26
Notre Dame Journal of Formal Logic
Volume 31, Number 1, Winter 1990

Characters and Fixed Points in Provability Logic

ZACHARY GLEIT and WARREN GOLDFARB*

Abstract Some basic theorems about provability logic—the system of modal logic that reflects the behavior of formalized provability predicates in theories such as arithmetic—are given simplified, model-theoretic proofs. The theorems include the Fixed Point Theorem of de Jongh and Sambin, the Craig Interpolation Theorem, and the Beth Definability Theorem. Attention is also paid to the complexity of models for formulas in this logic.

Provability logic is the modal logic whose axioms are the tautologies and all formulas of the forms $\Box(A \to B) \to (\Box A \to \Box B)$ and $\Box(\Box A \to A) \to \Box A$, and whose inference rules are modus ponens and necessitation. This system has been known variously as K4W, L, G, GL, and PRL; we adopt the next to last of these monikers, and use "$\vdash$" for provability in GL. GL is of interest since it reflects the behavior of formalized provability predicates in such theories as Peano Arithmetic (for details see [1] or [3]).

A formula A is said to be *modalized* in a sentence letter p iff every occurrence of p in A lies in the scope of a $\Box$. The Fixed Point Theorem, due to de Jongh and Sambin, states that if A is modalized in p then there exists a formula H containing only sentence letters of A aside from p such that $\vdash \Box(p \leftrightarrow A) \leftrightarrow \Box(p \leftrightarrow H)$. Via the connection of GL with provability in formal theories, this theorem implies that sentences in such formal theories constructed from formalized provability predicates by "self-referential" techniques are provably equivalent to sentences involving no self-reference.

In Section 1, we give a purely model-theoretic proof of the Fixed Point Theorem, which we think to be more perspicuous than the extant proofs. The

*We would like to thank George Boolos for much generous assistance and many helpful suggestions, and in particular for simplifying the counterexample found in the fourth paragraph of Section 3. We are also grateful to Lisa Reidhaar-Olson for catching several errors in the penultimate draft.

Received January 20, 1988; revised July 26, 1988