Taming first order logic: relating the semantic and the syntactic approach

Maarten Marx

Abstract

Out of the joint work of Johan van Benthem and the Hungarian group round Hajnal Andréka, István Németi and Ildikó Sain and their PhD students, two approaches for taming a logic evolved. With taming a logic we mean changing the logic in such a way that it becomes decidable. For first order logic, they took a semantic route using relativisation of models, and a syntactic route focusing on guarded fragments. The purpose of this paper is to show that these two routes are really two sides of the same coin. We do this by showing that a certain guarded fragment (called here the packed fragment) of first order logic forms precisely the set of first order sentences which are invariant for relativisation with a tolerance relation. Besides this technical contribution we provide an intuitive explanation of relativisation in terms of information transmission.

Contents

Dedicated to Johan van Benthem on his fiftieth birthday.

1 Introduction

The main purpose of this paper is to give a semantic characterisation of the guarded fragment (here called the packed fragment) in terms of invariance for relativisation (Theorem 4.11). We will argue that relativisation in a first order setting is the analogue of taking generated submodels in modal logic. Our tentative conclusion will be that the packed fragment is the true modal fragment of first order logic, because it has the same "local flavour" and is decidable for very similar reasons. Besides this technical result we provide an interpretation of relativisation as a sceptical information processing strategy in section 2. We show that for sentences in the packed fragment, this sceptical strategy leads to the same results (in terms of validity and satisfaction) as the classical first order interpretation. This result makes the sceptical strategy an interesting alternative because it has great computational advantages over the classical first order way of interpreting sentences.

The technical part of this paper is organised as follows. We introduce relativised semantics for first order logic in Section 3. We take a modal view on first order logic and arrive to the notion of admissible assignments. We show how relativised semantics can be obtained from two primitive concepts: context sets and tolerance relations. Then we introduce our version of the (loosely) guarded fragment, called the packed fragment, and relate it to relativised semantics. We conclude by arguing that the packed fragment is the true modal fragment of first order logic.

2 Relativisation interpreted as a sceptical information processing strategy

We employ a very simple model of information transmission: there are two participants of which one does all the talking and the other merely listens and interprets the incoming information. It will be convenient to ascribe gender to the two participants. We assume that the speaker is male and the listener is female. The language used by the speaker is a first order language, which will be interpreted dynamically, in the style of Dynamic Predicate Logic (DPL) of [3, 4].

The update–semantics version of DPL provides a model for the interpretation process performed by the listener

In the dynamic view the interpretation process consists of two components:

1. finding antecedents for anaphora (interpreting discourse information)
2. building a model of the world described by the speaker (interpretation of world–information).

The speaker however is not God or some other embodiment of the world just reporting what is the case, the speaker is an observer of the world reporting how he perceives it. Thus the interpretation of the world–information is more accurately described by

2' building a model of the world perceived by the speaker on the basis of his report of it.

Once the listener realises that she is building a model of the world relative to the perception of the speaker, a number of information processing strategies are open to her. In this paper, we will focus on one such a strategy: the listener does not accept universal statements made by the speaker unconditionally, but relativises them with the proviso

> "provided that the elements quantified over are perceived together by the speaker."

We will make this proviso more precise in due course. First an example

Example 2.1 A company like McDonald's can be modeled in many different ways. A natural way to think about it is as a collection of databases each containing the employees of one outlet, databases containing the managers of a region, databases containing the whatevers of some larger geographical unit, and so on, all the way up to the database containing the president and his companions. In other words, a hierarchical setup of partially overlapping databases.

Let's look at the following situation. The president of McDonald's gives his yearly address and says: "In this company, everybody loves each other." A logician who ows his money baking hamburgers raises his finger and says that he doesn't love his neighbour in the audience at all, since that man works in an outlet at the other side of the world. The president answers, rather annoyed, that he meant of course that "everyone within a unit in McDonald's loves each other".

The president and the logician interpret the first sentence differently because they amalgamated the databases in a different way. The logician used the classical way, while the president amalgamated them in a relativised manner. In this way he could keep the natural structure of the company. For the president, the natural way of interpreting his universal statement was using the proviso given above.

Before we go into the strategy, let's look at some more basic questions. Why would a listener employ such a strategy? And, supposing there are good reasons, how does she know when elements of discourse are perceived together by the speaker? In other words, can she practically perform such a strategy at all?

To start with the first question, why would a listener not want to accept universal statements unconditionally? The answer is that the computational costs connected to universal statements are very high. They can quickly lead to infinite models, moreover it is undecidable to find out whether a set of sentences has a model at all.

Given this it is reasonable to postulate that the listener uses some kind of mechanism to cope with these difficulties of interpretation. What kind of mechanism she uses is of course open to debate. Whatever mechanism she will use, we can expect some natural properties of it:

- It should be sound: "whatever can still be deduced, can also be deduced in the "classical setting".
- It shouldn't lead to too much loss of information: on a large natural fragment of natural language, the classical and the adjusted interpretation process should lead to the same results.
- It should have definite computational advantages.

The strategy we propose satisfies all these. So let's look at the second question: suppose the listener uses the "proviso–strategy", how does she implement it? In particular how does she know which elements in the domain of discourse are perceived together by the speaker, and what does that mean?

We start with the latter. Given a domain D of individuals, and a subset $X \subseteq D$, how can we describe when the elements in X are perceived together by the speaker? One natural way to do this is to postulate that there exists a "distance" function $f : D \times D \longrightarrow \mathbb{R}$, which describes for the speaker the distance between each two elements in D. With distance we mean something inherently vague with many dimensions. It has at least spatial, temporal, conceptual and cultural components, but also cognitive ones, individualised to each speaker. The listener could now state that the elements in X are perceived together by the speaker if for all a, b in X, the distance between a and b is less than some fixed value d. But if the listener cannot know the function f, she also has no idea about d. So what is left for her is just the abstract information that there exists a tolerance relation on D for the speaker. With the assumption that the distance between one object is arbitrarily small, the only thing she knows then is that there exists a binary relation $\delta(x, y)$ on D which is reflexive and symmetric.

So far so good, but how does she *know* which elements stand in this relation? She can't know anything but the discourse of the speaker provides her with clues. That is, from the discourse she gets information which makes it reasonable to assume that indeed the speaker can perceive certain individuals together. The following are examples of such clues:

Named individuals all individuals which the speaker gives a name are perceived together.

