Katedra Informatiky
Fakulta Matematiky, Fyziky a Informatiky
Univerzita Komenského, Bratislava

Paralelizmus v grafových gramatikách

(Diplomová práca)

Daniel Buchta 2008

Paralelizmus v grafových gramatikách

DIPLOMOVÁ PRÁCA

Daniel Buchta

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
KATEDRA INFORMATIKY

Informatika

Vedúci diplomovej práce
Prof. RNDr. Branislav Rovan, PhD.

BRATISLAVA 2008

Čestne vyhlasujem, že som túto diplomovú prácu vypracoval samostatne, s použitím uvedenej literatúry.

. . . .

Ďakujem môjmu diplomovému vedúcemu Prof. RNDr. Branislavovi Rovanovi, PhD. za odborné vedenie práce, cenné rady a pripomienky. Ďakujem mojej rodine za podporu pri písaní tejto práce.

Abstrakt

Autor:	Daniel Buchta
Názov diplomovej práce:	Paralelizmus v grafových gramatikách
Škola:	Univerzita Komenského v Bratislave
Fakulta:	Fakulta matematiky, fyziky a informatiky
Katedra:	Katedra informatiky
Vedúci diplomovej práce:	Prof. RNDr. Branislav Rovan, PhD.
Rozsah práce:	58 strán

Bratislava, máj 2008

Práca analyzuje problematiku zavedenia paralelizmu v grafových gramatikách. Podrobnejšie v nej skúmame dva paralelné modely (systémy gramatík), ktorých definície sme zovšeobecnili pre grafové gramatiky. Zaujíma nás porovnanie generatívnej sily týchto systémov vzhľadom na rôzne typy komponentových gramatík a platnosť analogických tvrdení pre zodpovedajúce paralelné modely definované pre stringové gramatiky.

Kľúčové slová

grafové gramatiky, paralelizmus, kooperujúci distribuovaný systém grafových gramatík, paralelný komunikujúci systém grafových gramatík

Predhovor

Cieľom tejto diplomovej práce je analyzovať existujúce typy grafových gramatík a preskúmať možnosti využitia paralelizmu v nich. Spomedzi rôznych tried grafových gramatík (s rôznym prístupom k transformácii grafov) v tejto práci skúmame triedu $edNCE$ gramatík a niektoré jej podtriedy. Tieto typy gramatík (hlavne trieda konfluentných $edNCE$ gramatík) predstavujú dobré zovšeobecnenie bezkontextových stringových gramatík (ktoré patria medzi hlavné oblasti výskumu v prípade rôznych paralelných modelov).

Ukazuje sa, že aplikovať niektoré hlavné myšlienky fungujúce pre paralelné modely definované pre stringové gramatiky je možné aj v prípade analogických modelov definovaných pre grafové gramatiky. Problémom (nie len technického charakteru) je napríklad spomínaná konfluentnosť, čo je vlastnosť, ktorá nie je garantovaná len formátom pravidiel (ako je to v prípade bezkontextovosti u stringových gramatík).

Obsah

Zoznam obrázkov

Kapitola 1

Úvod

Grafové gramatiky vznikli na konci šestdesiatych rokov. Motiváciou k ich vzniku boli problémy ako rozpoznávanie vzorov, konštrukcia kompilátorov a špecifikácia dátových typov. Neskôr sa počet oblastí ich aplikácie výrazne rozšíril. Grafové gramatiky sú zaujímavé aj z teoretického hľadiska, keďže predstavujú prirodzené zovšeobecnenie teórie formálnych jazykov založenej na reťazcoch a teórie prepisovania termov založenej na stromoch. Tento model vlastne poskytuje mechanizmus, ktorým lokálne transformácie na grafe môžu byť modelované matematicky presným spôsobom. Dva základné prístupy k transformácii grafov sú nahradzovanie vrcholov a nahradzovanie (hyper)hrán. V tejto práci sa budeme zaoberať prvým z týchto prístupov. Hlavnou časťou takejto grafovej gramatiky je konečná množina pravidiel. Pravidlo je vo všeobecnosti trojica (M, D, E), kde M a D sú grafy ("materský" respektíve "dcérsky" graf) a E je spôsob vnorenia (v anglických textoch sa používa výraz "embedding"). Takéto pravidlo môže byť aplikované na graf H ak sa graf M vyskytuje v H. V tomto prípade graf M je odstránený z H, nahradený izomorfnou kópiou grafu D, pričom je použitý spôsob vnorenia E na pripojenie grafu D ku zvyšku H^- grafu H.

Boli skúmané viaceré spôsoby vnorenia. V tejto práci sa budeme zaoberať výhradne spôsobom, keď určité nové hrany sú použité na spojenie D so zvyškom H^- t.j. hrany ktorých jeden vrchol patrí do D a druhý do H^-. Keď je M odstránené z H, všetky hrany medzi M a vrcholmi z H, ktoré nepatria do M, sú odstránené tiež.

Hlavným objektom nášho záujmu budú bezkontextové grafové gramatiky, ktoré majú pravidlá tvaru (M, D, E), kde M pozostáva len z jedného vrchola t.j. v tomto prípade sú vrcholy nahrádzané grafmi. Ďalšou vlastnosťou, ktorú potrebujeme, aby definovaná trieda grafových gramatík bola dobrým zovšeobecnením bezkontextových stringových gramatík je, aby výsledok postupnosti nahradzovania vrcholov nebol závislý na poradí, v ktorom sú pravidlá použité. Toto robí bezkontextové grafové gramatiky vhodnými na popis rekurzívne definovateľných množín grafov.

V spomínanom bezkontextovom prípade, keď materský graf pozostáva len z jedného uzla a vloženie dcérskeho grafu D je urobené lokálne (spojením grafu D s vrcholmi hostitelského grafu, ktoré sú "blízko" materského vrchola), tak krok prepísania vrchola je vlastne lokálna transformácia grafu. Iterácia takýchto krokov vedie ku globálnej transformácii grafu, ktorá je založená na lokálnych transformáciach. To je základná idea grafových gramatík založených na prepisovaní vrcholov.

Typickým, veľmi jednoduchým príkladom takéhoto mechanizmu sú gramatiky s odvodením riadeným symbolmi vrcholov (Node Label Controlled - NLC) definované napríklad v [5]. V NLC gramatike sa prepisujú neorientované grafy, ktorých vrcholom sú priradené symboly z nejakej (konečnej) abecedy, pravidlá gramatiky definujú nahradenie vrchola a spôsob, akým sa vrcholy dcérskeho grafu spoja so susedmi materského uzla - teda proces prepísania grafu je úplne lokálny. V NLC prístupe je všetko založené na symboloch vrcholov. NLC pravidlo je tvaru $X \to D$, kde X je (neterminálny) symbol vrchola a D je neorientovaný graf s (terminálnymi alebo neterminálnymi) symbolmi vrcholov. Takéto pravidlo môže byť aplikované na ľubovoľný vrchol v so symbolom X. Výsledkom aplikácie tohto pravidla je nahradenie (materského) vrchola v dcérskym grafom D. Všetky pravidlá zdielajú spoločnú množinu spojovacích inštrukcií. Spojovacia inštrukcia je tvaru (μ, δ), kde μ aj δ sú (terminálne alebo neterminálne) symboly vrcholov. Takáto inštrukcia hovorí, že pri kroku odvodenia sa má vytvoriť hrana medzi každým vrcholom so symbolom δ v dcérskom grafe D a každým vrcholom so symbolom μ, ktorý je susedom materského vrchola v. Gramatika žiadnym spôsobom "nekontroluje", či taká hrana existuje - ak v nemá žiadneho μ suseda, potom inštrukcia (μ, δ) zostane nevyužitá.

NLC grafové gramatiky predstavujú jeden z pokusov ako definovať triedu bezkontextových grafových gramatík t.j. grafových gramatík, ktoré sú podobné bezkontextovým stringovým gramatikám. NLC gramatiky su bezkontextové v tom, že krok odvodenia je úplne lokálny a nie sú kladené žiadne podmienky na aplikáciu pravidiel. To umožňuje napríklad modelovať ich odvodenie pomocou stromov odvodenia. Bohužiaľ, NLC gramatiky nemajú tú (dôležitú) bezkontextovú vlastnosť, že výsledok aplikácie viacerých pravidiel nie je závislý na poradí, v ktorom sú tieto pravidlá aplikované. Gramatika, ktorá túto vlastnosť má, sa nazýva konfluentá.

Myšlienka NLC gramatík môže byť zovšeobecnená viacerými spôsobmi. Je celkom prirodzené odkazovať sa v spojovacích inštrukciách priamo na vrcholy dcérskeho grafu namiesto jeho symbolu, čo umožňuje rozlišovať medzi vrcholmi dcérskeho grafu. Takýto typ gramatiky sa potom označuje NCE grafová gramatika (NCE - Neighbourhood Controlled Embedding). Ďalšou možnosťou zovšeobecnenia je rozšírenie NLC gramatík na orientované grafy. Spojovacia relácia bude potom obsahovať trojice (μ, δ, d), kde $d \in \{in, out\}$. Význam takejto inštrukcie v prípade, že $d = in$ je, že pri prepisovaní grafu sa má vytvoriť hrana do každého vrchola so symbolom δ v dcérskom grafe D z každého vrchola so symbolom μ, z ktorého vedie hrana do materského vrchola v v prepisovanom grafe (analogicky je to v prípade, že $d = out$). Spomínané rozšírenia NLC prístupu umožňujú rozlišovať medzi jednotlivými vrcholmi dcérskeho grafu D resp. medzi "in" a "out" susedmi materského vrchola v. Ešte väčšiu rozlišovaciu silu je možné gramatike dodať tým, že ju rošírime na (orientované alebo neorientované) grafy, ktoré okrem symbolov vrcholov budú mať aj hranové symboly. To rozdelí susedov materského vrchola v na viacero rozdielnych typov podľa symbolu hrany, ktorá spája suseda vrcholu v s vrcholom v. Prirodzená myšlienka je umožniť prepisovaciemu mechanizmu zmeniť typ týchto susedov, čo vedie k inštrukciám typu $(\mu, p/q, \delta)$, kde p a q sú hranové symboly. Význam takejto inštrukcie je, že sa má vytvoriť hrana so symbolom q medzi každým p susedom so symbolom μ materského vrchola v a každým vrcholom so symbolom δ dcérskeho grafu D. Teda hranový symbol p je zmenený na q.

Kombinácia všetkých spomínaných vlastností vedie ku triede C-$edNCE$ grafových gramatík t.j. konfluentných $edNCE$ grafových gramatík. Každé pravidlo $edNCE$ grafovej gramatiky je tvaru $X \to (D, C)$ a každá spojovacia inštrukcia C je tvaru $(\mu, p/q, x, d)$, kde μ je symbol vrchola, p a q sú hranové symboly, x je vrchol grafu D a $d \in \{in, out\}$. Ak je napríklad $d = in$, tak takáto inštrukcia je interpretovaná nasledovne: hrana so symbolom q má byť vytvorená do vrchola x grafu D z každého p suseda so symbolom μ materského vrchola v, z ktorého vedie hrana do vrchola v. Formálne zavedieme triedu C-$edNCE$ a niektoré jej významné podtriedy v ďalšej časti.

Jedným zo smerov, ktorým sa uberal vývoj klasických typov gramatík bolo skúmanie vlastností rôznych paralelných modelov. Za posledných päťdesiat rokov boli v oblasti gramatík definované rôzne typy paralelných modelov, spomeňme napríklad G systémy zavedené v [3] umožňujúce simulovať rôzne typy sekvenčných a paralelných modelov, kooperujúce distribuované systémy gramatík (prvý krát definované v [2]), ktoré v sebe kombinujú sekvenčný a paralelný výpočet, paralelné kooperujúce systémy gramatík definované v [1], ktoré umožňujú študovať rôzne typy komunikácie a komunikačných štruktúr, rôzne typy OL systémov . . . V tejto práci nás budú zaujímať vlastnosti niektorých paralelných modelov definovaných nad grafovými gramatikami (budeme analyzovať rôzne typy grafových gramatík, najmä C-$edNCE$, B-$edNCE$ a $edNCE$ grafové gramatiky), hlavne vzťahy platiace medzi rôznymi triedami jazykov definovaných pomocou týchto modelov.

Kapitola 2

Definície

V tejto kapitole formálne zadefinujeme základné typy grafových gramatík a niektoré ďalšie pojmy potrebné v ďalších kapitolách tejto diplomovej práce tak ako je to v [5].

2.1 Základné pojmy

Definícia 2.1.1. *Abeceda je ľubovoľná neprázdna konečná množina symbolov.*

Poznámka 2.1.2. *Abecedu budeme označovať veľkými gréckymi písmenami* Σ, Γ, Ω, Δ, ...

Definícia 2.1.3. *Nech* Σ *je abeceda symbolov vrcholov a* Γ *je abeceda hranových symbolov. Graf nad* Σ *a* Γ *je trojica* $H = (V, E, \lambda)$*, kde* V *je konečná množina vrcholov,* $E \subseteq \{(v, \gamma, w) | v, w \in V, v \neq w, \gamma \in \Gamma\}$ *je množina hrán a* $\lambda : V \to \Sigma$ *je funkcia priraďujúca symboly vrcholom grafu (t.j. uvažujeme orientované grafy bez slučiek, multihrany su povolené, ale musia mať rôzne symboly).*

Definícia 2.1.4. *Dva grafy sú izomorfné, ak existuje bijekcia* $f : V_H \to V_K$ *taká, že* $E_K = \{(f(v), \gamma, f(w)) \ / \ (v, \gamma, w) \in E_H\}$ *a pre všetky* $v \in V_H$ *platí* $\lambda_K(f(v)) = \lambda_H(v)$*. Pre graf* H*, množinu všetkých grafov izomorfných s* H *označujeme ako* $[H]$ *(*H *v tomto prípade nazývame konkrétny graf,* $[H]$ *abstraktný graf)*

Definícia 2.1.5. *Množinu všetkých (konkrétnych) grafov nad* Σ *a* Δ *označujeme* $GR_{\Sigma,\Gamma}$ *a množinu všetkých abstraktných grafov označujeme* $[GR_{\Sigma,\Gamma}]$*. Podmnožinu* $[GR_{\Sigma,\Gamma}]$ *nazývame grafový jazyk.*

Definícia 2.1.6. *Graf s vnorením* [1] *nad* Σ *a* Δ *je dvojica* (H, C)*, kde* $H \in GR_{\Sigma,\Gamma}$ *a* $C \subseteq \Sigma \times \Gamma \times \Gamma \times V_H \times \{in, out\}$*.* C *je spojovacia relácia grafu* (H, C) *a každý prvok* $(\sigma, \beta/\gamma, x, d)$ *relácie* C *nazývame spojovacia inštrukcia.*

Definícia 2.1.7. *Dva grafy s vnorením sú izomorfné, ak existuje izomorfizmus* f *z* H *do* K *taký, že* $C_K = \{(\sigma, \beta/\gamma, f(x), d) \mid (\sigma, \beta/\gamma, x, d) \in C_H\}$*.*

Definícia 2.1.8. *Množinu všetkých grafov s vnorením nad* Σ *a* Γ *označujeme* $GRE_{\Sigma,\Gamma}$*. Každý obyčajný graf môžeme považovať za graf s (prázdnym) vnorením (t.j.* $C = \emptyset$*), teda* $GR_{\Sigma,\Gamma} \subseteq GRE_{\Sigma,\Gamma}$*.*

Definícia 2.1.9. *Nech* (H, C_H) *a* (D, C_D) *sú dva grafy v* $GRE_{\Sigma,\Gamma}$ *také, že* H *a* D *sú disjunktné a nech* q *je vrchol* H*. Substitúcia* (D, C_D) *za vrchol* q *v grafe* (H, C_H) *je graf* (V, E, λ, C) *v* $GRE_{\Sigma,\Gamma}$ *taký, že*

$$
\begin{aligned}
V &= (V_H - \{q\}) \cup V_D \\
E &= E_D \cup \{(x, \gamma, y) \in E_H \mid x \neq q, y \neq q\} \\
& \cup \\
& \{(w, \gamma, x) \mid \exists \beta \in \Gamma : (w, \beta, q) \in E_H, (\lambda_H(w), \beta/\gamma, x, in) \in C_D\} \\
& \cup \\
& \{(x, \gamma, w) \mid \exists \beta \in \Gamma : (q, \beta, w) \in E_H, (\lambda_H(w), \beta/\gamma, x, out) \in C_D\} \\
\lambda(x) &= \lambda_H(x) \text{ ak } x \in V_H - \{q\} \text{ a} \\
\lambda(x) &= \lambda_D(x) \text{ ak } x \in V_D \\
C &= \{(\sigma, \beta/\gamma, x, d) \in C_H \mid x \neq q\} \\
& \cup \\
& \{(\sigma, \beta/\delta, x, d) \mid \exists \in \Gamma : (\sigma, \beta/\gamma, q, d) \in C_H, (\sigma, \gamma/\delta, x, d) \in C_D\}
\end{aligned}
$$

Označujeme ju $(H, C_H)\,[q/(D, C_D)]$*.*

2.2 Grafové gramatiky

Definícia 2.2.1. *edNCE gramatika je 6-tica* $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$*, kde*

- Σ *je abeceda symbolov vrcholov,*
- $\Delta \subseteq \Sigma$ *je abeceda terminálnych symbolov vrcholov,*

[1] Presnejšie povedané, ide graf s inštrukciami na vnorenie do iných grafov

- Γ *je abeceda hranových symbolov,*
- $\Omega \subseteq \Gamma$ *je abeceda terminálnych hranových symbolov,*
- P *je konečná množina pravidiel a*
- $S \in \Sigma - \Delta$ *je počiatočný neterminál.*

Pravidlá sú tvaru $X \to (D, C)$ *pričom* $X \in \Sigma - \Delta$ *a* $(D, C) \in GRE_{\Sigma,\Gamma}$.

Prvky množiny $\mathrm{N} = \Sigma - \Delta$ *nazývame neterminálne symboly vrcholov a prvky množiny* $\Gamma - \Omega$ *nekoncové hranové symboly.*

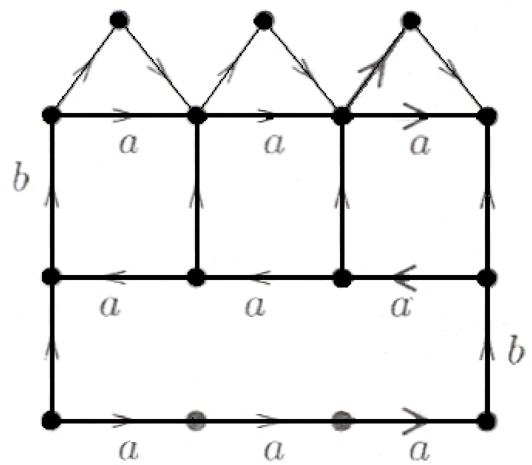
Graf nazývame terminálny, ak všetky jeho vrcholy sú terminálne (t.j. majú terminálne symboly) t.j. patrí do $GRE_{\Delta,\Gamma}$.

Nech $p : X \to (D, C)$ *je pravidlo,* X *nazývame jeho ľavou stranou,* (D, C) *pravou stranou pravidla* p*, označujeme ich ako* $lhs(p)$ *resp.* $rhs(p)$.

