WHAT DOES ONE NEED, WHEN SHE NEEDS "HIGHER-ORDER LOGIC"?

Jaroslav Peregrin[1]
www.cuni.cz/~peregrin
[*Proceedings of LOGICA'96*, FILOSOFIA, Praha, 1997, 75-92.]

1. First-Order Predicate Calculus

Formal languages, which are the medium of modern formal logic, have reached their definitive form during the first part of this century. The most substantial of them, the language of *predicate calculus*, is characterized by three types of syntactic rules. Two of these rules reflect, quite straightforwardly, basic syntactic structures of natural language: (i) the fact that an elementary natural language sentence typically consists of a verbal phrase complemented by several nominal phrases, and (ii) the fact, that sentences can be negated and joined into more complex sentences by means of certain connectives. Hence the rules

(i): An n-ary predicate plus n terms yield an (elementary) statement.
(ii) An n-ary logical operator plus n statements yield a statement.

The third kind of rule is of a different sort: it reflects no such general syntactic structure of natural language, but rather the structure of certain specific judgments which we can make *about language*. If we turn an expression into a „matrix" (a general scheme) by substituting *variables* - in the role of gap-markers - for some of its parts, we can then imagine the gaps filled with various concrete things (and the matrix thus being turned back into a statement) and scrutinize in which cases the resulting statement holds, and in which it does not not. On the basis of such observations we formulate judgments such as „the matrix M yields a true statement independently of whatever we fill its gaps" and „the gaps in M can be filled by something to yield a true statement"; or simply „for every x, N" and „for some x, M"[2]. This leads to an intuition (which is, in contrast to the previous two, very specific), that a statement might consist of a quantifier (stating whether we are talking about all possibilities, about the existence of at least one possibility, or, as the case may be, about some other pattern of existence of possibilities), a variable (stating which kind of gap the assertion is about), and a matrix:

(iii) A quantifier plus a variable plus a statement yield a statement (where we suppose that a variable is also a term and hence can enter the rule (i)).

This last intuition, however, has not always been understood and explicated quite uniformly: differences have occurred especially in whether variables can legitimately replace only terms, or also other kinds of expressions, particularly predicates. This ambiguity lead to the constitution

---

[1]Thanks to V. Švejdar, P. Kolář a P. Materna for inspiring comments on a previous draft of the paper.

[2]This view of quantification is discussed in Peregrin (1995).

of first-order predicate logic as a self-contained subsystem of general predicate logic[3]. The first-order paradigm means not only that we have no other variables than those substitutable for terms, but also that we cannot have a mechanism which would allow for a higher-order quantification in an indirect way.[4] Due to all the nice properties of this subsystem (completeness, compactness, Löwenheim-Skolem property), many logicians are convinced, that logic should concentrate on this system alone.

Besides the exponents of first-order logic, however, there are also many logicians regarding this as too restrictive. Such voices are heard not only from the camp of those involving themselves with the analysis of natural language (and thus requiring a far richer repertoire of syntactic means than can be offered by the first-order predicate calculus), but also from among those who are concerned with the foundations of mathematics (Barwise a Feferman, 1985; Shapiro, 1991). Unfortunately it seems that many discussions around this theme suffer both from the fact that their participants sometimes fail to acknowledge all the facts about the relationship between first-order and higher-order logic, and also from the fact that terms like *higher-order logic* are employed in various different senses.

The aim of this contribution is to clarify what is, and what is not, a legitimate subject to such discussions: it thus presents nothing essentially new, but it assembles some available relevant facts in a way which the author finds useful and which he has found lacking in the current literature.

2. Beyond the border of first-order logic

When you ask someone why she needs a logic of an order higher than one, why she cannot make do with first-order logic, she is likely to respond that she needs a *higher expressive power*, than first order-logic can offer. However, this may be ambiguous. It may be interpreted as expressing (i) the need of a syntactically richer language; which can, in turn, represent either (i.i) the simple need of higher-order predicates; or (i.ii) the need to quantify over the corresponding variables. However, the response might also be interpreted amounting to (ii) the need of a logic which would allow us to articulate some of the concepts which are not expressible within first-order logic, e.g. the concept of finiteness. These motifs, however interconnected they may be, cannot be identified. Let us discuss them in greater detail.

The need (i.i) may be motivated by an effort to apply a straightforward logical analysis to natural language statements like (1) or (2); the difficulty being that in order to form the formula (1'), we need the second-order predicate **La**, and in order to form (2'), we need the ‘predicate’ **Qu**, whose application to a predicate yields again a predicate.

To be brave is laudable (1)

[3]See Moore (1988).

[4]Such mechanism might be e.g. a rule of comprehension, which guaranteed that each predicate (or its denotation) possesses its unique ‘objectual correlate’: then we could emulate the quantification over predicates via quantification over their correlates. It was precisely this kind of mechanism which made the first systematic articulation of formal logic, presented by Frege, into more than first-order logic. (In fact, it was also one of the reasons why Frege’s system was *contradictory*, as pointed out by Russell.).

