Bridging the Gap Between Underspecification Formalisms: Minimal Recursion Semantics as Dominance Constraints

Abstract

Minimal Recursion Semantics (MRS) is the standard formalism used in large-scale HPSG grammars to model underspecified semantics. We present the first provably efficient algorithm to enumerate the readings of MRS structures. It is obtained by translating MRS into normal dominance constraints for which efficient algorithms exist.

1 Introduction

In the past few years there has been considerable activity in the development of formalisms for *underspecified semantics* (Alshawi and Crouch, 1992; Reyle, 1993; Bos, 1996; Copestake et al., 1999; Egg et al., 2001). The common idea is to delay the enumeration of all readings for as long as possible. Instead, they work with a compact *underspecified representation* for as long as possible, only enumerating readings from this representation by need.

Minimal Recursion Semantics (MRS) (Copestake et al., 1999) is the standard formalism for semantic underspecification used in large-scale HPSG grammars (Pollard and Sag, 1994; Copestake and Flickinger, 2000). Despite of this clear relevance, the most obvious questions about MRS are still open:

1. Is it possible to enumerate the readings of an MRS-structure efficiently? No algorithm is published so far. Existing implementations seem to be practical, even though the existence of readings of an MRS structure is already NP-complete (Theorem 10.1 of Althaus et al. (2003)).
2. What is the precise relationship to other underspecification formalism? Are all of them the same, or else, what are the differences?

We answer both questions in this paper. We distinguish the sublanguages of MRS-nets and normal dominance nets, and show that they can be intertranslated. We can therefore apply existing constraint solver for normal dominance constraints to enumerate the readings of MRS-nets in low polynomial time.

MRS-nets or equivalently normal dominance nets are sufficiently powerful for modeling scope underspecification: they generalize on chain-connected normal dominance constraints, which are already sufficient as argued by Koller et al. (2003). Nets can also be defined for Hole semantics (Bos, 1996); the results of the present paper show, that this yields another equivalent underspecification formalism.

Furthermore, this paper introduces a new proof technique: Arguing about MRS is reduced to reasoning about weakly normal dominance constraints (Bodirsky et al., 2003). Dealing with hole semantics is easier since it requires properties of normal dominance constraints only (Koller et al., 2003).

2 Minimal Recursion Semantics

We first define a simple version of Minimal Recursion Semantics (Copestake et al., 1999) that captures the essence of MRS, and then discuss the differences to the original version.

2.1 Definition

MRS is a description language for formulas of some first order object languages with generalized quantifiers. Underspecified representations in MRS consist of *elementary predications* and *handle constraints*. Roughly, elementary predications are object language formulas with "holes" into which other formulas can be plugged; handle constraints restrict the way these formulas can be plugged into each other.

Underspecified representations are built from the following vocabulary:

1. *Variables.* An infinite set of variables ranged over by h. Variables are also called *handles*.
2. *Constants.* An infinite set of constants ranged

over by x, y, z. Constants are the *individual variables* of the object language.

3. *Function symbols.* A set of function symbols ranged over by P and sets of quantifier symbols Q_x for all individual variables x, where Q stands for a quantifier. We call Q_x the *variable binder* of x.
4. The symbol $\leq$ for the outscopes relation.

Formulas of MRS have three kinds of literals:

1. $h : P(x_1, \ldots, x_n, h_1, \ldots, h_m)$
2. $h : Q_x(h_1, h_2)$
3. $h_1 \leq h_2$

The first two forms are called *elementary predications* (EPs) and the third form *handle constraints*.

The position on the left of a colon ':' in an EP is the *label position* and the positions on the right the *argument positions*. Let M be a set of literals. The *labels* of M are those handles occurring in label but not in argument positions in M. The *argument handles* of M are those that occur in argument but not in label position.

Definition 1 (MRS). An MRS is finite set M of MRS-literals such that:

M1 Every handle occurs at most once in label and at most once in argument position in M.

M2 Handle constraint $h_1 \leq h_2$ in M always relate argument handles h_1 to labels h_2 of M.

