



HAL
open science

Extended Lambek calculi and first-order linear logic

Richard Moot

► **To cite this version:**

Richard Moot. Extended Lambek calculi and first-order linear logic. Claudia Casadio, Bob Coecke, Michael Moortgat and Philip Scott. Categories and Types in Logic, Language and Physics: Essays Dedicated to Jim Lambek on the Occasion of His 90th Birthday, 8222, Springer, pp.297-330, 2014, Lecture Notes in Computer Science, 978-3-642-54788-1. hal-00826416

HAL Id: hal-00826416

<https://hal.science/hal-00826416>

Submitted on 27 May 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Extended Lambek calculi and first-order linear logic

Richard Moot

CNRS, LaBRI, University of Bordeaux
Richard.Moot@labri.fr

1 Introduction

The Syntactic Calculus (Lambek 1958) — often simply called the Lambek calculus, L , — is a beautiful system in many ways: Lambek grammars give a satisfactory syntactic analysis for the (context-free) core of natural language and, in addition, it provides a simple and elegant syntax-semantics interface.

However, since Lambek grammars generate only context-free languages (Pentus 1997), there are some well-known linguistic phenomena (Dutch verb clusters, at least if we want to get the semantics right (Huybregts 1984), Swiss-German verb clusters (Shieber 1985), etc.) which cannot be treated by Lambek grammars.

In addition, though the syntax-semantics interface works for many of the standard examples, the Lambek calculus does not allow a non-peripheral quantifier to take wide scope (as we would need for sentence (1) below if we want the existential quantifier to have wide scope, the so-called “de re” reading) or non-peripheral extraction (as illustrated by sentence (2) below); see (Moortgat 2011, Section 2.3) for discussion.

- (1) John believes someone left.
- (2) John picked up the package which Mary left yesterday.

To deal with these problems, several extensions of the Lambek calculus have been proposed. Though this is not the time and place to review them — I recommend (Moortgat 2010, Moortgat 2011) and the references cited therein for an up-to-date overview of the most prominent extensions; they include multi-modal categorial grammar (MMCG, Moortgat 1996), the Lambek-Grishin calculus (LG, Moortgat 2007) and the Displacement calculus (D, Morrill, Valentín & Fadda 2011) — I will begin by listing a number of properties which I consider desirable for such an extension. In essence, these desiderata are all ways of

keeping as many of good points of the Lambek calculus as possible while at the same time dealing with the inadequacies sketched above.¹

1. simple proof theory,
2. generate the mildly context-sensitive languages,
3. simple syntax-semantics interface giving a correct and simple account of medial scope for quantifiers and of medial extraction,
4. have a reasonable computational complexity.

None of these desiderata is absolute: there are matters of degree for each of them. First of all, it is often hard to distinguish familiarity from simplicity, but I think that having multiple equivalent proof systems for a single calculus is a sign that the calculus is a natural one: the Lambek calculus has a sequent calculus, natural deduction, proof nets, etc. and we would like its extensions to have as many of these as possible, each formulated in the simplest possible way.

The mildly context-sensitive languages (Joshi 1985) are a family of languages which extend the context-free language in a limited way, and opinions vary as to which of the members of this family is the most appropriate for the description of natural language. Throughout this article, I will only make the (rather conservative and uncontroversial) claim that any extension of the Lambek calculus should at least generate the tree adjoining languages, the multiple context-free languages (Seki, Matsumura, Fujii & Kasami 1991) (the well-nested MCFGs (Kanazawa 2009) are weakly equivalent to the tree adjoining languages) or the simple, positive range concatenation grammars (sRCG, weakly equivalent to MCFG, Boullier 1998).

With respect to the semantics, it generally takes the form of a simple homomorphism from proofs in the source logic to proofs in the Lambek-van Benthem calculus LP (which is multiplicative intuitionistic linear logic, MILL, for the linear logicians), though somewhat more elaborate continuation-based mappings (Bernardi & Moortgat 2007) have been used as well.

Finally, what counts as reasonable computational complexity is open to discussion as well: since theorem-proving for the Lambek calculus is NP complete (Pentus 2006), I will consider NP-complete to be “reasonable”, though polynomial parsing is generally considered a requirement for mildly context-sensitive formalisms (Joshi 1985). Since the complexity of the logic used corresponds to the *universal* recognition problem in formal language theory, NP completeness

¹ To the reader who is justifiably skeptical of any author who writes down a list of desiderata, followed by an argument by this same author arguing how well he scores on his own list, I say only that, in my opinion, this list is uncontroversial and at least implicitly shared by most of the work on extensions of the Lambek calculus and that the list still allows for a considerable debate as to *how well* each extension responds to each desideratum as well as discussion about the relative importance of the different items.

is not as bad as it may seem, since it corresponds to the complexity of the universal recognition problem for multiple context-free grammars (when we fix the maximum number of string tuples a non-terminal is allowed to have), which is a prototypical mildly context-sensitive formalism. Little is known on polynomial *fixed* recognition for extended Lambek calculi (though some partial results can be found in (Moot 2002, Moot 2008)). The fixed recognition problem for the Lambek calculus itself is known to be polynomial (Pentus 2010).

Calculus	Complexity	Languages	Scope	Extraction
L	NP complete	CFL	-	-
MMCG	PSPACE complete	CSL	+	+
LG	NP complete	\geq MCFL	+	-
D	NP complete	\geq MCFL	+	+
MILL1	NP complete	\geq MCFL	+	+

Table 1. The Lambek calculus and several of its variants/extensions, together with the complexity of the universal recognition problem, classes of languages generated and the appropriateness of the formalism for handling medial quantifier scope and medial extraction

Table 1 gives an overview of the Lambek calculus as well as several of its prominent extensions with respect to the complexity of the universal recognition problem, the class of languages generated and the facilities in the formalism for handling medial quantifier scope and medial extraction.

In this paper, I will present an alternative extension of the Lambek calculus: first-order multiplicative intuitionistic linear logic (MILL1) (Girard 1991, Moot & Piazza 2001). It generates the right class of languages (MCFG are a subset of the Horn clause fragment, as shown in Section 3.3), and embeds the simple Displacement calculus (D, as shown in Section 4 and 5). As can be seen in Table 1, it has the lowest complexity class among the different extensions, generates (at least) the right class of languages, but also handles medial scope and medial extraction in a very simple way (as shown already in (Moot & Piazza 2001)). In addition, as we will see in Section 2, MILL1 has a very simple proof theory, essentially a resource-conscious version of first-order logic, with a proof net calculus which is a simple extension of the proof nets of multiplicative linear logic (Danos & Regnier 1989, Girard 1991). Finally, the homomorphism from MILL1 to MILL for semantics consists simply of dropping the first-order quantifiers.

I will also look at the Displacement calculus from the perspective of MILL1 and give a translation of D into MILL1, indirectly solving two open problems from (Morrill 2011) by providing a proof net calculus for D and showing that D is NP-complete. In addition it is also worth mentioning briefly that the simpler proof theory of MILL1 (ie. proof nets) makes proving cut elimination for D very easy: as for the multiplicative case, cut elimination for MILL1 consists of simple,

local conversions only with only three distinct cases to verify (axiom, tensor/par and existential/universal).

The remainder of this paper is structured as follows. In the next section, I will briefly introduce MILL1 and its proof theory, including a novel correctness condition for first-order proof nets, which is a simple extension of the contraction criterion from Danos (1990). Section 3 will introduce the Displacement calculus, D, using a presentation of the calculus from (Morrill et al. 2011) which emphasizes the operations on string tuples and, equivalently, on string positions. Section 4 will present a translation from D to MILL1, with a correctness proof in Section 5. Section 6 will briefly mention some other possible applications of MILL1, which include agreement, non-associativity and island constraints and quantifier scope restrictions. Finally, I will reflect on the implications of the results in this paper, giving some interesting open problems.

2 MILL1

First-order multiplicative intuitionistic linear logic (MILL1) extends (multiplicative) intuitionistic linear logic with the first-order quantifiers \exists and \forall . The first-order multiplicative fragment shares many of the good properties of the propositional fragment: the decision problem is NP complete (Lincoln 1995) and it has a simple proof net calculus which is an extension of the proof net calculus for multiplicative linear logic.

Table 2 presents the natural deduction calculus for MILL1, which is without surprises, though readers familiar with intuitionistic logic should note that the $\otimes E$, $\multimap I$ and $\exists E$ rule discharge exactly one occurrence of each of the hypotheses with which it is coindexed.

I will present the proof net calculus in three steps, which also form a basic proof search procedure: for a given statement $\Gamma \vdash C$ (with C a formula and Γ a multiset of formulas) we form a *proof frame* by unfolding the formulas according to the logical links shown in the bottom two rows of Table 3, using the negative unfolding for the formulas in Γ and the positive unfolding for the formula C . We then connect the atomic formulas using the axiom link (shown on the top left of the table) until we have found a complete matching of the atomic formulas, forming a *proof structure*. Finally, we check if the resulting proof structure is a *proof net* (ie. we verify if $\Gamma \vdash C$ is derivable) by verifying it satisfies a correctness condition.

As is usual, I will use the following conventions, which will make formulating the proof net calculus simpler.

- dotted binary links are called *par* links, solid binary links are called *tensor* links,

$$\begin{array}{c}
[A]^i[B]^i \\
\vdots \\
\frac{A \otimes B}{C} \otimes E_i \quad \frac{A \quad B}{A \otimes B} \otimes I \\
\\
\frac{A \quad A \multimap B}{B} \multimap E \quad \frac{[A]^i}{A \multimap B} \multimap I \\
\\
\frac{\exists x.A \quad C}{C} \exists E_i^* \quad \frac{A[x := t]}{\exists x.A} \exists I \\
\\
\frac{\forall x.A}{A[x := t]} \forall E \quad \frac{A}{\forall x.A} \forall I^*
\end{array}$$

* no free occurrences of x in any of the free hypotheses

Table 2. Natural deduction rules for MILL1

- dotted unary links are called *universal* links, solid unary links are called *existential* links, the bound variables of these links are called *universally bound* and *existentially bound* respectively.
- each occurrence of a quantifier link uses a distinct bound variable,
- the variable of a positive \forall and a negative \exists link (ie. the universal links and universally quantified variables) are called its *eigenvariable*,
- following (Bellin & van de Wiele 1995), I require eigenvariables of existential links to be used *strictly*, meaning that replacing the eigenvariable throughout a proof with a special, unused constant will not result in a proof (in other words, we never unnecessarily instantiate an existentially quantified variable with the eigenvariable of a universal link).

The fact that par links and universal links are drawn with dotted lines is not a notational accident: one of the fundamental insights of focusing proofs and ludics (Andreoli 1992, Girard 2001) is that these two types of links naturally group together, as do the existential and tensor links, both drawn with solid lines. This property is also what makes the correctness proof of Section 5 work. When it is convenient to refer to the par and universal links together, I will call them *asynchronous* links, similarly I will refer to the existential and tensor links as *synchronous* links (following Andreoli (1992)).

In Table 3, the formulas drawn below the link are its conclusions (the axiom link, on the top left of the table, is the only multiple conclusion link, the cut

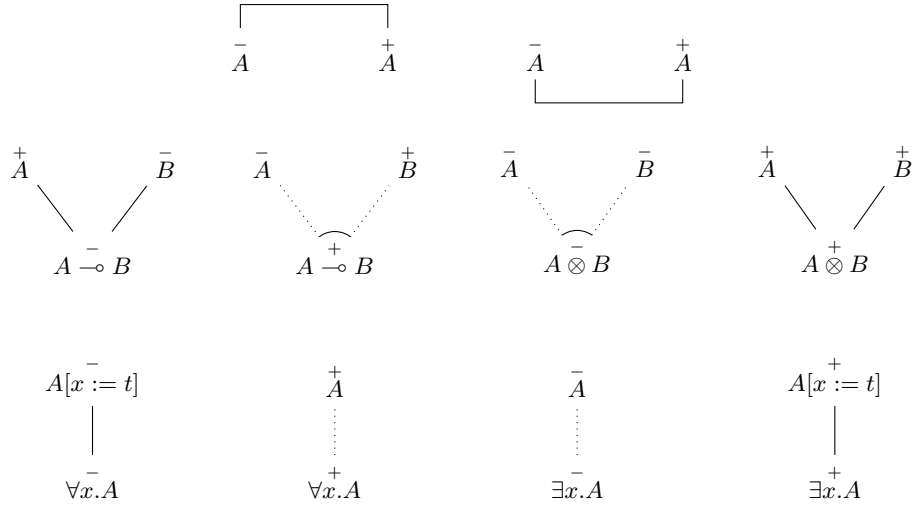


Table 3. Logical links for MILL1 proof structures

link, on the top right, does not have a conclusion, all logical links have a single conclusion), the formulas drawn above the link are its premisses.

Definition 1. A proof structure is a set of polarized formulas connected by instances of the links shown in Table 3 such that each formula is at most once the premiss of a link and exactly once the conclusion of a link. Formulas which are not the premiss of any link are called the conclusions of the proof structure. We say a proof structure with negative conclusions Γ and positive conclusions Δ is a proof structure of the statement $\Gamma \vdash \Delta$.

Definition 2. Given a proof structure Π a switching is

- for each of the par links a choice of one of its two premisses,
- for each of the universal links a choice either of a formula containing the eigenvariable of the link or of the premiss of the link.

Definition 3. Given a proof structure Π and a switching s we obtain a correction graph G by

- replacing each par link by an edge connecting the conclusion of the link to the premiss selected by s
- replacing each universal link by an edge connecting the conclusion of the link to the formula selected by s

Whereas a proof structure is a graph with some additional structure (paired edges, draw as connected dotted lines for the par links, and “universal” edges,

draw as dotted lines) a correction graph is a plain graph as used in graph theory: both types of special edges are replaced by normal edges according to the switching s .

