# What is a DRS?

Tim Fernando*
fernando@ims.uni-stuttgart.de

August 1994

**Abstract**

Discourse representation structures (DRSs) are characterized by the transitions they induce on states. Just as first-order logic can be presented either model-theoretically or proof-theoretically, DRS transitions can be described in either semantic or syntactic terms. A semantic conception of states formed from first-order models and variable assignments (or so-called embeddings) is related (through a determinization of the transitions and the notion of a bisimulation) to a syntactic conception of states given by the DRSs themselves.

## 1 Introduction

The notion of a *discourse representation structure* (DRS) was introduced in Kamp [11] to analyze anaphora (in a manner similar to Heim [8]). Unimpressed by this "extra level of explicit syntactic representations", Barwise [1] eliminated DRSs from the analysis, arguing that what is essential is to

> view the utterance of an expression more dynamically, as having an effect on the environment shared by speaker and hearer, the effect being represented by various sorts of changes in variable assignments. (p. 2)

Focussing on the first-order fragment of Barwise [1], Groenendijk and Stokhof [6] kept DRSs out by bringing in programs from (quantified) dynamic logic (e.g., Harel [7]). In the meantime, various formal accounts of natural language semantic phenomena have been encoded in DRSs, going well beyond Kamp [11], Barwise [1] or Groenendijk and Stokhof [6]. Many of these developments are reported in the textbook Kamp and Reyle [10], a glance through which may well suggest (to the reader) that DRSs are here to stay. Or, at the very least, before DRSs are expelled from formal semantics, they deserve a hearing. Indeed, the converse can be argued to hold as well: a reply to the question "what is a DRS?" ought to spell out an interpretation $[\![K]\!]$ of a DRS $K$ that does not refer circularly to the concept of a DRS. For example, equating $[\![K]\!]$ with the (term model) object $\{K' \mid K \equiv K'\}$ (built from DRSs) merely shoves the problem over to providing a semantic basis for the logic behind the equivalence $\equiv$. Saying that a DRS $K$ is a piece of syntax (expressing some notion that can perhaps be captured by Heim [8]'s files) begs the question, what does that piece of syntax denote?

Now, under so-called dynamic semantics, a syntactic entity $\varphi$ is interpreted as a binary relation $[\![\varphi]\!]$ on a set of "states", specifying the change induced by $\varphi$

$$s[\![\varphi]\!]s' \quad \text{iff} \quad \text{on input } s,\ \varphi\ \textit{can}\ \text{output } s'\ . \tag{1}$$

The first two sections of the present paper concentrate on an interpretation $P$ of DRSs $K$ as input/output relations $P(K)$. To explain how a DRS effects the changes associated with it — i.e., to say what the word "can" in (1) means —, section 3 transforms the interpretation $P$ to a deterministic form in which DRSs may also be viewed as states. This dual nature of DRSs (as, on the one hand, input/output relations, and as,

*My thanks to Hans Kamp, to Robin Cooper, to the participants of a seminar in Stuttgart (Steve Berman, Ed Keenan, and Robert van Rooy), to Emiel Krahmer and to Josef van Genabith for helpful discussions.

on the other hand, inputs or outputs) is employed to give non-representational views of the kind expressed above by Barwise computational content. In particular, the operationalizability of what Groenendijk and Stokhof [6] somewhat loosely call programs ("computing" the desired input/output relations) is described by appealing to "an extra level of explicit syntactic representations" — a move that is, with hindsight, hardly surprising, given that computation requires (syntactic) coding.

## 1.1 Related work

The present paper was prompted by Cooper [2], which considers DRSs as abstracts, with the emphasis, as in Barwise [1], on the role played by variable assignments. Without denying the validity of such a view (although stressing also the importance for that view of the first-order models that structure the ranges of the variable assignments), an alternative conception of DRSs as abstracts is developed below that mentions neither variable assignments nor first-order models but only DRSs. That conception is somewhat syntactic, in contrast to Zeevat [18], where a DRS $(U, C)$ (consisting of a set $U$ of variables or discourse markers, and a set $C$ of conditions) is interpreted semantically as $(U, F)$, with $F$ the set of *verifying embeddings* of $C$, or to Muskens [14], where DRSs are interpreted as input/output relations (which are encoded in a type theory). The states over which the inputs and outputs vary are taken there to be total functions — i.e., functions defined on all variables from which DRSs are built. In this respect, the account is not faithful to DRT, and, indeed, Muskens [14] concedes that "I leave a partialization of the theory presented in this paper for future research" (p. 469).

## 1.2 Summary of the present work

Beyond providing a partialization (left open in Muskens [14]) of the relational interpretation of DRSs, the present paper develops a dual interpretation of DRSs as states, in order to understand DRSs as programs (a basic challenge for which is to make sense computationally of the interpretation of negation as the complement of the halting problem). The key is to determinize the interpretation $P(K) \subseteq S \times S$ of a DRS $K$ along the lines of the classic "subset construction" that turns a non-deterministic finite automaton into a deterministic one accepting the same language (e.g. Hopcroft and Ullman [9]). Rather than constructing (determinized) transitions from arbitrary subsets of $S$, it is sufficient to consider sets $\theta \subseteq S$ generated by a fixed set $\theta_0 \subseteq S$ of "initial" states. These states $\theta$ are named by DRSs $K$, according to the transitions induced by $K$ from $\theta_0$. More precisely, for every accessible set $\theta$, there is a DRS $K$ that *names* $\theta$ inasmuch as $P^0(K) = \theta$, where[1]

$$P^0(K) \quad := \quad \{s \in S \mid (\exists s_0 \in \theta_0)\ s_0\ P(K)\ s\}\ .$$

The determinized transitions can then be presented syntactically as

$$K \stackrel{K'}{\to} K'' \quad \text{iff} \quad K'' = K^\frown K'$$

where the *merge*, $K^\frown K'$, of two DRSs is defined to combine discourse markers and conditions respectively,

$$(U, C)^\frown(U', C') \quad = \quad (U \cup U', C \cup C') \text{ modulo renaming of variables in } C \text{ (to avoid capture)}\ .$$

The match-up with the determinization of $P$ can be made precise in terms of the notion of *bisimulation* equivalence (Park [16]) between states that arise from the possibility (as opposed to the impossibility) of $P$-transitions. In particular, two DRSs $K$ and $K'$ are bisimulation equivalent iff $P^0(K) = P^0(K')$ (Theorem 10), which is denoted $K \equiv_0 K'$. The merge $K^\frown K'$ can be interpreted incrementally under $P$ as the relational composition $P(\hat{K}) \circ P(K')$, for an appropriate $\hat{K}$ that is $\equiv_0$-equivalent to $K$ (Proposition 2).

A certain set of (quantifier-free) formulas $\varphi$ called *updates* is isolated such that for every DRS $K$, there is an update $\varphi$ that induces a DRS $K_\varphi$ that is $\equiv_0$-equivalent to $K$ (Update lemma). A slightly altered definition of a DRS is discussed briefly that avoids complications arising from variable clashes, and yields updates $\varphi$ that not only support the Update Lemma, but also enjoy the property that $P(K_\varphi)$ can be determined from either $P^0(K_\varphi)$ or its fixed points (i.e., verifying embeddings). This property provides a

[1] Throughout the present paper, the notation := will be used for definitional equality, ::== for a Backus-Naur-Form/rewrite/grammar arrow (to alternatives separated by |), and = for ordinary equality.

pleasing bridge between static (fixed point) semantics and dynamic (relational) semantics, overcoming the apparent discrepancy in complexity between the meaning of a formula being a binary relation on $S$ (i.e., the state transitions induced by the formula) and the meaning being simply a subset of the set $S$ of states (i.e., the states supporting the formula).