M3 For every constant (individual variable) x in argument position in M there is a unique literal of the form $h : Q_x(h_1, h_2)$ in M.

We call an MRS *compact* if it satisfies:

M4 Every handle of M occurs exactly once in an elementary predication of M.

We say that a handle h *immediately outscopes* a handle h' in an MRS M iff there is an EP E in M such that h occurs in label and h' in argument position of E. The outscopes relation is the reflexive, transitive closure of the immediate outscopes relation.

We often represent MRSs by graphs. For instance, the graph in Figure 1 represents the following MRS for the sentence "Every student reads a book."

$$\{h_1 : \text{every}_x(h_2, h_4), h_3 : \text{student}(x), h_5 : \text{some}_y(h_6, h_8), h_7 : \text{book}(y), h_9 : \text{read}(x, y), h_2 \leq h_3, h_6 \leq h_7\}$$

Elementary predications are drawn with solid edges and handle constraints are represented by dotted

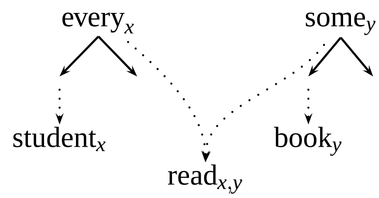


Figure 1: Graph for "Every student reads a book."

lines. Note that we make the relation between bound variables and their binders explicit by dotted lines; however, redundant "binding-edges" are omitted.

A solution for an underspecified MRS is called a *configuration*, or *scope-resolved MRS*.

Definition 2 (Configuration). An MRS M is a *configuration* if it satisfies the following conditions.

C1 The graph of M is a tree, i.e. no handle properly outscopes itself, no handle occurs in different argument positions and all handles are pairwise connected by elementary predications.

C2 If two EPs $h : P(\ldots, x, \ldots)$ and $h_0 : Q_x(h_1, h_2)$ belong to M then h_0 outscopes h in M.

We call M a *configuration for* another MRS M' if there exists some substitution $\sigma : \mathsf{arg}(\mathsf{M}') \mapsto \mathsf{lab}(\mathsf{M}')$ which states how to identify argument handles of M' with labels of M', so that:

C3 $M = \{\sigma(E) \mid E \text{ is EP in } M'\}$, and

C4 $\sigma(h_1)$ outscopes h_2 in M for all $h_1 \leq h_2 \in M'$.

The value $\sigma(E)$ is obtained by substituting all argument handles in E, leaving all others unchanged.

2.2 Remarks

Our version of MRS differs in some respects from the original version in Copestake et al. (1999).

First, we assume that different EPs must be labeled with different handles, and that labels cannot be identified. In standard MRS, however, conjunctions are encoded by labeling different EPs with the same handle. If we make the plausible assumption that different labels cannot be identified dynamically, i.e. if different labels in an MRS M remain different in every configuration of M, then these EP-conjunctions can be resolved by introducing an additional EP which makes the conjunction explicit.

Second, we use a slightly weaker form of handle constraints. Standard MRS uses "qeq" constraints instead of our outscopes constraints. A handle h

is qeq h' (written $h =_q h'$) in an MRS M if either $h = h'$ or $h : Q_x(h_1, h_2)$ occurs in M and h_2 is qeq h'. Thus, $h =_q h'$ implies $h \leq h'$, but not the other way round. Although qeq-constraints are stronger than outscopes constraints, we don't konw of any example, where this additional strength is needed.

Apart from these differences, we also depart from standard MRS in some minor details, e.g. we use sets instead of multi-sets and omit the usual top-handle, which is useful only during semantics construction.

3 Dominance Constraints

Dominance constraints are a general framework for the partial description of trees, and thus of the syntax trees of logic formulas. Dominance constraints are the core language underlying CLLS (Egg et al., 2001), which adds parallelism and binding constraints.

3.1 Syntax and Semantics

We assume a finite or infinite signature Σ of function symbols with fixed arities and an infinite set Var of variables ranged over by X, Y, Z. We write f, g for function symbols and $\mathsf{ar}(f)$ for the arity of f.

