Set Theory and Nominalisation, Part I
Journal of Logic and Computation, 2(5), 579-604, 1992

Fairouz Kamareddine
Department of Mathematics and Computing Science
Eindhoven University of Technology
Den Dolech 2, postbus 513
5600 MB Eindhoven, the Netherlands
email: fairouz@info.win.tue.nl
tel: +31 40 474319

November 30, 1996

Abstract

This paper argues that the basic problems of nominalisation are those of set theory. We shall therefore overview the problems of set theory, the various solutions and assess the influence on nominalisation. We shall then discuss Aczel's Frege structures and compare them with Scott domains. Moreover, we shall set the ground for the second part which demonstrates that Frege structures are a suitable framework for dealing with nominalisation.

Keywords: *Frege structures, Nominalisation, Logic and Type freeness.*

1 The Problems

We shall examine the problem of the semantics of nominalised terms from two angles: the formal theory and the existence of models.

1.1 The problem of the formal theory

Any theory of nominalisation should be accompanied by some ontological views on concepts — for predicates and open well-formed formulae act semantically as concepts. This is vague, however, if only because where I use the word *concept*, someone else might use *class, predicate, set, property* or even *system*(Dedekind). This terminological profusion is hardly surprising, for we are touching on the problem of *universals*, a problem philosophers have been debating for hundreds of years. (This new term — universal — may be more confusing than any of the others, but we may use Aristotle as a preliminary guide and define a universal to be *that which can be predicated of things.*) The aim of this section is not to take a standpoint on any of the philosophical theories of universals; rather it is to show that, no matter what approach we adopt, nominalisation is going to generate a problem.

1.1.1 Ontology, concepts, predicates, properties and sets

According to Quine in [Quine 1969], page 1, "*the notion of a class is such that there is supposed to be, to the various things of which that sentence is true, a further thing which is the class having each of those things and no others as member.*" As an example we take the sentence *being an x such that the colour of x is red*. We have in our universe various things of which this sentence is true; but perhaps we can also say that the class of all those things which are red also exists in our universe. I say perhaps because it will be shown shortly that if we let any open sentence determine an object which is the class of all those things of which the sentence is true, we run into difficulties. To see this clearly it is important that the reader bear in mind the following four notions: *the Comprehension Principle, Quantification, Interpretation* and *Russell's Paradox*. I shall comment here on how each such notion is to be understood in the present context.

The Comprehension Principle This is the principle which decides which open sentence in our theory determines a class (or set) of precisely those entities that satisfy it.

Quantification Take a class which stands for an open sentence (i.e. the class of all those objects which when substituted for the free variables in the open sentence returns true). Does this class act exactly like any other object in our universe? If so, should we be able to quantify over it?

Interpretation Should we keep to a full classical interpretation or use a non-classical one? If we keep to a full classical interpretation, and assume that the comprehension principle applies to each open sentence and that we have full quantification, we will fall foul of Russell's paradox.[1]

Russell's paradox The paradox derives from assumptions similar to the following: Let S be the set of all sets that do not contain themselves. Such an assumption is contradictory for we can deduce from it that S is in S iff S is not in S.

The important point to concentrate on is how these four notions interact, and in particular to note that an assumption of full comprehension (i.e. every open sentence determines a class) and of full quantification (i.e every class acts exactly like any object and can be quantified over) will, under some interpretations, lead to Russell's paradox. This point will be presented in more detail in the following section. We will now describe the four main conceptions of universals, all of which will have to face up to this sort of problem.

1. **Realistic conception (Platonism)** Platonists take concepts to be real properties. That is, concepts are language/observer independent entities. Platonists also subscribe to an unrestricted (or full) comprehension principle, i.e. to each well defined condition, there exists a set (or class) of all entities satisfying the condition. Moreover, this set is an entity in its own right and can be quantified over. According to this conception, interpretation is much more important than language and therefore it seems obligatory to use the referential (and not the substitutional) interpretation where a substitutional interpretation of the quantifiers $[[\,]]_{S,g}$ involves truth clauses of the following kind:
   $[[\exists x \Phi]]_{S,g}$ is true $\Leftrightarrow$ for some name a in the language, $[[\Phi[a/x]]]_{S,g}$ is true.
   $[[\forall x \Phi]]_{S,g}$ is true $\Leftrightarrow$ for every name a in the language, $[[\Phi[a/x]]]_{S,g}$ is true.
   By contrast, Referential interpretation $[[]]_{R,g}$ treats quantifiers as follows:
   $[[\exists x \Phi]]_{R,g}$ is true $\Leftrightarrow$ for some object a in the model, $[[\Phi[a/x]]]_{R,g}$ is true.
   $[[\forall x \Phi]]_{R,g}$ is true $\Leftrightarrow$ for every object a in the model, $[[\Phi[a/x]]]_{R,g}$ is true.

[1] It should be noted that the paradox occurs even in intuitionistic theories.

2. **Formalist conception (Nominalism)** Formalists, of whom Hilbert was the father, insist on the paramount importance of language. Hilbert's program, as it is well known, consisted of separating signs and meaning and only allowing finitary arguments in the proof theory. Had the program worked, it would have made it easy to prove things about the theory inside the theory itself. Gödel's result made apparent the impossibility of carrying out this aim — and as has been said by Quine ([Quine 1969]): "*Gödel's proof is beyond doubt, we can philosophise about it but we can not philosophise it away*". According to the formalists, concepts are predicate expressions which do not exist beyond our linguistic expressions. Open sentences are excluded from standing for concepts, and furthermore the comprehension principle is restricted. As language is the most important thing for them, interpretation is secondary. Thus it seems that the obvious semantics should be based on a substitutional interpretation.

3. **Conceptualism** Borrowing a sentence from Fraenkel (at the end of [Fraenkel 1973], page 336): Conceptualists are "*attracted neither by the luscious jungle flora of platonism nor by the ascetic desert landscape of neo-nominalism*". Concepts here are neither predicate expressions nor real properties. They are not objects but unsaturated entities, the saturation of which results in a mental act and not necessarily a truth value. Some conceptualists are constructive and construct only those sets that correspond to predicative conditions; some others accept an unrestricted comprehension principle. However all of them care for interpretation, and in a semantics for a conceptualistic theory one should consider a referential interpretation where the meaning of a concept applied to an object does not necessarily have to be a truth value.

4. **Fregean conception** It might be said that Frege is both a realist and a conceptualist but; he is anti-formalist and tends to lean towards conceptualism. The ontology assumed by Frege of concepts was that they are functions of one argument whose values are always truth-values. Concepts, according to him, are unsaturated, and the behaviour of a concept is predicative even if something is being asserted about it. The unsaturation of a concept comes from the fact that concepts can never themselves be objects and only by applying the concept to an object can we obtain a saturated element (an object which is a truth-value). Assertions that are made about concepts do not apply to objects: for example, existence is a property of concepts and not of objects. However, the way we attach properties to concepts consists in predicating the property not of the concept but of the concept-correlate. This concept-correlate is the extension of the concept, according to Frege, and is an object. We said that concepts here are functions: thus the graphs of functions are objects even though functions themselves are not. This is exactly the case with concepts and their extensions. The extensions are objects but the concepts themselves are not. The extension of a concept does not fully determine the concept, for we can have two extensions which are the same while the concepts themselves are not. Frege always warned against confusing a concept with its extension and defined sets and classes to be the extensions of the concepts, not the concepts themselves: "*sets and classes are objects whereas concepts are anything but objects*". Something falls under a concept and the grammatical predicate stands for this concept. A name of an object is incapable of being used as a grammatical predicate. For Frege, the saturation of a concept results in a truth-value and according to him each open sentence denotes a class. Those classes are objects and can be quantified over. Being

an anti-formalist he insisted on interpretation, but as is well known he paid a high price for these relaxed conditions: his theory, known as the naive theory, was found to be subject to Russell's paradox, since the concept *the set of all those things that do not belong to themselves* has an extension K which is a proper object. Thus his theory is contradictory.

To avoid inconsistency, some people restricted their comprehension principle but still allowed unlimited quantification; others restricted both quantification and comprehension. Let us here examine in detail how the Russell paradox can threaten theories of nominalisation; and then we shall meet some solutions to the problem.

1.1.2 A language of nominalisation

If we are going to assume a first order language of nominalisation and we are going to let any open well-formed formula stand for a concept, then we might fall into the paradox. This is shown as follows: take a first order calculus and add to it a new primitive relation $\in$ and the axiom:

Comprehension For each open well-formed formula Φ, $\exists y \forall x[(x \in y) \Leftrightarrow \Phi(x)]$ where y is not free in $\Phi(x)$.

This theory is obviously inconsistent, for take $\Phi(x)$ to be $\neg(x \in x)$. Then we get:
$\exists y \forall x[(x \in y) \Leftrightarrow \neg(x \in x)] \Longrightarrow \forall x[(x \in y) \Leftrightarrow \neg(x \in x)] \Longrightarrow [(y \in y) \Leftrightarrow \neg(y \in y)]$.

In this theory of nominalisation, we assumed that each open well-formed expression determines a concept whose extension exists and is the set of all those elements which satisfy the concept. We could restrict our comprehension principle so that $\Phi(x)$ stands for everything except $\neg(x \in x)$, but this will not save us from paradox. To see this let $\Phi(x)$ stand for $\neg(x \in_2 x)$ where $(x \in_2 y)$ abbreviates $(\exists z)((x \in z) \wedge (z \in y))$. Again, ruling out this instance is not enough for we will still get the paradox if we take $\Phi(x)$ to be $\neg(x \in_3 y)$. This process continues ad infinitum. We could rule out all such instances —but the problem will persist, for take a sentence $\Phi(x)$ like: $\neg(\exists z_1, z_2, \ldots)[\ldots(z_3 \in z_2) \wedge (z_2 \in z_1) \wedge (z_1 \in x)]$ and let y be the class obtained from the comprehension axiom for $\Phi(x)$.

