija

Kontekstno-neodvisna gramatika je četverica $G = (V, \Sigma, P, S)$, kjer je

- V končna množica *nekončnih simbolov*,
- abeceda Σ množica *končnih simbolov*,
- $P \subseteq V \times (V \cup \Sigma)^*$ celovita relacija,
- $S \in V$ *začetni simbol*.

Kontekstno-neodvisne gramatike

Definicija

Kontekstno-neodvisna gramatika je četverica $G = (V, \Sigma, P, S)$, kjer je

- V končna množica *nekončnih simbolov*,
- abeceda Σ množica *končnih simbolov*,
- $P \subseteq V \times (V \cup \Sigma)^*$ celovita relacija,
- $S \in V$ *začetni simbol*.

Definicija

Jezik kontekstno-neodvisne gramatike G je množica vseh nizov, ki jih lahko izpeljemo z gramatiko G, označimo ga z $L(G)$.

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A, B, C\}$

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A, B, C\}$
- $\Sigma = \{a, b, c\}$

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A, B, C\}$
- $\Sigma = \{a, b, c\}$
- $P = \{S \to cCCA, A \to ab, B \to ccc, C \to AAB\}$

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A, B, C\}$
- $\Sigma = \{a, b, c\}$
- $P = \{S \to cCCA, A \to ab, B \to ccc, C \to AAB\}$
- $S = S$

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A, B, C\}$
- $\Sigma = \{a, b, c\}$
- $P = \{S \to cCCA, A \to ab, B \to ccc, C \to AAB\}$
- $S = S$

$S = cCCA$

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A, B, C\}$
- $\Sigma = \{a, b, c\}$
- $P = \{S \to cCCA, A \to ab, B \to ccc, C \to AAB\}$
- $S = S$

$S = cCCA \overset{C}{\Rightarrow} cAABAAB$

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A, B, C\}$
- $\Sigma = \{a, b, c\}$
- $P = \{S \to cCCA, A \to ab, B \to ccc, C \to AAB\}$
- $S = S$

$$S = cCCA \overset{C}{\Rightarrow} cAABAAB \overset{B}{\Rightarrow} cAAcccAAccc$$

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A, B, C\}$
- $\Sigma = \{a, b, c\}$
- $P = \{S \to cCCA, A \to ab, B \to ccc, C \to AAB\}$
- $S = S$

$$S = cCCA \overset{C}{\Rightarrow} cAABAAB \overset{B}{\Rightarrow} cAAcccAAccc \overset{A}{\Rightarrow} cababcccababccc$$

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A, B, C\}$
- $\Sigma = \{a, b, c\}$
- $P = \{S \to cCCA, A \to ab, B \to ccc, C \to AAB\}$
- $S = S$

$$S = cCCA \stackrel{C}{\Rightarrow} cAABAAB \stackrel{B}{\Rightarrow} cAAcccAAccc \stackrel{A}{\Rightarrow} cababcccababccc = w$$

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

Naj bo $G_w = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A, B, C\}$
- $\Sigma = \{a, b, c\}$
- $P = \{S \to cCCA, A \to ab, B \to ccc, C \to AAB\}$
- $S = S$

$S = cCCA \overset{C}{\Rightarrow} cAABAAB \overset{B}{\Rightarrow} cAAcccAAccc \overset{A}{\Rightarrow} cababcccababccc = w$

Velja, da je

$$L(G_w) = \{w\}.$$

Deterministična kontekstno-neodvisna gramatika

Definicija

Kontekstno-neodvisna gramatika G je *deterministična*, če vsak nekončen simbol $A \in V$, nastopa natanko enkrat kot levi član nekega produkcijskega pravila. Kontekstno-neodvisna gramatika, ki ni deterministična, je *nedeterministična*.

Deterministična kontekstno-neodvisna gramatika

Definicija

Kontekstno-neodvisna gramatika G je *deterministična*, če vsak nekončen simbol $A \in V$, nastopa natanko enkrat kot levi član nekega produkcijskega pravila. Kontekstno-neodvisna gramatika, ki ni deterministična, je *nedeterministična*.