Charles runs quickly (2)

La(**Br**) (1')

(**Qu**(**Ru**))(**Ch**) (2')

Similarly we may want to capture some mathematical concepts by means of second-order predicates; we may, e.g., want to regiment (3) as (3'), where **Pr** is a unary predicate.

There are infinitely many primes (3)

Inf(**Pr**) (3')

The need (i.ii) may be motivated by the effort of capturing the statements like (4) and (5) as something like (4') and (5'):

Charles and Peter share a property (4)

To do something quickly is not to do it slowly (5)

$\exists p.p(\mathbf{Ch})\&p(\mathbf{Pe})$ (4')

$\forall p\forall x.(\mathbf{Qu}(p))(x)\rightarrow\neg(\mathbf{Sl}(p))(x)$ (5')

Similarly, the Dedekind definition of an infinite range, (6), can be straightforwardly regimented as (6'), while the axiom of induction, (7), as (7').

The range P is infinite, if it can be injectively mapped on its proper subset. (6)

$\mathbf{Inf}(\mathrm{P})\equiv\exists f.\forall x\forall y(fx{=}fy\rightarrow x{=}y)\&\forall x(\mathrm{P}(x)\rightarrow\mathrm{P}(fx))\&\exists y(\mathrm{P}(y)\&\forall x(\mathrm{P}(x)\rightarrow fx\neq y))$ (6')

If 0 has a property, and if, moreover, a successor n' of every number n has the property if n has the property, then every number has the property. (7)

$\forall p.(p(\mathbf{0})\&\forall n(p(n)\rightarrow p(n')))\rightarrow\forall n.p(n)$ (7')

In contradistinction, the need (ii) then can be motivated simply by the desire for a logic in which it would be possible to express concepts which are arguably not expressible within first-order logic (*infinity*, for instance), but not necessarily by means of an explicit definition in the object language (such as (6')). If we accept the model-theoretic notion of logic (viz. Barwise & Feferman, 1985), we can legitimately define, for example, the quantifier $\exists_\infty$ by the following metalinguistic prescription (where $\|...\|_{[x\leftarrow d]}$ is that interpretation which differs from $\|...\|$ at most in that $\|x\|_{[x\leftarrow d]}=d$):

$$\|\exists_\infty x\mathrm{P}(x)\| = 1 \text{ iff there is an infinite number of such objects d that } \|\mathrm{P}(x)\|_{[x\leftarrow d]} = 1$$

Such a definition does not necessarily carry us over the borders of the syntax of first-order logic - $\exists_\infty$ is an expression of the very same syntactic category as $\exists$ and $\forall$, and to introduce new expressions of this category does not cause any paramount problems for the first-order framework (*viz.* $\exists!$).

This indicates that (i) and (ii), however interconnected, differ substantially in character. In the case of (i) the requirement is a larger repertoire of syntactic means, which alone need not involve any nontrivial step beyond the boundaries of first-order logic. For there is a strategy for both directly "emulating" such means within the framework of first-order logic, or alternatively for extending the language of first-order logic so as to bring us the needed syntactic means without crossing the boundaries of first-order logic. Let me now sketch two variants of the strategy for reducing higher-order quantification to first-order one.

3. Predicates as individuals

The first of the variants stems from the conviction that a subject of predication is always necessarily an individual. Frege (1892, s.197) says: "[der Begriff] kann wegen seiner prädikativen Natur nicht ohne weiteres [als Bedeutung des grammatischen Subjekts] erscheinen, sondern muβ erst in einen Gegenstand verwandelt werden, oder, genauer gesprochen, er muβ durch einen Gegenstand vertreten werden." This means that what we may see as a property of properties applied to a property should be viewed as a property of individuals applied not to a property, but rather to some kind of its "objectual correlate" (its extension, within Frege's framework). In natural language the situation is indeed such that a predicate can be conjoined only with a nominal form (*nominalization*) of another predicate (typically a verbal noun, an infinitive, or a gerund). This consideration leads to the situation that, for example, the statement (1) is regimented as the application of the predicate **La** to the term **BBr**, which denotes the „objectual correlate“ of the predicate **Br**.

Obviously, however, the systematic relationship between predicates and their objectual correlates is logically significant: the inference such as (8) is generally valid.

Charles is brave (8)
To be brave is laudable

---

hence *Charles has a laudable property*

Such inferences are nevertheless easy to render: it is only necessary to take the predicate *to have a property* seriously (to regiment it as a binary predicate constant), and further to understand **BBr** not as a primitive term, but rather as a „nominalizing“ operator **B** applied to the predicate **Br** - i.e. to understand **BBr** as **B**(**Br**). Then we can articulate the general inference rule

P(*T*) (9)

---

hence **HasPr**(*T*,**B**(*P*));

and regiment the inference (8) with its help:

Br(**Ch**) (8')
La(**B**(**Br**))

---

hence **HasPr**(**Ch**,**B**(**Br**))**&La**(**B**(**Br**))
and hence ∃*x*.**HasPr**(**Ch**,*x*)**&La**(*x*)