The present work is confined to the first-order DRSs in Kamp and Reyle [10]. The first-order character of these DRSs is exposed relative to the measure $P^0$ by a characterization lemma associating first-order formulas with DRSs (Lemma 1) and DRSs with first-order formulas (Lemma 3). A brief argument is presented in the concluding section suggesting that first-order DRSs capture the "logical core" of DRT, although a careful analysis of the matter will have to be taken up elsewhere.

# 2 First-order DRSs

Fix a signature $\mathbf{L}$ (with equality) and an infinite set $X$ of variables. *First-order discourse representation structures* (DRSs) $K$ and *first-order conditions* $\gamma$ are generated simultaneously from finite subsets $U$ of $X$, and atomic $\mathbf{L}$-formulas $A$ with variables from $X$ according to

$$\begin{array}{rcl} K & ::== & (U,\emptyset) \mid K^{\frown}\gamma \\ \gamma & ::== & A \mid \neg K \ . \end{array}$$

A DRS $K$ and a (first-order) condition $\gamma$ are interpreted relative to an $\mathbf{L}$-model $M$ by $P_M(K) \subseteq S_M \times S_M$ and $v_M(\gamma) \subseteq S_M$ respectively, where $S_M$ is the set of $M$-assignments, or (following Kamp [11]) "embeddings" — i.e., partial functions $f$ from $X$ to the universe of $M$. The semantic clauses corresponding to the syntactic clauses above are

$$\begin{array}{rcl} P_M(U,\emptyset) & := & \{(f,g) \in S_M \times S_M \mid f \subseteq g \text{ and } \mathrm{dom}(g) = \mathrm{dom}(f) \cup U\} \\ v_M(A) & := & \{f \in S_M \mid M \models A[f]\} \\ P_M(K^{\frown}\gamma) & := & \{(f,g) \in S_M \times S_M \mid f\ P_M(K)\ g \text{ and } g \in v_M(\gamma)\} \\ v_M(\neg K) & := & \{f \in S_M \mid \text{there is } \textit{no } g \text{ such that } f\ P_M(K)\ g\} \\ & (= & S_M - \mathrm{dom}(P(K))) \ . \end{array}$$

The overall interpretations are the disjoint sums

$$\begin{array}{rcl} P(K) & := & \{((M,f),(M,g)) \mid M \in \hat{\mathcal{M}} \text{ and } f\ P_M(K)\ g\} \\ v(\gamma) & := & \{(M,f) \mid M \in \hat{\mathcal{M}} \text{ and } f \in v_M(\gamma)\} \end{array}$$

over the class $\hat{\mathcal{M}}$ of all $\mathbf{L}$-models. (The reader worried that $\hat{\mathcal{M}}$ is a proper class can replace it by a sufficiently large set of $\mathbf{L}$-models — e.g., models, by Löwenheim-Skolem, with objects drawn from $\max(|\mathbf{L}|, \aleph_0)$.)

## 2.1 DRS normal form

A DRS $K$ can be flattened into a pair $(U,C)$ of finite sets $U$ of variables and $C$ of conditions, by setting

$$(U,C)^{\frown}\gamma \quad := \quad (U, C \cup \{\gamma\}) \ ,$$

from which it then follows that

$$P_M(U,C) \quad = \quad \{(f,g) \in S_M \times S_M \mid f\ P_M(U,\emptyset)\ g \text{ and } (\forall \gamma \in C)\ g \in v_M(\gamma)\} \ .$$

As relations, DRSs have the normal form

$$P(U,C) \quad = \quad P(U,\emptyset) \circ P(\emptyset,C) \ ,$$

where the second relation,

$$P(\emptyset,C) \quad = \quad \{(s,s) \mid (\forall \gamma \in C)\ s \in v(\gamma)\} \ ,$$

is independent of $U$, in contrast to the set of fixed points

$$V(U,C) \quad := \quad \{(s,s) \mid s\ P(U,C)\ s\} \ ,$$

relative to which $P(U,C)$ can also be characterized as $P(U,\emptyset) \circ V(U,C)$. In either case, $P(U,C)$ cannot be determined from the second relation (viz., $P(\emptyset,C)$ or $V(U,C)$), as can be seen by considering the pair of DRSs $(\{x\},\{x=x\})$ and $(\emptyset,\{x=x\})$. Elements of the set $V(K)$ are precisely the verifying embeddings of $K$ in Kamp and Reyle [10]. The truth of $K$ at a model $M$ is equated there with the assertion $\emptyset \in \mathrm{dom}(P_M(K))$, which (as the same pair of DRSs above shows) cannot be determined from the fixed points (i.e., verification set) of $K$ alone. The complication, however, would seem to be easy enough to overcome, arising only because the set $U$ in a DRS $K = (U,C)$ is independent from the set of discourse markers referred to in $C$. This matter is pursued in section 3, but first, let us consider another subset of $S$ that is naturally associated with $K$.

### 2.2 The moment of a DRS

Concentrating on transitions from the empty (i.e., nowhere defined) embedding $\emptyset$, define the *moment of* a DRS $K$ to be

$$P^0(K) \quad := \quad \{(M,f) \mid \emptyset\ P_M(K)\ f\} \ .$$

Given a DRS $K = (U,C)$, call a first-order formula $\chi$ a *characteristic formula of* $K$ if the set of variables occurring freely in $\chi$ is $U$, and

$$P^0(K) \quad = \quad \{(M,f) \mid \mathrm{dom}(f) = U \text{ and } M \models \chi[f]\} \ .$$

**Lemma 1** (Characterization lemma, Part 1). *Every DRS has a characteristic formula.*

**Proof**. An inductive argument is facilitated by strengthening the assertion as follows

> for every DRS $(U,C)$ and every finite set $U_0$ of variables, there is a first-order formula $\chi$ such that for all $M$, $f : U_0 \to |M|$, and $g$,
>
> $$f\ P_M(U,C)\ g \quad \text{iff} \quad f\ P_M(U,\emptyset)\ g \text{ and } M \models \chi[g] \ .$$

The cases $(U,\emptyset)$ and $K^\frown A$ are trivial. As for $(U,C)^\frown\neg K$, appeal to the inductive hypothesis, as strengthened above, to form a conjunction of formulas, the second of which is (ignoring pesky subscripts for variables) $\neg\exists\overline{x}\ \chi_K$, where $\overline{x}$ is a list of variables declared in $K$ but which are not (already) in $U \cup U_0$. $\dashv$

The restriction to transitions from the empty embedding $\emptyset$ cannot simply be dropped, in view of the existence of DRSs $K$ such that for every first-order **L**-formula $\chi$ and every finite set $U$ of variables, there is an **L**-model $M$ satisfying

$$P_M(K) \quad \neq \quad \{(f,g) \mid f\ P_M(U,\emptyset)\ g \text{ and } M \models \chi[g]\} \ .$$

A trivial example is provided by the DRS $(\emptyset,\{x=x\})$, although more interesting examples can be drawn from a certain special family of DRSs (to which $(\emptyset,\{x=x\})$ does not belong), specified in the next section. On the positive side, the converse to Lemma 1 will also be established there, with $K$ chosen from that family.