Trditev

Naj bo G deterministična kontekstno-neodvisna gramatika. Potem je jezik gramatike G enojec ali pa prazna množica.

Deterministična kontekstno-neodvisna gramatika

Deterministična kontekstno-neodvisna gramatika z praznim jezikom

Naj bo $G = (V, \Sigma, P, S)$, kjer je

Deterministična kontekstno-neodvisna gramatika

Deterministična kontekstno-neodvisna gramatika z praznim jezikom

Naj bo $G = (V, \Sigma, P, S)$, kjer je

- $V = \{S\}$

Deterministična kontekstno-neodvisna gramatika

Deterministična kontekstno-neodvisna gramatika z praznim jezikom

Naj bo $G = (V, \Sigma, P, S)$, kjer je

- $V = \{S\}$
- $\Sigma = \{a\}$

Deterministična kontekstno-neodvisna gramatika

Deterministična kontekstno-neodvisna gramatika z praznim jezikom

Naj bo $G = (V, \Sigma, P, S)$, kjer je

- $V = \{S\}$
- $\Sigma = \{a\}$
- $P = \{S \to S\}$

Deterministična kontekstno-neodvisna gramatika

Deterministična kontekstno-neodvisna gramatika z praznim jezikom

Naj bo $G = (V, \Sigma, P, S)$, kjer je

- $V = \{S\}$
- $\Sigma = \{a\}$
- $P = \{S \to S\}$
- $S = S$

Deterministična kontekstno-neodvisna gramatika

Deterministična kontekstno-neodvisna gramatika z praznim jezikom

Naj bo $G = (V, \Sigma, P, S)$, kjer je

- $V = \{S\}$
- $\Sigma = \{a\}$
- $P = \{S \to S\}$
- $S = S$

Gramatika je očitno deterministična in velja $L(G) = \emptyset$.

Deterministična kontekstno-neodvisna gramatika

Deterministična kontekstno-neodvisna gramatika z odvečnimi simboli

Naj bo $G = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A\}$
- $\Sigma = \{a\}$
- $P = \{S \to a, A \to a\}$
- $S = S$

Deterministična kontekstno-neodvisna gramatika

Deterministična kontekstno-neodvisna gramatika z odvečnimi simboli

Naj bo $G = (V, \Sigma, P, S)$, kjer je

- $V = \{S, A\}$
- $\Sigma = \{a\}$
- $P = \{S \to a, A \to a\}$
- $S = S$

Gramatika je očitno deterministična in velja $L(G) = \{a\}$.

Neuporabni simboli in prazen jezik

Definicija

Pravimo, da kontekstno-neodvisna gramatika G *ne vsebuje neuporabnih simbolov*, ko za vsak simbol $y \in V \cup \Sigma,\ y \neq S$, obstaja končno mnogo nizov $\alpha_1, \alpha_2, \ldots, \alpha_n \in (V \cup \Sigma)^*$ tako, da je y vsebovan vsaj v enem izmed nizov in velja

$$S = \alpha_1 \Rightarrow \alpha_2 \Rightarrow \cdots \alpha_n \in L(G).$$

Dopustne gramatike

Definicija

Kontekstno-neodvisna gramatika G je *dopustna*, če je

- Deterministična,
- Ne vsebuje neuporabnih simbolov,
- $L(G) \neq \emptyset$ in prazen niz ne nastopa kot desni član kateregakoli produkcijska pravila v P.

Dopustne gramatike

Definicija

Kontekstno-neodvisna gramatika G je *dopustna*, če je

- Deterministična,
- Ne vsebuje neuporabnih simbolov,
- $L(G) \neq \emptyset$ in prazen niz ne nastopa kot desni član kateregakoli produkcijska pravila v P.

Posledica

Jezik dopustne kontekstno-neodvisne gramatike je enojec.