Definition 4. *A proof structure is a proof net iff for all switchings s the corresponding correction graph G is acyclic and connected.*

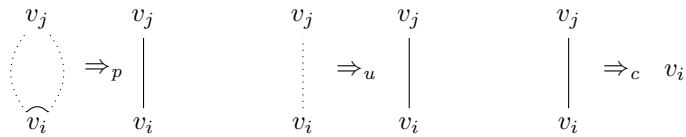
Remarkably, the proof nets correspond exactly to the provable statements in MILL1 (Girard 1991).

The basic idea of (Moot & Piazza 2001) is very simple: instead of using the well-known translation of Lambek calculus formulas into first-order logic (used for model-theory, see e.g. Došen 1992), we use this same translation to obtain formulas of first-order multiplicative *linear* logic. In this paper, I extend this result to the discontinuous Lambek calculus D, while at the same time sketching some novel applications of the system which correspond more closely to analyses in multimodal categorial grammars.

2.1 A Danos-style correctness condition

Though the correctness condition is conceptually simple, a proof structure has a number of correction graphs which is exponential in the number of asynchronous links, making the correctness condition hard to verify directly (though linear-time algorithms for checking the correctness condition exist in the quantifier-free case, eg. (Guerrini 1999, Murawski & Ong 2000)).

Here, I present an extension of the correctness condition of (Danos 1990) to the first-order case, which avoids this exponential complexity. Let G be a proof structure, where each vertex of the proof structure is assigned the set of eigenvariables which occur in the corresponding formula. Then we have the following contractions.



There is one contraction for the par links (p), one contraction for the universal links (u) and a final contraction which contracts components (connected subgraphs consisting only of synchronous, axiom and cut links) to a single vertex (c). The u contraction has the condition that there are no occurrences of the eigenvariable of the universal variable corresponding to the link outside of v_j . The c contraction has as condition that $i \neq j$; it contracts the vertex connecting i and j and the set of eigenvariables of v_i on the right hand side of the contraction corresponds to the set union of the eigenvariables of v_i and v_j on the left hand side of the contraction.

The following proposition is easy to prove using induction on the number of asynchronous links in the proof structure, using a variant of the “splitting par” sequentialization proof of Danos (1990)

Proposition 1. *A proof structure is a proof net iff it contracts to a single vertex using the contractions p , u and c .*

It is also easy to verify that the contractions are confluent, and can therefore be applied in any desired order.

To give an idea of how these contractions are applied, Figure 1 shows (on the left) a proof structure for the underivable statement $\forall x \exists y. f(x, y) \vdash \exists v \forall w. f(w, v)$. In the middle of the Figure, we see the proof structure with each formula replaced by the set of its free variables and before any contractions, with the eigenvariables shown next to their universal links. On the right, we see the structure after all c contractions have been applied. It is clear that we cannot apply the u contraction for y , since y occurs at a vertex other than the top vertex. Similarly, we cannot apply the u contraction for w either, meaning the proof structure is not contractible and therefore not a proof net.

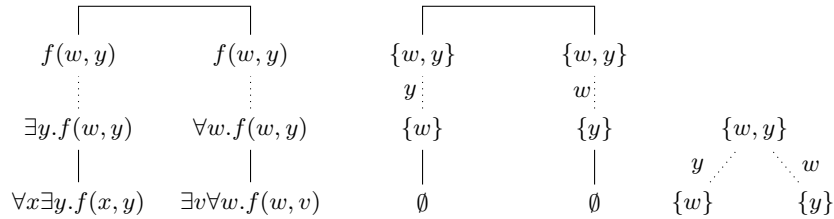


Fig. 1. Proof structure and partial contraction sequence for the underivable statement $\forall x \exists y. f(x, y) \vdash \exists v \forall w. f(w, v)$

Figure 2 shows the proof structure and part of the contraction sequence for the derivable statement $\exists x \forall y. f(x, y) \vdash \forall v \exists w. f(w, v)$. In this case, the structure on the right *does* allow us to perform the u contractions (in any order), producing a single vertex and thereby showing the proof structure is a proof net.

2.2 Eager Application of the Contractions

Though the contraction condition can be efficiently implemented, when verifying whether or not a given statement is a proof net it is often possible to disqualify partial proof structures (that is, proof structures where only some of the axiom links have been performed). Since the number of potential axiom links is enormous ($n!$ in the worst case), efficient methods for limiting the combinatorial

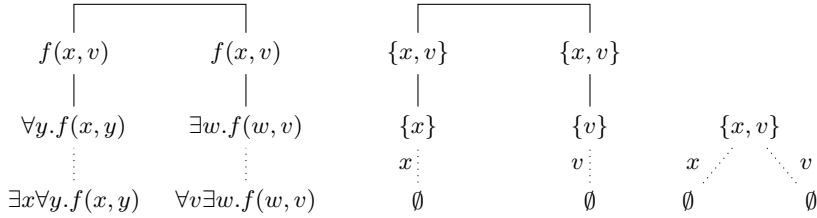


Fig. 2. Proof net and partial contraction sequence for the derivable statement $\exists x \forall y. f(x, y) \vdash \forall v \exists w. f(w, v)$

explosion as much as possible are a prerequisite for performing proof search on realistic examples.

The contractions allow us to give a compact representation of the search space by reducing the partial proof structure produced so far. When each vertex is assigned a multiset of literals (in addition to the set of eigenvariables already required for the contractions), the axiom rule corresponds to selecting, if necessary, while unifying the existentially quantified variables, two conjugate literals $+A$ and $-A$ from two different vertices (since the axiom rule corresponds to an application of the c contraction), identifying the two vertices and taking the multiset union of the remaining literals from the two vertices, in addition to taking the set union of the eigenvariables of the vertices. When the input consists of (curried) Horn clauses, each vertex will correspond to a Horn clause; therefore this partial proof structure approach generalizes resolution theorem proving. However, it allows for a lot of freedom in the strategy of literal selection, so we can apply “smart backtracking” strategies such as selecting the literal which has the smallest number of conjugates (Moot 2007). The contraction condition immediately suggest the following.

- never connect a literal to a descendant or an ancestor (generalizes “formulas from different vertices” for the Horn clause case); failure to respect this constraint will result in a cyclic proof structure,
- if the premiss of an asynchronous link is a leaf with the empty set of literals, then we must be able to contract it immediately ; failure to respect this constraint will result in a disconnected proof structure.
- similarly, if an isolated vertex which is not the only vertex in the graph has the empty set of literals, then the proof structure is disconnected.

3 The Displacement calculus

The Displacement calculus (Morrill et al. 2011) is an extension of the Lambek calculus using tuples of strings as their basic units.

3.1 String tuples

Whereas the Lambek calculus is the logic of strings, several formalisms are using *tuples* of strings as their basic units (eg. MCFGs, RCGs).

In what follows I use $s, s_0, s_1, \dots, s', s'', \dots$ to refer to *simple* strings (ie. the 1-tuples) with the constant ϵ for the empty string. The letters t, u, v etc. refer to i -tuples of strings for $i \geq 1$. I will write a i -tuple of strings as s_1, \dots, s_i , but also (if $i \geq 2$) as s_1, t or t', s_i where t is understood to be the string tuple s_2, \dots, s_i and t' the string tuple s_1, \dots, s_{i-1} , both $(i - 1)$ -tuples.

The basic operation for simple strings is concatenation. How does this operation extend to string tuples? For our current purposes, the natural extension of concatenation to string tuples is the following

$$(s_1, \dots, s_m) \circ (s'_1, \dots, s'_n) = s_1, \dots, s_m s'_1, \dots, s'_n$$

where $s_m s'_1$ is the string concatenation of the two simple strings s_m and s'_1 . In other words, the result of concatenating an m -tuple t and an n -tuple u is the $n + m - 1$ tuple obtained by first taking the first $m - 1$ elements of t , then the simple string concatenation of the last element of t with the first element of u and finally the last $n - 1$ elements of u . When both t and u are simple strings, then their concatenation is the string concatenation of their single element.² In what follows, I will simply write tu for the concatenation of two string tuples t and u and $u[t]$ to abbreviate $u_1 t u_2$.

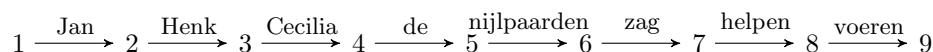
3.2 Position pairs

As is common in computational linguistics, it is sometimes more convenient to represent a simple string as a pair of string positions, the first element of the pair representing the leftmost string position and the second element its rightmost position. These positions are commonly represented as integers (to make the implicit linear precedence relation more easily visible). Likewise, we can represent an n -tuple of strings as a $2n$ tuple of string positions. This representation has the advantage that it makes string concatenation trivial: if x_0, x_1 is a string starting at position x_0 and ending at position x_1 and x_1, x_2 is a string starting at position x_1 and ending at position x_2 then the concatenation of these two strings is simply x_0, x_2 (this is the familiar difference list concatenation from Prolog (Pereira & Shieber 1987)).

Definition 5. *We say a grammar is simple in the input string if for each input string w_1, \dots, w_n we have that w_i spans positions $i, i + 1$*

² Another natural way to define concatenation is as point-wise concatenation of the different elements of two (equal-sized) tuples, as done by Stabler (2003).

Much of the work in parsing presupposes grammars are simple in the input string (Nederhof & Satta 2010), since it makes the definition of the standard parsing algorithms much neater. However, the original construction of Bar-Hillel, Perles & Shamir (1964) on which it is based is much more general: it computes the intersection of a context-free grammar and a finite-state automaton (FSA), where each non-terminal is assigned an input state and an output state of the FSA. For grammars which are not simple in the input string, this FSA can have self-loops and complex cycles, whereas the input string for a simple grammar is an FSA with a simple, deterministic linear path as shown in the example below. With the exception of Section 4.2, where I discusses the possibility of abandoning this constraint, the grammars I use will be simple in the input string. A simple example is shown below.



Suppose “nijlpaarden” (hippos) above is assigned the category n , for noun. Incorporating its string positions produces $n(5, 6)$. It gets more interesting with the determiner “de” (the): we assign it the formula $\forall x.n(5, x) \multimap np(4, x)$, which says that whenever it finds an n to its immediate right (starting at position 5 and ending at any position x) it will return an np from position 4 to this same x (this is the MILL1 translation of np/n at position 4, 5). In a chart parser, we would indicate this by adding an np arc from 4 to 6. There is an important difference with a standard chart parser though: since we are operating in a resource-conscious logic, we know that in a correct proof each rule is used exactly once (though their order is only partially determined).

Figure 3 shows the three elementary string operations of the Displacement calculus both in the form of operations of string tuples and in the form of operations of string position pairs.

Concatenation takes an i -tuple t (shown in the top of the figure as the white blocks, with corresponding string positions x_0, \dots, x_n for $n = 2i - 1$) and a j -tuple u (shown in the top of the figure as the gray blocks, with corresponding string positions x_n, \dots, x_{n+m} for $m = 2j - 1$) and the resulting concatenation tu (with the last element of t concatenated to the first element of u , indicated as the gray-white block x_{n-1}, x_{n+1} ; x_n is not a string position in the resulting $i + j - 1$ -tuple tu , which consists of the string positions $x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}$).

Left wrap takes an $i + 1$ -tuple s, t (with s a simple string and t an i -tuple, with string positions $x_0, x_1, x_n, \dots, x_{n+m}$) and a j -tuple u (with string positions x_1, \dots, x_n) and wraps s, t around u producing an $i + j - 1$ -tuple sut with string positions $x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}$, with positions x_1 and x_n removed because of the two concatenations.

Symmetrically, *right wrap* takes an $i + 1$ -tuple t, s (with s a simple string and t an i -tuple) and a j -tuple u and wraps t, s around u producing tus .

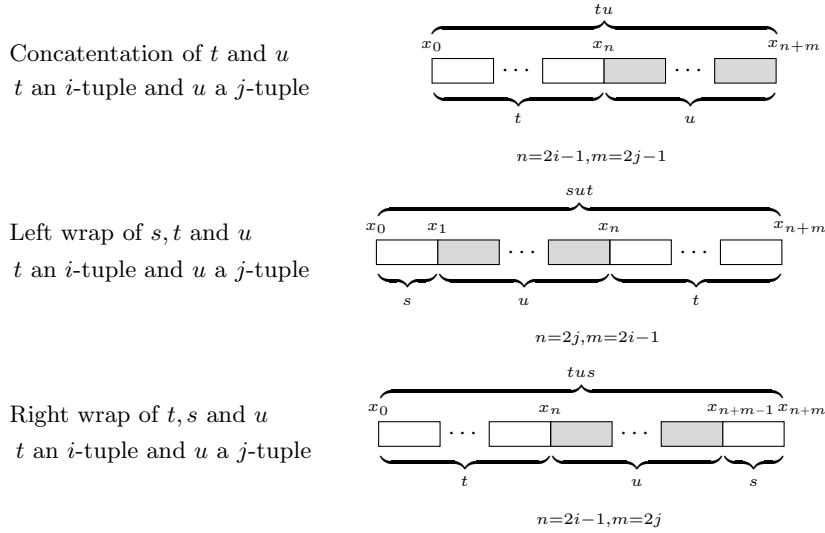


Fig. 3. String operations and their equivalent string positions operations

Given these operations, the proof rules for D are simple to state. I give a notational variant of the natural deduction calculus of (Morrill et al. 2011). As usual, natural deduction proofs start with a hypothesis $t : A$ (for $t : A$ a member of the lexicon of the grammar, in which case t is a constant from the lexicon, or for a hypothesis discharged by the product elimination and implication introduction rules). In each case the constant string t is unique in the proof.

For a given (sub-)proof, the *active* hypotheses are all hypotheses which have not been discharged by a product elimination or implication introduction rule in this (sub-)proof.

For the logical rules, we can see that the different families of connectives correspond to the three basic string tuple operations: with concatenation for $/$, \bullet and \backslash (the rules are shown in Figure 4 and Figure 7 with the corresponding string tuples), left wrap for $\uparrow_>$, $\odot_>$ and $\downarrow_>$ (shown in Figure 5 and Figure 8) and right wrap for $\uparrow_<$, $\odot_<$ and $\downarrow_<$ (shown in Figure 6 and Figure 9).

