On the Logic of LGB Type Structures. Part I: Multidominance Structures

MARCUS KRACHT

ABSTRACT. The present paper is the first part of a sequence of papers devoted to the modal logics of structures that arise from Government and Binding theory. It has been shown in [9] that they can be modeled by so-called multidominance structures (MDSs). The result we are going to prove here is that the dynamic logic of the MDSs is decidable. Moreover, we shall indicate how the theory of Government and Binding as well as the Minimalist Program can be coded in modal logic. Some preliminary decidability results for GB are obtained, which will be followed up in the sequel to this paper.

1. INTRODUCTION

In recent years, the idea of model theoretic syntax has been getting more attention. One of the advantages of model theoretic syntax is that it describes syntactic structures using a logical language so that fundamental theoretical questions can receive a precise formulation and can—hopefully—be answered. This idea can be found already in the work by ED STABLER (see [15]), where it was argued that questions of dependency among different modules of grammar, or independence questions for principles can be translated into logical questions. STABLER chose a translation into predicate logic, accompanied by an implementation in Prolog. Thus, questions could be posed to a computer, which would then answer them. The problem with this procedure is twofold. Often the predicate logic of a class of structures is undecidable and so not all questions can be effectively answered. Second, even if the logic is decidable we need to know about its complexity so that we know how long we have to wait until we get an answer. Thus, the best possible result would be one where we had not only a decidability result but also a complexity result, preferably showing that complexity is low.

Rabin has shown that the (weak) monadic second order logic (MSO) of trees is decidable, a result that JAMES ROGERS [14] has applied to syntactic theory. The main disadvantage of this approach is that it does not cover

LGB type structures.[1] The obvious step was to reduce the latter to the former. This is not always possible, but it led to a result (independently proved by JAMES ROGERS and myself) that if head movement is bounded then Minimality in the sense of LUIGI RIZZI [13] or Locality in the sense of RITA MANZINI [12] come down to the theory that the language is strongly context free. However, nothing could be said about the case when head movement was unbounded, because the reduction fails in this case. Now, ROGERS remarks that adding free indexation makes the second order theory undecidable (it is no longer monadic), thus implicitly claiming that the monadic second order theory of LGB type structures might be undecidable.

The good news however is that this need not always be a concern. In this paper I shall show that the dynamic logic of a good many classes of structures is decidable. An application to non-context free languages will be given. Moreover, I shall describe how GB type structures as well as MP type structures can be described using dynamic logic. The sequel to this paper will generalise the result of this paper still further. It will emerge that many theories of generative grammar are effectively decidable. This is hopefully the beginning of a general decidability proof that covers the linguistically relevant structures. The applications of the present results are manifold. We are given a decision procedure to see whether certain principles of grammar are independent or not, and we are given a decision procedure whether or not a sentence is in the language.

2. MULTIDOMINANCE STRUCTURES

In generative grammar, structures are derived from deep structure trees. In [9] I considered three kinds of structures: **trace chain structures** (**TCS**s), **copy chain structures** (**CCS**s) and **multidominance structures** (**MDS**s). TCSs are the kind of entities most popular in linguistics. When an element moves, it leaves behind a trace and forms a chain together with the trace. The technical implementation is a little different, but the idea is very much the same. CCSs are different in that the moving element does not leave just a trace behind but a full copy of itself. This type of chain structures is more in line with recent developments (the Minimalist Program), rather than with the GB. MDSs, however, are different from both. In an MDS, there are no traces. Instead, movement to another position is represented by the addition of a link to that position. Thus, as soon as there is movement, there are elements which have more than one mother. Moreover, it was shown in [9] that MDSs contain exactly the same information as TCSs, since there is an algorithm that converts one into the other. MDSs, like TCSs, are based

---

[1]The shorthand 'LGB' refers to [2] as a generic source for the kinds of structures that Government and Binding uses.

on an immediate dominance relation, written $\succ$. (The converse of this relation is denoted by $\prec$.) In what is to follow, we assume that structures are downward binary branching. Every node has at most two daughters. To install this, we shall assume two relations, $\succ_0$ and $\succ_1$ each of which is a partial function, and $\succ \; = \; \succ_0 \cup \succ_1$. We do not require the two relations to be disjoint.

Definition 1. *A **preMDS** is a structure $\langle M, \succ_0, \succ_1 \rangle$, where the following holds:*

(P1) *If $y \succ_0 x$ and $y \succ_0 x'$ then $x = x'$.*
(P2) *If $y \succ_1 x$ and $y \succ_1 x'$ then $x = x'$.*
(P3) *If $y \succ_1 x$ then there is a z such that $y \succ_0 z$.*
(P4) *There is exactly one x such that for no y, $y \succ x$ (this element is called the **root**).*
(P5) *The set $M(x) := \{y : x \prec y\}$ is linearly ordered by $\prec^+$.*

We call a pair $\langle x, y \rangle$ such that $x \prec y$ a **link**. We shall also write $x ; y$ to say that $\langle x, y \rangle$ is a link. The link $\langle x, y \rangle$ such that y is minimal in $M(x)$ is called a **root link**. A link that is not a root link is called **derived**. A **leaf** is a node without daughters. Recall the definition of the transitive closure R^+ of a binary relation R.

For technical reasons we shall split $\prec_0$ as follows.

$$(1) \qquad\qquad \prec_0 \; = \; \prec_{00} \cup \prec_{01}$$

where $x \prec_{00} y$ iff $x \prec_0 y$ and y is minimal in $M(x)$. Alternatively, $x \prec_{00} y$ if $x \prec_0 y$ and $\langle x, y \rangle$ is a root link. Similarly,

$$(2) \qquad\qquad \prec_1 \; = \; \prec_{10} \cup \prec_{11}$$

where $x \prec_{10} y$ iff $x \prec_1 y$ and y is minimal in $M(x)$ (or, equivalently, $\langle x, y \rangle$ is a root link). The definition above is suitably amended. We shall define

$$(3) \qquad\qquad \prec_{\bullet 0} := \prec_{00} \cup \prec_{10}$$

$$(4) \qquad\qquad \prec_{\bullet 1} := \prec_{01} \cup \prec_{11}$$

The structures we get are called *PMDSs*.

Definition 2. *A **PMDS** is a structure $\langle M, \succ_{00}, \succ_{01}, \succ_{10}, \succ_{11} \rangle$ which, in addition to (P1) – (P5) of Definition 1 satisfies*

(P6) *If $y \in M(x)$ then $x \prec_{\bullet 0} y$ iff y is the least element of $M(x)$ with respect to $\prec$.*

We note here that every MDS can be turned into a unique PMDS, and every PMDS defines exactly one MDS. We shall work with PMDSs henceforth (but continue to call them MDSs). We assume that the leaves are

linearly ordered in the following way.

$$(5) \qquad x \sqsubset y :\Leftrightarrow (\exists z)(\exists u)(\exists v)(x <^*_{\bullet 0} z \prec_{00} u >_{10} v >^*_{\bullet 0} y)$$

(This is not the only possible ordering; this establishes in fact the order at D-structure. This is enough for the present purposes, though.)

Definition 3. *An **ordered MDS** (**OMDS**) is a PMDS in which $\sqsubset$ is linear on the leaves.*

Now, since $<^+_{\bullet 0}$ is a tree ordering, we can extend $\sqsubset$ to an ordering between any two incomparable nodes (where x and y are incomparable if neither $x <^+_{\bullet 0} y$ nor $y <^+_{\bullet 0} x$ nor $x = y$). In fact, the extension is exactly as defined by (5). Details can be found, for example, in [11]. Notice that in an OMDS, $\prec_0 \cap \prec_1 = \emptyset$. For suppose otherwise. Then for some x and y we have $x \prec_0 y$ and $x \prec_1 y$ and therefore $z \sqsubset z$ for any leaf $z \leq x$, by definition of $\sqsubset$.

In presence of the ordering postulate, the condition (P5) and (P6) can be replaced by the following

$$(6) \qquad \text{The set } M(x) := \{y : x \prec y\} \text{ is linearly ordered by } <^+_{\bullet 0}.$$

This is easy to see. First we prove a

Lemma 4. *Suppose that $y \prec y'$ and that there is no x such that $y \prec^+ x \prec^+ y'$. Then $y \prec_{\bullet 0} y'$.*

The proof of the claim is in the fact that $y' \in M(y)$. If the link is derived, it is not minimal, so there is a z such that $y' <_{\bullet 0} z \prec^+ y'$. And conversely.

Suppose now that $x \prec y$. Then there is a chain $y = y_0 \prec y_1 \prec y_2 \prec \cdots \prec y_n = y'$. The longest such chain contains only nonderived links, by Lemma 4. This means that $x <^+_{\bullet 0} y$. Now, $<^+_{\bullet 0}$ is a tree ordering, so that if $y' \in M(x)$, then $x <^+_{\bullet 0} y'$ as well, and so either $y = y'$ or $y <^+_{\bullet 0} y'$ or $y' <^+_{\bullet 0} y$, as promised.

Proposition 5. *Let $\mathfrak{M}$ be a PMDS. $\mathfrak{M}$ is an OMDS iff the following holds: if x is not the root, $\prec_{10}$ is defined iff $\prec_{00}$ is undefined on x.*

We shall prove the theorem and exhibit some useful techniques. We code the elements of $\mathfrak{M}$ by sequences in the following way. Let I be a chain $\{x_i : i < n + 1\}$ such that x_0 is the root, and $x_i >_{\bullet 0} x_{i+1}$ for every $i < n$. We call I a **standard identifier** for x and denote it by $I(x)$. n is called the **standard depth** of x_n and we write $\mathrm{sd}(x_n)$ to denote it.

Lemma 6. *In an OMDS, every x has exactly one standard identifier. Hence, the standard depth of x is uniquely defined.*

(See also [9] on the notion of an identifier.) Let us see why the standard identifier is unique.

We translate the identifier into a binary sequence $b_0 b_1 \cdots b_n$ defined by

$$(7) \qquad b_i = \begin{cases} 0 & \text{if } x_i >_{00} x_{i+1}, \\ 1 & \text{if } x_i >_{10} x_{i+1}. \end{cases}$$

In this way, we associate a binary sequence with each node. Now suppose that (5) defines a linear ordering on the leaves. First: this number is unique. For if not, there are two sequences, $b_0 b_1 \cdots b_n$ and $c_0 c_1 \cdots c_m$ of this kind for x_n. As they are distinct there is a least index j such that $b_j \neq c_j$, say $b_j = 0$ and $c_j = 1$. Then, by (5), if $z \leq x_n$ is a leaf, $z \sqsubset z$. Contradiction. Now, let x be given. It has a sequence $b_0 b_1 \cdots b_n$ associated with it. Let $y >_{\bullet 0} x$. Then y is defined by $b_0 b_1 \cdots b_{n-1}$, which is unique. So, $<_{\bullet 0}$ is a partial function. Conversely, if $<_{\bullet 1}$ is a partial function, then the translation into binary sequences is unique. Now define $\sqsubset$ for sequences by $b_0 b_1 \cdots b_n$ and $c_0 c_1 \cdots c_m$ iff for the first j such that $b_j \neq c_j$, $b_j = 0 < c_j = 1$. This is exactly the order (5), spelled out for the representing sequences. This order is loop free, transitive and linear on the maximal sequences (which correspond to the leaves). We add that $b_0 b_0 \cdots b_m$ is immediately to the left of $c_0 c_1 \cdots c_n$ if

$$(8) \qquad \begin{aligned} b_0 b_0 \cdots b_m &= b_0 b_1 \cdots b_{j-1} 01 \cdots 1, \\ c_0 c_1 \cdots c_n &= b_0 b_1 \cdots b_{j-1} 10 \cdots 0 \end{aligned}$$

(The lengths of these sequences need not be equal.)