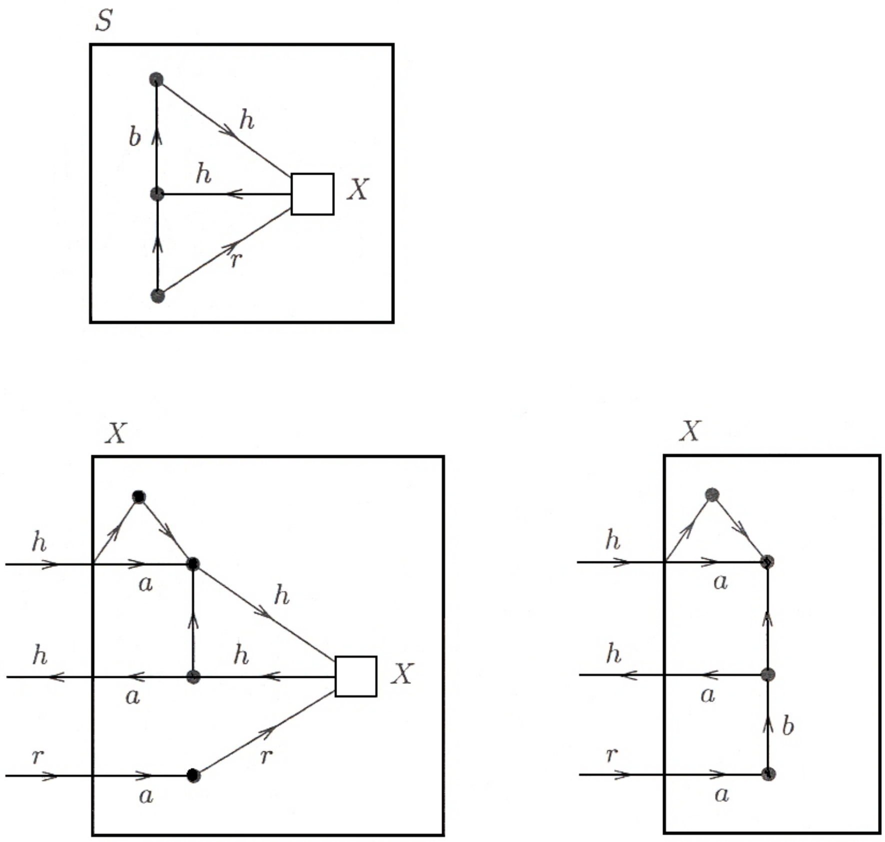
Dve pravidlá $X_1 \to (D_1, C_1)$ *a* $X_2 \to (D_2, C_2)$ *sú izomorfné, ak* $X_1 = X_2$ *a* (D_1, C_1) *je izomorfné s* (D_2, C_2).

Nekonečnú množinu všetkých pravidiel izomorfných s pravidlom P *označujeme* $copy(P)$*, ľubovoľný jej prvok nazývame kópia pravidla.*

Príklad pravidiel grafovej gramatiky, ktoré generuje množinu všetkých "ulíc" tvaru 2.1 je na obrázku 2.2. $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$, kde $\Sigma = \{S, X, +\}$, $\Delta = \{+\}$, $\Gamma = \{h, r, a, b, *\}$, $\Omega = \{a, b, *\}$. Neoznačené vrcholy (hrany) majú symboly $+$ (resp. $*$).



Obrázok 2.1: Ulica



Obrázok 2.2: Pravidlá gramatiky G

Definícia 2.2.2. *Nech $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$ je edNCE gramatika , H a H' sú grafy v $GRE_{\Sigma,\Gamma}$, nech $m \in V_H$ a nech $p : X \to (D, C)$ je kópia pravidla v G taká, že D a H sú disjunktné. Potom píšeme $H \Rightarrow_{m,p} H'$ alebo $H \Rightarrow H'$, ak $\lambda_H(m) = X$ a $H' = H[m/(D, C)]$. $H \Rightarrow H'$ nazývame krok odvodenia a postupnosť týchto krokov nazývame odvodenie.*

Definícia 2.2.3. *Odvodenie $H_0 \Rightarrow_{v_1,p_1} H_1 \Rightarrow_{v_2,p_2} \ldots \Rightarrow_{v_n,p_n} H_n$, $n \geq 0$ je kreatívne, ak graf H_0 a $rhs(p_i)$ pre $1 \leq i \leq n$ sú vzájomne disjunktné (budeme ďalej uvažovať len kreatívne odvodenia). Píšeme $H \Rightarrow^* H'$ ak existuje kreatívne odvodenie uvedené vyššie, kde $H_0 = H$ a $H_n = H'$*

Definícia 2.2.4. *Vetná forma gramatiky G je graf H taký, že $sn(S, z) \Rightarrow^* H$, kde $sn(S, z)$ označuje graf s jediným vrcholom z so symbolom S.*

Definícia 2.2.5. *Grafový jazyk generovaný gramatikou G je množina*

$$L(G) = \{[H] \mid H \in GR_{\Delta,\Omega}, \exists z : sn(S, z) \Rightarrow^* H\}$$

Definícia 2.2.6. *Dve edNCE gramatiky G a G' sú ekvivalentné, ak $L(G) - \{\Lambda\} = L(G') - \{\Lambda\}$, kde Λ je prázdny graf.*

Definícia 2.2.7. *edNCE gramatiku $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$ nazývame konfluentnou alebo C-edNCE gramatikou, ak pre všetky pravidlá $X_1 \to (D_1, C_1)$ a $X_2 \to (D_2, C_2)$ v P, pre všetky vrcholy $x_1 \in V_{D_1}$ a $x_2 \in V_{D_2}$ a všetky hranové symboly $\alpha, \delta \in \Gamma$ platí nasledujúca ekvivalencia*

$$\begin{array}{lcl} \exists\beta \in \Gamma & : & (X_2, \alpha/\beta, x_1, out) \in C_1 \text{ a } (\lambda_{D_1}(x_1), \beta/\delta, x_2, in) \in C_2 \\ \Longleftrightarrow & & \\ \exists\gamma \in \Gamma & : & (X_1, \alpha/\gamma, x_2, in) \in C_2 \text{ a } (\lambda_{D_2}(x_2), \gamma/\delta, x_1, out) \in C_1. \end{array}$$

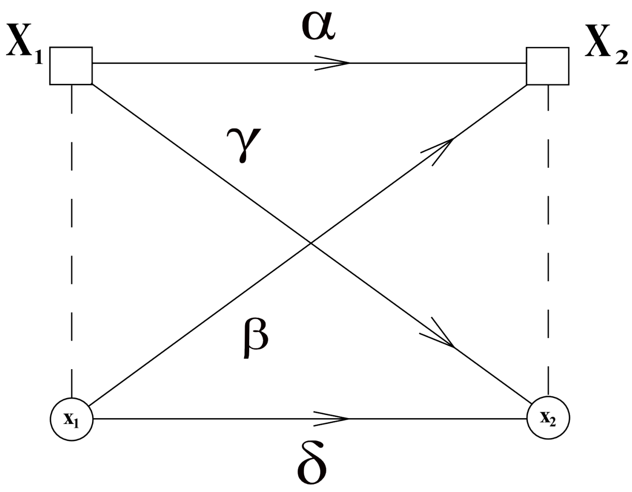
Symbolické znázornenie definície 2.2.7 je na obrázku 2.3.

Veta 2.2.8. *edNCE gramatika $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$ je konfluentná práve vtedy, keď pre každý graf $H \in GR_{\Sigma,\Gamma}$ (resp. $H \in GRE_{\Sigma,\Gamma}$), $u_1, u_2 \in V_H$ a $u_1 \neq u_2$, platí nasledujúca implikácia:*
ak

$$H \Rightarrow_{u_1,p_1} H_1 \Rightarrow_{u_2,p_2} H_{12} \quad \text{a} \quad H \Rightarrow_{u_2,p_2} H_2 \Rightarrow_{u_1,p_1} H_{21}$$

sú odvodenia v G, potom

$$H_{12} = H_{21}$$



Obrázok 2.3: Konfluentnosť

Dôkaz. Pozri tvrdenie 1.3.6 v [5]. □

Definícia 2.2.9. *edNCE gramatiku* $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$ *nazývame dynamicky konfluentnou, ak nasledujúca implikácia platí pre všetky vetné formy* H *gramatiky* G*: ak*

$$H \Rightarrow_{u_1,p_1} H_1 \Rightarrow_{u_2,p_2} H_{12}$$

a

$$H \Rightarrow_{u_2,p_2} H_2 \Rightarrow_{u_1,p_1} H_{21}$$

sú (kreatívne) odvodenia v G*,* $u_1, u_2 \in V_H$ *a* $u_1 \neq u_2$*, potom* $H_{12} = H_{21}$*.*

Definícia 2.2.10. *Usporiadaný graf (s vnorením) je graf spolu s lineárnym usporiadaním jeho vrcholov. Ak grafy* H *a* D *sú disjunktné grafy (s vnorením), ktoré sú usporiadané a* q *je vrchol grafu* H*, potom usporiadanie substitúcie* $H[q/D]$ *je získané (neformálne povedané) substituovaním usporiadania grafu* D *za* q *v usporiadaní grafu* H*, tak ako pre reťazce. Formálne, ak vrcholy grafu* H *sú usporiadané ako* $(v_1, v_2, \ldots, v_h)$*, kde* $q = v_i$ *a vrcholy grafu* D *sú usporiadané ako* $(w_1, w_2, \ldots, w_d)$*, potom usporiadanie grafu* $H[q/D]$ *je* $(v_1, v_2, \ldots, v_{i-1}, w_1, \ldots, w_d, v_{i+1}, \ldots, v_h)$*.*

Poznámka 2.2.11. *Ak* G *je edNCE grafová gramatika môžeme pravé strany jej pravidiel považovať za usporiadané grafy (s nejakým lineárnym usporiadaním ich vrcholov). Keďže kroky odvodenia gramatiky* G *sú definované pomocou substitúcie, vetné formy gramatiky* G *sú tak isto usporiadané grafy. Zrejme usporiadanie vetných foriem nemá vplyv na jazyk* $L(G)$ *generovaný* *gramatikou G. Pre odvodenie $H_0 \Rightarrow_{v_1,p_1} H_1 \Rightarrow_{v_2,p_2} \ldots \Rightarrow_{v_n,p_n} H_n$, kde H_0 je usporiadaný graf, predpokladáme, že usporiadanie H_1 až H_n je získané tak, ako je uvedené vyššie.*

Definícia 2.2.12. *Nech G je edNCe gramatika a nech H je usporiadaný graf. Krok odvodenia $H \Rightarrow_{v,p} H'$ v G nazývame ľavým krajným krokom odvodenia, ak v je prvým neterminálnym vrcholom v usporiadaní grafu H. Odvodenie nazývame ľavé krajné odvodenie, ak sú také všetky jeho kroky. Píšeme $H \Rightarrow_{lm}^{*} H'$, ak existuje ľavé krajné odvodenie grafu H' z grafu H.*

Definícia 2.2.13. *edNCE gramatiku $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$ nazývame hraničnou alebo B-edNCE gramatikou, ak pre každé pravidlo $X \to (D, C)$ platí*

1. *D neobsahuje hranu medzi neterminálnymi vrcholmi a*
2. *C neobsahuje inštrukciu $(\sigma, \beta/\gamma, x, io)$, kde σ je neterminál*

Poznámka 2.2.14. *V definícii B-edNCE je možné jednu z podmienok (1) alebo (2) vynechať v tom zmysle, že ku každej edNCE gramatike spĺňajúcu jednu z podmienok môže byť skonštruovaná ekvivalentná gramatika spĺňajúca obidve podmienky.*

Definícia 2.2.15. *edNCE gramatiku $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$ nazývame lineárnou alebo LIN-edNCE gramatikou, ak pre každé pravidlo $X \to (D, C)$ platí, že D má najviac jeden neterminálny vrchol.*

Označenie 2.2.16. *Trieda jazykov generovaných triedou X-edNCE gramatík označujeme tiež ako X-edNCE.*

Poznámka 2.2.17. *Ľahko vidieť, že LIN-edNCE $\subseteq$ B-edNCE, dokonca každá LIN-edNCE gramatika je B-edNCE gramatika.*

Kapitola 3

Kooperujúce Distribuované Systémy Grafových Gramatík

3.1 Úvod

Kooperujúci distribuovaný systém grafových gramatík - model, ktorý budeme v tejto kapitole skúmať, je analógiou kooperujúceho distribuovaného systému gramatík ($CDGS$) definovaného pre klasické "stringové" gramatiky. Ide o systém gramatík nejakého typu, ktoré pristupujú k vetnej forme podľa istých pravidiel tak, že v jednom kroku s ňou pracuje iba jedna gramatika. Podobne ako v prípade $CDGS$ zadefinujeme tento model všeobecne pre $edNCE$ gramatiky, pričom budeme rozlišovať rôzne typy $CDGGS$ podľa typu komponentových gramatík. Pri skúmaní tohto modelu nás bude zaujímať, či preň budú platiť analogické tvrdenia o vzťahu medzi rôznymi triedami jazykov ako v prípade $CDGS$ (inšpiráciou pre niektoré z nich bola [4]). Ukazuje sa, že v prípade $CDGGS$ s komponentovými gramatikami, ktoré sú typu C-$edNCE$ je problémom resp. komplikáciou pri dôkaze tychto tvrdení (okrem faktu, že grafy majú samy o sebe zložitejšiu štruktúru ako reťazce) konfluentnosť, keďže táto vlastnosť nie je garantovaná len formou jednotlivých pravidiel ale aj istým pomerne komplikovaným vzťahom medzi všetkými pravidlami gramatiky.

3.2 Definície a označenia

Definícia 3.2.1. (CDGGS) *Kooperujúci distribuovaný systém grafových gramatík s n komponentami je $(n+3)$-tica*

$$\Theta = (\Delta, \Omega, G_1, G_2, \dots, G_n, S),\ kde$$

$$\begin{array}{rcl} G_i & = & (\Sigma_i, \Delta_i, \Gamma_i, \Omega_i, P_i) \text{ je } edNCE \text{ gramatika bez počiatočného} \\ & & \text{neterminálu pre } i \in \{1, \dots, n\} \\ \Delta & = & \Delta_1 \cup \Delta_2 \cup \ldots \cup \Delta_n \text{ je abeceda terminálnych symbolov vrcholov,} \\ \Omega & = & \Omega_1 \cup \Omega_2 \cup \ldots \cup \Omega_n \text{ je abeceda terminálnych symbolov hrán a} \\ S & \in & \bigcup_{1 \le i \le n} (\Sigma_i - \Delta_i) \text{ je počiatočný neterminál.} \end{array}$$

Definícia 3.2.2. (Krok odvodenia) *Nech* $\Theta = (\Delta, \Omega, G_1, G_2, \dots, G_n, S)$, *je CDGGS. Krokom odvodenia CDGGS* Θ *pri spôsobe prepisovania* $\le k$, *kde* $k \in N$ *je relácia*

$$\Longrightarrow_{\Theta}^{\le k} = \bigcup_{1 \le i \le n} \Longrightarrow_{G_i}^{\le k},\ kde \Longrightarrow_{G_i}^{\le k} = \bigcup_{1 \le j \le k} \Longrightarrow_{G_i}^{j}$$

Analogicky je definovaný krok odvodenia pri spôsobe prepisovania $= k$ *a* $\ge k$ *pre* $k \in N$.

Ďalej definujeme $\Longrightarrow_{\Theta}^{t} = \cup \Longrightarrow_{G_i}^{t}$, *kde* $K \Longrightarrow_{G_i}^{t} L$ *platí práve vtedy, keď* $K \Longrightarrow_{G_i}^{*} L$ *a neexistuje také* M, *že by platilo* $M \ne L \wedge L \Longrightarrow_{G_i} M$. *Spôsob prepisovania* t *nazývame terminálny spôsob prepisovania.*

Poznámka 3.2.3. *Terminálny spôsob prepisovania určuje, že gramatika, ktorá práve pracuje na vetnej forme, na nej musí pracovať dovtedy, kým obsahuje nejaké pravidlo, ktorým môže predĺžiť odvodenie.*

Spôsobo prepisovania $= k$ $(\ge k)$ *hovori, že každá gramatika, ktorá pracuje na vetnej forme, na nej musí pradovať práve (aspoň)* k *krokov pred tým, ako je (potenciálne) vystriedaná.*

Spôsobo prepisovania $\le k$ *v skutočnosti nie je nijak obmedzujúci, keďže nepožadujeme, aby sa gramatiky museli pri práci na vetnej forme striedať.*

Definícia 3.2.4. (Grafový jazyk) *Nech* $f \in \{t, *\} \cup \{= k, \le k, \ge k | k \in N \wedge k > 0\}$ *a nech* Θ *je* $CDGGS$. *Potom jazyk grafov (grafový jazyk) definovaný pri spôsobe prepisovania* f *je množina*

$$\begin{array}{rcl} L_f(\Theta) & = & \{[H] \ / \ H \in G_{\Delta\Omega},\ \exists z, r, i_1, ..., i_r \colon sn(S, z) \Longrightarrow_{G_{i_1}}^{f} H_1 \\ & & \Longrightarrow_{G_{i_2}}^{f} H_2 \Longrightarrow_{G_{i_3}}^{f} \dots \Longrightarrow_{G_{i_r}}^{f} H_r \equiv H\} \end{array}$$

Poznámka 3.2.5. *Neformálne si je možné tento model predstaviť ako systém n gramatík, ktoré sa striedajú pri (sekvenčnej) práci na jednej vetnej forme. Spôsob prepisovania vlastne predstavuje obmedzenie na to, kedy sa môžu gramatiky vystriedať.*

Označenie 3.2.6. *Nech X-edNCE je trieda grafových gramatík, potom triedu jazykov generových CDGGS gramatikami s n X-edNCE komponentovými gramatikami pri spôsobe prepisovania f označujeme $\mathcal{L}(X\text{-}CDGGS_n, f)$ (v prípade, že komponentové gramatiky sú (všeobecné) edNCE gramatiky, budeme kvôli jednoduchosti zápisu tento systém označovať ED-CDGGS).*

3.3 Generatívna sila CDGGS

Veta 3.3.1. *Nech $f \in \{*, 1, \geq 1, \leq 1, \leq 2, ...\}$, $X \in \{ED, B, LIN\}$, potom $\mathcal{L}(X\text{-}CDGGS_n, f) = X\text{-}edNCE$.*

Dôkaz. Toto tvrdenie hovorí, že generatívna sila X-$CDGGS$ je rovnaká ako X-$edNCE$, ak sa obmedzíme na spôsob prepisovania, ktorý "nenúti" jednotlivé gramatiky pracovať na vetnej forme dlhšie ako jeden krok odvodenia. Potrebujeme dokázať 2 inklúzie:

$\supseteq$: Nech $L \in X$-$edNCE$, $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$ je X-$edNCE$ gramatika taká, že $L(G) = L$, definujme X-$CDGGS$ systém gramatík Θ pre jazyk L nasledovne:

$$\begin{aligned} \Theta &= (\Delta, \Omega, G_1, S), \text{ kde} \\ G_1 &= (\Sigma, \Delta, \Gamma, \Omega, P) \end{aligned}$$

Potrebujeme dokázať, že platí $L_f(\Theta) = L(G)$, t.j. $\forall H([H] \in L_f(\Theta) \Leftrightarrow [H] \in L(G))$.

Nech $[H] \in L_f(\Theta)$, z definície jazyka $L_f(\Theta)$ platí

$$sn(S, z) \Longrightarrow^f_{G_1} H_1 \Longrightarrow^f_{G_1} H_2 \Longrightarrow^f_{G_1} ... \Longrightarrow^f_{G_1} H_r \equiv H,$$

a teda

$$sn(S, z) \Longrightarrow^*_{G_1} H$$

Keďže gramatiky G a G_1 majú rovnakú množinu pravidiel a sú nad rovnakými abecedami, tak platí $sn(S, z) \Longrightarrow^*_G H$. Z definície jazyka $L(G)$ ďalej platí $[H] \in L(G)$. Tým sme dokázali, že $[H] \in L_f(\Theta) \Rightarrow [H] \in L(G)$. Ostáva nám dokázať druhú implikáciu.

Nech $[H] \in L(G)$, z definície jazyka $L(G)$ platí $sn(S,z) \Longrightarrow_G^* H$. Keďže gramatiky G a G_1 majú rovnakú množinu pravidiel a sú nad rovnakými abecedami, tak $sn(S,z) \Longrightarrow_{G_1}^* H$ a pre ľubovoľné $f \in \{*, 1, \geq 1, \leq 1, \leq 2, ...\}$ platí, že

$$sn(S,z) \Longrightarrow_{G_1}^f H_1 \Longrightarrow_{G_1}^f H_2 \Longrightarrow_{G_1}^f ... \Longrightarrow_{G_1}^f H_r \equiv H$$

pre nejaké r. Tým sme dokázali, že $[H] \in L_f(\Theta)$, a teda aj $[H] \in L(G) \Rightarrow [H] \in L_f(\Theta)$.