- If $(y \in y)$ then $\neg(\exists z_1, z_2, ...)[... \in z_2) \wedge (z_2 \in z_1) \wedge (z_1 \in y)]$. But we can take $z_1 = z_2 = ... = y$, and get a contradiction.
- If $\neg(y \in y)$ then $(\exists z_1, z_2, ...)$ $[... \in z_2) \wedge (z_2 \in z_1) \wedge (z_1 \in y)]$. But as $(z_1 \in y)$ then $\Phi(y)$; however we have that $\neg\Phi(y)$. Contradiction.

We have assumed above a first order language of nominalisation but we do not take a standpoint on whether we need higher order languages.[2] We shall hence show that we also face the problem with higher order languages. For this I shall use a second order theory due to Cocchiarella's formulation of second order logic with nominalised predicates which appears in [Cocchiarella 1984]. This language essentially embodies Frege's conceptions of concepts and objects summarised above, according to which we need to quantify over our predicates.

[2]We will however, remark en passant that it seems we do not need to go higher than second order languages for the semantics of nominalisation. In fact, according to Frege's conception, we stop at second level concepts, but these can be mapped into first order concepts which in turn can be mapped into objects. So when we come to quantify over properties, we really quantify over their extensions which are objects.

Moreover, predicate quantifiers have a referential significance, even though predicates themselves are not singular terms. I shall start by writing down the axioms and rules of a second order language which will accommodate nominalised predicates. If this language is to allow us to talk about nominalisation, it should have a device which can turn any open wff (well formed formula) or predicate into a singular term. For example, we should turn *run* into *to run*, *the sun is grey* into *that the sun is grey* and so on. Clauses (1), (8) and (9) below will see to this.

The typing of the language is as follows:

0 represents the type of all singular terms,

1 represents the type of propositions,

n+1 represents the type of n-place predicates, for $n > 0$.

For each $n > 0$ assume the existence of denumerably many variables. I shall use the following metavariables:

$x, y, \ldots$ refer to individual variables

$u, \ldots$ refer to both individual and predicate variables

$F_n, G_n \ldots$ refer to n-place predicate variables. We can get rid of the subscript when no confusion occurs

$x, y, z, w, \ldots$ refer to individual variables.

$a, b, \ldots$ refer to singular terms.

The primitive symbols of the language are: $\Rightarrow, \neg, =, \forall, \lambda$. The others are defined in the metalanguage. The meaningful expressions of any type n, ME_n are defined recursively as:

1. Every individual variable (or constant) is in ME_0 and every n-place predicate is in both ME_0 and ME_{n+1}, for $n > 0$.

2. For a, b in ME_0, $(a = b)$ is in ME_1.

3. If Π is in ME_{n+1}, and $a_1, \ldots, a_n$ are in ME_0 then $\Pi(a_1, \ldots, a_n)$ is in ME_1.

4. If Φ is in ME_1 and $x_1, .., x_n$ are pairwise distinct variables, where $n \geq 1$, then $[\lambda x_1, \ldots, x_n F]$ is in ME_{n+1}.

5. If Φ in ME_1 then $\neg\Phi$ is in ME_1.

6. If Φ, Ψ are in ME_1 then $(\Phi \Rightarrow \Psi)$ is in ME_1.

7. If Φ is in ME_1, u is an individual or a predicate variable then $\forall u\Phi$ is in ME_1.

8. If Φ is in ME_1 then $[\lambda\Phi]$ is in ME_0.

9. For all $n > 1$, ME_n is included in ME_0. (9 does not follow from 1.)

Axioms:

- (A0*) All tautologous (classical) well formed formulae.
- (A1*) $\forall u(\Phi \Rightarrow \Psi) \Rightarrow (\forall u\Phi \Rightarrow \forall u\Psi)$, for u an individual or a predicate variable.
- (A2*) $\Phi \Rightarrow \forall u\Phi$, for u an individual or a predicate variable not free in Φ.
- (A3*) $\exists x(a = x)$, for a singular term in which x is not free.

- (λ^*) $(a = b) \Rightarrow (\Phi \Leftrightarrow \Psi)$ where a, b are singular terms and Ψ comes from Φ by replacing one or more free occurrences of b by free occurrences of a.
- (CP*) $\exists F_n \forall x_1, .., x_n [F_n(x_1, .., x_n) \Leftrightarrow \Phi]$ where F_n does not occur free in Φ and $x_1, .., x_n$ are distinct individual variables.
- (λ-CONV*) $[\lambda x_1, \ldots, x_n \Phi](a_1, \ldots, a_n) \Leftrightarrow \Phi(a_1/x_1, \ldots, a_n/x_n)$ where $a_1, \ldots, a_n$ are singular terms and each a_i is free for x_i in Φ.
- (ID λ^*) $[\lambda x_1, \ldots, x_n R(x_1, .., x_n)] = R$ for R an n-place predicate variable or constant.

Inference Rules: The two inference rules are MP and UG, where

- MP is: infer from $\Phi \Rightarrow \Psi$ and Φ that Ψ.
- UG is: infer from Φ that $\forall u \Phi$ where u is an individual or a predicate variable.

The system (just described) is subject to Russell's paradox, for take the special instance of (CP*): $\exists F \forall x [F(x) \Leftrightarrow \exists G [x = G \wedge \neg G(x)]]$. The presence however of (CP*) is necessary for second-order logics with nominalised predicates and the problem comes from (CP*) together with (A3*) under various logical laws.

From our above discussion, it seems that set theory is very basic to nominalisation. Let us hence, comment on the ontological status of sets and on the nature of Russell's paradox, as the solutions depend on both issues.

1.1.3 The ontological status of sets

There are two main views of sets: *the mathematical conception* of set and *the logical conception*. According to the mathematical conception, a set is determined by the elements that belong to it. E.g. {1,2,3} is the set of the numbers 1, 2 and 3. The logical conception, on the other hand, regards sets as existing according to their defining concepts, and not their constituent objects; so here {1,2,3} might be the set of positive integers less than 4. Frege's conception of set was a logical one, and is known in the literature as the naive conception of set. According to this view, any predicate has an extension and sets are extensions of predicates. However, under the classical laws of logic and especially the law of excluded middle (LEM) and non-free logic (where not necessarily each element denotes), this notion of set is subject to Russell's paradox. However, the paradox holds even in minimal logic and other non-classical logics, e.g. we can derive the paradox without the use of LEM which means that the paradox is intuitionistically derivable. I shall illustrate the occurrence of the paradox by assuming both LEM and that every predicate has an extension. Now, if one chooses $P(x)$ to be $\neg(x \in x)$, then $\{x : \neg(x \in x)\}$ is an r to which LEM applies. So we have either $(r \in r)$ or $\neg(r \in r)$. In both cases we get a contradiction.

After Frege's naive set theory was shown to be inconsistent, set theorists were anxious to solve the problem, and many directions were followed to overcome the paradox. Frege himself had something to say about the paradox. He stated that if one abandoned the naive conception and the use of full comprehension, it would not be obvious how to define numbers (see [Frege 1970], Frege on Russell's paradox). This follows because the essential definition of numbers in Frege's theory was based on the existence of extensions of concepts — thus the paradox shook Frege's whole theory. Frege suggested that the solution lay in either banishing

LEM for classes, or forbiding some concepts from having extensions. He was not satisfied with the first solution because he wanted classes to be full objects - and full objects obey LEM. If classes are to be considered as improper objects then this will create an infinite number of types in the theory, for we are going to have functions that apply to proper, improper or mixed arguments. Frege was not in favour of that solution, and preferred to acknowledge the existence of concepts that have no extensions. This would affect axiom (V) and in particular (Vb) where (V) and (Vb) are as follows:

(V) $z'f(z) = z'g(z) \Longleftrightarrow \forall x(f(x) \Leftrightarrow g(x))$, where $z'f(z)$ is the extension of f

(Vb) $z'f(z) = z'g(z) \Longrightarrow \forall x(f(x) \Leftrightarrow g(x))$.

This axiom states that if two concepts are equal in extension then whatever falls under one falls under the other (see [Frege 1970], pages 214-224). Frege made only general remarks about the inconsistency and did not pin down what caused the problem. He sometimes felt the problem lay in (Vb) and at other times thought that the assumption of the existence of an extension to each concept was to blame. However, (Va) the opposite direction of (Vb), is acceptable as it takes us from equality that holds in general to an equality that holds of graphs (or extensions). But according to Frege (in [Frege 1970] page 219), "*We cannot in general take the words 'the function $\Phi(c)$ has the same graph as the function $\Psi(c)$' to mean the same thing as the words 'the functions $\Phi(c)$ and $\Psi(c)$ always have the same value for the same argument'; and we must take into account the possibility that there are concepts with no extension (...)*". This is true; however, Frege did not realise that his domain of concepts was far too big. Concepts are propositional functions but according to Frege's conception, there are far more propositions than there should be. For each object a, $-a$ (the content of a) is a proposition even though a was not. Thus Frege has far too many concepts and some paradoxical sentences stand for concepts when they should not do. Accordingly, a way of ruling out the paradox might be to restrict the number of concepts. Let us look again at the paradoxical sentence: *the set of all things that do not belong to themselves.* Under the restriction strategy, we cannot tell whether this sentence stands for a concept or not, as we do not know if this is a propositional function or not so we cannot think of its extension. We could say that there were two ways of reformulating set theory. One is to abandon Frege's definition of set and use the mathematical notion instead. The second is to keep to the logical definition of set and try to make it consistent. To conclude this section, it is worth drawing attention to the role self reference plays in these set theoretic paradoxes. Paradoxes involving self reference are well known in the literature, and are of two kinds: logical and semantical paradoxes. Russell's paradox has been classified under the logical category, as have the barber's paradox and Cantor's paradox. Of the semantical paradoxes, we mention the Grelling's [3] and The liar's paradoxes.

1.2 The problem of the existence of models

The theory discussed in 1.1 is inconsistent, so it does not have models. But even in the case of a theory whose consistency we are sure of, we still sometimes cannot imagine what the models look like. This section describes what a model of nominalisation should be, and what the difficulties of constructing such models are.

