A Kripkean Solution to Paradoxes of Denotation

Casper Storm Hansen*

University of Aberdeen, Aberdeen, UK
casper_storm_hansen@hotmail.com

Abstract. Kripke's solution to the Liar Paradox and other paradoxes of truth (1975) is generalized to the paradoxes of denotation. Berry's Paradox and Hilbert and Bernays' Paradox are treated in detail.

Keywords: Paradoxes of Denotation, Kripke, Berry's Paradox, Hilbert and Bernays' Paradox, Fixed Point

1 Introduction

Priest has demonstrated (2002) that all of the semantical paradoxes share a common structure and has argued, that the solution to this class of paradoxes should therefore also be shared. According to him, this is a reason to reject Kripke's famous solution to the paradoxes of truth (1975), as it is indeed *only* a solution to these paradoxes and not to the paradoxes of denotation. In this paper I will show that this critique is misplaced. Kripke's solution can be generalized. I will just treat two of the paradoxes of denotation, namely Berry's and Hilbert and Bernays', but the approach can be applied to them all.

Berry's Paradox (Russell 1908) results from the definite description

> Berry's description: the least integer not describable in fewer than twenty syllables

which is a description of nineteen syllables. So the least integer not describable in fewer than twenty syllables is describable in only nineteen syllables.

Hilbert and Bernays' Paradox (originally presented in (Bernays 1939), natural language formulation in (Priest 2006)) also results from a definite description, namely this:

> Hilbert and Bernays' description: the sum of 1 and the reference of Hilbert and Bernays' description

If we let n be the reference of Hilbert and Bernays' description, then it also refers to $n+1$. As the reference of a definite description is unique, it follows that $n = n + 1$.

I will assume familiarity with Kripke's paper.

* I would like to express my gratitude to Vincent Hendricks for his encouragement and assistance in the work that lead to this paper, and to the three anonymous reviewers for their valuable comments.

2 Informal Presentation of the Theory

In Kripke's theory sentences become true and false in a recursive process, where a sentence is given a truth value when there is, so to speak, enough information to do so. For instance a sentence of the form "sentence S is true" is made true after it has been decided that S is true, false after it has been decided that S is false, and is left undecided as long as S is. And a disjunction is made true at such time as one of the disjuncts is, since the information about the (eventual) truth value of the other disjunct is irrelevant.

To formulate Berry's description we need two linguistic resources that are not in the formal language of Kripke's paper: The ability to form definite descriptions and a binary predicate expressing that a given term refers to a given object. But when we equip the formal language with these resources, the principle in Kripke's theory can be transfered to these. We let a definite description refer to a given object, when it is determined that this is the unique object which satisfies the description. And if it is decided at some point in the iterative process that there are no objects or more than one object which satisfy the description, it is decided that the definite description fails to refer. And a sentence of the form "T refers to O" is made true, if at some point it is decided that the term T indeed does refer to the object O, and made false, if it is decided that T refers to something different from O or fails to refer.

In Kripke's theory the Liar Sentence "this sentence is false" is neither true nor false, it is "undefined". The reason is that it could only receive a truth value after it itself had received a truth value, so at no point in the iterative process does that happen. When the semantics of definite descriptions and the object-language reference predicate works as described, something similar is the case for Berry's description. Prior to the determination of the reference of Berry's description, the predicate "is an integer not describable in fewer than twenty syllables" is false of a lot of integers, for example 3 and 11 which are the referents of "the square root of 9" and "the number of letters in 'phobophobia'" respectively. But it is not true of any integers, for given any integer for which the predicate is not yet false, it is not yet ruled out that Berry's description might refer to that integer. Ergo the unique object satisfying Berry's description cannot be identified prior to this identification itself, so Berry's description is never assigned a referent and is hence undefined in the fixed point.

In formalizing Hilbert and Bernays' Paradox we will also use definite descriptions and the reference relation. But we need one more thing, namely functions. As is standard, the interpretation of a function symbol will be specified by the interpretation function, and the function symbol can take terms as its arguments. But the value of a function for given arguments may be undetermined for a while in the evaluation process, since it may be undetermined what the terms acting as arguments refer to. We will treat this similarly to the truth functions which constitute the semantics of the connectives and the quantifiers; when there is sufficient information, the function value will be determined. To take an example, consider $f(t_1, t_2, t_3)$ where f is a function symbol and t_1, t_2, and t_3 are terms, and suppose that at some stage in the evaluation process, the reference of t_1 and

t_2 but not t_3 has been determined. Then $f(t_1, t_2, t_3)$ will get a reference at this stage, iff the reference of t_3 does not matter, i.e. if $I(f)(r_1, r_2, d)$, where I is the interpretation function, and r_1 and r_2 are the referents of t_1 and t_2 respectively, has the same value for every value of d.