$\subseteq$: Nech $L \in \mathcal{L}(X\text{-}CDGGS_n, f)$, $\Theta = (\Delta, \Omega, G_1, G_2, ..., G_n, S)$ (kde $G_i = (\Sigma_i, \Delta_i, \Gamma_i, \Omega_i, P_i)$) je X-$CDGGS$ systém gramatík taký, že $L_f(\Theta) = L$. X-$edNCE$ gramatika G pre jazyk L bude jednoducho simulovať všetky gramatiky G_i tým, že bude môcť použiť ľubovoľné z ich pravidiel. Gramatiku G teda definujeme nasledovne:

$$\begin{aligned} G &= (\Sigma, \Delta, \Gamma, \Omega, P, S), \text{ kde} \\ \Sigma &= \bigcup_{1 \leq i \leq n} \Sigma_i \\ \Gamma &= \bigcup_{1 \leq i \leq n} \Gamma_i \\ P &= \bigcup_{1 \leq i \leq n} P_i \end{aligned}$$

Potrebujeme dokázať, že $L_f(\Theta) = L(G)$, t.j. pre každú vetnú formu - graf H platí $[H] \in L_f(\Theta) \Leftrightarrow [H] \in L(G)$.

Nech $[H] \in L_f(\Theta)$, potom z definície jazyka $L_f(\Theta)$ platí

$$sn(S,z) \equiv H_0 \Longrightarrow_{G_{i_1}}^f H_1 \Longrightarrow_{G_{i_2}}^f H_2 \Longrightarrow_{G_{i_3}}^f ... \Longrightarrow_{G_{i_r}}^f H_r \equiv H$$

Keďže $H_j \in GRE_{\Sigma\Gamma}$ pre ľubovoľné j a gramatika G obsahuje pravidlá gramatiky G_{i_j}, tak $H_{j-1} \Longrightarrow_G^f H_j$, a teda $sn(S,z) \Longrightarrow_G^* H$. Tým sme dokázali, že $[H] \in L_f(\Theta) \Rightarrow [H] \in L(G)$. Ostáva nám dokázať opačnú inklúziu.

Nech $[H] \in L(G)$, potom

$$sn(S,z) \equiv H_0 \Longrightarrow_{G,p_1} H_1 \Longrightarrow_{G,p_2} H_2 \Longrightarrow_G ... \Longrightarrow_{G,p_r} H_r \equiv H$$

Z definície gramatiky vyplýva, že $(H_{j-1} \Longrightarrow_{G,p_j} H_j) \Rightarrow (H_{j-1} \Longrightarrow_{G_k}^f H_j)$, ak p_j je pravidlo gramatiky G_k (pričom taká gramatika existuje). Teda platí

$[H] \in L(G) \Rightarrow [H] \in L_f(\Theta)$, čím je dokázaná posledná inklúzia $L(G) \subseteq L_f(\Theta)$ a aj celá veta. $\square$

Poznámka 3.3.2. *Otvorenou otázkou ostáva prípad C-$CDGGS$. Zrejme ak by sme našli grafový jazyk mimo triedy C-$edNCE$, pre ktorý by existovala $edNCE$ gramatika, ktorej každé pravidlo by samostatne spĺňalo podmienku konfluentnosti, tak C-$edNCE \subsetneq \mathcal{L}(C$-$CDGGS_*, f)$.*

Poznámka 3.3.3. *Nasledujúce tri tvrdenia hovoria, že "donucovací" spôsob prepisovania nepridáva systému gramatík na generatívnej sile, ak uvažujeme len systémy gramatík s jednou komponentovou gramatikou.*

Veta 3.3.4. *$\mathcal{L}(X$-$CDGGS_1, t) = X$-$edNCE$ pre $X \in \{ED, C, B, LIN\}$.*

Dôkaz. Vzhľadom na to, že uvažujeme terminálny spôsob prepisovania a len jednu komponentovú gramatiku, tak gramatiky, ktoré potrebujeme zostrojiť na dôkaz oboch inklúzií, sú principiálne rovnaké so svojími vzormi. Dokážeme najprv inklúziu

$$\mathcal{L}(X\text{-}CDGGS_1, t) \subseteq X\text{-}edNCE$$

Nech $L \in \mathcal{L}(X$-$CDGGS_1, t)$, $\Theta = (\Delta, \Omega, G_1, S)$, kde $G_1 = (\Sigma, \Delta, \Gamma, \Omega, P)$, je gramatika pre jazyk L, t.j. $L_t(\Theta) = L$. Definujme X-$edNCE$ gramatiku G pre jazyk L nasledovne: $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$. Potrebujeme dokázať, že $L_t(\Theta) = L(G)$, t.j. $\forall H([H] \in L_t(\Theta) \Leftrightarrow [H] \in L(G))$.

Nech $[H] \in L_t(\Theta)$, z defínie L_t vyplýva, že

$$sn(S, z) \equiv H_0 \Longrightarrow_{G_1}^t H_1 \Longrightarrow_{G_1}^t H_2 \Longrightarrow_{G_1}^t \dots \Longrightarrow_{G_1}^t H_r \equiv H$$

Keďže uvažujeme spôsob prepisovania t, tak $r = 1$, a teda $sn(S, z) \Longrightarrow_{G_1}^t H$, z čoho zjavne vyplýva, že $sn(S, z) \Longrightarrow_{G_1}^* H$. Keďže gramatiky G a G_1 majú rovnakú množinu pravidiel a sú nad rovnakými abecedami, tak $sn(S, z) \Longrightarrow_G^* H$, a teda $[H] \in L(G)$.

Naopak, nech $[H] \in L(G)$, z definície jazyka platí, že

$$H \in GR_{\Delta\Omega} \text{ a } sn(S, z) \Longrightarrow_G^* H$$

Keďže $H \in GR_{\Delta\Omega}$, tak neexistuje graf K taký, že by platilo $H \Longrightarrow_G K$, a teda $sn(S, z) \Longrightarrow_{G_1}^t H$, z čoho vyplýva, že $[H] \in L_t(\Theta)$. Tým sme dokázali, že $\forall H([H] \in L(\Theta) \Leftrightarrow [H] \in L(G))$ a teda aj $L(\Theta) = L(G)$. Potrebujeme teraz dokázať opačnú inklúziu.

$$\mathcal{L}(X\text{-}CDGGS_1, t) \supseteq X\text{-}edNCE$$

Nech $L \in X\text{-}edNCE$, $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$ je $X\text{-}edNCE$ gramatika pre jazyk L, t.j. $L(G) = L$. Definujme $X\text{-}CDGGS_1$ systém gramatík Θ generujúci jazyk L pri terminálnom spôsobe prepisovania nasledovne:

$$\begin{aligned} \Theta &= (\Delta, \Omega, G_1, S), \text{ kde} \\ G_1 &= (\Sigma, \Delta, \Gamma, \Omega, P) \end{aligned}$$

Potrebujeme dokázať, že $L_t(\Theta) = L(G)$, t.j. $\forall H([H] \in L_t(\Theta) \Leftrightarrow [H] \in L(G))$.

Nech $[H] \in L_t(\Theta)$, z defínie jazyka $L_t(\Theta)$ vyplýva, že

$$sn(S, z) \equiv H_0 \Longrightarrow^t_{G_1} H_1 \Longrightarrow^t_{G_1} H_2 \Longrightarrow^t_{G_1} ... \Longrightarrow^t_{G_1} H_r \equiv H$$

Keďže uvažujeme spôsob prepisovania t, tak $r = 1$, a teda $sn(S, z) \Longrightarrow^t_{G_1} H$, z čoho vyplýva, že $sn(S, z) \Longrightarrow^*_{G_1} H$. Keďže gramatiky G a G_1 majú rovnakú množinu pravidiel a sú nad rovnakými abecedami, tak $sn(S, z) \Longrightarrow^*_G H$, a teda $[H] \in L(G)$.

Naopak, nech $[H] \in L(G)$, z definície jazyka platí, že

$$H \in GR_{\Delta\Omega} \text{ a } sn(S, z) \Longrightarrow^*_G H$$

Keďže $H \in GR_{\Delta\Omega}$, tak neexistuje graf K taký, že by platilo $H \Longrightarrow_G K$, a teda $sn(S, z) \Longrightarrow^t_{G_1} H$, z čoho vyplýva, že $[H] \in L_t(\Theta)$. Tým sme dokázali, že $\forall H([H] \in L_t(\Theta) \Leftrightarrow [H] \in L(G))$ a teda aj $L(\Theta) = L(G)$. □

Veta 3.3.5. *$\mathcal{L}(X\text{-}CDGGS_1, \geq k) = X\text{-}edNCE$ pre $X \in \{ED, C, B, LIN\}$ a $k \geq 2$.*

Dôkaz. Dokážeme najprv inklúziu $\mathcal{L}(X\text{-}CDGGS_1, \geq k) \supseteq X\text{-}edNCE$. Nech $L \in X\text{-}edNCE$ a $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$ je $X\text{-}edNCE$ gramatika pre jazyk L, t.j. $L(G) = L$. $X\text{-}CDGGS_1$ systém gramatík generujúci jazyk L pri spôsobe prepisovania $\geq k$ bude pracovať tak, že na $k - 1$ krokov odvodí graf s jedným vrcholom so symbolom S a potom bude pokračovať odvodením ako v gramatike G – to znamená, že každé odvodenie terminálneho grafu H bude dĺžky aspoň k, čo vyhovuje definícii spôsobu prepisovania $\geq k$.

Definujme teda X-$CDGGS_1$ systém gramatík Θ generujúci jazyk L pri spôsobe prepisovania $\geq k$ nasledovne:

$$\begin{aligned}
\Theta &= (\Delta, \Omega, G_1, S_0), \text{ kde} \\
G_1 &= (\Sigma_1, \Delta, \Gamma, \Omega, P_1) \\
\Sigma_1 &= \Sigma \cup \{S_i \mid i \in \{0, ..., k-2\}\} \\
P_1 &= P \cup \{S_i \rightarrow (sn(S_{i+1}), \emptyset) \mid i \in \{0, ..., k-2\}\}
\end{aligned}$$

Predpokladáme, že S_i sú všetko nové symboly vrcholov, t.j. že $\Sigma \cap \{S_i \mid i \in \{0, ..., k-2\}\} = \emptyset$ a $S_{k-1} \equiv S$.

Ľahko vidno, že pre ľubovoľné $X \in \{ED, C, B, LIN\}$ gramatika G_1 je X-$edNCE$ gramatika, vzhľadom na to, že nové pravidlá su lineárne (čo je dôležité v prípade, že $X \in \{B, LIN\}$) a majú prázdnu spájaciu reláciu C (z čoho vyplýva prípad, keď $X \in \{B, C\}$).

Nové symboly a pravidlá tejto gramatiky využijeme na to, aby sme vygenerovali dostatočný počet vetných foriem pred tým, ako vygenerujeme vrchol s počiatočným symbolom pôvodnej gramatiky.

Potrebujeme dokázať, že platí $L_{\geq k}(\Theta) = L(G)$, t.j. $\forall H([H] \in L_{\geq k}(\Theta) \Leftrightarrow [H] \in L(G))$.

Nech $[H] \in L(G)$, z definície jazyka platí, že

$$H \in GR_{\Delta\Omega} \text{ a } sn(S, z) \Longrightarrow_G^* H$$

Keďže $\Sigma \subseteq \Sigma_1$ a $P \subseteq P_1$ tak odvodenie grafu H z grafu $sn(S, z)$ existuje aj v gramatike G_1, t.j. $sn(S, z) \Longrightarrow_{G_1}^* H$. Ľahko vidno, že v gramatike G_1 vieme z počiatočného grafu $sn(S_0, z)$ (na $k-1$ krokov) odvodiť graf s jediným vrcholom so symbolom S: $sn(S_0, z) \Longrightarrow_{G_1}^{k-1} sn(S, z_1)$. Spojením týchto dvoch faktov dostávame, že

$$sn(S_0, z) \Longrightarrow_{G_1}^{k-1} sn(S, z_1) \Longrightarrow_{G_1}^* H$$

Keďže odvodenie grafu H z grafu $sn(S, z_1)$ je dĺžky aspoň 1, tak sme dokázali, že $sn(S_0, z) \Longrightarrow_{G_1}^{\geq k} H$, a teda $[H] \in L(\Theta)$. Na dokázanie prvej inklúzie potrebujeme ešte dokázať opačnú implikáciu.

Nech $[H] \in L_{\geq k}(\Theta)$, z definície jazyka platí, že

$$H \in GR_{\Delta\Omega} \text{ a } sn(S, z) \Longrightarrow_{G_1}^{\geq k} H_1 \Longrightarrow_{G_1}^{\geq k} H_2 \Longrightarrow_{G_1}^{\geq k} ... \Longrightarrow_{G_1}^{\geq k} H_r \equiv H,$$

a teda aj $sn(S_0, z) \Longrightarrow_{G_1}^{\geq k} H$. Ľahko vidno, že pre (každé) odvodenie ľubovoľného (terminálneho) grafu H platí:

$$sn(S_0, z) \Longrightarrow_{G_1}^{k-1} sn(S, z_1) \Longrightarrow_{G_1}^* H$$

Keďže platí, že ľubovoľná vetná forma, ktorú možno odvodiť z grafu $sn(S, z_1)$ je graf v $H \in GRE_{\Sigma\Gamma}$ a pri jeho odvodzovaní boli použité len pravidlá z množiny P, tak dostávame $sn(S, z) \Longrightarrow_G^* H$ pre nejaké z, a teda aj $[H] \in L(G)$.

Ostáva nám dokázať opačnú inklúziu

$$\mathcal{L}(X\text{-}CDGGS_1, \geq k) \subseteq X\text{-}edNCE$$

Nech $L \in \mathcal{L}(X\text{-}CDGGS_1, \geq k)$, $\Theta = (\Delta, \Omega, G_1, S)$, $G_1 = (\Sigma, \Delta, \Gamma, \Omega, P)$ je $X\text{-}CDGGS_1$ systém gramatík pre jazyk L, t.j. $L_{\geq k}(\Theta) = L$. Gramatika G pre jazyk L bude pracovať tak, že z počiatočného grafu odvodí na jeden krok ľubovoľný graf, ktorý vie gramatika G_1 odvodiť (z počiatočného grafu) na k krokov. Takto dokáže gramatika G vygenerovať len tie (a práve) tie grafy, ktoré dokáže vygenerovať gramatika G_1 aspoň k krokmi. Definujme teda gramatiku G takto:

$$\begin{aligned} G &= (\Sigma', \Delta, \Gamma, \Omega, P', S'), \text{ kde} \\ \Sigma' &= \Sigma \cup \{S'\} \text{ (predpokladáme, že } S' \text{ je nový symbol)} \\ P' &= P \cup \{S' \to (D, \emptyset) \mid D \in [S]_{G_1}^k\} \\ [S]_{G_1}^k &= \{D \mid sn(S, z) \Rightarrow_{G_1}^k D\} \end{aligned}$$

Z koštrukcie gramatiky G vyplýva, že ak gramatika G_1 je $X\text{-}edNCE$, tak gramatika G je rovnakého typu:

- prípad $X = ED$ vyplýva z formátu pravidiel
- prípad $X = C$ vyplýva z konfluentnosti gramatiky G a z faktu, že spojovacia relácia C nových pravidiel je prázdna
- prípad $X = B$ vyplýva z toho, že ľubovoľná vetná forma $B\text{-}edNCE$ spĺňa prvú podmienku z definície $B\text{-}edNCE$ a spojovacia relácia C nových pravidiel je prázdna
- pre $X = LIN$ máme linearitu gramatiky zabezpečenú tým, že vetné formy lineárnej gramatiky obsahujú (najviac) jeden neterminálny vrchol, a teda nové pravidlá sú lineárne.

Ľahko vidno, že pri prvom kroku odvodenia v gramatike G musí byť použité nové pravidlo $S' \to (D, C)$ pre nejaké pravidlo $(D, C) \in [S]_{G_1}^k$ a takisto, že nové pravidlá môžu byť použité len pri prvom kroku odvodenia (čo vyplýva z toho, že nový neterminálny symbol vrchola S' sa nevyskytuje ako symbol vrchola žiadneho grafu na pravej strane ľubovoľného pravidla a teda sa vrchol s týmto symbolom nemôže vyskytovať v žiadnej vetnej forme vygenerovanej z $sn(S, z)$).

Potrebujeme dokázať, že $L_{\geq k}(\Theta) = L(G)$. Nech $[H] \in L_{\geq k}(\Theta)$, z definície jazyka platí, že

$$H \in GR_{\Delta\Omega} \text{ a } sn(S, z) \Longrightarrow_{G_1}^{\geq k} H_1 \Longrightarrow_{G_1}^{\geq k} H_2 \Longrightarrow_{G_1}^{\geq k} \ldots \Longrightarrow_{G_1}^{\geq k} H_r \equiv H$$

Toto odvodenie môžeme zapísať aj v takomto tvare

$$sn(S, z) \Longrightarrow_{G_1}^{k} K \Longrightarrow_{G_1}^{*} H$$

Podľa definície takéto odvodenie existuje práve vtedy a len vtedy, keď existuje v gramatike G takéto odvodenie grafu H:

$$sn(S', z) \Longrightarrow_{G} K \Longrightarrow_{G}^{*} H$$

čo je podľa horeuvedených argumentov ekvivalentné s tým, že $[H] \in L(G)$. Týmto sme dokázali, že $L_{\geq k}(\Theta) = L(G)$. $\square$

Veta 3.3.6. *$\mathcal{L}(X\text{-}CDGGS_1, = k) = X\text{-}edNCE$ pre $X \in \{ED, C, B, LIN\}$ a $k \geq 2$.*

Dôkaz. Z trojice podobných nám ostavá dokázať posledné (a zároveň najzaujímavejšie). Začneme znovu s dokazom jednoduchšej inklúzie

$$\mathcal{L}(X\text{-}CDGGS_1, = k) \supseteq X\text{-}edNCE.$$

Nech $L \in X\text{-}edNCE$, $G = (\Sigma, \Delta, \Gamma, \Omega, P, S)$ je $X\text{-}edNCE$ gramatika pre jazyk L, t.j. $L(G) = L$. $X\text{-}CDGGS_1$ systém gramatík generujúci jazyk L pri spôsobe prepisovania $= k$ bude pracovať tak, že každý krok odvodenia gramatiky G bude simulovať k krokmi, čo je potrebné v prípade spôsobu prepisovania $= k$.