In the discontinuous Lambek calculus, we define the *sort* of a formula F , written $s(F)$ as the number of items in its string tuple minus 1. Given sorts for the atomic formulas, we compute the sort of a complex formula as shown in Table 4 (the distinction between the left wrap and right wrap connectives is irrelevant for the sorts).

$$\begin{array}{c}
\frac{t : A \quad u : A \setminus C}{tu : C} \setminus E \qquad \frac{[t : A]^i \quad \dots \quad tu : C}{u : A \setminus C} \setminus I_i \\
\frac{t : C / B \quad u : B}{tu : C} / E \qquad \frac{[u : B]^i \quad \dots \quad tu : C}{t : C / B} / I_i \\
\frac{t : A \bullet B \quad u [t_1 t_2] : C}{u[t] : C} \bullet E_i \quad \frac{[t_1 : A]^i \quad [t_2 : B]^i \quad \dots \quad u [t_1 t_2] : C}{tu : A \bullet B} \bullet I
\end{array}$$

Fig. 4. Proof rules – Lambek calculus

$$\begin{array}{c}
\frac{s, t : A \quad u : A \downarrow C}{sut : C} \downarrow E \qquad \frac{[s, t : A]^i \quad \dots \quad sut : C}{u : A \downarrow C} \downarrow I_i \\
\frac{s, t : C \uparrow B \quad u : B}{sut : C} \uparrow E \qquad \frac{[u : B]^i \quad \dots \quad sut : C}{s, t : C \uparrow B} \uparrow I_i \\
\frac{t : A \odot B \quad u [st_1 t_2] : C}{u[t] : C} \odot E_i \quad \frac{[s, t_2 : A]^i \quad [t_1 : B]^i \quad \dots \quad u [st_1 t_2] : C}{sut : A \odot B} \odot I
\end{array}$$

Fig. 5. Proof rules — leftmost infixation, extraction

$$\begin{aligned}
s(A \bullet B) &= s(A) + s(B) \\
s(A \setminus C) &= s(C) - s(A) \\
s(C / B) &= s(C) - s(B) \\
s(A \odot B) &= s(A) + s(B) - 1 \\
s(A \downarrow C) &= s(C) + 1 - s(A) \\
s(C \uparrow B) &= s(C) + 1 - s(B)
\end{aligned}$$

Table 4. Computing the sort of a complex formula given the sort of its immediate subformulas

$$\begin{array}{c}
\frac{t, s : A \quad u : A \downarrow_{<} C}{sut : C} \downarrow_{<} E \qquad \frac{[t, s : A]^i \quad \dots \quad tus : C}{u : A \downarrow_{<} C} \downarrow_{<} I_i \\
\frac{t, s : C \uparrow_{<} B \quad u : B}{sut : C} \uparrow_{<} E \qquad \frac{[u : B]^i \quad \dots \quad tus : C}{t, s : C \uparrow_{<} B} \uparrow_{<} I_i \\
\frac{[t_1, s : A]^i \quad \dots \quad [t_2 : B]^i \quad t : A \odot_{<} B \quad u[t_1 t_2 s] : C}{u[t] : C} \odot_{<} E_i \quad \frac{t, s : A \quad u : B}{tus : A \odot_{<} B} \odot_{<} I
\end{array}$$

Fig. 6. Proof rules — rightmost infixation, extraction

3.3 MILL1 and Multiple Context-Free Grammars

It is fairly easy to see that MILL1 generates (at least) the multiple context-free languages (or equivalently, the languages generated by simple, positive range concatenation grammars (Boullier 1998)) by using a lexicalized form of the grammars as defined below.

Definition 6. *A grammar is lexicalized if each grammar rule uses exactly one non-terminal symbol.*

Lexicalization is one of the principal differences between traditional phrase structure grammars and categorial grammars: categorial grammars generally require a form of lexicalization, whereas phrase structure grammars do not. The most well-known lexicalized form is the Greibach normal form for context-free grammars (Greibach 1965). Wijnholds (2011) shows that any (ϵ -free) simple, positive range concatenation grammar has a lexicalized grammar generating the same language. Since ranges are simply pairs of non-negative integers (see Section 3.2 and (Boullier 1998)) these translate directly to Horn clauses in MILL1. The following rules are therefore both a notational variant of a (lexicalized) MCFG/sRCG and an MILL1 lexicon (corresponding to the verbs of the example on page 11). See Appendix C for the construction of a proof net using (the curried versions of) these formulas.

$$\begin{array}{l}
\forall x_0 x_1 x_2 x_3. np(x_0, x_1) \otimes np(x_1, x_2) \otimes inf(x_2, 6, 7, x_3) \multimap s(x_0, x_3) \quad zag \\
\forall x_0 x_1 x_2 x_3. np(x_0, x_1) \otimes inf(x_1, x_2, 8, x_3) \multimap inf(x_0, x_2, 7, x_3) \quad helpen \\
\forall x_0 x_1. np(x_0, x_1) \multimap inf(x_0, x_1, 8, 9) \quad voeren
\end{array}$$

In Section 4.2, we will see how we can obtain these formulas via a translation of D.

4 Translations

The proof rules and the corresponding string tuple operations (shown in Figures 7, 8 and 9) suggest the translation shown in Table 5 of D formulas into MILL1 formulas. It is an extension of the translation of Lambek calculus formulas of (Moot & Piazza 2001), while at the same time extending the translation of (Morrill & Fadda 2008) for the simple displacement calculus.

$$\begin{array}{c}
 \overbrace{\underbrace{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}_{A \bullet B}} \\
 \underbrace{\phantom{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}}_A \quad \underbrace{\phantom{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}}_B \\
 n=2s(A)+1, m=2s(B)+1
 \end{array}
 \quad
 \frac{[t_1 : A]^i \quad [t_2 : B]^i}{u[t_1 t_2] : C} \bullet E_i \quad \frac{t : A \bullet B \quad u[t_1 t_2] : C}{u[t] : C} \bullet I$$

$$\begin{array}{c}
 \overbrace{\underbrace{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}_C} \\
 \underbrace{\phantom{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}}_{C/B} \quad \underbrace{\phantom{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}}_B \\
 n=2s(C/B)+1, m=2s(B)+1
 \end{array}
 \quad
 \frac{t : C / B \quad u : B}{tu : C} / E \quad \frac{[u : B]^i}{tu : C} / I_i$$

$$\begin{array}{c}
 \overbrace{\underbrace{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}_C} \\
 \underbrace{\phantom{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}}_A \quad \underbrace{\phantom{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}}_{A \setminus C} \\
 n=2s(A)+1, m=2s(A \setminus C)+1
 \end{array}
 \quad
 \frac{t : A \quad u : A \setminus C}{tu : C} \setminus E \quad \frac{[t : A]^i}{u : A \setminus C} \setminus I_i$$

Fig. 7. String positions – Lambek calculus

The reader intimidated by the number variable indices in Table 5 is invited to look at Figures 7, 8 and 9 for the correspondence between the string position numbers and the strings components of the different formulas in the translation. Section 4.1 will illustrate the translation using some examples, whereas Section 5 will make the correspondence the rules from the Displacement calculus more precise.

Note: the sequence x_i, \dots, x_i is of course simply the unit sequence x_i whereas the sequence x_i, \dots, x_{i-1} is the empty sequence.

If there are at most two string tuples, both $C \uparrow_{>} B$ (Equation 5 with $n = 2, m = 1$, remembering that $x_2, \dots, x_{n-1} \equiv x_2, \dots, x_1$ which is equivalent to the empty

$$\begin{array}{c}
\begin{array}{c}
\overbrace{x_0 \quad x_1 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}^{A \odot > B} \\
\begin{array}{c}
\boxed{x_0} \quad \boxed{x_1} \quad \dots \quad \boxed{x_n} \quad \dots \quad \boxed{x_{n+m}} \\
A \quad B \quad A \\
n=2s(B)+2, m=2s(A)-1
\end{array}
\end{array} \\
\begin{array}{c}
\overbrace{x_0 \quad x_1 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}^C \\
\begin{array}{c}
\boxed{x_0} \quad \boxed{x_1} \quad \dots \quad \boxed{x_n} \quad \dots \quad \boxed{x_{n+m}} \\
C \uparrow > B \quad B \quad C \uparrow > B \\
n=2s(B)+2, m=2s(C \uparrow > B)-1
\end{array}
\end{array} \\
\begin{array}{c}
\overbrace{x_0 \quad x_1 \quad \dots \quad x_n \quad \dots \quad x_{n+m}}^C \\
\begin{array}{c}
\boxed{x_0} \quad \boxed{x_1} \quad \dots \quad \boxed{x_n} \quad \dots \quad \boxed{x_{n+m}} \\
A \quad A \downarrow > C \quad A \\
n=2s(A \downarrow > C)+2, m=2s(A)-1
\end{array}
\end{array}
\end{array}
\quad
\begin{array}{c}
\begin{array}{c}
[s, t_2 : A]^i \quad [t_1 : B]^i \\
\vdots \\
\frac{t : A \odot > B \quad u[st_1t_2] : C}{u[t] : C} \odot > E_i \quad \frac{s, t : A \quad u : B}{sut : A \odot > B} \odot > I
\end{array} \\
\begin{array}{c}
\frac{s, t : C \uparrow > B \quad u : B}{sut : C} \uparrow > E \quad \frac{[u : B]^i \quad \dots \quad sut : C}{s, t : C \uparrow > B} \uparrow > I_i
\end{array} \\
\begin{array}{c}
\frac{s, t : A \quad u : A \downarrow > C}{sut : C} \downarrow > E \quad \frac{[s, t : A]^i \quad \dots \quad sut : C}{u : A \downarrow > C} \downarrow > I_i
\end{array}
\end{array}$$

Fig. 8. String positions – leftmost infix/extraction

$$\begin{array}{c}
\begin{array}{c}
\overbrace{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m-1} \quad x_{n+m}}^{A \odot < B} \\
\begin{array}{c}
\boxed{x_0} \quad \dots \quad \boxed{x_n} \quad \dots \quad \boxed{x_{n+m-1}} \quad \boxed{x_{n+m}} \\
A \quad B \quad A \\
n=2s(A)-1, m=2s(B)+2
\end{array}
\end{array} \\
\begin{array}{c}
\overbrace{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m-1} \quad x_{n+m}}^C \\
\begin{array}{c}
\boxed{x_0} \quad \dots \quad \boxed{x_n} \quad \dots \quad \boxed{x_{n+m-1}} \quad \boxed{x_{n+m}} \\
C \uparrow < B \quad B \quad C \uparrow < B \\
n=2s(C \uparrow < B)-1, m=2s(B)+2
\end{array}
\end{array} \\
\begin{array}{c}
\overbrace{x_0 \quad \dots \quad x_n \quad \dots \quad x_{n+m-1} \quad x_{n+m}}^C \\
\begin{array}{c}
\boxed{x_0} \quad \dots \quad \boxed{x_n} \quad \dots \quad \boxed{x_{n+m-1}} \quad \boxed{x_{n+m}} \\
A \quad A \downarrow < C \quad A \\
n=2s(A)-1, m=2s(A \downarrow < C)+2
\end{array}
\end{array}
\end{array}
\quad
\begin{array}{c}
\begin{array}{c}
[t_1, s : A]^i \quad [t_2 : B]^i \\
\vdots \\
\frac{t : A \odot < B \quad u[t_1t_2s] : C}{u[t] : C} \odot < E_i \quad \frac{t, s : A \quad u : B}{tus : A \odot < B} \odot < I
\end{array} \\
\begin{array}{c}
\frac{t, s : C \uparrow < B \quad u : B}{sut : C} \uparrow < E \quad \frac{[u : B]^i \quad \dots \quad tus : C}{t, s : C \uparrow < B} \uparrow < I_i
\end{array} \\
\begin{array}{c}
\frac{t, s : A \quad u : A \downarrow < C}{sut : C} \downarrow < E \quad \frac{[t, s : A]^i \quad \dots \quad tus : C}{u : A \downarrow < C} \downarrow < I_i
\end{array}
\end{array}$$

