Highly constrained unification grammars

Daniel Feinstein
Department of Computer Science
University of Haifa
31905 Haifa, Israel
dfeinstein@gmail.com

Shuly Wintner
Department of Computer Science
University of Haifa
31905 Haifa, Israel
shuly@cs.haifa.ac.il

April 30, 2008

Abstract

Unification grammars are widely accepted as an expressive means for describing the structure of natural languages. In general, the recognition problem is undecidable for unification grammars. Even with restricted variants of the formalism, *off-line parsable* grammars, the problem is computationally hard. We present two natural constraints on unification grammars which limit their expressivity and allow for efficient processing. We first show that *non-reentrant* unification grammars generate exactly the class of context-free languages. We then relax the constraint and show that *one-reentrant* unification grammars generate exactly the class of mildly context-sensitive languages. We thus relate the commonly used and linguistically motivated formalism of unification grammars to more restricted, computationally tractable classes of languages.

1 Introduction

Unification grammars (Shieber, 1986; Shieber, 1992; Carpenter, 1992; Wintner, 2006a) have originated as an extension of context-free grammars, the basic idea being to augment the context-free rules with non context-free annotations (feature structures) in order to express additional information. They can describe phonological, morphological, syntactic and semantic properties of languages simultaneously and are thus linguistically suitable for modeling natural languages. Several formulations of unification grammars have been proposed, and they are used extensively by computational linguists to describe the structure of a variety of natural languages.

Unification grammars (UGs) are Turing complete: determining whether a given string is generated by a given grammar is as hard as deciding whether a Turing machine halts on the empty input (Johnson, 1988). Therefore, the recognition problem for unification grammars is undecidable in the general case. In order to ensure its decidability, several constraints on unification grammars, commonly known as the *off-line parsability (OLP) constraints*, were suggested, such that the recognition problem is decidable for off-line parsable grammars (Jaeger, Francez, and Wintner, 2005). The idea behind all the OLP definitions is to rule out grammars which license trees in which an unbounded amount of material is generated without expanding the frontier word. This can happen due to two kinds of rules: ϵ-rules (whose bodies are empty) and unit rules (whose bodies consist of a single element). However, even for unification grammars with no such rules the recognition problem is NP-hard (Barton, Berwick, and Ristad, 1987).

In order for a grammar formalism to make predictions about the structure of natural language its generative capacity must be constrained. It is now generally accepted that Context-free Grammars (CFGs) lack the generative power needed for this purpose (Savitch et al., 1987), due to natural language constructions such as reduplication, multiple agreement and crossed agreement. Several linguistic formalisms have been proposed as capable of modeling these phenomena, including Linear Indexed Grammars (LIG) (Gazdar, 1988), Head Grammars (Pollard, 1984), Tree Adjoining Grammars (TAG) (Joshi, 2003) and Combinatory Categorial Grammars (Steedman, 2000). In a seminal work, Vijay-Shanker and Weir (1994) prove that all four formalisms are weakly equivalent. They all generate the class of *mildly context-sensitive languages*[1] (MCSL), for which recognition algorithms with time complexity $O(n^6)$ are known (Vijay-Shanker and Weir, 1993; Satta, 1994). As a result of the weak equivalence of four independently developed (and linguistically motivated) extensions of CFG, the class MCSL is considered to be linguistically meaningful, a natural class of languages for characterizing natural languages.

The main objective of this work is to define constraints on UGs which naturally limit their generative capacity. We define two natural and easily testable syntactic constraints on UGs which ensure that grammars satisfying them generate the context-free and the mildly context-sensitive languages, respectively. The contribution of this result is twofold:

- From a theoretical point of view, constraining unification grammars to generate exactly the class MCSL results in a grammatical formalism which is, on one hand, powerful enough for linguists to express linguistic generalizations in, and on the other hand as cognitively adequate as the other MCSL formalisms;

- Practically, such a constraint may provide efficient recognition algorithms for the limited class of unification grammars, although we do not explore such possibilities in this paper.

This result is closely related to the result of Keller and Weir (1995), who define a version of unification grammars which is more expressive than MCSL, yet has a polynomial time recognition algorithm. Our work attempts to address precisely the class MCSL, and the definitions we provide are inherently different.

We define some preliminary notions in section 2 and then show a constrained version of unification grammars which generates the class CFL of context-free languages in section 3. Section 4 presents the main result, namely a restricted version of unification grammars and a mapping of its grammars to LIG, establishing the proposition that such grammars generate exactly the class MCSL. We conclude with suggestions for future research. To facilitate readability, most of the proofs were moved to the appendices.

2 Preliminary notions

2.1 Linear indexed grammars

A CFG is a four-tuple $G^{cf} = \langle V_N, V_t, \mathcal{R}^{cf}, S \rangle$ where V_t is a set of *terminal*s, V_N is a set of *non-terminal*s, including the *start symbol* S, and $\mathcal{R}^{cf}$ is a set of productions, assumed to be in a normal form where each rule has either (zero or more) non-terminals or a single terminal in its body, and where the start symbol never occurs in the right hand side of rules. The set of all such context-free grammars is denoted CFGS.

In a linear indexed grammar (LIG),[2] strings are derived from nonterminals with an associated stack denoted $A[l_1 \ldots l_n]$, where A is a nonterminal, each l_i is a stack symbol, and l_1 is the top of the stack. Since stacks can grow to be of unbounded size during a derivation, some way of partially specifying unbounded

[1]The term *mildly context-sensitive* was coined by Joshi, Levy, and Takahashi (1975), who used it to refer to a more informal class of languages. In this paper we use this term to denote the class of tree adjoining languages, a trend which has become common in the literature.

[2]The definition is based on Vijay-Shanker and Weir (1994).

stacks in LIG productions is needed. We use $A[l_1 \ldots l_n \infty]$ to denote a nonterminal A associated with any stack η whose top n symbols are $l_1, l_2 \ldots, l_n$. The set of all nonterminals in V_N, associated with stacks whose symbols come from V_s, is denoted $V_N[V_s^*]$.

Definition 1 (LIG). *A **Linear Indexed Grammar** is a five tuple $G^{li} = \langle V_N, V_t, V_s, \mathcal{R}^{li}, S\rangle$ where V_t, V_N and S are as above, V_s is a finite set of indices (stack symbols) and $\mathcal{R}^{li}$ is a finite set of productions in one of the following two forms:*

- ***fixed** stack: $N[p_1 \ldots p_n] \rightarrow \alpha$*
- ***unbounded** stack: $N[p_1 \ldots p_n \infty] \rightarrow \alpha$ or $N[p_1 \ldots p_n \infty] \rightarrow \alpha N'[q_1 \ldots q_m \infty]\beta$*

where $N, N' \in V_N$, $p_1 \ldots p_n, q_1 \ldots q_m \in V_s$, $n, m \geq 0$ and $\alpha, \beta \in (V_t \cup V_N[V_s^*])^*$.

Example 1 (LIG). *$G_1^{li} = \langle V_N, V_t, V_s, \mathcal{R}^{li}, S\rangle$ is a LIG, where:*

- $V_N = \{S, N_2, N_3\}$
- $V_t = \{a, b\}$
- $V_s = V_t$
- $\mathcal{R}^{li} = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7\}$, *where*
  1. $r_1 = S[\,] \rightarrow N_2[\,]$
  2. $r_2 = N_2[\,\infty] \rightarrow N_2[a\,\infty]a$
  3. $r_3 = N_2[\,\infty] \rightarrow N_2[b\,\infty]b$
  4. $r_4 = N_2[\,\infty] \rightarrow N_3[\,\infty]$
  5. $r_5 = N_3[a\,\infty] \rightarrow aN_3[\,\infty]$
  6. $r_6 = N_3[b\,\infty] \rightarrow bN_3[\,\infty]$
  7. $r_7 = N_3[\,] \rightarrow \epsilon$

A crucial characteristic of LIG is that only *one* copy of the stack can be copied to a *single* element in the body of a rule. If more than one copy were allowed, the expressive power would grow beyond MCSL.

Definition 2 (LIG derivation). *Given a LIG $\langle V_N, V_t, V_s, \mathcal{R}^{li}, S\rangle$, the **derivation relation** '$\Rightarrow_{li}$' is defined as follows: for all $\Psi_1, \Psi_2 \in (V_N[V_s^*] \cup V_t)^*$ and $\eta \in V_s^*$,*

- *If $N_i[p_1 \ldots p_n] \rightarrow \alpha \in \mathcal{R}^{li}$ then*
$$\Psi_1 N_i[p_1 \ldots p_n]\Psi_2 \Rightarrow_{li} \Psi_1 \alpha \Psi_2$$
- *If $N_i[p_1 \ldots p_n \infty] \rightarrow \alpha \in \mathcal{R}^{li}$ then*
$$\Psi_1 N_i[p_1 \ldots p_n \eta]\Psi_2 \Rightarrow_{li} \Psi_1 \alpha \Psi_2$$
- *If $N_i[p_1 \ldots p_n \infty] \rightarrow \alpha N_j[q_1 \ldots q_m \infty]\beta \in \mathcal{R}^{li}$ then*
$$\Psi_1 N_i[p_1 \ldots p_n \eta]\Psi_2 \Rightarrow_{li} \Psi_1 \alpha N_j[q_1 \ldots q_m \eta]\beta\Psi_2$$

*The **language** generated by G^{li} is $L(G^{li}) = \{w \in V_t^* \mid S[\,] \overset{*}{\Rightarrow}_{li} w\}$, where '$\overset{*}{\Rightarrow}_{li}$' is the reflexive, transitive closure of '$\Rightarrow_{li}$'.*

Example 2 (LIG). *For G_1^{li} of example 1, $L(G_1^{li}) = \{ww \mid w \in \{a, b\}\}$.*

2.2 Unification grammars

We assume familiarity with theories of feature structures as formulated, e.g., by Wintner (2006a) or Wintner (2006b). We summarize below the concepts that are needed for the rest of this paper in order to set up notation.

Definition 3 (Signature)**.** *A **signature** is a structure $\mathcal{S} = \langle \textsc{Atoms}, \textsc{Feats}, \textsc{Tags} \rangle$, where* ATOMS *is a finite set of atoms,* FEATS *is a finite set of features and* TAGS *is an enumerable set of variables.*

Unless explicitly mentioned, the set TAGS of variables is assumed to be $\{\boxed{1}, \boxed{2}, \ldots\}$.

Definition 4 (Feature structures)**.** *Given a signature $\mathcal{S}$, the set $\mathcal{FS}(\mathcal{S})$ of **feature structures (FSs)** is the least set satisfying the following two clauses:*

1. $\mathsf{A} = Xa \in \mathcal{FS}(\mathcal{S})$ *for any $a \in$ ATOMS and $X \in$ TAGS; A is said to be **atomic** and X is the **tag** of A.*
2. $\mathsf{A} = X[f_1 : \mathsf{A}_1, \ldots, f_n : \mathsf{A}_n] \in \mathcal{FS}(\mathcal{S})$ *for $n \geq 0$, $X \in$ TAGS, $f_1, \ldots, f_n \in$ FEATS and $\mathsf{A}_1, \ldots, \mathsf{A}_n \in \mathcal{FS}(\mathcal{S})$, where $f_i \neq f_j$ if $i \neq j$. A is said to be **complex**, and X is the **tag** of A. If $n = 0$, $\mathsf{A} = X[]$ is an **empty FS**.*

Meta-variables $\mathsf{A}, \mathsf{B}, \mathsf{C}$, with or without subscripts, range over $\mathcal{FS}$.

Example 3 (Feature structures)**.** *Consider a signature consisting of* ATOMS $= \{a\}$, FEATS $= \{\text{F,G}\}$. *Then $\mathsf{A}_1 = \boxed{4}\ a$ is an FS by the first clause of the definition, $\mathsf{A}_2 = \boxed{2}\ [\,]$ is an empty FS by the second clause, $\mathsf{A}_3 = \boxed{3}\ [\text{F} : \boxed{4}\ a]$ is an FS by the second clause, and A_4 is a FS by the second clause:*

$$\mathsf{A}_4 = \boxed{1} \begin{bmatrix} \text{G} : \boxed{3} \begin{bmatrix} \text{F} : \boxed{4}\ a \end{bmatrix} \\ \text{F} : \boxed{2}\ [\,] \end{bmatrix}$$

Tags that occur only once in a FS can be omitted, so A_4 above can be written thus:

$$\begin{bmatrix} \text{G} : \begin{bmatrix} \text{F} : a \end{bmatrix} \\ \text{F} : [\,] \end{bmatrix}$$

Definition 5 (Paths)**.** *A **path** (over* FEATS*) is a finite sequence of features, and the set* PATHS $=$ FEATS* *is the collection of all paths.*

Meta-variables π, μ (with or without subscripts) range over paths. ϵ is the empty path, denoted also by '$\langle\rangle$'. Path concatenation is denoted using either '$\cdot$' or juxtaposition. A path is a purely syntactic notion: every sequence of features constitutes a path. Usually, interesting paths are those that can be interpreted as actual paths in some FS, starting from the outermost level; we use Π_{A} to denote the paths of a FS A. The value of the path π in A, which is a sub-structure of A, is denoted *pval*(A, π).

When two different paths in some FS A have the same value we say that they are *reentrant*.

Definition 6 (Reentrancy)**.** *Two paths π_1 and π_2 are **reentrant** in a FS A if pval$(\mathsf{A}, \pi_1) =$ pval(A, π_2), denoted also $\pi_1 \overset{\mathsf{A}}{\leftrightsquigarrow} \pi_2$. A FS A is reentrant if there exist two paths $\pi_1, \pi_2 \in \Pi_{\mathsf{A}}$ such that $\pi_1 \neq \pi_2$ and $\pi_1 \overset{\mathsf{A}}{\leftrightsquigarrow} \pi_2$.*

Definition 7 (FS subsumption)**.** *Let A_1, A_2 be FSs over the same signature. A_1 **subsumes** A_2, denoted $\mathsf{A}_1 \sqsubseteq \mathsf{A}_2$, if the following conditions hold:*

1. $\Pi_{\mathsf{A}_1} \subseteq \Pi_{\mathsf{A}_2}$; *furthermore, if pval*$(\mathsf{A}_1, \pi)$ *is an atomic FS then pval*(A_2, π) *is an atomic FS with the same atom;*

2. *if* $\pi_1 \overset{\mathsf{A}_1}{\leftrightsquigarrow} \pi_2$ *then* $\pi_1 \overset{\mathsf{A}_2}{\leftrightsquigarrow} \pi_2$.

Definition 8 (Unification)**.** *The **unification** of two FSs* A_1 *and* A_2*, denoted* $\mathsf{A}_1 \sqcup \mathsf{A}_2$*, is the least upper bound of* A_1 *and* A_2 *with respect to subsumption. If no upper bound exists, the unification **fails**, sometimes denoted* $\mathsf{A}_1 \sqcup \mathsf{A}_2 = \top$.

FSs can encode lists in a natural way, using a *head*|*tail* notation (dubbed HD|TL in the sequel).

Example 4 (Encoding lists as FSs)**.** *The list of three elements,* $\langle a, b, c\rangle$*, can be encoded as the following FS, which is over a signature including the features* HD, TL *and the atoms* $a, b, c, elist$*:*

$$\begin{bmatrix} \text{HD} : a \\ \text{TL} : \begin{bmatrix} \text{HD} : b \\ \text{TL} : \begin{bmatrix} \text{HD} : c \\ \text{TL} : elist \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

For the sake of brevity, we use standard list notation when FSs encode lists, with double angular brackets. The FS of example 4 is thus depicted as $\langle\langle a, b, c\rangle\rangle$. We also provide means for encoding *open-ended* lists, namely lists which do not terminate with *elist* (and can therefore be extended). We use the notation $\langle\langle a, b, c | \boxed{\text{i}} \rangle\rangle$ for the FS

$$\begin{bmatrix} \text{HD} : a \\ \text{TL} : \begin{bmatrix} \text{HD} : b \\ \text{TL} : \begin{bmatrix} \text{HD} : c \\ \text{TL} : \boxed{\text{i}} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

We now extend feature structures to *multi-rooted structures*; these are basically sequences of FSs, in which the scope of variables is extended to the entire sequence, enabling paths to be reentrant even if they leave different elements of the sequence.

Definition 9 (MRS)**.** *Given a signature* $\mathcal{S}$*, a **multi-rooted structure (MRS)** of **length*** $n \geq 0$ *is a sequence* $\langle \mathsf{A}_1, \ldots, \mathsf{A}_n\rangle$ *such that for each* i*,* $1 \leq i \leq n$*,* A_i *is a FS over the signature.*

Meta-variables σ, ρ range over MRSs. The length of σ is denoted $len(\sigma)$. We usually do not distinguish between a MRS of length 1 and a FS.

Example 5 (MRSs)**.** *Following is an MRS of length 3:*

$$\boxed{2}\ \begin{bmatrix}\text{F} : \boxed{9}\ \begin{bmatrix}\text{H} : \boxed{1}\ [\,]\end{bmatrix}\end{bmatrix} \qquad \boxed{1}\ \begin{bmatrix}\text{F} : \boxed{8}\ \begin{bmatrix}\text{G} : \boxed{7}\ a \\ \text{H} : \boxed{2}\ [\,]\end{bmatrix}\end{bmatrix} \qquad \boxed{6}\ \begin{bmatrix}\text{F} : \boxed{5}\ \begin{bmatrix}\text{H} : \boxed{2}\ [\,]\end{bmatrix}\end{bmatrix}$$

Note that the same variable can tag different sub-FSs of different elements in the sequence (e.g., $\boxed{1}$ or $\boxed{2}$ in example 5). In other words, the *scope* of variables is extended from single FSs to MRSs.

The definition of *paths* and *path values* is naturally extended from FSs to MRSs by adding a parameter denoting the index of the element in the sequence from which the path leaves. An MRS is *reentrant* if it has two distinct paths which share the same value; these two paths may well be "rooted" in two different elements of the MRS. Since MRSs are sequences, they can be concatenated; we use juxtaposition to denote MRS concatenation. If σ is an MRS, we use σ^i to denote the i-th element of σ.

Definition 10 (MRS subsumption)**.** *Let σ, ρ be two MRSs of the same length n and over the same signature. σ **subsumes** ρ, denoted $\sigma \sqsubseteq \rho$, if the following conditions hold:*

1. *for all i, $1 \leq i \leq n$, $\sigma^i \sqsubseteq \rho^i$;*
2. *if $\langle i, \pi_1 \rangle \overset{\sigma}{\leftrightsquigarrow} \langle j, \pi_2 \rangle$ then $\langle i, \pi_1 \rangle \overset{\rho}{\leftrightsquigarrow} \langle j, \pi_2 \rangle$.*

Example 6 (MRS subsumption)**.** *Let σ and σ' be the following two MRSs (of length 3):*

$$\sigma : \quad \boxed{1}\begin{bmatrix}\text{CAT} : np \\ \text{AGR} : \boxed{4}\end{bmatrix} \quad \boxed{2}\begin{bmatrix}\text{CAT} : vp \\ \text{AGR} : \boxed{4}\begin{bmatrix}\text{NUM} : sg \\ \text{PERS} : 3rd\end{bmatrix}\end{bmatrix} \quad \boxed{3}\begin{bmatrix}\text{CAT} : np \\ \text{AGR} : \boxed{6}\begin{bmatrix}\text{NUM} : sg \\ \text{PERS} : 3rd\end{bmatrix}\end{bmatrix}$$

$$\sigma' : \quad \boxed{1}\begin{bmatrix}\text{CAT} : np \\ \text{AGR} : \boxed{4}\end{bmatrix} \quad \boxed{2}\begin{bmatrix}\text{CAT} : vp \\ \text{AGR} : \boxed{4}\begin{bmatrix}\text{NUM} : sg \\ \text{PERS} : 3rd\end{bmatrix}\end{bmatrix} \quad \boxed{3}\begin{bmatrix}\text{CAT} : np \\ \text{AGR} : \boxed{4}\end{bmatrix}$$

Then $\sigma \sqsubseteq \sigma'$ but not $\sigma' \sqsubseteq \sigma$.