It is easy to see intuitively that also Hilbert and Bernays' description does not have a reference in the fixed point; a reference of the description cannot be determined prior to this determination itself.

As I plan on showing in a forthcoming longer paper, the Kripkean approach can be used to solve all the known paradoxes of denotation, such as for example the paradoxes of König and Richard. But here I will focus on the paradoxes of Berry and Hilbert and Bernays and present a formal language that has just the resources needed to formalize them.

In Kripke's theory the evaluations at the various levels consist of a set of true sentences and a set of false sentences. The extension of the theory here envisaged means that an evaluation must also contain a reference relation from the set of terms to the domain (supplemented with something to indicate that it has been decided that a given term fails to refer). But it is not necessary to complicate things by making an evaluation a triple. Instead we can take a cue from Frege (1892) and identify a sentence *being* true/false with the sentence *referring* to Truth/Falsity. That way an evaluation can simply *be* a reference relation – one from the union of the set of sentences and the set of terms to the union of the domain and $\{\top, \bot, *\}$, where $\top$, $\bot$, and $*$ are symbols for Truth, Falsity, and failing to refer respectively.

We will use a standard first-order predicate language with function symbols supplemented with three things: A unary predicate T for "is true", a binary predicate R for "refers to", and a definite description operator: "$\imath v(\phi)$" is to be read as "the v such that ϕ".

In order to keep the technical complexity at a minimum, self-reference is made possible simply by letting the domain include all sentences and terms of the language, and by making certain assumptions about the denotation of specific constants when the "paradoxical" terms are formalized. That way, the complications of Gödel coding and a diagonal lemma can be avoided.

3 Syntax

We now turn to the precise specification of the syntax (this section) and semantics (next section) of a formal language. For each $n \in \mathbb{N}$ let there be a countable set $\mathcal{P}_n$ of **ordinary n-ary predicates** and a countable set $\mathcal{F}_n$ of **n-ary function symbols**. In addition there are two **extra-ordinary predicates**, one unary, T, and one binary, R. We also have a set $\mathcal{C}$ of **constants** and a set of **variables**, both of cardinality $\aleph_0$.

The set of **well-formed formulas (wff's)** and the set of **terms** are defined recursively thus:

- Every constant and variable is a term.

– If P is an ordinary n-ary predicate and $t_1, \ldots, t_n$ are terms, then $P(t_1, \ldots, t_n)$ is a wff.
– If ϕ and ψ are wff's, then $\neg\phi$ and $(\phi \wedge \psi)$ are wff's.
– If ϕ is a wff and v a variable, then $\forall v \phi$ is a wff's.
– If t_1 and t_2 are terms, then $T(t_1)$ and $R(t_1, t_2)$ are wff's.
– If ϕ is a wff and v a variable, then $\imath v(\phi)$ is a term.
– If f is an n-ary function symbol and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is a term.
– Nothing is a wff or term except by virtue of the above clauses.

The connective $\rightarrow$ is used as an abbreviation in the usual way.

Variables, constants, predicates (ordinary as well as extra-ordinary), function symbols, connectives, quantifiers, parenthesis, and commas are called **primitive symbols**.

When ϕ is a wff, v a variable, and c a constant, $\phi(v/c)$ is the wff which is identical with ϕ with the possible exception that all free occurrences of v are replaced with c.

A wff is a **sentence**, and a term is **closed**, if it does not contain any free variables. Let $\mathcal{S}$ and $\mathcal{CT}$ be the set of sentences and the set of closed terms respectively.

We will make use of a notion of **complexity** of a formula, but a precise definition can be dispensed with. Any reasonable definition will do.