Fig. 9. String positions – rightmost infix/extraction

sequence of string positions and the empty sequence of quantifier prefixes, and that $x_{n+1}, \dots, x_{n+m} \equiv x_3, \dots, x_3 \equiv x_3$) and $C \uparrow < B$ (Equation 8 with $n = 1$, $m = 2$) translate to the following

$$\begin{aligned}
(1) \quad & \left. \begin{aligned} & \|A \bullet B\|^{x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}} \\ & = \exists x_n \|A\|^{x_0, \dots, x_n} \otimes \|B\|^{x_n, \dots, x_{n+m}} \end{aligned} \right\} n=2s(A)+1, m=2s(B)+1 \\
(2) \quad & \left. \begin{aligned} & \|C / B\|^{x_0, \dots, x_n} \\ & = \forall x_{n+1}, \dots, x_{n+m} \|B\|^{x_n, \dots, x_{n+m}} \multimap \\ & \quad \|C\|^{x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}} \end{aligned} \right\} n=2s(C/B)+1, m=2s(B)+1 \\
(3) \quad & \left. \begin{aligned} & \|A \setminus C\|^{x_n, \dots, x_{n+m}} \\ & = \forall x_0, \dots, x_{n-1} \|A\|^{x_0, \dots, x_n} \multimap \\ & \quad \|C\|^{x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}} \end{aligned} \right\} n=2s(A)+1, m=2s(A \setminus C)+1 \\
(4) \quad & \left. \begin{aligned} & \|A \odot > B\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}} \\ & = \exists x_1, x_n \|A\|^{x_0, x_1, x_n, \dots, x_{n+m}} \otimes \|B\|^{x_1, \dots, x_n} \end{aligned} \right\} n=2s(B)+2, m=2s(A)-1 \\
(5) \quad & \left. \begin{aligned} & \|C \uparrow > B\|^{x_0, x_1, x_n, \dots, x_{n+m}} \\ & = \forall x_2, \dots, x_{n-1} \|B\|^{x_1, \dots, x_n} \multimap \\ & \quad \|C\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}} \end{aligned} \right\} n=2s(B)+2, m=2s(C \uparrow > B)-1 \\
(6) \quad & \left. \begin{aligned} & \|A \downarrow > C\|^{x_1, \dots, x_n} \\ & = \forall x_0, x_{n+1}, \dots, x_{n+m} \|A\|^{x_0, x_1, x_n, \dots, x_{n+m}} \multimap \\ & \quad \|C\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}} \end{aligned} \right\} n=2s(A \downarrow > C)+2, m=2s(A)-1 \\
(7) \quad & \left. \begin{aligned} & \|A \odot < B\|^{x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m-2}, x_{n+m}} \\ & = \exists x_n, x_{n+m-1} \|A\|^{x_0, \dots, x_n, x_{n+m-1}, x_{n+m}} \otimes \|B\|^{x_n, \dots, x_{n+m-1}} \end{aligned} \right\} n=2s(A)-1, m=2s(B)+2 \\
(8) \quad & \left. \begin{aligned} & \|C \uparrow < B\|^{x_0, \dots, x_n, x_{n+m-1}, x_{n+m}} \\ & = \forall x_{n+1}, \dots, x_{n+m-2} \|B\|^{x_n, \dots, x_{n+m-1}} \multimap \\ & \quad \|C\|^{x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m-2}, x_{n+m}} \end{aligned} \right\} n=2s(C \uparrow < B)-1, m=2s(B)+2 \\
(9) \quad & \left. \begin{aligned} & \|A \downarrow < C\|^{x_n, \dots, x_{n+m-1}} \\ & = \forall x_0, \dots, x_{n-1}, x_{n+m} \|A\|^{x_0, \dots, x_n, x_{n+m-1}, x_{n+m}} \multimap \\ & \quad \|C\|^{x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m-2}, x_{n+m}} \end{aligned} \right\} n=2s(A)-1, m=2s(A \downarrow < C)+2
\end{aligned}$$

Table 5. Translation of D formulas to MILL1 formulas

$$\|C \uparrow B\|^{x_0, x_1, x_2, x_3} = \|B\|^{x_1, x_2} \multimap \|C\|^{x_0, x_3}$$

Similarly, it is easy to verify that both $A \downarrow_{>} C$ (Equation 6 with $n = 2$, $m = 1$, remember that $x_2, \dots, x_{n-1} \equiv x_2, \dots, x_1$ and therefore equal to the empty sequence and that $x_{n+1}, \dots, x_{n+m} \equiv x_3, \dots, x_3 \equiv x_3$) and $A \downarrow_{<} C$ (Equation 9 with $n = 1$, $m = 2$) produce the following translation for D formulas with at most two string tuples.

$$\|A \downarrow C\|^{x_1, x_2} = \forall x_0, x_3 \|A\|^{x_0, x_1, x_2, x_3} \multimap \|C\|^{x_0, x_3}$$

In the Lambek calculus, all sorts are zero, therefore instantiating Equation 3 with $n=1$, $m=1$ produces the following

$$\|A \setminus C\|^{x_1, x_2} = \forall x_0 \|A\|^{x_0, x_1} \multimap \|C\|^{x_0, x_2}$$

and therefore has the translation of (Moot & Piazza 2001) as a special case.

4.1 Examples

As an illustration, let's look at the formula unfolding of $((vp \uparrow vp)/vp) \setminus (vp \uparrow vp)$, which is the formula for “did” assigned to sentences like

(3) John slept before Mary did.

by (Morrill et al. 2011). This lexical entry for “did” is of sort 0 and therefore has two string positions (I use 4 and 5) to start off its translation. However, since both direct subformulas are of sort 1 its these subformulas have four position variables each. Applying the translation for \setminus shown in Equation 3 with $n = 3$ ($= 2s((vp \uparrow vp)/vp) + 1$), $m = 1$ (the sort of the complete formula being 0) gives us the following partial translation.

$$\forall x_0 x_1 x_2 \|((vp \uparrow vp)/vp)\|^{x_0, x_1, x_2, 4} \multimap \|vp \uparrow vp\|^{x_0, x_1, x_2, 5}$$

I first translate the leftmost subformula, which is of sort 1, and apply the $/$ rule (Equation 2) with $n = 3$ ($= 2s((vp \uparrow vp)/vp) + 1$) and $m = 1$ ($= 2s(vp) + 1$) giving the following partial translation.

$$\forall x_0 x_1 x_2 [\forall x_3 [\|vp\|^{4, x_3} \multimap \|vp \uparrow vp\|^{x_0, x_1, x_2, x_3}] \multimap \|vp \uparrow vp\|^{x_0, x_1, x_2, 5}]$$

Applying the translation rule for $C \uparrow B$ (Equation 5) twice produces.

$$\forall x_0 x_1 x_2 [\forall x_3 [\|vp\|^{4, x_3} \multimap \|vp\|^{x_1, x_2} \multimap \|vp\|^{x_0, x_3}] \multimap \|vp\|^{x_1, x_2} \multimap \|vp\|^{x_0, 5}]$$

Appendix A gives a step-by-step derivation of the proof net for sentence (3) using the contractions of Section 2.1 (see also Appendix B for a slightly abbreviated alternative proof net derivation). The intelligent backtracking solution of Section 2.2 and (Moot 2007) guarantee that at each step of the computation we can make a deterministic choice for literal selection, though the reader is invited to try and find a proof by hand to convince himself that this is by no means a trivial example!

As a slightly more complicated example translation, which depends on the distinction between left wrap and right wrap, Morrill et al. (2011) give the following formula for an object reflexive.

$$((vp \uparrow_{>} np) \uparrow_{<} np) \downarrow_{<} (vp \uparrow_{>} np)$$

Translating the $\downarrow_{<}$ connective, with input positions 3 and 4 using Equation 9 with $n = 3$ (since $s((vp \uparrow_{>} np) \uparrow_{<} np) = 2$) and $m = 2$ gives the following partial translation.

$$\forall x_0, x_1, x_2, x_5 \|(vp \uparrow_{>} np) \uparrow_{<} np\|^{x_0, x_1, x_2, 3, 4, x_5} \multimap \|(vp \uparrow_{>} np)\|^{x_0, x_1, x_2, x_5}$$

Translating the $\uparrow_{<}$ connective using Equation 8 with $n = 3$ and $m = 2$ gives.

$$\forall x_0, x_1, x_2, x_5 [\|np\|^{3, 4} \multimap \|(vp \uparrow_{>} np)\|^{x_0, x_1, x_2, x_5}] \multimap \|(vp \uparrow_{>} np)\|^{x_0, x_1, x_2, x_5}$$

Finally, unfolding the two $\uparrow_{>}$ connectives (using Equation 5) gives.

$$\forall x_0, x_1, x_2, x_5 [np(3, 4) \multimap np(x_1, x_2) \multimap \|(vp)\|^{x_0, x_5}] \multimap np(x_1, x_2) \multimap \|(vp)\|^{x_0, x_5}$$

Indicating that an object reflexive takes a ditransitive verb (with an object argument spanning the positions of the reflexive) to produce a transitive verb, which corresponds to the intuitive meaning of the lexical entry.

4.2 Synthetic connectives

Morrill et al. (2011) introduce the synthetic connectives — though note that from the point of view of MILL1, *all* D-connectives a synthetic MILL1-connectives, see Section 5. These connectives can be seen as abbreviations of combinations of a connective and an identity element (I denoting the empty string ϵ and J denoting ϵ, ϵ) :

$\tilde{A} =_{def} A \uparrow I$	Split
$\hat{A} =_{def} A \odot I$	Bridge
$\triangleright^{-1}A =_{def} J \setminus A$	Right projection
$\triangleright A =_{def} J \bullet A$	Right injection
$\triangleleft^{-1}A =_{def} A / J$	Left projection
$\triangleleft A =_{def} A \bullet J$	Left injection

Figures 10 and 11 show the proof rules for leftmost bridge/split and right projection/injection (the proof rules for left projection and injection as well as the proof rules for rightmost bridge and split are symmetric).

$$\begin{array}{c}
\frac{s, t : \tilde{A}}{st : A} \tilde{E} \quad \frac{st : A}{s, t : \tilde{A}} \tilde{I} \\
\\
\frac{t : \hat{A} \quad \begin{array}{c} u[st'] : C \\ \vdots \\ u[t] : C \end{array}}{u[t] : C} \hat{E} \quad \frac{s, t : A}{st : \hat{A}} \hat{I}
\end{array}$$

Fig. 10. Proof rules — leftmost split, bridge

$$\begin{array}{c}
\frac{t : \triangleright^{-1}A}{\epsilon, t : A} \triangleright^{-1}E \quad \frac{t : A}{\epsilon, t : \triangleright^{-1}A} \triangleright^{-1}I \\
\\
\frac{v : \triangleright A \quad \begin{array}{c} u, tu' : C \\ \vdots \\ uvu' : C \end{array}}{uvu' : C} \triangleright E \quad \frac{t : A}{\epsilon, t : \triangleright A} \triangleright I
\end{array}$$

Fig. 11. Proof rules — right projection, injection

The synthetic connectives are translated as follows (only the leftmost split and wedge are shown, the rightmost versions are symmetric in the variables):

$$\begin{aligned}
(10) \quad & \| \check{A} \|^{x_0, x_1, x_1, x_2, \dots, x_n} = \| A \|^{x_0, x_2, \dots, x_n} \\
(11) \quad & \| \wedge A \|^{x_0, x_2, \dots, x_n} = \exists x_1. \| A \|^{x_0, x_1, x_1, x_2, \dots, x_n} \\
(12) \quad & \| \triangleright A \|^{x_0, x_0, x_1, \dots, x_n} = \| A \|^{x_1, \dots, x_n} \\
(13) \quad & \| \triangleright^{-1} A \|^{x_1, \dots, x_n} = \forall x_0. \| A \|^{x_0, x_0, x_1, \dots, x_n} \\
(14) \quad & \| \triangleleft A \|^{x_0, \dots, x_n, x_{n+1}, x_{n+1}} = \| A \|^{x_0, \dots, x_n} \\
(15) \quad & \| \triangleleft^{-1} A \|^{x_0, \dots, x_n} = \forall x_{n+1}. \| A \|^{x_0, \dots, x_n, x_{n+1}, x_{n+1}}
\end{aligned}$$

In (Morrill et al. 2011), the bridge connective appears exclusively in (positive) contexts $\check{\cdot}(A \uparrow B)$ where it translates as.

$$\begin{aligned}
\| \check{\cdot}(A \uparrow B) \|^{x_0, x_2} &= \exists x_1. \| A \uparrow B \|^{x_0, x_1, x_1, x_2} \\
&= \exists x_1. [\| B \|^{x_1, x_1} \multimap \| A \|^{x_0, x_2}]
\end{aligned}$$

The resulting formula indicates that it takes a B argument spanning the empty string (anywhere) to produce an A covering the original string position x_0 and x_2 . Intuitively, this formalizes (in positive contexts) an A constituent with a B trace. The final translation is positive subformula of the extraction type used in (Moot & Piazza 2001).

The split connective ($\check{\cdot}$, but also \triangleleft and \triangleright) is more delicate, since it identifies string position variables. This can force the identification of variables, which means that direct application of the translation above can produce formulas which have “vacuous” quantifications, though this is not harmful (and these are easily removed in a post-processing step if desired). However, this identification of variables means that the grammars are no longer necessarily simple in the input string as discussed in Section 3.2. As an example, unfolding the formula below (which is patterned after the formula for “unfortunately” from (Morrill et al. 2011)) with input variables x_i and x_j forces us to identify x_i and x_j as shown below, hence producing a self-loop in the input FSA.

$$\begin{aligned}
\| \check{A} \downarrow B \|^{x_i, x_i} &= \forall x_0, x_2. \| \check{A} \|^{x_0, x_i, x_i, x_2} \multimap \| B \|^{x_0, x_2} \\
&= \forall x_0, x_2. \| A \|^{x_0, x_2} \multimap \| B \|^{x_0, x_2}
\end{aligned}$$

Intuitively, this translation indicates that a formula of the form $\check{A} \downarrow B$ takes its A argument at any span of the string and produces a B at the same position, with the complete formula spanning the empty string. It is, in essence, a translation of the (commutative) linear logic or LP implication into our current context. The MIX language can easily be generated using this property (Morrill & Valentín 2010).

It is easy to produce a complex formulas which, together with its argument, produce a complex cycle. The following formula spans the empty string after it combines with its C argument.

$$\begin{aligned}
\|(\check{A} \downarrow B)/C\|^{x_i, x_j} &= \forall x_1 \|C\|^{x_j, x_1} \multimap \| \check{A} \downarrow B \|^{x_i, x_1} \\
&= \forall x_1 [\|C\|^{x_j, x_1} \multimap \forall x_0, x_2. [\| \check{A} \|^{x_0, x_i, x_1, x_2} \multimap \|B\|^{x_0, x_2}]] \\
&= \forall x_1 [\|C\|^{x_j, x_i} \multimap \forall x_0, x_2. [\|A\|^{x_0, x_2} \multimap \|B\|^{x_0, x_2}]] \\
&= \|C\|^{x_j, x_i} \multimap \forall x_0, x_2. [\|A\|^{x_0, x_2} \multimap \|B\|^{x_0, x_2}]
\end{aligned}$$

The final line in the equation simply removes the x_1 quantifier. Since there are no longer any occurrences of the x_1 variable in the rest of the formula, this produces the equivalent formula shown. The translation specifies that the preposition, which spans positions x_i to x_j takes an np argument spanning positions x_j to x_i , ie. the rightmost position of the np argument is the leftmost position of the preposition.