When MRSs are concerned, two variants of unification are defined, one which unifies two same-length structures and produces their least upper bound with respect to subsumption, and one, called *unification in context*, which combines the information in two feature structures, each of which may be an element in a larger structure.

Definition 11 (MRS unification)**.** *Let σ, ρ be MRSs of the same length, n. The **unification** of σ and ρ, denoted $\sigma \sqcup \rho$, is the least upper bound of σ and ρ with respect to MRS subsumption, if it exists.*

Definition 12 (Unification in context)**.** *Let σ, ρ be two MRSs and i, j be indexes such that $i \leq len(\sigma)$ and $j \leq len(\rho)$. Then $\langle \sigma', \rho' \rangle = (\sigma, i) \sqcup (\rho, j)$ iff $\sigma' = \min_{\sqsubseteq}\{\sigma'' \mid \sigma \sqsubseteq \sigma''$ and $\rho^j \sqsubseteq \sigma''^i\}$ and $\rho' = \min_{\sqsubseteq}\{\rho'' \mid \rho \sqsubseteq \rho''$ and $\sigma^i \sqsubseteq \rho''^j\}$.*

Lemma 1. *If $\langle \sigma', \rho' \rangle = (\sigma, i) \sqcup (\rho, j)$ then $\sigma'^i = \rho'^j = \sigma^i \sqcup \rho^j$.*

Example 7 (Unification in context)**.** *Let*

$$\sigma = \begin{bmatrix}\text{LIST} : \boxed{3}\ [\,]\end{bmatrix} \quad \begin{bmatrix}\text{LIST} : \boxed{3}\end{bmatrix}$$

$$\rho = \begin{bmatrix}\text{LIST} : \begin{bmatrix}\text{HD} : \boxed{1} \\ \text{TL} : \boxed{2}\end{bmatrix}\end{bmatrix} \quad \begin{bmatrix}\text{LIST} : \boxed{2}\end{bmatrix} \quad \begin{bmatrix}\text{LIST} : \begin{bmatrix}\text{HD} : \boxed{1} \\ \text{TL} : elist\end{bmatrix}\end{bmatrix}$$

The unification in context of the second element of σ with the first element of ρ is $(\sigma, 2) \sqcup (\rho, 1) = \langle \sigma', \rho' \rangle$, where:

$$\sigma' = \begin{bmatrix}\text{LIST} : \boxed{3}\end{bmatrix} \quad \begin{bmatrix}\text{LIST} : \boxed{3}\begin{bmatrix}\text{HD} : \boxed{1} \\ \text{TL} : \boxed{2}\end{bmatrix}\end{bmatrix}$$

$$\rho' = \begin{bmatrix}\text{LIST} : \begin{bmatrix}\text{HD} : \boxed{1} \\ \text{TL} : \boxed{2}\end{bmatrix}\end{bmatrix} \quad \begin{bmatrix}\text{LIST} : \boxed{2}\end{bmatrix} \quad \begin{bmatrix}\text{LIST} : \begin{bmatrix}\text{HD} : \boxed{1} \\ \text{TL} : elist\end{bmatrix}\end{bmatrix}$$

Definition 13 (Unification grammars)**.** *A **Unification grammar** over a signature $\mathcal{S}$ and a finite set* WORDS *of words is a tuple $G^u = \langle \mathcal{R}^u, \mathcal{L}, \mathsf{A}^s \rangle$ where:*

- $\mathcal{R}^u$ *is a finite set of rules, each of which is an MRS of length* $n \geq 1$*, with a designated first element, the* ***head*** *of the rule, followed by its* ***body****. The head and body are separated by an arrow (*$\rightarrow$*).*
- $\mathcal{L}$ *is a* ***lexicon****, which associates with every word* $w \in$ WORDS *a finite set of feature structures,* $\mathcal{L}(w)$.
- A^s *is a feature structure, the* ***start symbol****.*

We use meta-variables G^u (with or without subscripts) to denote unification grammars.

Example 8 (Unification grammar)**.** *Let* G^u_{ww} *be the unification grammar over the signature* $\langle$ATOMS, FEATS, TAGS, WORDS$\rangle$*, where* FEATS $= \{$LIST, HD, TL$\}$, ATOMS $= \{s, elist, ta, tb\}$ *and* WORDS $= \{a, b\}$*, defined as:*

$$\mathsf{A}^s = \left[\text{LIST}: \begin{bmatrix}\text{HD}: s\\ \text{TL}: elist\end{bmatrix}\right]$$

$$\mathcal{R}^u = \left\{ \begin{array}{l} \left[\text{LIST}: \begin{bmatrix}\text{HD}: s\\ \text{TL}: elist\end{bmatrix}\right] \rightarrow \left[\text{LIST}: \boxed{3}\right] \quad \left[\text{LIST}: \boxed{3}\right] \\ \left[\text{LIST}: \begin{bmatrix}\text{HD}: \boxed{1}\\ \text{TL}: \boxed{2}\end{bmatrix}\right] \rightarrow \left[\text{LIST}: \boxed{2}\right] \quad \left[\text{LIST}: \begin{bmatrix}\text{HD}: \boxed{1}\\ \text{TL}: elist\end{bmatrix}\right] \end{array} \right\}$$

$$\mathcal{L}(a) = \left\{ \left[\text{LIST}: \begin{bmatrix}\text{HD}: ta\\ \text{TL}: elist\end{bmatrix}\right] \right\} \qquad \mathcal{L}(b) = \left\{ \left[\text{LIST}: \begin{bmatrix}\text{HD}: tb\\ \text{TL}: elist\end{bmatrix}\right] \right\}$$

To define the *language* generated by a unification grammar G^u, we define *forms* as MRSs. A form $\sigma_A = \langle \mathsf{A}_1, \ldots, \mathsf{A}_k \rangle$ *immediately derives* another form $\sigma_B = \langle \mathsf{B}_1, \ldots, \mathsf{B}_m \rangle$ (denoted by $\sigma_A \overset{1}{\Rightarrow}_u \sigma_B$) iff there exists a rule $\rho^u \in \mathcal{R}^u$ of length n that licenses the derivation. The head of the rule is matched against some element A_i in σ_A using unification in context: $(\sigma_A, i) \sqcup (\rho^u, 0) = (\sigma'_A, \rho')$. If the unification does not fail, σ_B is obtained by replacing the i-th element of σ'_A with the body of ρ'. The reflexive transitive closure of '$\overset{1}{\Rightarrow}_u$' is denoted by '$\overset{*}{\Rightarrow}_u$'. An empty derivation sequence means that an empty sequence of rules is applied to the source MRS and is denoted by '$\overset{0}{\Rightarrow}_u$', for example $\sigma_A \overset{0}{\Rightarrow}_u \sigma_A$. A form is *sentential* if it is derivable from the start symbol of the grammar.

Definition 14. *The* ***language*** *of a unification grammar* G^u *is* $L(G^u) = \{w_1 \cdots w_n \in \text{WORDS}^* \mid \mathsf{A}^s \overset{*}{\Rightarrow}_u \sigma_l$ *and* σ_l *is unifiable with* $\langle \mathsf{A}_1, \ldots, \mathsf{A}_n \rangle\}$*, where* $\mathsf{A}_i \in \mathcal{L}(w_i)$ *for* $1 \leq i \leq n$.

Example 9 (Derivation sequence)**.** *As an example, consider again the grammar* G^u_{ww} *of example 8. The following is a derivation sequence for the string '*$baba$*' with this grammar. Note that the scope of variables is limited to a single MRS (so that multiple occurrences of the same tag in a single form denote reentrancy, whereas across forms they are unrelated).*

$$\mathsf{A}^s = \left[\text{LIST}: \begin{bmatrix}\text{HD}: s\\ \text{TL}: elist\end{bmatrix}\right] \qquad \textit{apply rule 1 to the single element of the form}$$

$$\sigma_1 = \left[\text{LIST}: \boxed{3}\right] \left[\text{LIST}: \boxed{3}\right] \qquad \textit{apply rule 2 to the second element}$$

$$\sigma_2 = \left[\text{LIST}: \boxed{3} \begin{bmatrix}\text{HD}: \boxed{1}\\ \text{TL}: \boxed{2}\end{bmatrix}\right] \left[\text{LIST}: \boxed{2}\right] \left[\text{LIST}: \begin{bmatrix}\text{HD}: \boxed{1}\\ \text{TL}: elist\end{bmatrix}\right] \textit{apply rule 2 to the first element}$$

$$\sigma_3 = \left[\text{LIST}: \boxed{2}\right] \left[\text{LIST}: \begin{bmatrix}\text{HD}: \boxed{1}\\ \text{TL}: elist\end{bmatrix}\right] \left[\text{LIST}: \boxed{2}\right] \left[\text{LIST}: \begin{bmatrix}\text{HD}: \boxed{1}\\ \text{TL}: elist\end{bmatrix}\right]$$

Now consider the MRS obtained by concatenating (the single elements of) $\langle \mathcal{L}(b), \mathcal{L}(a), \mathcal{L}(b), \mathcal{L}(a)\rangle$:

$$\sigma_l = \begin{bmatrix}\text{LIST}: \begin{bmatrix}\text{HD}: tb\\ \text{TL}: elist\end{bmatrix}\end{bmatrix} \begin{bmatrix}\text{LIST}: \begin{bmatrix}\text{HD}: ta\\ \text{TL}: elist\end{bmatrix}\end{bmatrix} \begin{bmatrix}\text{LIST}: \begin{bmatrix}\text{HD}: tb\\ \text{TL}: elist\end{bmatrix}\end{bmatrix} \begin{bmatrix}\text{LIST}: \begin{bmatrix}\text{HD}: ta\\ \text{TL}: elist\end{bmatrix}\end{bmatrix}$$

Since σ_l and σ_3 are unifiable, the string 'baba' is in $L(G^u_{ww})$. In fact, $L(G^u_{ww}) = \{ww \mid w \in \{a, b\}^+\}$.

In order to limit the generative capacity of unification grammars we define two constrained versions of the formalism below. Both limit the number of reentrancies which are allowed between the head of each grammar rule and its body. Informally, a rule is non-reentrant if no reentrancy tags occur in it. A rule is one-reentrant if at most one reentrancy tag occurs in it, exactly twice: once in the head of the rule and once in an element of its body.

Definition 15 (Non-/One-reentrant unification grammars)**.** *A unification grammar $\langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L}\rangle$ over the signature $\langle \text{ATOMS}, \text{FEATS}, \text{TAGS}, \text{WORDS}\rangle$ is **non-reentrant** iff for any rule $\rho^u \in \mathcal{R}^u$, ρ^u is non-reentrant. It is **one-reentrant** iff for every rule $\rho^u \in \mathcal{R}^u$, ρ^u includes at most one reentrancy, between the head of the rule and some element of the body. Formally, it can have at most one (non-trivial, i.e., non-identity) reentrancy $\langle 1, \pi_1\rangle \overset{\rho^u}{\leftrightsquigarrow} \langle i, \pi_2\rangle$, where $i > 1$. Let UG_{nr}, UG_{1r} be the sets of all non-reentrant and one-reentrant unification grammars, respectively.*

One-reentrant unification grammars induce highly constrained (sentential) forms: in such forms, there are no reentrancies whatsoever, neither between distinct elements nor within a single element. The following lemma can be proven by a simple induction on the length of a derivation sequence; it follows directly from the fact that rules in a one-reentrant unification grammar have no reentrancies between elements of their bodies.

Lemma 2. *If τ is a sentential form induced by a one-reentrant grammar then there are no reentrancies between elements of τ or within an element of τ.*

Since all the feature structures in forms induced by one-reentrant unification grammars are non-reentrant, unification is simplified.

Lemma 3. *Let A and B be unifiable non-reentrant feature structures. Then $\mathsf{C} = \mathsf{A} \sqcup \mathsf{B}$ is also a non-reentrant feature structure, and $\Pi_C = \Pi_A \cup \Pi_B$.*

To simplify some of the constructions, we define a simplified variant of one-reentrant unification grammars, which is equivalent to the original definition. In the sequel we assume that all one-reentrant unification grammars are *simplified*.

Definition 16 (Simplified one-reentrant unification grammars)**.** *A one-reentrant unification grammar $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L}\rangle$ over the signature $\sigma = \langle \text{ATOMS}, \text{FEATS}, \text{TAGS}, \text{WORDS}\rangle$ is **simplified** iff the lexical categories of words are inconsistent with any feature structure (except themselves). Formally, if τ is a sentential form induced by G^u and τ^i is an element of τ then for each word $a \in \text{WORDS}$, $\mathcal{L}(a) = \{\mathsf{A}\}$, where $\mathsf{A} \sqcup \tau^i \neq \top$ iff $\mathsf{A} = \tau^i$.*

3 Context-free unification grammars

In this section we define a constraint on unification grammars which ensures that grammars satisfying it generate the class CFL. The constraint disallows *any* reentrancies in the rules of the grammar. When rules are non-reentrant, applying a rule implies that an exact copy of the body of the rule is inserted into the generated (sentential) form, not affecting neighboring elements of the form the rule is applied to. The only difference between rule application in UG_{nr} and the analog operation in CFGS is that the former requires unification whereas the latter only calls for identity check. This small difference does not affect the generative power of the formalism, since unification can be pre-compiled in this simple case.

The trivial direction is to map a CFG to a non-reentrant unification grammar, since every CFG is, trivially, such a grammar (where terminal and non-terminal symbols are viewed as atomic feature structures). For the reverse direction, we define a mapping from UG_{nr} to CFGS. The non-terminals of the CFG in the image of the mapping are the set of all feature structures defined in the source unification grammar.

Definition 17. *Let ug2cfg* : $\mathrm{UG}_{nr} \mapsto$ CFGS *be a* ***mapping of*** UG_{nr} ***to*** CFGS, *such that if* $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L} \rangle$ *is over the signature* $\langle \text{ATOMS}, \text{FEATS}, \text{TAGS}, \text{WORDS} \rangle$ *then ug2cfg*$(G^u) = \langle V_N, V_t, \mathcal{R}^{cf}, S^{cf} \rangle$, *where:*

- $V_N = \{\mathsf{A}_i \mid \mathsf{A}_0 \to \mathsf{A}_1 \ldots \mathsf{A}_n \in \mathcal{R}^u, i \geq 0\} \cup \{\mathsf{A} \mid \mathsf{A} \in \mathcal{L}(a), a \in \text{ATOMS}\} \cup \{\mathsf{A}^s\}$. *$V_N$ is the set of all the feature structures occurring in any of the rules or the lexicon of G^u.*
- $S^{cf} = \mathsf{A}^s$
- $V_t = \text{WORDS}$
- *$\mathcal{R}^{cf}$ consists of the following rules:*
  1. *Let* $\mathsf{A}_0 \to \mathsf{A}_1 \ldots \mathsf{A}_n \in \mathcal{R}^u$ *and* $\mathsf{B} \in \mathcal{L}(b)$. *If for some* i, $1 \leq i \leq n$, $\mathsf{A}_i \sqcup \mathsf{B} \neq \top$, *then* $\mathsf{A}_i \to b \in \mathcal{R}^{cf}$
  2. *If* $\mathsf{A}_0 \to \mathsf{A}_1 \ldots \mathsf{A}_n \in \mathcal{R}^u$ *and* $\mathsf{A}^s \sqcup \mathsf{A}_0 \neq \top$ *then* $S^{cf} \to \mathsf{A}_1 \ldots \mathsf{A}_n \in \mathcal{R}^{cf}$.
  3. *Let* $\rho_1^u = \mathsf{A}_0 \to \mathsf{A}_1 \ldots \mathsf{A}_n$ *and* $\rho_2^u = \mathsf{B}_0 \to \mathsf{B}_1 \ldots \mathsf{B}_m$, *where* $\rho_1^u, \rho_2^u \in \mathcal{R}^u$. *If for some* i, $1 \leq i \leq n$, $\mathsf{A}_i \sqcup \mathsf{B}_0 \neq \top$, *then the rule* $\mathsf{A}_i \to \mathsf{B}_1 \ldots \mathsf{B}_m \in \mathcal{R}^{cf}$

Since $\mathcal{R}^u$ and $\mathcal{L}$ are finite, so is V_N. V_t is finite because WORDS is, and $\mathcal{R}^{cf}$ is finite because $\mathcal{R}^u$ and $\mathcal{L}$ are. The size of *ug2cfg*(G^u) is polynomial in the size of G^u.

Example 10 (Mapping from UG_{nr} to CFGS)**.** *Let* $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L} \rangle$ *be a non-reentrant unification grammar for the language* $\{a^n b^n \mid 0 \leq n\}$ *over the signature* $\langle \text{ATOMS}, \text{FEATS}, \text{TAGS}, \text{WORDS} \rangle$, *such that:*

- $\text{ATOMS} = \{v, u, w\}$
- $\text{FEATS} = \{\text{F}_1, \text{F}_2\}$
- $\text{WORDS} = \{a, b\}$
- $\mathsf{A}^s = \begin{bmatrix} \text{F}_1 : w \\ \text{F}_2 : w \end{bmatrix}$
- $\mathcal{L}(a) = \{ \begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : v \end{bmatrix} \}$ *and* $\mathcal{L}(b) = \{ \begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : u \end{bmatrix} \}$
- *The set of rules $\mathcal{R}^u$ is defined as:*
  1. $\begin{bmatrix} \text{F}_1 : w \\ \text{F}_2 : w \end{bmatrix} \to \varepsilon$

2. $\begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : w \end{bmatrix} \rightarrow \begin{bmatrix} \text{F}_1 : u \\ \text{F}_2 : v \end{bmatrix} \begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : w \end{bmatrix} \begin{bmatrix} \text{F}_1 : v \\ \text{F}_2 : u \end{bmatrix}$

Then the context-free grammar $G^{cf} = \langle V_N, V_t, \mathcal{R}^{cf}, S^{cf} \rangle = ug2cfg(G^u)$ *is:*

- $V_N = \left\{ \begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : v \end{bmatrix}, \begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : u \end{bmatrix}, \begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : w \end{bmatrix}, \begin{bmatrix} \text{F}_1 : w \\ \text{F}_2 : w \end{bmatrix}, \begin{bmatrix} \text{F}_1 : u \\ \text{F}_2 : v \end{bmatrix}, \begin{bmatrix} \text{F}_1 : v \\ \text{F}_2 : u \end{bmatrix} \right\}$
- $V_t = \text{WORDS} = \{a, b\}$
- $S^{cf} = \mathsf{A}^s = \begin{bmatrix} \text{F}_1 : w \\ \text{F}_2 : w \end{bmatrix}$
- *The set of rules* $\mathcal{R}^{cf}$ *is defined as:*

  1. $\begin{bmatrix} \text{F}_1 : u \\ \text{F}_2 : v \end{bmatrix} \rightarrow a$

  2. $\begin{bmatrix} \text{F}_1 : v \\ \text{F}_2 : u \end{bmatrix} \rightarrow b$

  3. $\begin{bmatrix} \text{F}_1 : w \\ \text{F}_2 : w \end{bmatrix} \rightarrow \epsilon$

  4. $\begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : w \end{bmatrix} \rightarrow \epsilon$

  5. $\begin{bmatrix} \text{F}_1 : w \\ \text{F}_2 : w \end{bmatrix} \rightarrow \begin{bmatrix} \text{F}_1 : u \\ \text{F}_2 : v \end{bmatrix} \begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : w \end{bmatrix} \begin{bmatrix} \text{F}_1 : v \\ \text{F}_2 : u \end{bmatrix}$

  6. $\begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : w \end{bmatrix} \rightarrow \begin{bmatrix} \text{F}_1 : u \\ \text{F}_2 : v \end{bmatrix} \begin{bmatrix} \text{F}_1 : [\,] \\ \text{F}_2 : w \end{bmatrix} \begin{bmatrix} \text{F}_1 : v \\ \text{F}_2 : u \end{bmatrix}$

By induction on the lengths of the derivation sequences, we prove the following theorem (the full proof is deferred to appendix A):

Theorem 4. *If* $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L} \rangle$ *is a non-reentrant unification grammar and* $G^{cf} = ug2cfg(G^u)$*, then* $L(G^{cf}) = L(G^u)$.