An operator such as **B** cannot, of course, be squeezed directly into the first-order framework; but nevertheless its introduction requisites a modification of the framework other than the introduction of higher-order predicates. Problems connected with operators of this kind, and, more generally, problems concomitant with logical analysis of the phenomenon of nominalization in natural language are discussed in detail by Chierchia (1982) and Turner (1983).[5]

A variation on the same theme is the Davidsonian approach to the regimentation of sentences of type (2): Davidson (1980) proposes to enrich each predicate by a new, in natural language covert, argument place, fillable with something like „events“: the statement (2) will thus be understood as *There is an "event of running" the agent of which is Charles and this event is quick* (see also Parsons, 1990).

$$\exists e.\mathbf{Ru}(e,\mathbf{Ch})\&\mathbf{Qu}(e). \qquad (2'')$$

In a certain sense, both model theory and set theory can be seen as general expressions of this strategy: a model-theoretic interpretation of a formal language can be seen as a sort of a translation of this language into the language of set theory - hence into a first-order language.[6] For the usual model theory can be seen as - in effect - a means of translating, e.g., the statement $P(T)$ into the "metastatement" $\|T\|\in\|P\|$ (similarly, with some insubstantial complications, for predicates of higher arities); and hence the reduction of the truth of the former to the truth of the latter - if we understand $\|...\|$ simply as a nominalizing device (so that $\|T\|=T$, for *T is* nominal in itself, and $\|P\|=\mathbf{B}(P)$), and if we write **HasPr** instead of $\in$, we turn $\|T\|\in\|P\|$ into our familiar $\mathbf{HasPr}(T,\mathbf{B}(P))$.[7]

4. Henkinian Understanding of Higher-Order Logics

The second variant of this strategy is based on the idea of accepting the syntactic means of higher-order logics without any limitations, but semantically interpreting them in the spirit of first-order logic; i.e. taking them as mere ‘notational variants’ of first-order means. If we concentrate on second-order logic, this strategy amounts, informally speaking, to taking relations as a peculiar kind of individuals (the relations will hence be elements of the domain of individuals). This causes the expression $P(T)$ to be seen as expressing a relation between two individuals: the relation-taken-as-individual $\|P\|$ and the ("classical") individual $\|T\|$. In this way, quantification over relations becomes quantification over a certain kind of individuals.

The interpretation of a second-order language consists of a universe U and an interpretation function which maps individual constants on the elements of U and predicate constants on the relations over U; the range of the individual variables is then U and the ranges of predicate variables are the corresponding sets of relations (subsets of Cartesian powers of U). Hence the difference between a first-order interpretation and a second-order one is that the latter works not only with the range U, but also with the ranges Pow(U), Pow(U^2), ... Nevertheless, multiple ranges can also be readily accommodated even within first-order semantics: directly within *sorted* first-order logic (which is a straightforward and formally unproblematic variety of

[5]See also Peregrin (1990).

[6]There is, of course, also a higher-order set theory, but model theory is usually considered as based on the first-order version.

[7]See also Peregrin (1992).

standard first-order logic, in which we have terms of multiple categories and hence multiple universes of individuals, or multiple compartments of the single universe), and indirectly even within standard (unsorted) first-order logic; the idea is that of „modeling“ the various ranges as various parts of the single universe. This can be done in such a way that the quantification over a specific range is replaced by the quantification over the whole universe, but each quantified formula is interpreted as a conditional whose antecedent restricts the quantification to that part of the universe which models the range in question: $\forall p.p(x)$ is thus interpreted as $\forall y.\mathbf{P}(y)\rightarrow\mathbf{PR}(y,x)$, where **P** is the characteristic function of that part of the universe which models the range of unary predicate variables, and **PR** is the binary predicate which renders predication as the relation between a pair of individuals.

Each second-order interpretation thus straightforwardly „induces“ a certain first-order interpretation, and the relation between the inducing and the induced preserves satisfiability. We may single out a certain class of first-order interpretations which are of the kind of those induced by second-order interpretations and call them *quasisecond-order interpretations*. Such a class can be characterized by a certain first-order theory; however, doing this suscitates the problem that although each second-order interpretation induces a quasisecond-order interpretation, not every quasisecond-order interpretation is induced by a second-order interpretation. Hence there will be no guarantee that each formula valid under every second-order interpretation will be valid also under every quasisecond-order interpretation; and Gödel’s theorem implies that there will indeed be formulas, which are second-order valid, but not quasisecond-order valid.

It is interesting that instead of interpreting second-order logic in the first-order way we can equally accept second-order interpretations in which the ranges of the predicate variables need not necessarily contain *all* the relations of the corresponding arities. The point is that it can be easily proved that there is a one-to-one correspondence between these so-called Henkinian interpretations[8] and quasisecond-order interpretations.

It is also worth noting that the difference between this kind of „reducing“ second-order logic to the first-order one, and the procedure discussed in the previous section is in fact only an „ideological“ one: while in the previous case we first translated the second-order language into a first-order language which we then interpret in the normal way, in the present case these two steps get mingled together: second-order logic is directly interpreted in the first-order way, and there is no intervening first-order language. (Neither do we introduce the problematical nominalization operator of the kind of **B** of the previous section.) In the previous section we said that we first „translate“ predicate into terms and then interpreted these by individuals; now we say that we interpret predicates directly by individuals - the difference is obviously not substantial.