If we want a displacement grammar to be simple in the input string, we can restrict the synthetic connectives used for its lexical entries to $\hat{\ } , \triangleright^{-1}$ and \triangleleft^{-1} ; in addition, no formulas contain the units I and J except where these occurrences are instances of the allowed synthetic connectives.³ The only proposed lexical entries which are not simple in this sense are those of the MIX grammar and the type for “supposedly” discussed above.

The \triangleright^{-1} and \triangleleft^{-1} connectives, together with atomic formulas of sort greater than 0, allow us to encode MCFG-style analyses, as we have seen them in Section 3.3, into D. As an example, let’s look at the unfolding of “lezen” which is assigned formula $\triangleright^{-1} np \setminus (np \setminus si)$ and assume “lezen” occupies the string position 4,5.

$$\forall x_2 \|np \setminus (np \setminus si)\|^{x_2, x_2, 4, 5}$$

Given that $s(np) = 0$ this reduces further to.

$$\forall x_2 \forall x_1 \|np\|^{x_1, x_2} \multimap \|np \setminus si\|^{x_1, x_2, 4, 5}$$

If we combine this entry with “boeken” from positions 1,2 (ie. the formula $np(1, 2)$, instantiating x_1 to 1 and x_2 to 2, this gives the following partial translation for “boeken lezen”

³ Alternatively, we can allow the $\check{\ } , \triangleright$ and \triangleleft connectives but restrict them to cases where there is strict identity of the two string positions (disallowing instantiation of variables to obtain identity). Note that this means that formulas of the form $\check{A} \downarrow B$ are valid only in contexts spanning the empty string.

$$\|np \setminus si\|^{1,2,4,5}$$

Similarly, “kunnen” with formula $\triangleright^{-1}(np \setminus si) \downarrow (np \setminus si)$ reduces as follows when occupying string position 3,4.

$$\forall x_2 \|(np \setminus si) \downarrow (np \setminus si)\|^{x_2, x_2, 3, 4}$$

Which unfolds further as.

$$\forall x_2 \forall x_0 \forall x_5 \|np \setminus si\|^{x_0, x_2, 4, x_5} \multimap \|np \setminus si\|^{x_0, x_2, 3, x_5}$$

This combines with the previous translation of “boeken lezen”, instantiating x_2 to 2, x_0 to 1 and x_5 to 5, giving “boeken kunnen lezen” with translation $\|np \setminus si\|^{1,2,3,5}$.

Finally, the tensed verb “wil” with formula $(np \setminus si) \downarrow (np \setminus s)$ unfolds at position 2,3 as.

$$\forall x_0 \forall x_3 \|np \setminus si\|^{x_0, 2, 3, x_3} \multimap \|np \setminus s\|^{x_0, x_3}$$

Instantiating x_0 to 1 and x_3 to 5 and combining this with the previously computed translation of “boeken kunnen lezen” produces $\|np \setminus s\|^{1,5}$ for “boeken wil kunnen lezen”. Appendix C shows a proof net derivation of the slightly more complex “(dat) Jan Henk Cecilia de nijlpaarden zag helpen voeren”. Note that the axiom linkings are again fully deterministic.

5 Correctness of the translation

The basic idea of the correctness proof is again very simple: we use the property of focused proof search and of ludics that combinations of synchronous connectives can always be seen as instances of a synthetic synchronous connective, whereas the same holds for the asynchronous connectives. Since the translations either use a combination of \exists and \otimes (both synchronous) or a combination of \forall and \multimap (both asynchronous), it follows immediately that we can treat these combinations as synthetic connectives, giving a rule to (synthetic) rule translation.

Lemma 1. *For every proof of $t_1 : A_1, \dots, t_n : A_n \vdash t : C$ in D , there is a proof of its translation in $MILL1$*

Proof Refer back to Figure 8 to see the correspondence between pairs of string positions and tuples of strings more clearly. The rules are simply the translation of the natural deduction rules of D, where the string tuples have been replaced by pairs of string positions.

For the case of $\setminus E$ we are in the following situation (let $i = \frac{1}{2}(n-1)$, $j = \frac{1}{2}(m-1)$, then x_0, \dots, x_n corresponds to a i -tuple t , x_n, \dots, x_{n+m} to a j -tuple u and $x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}$ to their concatenation tu). The translations of A and $A \setminus C$ share point x_n and we can instantiate the universally quantified variables of the other points of A (x_0 to x_{n-1}) applying the $\forall E$ rule n times ($/$ is symmetric).

$$\frac{\frac{\frac{\|A \setminus C\|^{x_n, \dots, x_{n+m}}}{\forall y_0, \dots, y_{n-1} \|A\|^{y_0, \dots, y_{n-1}, x_n} \multimap \|C\|^{y_0, \dots, y_{n-1}, x_{n+1}, \dots, x_{n+m}}} =_{def}}{\|A\|^{x_0, \dots, x_n} \multimap \|C\|^{x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}}} \forall E (n \text{ times})}{\|C\|^{x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}}} \multimap E$$

For the introduction rule, we again set i to $\frac{1}{2}(n-1)$ and j to $\frac{1}{2}(m-1)$, making x_0, \dots, x_n corresponds to a i -tuple t , x_n, \dots, x_{n+m} to a j -tuple u and $x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}$ to their concatenation tu . In this case, induction hypothesis gives us a MILL1 proof corresponding to $\Gamma, t : A \vdash tu : C$. To extend this proof to a MILL1 proof corresponding to $\Gamma \vdash u : A \setminus C$ ($/$ is again symmetric). which can be extended to a proof for $\Gamma, t : A \vdash tu : C$ as follows.

$$\frac{\frac{\frac{\|A\|^{x_0, \dots, x_n}]^i \dots \Gamma}{\vdots} \|C\|^{x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}} \multimap I_i}{\|A\|^{x_0, \dots, x_n} \multimap \|C\|^{x_0, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}}} \forall I (n \text{ times})}{\|A \setminus C\|^{x_n, \dots, x_{n+m}}} =_{def} \forall E (n \text{ times})$$

The cases for $\uparrow_>$ are shown below ($\downarrow_>$ is easily verified).

$$\frac{\frac{\frac{\|C \uparrow_> B\|^{x_0, x_1, x_n, \dots, x_{n+m}}}{\forall y_2, \dots, y_{n-1} \|B\|^{x_1, y_2, \dots, y_{n-1}, x_n} \multimap \|C\|^{x_0, y_2, \dots, y_{n-1}, x_{n+1}, \dots, x_{n+m}}} =_{def}}{\|B\|^{x_1, \dots, x_n} \multimap \|C\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}}} \forall E (n-2 \text{ times})}{\|C\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}}} \multimap E$$

$$\frac{\frac{\frac{[[B]^{x_1, \dots, x_n}]^i \dots \Gamma}{\vdots}}{\|C\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}} \multimap I_i}}{\|B\|^{x_1, \dots, x_n} \multimap \|C\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}}} \forall I \ (n-2 \text{ times})}{\frac{\forall x_2, \dots, x_{n-1} \|B\|^{x_1, \dots, x_n} \multimap \|C\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}}}{\|C \uparrow_{>} B\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}}} =_{def}}$$

Finally, the cases for $\odot_{>}$ are as follows.

$$\frac{\frac{\|A \odot_{>} B\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}}{\exists x_1 \exists x_n \|A\|^{x_0, \dots, x_n} \otimes \|B\|^{x_n, \dots, x_{n+m}}} =_{def} \frac{\frac{[[A]^{x_0, x_1, x_n, \dots, x_{n+m}} \otimes \|B\|^{x_1, \dots, x_n}]^i}{C} \otimes E_j}{\frac{[[A]^{x_0, x_1, x_n, \dots, x_{n+m}}]^j \quad [[B]^{x_1, \dots, x_n}]^i}{\vdots} C \otimes E_j}}{C} \exists E_i \text{ twice}}$$

$$\frac{\frac{\frac{\frac{\|A\|^{x_0, x_1, x_n, \dots, x_{n+m}} \quad \|B\|^{x_1, \dots, x_n}}{\|A\|^{x_0, x_1, x_n, \dots, x_{n+m}} \otimes \|B\|^{x_1, \dots, x_n}} \otimes I}{\exists x_n \|A\|^{x_0, x_1, x_n, \dots, x_{n+m}} \otimes \|B\|^{x_1, \dots, x_n}} \exists I}{\exists x_1 \exists x_n \|A\|^{x_0, x_1, x_n, \dots, x_{n+m}} \otimes \|B\|^{x_1, \dots, x_n}} \exists I}{\|A \odot_{>} B\|^{x_0, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_{n+m}}} =_{def}}$$

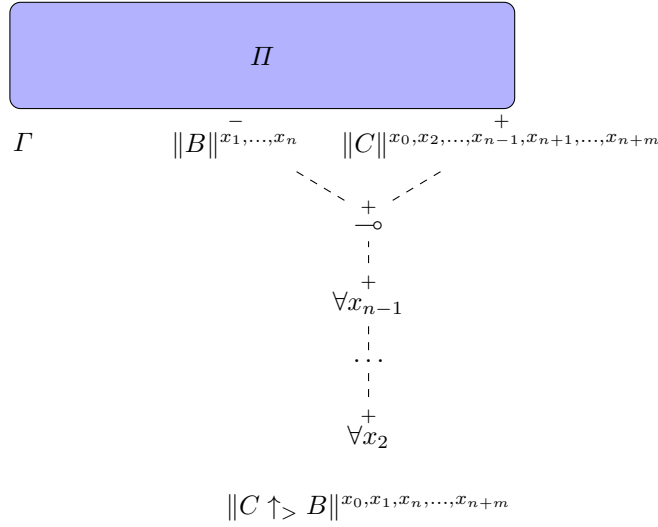
□

Lemma 2. *If the translation of a D sequent $t_1 : A_1, \dots, t_n : A_n \vdash t : C$ is provable, then there is a D proof of $t_1 : A_1, \dots, t_n : A_n \vdash t : C$.*

Proof This is most easily shown using proof nets, using induction on the number of links while removing them in groups of synchronous or asynchronous links corresponding to a D connective.

If there are terminal asynchronous links, then we proceed by case analysis knowing that we are dealing the result of the translation of D formulas.

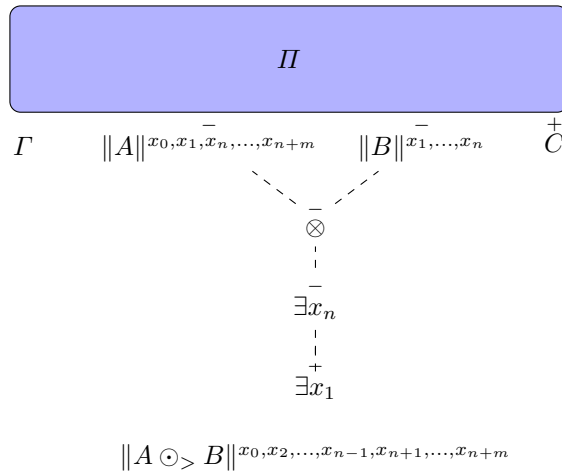
The case for $C \uparrow_{>} B$ looks as follows.



Given that removing the portrayed links produces a proof net Π of $\Gamma, B \vdash C$, we can apply the induction hypothesis, which gives a proof δ of $\Gamma, u : B \vdash sut : C$, which we can extend as follows.

$$\frac{\Gamma \quad u : B \quad \vdots \quad \delta \quad \frac{C : sut}{s, t : C \uparrow_{>} B} \uparrow_{>} I}{s, t : C \uparrow_{>} B} \uparrow_{>} I$$

Similarly, the par case for $\odot_{>}$ looks as follows.

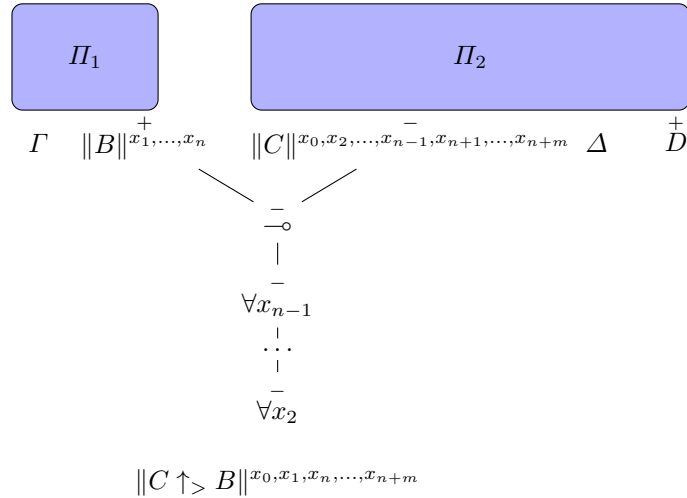


Again, we know by induction hypothesis that there is a proof δ of $\Gamma, s, t : A, u : B \vdash v[sut] : C$ and we need to show that there is a proof of $\Gamma, sut : A \odot_{>} B \vdash v[sut] : C$, which we do as follows.

$$\frac{\Gamma \quad [s, t : A]^i \quad [u : B]^i \quad \vdots \quad \delta \quad sut : A \odot_{>} B \quad v[sut] : C}{v[sut] : C} \odot_{>} E^i$$

Suppose there are no terminal asynchronous links, then we know there must be a group of splitting synchronous links corresponding to a D connective (a series of universal links ended by a tensor link which splits the proof net into two subnets, though the synthetic connectives of Section 4.2 allow for a single universal link, which is splitting by definition, since after removal of the link, all premisses of the link are the conclusion of disjoint subnets), using the standard splitting tensor argument (Bellin & van de Wiele 1995).