Corollary 5. *Non-reentrant unification grammars are weakly equivalent to* CFGS.

4 Mildly context-sensitive unification grammars

In this section we show that *one-reentrant unification grammars* generate exactly the class MCSL. In such grammars each rule can have at most one reentrancy, reflecting the LIG situation where stacks can be copied to exactly one daughter in each rule.

4.1 Mapping LIG to UG_{1r}

In order to simulate a given LIG with a unification grammar, a dedicated signature is defined based on the parameters of the LIG.

Definition 18. *Given a LIG* $\langle V_N, V_t, V_s, \mathcal{R}^{li}, S\rangle$, *let* $\tau = \langle \text{ATOMS}, \text{FEATS}, \text{TAGS}, \text{WORDS}\rangle$, *where* $\text{ATOMS} = V_N \cup V_s \cup \{elist\}$, $\text{FEATS} = \{\text{HD}, \text{TL}\}$, $\text{TAGS} = \{\boxed{1}, \boxed{2}, \ldots\}$, *and* $\text{WORDS} = V_t$.

We use τ throughout this section as the signature over which UGs are defined. We use FSs over the signature τ to represent and simulate LIG symbols. In particular, FSs will encode lists in the natural way (see example 4), hence the features HD and TL. With this notation in mind, LIG symbols are mapped to FSs thus:

Definition 19. *Let toFs be a mapping of LIG symbols to feature structures, such that:*

1. *If* $t \in V_t$ *then* $\textit{toFs}(t) = \begin{bmatrix} \text{HD} : t \\ \text{TL} : elist \end{bmatrix} = \langle\langle t\rangle\rangle$
2. *If* $N \in V_N$ *and* $p_i \in V_s, 1 \le i \le n$, *then*

$$\textit{toFs}(N[p_1, \ldots, p_n]) = \begin{bmatrix} \text{HD} : N \\ \text{TL} : \begin{bmatrix} \text{HD} : p_1 \\ \text{TL} : \cdots \begin{bmatrix} \text{HD} : p_n \\ \text{TL} : elist \end{bmatrix} \cdots \end{bmatrix} \end{bmatrix} = \langle\langle N, p_1, \ldots, p_n\rangle\rangle$$

The mapping *toFs* is extended to sequences of symbols by setting $\textit{toFs}(\alpha\beta) = \textit{toFs}(\alpha)\textit{toFs}(\beta)$. Note that *toFs* is one to one.

When FSs that are images of LIG symbols are concerned, unification is reduced to identity:

Lemma 6. *Let* $X_1, X_2 \in V_N[V_s^*] \cup V_t$. *If* $\textit{toFs}(X_1) \sqcup \textit{toFs}(X_2) \neq \top$ *then* $\textit{toFs}(X_1) = \textit{toFs}(X_2)$.

When a feature structure which encodes an open-ended list (a list that is not terminated by $elist$, refer back to example 4) is unifiable with an image of a LIG symbol, the former is a prefix of the latter.

Lemma 7. *Let* $\mathsf{C} = \langle\langle p_1, \ldots, p_n | \boxed{i} \rangle\rangle$ *be a non-reentrant feature structure, where* $p_1, \ldots, p_n \in V_s$, *and let* $X \in V_N[V_s^*] \cup V_t$. *Then* $\mathsf{C} \sqcup \textit{toFs}(X) \neq \top$ *iff* $\textit{toFs}(X) = \langle\langle p_1, \ldots, p_n, \alpha\rangle\rangle$, *for some* $\alpha \in V_s^*$.

To simulate LIGs with UGs we represent each symbol in the LIG as a feature structure, encoding the stacks of LIG non-terminals as lists. Rules that propagate stacks (from mother to daughter) are simulated by means of reentrancy in the unification grammar.

Definition 20. *Let lig2ug be a mapping of* LIGS *to* UG_{1r}, *such that if* $G^{li} = \langle V_N, V_t, V_s, \mathcal{R}^{li}, S\rangle$ *and* $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L}\rangle = \textit{lig2ug}(G^{li})$ *then* G^u *is defined over the signature* τ *(definition 18),* $\mathsf{A}^s = \textit{toFs}(S[\,])$, *for all* $t \in V_t$, $\mathcal{L}(t) = \{\textit{toFs}(t)\}$ *and* $\mathcal{R}^u$ *is defined by (refer back to definition 1 for the format of LIG rules):*

- *A LIG rule of the form* $X_0 \to \alpha$ *is mapped to the unification rule* $\textit{toFs}(X_0) \to \textit{toFs}(\alpha)$
- *A LIG rule of the form* $N[p_1, \ldots, p_n \infty] \to \alpha\ N'[q_1, \ldots, q_m \infty]\ \beta$ *is mapped to the unification rule* $\langle\langle N, p_1, \ldots, p_n | \boxed{1} \rangle\rangle \to \textit{toFs}(\alpha)\ \langle\langle N', q_1, \ldots, q_m | \boxed{1} \rangle\rangle\ \textit{toFs}(\beta)$

Evidently, $\textit{lig2ug}(G^{li}) \in \text{UG}_{1r}$ for any LIG G^{li}. Also, the mapping *lig2ug* of definition 20 is one to one.

Example 11 (Mapping from LIGS to UG_{1r}). *We map the LIG* G_1^{li} *of example 1 above to* $G^u = \textit{lig2ug}(G^{li})$ *defined over the signature* τ *of definition 18, with the start symbol* $\textit{toFs}(S[\,])$. *The lexicon is defined for the words* a *and* b *as* $\mathcal{L}(a) = \{\langle a\rangle\}$ *and* $\mathcal{L}(b) = \{\langle b\rangle\}$. *The set of productions* $\mathcal{R}^{li}$, *is defined as follows:*

1. $\rho_1^u = \langle\langle S\rangle\rangle \to \langle\langle N_2\rangle\rangle$, *where the LIG rule is* $r_1 = S[\,] \to N_2[\,]$
2. $\rho_2^u = \langle\langle N_2|\boxed{1}\rangle\rangle \to \langle\langle N_2, a|\boxed{1}\rangle\rangle\langle\langle a\rangle\rangle$, *where the LIG rule is* $r_2 = N_2[\,\infty] \to N_2[a\,\infty]a$
3. $\rho_3^u = \langle\langle N_2|\boxed{1}\rangle\rangle \to \langle\langle N_2, b|\boxed{1}\rangle\rangle\langle\langle b\rangle\rangle$, *where the LIG rule is* $r_3 = N_2[\,\infty] \to N_2[b\,\infty]b$
4. $\rho_4^u = \langle\langle N_2|\boxed{1}\rangle\rangle \to \langle\langle N_3|\boxed{1}\rangle\rangle$, *where the LIG rule is* $r_4 = N_2[\,\infty] \to N_3[\,\infty]$
5. $\rho_5^u = \langle\langle N_3, a|\boxed{1}\rangle\rangle \to \langle\langle a\rangle\rangle\langle\langle N_3|\boxed{1}\rangle\rangle$, *where the LIG rule is* $r_5 = N_3[a\,\infty] \to aN_3[\,\infty]$
6. $\rho_6^u = \langle\langle N_3, b|\boxed{1}\rangle\rangle \to \langle\langle b\rangle\rangle\langle\langle N_3|\boxed{1}\rangle\rangle$, *where the LIG rule is* $r_6 = N_3[b\,\infty] \to bN_3[\,\infty]$
7. $\rho_7^u = \langle\langle N_3\rangle\rangle \to \epsilon$, *where the LIG rule is* $r_7 = N_3[\,] \to \epsilon$

The following theorem, whose proof is deferred to appendix B, summarizes this direction of the result:

Theorem 8. *If G^{li} is a LIG and $G^u = lig2ug(G^{li})$ then $L(G^u) = L(G^{li})$.*

4.2 Mapping UG_{1r} to LIG

We are now interested in the reverse direction, namely mapping (one-reentrant) UGs to LIG. The differences between the two formalisms can be summarized along three dimensions:

The basic elements Unification grammars manipulate feature structures, and rules (and forms) are MRSs; whereas LIG manipulates terminals and non-terminals with stacks of elements, and rules (and forms) are sequences of such symbols.

Rule application In UGs a rule is applied by *unification in context* of the rule and a sentential form, both of which are MRSs, whereas in LIG, the head of a rule and the selected element of a sentential form must have the same non-terminal symbol and consistent stacks.

Propagation of information in rules In UGs information is shared through reentrancies, whereas in LIG, information is propagated by copying the stack from the head of the rule to one element of its body.

We show that one-reentrant UGs can all be correctly mapped to LIGs. For the rest of this section we fix a signature $\langle \textsc{Atoms}, \textsc{Feats}, \textsc{Tags}, \textsc{Words}\rangle$ over which UGs are defined. Let $\textsc{NrFss}$ be the set of all non-reentrant FSs over this signature.

One-reentrant UGs induce highly constrained (sentential) forms: in such forms, there are no reentrancies whatsoever, neither between distinct elements nor within a single element. Hence all the FSs in forms induced by a one-reentrant UG are non-reentrant (lemma 2).

Definition 21 (Height)**.** *Let A be a feature structure with no reentrancies. The* ***height*** *of A, denoted $|\mathsf{A}|$, is the length of the longest path in A. This is well-defined since non-reentrant feature structures have finitely many paths. Let $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L}\rangle \in \mathrm{UG}_{1r}$ be a one-reentrant unification grammar. The* ***maximum height*** *of the grammar, maxHt(G^u), is the height of the highest feature structure in the grammar. This is well defined since all the feature structures of one-reentrant grammars are non-reentrant.*

The following lemma (which is proven in appendix C) indicates an important property of one-reentrant UGs. Informally, in any FS that is an element of a sentential form induced by such grammars, if two paths are long (specifically, longer than the maximum height of the grammar), then they must have a long common prefix.

Lemma 9. *Let $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L}\rangle \in \mathrm{UG}_{1r}$ be a one-reentrant unification grammar. Let* A *be an element of a sentential form induced by G^u. If $\pi \cdot \langle \mathrm{F}_j\rangle \cdot \pi_1, \pi \cdot \langle \mathrm{F}_k\rangle \cdot \pi_2 \in \Pi_A$, where $\mathrm{F}_j, \mathrm{F}_k \in$ FEATS, $j \neq k$ and $|\pi_1| \leq |\pi_2|$, then $|\pi_1| \leq$ maxHt(G^u).*

Lemma 9 facilitates a view of all the FSs induced by such a grammar as (unboundedly long) lists of elements drawn from a finite, predefined set. The set consists of all features in FEATS and all the non-reentrant feature structures whose height is limited by the maximal height of the unification grammar. Note that even with one-reentrant UGs, feature structures can be unboundedly deep. What lemma 9 establishes is that if a feature structure induced by a one-reentrant unification grammar is deep, then it can be represented as a *single* "trunk" path which is long, and all the sub-structures which "hang" from this trunk are depth-bounded. We use this property to encode such feature structures as *cords*.

Definition 22 (Cords)**.** *Let $\Psi : \mathrm{NRFSS} \times \mathrm{PATHS} \mapsto (\mathrm{FEATS} \cup \mathrm{NRFSS})^*$ be a mapping such that if* A *is a non-reentrant FS and $\pi = \langle \mathrm{F}_1, \ldots, \mathrm{F}_n\rangle \in \Pi_A$, then the **cord** $\Psi(\mathsf{A}, \pi)$ is $\langle \mathsf{A}_1, \mathrm{F}_1, \ldots, \mathsf{A}_n, \mathrm{F}_n, \mathsf{A}_{n+1}\rangle$, where for $1 \leq i \leq n+1$, A_i is the non-reentrant FSs obtained from* A *by removing from $pval(\mathsf{A}, \langle \mathrm{F}_1, \ldots, \mathrm{F}_{i-1}\rangle)$ the feature F_i and its value.*

*We also define last$(\Psi(\mathsf{A}, \pi)) = \mathsf{A}_{n+1}$. The **height** of a cord is defined as $|\Psi(\mathsf{A}, \pi)| = \max_{1 \leq i \leq n+1}(|\mathsf{A}_i|)$. For each cord $\Psi(\mathsf{A}, \pi)$ we refer to* A *as the **base feature structure** and to π as the **base path**. The **length** of a cord is the length of the base path.*

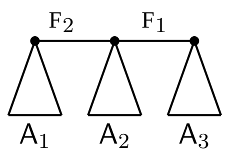
Example 12 (Cords)**.** *Let* A *be a non-reentrant feature structure over the signature* $\mathrm{FEATS} = \{\mathrm{F}_1, \mathrm{F}_2, \mathrm{F}_3\}$, $\mathrm{ATOMS} = \{a, b\}$*:*

$$\mathsf{A} = \begin{bmatrix} \mathrm{F}_1 : b \\ \mathrm{F}_2 : \begin{bmatrix} \mathrm{F}_1 : \begin{bmatrix} \mathrm{F}_2 : \begin{bmatrix} \mathrm{F}_3 : a \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \mathrm{F}_3 : \begin{bmatrix} \mathrm{F}_1 : [\,] \\ \mathrm{F}_2 : a \\ \mathrm{F}_3 : \begin{bmatrix} \mathrm{F}_1 : [\,] \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

If $\pi = \langle \mathrm{F}_2, \mathrm{F}_1\rangle$ then the cord representation of A *on the path π is $\Psi(\mathsf{A}, \pi) = \langle \mathsf{A}_1, \mathrm{F}_2, \mathsf{A}_2, \mathrm{F}_1, \mathsf{A}_3\rangle$, where*

$$\mathsf{A}_1 = \begin{bmatrix} \mathrm{F}_1 : b \\ \mathrm{F}_3 : \begin{bmatrix} \mathrm{F}_1 : [\,] \\ \mathrm{F}_2 : a \\ \mathrm{F}_3 : \begin{bmatrix} \mathrm{F}_1 : [\,] \end{bmatrix} \end{bmatrix} \end{bmatrix} ;\ \mathsf{A}_2 = [\,];\ \mathsf{A}_3 = \begin{bmatrix} \mathrm{F}_2 : \begin{bmatrix} \mathrm{F}_3 : a \end{bmatrix} \end{bmatrix}$$

$\Psi(\mathsf{A}, \pi)$ *can be graphically depicted as:*



Note that the function Ψ is one to one. In other words, given $\Psi(\mathsf{A}, \pi)$, both A and π are uniquely determined. The path π is determined by the sequence of the features on the cord $\Psi(\mathsf{A}, \pi)$, in the order they occur in the cord. Since A is non-reentrant, all A_i in $\Psi(\mathsf{A}, \pi)$ are non-reentrant feature structures, i.e., trees. To see that A is uniquely determined, simply view π as a branch of a tree and "hang" the subtrees A_i on π, in the order determined by the features in the cord, to obtain a unique feature structure.

Lemma 10. *Let* G^u *be a one-reentrant unification grammar and let* A *be an element of a sentential form induced by* G^u*. Then there is a path* $\pi \in \Pi_A$ *such that* $|\Psi(\mathsf{A}, \pi)| < maxHt(G^u)$.

Lemma 10 (which is an immediate corollary of lemma 9) implies that every FS induced by one-reentrant grammars (such FSs are necessarily non-reentrant by lemma 2) can be represented as a height-limited cord. This mapping resolves the first difference between LIG and unification grammars, by providing a representation of the *basic elements*. We use cords as the stack contents of LIG non-terminal symbols: cords can be unboundedly long, but so can LIG stacks; the crucial point is that cords are height limited, implying that they can be represented using a *finite* number of elements.

We now show how to simulate, in LIG, the unification in context of a rule and a sentential form. The first step is to have exactly one non-terminal symbol (in addition to the start symbol); when all non-terminal symbols are identical, only the content of the stack has to be taken into account. Recall that in order for a LIG rule to be applicable to a sentential form, the stack of the rule's head must be a *prefix* of the stack of the selected element in the form. The only question is whether the two stacks are equal (fixed rule head) or not (unbounded rule head). Since the contents of stacks are cords, we need a property relating two cords, on one hand, with unifiability of their base feature structures, on the other. Lemma 11 (whose proof is deferred to appendix C) establishes such a property. Informally, if the base path of one cord is a prefix of the base path of the other cord and all feature structures along the common path of both cords are unifiable, then the base feature structures of both cords are unifiable. The reverse direction also holds.

Lemma 11. *Let* $\mathsf{A}, \mathsf{B} \in \textsc{NrFss}$ *be non-reentrant feature structures and* $\pi_1, \pi_2 \in \textsc{Paths}$ *be paths such that* $\pi_1 \in \Pi_B$, $\pi_1 \cdot \pi_2 \in \Pi_A$, $\Psi(\mathsf{A}, \pi_1 \cdot \pi_2) = \langle \mathsf{t}_1, \text{F}_1, \ldots, \text{F}_{|\pi_1|}, \mathsf{t}_{|\pi_1|+1}, \text{F}_{|\pi_1|+1}, \ldots, \mathsf{t}_{|\pi_1 \cdot \pi_2|+1} \rangle$, $\Psi(\mathsf{B}, \pi_1) = \langle \mathsf{s}_1, \text{F}_1, \ldots, \mathsf{s}_{|\pi_1|+1} \rangle$*, and* $\langle \text{F}_{|\pi_1|+1} \rangle \notin \Pi_{s_{|\pi_1|+1}}$*. Then* $\mathsf{A} \sqcup \mathsf{B} \neq \top$ *iff for all* i*,* $1 \leq i \leq |\pi_1| + 1$*,* $\mathsf{s}_i \sqcup \mathsf{t}_i \neq \top$.

The length of a cord of an element of a sentential form induced by the grammar need not be bounded, but the length of any cord representation of a rule head is limited by the grammar height. By lemma 11, unifiability of two feature structures can be reduced to a comparison of two cords representing them and only the prefix of the longer cord (as long as the shorter cord) affects the result. Since the cord representation of any grammar rule's head is limited by the height of the grammar we always choose it as the shorter cord in the comparison.

We now define, for a feature structure C (which is a head of a rule) and some path π, the set that includes all feature structures that are both unifiable with C and can be represented as a cord whose height is limited by the grammar height and whose base path is π. We call this set the *compatibility set* of C and π and use it to define the set of all possible prefixes of cords whose base FSs are unifiable with C (see definition 24). Crucially, the compatibility set of C is finite for any feature structure C since the heights and the lengths of the cords are limited.

Definition 23 (Compatibility set)**.** *Given a non-reentrant feature structure* C*, a path* $\pi = \langle \text{F}_1, \ldots, \text{F}_n \rangle \in \Pi_C$ *and a natural number* h*, the* ***compatibility set****,* $\Gamma(\mathsf{C}, \pi, h)$*, is defined as the set of all feature structures* A *such that* $\mathsf{C} \sqcup \mathsf{A} \neq \top$*,* $\pi \in \Pi_A$*, and* $|\Psi(\mathsf{A}, \pi)| \leq h$.