5. The Principles Of Second-to-First-Order Translation

Let us analyze more closely how higher-order logic is reduced to lower-order logic - we sketch an algorithm for transforming any second-order language into a first-order one and each second-order theory into a first-order one. For simplicity we restrict ourselves to *monadic* second-order

[8] According to Henkin (1950).

logic, i.e. to second order logic the language of which contains no predicates of any arity greater than 1; and we shall further limit ourselves to languages not containing functors.

The language of monadic second-order predicate calculus (MPC2) thus consists of individual and predicate constants (ic, pc), individual and predicate variables (iv, pv), logical operators and quantifiers. The language of two-sorted first-order predicate calculus (PC1(2)) has no predicate variables, but rather only individual ones, and its individual constants and variables are divided into two *sorts* (hence we have ic^1, ic^2 and iv^1, iv^2).

Let us now have a language L_1 of MPC2. Let us construct the language L_2 of PC1(2) in such a way that:

- the set of ic^1 of L_2 is identical with the set of ic of L^1
- the set of ic^2 of L_2 is identical with the set of pc of L^1
- the set of iv^1 of L_2 is identical with the set of iv of L^1
- the set of iv^2 of L_2 is identical with the set of pv of L^1
- the set of pc of L_2 contains a single expression, the binary pc **PR** of the type <2,1> (i.e. such that it yields a statement together with a term of the sort 2 and a term of the sort 1).

Let us define, by induction, the translation of expressions of L_1 into those of L_2 - if X is an expression of L_1, let us denote its translation into L_2 as X^*:

$$\begin{array}{ll} X^* & = X \text{ if X is ic, pc, iv or pv} \\ (P(T))^* & = \mathbf{PR}(P^*,T^*) \\ (F_1 \mathbin{\&} F_2)^* & = F_1^* \mathbin{\&} F_2^* \\ (F_1 \vee F_2)^* & = F_1^* \vee F_2^* \\ (F_1 \rightarrow F_2)^* & = F_1^* \rightarrow F_2^* \\ (\neg F)^* & = \neg(F^*) \\ (\forall x F)^* & = \forall x^* F^* \\ (\forall p F)^* & = \forall f^* F^* \end{array}$$

Let us first note that the translation defined in this way could be interpreted as simply an *introduction of new notation* for MPC2 - as the trivial replacement of the notation $p(t)$ by the notation $\mathbf{PR}(p,t)$. From this vantage point, **PR** is nothing more than an auxiliary symbol on par with brackets. What is going to change, then, when we begin to see these notational variants of formulas of MPC2 as formulas of PC1(2) (and so also **PR** as a fully-fledged binary predicate)? The specific axioms of MPC2 concerning quantification over predicates obviously emerge as instances of the axioms of PC1(2) concerning quantification over terms of the sort 2; and similarly the rule for second-order generalization. As we include also the instances of the rule of comprehension (i.e. statements of the form $\exists p \forall x(p(x) \leftrightarrow F$, where x is the only variable free in F) among the axioms of MPC2, we have to add also the translation of this rule to the axioms of PC1(2). The translation is as follows (where x is a variable of the sort 1, y a variable of the sort 2, and x is again the only variable free in F):

$$\exists y \forall x(\mathbf{PR}(y,x) \leftrightarrow F) \qquad \text{(Compr)}$$

If, in L_2, the identity sign = is applicable only to terms of the sort 1, the translation thus defined will be a *one-to-one* function (to each formula of L_1 there corresponds a unique formula of L_2 and vice versa), and, moreover, it will obviously be the case that a formula of L_1 is provable in MPC2 just where its translation in L_2 is provable in PC1(2)+(Compr). If we admit the identity sign = between the terms of the sort 2, we will obtain formulas in L_2 which are not translations of any formulas of L_1 (the translation relation will not be surjective); and it will be reasonable to add the following axiom:

$$\forall y \forall z(\forall x(\mathbf{PR}(y,x) \leftrightarrow \mathbf{PR}(z,x)) \rightarrow (y=z)) \qquad \text{(Ext)}$$

Clearly, it will hold that a formula of L_1 is provable in MPC2 just where its translation in L_2 is provable in PC1(2)+(Compr)+(Ext).

Let now T be a theory in L_1; we define the theory T^* in L_2 in such a way that it contains the translation A^* of every axiom A of T, plus (Ext) and (Compr). Let $I=\langle U,P\rangle$ (where U is a set and P is an assignment of elements of U to ic's of L_1 and subsets of U to pc's of L_1) a model of the theory T. Let $U_1=U$, $U_2=Pow(U)$ and let P^* be such minimal extension of the interpretation function P that $P^*(\mathbf{PR})=\{\langle y,x\rangle \mid x \in y\}$. Then, $I^*=\langle U_1,U_2,P^*\rangle$ is obviously an interpretation of L_2. It is easy to check that I satisfies a statement F of L_1 if and only if I^* satisfies the translation F^* of F into L_2; and as I^* obviously satisfies both (Ext) and (Compr), I^* is a model of T^*. It follows that to every interpretation of a theory in MPC2 there corresponds a certain unique interpretation of the translation of the theory into PC1(2); especially to each interpretation of MPC2 there corresponds some unique interpretation of PC1(2)+(Ext)+(Compr).