## 3 DRSs as updates

The present section isolates a certain family of DRSs given by quantifier-free first-order formulas that includes $\equiv_0$-representatives of all DRSs. As a first approximation, consider the set of quantifier-free formulas $\varphi$ obtained by closing atomic formulas $A$ under negation and conjunction

$$\varphi \quad ::== \quad A \mid \neg\varphi \mid \varphi \& \psi \ ,$$

inducing DRSs $K_\varphi$ according to

$$\begin{aligned} K_A &:= (\mathrm{VAR}(A), \{A\}) \\ K_{\neg\varphi} &:= (\emptyset, \{\neg K_\varphi\}) \\ K_{\varphi\&\psi} &:= K_\varphi \bullet K_\psi \ , \end{aligned}$$

where $\mathrm{VAR}(A)$ is the set of variables occurring in $A$, and the operation $\bullet$ combines discourse markers and conditions respectively,

$$(U, C) \bullet (U', C') \quad := \quad (U \cup U', C \cup C') \ .$$

A sequential incremental interpretation of DRSs suggests that $P(K \bullet K')$ be the relational composition $P(K) \circ P(K')$. But how does $P(K) \circ P(K')$ square with the normal form of $P(K \bullet K')$, which requires all variables initializations to precede all tests of conditions? Irreconcilable differences may lurk, but only if some variable in (a condition in) $K$ left uninitialized by $P(K)$ is initialized by $P(K')$. (For example, consider $P(\emptyset, \{x = x\}) \circ P(\{x\}, \emptyset)$.) Accordingly, define

$$(U, C)^\frown(U', C') \quad := \quad (U \cup U', C[U, U'] \cup C')$$

where $C[U, U']$ is the set $C$ of conditions with variables in $U' - U$ renamed to be disjoint from $U \cup U'$ (so as to avoid being captured by the DRS $(U', \emptyset)$). A definite scheme for renaming variables can be fixed according to a well-ordering on the set $X$ of variables, but we will not worry about just what that is. Defining

$$K \equiv_0 K' \quad \text{iff} \quad P^0(K) = P^0(K') \ ,$$

the important point is that $(U, C[U, U']) \equiv_0 (U, C)$, which is to say that $^\frown$ is just $\bullet$, with perhaps a happier choice of $\equiv_0$-equivalent DRSs.[2] Let us record this fact as

**Proposition 2**. *For all DRSs $K$ and $K'$, there is a DRS $\hat{K} \equiv_0 K$ such that $P(K^\frown K') = P(\hat{K}) \circ P(K')$. Furthermore, in the absence of variable clashes (i.e., in case $K^\frown K' = K \bullet K'$), $\hat{K}$ can be chosen to be $K$.*

Now, define the set of *(first-order) updates* $\varphi$ inductively by

$$\varphi \quad ::== \quad A \mid \neg\varphi \mid \varphi^\frown\psi \ ,$$

subject to the previous associations $K_\varphi$ of DRSs, with $K_{\varphi^\frown\psi}$ arranged to be $K_\varphi{}^\frown K_\psi$ either directly or by defining $\varphi^\frown\psi$ to be $\varphi_\psi\&\psi$ for the appropriate alphabetic variant $\varphi_\psi$ of $\varphi$. In either case, let us identify $\varphi^\frown\psi$ with $\varphi\&\psi$ if $K_{\varphi^\frown\psi} = K_\varphi \bullet K_\psi$. Notice that if $\varphi$ and $\psi$ are updates, then so is $\neg(\varphi\&\neg\psi)$, which can be understood (following their equivalence in classical logic) as a definition of $\varphi \supset \psi$. In terms of DRSs and conditions, the condition $\neg(K^\frown\neg K')$ can be rewritten as $K \supset K'$ so that

$$v(K \supset K') \quad = \quad \{s \mid (\forall s' \text{ s.t. } s \ P(K) \ s') \ s' \in \mathrm{dom}(P(K'))\} \ .$$

Next, define the interpretation $[\![\varphi]\!]$ of an update $\varphi$ to be $P(K_\varphi)$. By Proposition 2, if $\varphi^\frown\psi$ is $\varphi\&\psi$, then $[\![\varphi^\frown\psi]\!]$ is equal to the relational composition $[\![\varphi]\!] \circ [\![\psi]\!]$. Moreover, the converse to Lemma 1 can be established (as promised).

**Lemma 3** (Characterization lemma, Part 2). *For every first-order formula $\chi$ with set $U$ of free variables, there is an update $\varphi$ such that*

$$P^0(K_\varphi) \quad = \quad \{(M, f) \mid \mathrm{dom}(f) = U \text{ and } M \models \chi[f]\} \ .$$

[2]For this reason, the alternative definition of $(U, C)^\frown(U', C')$ in terms of $\bullet$ where $(U', C')$ is modified rather than $C$ has not been adopted, as it seems more reasonable to rename variables buried inside a DRS than free variables visible in the domain of a verifying embedding.

**Proof**. Assume without loss of generality that a variable $x$ cannot occur both free and bound in $\chi$, and that it is bound (quantificationally) at most once in $\chi$. Then build an update $\varphi_\chi$ inductively from $\chi$ according to

$$\begin{array}{rcl} A & \mapsto & A \\ \chi \& \chi' & \mapsto & \varphi_\chi {}^\frown \varphi_{\chi'} \\ \neg\chi & \mapsto & \mathrm{init}(U^{\varphi_\chi})^\frown \neg\varphi_\chi \\ \exists x\ \chi & \mapsto & \mathrm{init}(U^{\varphi_\chi} - \{x\})^\frown \neg\neg((x = x)^\frown \varphi_\chi) \ , \end{array}$$

where $\mathrm{init}(U)$ is the conjunction $(x = x)^\frown \cdots {}^\frown (x' = x')$ over all variables $x, \ldots, x' \in U$, and where $U^\varphi$ is the set of discourse markers such that $K_\varphi = (U^\varphi, C)$ for some set $C$ of conditions. $\dashv$

An immediate corollary of the characterization lemma (parts 1 and 2) is

**Lemma 4** (Update lemma). *For every DRS $K$, there is an update $\varphi$ such that $K \equiv_0 K_\varphi$.*

The update lemma can also be proved directly by an inductive argument.

### 3.1 An example

To illustrate the idea behind updates, let us consider (what else?) the notorious "donkey" sentence

If a farmer owns a donkey, he beats it.

A naive translation into first-order logic (based on a decomposition of the sentence into an implication between two sentences) is

$$(\exists x)(\exists y)(\mathtt{farmer}(x)\ \&\ \mathtt{donkey}(y)\ \&\ \mathtt{own}(x,y)) \quad \supset \quad \mathtt{beat}(x,y)\ , \tag{2}$$

which corresponds (in a sense to be explained below) to the DRS $(\emptyset, \{K \supset K'\})$, where[3]

$$\begin{array}{rcl} K & := & (\{x,y\}, \{\mathtt{farmer}(x), \mathtt{donkey}(y), \mathtt{own}(x,y)\}) \\ K' & := & (\emptyset, \{\mathtt{beat}(x,y)\})\ . \end{array}$$

Observe that the DRS $K'$ cannot be induced by an update because the variables $x$ and $y$ are not declared in $K'$. On the other hand, the DRS $(\emptyset, \{K \supset (\{x,y\}, \{\mathtt{beat}(x,y)\})\})$ induced by the (quantifier-free) update

$$(\mathtt{farmer}(x)\ \&\ \mathtt{donkey}(y)\ \&\ \mathtt{own}(x,y)) \quad \supset \quad \mathtt{beat}(x,y) \tag{3}$$

is $\equiv_0$-equivalent to $(\emptyset, \{K \supset K'\})$, with which it shares the characteristic formula

$$(\forall x)(\forall y) \quad ((\mathtt{farmer}(x)\ \&\ \mathtt{donkey}(y)\ \&\ \mathtt{own}(x,y)) \ \supset\ \mathtt{beat}(x,y))\ .$$

(Indeed, $P(\emptyset, \{K \supset (\{x,y\}, \{\mathtt{beat}(x,y)\})\}) = P(\emptyset, \{K \supset K'\})$.) Thus, by restricting ourselves to updates, we lose the faithful formulation of the succedent of (2) by a DRS $K'$ with an empty set of discourse markers, reflecting the absence of quantification in the succedent. That is to say, the difference between (2) and (3) — or, more precisely, between their corresponding DRSs — concerns explicit quantification. (The quantifier-free character of updates here is reminiscent of Pagin and Westersthl [15].) But if that bit of explicitness is so important, it is rather curious that the semantics of DRSs is defined so that even if $U$ is non-empty, the DRS $(U, \emptyset)$ cannot change an input assignment whose domain already includes $U$. That is, while it is tempting to record within the component set $U$ of a DRS $(U, C)$, the "novel" character of the indefinite "a" in "a farmer," the fact is that even if a variable is in $U$, $P(U, \emptyset)$ may have no effect on it.