Definujme teda $X\text{-}CDGGS_1$ systém gramatík Θ generujúci jazyk L pri spôsobe prepisovania $= k$ takto:

$$\begin{aligned}
\Theta &= (\Delta, \Omega, G_1, S), \text{ kde} \\
G_1 &= (\Sigma_1, \Delta, \Gamma, \Omega, P_1) \\
\Sigma_1 &= \Sigma \cup \{Y_i \mid Y \to (D, C) \in P, i \in \{1, \dots, k-1\}\} \\
P_1 &= \{Y_i \to (sn(Y_{i+1}, y_{i+1}), C_{full}(y_{i+1})) \mid Y \to (D, C) \in P, \\
&\quad i \in \{0, \dots, k-2\}\} \\
&\quad \cup \\
&\quad \{Y_{k-1} \to (D, C') | Y \to (D, C) \in P\} \\
C_{full}(y) &= \{(Y, \alpha/\alpha, y, io) \mid Y \in \Sigma_1, \alpha \in \Gamma, io \in \{in, out\}\}, \\
&\quad \text{ak } X \in \{C, ED\} \\
C_{full}(y) &= \{(Y, \alpha/\alpha, y, io) \mid Y \in \Delta, \alpha \in \Gamma, io \in \{in, out\}\}, \\
&\quad \text{ak } X \in \{B, LIN\} \\
C' &= \{(Y_i, \alpha/\beta, y, io) \mid (Y, \alpha/\beta, y, io) \in C,\ Y \in \mathrm{N},\ \mathrm{i} \in \{0, ..., \mathrm{k}-1\}\} \\
&\quad \cup \\
&\quad \{(a, \alpha/\beta, y, io) \mid (a, \alpha/\beta, y, io) \in C,\ a \in \Delta,\ i \in \{0, ..., k-1\}\}, \\
&\quad \text{ak } X \in \{C, ED\} \\
C' &= C, \text{ ak } X \in \{B, LIN\}
\end{aligned}$$

Predpokladáme, že $Y_0 \equiv Y$ (kvôli jednoduchosti zápisu) a že Y_i sú všetko (navzájom rôzne) nové symboly vrcholov, t.j.

$$\Sigma \cap \{Y_i \mid Y \to (D, C),\ i \in \{1, \dots, k-1\}\} = \emptyset$$

Najprv potrebujeme dokázať, že skonštruovaná gramatika je X-$edNCE$:

- prípad $X = ED$ vyplýva z formátu pravidiel
- $X = LIN$ je tiež triviálne, keďže grafy na pravej strane pravidiel buď obsahujú jeden vrchol, alebo sú lineárne v dôsledku linearity vzorovej gramatiky
- pre $X = B$ je prvá podmienka z definície B-$edNCE$ splnená podľa podobného argumentu ako v predchádzajúcom prípade, druhá vyplýva z konštrukcie gramatiky
- posledná možnosť je $X = C$, konfluentnosť gramatiky dokážeme explicitne. Uvažujme teda pravidlá $Z_1 \to (D_1, C_1')$, $Z_2 \to (D_2, C_2')$ v P_1,

ľubovoľný uzol $x_1 \in V_{D_1}$ a $x_2 \in V_{D_2}$ a ľubovoľnú hranu $\alpha, \delta \in \Gamma$. Môžu nastať tieto 3 možnosti:

1. $Z_1 = X_{k-1}$ a $Z_2 = Y_{k-1}$ pre nejaké X a Y. Z definície množiny P_1 vyplýva, že $X \to (D_1, C_1) \in P$ a $Y \to (D_2, C_2) \in P$. Z konfluentnosti gramatiky G vyplýva, že platí nasledujúca ekvivalencia

$$\begin{array}{lcl} \exists \beta \in \Gamma & : & (Y, \alpha/\beta, x_1, out) \in C_1 \text{ a } (\lambda_{D_1}(x_1), \beta/\delta, x_2, in) \in C_2 \\ \Longleftrightarrow & & \\ \exists \gamma \in \Gamma & : & (X, \alpha/\gamma, x_1, in) \in C_2 \text{ a } (\lambda(x_2), \gamma/\delta, x_1, out) \in C_1 \end{array}$$

   Zjavne

$$\forall \beta \in \Gamma : (Y, \alpha/\beta, x_1, out) \in C_1 \Longleftrightarrow (Y_i, \alpha/\beta, x_1, out) \in C_1'$$

   pre ľubovoľné $i \in \{1, \ldots, k-1\}$, podobne platí analogická ekvivalencia pre ostatné 3 inklúzie. Ľahko potom vidno, že

$$\begin{array}{lcl} \exists \beta \in \Gamma & : & (Y_{k-1}, \alpha/\beta, x_1, out) \in C_1' \text{ a } (\lambda_{D_1}(x_1), \beta/\delta, x_2, in) \in C_2' \\ \Longleftrightarrow & & \\ \exists \gamma \in \Gamma & : & (X_{k-1}, \alpha/\gamma, x_2, in) \in C_2' \text{ a } (\lambda(x_2), \gamma/\delta, x_1, out) \in C_1' \end{array}$$

   čo bolo treba dokázať.

2. $Z_1 = X_{k-1}$ a $Z_2 = Y_i$ pre nejaké $X, Y \in \Sigma$ a $i < k-1$. Potom platí, že $D_2 = sn(Y_{i+1}, y_{i+1})$, $C_2' = C_{full}(y_{i+1})$ a $x_2 = y_{i+1}$. Potrebujeme teda dokázať, že

$$\begin{array}{lcl} \exists \beta \in \Gamma & : & (Y_i, \alpha/\beta, x_1, out) \in C_1' \text{ a} \\ & & (\lambda_{D_1}(x_1), \beta/\delta, y_{i+1}, in) \in C_{full}(y_{i+1}) \\ \Longleftrightarrow & & \\ \exists \gamma \in \Gamma & : & (X_{k-1}, \alpha/\gamma, y_{i+1}, in) \in C_{full}(y_{i+1}) \text{ a} \\ & & (Y_{i+1}, \gamma/\delta, x_1, out) \in C_1' \end{array}$$

   Obe inklúzie vyplývajú z argumentu uvedeného v prvej časti, keď položíme $\beta = \gamma = \delta$.

3. $Z_1 = X_i$ a $Z_2 = Y_j$ pre nejaké $X, Y \in \Sigma$ a $i, j < k-1$. Konfluentnosť v tomto prípade triviálne platí.

Potrebujeme dokázať, že platí $L_{=k}(\Theta) = L(G)$, t.j. $\forall H([H] \in L_{=k}(\Theta) \Leftrightarrow [H] \in L(G))$.

Nech $[H] \in L(G)$, z definície jazyka platí, že $H \in GR_{\Delta\Omega}$ a $sn(S,z) \Longrightarrow^*_G H$ t.j.

$$sn(S,z) \equiv H_0 \Longrightarrow_G H_1 \Longrightarrow_G \dots \Longrightarrow_G H_n \equiv H$$

Ľahko vidno, že jeden krok odvodenia v gramatike G môžme nahradiť k krokmi v gramatike G_1, teda

$$H_i \Longrightarrow_G H_{i+1} \Longleftrightarrow H_i \Longrightarrow^{=k}_{G_1} H_{i+1}$$

Tým sme dokázali, že $[H] \in L_{=k}(\Theta)$. Potrebujeme ešte dokázať opačnú inklúziu.

Nech $[H] \in L_{=k}(\Theta)$, z definície jazyka platí, že $H \in GR_{\Delta\Omega}$ a

$$sn(S,z) \Longrightarrow^{=k}_{G_1} H_1 \Longrightarrow^{=k}_{G_1} H_2 \Longrightarrow^{=k}_{G_1} \dots \Longrightarrow^{=k}_{G_1} H_r \equiv H$$

Na to, aby sme ukázali, že graf H možno odvodiť aj v gramatike G nám stačí ukázať nasledovné tvrdenie:

Ak platí

$$sn(S,z) \Longrightarrow^n_{G_1} H_1(\in GRE_{\Sigma_1\Gamma}), \text{ tak } sn(S,z) \Longrightarrow^*_G H(\in GRE_{\Sigma\Gamma})$$

pričom existuje bijektívna funkcia

$$f : V_{H_1} \to V_H \text{ taká, že } E_H = \{(f(v), \gamma, f(w)) | (v, \gamma, w) \in E_{H_1}\}$$

a pre všetky $v \in V_{H_1}$ platí

$$\begin{aligned} \lambda_H(f(v)) &= \lambda_{H_1}(v), \text{ ak } \lambda_{H_1}(v) \in \Sigma \\ \lambda_H(f(v)) &= X, \text{ ak } \lambda_{H_1}(v) = X_i \in \Sigma_1 - \Sigma \end{aligned}$$

(t.j. grafy H_1 a H sú izomorfné až na symboly ich neterminálnych vrcholov, pre ktoré platí uvedená podmienka).

Toto tvrdenie dokážeme indukciou na n:

1. ak $n = 0$, tak $H_1 \equiv sn(S,z)$ a tvrdenie triviálne platí

2. predpokladáme, že tvrdenie platí pre každé $n_1 < n$, musíme ukázať, že platí aj pre n. Uvažujme odvodenie dĺžky $n > 0$:

$$sn(S,z) \Longrightarrow_{G_1}^{n-1} H_1' \Longrightarrow_{G_1} H_1$$

z IP pre $n_1 = n-1$ dostaneme, že $sn(S,z) \Longrightarrow_G^* H'$ a existuje bijektívna funkcia

$$f : V_{H_1'} \to V_{H'} \text{ taká, že } E_{H'} = \{(f(v),\gamma,f(w)) \mid (v,\gamma,w) \in E_{H_1'}\}$$

a pre všetky $v \in V_{H_1'}$ platí

$$\begin{aligned} \lambda_{H'}(f(v)) &= \lambda_{H_1'}(v), \text{ ak } \lambda_{H_1'}(v) \in \Sigma \\ \lambda_{H'}(f(v)) &= X, \text{ ak } \lambda_{H_1'}(v) = X_i \in \Sigma_1 - \Sigma \end{aligned}$$

Nech pri poslednom (n-tom) kroku odvodenia bolo použité pravidlo p aplikované na vrchol v so symbolom X_i, mohli nastať tieto dve možnosti:

(a) $i < k-1$, v tomto prípade je pravidlo p tvaru

$$X_i \to (sn(X_{i+1}, x_{i+1}), C_{full}(x_{i+1}))$$

Ľahko vidno, že graf H_1 sa od grafu H_1' líši len symbolom vrchola, a teda dokazované tvrdenie platí aj pre graf H_1.

(b) $i = k-1$, v tomto prípade je pravidlo p tvaru $X_{k-1} \to (D, C')$ a $X \to (D,C) \in P$. Pre vrchol $f(v)$ grafu H' platí $\lambda_{H'}(f(v)) = X$, teda môžme naňho aplikovať pravidlo $X \to (D,C)$ a dostaneme graf H izomorfný s H_1 až na symboly ich vrcholov, pre ktoré platí uvedená podmienka, teda H dokazuje uvedené tvrdenie.

Týmto sme dokázali, že pre graf H zo začiatku dôkazu tejto inklúzie existuje aj odvodenie v gramatike G t.j. $[H] \in L(G)$, čo bolo treba dokázať.

Na uzavretie dôkazu vety potrebujeme ešte dokázať opačnú inklúziu

$$\mathcal{L}(X\text{-}CDGGS_1, = k) \subseteq X\text{-}edNCE$$

Nech $L \in \mathcal{L}(X\text{-}CDGGS_1, = k)$, $\Theta = (\Delta, \Omega, G_1, S)$, $G = (\Sigma_1, \Delta, \Gamma, \Omega, P_1)$ je X-$CDGGS$ systém gramatík pre jazyk L t.j. $L_{=k}(\Theta) = L$. Potrebujeme zostrojiť gramatiku G, ktorá generuje práve tie (terminálne) grafy, ktoré gramatika G_1 generuje na mk krokov pre nejaké m. Pre tento účel si budeme vo vetnej forme pamätať akési (distribuované) počítadlo, v ktorom budeme udržovať informáciu o tom, koľko krokov odvodenia (modulo k) ešte môžeme použiť. Neterminálny symbol priradený nejakému vrcholu bude okrem neterminálneho symbolu pôvodnej gramatiky obsahovať informáciu o počte krokov odvodenia, ktoré možno použiť pri odvodzovaní terminálneho grafu z daného neterminálneho vrcholu.

Definujme X-$edNCE$ gramatiku pre jazyk L nasledovne:

$$
\begin{array}{rcl}
G & = & (\Sigma, \Delta, \Gamma, \Omega, P, [S, 0]), \text{ kde} \\
\Sigma & = & \Delta \cup \{[Y, i] \mid Y \in \mathrm{N}_1, \mathrm{i} \in \{0 \ldots \mathrm{k}-1\}\} \\
P & = & \{[Y, i] \to (D', C') \mid Y \to (D, C) \in P_1, i \in \{0 \ldots k-1\}, \\
 & & (D', C') \in \kappa(D, C, i \ominus_k 1), \\
 & & D \text{ obsahuje neterminálny vrchol}\} \\
 & & \cup \\
 & & \{[Y, 1] \to (D, C') \mid Y \to (D, C) \in P_1, \\
 & & D \text{ neobsahuje neterminálny vrchol}\} \\
\kappa(D, C, i) & = & \{(D', C') \mid D' = (V_D, E_D, \lambda_{D'}), \lambda_{D'} \in \Lambda(D, i)\} \\
\Lambda(D, i) & = & \{\lambda \mid \lambda\colon V_D \to \Sigma, \\
 & & \lambda(v) = \lambda_D(v) \text{ ak } \lambda_D(v) \in \Delta, \\
 & & \lambda(v) = [\lambda_D(v), j] \text{ ak } \lambda_D(v) \in \mathrm{N}_1, \\
 & & j \in \{0 \ldots k-1\}, \\
 & & \displaystyle\sum_{v, \lambda_D(v) \in \mathrm{N}} \mathit{counter}(\lambda(v)) \mod k = i\} \\
\mathit{counter}([Y, i]) & = & i \text{ pre } [Y, i] \in \mathrm{N} \\
C' & = & \{([Y, i], \alpha/\beta, v, io) \mid (Y, \alpha/\beta, v, io) \in C,\ Y \in \Sigma_1 - \Delta, \\
 & & i \in \{0, \ldots k-1\}\} \\
 & & \cup \\
 & & \{(Y, \alpha/\beta, v, io) \mid (Y, \alpha/\beta, v, io) \in C,\ Y \in \Delta\}
\end{array}
$$

Neterminánymi symbolmi vrcholov tejto gramatiky sú dvojice $[Y, i]$, kde Y je neterminálny symbol vrchola pôvodnej gramatiky a i je prirodzené číslo od 0 po $k-1$, ktorého význam je ten, že odvodenie z vrchola označeného týmto neterminálom musí mať dĺžku $mk + i$. Množinu neterminálnych symbolov vrcholov pôvodnej gramatiky budeme označovať N_1, množinu neterminálnych symbolov vrcholov zostrojenej gramatiky budeme označovať N. Každému pravidlu pôvodnej gramatiky zodpovedá jeden z dvoch typov pravidiel v závislosti na tom, či graf na pravej strane pôvodného pravidla obsahuje neterminálny vrchol. V prípade, že nejaký taký vrchol existuje, definujeme pre dané pravidlo v novej gramatike skupinu pravidiel spĺňajúce podmienku, že súčet počítadiel neterminálnych vrcholov grafu na pravej strane nového pravidla je rovný $(i \mod k) \ominus_k -1$, kde i je počítadlo vrcholu na ľavej strane nového pravidla.

Pre $X \in \{ED, B, LIN\}$, z konštrukcie gramatiky G_1 triviálne vyplýva, že je X-$edNCE$, nech teda $X = C$, dokážeme konfluentnosť gramatiky. Uvažujme pravidlá $p'_1 : [Y_1, i] \to (D'_1, C'_1)$, $p'_2 : [Y_2, j] \to (D'_2, C'_2)$ v P, ľubovoľný uzol $y_1 \in V_{D'_1}$ a $y_2 \in V_{D'_2}$ a ľubovoľné hranové symboly $\alpha, \delta \in \Gamma$, potrebujeme dokázať, že platí

$$\begin{aligned} &\exists\beta \in \Gamma \quad : \quad ([Y_2, j], \alpha/\beta, x_1, out) \in C'_1 \text{ a } (\lambda_{D'_1}(x_1), \beta/\delta, y_2, in) \in C'_2 \\ &\quad\Longleftrightarrow \\ &\exists\gamma \in \Gamma \quad : \quad ([Y_1, i], \alpha/\gamma, y_2, in) \in C'_2 \text{ a } (\lambda_{D'_2}(y_2), \gamma/\delta, y_1, out) \in C'_1 \end{aligned}$$

Označme vzory pravidiel p'_1 a p'_2 v gramatike G_1 ako $p_1 : Y_1 \to (D_1, C_1)$ resp. $p_2 : Y_2 \to (D_2, C_2)$. Z predpokladu konfluentnosti gramatiky G_1 pre vrcholy y_1, y_2 a hranové symboly $\alpha, \delta \in \Gamma$ platí

$$\begin{aligned} &\exists\beta \in \Gamma \quad : \quad (Y_2, \alpha/\beta, y_1, out) \in C_1 \text{ a } (\lambda_{D_1}(y_1), \beta/\delta, y_2, in) \in C_2 \\ &\quad\Longleftrightarrow \\ &\exists\gamma \in \Gamma \quad : \quad (Y_1, \alpha/\gamma, y_2, in) \in C_2 \text{ a } (\lambda_{D_2}(y_2), \gamma/\delta, y_1, out) \in C_1 \end{aligned}$$

Keďže pre ľubovoľný neterminálny symbol vrchola Y platí

$$(Y, \alpha/\beta, y_1, out) \in C \Leftrightarrow ([Y, i], \alpha/\beta, y_1, out) \in C$$

tak je konfluentnosť gramatiky G dokázaná.

Musíme ukázať, že $L_{=k}(\Theta) = L(G)$, t.j. $\forall H([H] \in L_{=k}(\Theta) \Leftrightarrow [H] \in L(G))$.

Nech $[H] \in L_{=k}(\Theta)$, z definície jazyka platí, že

$$sn(S, z) \equiv H_0 \Longrightarrow^{=k}_{G_1} H_1 \Longrightarrow^{=k}_{G_1} H_2 \Longrightarrow^{=k}_{G_1} \ldots \Longrightarrow^{=k}_{G_1} H_m \equiv H$$

a

$$H_i \equiv H_{i_0} \Longrightarrow_{v_{i_1},p_{i_1}} H_{i_1} \Longrightarrow_{v_{i_2},p_{i_2}} H_{i_2} \Longrightarrow_{v_{i_3},p_{i_3}} \ldots \Longrightarrow_{v_{i_k},p_{i_k}} H_{i_k} \equiv H_{i+1}$$

Uvažujme najkratšie odvodenie grafu H. Potrebujeme nájsť odvodenie grafu H v gramatike G - toto odvodenie bude "skoro rovnaké" ako odvodenie v pôvodnej gramatike (ak odhliadneme od počítadla v každom neterminálnom vrchole).

Definujme pomocnú funkciu *removeCounter* pre graf D takto:

$$\begin{aligned} removeCounter(D) &= (V_D, E_D, \lambda), \text{ kde} \\ \lambda(v) &= X, \text{ ak } \lambda_D(v) = [X, i] \in \mathrm{N} \\ \lambda(v) &= X, \text{ ak } \lambda_D(v) = X \in \Delta \end{aligned}$$

Pre neterminál $\alpha = [X, i]$ z abecedy N označujeme jeho druhú zložku i ako $counter(\alpha)$

Definujme funkcie *index*, ktorá označuje poradie grafu v odvodení grafu H, t.j.

$$index(H_{i_j}) = ik + j \text{ (zjavne } index(H_{i_j}) \in \{0, \ldots, mk\})$$

a funkciu *dLength*, ktorá pre vrchol v_i vracia počet krokov potrebných na odvodenie terminálneho grafu z tohoto vrchola t.j.

$$\begin{aligned} dLength(v_i) &= 1, \text{ ak } p_i \text{ je terminálne pravidlo} \\ dLength(v_i) &= 1 + \sum_{v, v \in V_{D_i}, \lambda_{D_i}(v) \in \mathrm{N}_1, \mathrm{D_i = rhs(p_i)_{[1]}}} dLength(v), \text{ inak} \end{aligned}$$

Medzi funkciami *index* a *dLength* platí nasledujúci vzťah:

$$mk - index(D) = \sum_{v, v \in V_D, \lambda_D(v) \in \mathrm{N}_1} dLength(v)$$

Toto tvrdenie dokážeme indukciou na $mk - index(D)$.

1. $mk - index(D) = 0$: v tomto prípade je $D \equiv H \equiv H_m \equiv H_{m-1_k}$. Pravá strana nerovnosti je potom rovná 0:

$$\sum_{v, v \in V_D, \lambda_D(v) \in \mathrm{N}_1} dLength(v) = 0$$

keďže D neobsahuje neterminálne vrcholy.