[3]Some adjectives possess the proeprty that they denote (e.g. English, Polysyllabic) and some do not (e.g. French). Call the second type *heterological*; then *heterological is heterological* iff *heterological is not heterological.* Another example of this paradox is: A concept is *predicable* if it can be predicated of itself, otherwise it is *impredicable.* Hence *impredicable is impredicable* iff *impredicable is not impredicable.*

1.2.1 What a model should look like

A model of nominalisation will be roughly as follows: $M = (U, P, f)$ where U is the domain of objects, P is the domain of functions from U into $\{0, 1\}$ and f is the nominalisation function which is needed for nominalisation as predicates should be turned to objects in order that they can be applied to themselves. f is a function from P into U which should be injective. This implies that P is a subset of U up to an isomorphism. Let me describe in more detail what this means. In trying to build our semantic function which maps each syntactic entity into a semantic one, we should do the following:

(1) Map individual variables and singular terms into objects in U.

(2) Map the predicates into P, the domain of the first order properties. The nominalised items are singular terms and they are mapped into U. The function f acts as a nominalisation function, assigning to each element p of P, an element in U called the correlate of p. This correlate is the denotation of the nominalised item that corresponds to the predicate.

$f : P \longrightarrow f(P)$ is an isomorphism because:

- f is well defined: We assume that each property has a single correlate.
- f is injective: We assume that each two distinct properties in P have distinct correlates in $f(P)$.
- f is surjective: Because every element in $f(P)$ corresponds to an element in P.

So in constructing a model of nominalisation, we should construct three domains such as U, P and $f(P)$ satisfying the condition that P (or $f(P)$) is a subset of U. According to Cantor's diagonal theorem, we cannot take P to be the set of all functions from U to $\{0, 1\}$. We have to restrict P, but we should not restrict it too much, for we would like to obtain the nominal of all the desired items.

1.2.2 Difficulties with such models

Cantor's Theorem will pose a difficulty to any theory which aims to make functions play the role of objects. According to Cantor's theorem which states that if S is any set, then the power set of S has a greater cardinality than S, the cardinality of a function space is bigger than the cardinality of the domain itself.

1.2.3 Existence of models

The above shows that we are going to have problems constructing models of nominalisation —recall that we previously wanted P to be a subset of U, but by Cantor's theorem the cardinality of P is greater than that of U. In essence, we need to find ways of restricting P without either lapsing into triviality or running foul of Cantor's theorem. That is, we are looking for interesting restrictions —restrictions which leave us with enough functions for nominalisation. We must break the ties created by the old tradition and build somewhat more original models. We shall in the following paragraph talk about different ways of proving the existence of non-trivial models which are not susceptible to Cantor's argument. Those models will contain denotations for all nominalised items. Scott models and Frege structures both possess this property; but as we shall see, the former have a difficulty regarding quantification, while the latter do not. Non well founded sets on the other hand are a third kind of model that should be looked at from a different angle.

2 The different solutions to the theoretical problem

We said that the theoretical problem is mainly a problem of set theory and of predication theory. The following is a summary of various set theories and their application to the development of theories of nominalisation.

2.1 Notes on set theory

2.1.1 Altering the language

Since Russell's letter to Frege, concerning the inconsistency of Frege's system, there have been many attempts at overcoming the paradox. The first two accounts of avoiding the paradox by restricting the language were due to Russell and Poincaré. They both disallowed impredicative specification: only predicative specification (as will be defined below) was to be permitted. Russell's own solution (in [Russell 1908]) was to adopt the vicious circle principle which can be roughly stated as follows: "*No entity determined by a condition that refers to a certain totality should belong to this totality*". Poincaré (in [Poincaré 1900]) took refuge in banning "*les définitions non prédicatives*" which were taken by him to be: *Definitions by a relation between the object to be defined and all individuals of a kind of which either the object itself to be defined is supposed to be a part or other things that cannot be themselves defined except by the object to be defined.* So both Russell and Poincaré required only predicative sets to be considered, where $A = \{x : \Phi(x)\}$ is predicative iff Φ contains no variable which can take A as a value. This helps because it is otherwise very easy to get a vicious circle fallacy if we let the arguments of a certain propositional function (or the elements of a set) presuppose the function (or the set) itself. Russell's and Poincaré's solution was to use predicative comprehension, instances of which start with individuals, then generate sets, then new sets and so on as in the following example: Take 0 at level $0, \{0, \{0\}\}$ at level $1, \{0, \{\{0\}\}, \{0, \{0\}\}\}$ at level 2 and so on. Russell's simple theory of types in Principia Mathematica applied the vicious circle principle, assuming all the elements of the set before constructing it. This theory obviously overcomes the paradox for the sentence Φ denoting $\neg(y \in y)$ is not stratified. Let us recall that the concept of stratification for it is going to form an important step in our discussion and assessment of our theory in terms of the others. There are two types of stratification: *homogeneous* stratification and *heterogeneous* one. A well formed formula Φ is said to be *heterogeneously stratified* if there is a function f from the variables and constants of Φ to the natural numbers such that for each atomic well formed formula $F(x_1, ..., x_n)$ of $\Phi, f(F) = 1 + max[f(x_i)]$. Φ is said to be *homogeneously stratified* if the function f is further restricted so that $f(x_i) = f(x_j)$ for $0 \leq i, j \leq n$.

This approach (of Russell and Poincaré) is rather unsatisfactory from the point of view of nominalisation, for the following reasons:

1. We need formulas which are not stratified (i.e. where we have impredicativity), such as the sentence it is nice to be nice. In fact sentences that involve self application or self reference are not there because they are not stratified. However self application and self reference are fundamental to nominalisation.

2. A class can have members only of uniform type. Also, sets here can neither belong to themselves, nor contain other sets from the same level. Hence again no self reference.

3. Also most of our structures get reproduced at each level. For instance, universal classes, numbers and Boolean algebras. There is however, another approach which falls under this category, that of Cocchiarella's λHST* in [Cocchiarella 1986]. λHST* uses homogeneous stratification where the paradox is avoided by disallowing problematic λ-abstracts from being well-formed. λHST* however, allows for formulas such as *Nice(Nice)* to be expressed and to be provable. This approach though, is rather different from that of Russell's type theory. Predicates are not typed in this system as part of the object language the way they are in Russell's type theory.

2.1.2 Altering the axioms

- *Iterative sets* We can avoid the paradox by altering not the language but the axioms of the theory. The most straightforward such theory is ZF (Zermelo-Fraenkel) where the axioms are made to fit the limitation of size doctrine; that is, sets are not allowed to get too big too quickly. Take the system of first order logic provided in 1.1, and alter comprehension to the following axiom which is known as the separation axiom:

  For each open well formed formula $\Phi, \exists x \forall y[(y \in x) \Leftrightarrow (y \in z) \wedge \Phi]$ where x does not occur in Φ.

  It is exactly this new axiom which is responsible for the elimination of the paradoxes. Take Russell's paradox: to prove the existence of $\{x : \neg(x \in x)\}$ we need a z big enough so that $\{x : \neg(x \in x)\}$ is included in z. But we cannot show the existence of such a z. More precisely Russell's paradox is restricted in ZF as follows:

  Take $\Phi(z)$ to be $\neg(z \in z)$,

  take $x = \{y : (y \in z) \wedge \neg(y \in y)\}$

  - If $(x \in x) \Longrightarrow (x \in z)$ and $\neg(x \in x)$ contradiction,
  - If $\neg(x \in x) \Longrightarrow$
    * if $(x \in z) \Longrightarrow (x \in x)$ contradiction,
    * if$\neg(x \in z)$ then we are fine.

  So the limitation of size doctrine exemplified by the above axiom is how we avoid the paradox. Note however that we still have sets which belong to themselves, for if we take $\Phi(x)$ to be $x \in x$ then $\exists z \forall x[(x \in z) \Leftrightarrow (x \in \alpha \wedge x \in x)]$, yet those sets have to belong to some already existing sets. Throughout the development of ZF, it was felt that the following foundation axiom (which is independent of and consistent with all other axioms of ZF) has to be added:

  (FA) $(\exists x)(x \in a) \Rightarrow (\exists x \in a)(\forall y \in x)\neg(y \in a)$

  As a corollary of (FA), we have that there is no set a which has itself as its only element, for if there was then take $x = a$ in (FA) above and you get $(\exists x \in a)(\forall y \in x)\neg(y \in a)$, which is absurd. In fact, it is easy to prove the following corollaries of (FA):

  1. No solution for $a = \{\ldots\{a\}\ldots\}$ exists, where $\{\ldots\{n$ times for $n > 0$.
  2. No solution for $a = \{a, b\}, b \neq a$ exists.
  3. No solution for $a = \{b, \{a\}\}$ exists.

4. No solution for $a = \{\{b\}, \{a, b\}\}$ (i.e. $a = (b, a)$) exists.
5. No solution for $a \in_i a$ for $i \geq 1$, where $x \in_i y \Leftrightarrow (\exists z_1, \ldots, z_{i-1})(x \in z_{i-1} \in \ldots \in z_1 \in y)$.

It is worth pointing out that although very different conceptually, both the simple theory of types and ZF which of course includes (FA), give rise to an iterative concept of set. That is, both require the elements of a set be present before a new set can be constructed [Boolos 1971]. This implies that ZF is not adequate for nominalisation.

- *2 non iterative sets, NF and ML* ZF is not the only axiomatic approach aimed at restricting the paradoxes. In NF (New Foundations), Quine restricts the axiom of comprehension of 1.1.2, to obtain the following:

  (SCP) $\exists x \forall y[(y \in x) \Leftrightarrow \Phi(y)]$ where x is not free in $\Phi(y)$ and $\Phi(y)$ is stratified.

  Thus it applies only to stratified formulae and now the only concepts that are allowed to have extensions are the concepts that correspond to these stratified formulae. In ZF, we did not have a universal set whereas in NF we do, for take $x = x$, this is a stratified formula and hence NF can have nominalisation. Moreover, NF has only one universal set, one complement of each set, and one null set. Furthermore, Cantor's theorem does not hold in NF (the universal set is equinumerous to its power set). However, NF is weak for mathematical induction and the axiom of choice is not compatible with NF. We cannot prove Peano's axiom $[s(n) = s(m) \Rightarrow n = m]$ in it, unless we assume the existence of a class with $m+1$ elements. Also, NF is said to lack motivation because its axiom of comprehension is justified only on technical grounds and one's mental image of set theory does not lead to such an axiom. To overcome some of the difficulties, Quine adopted similar measures to B-G (Bernays-Gödel) set theory. Like B-G, ML contains a bifurcation of classes into elements and non-elements. Sets can enjoy the property of being full objects whereas classes cannot. ML was obtained from NF by replacing (SCP) by two axioms, one for class existence and one for elementhood. The rule of class existence provides for the existence of the classes of all elements satisfying any condition Φ, stratified or not. The rule of elementhood is such as to provide the elementhood of just those classes which exist for NF. Therefore, the two axioms of comprehension of ML are The axiom of comprehension by a set:

  (1) $\exists y \forall x(x \in y \Leftrightarrow \Phi(x))$, where $\Phi(x)$ is a stratified formula with set variables only in which y does not occur free.