Existentially introduced individuals If the speaker introduces two individuals existentially in the same discourse, then their denotation can be perceived together by him.

Primitive relations If the speaker puts certain individuals together in a primitive relation, then he can perceive them together. (We can view this as "naming" a group.)

Related to named individuals Quite a bit stronger is to postulate that the distance between any individual and any named individual is arbitrarily small.

Macho or classical δ is the universal relation on the domain of discourse. That is, the speaker can perceive each two individuals in the domain of discourse together.

The last clue brings us back to the classical interpretation of a discourse. Clearly such clues can be provided by the speaker. We will assume though that he can only do this *outside* the object language that we are studying. So the speaker should make a kind of meta–statement in order to effectuate this information. (Think of a math teacher who starts a class with: "Everything I will say holds for all the natural numbers and for them only".)

This concludes our view on relativisation in a discourse setting. We wil now turn to the technical work. The clues presented above will re-occur there.

3 Relativised semantics for first order logic

We provide first order logic with a different semantics than the standard semantics. We assume we are working with a standard first order language without function symbols: thus the language contains equality, the usual first order connectives, a countable stock of variables and individual constants, and a countable stock of n-ary relation symbols, for every n. In addition we will assume that we have as primitive symbols also $\exists \bar{v}$, where $\bar{v}$ is a *finite* set of variables. When $\bar{v} = \{v_1, \ldots, v_k\}$, then $\exists \bar{v}\varphi$ just means $\exists v_1 \ldots \exists v_k \varphi$ in classical logic. For φ a formula in this language, $FV(\varphi)$ denotes the set of *free variables* of φ, defined in the standard way.

In classical first order logic, the interpretation of a formula in a model (D, I) is given relative to an assignment of the variables s. Given a domain D, the set of assignments consists of all functions from the set of variables into D. The key idea of relativised semantics is that meaning of formulas becomes relativised to a subset of the set of all assignments. We call such a set the *admissible assignments*.

In what follows we will first define this relativised semantics. Then we see what intuitions one can develop about the set of admissible assignments. We finish with providing the connection with standard first order logic and establish a notion of bisimulation.

3.1 Admissible assignments

Assignments. Given a model (D, I), an assignment is a function from the set of variables into D. We assume our language has ω many variables $v_0, v_1, \ldots,$. An assignment g can then be viewed as a sequence from ${}^{\omega}D$: $g(i)$ then gives the value of v_i according to g.

To define the meaning of the existential quantifier, it is handy to create the following relation between assignments:

$$s \equiv_i t \text{ iff for all } j \neq i\text{: } s(j) = t(j). \tag{1}$$

That is: two assignments s and t are $\equiv_i$ related iff they agree on all values of the variables except possibly for v_i. We can also define these relations for sets of variables $\bar{v}$:

$$s \equiv_{\bar{v}} t \text{ iff for all } j \notin \bar{v}\text{: } s(j) = t(j). \tag{2}$$

Using the relation $\equiv_i$, we can give an alternative equivalent definition of the meaning of the existential quantifier. Given a model $\mathfrak{M} = (D, I)$ and an assignment $s \in {}^{\omega}D$, define

$$\mathfrak{M} \models \exists v_i \varphi[s] \iff \text{there exists a } t \equiv_i s \text{ such that } \mathfrak{M} \models \varphi[t].$$

Let us now look at this definition from a modal perspective. We view the assignments as worlds and $\equiv_i$ as an accessibility relation. Then this definition is just the standard modal truth–definition of the "diamond" $\exists v_i$.

Given a first order model (D, I), the set of assignments (worlds) is uniquely determined: it is the set ${}^{\omega}D$. The theory we will develop below abandons this classical rigidness: we will allow other subsets of ${}^{\omega}D$ to be set of "worlds" of our first order models.

Before we can start we have to solve a technical difficulty. First order logic satisfies the following appealing *locality* condition. It says that the meaning of a formula depends only on the model and the variables occurring free in the formula.

Fact 3.1 [Locality] Let φ be a first order formula. Let $\mathfrak{M} = (D, I)$ be a model, and s, t be two assignments such that $s(i) = t(i)$ for all $v_i \in FV(\varphi)$. Then

$$\mathfrak{M} \models \varphi[s] \text{ if and only if } \mathfrak{M} \models \varphi[t].$$

When we give first order logic a relativised semantics, locality does not necessarily hold, cf [1, 6]. We will give meaning to the existential quantifier using the dual of the relation $\equiv_{\bar{v}}$. This relation was introduced into cylindric algebra theory by Y. Venema. On relativised models, this will ensure locality, and on standard models, the meaning of the existential quantifier is just the same. Define for s, t assignments, $\bar{v}$ a set of variables,

$$s \equiv^{\partial}_{\bar{v}} t \text{ iff for all } v_i \in \bar{v}\ s(i) = t(i). \tag{3}$$

Admissible assignments. The key idea of relativised semantics for first order logic is that given a model $\mathfrak{M} = (D, I)$, only a subset of the set of all assignments ${}^{\omega}D$ is available for the interpretation of the formulas. We will now provide the truth definition for a first order language relative to such an admissible set of assignments $V \subseteq {}^{\omega}D$. First we give meaning to terms: let s be an assignment, and $\mathfrak{M} = (D, I)$ a model. We define a function i from the set of terms into D as

$$i(t) = \begin{cases} I(t) & \text{if } t \text{ is a constant} \\ s(i) & \text{if } t \text{ is the variable } v_i \end{cases} \tag{4}$$

Now for $\mathfrak{M} = (D, I)$ a model, and $V \subseteq {}^{\omega}D$ a set of assignments, we define truth of a formula in $\mathfrak{M}$ relative to assignments in V. For $s \in V$,

$$\begin{array}{lcl}
\mathfrak{M} \models_V R(t_1, \ldots, t_n)[s] & \iff & (i(t_1), \ldots, i(t_n)) \in I(R) \\
\mathfrak{M} \models_V t_1 = t_2[s] & \iff & i(t_1) = i(t_2) \\
\mathfrak{M} \models_V \neg\varphi[s] & \iff & \mathfrak{M} \not\models_V \varphi[s] \\
\mathfrak{M} \models_V \varphi \wedge \psi[s] & \iff & \mathfrak{M} \models_V \varphi[s] \text{ and } \mathfrak{M} \models_V \psi[s] \\
\mathfrak{M} \models_V \exists\bar{v}\varphi[s] & \iff & \text{there exists a } t \in V \text{ such that} \\
 & & s \equiv^{\partial}_{FV(\exists\bar{v}\varphi)} t \text{ and } \mathfrak{M} \models_V \varphi[t].
\end{array}$$