So why not redefine the semantics of $P(U, \emptyset)$ so that new values can always be assigned to all variables in $U$? This is one of the features of the formalism *Dynamic Predicate Logic* (DPL) of Groenendijk and Stokhof

[3] We concentrate here on the so-called "strong" reading, although much the same applies to the "weak" reading which is obtained by taking $K$ to be $(\{x\}, \{\mathtt{farmer}(x), \neg(\emptyset, \{\neg(\{y\}, \{\mathtt{donkey}(y), \mathtt{own}(x,y)\})\})\})$, and $K'$ to be $(\{y\}, \{\mathtt{donkey}(y), \mathtt{own}(x,y), \mathtt{beat}(x,y)\})$.

[6] that sets it apart from DRT.[4] Rather than introducing DRSs and conditions, DPL translates first-order formulas into programs from (quantified) dynamic logic (e.g., Harel [7]), in which variable assignments are *total* functions defined on the full set $X$ of variables. This use of total functions means that DPL programs (i.e., so-called random assignments) corresponding to $P(U, \emptyset)$ must be free to overwrite values to variables in $U$. By contrast, information is never destroyed by a DRS in the sense that variable assignments can only be extended. Even if DPL were to abandon the requirement that variable assignments be total, the question is whether the semantics of $P(U, \emptyset)$ ought to be redefined, concerning which the present author offers three arguments for the status quo.

(i) Forcing an assignment of a value to a variable to model the utterance of an indefinite such as "a" in "a farmer" can be arranged without changing the semantics of $P(U, C)$, but simply by a suitable choice of a variable (for the indefinite farmer) to put into the set $U$. So-called "novelty" and "familiarity" conditions related to this choice belong to part of the passage from natural language utterances to formulas that falls outside the scope of the analysis given by the interpretation $P$.

(ii) The non-destructive nature of $P(U, \emptyset)$ is exploited heavily in Fernando [5] to define a Boolean-valued notion of truth for a relational interpretation that extends DRT updates conservatively with witness constructs $\epsilon x : \varphi$ for explicit existential quantification $\exists x\ \varphi$. In addition to the update (3), that interpretation supports a translation of the donkey sentence as

$$(\exists x)(\exists y)(\mathtt{farmer}(x)\ \&\ \mathtt{own}(x,y)\ \&\ \mathtt{donkey}(y)) \quad \supset \quad \mathtt{beat}(\epsilon x : \exists y \varphi, \epsilon y : \psi)\ ,$$

where $\varphi$ is $\mathtt{farmer}(x)\ \&\ \mathtt{own}(x,y)\ \&\ \mathtt{donkey}(y)$, and $\psi$ is $\varphi$ with $x$ replaced by $\epsilon x : \exists y \varphi$. The formula $\varphi$ in an $\epsilon$-term $\epsilon x : \varphi$ can be understood situation-theoretically (e.g., Cooper [2]) as a restriction on the parameter $x$.

(iii) A final point that the present author owes to a remark by C. Gardent is that the initialization effects of updates (from the clause $K_A = (\mathrm{VAR}(A), \{A\})$) appear useful for treating kataphora.

### 3.2 Some technical points concerning variables and compositionality

Before proceeding in the next section to provide further evidence for identifying the essence of a DRS $K$ with its moment $P^0(K)$, let us pause here to note some wrinkles in such a proposal that have to do with variables.

It is easy to see that $\equiv_0$ does not imply equivalence relative to $P$, because of variables that occur internally in a DRS. Let $A$ be a unary relation symbol, and consider the two $\equiv_0$-equivalent updates $\neg A(x)$ and $\neg A(y)$, with characteristic formula $\forall z \neg A(z)$. This pair shows that $P(K)$ is not (in general) determined by $P^0(K)$ or by $\{(M, f) \in V(K) \mid \mathrm{dom}(f) = U\}$ where $K = (U, C)$. It also provides a counter-example to the strengthening of Lemma 1 mentioned in the end of §2.2. Furthermore, the pair refutes the claim that $\equiv_0$ is a congruence with respect to the merge operation $^\frown$ (or, in other words, that $P^0$ is compositional with respect to $^\frown$)

$$K_{\neg A(x)} \equiv_0 K_{\neg A(y)} \quad \text{but} \quad K_{x=x}{}^\frown K_{\neg A(x)} \not\equiv_0 K_{x=x}{}^\frown K_{\neg A(y)}\ .$$

This defect can be corrected by modifying the operation $^\frown$ to $+$, just as $\bullet$ was modified to $^\frown$

$$(U, C) + (U', C') \quad := \quad (U \cup U', C[U, U'] \cup C'[U', U])\ .$$

Under this more symmetric revision of $\bullet$, we again have $P(K + K') = P(\hat{K}) \circ P(\hat{K'})$, for some $\hat{K} \equiv_0 K$ and $\hat{K'} \equiv_0 K'$, where, in the absence of variable conflicts, $\hat{K}$ and $\hat{K'}$ can be taken to be $K$ and $K'$ respectively. Moreover,

**Lemma 5** (Congruence lemma). *$K_1 \equiv_0 K_1'$ and $K_2 \equiv_0 K_2'$ imply $K_1 + K_2 \equiv_0 K_1' + K_2'$.*

**Proof**. The lemma, in fact, holds with $\equiv_0$ weakened to $\equiv_0^M$, where $K \equiv_0^M K'$ means that for all $f$, $\emptyset\ P_M(K)\ f$ iff $\emptyset\ P_M(K')\ f$. Assume without loss of generality that $K_1 + K_2 = K_1 \bullet K_2$. (Otherwise, replace the DRSs

[4] A comparative study of variables in dynamic semantics is undertaken in Vermeulen [17]. A more recent paper, closer in spirit to that of the present work, is Dekker [3], which emphasizes the similarities with first-order logic.

by $\equiv_0$-equivalent DRSs giving the same sums.) Under this assumption, it is easy to see that $\chi_{K_1+K_2}$ is equivalent to $\chi_{K_1}\&\chi_{K_2}$, where $\chi_K$ is the characteristic formula of $K$ given by Lemma 1. Applying the same reasoning to $K_1'$ and $K_2'$, conclude $\chi_{K_1'+K_2'}$ is equivalent to $\chi_{K_1'}\&\chi_{K_2'}$, and hence to $\chi_{K_1}\&\chi_{K_2}$, provided $K_1 \equiv_0 K_1'$ and $K_2 \equiv_0 K_2'$. $\dashv$

Lemma 5 and the failure of $^\frown$ to be a congruence suggest the following notion of a safe merge: $(U, C)$ *safely merges with* $(U', C')$ if $C'[U', U] = C'$ — i.e., if $K^\frown K' = K + K'$. Now, it is natural to inquire: why not ban merges that are dangerous? Unfortunately, Lemmas 3 and 4 would not survive a substitution of $^\frown$ by $+$; consider, for example, the first-order formula $\forall y \neg R(x, y)$ (where $R$ is a binary relation symbol), which is the characteristic formula of the update $(x = x)^\frown \neg R(x, y)$, but *not* of $x = x + \neg R(x, y)$ or of any update built from atomic formulas using at most $\neg$ and $+$.