  The axiom of impredicative comprehension by a class:

  (2) $\exists y \forall x(x \in y \Leftrightarrow \Phi(x))$, where $\Phi(x)$ is any formula in which y does not occur free.

  ML was liked both for the manipulative convenience we regain in it and the symmetrical universe it furnishes. It was however proved subject to the Burali-Forti paradox — The well ordered set Ω of all ordinals has an ordinal which is greater than any member of Ω and hence is greater than Ω. Our description above of Russell's type theory, ZF set theory and Quine's NF and ML, has been brief, but should suffice to convince the reader of the need to have as many sets as one can. It has been argued by those who favour the iterative conception of set that we do not need self-application ([Boolos 1971]). But we have seen the necessity of type-free theories and the development of many type free systems such as Feferman's (in [Feferman 1979] and [Feferman' 1984]). Kripke's work on the theory of Truth [Kripke 1975] is further evidence that we should not rule out self referential statements and that we must look for a theory which allows for it. Gödel's work and especially his proof of the incompleteness theorems, showed that self-referential statements are as legitimate as arithmetic. Natural language is full of self-reference and self-application like: *There is nothing more beautiful than beauty.* All this points to the need for as many sets as possible, including sets that belong to themselves. All the above set theories reject the impredicative specifications and assumptions of classes and class existence, except ML which assumes impredicative clauses due to axiom (2) above. Also, from above, iterative sets are well founded, but NF and ML are non well founded. where a set a is non well founded iff $(\exists a_0, a_1, \ldots)(\ldots \in a_{n+1} \in a_n \in a_1 \in a_0 \in a)$. Now one can prove that (FA) implies the existence of only well founded sets.

- *Non well founded sets* Can we exchange (FA) for another which allows non well founded sets? Would this axiom remain consistent with or independent of other axioms? The answer is yes and many people have worked on various Anti Foundation Axioms ([Aczel 1984]). But what is the (AFA)? In his account, Aczel looks at sets in terms of pictures, where a picture of a set is an apg (accessible, pointed graph) with a decoration d such that d (the node) = the set itself and d: Nodes $\rightarrow$ sets where $d(n) = \{dn' : n \rightarrow n'\}$. For example, $\Omega\bullet$ is an apg and is a decoration of $\Omega = \{\Omega\}$. Ω will exist due to the anti foundation axiom, where

  (AFA) Every graph has a unique decoration.

  As a corollary of (AFA), one can prove that non well founded sets exist. In fact with (AFA), all possible non well founded sets exist. (AFA) is consistent with ZF, and we do not get the paradoxes with it. This is because it is not (FA) which was responsible for avoiding contradictions but it was the separation axiom. In fact, here the same proof as above will hold for the avoidance of Russell's paradox in ZF, and we have seen when we explained how Russell's paradox is avoided in ZF (section 2.1.2), that we only used the separation axiom and no mention was made for (FA) or (AFA).

2.1.3 Altering the logic

- *Rejection of the law of excluded middle* The paradox we faced was of the form:

  $(x \in x) \Leftrightarrow \neg(x \in x)$. Clearly the paradox can be avoided by dropping the assumption of LEM that any one place predicate either applies to a given object or does not. Note that here we can stick to two valued logics and that this system is not necessarily intuitionistic. If we go back to the example of impredicative specification given at the beginning of this section, according to this approach we can assume the existence of R, the set of all elements which do not contain themselves. What we cannot do though is assume that we have either $(R \in R)$ or $\neg(R \in R)$.

- *Many valued logics* $(x \in x) \Leftrightarrow \neg(x \in x)$ would not be contradictory if a consistent set of truth values was chosen. Consider as an illustration a three valued logic where the truth values are 0 (truth), 1(false) and u(undefined). The above sentence is not contradictory for we associate with $(x \in x)$ the value u and we define in the semantics that the negation of u is u. Therefore $u \Leftrightarrow \neg u$ is not contradictory and the paradox is avoided. Note here that there are many three valued interpretations and that the status of u varies from one interpretation to another. For some, u acts as not yet known, for others it is undefined. If we take the view that u is not yet known then we can order our models according to the state of our knowledge. Knowledge is cumulative whereas ignorance is not. What we know up to a stage, will always remain known after that stage, but we will also know more things. Domains looked at in this way are ordered and the fixed point theorem is applicable; this enables the construction of the limit model which is a model of the limit of our knowledge.

- *Frege structures* Frege structures are not only solutions to the problem of model existence, but are also systems of set theory in their own right: they single out that part of Frege's theory which is consistent. Frege structures could be classified as a restriction of logic, and they free Frege's notion of set from the paradox in the following way: the logical constants can apply to any object, but the result will never be a truth value unless the object itself was a proposition. The condition $x \in x$ is not necessarily a proposition and so $(x \in x) \Leftrightarrow \neg(x \in x)$ is not contradictory. In fact, the axioms of a Frege structure only enable one to derive propositions from previously known ones. x here is however arbitrary and so no deduction will give us $x \in x$ to be a proposition. The logic is weak in this way: the logical constants still apply to any object as with Frege but the result is a truth-value only if the object itself is one. With Frege this was not the case; he had the operator $-$ (which stands for content) and which gives the content of each object. So $-A$ is always a truth value whether or not the object A itself was a truth value. All the other logical constants in Frege's theory were applied to the content of the object and so always resulted in a truth-value. So in particular -||- A (not A) is always a truth value whether or not A was. Realising this about Frege's theory, Aczel reduced the logic to a weaker one where the logical constants only give truth values for truth values. In Aczel's Frege structures, the axiom (Vb) is not rejected. In fact the whole of axiom (V) is proven as a theorem in Frege structures and does not need to be asserted as an axiom as with Frege. Also, each concept has an extension, and decidable sets (the extensions of decidable concepts) are objects to which LEM applies. In a Frege structure you can prove that a set belongs to itself, (take $R = \{x : (x = x)\}$) and so it seems quite convenient to think of Frege structures as models for nominalisation. Before we move on, we give a summary of the work that was carried out by Feferman in the foundations of set theory. This is because Feferman's work investigates all of these restrictions (i.e. restricting the axioms, the logic or the language) and plays a crucial role in the area of nominalisation.

2.1.4 Feferman and the foundational issues

Feferman, in many of his papers, has worked on the question of the paradoxes and the possible solutions. He investigated for instance in [Feferman' 1984] the strategies of restricting the axioms, the logic or the language. He also investigated in [Feferman 1979] a theory T_0 which I believe is worth more attention than it has received. Feferman's T_0 was a formulation of Bishop's constructive mathematics, as are the theories of Martin-Löf's and Myhill. Yet Martin-Löf's is the theory which has been mostly used by Computer Scientists because it is more related to notions such as computation, program specifications and constructive proofs. Maybe it is the presence of canonical/noncanonical elements in Martin-Löf's theory and the notion of types which are very attractive to computer scientists. Yet I believe that Feferman's theory is simpler mainly because it is more flexible so that we do not commit ourselves to particular typing as we do with Martin-Löf's type theory. Of course here there is no room to discuss either T_0 or any other of Feferman's theories which avoid the paradoxes by various means. We must still however introduce the comprehension principles that Feferman uses in two of his theories. In T_0, the comprehension principle is restricted to *elementary* formulae where a formula is *elementary* if it is both stratified and has no bound class variables. Hence the principle looks like:

(ECA) $(\exists X)(\{x : \Phi(x, y, z)\} = X \wedge \forall x(x \in X \Leftrightarrow \Phi(x, y, z)))$, where $\Phi(x, y, z)$ can only be an elementary formula.

T_0 was a constructive theory. Feferman, before T_0 had investigated the use of full classical logic. Yet the paradox is avoided by having positive and negative formulae. The membership relation is now split into two partial predicates $\in$ and $\in'$ with the axiom:

Dis($\in, \in'$) $\neg(x \in \{u : \Phi(u, y_1, \ldots, y_n)\}) \wedge x \in' \{u : \Phi(u, y_1, \ldots, y_n)\})$

The comprehension principle is then divided into two comprehension principles: one for the positive formulas and the other is for the negative formulas as follows:

(CA)(+/-)

- $x \in \{u : \Phi(u, y_1, \ldots, y_n)\} \Leftrightarrow \Phi_+(x, y_1, \ldots, y_n)$
- $x \in' \{u : \Phi(u, y_1, \ldots, y_n)\} \Leftrightarrow \Phi_-(x, y_1, \ldots, y_n)$

Now of course Russell's paradox is avoided here because if we take $R = \{x : \neg x \in x\}$, then $R \in R \Leftrightarrow (\neg R \in R)_+ = (R \in R)_- = R \in' R$.

These are two of the ways that Feferman uses to avoid the paradoxes. However none of them as we see has a full comprehension principle, whereas Frege structures provide us with a full one.

2.2 Effects of set theory on nominalisation

2.2.1 Language and nominalisation

The reform of set theory by following the route of altering the language was based on the vicious circle principle, and resulted in Russell's theory of types. The language here becomes typed and the ladder of types has to be climbed step by step. Russell's theory of types was made simpler by Church and this is essentially the language used by Montague (in [Thomason 1974]) as an application to natural language. However, Montague did not himself deal with nominalisation and his account is very problematic from the nominalisation point of view. There have been few attempts at dealing with nominalisation within the Montague tradition. Examples are Carlson's work and Parson's floating types (in [Parsons 1979]). The main problem with Montague semantics is the typing constraints and the existence of the function f which has to associate once and for all the syntactic type of each syntactic category. This could be dealt with by changing the function f, but the approach is cumbersome and leads to difficulties.