Note that the only difference with the definition in any textbook on first order logic is in the clause for the existential quantifier: the assignment t witnessing φ must be admissible, and we use the dual relation $\equiv^\partial$. For comparison, let us define $\models^c_V$ —for "classical $\models$", where all clauses are the same as for $\models_V$ above, except the existential quantifier is defined as

$$\mathfrak{M} \models^c_V \exists \bar{v}\varphi[s] \iff \text{there exists a } t \in V \text{ such that } s \equiv_{\bar{v}} t \text{ and } \mathfrak{M} \models^c_V \varphi[t].$$

The next fact states that the two definitions are equivalent on classical models.

Fact 3.2 Let $\mathfrak{M} = (D, I)$ be a model and let $V = {}^\omega D$. Then for any $s \in V$, for any formula φ,

$$\mathfrak{M} \models_V \varphi[s] \text{ if and only if } \mathfrak{M} \models^c_V \varphi[s].$$

PROOF. The only difference is in the meaning of the existential quantifier. That case goes through by the facts that locality holds on classical models and $s \equiv_{\bar{v}} t$ implies that $s \equiv^\partial_{FV(\exists\bar{v}\varphi)} t$. QED

The next fact states that on relativised models, locality holds as well.

Fact 3.3 Let $\mathfrak{M} = (D, I)$ be a model and $V \subseteq {}^\omega D$ a set of admissible assignments. For any formula φ, for any $s, t \in V$ such that $s \equiv^\partial_{FV(\varphi)} t$ (that is, s and t assign the same values to the free variables in φ):

$$\mathfrak{M} \models_V \varphi[s] \text{ if and only if } \mathfrak{M} \models_V \varphi[t].$$

Truth at non–admissible assignments. Given a model $\mathfrak{M} = (D, I)$ and a set of admissible assignments $V \subseteq {}^\omega D$, we have defined what it means for a formula to be true in $\mathfrak{M}$ at assignments in V. But what about the assignments in ${}^\omega D \setminus V$? The obvious way to do this for $s \in {}^\omega D \setminus V$ is as follows,

$$\mathfrak{M} \models_V \varphi[s] \iff \text{there exists } t \in V \text{ such that } t \equiv^\partial_{FV(\varphi)} s \text{ and } \mathfrak{M} \models_V \varphi[t].$$

Note that for formulas φ, $\mathfrak{M} \models_V \varphi[s]$ can still be undefined. If we assume that V is always non–empty, then for sentences however it is always defined. We now have two ways of defining truth in a model for sentences:

$\mathfrak{M} \models^1_V \varphi$ iff $(\forall s \in {}^\omega D) : \mathfrak{M} \models_V \varphi[s]$
$\mathfrak{M} \models^2_V \varphi$ iff $(\forall s \in V) : \mathfrak{M} \models_V \varphi[s]$.

For sentences, these two definitions give the same result:

Fact 3.4 For every model $\mathfrak{M}$, for every non–empty $V \subseteq {}^\omega D$, for every sentence φ,

$$\mathfrak{M} \models^1_V \varphi \text{ if and only if } \mathfrak{M} \models^2_V \varphi.$$

Since $\models^2_V$ is more economical, we will use this notion from now on and delete the superscript.

Finally we define the notion of validity and of valid consequence. As usual we will overload the meaning of the symbol $\models$. Let Σ be a set of conditions on sets of admissible assignments. Let φ be a sentence and Γ a set of sentences. We define

$\models_\Sigma \varphi$	iff	for every first order model $\mathfrak{M} = (D, I)$,
		for every $V \subseteq {}^\omega D$ satisfying Σ,$\mathfrak{M} \models_V \varphi$.
$\Gamma \models_\Sigma \varphi$	iff	for every first order model $\mathfrak{M} = (D, I)$,
		for every $V \subseteq {}^\omega D$ satisfying Σ, $\mathfrak{M} \models_V \Gamma$ implies $\mathfrak{M} \models_V \varphi$.

3.2 Different relativisations

Above we have defined relativised semantics for any choice of $V \subseteq {}^\omega D$. In the literature several restrictions on V have been proposed.

Here we will show how one can define a set of admissible assignments from a tolerance relation on the domain of the model and from the notion of a context set [9]. Tolerances were introduced in the previous section. For $\mathfrak{M} = (D, I)$ a model, a context is just a subset of D. The intuitive meaning of a context $X \subseteq D$ is

all elements in X can be perceived together by the speaker.

We will now define these notions and investigate their effects.

Let $\mathfrak{M} = (D, I)$ be a first order model. Suppose $f : D \times D \longrightarrow \mathbb{R}_0^+$ is a function associating with every pair of elements in the model a value, which we think of as the distance between the elements. We can then define $\delta(x, y) \iff f(x, y) \leq d$ for some fixed positive d. If we assume that $f(x, x) = 0$, then δ satisfies

$\delta 1$ $(\forall x \in D) : \delta(x, x)$

$\delta 2$ $(\forall xy \in D) : (\delta(x, y) \rightarrow \delta(y, x))$.

We call a relation $\delta \subseteq D \times D$ satisfying these two requirements a *tolerance on* D.

Let C be a collection of subsets of D. Let C satisfy the following two conditions:

$C1$ all singleton sets belong to C

$C2$ C is closed under subsets.

If in addition C satisfies the following *packed dense* condition,

CPD if for all $x, y \in X$, $\{x, y\} \in C$, then also $X \in C$,

we call C a *context set*. Tolerances and context sets are closely related.

Fact 3.5 (i) If δ is a tolerance on D, then the set C defined by

$$X \in C \text{ iff for all } x, y \in X,\ \delta xy,$$

satisfies $C1$, $C2$ and CPD.