A dominance constraint φ is a conjunction of dominance, inequality, and labeling literals of the following form, where $\mathsf{ar}(f) = n$:

$$\varphi ::= X \triangleleft^* Y \mid X \neq Y \mid X : f(X_1, \ldots, X_n) \mid \varphi \wedge \varphi'$$

Dominance constraints are interpreted over finite constructor trees over signature Σ, i.e. ground terms constructed from the function symbols in Σ. We identify ground terms with trees that are rooted, ranked, edge-ordered and labeled. A solution for a dominance constraint consists of a tree τ and a variable assignment α that maps variables to nodes of τ such that all constraints are satisfied: a labeling literal $X : f(X_1, \ldots, X_n)$ is satisfied iff the node $\alpha(X)$ is labeled with f and has daughters $\alpha(X_1), \ldots, \alpha(X_n)$ in this order; a dominance literal $X \triangleleft^* Y$ is satisfied iff there is a path from $\alpha(X)$ to $\alpha(Y)$ in τ; and an inequality literal $X \neq Y$ is satisfied iff $\alpha(X)$ and $\alpha(Y)$ are distinct nodes.

Note that a solution may contain additional material. For instance, the tree $f(a, b)$ satisfies the constraint $X \triangleleft^* Y \wedge X \triangleleft^* Z \wedge Y : a \wedge Z : b$.

3.2 Normality and Weak Normality

The satisfiability problem of arbitrary dominance constraints is NP-complete (Koller et al., 2001) in general. However, Althaus et al. (2003) identify a natural fragment of dominance constraints, *normal dominance constraints*, which have a polynomial time satisfiability problem. Bodirsky et al. (2003) generalize this notion to *weakly normal dominance constraints*.

Definition 3. A dominance constraint φ is *normal* if it satisfies the following conditions.

- N1 (a) each variable of φ occurs at most once in the labeling literals of φ. A variable Y_i is a *hole* of φ whenever it occurs as an argument in a labeling $X : f(Y_1, \ldots, Y_n)$ of φ; else it is a *root* of φ.
  (b) each variable of φ occurs at least once in the labeling literals of φ.
- N2 if X and Y are distinct roots in φ, $X \neq Y$ occurs in φ.
- N3 (a) if $X \triangleleft^* Y$ occurs in φ, Y is a root in φ.
  (b) if $X \triangleleft^* Y$ occurs in φ, X is a hole in φ.

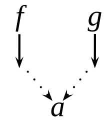
A dominance constraint is *weakly normal* if it satisfies all above properties except for N1(b) and N3(b).

The idea behind (weak) normality is that the constraint graph (see below) of a dominance constraint consists of solid fragments which are connected by dominance constraints; these fragments may not overlap in a solution.

Note that this definition applies only to compact constraints. As for MRS, all dominance constraints can be compactified, provided that dominance links relate either roots or holes with roots.

Dominance Graphs. We often represent dominance constraints as *graphs*. A *dominance graph* is the digraph $(V, \triangleleft^* \uplus \triangleleft)$. The graph of a weakly normal constraint φ is defined as follows: The nodes of the graph of φ are the variables of φ. A labeling literal $X : f(X_1, \ldots, X_n)$ of φ contributes *tree edges* $(X, X_i) \in \triangleleft$ for $1 \leq i \leq n$ that we draw as $X \rightarrow X_i$; we freely omit the label f and the edge order in the graph. A dominance literal $X \triangleleft^* Y$ contributes a dominance edge $(X, Y) \in \triangleleft^*$ that we draw as $X \dashrightarrow Y$. Inequality literals in φ are also omitted in the graph.

For example, the constraint graph on the right represents the dominance constraint $X:f(X') \wedge Y:g(Y') \wedge X' \lhd^* Z \wedge Y' \lhd^* Z \wedge Z:a \wedge X \neq Y \wedge X \neq Z \wedge Y \neq Z$



A dominance graph is *weakly normal* or a *wnd-graph* if it does not contain any forbidden subgraph of the following forms:

Dominance graphs of a weakly normal dominance constraints are obviously weakly normal.