2.2.2 Axioms and nominalisation

In ZF, we cannot have a set that contains itself and hence ZF is not suitable for nominalisation. NF or ML contain sets that belong to themselves, and so they should be promising candidates for the semantics of nominalisation. In fact they have already been applied by Cocchiarella who altered the system of non-standard second order logic shown in 1.1 to obtain two systems which he proved to be equivalent to NF and ML respectively ([Cocchiarella 1986]). The two systems are as follows:

1. **Altering (CP*)** Here, the paradox is avoided by restricting the formulae in (CP*) to stratified formulae (see previous section). In (CP*), Cocchiarella does not take F_n to be simply free in F, but imposes in addition the constraint that the whole bivalence be stratified. To return to our example, $[F(x) \Leftrightarrow \exists G[x = G \wedge \neg G(x)]]$ is not a stratified formula and so the comprehension principle cannot assure us of the existence of the predicate F, and hence there is no contradiction.

2. **Altering (A*)** Instead of altering (CP*), we alter (A3*) to (A3**) where

   (A3**) $\forall x \exists y(x = y)$.

   We then have to add $(a = a)$ as an axiom and replace (λ-CONV*) to:

   (E/λ-CONV*) $[\lambda x_1, \ldots x_n \Phi](a_1, \ldots, a_n) \Leftrightarrow \exists x_1, \ldots x_n (a_1 = x_1) \wedge \ldots \wedge (a_n = x_n) \wedge \Phi)$ where no x_i occurs free in any a_j, for $1 \leq i, j \leq n$.

   Note here that because of the elimination of (A3*), we can no longer prove the theorem $\forall x \Phi \Rightarrow \Phi(a/x)$. Therefore, we cannot substitute F for x in the special instance of (CP*) and so we cannot derive the paradox. The disadvantages of Cocchiarella's two systems are that the models are not easy to imagine.

2.2.3 Logic and nominalisation

The last category is the use of non-standard logics. Take for instance the use of a three-valued logic, rather than the classical two-valued one. $F(F) \Leftrightarrow \neg F(F)$ would not be inconsistent any more, for we can give $F(F)$ the value u (undefined) and in the interpretation of $\neg$ and $\Leftrightarrow$, we take: $\neg u \Leftrightarrow u$. This solution has been applied to nominalisation by Turner ([Turner 1984, 1987]). Turner used three valued logics and this allowed him to have an untyped language which could deal with nominalisation without falling into the paradox. This approach has been successful as far as predication is concerned, for one can nominalise all formulae. However it has a problem with quantification, since it is only to quantify over ideal elements (i.e. the limits of the finite ones as we shall see in part II). We have talked about the set theoretical approaches that have been offered. We looked at the theory of types and nominalisation and although we did not claim it was impossible to work out a theory of nominalisation based on Montague's semantics, we did say that it was difficult and cumbersome. We recall here that Russell's theory of types was unsatisfactory and so other theories came into being. The same applies to nominalisation, for Turner's and Cocchiarella's systems are less problematic than Montague's approach, because systems like NF and ML, or logics which are non-standard, were better attempts to provide a system without paradox than Russell's theory of types. Our criticism of Cocchiarella is that his models are difficult to imagine. It seems therefore that all the theories of nominalisation that have been worked out so far face some problems. There still are many solutions for set theory that have not hitherto been applied to the semantics of nominalisation, two of these being the notions of Frege structures and non well founded sets. It seems at this stage that most of the disadvantages of the theories that have been worked out so far can be circumvented by the use of these two notions. The use of Frege structures will allow us to keep to two-valued logic; also we can quantify over all our nominalised items, which is another advantage. The use of non well founded sets on the other hand, seems to be natural in the sense that they model the self reference that might be involved in nominalisation. For example, we might consider *nice* to be a solution to the equation $x = \{x\}$ and hence we get *nice is nice* to be true.

3 Solution to model existence

The problem discussed in 1.2 is not specific to nominalisation. It is the problem of finding models of the λ-calculus. Therefore I shall start by describing some of those models, and then I shall discuss how they have been used for the semantics of nominalisation.

3.1 λ-calculus and its models

We can forget about the formal axiomatisation of the λ-calculus with logic on the top of it and just remember that the λ-calculus with logic is a formal system which has 2 important operations: *abstraction* and *application* together with *λ-conversion*. Until recently, models of the λ-calculus have been problematic: do they really exist, and what are they like? One answer can be that the model itself is a structure which has two operations (abstraction and application); but this is an unsatisfactory answer. First, we could abstract the formula $\neg P(x)$ and then apply the abstract to itself which would yield Russell's paradox. Second, not every structure which has the two operations can be a model of the λ-calculus. Take for instance any combinatory algebra (which has K, S and '.'). We could prove in a combinatory algebra that the axiom of abstraction $(\exists F)(\forall y_1), \ldots, (\forall y_n)[F(y_1, \ldots, y_n) = A]$ holds, but that does not mean that the combinatory algebra is a model of the λ-calculus. It will be if we consider the extensional λ-calculus, but in the absence of extensionality we will have many choices for the function F in the axiom of abstraction and so the structure cannot be a model. What we should really require from the model is that if two wffs are equivalent or convertible in the λ-calculus then their values in the model must be the same. The other problem with defining models of the λ-calculus is that some λ-terms denote functions and so they have to take the elements of the structure M itself as argument. But again they themselves are terms and must take elements of M as values. We could take *term models* as models of the λ-calculus. Term models are just trivial formulations because all they do is translate the syntax step by step. Two other formulations of models are *environment models* and *combinatory models*. The environment models include in them two embedding functions Φ and Ψ which belong to $D \longrightarrow [D \longrightarrow D]$ and $[D \longrightarrow D] \longrightarrow D$ respectively. $[D \longrightarrow D]$ is not the set of all functions and it usually is the case that certain mathematical properties play a role in choosing $[D \longrightarrow D]$. Usually, $[D \longrightarrow D]$ is the set of all the continuous functions and is closed under the standard operations (such as composition, abstraction, application, ...). The combinatory model is exactly the combinatory algebra we talked about above but with the very important element ε which obeys some axioms. What ε does is to single out the functional part of every element. In the presence of extensionality we do not need ε and that is why in the case of extensionality, combinatory algebras are models of the λ-calculus. Both environment models and combinatory models are equivalent to each other and for a proof of this, the reader is referred to [Meyer 1981]. These are not the only kinds of models provided for the λ-calculus. The two kinds of models cited above together with the term models are algebraic, there are others which have a built-in structure. (It is easy to work with such models as one does not get involved with the cumbersome syntax). The two main models that I shall talk about throughout are: *Scott domains* and *Frege structures.*

3.2 Scott domains

We will be concerned with Scott domains[4] built as semantic domains where a semantic domain is a domain D with a binary relation $\subseteq$ such that
(i) D has a bottom element u satisfying $(\forall x \in D)[u \subseteq x]$.
(ii) $\subseteq$ is a partial ordering on D
(iii) every ω-sequence has a least upper bound in D where

1. An ω-sequence is a sequence $(x_n)_{n\in\omega}$ of elements of D such that $(\forall n \geq 0)[x_n \subseteq x_{n+1}]$.

2. An element d in D is the least upper bound of a subset X of D, iff

   - $(\forall d' \in X)[d' \subseteq d]$.
   - $(\forall d' \in D)[(\forall x \in X)[x \subseteq d'] \Longrightarrow d \subseteq d']$

We denote the least upper bound of $(x_n)_{n\in\omega}$ by $\cup_{n\in\omega} x_n$ and when no confusion occurs, we write $\cup x_n$. Basic to Scott domains is the notion of a *continuous function* where a function f from a semantic domain D into another semantic domain D' is *continuous* iff (for each ω-sequence $(d_n)_n \in D)[f(\cup d_n) = \cup f(d_n)]$[5]

New domains are built out of old ones using the following three notions: Let $(D_1, \subseteq_1)$ and $(D_2, \subseteq 2)$ be two semantic domains

1. Define $D_1 + D_2 = \{(d_i, i) \text{ such that } d_i \in D_i*\} \cup \{u\}$ where $(u \notin D_1 \cup D_2)$, and $(\forall d = (d_i, i), d' = (d'_j, j) \in D_1 + D_2)[d \subseteq d' \Longleftrightarrow (d = u \text{ or } (i = j \text{ and } d_i \subseteq_i d'_j))]$.

2. Define $D = D_1 \times D_2 = \{< d_1, d_2 > \text{ where } d_1 \in D_1 \text{ and } d_2 \in D_2\}$
   and $(\forall < d_{10}, d_{20} >, < d_{11}, d_{21} > \in D_1 \times D_2)$,
   $[< d_{10}, d_{20} > \subseteq < d_{11}, d_{21} > \Longleftrightarrow d_{10} \subseteq_1 d_{11} \text{ and } d_{20} \subseteq_2 d_{21}]$

3. Let $[D_1 \longrightarrow D_2]$ be the set of continuous functions from the domain $(D_1, \subseteq_1)$ to the domain $(D_2, \subseteq_2)$. Define a binary relation on $[D_1 \longrightarrow D_2]$ as follows:
   $(\forall f, g \text{ in } [D_1 \longrightarrow D_2])[f \subseteq g \Longleftrightarrow (\forall d \in D)[f(d) \subseteq_2 g(d)]]$.

Lemma 3.1 *$(D_1 + D_2, \subseteq), (D_1 \times D_2, \subseteq)$ and $([D_1 \longrightarrow D_2], \subseteq)$ are semantic domains.*

We are interested in domains E which satisfy an equation of the form: $E \approx [E \longrightarrow E]$. We define B the set of truth values, i.e. $B = \{0, 1, u_0\}$ where $u_0 \subseteq 1$ and $u_0 \subseteq 0$ (B is a semantic domain). We build our domain E by building a sequence of domains (by induction). We start with $E_0 = B$ and build $E_{n+1} = B + [En \longrightarrow E_n]$ for $n \geq 0$ such that for all n, E_n is a semantic domain. We would like, however, to relate all those domains with an ordering relation and find the limit of such a sequence. This limit is going to be the required E. We start with some definitions:

[4]see [Barendregt 1981] and [Barendregt' 1981]
[5]From now on, we will use $D*$ to denote $D - \{u\}$.