(ii) If $C \subseteq \mathcal{P}(D)$ satisfies $C1$, then δ defined by

$$\delta xy \text{ iff there exists a set } X \in C \text{ such that } \{x, y\} \subseteq X,$$

defines a tolerance on D.

(iii) Tolerances and context sets are inter definable by the above definitions.

We now relate tolerances and context sets to sets of admissible assignments.

Definition 3.6 *Let (D, I) be a model, and let δ be a tolerance and C a context set on (D, I). We define two sets of admissible assignments V_δ and V_C as the smallest subsets of ${}^\omega D$ satisfying*

$$\begin{array}{rcl} s \in V_\delta & \textit{iff} & (\forall i, j) : \delta(s(i), s(j)) \\ s \in V_C & \textit{iff} & \{s(i) \mid i \in \omega\} \in C, \end{array}$$

respectively.

The following fact is immediate by Fact 3.5.

Fact 3.7 Let (D, I) be a model, and let δ be a tolerance and C a context set on (D, I). Then $V_\delta = V_C$.

Until now we have given only minimal requirements on the notions of context sets and tolerances. With our intended interpretation it makes sense to make them language–dependent as well (cf., the clues provided in Section 2).

The following extra conditions make sense in a language with constants. ¿From the perspective of a distance function, it says that the distance between any element and a named element is arbitrarily small.

$\delta 3$ $(\forall x \in D) : \delta(x, I(m))$ for all constants m

$C3$ $(\forall x \in D) : \{x, I(m)\} \in C$ for all constants m.

A further restriction on δ (and hence C) is to ask that the distance between two elements which stand together in a primitive relation is arbitrarily small. This would lead to the following extra conditions on δ and C:

$\delta 4$ $(\forall x, y \in D)$: if $(\exists z_1 \dots z_k(x = z_i \wedge y = z_j \wedge (z_1, \dots, z_k) \in I(R)$ for some R, i, j, then $\delta(x, y)$.

$C4$ $(\forall x, y \in D)$: if $(\exists z_1 \dots z_k(x = z_i \wedge y = z_j \wedge (z_1, \dots, z_k) \in I(R)$ for some R, i, j, then $\{x, y\} \in C$.

We now look at the effect of the extra restrictions about constants and primitive relations on the logic. The condition $\delta 3$ does have an effect (of course only in languages with constants in the signature), while $\delta 4$ does not.

Fact 3.8 There exists a sentence which can be falsified on a model with a set of admissible assignments V_δ defined by a tolerance δ, but which holds on all models where V_δ is defined by a tolerance satisfying $\delta 3$.

Proof. Consider the first order tautology

$$[\forall x \exists y Rxy \wedge \forall xyz((Rxy \wedge Ryz) \to Sxz)] \to \exists xy Sxy.$$

The counter model has the natural numbers as its domain, R is interpreted as successor, $S = \emptyset$ and δ is generated by the successor relation (that is $\delta(x, y)$ iff $x = y$ or Rxy or Ryx holds). Clearly δ is a tolerance and the consequent fails on this model. To see that the antecedent holds, observe that for no assignment

$s \in V_\delta$, the range of s contains more than two elements. Thus the second conjunct in the antecedent cannot be falsified.

On the other hand, consider any model $\mathfrak{M}$ where V_δ is defined from a tolerance satisfying $\delta 3$. Assume the antecedent holds in $\mathfrak{M}$. Let a be the element named by some constant a. Then by the first conjunct, there exists a $b \in D$ such that Rab. Let s be the assignment sending every variable to b. Then $\mathfrak{M} \models_V \exists y Rby[s]$, whence there exists a c such that Rbc and $\delta(b,c)$ holds. But by $\delta 3$, also $\delta(a,b)$ and $\delta(b,c)$. But then $\{a,b,c\} \in V_\delta$, whence by the truth of the second conjunct of the antecedent, Sac must hold. QED

For sentences, condition $\delta 4$ does not lead to extra validities.

Fact 3.9 For every sentence φ,

$$\models_{\{\delta 1, \delta 2, \delta 3\}} \varphi \text{ if and only if } \models_{\{\delta 1, \delta 2, \delta 3, \delta 4\}} \varphi.$$

The same holds when we disregard condition $\delta 3$.

Proof. From left to right is obvious. For the other direction, assume $\mathfrak{M} \models_V \varphi$ where V is defined from a tolerance not satisfying $\delta 4$. Change the valuation of the relation symbols such that $\delta 4$ holds as well, by deleting any tuple $\bar{a}$ containing elements a_i, a_j which are *not* δ–related from the interpretation of every relation symbol. Call this model $\mathfrak{M}'$. But then still $\mathfrak{M}' \models_V \varphi$, since to determine the truth of a sentence at an assignment in V_δ one only needs to consider assignments in V_δ. QED

Summing up. We have given several ways of defining relativised semantics. Now it is time to make a choice. In a language without constants this would be easy: we only allow admissible assignments defined from a tolerance δ. Then, just because it is handy, we can ask for condition $\delta 4$ as well, since it does not alter the logic anyway. With constants in the language we should make a decision about $\delta 3$. Since it seems a natural condition and it makes the logic stronger, we have chosen to include that as well. So from now on we only use relativised models where the set of admissible assignments is defined from a tolerance δ satisfying $\delta 3$ and $\delta 4$ (or equivalently, from a context set C which satisfies $C3$ and $C4$). ¿From now on a tolerance means a tolerance satisfying $\delta 3$ and $\delta 4$.

Definition 3.10 *Let $\mathfrak{M} = (D, I)$ be a model, and $\delta \subseteq D \times D$. The relation δ is called a* tolerance *if it satisfies $\delta_1, \delta_2, \delta_3$ and δ_4.*

3.3 Bisimulations

Explicit relativisations. Let $\mathfrak{M} = (D, I)$ be a model and $C \subset \mathcal{P}(D)$ a context set on it. Recall that the intuitive meaning of C was as follows:

$X \in C$ if and only if all elements in X can be perceived together by the speaker.