I should emphasize that the identifiers do not necessarily form a tree domain. Recall that a tree domain T is a subset of $\mathbb{N}^*$ such that the following holds: (a) if $\vec{x} i \in T$ then $\vec{x} \in T$, and (b) if $\vec{x} j \in T$ and $i < j$ then also $\vec{x} i \in T$. Property (a) holds but (b) does not hold in general. For suppose that $x >_{01} y$ and $x >_{10} z$. Then $I(z) = I(x)1$. However since the link $y; x$ is derived there is no standard identifier of the form $I(x)0$. The identifier $I(y)$ contains $I(z) = I(x)1$ as a prefix.

3. Dynamic Logic

The language of dynamic propositional logic (**PDL**) is defined as follows. Given any set Π_0 of so-called **basic programs**, a set Γ of propositional constants, and V of variables, the set of formulae is the closure under the following:

   ☞ If χ is a formula, $\chi?$ is a program.
   ☞ If χ, χ' are formulae, so are $\neg\chi$ and $\chi \wedge \chi'$.
   ☞ If α, α' are programs, so is $\alpha; \alpha'$ and $\alpha \cup \alpha$; and α^*.
   ☞ If α is a program and χ a formula, $\langle\alpha\rangle\chi$ is a formula.

We put $\chi \vee \chi' := \neg(\neg\chi \wedge \neg\chi')$ and $[\alpha]\chi := \neg\langle\alpha\rangle\neg\chi$, and similarly for other boolean connectives. The minimal logic, also denoted **PDL**, is the least set of formulae with the following properties:

(1) All propositional tautologies are in **PDL**.

(2) $[\alpha](\chi \to \chi') \to ([\alpha]\chi \to [\alpha]\chi') \in$ **PDL**.

(3) $\langle\chi?\rangle\chi' \leftrightarrow (\chi \wedge \chi') \in$ **PDL**.

(4) $\langle\alpha \cup \alpha'\rangle\chi \leftrightarrow \langle\alpha\rangle\chi \vee \langle\alpha'\rangle\chi \in$ **PDL**.

(5) $\langle\alpha; \alpha'\rangle\chi \leftrightarrow \langle\alpha\rangle\langle\alpha'\rangle\chi \in$ **PDL**.

(6) $\chi \wedge [\alpha^*](\chi \to [\alpha]\chi) \to [\alpha^*]\chi \in$ **PDL**.

(7) If $\chi \in$ **PDL** then $[\alpha]\chi \in$ **PDL**.

(8) If $\chi \to \chi' \in$ **PDL** and $\chi \in$ **PDL** then $\chi' \in$ **PDL**.

(9) If $\chi \in$ **PDL**, then $s(\chi) \in$ **PDL** for every substitution s.

Here, a substitution is defined by a function s that assigns a formula $s(p)$ to every variable p. The formula $s(\chi)$ is obtained by replacing every occurrence of a variable p by $s(p)$, for every variable p. A **dynamic logic** is a set $L \subseteq$ **PDL** which has the properties (7) – (9). Let χ be a formula and L a dynamic logic; then $L \oplus \chi$ denotes the least dynamic logic containing L and χ. Similarly with a set Δ in place of χ.

Model structures are of the form $\mathfrak{F} = \langle W, C, R \rangle$, where W is a set (the set of **worlds** or **points**), $C : \Gamma \to \wp(W)$ a function assigning each constant a set of worlds, and $R : \Pi_0 \to \wp(W \times W)$ a function assigning each basic program a relation on W. A **valuation** is a function $\beta : V \to \wp(W)$. Based on this we define the interpretation of complex programs as relations in the following way.

$$(9) \quad \begin{aligned} R(\alpha \cup \alpha') &:= R(\alpha) \cup R(\alpha') \\ R(\alpha; \alpha') &:= R(\alpha) \circ R(\alpha') \\ R(\alpha^*) &:= R(\alpha)^* \\ R(\chi?) &:= \{\langle w, w\rangle : \langle\mathfrak{F}, \beta, w\rangle \vDash \chi\} \end{aligned}$$

The truth of a formula at a world is defined thus.

$$(10) \quad \begin{aligned} \langle\mathfrak{F}, \beta, w\rangle &\vDash \neg\chi :\Leftrightarrow \langle\mathfrak{F}, \beta, w\rangle \nvDash \chi \\ \langle\mathfrak{F}, \beta, w\rangle &\vDash \chi \wedge \chi' :\Leftrightarrow \langle\mathfrak{F}, \beta, w\rangle \vDash \chi; \chi' \\ \langle\mathfrak{F}, \beta, w\rangle &\vDash \langle\alpha\rangle\chi :\Leftrightarrow \text{there is } u: w\, R(\alpha)\, u \text{ and } \langle\mathfrak{F}, \beta, u\rangle \vDash \chi \end{aligned}$$

We write $\mathfrak{F} \vDash \varphi$ if for all valuations β and all worlds w: $\langle\mathfrak{F}, \beta, w\rangle \vDash \varphi$. The logic of a class $\mathcal{K}$ of structures is

$$(11) \qquad\qquad \mathrm{Th}(\mathcal{K}) := \{\varphi : \text{for all } \mathfrak{F} \in \mathcal{K}: \mathfrak{F} \vDash \varphi\}$$

It has been shown that **PDL** is the logic of all structures and that it is also the logic of the finite structures. From this follows the decidability of **PDL**. However, more is known.

Theorem 7. **PDL** *is EXPTIME-complete.*

This means that there are constants c and b and a polynomial $p(x)$ such that for every formula φ of length $n > c$ the time needed to solve the problem whether $\varphi \in$ **PDL** takes $b^{p(n)}$ time. (Additionally, any problem of this complexity can be coded as such a problem in polynomial time.)

4. GRAMMARS AS LOGICS

In context free grammars one distinguishes the terminal alphabet from the rules. A similar distinction is made here as well. Nodes that have no daughters are called **terminal**. The lexicon is a set of declarations which state which labels terminal nodes may have. This is typically done by introducing a finite set of constants and the statement that all and only those nodes may be terminal at which one of the constants is true. With the constants a part of the language the lexicon is effectively identified with a specific nonmodal formula. In fact, we are more generous here and assume that the lexicon is a constant formula λ, which may involve modal operators. This is useful when we want to assume that the lexicon also contains complex items, as is often the case in generative grammar. The grammar is a (finite) set of formulae expressed in the above language. While the grammar is believed to be the same for all languages, the lexicon is subject to variation.

The logic **DPDL** is obtained by adding the formula $\langle\alpha\rangle\chi \rightarrow [\alpha]\chi$ for every basic program. A frame is a frame for **DPDL** iff for every basic program α: if $x\ R(\alpha)\ y$ and $x\ R(\alpha)\ y'$ then $y = y'$. (One says that the program α is **deterministic**, and this is the reason the logic is called **DPDL**.) Furthermore, the logic of finite deterministic computations is obtained by adding the formula

$$[\alpha^+]([\alpha^+]p \rightarrow p) \rightarrow [\alpha^+]p \tag{12}$$

where α is the union of all basic programs (hence this definition requires that Π_0 is finite). If we want to mention the number n of programs, we write **DPDL$_n$.f**. The following is proved in [7] (fmp and decidability) and [16] (EXPTIME-completeness).

Theorem 8. *For every n,* **DPDL$_n$.f** *is the logic of all finite structures with n basic programs, where the basic programs are deterministic and their union is loop free.* **DPDL$_n$.f** *is decidable, it is EXPTIME-complete and complete with respect to finite trees.*

Theorem 9. *For every n, the* **PDL***-logic of n-branching trees has the finite model property and is decidable.*

Many of the basic results can also be obtained by using a translation of dynamic logic into monadic second logic (MSO). The disadvantage is of

using MSO is that the complexity of the logic is for the most part nonelementary (in the sense of recursion theory), while **PDL** is elementary (it is EXPTIME complete). Second, the main result that we shall establish here, the decidability of the dynamic logic of multidominance structures, cannot be established in this way, as far as we can see. For this reason we shall use dynamic logic.

5. The Logic of Multidominance Structures

For notation, let us agree on the following. For each of the relations $\succ_{ij}$ we introduce a program ∇_{ij}, which is interpreted by a relation that we write $\succ_{ij}$ or $\lhd_{ij}$ rather than $R(\nabla_{ij})$. Structures are of the form

$$(13) \qquad \langle M, \succ_{00}, \succ_{01}, \succ_{10}, \succ_{11} \rangle.$$

We use ∇_0 in place of $\nabla_{00} \cup \nabla_{01}$, ∇_1 for $\nabla_{10} \cup \nabla_{11}$ and ∇ for $\nabla_0 \cup \nabla_1$. The relations ∇_0 and ∇_1 are partial functions. Also, the notation $\nabla_{\bullet 0} := \nabla_{00} \cup \nabla_{10}$ and $\nabla_{\bullet 1} := \nabla_{01} \cup \nabla_{11}$ is frequently used. Finally, let us write

$$(14) \qquad u := \nabla^*$$

A structure is called **generated** if there is a single element w such that the least set containing w which is closed under taking successors along all basic relations is the entire set of worlds. (In our case this is exactly true if the structure is a constituent.) The following is easy to see.

Lemma 10. *Let $\mathfrak{M}$ be a generated $\mathbf{PDL}_n$-structure with root x. Then we have $\langle \mathfrak{M}, \beta, x \rangle \vDash [u]\varphi$ iff for all w: $\langle \mathfrak{M}, \beta, w \rangle \vDash \varphi$.*

Our first goal is axiomatise the logic of all PMDSs. There is an important tool that we shall use over and over.

Theorem 11. *Suppose that L is a logic containing $\mathbf{PDL}_n$ which has the finite model property, and let χ be a formula. Then the logic $L \oplus \chi$ also has the finite model property.*

Proof. Suppose that φ is consistent with $L \oplus \chi$. Then $\varphi; [u]\chi$ also is $L \oplus \chi$-consistent, and a fortiori L-consistent. Thus it has a finite model $\langle \langle F, R \rangle, \beta, x \rangle$. We may actually assume that for every y, $x \overset{u}{\to} y$. Then $y \vDash \chi$, and so the frame is a frame for $L \oplus \chi$, since χ is constant. $\square$

This theorem has interesting consequences worth pointing out. It allows us to focus on the grammar rather than the lexicon. This reduces the problem to some degree.

Definition 12. *Let $\mathbf{PM} := \mathbf{DPDL}_4.\mathbf{f} \oplus \langle \nabla_{10} \rangle \top \to [\nabla_{11}]\top \oplus \langle \nabla_{00} \rangle p \to [\nabla_{01}]\top \oplus \langle \nabla_{\bullet 1} \rangle p \to \langle \nabla_{\bullet 0}^+; \nabla \rangle p.$*

The first two formulae make sure that each node has at most one left hand daughter and at most one right hand daughter. The postulates are constant and can be added without sacrificing decidability, by Theorem 11.

The third postulate ensures that the structures are trees, more exactly, they are generated from a single node via $\succ_{\bullet 0}$ (rather than all four relations). This is because if $z \prec_{\bullet 1} y$ then there is a path along nonderived links to y, as we shall show.

Lemma 13. *Suppose $\mathfrak{F}$ is a structure in which every basic program is deterministic and loop free and satisfies*

> *For all w, u: if $w \succ_{\bullet 1} u$ then there is a y such that $y \prec^+_{\bullet 0} w$ and $w \prec y$.*

*Then $\mathfrak{F}$ is a structure for **PM**.*

Proof. It suffices to show that this structure satisfies the additional axiom of **PM**. Choose a valuation β and a point w such that

$$(15) \qquad\qquad \langle \mathfrak{F}, \beta, w \rangle \vDash \langle \nabla_{\bullet 1} \rangle p$$

Then either $w \vDash \langle \nabla_{01} \rangle p$ or $w \vDash \langle \nabla_{11} \rangle p$. Then there is a u such that $u \prec_{01} w$ and $u \vDash p$. By assumption on $\mathfrak{F}$, there is a y such that $y \prec^+_{\bullet 0} w$ and $u \prec y$. From the latter we get $y \vDash \langle \nabla \rangle p$, and from the first

$$(16) \qquad\qquad \langle \mathfrak{F}, \beta, y \rangle \vDash \langle \nabla^+_{\bullet 0} \rangle \langle \nabla \rangle p$$

This shows the claim. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

From this we prove that the axioms of **PM** are valid in all PMDSs (Lemma 14). This is one half of the characterization, Theorem 20, which asserts that if a finite structure satisfies the axioms of **PM** then it is actually a PMDS. The other half is constituted by Lemma 19.