2. predpokladáme, že tvrdenie platí pre každý graf H_{i_j}, pre ktorý $mk - index(H_{i_j}) < n$, potrebujeme dokázať, že platí aj pre graf $H_{i_j} \equiv D$ s $mk - index(D) = n$. Prepokladajme, že n je aspoň 1 (keďže prípad $n = 0$ sme dokázali v báze indukcie), potom graf D nie je posledným grafom v odvodení. Uvažujme graf E (priamo) odvodený z grafu D, z definície funkcie *index* platí, že $index(D) + 1 = index(E)$, a teda graf E spĺňa indukčný predpoklad:

$$n - 1 = mk - index(E) = \sum_{v, v \in V_E, \lambda_E(v) \in \mathrm{N}_1} dLength(v)$$

Graf E vznikol z grafu D nahradením nejakého vrchola (označme ho v) grafom z pravej strany daného pravidla (označme ho I). Uvažujme, ako sa líši suma pre grafy D a E: v sume pre D zarátavame $dLength(v)$, v sume pre E zarátavame namiesto toho $dLength(w)$ pre neterminálne vrcholy grafu I, čiže medzi sumami pre D a E platí nasledujúci vzťah

$$dLength(v) + \sum_{u, u \in V_E, \lambda_E(u) \in \mathrm{N}_1} dLength(u)$$
$$=$$
$$\sum_{u, u \in V_I, \lambda_I(u) \in \mathrm{N}_1} dLength(u) + \sum_{u, u \in V_D, \lambda_D(u) \in \mathrm{N}_1} dLength(u)$$

Z definície funkcie *dLength* zase platí, že

$$dLength(v) = 1 + \sum_{u, u \in V_I, \lambda_I(u) \in \mathrm{N}_1} dLength(u)$$

Spojením posledných troch vzťahov dostávame, že

$$\sum_{v, v \in V_D, \lambda_D(v) \in \mathrm{N}_1} dLength(v) = n$$

čo bolo treba dokázať.

Na dôkaz existencie odvodenia grafu H nám stačí ukázať, že pre ľubovoľný graf H_{i_j} platí, že v gramatike G existuje odvodenie grafu D s vlastnosťou

$$removeCounter(D) = H_{i_j}$$

a pre ľubovoľný neterminálny vrchol v grafov H_{i_j} resp. D platí

$$counter(\lambda_D(v)) = dLength(v) \mod k$$

Tvrdenie dokážeme indukciou na $index(H_{i_j})$

1. $index(H_{i_j}) = 0$: $H_{i_j} = H_0 \equiv sn(S, z)$ je graf s jediným vrcholom so symbolom S, ľahko vidieť, že hľadaný graf D je rovný $sn([S, 0], z)$, keďže
$$removeCounter(sn([S, 0], z)) = sn(S, z)$$
Priamočiarým overením dostaneme, že
$$counter([S, 0]) = 0 = mk \mod k$$
Keďže $index(sn(S, z)) = 0$, tak podľa vyššie dokázaného tvrdenia je
$$mk = \sum_{v, v \in V_{sn(S,z)}, \lambda_{sn(S,z)}(v) \in N_1} dLength(v) = dLength(z)$$
čím je báza indukcie dokázaná.

2. predpokladáme, že tvrdenie platí pre každý graf s indexom menším ako n (IP), potrebujeme dokázať, že platí aj pre graf H_{i_j} s indexom rovným n. Ďalej môžeme predpokladať, že n je aspoň 1, a teda existuje predchodca grafu H_{i_j} v uvažovanom odvodení (označme ho H_{k_l}). Keďže $index(H_{k_l}) = index(H_{i_j}) - 1$, tak graf H_{k_l} spĺňa indukčný predpoklad, čiže v gramatike G existuje odvodenie grafu E s vlastnosťou
$$removeCounter(E) = H_{k_l}$$
a pre ľubovoľný vrchol v grafov H_{k_l} resp. E platí
$$counter(\lambda_D(v)) = dLength(v) \mod k$$
Graf H_{i_j} vznikol z grafu H_{k_l} aplikovaním pravidla $p : X \to (D, C)$ na vrchol v grafu H_{k_l}. Vrchol v má v grafe E symbol $[X, dLength(v)]$, aplikujeme naňho také pravidlo $p' : [X, dLength(v) \mod k] \to (D', C')$, pre ktoré platí, že $D = removeCounter(D')$ a symbol každého neterminálneho vrchola
$$\lambda_{D'}(w) = [\lambda_D(w), dLength(w) \mod k]$$
Z definície množiny pravidiel P a funkcie *dLength* vidieť, že také pravidlo existuje. Výsledok aplikácie tohto pravidla označme F. Zrejme graf F spĺňa požadované podmienky, čím je toto tvrdenie dokázané.

Priamočiarým dôsledkom tohto tvrdenia je fakt, že v gramatike G existuje odvodenie terminálneho grafu H, keďže podmienka $H = removeCounter(H')$ platí pre graf H' len ak $H = H'$. Teda $[H] \in L(G)$, $L(\Theta) \subseteq L(G)$.

Potrebujeme ešte dokázať opačnú inklúziu. Zrejme pre graf $H \in L(G)$ stačí ukázať, že jeho (ľubovoľné) odvodenie v gramatike G je dĺžky mk pre vhodné m (v gramatike G_l bude potom existovať jeho odvodenie, kde namiesto previdla p použijeme jeho vzor, ktorý je jednoznačne určený). Uvažujme ľubovoľné odvodenie dĺžky k v gramatike G

$$H_k \Rightarrow H_{k-1} \Rightarrow \dots H_1 \Rightarrow H_0$$

kde

$$\sum_{v; v \in V_{H_k}, \lambda_{H_k}(v) \in \mathrm{N}} counter(v) \mod k = 0$$

Z defininície množiny pravidiel P vyplýva, že

$$\sum_{v; v \in V_{H_j}, \lambda_{H_j}(v) \in \mathrm{N}} counter(v) \mod k = j$$

Z tohoto faktu potom priamo vyplýva, že horeuvedená suma je rovná 0 len pre každú k-tu vetnú formu v odvodení, čiže pre terminálny graf H platí, že ľubovoľné jeho odvodenie je dĺžky mk pre vhodné m. $\square$

Veta 3.3.7. *$\mathcal{L}(X\text{-}CDGGS_n, f) \subseteq \mathcal{L}(X\text{-}CDGGS_{n+1}, f)$ pre ľubovoľné f.*

Dôkaz. Toto tvrdenie vyplýva priamo z definície. K systému n gramatík priradíme systém $n+1$ gramatík, kde $(n+1)$-vá gramatika nebude môcť participovať na odvodzovaní vetnej formy (napríklad bude mať prázdnu množinu pravidiel). $\square$

Veta 3.3.8. *$\mathcal{L}(B\text{-}CDGGS_2, t) = B\text{-}edNCE$.*

Dôkaz. Inklúzia $\supseteq$ vyplýva z predchádzajúcej vety, potrebujeme teda dokázať, že systém dvoch *B-edNCE* gramatík (uvažujeme samozrejme spôsob prepisovania t) vieme simulovať obyčajnou *B-edNCE* gramatikou. Nech $L \in \mathcal{L}(B\text{-}CDGGS_2, t)$, $\Theta = (\Delta, \Omega, G_0, G_1, S)$ je systém gramatík generujúci jazyk L pri terminálnom spôsobe prepisovania t.j. $L = L_t(\Theta)$, nech $G_i = (\Sigma_i, \Delta_i, \Gamma_i, \Omega_i, P_i)$ sú *B-edNCE* gramatiky bez počiatočného neterminálu (predpokladajme, že abecedy neterminálnych symbolov vrcholov Σ_0 a

Σ_1 neobsahujú symboly, ktoré by sa nenachádzali na ľavej strane nejakého pravidla z P_0 alebo P_1). Definujme *B-edNCE* gramatiku pre jazyk L takto:

$$
\begin{aligned}
G &= (\Sigma, \Delta, \Gamma, \Omega, P, S'), \text{ kde} \\
S' &\equiv S \text{ ak } S \notin L_{P_0} \cap L_{P_1} \\
& \quad \text{inak } S' \text{ je nový symbol} \\
\Sigma &= \Delta \cup \mathrm{M}_0 \cup \mathrm{M}_1 \cup \mathrm{E}_{\mathrm{P}_0} \cup \mathrm{E}_{\mathrm{P}_1} \cup \{\mathrm{S}' \mid \text{ak } \mathrm{S} \in \mathrm{L}_{\mathrm{P}_0} \cap \mathrm{L}_{\mathrm{P}_1}\} \\
\Gamma &= \Gamma_0 \cup \Gamma_1 \\
\mathrm{M_i} &= \{X_i \mid X \in L_{P_0} \cap L_{P_1}\} \text{ pre } i = \{0,1\} \\
E_{P_i} &= \Sigma_i - \Delta - L_{P_0} \cap L_{P_1}, \text{ pre } i = \{0,1\} \\
L_{P_i} &= \{X \mid X \to (D,C) \in P_i \text{ pre nejaký graf } (D,C)\} \\
P &= \{X_i \to (num(D,i),C) \mid X \to (D,C) \in P_i, i \in \{0,1\} \\
& \quad X \in L_{P_0} \cap L_{P_1},\} \\
& \quad \cup \\
& \quad \{X \to (num(D,0),C) \mid X \in E_{P_0}, X \to (D,C) \in P_0\} \\
& \quad \cup \\
& \quad \{X \to (num(D,1),C) \mid X \in E_{P_1}, X \to (D,C) \in P_1\} \\
& \quad \cup \\
& \quad \{S' \to (sn(S_i,z),\emptyset) \mid i \in \{0,1\}; \text{ ak } S \in L_{P_0} \cap L_{P_1}\} \\
num(D,i) &= (V_D, E_D, \lambda_{num(D,i)}) \\
\lambda_{num(D,i)}(v) &= \lambda_D(v) \text{ ak } \lambda_D(v) \in \Sigma_i - L_{P_0} \cap L_{P_1} \\
\lambda_{num(D,i)}(v) &= X_i \text{ ak } \lambda_D(v) = X \in L_{P_0} \cap L_{P_1}
\end{aligned}
$$

Predpokladáme, že prvky množín $\mathrm{M_i}$ a aj S' sú všetko nové symboly a teda, že množiny Δ, M_0, M_1, E_{P_0}, E_{P_1}, $\{S'\}$ sú po dvoch disjunktné.

Hlavnou myšlienkou uvedenej konštrukcie je to, že pre každý neterminál, ktorý je spoločný pre obe gramatiky, máme dve jeho verzie. To nám spolu s príslušnou úpravou pravidiel umožní rozlišovať medzi odvodeniami v jednotlivých gramatikách, a tým vlastne simulovať odvodenie v Θ.

Uvažujme ako vyzerá odvodenie ľubovoľného terminálneho grafu H v Θ

$$sn(S,z) \Rightarrow^t_{G_0} H_{1_0} \Rightarrow^t_{G_1} H_{1_1} \Rightarrow^t_{G_0} H_{2_0} \Rightarrow^t_{G_1} H_{2_1} \Rightarrow^t_{G_0} \ldots \Rightarrow^t_{G_1} H_{n_1} \equiv H$$

Ľahko vidno, že žiadny graf H_{i_j} neobsahuje vrchol s neterminálnym symbolom z abecedy $L_{P_0} \cap L_{P_1}$ (to by bolo v spore s terminálnym spôsobom prepisovania) a všetky symboly neterminálnych vrcholov grafu H_{i_0} sú z abecedy E_{P_1} (resp. z abecedy E_{P_0} pre grafy H_{i_1}). Na overenie, že graf H možno vygenerovať aj v gramatike G, stačí dokázať, že ak pre ľuboľné dva grafy H_1 a H_2 z horeuvedeného odvodenia platí $H_1 \Rightarrow^t_{G_i} H_2$, tak $H_1 \Rightarrow^*_G H_2$. Nech teda

$$H_1 \equiv H_{j_0} \Rightarrow_{v_1,p_1} H_{j_1} \Rightarrow_{v_2,p_2} H_{j_2} \Rightarrow_{v_3,p_3} \ldots \Rightarrow_{v_k,p_k} H_{j_k} \equiv H_2$$

z definície gramatiky G vyplýva, že ak $H_{j_k} \Rightarrow_{v_k,p_k} H_{j_{k+1}}$ (v G_i), tak

$$num(H_{j_k}, i) \Rightarrow_{v_k,num(p_k,i)} num(H_{j_{k+1}}, i) \text{ (v } G)$$

dôsledkom čoho v G existuje takéto odvodenie grafu H_2:

$$\begin{gathered} H_1 \equiv H_{j_0} \Rightarrow_{v_1,num(p_1,i)} \\ num(H_{j_1}, i) \Rightarrow_{v_2,num(p_2,i)} \\ num(H_{j_2}, i) \Rightarrow_{v_3,num(p_3,i)} \\ \ldots \\ num(H_{j_{k-1}}, i) \Rightarrow_{v_k,num(p_k,i)} \\ H_{j_k} \equiv H_2 \end{gathered}$$

čo bolo treba dokázať.

Ostáva nám ešte ukázať, že ľubovoľný graf, ktorý vygenerujeme v gramatike G, vieme vygenerovať aj v systéme gramatík Θ. Neterminálne symboly vrcholov a pravidlá gramatiky G vieme rozdeliť na dve disjunktné množiny tak, že pravidlá prvej skupiny pravidiel sú aplikovateľné len na neterminálne symboly z prvej skupiny neterminálov a pravidlá druhej skupiny pravidiel sú aplikovateľné len na neterminálne symboly z druhej skupiny neterminálov. Presnejšie,

$$N_i = M_i \cup E_{P_i}$$

a

$$\begin{aligned} R_i \;=\; & \{X_i \to (num(D,i), C) \mid X \to (D, C) \in P_i, X \in L_{P_0} \cap L_{P_1},\} \\ & \cup \\ & \{X \to (num(D,i), C) \mid X \in E_{P_i}, X \to (D, C) \in P_i\} \\ & \cup \\ & \{S' \to (sn(S_i, z), \emptyset) \mid \text{ak } S \in L_{P_0} \cap L_{P_1}\} \end{aligned}$$

Uvažujme odvodenie α ľubovoľného terminálneho grafu H v G, k tomuto odvodeniu nájdeme také odvodenie grafu H pre ktoré bude platiť, že na každú vetnú formu, ktorá vznikla z predchádzajúcej vetnej formy aplikovaním pravidla zo skupiny R_i, je aplikované pravidlo z druhej skupiny pravidiel $R_{i \oplus_2 1}$ len ak na ňu nie je možné aplikovať pravidlo zo skupiny R_i. Ak odvodenie α nespĺňa uvedenú podmienku, existuje v ňom vetná forma I , ktorá ju porušuje (t.j. je na ňu aplikované pravidlo p z $R_{i \oplus_2 1}$). Predpokladajme, že I je najľavejšia taká vetná forma. Keďže I porušuje uvedenú podmienku, existuje v ňom neterminálny vrchol v, na ktorý je možné aplikovať pravidlo z R_i.

Vzhľadom na to, že α je terminálne odvodenie, na vrchol v je neskôr v tomto odvodení aplikované pravidlo r z R_i. Keďže gramatika G je B-$edNCE$ gramatika a teda aj konfluentná, tak môžeme použiť pravidlo r pred požitím pravidla p (t.j. namiesto postupnosti použitia pravidiel $p_1,\ldots,\ p_{j-1},p,$ $p_{j+1},\ldots,\ p_{k-1},r,p_{k+1},\ \ldots,\ p_n$ použijeme pravidlá v poradí $p_1,\ldots,\ p_{j-1},r,p,$ $p_{j+1},\ldots,\ p_{k-1},p_{k+1},\ \ldots,\ p_n$). Opakovaným použitím tohto algoritmu sa zväčšuje (maximálny) prefix odvodenia spĺňajúci uvedenú podmienku, teda po konečnom počte krokov dostaneme požadované odvodenie.

Z definície gramatiky vyplýva, že ak graf K obsahuje neterminálne vrcholy so symbolmi len z jednej skupiny neterminálov N_i, tak postupným aplikovaním pravidiel len z množiny R_i nemôžeme vygenerovať graf, ktorý by obsahoval vrcholy s neterminálnym symbolom z množiny M_i.

Uvažujme teraz odvodenie terminálneho grafu H, ktoré spĺňa hore uvedené podmienku, toto odvodenie môžeme zapísať nasledovne (ukážeme tu prípad, keď $S \in E_{P_i}$ pre nejaké $i \in \{0, 1\}$, v opačnom prípade $S \in L_{P_0} \cap L_{P_1}$

by bol dôkaz skoro rovnaký):

$$
\begin{gathered}
sn(S,z) \equiv I_{1_0} \\
\Rightarrow_{u_{1_0},p_{1_0}} I_{1_1} \Rightarrow_{u_{1_1},p_{1_1}} I_{1_2} \Rightarrow_{u_{1_2},p_{1_2}} \ldots \Rightarrow_{u_{1_{y_1-1}},p_{1_{y_1-1}}} I_{1_{y_1}} \equiv J_{1_0} \\
\Rightarrow_{v_{1_0},r_{1_0}} J_{1_1} \Rightarrow_{v_{1_1},r_{1_1}} J_{1_2} \Rightarrow_{v_{1_2},r_{1_2}} \ldots \Rightarrow_{v_{1_{z_1-1}},r_{1_{z_1-1}}} J_{1_{z_1}} \equiv I_{2_0} \\
\Rightarrow_{u_{2_0},p_{2_0}} I_{2_1} \Rightarrow_{u_{2_1},p_{2_1}} I_{2_2} \Rightarrow_{u_{2_2},p_{2_2}} \ldots \Rightarrow_{u_{2_{y_2-1}},p_{2_{y_2-1}}} I_{2_{y_2}} \equiv J_{2_0} \\
\Rightarrow_{v_{2_0},r_{2_0}} J_{2_1} \Rightarrow_{v_{2_1},r_{2_1}} J_{2_2} \Rightarrow_{v_{2_2},r_{2_2}} \ldots \Rightarrow_{v_{2_{z_2-1}},r_{2_{z_2-1}}} J_{2_{z_2}} \equiv I_{3_0} \\
\Rightarrow_{u_{3_0},p_{3_0}} \ldots \\
\Rightarrow_{u_{k_0},p_{k_0}} I_{k_1} \Rightarrow_{u_{k_1},p_{k_1}} I_{k_2} \Rightarrow_{u_{k_2},p_{k_2}} \ldots \Rightarrow_{u_{k_{y_k-1}},p_{k_{y_k-1}}} I_{k_{y_k}} \equiv J_{k_0} \\
\Rightarrow_{v_{k_0},r_{k_0}} J_{k_1} \Rightarrow_{v_{k_1},r_{k_1}} J_{k_2} \Rightarrow_{v_{k_2},r_{k_2}} \ldots \Rightarrow_{v_{k_{z_k-1}},r_{k_{z_k-1}}} J_{k_{z_k}} \equiv I_{k+1_0} \\
(\Rightarrow_{u_{k+1_0},p_{k+1_0}} I_{k+1_1} \Rightarrow_{u_{k+1_1},p_{k+1_1}} \ldots \Rightarrow_{u_{k+1_{y_{k+1}-1}},p_{k+1_{y_{k+1}-1}}} I_{k+1_{y_{k+1}}}) \\
\equiv H
\end{gathered}
$$

pričom platí, že všetky pravidlá p_{s_t} sú z množiny R_i, pravidlá r_{s_t} sú z množiny $R_{i\oplus_2 1}$, grafy $J_{m_{z_m}}$ obsahujú neterminálne vrcholy len z množiny E_{P_i}, grafy $I_{m_{y_m}}$ obsahujú neterminálne vrcholy len z množiny $E_{P_{i\oplus_2 1}}$ a žiadna vetná forma v odvodení neobsahuje súčasne neterminálne vrcholy so symbolmi z množín M_0 aj M_1. Z toho vyplýva, že pre každý graf I_{s_t} existuje taký graf K_{s_t}, že $I_{s_t} = num(K_{s_t}, i)$, podobne pre J_{s_t} existuje taký graf L_{s_t}, že $J_{s_t} = num(L_{s_t}, i \oplus_2 1)$, označme vzor pravidla p_{s_t} ako p'_{s_t} a vzor pravidla r_{s_t} ako r'_{s_t}. Z definície gramatiky potom platí

$$L_{s_0} \Rightarrow_{v_{s_0},r'_{s_0}} L_{s_1} \Rightarrow_{v_{s_1},r'_{s_1}} L_{s_2} \Rightarrow_{v_{s_2},r'_{s_2}} \ldots \Rightarrow_{v_{s_{z_s-1}},r'_{s_{z_s-1}}} L_{s_{z_s}}$$

v $G_{i\oplus_2 1}$, zjavne odvodenie už nemôže v $G_{i\oplus_2 1}$ ďalej pokračovať, teda

$$L_{s_0} \Rightarrow^t_{G_{i\oplus_2 1}} L_{s_{z_s}}$$

(a analogicky pre K_{s_t} a p'_{s_t} v gramatiku G_1), čo bolo treba dokázať. □

Poznámka 3.3.9. *Horeuvedené tvrdenie platí aj pre lineárne gramatiky (s analogickou konštrukciou a dôkazom).*

Veta 3.3.10. *Nech* $X \in \{ED, B, LIN\}$*, potom platí nasledujúca rovnosť* $\mathcal{L}(X\text{-}CDGGS_*, t) = \mathcal{L}(X\text{-}CDGGS_3, t)$.