The compatibility set is defined for a feature structure and a given path (when h is taken to be the grammar height). We now define two similar sets, FH (fixed head) and UH (unbounded head), for a given FS, independently of a path. When rules of a one-reentrant unification grammar are mapped to LIG rules (definition 25), FH and UH are used to define heads of fixed and unbounded LIG rules, respectively. A single unification rule is mapped to a *set* of LIG rules, each with a different head. The stack of the head is some member of the sets FH and UH. Each such member is a prefix of the stack of potential elements of sentential forms that the LIG rule can be applied to.

Definition 24. *Let* C *be a non-reentrant feature structure and* h *be a natural number. Then:*

$$\begin{array}{rcl} \mathrm{FH}(\mathsf{C},h) & = & \{\Psi(\mathsf{A},\pi) \mid \pi \in \Pi_C, \mathsf{A} \in \Gamma(\mathsf{C},\pi,h)\} \\ \mathrm{UH}(\mathsf{C},h) & = & \{\Psi(\mathsf{A},\pi)\cdot\langle\mathrm{F}\rangle \mid \Psi(\mathsf{A},\pi) \in \mathrm{FH}(\mathsf{C},h), pval(\mathsf{C},\pi) \textit{ is not atomic,} \\ & & \mathrm{F} \in \textsc{Feats}, \textit{ and } \mathrm{F} \textit{ is not defined in } last(\Psi(\mathsf{C} \sqcup \mathsf{A},\pi))\} \end{array}$$

This accounts for the second difference between LIG and one-reentrant unification grammars, namely *rule application*. We now briefly illustrate our account of the last difference, *propagation of information in rules*. In UG_{1r} information is shared between the rule's head and a single element in its body. Let $\rho^u = \langle \mathsf{C}_0, \ldots, \mathsf{C}_n \rangle$ be a reentrant unification rule in which the path μ_e, leaving the e-th element of the body, is reentrant with the path μ_0 leaving the head. This rule is mapped to a *set* of LIG rules, corresponding to the possible rule heads (the sets FH and UH) induced by the compatibility set of C_0. Let r be a member of this set, and let X_0 and X_e be the head and the e-th element of r, respectively. Reentrancy in ρ^u is modeled in the LIG rule by copying the stack from X_0 to X_e. The major complication is the contents of this stack, which varies according to the cord representations of C_0 and C_e and to the reentrant paths.

Summing up, in a LIG simulating a one-reentrant unification grammar, FSs are represented as stacks of symbols. The set of stack symbols V_s, therefore, is defined as a set of height bounded non-reentrant FSs. Also, all the features of the unification grammar are stack symbols. V_s is finite due to the restriction on FSs (no reentrancies and height-boundedness over the fixed signature). The set of terminals, V_t, is the words of the unification grammar. There are exactly two non-terminal symbols, S (the start symbol) and N.

The set of rules is divided to four. The *start rule* only applies once in a derivation; it simulates the sentential form which is obtained by a zero-length derivation sequence in the unification grammar. *Terminal rules* are a straight-forward implementation of the lexicon in terms of LIG. *Non-reentrant rules* are simulated in a similar way to how rules of a non-reentrant UG are simulated by CFG (section 3). The major difference is the head of the rule, X_0, which is defined as explained above. *One-reentrant rules* are simulated similarly to non-reentrant ones, the only difference being the selected element of the rule body, X_e, which is defined as follows.

Definition 25. *Let ug2lig be a* ***mapping of*** UG_{1r} ***to*** $\textsc{Ligs}$, *such that if* $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L} \rangle \in \mathrm{UG}_{1r}$ *then* *ug2lig*$(G^u) = \langle V_N, V_t, V_s, \mathcal{R}^{li}, S \rangle$, *where* $V_N = \{N, S\}$ *(fresh symbols),* $V_t = \textsc{Words}$, $V_s = \textsc{Feats} \cup \{\mathsf{A} \mid \mathsf{A} \in \textsc{NrFss}, |\mathsf{A}| \leq \textit{maxHt}(G^u)\}$, *and* $\mathcal{R}^{li}$ *is defined as follows:*

1. $S[\,] \rightarrow N[\Psi(\mathsf{A}^s, \varepsilon)]$

2. *For every* $w \in \textsc{Words}$ *such that* $\mathcal{L}(w) = \{\mathsf{C}_0\}$ *and for every* $\pi_0 \in \Pi_{\mathsf{C}_0}$, *the rule* $N[\Psi(\mathsf{C}_0, \pi_0)] \rightarrow w$ *is in* $\mathcal{R}^{li}$.

3. *Let* $\textsc{ligHead}(\mathsf{C})$ *be* $\{N[\eta] \mid \eta \in \mathrm{FH}(\mathsf{C}, \textit{maxHt}(G^u))\} \cup \{N[\eta\, \infty] \mid \eta \in \mathrm{UH}(\mathsf{C}, \textit{maxHt}(G^u))\}$

4. *If* $\langle \mathsf{C}_0, \ldots, \mathsf{C}_n \rangle \in \mathcal{R}^u$ *is a non-reentrant rule, then for every* $X_0 \in \textsc{ligHead}(\mathsf{C}_0)$ *the rule* $X_0 \rightarrow N[\Psi(\mathsf{C}_1, \varepsilon)] \ldots N[\Psi(\mathsf{C}_n, \varepsilon)]$ *is in* $\mathcal{R}^{li}$.

5. *Let* $\rho^u = \langle \mathsf{C}_0, \ldots, \mathsf{C}_n \rangle \in \mathcal{R}^u$ *and* $(0, \mu_0) \overset{\rho^u}{\rightsquigarrow} (e, \mu_e)$, *where* $1 \leq e \leq n$. *Then for every* $X_0 \in \textsc{ligHead}(\mathsf{C}_0)$ *the rule*

$$\begin{array}{rcl} X_0 & \rightarrow & N[\Psi(\mathsf{C}_1, \varepsilon)] \ldots N[\Psi(\mathsf{C}_{e-1}, \varepsilon)] \\ & & X_e \\ & & N[\Psi(\mathsf{C}_{e+1}, \varepsilon)] \ldots N[\Psi(\mathsf{C}_n, \varepsilon)] \end{array}$$

is in $\mathcal{R}^{li}$, where X_e is defined as follows. Let π_0 be the base path of the stack of X_0 and A *be the base feature structure of X_0. Applying the rule ρ^u to* A*, define*

$$(\langle \mathsf{A} \rangle, 0) \sqcup (\rho^u, 0) = (\langle \mathsf{P}_0 \rangle, \langle \mathsf{P}_0, \ldots, \mathsf{P}_e, \ldots, \mathsf{P}_n \rangle)$$

(a) If μ_0 is not a prefix of π_0 then $X_e = N[\Psi(\mathsf{P}_e, \mu_e)]$.

(b) If $\pi_0 = \mu_0 \cdot \nu$, $\nu \in$ PATHS then

i. If $X_0 = N[\Psi(\mathsf{A}, \pi_0)]$ then $X_e = N[\Psi(\mathsf{P}_e, \mu_e \cdot \nu)]$.

ii. If $X_0 = N[\Psi(\mathsf{A}, \pi_0), \mathrm{F}\,\infty]$ then $X_e = N[\Psi(\mathsf{P}_e, \mu_e \cdot \nu), \mathrm{F}\,\infty]$.

By inductions on the lengths of the derivations we prove (see appendix C) that the mapping is correct.

Theorem 12. *If $G^u \in \mathrm{UG}_{1r}$, then $L(G^u) = L(ug2lig(G^u))$.*

5 Conclusions

The main contribution of this work is the definition of two constraints on unification grammars which dramatically limit their expressivity. We prove that non-reentrant unification grammars generate exactly the class of context-free languages; and that one-reentrant unification grammars generate exactly the class of mildly context-sensitive languages. We thus obtain two linguistically plausible constrained formalisms whose computational processing is tractable.

This main result is primarily a formal grammar result, which may or may not be relevant to computational linguists. However, we maintain that it can be easily adapted such that its consequences to (practical) computational linguistics are more evident. The motivation behind this observation is that reentrancy only adds to the expressivity of a grammar formalism when it is potentially *unbounded*, i.e., when infinitely many feature structures can be the possible values at the end of the reentrant paths. It is therefore possible to modestly extend the class of unification grammars which can be shown to generate exactly the class of mildly context-sensitive languages, and allow also a limited form of more than one reentrancy among the elements in a rule, which on one hand can be most useful for grammar writers, but on the other hand adds nothing to the expressivity of the formalism. We leave the formal details of such an extension to future work.

This work can also be extended in other directions. The mapping of one-reentrant UGs to LIG is highly verbose, resulting in LIGs with a huge number of rules. We believe that it should be possible to optimize the mapping such that much smaller LIGs are generated. It would be interesting to experiment with a mapping of one-reentrant UGs to some other MCSL formalism, notably TAG.

The two constraints on unification grammars (non-reentrant and one-reentrant) are parallel to the first two classes of the Weir (1992) hierarchy of languages. A possible extension of this work could be a definition of constraints on unification grammars that would generate all the classes of the hierarchy. Another direction is an extension of one-reentrant unification grammars, where the reentrancy inside a grammar rule does not have to be between the head and one element in the body, but can also be, for example, between two elements of the body or within an element. Then it is interesting to explore the power of two-reentrant unification grammars, with limited kinds of reentrancies (as two arbitrary reentrancies suffice to model a Turing Machine).

Acknowledgments

This research was supported by The Israel Science Foundation (grant no. 136/01). We are grateful to Yael Cohen-Sygal, Nissim Francez and James Rogers for their comments and help. We greatly benefited from

comments made by an anonymous JoLLI reviewer. This paper is an extended and revised version of Feinstein and Wintner (2006).

References

Barton, Jr., G. Edward, Robert C. Berwick, and Eric Sven Ristad. 1987. The complexity of LFG. In G. Edward Barton, Jr., Robert C. Berwick, and Eric Sven Ristad, editors, *Computational Complexity and Natural Language*, Computational Models of Cognition and Perception. MIT Press, Cambridge, MA, chapter 3, pages 89–102.

Carpenter, Bob. 1992. *The Logic of Typed Feature Structures*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press.

Feinstein, Daniel and Shuly Wintner. 2006. Highly constrained unification grammars. In *Proceedings of Coling–ACL 2006*, pages 1089–1096, Sydney, Australia, July.

Gazdar, Gerald. 1988. Applicability of indexed grammars to natural languages. In Uwe Reyle and Christian Rohrer, editors, *Natural Language Parsing and Linguistic Theories*. Reidel Publishing Company, Dordrecht, pages 69–94.

Jaeger, Efrat, Nissim Francez, and Shuly Wintner. 2005. Unification grammars and off-line parsability. *Journal of Logic, Language and Information*, 14(2):199–234.

Johnson, Mark. 1988. *Attribute-Value Logic and the Theory of Grammar*, volume 16 of *CSLI Lecture Notes*. CSLI, Stanford, California.

Joshi, Aravind K. 2003. Tree-adjoining grammars. In Ruslan Mitkov, editor, *The Oxford handbook of computational linguistics*. Oxford university Press, chapter 26, pages 483–500.

Joshi, Aravind K., L. Levy, and M. Takahashi. 1975. Tree Adjunct Grammars. *Journal of Computer and System Sciences*.

Keller, Bill and David Weir. 1995. A tractable extension of linear indexed grammars. In *Proceedings of the seventh meeting of the European chapter of the Association for Computational Linguistics*, pages 75–82.

Pollard, Carl. 1984. *Generalized phrase structure grammars, head grammars and natural language*. Ph.D. thesis, Stanford University.

Satta, Giorgio. 1994. Tree-adjoining grammar parsing and boolean matrix multiplication. In *Proceedings of the 20st Annual Meeting of the Association for Computational Linguistics*, volume 20.

Savitch, Walter J., Emmon Bach, William Marsh, and Gila Safran-Naveh, editors. 1987. *The formal complexity of natural language*, volume 33 of *Studies in Linguistics and Philosophy*. D. Reidel, Dordrecht.

Shieber, Stuart M. 1986. *An Introduction to Unification Based Approaches to Grammar*. Number 4 in CSLI Lecture Notes. CSLI.

Shieber, Stuart M. 1992. *Constraint-Based Grammar Formalisms*. MIT Press, Cambridge, Mass.

Steedman, Mark. 2000. *The Syntactic Process*. Language, Speech and Communication. The MIT Press, Cambridge, Mass.

Vijay-Shanker, K. and David J. Weir. 1993. Parsing some constrained grammar formalisms. *Computational Linguistics*, 19(4):591 – 636.

Vijay-Shanker, K. and David J. Weir. 1994. The equivalence of four extensions of context-free grammars. *Mathematical systems theory*, 27:511–545.

Weir, David J. 1992. A geometric hierarchy beyond context-free languages. *Theoretical Computer Science*, 104:235–261.

Wintner, Shuly. 2006a. Introduction to unification grammars. In Zoltán Ésik, Carlos Martín-Vide, and Victor Mitrana, editors, *Recent Advances in Formal Languages and Applications*, volume 25 of *Studies in Computational Intelligence*. Springer, chapter 13, pages 321–342.

Wintner, Shuly. 2006b. Unification: Computational issues. In Keith Brown, editor, *Encyclopedia of Language and Linguistics*, volume 13. Elsevier, Oxford, second edition, pages 238–250.

A Mapping UG_{nr} to CFG

Lemma 13. *Let* $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L}\rangle$ *be a non-reentrant unification grammar over the signature* $\langle \textsc{Atoms}, \textsc{Feats}, \textsc{Words}\rangle$ *and* $\mathsf{A}^s \stackrel{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ *be a derivation sequence. Then for all* A_i *there exist a rule* $\rho^u \in \mathcal{R}^u$ *such that* $\rho^u = \mathsf{B}_0 \rightarrow \mathsf{B}_1 \ldots \mathsf{B}_m$ *and an index* j*,* $0 < j \leq m$*, for which* $\mathsf{B}_j = \mathsf{A}_i$*.*

Proof. We prove by induction on the length of the derivation sequence. The induction hypothesis is that if $\mathsf{A}^s \stackrel{k}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ then for all A_i, where $1 \leq i \leq n$, there are a rule $\rho^u = \mathsf{B}_0 \rightarrow \mathsf{B}_1 \ldots \mathsf{B}_m$, $\rho^u \in \mathcal{R}^u$ and an index j such that $\mathsf{B}_j = \mathsf{A}_i$. If $k = 1$, then there is a rule $\mathsf{C} \rightarrow \mathsf{A}_1 \ldots \mathsf{A}_n$, $\mathsf{A}^s \sqcup \mathsf{C} \neq \top$, and all A_i are part of the rule's body because a non-reentrant rule does not propagate information from the rule head to the body. Assume that the hypothesis holds for every l, $0 < l < k$; let the length of the derivation sequence be k. If $\mathsf{A}^s \stackrel{k-1}{\Rightarrow}_u \mathsf{D}_1 \ldots \mathsf{D}_m \stackrel{1}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ then there exist an index j and a rule $\rho^u = \mathsf{C} \rightarrow \mathsf{A}_j \ldots \mathsf{A}_{n-m+j} \in \mathcal{R}^u$ such that:

1. $\mathsf{C} \sqcup \mathsf{D}_j \neq \top$
2. $\mathsf{D}_i = \begin{cases} \mathsf{A}_i & i < j \\ \mathsf{A}_{i+n-m} & i > j \end{cases}$

By the induction hypothesis for all A_i, where $i < j$ or $i > n - m + j$, there is a rule that contains A_i in its body. For A_i, where $j \leq i \leq n - m + j$, the rule ρ^u completes the proof. □

Theorem 14 (Theorem 4)**.** *Let* $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L}\rangle$ *be a non-reentrant unification grammar over the signature* $\langle \textsc{Atoms}, \textsc{Feats}, \textsc{Words}\rangle$ *and* $G^{cf} = \langle V_N, V_t, \mathcal{R}^{cf}, S^{cf}\rangle =$ *ug2cfg*(G^u)*. Then* $L(G^{cf}) = L(G^u)$*.*

Proof. We prove by induction on the length of a derivation sequence that $\mathsf{A}^s \stackrel{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ iff $S^{cf} \stackrel{*}{\Rightarrow}_{cf} \mathsf{A}_1 \ldots \mathsf{A}_n$.

Assume that $\mathsf{A}^s \stackrel{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$. The induction hypothesis is that if $\mathsf{A}^s \stackrel{k}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ then $S^{cf} \stackrel{k}{\Rightarrow}_{cf} \mathsf{A}_1 \ldots \mathsf{A}_n$. If $k = 1$, then there is a rule $\mathsf{C} \rightarrow \mathsf{A}_1 \ldots \mathsf{A}_n$, $\mathsf{A}^s \sqcup \mathsf{C} \neq \top$, and by the definition of *ug2cfg*,

$S^{cf} \rightarrow \mathsf{A}_1 \ldots \mathsf{A}_n \in \mathcal{R}^{cf}$. Then $S^{cf} \overset{k=1}{\Rightarrow}_{cf} \mathsf{A}_1 \ldots \mathsf{A}_n$. Assume that the hypothesis holds for every l, $0 < l < k$; let the length of the derivation sequence be k. If $\mathsf{A}^s \overset{k-1}{\Rightarrow}_u \mathsf{D}_1 \ldots \mathsf{D}_m \overset{1}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ then there exist an index j and a rule $\rho_1^u = \mathsf{C} \rightarrow \mathsf{A}_j \ldots \mathsf{A}_{n-m+j} \in \mathcal{R}^u$ such that:

1. $\mathsf{C} \sqcup \mathsf{D}_j \neq \top$
2. $\mathsf{D}_i = \begin{cases} \mathsf{A}_i & i < j \\ \mathsf{A}_{i+n-m} & i > j \end{cases}$

By lemma 13 there is some rule $\rho_2^u \in \mathcal{R}^u$ such that D_j is an element of its body. Hence, by definition 17 there is a rule $r_3 = \mathsf{D}_j \rightarrow \mathsf{A}_j \ldots \mathsf{A}_{n-m+j} \in \mathcal{R}^{cf}$ which is a result of combining ρ_1^u and ρ_2^u. By the induction hypothesis $S^{cf} \overset{k-1}{\Rightarrow}_{cf} \mathsf{D}_1 \ldots \mathsf{D}_n$, and by application of the rule r_3 we obtain:

$$S^{cf} \overset{k}{\Rightarrow}_{cf} \mathsf{D}_1 \ldots \mathsf{D}_{j-1}\mathsf{A}_j \ldots \mathsf{A}_{j+n-m}\mathsf{D}_{j+1} \ldots \mathsf{D}_m = \mathsf{A}_1 \ldots \mathsf{A}_n$$

Assume $S^{cf} \overset{*}{\Rightarrow}_{cf} \mathsf{A}_1 \ldots \mathsf{A}_n$.[3] The induction hypothesis is that if $S^{cf} \overset{k}{\Rightarrow}_{cf} \mathsf{A}_1 \ldots \mathsf{A}_n$ then $\mathsf{A}^s \overset{k}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$. If $k = 1$, then there is a rule $S^{cf} \rightarrow \mathsf{A}_1 \ldots \mathsf{A}_n \in \mathcal{R}^{cf}$ and by definition of *ug2cfg* (note that S^{cf} is not a part of any rule body in $\mathcal{R}^{cf}$), $\mathsf{C} \rightarrow \mathsf{A}_1 \ldots \mathsf{A}_n \in \mathcal{R}^u$, where $\mathsf{A}^s \sqcup \mathsf{C} \neq \top$. Then $\mathsf{A}^s \overset{k=1}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$. Assume that the hypothesis holds for every i, $0 < i < k$; let the length of the derivation sequence be k. If $S^{cf} \overset{k-1}{\Rightarrow}_{cf} \mathsf{D}_1 \ldots \mathsf{D}_m \overset{1}{\Rightarrow}_{cf} \mathsf{A}_1 \ldots \mathsf{A}_n$ then there exist an index j and a rule $r_1 = \mathsf{D}_j \rightarrow \mathsf{A}_j \ldots \mathsf{A}_{n-m+j} \in \mathcal{R}^{cf}$ such that:

$$\mathsf{D}_i = \begin{cases} \mathsf{A}_i & i < j \\ \mathsf{A}_{i+n-m} & i > j \end{cases}$$

By the definition of *ug2cfg* there are rules $\rho_2^u = \mathsf{B}_0 \rightarrow \mathsf{B}_1 \ldots \mathsf{B}_p$, $\rho_3^u = \mathsf{C} \rightarrow \mathsf{A}_j \ldots \mathsf{A}_{n-m+j}$ in $\mathcal{R}^u$ and an index t, $1 \leq t \leq p$, such that $\mathsf{B}_t = \mathsf{D}_j$ and $\mathsf{C} \sqcup \mathsf{D}_j \neq \top$.