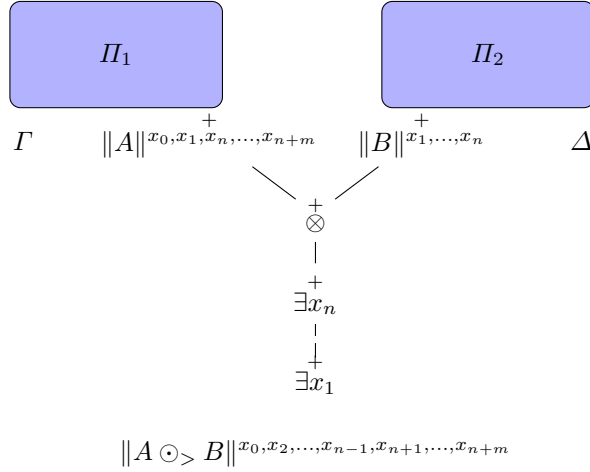
Suppose this group of splitting links is the translation of $\uparrow_{>}$, then the proof net is of the following form. Note that the translation of B corresponds to the string tuple u (with $i = \frac{1}{2}n$ components), the translation of $C \uparrow_{>} B$ to the string tuple sut and the translation of C to the string tuple s, t .



Therefore, we know by induction hypothesis that there is a proof δ_1 of $\Gamma \vdash u : B$ and a proof δ_2 of $\Delta, sut : C \vdash D$. We need to show that there is a proof $\Gamma, \Delta, s, t : C \uparrow_{>} B \vdash D$, which we can do as follows.

$$\Delta \frac{\frac{\Gamma \vdots \delta_1 \quad u : B \quad s, t : C \uparrow_{>} B}{sut : C} \uparrow_{>} E}{\vdots \delta_2 \quad D}$$

In case the hereditary splitting tensor link is the translation of a $\odot_{>}$ formula, we are in the following case.



Induction hypothesis gives us a proof δ_1 of $\Gamma \vdash s, t : A$ and a proof δ_2 of $\Delta \vdash u : B$, which we combine as follows.

$$\frac{\frac{\Gamma \vdots \delta_1 \quad s, t : A \quad \Delta \vdots \delta_2 \quad u : B}{sut : A \odot_{>} B} \odot_{>} I}{\vdots \delta_2 \quad D}$$

□

Theorem 1. *Derivability of a statement in D and derivability of the translation of this statement into MILL1 coincide.*

Proof Immediate from Lemma 1 and 2.

The main theorem gives a simple solution to two of the main open problems from (Morrill 2011).

Corollary 1. *D is NP-complete*

Proof We have that the derivability of L, D and MILL1 are related as follows (given the translations of L and D into MILL1) $L \subset D \subset MILL1$. Therefore NP-completeness of L and MILL1 gives us NP-completeness of D.

Corollary 2. *MILL1 provides a proof net calculus for D.*

Corollary 3. *D satisfies cut elimination.*

Using the translation into MILL1 gives us a very easy cut elimination proof.

6 Agreement, non-associativity and scope restrictions

Though I have focused only on using the first-order terms for representing string positions, I will sketch a number of other applications of the first-order terms which are orthogonal to their use for string positions, for which other extension of the Lambek calculus have introduced additional connectives and logical rules, such as the unary modalities of multimodal categorial grammar (Kurtonina & Moortgat 1997).

The most obvious of these applications is for the use of linguistic features, allowing us, for example, to distinguish between nominative and accusative noun phrases $np(nom)$ and $np(acc)$ but also allowing a lexical entry to fill either role by assigning it the formula $\forall x.np(x)$.

Until now, we have only seen variables and constants as arguments of predicate symbols. When we allow more complex terms, things get more interesting. Let's only consider complex terms of the form $s(T)$ — the well-known successor term from unary arithmetic not to be confused with the predicate symbol s for sentence — where T is itself a term (complex, a variable or a constant). These complex terms allow us to implement non-associativity when we need it, using the following translation (remember that the string positions are orthogonal and can be included if needed).

$$\begin{aligned} \|A \bullet B\|^x &= \|A\|^{s(x)} \otimes \|B\|^{s(x)} \\ \|C/B\|^s(x) &= \|B\|^{s(x)} \multimap \|C\|^x \\ \|A \setminus C\|^s(x) &= \|A\|^{s(x)} \multimap \|C\|^x \end{aligned}$$

The translation is parametric in a single variable x unique to the formula, which can get partially instantiated during the translation, producing a formula with a single free variable which is universally quantified to complete the translation. For example, a prototypical statement whose derivability presupposes associativity

$$a/b, b/c \vdash a/c$$

translates as

$$\forall x[b(s(x)) \multimap a(x)], \forall y[c(s(y)) \multimap b(y)] \vdash \forall z[c(s(z)) \multimap a(z)]$$

which the reader can easily verify to be underivable. This translation generalizes both the translation of NL to MILL1 and the implementation of island constraints of (Moot & Piazza 2001).

In addition, we can handle scope restrictions in the same spirit as (Bernardi & Moot 2003), by translating s_1 as $\forall x.s(x)$, s_2 as $\forall x.s(s(x))$ and s_3 as $\forall x.s(s(s(x)))$, which are easily verified to satisfy $s_i \vdash s_j$ for $i \leq j$ and $s_i \not\vdash s_j$ for $i > j$.

Scope restrictions and island constraints are some of the iconic applications of the unary modalities of multimodal categorial grammars and I consider it an attractive feature of MILL1 they permit a transparent translation of these applications.

The use of complex terms moves us rather close to the indexed grammars (Aho 1968), where complex unary term symbols play the role of a stack of indices. The linear indexed grammars (Gazdar 1988) would then correspond to the restriction of quantified variables to two occurrences of opposite polarity⁴ (or a single occurrence of any polarity; for the string position variables, they occur twice: either once as a left (resp. right) position of a positive atomic formula and once as a left (resp. right) position of a negative atomic formula or once as a left position and once as a right position of atomic formulas of the same polarity). If we restrict variables to at most two occurrences of each variable, without any restriction on the polarities, we are closer to an extension of linear indexed grammars proposed by (Keller & Weir 1995), which they call partially linear PATR, and thereby closer to unification-based grammars. This restriction on quantified variables seems very interesting and naturally encompasses the restriction on string position variables.

These are of course only suggestions, which need to be studied in more detail in future work.

7 Conclusions and Open Questions

First-order multiplicative intuitionistic linear logic includes several interesting subsystems: multiple context-free grammars, the Lambek calculus and the Displacement calculus. In spite of this, the computational complexity of MILL1 is

⁴ The encoding of non-associativity above is a clear violation of this constraint, since the quantified variable will occur in all atomic subformulas.

the same as the complexity of the universal recognition problem for each of these individual systems. In addition, it gives a natural implementation of several additional linguistic phenomena, which would require further machinery in each of the other calculi.

MILL1 satisfies all conditions of extended Lambek calculi: it has a simple proof theory, which includes a proof net calculus, it generates the mildly context-free languages, it is NP-complete and the homomorphism for semantics consists of simply dropping the quantifiers to obtain an MILL proof — though it is conceivable to use the first-order quantifiers for *semantic* features which would have a reflection in the homomorphism.

Many important questions have been left open. Do MILL1 grammars without complex terms (simple in the input string or not) generate exactly the MCFs or strictly more? Do MILL1 grammars *with* complex terms generate exactly the indexed languages and can we get interesting subclasses (eg. partially linear PATR) by restricting the variables to occur at most twice? Are there fragments of MILL1 grammars which have a polynomial fixed recognition problem? I hope these questions will receive definite answers in the future.

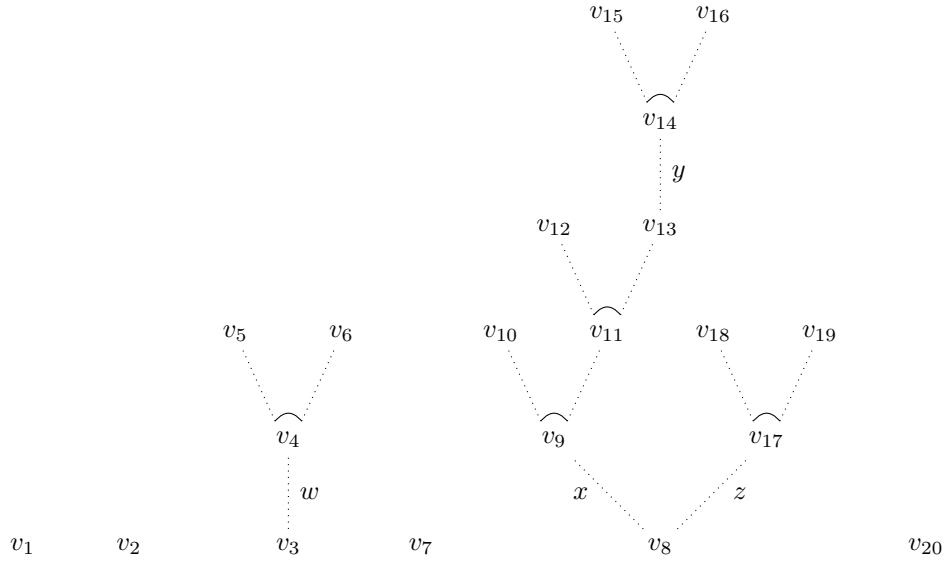
References

- Aho, A. (1968), ‘Indexed grammars: an extension of context-free grammars’, *Journal of the ACM* **15**(4), 647–671.
- Andreoli, J.-M. (1992), ‘Logic programming with focussing proofs in linear logic’, *Journal of Logic and Computation* **2**(3).
- Bar-Hillel, Y., Perles, M. & Shamir, E. (1964), On formal properties of simple phrase structure grammars, in Y. Bar-Hillel, ed., ‘Language and Information. Selected Essays on their Theory and Application’, Addison-Wesley, New York, pp. 116–150.
- Bellin, G. & van de Wiele, J. (1995), Empires and kingdoms in MLL, in J.-Y. Girard, Y. Lafont & L. Regnier, eds, ‘Advances in Linear Logic’, Cambridge University Press, pp. 249–270.
- Bernardi, R. & Moortgat, M. (2007), Continuation semantics for symmetric categorial grammar, in ‘Proceedings of WoLLIC 2007’, Vol. 4567 of *LNCS*, Springer, pp. 53–71.
- Bernardi, R. & Moot, R. (2003), ‘Generalized quantifiers in declarative and interrogative sentences’, *Logic Journal of the IGPL* **11**(4), 419–434.
- Boullier, P. (1998), Proposal for a natural language processing syntactic backbone, Technical Report 3342, INRIA, Rocquencourt.
- Danos, V. (1990), La Logique Linéaire Appliquée à l’étude de Divers Processus de Normalisation (Principalement du λ -Calcul), PhD thesis, University of Paris VII.
- Danos, V. & Regnier, L. (1989), ‘The structure of multiplicatives’, *Archive for Mathematical Logic* **28**, 181–203.
- Došen, K. (1992), ‘A brief survey of frames for the Lambek calculus’, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* **38**, 179–187.
- Gazdar, G. (1988), Applicability of indexed grammars to natural languages, in U. Reyle & C. Rohrer, eds, ‘Natural Language Parsing and Linguistic Theories’, D. Reidel, Dordrecht, pp. 69–94.

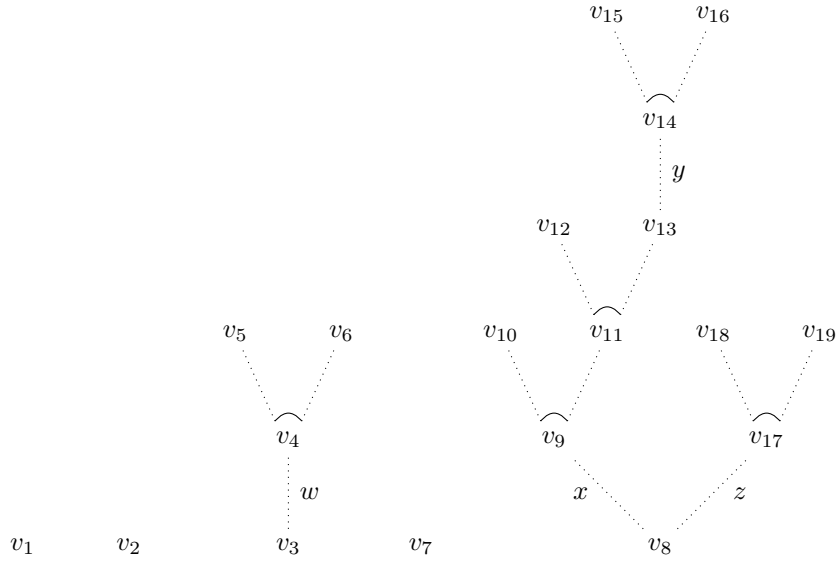
- Girard, J.-Y. (1991), Quantifiers in linear logic II, in G. Corsi & G. Sambin, eds, 'Nuovi problemi della logica e della filosofia della scienza', Vol. II, CLUEB, Bologna, Italy. Proceedings of the conference with the same name, Viareggio, Italy, January 1990.
- Girard, J.-Y. (2001), 'Locus solum: from the rules of logic to the logic of rules', *Mathematical Structures in Computer Science* **11**, 301–506.
- Greibach, S. A. (1965), 'A new normal-form theorem for context-free phrase structure grammars', *Journal of the ACM* **12**(1), 42–52.
- Guerrini, S. (1999), Correctness of multiplicative proof nets is linear, in 'Fourteenth Annual IEEE Symposium on Logic in Computer Science', IEEE Computer Science Society, pp. 454–263.
- Huybregts, R. (1984), The weak inadequacy of context-free phrase structure grammars, in G. de Haan, M. Trommelen & W. Zonneveld, eds, 'Van Periferie naar Kern', Foris, Dordrecht.
- Joshi, A. (1985), Tree-adjointing grammars: How much context sensitivity is required to provide reasonable structural descriptions, in D. Dowty, L. Karttunen & A. Zwicky, eds, 'Natural Language Processing: Theoretical, Computational, and Psychological Perspectives', Cambridge University Press.
- Kanazawa, M. (2009), The pumping lemma for well-nested multiple context-free languages, in V. Diekert & D. Nowotka, eds, 'Developments in Language Theory: 13th International Conference (DLT 2009)', Vol. 5583 of *Lecture Notes in Computer Science*, Springer, pp. 312–325.
- Keller, B. & Weir, D. (1995), A tractable extension of linear indexed grammars, in 'Proceedings of the Seventh Meeting of the European Chapter of the Association for Computational Linguistics', pp. 75–82.
- Kurtonina, N. & Moortgat, M. (1997), Structural control, in P. Blackburn & M. de Rijke, eds, 'Specifying Syntactic Structures', CSLI, Stanford, pp. 75–113.
- Lambek, J. (1958), 'The mathematics of sentence structure', *American Mathematical Monthly* **65**, 154–170.
- Lincoln, P. (1995), Deciding provability of linear logic formulas, in J.-Y. Girard, Y. Lafont & L. Regnier, eds, 'Advances in Linear Logic', Cambridge University Press, pp. 109–122.
- Moortgat, M. (1996), 'Multimodal linguistic inference', *Journal of Logic, Language and Information* **5**(3–4), 349–385.
- Moortgat, M. (2007), Symmetries in natural language syntax and semantics: the Lambek-Grishin calculus, in 'Proceedings of WoLLIC 2007', Vol. 4567 of *LNCS*, Springer, pp. 264–284.
- Moortgat, M. (2010), 'Typological grammar', Stanford Encyclopedia of Philosophy Website. <http://plato.stanford.edu/entries/typological-grammar/>.
- Moortgat, M. (2011), Categorical type logics, in J. van Benthem & A. ter Meulen, eds, 'Handbook of Logic and Language', Elsevier/MIT Press, chapter 2, pp. 95–179.
- Moot, R. (2002), Proof Nets for Linguistic Analysis, PhD thesis, Utrecht Institute of Linguistics OTS, Utrecht University.
- Moot, R. (2007), Filtering axiom links for proof nets, in L. Kallmeyer, P. Monachesi, G. Penn & G. Satta, eds, 'Proceedings of Formal Grammar 2007'. to appear with CSLI.
- Moot, R. (2008), Lambek grammars, tree adjoining grammars and hyperedge replacement grammars, in C. Gardent & A. Sarkar, eds, 'Proceedings of TAG+9, The ninth international workshop on tree adjoining grammars and related formalisms', pp. 65–72.
- Moot, R. & Piazza, M. (2001), 'Linguistic applications of first order multiplicative linear logic', *Journal of Logic, Language and Information* **10**(2), 211–232.