Lemma 14. *PMDSs are **PM**-structures.*

Proof. To see this, we shall verify that they satisfy the property given in Lemma 13. To this end, take a PMDS $\langle M, \succ_{00}, \succ_{01}, \succ_{10}, \succ_{11} \rangle$. Suppose that $x \succ_{\bullet 1} y$. Then $x \in M(y)$, and there is, by assumption, an element $u \in M(y)$ such that $u \prec^+ x$. (Notice that by (P6) of Definition 2, x cannot be the least element in $M(y)$ with respect to $\prec^+$ since the link $\langle x, y \rangle$ is derived.) Choose a path $\Pi_0 = u; \cdots; x$. If this path contains only underived links, we are done. Otherwise, let the path contain $v; v'$, a derived link. Then there is a path $\Delta = v; \cdots; w; v'$ such that $w \prec_{\bullet 0} v'$, by a similar argument. Replace the pair $v; v'$ in Π_0 by Δ. This gives a path which is longer than Π_0. Thus, as long as we have derived links we can replace them, increasing the length of the path. Hence, as the procedure must end, it will deliver a path without derived links, as promised. $\qquad\qquad\square$

In the connection of the following lemma, we say that $R(\alpha)$ satisfies a **fixed point property** if for all formulae φ, frames $\mathfrak{F}$, valuations β and points x:

$$(17) \qquad \langle \mathfrak{F}, \beta, x \rangle \vDash \langle \alpha^* \rangle \varphi \leftrightarrow \varphi \vee \langle \alpha; \alpha^* \rangle \varphi$$

Lemma 15. *Let $\langle F, R \rangle$ be a finite frame β a valuation, and $R(\alpha)$ be loop free. Then for all x and φ:*

$$(18) \qquad \langle \mathfrak{F}, \beta, x \rangle \vDash \langle \alpha^* \rangle \varphi \leftrightarrow \varphi \vee \langle \alpha; \alpha^* \rangle \varphi$$

Proof. It follows from the axioms that $\varphi \to \langle \alpha^* \rangle \varphi$ and $\langle \alpha; \alpha^* \rangle \varphi \to \langle \alpha^* \rangle \varphi$ are generally valid. Hence we only have to establish

$$(19) \qquad \langle \mathfrak{F}, \beta, x \rangle \vDash \langle \alpha^* \rangle \varphi \to \varphi \vee \langle \alpha; \alpha^* \rangle \varphi$$

By assumption on $R(\alpha)$, for every x there is a sequence $x = x_0 \xrightarrow{\alpha} x_1 \xrightarrow{\alpha} x_2 \cdots \xrightarrow{\alpha} x_n$ where x_n has no $R(\alpha)$-successor. We proceed by induction on maximum length of such a chain starting at x. Call this the **height** of x. If the height is 0, x has no $R(\alpha)$-successors. Then $\langle \alpha; \alpha^* \rangle \varphi$ is false, and so the claim reduces to

$$(20) \qquad \langle \mathfrak{F}, \beta, x \rangle \vDash \langle \alpha^* \rangle \varphi \to \varphi$$

which is correct. Now let x be of height $n + 1$ and the claim proved for all points of height $\leq n$. Suppose $\langle \alpha^* \rangle \varphi$ is true at x. Then there is a chain of length $\leq n + 1$: $x = x_0 \xrightarrow{\alpha} x_1 \xrightarrow{\alpha} x_2 \cdots \xrightarrow{\alpha} x_k$, and φ is true at x_k. Two cases arise. $k = 0$, in which case $x \vDash \varphi$ and we are done. Or $k > 0$. Then, by inductive hypothesis, since x_1 has height $\leq n$, $\langle \mathfrak{F}, \beta, x_1 \rangle \vDash \langle \alpha^* \rangle \varphi$ and so we have $x \vDash \langle \alpha; \alpha^* \rangle \varphi$, as promised. $\square$

Say that a program α is **progressive** in L if $R(\alpha)$ is loop free in every structure for L. In that case we say that a node x has α-height n if there is no sequence $x \xrightarrow{\alpha} x_1 \xrightarrow{\alpha} x_2 \cdots \xrightarrow{\alpha} x_{n+1}$. If x has α-height 0 it means that it has no α-successors. The important fact to note is that we can restrict ourselves in the present context to progressive programs, and these are the programs for which the fixed point equation holds. We say that α is contained in β, in symbols $\alpha \subseteq \beta$, if $L \vdash \langle \alpha \rangle p \to \langle \beta \rangle p$. If L has the finite model property this is equivalent to $R(\alpha) \subseteq R(\beta)$ in every finite L-structure. If $L' \supseteq L$ and $\alpha \subseteq \beta$ in L, then this holds also in L', so this does not dependent much on L. α and β are **equivalent in** L if $\alpha \subseteq \beta$ as well as $\beta \subseteq \alpha$ in L. If α is progressive then so are α^n ($n > 0$) and α^+. The following theorem rests on the fact that the logic of finite computations has a maximal progressive program.

Lemma 16. *In $\mathbf{PDL}_n.\mathbf{f}$ every program is equivalent to a program of the form $\varphi?$, $\varphi? \cup \alpha$, where α is progressive.*

Proof. Notice that α is equivalent to $\bot ? \cup \alpha$, so we need no separate case for progressive programs. Let ζ_i, $i < n$, be the basic modalities. Put

$$(21) \qquad \gamma = (\zeta_0 \cup \zeta_1 \cup \cdots \cup \zeta_{n-1})^+$$

In $\mathbf{PDL_n.f}$, γ is progressive. Then $\gamma; \gamma$ as well γ^+ are likewise progressive. Every η that is contained in a progressive program is also progressive. What we shall show is that every program η that is not a test can be written as $\varphi ? \cup \alpha$ where α is contained in γ. Before we start notice that if χ is a test and $\alpha \subseteq \gamma$ then $\chi ?; \alpha \subseteq \alpha \subseteq \gamma$ and likewise $\alpha; \chi ? \subseteq \alpha \subseteq \gamma$.

We note that $\varphi ?; \chi ?$ is equivalent to $(\varphi \wedge \chi)?$ and that $\varphi ? \cup \chi ?$ is equivalent to $(\varphi \vee \chi)?$. Finally, $(\varphi ?)^*$ is equivalent to $\varphi ?$, so that the operators reduce on tests to a single test. Now, suppose that $\eta_1 = \varphi_1 ? \cup \alpha_1$ and $\eta_2 = \varphi_2 ? \cup \alpha_2$ with α_1, α_2 contained in γ. Then

$$(22) \qquad \begin{aligned} \eta_1 \cup \eta_2 &= (\varphi_1 ? \cup \alpha_1) \cup (\varphi_2 ? \cup \alpha_2) \\ &= (\varphi_1 \vee \varphi_2)? \cup (\alpha_1 \cup \alpha_2) \end{aligned}$$

is of the desired form.

$$(23) \qquad \begin{aligned} \eta_1; \eta_2 &= (\varphi_1 ? \cup \alpha_1); (\varphi_2 ? \cup \alpha_2) \\ &= (\varphi_1 ?; \varphi_2 ?) \cup (\varphi_1 ?; \alpha_2) \cup (\varphi_1 ?; \alpha_2) \cup (\alpha_1; \alpha_2) \\ &\subseteq (\varphi_1 \wedge \varphi_2)? \cup (\varphi_1 ?; \alpha_1 \cup \alpha_2; \varphi_2 ? \cup \alpha_1; \alpha_2) \end{aligned}$$

which is again of the desired form. Finally, let $\eta = \varphi ? \cup \alpha$. We observe that $\eta \subseteq \top ? \cup \alpha$. Furthermore, since star is monotone, $\eta^* \subseteq (\top ? \cup \alpha)^* = \top ? \cup \alpha^+$. Now, $\alpha \subseteq \gamma$, and so $\alpha^+ \subseteq \gamma^+ \subseteq \gamma$, since γ is progressive.

$\square$

Definition 17. *The **Fisher Ladner closure** of a formula is defined as follows.*

    (1) *If $\chi \wedge \psi \in \mathrm{FL}(\varphi)$ then $\chi, \psi \in \mathrm{FL}(\varphi)$.*
    (2) *If $\neg\chi \in \mathrm{FL}(\varphi)$ then $\chi \in \mathrm{FL}(\varphi)$.*
    (3) *If $\langle \alpha \cup \beta \rangle \chi \in \mathrm{FL}(\varphi)$ then $\langle \alpha \rangle \chi, \langle \beta \rangle \chi \in \mathrm{FL}(\varphi)$.*
    (4) *If $\langle \alpha; \beta \rangle \chi \in \mathrm{FL}(\varphi)$ then $\langle \alpha \rangle \langle \beta \rangle \chi \in \mathrm{FL}(\varphi)$.*
    (5) *If $\langle \alpha^* \rangle \chi \in \mathrm{FL}(\varphi)$ then $\chi, \langle \alpha \rangle \langle \alpha^* \rangle \chi \in \mathrm{FL}(\varphi)$.*
    (6) *If $\langle \psi ? \rangle \chi \in \mathrm{FL}(\varphi)$ then $\psi, \chi \in \mathrm{FL}(\varphi)$.*
    (7) *If $\langle \alpha \rangle \chi \in \mathrm{FL}(\varphi)$, α basic, then $\chi \in \mathrm{FL}(\varphi)$.*

We remark that $|\mathrm{FL}(\varphi)|$ is linear in the length of φ. This is shown by induction on φ. This means that complexity can be measured either in terms of the size of the formula or in terms of the size of $\mathrm{FL}(\varphi)$.

By collecting all the previous facts together one can actually show with a little effort the following (see [16]).