Definition 3.2 *A projection pair of D_1 on D_2 is a pair $< \Phi, \Psi >$ such that: $\Phi : D_1 \longrightarrow D_2$, $\Psi : D_2 \longrightarrow D_1$ and*

- *Φ, Ψ are both continuous,*
- $(\forall x \in D_1)[\Psi(\Phi(x)) = x]$
- $(\forall x \in D_2)[\Phi(\Psi(x)) \subseteq_2 x]$

For each $n \geq 0$, we define a projection pair $< \psi_n, \Phi_n >$. The aim of each Φ_n is to embed E_n into E_{n+1}, whereas Ψ_n is a surjection from E_{n+1} to E_n. Our construction of $(\Phi_n)_{n\in\omega}$ is done by induction as follows:

$\Phi_0 : E_0 \longrightarrow E_1$ such that $\Phi_0(x) = x \in B* \hookrightarrow x, u_1$[6]

$\Psi_0 : E_1 \longrightarrow E_0$ such that $\Psi_0(x) = x \in B* \hookrightarrow x, u_0$

Assume that Φ_n and Ψ_n have been defined such that $< \Phi_n, \Psi_n >$ is a projection pair of E_n on E_{n+1}, we build Φ_{n+1} and Ψ_{n+1} as follows:

$\Phi_{n+1} : E_{n+1} \longrightarrow E_{n+2}$ so that $\Phi_{n+1}(x) = x \in B* \hookrightarrow x, (x = u_{n+1} \hookrightarrow u_{n+2}, \Phi_n \circ x \circ \Psi_n)$

$\Psi_{n+1} : E_{n+2} \longrightarrow E_{n+1}$ so that $\Psi_{n+1}(x) = x \in B* \hookrightarrow x, (x = u_{n+2} \hookrightarrow u_{n+1}, \Psi_n \circ x \circ \phi_n)$

One can easily prove that $< \Phi_{n+1}, \Psi_{n+1} >$ is a projection pair of E_{n+1} on E_{n+2}. Now we construct a domain E_∞ which will contain all the E_n for $n \in \omega$.

$E_\infty = \{< f_n >: f_n \in En \text{ and } \Psi_n(f_{n+1}) = f_n\}$.

The $\subseteq$ on E_∞ is:

$(\forall < f_n >, < g_n > \in E_\infty)[< f_n >_{n\in\omega} \subseteq < g_n >_{n\in\omega} \Longleftrightarrow (\forall n \in \omega)[f_n \subseteq_n g_n]]$

Lemma 3.3 *$(E_\infty, \subseteq)$ is a semantic domain.*

Now we define application in E_∞. Let f, e be in E_∞ and define $f \bullet e = \cup f_{n+1}(e_n)$. Again the following proofs are left to the reader.

Lemma 3.4 *$\bullet : E_\infty \times E_\infty \longrightarrow E_\infty$ is continuous.*

Theorem 3.5 $(\forall f \in [E_\infty \longrightarrow E_\infty])(\exists X_f \in E_\infty)[(\forall e \in E_\infty)\ [f(e) = X_f \bullet e]]$.

Theorem 3.6 $E_\infty \approx [E_\infty \longrightarrow E_\infty]$[7].

3.3 Frege structures

Before launching into this section, let us introduce some convenient notation:

Notation 3.7 If f is a function of 2 arguments then we will sometimes write afb for $f(a, b)$. For example, we write $a \wedge b$ for $\wedge(a, b)$.

Until we give the exact definition of an **F**-functional, let us understand it to be a function which takes functions as arguments and returns functions as values.

Notation 3.8 $\mathbf{F_0^n}$ stands for: $\mathbf{F_0} \times \mathbf{F_0} \times \ldots \times \mathbf{F_0}, n$ times.

[6]The notation $b \hookrightarrow a_1, a2$ is to be understood as: If b then a_1 else a_2.

[7]Actually E_∞ is the least upper bound of the sequence of domains $(E_n)_n$.

Notation 3.9 *(Metalanguage abstraction)* For every expression $e[x_1, \ldots, x_n]$ of the metalanguage built up in the usual way from variables ranging over $\mathbf{F_0}$ and constants ranging over $\cup_n \mathbf{F_n}$, the expression $< e[x_1, \ldots, x_n]/x_1, \ldots, x_n >$ denotes the n-place function f: $\mathbf{F_0} \times \ldots \times \mathbf{F_0} \longrightarrow \mathbf{F_0}$ such that for each a_i in $\mathbf{F_0}$, $1 \leq i \leq n, f(a_1, \ldots, a_n)$ is the value of $e[a_1, \ldots, a_n]$, the expression e in which x_i has been replaced by a_i for $i = 1, \ldots, n$. For each expression $e[\xi_1, \xi_2, \ldots, \xi_n]$ of the metalanguage built in the usual way out of variables (ranging over $\mathbf{F_n}$ for $n \geq 0$) and constants (ranging over $\mathbf{F_n}$ for $n \geq 0$ and over $\mathbf{F}$-functionals), the expression $< e[\xi_1, \xi_2, \ldots, \xi_n]/\xi_1, \xi_2, \ldots, \xi_n >$ denotes the n-place function obtained by abstracting $\xi_1, \xi_2, .., \xi_n$ in e.

Notation 3.10 *(Def*)* If F is a 1-place $\mathbf{F}$-functional and $< e[x]/x >$ is in the domain of F, we write $Fxe[x]$ for $F(< e[x]/x >)$.

For example, $\forall : \mathbf{F_1} \longrightarrow \mathbf{F_0}$ and $\lambda : \mathbf{F_1} \longrightarrow \mathbf{F_0}$ are $\mathbf{F}$-functionals; we write $\forall < f(x)/x >$ and $\lambda < f(x)/x >$ as $\forall x f(x)$ and $\lambda x f(x)$ respectively.

It should be noted here that λ does not represent implication in logic. For implication we have another sign which is unrelated to λ. Furthermore, even though λ and $\forall$ have the same functional space, they are different. The first has the property that if we apply λf to a, we get $f(a)$, whereas the second satisfies that if f is a propositional function then $\forall f$ is a proposition. (Here we understand a *propositional function*, to be a function of the Frege structure which takes propositions as values, i.e. $f(x)$ is a proposition for every x.) In fact, $\mathbf{app}(\lambda f, a)$ makes sense but $\mathbf{app}(\forall f, a)$ does not. Moreover, if f is a propositional function than λf is a set and not a proposition. Furthermore, we will only work with models where the collection of sets and the collection of propositions are disjoint.

3.3.1 Informal introduction

The existing models of the λ-calculus did not deal with logic added on top of the λ-calculus, since once logic is added, consistency might be threatened. Also, if one constructs a theory which will have logic, λ-abstraction and predication, then one has to show the existence of the models of this theory. This is the work we find with Feferman for instance, yet his models are not tidy and clear. Hence one would like to have a clear idea of a model of the λ-calculus with logic on it, and Frege structure is such a model. However, such a construction was not obvious for a long time. It was initiated by Scott in [Scott 1975] yet the work was incomplete and hence such a model was not achieved. Then came the construction of Frege structures where simply the idea is to start from any model of the λ-calculus and build logic on top by inductively constructing two collections (of the possible propositions and the possible truths) and taking the limit of these two collections which actually draw the logic we now have on the top of the initially considered model of the λ-calculus. As it sounds, the process is quite simple, yet it depends on having a clear idea of the structure and on proving some theorems which will ensure the existence of the various logical connectives in the model considered. Now that logic has been constructed on the top of a model of the λ-calculus, we can consider the structure only in terms of its objects and functions. The objects include propositions and truths and the functions obey the condition that propositional functions can be projected in the domain of objects (i.e. as sets). Those sets can be applied to any object (hence we now have not only functional application such as $f(x)$[8], but also the application of one object to

[8]Note here that $f(x)$ means 'f applied to x' and that if one wanted to write 'f is an expression which depends on x', one would use $f[x]$. See Def* and the definition of metalanguage abstraction, especially the

another as in $\mathbf{app}(a,b)$), and set application to an object results in a proposition. This is the simple idea of a Frege structure. Next, the reader finds the various steps used to construct a Frege structure.

A Frege structure consists of a denumerably infinite number of collections $(\mathbf{F_n})_{n\geq 0}$ such that:

1. $\mathbf{F_0}$ is a collection of objects which has three very important subcollections **PROP**, **TRUTH** and **SET** where, **PROP** is a subcollection of $\mathbf{F_0}$ which can be thought of as the collection of propositions and **TRUTH** is a subcollection of **PROP** which can be thought of as the collection of true propositions. **SET** is a subcollection of $\mathbf{F_0}$ which can be thought of as the collection of objects which are nominals of propositional functions.

2. For each $n > 0$, $\mathbf{F_n}$ is a collection of n-ary functions which take all their arguments in $\mathbf{F_0}$.

3. There is a set of **F**-functionals that operate over $(\mathbf{F_n})_{n\geq 0}$ and which ensure important closure properties on $(\mathbf{F_n})_{n\geq 0}$. For example:

   $\forall : \mathbf{F_1} \longrightarrow \mathbf{F_0}$ is a functional such that: If f in $\mathbf{F_1}$ is a propositional function then $\forall f$ is in **PROP** and $\forall f$ is in **TRUTH** iff $f(a)$ is in **TRUTH** for each a in $\mathbf{F_0}$

   $\lambda : \mathbf{F_1} \longrightarrow \mathbf{F_0}$ and $\mathbf{app} : \mathbf{F_0} \times \mathbf{F_0} \longrightarrow \mathbf{F_0}$ are two other functionals which possess the very important property: $\mathbf{app}(\lambda f, a) = f(a)$ for every a in $\mathbf{F_0}$ and every f in $\mathbf{F_1}$. Note here that **app** is different from *real application*. In fact, $\mathbf{app} : \mathbf{F_0} \times \mathbf{F_0} \longrightarrow \mathbf{F_0}$ whereas *real application*: $\mathbf{F_1} \times \mathbf{F_0} \longrightarrow \mathbf{F_0}$. Still, **app** and real application are related in that $\mathbf{app}(\lambda f, a) = f(a)$. **app** is really introduced to capture that when turning a function f into an object λf, we preserve the information of f's functionality. That is λf applies to an object a and gives the same result as applying f to a.