Solved Forms and Configurations. The main difference between MRS and dominance constraints lies in their notion of interpretation: solutions versus configurations.

Every satisfiable dominance constraint has infinitely many solutions. Algorithms for dominance constraints therefore do not enumerate solutions but *solved forms*. We say that a wnd-graph Φ is in solved form iff Φ is a forest. The *solved forms of* Φ are solved forms Φ' that are more specific than Φ, i.e. Φ and Φ' differ only in their dominance edges and the reachability relation of Φ extends the reachability of Φ'. A *minimal solved form of* Φ is a solved form of Φ that is minimal with respect to specificity.

The notion of configurations from MRS applies to dominance constraints as well. Here, a *configuration* is a dominance constraint whose graph is a tree without dominance edges. A configuration of a constraint φ is a configuration that solves φ in the obvious sense. Configurations correspond precisely to *simple solved forms*, which are tree-shaped solved forms where every hole has exactly one outgoing dominance edge.

Lemma 1. Simple solved forms and configurations correspond: Every simple solved form has exactly one configuration, and for every configuration there is exactly one solved form that it configures.

Unfortunately, Lemma 1 does not hold for nominal instead of simple solved forms: there are constraints that have minimal solved forms which do not have a configuration. For instance, consider the schema of the constraint on the left in Figure 2. The graph on the right schematically represents a minimal solved form for it. Obviously, this solved form does not have a configuration.

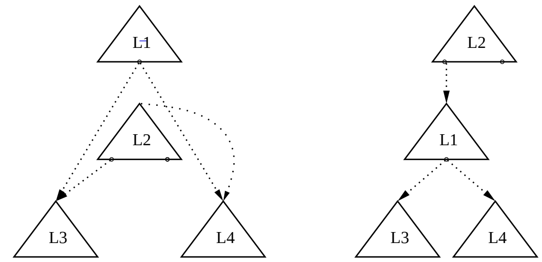


Figure 2: A dominance constraint (left) with a minimal solved form (right) that has no configuration

4 Representing MRSs

We next map MRSs to weakly normal dominance constraints so that configurations are preserved. Unfortunately, this translation is based on a non-standard semantics for dominance constraints, namely configurations. We address this problem in the following sections.

The translation of an MRS M to a dominance constraint φ_M is quite trivial. The variables of φ_M are the handles of M and its literal set is:

$$\begin{array}{ll} & \{h:P_{x_1,\ldots,x_n}(h_1,\ldots) \mid h:P(x_1,\ldots,x_n,h_1,\ldots) \in M\} \\ \cup & \{h:Q_x(h_1,h_2) \mid h:Q_x(h_1,h_2) \in M\} \\ \cup & \{h_1 \lhd^* h_2 \mid h_1 \leq h_2 \in M\} \\ \cup & \{h \lhd^* h_0 \mid h:Q_x(h_1,h_2), h_0:P(\ldots,x,\ldots) \in M\} \\ \cup & \{h \neq h' \mid h,h' \text{ in distinct label positions of } M\} \end{array}$$

Compact MRSs M are clearly translated into weakly normal dominance constraints. Labels of M become roots in φ_M while argument handles become holes. Weak root-to-root dominance literals are needed to encode MRS's variable binding condition C2. It could be formulated equivalently through lambda binding constraints of CLLS (but this is not necessary here in the absence of parallelism).

The compactness assumption does not seriously limit the results of this paper, but simplifies its presentation. Arbitrary MRSs could either be handled by "compressing" several predicate symbols into a more complex predicate symbol, or allowing for a slightly more general notion of weakly normal dominance constraints.

Proposition 1. The translation of a compact MRS M into a weakly normal dominance constraint φ_M preserves configurations.

This weak correctness property follows straightforwardly from the analogy in the definitions.

5 Constraint Solving

We recall an efficient algorithm from (Bodirsky et al., 2003) that enumerates all minimal solved forms of a wnd-graph or constraint. All results of this section are proved there.