4 Semantics

A **model** is defined as a pair $\mathfrak{M} = (D, I)$, where D, the **domain**, and I, the **interpretation function**, satisfy the following:

– D is a superset of $\mathcal{S} \cup \mathcal{CT} \cup \mathbb{N}$ such that
  - $* \notin D$, and
– I is a function defined on $\bigcup_{n \in \mathbb{N}} (\mathcal{P}_n \cup \mathcal{F}_n) \cup \mathcal{C}$ such that
  - for every $P \in \mathcal{P}_n$, $I(P) \subseteq D^n$,
  - for every $f \in \mathcal{F}_n$, $I(f)$ is a function from D^n to D,
  - for every $c \in \mathcal{C}$, $I(c) \in D$, and
  - $I[\mathcal{C}] = D$.

Let a model be fixed for the remainder of this paper. We now define an **evaluation** to be a relation $\mathcal{E}$ from $\mathcal{S} \cup \mathcal{CT}$ to $D \cup \{\top, \bot, *\}$ such that elements of $\mathcal{S}$ are only related to elements of $\{\top, \bot\}$ and elements of $\mathcal{CT}$ are only related to elements of $D \cup \{*\}$. $\mathcal{E}$ is **consistent** if every sentence and closed term is related by $\mathcal{E}$ to at most one element. An evaluation $\mathcal{E}'$ **extends** $\mathcal{E}$ if $\mathcal{E} \subseteq \mathcal{E}'$.

The semantics is build up in levels as in Kripke's theory. We first specify how to "get from one level to the next": The **evaluation with respect to the evaluation** $\mathcal{E}$, $\mathrm{E}_{\mathcal{E}}$, is defined by recursion on the complexity of the formula[1]:

[1] The clauses make reference to $\mathrm{E}_{\mathcal{E}}$, but only with respect to less complex formulas than the one under consideration. By clause 6 and 7, a formula may "gain" its