Theorem 18. $\mathbf{DPDL_n.f}$ *has the finite model property, and is decidable in EXPTIME.*

Now let $At(\varphi)$ be the set of all conjunctions of formulae (or their negations) from the Fisher Ladner closure of φ. Set

$$(24) \qquad X(\varphi) := \{\langle \nabla_{\bullet 1} \rangle \delta \rightarrow \langle \nabla_{\bullet 0}^+; \nabla \rangle \delta : \delta \in At(\varphi)\}$$

Lemma 19. φ *is consistent with* **PM** *iff* $\varphi; [u]X(\varphi)$ *is consistent with* **DPDL**$_4$**.f.**

Proof. ($\Rightarrow$). If $\varphi; [u]X(\varphi)$ is inconsistent in **DPDL**$_\mathbf{n}$**.f**, $\neg\varphi$ can be proved from $[u]X(\varphi)$ in **DPDL**$_4$**.f**. However, $[u]X(\varphi)$ can be proved in **PM**. Hence $\neg\varphi$ is provable in **PM**. ($\Leftarrow$). Now let us suppose that $\varphi; [u]X(\varphi)$ is **DPDL**$_4$**.f**-consistent. Then by Theorem 8 it has a finite model based on a frame

$$(25) \qquad \mathfrak{M} = \langle M, \succ_{00}, \succ_{01}, \succ_{10}, \succ_{11} \rangle$$

with root w_0 and valuation β. So,

$$(26) \qquad \langle \mathfrak{M}, \beta, w_0 \rangle \vDash \varphi; [u]X(\varphi)$$

We may assume that the relation $\prec_{\bullet 0}$ induces a tree ordering on the set of worlds, though with multiple roots (this we have what is known as a forest). We shall construct a finite **PM**-model from this. Let S be the closure of w_0 under the relation $\succ_{\bullet 0}$, that is, S is the least set which contains w_0 and is closed under $\succ_{\bullet 0}$. Members of S are called **standard points**. Let

$$(27) \qquad E := \{w : \text{there is } v \in S \text{ such that } w \prec_{\bullet 1} v\}$$

For a point w, let $a(w)$ be the unique $\delta \in At(\varphi)$ such that

$$(28) \qquad \langle \mathfrak{M}, \beta, w \rangle \vDash \delta$$

Now choose a $w \in E$. Let v be a standard world such that $w \prec_{\bullet 1} v$. By choice of $X(\varphi)$,

$$(29) \qquad \langle \mathfrak{M}, \beta, w_0 \rangle \vDash [u](\langle \nabla_{\bullet 1} \rangle a(w) \rightarrow \langle \nabla_{\bullet 0}^+; \nabla \rangle a(w))$$

where w_0 is the root, so

$$(30) \qquad \langle \mathfrak{M}, \beta, v \rangle \vDash \langle \nabla_{\bullet 1} \rangle a(w) \rightarrow \langle \nabla_{\bullet 0}^+; \nabla \rangle a(w)$$

Since $a(w)$ is true at w and since $w \prec_{\bullet 1} v$, we have

$$(31) \qquad \langle \mathfrak{M}, \beta, v \rangle \vDash \langle \nabla_{\bullet 0}^+; \nabla \rangle a(w)$$

Hence there is a standard $u \prec_{\bullet 0}^+ v$ and $u^* \prec u$ such that $a(u^*) = a(w)$. For each w, pick such a point and say that it is **linked from** w and write $w L u^*$. Thus, L establishes a relation from E to $E \cup S$. We note the following. $w L u^*$ does not mean that u^* is standard. However, u^* has greater standard depth as w, and if $u \notin S$ then $u^* \in E$ and so u^* can in turn be linked to some node. It follows that for every $w \in E$ there is a standard u such that $w L^+ u$. For suppose not. Then there is a $w \in E$ of maximal depth which cannot be linked to a standard point. But it can be linked to a point in E. The latter has greater depth. Contradiction.

Now we define a new frame $\mathfrak{S}$ as follows. The set of points is S. Put $x \lhd_{00} y$ iff $x \prec_{00} y$, $x \lhd_{10} y$ iff $x \prec_{10} y$; put $x \lhd_{01} y$ iff there is a u such that $u \prec_{01} y$ and $u\ L^+\ x$; $x \lhd_{11} y$ iff there is a u such that $u \prec_{11} y$ and $u\ L^+\ x$. Finally,

$$(32) \qquad\qquad \mathfrak{S} := \langle S, \rhd_{00}, \rhd_{01}, \rhd_{10}, \rhd_{11} \rangle$$

The valuation β' is defined by $\beta'(p) := \beta(p) \cap S$. (If constants are present, the value of a constant c in $\mathfrak{S}$ is the value of c intersected with S.) We shall prove for every $w \in S$ and every $\chi \in \mathrm{FL}(\varphi)$:

$$(33) \qquad\qquad \langle \mathfrak{S}, \beta', w \rangle \vDash \chi \qquad \Leftrightarrow \qquad \langle \mathfrak{M}, \beta, w \rangle \vDash \chi$$

The basic clause is
(Case 1.) $\chi = p$, a variable. Then $\langle \mathfrak{S}, \beta', w \rangle \vDash p$ iff $w \in \beta'(p)$ iff $w \in \beta(p)$ iff $\langle \mathfrak{M}, \beta, w \rangle \vDash p$, by definition of β'.
(Case 2.) $\chi = \neg \vartheta$.

$$
\begin{aligned}
(34) \qquad\qquad \langle \mathfrak{S}, \beta', w \rangle \vDash \neg\vartheta \text{ iff } & \langle \mathfrak{S}, \beta', w \rangle \nvDash \vartheta \\
\text{iff } & \langle \mathfrak{M}, \beta, w \rangle \nvDash \vartheta \\
\text{iff } & \langle \mathfrak{M}, \beta, w \rangle \vDash \neg\vartheta
\end{aligned}
$$

(Case 3.) $\chi = \vartheta \wedge \vartheta'$.

$$
\begin{aligned}
(35) \qquad\qquad \langle \mathfrak{S}, \beta', w \rangle \vDash \vartheta \wedge \vartheta' \text{ iff } & \langle \mathfrak{S}, \beta', w \rangle \vDash \vartheta; \vartheta' \\
\text{iff } & \langle \mathfrak{M}, \beta, w \rangle \vDash \vartheta; \vartheta' \\
\text{iff } & \langle \mathfrak{M}, \beta, w \rangle \vDash \vartheta \wedge \vartheta'
\end{aligned}
$$

Now let $\chi = \langle \alpha \rangle \vartheta$. The claim will be proved by induction on the syntactic complexity of α.
(Case 4.) $\alpha = \alpha' \cup \alpha''$.

$$
\begin{aligned}
(36) \qquad \langle \mathfrak{S}, \beta', w \rangle \vDash \langle \alpha' \cup \alpha'' \rangle \vartheta' \text{ iff } & \langle \mathfrak{S}, \beta', w \rangle \vDash \langle \alpha' \rangle \vartheta \vee \langle \alpha'' \rangle \vartheta \\
\text{iff } & \langle \mathfrak{M}, \beta, w \rangle \vDash \langle \alpha' \rangle \vartheta \vee \langle \alpha'' \rangle \vartheta \\
\text{iff } & \langle \mathfrak{M}, \beta, w \rangle \vDash \langle \alpha' \cup \alpha'' \rangle \vartheta
\end{aligned}
$$

(Case 5.) $\alpha = \alpha'; \alpha''$.

$$
\begin{aligned}
(37) \qquad \langle \mathfrak{S}, \beta', w \rangle \vDash \langle \alpha'; \alpha'' \rangle \vartheta \text{ iff } & \langle \mathfrak{S}, \beta', w \rangle \vDash \langle \alpha' \rangle \langle \alpha'' \rangle \vartheta \\
\text{iff } & \langle \mathfrak{M}, \beta, w \rangle \vDash \langle \alpha' \rangle \langle \alpha'' \rangle \vartheta \\
\text{iff } & \langle \mathfrak{M}, \beta, w \rangle \vDash \langle \alpha'; \alpha'' \rangle \vartheta
\end{aligned}
$$

We use (i) the fact that α' is syntactically less complex than $\alpha'; \alpha''$ and (ii) the inductive hypothesis for $\langle \alpha'' \rangle \vartheta$.

(Case 6.) $\alpha = \psi?$.

$$(38) \qquad \langle \mathfrak{S}, \beta', w \rangle \vDash \langle \psi? \rangle \vartheta \text{ iff } \langle \mathfrak{S}, \beta', w \rangle \vDash \psi; \vartheta$$
$$\text{iff } \langle \mathfrak{M}, \beta, w \rangle \vDash \psi; \vartheta$$
$$\text{iff } \langle \mathfrak{M}, \beta, w \rangle \vDash \langle \psi? \rangle \vartheta$$

Using the inductive assumptions on ψ and ϑ.

(Case 7.) $\alpha = \alpha'^*$. Now, in virtue of Lemma 16 we may assume that α' is progressive, so

$$(39) \qquad \langle \alpha'^* \rangle \chi \leftrightarrow \chi \vee \langle \alpha' \rangle \langle \alpha'^* \rangle \chi$$

is a theorem of **PDL**. Further, α' is of lesser complexity than α'^*.

$$(40) \qquad \langle \mathfrak{S}, \beta', w \rangle \vDash \langle \alpha'^* \rangle \vartheta \text{ iff } \langle \mathfrak{S}, \beta', w \rangle \vDash \vartheta \vee \langle \alpha' \rangle \langle \alpha'^* \rangle \vartheta$$
$$\text{iff } \langle \mathfrak{M}, \beta, w \rangle \vDash \vartheta \vee \langle \alpha' \rangle \langle \alpha'^* \rangle \vartheta$$
$$\text{iff } \langle \mathfrak{M}, \beta, w \rangle \vDash \langle \alpha'^* \rangle \vartheta$$

(Case 8.) $\alpha = \nabla_{00}$. Then the claim follows since $\prec_{00} = \lhd_{00}$.

(Case 9.) $\alpha = \nabla_{10}$. Likewise.

(Case 10.) $\alpha = \nabla_{01}$. We show first ($\Rightarrow$) in (33). $\langle \mathfrak{S}, \beta', w \rangle \vDash \langle \nabla_{01} \rangle \vartheta$ implies that there is a $v \lhd_{01} w$ such that $\langle \mathfrak{S}, \beta', v \rangle \vDash \vartheta$. v is standard, and by induction hypothesis, $\langle \mathfrak{M}, \beta, v \rangle \vDash \vartheta$. By construction, $w >_{01} u$ for a $u \in E$ such that $u \ L^+ \ v$. This means that $a(u) = a(v)$ and so $\langle \mathfrak{M}, \beta, u \rangle \vDash \vartheta$; hence $\langle \mathfrak{M}, \beta, w \rangle \vDash \langle \nabla_{01} \rangle \vartheta$. Now we show ($\Leftarrow$) in (33). Assume $\langle \mathfrak{M}, \beta, v \rangle \vDash \langle \nabla_{01} \rangle \vartheta$ and $v \in S$. Then there is a $w \in E$ such that $w \prec_{01} v$ and $\langle \mathfrak{M}, \beta, w \rangle \vDash \vartheta$. By construction there is a standard u such that $w \ L^+ \ u$, and so $\langle \mathfrak{M}, \beta, u \rangle \vDash \vartheta$, since $a(u) = a(w)$. By inductive hypothesis, $\langle \mathfrak{S}, \beta, u \rangle \vDash \vartheta$. Again by construction, $v \rhd_{01} u$, so $\langle \mathfrak{S}, \beta, v \rangle \vDash \langle \nabla_{01} \rangle \vartheta$.

(Case 11.) $\alpha = \nabla_{11}$. Similar.

The next step is to verify that $\mathfrak{S}$ is an **PM**-frame. To that effect we have to ensure that the union of the basic programs is deterministic and loop free and that the structure satisfies the new axiom. First, let $w \in S$. Recall the definition of the standard depth. It is easy to see that the standard depth of points is the same in both structures. Now suppose that $w \lhd u$. We claim that $sd(w) > sd(u)$. (Case 1.) $w \lhd_{\bullet 0} u$. Then $w \prec_{00} u$ or $w \prec_{10} u$, and by definition of standard depth, $sd(w) = 1 + sd(u)$. (Case 2.) $w \lhd_{01} u$ or $w \lhd_{11} u$. In this case there is a y such that $w >_{01} y$ or $w >_{11} y$ such that $y \ L^+ \ u$ and $w >^+ u'$ for some standard u'. This means that $sd(u) \geq 2 + sd(w)$. Next, to show that the programs are deterministic, observe that the original programs were deterministic, and each link was replaced by just one link. Finally, from Lemma 13 it follows that the constructed structure satisfies **PM**.

Now, from (33) it follows that

$$(41) \qquad \langle \mathfrak{S}, \beta, w_0 \rangle \vDash \varphi$$

This shows the claim.   □

Theorem 20. *The logic of PMDSs is* **PM**. *Moreover, this logic has the finite model property, if finitely axiomatisable and therefore decidable. Its complexity is 2EXPTIME.*

The complexity follows from the fact that the formula to be satisfied has length $O(2^n)$, and that **DPDL$_4$.f** is in EXPTIME.