The same algorithm can be used to enumerate configurations for large subclasses of wnd-constraints or MRSs, as we will see in Section 6. But equally importantly, this algorithm provides a powerful proof method to reason about solved forms and configurations on which all our results rely.

5.1 Weak Connectedness

Two nodes X and Y of a wnd-graph $\Phi = (V,E)$ are *weakly connected* if there is an undirected path from X to Y in (V,E). We call Φ weakly connected if all its nodes are weakly connected. A weakly connected component (wcc) of Φ is a maximal weakly connected subgraph of Φ. The wccs of $\Phi = (V,E)$ form proper partitions of V and E.

Proposition 2. The graph of a solved form of a weakly connected wnd-graph is a tree.

5.2 Freeness

The idea of the enumeration algorithm is based on the notion of *freeness*.

Definition 4. A node X of a wnd-graph Φ is called *free* in Φ if there exists a solved form of Φ which is a tree with root X.

A weakly connected wnd-graph without free nodes is unsolvable. Otherwise, it has a solved form whose digraph is a tree by Prop. 2, and the root of this tree is free for Φ.

Given a set of nodes $V' \subseteq V$, we write $\Phi|_{V'}$ for the restriction of Φ to nodes in V' and edges in $V' \times V'$. The following lemma characterizes freeness:

Lemma 2. A wnd-graph Φ with free node X satisfies:

- F1 node X has indegree zero in graph Φ, and
- F2 no distinct children Y and Y' of X in Φ that are linked to X by immediate dominance edges are weakly connected in the remainder $\Phi|_{V\setminus\{X\}}$.

5.3 Algorithm

The algorithm for enumerating the minimal solved forms of a wnd-graph (or equivalently constraint) is given in Figure 3. We illustrate this algorithm at the problematic wnd-graph Φ in Fig. 2. The dominance net Φ is weakly connected, so that we can call solve(Φ). This procedure guesses the topmost fragment in the solved form of Φ (which exists by Proposition 2 as the solved form is a tree).

The only candidates are L1 or L2 since L3 and L4 have incoming dominance edges violating property F1. Let us choose fragment L2 to be topmost. The graph which remains when removing L2 is still weakly connected. It has a single minimal solved form computed by a recursive call of the solver, where L1 dominates L3 and L4. The solved form of the restricted graph is then put below the left hole of L2, since it is connected to this hole. As a result, we obtain the solved form on the right of Fig. 3.

Normalizing Φ has an interesting consequence: norm(Φ) can not be satisfied while placing L2 topmost. Our algorithm detects this correctly: the normalization of fragment L2 is not free in norm(Φ) since it violates property F2.

Theorem 1. The function solved_form(Φ) computes all minimal solved forms of a weakly normal dominance graph Φ; it runs in quadratic time per solved from.

6 Correct Encoding

We next explain how to encode a large fragment of MRSs into wnd-constraints such that configurations correspond precisely to minimal solved forms. The result of the translation will indeed be normal.

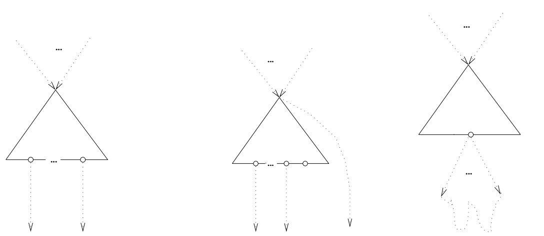
6.1 Problems and Examples

The naive representation of MRS formulas as weakly normal dominance constraints is only correct in a weak sense. The encoding fails in that some MRSs without configurations are mapped to solvable wnd-constraints. For instance, this holds for both MRSs in Fig 2.