By the induction hypothesis, $\mathsf{A}^s \overset{k-1}{\Rightarrow}_u \mathsf{D}_1 \ldots \mathsf{D}_n$, and by application of the rule ρ_3^u we obtain:

$$\mathsf{A}^s \overset{k}{\Rightarrow}_u \mathsf{D}_1 \ldots \mathsf{D}_{j-1}\mathsf{A}_j \ldots \mathsf{A}_{j+n-m}\mathsf{D}_{j+1} \ldots \mathsf{D}_m = \mathsf{A}_1 \ldots \mathsf{A}_n$$

In sum, $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ iff $S^{cf} \overset{*}{\Rightarrow}_{cf} \mathsf{A}_1 \ldots \mathsf{A}_n$. Hence, $L(G^{cf}) = L(\textit{ug2cfg}(G^u))$. ☐

B Mapping LIG to UG_{1r}

To show that the unification grammar *lig2ug*(G^{li}) correctly simulates the LIG grammar G^{li} we first prove that every derivation in the latter has a corresponding derivation in the former (theorem 15). Theorem 16 proves the reverse direction.

Theorem 15. *Let $G^{li} = \langle V_N, V_t, V_s, \mathcal{R}^{li}, S \rangle$ be a LIG and $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L} \rangle$ be lig2ug(G^{li}). If $S[\,] \overset{*}{\Rightarrow}_{li} \alpha$ then $\mathsf{A}^s \overset{*}{\Rightarrow}_u$ toFs(α), where $\alpha \in (V_N[V_s^*] \cup V_t)^*$.*

Proof. We prove by induction on the length of the derivation sequence. The induction hypothesis is that if $S[\,] \overset{k}{\Rightarrow}_{li} \alpha$, then $\mathsf{A}^s \overset{k}{\Rightarrow}_u$ *toFs*(α). If $k = 1$, then

[3] Recall that all elements of V_N are feature structures, and therefore all the elements of a (CFG) sentential form can be represented as A_i, where A_i is a feature structure.

1. $S[\,] \overset{k=1}{\Rightarrow}_{li} \alpha$;

2. Hence, $S[\,] \rightarrow \alpha \in \mathcal{R}^{li}$;

3. By definition 20, *toFs*$(S) \rightarrow$ *toFs*$(\alpha) \in \mathcal{R}^{u}$;

4. Since $\mathsf{A}^s =$ *toFs*(S) we obtain that $\mathsf{A}^s \rightarrow$ *toFs*$(\alpha) \in \mathcal{R}^{u}$;

5. Therefore, $\mathsf{A}^s \overset{k=1}{\Rightarrow}_{u}$ *toFs*(α)

Assume that the hypothesis holds for every i, $0 < i < k$; let the length of the derivation sequence be k.

1. Let $S[\,] \overset{k-1}{\Rightarrow}_{li} \gamma_1\ N[p_1, \ldots, p_n]\ \gamma_2 \overset{1}{\Rightarrow}_{li} \gamma_1\ \alpha\ \gamma_2$, where $\gamma_1, \gamma_2, \alpha \in (V_N[V_s^*] \cup V_t)^*$. Let $r \in \mathcal{R}^{li}$ be a LIG rule that is applied to $N[p_1, \ldots, p_n]$ at step k of the derivation.

2. By the induction hypothesis, $\mathsf{A}^s \overset{k-1}{\Rightarrow}_{u}$ *toFs*$(\gamma_1\ N_i[p_1, \ldots, p_n]\ \gamma_2)$.

3. By definition 19,

$$\begin{aligned} \textit{toFs}(\gamma_1 N[p_1, \ldots, p_n]\gamma_2) &= \textit{toFs}(\gamma_1)\ \textit{toFs}(N[p_1, \ldots, p_n])\ \textit{toFs}(\gamma_2) \\ &= \textit{toFs}(\gamma_1)\ \langle\langle N, p_1, \ldots, p_n\rangle\rangle\ \textit{toFs}(\gamma_2) \end{aligned}$$

4. From (2) and (3), $\mathsf{A}^s \overset{k-1}{\Rightarrow}_{u}$ *toFs*$(\gamma_1)\ \langle\langle N, p_1, \ldots, p_n\rangle\rangle$ *toFs*(γ_2).

5. The rule r can be of either of two forms as follows:

   (a) Let r be $N[p_1, \ldots, p_n] \rightarrow \alpha$.

      i. By definition 20, $\mathcal{R}^u$ includes the rule *toFs*$(N[p_1, \ldots, p_n]) \rightarrow$ *toFs*(α).

      ii. This rule is applicable to the form in (4), providing $\mathsf{A}^s \overset{k}{\Rightarrow}_{u}$ *toFs*(γ_1) *toFs*(α) *toFs*(γ_2).

      iii. By definition 19, *toFs*(γ_1) *toFs*(α) *toFs*$(\gamma_2) =$ *toFs*$(\gamma_1\ \alpha\ \gamma_2)$. Hence $\mathsf{A}^s \overset{k}{\Rightarrow}_{u}$ *toFs*$(\gamma_1\ \alpha\ \gamma_2)$.

   (b) Let r be $N[p_1, \ldots, p_x\ \infty] \rightarrow \alpha_1\ M[q_1, \ldots, q_m\ \infty]\ \alpha_2$, where $x \leq n$, $M \in V_N$, $q_1, \ldots, q_m \in V_s$ and $\alpha_1, \alpha_2 \in (V_N[V_s^*] \cup V_t)^*$.

      i. By applying the rule r at the last derivation step in (1) we obtain:

$$S[\,] \overset{k-1}{\Rightarrow}_{li} \gamma_1\ N[p_1, \ldots, p_n]\ \gamma_2 \overset{1}{\Rightarrow}_{li} \gamma_1\ \alpha_1\ M[q_1, \ldots, q_m, p_{x+1}, \ldots, p_n]\ \alpha_2\ \gamma_2$$

      ii. By definition 20, $\mathcal{R}^u$ includes the rule

$$\langle\langle N, p_1, \ldots, p_x | \boxed{1}\ \rangle\rangle \rightarrow \textit{toFs}(\alpha_1)\ \langle\langle M, q_1, \ldots, q_m | \boxed{1}\ \rangle\rangle\ \textit{toFs}(\alpha_2)$$

      iii. By applying this rule to the form in (4) we obtain

$$\begin{aligned} \mathsf{A}^s \overset{k-1}{\Rightarrow}_{u} &\quad \textit{toFs}(\gamma_1)\ \langle\langle N, p_1, \ldots, p_n\rangle\rangle\ \textit{toFs}(\gamma_2) \\ \overset{1}{\Rightarrow}_{u} &\quad \textit{toFs}(\gamma_1)\ \textit{toFs}(\alpha_1)\ \langle\langle M, q_1, \ldots, q_m, p_{x+1}, \ldots, p_n\rangle\rangle\ \textit{toFs}(\alpha_2)\ \textit{toFs}(\gamma_2) \end{aligned}$$

iv. By definition 19,

$$\textit{toFs}(M[q_1,\ldots,q_m,p_{x+1},\ldots,p_n]) = \langle\langle M, q_1,\ldots,q_m,p_{x+1},\ldots,p_n\rangle\rangle$$

Hence

$$\mathsf{A}^s \overset{k}{\Rightarrow}_u \textit{toFs}(\gamma_1)\ \textit{toFs}(\alpha_1)\ \textit{toFs}(M[q_1,\ldots,q_m,p_{x+1},\ldots,p_n])\ \textit{toFs}(\alpha_2)\ \textit{toFs}(\gamma_2)$$

v. Therefore, $\mathsf{A}^s \overset{k}{\Rightarrow}_u \textit{toFs}(\gamma_1\ \alpha_1\ M[q_1,\ldots,q_m,p_{x+1},\ldots,p_n]\ \alpha_2\ \gamma_2)$.

□

Theorem 16. *Let $G^{li} = \langle V_N, V_t, V_s, \mathcal{R}^{li}, S\rangle$ be a LIG and $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L}\rangle = lig2ug(G^{li})$ be a one-reentrant unification grammar. If $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ then $S[\,] \overset{*}{\Rightarrow}_{li} X_1 \ldots X_n$ such that for every i, $1 \leq i \leq n$, $\mathsf{A}_i = \textit{toFs}(X_i)$.*

Proof. We prove by induction on the length of the (unification) derivation sequence. The induction hypothesis is that if $\mathsf{A}^s \overset{k}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$, then $S[\,] \overset{k}{\Rightarrow}_{li} X_1 \ldots X_n$ such that for every i, $1 \leq i \leq n$, $\mathsf{A}_i = \textit{toFs}(X_i)$. If $k = 1$, then $\mathsf{A}^s \overset{k=1}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$. Hence, $\mathsf{A}^s \rightarrow \mathsf{A}_1 \ldots \mathsf{A}_n \in \mathcal{R}^u$. By definition 20, $\mathsf{A}^s = \textit{toFs}(S[\,])$. Since *toFs* is a one-to-one mapping we obtain that the unification rule is created from the LIG rule $S^{li}[\,] \rightarrow X_1 \ldots X_n \in \mathcal{R}^{li}$, where for every i, $1 \leq i \leq n$, $\mathsf{A}_i = \textit{toFs}(X_i)$. Therefore, $S^{li}[\,] \overset{k}{\Rightarrow}_{li} X_1 \ldots X_n$ and for every i, $1 \leq i \leq n$, $\mathsf{A}_i = \textit{toFs}(X_i)$.

Assume that the hypothesis holds for every l, $0 < l < k$; let the length of the derivation sequence be k.

1. Assume that $\mathsf{A}^s \overset{k}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$. Then $\mathsf{A}^s \overset{k-1}{\Rightarrow}_u \mathsf{B}_1 \ldots \mathsf{B}_m \overset{1}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$.

2. The last step of the unification derivation is established through a rule $\rho^u = \mathsf{C}_0 \rightarrow \mathsf{C}_1 \ldots \mathsf{C}_{n-m+1}$, $\rho^u \in \mathcal{R}^u$, and an index j, such that:

$$(\langle \mathsf{B}_1,\ldots,\mathsf{B}_m\rangle, j) \sqcup (\langle \mathsf{C}_0,\ldots,\mathsf{C}_{n-m+1}\rangle, 0) =$$
$$(\langle \mathsf{B}_1,\ldots,\mathsf{B}_{j-1},\mathsf{Q},\mathsf{B}_{j+1},\ldots,\mathsf{B}_m\rangle, \langle \mathsf{Q},\mathsf{A}_j,\ldots,\mathsf{A}_{j+n-m}\rangle)$$

3. By lemma 2, the sentential form $\langle \mathsf{A}_1,\ldots,\mathsf{A}_n\rangle$ has no reentrancies between its elements, hence for every i, $1 \leq i < j$, $\mathsf{A}_i = \mathsf{B}_i$ and for i, $j < i \leq m$, $\mathsf{A}_{i+n-m} = \mathsf{B}_i$.

4. By the induction hypothesis, if $\mathsf{A}^s \overset{k-1}{\Rightarrow}_u \mathsf{B}_1 \ldots \mathsf{B}_m$ then $S^{li}[\,] \overset{k-1}{\Rightarrow}_{li} Y_1 \ldots Y_m$ and

$$\langle \mathsf{B}_1,\ldots,\mathsf{B}_m\rangle = \langle \textit{toFs}(Y_1),\ldots,\textit{toFs}(Y_m)\rangle$$

5. Hence, $\mathsf{A}^s \overset{k-1}{\Rightarrow}_u \textit{toFs}(Y_1)\ldots\textit{toFs}(Y_m) \overset{1}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ and from (3), for every i, $1 \leq i < j$, $\mathsf{A}_i = \textit{toFs}(Y_i)$ and for i, $j < i \leq m$, $\mathsf{A}_{i+n-m} = \textit{toFs}(Y_i)$.

6. By definition 20, the rule ρ^u is created from a LIG rule r. We now show that the rule r can be applied to the element Y_j of the LIG sentential form, $\langle Y_1,\ldots,Y_m\rangle$, and the resulting sentential form, $\langle X_1,\ldots,X_n\rangle$, for every i, $1 \leq i \leq n$, satisfies the equation $\mathsf{A}_i = \textit{toFs}(X_i)$. Since from (5), for every i, $1 \leq i < j$, $\mathsf{A}_i = \textit{toFs}(Y_i)$ and for i, $j < i \leq m$, $\mathsf{A}_{i+n-m} = \textit{toFs}(Y_i)$, we just need to show that $\mathsf{A}_i = \textit{toFs}(X_i)$ for every i, $j \leq i \leq n-m+j$.

7. By definition of LIG the rule r has one of the following forms:

   (a) Let $r = N_i[p_1, \ldots, p_x] \to Z_1 \ldots Z_{n-m+1}$. Hence, by definition 20, the unification rule ρ^u is

$$\textit{toFs}(N_i[p_1, \ldots, p_x]) \to \textit{toFs}(Z_1) \ldots \textit{toFs}(Z_{n-m+1})$$

   where $\mathsf{C}_0 = \textit{toFs}(N_i[p_1, \ldots, p_x])$ and for every i, $1 \leq i \leq n-m+1$, $\mathsf{C}_i = \textit{toFs}(Z_i)$. Note that there are no reentrancies between the elements of the unification rule ρ^u and hence $\langle \mathsf{A}_j, \ldots, \mathsf{A}_{n-m+j} \rangle = \langle \mathsf{C}_1, \ldots, \mathsf{C}_{n-m+1} \rangle$.
   We now show that the rule r can be applied to the element Y_j of the LIG sentential form. Since $\mathsf{C}_0 \sqcup \mathsf{B}_j = \mathsf{C}_0 \sqcup \textit{toFs}(Y_j) = \textit{toFs}(N_i[p_1, \ldots, p_x]) \sqcup \textit{toFs}(Y_j) \neq \top$ we obtain, by lemma 6, that

$$\textit{toFs}(Y_j) = \textit{toFs}(N_i[p_1, \ldots, p_x])$$

   Since *toFs* is one-to-one mapping we obtain that $Y_j = N_i[p_1, \ldots, p_x]$. Hence the LIG rule r can be applied to Y_j.
   We now show that $\mathsf{A}_i = \textit{toFs}(X_i)$ for every i, $j \leq i \leq n-m+j$. We apply the rule r to Y_j as follows:

$$Y_1 \ldots Y_j \ldots Y_m \overset{1}{\Rightarrow}_{li} X_1 \ldots X_{j-1} Z_1 \ldots Z_{n-m+1} X_{n-m+j+1} \ldots X_n$$

   Hence $\langle X_j, \ldots, X_{n-m+j} \rangle = \langle Z_1, \ldots, Z_{n-m+1} \rangle$. Therefore,

$$\begin{aligned} \langle \mathsf{A}_j, \ldots, \mathsf{A}_{n-m+j} \rangle &= \langle \mathsf{C}_1, \ldots, \mathsf{C}_{n-m+1} \rangle \\ &= \langle \textit{toFs}(Z_1), \ldots, \textit{toFs}(Z_{n-m+1}) \rangle \\ &= \langle \textit{toFs}(X_j), \ldots, \textit{toFs}(X_{n-m+j}) \rangle \end{aligned}$$

   (b) Let $r = N_i[p_1, \ldots, p_x \; \infty] \to Z_1 \ldots Z_{e-1} \; N_f[q_1, \ldots, q_y \; \infty] \; Z_{e+1} \ldots Z_{n-m+1}$, where $1 \leq e \leq n-m+1$. Hence, by definition 20, the unification rule ρ^u is defined as

$$\langle \langle N_i, p_1, \ldots, p_x | \boxed{1} \rangle \rangle \to \textit{toFs}(Z_1 \ldots Z_{e-1}) \; \langle \langle N_f, q_1, \ldots, q_y | \boxed{1} \rangle \rangle \; \textit{toFs}(Z_{e+1} \ldots Z_{n-m+1})$$

   where $\mathsf{C}_0 = \langle \langle N_i, p_1, \ldots, p_x | \boxed{1} \rangle \rangle$, $\mathsf{C}_e = \langle \langle N_f, q_1, \ldots, q_y | \boxed{1} \rangle \rangle$ and for every i, $i \neq e$, $\mathsf{C}_i = \textit{toFs}(Z_i)$. Note that there is a reentrancy between C_0 and C_e. We now calculate the information propagated from B_j to A_{j+e-1} during the last step of the unification derivation (see 2). Since $\mathsf{C}_0 \sqcup \mathsf{B}_j = \mathsf{C}_0 \sqcup \textit{toFs}(Y_j) \neq \top$ we obtain by lemma 7, that $\textit{toFs}(Y_j) = \langle \langle N_i, p_1, \ldots, p_x, \gamma \rangle \rangle$, where $\gamma \in V_s^*$. Therefore, $\mathsf{A}_{j+e-1} = \langle \langle N_f, q_1, \ldots, q_y, \gamma \rangle \rangle$.
   We now show that the LIG rule r can be applied to the element Y_j of the LIG sentential form. Since *toFs* is one-to-one and $\textit{toFs}(Y_j) = \langle \langle N_i, p_1, \ldots, p_x, \gamma \rangle \rangle$, $Y_j = N_i[N_i, p_1, \ldots, p_x, \gamma]$. Hence the LIG rule r can be applied to Y_j.
   We now show that $\mathsf{A}_i = \textit{toFs}(X_i)$ for every i, $j \leq i \leq n-m+j$. We apply the rule r to Y_j as follows:

$$\begin{aligned} &Y_1 \ldots Y_j \ldots Y_m \overset{1}{\Rightarrow}_{li} \\ &X_1 \ldots X_{j-1} \; Z_1 \ldots Z_{e-1} \; N_f[q_1, \ldots, q_y, \gamma] \; Z_{e+1} \ldots Z_{n-m+1} \; X_{n-m+j+1} \ldots X_n \end{aligned}$$

   Hence $\langle X_j, \ldots, X_{n-m+j} \rangle = \langle Z_1, \ldots, Z_{e-1}, N_f[q_1, \ldots, q_y, \gamma], Z_{e+1}, \ldots, Z_{n-m+1} \rangle$. Therefore,

$$\begin{aligned} &\langle \mathsf{A}_j, \ldots, \mathsf{A}_{j+e-1}, \ldots, \mathsf{A}_{n-m+j} \rangle \\ &= \mathsf{C}_1 \ldots \mathsf{C}_{e-1} \; \langle N_f[q_1, \ldots, q_y, \gamma] \rangle \; \mathsf{C}_{e+1} \ldots \mathsf{C}_{n-m+1} \\ &= \textit{toFs}(Z_1 \ldots Z_{e-1}) \; \textit{toFs}(N_f[q_1, \ldots, q_y, \gamma]) \; \textit{toFs}(Z_{e+1} \ldots Z_{n-m+1}) \\ &= \langle \textit{toFs}(X_j), \ldots, \textit{toFs}(X_{n-m+j}) \rangle \end{aligned}$$

□

Theorem 17 (Theorem 8). *If $G^{li} = \langle V_N, V_t, V_s, \mathcal{R}^{li}, S^{li}\rangle$ is a LIG then there exists a unification grammar $G^u = lig2ug(G^{li})$ such that $L(G^u) = L(G^{li})$.*

Proof. Let $G^{li} = \langle V_N, V_t, V_s, \mathcal{R}^{li}, N\rangle$ be a LIG and $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L}\rangle = lig2ug(G^{li})$. Then by theorem 15, if $S[\,] \overset{*}{\Rightarrow}_{li} \alpha$ then $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathit{toFs}(\alpha)$, where $\alpha = w_1, \ldots, w_n \in V_t^*$. By definition 20, for every i, $\mathcal{L}(w_i) = \{\mathit{toFs}(w_i)\}$, hence $\mathit{toFs}(\alpha) = \mathit{toFs}(w_1), \ldots, \mathit{toFs}(w_n)$. Hence $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathit{toFs}(w_1), \ldots, \mathit{toFs}(w_n) \in L(G^u)$.