Dôkaz. Inklúzia $\supseteq$ vyplýva priamo z definície, stačí nám teda dokázať, že

$$\mathcal{L}(X\text{-}CDGGS_*(t) \subseteq \mathcal{L}(X\text{-}CDGGS_3(t)$$

Nech $L \in \mathcal{L}(X\text{-}CDGGS_*, t)$, $\Theta = (\Delta, \Omega, G_1, G_2, \ldots, G_n, S)$ je X-$CDGGS_*$ systém gramatík generujúci jazyk L pri terminálnom spôsobe prepisovania (kde $G_i = (\Sigma_i, \Delta_i, \Gamma_i, \Omega_i, P_1)$) t.j. $L = L_t(\Theta)$ (bez ujmy na všeobecnosti predpokladáme, že n je párne a väčšie ako 3). X-$CDGGS_3$ systém gramatík generujúci jazyk L pri terminálnom spôsobe prepisovania bude pozostávať z troch komponent, z ktorých jedna bude simulovať odvodenie v gramatikách G_i, zvyšné dve budú len "prepínať" simulovanú gramatiku. Neterminálne symboly vrcholov vetných foriem tohoto systému budú obsahovať *index* určujúci, ktorú gramatiku systému Θ má prvá gramatika simulovať. Táto gramatika bude vždy pracovať na grafe, ktorého všetky neterminálne vrcholy majú rovnaké indexy a bude teda pracovať dovtedy, kým by pracovala simulovaná gramatika (na vetnej forme bez indexov). Druhá gramatika začína pracovať na vetných formách, ktorých všetky neterminálne vrcholy majú index rovnaké nepárne číslo a postupne zväčšuje indexy všetkých neterminálnych vrcholov o 1, čiže po skončení jej práce majú všetky neterminálne vrcholy vygenerovaného grafu index rovnaké párne číslo. Na takejto vetnej forme môže začať pracovať tretia gramatika, ktorá podobne ako druhá len postupne zväčší indexy všetkých neterminálnych vrcholov o 1, čiže po skončení jej práce majú všetky neterminálne vrcholy vygenerovaného grafu index rovnaké nepárne číslo.

Definujme teda gramatiku Θ' nasledovne

$$
\begin{aligned}
\Theta' &= (\Delta, \Omega, G_1', G_2', G_3', [S,1]) \\
\\
G_1' &= (\Sigma', \Delta, \Gamma', \Omega, P_1') \\
\Sigma' &= \Delta \cup \bigcup_{1 \le i \le n} \{[Y,i] \mid Y \in \Sigma_i - \Delta_i\} \\
\Gamma' &= \bigcup_{1 \le i \le n} \Gamma_i \\
P_1' &= \{[Y,i] \to ([D,i], C(i)) \mid Y \to (D,C) \in P_i, 1 \le i \le n\} \\
[D,i] &= (V_D, E_D, \lambda), \text{ kde} \\
\lambda(v) &= \lambda_D(v), \text{ ak } \lambda_D(v) \in \Delta_i \\
\lambda(v) &= [\lambda_D(v), i], \text{ ak } \lambda_D(v) \in \Sigma_i - \Delta_i \\
C(i) &= \{([Y,i], \alpha/\beta, y, io) \mid (Y, \alpha/\beta, y, io) \in C, Y \in \Sigma_i - \Delta_i\} \\
&\quad \cup \\
&\quad \{(Y, \alpha/\beta, y, io) \mid (Y, \alpha/\beta, y, io) \in C, Y \in \Delta_i\} \\
\\
G_2' &= (\Sigma', \Delta, \Gamma', \Omega, P_2') \\
P_2' &= \{[Y,i] \to (sn([Y,i+1], y), C(y)) \mid Y \in \Sigma', \\
&\quad 1 \le i \le n, i \text{ je nepárne}\} \\
C(y) &= \{(Y, \alpha/\alpha, y, i) \mid Y \in \Sigma', \alpha \in \Gamma', i \in \{in, out\}\} \text{ ak } X \equiv ED \\
C(y) &= \{(Y, \alpha/\alpha, y, i) \mid Y \in \Delta, \alpha \in \Gamma', i \in \{in, out\}\} \text{ ak } X \in \{B, LIN\} \\
\\
G_3' &= (\Sigma', \Delta, \Gamma', \Omega, P_3') \\
P_3' &= \{[Y,i] \to (sn([Y,i+1], y), C(y)) \mid Y \in \Sigma', 1 \le i < n, i \text{ je párne}\} \\
&\quad \cup \\
&\quad \{[Y,n] \to (sn([Y,1], y), C(y)) \mid Y \in \Sigma'\}
\end{aligned}
$$

To, že je definovaný systém X-$CDGGS_3$, vyplýva triviálne z jeho definície (konštrukcie).

Pre neterminálny vrchol v grafu $H \in GRE_{\Sigma',\Gamma'}$ definujme funkciu *index* takto:

$$index(v) = i \text{ ak } \lambda_H(v) = [Y, i]$$

Najskôr ukážeme, ako vyzerajú vetné formy tohto systému gramatík. Ľahko vidno, že ak pre graf H platí, že indexy všetkých jeho neterminálnych vrcholov sú rovné i, potom po odvodení $H \Rightarrow^*_{G'_1} H'$ platí pre graf H' rovnaká rovnosť. Formálne, ak $H \Rightarrow^*_{G'_1} H'$ a existuje $i \in \{1, \ldots, n\}$ také, že pre každý neterminálny vrchol $v \in V_H$ platí $index(v) = i$, tak pre každý neterminálny vrchol $v \in V_{H'}$ platí $index(v) = i$.

Podobné tvrdenia platia pre (terminálne) odvodenia v gramatikách G'_2 a G'_3. T.j. ak $H \Rightarrow^t_{G'_2} H'$ (resp. $H \Rightarrow^t_{G'_3} H'$) a existuje $i \in \{1, \ldots, n\}$ také, že pre každý neterminálny vrchol $v \in V_H$ platí $index(v) = i$, tak pre každý neterminálny vrchol $v \in V_{H'}$ platí $index(v) = ((i-1) \oplus_n 1) + 1$ (neformálne povedané indexy vrcholov v grafe H' sú rovné $i+1$, až na prípad, keď $i = n$, vtedy sú rovné 1).

Ak teda uvažujeme odvodenie ľubovoľnej vetnej formy H v Θ takej, že

$$sn([S,1], z) \Rightarrow^t_{G'_{i_1}} H_1 \Rightarrow^t_{G'_{i_2}} H_2 \ldots \Rightarrow^t_{G'_{i_k}} H_k \equiv H,$$

tak v dôsledku horeuvedených tvrdení majú neterminálne vrcholy grafu H rovnaké indexy.

Ďalej platí nasledujúci vzťah medzi odvodeniami v Θ a Θ':

$$\begin{gathered} H \Rightarrow_{v,p} I \text{ je krok odvodenia v gramatike } G_i \\ \Leftrightarrow \\ [H,i] \Rightarrow_{v,[p,i]} [I,i] \text{ je krok odvodenia v gramatike } G'_1 \end{gathered}$$

Potrebujeme vlastne ukázať, že pre graf J taký, že $[H,i] \Rightarrow_{v,[p,i]} J$ platí $J = [I,i]$. Z definície kroku odvodenia je $I = (H, \emptyset)\,[v/(D,C)]$, teda

$$\begin{array}{rcl}
I &=& (V, E, \lambda, \emptyset), \text{ kde} \\
V &=& (V_H - \{v\}) \cup V_D \\
E &=& E_D \cup \{(x,\gamma,y) \in E_H \mid x \neq v, y \neq v\} \\
&& \cup \\
&& \{(w,\gamma,x) \mid \exists \beta \in \Gamma : (w,\beta,v) \in E_H, (\lambda_H(w), \beta/\gamma, x, in) \in C_D\} \\
&& \cup \\
&& \{(x,\gamma,w) \mid \exists \beta \in \Gamma : (v,\beta,w) \in E_H, (\lambda_H(w), \beta/\gamma, x, out) \in C_D\} \\
\lambda(x) &=& \lambda_H(x) \text{ ak } x \in V_H - \{v\} \text{ a} \\
\lambda(x) &=& \lambda_D(x) \text{ ak } x \in V_D
\end{array}$$

Z definície gramatiky zase platia takéto vzťahy medzi grafmi H a $[H, i]$:

$$\begin{aligned} [H,i] &= (V_H, E_H, \lambda), \text{ kde} \\ \lambda(v) &= \lambda_H(v) \text{ ak } \lambda_H(v) \in \Delta_i \\ \lambda(v) &= [\lambda_H(v), i] \text{ ak } \lambda_H(v) \in \Sigma_i - \Delta_i \end{aligned}$$

a medzi pravidlami $p : X \to (D, C)$ a $[p, i] : [X, i] \to ([D, i], C(i))$:

$$\begin{aligned} [D,i] &= (V_D, E_D, \lambda), \text{ kde} \\ \lambda(v) &= \lambda_D(v), \text{ ak } \lambda_D(v) \in \Delta_i \\ \lambda(v) &= [\lambda_D(v), i], \text{ ak } \lambda_D(v) \in \Sigma_i - \Delta_i \\ C(i) &= \{([X,i], \alpha/\beta, x, io) \mid (X, \alpha/\beta, x, io) \in C\} \end{aligned}$$

z čoho vieme (z definície kroku odvodenia) určiť vzťah medzi H a grafom J, ktorý odvodíme z $[H, i]$:

$$\begin{aligned} J &= (V', E', \lambda', \emptyset), \text{ kde} \\ V' &= (V_H - \{v\}) \cup V_D \\ E' &= E_D \cup \{(x, \gamma, y) \in E_H \mid x \neq v, y \neq v\} \\ &\quad \cup \\ &\quad \{(w, \gamma, x) \mid \exists \beta \in \Gamma : (w, \beta, v) \in E_H, (\lambda_H(w), \beta/\gamma, x, in) \in C_D\} \\ &\quad \cup \\ &\quad \{(x, \gamma, w) \mid \exists \beta \in \Gamma : (v, \beta, w) \in E_H, (\lambda_H(w), \beta/\gamma, x, out) \in C_D\} \\ \lambda(x) &= \lambda_H(x) \text{ ak } x \in V_H - \{v\} \text{ a } \lambda_H(v) \in \Delta_i \\ \lambda(x) &= [\lambda_H(x), i] \text{ ak } x \in V_H - \{v\} \text{ a } \lambda_H(v) \in \Sigma_i - \Delta_i \\ \lambda(x) &= \lambda_D(x) \text{ ak } x \in V_D \text{ a } \lambda_D(v) \in \Delta_i \\ \lambda(x) &= [\lambda_D(x), i] \text{ ak } x \in V_D \text{ a } \lambda_D(v) \in \Sigma_i - \Delta_i \end{aligned}$$

Z toho už ľahko vidno, že $J = [I, i]$.

Ostáva teda ešte dokázať rovnosť $L_t(\Theta) = L_t(\Theta')$. Nech $[H] \in L_t(\Theta)$, z definície jazyka $L_t(\Theta)$ existuje takéto odvodenia grafu H v systéme gramatík Θ:

$$sn(S, z) \Rightarrow^t_{G_{i_1}} H_1 \Rightarrow^t_{G_{i_2}} H_2 \ldots \Rightarrow^t_{G_{i_k}} H_n \equiv H$$

Potrebujeme dokázať, že existuje odvodenie rovnakého grafu v Θ' z počiatočného grafu $sn([S, 1], z)$. Vzhľadom na vyššie dokázané tvrdenia stačí ukázať, že pre ľubovoľný graf H, pre ľubovoľné $i, j \in \{1, \ldots, n\}$ platí:

$$[H, i] \Rightarrow^t_{G'_{i_1}} [H, ((i-1) \oplus_n 1) + 1] \Rightarrow^t_{G'_{i_2}} \ldots \Rightarrow^t_{G'_{i_k}} [H, j],$$

kde $i_l \in \{2,3\}$. Uvažujme graf $[H,i]$, odvodenie $[H,i] \Rightarrow_{G'_2} I$, predpokladajme bez ujmy na všeobecnosti, že i je párne (pre nepárne i by zdôvodnenie vyzeralo velmi podobne, len by sme uvažovali gramatiku G'_3 namiesto G'_2 a vymenila by sa párnosť za nepárnosť) a pozrime sa, ako bude vyzerať graf I odvodený z $[H,i]$. Bez toho, aby sme museli rozpisovať definíciu grafu I pomocou substitúcie použitej pri jeho odvodení, vidíme, že graf I je rovnaký ako graf $[H,i]$, až na symbol σ nahradzovaného vrchola v, ktorý je rovný $[X,((i-1)\oplus_n 1)+1]$, ak $\sigma \equiv [X,i]$. Z definície gramatiky G'_2 je jasné, že na vrchol v už nemôže byť žiadne iné pravidlo tejto gramatiky aplikované, keďže jeho index je párne číslo. Ak $[H,i]$ má m neterminálnych vrcholov, tak po m krokoch vygenerujeme graf $[H,((i-1)\oplus_n 1)+1]$ a odvodenie už v gramatike G'_2 nemôže ďalej pokračovať. Dôsledkom je, že pre ľubovoľné $i,j \in \{1,\ldots,n\}$, $i<j$ existuje odvodenie

$$[H,i] \Rightarrow^t_{G'_{i_1}} [H,((i-1)\oplus_n 1)+1] \Rightarrow^t_{G'_{i_2}} \ldots \Rightarrow^t_{G'_{i_{j-i}}} [H,j]$$

a pre $i>j$ existuje odvodenie

$$[H,i] \Rightarrow^t_{G'_{i_1}} [H,((i-1)\oplus_n 1)+1] \Rightarrow^t_{G'_{i_2}} \ldots \Rightarrow^t_{G'_{i_{j+n-i}}} [H,j].$$

Pre odvodenie v systéme gramatík Θ:

$$sn(S,z) \Rightarrow^t_{G_{i_1}} H_1 \Rightarrow^t_{G_{i_2}} H_2 \Rightarrow^t_{G_{i_3}} \ldots \Rightarrow^t_{G_{i_k}} H_k \equiv H$$

teda existuje takéto odvodenie v systéme gramatík Θ':

$$\begin{gathered}
sn([S,1],z) \Rightarrow^t_{G'_{r_{1_1}}} \\
sn([S,(0\oplus_n 1)+1],z) \Rightarrow^t_{G'_{r_{1_2}}} \ldots \Rightarrow^t_{G'_{r_{1_q}}} \\
sn([S,i_1],z) \Rightarrow^t_{G'_1} [H_1,i_1] \Rightarrow^t_{G'_{r_{2_1}}} \\
[H_1,((i_1-1)\oplus_n 1)+1] \Rightarrow^t_{G'_{r_{2_2}}} \ldots \Rightarrow^t_{G'_{r_{2_q}}} \\
[H_1,i_2] \Rightarrow^t_{G'_1} [H_2,i_2] \Rightarrow^t_{G'_{r_{3_1}}} \\
[H_2,((i_2-1)\oplus_n 1)+1] \Rightarrow^t_{G'_{r_{3_2}}} \ldots \Rightarrow^t_{G'_{r_{3_q}}} \\
\ldots \Rightarrow^t_{G'_1} [H_{k-1},i_{k-1}] \Rightarrow^t_{G'_{r_{k_1}}} \\
[H_{k-1},((i_{k-1}-1)\oplus_n 1)+1] \Rightarrow^t_{G'_{r_{k_2}}} \ldots \Rightarrow^t_{G'_{r_{k_q}}} \\
[H_{k-1},i_k] \Rightarrow^t_{G'_1} [H_k,i_k] \equiv H
\end{gathered}$$

kde $r_{s_t} \in \{2,3\}$, čo bolo treba dokázať.

Naopak, uvažujme (najkratšie) odvodenie ľubovoľného terminálneho grafu H v $L_t(\Theta')$, podľa argumentov uvedených v tomto dôkaze, musí byť toto odvodenie tvaru

$$
\begin{gathered}
sn([S,1],z) \Rightarrow^t_{G'_{r_{1_1}}} \\
sn([S,(0 \oplus_n 1)+1],z) \Rightarrow^t_{G'_{r_{1_2}}} \ldots \Rightarrow^t_{G'_{r_{1_q}}} \\
sn([S,i_1],z) \Rightarrow^t_{G'_1} [H_1,i_1] \Rightarrow^t_{G'_{r_{2_1}}} \\
[H_1,((i_1-1) \oplus_n 1)+1] \Rightarrow^t_{G'_{r_{2_2}}} \ldots \Rightarrow^t_{G'_{r_{2_q}}} \\
[H_1,i_2] \Rightarrow^t_{G'_1} [H_2,i_2] \Rightarrow^t_{G'_{r_{3_1}}} \\
[H_2,((i_2-1) \oplus_n 1)+1] \Rightarrow^t_{G'_{r_{3_2}}} \ldots \Rightarrow^t_{G'_{r_{3_q}}} \\
\ldots \Rightarrow^t_{G'_1} [H_{k-1},i_{k-1}] \Rightarrow^t_{G'_{r_{k_1}}} \\
[H_{k-1},((i_{k-1}-1) \oplus_n 1)+1] \Rightarrow^t_{G'_{r_{k_2}}} \ldots \Rightarrow^t_{G'_{r_{k_q}}} \\
[H_{k-1},i_k] \Rightarrow^t_{G'_1} [H_k,i_k] \equiv H
\end{gathered}
$$

kde $r_{s_t} \in \{2,3\}$ a zodpovedajúce odvodenie grafu H v systéme gramatík $L(\Theta)$ je

$$sn(S,z) \Rightarrow^t_{G_{i_1}} H_1 \Rightarrow^t_{G_{i_2}} H_2 \Rightarrow^t_{G_{i_3}} \ldots \Rightarrow^t_{G_{i_k}} H_k \equiv H$$

$\square$

Kapitola 4

Paralelné Komunikujúce Systémy Grafových Gramatík

4.1 Úvod

Tento model predstavuje analógiu $PCGS$ systémov definovaných pre klasické stringové gramatiky. Paralelný komunikujúci systém grafových gramatík v sebe integruje viacero gramatík rovnakého typu, ktoré sú zosynchronizované podľa akýchsi globálnych hodín, ktoré pracujú v taktoch. Každá gramatika pracuje na svojej vetnej forme podľa vlastných pravidiel. Naviac týmto gramatikám dodávame špeciálny komunikačný symbol. Ak sa vo vetnej forme nejakej gramatiky vyskytne vrchol s komunikačným symbolom Q_i, znamená to, že v ďalšom takte bude komunikačný vrchol nahradený vetnou formou vygenerovanou i-tou gramatikou (pričom táto začne generovať odznova). Toto nahradenie sa vykoná, len ak táto gramatika neobsahuje sama nejaký komunikačný vrchol. Formálne zadefinujeme tento model v prvej časti tejto kapitoly.