We even cannot hope to translate arbitrary MRSs correctly into wnd-constraints: the configurability problem of MRSs is NP-complete, while satisfiability of wnd-constraints can be solved in polynomial time. We next introduce the fragment of MRS-nets

solved_form(Φ) $\equiv$
  Let $\Phi_1, \ldots, \Phi_k$ be the wccs of $\Phi = (V, E)$
  Let (V_i, E_i) be the result of solve(Φ_i)
  return $(V, \cup_{i=1}^{k} E_i)$
solve(Φ) $\equiv$
  precond: $\Phi = (V, \triangleleft \uplus \triangleleft^*)$ is weakly connected
  choose a node X satisfying (F1) and (F2) in Φ **else fail**
  Let $Y_1, \ldots, Y_n$ be all nodes s.t. $X \triangleleft Y_i$
  Let $\Phi_1, \ldots, \Phi_k$ be the weakly connected components of $\Phi|_{V-\{X, Y_1, \ldots, Y_n\}}$
  Let (W_j, E_j) be the result of solve(Φ_j), and $X_j \in W_j$ its root
  return $(V, \cup_{j=1}^{k} E_j \cup \triangleleft \cup \triangleleft_1^* \cup \triangleleft_2^*)$ where
    $\triangleleft_1^* = \{(Y_i, X_j) \mid \exists X' : (Y_i, X') \in \triangleleft^* \wedge X' \in W_j\}$,
    $\triangleleft_2^* = \{(X, X_j) \mid \neg\exists X' : (Y_i, X') \in \triangleleft^* \wedge X' \in W_j\}$

Figure 3: Enumerating the minimal solved-forms of a wnd-graph.



(a) strong (b) weak (c) island

Figure 4: Fragment Schemas of Nets

and equivalent wnd-nets, and show that they can be intertranslated in quadratic time.

6.2 Dominance and MRS-Nets

A hypernormal path (Althaus et al., 2003) in a wnd-graph is a sequence of adjacent edges that does not traverse two outgoing dominance edges of some hole X in sequence, i.e. a wnd-graph without situations $Y_1 \dashrightarrow X \dashleftarrow Y_2$.

A dominance net φ is weakly normal dominance constraint whose fragments all satisfy one of the three schemas in Fig. 4. MRS-nets can be defined in analogy. This means that all roots of φ are labeled in φ, and that for all fragments $X : f(X_1, \ldots, X_n)$ of φ where $n \geq 0$ one of the following three conditions holds:

strong. $n \geq 0$ and for all $Y \in \{X_1, \ldots, X_n\}$ there exists a unique Z such that $Y \triangleleft^* Z$ in φ, and not exists Z such that $X \triangleleft^* Z$ in φ.

weak. $n \geq 1$ and for all $Y \in \{X_1, \ldots, X_{n-1}, X\}$ there exists a unique Z such that $Y \triangleleft^* Z$ in φ, and not exists Z such that $X_n \triangleleft^* Z$ in φ.

island. $n = 1$ and all variables in $\{Y \mid X_1 \triangleleft^* Y\}$ are connected by a hypernormal path in the graph of the restricted constraint $\varphi_{|V-\{X_1\}}$, and not exists Z such that $X \triangleleft^* Z$ in φ.

As mentioned before, dominance nets generalize chain-connectedness (Koller et al., 2003). In short, the important property of a chain-connected constraint is that whenever a hole has two outgoing dominance edges (say, to X and Y), then either X dominates Y or Y dominates X in every solved form. This property is reflected directly by the island property above.

6.3 Normalizing Dominance Nets

Dominance nets are wnd-constraints. We next show how to translate configurability of dominance nets correctly to satisfiability of normal dominance constraints.

The trick is normalization of weak dominance edges. The normalization $\mathsf{norm}(\varphi)$ of a weakly normal dominance constraint φ is obtained by converting all root-to-root dominance literals $X \triangleleft^* Y$ as follows:

$$X \triangleleft^* Y \Rightarrow X_n \triangleleft^* Y$$

if X roots a fragment of φ that satisfies schema weak of net fragments. If φ is a dominance net then

$\mathsf{norm}(\varphi)$ is indeed a normal dominance net.