Assume that $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathsf{A}_1, \ldots, \mathsf{A}_n$, where $\mathsf{A}_1, \ldots, \mathsf{A}_n$ is a pre-terminal sequence and $\mathsf{A}_1, \ldots, \mathsf{A}_n \overset{*}{\Rightarrow}_u w_1, \ldots, w_n$. By theorem 16, there is the LIG derivation sequence such that $S[\,] \overset{*}{\Rightarrow}_{li} X_1, \ldots, X_n$ and for all i, $\mathit{toFs}(X_i) = \mathsf{A}_i$. By definition 20, each entry $\mathcal{L}(w_i) = \{\mathsf{A}_i\}$ in the lexicon of G^u is created from a terminal rule $X_i \to w_i$ in $\mathcal{R}^{li}$. Therefore, $S[\,] \overset{*}{\Rightarrow}_{li} X_1, \ldots, X_n \overset{*}{\Rightarrow}_{li} w_1, \ldots, w_n$. □

C Mapping UG_{1r} to LIG

Lemma 18 (Lemma 9). *Let $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L}\rangle \in \mathrm{UG}_{1r}$ be a one-reentrant unification grammar. Let λ be a sentential form derived by G^u and A be an element of λ. If $\pi \cdot \langle \mathrm{F}_j\rangle \cdot \pi_1, \pi \cdot \langle \mathrm{F}_k\rangle \cdot \pi_2 \in \Pi_A$, where $\mathrm{F}_j, \mathrm{F}_k \in \mathrm{FEATS}$, $\mathrm{F}_j \neq \mathrm{F}_k$ and $|\pi_1| \leq |\pi_2|$, then $|\pi_1| \leq \mathit{maxHt}(G^u)$.*

Proof. We prove by induction on the length of the derivation sequence that if $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$, then the lemma conditions hold. Let $h = \mathit{maxHt}(G^u)$.

The induction hypothesis is that if $\mathsf{A}^s \overset{k}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$, then the lemma conditions hold for any A_l, where $1 \leq l \leq n$. If $k = 0$, then by definition $\mathsf{A}^s \overset{0}{\Rightarrow}_u \mathsf{A}^s$. Since $|\mathsf{A}^s| \leq h$ then for any j,π and π_1, such that $\pi \cdot \langle \mathrm{F}_j\rangle \cdot \pi_1 \in \Pi_{\mathsf{A}^s}$, $|\pi \cdot \langle \mathrm{F}_j\rangle \cdot \pi_1| \leq h$. Therefore, $|\pi_1| < h$.

Assume that the hypothesis holds for every i, $0 \leq i < k$; let the length of the derivation sequence be k. Let $\mathsf{A}^s \overset{k-1}{\Rightarrow}_u \mathsf{B}_1 \ldots \mathsf{B}_m \overset{1}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$. Then by definition of UG_{1r} derivation, there are an index j and a rule $\rho^u = \mathsf{C}_0 \to \mathsf{C}_1 \ldots \mathsf{C}_{n-m+1}$, $\rho^u \in \mathcal{R}^u$, such that

$$(\langle \mathsf{C}_0, \ldots, \mathsf{C}_{n-m+1}\rangle, 0) \sqcup (\langle \mathsf{B}_1, \ldots, \mathsf{B}_m\rangle, j) = (\langle \mathsf{Q}_0, \ldots, \mathsf{Q}_{n-m+1}\rangle, \langle \mathsf{B}_1, \ldots, \mathsf{B}_{j-1}, \mathsf{Q}_0, \mathsf{B}_{j+1}, \ldots, \mathsf{B}_m\rangle)$$

where

1. $\langle \mathsf{A}_1, \ldots, \mathsf{A}_{j-1}\rangle = \langle \mathsf{B}_1, \ldots, \mathsf{B}_{j-1}\rangle$
2. $\langle \mathsf{A}_j, \ldots, \mathsf{A}_{n-m+j}\rangle = \langle \mathsf{Q}_1, \ldots, \mathsf{Q}_{n-m+1}\rangle$
3. $\langle \mathsf{A}_{n-m+j+1}, \ldots, \mathsf{A}_n\rangle = \langle \mathsf{B}_{j+1}, \ldots, \mathsf{B}_m\rangle$

By the induction hypothesis, in cases (1) and (3) the lemma conditions hold for A_l, where $1 \leq l < j$ or $n - m + j + 1 \leq l \leq n$. We now analyze case (2). Since G^u is one-reentrant there are only two options for the rule ρ^u:

1. ρ^u has no reentrancies;
2. $(0, \pi_0) \overset{\rho^u}{\leftrightsquigarrow} (e, \pi_e)$, where $1 \leq e \leq n - m + 1$;

If ρ^u is non-reentrant, $\langle \mathsf{C}_1, \ldots, \mathsf{C}_{n-m+1} \rangle = \langle \mathsf{Q}_1, \ldots, \mathsf{Q}_{n-m+1} \rangle = \langle \mathsf{A}_j, \ldots, \mathsf{A}_{n-m+j} \rangle$. Hence for any l, $j \leq l \leq n - m + j$, $|\mathsf{A}_l| \leq h$. Hence, for any F, π and π_1, such that $\pi \cdot \langle \mathrm{F}_j \rangle \cdot \pi_1 \in \Pi_{\mathsf{A}_i}$, $|\pi \cdot \langle \mathrm{F}_j \rangle \cdot \pi_1| \leq h$. Therefore, $|\pi_1| < h$.

If $(0, \pi_0) \overset{\rho^u}{\rightsquigarrow} (e, \pi_e)$ then by the definition of unification, $\mathsf{Q}_l = \mathsf{C}_l$ if $1 \leq l < e$ or $e < l \leq n - m + 1$, hence $|\mathsf{Q}_l| \leq h$. Therefore, the lemma conditions hold for any Q_l, where $l \neq e$. We now check whether the lemma conditions hold for Q_e. The rule ρ^u, when applied to B_j, can result in modifying the body of the rule, $\mathsf{C}_1 \ldots \mathsf{C}_{n-m+1}$. However, due to the fact that ρ^u is one-reentrant, only a single element C_e can be modified. Furthermore, the only possible modifications to C_e are addition of paths and further specification of atoms (lemma 3). The latter has no effect on path length, so we focus on the former. The only way for a path $\pi_e \cdot \pi$ to be added is if some path $\pi_0 \cdot \pi$ already exists in B_j. Hence, let P be a set of paths such that:

$$P = \{\pi_e \cdot \pi \mid \pi_0 \cdot \pi \in \Pi_{\mathsf{B}_j}\}$$

By definition of unification $\Pi_{\mathsf{Q}_e} = P \cup \Pi_{\mathsf{C}_e}$. To check the lemma conditions we only need to check the pairs of paths where both members are longer than h, otherwise the conditions trivially hold. Since for any path π, $\pi \in \Pi_{\mathsf{C}_e}$, $|\pi| \leq h$, we check only the pairs of paths from P to evaluate the lemma conditions. Let $\pi_e \cdot \pi_1, \pi_e \cdot \pi_2 \in P \subseteq \Pi_{\mathsf{Q}_e}$, where $|\pi_1| \leq |\pi_2|$, π_1 and π_2 differ in the first feature. By definition of P, $\pi_0 \cdot \pi_1, \pi_0 \cdot \pi_2 \in \Pi_{\mathsf{B}_j}$. Hence, by the induction hypothesis $|\pi_1| \leq h$. Therefore, for any pair of paths in Π_{Q_e} the lemma conditions hold. ☐

Lemma 19. *Let* A *and* B *be two non-reentrant feature structures. Let* π_A, π_B *be paths such that* $\pi_A \in \Pi_A$, $\pi_B \in \Pi_B$ *and* $\mathit{last}(\Psi(\mathsf{A}, \pi_A)) \notin$ Atoms. *And let* G *be a feature such that* $\langle \mathrm{G} \rangle \notin \Pi_{\mathit{last}(\Psi(\mathsf{A},\pi_A))}$. *Then* $\Psi(\mathsf{A}, \pi_A) \cdot \langle \mathrm{G} \rangle \cdot \Psi(\mathsf{B}, \pi_B)$ *is a cord.*

Let $\Psi(\mathsf{A}, \pi_A) = \langle \mathsf{A}_1, \mathrm{F}_1, \ldots, \mathsf{A}_i, \mathrm{F}_i, \mathsf{A}_{i+1}, \ldots, \mathrm{F}_n, \mathsf{A}_{n+1} \rangle$. *Then for any* i, $1 \leq i \leq n$, *the sequences* $\langle \mathsf{A}_1, \mathrm{F}_1, \ldots, \mathsf{A}_i \rangle$ *and* $\langle \mathsf{A}_{i+1}, \ldots, \mathrm{F}_n, \mathsf{A}_{n+1} \rangle$ *are cords.*

Proof. Immediate from the definition of cords. ☐

Lemma 20 (Lemma 11)**.** *Let* $\mathsf{A}, \mathsf{B} \in$ NrFss *be non-reentrant feature structures and* $\pi_1, \pi_2 \in$ Paths *be paths such that*

- $\pi_1 \in \Pi_B$,
- $\pi_1 \cdot \pi_2 \in \Pi_A$,
- $\Psi(\mathsf{A}, \pi_1 \cdot \pi_2) = \langle \mathsf{t}_1, \mathrm{F}_1, \ldots, \mathrm{F}_{|\pi_1|}, \mathsf{t}_{|\pi_1|+1}, \mathrm{F}_{|\pi_1|+1}, \ldots, \mathsf{t}_{|\pi_1 \cdot \pi_2|+1} \rangle$,
- $\Psi(\mathsf{B}, \pi_1) = \langle \mathsf{s}_1, \mathrm{F}_1, \ldots, \mathsf{s}_{|\pi_1|+1} \rangle$, *and*
- $\langle \mathrm{F}_{|\pi_1|+1} \rangle \notin \Pi_{s_{|\pi_1|+1}}$

then for all i, $1 \leq i \leq |\pi_1| + 1$, $\mathsf{s}_i \sqcup \mathsf{t}_i \neq \top$ *iff* $\mathsf{A} \sqcup \mathsf{B} \neq \top$.

Proof. Assume that for every $1 \leq i \leq |\pi_1| + 1$, $\mathsf{s}_i \sqcup \mathsf{t}_i \neq \top$. Since the prefixes of $\Psi(\mathsf{B}, \pi_1)$ and $\Psi(\mathsf{A}, \pi_1 \cdot \pi_2)$ are consistent up to $\mathrm{F}_{|\pi_1|+1}$ and the suffix of the cord $\Psi(\mathsf{A}, \pi_1 \cdot \pi_2)$ does not occur in $\Psi(\mathsf{B}, \pi_1)$, and hence does not contradict with B, the feature structures A and B are unifiable.

Assume that $\mathsf{A} \sqcup \mathsf{B} \neq \top$. Then all subtrees of the feature structures are consistent. Therefore, $\mathsf{s}_i \sqcup \mathsf{t}_i \neq \top$, for every $1 \leq i \leq |\pi_1| + 1$. ☐

Corollary 21. *Let $\Psi(\mathsf{A}, \pi_A)$, $\Psi(\mathsf{B}, \pi_B)$, $\Psi(\mathsf{C}, \pi_A)$ be cords, where pval(A, π_a) is not atomic. Let* G *be a feature such that:*

- $\langle \mathrm{G} \rangle \notin \Pi_{last(\Psi(\mathsf{A},\pi_A))}$ *and*
- $\langle \mathrm{G} \rangle \notin \Pi_{last(\Psi(\mathsf{C},\pi_A))}$

Consider the cord $\Psi(\mathsf{A}, \pi_A) \cdot \langle \mathrm{G} \rangle \cdot \Psi(\mathsf{B}, \pi_B)$ (by lemma 19, this is well defined) and write it as $\Psi(\mathsf{D}, \pi_A \cdot \langle \mathrm{G} \rangle \cdot \pi_B)$. Then $\mathsf{C} \sqcup \mathsf{A} \neq \top$ iff $\mathsf{C} \sqcup \mathsf{D} \neq \top$.

Theorem 22. *Let $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L} \rangle$ be a one-reentrant unification grammar and $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ be a derivation sequence. If $G^{li} = \langle V_N, V_t, V_s, \mathcal{R}^{li}, S \rangle =$ ug2lig(G^u) then there is a sequence of paths $\langle \pi_1, \ldots, \pi_n \rangle$, such that $S[\,] \overset{*}{\Rightarrow}_{li} N[\Psi(\mathsf{A}_1, \pi_1)] \ldots N[\Psi(\mathsf{A}_n, \pi_n)]$.*

Proof. We prove by induction on the length of the derivation sequence. The induction hypothesis is that if $\mathsf{A}^s \overset{k}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$, then there is a sequence of paths $\langle \pi_1, \ldots, \pi_n \rangle$, such that

$$S[\,] \overset{k+1}{\Rightarrow}_{li} N[\Psi(\mathsf{A}_1, \pi_1)] \ldots N[\Psi(\mathsf{A}_n, \pi_n)]$$

If $k = 0$, then

1. By the definition of derivation in UG, $\mathsf{A}^s \overset{0}{\Rightarrow}_u \mathsf{A}^s$;
2. By definition 25 case (1), the rule $S[\,] \rightarrow N[\Psi(\mathsf{A}^s, \varepsilon)]$ is in $\mathcal{R}^{li}$.
3. Hence, $S[\,] \overset{1}{\Rightarrow}_{li} N[\Psi(\mathsf{A}^s, \varepsilon)]$ and $N[\Psi(\mathsf{A}^s, \varepsilon)]$ is well defined since $\Psi(\mathsf{A}^s, \varepsilon) = \langle \mathsf{A}^s \rangle$, $|\mathsf{A}^s| \leq$ *maxHt*(G^u).

Assume that the hypothesis holds for every i, $0 \leq i < k$. Assume further that $\mathsf{A}^s \overset{k-1}{\Rightarrow}_u \mathsf{D}_1 \ldots \mathsf{D}_m \overset{1}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$.

1. By definition of UG derivation, there are an index j and a rule $\rho^u = \mathsf{C}_0 \rightarrow \mathsf{C}_1 \ldots \mathsf{C}_{n-m+1}$, $\rho^u \in \mathcal{R}^u$, such that ρ^u is applicable to D_j:

   $$(\langle \mathsf{C}_0, \ldots, \mathsf{C}_{n-m+1} \rangle, 0) \sqcup (\langle \mathsf{D}_1, \ldots, \mathsf{D}_m \rangle, j) = (\langle \mathsf{Q}_0, \ldots, \mathsf{Q}_{n-m+1} \rangle, \langle \mathsf{D}_1 \ldots \mathsf{D}_{j-1} \mathsf{Q}_0 \mathsf{D}_{j+1} \ldots \mathsf{D}_m \rangle)$$

   where

   - $\langle \mathsf{A}_1, \ldots, \mathsf{A}_{j-1} \rangle = \langle \mathsf{D}_1, \ldots, \mathsf{D}_{j-1} \rangle$
   - $\langle \mathsf{A}_j, \ldots, \mathsf{A}_{n-m+j} \rangle = \langle \mathsf{Q}_1, \ldots, \mathsf{Q}_{n-m+1} \rangle$
   - $\langle \mathsf{A}_{n-m+j+1}, \ldots, \mathsf{A}_n \rangle = \langle \mathsf{D}_{j+1}, \ldots, \mathsf{D}_m \rangle$

   Note that it is only possible to write the MRS $\langle \mathsf{A}_1, \ldots, \mathsf{A}_n \rangle$ in such a way due to the fact that the grammar G^u is one-reentrant: by lemma 2, no reentrancies can occur among two elements in a sentential form.

2. Hence, $\mathsf{A}^s \overset{k}{\Rightarrow}_u \mathsf{D}_1 \ldots \mathsf{D}_{j-1} \mathsf{Q}_1 \ldots \mathsf{Q}_{n-m+1} \mathsf{D}_{j+1} \ldots \mathsf{D}_m$
3. By the induction hypothesis there is a sequence of paths $\langle \nu_1, \ldots, \nu_m \rangle$ such that

   $$S[\,] \overset{k}{\Rightarrow}_{li} N[\Psi(\mathsf{D}_1, \nu_1)] \ldots N[\Psi(\mathsf{D}_m, \nu_m)]$$

4. We denote $\Psi(\mathsf{D}_j, \nu_j)$ as $\langle \mathsf{B}_1, \mathrm{F}_1, \ldots, \mathsf{B}_{|\nu_j|+1} \rangle$ (recall that j is the index of the selected element in the sentential form).

   We now want to show the existence of a rule $r \in \mathcal{R}^{li}$, created from ρ^u by the mapping *ug2lig*, which can be applied to j-th element of the LIG sentential form, $N[\Psi(\mathsf{D}_j, \nu_j)]$. We define the feature structure A to be a "bridge" between D_j and C_0 which together with a path π_0 (a prefix of the path ν_j) defines the head of the rule r.

5. Let π_0 be a maximal prefix of ν_j such that $\pi_0 \in \Pi_{C_0}$. Recall that $\langle \mathsf{B}_1, \mathrm{F}_1, \ldots, \mathsf{B}_{|\pi_0|+1} \rangle$ is a prefix of $\Psi(\mathsf{D}_j, \nu_j)$ because π_0 is a prefix of ν_j. Let A be such that $\Psi(\mathsf{A}, \pi_0) = \langle \mathsf{B}_1, \mathrm{F}_1, \ldots, \mathsf{B}_{|\pi_0|+1} \rangle$. By the induction hypothesis, $\mathsf{B}_i \leq \mathit{maxHt}(G^u)$, $1 \leq i \leq |\nu_j| + 1$. We will show that A is unifiable with both D_j and C_0.

6. We first show that $\mathsf{A} \in \Gamma(\mathsf{C}_0, \mathit{maxHt}(G^u))$. Since $\mathsf{D}_j \sqcup \mathsf{C}_0 \neq \top$ and A is a substructure of D_j we obtain that $\mathsf{A} \sqcup \mathsf{C}_0 \neq \top$. Since $\pi_0 \in \Pi_A$ and $|\mathsf{B}_i| \leq \mathit{maxHt}(G^u)$, $1 \leq i \leq |\nu_j| + 1$, $\mathsf{A} \in \Gamma(\mathsf{C}_0, \pi_0, \mathit{maxHt}(G^u))$.