6. Single Movement MDSs

There is an important class of MDSs, those where $M(x)$ has at most two elements. This means in practice that each element is allowed to move only once. This class of structures is very important, since the now current Minimalist Program requires each movement step to be licensed. These structures be the topic of Part II. Here we are interested only the axiomatisation of these structures. We have noted earlier that root links are always the lowest links. Therefore, for every node x there is at most one y such that $x \prec_{\bullet 0} y$. On the other hand there can be any number of non-root links. The narrowness determines the maximum number of non-root links.

$$(42) \qquad v(p) := (p \to [\nabla^+]\neg p) \wedge \neg(\langle \nabla_{00}; \nabla^* \rangle p \wedge \langle \nabla_{10}; \nabla^* \rangle p)$$

Lemma 21. *Let β be a valuation such that $\langle \mathfrak{F}, \beta \rangle \vDash [u]v(p)$. Then $|\beta(p)| \le 1$.*

Proof. Suppose that $x, y \in \beta(p)$. Then $x \prec_{\bullet 0}^+ y$ cannot hold; for then $y \vDash p$ but $y \nvDash [\nabla^+]\neg p$. If however x and y are incomparable there are points u, v and v' such that $v \ne v'$ and $x \prec^+ v \prec u$ as well as $y \prec^+ v' \prec u$. Then however $u \vDash \langle \nabla_{00}; \nabla^+ \rangle p; \langle \nabla_{10}; \nabla^+ \rangle p$.   □

Definition 22. *An MDS is called n-**narrow** if $|M(x)| \le n - 1$ for all x. An MDS is called **narrow** if it is 1-narrow.*

$$(43) \qquad \begin{aligned} \xi(p) :=& [u]v(p) \\ &\to [u](\langle \nabla_{\bullet 1} \rangle p \to [\nabla_{\bullet 0}; \nabla^*](\langle \nabla \rangle p \to \langle \nabla_{\bullet 0} \rangle p)) \end{aligned}$$

Lemma 23. *A MDS satisfies $\xi(p)$ iff it is narrow.*

Proof. Suppose the MDS is not narrow. Then there is a y and $z, z' \in M(y)$ such that $z \prec^+ z'$ and both links $y; z$ and $y; z'$ are not root links. Then put $\beta(p) := \{y\}$. Then throughout the MDS, $p \to [d^+]\neg p$ holds. Also, there is no point u such that $u >_{00} v$, $u >_{10} v'$ and $y \prec^* v; v'$. It follows that $z \vDash \langle \nabla_{\bullet 1} \rangle p; \neg \langle \nabla \rangle p$ and $z' \vDash \langle \nabla_{\bullet 1} \rangle p$. However, $z' \, R(\nabla_{\bullet 0}; \nabla^*) \, z$. So the formula

is false under this valuation. Now assume that the MDS is narrow. Take a valuation such that $\nu(p)$ everywhere. Let y be a node such that $y \vDash \langle \nabla_{\bullet 1} \rangle p$, say $u \prec_{\bullet 1} y$ is such that $u \vDash p$. Then $\beta(p) = \{u\}$ by Lemma 21. Now let z and z' be such that $z \leq z' \prec_{\bullet 0} y$ and $z \vDash \langle \nabla \rangle p$. Then $z > u$. Since the structure is narrow, $u \prec_{\bullet 0} z$, showing $z \vDash \langle \nabla_{\bullet 0} \rangle p$. □

7. Extending and Reducing the Language

The fact that we are dealing with cycle free structures has great effect on the expressivity of the language; basically, using implicit definitions all programs constructors of PDL can be eliminated; conversely, many seemingly more powerful constructs can be effectively mimicked. We consider here two devices: nominals and the converse. A **nominal** is a variable that can be true only at a single world. It emerges from the discussion above that nominals actually do not add any expressive strength to our language. Consider a formula $\varphi(i)$ which contains a nominal i. Now consider the formula

$$(44) \qquad \nu(p) \wedge \langle \nabla^+ \rangle p \to \varphi(p/i)$$

This formula has a model $\langle \mathfrak{F}, \beta, x \rangle$ only if $\beta(p)$ is a singleton. The consequence of this is the following

Theorem 24. *For every first-order universal formula ζ using atomic formulae of the form $x\, R(\alpha)\, y$ or $x = y$ there is a modal formula φ such that for any MDS, $\mathfrak{F} \vDash \zeta$ iff $\mathfrak{F} \vDash \varphi$.*

Proof. Let $\zeta = (\forall x_0 x_1 \cdots x_{n-1})\alpha$. Introduce nominals $i_0, i_1, \cdots, i_{n-1}$ and define the following translation:

$$(45) \qquad \begin{aligned} (x_p = x_q)^\dagger &:= \langle \nabla^* \rangle (i_p \wedge i_q) \\ (x_p\, R(\alpha)\, x_q) &:= \langle \nabla^* \rangle (i_p \wedge \langle \alpha \rangle i_q) \\ (\neg \alpha)^\dagger &:= \neg \alpha^\dagger \\ (\alpha \wedge \alpha')^\dagger &:= \alpha^\dagger \wedge \alpha'^\dagger \end{aligned}$$

It is not hard to see that $\langle \mathfrak{F}, \beta, x \rangle \vDash \neg \alpha^\dagger$ iff $\mathfrak{F} \nvDash \zeta$. The sought after formula is

$$(46) \qquad \nu(p_0) \wedge \nu(p_1) \wedge \cdots \nu(p_{n-1}) \to \alpha^\dagger([p_k/i_k : i < n])$$

This completes the proof. □

Also, let me recall a few other reductions that we have achieved. The following equivalences hold:

$$(47) \qquad \langle \alpha \cup \alpha' \rangle p \leftrightarrow \langle \alpha \rangle p \vee \langle \alpha' \rangle p$$

$$(48) \qquad \langle \alpha ; \alpha' \rangle p \leftrightarrow \langle \alpha \rangle \langle \alpha' \rangle p$$

$$(49) \qquad \langle \varphi? \rangle \chi \leftrightarrow \varphi \wedge \chi$$

This means that the program constructs union, concatenation and test are eliminable if they occur as outermost program constructors. However, we have also shown that every program is a union of a test and a progressive program and that for progressive programs the following holds in finite structures:

$$(50) \qquad \langle \alpha^* \rangle p \leftrightarrow p \vee \langle \alpha \rangle \langle \alpha^* \rangle p$$

This allows to eliminate the star as follows:

Lemma 25. *Let α be progressive in $\mathfrak{F}$ and χ a formula. Then*

$$(51) \qquad \langle \mathfrak{F}, x \rangle \vDash \chi(\langle \alpha^* \rangle \delta / q) \Leftrightarrow \langle \mathfrak{F}, x \rangle \vDash [u](q \leftrightarrow \delta \vee \langle \alpha \rangle q) \rightarrow \chi$$

Proof. Suppose $\langle \mathfrak{F}, \beta, x \rangle \nvDash [u](q \leftrightarrow \delta \vee \langle \alpha \rangle q)$. Then by induction on the α-height it is shown that $\beta(q) = \{x : x \vDash \langle \alpha^* \rangle \delta\}$. (Assume that the claim is shown for all point of height less than the height of x. $x \vDash q$ iff either $x \vDash \delta$ in which case $x \vDash \langle \alpha^* \rangle \delta$; or $x \vDash \langle \alpha \rangle q$, in which case there is an y of lesser height such that $y \vDash q$. By inductive hypothesis, $y \vDash \langle \alpha^* \rangle \delta$ and so $x \vDash \langle \alpha^* \rangle \delta$ in this case as well.) Now assume that $\langle \mathfrak{F}, x \rangle \nvDash [u](q \leftrightarrow \delta \wedge \langle \alpha \rangle q) \rightarrow \chi$. Then $x \vDash [u](q \leftrightarrow \delta \wedge \langle \alpha \rangle q); \neg \chi$, from which we get that $\beta(q) = \beta(\langle \alpha \rangle \delta)$. So we substitute and get $x \vDash \neg \chi(\langle \alpha^* \rangle \delta / q)$. This argument is reversible. $\qquad \square$

Notice that it is irrelevant whether q actually occurs in χ or not. We shall strengthen this language further by adding an operator on programs, the **converse**. This will allow to talk about going up the tree. This makes the statement of some restrictions easier. We shall show that for a large enough portion of the newly added formulae, they do not add expressive power, they just make life easier. The good news about them is that they can be added without having to redo the proofs.

Recall that for a binary relation R,

$$(52) \qquad R^\smile := \{\langle y, x \rangle : x \, R \, y\}$$

The language **PDL**$^\smile$ has in addition a unary operator $\smile$, and the interpretation of $\alpha^\smile$ is $R(\alpha)^\smile$, the converse of $R(\alpha)$. It is axiomatised by **PDL** for all programs plus for every program α:

$$(53) \qquad p \rightarrow [\alpha] \langle \alpha^\smile \rangle p, \qquad p \rightarrow [\alpha^\smile] \langle \alpha \rangle p$$

It turns out that it is enough to just add the converse for every elementary program, for we have

$$(54) \qquad (R \cup S)^\smile := R^\smile \cup S^\smile$$

$$(55) \qquad (R \circ S)^\smile := S^\smile \circ R^\smile$$

$$(56) \qquad (R^*)^\smile := (R^\smile)^*$$

Also, notice that

$$(57) \qquad R((\varphi?)^\smile) = R(\varphi?)$$

Thus, rather than looking at $\mathbf{PDL}_4^{\smile}$ (four basic programs and a converse operator) we may look at $\mathbf{PDL}_8$ (eight basic programs, no converse), where the postulates (53) have been added just for the basic programs. We shall not take that route, though, since it produces needless complications. Rather, we shall make the following observation.

Lemma 26. *Let $\mathfrak{F} = \langle F, R \rangle$ be a frame, $x \in F$ a world, and $\square$, $\boxminus$, and $\boxtimes$ modalities such that $R(\boxminus) = R(\square)^{\smile}$ is a partial function and $\boxtimes$ an operator such that $x\ R(\boxtimes)\ y$ for all y. Then for any two formulas χ and δ and any valuation β:*

$$(58) \quad \langle \mathfrak{F}, \beta, x \rangle \vDash \boxtimes((\boxminus\bot \to q) \wedge \Diamond\top \to (\delta \to \square q) \wedge (\neg\delta \to \square\neg q)) \to \chi$$

iff

$$(59) \qquad\qquad \langle \mathfrak{F}, \beta, x \rangle \vDash \chi(\boxminus\delta/q)$$

Proof. Assume (58). We claim that $\beta(q) = \{u : u \vDash \boxminus\delta\}$. Two cases arise. (Case 1.) z has no $R(\boxminus)$-successor. Then $\boxminus\bot$ is true at z and so is q. (Case 2.) z has a $R(\boxminus)$-successor. Then this successor is unique by assumption. Call it y. By assumption we have $y\ R(\square)\ z$. Furthermore, as $x\ R(\boxtimes)\ y$, we have $y \vDash \delta \to \square q$ as well as $y \vDash \neg\delta \to \square\neg q$. Suppose $z \vDash \boxminus\delta$. Then $y \vDash \delta$, from which $y \vDash \square q$, and so $z \vDash q$. If $z \vDash \neg\boxminus\delta$ then $y \vDash \neg\delta$, by functionality of $R(\boxminus)$. Hence $y \vDash \square\neg q$ and so $z \vDash \neg q$. This argument is reversible. $\square$

This lemma can be used to introduce converses for the programs $\triangledown_{00}$ and $\triangledown_{10}$, since they are backwards deterministic. This seemingly allows for the reduction of any program to a forward looking program. However, recall that the elimination of star used the fact that every program is basically progressive. With converses added this is no longer the case. So, star is eliminable only if the program either contains only downward looking modalities or only upward looking modalities. Tests may be considered to belong to either class. Call such a formula **finite turn**.

