The Theory of Abstract Objects

E. Badstübner, D. Kirchner, P. Schießl

Project of Computational Metaphysics 2016

C. Benzmüller, A. Steen, M. Wisniewski

Freie Universität Berlin, Institute of Computer Science, Germany

Introduction

The *Theory of Abstract Objects* is the core of Edward Zalta's upcoming *Principia Metaphysica* [2], a foundational theory of metaphyiscs in the spirit of Whitehead and Russel's *Principia Mathematica*.

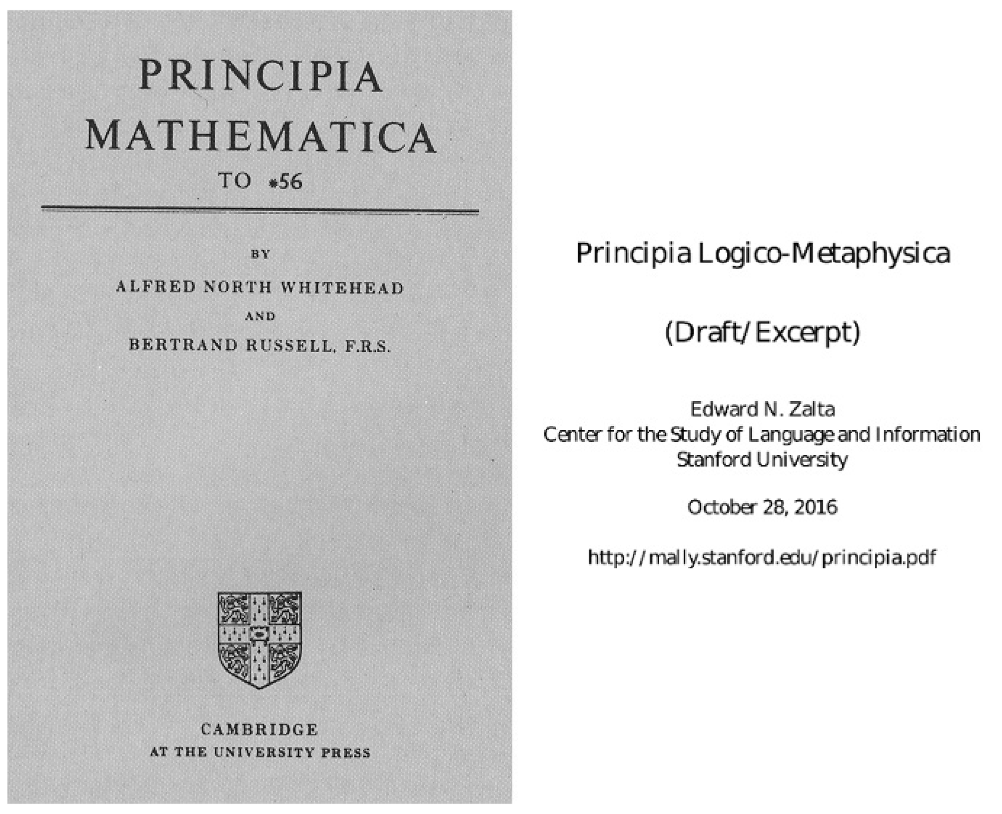

Figure 1: Principia Mathematica and the current draft of Principia Metaphysica

The theory postulates the existence of *abstract objects* (e.g. mathematical objects like numbers) that can *encode* properties 'xF' in contrast to concrete *ordinary objects* (e.g. people, trees, posters) that can only *exemplify* properties 'Fx'.

Exemplification is used for classical predication, e.g. 'John is happy', whereas *encoding* is a second mode of predication used for abstract objects (e.g the fictional character Sherlock Holmes).

The idea is that if an object *exemplifies* a property it must have a spatiotemporal location, a body with a shape, a mass, etc. None of that is true for the fictional character Sherlock Holmes. Therefore the property of 'being a detective' is not *exemplified* by Sherlock Holmes. On the other hand 'being a detective' is a property we use to identify Sherlock Holmes and distinguish him from other fictional characters. To account for that we say that Sherlock Holmes *encodes* the property of being a detective.

The Theory of Abstract Objects

(I) $\exists x(A!x \,\&\, \forall F(xF \equiv \Phi))$

(II) $x = y \;\equiv\; \Box\forall F(xF \equiv yF)$

The equations above are the two most important principles of the Theory of Abstract Objects.

Flammarions Holzstich – in L'atmosphère, Paris 1888.

Figure 2: Metaphysics goes beyond real-world experience

Their intention can informally be read as:

(I) For each group of properties, there is an abstract object that encodes exactly the properties in that group.

(II) Two abstract objects are identical if and only if they (necessarily) encode the same properties.

Principia Metaphysica constructs a formal axiomatic system around these principles that makes it possible to describe and analyse all kinds of *abstract objects* within a single framework.

This includes theoretical mathematical objects such as natural numbers, as well as philosophical objects such as Forms (Plato), concepts (Leibniz), possible worlds (Leibniz), senses (Frege), the world as state of affairs (Wittgenstein), etc.