V prípade grafových gramatík ale nastáva pri definícii jeden problém, ktorý znemožňuje zadefinovať tento model pre všeobecné (C-)$edNCE$ gramatiky - pri komunikačom kroku môže nejaká komponenta vetnej formy obsahovať viacej komunikačných vrcholov, čo znamená, že sa do tejto vetnej formy v jednom kroku substituuje viacero grafov. V prípade stringových gramatík s týmto nebol problém, keďže vysledok paralelného nahradenia viacerých neterminálov je jednoznačne určený (je vlastne rovný sekvenčnému prepisovaniu neterminálov v ľubovoľnom poradí). V prípade grafových gramatík, aj keď by sme uvažovali systém, čo má konfluentné komponentové gramatiky, nie je možné paralelnú substitúciu rozumne definovať aj vzhľadom na to, že výsledok nahradenia viacerých komunikačných vrcholov môže byť závislý na poradí týchto krokov. Preto definujeme $PCGGS$, ktorého komponentové gramatiky sú B-$edNCE$ (samozrejme má zmysel uvažovať aj niektoré podtriedy B-$edNCE$), kde problém s jednoznačnosťou paralelného nahradenia viacerých komunikačných vrcholov nenastáva. Ďalšou možnosťou by mohlo byť, ak by sme požadovali, aby boli konfluentné všetky pravidlá navzájom, nie len v jednotlivých gramatikách.

4.2 Definície a označenia

Definícia 4.2.1. (PCGGS) *Paralelný Komunikujúci Systém Grafových Gramatík (PCGGS) stupňa n je $(n+5)$-tica*

$$\Theta = (\Sigma, \Psi, \Delta, \Gamma, \Omega, G_1, G_2, \ldots, G_n)$$

kde

- Σ *je abeceda neterminálnych symbolov vrcholov,*
- Ψ *je abeceda komunikačných symbolov; štandardne ich označujeme ako* $\Psi = \{Q_1, Q_2, \ldots, Q_n\}$,
- Δ *je abeceda terminálnych symbolov vrcholov,*
- Σ, Ψ *a* Δ *sú po dvoch disjunktné,*
- Γ *je abeceda neterminálnych hranových symbolov,*
- Ω *je abeceda terminálnych hranových symbolov,*
- $\Gamma \cap \Omega = \emptyset$
- *pre každé* $i \in \{1, \ldots, n\}$ *je* $G_i = (\Sigma \cup \Psi \cup \Delta, \Delta, \Gamma \cup \Omega, \Omega, P_i, S_i, C_i(z_i))$ B-$edNCE$ *gramatika s počiatočnou spojovacou reláciou, pre ktorej ľubovoľné pravidlo* $X \rightarrow (D, C)$ *platí, že v* P_i *nie sú pravidlá obsahujúce na ľavej strane komunikačný symbol.*

Označenie 4.2.2. (Komunikačný vrchol) *Vrchol so symbolom* $Q_i \in \Psi$ *nazývame komunikačný vrchol.*

Definícia 4.2.3. *Množinu* $\Sigma \cup \Psi \cup \Delta$ *označujeme ako* Λ_V*, množinu* $\Gamma \cup \Omega$ *označujeme ako* Λ_E*.*

Definícia 4.2.4. (n-vetná forma) *Nech* Θ *je PCGGS,* n*-vetná forma resp. konfigurácia systému* Θ *je* n*-tica grafov* $(H_1, H_2, \ldots, H_n)$*, kde* $H_i \in GRE_{\Lambda_V, \Lambda_E}$*.*

Poznámka 4.2.5. *Namiesto pojmu* n*-vetná forma budeme niekedy používať pojem vetná forma.*

Definícia 4.2.6. (Krok odvodenia) *Nech* Θ *je PCGGS, krok odvodenia systému* Θ *je relácia na* n*-vetných formách definovaná nasledovne:*

$$(H_1, H_2, \ldots, H_n) \Rightarrow_\Theta (K_1, K_2, \ldots, K_n)$$

práve vtedy, keď nastane práve jeden z týchto prípadov:

1. *(prepisovací krok) žiaden graf* H_i *neobsahuje vrchol s komunikačným symbolom*
   - *ak* H_i *obsahuje neterminálny vrchol, tak* $H_i \Rightarrow_{G_i} K_i$
   - *ak* H_i *neobsahuje neterminálny vrchol, tak* $H_i = K_i$
2. *(komunikačný krok) niektoré grafy* H_i *obsahujú vrcholy s komunikačnými symbolmi: nech* H_i *obsahuje vrcholy s komunikačnými symbolmi* $Q_{j_1}, Q_{j_2}, \ldots, Q_{j_s}$
   - *ak žiaden z grafov* H_{j_k} *neobsahuje komunikačný vrchol, tak v* H_i *nahradíme* Q_{j_k} *(substitúciou, tak ako je definovaná v 2.1.9) grafom izomorfným s* H_{j_k}*, ktorý je disjunkný s grafom, do ktorého substituujeme a* $K_{j_k} = sn(S_{j_k}, z, C_{j_k}(z))$
   - *ak nejaký* H_{j_k} *obsahuje komunikačný vrchol, tak* $H_i = K_i$

   pre všetky ostatné H_i*, ktoré neobsahujú komunikačné symboly platí* $H_i = K_i$*.*

Definícia 4.2.7. (Grafový jazyk) *Jazyk generovaný PCGGS systémom* Θ *je množina*

$$\begin{aligned} L(\Theta) &= \{H \in G_{\Delta\Omega} \,/\, (sn(S_1, z_1, C_1(z_1)), sn(S_2, z_2, C_2(z_2)), \ldots, \\ &\quad sn(S_n, z_n, C_n(z_n))) \Rightarrow^* (H, H_2, \ldots, H_n)\} \end{aligned}$$

Označenie 4.2.8. *Nech X-edNCE je trieda grafových gramatík (budeme uvažovať B-edNCE a LIN-edNCE), potom $PCGGS_nX$ označuje triedu PCGGS systémov s n X-edNCE komponentovými gramatikami. Triedu jazykov generových takýmito systémami gramatík označujeme $\mathcal{L}(PCGGS_nX)$.*

Poznámka 4.2.9. *Ďalšou možnosťou, ako parametrizovať PCGGS systém, je uvažovať rôznu komunikačnú štruktúru - t.j. obmedzenie na to, ktoré gramatiky môžu generovať komunikačné vrcholy. V nasledujúcej časti dokážeme vetu ktorá hovorí, že pri použití všeobecnej komunikačnej štruktúry vieme s lineárnymi gramatikami vygenerovať grafový jazyk, ktorý nie je v C-edNCE.*

Definícia 4.2.10. (Komunikačná štruktúra) *Nech*

$$\Theta = (\Sigma, \Psi, \Delta, \Gamma, \Omega, G_1, \dots, G_n)$$

je PCGGS, komunikačná štruktúra systému Θ je orientovaný graf $G = (\{G_1, \dots, G_n\}, E)$ definovaný takto: $(G_i, G_j) \in E$, ak gramatika G_i obsahuje pravidlo, na ktorého pravej strane je graf obsahujúci komunikačný vrchol so symbolom Q_j.

4.3 Generatívna sila PCGGS

Najpv ukážeme, ako $PCGGS$ s rôznymi typmi a počtom komponentových gramatík dokáže vygenerovať grafový jazyk mimo triedy C-$edNCE$.

Veta 4.3.1. $\mathcal{L}(PCGGS_3LIN) - C\text{-}edNCE \neq \emptyset$

Dôkaz. Stačí ukázať (pozri str. 33 v [5]), že v $\mathcal{L}(PCGGS_3LIN)$ existuje grafový jazyk L taký, že jazyk $\{a^n \mid n = |V_H| \text{ pre nejaký graf } [H] \in L\}$ nie je bezkontextový. Definujme jazyk L takto:

$$\begin{aligned} L \;=\; & \{[H] \in GRE_{\{a\}\{*\}} \mid H = (V, E, \lambda), \text{ kde} \\ & V = \{1, 2, \dots, n\}, \\ & E = \{(i, *, i+1) \mid i \in \{1, 2, \dots, n-1\}\}, \\ & \lambda(i) = a \text{ pre ľubovoľné } i \in V, \\ & n = 2^m \text{ pre nejaké } m \geq 0\} \end{aligned}$$

Tento jazyk vieme generovať nasledujúcim $PCGGS_3LIN$ systémom gramatík Θ:

$$\begin{aligned}
\Theta &= (\Sigma, \Psi, \Delta, \Gamma, \Omega, G_1, G_2, G_3), \text{ kde}\\
\Sigma &= \{S_1, S_2, S_3, B, B_1\}\\
\Psi &= \{Q_1, Q_2, Q_3\}\\
\Delta &= \{a\}\\
\Gamma &= \emptyset\\
\Omega &= \{*\}
\end{aligned}$$

$$G_i = (\Sigma \cup \Psi \cup \Delta, \Delta, \Gamma \cup \Omega, \Omega, P_i, S_i, C(z)) \text{ pre } i \in \{1, 2, 3\}$$

$$\begin{aligned}
P_1 = \; &\{S_1 \to ((V = \{1, 2\}, E = \{(1, *, 2)\}, \lambda = \{(1, a), (2, B)\}),\\
&\{(a, */*, 1, in)\}),\\
&S_1 \to ((V = \{1\}, E = \emptyset, \lambda = \{(1, Q_2)\}), \emptyset),\\
&B_1 \to ((V = \{1\}, E = \emptyset, \lambda = \{(1, B)\}), \{(a, */*, 1, in)\}),\\
&B \to ((V = \emptyset, E = \emptyset, \lambda = \emptyset), \emptyset)\}
\end{aligned}$$

$$\begin{aligned}
P_2 = \; &\{S_2 \to ((V = \{1\}, E = \emptyset, \lambda = \{(1, Q_1)\}), \{(a, */*, 1, in)\}),\\
&B \to ((V = \{1\}, E = \emptyset, \lambda = \{(1, Q_3)\}), \{(a, */*, 1, in)\})\}
\end{aligned}$$

$$\begin{aligned}
P_3 = \; &\{S_3 \to ((V = \{1\}, E = \emptyset, \lambda = \{(1, Q_1)\}), \{(a, */*, 1, in)\}),\\
&B \to ((V = \{1\}, E = \emptyset, \lambda = \{(1, B_1)\}), \{(a, */*, 1, in)\})\}
\end{aligned}$$

$$C(z) = \{(a, */*, z, in)\} \text{ pre } i \in \{1, 2, 3\}$$

Uvažujme graf $[H] \in L$ s $n = 2^m$ vrcholmi, (takýto graf existuje len jeden), potrebujeme nájsť jeho odvodenie v Θ. Zrejme prvý krok odvodenia (terminálneho grafu) v Θ vyzerá takto:

$$\begin{gathered}
(sn(S_1, z_1, C(z_1)), sn(S_2, z_2, C(z_2)), sn(S_3, z_3, C(z_3)))\\
\Rightarrow\\
(H_{1_1}, sn(Q_1, w_2, C(w_2)), sn(Q_1, w_3, C(w_3)))
\end{gathered}$$

kde

$$H_{1_1} = (\{w_{1_1}, w_{1_2}\}, \{(w_{1_1}, *, w_{1_2})\}, \{(w_{1_1}, a), (w_{1_2}, B_1)\}, \{(a, */*, w_{1_1}, in)\})$$

t.j. H_{1_1} je graf s dvoma vrcholmi so symbolmi a a B spojenými hranou (keďže síce pre prvú komponentu n-vetnej formy máme dve pravidlá, ktoré možno použiť, avšak ak by sme použili druhé pravidlo, tak by odvodenie nemohlo ďalej pokračovať, pretože každá komponenta n-vetnej formy by obsahovala komunikačný symbol).

Predpokladajme teraz, že vieme vygenerovať n-vetnú formu

$$(H_1, sn(Q_1, y_2, C(y_2)), sn(Q_1, y_3, C(y_3)))$$

kde

$$\begin{aligned}
H_1 &= (V, E, \lambda, C) \\
V &= \{1, 2, \ldots, k+1\}, \\
E &= \{(i, *, i+1) \mid i \in \{1, 2, \ldots, k\}\}, \\
\lambda(i) &= a \text{ pre } i \in V - \{k+1\}, \\
\lambda(k+1) &= B \\
C &= \{(a, */*, 1, in)\}
\end{aligned}$$

Potom existuje v Θ odvodenie n-vetnej formy

$$(H_6, sn(Q_1, y_2, C(y_2)), sn(Q_1, y_3, C(y_3)))$$

kde

$$\begin{aligned}
H_6 &= (V, E, \lambda, C) \\
V &= \{1, 2, \ldots, 2k+1\}, \\
E &= \{(i, *, i+1) \mid i \in \{1, 2, \ldots, 2k\}\}, \\
\lambda(i) &= a \text{ pre } i \in V - \{2k+1\}, \\
\lambda(2k+1) &= B \\
C &= \{(a, */*, 1, in)\}
\end{aligned}$$

Spomínané odvodenie vyzerá nasledovne

$$
\begin{gathered}
(H_1, sn(Q_1, y_2, C(y_2)), sn(Q_1, y_3, C(y_3))) \\
\Rightarrow \\
(sn(S_1, z_1, C(z_1)), H_1, H_1) \\
\Rightarrow \\
(sn(Q_2, w_1, C(w_1)), H_{3_2}, H_{3_3}) \\
\Rightarrow \\
(sn(Q_2, w_1, C(w_1)), H_{4_2}, sn(S_3, y_3, C(y_3))) \\
\Rightarrow \\
(H_{5_1}, sn(S_2, y_2, C(y_2)), sn(S_3, y_3, C(y_3))) \\
\Rightarrow \\
(H_6, sn(Q_1, v_2, C(v_2)), sn(Q_1, v_3, C(v_3)))
\end{gathered}
$$

kde

$$
\begin{aligned}
H_{3_2} &= (V_{3_2}, E_{3_2}, \lambda_{3_2}, C_{3_2}) \\
V_{3_2} &= \{1, 2, \ldots, k+1\}, \\
E_{3_2} &= \{(i, *, i+1) \mid i \in \{1, 2, \ldots, k\}\}, \\
\lambda_{3_2}(i) &= a \text{ pre } i \in V - \{k+1\}, \\
\lambda_{3_2}(k+1) &= Q_3 \\
C_{3_2} &= \{(a, */*, 1, in)\}
\end{aligned}
$$

$$
\begin{aligned}
H_{3_3} &= (V_{3_3}, E_{3_3}, \lambda_{3_3}, C_{3_3}) \\
V_{3_3} &= \{1, 2, \ldots, k+1\}, \\
E_{3_3} &= \{(i, *, i+1) \mid i \in \{1, 2, \ldots, k\}\}, \\
\lambda_{3_3}(i) &= a \text{ pre } i \in V - \{k+1\}, \\
\lambda_{3_3}(k+1) &= B_1 \\
C_{3_3} &= \{(a, */*, 1, in)\}
\end{aligned}
$$

$$\begin{aligned}
H_{4_2} &= (V_{4_2}, E_{4_2}, \lambda_{4_2}, C_{4_2}) \\
V_{4_2} &= \{1, 2, \ldots, 2k+1\}, \\
E_{4_2} &= \{(i, *, i+1) \mid i \in \{1, 2, \ldots, 2k\}\}, \\
\lambda_{4_2}(i) &= a \text{ pre } i \in V - \{2k+1\}, \\
\lambda_{4_2}(2k+1) &= B_1 \\
C_{4_2} &= \{(a, */*, 1, in)\}
\end{aligned}$$

$$\begin{aligned}
H_{5_1} &= (V_{5_1}, E_{5_1}, \lambda_{5_1}, C_{5_1}) \\
V_{5_1} &= \{1, 2, \ldots, 2k+1\}, \\
E_{5_1} &= \{(i, *, i+1) \mid i \in \{1, 2, \ldots, 2k\}\}, \\
\lambda_{5_1}(i) &= a \text{ pre } i \in V - \{2k+1\}, \\
\lambda_{5_1}(2k+1) &= B \\
C_{5_1} &= \{(a, */*, 1, in)\}
\end{aligned}$$

Z vyššie uvedených faktov vyplýva, že v gramatike Θ existuje odvodenie n-vetnej formy, kde prvá komponenta je graf H_1 s 2^m+1 vrcholmi pre ľubovoľné $m \geq 0$. Aplikovaním štvrtého pravidla z P_1 dostaneme odvodenie hľadaného grafu H.

Potrebujeme ešte ukázať, že všetky grafy, ktoré vieme v Θ odvodiť sú z jazyka L. Zrejme stačí ukázať, že horeuvedené odvodenie je jediné možné (resp. že všetky ostatné varianty použitia pravidiel vedú buď hneď k odvodeniu platného terminálneho grafu alebo k zaseknutiu odvodenia).

Prvý krok je komunikačný a druhá n-vetná forma je jednoznačne určená. V druhom kroku môžeme na prvú komponentu namiesto druhého pravidla použiť prvé pravidlo. V tomto prípade sa odvodenie za dva kroky zasekne, keďže druhá komponenta n-vetnej formy obsahuje graf s vrcholom so symbolom B_1 (pre ktorý nie je v druhej gramatike pravidlo) a žiadna komponenta n-vetnej formy neobsahuje vrchol s komunikačným symbolom. Tretí a štvrtý krok su komunikačné, a teda štvrtá a piata n-vetná forma sú jednoznačne určené. Posledný krok je tak isto jednoznačný, keďže je iba jediná trojica pravidiel, ktoré možno použiť. □

Podobný jazyk ako v predchádzajúcej vete vieme vygenerovať aj s dvoma komponentovými gramatikami (v tomto prípade typu B-$edNCE$.).

Veta 4.3.2. $\mathcal{L}(PCGGS_2B) - C\text{-}edNCE \neq \emptyset$

Dôkaz. Podobne ako v predchádzajúcom dôkaze nám stačí nájsť v triede jazykov $\mathcal{L}(PCGGS_2B)$ taký grafový jazyk L, že stringový jazyk $\{a^n \mid n = |V_H|$ pre nejaký graf $[H] \in L\}$ nie je bezkontextový. Definujme jazyk L takto:

$$\begin{aligned} L \;=\; & \{[H] \in GRE_{\{a,b\}\{*\}} \mid H = (V, E, \lambda), \text{ kde} \\ & V = \{1, 2, \ldots, n\}, \\ & E = \{(i, *, i+1) \mid i \in \{1, 2, \ldots, n-1\}\}, \\ & \lambda(i) = a \text{ pre ľubovoľné nepárne } i \in V, \\ & \lambda(i) = b \text{ pre ľubovoľné párne } i \in V, \\ & n = 2^m - 1 \text{ pre nejaké } m \geq 1\} \end{aligned}$$

Tento jazyk vieme generovať nasledujúcou $PCGGS_2B$ gramatikou Θ:

$$\begin{aligned} \Theta &= (\Sigma, \Psi, \Delta, \Gamma, \Omega, G_1, G_2), \text{ kde} \\ \Sigma &= \{S_1, S_2\} \\ \Psi &= \{Q_1, Q_2\} \\ \Delta &= \{a, b\} \\ \Gamma &= \emptyset \\ \Omega &= \{*\} \end{aligned}$$

$$G_i = (\Sigma \cup \Psi \cup \Delta, \Delta, \Gamma \cup \Omega, \Omega, P_i, S_i, C(z)) \text{ pre } i \in \{1, 2\}$$

$$\begin{aligned} P_1 = \; & \{S_1 \to (V = \{1\}, E = \emptyset, \lambda = \{(1, a)\}, \\ & \{(b, */*, 1, in), (b, */*, 1, out)\}), \\ & S_1 \to (V = \{1\}, E = \emptyset, \lambda = \{(1, Q_2)\}, \\ & \{(b, */*, 1, in), (b, */*, 3, out)\})\} \end{aligned}$$

$$\begin{aligned} P_2 = \; & \{S_2 \to (V = \{1, 2, 3\}, E = \{(1, *, 2), (2, *, 3)\}, \\ & \lambda = \{(1, Q_1), (2, b), (3, Q_1)\}, \{(b, */*, 1, in), (b, */*, 3, out)\})\} \end{aligned}$$

$$C(z) = \{(b, */*, z, in), (b, */*, z, out)\} \text{ pre } i \in \{1, 2\}$$

Uvažujme graf $[H] \in L$ s $n = 2^m - 1$ vrcholmi, (takýto graf existuje len jeden), potrebujeme nájsť jeho odvodenie v Θ. Zrejme prvý krok odvodenia (terminálneho grafu) v Θ vyzerá takto:

$$(sn(S_1, z_1, C(z_1)), sn(S_2, z_2, C(z_2))) \Rightarrow (sn(a, z_1, C(z_1)), H_{1_2})$$

kde

$$\begin{aligned} H_{1_2} \;=\; & (V = \{1,2,3\}, E = \{(1,*,2),(2,*,3)\}, \lambda = \{(1,Q_1),(2,b),(3,Q_1)\}, \\ & C = \{(b,*/*,1,in),(b,*/*,3,out)\}) \end{aligned}$$

t.j. H_{1_2} je graf s troma vrcholmi so symbolmi Q_1, b a Q_1 spojenými postupne hranou (pre prvú komponentu n-vetnej formy máme dve pravidlá, ktoré možno použiť, avšak ak by sme použili druhé pravidlo, tak by odvodenie nemohlo ďalej pokračovať, pretože každá komponenta n-vetnej formy by obsahovala komunikačný symbol).