1. If t is a constant then $t \, \mathrm{E}_{\mathcal{E}} \, I(t)$.
2. If s is of the form $P(t_1, \ldots, t_n)$ where P is an ordinary n-ary predicate and $t_1, \ldots, t_n$ are closed terms, then
   - $s \, \mathrm{E}_{\mathcal{E}} \top$ if there are $d_1, \ldots, d_n \in D$ satisfying $t_1 \, \mathrm{E}_{\mathcal{E}} \, d_1, \ldots, t_n \, \mathrm{E}_{\mathcal{E}} \, d_n$ such that $(d_1, \ldots, d_n) \in I(P)$, and
   - $s \, \mathrm{E}_{\mathcal{E}} \perp$ if there are $d_1, \ldots, d_n \in D$ satisfying $t_1 \, \mathrm{E}_{\mathcal{E}} \, d_1, \ldots, t_n \, \mathrm{E}_{\mathcal{E}} \, d_n$ such that $(d_1, \ldots, d_n) \notin I(P)$.
3. If s is of the form $\neg\phi$ where ϕ is a sentence, then
   - $s \, \mathrm{E}_{\mathcal{E}} \top$ if $\phi \, \mathrm{E}_{\mathcal{E}} \perp$, and
   - $s \, \mathrm{E}_{\mathcal{E}} \perp$ if $\phi \, \mathrm{E}_{\mathcal{E}} \top$.
4. If s is of the form $(\phi \wedge \psi)$ where ϕ and ψ are sentences, then
   - $s \, \mathrm{E}_{\mathcal{E}} \top$ if $\phi \, \mathrm{E}_{\mathcal{E}} \top$ and $\psi \, \mathrm{E}_{\mathcal{E}} \top$, and
   - $s \, \mathrm{E}_{\mathcal{E}} \perp$ if $\phi \, \mathrm{E}_{\mathcal{E}} \perp$ or $\psi \, \mathrm{E}_{\mathcal{E}} \perp$.
5. If s is of the form $\forall v \phi$ where v a variable and ϕ is a wff with at most v free, then
   - $s \, \mathrm{E}_{\mathcal{E}} \top$ if for all $c \in \mathcal{C}$, $\phi(v/c) \, \mathrm{E}_{\mathcal{E}} \top$, and
   - $s \, \mathrm{E}_{\mathcal{E}} \perp$ if there exists a $c \in \mathcal{C}$ such that $\phi(v/c) \, \mathrm{E}_{\mathcal{E}} \perp$.
6. If s is of the form $T(t)$ where t is a closed term, then
   - $s \, \mathrm{E}_{\mathcal{E}} \top$ if there is a $s' \in \mathcal{S}$ such that $t \, \mathrm{E}_{\mathcal{E}} \, s'$ and $s' \mathcal{E} \top$,
   - $s \, \mathrm{E}_{\mathcal{E}} \perp$ if there is a $s' \in \mathcal{S}$ such that $t \, \mathrm{E}_{\mathcal{E}} \, s'$ and $s' \mathcal{E} \perp$, and
   - $s \, \mathrm{E}_{\mathcal{E}} \perp$ if there is a $d \in D$ such that $t \, \mathrm{E}_{\mathcal{E}} \, d$, but no $s' \in \mathcal{S}$ such that $t \, \mathrm{E}_{\mathcal{E}} \, s'$.
7. If s is of the form $R(t_1, t_2)$ where t_1 and t_2 are closed terms, then
   - $s \, \mathrm{E}_{\mathcal{E}} \top$ if there is a $d \in D \cup \{*\}$ and a closed term t_1' such that $t_1 \, \mathrm{E}_{\mathcal{E}} \, t_1'$, $t_1' \mathcal{E} d$ and $t_2 \, \mathrm{E}_{\mathcal{E}} \, d$,
   - $s \, \mathrm{E}_{\mathcal{E}} \perp$ if there are $d_1, d_2 \in D \cup \{*\}$, such that $d_1 \neq d_2$, and a closed term t_1' such that $t_1 \, \mathrm{E}_{\mathcal{E}} \, t_1'$, $t_1' \mathcal{E} d_1$ and $t_2 \, \mathrm{E}_{\mathcal{E}} \, d_2$, and
   - $s \, \mathrm{E}_{\mathcal{E}} \perp$ if there is a $d' \in D \cup \{*\}$ such that $t_1 \, \mathrm{E}_{\mathcal{E}} \, d'$, but no closed term t_1' such that $t_1 \, \mathrm{E}_{\mathcal{E}} \, t_1'$.
8. If t is of the form $\imath v(\phi)$ where v is a variable and ϕ is a wff with at most v free, then
   - $t \, \mathrm{E}_{\mathcal{E}} \, d$ if d is an element of D such that for some $c \in \mathcal{C}$, $I(c) = d$ and $\phi(v/c) \, \mathrm{E}_{\mathcal{E}} \top$ and for all other elements d' of D, every $c' \in \mathcal{C}$, such that $I(c') = d'$, satisfies $\phi(v/c') \, \mathrm{E}_{\mathcal{E}} \perp$,
   - $t \, \mathrm{E}_{\mathcal{E}} \, *$ if there are two different elements d_1 and d_2 of D such that for some $c_1, c_2 \in \mathcal{C}$, $I(c_1) = d_1$, $I(c_2) = d_2$, $\phi(v/c_1) \, \mathrm{E}_{\mathcal{E}} \top$ and $\phi(v/c_2) \, \mathrm{E}_{\mathcal{E}} \top$, and
   - $t \, \mathrm{E}_{\mathcal{E}} \, *$ if for all elements d of D, there is a $c \in \mathcal{C}$ such that $I(c) = d$ and $\phi(v/c) \, \mathrm{E}_{\mathcal{E}} \perp$.
9. If t is of the form $f(t_1, \ldots, t_n)$ where f is a n-ary function symbol and $t_1, \ldots, t_n$ are closed terms, then $t \, \mathrm{E}_{\mathcal{E}} \, d$ if d is an element of D for which every n-tuple $(d_1, \ldots, d_n)$ such that for each $i \in \{1, \ldots, n\}$ either $t_i \, \mathrm{E}_{\mathcal{E}} \, d_i$ or t_i is not related to anything by $\mathrm{E}_{\mathcal{E}}$, satisfy $I(f)(d_1, \ldots, d_n) = d$.

---

reference from a more complex formula, but here it is only the relation $\mathcal{E}$ that is used. In short, the reference of a formula only depends on the previous level and formulas of lower complexity. Hence, as stated, the definition is simply by recursion on the complexity of the formula.