Za dopustno gramatiko je dovolj podati produkcijska pravila

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$

Za dopustno gramatiko je dovolj podati produkcijska pravila

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$

Nekončni simboli gramatike so

$$V = \{A_0, A_1, A_2, A_3\}.$$

Za dopustno gramatiko je dovolj podati produkcijska pravila

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$

Nekončni simboli gramatike so

$$V = \{A_0, A_1, A_2, A_3\}.$$

Končni simboli gramatike so

$$\Sigma = \{a, b\}.$$

Za dopustno gramatiko je dovolj podati produkcijska pravila

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$

Nekončni simboli gramatike so

$$V = \{A_0, A_1, A_2, A_3\}.$$

Končni simboli gramatike so

$$\Sigma = \{a, b\}.$$

Začetni simbol je

$$S = A_0.$$

Endomorfizem

Definicija

Naj bo Σ abeceda. *Endomorfizem množice* Σ^* je preslikava $f : \Sigma^* \to \Sigma^*$ tako, da je

$$f(\varepsilon) = \varepsilon \text{ in } \forall w, u \in \Sigma^* : f(wu) = f(w)f(u).$$

Endomorfizem

Definicija

Naj bo Σ abeceda. *Endomorfizem množice* Σ^* je preslikava $f : \Sigma^* \to \Sigma^*$ tako, da je

$$f(\varepsilon) = \varepsilon \text{ in } \forall w, u \in \Sigma^* : f(wu) = f(w)f(u).$$

Induktivno definiramo

$$\begin{aligned} f^0(w) &= w, \\ f^1(w) &= f(w), \\ f^k(w) &= f(f^{k-1}(w)), \end{aligned}$$

kjer je $w \in \Sigma^*$ in $k \geq 2$ celo število.

D0L sistem

Definicija

D0L sistem je trojica $D = (\Sigma, f, w)$, kjer je

D0L sistem

Definicija

D0L sistem je trojica $D = (\Sigma, f, w)$, kjer je

- Σ abeceda,

D0L sistem

Definicija

D0L sistem je trojica $D = (\Sigma, f, w)$, kjer je

- Σ abeceda,
- f endomorfizem množice Σ^*,

D0L sistem

Definicija

D0L sistem je trojica $D = (\Sigma, f, w)$, kjer je

- Σ abeceda,
- f endomorfizem množice Σ^*,
- $w \in \Sigma^*$ *aksiom*.

D0L sistem

Definicija

D0L sistem je trojica $D = (\Sigma, f, w)$, kjer je

- Σ abeceda,
- f endomorfizem množice Σ^*,
- $w \in \Sigma^*$ *aksiom*.

Sistem tvori zaporedje nizov $\{f^k(w) \mid k = 0, 1, 2 \ldots\}$, ki ima *fiksno točko* w^*, če velja

$$w^* \in \{f^k(w) \mid k = 0, 1, 2 \ldots\},$$
$$f(w^*) = w^*.$$

D0L sistem prirejen gramatiki

Definicija

Naj bo G deterministična kontekstno-neodvisna gramatika v kateri prazen niz ne nastopa kot desni član kateregakoli produkcijska pravila. Na $(V \cup \Sigma)^*$ definiramo endomorfizem f_G tako, da

$$\forall a \in \Sigma \colon f_G(a) = a;$$

$$\text{če je } A \to \alpha \text{ produkcijsko pravilo, potem je } f_G(A) = \alpha.$$

D0L sistem prirejen gramatiki

Definicija

Naj bo G deterministična kontekstno-neodvisna gramatika v kateri prazen niz ne nastopa kot desni član kateregakoli produkcijska pravila. Na $(V \cup \Sigma)^*$ definiramo endomorfizem f_G tako, da

$$\forall a \in \Sigma \colon f_G(a) = a;$$

$$\text{če je } A \to \alpha \text{ produkcijsko pravilo, potem je } f_G(A) = \alpha.$$

D0L sistem $(V \cup \Sigma,\ f_G,\ S)$ imenujemo *D0L sistem prirejen gramatiki G*.