Predpokladajme teraz, že vieme vygenerovať n-vetnú formu (H_1, H_{1_2}), kde

$$\begin{aligned} H_1 &= (V, E, \lambda, C) \\ V &= \{1,2,\ldots,k\}, \\ E &= \{(i,*,i+1) \mid i \in \{1,2,\ldots,k-1\}\}, \\ \lambda(i) &= a \text{ pre ľubovoľné nepárne } i \in V, \\ \lambda(i) &= b \text{ pre ľubovoľné párne } i \in V, \\ k &= 2^j - 1 \text{ pre nejaké } j \geq 1 \\ C &= \{(b,*/*,1,in),(b,*/*,k,out)\} \end{aligned}$$

Potom existuje v Θ odvodenie n-vetnej formy (H_5, H_{1_2}), kde

$$\begin{aligned} H_5 &= (V, E, \lambda, C) \\ V &= \{1,2,\ldots,2k+1\}, \\ E &= \{(i,*,i+1) \mid i \in \{1,2,\ldots,2k\}\}, \\ \lambda(i) &= a \text{ pre ľubovoľné nepárne } i \in V, \\ \lambda(i) &= b \text{ pre ľubovoľné párne } i \in V, \\ k &= 2^j - 1 \text{ pre nejaké } j \geq 1 \\ C &= \{(b,*/*,1,in),(b,*/*,2k+1,out)\} \end{aligned}$$

Spomínané odvodenie vyzerá nasledovne

$$\begin{gathered}(H_1, H_{1_2})\\ \Rightarrow\\ (sn(S_1, z_1, C(z_1)), H_5)\\ \Rightarrow\\ (sn(Q_2, z_1, C(z_1)), H_5)\\ \Rightarrow\\ (H_5, sn(S_2, z_2, C(z_2)))\\ \Rightarrow\\ (H_5, H_{1_2})\end{gathered}$$

Z vyššie uvedených faktov vyplýva, že v gramatike Θ existuje odvodenie n-vetnej formy, kde prvá komponenta je graf H_5 s $2^m - 1$ vrcholmi pre ľubovoľné $m \geq 1$. Graf H_5 je terminálny graf, ktorý patrí do jazyka L.

Potrebujeme ešte ukázať, že všetky grafy ktoré vieme v Θ odvodiť, sú z jazyka L. Zrejme stačí ukázať, že horeuvedené odvodenie je jediné možné (resp. že všetky ostatné varianty použitia pravidiel vedú buď hneď k odvodeniu platného terminálneho grafu alebo k zaseknutiu odvodenia).

Použiť iné pravidlo môžeme len v druhom kroku. Vtedy vygenerujeme v prvej komponente n-vetnej formy terminálny graf $sn(a, z_1, C(z_1))$ a odvodenie už nebude ďalej pokračovať (presnejšie, vzhľadom na to, že obidve komponenty vetnej formy sú terminálne grafy, tak sa v nich už nič nemôže zmeniť). $\square$

Poznámka 4.3.3. *Nasledujúce dve vety dokazujú, že keď sa obmedzíme na PCGGS s centralizovanou komunikačnou štruktúrou, tak nebudeme môcť generovať grafové jazyky, ktorých grafy by boli od seba "príliš vzdialené".*

Definícia 4.3.4. *PCGGS systém Θ nazývame centralizovaný PCGGS alebo centr-PCGGS, ak komunikačná štruktúra systému Θ je graf $G = (V, E)$, kde $V = \{G_1, \ldots, G_n\}$ a množina hrán je definovaná takto: $E = \{(G_1, G_i) \mid i \in \{2, \ldots, n\}\}$ (inak povedané, ak len prvá gramatika obsahuje pravidlo(á), na ktorého pravej strane je komunikačný vrchol).*

Veta 4.3.5. *Nech*

- $\Theta = (\Sigma, \Psi, \Delta, \Gamma, \Omega, G_1, G_2)$ *je centr-PCGGS s dvoma lineárnymi komponentovými gramatikami,*

- *$L(\Theta)$ je grafový jazyk generovaný systémom Θ,*
- *$\{v_i\}$ je rastúca postupnosť prirodzených čísel taká, že $v \in \{v_i\}$ práve vtedy, keď $v = |H|$ pre nejaký graf $H \in L(\Theta)$,*
- *M_i označuje maximum počtu vrcholov grafov na pravých stranách pravidiel z množiny P_i.*

Potom pre dva susedné prvky v_j, v_{j+1} postupnosti $\{v_i\}$ platí

$$v_{j+1} - v_j \leq (M_1 + M_2)(|\Sigma|(|\Sigma|^2 + 1) + 2)$$

Dôkaz. Neformálne povedané, potrebujeme dokázať, že ak zoradíme grafy jazyka $L(G)$ do postupnosti podľa počtu vrcholov, tak rozdiel v počte vrcholov medzi dvoma susednými grafmi nebude väčší ako konštanta určená parametrami gramatiky.

Budeme predpokladať, že P_2 neobsahuje na pravej strane žiadneho pravidla terminálny graf. Ak by Θ túto podmienku porušovalo, môžeme ho upraviť takýmto spôsobom:

- do Σ pridáme nový neterminálny symbol vrchola Y
- do každého terminálneho grafu na pravej strane pravidla p z P_2 (ktorý porušuje uvedenú podmienku) pridáme vrchol so symbolom Y (toto pravidlo budeme označovať ako p')
- do P_2 pridáme pravidlo $Y \to (sn(Y, y), \emptyset)$
- do P_1 pridáme pravidlo $Y \to ((\emptyset, \emptyset, \emptyset), \emptyset)$

Treba si uvedomiť, že terminálne pravidlo z P_2 môže byť v ľubovoľnom odvodení použité len raz. Môžu nastať dve možnosti:

- po použití terminálneho pravidla p už v odvodení nebol žiadny komunikačný krok. V tomto prípade v upravenom systéme môžeme (a zároveň musíme) namiesto p použiť pravidlo p' a následne používať pravidlo $Y \to (sn(Y, y), \emptyset)$ na pokračovanie v odvodení. Očividne vygenerujeme rovnaký graf ako v neupravenom systéme.
- po použití terminálneho pravidla p bol v odvodení (práve jeden) komunikačný krok. Zjavne tento krok bol posledným krokom odvodenia,

keďže prvá komponenta vetnej formy, ktorá bola výsledkom komunikačného kroku, už neobsahuje neterminálny vrchol. V upravenom systéme použijeme rovnako ako v prvom prípade namiesto pravidla p pravidlo p' a následne pravidlo $Y \rightarrow (sn(Y, y), \emptyset)$ na pokračovanie v odvodení. Po komunikačnom kroku môžeme použiť už len pravidlo $Y \rightarrow ((\emptyset, \emptyset, \emptyset), \emptyset)$, čím vygenerujeme rovnaký graf ako v neupravenom systéme.

Z uvedeného predpokladu (a samozrejme z linearity gramatiky) vyplýva, že v ľubovoľnom odvodení pre každú vetnú formu odvodenia, až na poslednú, platí, že obidve jej komponenty obsahujú neterminálny vrchol.

Uvažujme dva susedné prvky v_j, v_{j+1} postupnosti $\{v_i\}$. Ak $v_{j+1} \leq (M_1 + M_2)(|\Sigma|(|\Sigma|^2 + 1) + 1)$, tak niet čo dokazovať (keďže rozdiel $v_{j+1} - v_j$ je určite menší nanajvýš rovný $(M_1 + M_2)(|\Sigma|(|\Sigma|^2 + 1) + 2)$). Nech teda $v_{j+1} > (M_1 + M_2)(|\Sigma|(|\Sigma|^2 + 1) + 2)$ a $H \in L(\Theta)$ je graf s v_{j+1} vrcholmi, uvažujme (ľubovoľné) jeho najkratšie odvodenie (označme ho α). Toto odvodenie musí mať dĺžku aspoň $|\Sigma|(|\Sigma|^2 + 1) + 1$ (a teda obsahuje aspoň $|\Sigma|(|\Sigma|^2 + 1) + 2$ vetných foriem), keďže systém gramatík odvodí v jednom kroku najviac $M_1 + M_2$ terminálnych vrcholov a zároveň obidve komponenty všetkých vetných foriem odvodenia, až na poslednú vetnú formu, obsahujú neterminálne vrcholy. Vo zvyšku dôkazu budeme α uvažovať bez poslednej vetnej formy.

To, čo potrebujeme dokázať je, že v odvodení α môžeme nejakú jeho (dostatočne malú) časť vynechať, čím dostaneme odvodenie grafu s menším počtom vrcholov ako v_{j+1}. Pre odvodenie α platí práve jedna z nasledujúcich dvoch možností:

- v α existuje súvislá postupnosť krokov odvodenia dĺžky aspoň $|\Sigma|^2 + 1$, v ktorej nie je použitý komunikačný krok. Potom z jednoduchej kombinatorickej úvahy vidieť, že v tejto postupnosti existujú dve ekvivalentné konfigurácie (vetné formy) t.j. také, že grafy v prvej resp. v druhej komponente týchto vetných foriem obsahujú vrchol s rovnakým neterminálnym symbolom. Formálnejšie, ukázali sme, že v odvodení α existuje úsek

$$((H_1, C_1), (H_2, C_2)) \Rightarrow^k ((I_1, D_1), (I_2, D_2))$$

kde $k \leq |\Sigma|^2 + 1$ a grafy H_i a I_i obsahujú vrchol s rovnakým neterminálnym symbolom X_i (zároveň I_i obsahuje ako svoj podgraf graf izomorfný s H_i). Uvedený úsek teda môžme z odvodenia vynechať (samozrejme až na prvú vetnú formu) a dostaneme korektné odvodenie α' v Θ. Keďže sme predpokladali, že odvodenie α je najkratšie, tak generovaný graf H' nemôže byť rovný (presnejšie izomorfný) s H. Vynechaním niekoľkých krokov odvodenia sme zmenšili počet vrcholov vygenerovaného grafu, nanajvýš však o $(|\Sigma|^2 + 1)(M_1 + M_2)$.

- v α neexistuje súvislá postupnosť krokov odvodenia dĺžky aspoň $|\Sigma|^2 + 1$, v ktorej nie je použitý komunikačný krok, v odvodení α ale existuje úsek, ktorý
  - má dĺžku $|\Sigma|(|\Sigma|^2 + 1) + 1$ (počet jeho vetných foriem)
  - začína vetnou formou tvaru $((D, C), sn(S_2, z_2, C_2(z_2)))$

  Uvedené podmienky spĺňa napríklad úsek začínajúci prvou vetnou formou v odvodení. V tomto úseku odvodenia museli byť nejaké komunikačné kroky, ich počet je aspoň $|\Sigma|$, čoho dôsledkom je, že uvažovaný úsek obsahuje aspoň $|\Sigma| + 1$ vetných foriem, ktorých druhá komponenta je graf (izomorfný s) $sn(S_2, z_2, C_2(z_2))$. Z Dirichletovho princípu vyplýva, že v tomto úseku znovu existujú dve ekvivalentné konfigurácie, navyše s vlastnosťou, že ich druhé komponenty sú izomorfné grafy. Formálnejšie, v odvodení α existuje úsek

  $$((H, C), sn(S_2, z_2, C_2(z_2))) \Rightarrow^k ((I, D), sn(S_2, z_2', C_2(z_2')))$$

  kde $k \leq |\Sigma|(|\Sigma|^2 + 1)$, grafy H_1 a I_1 obsahujú vrchol s rovnakým neterminálnym symbolom X a I_1 obsahuje ako svoj podgraf graf izomorfný s H_1. Vynechaním uvedeného úseku (okrem prvej vetnej formy) dostaneme korektné odvodenie α' v Θ. Keďže sme predpokladali, že odvodenie α je najkratšie, tak generovaný graf H' nemôže byť rovný (presnejšie izomorfný) s H. Vynechaním niekoľkých krokov odvodenia sme zmenšili počet vrcholov vygenerovaného grafu, nanajvýš však o $(M_1 + M_2)(|\Sigma|(|\Sigma|^2 + 1) + 1)$.

V oboch prípadoch musí mať graf H' menej vrcholov ako H, lebo inak by bol H' izomorfný s H a zároveň $v_{j+1} - |V_{H'}| \leq (M_1 + M_2)(|\Sigma|(|\Sigma|^2 + 1) + 1)$. Označme $|V_{H'}| = v_k$, môžu nastať tri moťnosti:

- $v_k < v_j$, potom dokazovaná nerovnosť triviálne platí
- $v_k = v_j$, dokazovaná nerovnosť platí ako v prvom prípade

- $v_k > v_j$ je v spore s tým, že v_{j+1} a v_j sú susedné prvky

Týmto je veta dokázaná. □

Poznámka 4.3.6. *Uvedený odhad nie je príliš tesný, v skutočnosti nám v odvodení stačí nájsť dve konfigurácie s rovnakými neterminálmi v i-tych komponentách a časť odvodenia medzi nimi vynechať. V nasledujúcej vete zovšeobecníme predchádzajúce tvrdenie pre centr-PCGGS s n lineárnymi komponentami a spresníme odhad pre rozdiel v počte vrcholov medzi "susednými" grafmi.*

Veta 4.3.7. *Nech*

- $\Theta = (\Sigma, \Psi, \Delta, \Gamma, \Omega, G_1, G_2, \ldots, G_n)$ *je centr-PCGGS s n lineárnymi komponentovými gramatikami,*
- $L(\Theta)$ *je grafový jazyk generovaný systémom Θ,*
- $\{v_i\}$ *je rastúca postupnosť prirodzených čísel taká, že $v \in \{v_i\}$ práve vtedy, keď $v = |H|$ pre nejaký graf $H \in L(\Theta)$,*
- M_i *označuje maximum počtu vrcholov grafov na pravých stranách pravidiel z množiny P_i.*

Potom pre dva susedné prvky v_j, v_{j+1} postupnosti $\{v_i\}$ platí

$$v_{j+1} - v_j \leq (\sum_{i=1}^{n} M_i)(2|\Sigma|^n + 2)$$

Dôkaz. Podobne ako v dôkaze predchádzajúceho tvrdenia budeme predpokladať, že P_i neobsahuje na pravej strane žiadneho pravidla terminálny graf pre $i \geq 2$. Ak by Θ túto podmienku porušovalo, môžeme ho upraviť analogickým spôsobom ako v predchádzajúcom dôkaze. Z linearity gramatiky a z uvedeného predpokladu vyplýva, že v ľubovoľnom odvodení pre každú vetnú formu odvodenia, až na poslednú, platí, že všetky jej komponenenty obsahujú neterminálny vrchol.

Nech v_j a v_{j+1} sú dva susedné prvky postupnosti $\{v_i\}$, nech

$$v_{j+1} > (\sum_{i=1}^{n} M_i)(2|\Sigma|^n + 2)$$

(inak niet čo dokazovať), nech $H \in L(\Theta)$ je graf s v_{j+1} vrcholmi, pre ktorý existuje v Θ najkratšie odvodenie (spomedzi grafov s v_{j+1} vrcholmi). Ľubovoľné jeho najkratšie odvodenie musí mať dĺžku aspoň $2|\Sigma|^n + 1$ a obsahuje minimálne $2|\Sigma|^n + 2$ vetných foriem. Uvažujme jeho prefix α dĺžky $2|\Sigma|^n$ s $2|\Sigma|^n + 1$ vetnými formami. Vzhľadom na to, že Θ je *centr-PCGGS*, nemôžu v α za sebou nasledovať dva komunikačné kroky (resp. inak povedané, nemôžu za sebou nasledovať dve vetné formy, ktorých prvá komponenta by obsahovala vrchol s komunikačným symbolom). Keďže prvá vetná forma odvodenia neobsahuje komunikačný symbol, tak α obsahuje aspoň $|\Sigma|^n + 1$ vetných foriem bez komunikačného symbolu, medzi ktorými musia existovať dve, ktorých i-te komponenty obsahujú rovnaké neterminálne symboly. Úsek odvodenia medzi týmito dvoma vetnými formami môžeme vynechať (vlastne tieto dve vetné formy stotožníme) a dostaneme korektné odvodenie grafu H'. H' nemôže mať:

- viacej vrcholov ako H, keďže v lineárnej gramatike nie je možné vymazať terminálny vrchol.
- rovnaký počet vrcholov ako H, keďže sme uvažovali najkratšie odvodenie grafu, ktorý má spomedzi všetkých grafov generovaných Θ s rovnakým počtom vrcholov najkratšie odvodenie.

Teda H' má menej vrcholov ako H, označme $|H'| = v_k$. Z odvodenia sme vynechali najviac $2|\Sigma|^n$ krokov, počas ktorých mohlo byť vygenerovaných najviac

$$(\sum_{i=1}^{n} M_i)(2|\Sigma|^n)$$

terminálnych vrcholov, teda

$$v_{j+1} - v_k \leq (\sum_{i=1}^{n} M_i)(2|\Sigma|^n),$$

čím je veta dokázaná. □

Poznámka 4.3.8. *Podobné tvrdenia (s analogickým dôkazom) možno vysloviť aj pre systémy, ktorých komunikačná štruktúra je strom, dag alebo cyklus*

Kapitola 5

Záver

V tejto práci sme skúmali možnosti využitia paralelizmu v grafových gramatikách. Skúmali sme dva paralelné modely, pôvodne definované pre stringové gramatiky, ktoré sme zadefinovali pre grafové gramatiky. Dokázali sme viaceré vzťahy medzi rôznymi triedami jazykov definovaných týmito modelmi, ktoré ukazujú zmysluplnosť týchto modelov. Niektoré idey, ktoré sme aplikovali pri dôkaze tvrdení v tejto práci, nie sú príliš odlišné od tých, ktoré platia pre skúmanané paralelné modely definované pre stringové gramatiky. Komplikáciou bola hlavne zložitejšia štruktúra grafov oproti reťazcom a v niektorých prípadoch aj konfluentnosť.

Otvorenou otázkou zostáva sformulovanie ďalších vzťahov medzi spomínanými triedami jazykov a najmä zistiť, kde platia nevlastné inklúzie.

Literatúra

[1] E. Csuhaj-Varjú, J. Dassow, J. Kelemen, and Gh. Paun. *Grammar Systems: A Grammatical Approuch to Distribution and Cooperation*. Gordon and Breach Science Publishers Ltd., London, 1994.

[2] R. Meersman and G. Rozenberg. Cooperating grammar systems. *Proceedings Mathematical Foundations of Computer Science '78, Lecture Notes in Computer Science*, 64:364–373, 1978.

[3] Branislav Rovan. *A framework for studying grammars*. Springer Verlag, 1981.

[4] Branislav Rovan. Teória paralelných výpočtov. Elektronické materiály spísané poslucháčmi fakulty FMFI UK na základe prednášok k predmetu Teória paralelných výpočtov, 2000.

[5] Grzegorz Rozenberg. *Handbook of Graph Grammar and Computing by Graph Transformation*. World Scientific Publishing, first edition, 1997.