The language does not have explicit means to state that two elements are perceived together. So we could add constants δ_{ij} for every $i, j \in \omega$, and provide them with the following meaning. For any $s \in {}^{\omega}D$,

$$\mathfrak{M} \models \delta_{ij}[s] \text{ if and only if } \{s(i), s(j)\} \in C. \tag{5}$$

Following Tarski, we call an operation *logical* if it's truth is preserved under automorphisms. Clearly δ_{ij} is not a logical constant. But intuitively it should not be one in this sense. δ_{ij} indicates that the elements denoted by $s(i)$ and $s(j)$ are part of a group of elements which can be perceived together. Arbitrary automorphisms can destroy this intuitive meaning of δ_{ij}. On the other hand, the truth of δ_{ij} is preserved under automorphisms which respect C.

Fact 3.11 Let $\mathfrak{M} = (D, I)$ be a model and $C \subset \mathcal{P}(D)$ a context set on it. Let g be an automorphism of $\mathfrak{M}$ such that for any set $X \subseteq D$, $X \in C$ if and only if $\{g(x) \mid x \in X\} \in C$. Then for any $s \in {}^{\omega}D$,

$$\mathfrak{M} \models \delta_{ij}[s] \text{ if and only if } \mathfrak{M} \models \delta_{ij}[g(s)].$$

Note that δ_{ij} is true on every $s \in V_C$, so on admissible assignments it is equivalent to $\top$. In particular the following equivalence holds.

$$\mathfrak{M} \models_{V_C} \exists\bar{v}\varphi[s] \text{ if and only if } \mathfrak{M} \models_{V_C} \exists\bar{v}(\bigwedge\{\delta_{ij} \mid v_i, v_j \in FV(\varphi)\} \wedge \varphi)[s].$$

What we just did is to make the implicit relativisation to admissible assignments in the meaning definition of the quantifiers explicit in the object language. This provides us with a translation to ordinary first order logic as follows.

Define recursively the following translation function $(\cdot)^\delta$ from first order formulas to first order formulas. $(\cdot)^\delta$ does nothing to atomic formulas, it commutes with the booleans and

$$(\exists\bar{v}\varphi)^\delta = \exists\bar{v}(\bigwedge\{\delta(v_i, v_j) \mid v_i, v_j \in FV(\varphi)\} \wedge \varphi^\delta).$$

Here δ is just a binary predicate. As expected we have,

Fact 3.12 Let $\mathfrak{M} = (D, I)$ be a model and δ a tolerance on it. Then for every formula φ, for all $s \in V_\delta$,

$$\mathfrak{M} \models_{V_\delta} \varphi[s] \text{ if and only if } (D, I, \overline{\delta}) \models \varphi^\delta[s].$$

Here $\overline{\delta}$ is defined as the set

$$\{(x, y) \in D \times D \mid x, y \text{ stand in the tolerance relation } \delta\},$$

and forms the interpretation of the binary predicate δ.

Bisimulations and packed sets. Let (D, I) be a model and δ a tolerance on it. We call a set $X \subseteq D$ *δ–packed* if $\delta(x, y)$ holds for all $x, y \in X$. Then V_δ —the set of admissible assignments defined from δ— is just the set of all sequences whose elements form a δ–packed subset of D. Using this we can define the appropriate notion of bisimulation for this logic. Note that the definition is very close to the one for the guarded fragment in [2].

Two pieces of notation come handy: define for $s \in {}^\omega D$,

$$\mathcal{R}(s) = \{s(i) \mid i \in \omega\}.$$

Also for g a function from D to D', and $s \in {}^\omega D$, define

$$g(s) = \text{ that sequence in } {}^\omega D' \text{ such that for all } i,\ g(s)(i) = g(i).$$

Definition 3.13 (Bisimulation) *Let $\mathfrak{M} = (D, I)$ and $\mathfrak{N} = (D', I')$ be two models for the same signature. Let $\delta_\mathfrak{M}$ and $\delta_\mathfrak{N}$ be tolerances on them respectively. A family F of finite partial isomorphisms between D to D' is called a δ–bisimulation if F satisfies the following conditions:*

- *if $f \in F$ and $g \subseteq f$, then also $g \in F$*
- *(totality)*
  - *for every δ–packed set $X \subseteq D$, there exists an $f \in F$ whose domain is X*
  - *similar for δ–packed subsets of D'*
- *(forth) if $f \in F$ and $dom(f) \subseteq X$ for some δ–packed set X, then there exists a $g \in F$ which extends f and whose domain is X*
- *(back) a similar condition in the other direction.*

Note that bisimulations are always non–empty, by totality and the fact that every singleton set is δ–packed. Of course we have the following

Fact 3.14 For every φ, for every $\mathfrak{M}, \mathfrak{N}$, for every δ–bisimulation F between them, for every $f \in F$, and for every assignment s such that $\mathcal{R}(s) = dom(f)$,

$$\mathfrak{M} \models_V \varphi[s] \text{ if and only if } \mathfrak{N} \models_V \varphi[f(s)].$$

Proof. The proof is by induction on formulas. We only consider the case for the existential quantifier. So let $\mathfrak{M} \models_V \exists\bar{v}\varphi[s]$ and $f \in F$ with $dom(f) = \mathcal{R}(s)$. The case when $\exists\bar{v}\varphi$ is a sentence is easy and left to the reader (use totality). So suppose otherwise. Then there exists a $t \in V_\delta$ such that $t \equiv^\partial_{FV(\exists\bar{v}\varphi)} s$ and $\mathfrak{M} \models_V \varphi[t]$. Let s' be such that $\mathcal{R}(s') = \{s(i) \mid v_i \in FV(\exists\bar{v}\varphi)\}$. Then by locality also $\mathfrak{M} \models_V \exists\bar{v}\varphi[s']$. Since F is closed under subsets, also $f_{\restriction\mathcal{R}(s')} \in F$. From $t \equiv^\partial_{FV(\exists\bar{v}\varphi)} s$ it follows that $\mathcal{R}(s') \subseteq \mathcal{R}(t)$. Whence by forth, there exists a $g \supseteq f_{\restriction\mathcal{R}(s')}$ with $dom(g) = \mathcal{R}(t)$. Thus by induction hypothesis, $\mathfrak{N} \models_V \varphi[g(t)]$. But $g \supseteq f_{\restriction s'}$ implies $f_{\restriction\mathcal{R}(s')}(s') \equiv^\partial_{FV(\exists\bar{v}\varphi)} g(t)$. Thus $\mathfrak{N} \models_V \exists\bar{v}\varphi[f_{\restriction\mathcal{R}(s')}(s')]$, whence by locality $\mathfrak{N} \models_V \exists\bar{v}\varphi[f(s)]$. QED

4 Packed fragment

In this section, we look for sentences whose truth in a model is unaffected by adding or deleting a tolerance. The syntactic characterisation of this fragment forms a slight generalisation of van Benthem's loosely guarded fragment. We first define the fragment. We work in a standard first order language with equality with one restriction: terms are variables or constant symbols.