- Morrill, G. (2011), *Categorial Grammar: Logical Syntax, Semantics, and Processing*, Oxford University Press.
- Morrill, G. & Fadda, M. (2008), ‘Proof nets for basic discontinuous Lambek calculus’, *Journal of Logic and Computation* **18**(2), 239–256.
- Morrill, G. & Valentín, O. (2010), On calculus of displacement, in ‘Proceedings of TAG+Related Formalisms’, University of Yale.
- Morrill, G., Valentín, O. & Fadda, M. (2011), ‘The displacement calculus’, *Journal of Logic, Language and Information* **20**(1), 1–48.
- Murawski, A. S. & Ong, C.-H. L. (2000), Dominator trees and fast verification of proof nets, in ‘Logic in Computer Science’, pp. 181–191.
- Nederhof, M.-J. & Satta, G. (2010), Theory of parsing, in A. Clark, C. Fox & S. Lappin, eds, ‘The Handbook of Computational Linguistics and Natural Language Processing’, Wiley-Blackwell, pp. 105–130.
- Pentus, M. (1997), ‘Product-free Lambek calculus and context-free grammars’, *Journal of Symbolic Logic* **62**, 648–660.
- Pentus, M. (2006), ‘Lambek calculus is NP-complete’, *Theoretical Computer Science* **357**(1), 186–201.
- Pentus, M. (2010), ‘A polynomial-time algorithm for Lambek grammars of bounded order’, *Linguistic Analysis* **36**(1–4), 441–471.
- Pereira, F. & Shieber, S. (1987), *Prolog and Natural Language Analysis*, CSLI, Stanford.
- Seki, H., Matsumura, T., Fujii, M. & Kasami, T. (1991), ‘On multiple context-free grammars’, *Theoretical Computer Science* **88**, 191–229.
- Shieber, S. (1985), ‘Evidence against the context-freeness of natural language’, *Linguistics & Philosophy* **8**, 333–343.
- Stabler, E. (2003), Tupled pregroup grammars, Technical report, University of California, Los Angeles.
- Wijnholds, G. (2011), Investigations into categorial grammar: Symmetric pregroup grammar and displacement calculus, Master’s thesis, Utrecht University.

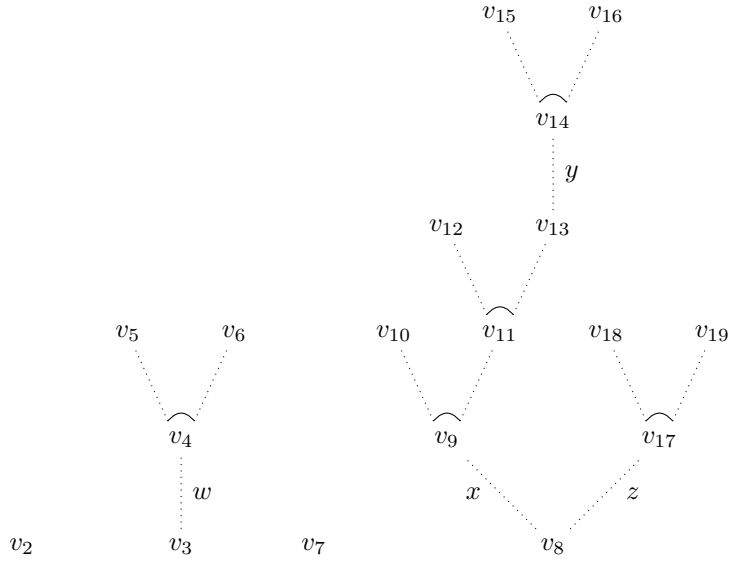
A Example: “John left before Mary did”



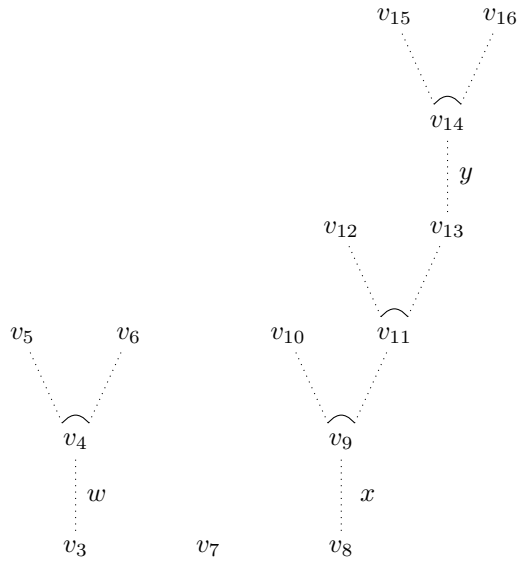
v_1	$\{np(\bar{0}, 1)\}$	\emptyset	v_{11}	\emptyset
v_2	$\{np(\bar{+}, 1), s(\bar{A}, 2)\}$	\emptyset	v_{12}	$\{np(\bar{+}, Y), s(\bar{E}, Z)\}$
v_3	$\{s(\bar{3}, W), np(\bar{B}, V), s(\bar{B}, W)\}$	\emptyset	v_{13}	\emptyset
v_4	\emptyset	\emptyset	v_{14}	\emptyset
v_5	$\{np(\bar{w}, V)\}$	$\{w\}$	v_{15}	$\{np(\bar{y}, X)\}$
v_6	$\{s(\bar{+}, 2)\}$	$\{w\}$	v_{16}	$\{s(\bar{+}, x)\}$
v_7	$\{np(\bar{3}, 4)\}$	\emptyset	v_{17}	\emptyset
v_8	$\{s(\bar{Y}', 5), np(\bar{Y}', X)\}$	\emptyset	v_{18}	$\{np(\bar{z}, Y)\}$
v_9	\emptyset	\emptyset	v_{19}	$\{s(\bar{+}, Z)\}$
v_{10}	$\{s(\bar{X}', x), np(\bar{X}', 4)\}$	$\{x\}$	v_{20}	$\{s(\bar{+}, 5)\}$



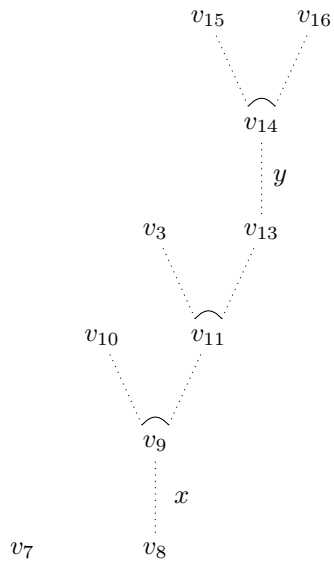
v_1	$\{np(\bar{0}, 1)\}$	\emptyset	v_{11}	\emptyset
v_2	$\{np(\overset{+}{A}, 1), s(\bar{A}, 2)\}$	\emptyset	v_{12}	$\{np(\overset{+}{E}, Y), s(\bar{E}, Z)\}$
v_3	$\{s(\overset{+}{3}, W), np(\overset{+}{B}, V), s(\bar{B}, W)\}$	\emptyset	v_{13}	\emptyset
v_4	\emptyset	\emptyset	v_{14}	\emptyset
v_5	$\{np(\bar{w}, V)\}$	$\{w\}$	v_{15}	$\{np(\bar{y}, X)\}$
v_6	$\{s(\overset{+}{w}, 2)\}$	$\{w\}$	v_{16}	$\{s(\overset{+}{y}, x)\}$
v_7	$\{np(\bar{3}, 4)\}$	\emptyset	v_{17}	\emptyset
v_8	$\{np(\overset{+}{0}, X)\}$	\emptyset	v_{18}	$\{np(\bar{z}, Y)\}$
v_9	\emptyset	\emptyset	v_{19}	$\{s(\overset{+}{z}, Z)\}$
v_{10}	$\{s(\bar{X}', x), np(\overset{+}{X'}, 4)\}$	$\{x\}$		$\{z\}$



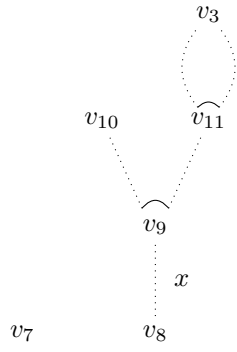
v_2	$\{np(A^+, 1), s(\bar{A}, 2)\}$	\emptyset	v_{11}	\emptyset
v_3	$\{s(\bar{3}, W^+), np(\bar{B}, V^+), s(\bar{B}, W^-)\}$	\emptyset	v_{12}	$\{np(\bar{E}^+, Y), s(\bar{E}, Z^-)\}$
v_4	\emptyset	\emptyset	v_{13}	\emptyset
v_5	$\{np(\bar{w}, V^-)\}$	$\{w\}$	v_{14}	\emptyset
v_6	$\{s(\bar{w}, 2^+)\}$	$\{w\}$	v_{15}	$\{np(\bar{y}, 1^-)\}$
v_7	$\{np(\bar{3}, 4^-)\}$	\emptyset	v_{16}	$\{s(\bar{y}, x^+)\}$
v_8	\emptyset	\emptyset	v_{17}	\emptyset
v_9	\emptyset	\emptyset	v_{18}	$\{np(\bar{z}, Y^-)\}$
v_{10}	$\{s(\bar{X}', x^-), np(\bar{X}', 4^+)\}$	$\{x\}$	v_{19}	$\{s(\bar{z}, Z^+)\}$



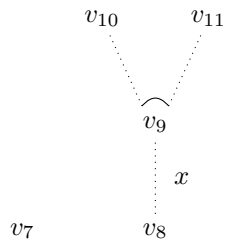
v_3	$\{s(3, W)^+, np(\bar{B}, V)^+, s(B, W)^-\}$	\emptyset	v_{10}	$\{s(\bar{X}', x)^-, np(\bar{X}', 4)^+\}$	$\{x\}$
v_4	\emptyset	\emptyset	v_{11}	\emptyset	\emptyset
v_5	$\{np(w, V)^-\}$	$\{w\}$	v_{12}	$\{np(\bar{E}, 1)^+, s(\bar{E}, 2)^-\}$	\emptyset
v_6	$\{s(w, 2)^+\}$	$\{w\}$	v_{13}	\emptyset	\emptyset
v_7	$\{np(\bar{3}, 4)^-\}$	\emptyset	v_{14}	\emptyset	\emptyset
v_8	\emptyset	\emptyset	v_{15}	$\{np(y, 1)^-\}$	$\{y\}$
v_9	\emptyset	\emptyset	v_{16}	$\{s(y, x)^+\}$	$\{x, y\}$



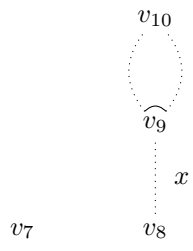
v_3	$\{s(3, W)^+, np(B, 1)^+, s(B, W)^-\}$	\emptyset	v_{10}	$\{s(X', x)^-, np(X', 4)^+\}$	$\{x\}$
			v_{11}	\emptyset	\emptyset
			v_{12}	$\{np(w, 1)^+, s(w, 2)^-\}$	\emptyset
			v_{13}	\emptyset	\emptyset
v_7	$\{np(3, 4)^-\}$	\emptyset	v_{14}	\emptyset	\emptyset
v_8	\emptyset	\emptyset	v_{15}	$\{np(y, 1)^-\}$	$\{y\}$
v_9	\emptyset	\emptyset	v_{16}	$\{s(y, x)^+\}$	$\{x, y\}$