D0L sistem prirejen gramatiki

Izrek

Naj bo G dopustna kontekstno-neodvisna gramatika. Potem jezik gramatike G ustreza fiksni točki D0L sistema $(V \cup \Sigma,\ f_G,\ S)$.

D0L sistem prirejen gramatiki

Izrek

Naj bo G dopustna kontekstno-neodvisna gramatika. Potem jezik gramatike G ustreza fiksni točki D0L sistema $(V \cup \Sigma,\ f_G,\ S)$.

Posledica

Jezik dopustne kontekstno-neodvisna gramatike G je

$$L(G) = \{f_G^{|V|}(S)\}.$$

Karakterizacija dopustnih gramatik preko D0L sistema

Izrek

Naj bo G kontekstno-neodvisna gramatika v kateri prazen niz ne nastopa kot desni član kateregakoli produkcijska pravila. Gramatika G je dopustna natanko takrat, ko je $f_G^{|V|}(S) \in \Sigma^+$ in vsak simbol iz $V \cup \Sigma$ nastopa v vsaj enem izmed nizov $f_G^i(S)$, kjer je $i = 0, 1, \ldots, |V|$.

Preverimo, da je gramatika dopustna in izračunajmo jezik

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$

Preverimo, da je gramatika dopustna in izračunajmo jezik

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$
- $V = \{A_0, A_1, A_2, A_3\}$ in $\Sigma = \{a, b\}$

Preverimo, da je gramatika dopustna in izračunajmo jezik

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$
- $V = \{A_0, A_1, A_2, A_3\}$ in $\Sigma = \{a, b\}$

$$f_G(A_0) = aA_1A_2A_3$$

Preverimo, da je gramatika dopustna in izračunajmo jezik

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$
- $V = \{A_0, A_1, A_2, A_3\}$ in $\Sigma = \{a, b\}$

$$f_G(A_0) = aA_1A_2A_3$$
$$f_G^2(A_0) = aabA_1bA_2b$$

Preverimo, da je gramatika dopustna in izračunajmo jezik

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$
- $V = \{A_0, A_1, A_2, A_3\}$ in $\Sigma = \{a, b\}$

$$
\begin{aligned}
f_G(A_0) &= aA_1A_2A_3 \\
f_G^2(A_0) &= aabA_1bA_2b \\
f_G^3(A_0) &= aababbA_1bb \\
f_G^4(A_0) &= aababbabbb
\end{aligned}
$$

Preverimo, da je gramatika dopustna in izračunajmo jezik

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$
- $V = \{A_0, A_1, A_2, A_3\}$ in $\Sigma = \{a, b\}$

$$
\begin{aligned}
f_G(A_0) &= aA_1A_2A_3 \\
f_G^2(A_0) &= aabA_1bA_2b \\
f_G^3(A_0) &= aababbA_1bb \\
f_G^4(A_0) &= aababbabbb
\end{aligned}
$$

Vsak od simbolov A_1, A_2, A_3, a, b se pojavi vsaj enkrat v zgoraj izračunanih nizih

Preverimo, da je gramatika dopustna in izračunajmo jezik

Podana imamo produkcijska pravila

- $P = \{A_0 \to aA_1A_2A_3, A_1 \to ab, A_2 \to A_1b, A_3 \to A_2b\}$
- $V = \{A_0, A_1, A_2, A_3\}$ in $\Sigma = \{a, b\}$

$$
\begin{aligned}
f_G(A_0) &= aA_1A_2A_3 \\
f_G^2(A_0) &= aabA_1bA_2b \\
f_G^3(A_0) &= aababbA_1bb \\
f_G^4(A_0) &= aababbabbb
\end{aligned}
$$

Vsak od simbolov A_1, A_2, A_3, a, b se pojavi vsaj enkrat v zgoraj izračunanih nizih in $f_G^4(A_0) \in \Sigma^+$.