Automation

The theory is a viable option for a foundational theory not only of metaphysics, but also of mathematics. Its automation in an computer-assisted reasoning system is therefore a highly interesting challenge. Fig. 2 displays an exemplary formalization within the interactive proof assistant Isabelle/HOL.

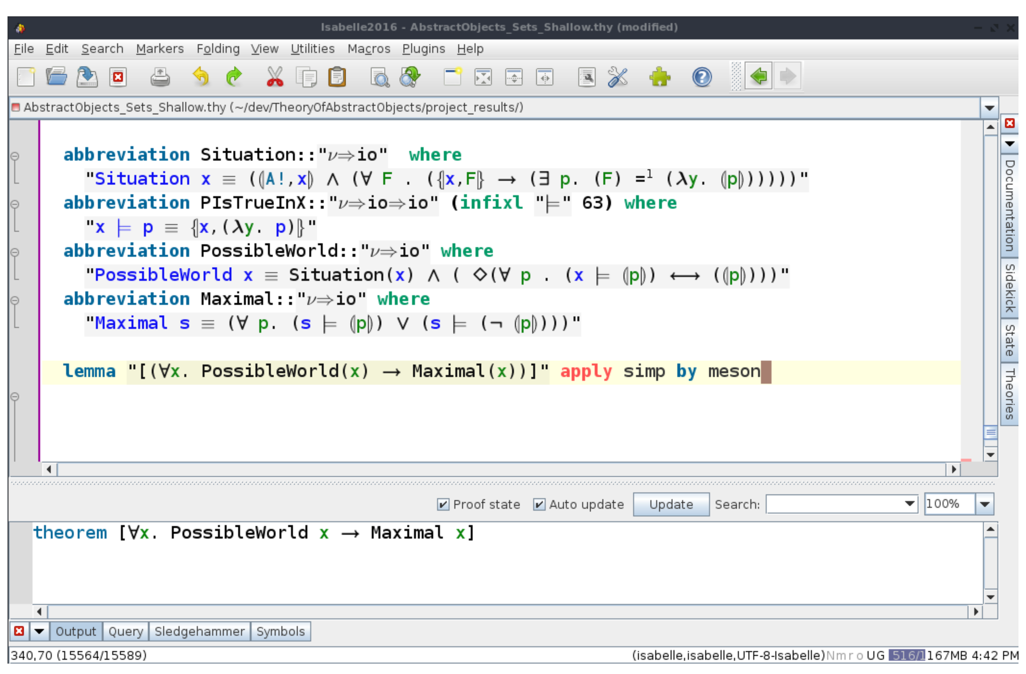

Figure 3: Proof of the theorem "Possible worlds are maximal" in Isabelle/HOL.

Relational vs. Functional Type Theory

The Theory of Abstract Objects is formulated using relational type theory, whereas higher order reasoning systems such as Isabelle/HOL are based on a functional type theory. The direct translation between the two is notoriously problematic [1].

As an example, the Theory of Abstract Objects does not allow encoding subformulas in lambda expressions (the first principle mentioned above with Φ being $F = [\lambda x \exists F(xF \,\&\, \neg Fx)]$ would lead to a contradiction similar to Russel's paradox). Reproducing this restriction in functional type theory is challenging.

Hyper-Intensionality

The Theory of Abstract Object uses a *defined* equality for objects and relations, e.g. two one-place relations are considered equal if they are *encoded* by the same objects. This leads to a hyper-intensional logic for which the material equality for properties (two properties are equal if and only if the same objects *exemplify* them) no longer holds. In the absence of encoding formulas, on the other hand, the logic still collapses to a classical extensional logic. As Boolean extensionality is a built-in feature of systems like Isabelle/HOL, this represents another challenge.

Summary and Results

Several solutions to the challenges mentioned above were analysed and found to be unsatisfactory (i.e. they either resulted in inconsistencies or showed major deviations from the intended logic of PM).

A promising new approach for a complete embedding of the theory based on a set-theoretic model has been proposed and is currently developed further in the context of a master thesis.

References

[1] P. E. Oppenheimer and E. N. Zalta. Relations versus functions at the foundations of logic: Type-theoretic considerations. *Journal of Logic and Computation*, 21(2):351–374, Jun 2010.

[2] Edward Zalta. Principia logico-metaphysica (draft/excerpt). `https://mally.stanford.edu/principia.pdf`, 2016. [Online; accessed 30-November-2016].

Leibniz: *Calculemus!*

Computational Metaphysics is a interdisciplinary lecture course designed for advanced students of computer science, mathematics and philosophy. The main objective of the course is to teach the students how modern proof assistants based on expressive higher-order logic support the formal analysis of rational arguments in philosophy (and beyond). In our first course in Summer 2016 the focus has been on ontological arguments for the existence of God. However, some students picked formalisation projects also from other areas (including maths).

Computational Metaphysics was awarded the *Central Teaching Award 2015* of the FU Berlin.