We say that a formula φ packs a set of variables $\{x_1, \ldots, x_k\}$ if φ is a conjunction of formulas of the form $t_i = t_j$ or $R(t_1, \ldots, t_n)$ or $\exists \bar{y} R(t_1, \ldots, t_n)$ such that for every $x_i \neq x_j$, there is a conjunct in φ in which x_i and x_j both occur free.

In the definition of the packed fragment we use generalised quantifiers $\forall \bar{x}(\varphi, \psi)$ where $\bar{x} = x_1, x_2, \ldots, x_n$ is a sequence of variables. The meaning of this quantifier is nothing but the meaning of $\forall x_1 \ldots \forall x_n (\varphi \to \psi)$ in first order logic.

A generalised quantifier $\forall \bar{x}(\varphi, \psi)$ is called *packed* if φ packs all free variables of ψ. We call φ the *guard* of $\forall \bar{x}(\varphi, \psi)$. Note that if ψ contains only one free variable, then the first argument of the universal quantifier can be anything: packedness only speaks about *pairs* of variables.

The *packed fragment* is defined as follows: a packed formula is constructed from atoms using the booleans and packed universal quantification $\forall \bar{x}(\varphi, \psi)$, where ψ must be a packed formula.

It will be useful to define two more fragments. A packed existential quantification is nothing but $\neg \forall \bar{v}(\varphi, \neg \psi)$, where $\forall \bar{v}(\varphi, \neg \psi)$ is a packed universal quantification (i.e., it is of the form $\exists \bar{v}(\varphi \wedge \psi)$, where φ packs all free variables of ψ).

The $\forall$–packed fragment is defined as follows: formulas are constructed from atoms and their negations using $\wedge, \vee, \exists$ and packed universal quantification $\forall \bar{x}(\varphi, \psi)$, where ψ must be a $\forall$–packed formula. The $\exists$–packed fragment is defined dually: so we may use unpacked $\forall$, but only packed $\exists$.

The three fragments are of course closely related

Fact 4.1 A first order sentence φ is equivalent to a packed sentence if and only if it is equivalent to a $\forall$– and a $\exists$–packed sentence.

We will now related the packed fragment to relativised semantics. One direction is obvious, and observed in [8].

Fact 4.2 The translation $(\cdot)^{\delta}$ goes to the packed fragment.

Just as all first order sentences are invariant for δ–bisimulations when they are interpreted relativised to a set of admissible assignments, all packed sentences are invariant for δ–bisimulations when they are classically interpreted.

Definition 4.3 *A sentence φ is invariant for δ–bisimulations if for all models $\mathfrak{M}$, $\mathfrak{N}$, for all tolerances $\delta_{\mathfrak{M}}, \delta_{\mathfrak{N}}$, and for all δ–bisimulations F: $\mathfrak{M} \simeq_F \mathfrak{N}$,*

$$\mathfrak{M} \models \varphi \textit{ if and only if } \mathfrak{N} \models \varphi.$$

Fact 4.4 All packed sentences are invariant for δ–bisimulations.

PROOF. Let $\mathfrak{M}, \mathfrak{N}$ be models, $\delta_{\mathfrak{M}}, \delta_{\mathfrak{N}}$ tolerances and F: $\mathfrak{M} \simeq_F \mathfrak{N}$ a δ–bisimulation. We show by induction that for every packed formula φ, for every $f \in F$, and for every assignment s such that $\mathcal{R}(s) = dom(f)$,

$$\mathfrak{M} \models \varphi[s] \text{ if and only if } \mathfrak{N} \models \varphi[f(s)].$$

Then the result follows by the non–emptyness of δ–bisimulations.

The inductive proof goes through for the atomic cases because the functions in F are partial isomorphisms. The boolean cases are trivial. For packed existential quantification we reason as follows. Let us call an assignment s *small for* φ if $(\forall x \notin FV(\varphi))(\exists y \in FV(\varphi)) : s(x) = s(y)$. Suppose $\mathfrak{M} \models \exists \bar{x}(\varphi, \psi)[s]$. By locality we may assume that s is small for $\exists \bar{x}(\varphi, \psi)$. Then there exists a $t \equiv_{\bar{x}} s$ such that $\mathfrak{M} \models \varphi \wedge \psi[t]$. Again we may assume that t is small for $\varphi \wedge \psi$. Since φ packs all free variables of ψ, the set $\mathcal{R}(t)$ is packed by $\delta_{\mathfrak{M}}$. Moreover $\mathcal{R}(s) \subseteq \mathcal{R}(t)$ because both are small. But then by *forth*, there exists a $g \in F$ extending f whose domain is $\mathcal{R}(t)$. Thus by inductive hypothesis $\mathfrak{N} \models \varphi \wedge \psi[g(t)]$, and finally $\mathfrak{N} \models \exists \bar{x}(\varphi, \psi)[f(s)]$, since g extends f. QED

Facts 3.12 and 4.2 show that on any model $\mathfrak{M}$, and any relativisation defined by a tolerance δ, the question whether $\mathfrak{M} \models_V \varphi[s]$ is equivalent to the question whether the packed formula φ^δ is classically satisfied in $\mathfrak{M}$ at s. The other direction was shown in [5], to give an alternative proof of the decidability of the loosely guarded fragment. The following fact is a generalisation of that result. It reduces the question of first order satisfiability of $\forall$–packed formulas to that of relativised satisfiability. We present the fact and its proof here since the construction used in the proof is typical for the packed fragment, and will later be used to give a semantic characterisation of it.

Fact 4.5 Every $\forall$–packed formula is classically satisfiable iff it is satisfiable on a relativised model.