Theorem 2. The configurations of a weakly connected dominance net φ correspond bijectively to the minimal solved forms of its normalization $\mathsf{norm}(\varphi)$.

The proof captures the rest of this section. We will show in a first step (Prop.3 below) that the configurations are preserved when normalizing weakly connected nets. In the second step, we then show that minimal solved forms of normalized nets, and thus of $\mathsf{norm}(\varphi)$, can always be configured (Prop. 4 below).

Corollary 1. Configurability of weakly connected MRS-nets can be decided in polynomial time; configurations of weakly connected MRS-nets can be enumerated in quadratic time per configuration.

6.4 Correctness Proof

Most importantly, nets can be recursively decomposed into nets:

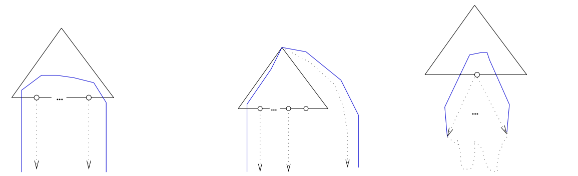
Lemma 3. If a dominance net φ has a configuration whose top-most fragment is $X : f(X_1, \ldots, X_n)$, then the restriction $\varphi_{|V-\{X,X_1,\ldots,X_n\}}$ is a dominance net.

Proof. First note that X is free in φ so that it does not have incoming edges (condition F1). This means, that restriction deletes only such dominance edges that depart from nodes in $\{X, X_1, \ldots, X_n\}$. Other fragments thus only loose ingoing dominance edges by normality condition N3. Such deletions preserve the validity of the schemas weak and strong.

The island schema is more problematic. We have to show that the hypernormal connections in this schema can never be cut. So suppose that $Y : f(Y_1)$ is an island fragment with outgoing dominance edges $Y_1 \lhd^* Z_1$ and $Y_1 \lhd^* Z_2$, so that Z_1 and Z_2 are connected by some hypernormal path traversing the deleted fragment $X : f(X_1, \ldots, X_n)$. We distinguish the three possible schemata of this fragment:

strong. Since X does not have incoming dominance edges, there is only a single non-trival kind of traversals, drawn in Fig. 5(a). But such travesals contradict the freeness of X according to F2.

weak. There is one other possible way of traversing weak fragment, shown in Fig. 5(b) Let $X \lhd^* Y$ be the weak dominance edge. The traversal proves that Y belongs to the weakly connected components of one of the X_i, so the $\varphi \wedge X_n \lhd^* Y$ is unsatisfiable.



(a) strong (b) weak (c) island

Figure 5: Traversals through fragments of free roots

This shows that the hole X_n cannot be identified with any root, i.e., φ does not have any configuration in contrast to our assumption.

island. Free island fragments permit only one single non-trivial form of traversal, depicted in. Fig. 5(c). But such traversals are not hypernormal.

□

Proposition 3. A configuration of weakly connected dominance net φ also configure its normalization $\mathsf{norm}(\varphi)$, and vice versa of course.

Proof. Let C be a configuration of φ. We show that it also configures $\mathsf{norm}(\varphi)$. Let S be the simple solved form of φ that is configured by C (Lemma 1), and S' a minimal solved form of φ more general than S.

Let $X : f(Y_1, \ldots, Y_n)$ be the top-most fragment of the tree S. This fragment must also be the top-most fragment of S', which is a tree since φ is assumed to be weakly connected (Prop. 2). S' is constructed by our algorithm (Theorem 1), so that the evaluation of solve(φ) must choose X as free root in φ.