Pretvorba niza v gramatiko

Z $\mathcal{A}$ označimo poljubno abecedo z vsaj dvema črkama in fiksiramo števno neskončno množico simbolov

$$\{A_0, A_1, A_2, \ldots\},$$

Predpostavimo, da nobeden od simbolov $A_0, A_1, A_2, \ldots$ ne nastopa v abecedi $\mathcal{A}$.

Pretvorba niza v gramatiko

Definicija

Naj bo $\mathcal{A}$ abeceda. Z $\mathcal{G}(\mathcal{A})$ označimo podmnožico vseh kontekstno neodvisnih gramatik G, ki izpolnjujejo naslednje pogoje:

1. G je dopustna;

Pretvorba niza v gramatiko

Definicija

Naj bo $\mathcal{A}$ abeceda. Z $\mathcal{G}(\mathcal{A})$ označimo podmnožico vseh kontekstno neodvisnih gramatik G, ki izpolnjujejo naslednje pogoje:

1. G je dopustna;
2. $\Sigma(G) \subseteq \mathcal{A}$;

Pretvorba niza v gramatiko

Definicija

Naj bo $\mathcal{A}$ abeceda. Z $\mathcal{G}(\mathcal{A})$ označimo podmnožico vseh kontekstno neodvisnih gramatik G, ki izpolnjujejo naslednje pogoje:

1. G je dopustna;
2. $\Sigma(G) \subseteq \mathcal{A}$;
3. $S(G) = A_0$;

Pretvorba niza v gramatiko

Definicija

Naj bo $\mathcal{A}$ abeceda. Z $\mathcal{G}(\mathcal{A})$ označimo podmnožico vseh kontekstno neodvisnih gramatik G, ki izpolnjujejo naslednje pogoje:

1. G je dopustna;
2. $\Sigma(G) \subseteq \mathcal{A}$;
3. $S(G) = A_0$;
4. $V(G) = \{A_0, A_1, A_2, \dots, A_{|V(G)|-1}\}$;

Pretvorba niza v gramatiko

Definicija

Naj bo $\mathcal{A}$ abeceda. Z $\mathcal{G}(\mathcal{A})$ označimo podmnožico vseh kontekstno neodvisnih gramatik G, ki izpolnjujejo naslednje pogoje:

1. G je dopustna;
2. $\Sigma(G) \subseteq \mathcal{A}$;
3. $S(G) = A_0$;
4. $V(G) = \{A_0, A_1, A_2, \ldots, A_{|V(G)|-1}\}$;
5. Če naštejemo nekončne simbole v vrstnem redu pojavitve v nizu

$$f_G^0(A_0) f_G^1(A_0) f_G^2(A_0) \ldots f_G^{|V(G)|-1}(A_0),$$

dobimo urejen seznam $A_0, A_1, A_2, \ldots, A_{|V(G)|-1}$.

Pretvorba niza v gramatiko

Opomba

Naj bo $G \in \mathcal{G}(\mathcal{A})$ in sta izpolnjena prva dva pogoja zgornje definicije. Potem lahko enolično preimenujemo spremenljivke iz $V(G)$ tako, da dobimo gramatiko, ki izpolnjuje tudi ostale tri pogoje. Pridelano gramatiko označimo z $[G]$ in očitno velja $[G] \in \mathcal{G}(\mathcal{A})$.

Pretvorba niza v gramatiko

Definicija

Naj bo $\mathcal{A}$ abeceda in G kontekstno-neodvisna gramatika. *Pretvorba niza v gramatiko* je preslikava

$$\mathcal{A}^+ \to \mathcal{G}(\mathcal{A}),$$
$$w \mapsto G_w.$$

Pretvorba niza v gramatiko

Definicija

Naj bo $\mathcal{A}$ abeceda in G kontekstno-neodvisna gramatika. *Pretvorba niza v gramatiko* je preslikava

$$\begin{aligned} \mathcal{A}^+ &\to \mathcal{G}(\mathcal{A}), \\ w &\mapsto G_w. \end{aligned}$$