PROOF. ¿From left to right is obvious, just take the tolerance to be the universal relation. For the other direction, let $\mathfrak{M} = (D, I)$ be a model, δ a tolerance on D and suppose $\mathfrak{M} \models_{V_\delta} \varphi$. We prove by induction for every $s \in V_\delta$, for every formula ψ,

$$(*) \quad \mathfrak{M} \models_{V_\delta} \psi[s] \Rightarrow \mathfrak{M} \models \psi[s],$$

from which the result follows.
For literals and formulas of the form $\exists \bar{x}(R(t_1, \ldots t_k))$ $(*)$ holds in both directions. For $\wedge, \vee$ and $\exists$, the proof is trivial.

For the universally quantified formulas, suppose $\mathfrak{M} \not\models \forall \bar{v}(\varphi, \psi)[s]$. Then there exists a $t \equiv_{\bar{v}} s$ such that $\mathfrak{M} \models \varphi[t]$ but $\mathfrak{M} \not\models \psi[t]$. By locality we may assume that $\mathcal{R}(t) = \{t(i) \mid v_i \in FV(\varphi \to \psi)\}$. Since φ packs all free variables of ψ, $\mathcal{R}(t)$ is δ–packed (since we assume condition $\delta 4$), so $t \in V_\delta$. Then $\mathfrak{M} \models_{V_\delta} \varphi[t]$, since $(*)$ holds in both directions for all conjuncts of φ. An application of the induction hypothesis leads to $\mathfrak{M} \not\models_{V_\delta} \psi[t]$. Whence $\mathfrak{M} \not\models_{V_\delta} \forall \bar{v}(\varphi, \psi)[s]$, because $s \equiv^{\partial}_{FV(\forall \bar{v}(\varphi, \psi))} t$ holds as a consequence of $s \equiv_{\bar{v}} t$. QED

Because he satisfaction problem for first order formulas on relativised models is decidable [5], we obtain

Corollary 4.6 *The satisfaction problem for $\forall$–packed formulas is decidable.*

The reader familiar with modal logic might recognise a basic construction from modal logic. If we view the elements of V_δ as worlds, then essentially we added worlds to the model. The reverse action would correspond roughly to the generated submodel construction from modal logic. Reading the last proof dually, it shows that $\exists$–packed sentences are preserved under adding a tolerance and interpreting sentences in a relativised manner. Formally,

Definition 4.7 *Let $\mathfrak{M} = (D, I)$ be a model, and δ any tolerance on D. We say that a first order sentence φ is preserved under δ–relativisation if $\mathfrak{M} \models \varphi$ in the classical sense only if $\mathfrak{M} \models_{V_\delta} \varphi$. A sentence φ is* invariant *for δ–relativisation if $\mathfrak{M} \models \varphi$ in the classical sense if and only if $\mathfrak{M} \models_{V_\delta} \varphi$.*

The last proof showed the following preservation result.

Fact 4.8 The $\exists$–packed fragment is preserved under δ–relativisation.

The converse of this preservation result does not hold. In fact this is to be desired since

Fact 4.9 The validity problem for the first order fragment preserved under δ–relativisation is undecidable.

Proof. Consider arrow logic interpreted on relativised pair–frames, and expand it with a coordinate wise difference operator D_1, with the following meaning, interpreted on a model with relativised domain V:

$$\mathfrak{M} \models D_1\varphi(x, y) \text{ iff there exists a } (x, z) \in V \text{ and } z \neq y \text{ and } \mathfrak{M} \models \varphi(x, z).$$

This logic is introduced in [5] and shown to be undecidable. It is easy to check that the range of the translation of this logic to first order logic is preserved under δ–relativisation. QED

Corollary 4.10 *The $\exists$–packed fragment does not capture all first order sentences preserved under δ–relativisation.*

Proof. An $\exists$–packed sentence is valid iff its negation (which is equivalent to a $\forall$–packed sentence is not satisfiable. By Fact 4.5 and the fact that the satisfaction problem for relativised semantics is decidable, the validity problem for $\exists$–packed sentences is decidable. This contradicts Fact 4.9. QED

We come to the central result of this paper: a semantic characterisation of the packed fragment in terms of relativised semantics.

Theorem 4.11 *Let φ be a first order sentence. The following are equivalent.*
(i) φ is equivalent to a packed sentence.
(ii) φ is invariant for δ–bisimulation.
(iii) φ is invariant for δ–relativisation.

Proof. From (i) to (ii) was shown in Fact 4.4. From (ii) to (iii) is immediate since every model δ–bisimulates with itself for every tolerance defined on it. For the proof of (iii) implies (i), we use a "diagram–chasing" argument well–known from van Benthem's work (see e.g., [7, 8]).

Let φ be a sentence which is invariant under δ–relativisation. Define $PF(\varphi)$ as the set of all packed sentences (in the same signature as φ) which classically follow from φ. We will show that $PF(\varphi) \models \varphi$, from which the result follows by compactness. If $PF(\varphi)$ is inconsistent, there is nothing to prove. So let $\mathfrak{M} = (M, I)$ be a model for $PF(\varphi)$, in the signature of φ. We will show that $\mathfrak{M} \models \varphi$. Consider the complete packed theory $PF(\mathfrak{M})$ of $\mathfrak{M}$, together with φ. This set of sentences is finitely satisfiable by a simple argument. Thus by compactness it has some model $\mathfrak{N} = (N, I')$. Again we can equate the signature of $\mathfrak{N}$ with the signature of φ.