4. $(\mathbf{F_n})_{n\geq 0}$ is *super explicitly closed*: i.e. for each expression $e[\xi_1, \xi_2, .., \xi_n]$ of the metalanguage built in the usual way out of variables (ranging over $\mathbf{F_n}$ for $n \geq 0$) and constants (ranging over $\mathbf{F_n}$ for $n \geq 0$ and over **F**-functionals), the n-place function denoted by $< e[\xi_1, \xi_2, \ldots, \xi_n]/\xi_1, \xi_2, \ldots, \xi_n >$ is an **F**-functional[9].

Now that we have some idea of the structures' form, let us try to give an intuitive picture. A Frege structure is a collection of both objects and functions (which are distinct) where we can map any function f into an object a and this object will preserve some of the properties of the function. For instance if the function f is a propositional function, then the nominal of the function, λf, is an object which belongs to the category **SET** (recall here that $\forall f$ is an object which belongs to the category **PROP** and that we will be interested in models where $\mathbf{PROP} \cap \mathbf{SET} = \emptyset$). Moreover **SET** contains only those objects which are nominals of propositional functions. Thus, if a is in **SET** then there must be a k-ary propositional function f such that $a = \lambda_{\mathbf{0}}^{\mathbf{n}} f$, where: $\lambda_{\mathbf{0}}^{\mathbf{1}}$ is λ and maps 1-ary functions into objects (i.e. into $\mathbf{F_0}$); $\lambda_{\mathbf{0}}^{\mathbf{2}}$ maps 2-ary functions into objects;... $\lambda_{\mathbf{0}}^{\mathbf{n}}$ maps n-ary functions into objects. By induction, we can define $\lambda_{\mathbf{m}}^{\mathbf{n}}$ which maps n-ary functions into $\mathbf{F_m}$.

It is natural to ask whether the intersection of **SET** and **PROP** is empty or not; some elements of **PROP** are elements of elements of **SET**, yet the intersection between **SET** and

notation $e[\xi_1, \xi_2, \ldots, \xi_n]$ and $< e[\xi_1, \xi_2, \ldots, \xi_n]/\xi_1, \xi_2, \ldots, \xi_n >$.

[9]This means that Frege structures are closed under composition, projection, etc.

PROP is not certain to be empty. Take for example an element a of **PROP** and consider b to be the set $\{x : (x = a)\}$. Obviously b is in **SET** because $< (x = a)/x >$ is a propositional function, and we have a is in b. **SET** and **PROP** are not necessarily disjoint. Take for example, a in **PROP** and assume the following principle:

$\forall x(\mathbf{app}(t, x) = \mathbf{app}(t', x)) \Rightarrow t = t'$.

Define moreover $\| \|$: to be the counterpart of λ. That is $\|a\|(x) = \mathbf{app}(a, x)$. Then $\lambda\|a\| = a$ can be seen as follows:

$\forall x, \mathbf{app}(\lambda\|a\|, x) = \|a\|(x) = \mathbf{app}(a, x)$.

Therefore $\forall x \mathbf{app}(\lambda\|a\|, x) = \mathbf{app}(a, x)$ and hence $a = \lambda\|a\|$. Now, if $\|\lambda a\|$ is a propositional function, then $\mathbf{SET} \cap \mathbf{PROP}$ is not empty. The question here is whether $\|a\|$ is a propositional function when a is a proposition. We do not need to answer this question here and independently of whether **SET** and **PROP** are disjoint, there is an important relation between them which is the following; they both have strong links with propositional functions. Let us consider 1-ary functions to illustrate the argument and take a propositional function f. For any object a, $f(a)$ is a proposition (i.e. is in **PROP**). λf is a set and $\mathbf{app}(\lambda f, a) = f(a)$. We can always jump from propositional functions to sets (and from sets to propositions). But we can also jump from sets to propositional functions. Take the operation $\| \|_1$ defined as: For each object a of the Frege structure, $\|a\|_1 =< \mathbf{app}(a, x)/x >$. Obviously for each $a, \|a\|_1$ is in $\mathbf{F_1}$ and if, in particular, we take a to be in **SET** (say a is λf) then we have that $\|a\|_1 = \|\lambda f\|_1 = f$. Therefore we have an equivalence between sets and propositional functions; each set corresponds to a propositional function and each propositional function corresponds to a set. This is important and it is this strong link that I am trying to emphasise between **SET** and propositional functions. Note that for each n, this bivalent path holds between P$\mathbf{F_n}$ and **SET**, through $\lambda_\mathbf{0}^\mathbf{n}$ and $\| \|_n$ where again we have $\mathbf{app_n}(\lambda_\mathbf{0}^\mathbf{n} f, \hat{a}) = f(\hat{a})$, for $\hat{a}$ in $\mathbf{F_0^n}$, and f in $\mathbf{F_n}$. The functionals $\lambda_\mathbf{0}^\mathbf{n}$, $\mathbf{app_n}$ and the operation $\| \|_n$ (the counterpart of $\lambda_\mathbf{0}^\mathbf{n}$) could be defined recursively as follows:

Take $\|a\|_n =< \mathbf{app_n}(a, x), x >$, and $(\lambda_\mathbf{n}^\mathbf{n+1} f)(\hat{a}) = \lambda(< f(\hat{a}, x)/x >)$, and assume $\lambda_\mathbf{n}^\mathbf{n+m} f$ has been defined. Then take $\lambda_\mathbf{n}^\mathbf{n+m+1} f = \lambda_\mathbf{n}^\mathbf{n+1}(\lambda_\mathbf{n+1}^\mathbf{n+m+1} f)$.

$\mathbf{app_n}$ is also defined by recursion where: $\mathbf{app_1} = \mathbf{app}$ and assume we have defined up to $\mathbf{app_n}$. Then $\mathbf{app_{n+1}}(a, b, b') = \mathbf{app_n}(\mathbf{app}(a, b), b')$. One can prove that $\mathbf{app_n}(\lambda_\mathbf{0}^\mathbf{n} f, \hat{a}) = f(\hat{a})$ for each n in ω and $\hat{a}$ in $\mathbf{F_0}$n.

So in a Frege structure, like in any (Fregean) calculus of functions and objects which has variable-binding (and application) through abstraction (resp. application) operators one can take functions into objects. In short, we do not lose information by mapping the function into an object. We can switch back from objects to functions using $\| \|_n$, the inverse operator of $\lambda_\mathbf{0}^\mathbf{n}$ where we have the following theorem: $\|\lambda_\mathbf{0}^\mathbf{n} f\|_n = f$ for any n-ary propositional function f.

The ability to switch back and forth between objects and functions is not the only important aspect of the program; the presence of **PROP**, **TRUTH** and of a logic in a Frege structure is also crucial. The logic is built in a way that allows us to talk about truths and propositions without falling into any contradictions. There are other accounts which can do this of course such as Martin-Löf's type theory with his judgements: *A type* and $a : A$. The difference here is that in a Frege structure, we have combined both the elegance of a simple structure (objects and functions without the typing strategies) together with the presence of a consistent logic.

3.3.2 Frege structures as models and comparison with Scott domains

λ-structures are models of the λ-calculus in an obvious way. For just take the interpretation of terms as follows over a defined Frege structure $\mathbf{F}$, where g is an assignment function which takes variables into objects of $\mathbf{F_0}$:

$[[x]]_{g,F} = g(x)$

$[[MN]]_{g,F} = \mathbf{app}([[M]]_{g,F}, [[N]]_{g,F})$

$[[\lambda x M]]_{g,F} = \lambda < [[M]]_{g[a/x]}, F/a >$.

Now it is easy to show that this interpretation has the property that:

$\lambda \vdash M = N \Longrightarrow [[M]]_{g,F} = [[N]]_{g,F}$.

Therefore, Frege structures are models of the λ-calculus and in turn we know that they solve the second problem. For the remainder of this section, we shall concentrate on the comparison between both Scott domains and Frege structures as models, and hence help justify our claim that Frege structures are better candidates for the semantics of natural languages than Scott domains. On Scott domains, one has a topology (Scott topology based on a partial ordering relation) and two special elements Top and Bottom. (Bottom is less than all the other elements and Top is greater than all of them.) We shall see in Part II that this ordering relation, together with the existence of Bottom and the requirement that the functions be continuous, make Scott domains problematic for the semantics of natural languages. On Frege structures, however, we have no ordering and no requirement on the continuity of functions.

What we have in a Frege structure is a collection of objects $\mathbf{F_0}$ together with, for each n, a collection $\mathbf{F_n}$ of n-ary functions which take elements of $\mathbf{F_0}$ as arguments and return elements of $\mathbf{F_0}$ as values. But although we do not consider all possible functions to be elements of the Frege structure, we still consider only structures which are explicitly closed. This explicit closure imposes the existence of some necessary functions such as projections, constants, etc, and requires the closure of our structure under some important functional operations such as composition. We have both constants and variables for functions, but the functionality on a Frege structure does not stop at those first order functions; we also have functionals. However, whereas for functions our language contains both variables and constants, for functionals it only contains constants.

One should bear in mind that none of the collections **PROP**, **TRUTH** or **SET** is internally definable. Intuitively, we say that a collection χ of objects is internally definable if we can talk about it through the object language and not just the metalanguage. An example of a collection which is not internally definable is the collection of truths in a theory which contains names for its wffs. If this collection was internally definable, then there must be a predicate T such that for any object a, $T(a)$ is true iff a is true. But according to Tarski, a theory cannot contain its own truth predicate (in the object language) without falling into inconsistency and therefore T is a predicate of the metalanguage. Now if we want to talk about truth in this metalanguage then again we have to have a truth predicate T' in the meta-metalanguage and this process iterates. Just as T is not an element of the object language in Tarski's approach, so inside a Frege structure the collection of truths is not internally definable. Aczel gives a more formal definition of internal definability and considers a collection χ of objects in $\mathbf{F_0}$ to be internally definable in the Frege structure iff there exists a propositional function C in $\mathbf{F_1}$ such that the following holds:

(**) For any object a in $\mathbf{F_0}, C(a)$ is in **TRUTH** iff a is in χ.