Z $|G|$ označimo skupno število vseh desnih članov produkcijskih pravil gramatike vključno s ponovitvami. Pretvorba niza v gramatiko je *asimptotsko kompaktna*, če je

$$\lim_{n\to\infty} \max_{w\in\mathcal{A}^n} \frac{|G_w|}{|w|} = 0.$$

Biskecijska pretvorba niza v gramatiko

Definicija

Naj bo $w = u_1 u_2 \cdots u_n \in \mathcal{A}^+$. Vpeljemo *bisekcijsko členitev niza*

$$B(w) = \{w\} \cup \{u_i \cdots u_j \mid \tfrac{i-1}{j-i+1} \text{ in } \log_2(j-i+1) \text{ sta celi števili}\}.$$

Biskecijska pretvorba niza v gramatiko

Definicija

Naj bo $w = u_1 u_2 \cdots u_n \in \mathcal{A}^+$. Vpeljemo *bisekcijsko členitev niza*

$$B(w) = \{w\} \cup \{u_i \cdots u_j \mid \tfrac{i-1}{j-i+1} \text{ in } \log_2(j-i+1) \text{ sta celi števili}\}.$$

Za vsak podniz $u \in B(w)$ z $(b(u,L), b(u,R))$ označimo členitev podniza, kjer sta niza $b(u,L)$ in $b(u,R)$ enake dolžine.

Biskecijska pretvorba niza v gramatiko

Biskecijska členitev niza

Naj bo $w = 0001010$, pote je

$$B(w) = \{w, 0001, 01, 0, 00, 1\}.$$

Definicija

Bisekcijska pretvorba niza v gramatiko je pretvorba niza v gramatiko $w \mapsto [G_w]$, kjer je

Definicija

Bisekcijska pretvorba niza v gramatiko je pretvorba niza v gramatiko $w \mapsto [G_w]$, kjer je

- $V(G_w) = \{A^u \mid u \in B(w)\}$,

Definicija

Bisekcijska pretvorba niza v gramatiko je pretvorba niza v gramatiko $w \mapsto [G_w]$, kjer je

- $V(G_w) = \{A^u \mid u \in B(w)\}$,
- $\Sigma(G_w) = \{u \in B(w) \mid |u| = 1\}$,

Definicija

Bisekcijska pretvorba niza v gramatiko je pretvorba niza v gramatiko $w \mapsto [G_w]$, kjer je

- $V(G_w) = \{A^u \mid u \in B(w)\}$,
- $\Sigma(G_w) = \{u \in B(w) \mid |u| = 1\}$,
- $S(G_w) = A^w$.

Definicija

Množica $P(G_w)$ vsebuje sledeča produkcijska pravila

- Če je $|u| = 1$, je produkcijsko pravilo oblike

$$A^u \to u.$$

Definicija

Množica $P(G_w)$ vsebuje sledeča produkcijska pravila

- Če je $|u| = 1$, je produkcijsko pravilo oblike

$$A^u \to u.$$

- Če je $\log_2 |u|$ celo število, je produkcijsko pravilo oblike

$$A^u \to A^{b(u,L)} A^{b(u,R)}.$$

Definicija

Množica $P(G_w)$ vsebuje sledeča produkcijska pravila

- Če je $|u| = 1$, je produkcijsko pravilo oblike

$$A^u \to u.$$

- Če je $\log_2 |u|$ celo število, je produkcijsko pravilo oblike

$$A^u \to A^{b(u,L)} A^{b(u,R)}.$$

- Če $\log_2 |u|$ ni celo število, torej je $u = w$, je produkcijsko pravilo oblike

$$A^u \to A^{u_1} A^{u_2} \cdots A^{u_t},$$

  kjer je $(u_1, u_2, \ldots, u_t)$ enolična členitev niza u, da je $\forall i \colon u_i \in B(w)$ in $|u| > |u_1| > |u_2| > \cdots > |u_t|$.