7. We now show that there is a LIG rule r, a mapping of ρ^u, which is applicable to $N[\Psi(\mathsf{D}_j, \nu_j)]$. There are two possibilities for the relation between π_0 and ν_j (recall that π_0 is a prefix of ν_j):

   - If $\nu_j = \pi_0$ then $\mathsf{A} = \mathsf{D}_j$ and $\Psi(\mathsf{A}, \pi_0) = \Psi(\mathsf{D}_j, \nu_j)$. Hence, every rule of the form $N[\Psi(\mathsf{A}, \pi_0)] \rightarrow \alpha$ is applicable to $\Psi(\mathsf{D}_j, \nu_j)$. Since $\mathsf{A} \in \Gamma(\mathsf{C}_0, \pi_0, \mathit{maxHt}(G^u))$ we obtain that $N[\Psi(\mathsf{A}, \pi_0)] \in \textsc{ligHead}(\mathsf{C}_0)$. Hence, the rule $N[\Psi(\mathsf{A}, \pi_0)] \rightarrow \alpha$ is in $\mathcal{R}^{li}$, where $\alpha \in (V_N[V_s^*] \cup V_t)^*$ is determined by ρ^u.
   - If $\nu_j \neq \pi_0$ then $\nu_j = \pi_0 \cdot \langle \mathrm{F}_{|\pi_0|+1}, \ldots, \mathrm{F}_{|\nu_j|} \rangle$. Recall that $\mathit{pval}\mathsf{B}_{|\pi_0|+1} \langle \mathrm{F}_{|\pi_0|+1} \rangle \uparrow$ because $\Psi(\mathsf{D}_j, \nu_j) = \langle \mathsf{B}_1, \mathrm{F}_1, \ldots, \mathsf{B}_{|\nu_j|+1} \rangle$ and $|\pi_0| + 1 < |\nu_j| + 1$. Since $\Psi(\mathsf{A}, \pi_0) = \langle \mathsf{B}_1, \mathrm{F}_1, \ldots, \mathsf{B}_{|\pi_0|+1} \rangle$, we obtain that every rule of the form $N[\Psi(\mathsf{A}, \pi_0), \mathrm{F}_{|\pi_0|+1}\ \infty] \rightarrow \alpha$ is applicable to $N[\Psi(\mathsf{D}_j, \nu_j)]$. Since $\mathsf{A} \in \Gamma(\mathsf{C}_0, \pi_0, \mathit{maxHt}(G^u))$ we obtain that

   $$N[\Psi(\mathsf{A}, \pi_0), \mathrm{F}_{|\pi_0|+1}\ \infty] \in \textsc{ligHead}(\mathsf{C}_0)$$

   Hence, the rule $N[\Psi(\mathsf{A}, \pi_0), \mathrm{F}_{|\pi_0|+1}\ \infty] \rightarrow \alpha$ is in $\mathcal{R}^{li}$, where $\alpha \in (V_N[V_s^*] \cup V_t)^*$ is determined by ρ^u.

8. The LIG rule r whose existence was established in (7) is applied to $N[\Psi(\mathsf{D}_j, \nu_j)]$ as follows:

$$\begin{aligned} S[\,] &\overset{k}{\Rightarrow}_{li} N[\Psi(\mathsf{D}_1, \nu_1)] \ldots N[\Psi(\mathsf{D}_m, \nu_m)] \\ &\overset{1}{\Rightarrow}_{li} N[\Psi(\mathsf{D}_1, \nu_1)] \ldots N[\Psi(\mathsf{D}_{j-1}, \nu_{j-1})]\ Y_1 \ldots Y_{n-m+1}\ N[\Psi(\mathsf{D}_{j+1}, \nu_{j+1})] \ldots N[\Psi(\mathsf{D}_m, \nu_m)] \end{aligned}$$

9. We now investigate the possible outcomes of applying the rule r to the selected element of the sentential form. Let $r = X_0 \rightarrow \alpha$, where $\alpha \in (V_N[V_s^*] \cup V_t)^*$. To complete the proof we have to show that for some sequence of paths $\langle \pi_1, \ldots, \pi_{n-m+1} \rangle$,

   $$\langle Y_1, \ldots, Y_{n-m+1} \rangle = \langle N[\Psi(\mathsf{Q}_1, \pi_1)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \pi_{n-m+1})] \rangle$$

   where $\mathsf{Q}_1, \ldots \mathsf{Q}_{n-m+1}$ are determined by the unification grammar, see (1) above.

- Assume that ρ^u has no reentrancies. Hence, $\mathsf{Q}_i = \mathsf{C}_i$, $1 \leq i \leq n-m+1$. By definition 25 case (4), the LIG rule body is

$$\alpha = \langle N[\Psi(\mathsf{C}_1), \varepsilon)], \ldots, N[\Psi(\mathsf{C}_{n-m+1}], \varepsilon)]\rangle = \langle N[\Psi(\mathsf{Q}_1), \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}], \varepsilon)]\rangle$$

Since the rule r does not copy the stack, $\alpha = \langle Y_1, \ldots, Y_{n-m+1}\rangle$. Therefore,

$$\langle Y_1, \ldots, Y_{n-m+1}\rangle = \langle N[\Psi(\mathsf{Q}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \varepsilon)]\rangle$$

- Assume that $(0, \mu_0) \overset{\rho^u}{\rightsquigarrow} (e, \mu_e)$, where $1 \leq e \leq n$. Hence, $\mathsf{Q}_i = \mathsf{C}_i$ and $Y_i = N[\Psi(\mathsf{Q}_i, \varepsilon)]$ is well defined for all i, $i \neq e$. By definition 25 case (5), the LIG rule body is

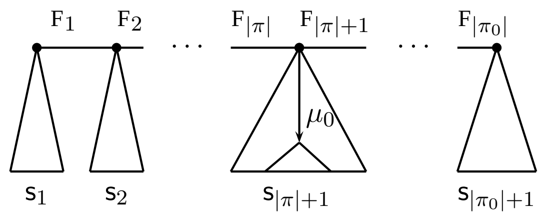
$$\begin{aligned} \alpha &= \langle N[\Psi(\mathsf{C}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{C}_{e-1}, \varepsilon)], X_e, N[\Psi(\mathsf{C}_{e+1}, \varepsilon)], \ldots, N[\Psi(\mathsf{C}_{n-m+1}, \varepsilon)]\rangle \\ &= \langle N[\Psi(\mathsf{Q}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_{e-1}, \varepsilon)], X_e, N[\Psi(\mathsf{Q}_{e+1}, \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \varepsilon)]\rangle \end{aligned}$$

This case is similar to the previous case, with the exception of X_e, which may be more complicated due to the propagation of the stack from X_0. We therefore focus on X_e (other elements of α are as above). Recall that by definition 25, $\langle \mathsf{P}_0, \ldots, \mathsf{P}_{n-m+1}\rangle$ is a sequence of feature structures such that

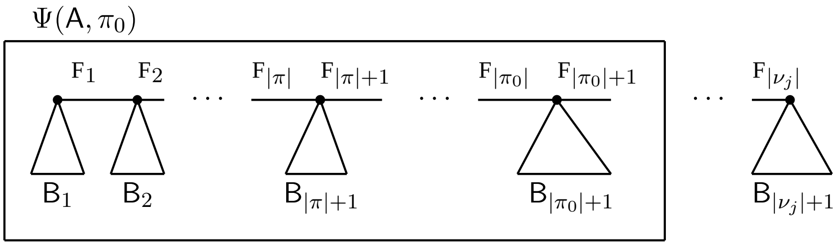
$$(\langle \mathsf{A}\rangle, 0) \sqcup (\rho^u, 0) = (\langle \mathsf{P}_0\rangle, \langle \mathsf{P}_0 \ldots \mathsf{P}_{n-m+1}\rangle)$$

We now analyze all the possible values of X_e, according to definition 25 case (5):

(a) Case 5a: if μ_0 is not a prefix of π_0 then by definition 25, $X_e = N[\Psi(\mathsf{P}_e, \mu_e)]$. Let π be the maximal prefix of π_0 and μ_0 such that $\mu_0 = \pi \cdot \mu_0'$. We denote $\Psi(\mathsf{C}_0, \pi_0)$ as $\langle \mathsf{s}_1, \mathrm{F}_1, \ldots, \mathsf{s}_{|\pi_0|+1}\rangle$, and graphically represent it as:

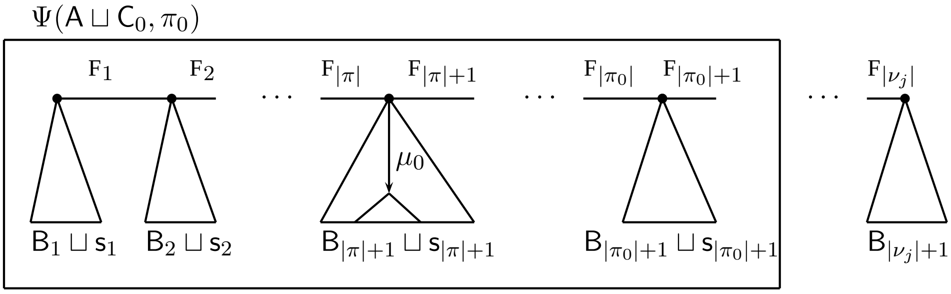


The cord $\Psi(\mathsf{D}_j, \nu_j)$ with its prefix $\Psi(\mathsf{A}, \pi_0)$ are represented as follows:



Note that the case $\pi_0 = \nu_j$ is just a special case of the figure above. The cord $\Psi(\mathsf{D}_j \sqcup \mathsf{C}_0, \nu_j)$

with its prefix $\Psi(\mathsf{A} \sqcup \mathsf{C}_0, \pi_0)$ are represented as follows:



Hence, $\mathit{pval}(\mathsf{A} \sqcup \mathsf{C}_0, \mu_0) = \mathit{pval}(\mathsf{B}_{|\pi|+1} \sqcup \mathsf{s}_{|\pi|+1}, \mu'_0) = \mathit{pval}(\mathsf{D}_j \sqcup \mathsf{C}_0, \mu_0)$. By definition of unification in context $\mathit{pval}(\mathsf{P}_e, \mu_e) = \mathit{pval}(\mathsf{A} \sqcup \mathsf{C}_0, \mu_0)$ and $\mathit{pval}(\mathsf{Q}_e, \mu_e) = \mathit{pval}(\mathsf{D}_j \sqcup \mathsf{C}_0, \mu_0)$. Hence, $\mathit{pval}(\mathsf{P}_e, \mu_e) = \mathit{pval}(\mathsf{Q}_e, \mu_e)$ and $\mathsf{Q}_e = \mathsf{P}_e$. Therefore,

$$\begin{aligned} \alpha = & \ \langle Y_1, \ldots, Y_{n-m+1} \rangle \\ = & \ \langle N[\Psi(\mathsf{Q}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_e, \mu_e)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \varepsilon)] \rangle \end{aligned}$$

(b) Case 5b: if μ_0 is a prefix of π_0, let $\pi_0 = \mu_0 \cdot \nu$, $\nu \in$ PATHS. Then by definition 25, the following holds:

- Case 5(b)i:
  If $X_0 = N[\Psi(\mathsf{A}, \pi_0)]$ then $X_e = N[\Psi(\mathsf{P}_e, \mu_e \cdot \nu)]$. Since $N[\Psi(\mathsf{A}, \pi_0)]$ is applicable to $N[\Psi(\mathsf{D}_j, \nu_j)]$ we obtain that $\pi_0 = \nu_j$ and $\mathsf{A} = \mathsf{D}_j$. Hence, $\mathsf{P}_e = \mathsf{Q}_e$. Therefore,

$$\begin{aligned} \alpha = & \ \langle Y_1, \ldots, Y_{n-m+1} \rangle \\ = & \ \langle N[\Psi(\mathsf{Q}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_e, \mu_e \cdot \nu)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \varepsilon)] \rangle \end{aligned}$$

- Case 5(b)ii:
  If $X_0 = N[\Psi(\mathsf{A}, \pi_0), \mathrm{F}_{|\pi_0|+1}\ \infty]$ then $X_e = N[\Psi(\mathsf{P}_e, \mu_e \cdot \nu), \mathrm{F}_{|\pi_0|+1}\ \infty]$. Let $\beta = \langle \mathsf{B}_{|\pi_0|+2}, \mathrm{F}_{|\pi_0|+2}, \ldots, \mathsf{B}_{|\nu_j|+1} \rangle$. By definition of A, $\Psi(\mathsf{D}_j, \nu_j) = \Psi(\mathsf{A}, \pi_0) \cdot \langle \mathrm{F}_{|\pi_0|+1} \rangle \cdot \beta$. We apply the LIG rule r to $N[\Psi(\mathsf{D}_j, \nu_j)]$ and obtain

$$\begin{aligned} & \langle Y_1, \ldots, Y_{n-m+1} \rangle \\ & = \langle N[\Psi(\mathsf{Q}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{P}_e, \mu_e \cdot \nu), \mathrm{F}_{|\pi_0|+1}, \beta], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \varepsilon)] \rangle \end{aligned}$$

  By definition of unification in context P_e differs from Q_e only in the value of the path $\mu_e \cdot \nu \cdot \langle \mathrm{F}_{|\pi_0|+1} \rangle$. The difference is in the value of the path $\mu_e \cdot \nu \cdot \langle \mathrm{F}_{|\pi_0|+1} \rangle$, it is not defined in P_e and equals β in Q_e. Hence, $\Psi(\mathsf{P}_e, \mu_e \cdot \nu) \cdot \langle \mathrm{F}_{|\pi_0|+1} \rangle \cdot \beta = \Psi(\mathsf{Q}_e, \mu_e \cdot \nu)$. Therefore,

$$\langle Y_1, \ldots, Y_{n-m+1} \rangle = \langle N[\Psi(\mathsf{Q}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_e, \nu_j)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \varepsilon)] \rangle$$

Note that in this case Y_e is well defined because it was created by applying a LIG rule to a well defined non-terminal symbol.

□

Theorem 23. *Let* $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L} \rangle$ *be a one-reentrant unification grammar and* $G^{li} = \langle V_N, V_t, V_s, \mathcal{R}^{li}, N \rangle = \mathit{ug2lig}(G^u)$ *be LIG. If* $S[\,] \stackrel{*}{\Rightarrow}_{li} Y_1 \ldots Y_n$*, where* $Y_i \in V_N[V_s^*]$*,* $1 \le i \le n$*, then there are a sequence of paths* $\langle \pi_1, \ldots, \pi_n \rangle$ *and a derivation sequence* $\mathsf{A}^s \stackrel{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$*, such that* $Y_i = N[\Psi(\mathsf{A}_i, \pi_i)]$.

Proof. We prove by induction on the length of the LIG derivation sequence. The induction hypothesis is that if $S[\,] \stackrel{k}{\Rightarrow}_{li} Y_1 \ldots Y_n$, where $Y_i \in V_N[V_s^*]$, $1 \leq i \leq n$, then $\mathsf{A}^s \stackrel{k-1}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$, such that for some sequence of paths $\langle \pi_1, \ldots, \pi_n \rangle$, $Y_i = N[\Psi(\mathsf{A}_i, \pi_i)]$, $1 \leq i \leq n$. If $k = 1$, then

1. By definition 25, the only rule that may be applied to the start symbol S in G^{li} is the rule defined by case (1) of the definition: $S[\,] \rightarrow N[\Psi(\mathsf{A}^s, \varepsilon)]$.

2. Hence, for $k = 1$, the only derivation sequence is $S[\,] \stackrel{1}{\Rightarrow}_{li} N[\Psi(\mathsf{A}^s, \varepsilon)]$

3. By the definition of derivation in UG, $\mathsf{A}^s \stackrel{0}{\Rightarrow}_u \mathsf{A}^s$.

Assume that the hypothesis holds for every i, $1 \leq i \leq k$; let the length of the derivation sequence be $k+1$.

1. Assume that $S[\,] \stackrel{k+1}{\Rightarrow}_{li} Y_1 \ldots Y_n$. Then $S[\,] \stackrel{k}{\Rightarrow}_{li} Y'_1 \ldots Y'_m \stackrel{1}{\Rightarrow}_{li} Y_1 \ldots Y_n$.

2. By the induction hypothesis, there exists a sequence of paths $\langle \nu_1, \ldots, \nu_m \rangle$ and feature structures $\mathsf{D}_1, \ldots, \mathsf{D}_m$, such that for $1 \leq i \leq m$, $Y'_i = N[\Psi(\mathsf{D}_i, \nu_i)]$, and $\mathsf{A}^s \stackrel{k-1}{\Rightarrow}_u \mathsf{D}_1 \ldots \mathsf{D}_m$. We therefore write:
$$S[\,] \stackrel{k}{\Rightarrow}_{li} N[\Psi(\mathsf{D}_1, \nu_1)] \ldots N[\Psi(\mathsf{D}_m, \nu_m)] \stackrel{1}{\Rightarrow}_{li} Y_1 \ldots Y_n$$

3. Furthermore, let $r = X_0 \rightarrow X_1 \ldots X_{n-m+1}$ be the G^{li} rule used for the last derivation step, and j be the index of the element to which r is applied, such that
$$N[\Psi(\mathsf{D}_1, \nu_1)] \ldots N[\Psi(\mathsf{D}_m, \nu_m)] \stackrel{1}{\Rightarrow}_{li}$$
$$N[\Psi(\mathsf{D}_1, \nu_1)] \ldots N[\Psi(\mathsf{D}_{j-1}, \nu_{j-1})] Y_j \ldots Y_{n-m+j} N[\Psi(\mathsf{D}_{n-m+j+1}, \nu_{n-m+j+1})] \ldots N[\Psi(\mathsf{D}_m, \nu_m)]$$

4. We denote $\Psi(\mathsf{D}_j, \nu_j)$ as $\langle \mathsf{t}_1, \mathrm{F}_1, \ldots, \mathsf{t}_{|\nu_j|+1} \rangle$

5. By definition 25, the rules that may be applied to $N[\Psi(\mathsf{D}_j, \nu_j)]$ are created by cases (4) and (5) of the definition, because the rule created by case (1) is headed by the non-terminal symbol S and the rules created by case (2) do not derive non-terminal symbols. Let $\rho^u = \mathsf{C}_0 \rightarrow \mathsf{C}_1 \ldots \mathsf{C}_{n-m+1}$ be a rule in $\mathcal{R}^u$ such that the rule r is created from ρ^u. Note that there may be more than one such rule.

6. We now show that $\mathsf{C}_0 \sqcup \mathsf{D}_j \neq \top$. In both cases (4) and (5) of definition 25 the head of the rule r, X_0, is a member of LIGHEAD(C_0). Since r is applicable to $N[\Psi(\mathsf{D}_j, \nu_j)]$ we obtain that X_0 has one of the following forms:

   (a) $X_0 = N[\Psi(\mathsf{D}_j, \nu_j)]$. By definition 24, $\Psi(\mathsf{D}_j, \nu_j) \in \mathrm{FH}(\mathsf{C}_0, \mathit{maxHt}(G^u))$. Since Ψ is a one-to-one mapping, we obtain that $\nu_j \in \Pi_{C_0}$ and $\mathsf{D}_j \in \Gamma(\mathsf{C}_0, \nu_j, \mathit{maxHt}(G^u))$. By definition of Γ, $\mathsf{D}_j \sqcup \mathsf{C}_0 \neq \top$.

   (b) $X_0 = N[\eta\, \infty]$, where η is a prefix of $\Psi(\mathsf{D}_j, \nu_j)$. Hence, we obtain that
   $$\eta = \langle \mathsf{t}_1, \mathrm{F}_1, \ldots, \mathsf{t}_{|\pi_0|+1}, \mathrm{F}_{|\pi_0|+1} \rangle$$
   where π_0 is a prefix of ν_j. By definition 24, $\eta \in \mathrm{UH}(\mathsf{C}_0, \mathit{maxHt}(G^u))$. Hence, there are a path $\pi_0 \in \Pi_{C_0}$ and a feature structure $\mathsf{A} \in \Gamma(\mathsf{C}_0, \pi_0, \mathit{maxHt}(G^u))$ such that $\eta = \Psi(\mathsf{A}, \pi_0) \cdot \langle \mathrm{F}_{|\pi_0|+1} \rangle$. By definition of Γ, $\mathsf{A} \sqcup \mathsf{C}_0 \neq \top$. Therefore, by corollary 21, $\mathsf{C}_0 \sqcup \mathsf{D}_j \neq \top$.