Considering now the inverse case, let us take $I^*=\langle U_1,U_2,P^*\rangle$ to be a model of T^*. Let every element y of the set U_2 be assigned the subset m(y) of U_1, so that $m(y)=\{x \in U_1 \mid \langle y,x\rangle \in P^*(\mathbf{PR})\}$. (In this way we see the elements of U_2, informally speaking, as 'objectual correlates' of the subsets of U_1 - the element y is the objectual correlate of the set m(y), or, we may say, it is 'this-set-understood-as-an-object'. The axiom (Ext) guarantees that m is injective, i.e. that every element of U_2 is the objectual correlate of at most one subset of U_1). Let now P be such a function that $P(i)=P^*(i^*)$ for every ic i of L_1 and $P(p)=m(P^*(p^*))$ for every pc p of L_1; then $I=\langle U_1,P\rangle$ is an interpretation of the language L_1. Let us distinguish two cases: first, if the range of the function m is the whole set $Pow(U_1)$ (i.e. if every $u \subseteq U_1$ is the value of m(y) for some $y \in U_2$), then it is again clear that every statement F of L_1 is satisfied by I if and only if F^* is satisfied by I^*, and especially that I is the model of T; and the interpretations I and I^* correspond to each other in this sense. Second, if this is not the case, i.e. if the range of m is the proper part of $Pow(U_1)$ (i.e. if there exists an $u \subseteq U_1$ which is m(y) for no $y \in U_2$), then we cannot exclude the possibility of the existence of a statement F of L_1 which is satisfied by I, although F^* is *not* satisfied by I^*, or vice versa. (Such a statement could, for example, state the existence of just such a subset of the universe, the objectual correlate of which is not in U_2.) Hence: to some, but not necessarily to every, interpretation of a theory in PC1(2)+(Ext)+(Compr) there corresponds a unique interpretation of the translation of the theory into MPC2; especially to some, but not necessarily to every, interpretation of PC1(2)+(Ext)+(Compr) there corresponds some unique interpretation of MPC2.

Let us further show that a theory in sorted first-order logic can be straightforwardly translated into a theory in unsorted first-order logic. For that purpose, let us construct the language L_3 of PC1 so that

- the set of ic of L_3 is identical with the union of the set of ic^1 and the set of ic^2 of L_2
- the set of iv of L_3 is identical with the union of the set of iv^1 and the set of iv^2 of L_2
- the set of pc of L_3 is constituted by the binary predicate **PR** and the unary predicates $\mathbf{S^1}$ and $\mathbf{S^2}$.

We define the translation of L_2 into L_3 by induction - if X is an expression of L_2, we shall denote its translation into L_3 as X^+:

$X^+ = X$ if X is ic or iv
$\mathbf{PR}(T,T')^+ = \mathbf{PR}(T^+,T'^+)$
$(F_1 \mathbin{\&} F_2)^+ = F_1^+ \mathbin{\&} F_2^+$
$(F_1 \vee F_2)^+ = F_1^+ \vee F_2^+$
$(F_1 \to F_2)^+ = F_1^+ \to F_2^+$
$(\neg F)^+ = \neg(F^+)$
$(\forall xF)^+ = \forall x^+(\mathbf{S^i}(x^+) \to F^+)$, where i is the sort of the variable x in L_2

However, this translation is surely not surjective: hence there exist formulas of L_3 which translate no formula of L_2, viz. formulas which quantify over the whole universe, rather than over one of its parts modeling the sorts L_2 (i.e. formulas of the shape $\forall xF$ or $\exists xF$, where F is not of the shape $\mathbf{S^i}(x) \to F'$), or formulas, which contain the predicate $\mathbf{S^i}$ elsewhere than within the antecedent of a quantified implication.

Consider a formula F of L_2 which is an axiom of PC1(2), and its translation F^+ into L_3. If F is an axiom of the propositional calculus, then F^+ is obviously an axiom of PC1; and if F is an axiom of quantification, F^+ will be a direct consequence of the corresponding general axiom of quantification of PC1 if we accept, for every ic^1 X occurring in F, the postulate

$$\mathbf{S^1}(X^+) \tag{IC1}$$

and for every such ic^2 the postulate

$$\mathbf{S^2}(X^+). \tag{IC2}$$

Let thus T_2 be a theory in L_2; we define the theory T_3 in L_3 in such a way that it contains the translation A^+ of every axiom A of the theory T_2, plus the corresponding instance of the axiom (IC1) resp. (IC2) for every ic^1 resp. ic^2 of L_2, plus the following axioms (which concern exclusively such formulas of L_3 which translate no formulas of L_2):

$$\exists x.\mathbf{S^1}(x) \tag{NEmpt1}$$
$$\exists x.\mathbf{S^2}(x) \tag{NEmpt2}$$
$$\forall x.\mathbf{S^1}(x) \vee \mathbf{S^2}(x) \tag{Exhst}$$
$$\neg\exists x.\mathbf{S^1}(x) \mathbin{\&} \mathbf{S^2}(x) \tag{Disj}$$
$$\mathbf{PR}(y,x) \to \mathbf{S^2}(y) \mathbin{\&} \mathbf{S^1}(x) \tag{PR}$$

The theory T_3 is obviously a first-order theory and it holds that a formula F of L_2 is provable in T_2 just where its translation F^+ into L_3 is provable in T_3.