It might be clearer if we set **FALSE** = **PROP** $\setminus$ **TRUTH**, and then replace (**) by the following:

(***) For any object a in $\mathbf{F_0}, C(a)$ is in **TRUTH** iff a is in χ and $C(a)$ is in **FALSE** otherwise.

Some might find it easier to draw a contrast with the following schematic definition, where C is not a propositional function:

(****) For any object a in $\mathbf{F_0}$, $C(a)$ is in **TRUTH** iff a is in χ and $C(a)$ is in $\mathbf{F_0}$ otherwise.

(***) makes χ decidable, while (****) only makes semi-decidable. It may seem unfortunate that the collection of **TRUTH**s is not internally definable, but it is essentially this that provides Frege structures with consistency. Notice that since elements of **SET** are the nominalisations of propositional functions, we have no way of talking about the nominalised items internally and **SET** is not internally definable. Moreover it may also seem that we will encounter a problem in defining second order quantifiers. I hope that it will become clear throughout the work that the inability to internally define quantifiers does not have any serious effects. On the contrary, we keep to simplicity while being able to formalise many concepts within the theory. The undefinability of **PROP** and of **SET** is due to the undefinability of **TRUTH**. The collection of propositions is not internally definable, for if it were (through a predicate P) we would find that **TRUTH** is also internally definable (through the propositional function $< P(x) \wedge (P(x) \rightarrow x)/x >$, which stands for a function in $\mathbf{F_1}$). That **PROP** is not internally definable implies that **SET** is not either. This is because if S were a propositional function in $\mathbf{F_1}$ internally defining **SET** then $< S(y/x)/x >$ is a propositional function in $\mathbf{F_1}$ internally defining the collection of propositions. Note also that, for each n, $\mathrm{P}\mathbf{F_n}$ (the collection of n-ary propositional functions) is not internally definable. For if it were, we would get that the collection of propositions would also be. The proof here needs an extension of the definition of internal definability so that instead of having a function we have a functional. Let us return to the comparison of Frege structures with Scott domains. Frege structures do not have any ordering or continuity problems and their restricted logic would allow us to solve the problems of Scott domains (and of Cocchiarella). But of course the solving part is not going to be easy. We have to do something about the non-internal definability of **SET**. There are a few ways to go here: we have to either see how the function domains (as with Scott) could be built inside Frege structures, or else show that we do not need second level quantifiers and therefore the problem does not arise. Now the word inside brings an uncomfortable feeling — especially after we pointed out that all the interesting collections are not internally definable. I assure the reader however that this difficulty is only temporary and that we can always find solutions to the problem. It is important for the reader to know that a Frege structure can be built on the top of a model where continuity and ordering play a very important role (such as E_∞). However the way quantifiers are constructed on a Frege structure using the fixed point, is not based on the ordering relation, and so the problem that faced Turner in his work based on E_∞ (where quantifiers depended on the ordering relation - see Part II) is not faced by the quantifier treatment on a Frege structure. The fact that functions, but not functionals, can be mapped into $\mathbf{F_0}$ in a Frege structure is not a disadvantage, indeed it may even be seen as a virtue, since there appears to be no justification in natural language semantics for nominalizing expressions — for example determiners — which would require a formalisation as functionals. Also, in Frege structures we have more possible elements than we do in Scott domains. We have propositions, truths and sets which are all legitimate elements of the Frege structure. We could not talk about them internally but that is how it should be. Tarski's undefinability of Truth and Gödel's famous result make it impossible for us to be able to internally define any of these collections.

So, our inability to internally define any of these collections is not a weakness in comparison with Scott domains; Scott domains could not talk about them at all, and therefore can not be adequate for natural language semantics. If we try to extend Scott domains in a way that will allow us to talk about truths and propositions, we obtain Frege structures.

4 Acknowledgments

The author is indebted to the anonymous referees whose suggestions played a crucial role in improving both the form and the contents of the paper.

References

[Aczel 1980] P. Aczel, Frege structures and the notions of proposition, truth and set; in: J. Barwise, ed., *The Kleene Synposium*, Studies in Logic 101, North-Holland, NewYork, 1980, pp. 31-60.

[Aczel 1984] P. Aczel, *Non-well founded sets*, CSLI Lecture notes No 14, 1984.

[Aczel 1985] P. Aczel, Properties and propositional functions, privately circulated note, Manchester University, 1985.

[Barendregt 1981] H. Barendregt, *The lambda calculus: its syntax and semantics*, North Holland, 1981.

[Barendregt' 1981] H. Barendregt, The type free lambda calculus, *Handbook of Mathematical logic*, Ed. J.Barwise, North Holland, 1981, pp. 1091-1132.

[Barwise 1987] J. Barwise and J. Etchemendy, *The liar: An Essay on Truth and Circular Propositions*, Oxford University Press, Oxford, 1987.

[Beeson 1987] M. Beeson, *Foundations of constructive Mathematics*, Springer-Verlag, Berlin, 1987.

[Bealer 1982] G. Bealer, *Quality and concept*, Clarendon press, Oxford, 1982.

[Boolos 1971] G. Boolos, The iterative conception of set, *Journal of Philosophy LXVIII*, 1971, pp. 215-231.

[Church 1940] A. Church, A formulation of the simple theory of types, *Journal of Symbolic Logic 5*, 1940, pp. 56-68.

[Cocchiarella 1984] N. Cocchiarella, Frege's Double Correlation Thesis and Quine's set theories NF and ML, *Journal of Philosophical Logic 13*, 1984.

[Cocchiarella 1986] N. Cocchiarella, Conceptualism, Ramified Logic and Nominalised predicates, *Topoi 5*, 1986, pp. 78-87.

[Cocchiarella' 1986] N. Cocchiarella, Philosophical Perspectives on Formal theories of Predication, *Handbook of Philosophical Logic 4*, 1986.

[Feferman 1975] S. Feferman, A language and axioms for explicit Mathematics, *Algebra & Logic*, lecture notes in mathematics 450, 1975, pp. 87-139.

[Feferman 1979] S. Feferman, Constructive theories of functions and classes, Ed. M.Boffa, *Logic Colloquium '78*, North-Holland, Amsterdam, 1979, pp. 159-224.

[Feferman 1981] S. Feferman, Working foundations, A revised version of a paper presented to the workshop The present state of the problem of foundations of Mathematics, Florence, 1981.

[Feferman' 1981] S. Feferman, A theory of variable types, privately circulated note, Stanford University, 1981.

[Feferman 1982] S. Feferman, Inductively presented systems and the formalisation of Meta-Mathematics, Eds. D. Lascar and J. Smiley, *Logic Colloquium '80*, North-Holland, Amsterdam, 1982, pp. 95-128.

[Feferman 1983] S. Feferman, Intensional Mathematics, *Logic Colloquium '83.*

[Feferman' 1983] S. Feferman, Between constructive and classical Mathematics, *Logic Colloquium '83.*

[Feferman 1984] S. Feferman, Intensionality in Mathematics, *Symposium on intensions and set theory*, Meeting of the Pacific Division of the American Philosophical Association, 1984.

[Feferman' 1984] S. Feferman, Towards useful type-free theories I, *Journal of Symbolic Logic 49*, 1984, pp. 75-111.

[Fraenkel 1966] A. Fraenkel, *Set theory and Logic*, Addison-Wesley, U.S.A, 1966.

[Fraenkel 1973] A. Fraenkel, Y. Bar-Hillel and A. Levy, *Foundations of set theory*, North-Holland, Amsterdam, 1973.

[Frege 1970] G.Frege, *Translations from the philosophical writings of Frege*, Ed. P. Geach and M. Black, Basil Blackwell, Oxford, 1970.

[Kamareddine 1988] F. Kamareddine, A polymorphic type theory, with Ewan Klein, Talk given at Titisee conference on unification, 1988.

[Kamareddine 1989] F. Kamareddine, *Semantics in a Frege Structure*, PhD thesis, University of Edinburgh, 1988.

[Kleene 1952] S. Kleene, *Introduction to Metamathematics*, D.Van Nostrand co, Princeton, 1952.

[Kripke 1963] S. Kripke, Semantical considerations on Modal Logic, Ed. L. Linsky, *Reference and Modality*, 1963.

[Kripke 1975] S. Kripke, Outline of a theory of Truth, *The Journal of Philosophy LXXII*, 1975, pp. 690-716.

[Meyer 1981] A. Meyer, What is a model of the Lambda Calculus? Unpublished ms., M.I.T. Lab., Computer Science, 1981.

[Parsons 1979] T. Parsons, The theory of types and ordinary language, Eds. S. Davies and M. Mithun, *Linguistics, Philosophy and Montague Grammar*, 1979, University of Texas Press, Austin.

[Poincaré 1900] H. Poincaré, Du role de l'intuition et de la logique en mathematiques, C.R. du II Congr. Intern. des Math., Paris 1900, pp. 200-202.

[Quine 1969] W. Quine, *Set theory and its Logic*, Belknap Press, Harvard, 1969.

[Russell 1908] B. Russell, Mathematical logic as based on the theory of types, *American Journal of of Math. 30*, 1908, pp. 222-262.

[Scott 1976] D. Scott, Data types as Lattices, Technical Monograph PRG-5, *Siam Journal on Computing 5*, 1976, pp. 522-587.

[Scott 1975] D. Scott, Combinators and classes, Ed. Böhm, *λ-calculus and Computer Science theory*, Lecture notes in Computer Science 37, Springer, 1975.

[Thomason 1974] R. Thomason, *Formal Philosophy; Selected papers by Richmond Montague*, Yale University, 1974.

[Turner 1984] R.Turner, Three Theories of Nominalized Predicates, *Studia Logica XLIV2*, 1984, pp. 165-186.

[Turner' 1984] R. Turner, Nominalization and Scott's Domains II, *Notre Dame Journal of Formal Logic 26*, 1984.

[Turner 1987] R. Turner, A theory of properties, *Journal of Symbolic Logic 52*, 1987.