7. Since $\mathsf{C}_0 \sqcup \mathsf{D}_j \neq \top$, the rule ρ^u is applicable to D_j as follows:

$$\begin{array}{rl} \mathsf{A}^s \overset{k-1}{\Rightarrow}_u & \mathsf{D}_1 \ldots \mathsf{D}_m \\ \overset{1}{\Rightarrow}_u & \mathsf{D}_1 \ldots \mathsf{D}_{j-1}\mathsf{Q}_1 \ldots \mathsf{Q}_{n-m+1}\mathsf{D}_{n-m+j+1} \ldots \mathsf{D}_m \end{array}$$

where $\mathsf{Q}_1, \ldots, \mathsf{Q}_{n-m+1}$ are feature structures.

8. From (6) above, X_0 uniquely defines π_0, A and $\mathrm{F}_{|\pi_0|+1}$. We denote $\Psi(\mathsf{C}_0, \pi_0)$ as $\langle \mathsf{s}_1, \mathrm{F}_1, \ldots, \mathsf{s}_{|\pi_0|+1}\rangle$. Recall that for every $1 \leq i \leq |\pi_0| + 1$, $\mathsf{s}_i \sqcup \mathsf{t}_i \neq \top$ because $\mathsf{A} \in \Gamma(\mathsf{C}_0, \pi_0, \mathit{maxHt}(G^u))$. Let $\langle \mathsf{P}_0, \ldots, \mathsf{P}_{n-m+1}\rangle$ be the sequence of feature structures such that

$$(\langle \mathsf{A}\rangle, 0) \sqcup (\rho^u, 0) = (\langle \mathsf{P}_0\rangle, \langle \mathsf{P}_0, \ldots, \mathsf{P}_{n-m+1}\rangle)$$

9. Now we show that there is a sequence of paths $\langle \pi_1, \ldots, \pi_{n-m+1}\rangle$ such that

$$\langle Y_j, \ldots, Y_{n-m+j}\rangle = \langle N[\Psi(\mathsf{Q}_1, \pi_1)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \pi_{n-m+1})]\rangle$$

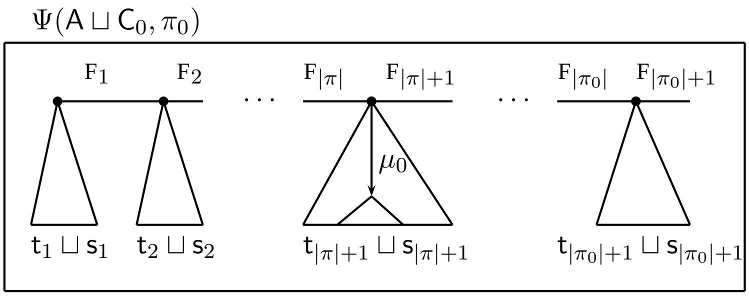
Without loss of generality, if ρ^u is reentrant we assume that its reentrant path is (e, μ_e), that is, $(0, \mu_0) \overset{\rho^u}{\leftrightsquigarrow} (e, \mu_e)$, where $1 \leq e \leq n$. By the definition of LIG there are two options for the rule r:

(a) The rule r does not copy the stack from the head to the body. Hence, $\langle X_1, \ldots, X_{n-m+1}\rangle = \langle Y_j, \ldots, Y_{n-m+j}\rangle$. Consider the possible sources of the rule r, according to definition 25:

- Case (4):
The rule ρ^u is non-reentrant. Hence, for $1 \leq i \leq n - m + 1$, $\mathsf{C}_i = \mathsf{Q}_i$ and $X_i = N[\Psi(\mathsf{C}_i, \varepsilon)] = N[\Psi(\mathsf{Q}_i, \varepsilon)]$. Therefore,

$$\langle Y_j, \ldots, Y_{n-m+j}\rangle = \langle N[\Psi(\mathsf{Q}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \varepsilon)]\rangle$$

- Case (5a):
If μ_0 is not a prefix of π_0 then for all i, $i \neq e$, $X_i = Y_{i+j-1} = N[\Psi(\mathsf{C}_i, \varepsilon)] = N[\Psi(\mathsf{Q}_i, \varepsilon)]$, and $X_e = N[\Psi(\mathsf{P}_e, \mu_e)]$. Let π be the maximal prefix of π_0 and μ_0 such that $\mu_0 = \pi \cdot \mu'_0$. The cord $\Psi(\mathsf{D}_j \sqcup \mathsf{C}_0, \nu_j)$ is graphically represented as:



Hence, $\mathit{pval}(\mathsf{A} \sqcup \mathsf{C}_0, \mu_0) = \mathit{pval}(\mathsf{t}_{|\pi|+1} \sqcup \mathsf{s}_{|\pi|+1}, \mu'_0) = \mathit{pval}(\mathsf{D}_j \sqcup \mathsf{C}_0, \mu_0)$. By definition of unification in context $\mathit{pval}(\mathsf{P}_e, \mu_e) = \mathit{pval}(\mathsf{A} \sqcup \mathsf{C}_0, \mu_0)$ and $\mathit{pval}(\mathsf{Q}_e, \mu_e) = \mathit{pval}(\mathsf{D}_j \sqcup \mathsf{C}_0, \mu_0)$. Hence, $\mathit{pval}(\mathsf{P}_e, \mu_e) = \mathit{pval}(\mathsf{Q}_e, \mu_e)$ and $\mathsf{Q}_e = \mathsf{P}_e$. Therefore, $X_e = N[\Psi(\mathsf{Q}_e, \mu_e)]$ and

$$\langle Y_j, \ldots, Y_{n-m+j}\rangle = \langle N[\Psi(\mathsf{Q}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_e, \mu_e)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \varepsilon)]\rangle$$

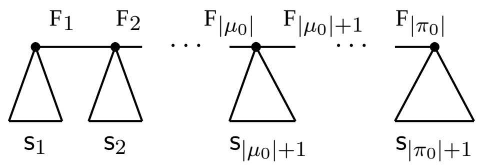
- Case (5(b)i):
  If $\pi_0 = \mu_0 \cdot \nu$, $\nu \in$ PATHS then $X_e = N[\Psi(\mathsf{P}_e, \mu_e \cdot \nu)]$. Since $N[\Psi(\mathsf{A}, \pi_0)]$ is applicable to $N[\Psi(\mathsf{D}_j, \nu_j)]$ we obtain that $\pi_0 = \nu_j$ and $\mathsf{A} = \mathsf{D}_j$. Hence $\mathsf{P}_e = \mathsf{Q}_e$. Therefore, $X_e = N[\Psi(\mathsf{Q}_e, \mu_e \cdot \nu)]$ and

$$\langle Y_j, \ldots, Y_{n-m+j} \rangle = \langle N[\Psi(\mathsf{Q}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_e, \mu_e \cdot \nu)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \varepsilon)] \rangle$$

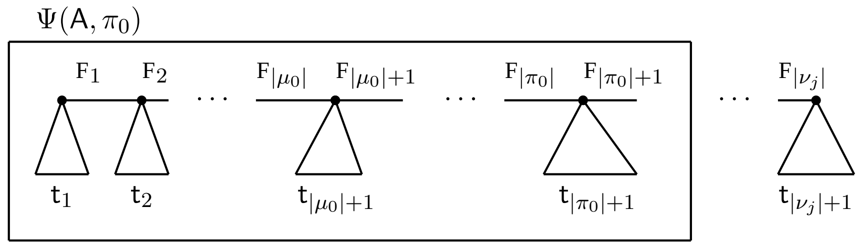
(b) The rule r copies the stack from X_0 to X_e. By definition 25, r is created from a reentrant unification rule, ρ^u, by case (5(b)ii) of the definition 25. Let $\nu_j = \pi_0 \cdot \nu'_j$ and $\pi_0 = \mu_0 \cdot \nu$, $\nu'_j, \nu \in$ PATHS. By the definition for all i, $i \neq e$, $X_i = Y_{i+j-1} = N[\Psi(\mathsf{C}_i, \varepsilon)] = N[\Psi(\mathsf{Q}_i, \varepsilon)]$ and $X_e = N[\Psi(\mathsf{P}_e, \mu_e \cdot \nu), \mathrm{F}_{|\pi_0|+1}\ \infty]$. Hence we just have to show that for some path $\pi_e \in$ PATHS, $Y_{j+e-1} = N[\Psi(\mathsf{Q}_e, \pi_e)]$. We will show that this equation holds for $\pi_e = \mu_e \cdot \nu \cdot \nu'_j$. Since π_0, A and $\mathrm{F}_{|\pi_0|+1}$ are uniquely defined by X_0 we obtain the following:

- $\Psi(\mathsf{C}_0, \pi_0) = \langle \mathsf{s}_1, \mathrm{F}_1, \ldots, \mathsf{s}_{|\pi_0|+1} \rangle$
- $\Psi(\mathsf{D}_j, \nu_j) = \langle \mathsf{t}_1, \mathrm{F}_1, \ldots, \mathsf{t}_{|\nu_j|+1} \rangle$
- $\Psi(\mathsf{D}_j \sqcup \mathsf{C}_0, \nu_j) = \Psi(\mathsf{Q}_0, \nu_j) = \langle \mathsf{s}_1 \sqcup \mathsf{t}_1, \mathrm{F}_1, \ldots, \mathsf{s}_{|\pi_0|+1} \sqcup \mathsf{t}_{|\pi_0|+1}, \mathrm{F}_{|\pi_0|+1}, \ldots, \mathsf{t}_{|\nu_j|+1} \rangle$
- $\Psi(\mathsf{A} \sqcup \mathsf{C}_0, \pi_0) = \Psi(\mathsf{P}_0, \pi_0) = \langle \mathsf{s}_1 \sqcup \mathsf{t}_1, \mathrm{F}_1, \ldots, \mathsf{s}_{|\pi_0|+1} \sqcup \mathsf{t}_{|\pi_0|+1} \rangle$
- $\Psi(\mathsf{P}_e, \mu_e \cdot \nu) = \mathit{butLast}(\Psi(\mathsf{C}_e, \mu_e)) \cdot \langle \mathsf{s}_{|\mu_0|+1} \sqcup \mathsf{t}_{|\mu_0|+1}, \mathrm{F}_{|\mu_0|+1}, \ldots, \mathsf{s}_{|\pi_0|+1} \sqcup \mathsf{t}_{|\pi_0|+1} \rangle$
- $\Psi(\mathsf{Q}_e, \pi_0 \cdot \nu'_j) =$
  $\mathit{butLast}(\Psi(\mathsf{C}_e, \mu_e)) \cdot \langle \mathsf{s}_{|\mu_0|+1} \sqcup \mathsf{t}_{|\mu_0|+1}, \mathrm{F}_{|\mu_0|+1}, \ldots, \mathsf{s}_{|\pi_0|+1} \sqcup \mathsf{t}_{|\pi_0|+1}, \mathrm{F}_{|\pi_0|+1}, \ldots, \mathsf{t}_{|\nu_j|+1} \rangle$

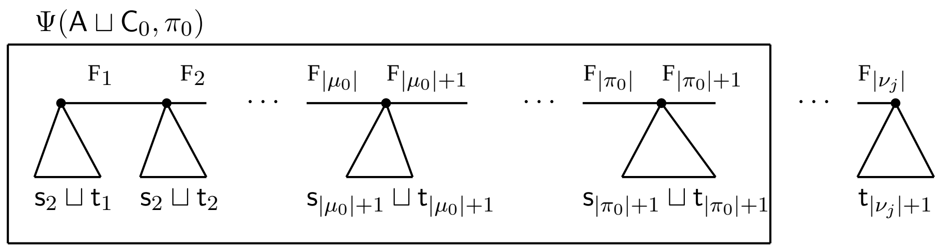
The cord $\Psi(\mathsf{C}_0, \pi_0) = \langle \mathsf{s}_1, \mathrm{F}_1, \ldots, \mathsf{s}_{|\pi_0|+1} \rangle$ is graphically represented as:



The cord $\Psi(\mathsf{D}_j, \nu_j) = \langle \mathsf{t}_1, \mathrm{F}_1, \ldots, \mathsf{t}_{|\nu_j|+1} \rangle$ whose prefix is the cord $\Psi(\mathsf{A}, \nu_j)$ is graphically represented as:



The cord $\Psi(\mathsf{D}_j \sqcup \mathsf{C}_0, \nu_j) = \Psi(\mathsf{Q}_0, \nu_j) = \langle \mathsf{s}_1 \sqcup \mathsf{t}_1, \mathrm{F}_1, \ldots, \mathsf{s}_{|\pi_0|+1} \sqcup \mathsf{t}_{|\pi_0|+1}, \mathrm{F}_{|\pi_0|+1}, \ldots, \mathsf{t}_{|\nu_j|+1} \rangle$ whose prefix is the cord $\Psi(\mathsf{A} \sqcup \mathsf{C}_0, \pi_0) = \Psi(\mathsf{P}_0, \pi_0) = \langle \mathsf{s}_1 \sqcup \mathsf{t}_1, \mathrm{F}_1, \ldots, \mathsf{s}_{|\pi_0|+1} \sqcup \mathsf{t}_{|\pi_0|+1} \rangle$ is graphically represented as:



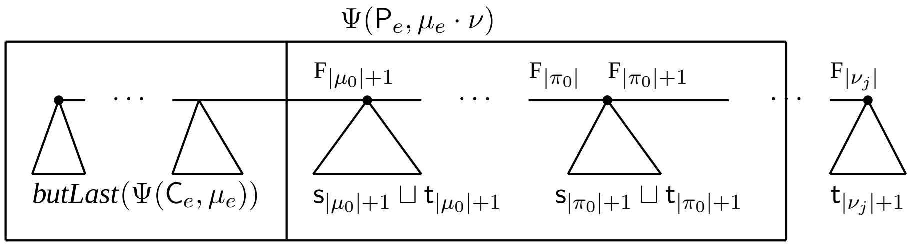
The relation between the cords

$$\Psi(\mathsf{P}_e, \mu_e \cdot \nu) = \textit{butLast}(\Psi(\mathsf{C}_e, \mu_e)) \cdot \langle \mathsf{s}_{|\mu_0|+1} \sqcup \mathsf{t}_{|\mu_0|+1}, \mathrm{F}_{|\mu_0|+1}, \ldots, \mathsf{s}_{|\pi_0|+1} \sqcup \mathsf{t}_{|\pi_0|+1} \rangle$$

and

$$\Psi(\mathsf{Q}_e, \pi_0 \cdot \nu'_j) =$$
$$\textit{butLast}(\Psi(\mathsf{C}_e, \mu_e)) \cdot \langle \mathsf{s}_{|\mu_0|+1} \sqcup \mathsf{t}_{|\mu_0|+1}, \mathrm{F}_{|\mu_0|+1}, \ldots, \mathsf{s}_{|\pi_0|+1} \sqcup \mathsf{t}_{|\pi_0|+1}, \mathrm{F}_{|\pi_0|+1}, \ldots, \mathsf{t}_{|\nu_j|+1} \rangle$$

is graphically represented as:



Hence,

$$\begin{aligned} Y_e =& N[\Psi(\mathsf{P}_e, \mu_e \cdot \nu), \mathrm{F}_{|\pi_0|+1}, \mathsf{t}_{|\pi_0|+2}, \mathrm{F}_{|\pi_0|+2}, \ldots, \mathsf{t}_{|\nu_j|+1}]] \\ =& N[\textit{butLast}(\Psi(\mathsf{C}_e, \mu_e)), \mathsf{s}_{|\mu_0|+1} \sqcup \mathsf{t}_{|\mu_0|+1}, \mathrm{F}_{|\mu_0|+1}, \ldots, \mathsf{s}_{|\pi_0|+1} \sqcup \mathsf{t}_{|\pi_0|+1}, \mathrm{F}_{|\pi_0|+1}, \ldots, \mathsf{t}_{|\nu_j|+1}] \\ =& N[\Psi(\mathsf{Q}_e, \pi_0 \cdot \nu'_j)] \end{aligned}$$

Therefore,

$$\langle Y_j, \ldots, Y_{n-m+j} \rangle = \langle N[\Psi(\mathsf{Q}_1, \varepsilon)], \ldots, N[\Psi(\mathsf{Q}_e, \pi_0 \cdot \nu'_j)], \ldots, N[\Psi(\mathsf{Q}_{n-m+1}, \varepsilon)] \rangle$$

□

Corollary 24 (Theorem 12)**.** *Let* $G^u \in \mathrm{UG}_{1r}$*, then* $L(G^u) = L(\textit{ug2lig}(G^u))$*.*

Proof. Let $G^u = \langle \mathcal{R}^u, \mathsf{A}^s, \mathcal{L} \rangle$ be a one-reentrant unification grammar and $G^{li} = \langle V_N, V_t, V_s, \mathcal{R}^{li}, N \rangle = \textit{ug2lig}(G^u)$. Then by theorem 22, there is a sequence of paths $\langle \pi_1, \ldots, \pi_n \rangle$ such that

$$\text{if } \mathsf{A}^s \overset{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n \text{ then } S[\,] \overset{*}{\Rightarrow}_{li} N[\Psi(\mathsf{A}_1, \pi_1)] \ldots N[\Psi(\mathsf{A}_n, \pi_n])$$

Where $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ is a pre-terminal sequence. Assume that $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n \overset{*}{\Rightarrow}_u w_1, \ldots w_n$, where $w_i \in$ WORDS, $1 \leq i \leq n$. Hence, $\mathcal{L}(w_i) = \{\mathsf{D}_i\}$ and $\mathsf{A}_i \sqcup \mathsf{D}_i \neq \top$. Since the grammar is a simplified unification grammar (definition 16), $\mathsf{A}_i = \mathsf{D}_i$. By definition 25 case (2), the rule $N[\Psi(\mathsf{A}_i, \pi_i)] \to w_i$ is in $\mathcal{R}^{li}$. Therefore, $S[\,] \overset{*}{\Rightarrow}_{li} N[\Psi(\mathsf{A}_1, \pi_1)] \ldots N[\Psi(\mathsf{A}_n, \pi_n]) \overset{*}{\Rightarrow}_{li} w_1, \ldots w_n$.

By theorem 23, if $S[\,] \overset{*}{\Rightarrow}_{li} Y_1 \ldots Y_n$ then there are a sequence of paths $\langle \pi_1, \ldots, \pi_n \rangle$, and a derivation sequence $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n$ such that for $0 \leq i \leq n$, $Y_i = N[\Psi(\mathsf{A}_i, \pi_i)]$. Assume that $S[\,] \overset{*}{\Rightarrow}_{li} N[\Psi(\mathsf{A}_1, \pi_1)] \ldots N[\Psi(\mathsf{A}_n, \pi_n]) \overset{*}{\Rightarrow}_{li} w_1, \ldots w_n$, $w_i \in V_t$. Then the rules $N[\Psi(\mathsf{A}_i, \pi_i)] \to w_i$ in $\mathcal{R}^{li}$, $1 \leq i \leq n$. By definition 25, each such rule is created from a lexicon entry $\mathcal{L}(w_i) = \{\mathsf{A}_i\}$. Hence, $\mathsf{A}^s \overset{*}{\Rightarrow}_u \mathsf{A}_1 \ldots \mathsf{A}_n \overset{*}{\Rightarrow}_u w_1, \ldots w_n$. □