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-neodivsnih gramatik želimo stisniti poljuben niz.

Spomnimo se

S kontekstno-neodivsnih 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-neodivsnih 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

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

Kontekstno-neodvisne gramatike

Definicija

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

- V končna množica *nekončnih simbolov*,

Kontekstno-neodvisne gramatike

Definicija

Kontektsno-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

Kontektsno-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

Kontektsno-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

Kontektsno-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

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

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

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

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

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

Stiskanje niza s kontekstno-neodvisno gramatiko

Stiskanje niza $w = cababcccababcccab$

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

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

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

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 člen 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 člen 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 produckcijska 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 produckcijska 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 produckcijska 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 produckcijska 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 $f\colon \Sigma^* \to \Sigma^*$ 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 generira 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

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

če je $A \to \alpha$ 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 $(V \cup \Sigma,\ f_G,\ S)$ označimo z $D(G)$ in ga 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 produckcijska 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 produckcijska 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 producksijska 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 produckcijska 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 produckcijska 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 $\Sigma = \{a, b\}$

$$f_G(A_0) = aA_1A_2A_3$$
$$f_G^2(A_0) = aabA_1bA_2b$$
$$f_G^3(A_0) = aababbA_1bb$$
$$f_G^4(A_0) = aababbabbb$$

Preverimo, da je gramatika dopustna in izračunajmo jezik

Podana imamo produckcijska 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 produckcijska 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 $\Sigma = \{a, b\}$

$$f_G(A_0) = aA_1A_2A_3$$
$$f_G^2(A_0) = aabA_1bA_2b$$
$$f_G^3(A_0) = aababbA_1bb$$
$$f_G^4(A_0) = aababbabbb$$

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

Posplošitev izreka

Izrek

Naj bo G dopustna kontekstno-neodvisna gramatika in naj bo $\alpha \in (V \cup \Sigma)^+$. Potem ima D0L sistem $(V \cup \Sigma, f_G, \alpha)$ fiksno točko $w^* \in \Sigma^+$ in velja formula

$$w^* = f_G^{|V|}(\alpha).$$

Uporaben endomorfizem

Definicija

Naj bo G dopustna kontekstno-neodvisna gramatika. Definiramo preslikavo $f_G^\infty : (V \cup \Sigma)^* \to (V \cup \Sigma)^*$, ki vsak $\alpha \in (V \cup \Sigma)^*$ preslika v fiksno točko D0L sistema $(V \cup \Sigma,\ f_G^\infty,\ \alpha)$.

Preobrazba niza

Z $\mathcal{A}$ označimo poljubno abecedo z vsaj dvema črkama in fiksiramo končno mnogo 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 $\mathcal{A}$.

Preobrazba niza

Definicija

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

1. G je dopustna;

Preobrazba niza

Definicija

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

1. G je dopustna;
2. $V(G) \subset \mathcal{A}$;

Preobrazba niza

Definicija

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

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

Preobrazba niza

Definicija

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

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

Preobrazba niza

Definicija

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

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

Preobrazba niza

Opomba

Naj bo $G \in \mathcal{G}(\mathcal{A})$ in sta izpoljena 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 $[G] \in \mathcal{G}(\mathcal{A})$.

Preobrazba niza

Definicija

Naj bo $\mathcal{A}$ abeceda. Z $\mathcal{G}^*(\mathcal{A})$ označimo pravo podmnožico množice $\mathcal{G}(\mathcal{A})$ za katero velja, da za vsaka nekončna simbola $A, B \in V(G),\ A \neq B$ velja

$$f_G^\infty(A) \neq f_G^\infty(B).$$

Preobrazba niza

Definicija

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

$$\begin{aligned} \mathcal{A}^+ &\to \mathcal{G}^*(\mathcal{A}), \\ x &\mapsto G_x. \end{aligned}$$

Preobrazba niza

Definicija

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

$$\begin{aligned} \mathcal{A}^+ &\to \mathcal{G}^*(\mathcal{A}), \\ x &\mapsto G_x. \end{aligned}$$

Z $|G|$ označimo skupno število vseh desnih členov produkcijskih pravil gramatike vključno s ponovitvami. Preobrazba niza je *asimptotsko kompaktna*, če je

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

Biskecijska preobrazba niza

Definicija

Naj bo $\boldsymbol{x} = x_1 x_2 \cdots x_n \in \mathcal{A}^+$. Vpeljemo *bisekcijsko členitev niza*

$$B(\boldsymbol{x}) = \{\boldsymbol{x}\} \cup \{x_i \cdots x_j \mid \tfrac{i-1}{j-i+1} \text{ in } \log_2(j-i+1) \text{ sta celi števili}\}.$$

Biskecijska preobrazba niza

Definicija

Naj bo $\boldsymbol{x} = x_1 x_2 \cdots x_n \in \mathcal{A}^+$. Vpeljemo *bisekcijsko členitev niza*

$$B(\boldsymbol{x}) = \{\boldsymbol{x}\} \cup \{x_i \cdots x_j \mid \tfrac{i-1}{j-i+1} \text{ in } \log_2(j-i+1) \text{ sta celi števili}\}.$$