Of course, Lemma 3 (whence Lemma 4) could be restored by admitting arbitrary first-order **L**-formulas (and not just atomic ones) as basic conditions. Indeed, if the only conditions allowed are first-order **L**-formulas (i.e., if all negations $\neg K$ of DRSs $K$ are replaced by first-order **L**-formulas, with verification conditions given according to the usual semantics of first-order logic), then the irksome variable clashes noted above evaporate. In particular, under this modification, the input/output relations $[\![\varphi]\!]$ of the resulting updates $\varphi$ can be determined from their moments $P^0(K_\varphi)$, or from their fixed points (i.e., verifying embeddings) as follows:

$$[\![\varphi]\!] \quad = \quad \{(s, s') \mid s\ P(U_\varphi, \emptyset)\ s' \text{ and } s'[\![\varphi]\!]s'\}\ ,$$

where (by definition)

$$U_\varphi \quad := \quad \bigcap\{\mathrm{dom}(f) \mid (M, f)[\![\varphi]\!](M, f)\}$$

(with $[\![\varphi]\!] = \emptyset$ if $[\![\varphi]\!]$ has no fixed point). Such harmony between static (fixed point) semantics and dynamic (relational) semantics (which simplifies Zeevat [18] by eliminating the set $U$ of dicourse markers from the interpretation of a DRS $(U, C)$) is not possible so long as conditions can be built from negations of DRSs under the semantics specified in section 2.[5] The approach taken in the present paper to overcoming variable conflicts has been to define merge operations $^\frown$ and $+$ that modify $\bullet$ by renaming the variables. But such maneuvers can be avoided by identifying conditions with arbitrary first-order formulas, leading to different input/output relations. (Note that none of the DRSs defined in section 2 has, for example, the input/output relation $\{((M, f), (M, f)) \mid f \in S_M \text{ and } M \models \forall x \neg A(x)\}$.) On the other hand, the same moments $P^0(K)$ would result (since these are still characterized by first-order formulas), and thus the DRSs remain essentially the same, provided the essence of a DRS $K$ is taken to be its moment $P^0(K)$. We will return to this matter briefly in the next section (§4.3).

# 4 DRSs as states

Assuming a DRS amounts to an input/output relation, what does it mean for the input/output pair $(s, s')$ to be in that relation? If there are two input/output pairs $(s, s_1')$ and $(s, s_2')$ with the same input state $s$, which state does the DRS output on input $s$? And secondly, whether or not there is a choice of outputs $s_1', s_2'$ possible, how is negation, which is interpreted as the complement of the halting problem, computed? (Notice that negation leads, via the reduction $\neg\exists x\ \neg A$, to universal formulas $\forall x\ A$, falling outside the realm of the recursively enumerable.) An attempt to understand the meaning of "can" in the equivalence

$$s[\![\varphi]\!]s' \quad \text{iff} \quad \text{on input } s,\ \varphi\ \textit{can} \text{ output } s'$$

runs against two questions: "which?" and "how?" The issues here of non-determinism[6] and operationalizability are not independent, and are best taken up in sequence, by identifying $[\![\cdot]\!]$ not with $P$ but with

[5] Consider the update described in the proof of Lemma 3 that corresponds to the unsatisfiable first-order sentence $\exists x \neg \exists y\ x = y$. That update, call it $\varphi$, induces a DRS with the same input/output relation as the DRS $(\emptyset, \{\neg(\emptyset, \{\neg K\})\})$, where $K$ is $(\{x\}, \{\neg(\{x, y\}, \{x = y\})\})$. Now, it follows (by considering an **L**-model $M$ with at least two elements, and arguing that $[\![\varphi]\!]$ cannot return an output if the input assignment is undefined on $y$) that $U_\varphi = \{y\}$, yielding the wrong input/output relation $[\![\varphi]\!]$.

[6] Non-determinism does *not* arise in the present context from ambiguity. The formula $\varphi$ is unambiguous, insofar as $[\![\cdot]\!]$ is a function. As a matter of methodology, ambiguity is shoved over to the (informal) stage of translating an utterance to a formula: an utterance may admit translations to several formulas.

$\mathcal{S} \circ \mathcal{D} \circ P$, where $\mathcal{D}$ and $\mathcal{S}$ "determinize" and "syntacticize" $P$ respectively. Actually, to understand the effects of determinization $\mathcal{D}$, it is useful to fix an $\mathbf{L}$-model $M$ and to determinize $P_M$ before summing over all $\mathbf{L}$-models (to get $P$).

## 4.1 Transitions determined by a model

Given a set $L$ of "labels" on transitions $s \xrightarrow{l} s'$ between "states" $s, s'$, it will often prove convenient to repackage the *$L$-transition predicate* $\rightarrow$, by which is mean any subset of $S \times L \times S$ for some set $S$ of states, as the $L$-model $\langle S, \{\xrightarrow{l}\}_{l \in L}\rangle$, where every $l \in L$ is understood to be a binary relation symbol. The chief example of $L$ that will concern us is the set $\mathcal{DRS}$ of DRSs (construed as binary relation symbols) over a fixed signature $\mathbf{L}$ (not to be confused with $L$), although we will have cause to expand that set later. The interpretation $P$ of DRSs provides a recipe for cooking up a $\mathcal{DRS}$-transition predicate $\rightarrow_M$ from an $\mathbf{L}$-model $M$; to wit,

$$f \xrightarrow{K}_M g \quad \text{iff} \quad f\ P_M(K)\ g\ .$$

Abusing notation, $P_M$ will be identified with the transition predicate it induces above; and similarly for $P$. Let us agree to understand isomorphism $\cong$ between $L$-transition predicates to mean isomorphism between the associated $L$-models.

**Theorem 6** (essentially Fernando [4]). *For all finite or countable* $\mathbf{L}$*-models $M$ and $N$, $P_M \cong P_N$ iff $M \cong N$.*

The proof of the non-trivial direction (from left to right) involves a notion that will interest us further below, and has come to be known in the computer science literature as a bisimulation (Park [16]). Given two $L$-transition predicates $\rightarrow \subseteq S \times L \times S$ and and $\rightharpoondown \subseteq S' \times L \times S'$, a binary relation $E \subseteq S \times S'$ is a *bisimulation* if whenever $sEs'$ then for every $l \in L$,

$$(\forall x \xleftarrow{l} s)\ (\exists x' \xleftharpoondown{l} s')\ xEx' \quad \text{and} \quad (\forall x' \xleftharpoondown{l} s')\ (\exists x \xleftarrow{l} s)\ xEx'\ .$$

Because $E$ occurs only positively in the condition above, there is a $\subseteq$-largest bisimulation (relative to $\rightarrow$ and $\rightharpoondown$), called *bisimilarity* and denoted $\underline{\leftrightarrow}$. When $s \underline{\leftrightarrow} s'$, the two states are said to be *bisimilar*, or to be more explicit, the *pointed* transition predicates $(\rightarrow, s)$ and $(\rightharpoondown, s')$ are *bisimilar*, denoted again $(\rightarrow, s) \underline{\leftrightarrow} (\rightharpoondown, s')$. Now, returning to Theorem 6, it is easy to see (appealing to DRSs of the form $(\emptyset, \{x = x\})$) that $P_M \cong P_N$ implies $(P_M, \emptyset) \underline{\leftrightarrow} (P_N, \emptyset)$. Moreover, so long as the label set $L$ includes the DRSs $(\{x\}, \emptyset)$ (for countably many $x \in X$) and $(\emptyset, \{A\})$ for all atomic $\mathbf{L}$-formulas over $X$, a bisimulation between $P_M$ and $P_N$ relating $\emptyset$ to itself yields a partial isomorphism family, supporting a standard "back-and-forth" argument in model theory (e.g., Keisler [13]) that sums, over finite or countable models, to an isomorphism.

The dependence of the transition predicate $P_M$ on its underlying model $M$ can be reduced by internalizing the non-determinism within "disjunctive" states by an operation $\mathcal{D}$, which is essentially the well-known subset construction reducing non-deterministic finite automata to deterministic ones (e.g., Hopcroft and Ullman [9]). It will be convenient for our purposes to define $\mathcal{D}$ on pointed transition predicates $(\rightarrow, s_0)$, restricting the newly formed states to non-empty sets accessible from $s_0$. More precisely, given a transition predicate $\rightarrow$ and a state $s_0$, let $\mathcal{D}(\rightarrow, s_0)$ be the pointed transition predicate $(\Rightarrow, \{s_0\})$ generated inductively from its point $\{s_0\}$ as follows: for all $l$ and $l' \in L$,

(i) if $U = \{s : s_0 \xrightarrow{l} s\} \neq \emptyset$, then $\{s_0\} \overset{l}{\Rightarrow} U$, and

(ii) if $U' \overset{l'}{\Rightarrow} U$ and $V = \{s : (\exists x \in U)\ x \xrightarrow{l} s\} \neq \emptyset$, then $U \overset{l}{\Rightarrow} V$.