Let now $I=\langle U_1,U_2,P\rangle$ be a model of T_2. Let $U=U_1 \cup U_2$ and let P^+ be such function that $P^+(X)=P(X)$ whenever X is ic or pc of L_2, and $P(\mathbf{S^i})=U_i$ for i=1,2; then $I^+=\langle U,P^+\rangle$ is clearly an interpretation of L_3 and it is easy to check that a formula F of L_2 is satisfied by I if and only if F^+ is satisfied by I^+. Moreover, as I^+ obviously satisfies (NEmpt), (Exhst), (Disj), (PR) and all instances of (IC1) and (IC2), I^+ is a model of T_3. Conversely, let $I^+=\langle U,P\rangle$ be a model of T_3. Let $U_i=P(\mathbf{S^i})$ for i=1,2, and let P^- be the restriction of the function P to the set of all ic and pc of L_2; then $I=\langle U_1,U_2,P^-\rangle$ is obviously an interpretation of L_2 and it holds that the formula F^+ of L_3 is satisfied by I^+ just where F is satisfied by I; and thus I is also a model of T_2. Hence: a formula F is satisfied by a model of a theory T just where F^+ is satisfied by a model of T^+; and F is satisfied by every model of T just where F^+ is satisfied by every model of T^+.

Assembling our results so far, we can conclude that there exists a class of first-order interpretations (namely those which satisfy the axioms (NEmpt), (Exhst1), (Exhst2), (Disj), (PR) and the translations (Ext+) and (Compr+) of the axioms (Ext) a (Compr)), which „model“ - in a certain, exactly specified sense - second-order interpretations within first-order logic.

$$\forall y.\mathbf{S^2}(y)\rightarrow\forall z.\mathbf{S^2}(z)\rightarrow(\forall x(\mathbf{S^1}(x)\rightarrow(\mathbf{PR}(y,x)\leftrightarrow\mathbf{PR}(z,x)))\rightarrow(y=z)) \qquad \text{(Ext+)}$$
$$\exists y.\mathbf{S^2}(y)\,\mathbf{\&}\,\forall x.\mathbf{S^1}(x)\rightarrow(\mathbf{PR}(y,x)\leftrightarrow F) \qquad \text{(Compr+)}$$

Let us call these interpretations *quasisecond-order*. If we call the first order theory which is constituted by the axioms (NEmpt1), (NEmpt2), (Exhst), (Disj), (PR), (Ext+) and (Compr+) *quasisecond-order predicate calculus (QPC2)*, a quasisecond-order interpretation will turn out to be a (first-order) interpretation of QPC2. The conclusion reached above then reads that there exists a one-to-one correspondence between the set of all second-order and a subset of the set of all quasi-second order interpretations such that a second-order interpretation I is a model of a second-order theory T just in case the corresponding quasisecond-order interpretation I' is a model of the translation of T into first-order logic; hence there is a one-to-one correspondence between the set of all second-order and the set of certain quasisecond-order interpretations such that it preserves satisfaction - *modulo* translation.

However, there exist also such quasisecond-order interpretations which correspond to no second-order interpretation. This means that every formula of MPC2 the translation of which is valid in QPC2 is valid in MPC2; but it is generally not the case that the translation of every formula valid in MPC2 is valid in QPC2. The reason is that a formula of QPC2 may be satisfied by every interpretation which corresponds to an interpretation of MPC2, but in the same time not be satisfied by an interpretation which corresponds to no interpretation of MPC2. This possibility could be excluded only if we were able to restrict the set of quasisecond-order interpretations in such a way that it contained only those interpretations corresponding to second-order interpretations. In case of *monadic* second-order logic this is arguably possible - it has been proved that the set of formulas valid in MPC2 is recursive (see, e.g., Dreben and Goldfarb, 1979, chapter 8.3).

Hence there exist quasisecond-order interpretations to which there correspond no second-order interpretations - and these cause the fact that the translations of some formulas which are second-order valid are not quasisecond-order valid. In this sense, the translation of second-order logic into first-order logic will never be ‘perfect’. The standard second-order logic is thus not in general reducible to first-order logic. However, the situation is different if we do not define second-order interpretations in the way we have done it above, i.e. *standardly* - if we allow for such interpretations in which the ranges of predicate variables are proper subsets of the sets of all the corresponding relations, i.e. if we allow for *Henkinian* interpretations. There is a one-to-one correspondence between quasisecond-order interpretations and Henkinian second-order interpretations which preserves satisfaction (viz, e.g., Shapiro, 1991, chap. 4.3) - each Henkinian interpretation can thus be seen as a quasisecond-order interpretation and vice versa. Therefore second-order logic interpreted in the Henkinian way *is* reducible to first-order logic.