Biskecijska pretvorba niza $w = 0001010$ v gramatiko

Vemo, da je $B(w) = \{w, 0001, 01, 0, 00, 1\}$

Biskecijska pretvorba niza $w = 0001010$ v gramatiko

Vemo, da je $B(w) = \{w, 0001, 01, 0, 00, 1\}$ in pravila G_w so

Biskecijska pretvorba niza $w = 0001010$ v gramatiko

Vemo, da je $B(w) = \{w, 0001, 01, 0, 00, 1\}$ in pravila G_w so

$$A^w \to A^{0001} A^{01} A^0$$

Biskecijska pretvorba niza $w = 0001010$ v gramatiko

Vemo, da je $B(w) = \{w, 0001, 01, 0, 00, 1\}$ in pravila G_w so

$$
\begin{aligned}
A^w &\to A^{0001}A^{01}A^0 \\
A^{0001} &\to A^{00}A^{01}
\end{aligned}
$$

Biskecijska pretvorba niza $w = 0001010$ v gramatiko

Vemo, da je $B(w) = \{w, 0001, 01, 0, 00, 1\}$ in pravila G_w so

$$
\begin{aligned}
A^w &\to A^{0001}A^{01}A^0 \\
A^{0001} &\to A^{00}A^{01} \\
A^{00} &\to A^0A^0
\end{aligned}
$$

Biskecijska pretvorba niza $w = 0001010$ v gramatiko

Vemo, da je $B(w) = \{w, 0001, 01, 0, 00, 1\}$ in pravila G_w so

$$
\begin{aligned}
A^w &\to A^{0001}A^{01}A^0 \\
A^{0001} &\to A^{00}A^{01} \\
A^{00} &\to A^0A^0 \\
A^{01} &\to A^0A^1
\end{aligned}
$$

Biskecijska pretvorba niza $w = 0001010$ v gramatiko

Vemo, da je $B(w) = \{w, 0001, 01, 0, 00, 1\}$ in pravila G_w so

$$
\begin{aligned}
A^{w} &\to A^{0001}A^{01}A^{0} \\
A^{0001} &\to A^{00}A^{01} \\
A^{00} &\to A^{0}A^{0} \\
A^{01} &\to A^{0}A^{1} \\
A^{0} &\to 0
\end{aligned}
$$

Biskecijska pretvorba niza $w = 0001010$ v gramatiko

Vemo, da je $B(w) = \{w, 0001, 01, 0, 00, 1\}$ in pravila G_w so

$$
\begin{aligned}
A^w &\to A^{0001}A^{01}A^0 \\
A^{0001} &\to A^{00}A^{01} \\
A^{00} &\to A^0A^0 \\
A^{01} &\to A^0A^1 \\
A^0 &\to 0 \\
A^1 &\to 1
\end{aligned}
$$

Biskecijska pretvorba niza $w = 0001010$ v gramatiko

Vemo, da je $B(w) = \{w, 0001, 01, 0, 00, 1\}$ in pravila G_w so

$$A^w \to A^{0001}A^{01}A^0 \tag{1}$$
$$A^{0001} \to A^{00}A^{01} \tag{2}$$
$$A^{00} \to A^0A^0 \tag{3}$$
$$A^{01} \to A^0A^1 \tag{4}$$
$$A^0 \to 0 \tag{5}$$
$$A^1 \to 1 \tag{6}$$

Biskecijska pretvorba niza v gramatiko w = 0001010

Vemo, da je $B(w) = \{w, 0001, 01, 0, 00, 1\}$ in pravila $[G_w]$ so

$$A_0 \to A_1A_2A_3 \tag{1}$$
$$A_1 \to A_4A_2 \tag{2}$$
$$A_2 \to A_3A_5 \tag{4}$$
$$A_3 \to 0 \tag{5}$$
$$A_4 \to A_3A_3 \tag{3}$$
$$A_5 \to 1 \tag{6}$$