Note that we have only included sets of states accessible from $s_0$, and, moreover, only *non*-empty sets. Adding the empty set would have amounted to creating a transition from the absence of a transition, trivializing the notion of a bisimulation relative to $\Rightarrow$. Observe that $\Rightarrow$ is *deterministic* in the sense that $s \overset{l}{\Rightarrow} s'$ and $s \overset{l}{\Rightarrow} s''$ imply $s' = s''$.

**Lemma 7**. *Over deterministic transition predicates $\rightarrow$ and $\rightharpoondown$, bisimilarity $\underline{\leftrightarrow}$ is just trace equivalence, or, in other words, $s \underline{\leftrightarrow} s'$ iff $tr_{\rightarrow}(s) = tr_{\rightharpoondown}(s')$, where*

$$tr_{\rightarrow}(s) \quad = \quad \{(l_1, l_2, \ldots, l_n) \mid n < \omega \text{ and } (\exists s_1, s_2, \ldots s_n)\ s \xrightarrow{l_1} s_1 \text{ and } s_1 \xrightarrow{l_2} s_2 \text{ and } \cdots \text{ and } s_{n-1} \xrightarrow{l_n} s_n\}$$

*and similarly for $tr_{\rightarrow}$.*

Now, it turns out that $\mathcal{D}$ extracts the first-order content of an **L**-model $M$,

$$Th(M) \quad := \quad \{\varphi \mid \varphi \text{ is a first-order } \mathbf{L}\text{-sentence true in } M\} \ ,$$

in the following sense.

**Theorem 8** (essentially Fernando [4]). *For all* **L***-models $M$ and $N$, the following are equivalent.*

*(i)* $\mathcal{D}(P_M, \emptyset) \cong \mathcal{D}(P_N, \emptyset)$.

*(ii)* $Th(M) = Th(N)$.

*(iii)* $\mathcal{D}(P_M, \emptyset) \underline{\leftrightarrow} \mathcal{D}(P_N, \emptyset)$.

Theorem 8 can be proved from parts 1 and 2 of the characterization lemma (helped along by Lemma 7). Recall that over finite models, elementary equivalence relative to a language with equality (i.e., condition (ii) of Theorem 8) is the same as isomorphism. Over infinite models, elementary equivalence is far weaker. In contrast to Theorem 6, Theorem 8 is very sensitive to additions to the label set. (In particular, it breaks down if labels are built from Kleene star $\cdot^*$, although the equivalences can be restored if negation $\neg$ is thrown out the same time that $\cdot^*$ is thrown in.)

## 4.2 Transitions determined by a family of models

It will be useful to strengthen Theorem 8 by passing from a single **L**-model $M$ to a family $\mathcal{M}$ of such. Towards that end, note that there is an obvious definition of a sum $\sum_{i \in I}(\rightarrow_i, s_i)$ of pointed deterministic transition predicates $(\rightarrow_i, s_i)$ such that

$$\sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset) \quad \cong \quad \hat{\mathcal{D}}(P, \{(M, \emptyset) \mid M \in \mathcal{M}\})$$

where $\hat{\mathcal{D}}(\rightarrow, U)$ is $\mathcal{D}(\rightarrow, s_0)$ except that the initial state $\{s_0\}$ is replaced by $U$. More precisely, define $\sum_{i \in I}(\rightarrow_i, s_i)$ to be the pointed transition predicate $(\Rightarrow, \{s_i \mid i \in I\})$ where $\Rightarrow$ is generated inductively as follows: (i) for all $l \in L$, if $U = \{s \mid (\exists i \in I)\ s_i \xrightarrow{l}_i s\} \neq \emptyset$, then $\{s_i \mid i \in I\} \overset{l}{\Rightarrow} U$, and (ii) for all $l, l' \in L$, if $U' \overset{l'}{\Rightarrow} U$ and $V = \{s \mid (\exists x \in U)\ x \xrightarrow{l}_i s\} \neq \emptyset$, then $U \overset{l}{\Rightarrow} V$. Now, Theorem 8 generalizes easily to families of **L**-models.

**Theorem 9** (essentially Fernando [4]). *For all families $\mathcal{M}$ and $\mathcal{N}$ of* **L***-models, the following are equivalent.*

*(i)* $\sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset) \cong \sum_{N \in \mathcal{N}} \mathcal{D}(P_N, \emptyset)$.

*(ii)* $\bigcap_{M \in \mathcal{M}} Th(M) = \bigcap_{N \in \mathcal{N}} Th(N)$.

*(iii)* $\sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset) \underline{\leftrightarrow} \sum_{N \in \mathcal{N}} \mathcal{D}(P_N, \emptyset)$.

Given that line (ii) of Theorem 9 can hold for very different families $\mathcal{M}$ and $\mathcal{N}$ of **L**-models, Theorem 9 suggests the possibility of building a much cheaper copy of the pointed transition predicate $\sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset)$.

## 4.3 Transitions determined syntactically

Insofar as $\sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset)$ is determined by $\bigcap_{M \in \mathcal{M}} Th(M)$, it is natural to seek a syntactic presentation of the transition predicate. With that in mind, given a first-order **L**-theory $T$, define a DRS $K = (U, C)$ to be $T$-*consistent* if $T \cup \{\exists x_0 \cdots \exists x_n\ \chi_K\}$ is consistent, where $U = \{x_0, \ldots, x_n\}$ and $\chi_K$ is a characteristic formula of $K$ given by Lemma 1. (The definition is clearly independent of the exact choice of $\chi_K$.) Now, let $\rightarrow_T$ be the following $\mathcal{DRS}$-transition predicate between $T$-consistent DRSs

$$\{(K, \hat{K}, K') \quad | \quad \hat{K} \in \mathcal{DRS}, \text{ and } K \text{ and } K' \text{ are } T\text{-consistent DRSs such that } K' = K^\frown \hat{K}\} \ .$$

**Theorem 10.** *Let $\mathcal{M}$ be a non-empty family of* **L**-*models, $T = \bigcap_{M \in \mathcal{M}} Th(M)$, and $P^0_{\mathcal{M}}$ be the function from DRSs obtained by relativizing $P^0$ to $\mathcal{M}$ as follows*

$$P^0_{\mathcal{M}}(K) \quad := \quad \{(M, f) \in P^0(K) \mid M \in \mathcal{M}\} \ .$$

(a) $(\rightarrow_T, (\emptyset, \emptyset)) \leftrightarrow \sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset)$.