Since φ is a net, some literal $X : f(Y_1, \ldots, Y_n)$ must belong φ. Let $\varphi' = \varphi_{|\{X,Y_1,\ldots,Y_n\}}$ be the restriction of φ to the lower fragments. The weakly connected components of all $Y_1, \ldots, Y_{n-1}$ must be pairwise disjoint by F2 (which holds by Lemma 2 since X is free in φ). The X-fragment of net φ must satisfy one of three possible schemata of net fragments:

weak. There exists a unique weak dominance edge $X \lhd^* Z$ in φ and unique hole Y_n without outgoing dominance edge. The variable Z must be a root in φ and thus labeled. If Z is equal to X then φ were unsatisfiable by normality condition N2, which is impossible. Hence, Z occurs in the restriction φ' but not in the weakly connected components of any $Y_1, \ldots, Y_{n-1}$. Otherwise, the minimal solved form S' could not be configured since the hole Y_n could not be identified with any root. Furthermore, the root of the Z-component must be identified with Y_n in any configuration of φ with root X. Hence, C satisfies $Y_n \lhd^* Z$ add by normalization.

The restriction φ' must be a dominance net by Lemma 3, and hence, all its weakly connected components are nets. For all $1 \leq i \leq n-1$, the component of Y_i in φ' is configured by the subtree of C at node Y_i, while the subtree of C at node Y_n configures the component of Z in φ'. We can thus apply the induction hypothesis to prove that the normalizations of all these components are configured by respective subconfigurations of C. This yields that $\mathsf{norm}(\varphi)$ is configured by C.

strong. If the fragment of X is strong, then it is not altered by normalization. We can thus recurse to the lower fragments of φ as we did in the previous case (if there exist any).

islands. Island remain unchanged by normalization, so we can recurse to the lower fragments. □

Proposition 4. Minimal solved forms of normal, weakly connected dominance nets have configurations.

Proof. By induction over the construction of minimal solved forms, we can show that all holes of minimal solved forms have a unique outgoing dominance edge at each hole. Furthermore, all minimal solved forms are trees since we assumed connectedness (Prop.2). Thus, all minimal solved forms are simple, so they have configurations (Lemma 1). □

7 Conclusion

We have related two undespecification formalism, MRS and normal dominance constraints. We have distinguished the sublanguages of MRS-nets and normal dominance nets that are sufficient to model scope underspecification, and proved their equivalence. Thereby, we have obtained the first provably efficient algorithm to enumerate the readings of underspecified semantic representations in MRS.

Finally, our encoding has the advantage that researchers interested in dominance constraints can soon benefit from the large grammar resources of MRS. This requires further work in order to deal with unrestricted versions of MRS used in practice.

References

H. Alshawi and R. Crouch. 1992. Monotonic semantic interpretation. In *Proc. 30th ACL*, pages 32–39.

E. Althaus, D. Duchier, A. Koller, K. Mehlhorn, J. Niehren, and S. Thiel. 2003. An efficient graph algorithm for dominance constraints. *Journal of Algorithms*. In press.

Manuel Bodirsky, Denys Duchier, Joachim Niehren, and Sebastian Miele. 2003. An efficient algorithm for weakly normal dominance constraints. Available at `www.ps.uni-sb.de/Papers`.

Johan Bos. 1996. Predicate logic unplugged. In *Amsterdam Colloquium*, pages 133–143.

Ann Copestake and Dan Flickinger. 2000. An open-source grammar development environment and broad-coverage english grammar using HPSG. In *Conference on Language Resources and Evaluation*.

Ann Copestake, Dan Flickinger, Ivan Sag, and Carl Pollard. 1999. Minimal Recursion Semantics: An Introduction. Manuscript, Stanford University.

Markus Egg, Alexander Koller, and Joachim Niehren. 2001. The Constraint Language for Lambda Structures. *Logic, Language, and Information*, 10:457–485.

Alexander Koller, Joachim Niehren, and Ralf Treinen. 2001. Dominance constraints: Algorithms and complexity. In *LACL'98*, volume 2014 of *LNAI*, pages 106–125.

Alexander Koller, Joachim Niehren, and Stefan Thater. 2003. Bridging the gap between underspecification formalisms: Hole semantics as dominance constraints. In *EACL'03*, April. In press.

Carl Pollard and Ivan Sag. 1994. *Head-driven Phrase Structure Grammar*. University of Chicago Press.

Uwe Reyle. 1993. Dealing with ambiguities by underspecification: Construction, representation and deduction. *Journal of Semantics*, 10(1).