Theorem 27. *Suppose a class of constituents is axiomatisable with some finite turn axioms using the operators $\triangle_{00}$ and $\triangle_{10}$ in addition to $\triangledown_{ij}$. Then it can be axiomatised without the use of $\triangle_{00}$ and $\triangle_{10}$.*

This can be used in the following way. We have said earlier that the PMDSs are not necessarily ordered in the standard sense. To enforce this we need to add another postulate. The linear order is modally definable by

$$(60) \qquad\qquad \alpha := \triangle^{\smile}; \triangle_0; \triangledown_1; \triangledown^*$$

In the definition we have made use of upward looking programs. It is straightforward to verify that

$$(61) \qquad\qquad x \sqsubset y \quad \Leftrightarrow \quad \langle x, y \rangle \in R(\alpha)$$

This would ordinarily involve adding the converse operator. We have seen, however, that there is a way to consider the converse operators as abbreviations. Thus we may define the following.

Definition 28. *Let*

$$(62) \qquad \mathbf{OL} := \quad \mathbf{PM} \oplus \langle \triangle_{00} \rangle \top \rightarrow [\triangle_{10}] \bot \oplus \langle \triangle_{10} \rangle \top \rightarrow [\triangle_{00}] \bot$$
$$\oplus \langle \triangle_{00} \rangle p \rightarrow [\triangle_{00}] p \oplus \langle \triangle_{10} \rangle p \rightarrow [\triangle_{10}] p$$

The way to understand this definition is to supplant the fomulae χ by $[u](H(\chi) \rightarrow p_\chi)$ if they contain the converse. (You may do that regardless of whether the converse occurs or not, but in the latter case this needlessly complicates the formulae.) Using Theorem 11 we see that

Theorem 29. OL *is decidable in EXPTIME.*

8. Nearness

The above results are encouraging, but not exactly what is needed. There typically is a restriction on the distance that an element can move in a single step. We take as our prime example the subjacency definition in [3]. As I have argued in [6], perhaps the best definition is this. The antecedent of a trace can be found within the next CP which contains the next IP properly containing the trace. One is tempted to cash this out as the following axiom.

$$(63) \qquad \langle \triangledown_{\bullet 1} \rangle p \rightarrow \langle \triangledown_{\bullet 0} \rangle \langle (\neg \mathrm{CP}?; \triangledown))^*; (\neg \mathrm{IP}; \triangledown)^* \rangle p$$

Here, CP, IP are constants, denoting phrasal nodes of category CP and IP. This formula says that for every node x, if there is a derived downward link from x to some y, then there is a path to y following first a nonderived link, then following down non-CP nodes and finally non-IP nodes. Unfortunately, matters are not that easy. I have shown in [8] that this gives the wrong results. Consider a VP and an NP that scrambles out of it. Consider a movement of the VP that passes the the NP, whereupon the NP moves to pass the VP again. Then the formula above maybe true even if there was a step that crossed a barrier. (It counts also nonexistent paths such as a movement that takes the NP for a ride inside the VP before it jumps out.) This structure can be complicated just a little bit so as to exclude the piggy-back movement entirely. Then the formula above is true even though no derivation of the desired kind exists.

So, nearness constraints are not easily captured in model theoretic terms because the structure does not explicitly say which link has been added before which other. Indeed, notice that one and the same MDS allows for quite different derivations. There is (up to inessential variations, see [10]) exactly one derivation that satisfies Freeze, and exactly one that satisfies Shortest Steps. As it turns out, however, at least Freeze derivations are

easy to characterize. The idea is that the longest path between two standard elements is actually the one following standard links. Suppose we want to define the subjacency domain for Freeze.

$$(64) \qquad \sigma = \langle \nabla_{\bullet 1} \rangle p \to \langle (\nabla_{\bullet 0}; \neg CP?)^{+}; (\nabla_{\bullet 0}; \neg IP?)^{+}; \nabla \rangle p$$

Lemma 30. $\mathfrak{M} \vDash \sigma$ *iff there is a Freeze derivation such that movement is within the IP $\circ$ CP-domain.*

Proof. Suppose that movement is such that each step is within the IP$\circ$CP-domain of the trace. Then in the MDS, every path between these nodes respects these domains. Conversely, let x be a node in an MDS and $y >_{\bullet 1} x$. Put $\beta(p) := \{x\}$. Then $y \vDash \langle \nabla_{\bullet 0} \rangle p$. Hence, by assumption,

$$(65) \qquad \langle \mathfrak{M}, \beta, y \rangle \vDash \langle (\nabla_{\bullet 0}; \neg CP?)^{+}; (\nabla_{\bullet 0}; \neg IP?)^{+}; \nabla \rangle p$$

which says that there is a standard path first along nodes that do no satisfy CP and then along nodes that do not satisfy IP to some node z which dominates x immediately. The standard path is the movement path in the Freeze derivation. This shows the theorem. $\square$

This can be generalized to any requirement that says that a path must respect a regular language, which is more general than the definable command relations of [4]. The general principle is therefore of the form

$$(66) \qquad \mathrm{Dist}(c; \oslash) = \langle \nabla_{\bullet 1} \rangle (c \wedge p) \to \langle \oslash; \nabla \rangle p$$

where c is a constant and $\oslash$ is an expression using only $\nabla_{\bullet 0}$ and constants. Moreover, as we shall see, one can mix these postulates to have a particular notion of distance for phrases and another one for heads, for example. In general, any mixture of distance postulates is fine, as long as it is finite.

Theorem 31. *The logic of MDSs which have a Freeze derivation satisfying a finite number of postulates of the form* $\mathrm{Dist}(R)$ *has the finite model property and is decidable.*

Proof. We replay the proof of Lemma 19. Let $\mathrm{Dist}(c_i; \oslash_i)$, $i < n$, be the distance postulates.

$$(67) \qquad Y(\varphi) := \{\langle \nabla_{\bullet 1} \rangle (c_i \wedge \delta) \to \langle \oslash_i; \nabla \rangle \delta : \delta \in \mathrm{At}(\varphi), i < n\}$$

Now define the linking in the following way. If $w <_{\bullet 1} u$ and $w \vDash c_i$, then

$$(68) \qquad u \vDash \langle \oslash_i; \nabla \rangle a(u)$$

Hence there are w', u' such that $u' <_{\bullet 0} u$, $w' < u'$ and the standard path from u to u' is contained in $\oslash_i$, and $a(w') = a(w)$. We then put $w\ L\ w'$. Thus, the condition on Freeze derivations is respected. The rest of the proof is the same. $\square$

9. First Example: Movement

We shall present an example of a language that is trans-context free and can be generated from a context free language through movement. Furthermore, it shall follow from our results that the logic of the associated structures is decidable. Take the following grammar.

$$
(69) \qquad
\begin{array}{ll}
S \to aT & S \to aX \\
T \to bU & X \to bc \\
U \to cS & S \to S
\end{array}
$$

This grammar generates the language $\{(\mathsf{abc})^n : n > 0\}$. Now, we shall allow for movement of any element into c-commanding position. Movement is only constrained by the fact that it is into c-commanding position, nothing else. Since we have added the rule $S \to S$, the base grammar freely generates sites to which a constituent can adjoin.

In order to implement this, we need to add constants. For each terminal and each nonterminal element there will be a constant denoted by underlining it; for example, $\underline{U}$ is the constant denoting nodes with label U. This will be our new language. We also add the condition that the constants from C are mutually exclusive:

$$
(70) \qquad \mathrm{Inc}(C) := \{\underline{X} \to \neg\underline{Y} : X \neq Y \text{ and } X, Y \in C\}
$$

Also, we express the fact at each node at least one constant from C must be true by

$$
(71) \qquad \mathrm{Suf}(C) := \bigvee \langle \underline{X} : X \in C \rangle
$$

These two together ensure that eacd node satisfies exactly one constant. Next the context free grammar is described by a set of rules:

$$
(72) \qquad
\begin{aligned}
\rho_S := \underline{S} \to \ & \langle \nabla_{00} \rangle \underline{a} \wedge \langle \nabla_{10} \rangle \underline{T} \\
& \vee \langle \nabla_{00} \rangle \underline{a} \wedge \langle \nabla_{10} \rangle \underline{X} \\
& \vee \langle \nabla_{00} \rangle \underline{S} \wedge \neg \langle \nabla_{10} \rangle \top \\
\rho_T := \underline{T} \to \ & \langle \nabla_{00} \rangle \underline{b} \wedge \langle \nabla_{10} \rangle \underline{U} \\
\rho_U := \underline{U} \to \ & \langle \nabla_{00} \rangle \underline{c} \wedge \langle \nabla_{10} \rangle \underline{S} \\
\rho_X := \underline{X} \to \ & \langle \nabla_{00} \rangle \underline{b} \wedge \langle \nabla_{10} \rangle \underline{c} \\
\rho_a := \underline{a} \to \ & \neg \langle \nabla_{\bullet 0} \rangle \top \\
\rho_b := \underline{b} \to \ & \neg \langle \nabla_{\bullet 0} \rangle \top \\
\rho_c := \underline{c} \to \ & \neg \langle \nabla_{\bullet 0} \rangle \top
\end{aligned}
$$

Now we are looking at the following logic **Mv**, where $C := \{\underline{S}, \underline{T}, \underline{U}, \underline{X}, \underline{a}, \underline{b}, \underline{c}\}$, with

(73) $$\mathbf{Mv} := \mathbf{OL} \oplus \mathrm{Inc}(C) \oplus \mathrm{Suc}(C) \oplus \{\rho_X : X \in C\}$$

Since the added postulates are constant, it is a matter of direct verification that the structures for this logic are the PMDSs in which the underlying tree (using the nonderived links) satisfies the context free grammar given in (69). Any constituent may move, and moves to any c-commanding position.

It is interesting to spell out which linear order we use for the surface constituents. To this end, let $x \prec_s y$ if y is the highest member of $P(x)$; we also call the link $x; y$ a **surface link**. It is not hard to show that $\prec_s^+$ defines a tree order on the worlds. Moreover, let $x \prec_{s0} y$ if $x \prec_s y$ and $x \prec_0 y$; similarly, $x \prec_{s1} y$ iff $x \prec_s y$ and $x \prec_1 y$. We say that for two leaves x and y that x **surface-precedes** y, in x symbols $x \propto y$.

(74) $$x \propto y :\Leftrightarrow (\exists u)(\exists v)(\exists w)(x \prec_s^+ u \prec_{s0} v \succ_{s1} w \succ_s^+ y)$$

This order is not modally definable. However, this does not defeat the usefulness of the present approach. There are two fixes; one is to introduce a surface relation. Like we did for the root links, we have to introduce the relations $\prec_{s0}$ and $\prec_{s1}$ explicitly. The proofs so far go through without a change. Decidability is again guaranteed.

10. Adjunction

The next generalization we are going to make concerns **adjunction**. Recall from [6] that it is not enough to leave adjunction implicit. We must add an explicit statement which nodes are maximal. An adjunction structure is therefore obtained by adding a subset $\mathcal{M}$ of M. (Intuitively, this set represents tha maximal nodes of a category.)

(75) $$x^\mu := \text{the least } y \in \mathcal{M} \text{ such that } y \succ^* x$$

The **category** of x is defined as follows.

(76) $$C(x) := \{y : y^\mu = x^\mu\}$$

A **category** is a subset of M of the form $C(x)$. y is a **segment** of $C(x)$ if $y \in C(x)$. Two categories are either equal or disjoint; hence the categories form a partition of M. Categories must also be linear. To ensure this we must have the following: if y and y' are distinct daughters of x then $y \in \mathcal{M}$ or $y' \in \mathcal{M}$. For otherwise $y^\mu = y;^\mu$ and so $C(y) = C(y') = C(x)$. Finally, in adjunction structures c-command is revised as follows. Say that y **includes** x if all segments dominate x. x **c-commands** y iff the least z including x dominates y. Now we require that chains are linearly ordered through ac-command. This is reflected in the following conditions.