(b) *DRSs $K$ and $K'$ are bisimilar relative to $\rightarrow_T$ (i.e., $K \leftrightarrow_T K'$) iff $P^0_{\mathcal{M}}(K) = P^0_{\mathcal{M}}(K')$.*

(c) *$\sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset)$ is "strongly extensional." That is, over the transition predicate of the pointed transition predicate $\sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset)$, bisimilarity is equality.*

(d) *Assume $T$ is r.e. Then there is an r.e. transition predicate $\rightsquigarrow_T$ with the same state set as $\rightarrow_T$ but with one more label than $\rightarrow_T$ such that two DRSs are bisimilar relative to $\rightsquigarrow_T$ iff they are bisimilar relative to $\rightarrow_T$.*

The bisimulation witnessing part (a) is given by the function $P^0_{\mathcal{M}}$, which also explains (b) and (c). (Note that the transition predicates are deterministic, whence, by Lemma 7, bisimilarity is just trace equivalence.) The equivalence asserted between the syntactic and semantic transition predicates is not terribly surprising, although a few words about the use of $^\cap$ rather than $\bullet$ or $+$ to define $\rightarrow_T$ are perhaps in order. To be more precise, why not define $K \xrightarrow{\hat{K}} K'$ to hold between $T$-consistent DRSs $K$ and $K'$ precisely if $K \bullet \hat{K} = K'$, or alternatively $K + \hat{K} = K'$? The point is that $\bullet$ may conflict with the interpretation of $K$ as $P^0_{\mathcal{M}}(K)$, while $+$ may conflict with the interpretation of $\hat{K}$ as $P(\hat{K})$. Brushing aside such fine points, the reader unimpressed by the syntactic reformulation of $\sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset)$ may very well ask: so what? While the strong extensionality of $\sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset)$ does not survive the passage to $\rightarrow_T$, the pay-off in syntacticization is the computational formulation of $\sum_{M \in \mathcal{M}} \mathcal{D}(P_M, \emptyset)$ described by (d), where $\rightsquigarrow_T$ is $\rightarrow_T$ with the $T$-consistency restrictions dropped, but an inconsistency check $\perp$ added. That is, $\rightsquigarrow_T$ is the $(\mathcal{DRS} \cup \{\perp\})$-transition predicate

$$\{(K, \hat{K}, K') \mid K^\cap \hat{K} = K'\} \quad \cup \quad \{(K, \perp, K) \mid K \text{ is } \textit{not } T\text{-consistent}\} \ ,$$

where $T$-consistency is co-r.e., provided $T$ is r.e. and consistency is understood in the sense of first-order logic. In summary, the computationally puzzling connective $\neg$ motivates a syntactic system of non-unique names, under which two names $K$ and $K'$ denote the same object exactly when they are bisimilar relative to a certain r.e. transition predicate.

The reference above to bisimulations may seem a bit of overkill, given that bisimilarity reduces to trace equivalence, since the relevant transition predicates are deterministic. Except, that is, for the transition predicate $P$, which is highly non-deterministic, and for which the notion of a bisimulation, rather than that of a trace, supplies the appropriate equivalence lying behind Theorem 6. While there is no denying that the notion of a trace is simpler than that of a bisimulation, the greater scope of bisimilarity as an extensional notion of equivalence can be said to make it that more interesting.

A final point worth mentioning concerns the relativization of the notion of a transition to a first-order **L**-theory $T$. Beyond providing a bridge between $P_M$ and $P$, the introduction of a background theory $T$ suggests relegating to that background, conditions in a DRS that are "static" in the sense that they can be replaced by first-order **L**-sentences (with no free variables). More precisely, given a DRS $(U, C)$, the static conditions referred to form a set $C' \subseteq C$ such that $P(U, C) = P(\emptyset, C') \circ (U, C - C')$. Transitions from the DRS $(U, C)$ can then be treated as $T$-transitions from the DRS $(U, C - C')$, where $T$ is the theory given by the characteristic formula of the DRS $(\emptyset, C')$. In addition to providing a separation between static and dynamic components, first-order formulas can also be employed to avoid pesky variable clashes (of the kind taken up in §3.2) if we go the whole hog and admit first-order formulas as conditions. (Recall §3.2.) Another way to avoid variable clashes is to keep disjoint sets of variables and parameters, requiring that

(i) in a DRS $(U, C)$, $U$ be a set of variables, whereas the discourse markers in $C$ not declared in $U$ be parameters, and

(ii) embeddings (i.e., variable assignments) be defined only on variables.

Whether this or any of the modifications to DRSs suggested above are worth pursuing are, in the final analysis, probably matters of taste, pure and simple.

# 5 Discussion

First, a brief review of the preceding sections. Inasmuch as a DRS $K$ is a pair $(U, C)$ of finite sets of discourse markers and of conditions, it is natural to build up a DRS from two DRSs $(U, C)$ and $(U', C')$ by merging their sets of discourse markers and conditions respectively, $(U, C) \bullet (U', C') = (U \cup U', C \cup C')$. But an incremental (sequential) interpretation of a DRS $K$ as an input/output relation $P(K) \subseteq S \times S$ suggests analyzing $\bullet$ as relational composition $\circ$ (by which is meant arranging the equation $P(K \bullet K') = P(K) \circ P(K')$) which may run afoul of the normal form of a DRS program $P(K)$ because of variable clashes. This defect is easily corrected by redefining the merge $K^\frown K'$ of two DRSs $K = (U, C)$ and $K' = (U', C')$ by

$$(U, C)^\frown(U', C') \quad := \quad (U \cup U', C[U, U'] \cup C') \ ,$$

where $C[U, U']$ is $C$ with variables occuring in $C$ that belong to $U' - U$ renamed (according to some fixed scheme) so as to be disjoint from $U \cup U'$. Focussing on transitions $P^0(K) = \{(M, f) \mid (M, \emptyset)\ P(K)\ (M, f)\}$ from the empty embedding $\emptyset$, every DRS is induced up to $P^0$ by a quantifier-free first-order formula. Unfortunately, the equivalence $\equiv_0$ modulo $P^0$ is not a congruence with respect to $^\frown$, although the blemish can again be repaired by defining $+$ through a careful choice of representatives of $\equiv_0$-equivalence classes

$$(U, C) + (U', C') \quad := \quad (U \cup U', C[U, U'] \cup C'[U', U]) \ .$$

Rather than revising the simple-minded merge $\bullet$ by renaming variables (yielding $^\frown$ or $+$), variable clashes can be avoided altogether by identifying DRS conditions with arbitrary first-order formulas (instead of closing the set of conditions under negations of DRSs). This modification to the definition of a DRS is a modest proposal, in view of the close correspondence between first-order formulas and DRSs (Lemmas 1 and 3). Under such an alteration or not, the dual nature of DRSs as states and as programs can be brought out by considering transitions between DRSs labelled by DRSs. More precisely, for the unaltered notion of a DRS, define a transition predicate $\rightarrow_T$ relative to a first-order theory $T$ by

$$K \stackrel{\hat{K}}{\rightarrow}_T K' \quad \text{iff} \quad K^\frown\hat{K} = K', \text{ and } K \text{ and } K' \text{ are } T\text{-consistent DRSs} \ . \tag{4}$$

These transitions can be analyzed semantically by interpreting the program $\hat{K}$ as $P(\hat{K})$, and the states $K$ and $K'$ as $P^0(K)$ and $P^0(K')$, respectively (Theorem 10). But as they stand, the transitions in (4) mention only DRSs, and make no reference to variable assignments, so long as consistency is understood syntactically (i.e., as the underivability of a contradiction). As such they are readily mechanizable in a slightly modified form (by $\rightsquigarrow_T$, Theorem 10 (d)), thereby providing an operational realization of dynamic intuitions about meaning. The "operational realization" here is completely analogous to (and, in fact, depends upon) the introduction in first-order logic of a (syntactic) proof theory to reason about Tarski's definition of truth relative to abstract objects (such objects being ill-suited for mechanical representation).

Inspecting line (4), observe that two ideas underly $\rightarrow_T$: a merge operation $^\frown$ and a notion of consistency. Let us conclude by taking up these ideas in turn.

## 5.1 DRSs as abstracts

Inasmuch as abstraction is meaningful only in relation to application, the idea that DRSs are abstracts[7] can be investigated by considering notions of application on DRSs. DRSs aside, notions of application are particularly interesting from a logical point of view in cases where self-application is permitted. Such a possibility is provided by interpreting application $\cdot$ on DRSs according to the merge operation $^\frown$ on DRSs. That is, the application $K \cdot K'$ of two DRSs $K$ and $K'$ can be interpreted as $K^\frown K'$ modulo a consistency restriction that leads to partiality either (i) by allowing (as in $\rightarrow_T$) for the possibility of undefinedness, or (ii) by introducing (as in $\rightsquigarrow_T$) a bottom element $\perp$. In either case, self-application is not on its face paradoxical because of the absence of lambda abstraction (and hence paradoxical combinators). In particular, there is no constant $k$-combinator meeting the requirement $kxy = x$. The fact is that as functions on DRSs, DRSs have a very special form.