Now we iterate the process by defining for all ordinals α the **evaluation with respect to the level** α, written E^α, by recursion:

$$\mathrm{E}^\alpha = \begin{cases} \emptyset & \text{if } \alpha = 0 \\ \mathrm{E}_{\mathrm{E}^{\alpha-1}} & \text{if } \alpha \text{ is a successor ordinal} \\ \bigcup_{\eta<\alpha} \mathrm{E}^\eta & \text{if } \alpha \text{ is a limit ordinal} \neq 0 \end{cases}$$

The following two lemmas show that the process is monotonic and does not result in any inconsistency:

Lemma 1. *For all ordinals α, β, if $\alpha < \beta$ then $\mathrm{E}^\alpha \subseteq \mathrm{E}^\beta$.*

Proof. By induction on the complexity of formulas it is seen that for each bullet in each of the nine clauses above, if the condition in that bullet is satisfied for some evaluation $\mathcal{E}$ it is also satisfied for every extension of $\mathcal{E}$. Ergo if $\mathcal{E} \subseteq \mathcal{E}'$ then $\mathrm{E}_\mathcal{E} \subseteq \mathrm{E}_{\mathcal{E}'}$. As it also holds that $\mathrm{E}^0 = \emptyset$ is a subset of every evaluation, the lemma follows. □

Lemma 2. *For every ordinal α, E^α is consistent.*

Proof. By outer induction on α and inner induction on the complexity of formulas, considering clause 1–9. □

For every ordinal α and every $x \in \mathcal{S} \cup \mathcal{CT}$ we define $[\![x]\!]^\alpha$ to be the unique y such that $x\, \mathrm{E}^\alpha\, y$, when there is a such. We say that x is **determined at level** α, if α is the first level where $[\![x]\!]^\alpha$ is defined.

We now come to the important fixed point theorem:

Theorem 3. *There is a unique consistent evaluation $\mathcal{E}$ such that for some ordinal α it holds that for all ordinals $\beta \geq \alpha$ that $\mathrm{E}^\beta = \mathcal{E}$.*

Proof. As there are only countable many sentences and closed terms, the monotonic process must reach a fixed point. Consistency of the fixed point follows from lemma 2. □

Letting $\mathcal{E}$ and α be as in the theorem, we define the **evaluation**, E, as $\mathcal{E}$, and for all $x \in \mathcal{S} \cup \mathcal{CT}$ set $[\![x]\!]$ equal to $[\![x]\!]^\alpha$ when this is defined. The value of $[\![x]\!]$ is to be thought of as *the* reference of x.

5 Expressibility of the Reference Relation

Kripke's theory is famous for validating the Tarskian T-schema in the sense that, if (in the notation of this paper) s is a sentence and c is a constant such that $I(c) = s$, then $[\![s]\!] = \top$ if and only if $[\![T(c)]\!] = \top$. In other words: If a sentence is true, this can be expressed in the object language. In this theory a similar result holds for reference; if a closed term refers to a given object, then this can be expressed in the language itself. That is the content of the following theorem.

Theorem 4. *Let t be a closed term, d an element of D, and c_1 and c_2 constants such that $I(c_1) = t$ and $I(c_2) = d$. The following biimplication holds: $[\![t]\!] = d$ iff $[\![R(c_1, c_2)]\!] = \top$.*

Proof. From clause 1 it is seen that for all ordinals α we have $c_1\, \mathrm{E}^\alpha\, t$ and $c_2\, \mathrm{E}^\alpha\, d$. So it follows from bullet 1 of clause 7 that $[\![t]\!] = d$ iff $t\, \mathrm{E}^\beta\, d$ for some ordinal β iff $R(c_1, c_2)\, \mathrm{E}^{\beta+1} \top$ iff $[\![R(c_1, c_2)]\!] = \top$. □

6 Solution to Berry's Paradox

In formalizing the Berry Description we have to get around the fact that in the formal language, any natural number can be defined with a definite description of just one symbol, namely a constant. We can do this by defining "length of a term" not in the obvious way as the number of primitive symbols in the term, but slightly differently. Reflecting the fact that in natural languages there are only finitely many primitive symbols, let Φ be a function from the set of primitive symbols of our formal language to $\mathbb{N}$ which sends only a finite number of primitive symbols to each $n \in \mathbb{N}$. Then define the length of a term to be the sum of $\Phi(x)$ for every occurrence x of a primitive symbol in the term.