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

Biskecijska preobrazba niza

Biskecijska členitev niza

Naj bo $\boldsymbol{x} = 0001010$, pote je

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

Definicija

Bisekcijska preobrazba niza je preobrazba niza $\boldsymbol{x} \mapsto [G_{\boldsymbol{x}}]$, kjer je

Definicija

Bisekcijska preobrazba niza je preobrazba niza $\boldsymbol{x} \mapsto [G_{\boldsymbol{x}}]$, kjer je

- $V(G_{\boldsymbol{x}}) = \{A^w \mid w \in B(\boldsymbol{x})\}$,

Definicija

Bisekcijska preobrazba niza je preobrazba niza $\boldsymbol{x} \mapsto [G_{\boldsymbol{x}}]$, kjer je

- $V(G_{\boldsymbol{x}}) = \{A^w \mid w \in B(\boldsymbol{x})\}$,
- $T(G_{\boldsymbol{x}}) = \{w \in B(\boldsymbol{x}) \mid |w| = 1\}$,

Definicija

Bisekcijska preobrazba niza je preobrazba niza $\boldsymbol{x} \mapsto [G_{\boldsymbol{x}}]$, kjer je

- $V(G_{\boldsymbol{x}}) = \{A^w \mid w \in B(\boldsymbol{x})\}$,
- $T(G_{\boldsymbol{x}}) = \{w \in B(\boldsymbol{x}) \mid |w| = 1\}$,
- $S(G_{\boldsymbol{x}}) = A^{\boldsymbol{x}}$.

Definicija

Množica $P(G_x)$ vsebuje sledeča produkcijska pravila

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

$$A^w \to w.$$

Definicija

Množica $P(G_x)$ vsebuje sledeča produkcijska pravila

- Če je $|u| = 1$, je produkcijsko pravilo oblike
$$A^w \to w.$$
- Če je $\log_2 |w|$ celo število, je produkcijsko pravilo oblike
$$A^w \to A^{b(w,L)} A^{b(w,R)}.$$

Definicija

Množica $P(G_{\boldsymbol{x}})$ vsebuje sledeča produkcijska pravila

- Če je $|u| = 1$, je produkcijsko pravilo oblike
$$A^w \to w.$$
- Če je $\log_2 |w|$ celo število, je produkcijsko pravilo oblike
$$A^w \to A^{b(w,L)} A^{b(w,R)}.$$
- Če $\log_2 |w|$ ni celo število, torej je $w = \boldsymbol{x}$, je produkcijsko pravilo oblike
$$A^w \to A_1^w A_2^w \cdots A_t^w,$$
kjer je $(w_1, w_2, \ldots, w_t)$ enolična členitev niza $\boldsymbol{x}$, da je $\forall i\colon w_i \in B(\boldsymbol{x})$ in $|x| > |u_1| > |u_2| > \cdots > |u_t|$.

Biskecijska preobrazba niza $\boldsymbol{x}$= 0001010

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

Biskecijska preobrazba niza $\boldsymbol{x}$= 0001010

Vemo, da je $B(\boldsymbol{x}) = \{\boldsymbol{x}, 0001, 01, 0, 00, 1\}$ in produkcijska pravila $G_{\boldsymbol{x}}$ so

Biskecijska preobrazba niza $\boldsymbol{x}$= 0001010

Vemo, da je $B(\boldsymbol{x}) = \{\boldsymbol{x}, 0001, 01, 0, 00, 1\}$ in produkcijska pravila $G_{\boldsymbol{x}}$ so

$$A^{\boldsymbol{x}} \to A^{0001} A^{01} A^{0}$$

Biskecijska preobrazba niza x= 0001010

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

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

Biskecijska preobrazba niza x= 0001010

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

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

Biskecijska preobrazba niza $\boldsymbol{x}$= 0001010

Vemo, da je $B(\boldsymbol{x}) = \{\boldsymbol{x}, 0001, 01, 0, 00, 1\}$ in produkcijska pravila $G_{\boldsymbol{x}}$ so

$$
\begin{aligned}
A^{\boldsymbol{x}} &\rightarrow A^{0001}A^{01}A^{0} \\
A^{0001} &\rightarrow A^{00}A^{01} \\
A^{00} &\rightarrow A^{0}A^{0} \\
A^{01} &\rightarrow A^{0}A^{1} \\
A^{0} &\rightarrow 0
\end{aligned}
$$

Biskecijska preobrazba niza $\boldsymbol{x}$= 0001010

Vemo, da je $B(\boldsymbol{x}) = \{\boldsymbol{x}, 0001, 01, 0, 00, 1\}$ in produkcijska pravila $G_{\boldsymbol{x}}$ so

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

Biskecijska preobrazba niza x= 0001010

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

$$
\begin{aligned}
A_0 &\to A_1A_2A_3\\
A_1 &\to A_4A_2\\
A_2 &\to A_3A_5\\
A_3 &\to 0\\
A_4 &\to A_3A_3\\
A_5 &\to 1
\end{aligned}
$$