

Definability in Self-Referential Systems

J. ZIMBARG SOBRINHO

Introduction Self-referential systems are theories formulated in a (typed) first-order language \mathcal{L} , in whose intended interpretation the predicates refer to objects which are themselves predicates of \mathcal{L} . In order to avoid the Russell-Zermelo paradox, Hiller and Zimbarg Sobrinho [1] introduced *self-referential systems with involution*: these are theories whose intended models admit an elementary embedding into itself, denoted by ‘*’, and called *involution map*. As a consequence, the universe of predicates inherits a structural hierarchy of objects, classified into countably many *types*.

A peculiar property of types refers to the size of the universe of their corresponding objects: the larger a type, the smaller its domain, and for this reason, types have been suggestively taken as negative integers: 0, -1 , -2 , and so on.

The main properties satisfied by self-referential systems with involution were outlined in [2], and are the following:

- (a) \mathcal{L} possesses unrestricted (or universal) variables
- (b) all predicates are extensional
- (c) the Comprehension axiom for starred formulas is true
- (d) (Definability condition) every element in the universe of a realization is definable by a one-free-variable formula of \mathcal{L} .

The first three clauses above can be directly expressed in our (typed) first-order language \mathcal{L} without any further ado. With respect to the Definability condition, however, it is not altogether clear how it could be formulated in a first-order version of self-reference, due to its obvious higher-order character. The purpose of this article is to present first-order axioms which, added to \mathcal{W} (see [1]), produce the same effect as the apparently stronger ‘Definability condition’.

1 Hiller’s problem Realizations of self-reference in which the Definability condition holds have been referred to as *intended models*. It is well-known that

Received July 7, 1986