6. Summary of Translatability

We have indicated how to translate monadic second order logic into first-order logic; the translation of the full (non-monadic) second-order logic into first-order logic is analogous. We only need to add other sorts (or „quasisorts") for the predicates with arities higher than 1. The case of nonmonadic second-order logic, however, is provably different in that it is no longer possible to delimit the range of quasi-second order interpretations in such a way that second-order validity implies quasisecond-order validity: it follows from Gödel's incompleteness theorem that the set of all second-order valid formulas is not recursively enumerable (and hence axiomatizable); hence there exists a formula valid in second-order logic the translation of which is not valid within quasisecond-order logic. (This follows directly from the fact that there is a finite categorical axiomatization of Peano arithmetic within second-order logic: if PA is the conjunction of the axioms and G Gödel's undecidable formula, then the formula PA→G is obviously second-order valid, whereas its translation into first-order logic is *not* valid). Thus, we can summarize:

1. There *exists* a translation of second-order logic into first-order logic such that it generally holds that if F is a formula of second-order logic and F' its translation into first-order logic, then if F' is valid, F is also valid; and, moreover, F is provable in second-order logic if and only if F' is provable in first-order logic.

2. There *does not exist* a translation of second-order logic into first-order logic such that it would generally hold that if F is a second-order formula and F' its first-order translation, then if F is valid, F' would be also valid.

In a similar way we can define the translation of any logic of order n into a logic of an order lower than n. However, once we begin to investigate the translation of third-order logic into second-order logic, we discover a fact which may surprise us: third-order logic, and in general any logic of an order higher than 2, is *perfectly* reducible to second-order logic; hence passing from second-order logic to a higher-order one does not provide, in contrast to passing from first-order logic to second-order one, an increase in 'expressive power'. The only substantial difference is between the first and the second order - any logic of an order higher than 2 can be, without any loss of generality, considered a mere 'notational variant' of second-order logic.

Why this is so can be seen if we return to our considerations of the impossibility of reducing second-order logic to first-order one. We have reached the conclusion that the problem is that the set of quasisecond-order interpretations, as we managed to define it, contains also some interpretations which have no equivalents among second-order interpretations; we noted, that this problem would be solved if we managed to characterize just the set of those quasisecond-order interpretations which do have such counterparts. Returning to the terminology of the previous chapter, we can say that they are such interpretations for which each subset of the universe has its 'objectual correlate'; that is if $<U_1,U_2,P>$ is a quasisecond-order interpretation, then it has a second-order equivalent if and only if for every subset u of U_1 there exists a y from U_2 so that $u=\{x\in U_1 \mid <y,x>\in P^*(\mathbf{PR})\}$. The desired subset of the set of quasisecond-order interpretations therefore could be delimited if we added the following axiom to the axioms of QPC2

$$\forall p\exists y.\mathbf{S}^2(y)\&\forall x.\mathbf{S}^1(x)\rightarrow(\mathbf{PR}(y,x)\leftrightarrow p(x))$$

The reason why we could *not* do this was that this is a *second*-order formula - p is a *predicate* variable (hence we were able to accept only the weaker axiom schema (Compr+)). However, the situation would be different if the language into which we translate were second-order - then such an axiom *could* be accepted. If what we desired were, as before, the reduction of second-order logic, then our effort would, of course, be futile (we would find ourselves 'reducing' second-order logic to second-order logic); nevertheless the procedure can be nontrivially employed if what we reduce to second-order logic is a logic of an order higher that 2. For details see Shapiro (1989, Chap. 6).

7. Discussion and Conclusion

The question now is to what extent we really need full second-order predicate calculus with all its power, i.e. with all its not-first-order-reducible validities, and to what extent we can make do with that part of it which *is* first-order reducible - in other words, how far we need construe the semantics of second-order logic in the standard way, and how far we can construe it in the Henkinian way. It seems plausible that if what we are pursuing is the analysis of natural language with its pronouncements like (1) or (4), then nothing stands in the way of accepting the Henkinian semantics and thus construing higher-order logics as mere notational variants of first-order logic.

The situation is, of course, more involved if what we are investigating are the foundations of mathematics. Let us take the definition of infinity, as it is expressed in (6'). It is clear that this definition can be articulated as soon as we have the *syntactic* means of second-order logic. This definition also always (respectable of whether we interpret the language in the standard, or in the Henkinian way) delimits those sets which can be mapped on their own proper subsets, and declares these very sets as infinite. It is usually assumed that such a definition is correct where we are interpreting the language standardly, but incorrect if our interpretation is Henkinian; i.e. that to define the 'real' infinity (and hence also finiteness) we need the fully-fledged second-order logic. This is because under the standard interpretation our 'nonexistence' necessarily means *real* nonexistence, and our definition thus delimits just those sets which can be *really* mapped onto their own proper parts and are thus *really* infinite; whereas under the Henkinian interpretation our 'nonexistence' can mean only 'nonexistence within the range of the interpretation considered', and a set which cannot be, apparently, mapped on its proper part may quite possibly be infinite - for it may be the case that the mapping of the set on its proper subset does indeed exist, but happen not to be included into the range of the interpretation considered.