Now we can formalize the Berry Description. Let n, m, and x be variables and let N and L be unary predicates and $\geq$ a binary predicate, such that $I(N)$ is the set of natural numbers, and $I(\geq)$ is the relation "larger than or equal to" on the set of natural numbers. L is to be interpreted as "long", but we postpone the precise specification of $I(L)$, until we know just what "long" should mean to make our formalization "paradoxical".

We can formalize "x is a definite description of the natural number n" thus:

$$N(n) \wedge R(x, n)$$

So "The natural number n does not have a short definite description" can be formalized

$$N(n) \wedge \forall x(R(x, n) \to L(x)) \ ,$$

and "n is the least natural number that does not have a short definite description"

$$(N(n) \wedge \forall x(R(x, n) \to L(x))) \wedge \\ \forall m((N(m) \wedge \forall x(R(x, m) \to L(x))) \to \geq(m, n)) \ .$$

Ergo, Berry's description in a version with length of formal expressions instead of number of syllables, "the least natural number that does not have a short definite description", can be formalized as (B):

$$\imath m((N(n) \wedge \forall x(R(x, n) \to L(x))) \wedge \\ \forall m((N(m) \wedge \forall x(R(x, m) \to L(x))) \to \geq(m, n))) \quad \text{(B)}$$

Now we can set $I(L)$ to be the set of terms which are longer than the length of (B).

That (B) fails to refer, i.e that there is no $d \in D$ such that $[\![(\mathrm{B})]\!] = d$, is proved as follows: Assume ad absurdum that there is such a $d \in D$. Then it follows by clause 8 that for a constant c with $I(c) = d$, we have

$$[\![(N(c) \wedge \forall x(R(x,c) \rightarrow L(x))) \wedge \\ \forall m((N(m) \wedge \forall x(R(x,m) \rightarrow L(x))) \rightarrow \geq(m,c))]\!] = \top \ .$$

Using clause 4 twice it can be inferred that

$$[\![\forall x(R(x,c) \rightarrow L(x))]\!] = \top \ ,$$

and consequently by clause 5 that

$$[\![R(c',c) \rightarrow L(c')]\!] = \top \ ,$$

where c' is a constant such that $I(c') = (B)$. It is already determined at level 1, that $L(c')$ is false. This follows from the specification of $I(L)$. Ergo we must have $[\![R(c',c)]\!] = \bot$. So at some level bullet 2 or 3 of clause 7 is satisfied. But bullet 3 can not be, for c' refers to (B) and since the referent of a constant is unique, not to some object which is not a term. And bullet 2 can not be either, for then (B) would have to refer to something different from d, but by assumption this is not the case. This is a contradiction.

7 Solution to Hilbert and Bernays' Paradox

The Hilbert and Bernays description can be formalized

$$+(\bar{1}, \imath v(R(h,v))) \ , \tag{HB1}$$

where v is a variable, h is a constant such that $I(h) =$ (HB1), and $+$ is a binary function symbol such that $I(+)$ is the function that sends every pair of numbers to their sum and every other pair to 0. $\bar{1}$ is a numeral for 1.

$[\![(\mathrm{HB1})]\!]$ is undefined, as we will proceed to prove. As the sum of 1 and n is not the same for every natural number n, (HB1) will get a reference, only if

$$\imath v(R(h,v)) \tag{HB2}$$

gets a reference (clause 9). By clause 8 this happens only if there is a constant c such that

$$R(h,c) \tag{HB3}$$

is related to $\top$. We have $h\,\mathrm{E}_{\emptyset}(\mathrm{HB1})$ from which it follows by bullet 1 of clause 7 that this can only be the case if (HB1) gets a reference. We have come full circle, and can conclude that neither (HB1), (HB2), nor (HB3) become related to anything.

References

Bernays, P. and Hilbert, D.: Grundlagen der Mathematik (zweiter Band). Verlag von Julius Springer (1939)

Frege, G.: Über Sinn und Bedeutung. Zeitschrift fur Philosophie und philosophische Kritik **100** (1892) 26–60

Kripke, S.: Outline of a theory of truth. The Journal of Philosophy **72** (1975) 690–716

Priest, G.: Beyond the limits of thought (second ed.). Oxford University Press (2002)

Priest, G.: The paradoxes of denotation. In T. Bolander, V. F. Hendricks and S. A. Pedersen (eds.): Self-reference. CSLI Publications (2006)

Russell, B.: Mathematical logic as based on the theory of types. American Journal of Mathematics **30** (1908) 222–262