The set $M(x)$ gets replaced by the set $P(x)$, which is formed as follows. Suppose that $x <^+ u$, where u is minimal in its category (so that the category is the least one that includes x), and there is a path Π from x to u going only through nonminimal nodes, and following derived links. Then $u \in P(x)$. As before, $P(x)$ reports about the movement history of x. But now that c-command is no longer defined using the one-node-up version (idc-command in the sense of [1]), we need to define a different set of nodes that need to be compared. This is why we chose $P(x)$ to be the mothers of the ultimate landing site of a complex formed through successive adjunction. The link that adjunction creates is always counted as derived. We shall see below an example of where this arises naturally.

In fact, adjunction has been taken to be more restrictive. Typically, when an element adjoins, it must adjoin to the maximal segment of the existing category. Whence, we shall simplify the task as follows. Call x **infimal** if there is no $y < x$ which is nonmaximal (that is to say, x is the least member in its category).

(77) $P(x) := \{y : y > x$ and x infimal or

there is a noninfimal z and $y >_{\bullet 0} z >_{\bullet 1} x\}$

Definition 32. *A **pseudo-ordered adjunction MDS (PAMDS)** is a structure* $\langle M, \mathcal{M}, >_{00}, >_{01}, >_{10}, >_{11} \rangle$, *where the following holds:*

(1) $\mathcal{M} \subseteq M$.
(2) *If $y >_0 x$ and $y >_0 x'$ then $x = x'$.*
(3) *If $y >_1 x$ and $y >_1 x'$ then $x = x'$.*
(4) *If $y >_1 x$ then there is a z such that $y >_0 x$.*
(5) *There is exactly one x such that for no y, $y > x$ (this element is called the **root**).*
(6) *If $x >_{\bullet 1} y$ then $y \in \mathcal{M}$. (Adjoining elements are maximal segments.)*
(7) *If $x < y$ and $x' < y$ and $x, x' \notin \mathcal{M}$ then $x = x'$. (Only one daughter is a nonmaximal segment. Categories are linear.)*
(8) *The set $P(x)$ is linearly ordered by $<^+_{\bullet 0}$ and if y is minimal with respect to $<^+$ then $y >_{\bullet 0} x$.*

As before, we need to define the logic of these structures and then show that the defined logic has the finite model property, which shall establish its decidability. First, let us notice a few facts about these structures. The idea of adjunction plays the biggest role in head adjunction, because here the new notion of c-command takes effect. A head adjoins to a higher head, but in the new position it does not idc-command its trace, it just c-commands it. The postulates are as follows. We shall introduce a constant μ whose

interpretation is the set $\mathfrak{M}$. First, let us agree on the following notation.

$$(78) \qquad A := \langle \nabla \rangle \neg \mu$$

$$(79) \qquad \blacktriangledown := (\neg A?; \nabla_{\bullet 1}) \cup (\nabla_{\bullet 0}; A?; \nabla_{\bullet 1})$$

A is true on the node to which one has adjoined; $\langle y, x \rangle \in R(\blacktriangledown)$ iff $y \in P(x)$.

Definition 33. *Let*

$$(80) \qquad \begin{aligned} \mathbf{PAM} = \quad & \mathbf{DPDL}_4.\mathbf{f} \\ & \oplus \langle \nabla_{00} \rangle \neg \mu \to [\nabla_{10}]\mu \\ & \oplus \langle \nabla_{10} \rangle \neg \mu \to [\nabla_{00}]\mu \\ & \oplus [\nabla_{\bullet 1}]\mu \\ & \oplus \langle \blacktriangledown \rangle p \to \langle \nabla^+; (\top? \cup \blacktriangledown) \rangle p \end{aligned}$$

Lemma 34. *Every finite PAMDS satisfies the postulates of* **PAM**.

Proof. (a) The postulates of $\mathbf{DPDL}_4.\mathbf{f}$ are satisfied, by similar arguments. (b) Suppose $\mathfrak{M}$ is a PAMDS, and let $x \in M$, $x \vDash \langle \nabla_{00} \rangle \neg \mu$. Then there is a $y \prec_{00} x$ which is not in $\mathfrak{M}$. By (7), if $z \prec_{10} x$, z must be maximal, whence $z \vDash \mu$. z was arbitrary (in fact, if it exists, it is unique). Therefore, $x \vDash [\nabla_{10}]\mu$. Similarly for the second axioms. (c) $x \vDash [\nabla_{\bullet 1}]\mu$. For let $y \prec_{\bullet 1} x$. Then by (6), $y \in \mathfrak{M}$, whence $y \vDash \mu$. (d) Suppose $x \vDash \langle \blacktriangledown \rangle p$. This means that there is a y such that $x \in P(y)$. By (8), if $x >_{\bullet 1} y$, then x is not minimal in $P(y)$. Hence, there is a z such that $x >^+ z$ and $z \in P(x)$. This means either that z is minimal in $P(x)$, in which case $z \vDash \langle \top? \rangle p$, or else that z is not minimal, but then $z \vDash \langle \blacktriangledown \rangle p$. By assumption on $P(y)$, that $x >_{\bullet 0}^+ z$. Hence $z \vDash \langle (\top?) \cup \blacktriangledown \rangle p$ and so $x \vDash \langle \nabla_{\bullet 0}^+; \blacktriangledown \rangle p$. $\qquad \square$

Now we turn to the converse. Put

$$(81) \qquad \begin{aligned} Z(\varphi) := \quad & \{[u](\langle \nabla_{\bullet 1} \rangle \delta \to \langle \nabla_{\bullet 0}^+; (\top? \cup \blacktriangledown) \rangle \delta : \delta \in \mathrm{At}(\varphi)\} \\ & \cup \{[u](\langle \nabla_{00} \rangle \neg \mu \to [\nabla_{10}]\mu), [u](\langle \nabla_{10} \rangle \neg \mu \to [\nabla_{00}]\mu)\} \\ & \cup \{[u][\nabla_{\bullet 1}]\mu\} \end{aligned}$$

Lemma 35. φ *is consistent with* **PAM** *iff* $\varphi; Z(\varphi)$ *is consistent with* $\mathbf{DPDL}_4.\mathbf{f}$.

Proof. ($\Rightarrow$.) Clear. ($\Leftarrow$). Let $Z(\varphi); \varphi$ be consistent with $\mathbf{DPDL}_4.\mathbf{f}$. Then it has a finite generated model based on $\mathfrak{M} = \langle M, \mathfrak{M}, >_{00}, >_{01}, >_{10}, >_{11} \rangle$, the valuation β and w_0 such that

$$(82) \qquad \langle \mathfrak{M}, \beta, w_0 \rangle \vDash Z(\varphi); \varphi$$

(a) By choice of $Z(\varphi)$, $w_0 \vDash [u](\langle \nabla_{00} \rangle \neg \mu \to [\nabla_{10}]\mu)$. Take $z \in M$. Then, by definition of u, $z \vDash \langle \nabla_{00} \rangle \neg \mu \to [\nabla_{10}]\mu$. Suppose now that y is nonmaximal and $z >_{00} y$. Then $z \vDash \langle \nabla_{00} \rangle \neg \mu$. Whence $z \vDash [\nabla_{10}]\mu$. So, if $z >_{10} u$, then u is maximal. Similarly it is seen that if $z >_{10} y$ and y is nonmaximal, and $z >_{00} u$

then u is maximal. This establishes linearity, (7). (b) $z \vDash [\nabla_{\bullet 1}]\mu$. Hence if $y \succ_{\bullet 1} z$, y is maximal. Thus, (6) is satisfied. (c) Now we deal with the most problematic formula, the last axiom. We replay the proof of Theorem 19. The only change is that we define the relation L differently. For as before, S is the set of standard points, and E the set of immediate, derived daughters of standard points. We shall have to verify that L is cycle free, and that the structure obtained by identifying all points L-related to each other is a **PAM**-structure and the resulting model satisfies φ. Basically, the proof of the latter is as in Theorem 19. So let us see why the structure is a **PAM**-structure. For, this we need to establish that $P(x)$ is linearly ordered by $\succ^+$. $\square$

There are typically other requirements that are placed on adjunction structures. The first is that head adjunction takes place to the right only. Thus, if y is a zero level projection and $x \succ_{\bullet 1} y$, then y must be to the right, so $\bullet = 1$. This is captured as follows. There is a constant H which is true of exactly the zero-level projections. So we say

$$(83) \qquad\qquad H \to [\triangle_{10}]\bot$$

Next, at least in the standard theory, the head-head complex cannot be taken apart by movement again. (The phenomenon is known as **excorporation**.) Structurally, it means that an adjoined element cannot have two mothers. Thus, if $x, x' \succ_{\bullet 1} y$ and y is zero level, then $x = x'$. This must be added to the list of requirements if needed. This is a universal first-order formula, so only have to appeal to Theorem 24 to see that it can be axiomatised modally.

11. Second Example: Swiss German

It is worth seeing a concrete example of how the present ideas can be made to work. We choose Swiss German to exemplify the interplay between movement and adjunction. Our analysis will be the cyclic head adjunction analysis put forward in the 80s for Dutch and German.

We shall assume that lexical items have internal structure, which is also binary branching. For simplicity, we denote the relations below the lexical level by another symbol ($\oslash$ and $\obar$). (For all those worried about decidability: they are dispensable. We could introduce a constant L, which is true of all sublexical nodes. Then we put $\obar = \nabla; L?$ and $\oslash = L?; \triangle$.) The lexicon contains complex nodes whose leftmost part is a string. The other nodes are auxiliary and carry phonetically empty material, here one of the following: α, δ and σ. They are mutually exclusive (just like the other labels). α is a feature for accusative case, δ for dative case and σ for the selection of an infinitival complement. The following are the lexical trees that we shall use; Figure 1 shows two of them in tree format. (By the way, we abandon now

FIGURE 1. Some Lexical Trees
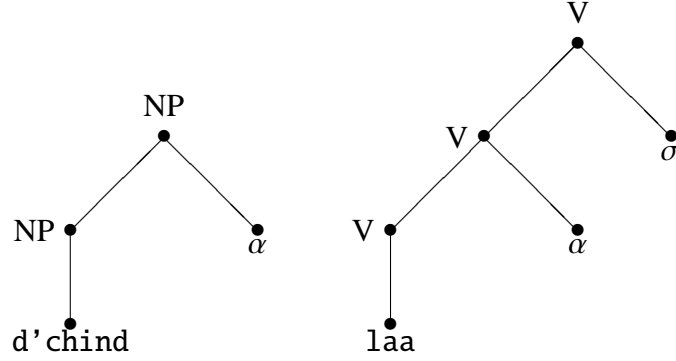
the underscore notation for constants.)

(84) $[\text{d'chind } \alpha]_{\text{NP}}$

(85) $[\text{em chind } \delta]_{\text{NP}}$

(86) $[\text{aastriche } \alpha]_{\text{V}}$

(87) $[[\text{hälfe } \delta]_{\text{V}} \; \sigma]_{\text{V}}$

(88) $[[\text{laa } \alpha]_{\text{V}} \; \sigma]_{\text{V}}$

The grammar for the deep structure is this:

(89) $\text{VP} \to \text{V}^1 \; \text{VP}$ $\text{VP} \to \text{V NP}$

(90) $\text{V}^1 \to \text{V NP}$ $\text{VP} \to \text{NP VP}$

We shall assume that the surface structure is created through successive cyclic head adjunction. That is to say, any head is allowed to move and adjoin to the next higher head; adjunction is always to the right, but it need not be cyclic. Suppose we have four heads $V_1 \; V_2 \; V_3 \; V_4$. Then we can first adjoin V_3 to V_4, giving $[V_4 \; V_3]$, then V_1 to V_2, giving $[V_2 \; V_1]$, and then finally $[V_2 \; V_1]$ to $[V_4 \; V_3]$ to give $[[V_4 \; V_3] \; [V_2 \; V_1]]$. This can be excluded, see below.