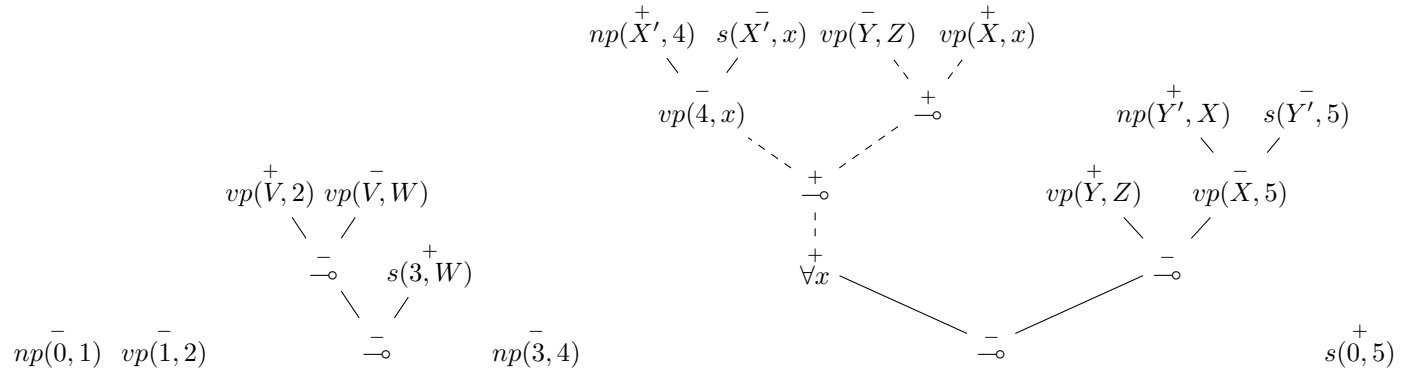
v_3	$\{s(3, x)^+\}$	$\{x\}$	v_{10}	$\{s(\bar{X}', x), np(\bar{X}', 4)\}$	$\{x\}$
			v_{11}	\emptyset	\emptyset
v_7	$\{np(\bar{3}, 4)\}$	\emptyset			
v_8	\emptyset	\emptyset			
v_9	\emptyset	\emptyset			



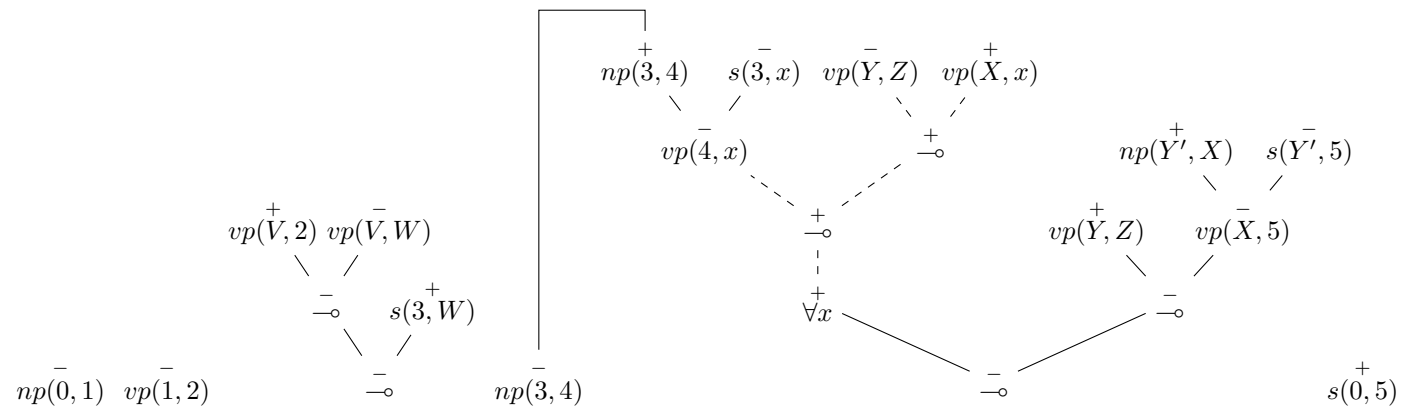
v_{10}	$\{s(\overline{X'}, x), np(\overline{X'}, 4)\}$	$\{x\}$
v_{11}	$\{s(\overline{3}, x)\}$	$\{x\}$
v_7	$\{np(\overline{3}, 4)\}$	\emptyset
v_8	\emptyset	\emptyset
v_9	\emptyset	\emptyset

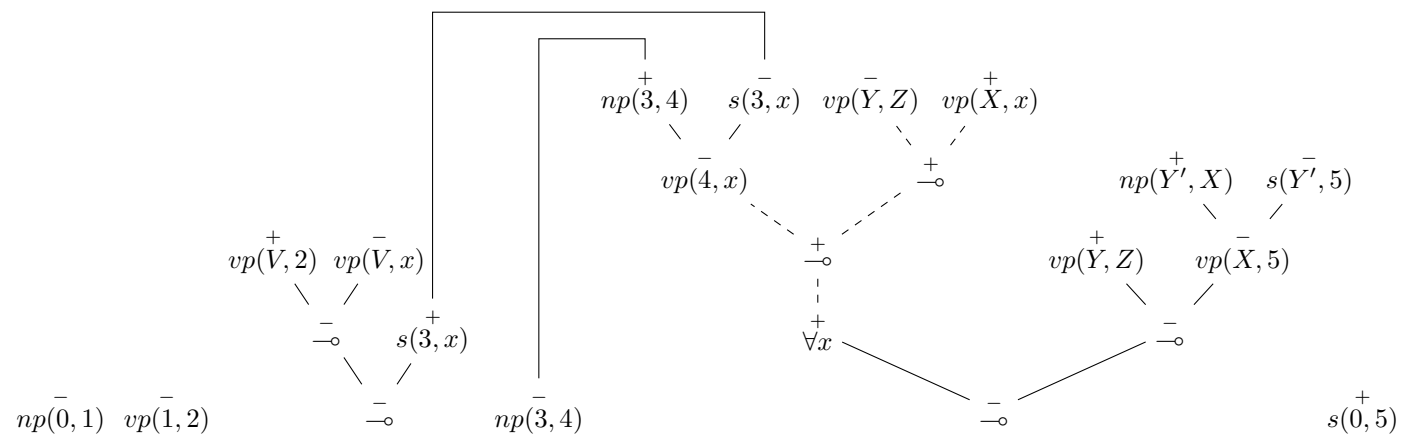


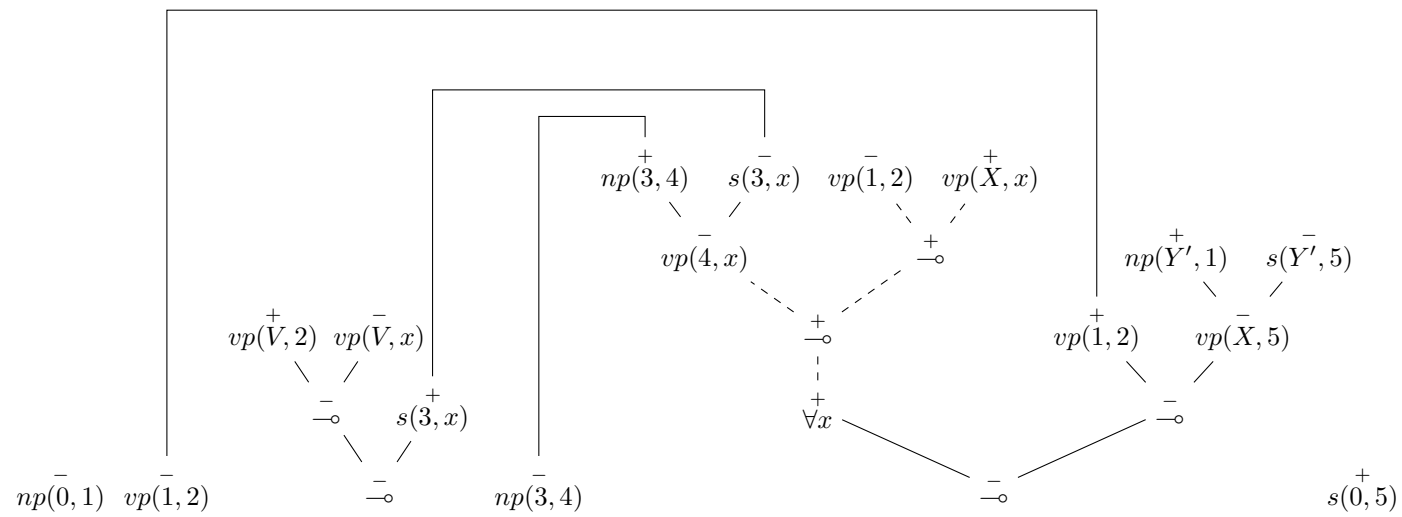
$$\begin{array}{l}
 v_7 \{np(\bar{3}, 4)\} \emptyset \\
 v_8 \emptyset \quad \emptyset \\
 v_9 \emptyset \quad \emptyset
 \end{array}
 \left|
 \begin{array}{l}
 v_{10} \{np(\overset{+}{3}, 4)\} \{x\}
 \end{array}
 \right.$$

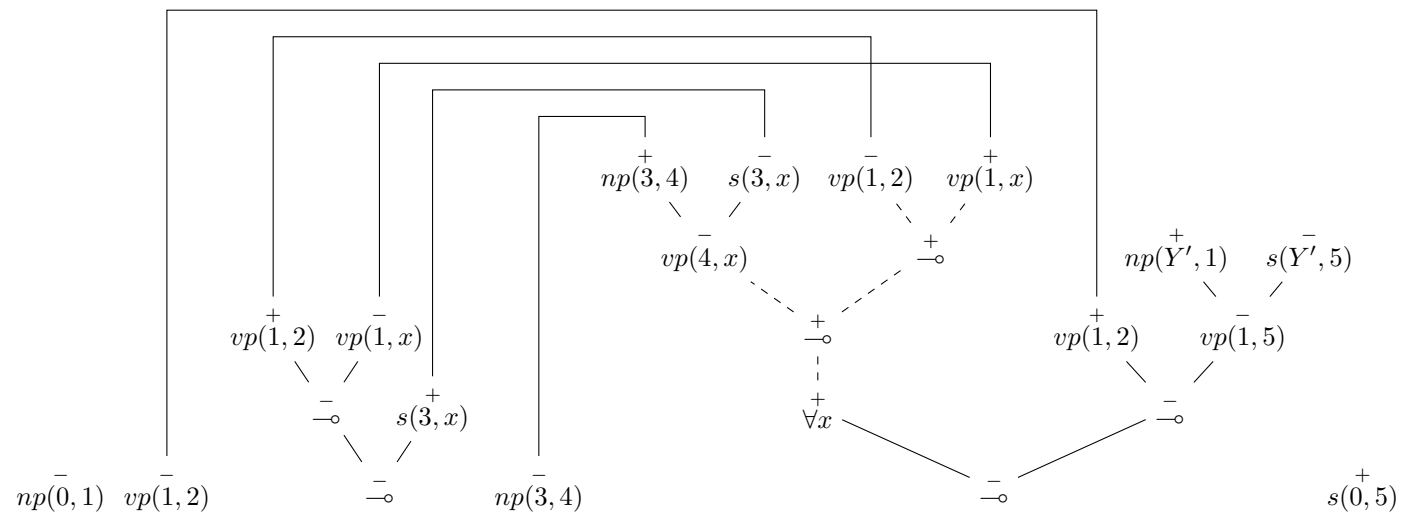


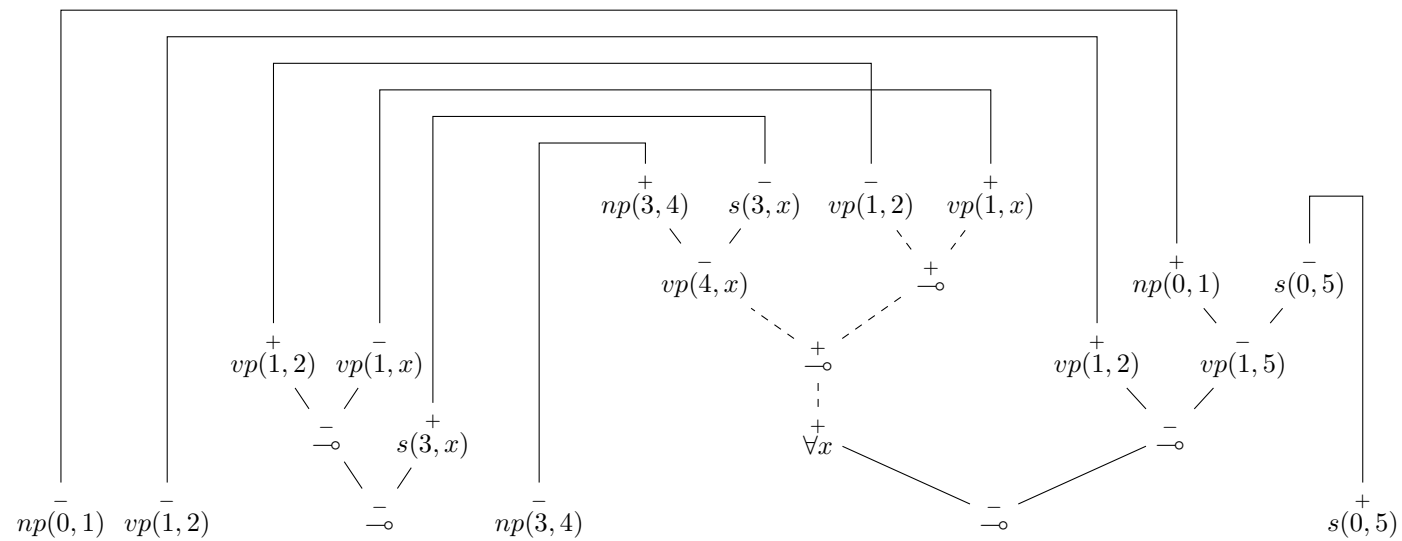
B Example: “John left before Mary did”











$np(\bar{1}, 2)$ $np(\bar{2}, 3)$ $np(\bar{3}, 4)$ $np(\bar{4}, 6)$
 Jan Henk Cecilia dn