[7]This theme was suggested by R. Cooper, and taken up in the DYANA meeting of April 1994.

Apart from the merge $^\frown$, there is also the matter of consistency in characterizing $\rightarrow_T$. If the logic underlying consistency is first-order logic, then the completeness theorem of first-order logic provides a purely syntactic analysis of consistency. In this sense, variable assignments are (pace Barwise [1]) completely dispensable notions (and indeed are, as pointed out in section 4, computationally problematic). This is not to say, of course, that these notions are useless, and indeed consistency might be most naturally studied on the basis of such concepts. But even under a semantic approach, states are more than variable assignments. A DRS applies uniformly to different **L**-models, and it is not so much $f$ that is interesting as it is the $(\mathbf{L} \cup \mathrm{dom}(f))$-model $(M, f)$. So long as this point is kept in mind, the present author sees little to choose in Cooper [2] between (i) viewing a DRS as a situation-theoretic relation $K$, with

$$M \models \langle\!\langle K, f; +\rangle\!\rangle \quad \text{iff} \quad (M, f) \in V(K) \ (= \{s \mid s\ P(K)\ s\}) \ ,$$

and (ii) interpreting a DRS as a situation-theoretic type $V(K)$, obtained by abstracting out $M$ as well as $f$ from the situation-theoretic (parametric) proposition $M \models \langle\!\langle K, f; +\rangle\!\rangle$.

The preceding has been predicated on a first-order notion of consistency. But can such a premiss always be defended? This question takes us to our final subsection.

## 5.2 Higher-order extensions

Although the terms "DRS" and "condition" were initially qualified in §2 as "first-order DRS" and "first-order condition", the designation "first-order" was then promptly dropped. Is there any reason other than laziness lying behind this practice? The clauses syntactically generating DRSs and conditions in §2 are strikingly simple compared to those in Kamp and Reyle [10] and Kamp [12], the latter of which describes seven different formal languages associated with DRT. By contrast, the expressiveness of the DRSs considered here is determined once the background signature **L** is fixed. Surely the matter cannot be left simply at that, can it?

One reason for optimism is the "adequacy" of first-order logic as a framework for formalizing mathematics. The adequacy referred to here concerns the possibility of encoding most (if not all) mathematical notions in set theory, which, in turn, admits a formulation in the first order language $\mathcal{L}(\in)$ with signature $\mathbf{L} = \{\in\}$ given by a single binary relation symbol $\in$. Of course, the simplicity of the "universal" language $\mathcal{L}(\in)$ has hardly led to its universal use in mathematics (universality in principle being one thing, and universality in practice, quite another). One type of complication when working in $\mathcal{L}(\in)$ is easily overcome. Rather than encoding all notions as sets (e.g., presenting real numbers as Cauchy sequences of rational numbers, which are reduced further to sets built from the empty set), the signature can be expanded beyond the binary relation symbol $\in$ (introducing, for example, a unary predicate symbol $\mathbf{R}$, and binary function symbols $+$ and $\times$ to talk about arithmetic over real numbers). A second, more delicate complication has to do with the notion of an arbitrary set (the root of many foundational headaches). The rather heavy ontological commitments of standard set theory (e.g., ZFC) can be weakened by making more modest demands, capturing the sets of immediate interest. Unfortunately, just what sets are of "immediate interest" may not be terribly clear. Nor is it obvious what is gained by describing such sets relative to some logic with mysterious higher-order notions built in. (But how else can we frame a notion of consistency stronger than first-order logic?) At any rate, if we agree to expel such set-theoretic complications from the province of logic, then it is plausible that first-order DRSs constitute the "logical core" of DRT.

To be more concrete, consider the analysis of plurals and generalized quantifiers in chapter 4 of Kamp and Reyle [10]. The abstraction and duplex conditions there are expressible in first-order logic in exactly the same sense that higher-order logic is first-orderized via Henkin generalized models — i.e., by expanding the signature to accomodate a sort for sets. (Observe that a well-defined generalized quantifier ceases to be "definable" only by banning references to sets.) Of course, if we insist on a particular notion of set given, for example, by the full power set of some infinite set of "atoms," then the Löwenheim-Skolem Theorem rules out the existence of a set of first-order sentences (over whatever signature) that is satisfied by precisely such models (and their isomorphic copies). Furthermore, assuming the atoms bear a modest resemblance to the natural numbers, the first-order theories of such models, equipped with an additional sort for *all* sets of atoms, are bound to be unaxiomatizable. But these points are old hat, and the reader is entitled to ask what's new? The basic thrust of the present paper has been to expose the close relationship between first-order logic and a fragment of DRT (which was, moreover, described both in semantic and syntactic terms inherited from first-order logic). Alas, if there is any substance to the move from a "static" (truth-centered) view of semantics (in which first-order logic occupies a singular position) to a "dynamic" one, then surely quite a bit remains to be said.

## References


[1] Jon Barwise. Noun phrases, generalized quantifiers and anaphora. In E. Engdahl and P. Gärdenfors, editors, *Generalized Quantifiers*, Studies in Language and Philosophy. Dordrecht: Reidel, 1987.

[2] Robin Cooper. Situation theoretic discourse representation theory. Manuscript, HCRC, Edinburgh, 1992.

[3] Paul Dekker. Predicate logic with anaphora. In *Proc. Fourth Annual Semantics and Linguistic Theory Conference*. Rochester, New York. May 1994.

[4] Tim Fernando. Transition systems and dynamic semantics. In D. Pearce and G. Wagner, editors, *Logics in AI*, LNCS 633 (subseries LNAI). Springer-Verlag, Berlin, 1992. A slightly corrected version has appeared as CWI Report CS-R9217, June 1992.

[5] Tim Fernando. A persistent notion of truth in dynamic semantics. In *Proc. Fourth Situation Theory and Appplications Conference*. Moraga, California. June 1994.

[6] J. Groenendijk and M. Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14, 1991.

[7] David Harel. Dynamic logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 2. Dordrecht: Reidel, 1984.

[8] Irene Heim. The semantics of definite and indefinite noun phrases. Dissertation, University of Massachusetts, Amherst, 1982.

[9] J.E. Hopcroft and J.D. Ullman. *Introduction to Automata Theory, Language and Computation.* Addison-Wesley, 1979.

[10] H. Kamp and U. Reyle. *From Discourse to Logic.* Kluwer Academic Publishers, Dordrecht, 1993.

[11] Hans Kamp. A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, and M. Stokhof, editors, *Formal Methods in the Study of Language.* Mathematical Centre Tracts 135, Amsterdam, 1981.

[12] Hans Kamp. Discourse Representation Theory as a knowledge representation language. Handout for a talk, 1993.

[13] H. Jerome Keisler. Fundamentals of model theory. In J. Barwise, editor, *Handbook of Mathematical Logic.* North-Holland, Amsterdam, 1977.

[14] Reinhard Muskens. A compositional discourse representation theory. In *Proc. Ninth Amsterdam Colloquium.* ILLC, University of Amsterdam, 1994.

[15] P. Pagin and D. Westersthl. Predicate logic with flexibly binding operators and natural language semantics. *J. Logic, Language and Information*, 2(2), 1993.

[16] David Park. Concurrency and automata on infinite sequences. In P. Deussen, editor, *Proc. 5th GI Conference*, LNCS 104. Springer-Verlag, Berlin, 1981.

[17] C.F.M. Vermeulen. Merging without mystery. Technical report, Logic Group Preprint Series Number 70, Philosophy Department, Utrecht, December 1991.

[18] Henk Zeevat. A compositional approach to discourse representation theory. *Linguistics and Philosophy*, 12, 1989.