The rules, together with the lexicon can be translated into constant axioms as follows. (Recall from (78) the definition $A := \langle \nabla \rangle \neg \mu$. Furthermore, $\circledcirc_0^2 := \circledcirc_2; \circledcirc_2$.)

(91) $\rho_{\text{VP}} := \text{VP} \to (\langle \nabla_{00} \rangle \text{V}^1 \wedge \langle \nabla_{10} \rangle \text{VP})$

$\vee \; (\langle \nabla_{00} \rangle \text{V} \wedge \langle \nabla_{10} \rangle \text{NP})$

$\vee \; (\langle \nabla_{00} \rangle \text{V} \wedge \langle \nabla_{10} \rangle \text{VP})$

(92) $\rho_{\text{V}'} := \text{V}^1 \to \langle \nabla_{00} \rangle \text{V} \wedge \langle \nabla_{10} \rangle \text{NP}$

$$(93) \qquad \rho_{\mathrm{NP}} := \mathrm{NP} \to (\langle \oslash_0^2 \rangle (\texttt{d'chind} \vee \texttt{Hans} \vee \cdots) \wedge \langle \oslash_1 \rangle \alpha))$$

$$\vee (\langle \oslash_0^2 \rangle (\texttt{em chind} \vee \texttt{em Hans} \vee \cdots) \wedge \langle \oslash_1 \rangle \delta))$$

$$(94) \qquad \rho_{\mathrm{V}}^{N} := (\mathrm{V} \wedge \neg A) \to (\langle \oslash_0^2 \rangle (\texttt{aastriche} \vee \cdots) \wedge \langle \oslash_1 \rangle \alpha)$$

$$\vee (\langle \oslash_0^2 \rangle (\langle \oslash_0 \rangle \texttt{hälfe} \vee \cdots) \wedge \langle \oslash_1 \rangle \delta) \wedge \langle \oslash_1 \rangle \sigma)$$

$$\vee (\langle \oslash_0 \rangle (\langle \oslash_0 \rangle \texttt{laa} \vee \cdots) \wedge \langle \oslash_1 \rangle \alpha) \wedge \langle \oslash_1 \rangle \sigma)$$

$$(95) \qquad \rho_{\mathrm{V}}^{A} := (\mathrm{V} \wedge A) \to \langle \nabla_{00} \rangle \mathrm{V} \wedge \langle \nabla_{11} \rangle (\mathrm{V} \wedge \mu)$$

$$(96) \qquad \rho_\alpha := \alpha \to [\oslash] \bot$$

$$(97) \qquad \rho_\delta := \delta \to [\oslash] \bot$$

$$(98) \qquad \rho_\sigma := \sigma \to [\oslash] \bot$$

Notice that it is possible to enforce cyclic head adjunction by issuing the following formula in place of ρ_{V}^{A}:

$$(99) \qquad \gamma_{\mathrm{V}}^{A} := (\mathrm{V} \wedge A) \to \langle \nabla_{00} \rangle (\mathrm{V} \wedge \neg A) \wedge \langle \nabla_{11} \rangle (\mathrm{V} \wedge \mu)$$

This says that the left hand daughter must be infimal, hence that daughter is lexical. The right hand daughter may however be complex.

Case government is implemented as follows.

$$(100) \qquad \kappa_\alpha := \mathrm{V} \wedge \langle \oslash \cup \oslash^2 \rangle \alpha \to \langle \triangle; \oslash \rangle \alpha$$

$$(101) \qquad \kappa_\delta := \mathrm{V} \wedge \langle \oslash \cup \oslash^2 \rangle \delta \to \langle \triangle; \oslash \rangle \delta$$

Selectional restriction concerning the infinitive is the formula

$$(102) \qquad \sigma := \mathrm{V} \wedge \langle \oslash \rangle \sigma \to \langle (\neg \mathrm{VP}?); \triangle \rangle^*; \nabla \rangle \mathrm{VP}$$

Notice that these formulae are all constant. They describe the restrictions that apply at D-structure.

The only derivational steps are head adjunction, as shown above. The crucial fact here is that head adjunction is local; so we restrict the condition (7) in Definition 32 by saying that the distance between two members of $P(x)$ must be small. The head movement constraint is embodied in the following formula

$$(103) \qquad \mu_h := \langle \blacktriangledown \rangle p \to \langle \nabla_{\bullet 0}^2; (\top? \cup \blacktriangledown) \rangle p$$

This formula is somewhat crude, saying that movement is only two steps up. It suffices for our purposes, thanks to the particular grammar chosen. It would be no problem to formulate a more sophisticated version which says that a head may only move to the next head.

Definition 36. *Call Swiss the logic*

$$(104) \qquad \mathbf{OL} \oplus \mathrm{Inc}(C) \oplus \mathrm{Suf}(C) \oplus \{\rho_{\mathrm{VP}}, \rho_{\mathrm{V}'}, \rho_{\mathrm{NP}}, \rho_{\mathrm{V}}^{N}, \rho_{\mathrm{V}}^{A}, \kappa_\alpha, \kappa_\gamma, \sigma, \mu_h\}$$

Swiss is decidable. This follows from our results. The language is trans-context free. To see this we must first define the surface order. This means that we have to spell out which of the links is a surface link. This is the standard link if the element is not a V, and it is not adjoined. Otherwise, it is a derived link.

(105) $\langle \prec_{s0} \rangle p \leftrightarrow ((\neg V \wedge \neg A) \rightarrow \langle \prec_{00} \rangle p)) \wedge ((V \vee A) \rightarrow \langle \prec_{01} \rangle p)$

(106) $\langle \prec_{s1} \rangle p \leftrightarrow ((\neg V \wedge \neg A) \rightarrow \langle \prec_{10} \rangle p)) \wedge ((V \vee A) \rightarrow \langle \prec_{11} \rangle p)$

Notice that although we have introduced new symbols, $\prec_{s0}$ and $\prec_{s1}$, they are eliminable, so they are in effect just shorthands.

After that we define the left-to-right order on the surface and finally the relation $\propto_s$, which is like the surface $\propto$, but it skips intervening empty heads.

(107)
$$\propto := \prec_s^*; \prec_{s0}; \succ_{s1}; \succ_s^*$$
$$c := \sigma \vee \alpha \vee \delta$$
$$\propto_s := \propto; (c?; \propto)^*; \neg c$$

Now, x is immediately to the left of y in surface order if $x \, R(\propto_s) \, y$. $x \, R(\Lambda_s) \, y$ if y is the next phonetically nonempty element to the right of x. So, the question whether the following sequence is derivable

(108)   `de chind em Hans es huus hälfe aastriche`

now becomes the question whether the following formula has a model:

(109) $[\propto_s^{\smile}]\bot \wedge \langle \propto_s \rangle (\underline{\texttt{de chind}} \wedge \langle \propto_s \rangle (\underline{\texttt{em Hans}} \wedge \langle \propto_s \rangle (\underline{\texttt{es huus}}$
            $\wedge \langle \propto_s \rangle (\underline{\texttt{hälfe}} \wedge \langle \propto_s \rangle (\underline{\texttt{aastriche}} \wedge \langle \propto_s \rangle [\Lambda]\bot)))))$

12. Conclusion

Let us briefly review what has been achieved and what remains to be done. We have established a way to reduce a grammar to a *logic L*, the lexicon to a *constant formula λ*. As a result, parsing becomes a *satisfiability problem* in a given logic (here $L \oplus \lambda$). (See [5, 8] for an extensive discussion.) Provided that the logic L is decidable, also the logic $L \oplus \lambda$ is decidable and the following questions become decidable:

- Given a string $\vec{x}$ and a particular lexicon λ, is $\vec{x}$ derivable in L plus λ?
- Does L plus λ imply a PDL-expressible principle α?
- Is a given regular language included in the language derived by L plus λ?

Since principles are axioms, our results establish decidability of these questions only on condition that L falls within the range of logics investigated here (or expansions by constant formulae). In particular, this means that

movement is assumed to satisfy *Freeze.* (This has consequences only for the formulation of nearness conditions.)

It should be said that there are questions that are known to be undecidable and so there is no hope of ever finding an algorithm that decides them once and for all. One problem is the question whether a given grammar generates less sentences than another one. This is undecidable already for context free grammars.

The reader might wonder what happened to surface structure and LF. These two pose no problems, as far as I can see. All that needs to be done is to split the relations $\prec_i$ into four different ones (which are not mutually exclusive). In this way, practically the full theory can be axiomatised within PDL. It is to be noted, however, that while the lexicon consists of constant formulae, the theory (consisting of general structural axioms) is phrased with formulae containing variables.

The results obtained in this paper support the claim that properties of generative grammars developed within GB or the Minimalist Program are in fact decidable as long as they can be expressed in PDL. In Part II of this sequence we shall show that this holds true also for the logic of narrow multidominance structures. These are structures where a given trigger licenses only one movement step. Decidability will be shown for theories that admit narrow structures with Freeze-style movement and command relations to measure distance. This will be exploited further in Part III, where we study Minimalism in depth.

References

[1] Chris Barker and Geoffrey Pullum. A theory of command relations. *Linguistics and Philosophy*, 13:1–34, 1990.

[2] Noam Chomsky. *Lecture Notes on Government and Binding*. Foris, Dordrecht, 1981.

[3] Noam Chomsky. *Barriers*. MIT Press, Cambrigde (Mass.), 1986.

[4] Marcus Kracht. Mathematical aspects of command relations. In *Proceedings of the EACL 93*, pages 241 – 250, 1993.

[5] Marcus Kracht. Is there a genuine modal perspective on feature structures? *Linguistics and Philosophy*, 18:401 – 458, 1995.

[6] Marcus Kracht. Adjunction Structures and Syntactic Domains. In Uwe Mönnich and Hans-Peter Kolb, editors, *The Mathematics of Sentence Structure. Trees and Their Logics*, number 44 in Studies in Generative Grammar, pages 259 – 299. Mouton–de Gruyter, Berlin, 1998.

[7] Marcus Kracht. *Tools and Techniques in Modal Logic*. Number 142 in Studies in Logic. Elsevier, Amsterdam, 1999.

[8] Marcus Kracht. Logic and syntax – a personal perspective. In Maarten de Rijke, Krister Segerberg, Heinrich Wansing, and Michael Zakharyaschev, editors, *Advances in Modal Logic '98*, pages 337 – 366. CSLI, 2001.

[9] Marcus Kracht. Syntax in Chains. *Linguistics and Philosophy*, 24:467 – 529, 2001.

[10] Marcus Kracht. Constraints on derivations. *Grammars*, 6:89 – 113, 2003.

[11] Marcus Kracht. *The Mathematics of Language*. Mouton de Gruyter, Berlin, 2003.

[12] Maria R. Manzini. *Locality – A Theory and Some of Its Empirical Consequences*. Number 19 in Linguistic Inquiry Monographs. MIT Press, 1992.

[13] Luigi Rizzi. *Relativized Minimality*. MIT Press, Boston (Mass.), 1990.

[14] James Rogers. *Studies in the Logic of Trees with Applications to Grammar Formalisms*. PhD thesis, University of Delaware, Department of Computer & Information Sciences, 1994.

[15] Edward P. Stabler. *The Logical Approach to Syntax. Foundation, Specification and Implementation of Theories of Government and Binding*. ACL-MIT Press Series in Natural Language Processing. MIT Press, Cambridge (Mass.), 1992.

[16] Moshe Vardi and P. Wolper. Automata theoretic techniques for modal logics of programs. *Journal of Computer and Systems Sciences*, 32:183 – 221, 1986.

DEPARTMENT OF LINGUISTICS, UCLA, 3125 CAMPBELL HALL, PO BOX 951543, LOS ANGELES, CA 90095-1543, `kracht@humnet.ucla.edu`