Without attempting at a real analysis of this state of affairs, let us note that this construal of the difference between the standard and the Henkinian second-order logic, although often taken for granted, is not wholly unproblematic. The difficulty is that it presupposes the picture according to which we apprehend directly the mathematical reality, and use the languages of logic only to *describe* it (let us stress that to assume this is *more* than to construe mathematical reality realistically, to assume that it exists independently of mathematicians). For the standard construal of second-order logic would not make good sense if we did not take such concepts as *all subsets of a given set* for granted. Against such a picture, a different picture can be opposed - the picture which was analyzed, for the first time, by Skolem (esp. 1958), and which is based on the view that mathematical concepts are inherently relative - that they make sense only within the context of a particular theory. If we thus say that a set is infinite, we have to ask *in which theory* - for the set can be finite according to one theory (perhaps according to the Henkinian second-order logic), whereas be infinite according to another theory (the standard second-order logic). The fact that from the viewpoint of a standard model the corresponding Henkinian model may appear as lacking something does not yet mean that the former is complete, and the latter incomplete, in an absolute sense.

However, this view makes the very notion of a *standard* interpretation problematic: for 'to be standard' means 'to include *all* subsets', and something thus can be called standard only from an absolute vantage point, from which it is possible to decide, when the subsets are *all*, and when not. Standard interpretations therefore cannot be delimited otherwise than by recoursing to a further unanalyzed concept of all subsets (which is quite straightforward for *finite* sets, but less so for infinite ones) - in contrast to Henkinian interpretations they cannot be gerrymandered via a recursive specification. It follows that to the same extent to which second-order logic is more plausible as the foundation of mathematics, it is - in a certain sense - more trivial. Exaggerating somewhat we can say that whereas we lack the means to characterize, e.g., infinite sets within first-order logic, in second-order logic we are enabled to do so, but with little more significance than when we simply said that infinite sets were those which were ('really') infinite.

Although the relationship between first-order and second-order logic may present problems, it is sure that the relation between second-order and higher-order logic is not problematic - any logic of an order higher than two can be considered as a 'notational' variant of second-order logic (which does not, of course, preclude this very notational variant from being sometimes quite useful).

I think that discussions of the relationship between first-order and higher-order logic often suffer both from the fact that its participants do not satisfactorily specify *what exactly* they mean by higher-order logic, and from the fact that they do not consider the whole depth of the problematicity of the relationship. This paper is meant to summarize some facts by which this problematicity is characterized.

References

Barwise, J., Feferman, S., eds. (1985): *Model-theoretic logics*, Springer, New York.

Davidson, D. (1980): *Essays on Actions and Events*, Clarendon Press, Oxford.

Dreben, B. and W. Goldfarb (1979): *The decision problem: solvable classes of quantificational formulas*, Addison-Wesley, London.

Frege, G. (1892): 'Über Begriff und Gegenstand', *Vierteljahrschrift für wissentschaftliche Philosophie* 16, pp. 192-205.

Henkin, L. (1950): 'Completness in the Theory of Types', *Journal of Symbolic Logic* 15, pp. 81-91.

Chierchia, G. (1982): 'Nominalization and Montague Grammar', *Linguistics and Philosophy* 5.

Church, A. (1940): 'A Formulation of the Simple Theory of Types', *Journal of Symbolic Logic* 5, pp. 56-68.

Kripke, S. (1963): 'Semantical Considerations on Modal Logic', *Acta Philosophica Fennica* 16, pp. 83-94.

Moore, G.H. (1988): 'The Emergence of First-Order Logic', *History and Philosophy of Modern Mathematics* (ed. W. Aspray and P.Kitcher), University of Minnesota Press, Minneapolis.

Parsons, T. (1990): *Events in the Semantics of English*, MIT Press, Cambridge (Mass.).

Peregrin, J. (1990): *Theory of Types: Good Servant Bad Master*, Prague Studies in Mathematical Linguistics 10, 159-176.

Peregrin, J. (1992): *Meaning, Truth and Models*, From the Logical Point of View 2/92, 67-75.

Peregrin, J. (1995): *Doing Worlds with Words*, Kluwer, Dordrecht.

Shapiro, S. (1991): *Foundations without Foundationalism*, Clarendon Press, Oxford.

Skolem, T. (1958): 'Une relativisation des notions mathématiques fondamentales', *Colloques internationaux du Centre de la Recherche Scientifique*, Paris, 13-18; reprinted in Skolem: *Selected Works in Logic* (ed. J.E. Fenstadt), Universitetsforlaget, Oslo, 633-38.

Turner, R. (1983): 'Montague Semantics, Nominalization and Scott's Domains', *Linguistics and Philosophy* 6.