Now take ω–saturated extensions $\mathfrak{M}^+$ and $\mathfrak{N}^+$, respectively. Let $\delta_{\mathfrak{M}}$ ($\delta_{\mathfrak{N}}$) be the smallest tolerance which can be defined on $\mathfrak{M}^+$ ($\mathfrak{N}^+$). Since $\mathfrak{M}$ and $\mathfrak{N}$ have the same finite signature, on the extensions δ can be defined as a finite disjunction of terms of the form

$$t_i = t_j, \;\; R(t_1, \ldots, t_n), \text{ and } \exists \bar{v} R(\bar{v}, v_i, v_j). \tag{6}$$

Since packed formulas are invariant under δ–relativisation, and φ as well, we have

$$\mathfrak{M}^+ \models_{V_{\delta_{\mathfrak{M}}}} PF(\varphi) \text{ and } \mathfrak{N}^+ \models_{V_{\delta_{\mathfrak{N}}}} PF(\mathfrak{M}) \cup \{\varphi\}.$$

We call finite δ–packed sets $X \subseteq M^+$, $Y \subseteq N^+$ *δ–relativised equivalent* if there exists a bijection $f : X \longrightarrow Y$, and for some $s \in {}^{\omega}M^+$, such that $\mathcal{R}(s) = X$, we have

$$\text{for all formulas } \varphi, \, \mathfrak{M}^+ \models_{V_{\delta_{\mathfrak{M}}}} \varphi[s] \iff \mathfrak{N}^+ \models_{V_{\delta_{\mathfrak{N}}}} \varphi[f(s)]. \tag{7}$$

Claim 1 The relation of δ–relativised equivalence is a δ–bisimulation between the models $\mathfrak{M}^+$ and $\mathfrak{N}^+$, with tolerances $\delta_{\mathfrak{M}}$ and $\delta_{\mathfrak{N}}$, respectively.

Proof of Claim. Obviously, the relation is closed under subsets. The proof that it is total needs some extra argumentation. First observe that by Fact 3.12, (7) is equivalent to

$$\text{for all formulas } \varphi, \, \mathfrak{M}^+ \models \varphi^{\delta}[s] \iff \mathfrak{N} \models^+ \varphi^{\delta}[f(s)], \tag{8}$$

where $(\cdot)^{\delta}$ is the translation function defined just above Fact 3.12. But since δ is the smallest tolerance, for every two variables v_i, v_j, $\delta(v_i, v_j)$ is equivalent with a disjunction of formulas of the form given in (6). But then, using distributivity of $\vee$ over $\wedge$ and of $\vee$ over $\exists\bar{v}$, every formula φ^δ can be equivalently written as a packed formula in the φ-signature. Whence (8) is equivalent to

$$\text{for all packed formulas } \varphi,\ \mathfrak{M}^+ \models \varphi[s] \iff \mathfrak{N}^+ \models \varphi[f(s)]. \tag{9}$$

But now the standard argument can be applied. Let $X \subseteq M^+$ be some finite δ–packed set, with $|X| = n$. Let $s \in {}^\omega M^+$ be such that $\mathcal{R}(s) = X$. Let $PF(s)$ be the set of packed formulas true at s in $\mathfrak{M}^+$. Then for every finite subset Ψ of $PF(s)$, $\mathfrak{M}^+ \models \exists v_1 \dots v_n(\varphi, \bigwedge \Psi)$, where φ describes how X is packed. Note that this is a packed sentence. So it also holds in $\mathfrak{M}$, whence also in $\mathfrak{N}+$. But then, by ω–saturation, $PF(s)$ is satisfied in $\mathfrak{N}^+$ at some t. Then $\mathcal{R}(t)$ is the required witness.

The other direction of totality uses a symmetric argument. The back and forth clauses are now proved similarly, again exploiting the fact that the models are saturated. ◀

Now we are almost finished. Since $\mathfrak{N}^+ \models_{V_{\delta_{\mathfrak{N}}}} \varphi$, by Fact 3.14, $\mathfrak{M}^+ \models_{V_{\delta_{\mathfrak{M}}}} \varphi$, whence since φ is invariant for δ–relativisation, also $\mathfrak{M}^+ \models \varphi$. Whence, since $\mathfrak{M}$ is an elementary submodel of $\mathfrak{M}$, $\mathfrak{M} \models \varphi$, as required. QED

5 Conclusion

Theorem 4.11 provides us with a semantic characterisation of the packed fragment. Moreover it indicates a strong connection between the packed fragment and relativised semantics. The packed fragment forms precisely the set of first order sentences for which it does not matter whether they are interpreted classically or relativised on a model.

Another perspective on the packed fragment is obtained by analogy with modal logic. In describing the difference between first order logic and modal logic, modal logicians often first come up with the fact that truth is determined locally in modal logic. What is meant is that truth of a modal formula at some state s only depends on the states which are finitely accessible from s. In other words, truth at s is invariant for adding or deleting states which are not accessible from s. Put differently, modal formulas are invariant under generated submodels.

Relativisation implements this modal "local evaluation" in first order logic by relativising the meaning of the quantifiers to admissible assignments. In modal terminology: in a model (D, I), every assignment $s \in {}^\omega D$ is accessible to every other assignment $t \in {}^\omega D$; relativisation restricts the number of accessible assignments to the admissible ones, and thus localises first order logic. ¿From this perspective, Theorem 4.11 tells us that the packed fragment is the "local" fragment of first order logic. Whence, if we are willing to equate local with modal, it states that the packed fragment is the true modal fragment of first order logic.

References

[1] N. Alechina. *Modal quantifiers*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 1995. ILLC Dissertation Series 1995–20.

[2] H. Andréka, J. van Benthem, and I. Németi. Modal languages and bounded fragments of predicate logic. *Journal of Philosophical Logic*, 27(3):217–274, 1998.

[3] J. Groenendijk and M. Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14:39–100, 1991.

[4] J. Groenendijk, M. Stokhof, and F. Veltman. Coreference and modality. In S. Lappin, editor, *The Handbook of Comtempory Semantic Theory*, pages 179–213. Blackwell, 1996.

[5] M. Marx. Complexity of modal logics of relations. Technical Report ML–97–02, Institute for Logic, Language and Computation, University of Amsterdam, May 1997.

[6] I. Németi. A fine-structure analysis of first-order logic. In M. Marx, L. Pólos, and M. Masuch, editors, *Arrow Logic and Multimodal Logics*, Studies in Logic, Language and Information, pages 221–247. CSLI Publications, Stanford, 1995.

[7] J. van Benthem. *Modal Logic and Classical Logic*. Bibliopolis, Naples, 1983.

[8] J. van Benthem. *Exploring Logical Dynamics*. Studies in Logic, Language and Information. CSLI Publications, Stanford, 1996.

[9] D. Westerstahl. Determiners and context sets. In J. van Benthem and A. ter Meulen, editors, *Generalized Quantifiers and Natural Language*, pages 45–71. Foris, Dordrecht, 1985.