Annals of Pure and Applied Logic 165 (2014) 199–224

Contents lists available at SciVerse ScienceDirect

Annals of Pure and Applied Logic

www.elsevier.com/locate/apal

The Paradox of the Knower revisited

Walter Dean [a,*], Hidenori Kurokawa [b]

[a] *Department of Philosophy, University of Warwick, Coventry, CV4 7AL, UK*
[b] *Kobe University, Department of Information Science, 1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan*

ARTICLE INFO

Article history:
Available online 19 August 2013

MSC:
03F30
03B42
03F45
03F40

Keywords:
Knower Paradox
Explicit modal logic
Justification logic
Reflection principles
Provability logic

ABSTRACT

The Paradox of the Knower was originally presented by Kaplan and Montague (1960) [26] as a puzzle about the everyday notion of knowledge in the face of self-reference. The paradox shows that any theory extending Robinson arithmetic with a predicate $K(x)$ satisfying the factivity axiom $K(\overline{A}) \rightarrow A$ as well as a few other epistemically plausible principles is inconsistent. After surveying the background of the paradox, we will focus on a recent debate about the role of epistemic closure principles in the Knower. We will suggest this debate sheds new light on the concept of knowledge which is at issue in the paradox – i.e. is it a "thin" notion divorced from concepts such as evidence or justification, or is it a "thick" notion more closely resembling mathematical provability? We will argue that a number of features of the paradox suggest that the latter option is more plausible. Along the way, we will provide a reconstruction of the paradox using a quantified extension of Artemov's (2001) [2] Logic of Proofs, as well as a series of results linking the original formulation of the paradox to reflection principles for formal arithmetic. On this basis, we will argue that while the Knower can be understood to motivate a distinction between levels of knowledge, it does not provide a rationale for recognizing a uniform hierarchy of knowledge predicates in the manner suggested by Anderson (1984) [1].

The Paradox of the Knower is traditionally understood as exposing an antinomy pertaining to the everyday concept of knowledge in the face of self-reference. But more than fifty years after its discovery by Kaplan and Montague (K&M) [26] in 1960, there appears to be little consensus as to its proper resolution. This owes in part to the fact that the paradox rests on a number of distinct assumptions about the logical properties of knowledge which may be contested individually. However, the paradox itself consists of nothing more than a derivation in an axiomatic system containing an otherwise uninterpreted predicate $K(x)$. And at a formal level, the reasoning it embodies is similar in form to inconsistency results which are traditionally understood to describe notions other than knowledge – e.g. informal provability [37], logical necessity [36], or "knowable truth" [14]. It thus seems that not only is there little consensus as to how the Knower should be resolved, but also that it is not entirely clear which of our pre-theoretical notions it is intended to cast into doubt.

* Corresponding author.
E-mail addresses: W.H.Dean@warwick.ac.uk (W. Dean), hidenori.kurokawa@gmail.com (H. Kurokawa).

0168-0072/$ – see front matter
http://dx.doi.org/10.1016/j.apal.2013.07.010

Most commentators on the Knower have understood $K(x)$ in what we will describe as a *schematic manner* – i.e. as implicitly defining a concept which satisfies certain "knowledge-like" sentential principles but which is not assumed to result from an analysis of knowledge in terms of more primitive notions such as justification or belief forming mechanisms.[1] However, a recent exchange between Charles Cross [5–7] and Gabriel Uzquiano [50] about the role of epistemic closure principles in the Knower highlights several connections between the interpretation of $K(x)$ in the Knower and formal provability predicates for arithmetic. This suggests that the technical setting of the paradox may itself make it more difficult to maintain a schematic interpretation of $K(x)$ than it might at first appear.

Our first goal in this paper will be to argue that by taking this observation to heart, it is possible to gain some insight into the range of possible interpretations of $K(x)$. After first presenting some preliminaries about the original formulation of the Knower and Cross's recasting thereof (i.e. what he calls the *Paradox of the Knowledge-plus Knower*) in Sections 1 and 2, we will argue in Section 3 that a number of considerations about the formulation of the Knower suggest that the notion of knowledge which is described by its premises is one on which mathematical proof is taken to be at least a *sufficient* condition for knowledge.

In Section 4 we will sketch how on this interpretation the reasoning of the Knower can be redeveloped in a quantified extension of Artemov's [2] *Logic of Proofs*. We will suggest that adopting a proof-based conception of knowledge reveals some additional logical structure which is latent in the original K&M derivation. On this basis we will also suggest that the paradox cannot be satisfactorily resolved by denying that $K(x)$ satisfies an appropriate epistemic closure principle in the manner proposed by Maitzen [33] (and tentatively seconded by Uzquiano).

In Section 5 we will highlight the relationship between the epistemic principles on which the Knower is based and formal reflection principles for arithmetic. On this basis, we will suggest that if we wish to understand statements attributing knowledge as being an *iterable* (i.e. so that $K(x)$ may hold of sentences which themselves contain instances of this predicate), then it becomes difficult to maintain a uniform interpretation of $K(x)$ in all syntactic contexts in which it might appear.

Relative to the proof-based interpretation of knowledge, we will suggest that the moral of the Knower is closely related to the difference between the sorts of evidence we might possess for provable arithmetical statements and formally independent ones. We will conclude in Section 6 by offering some observations which we take to detract from the appeal of taking such considerations to motivate the recognition of a uniform hierarchy of knowledge predicates in the style of Anderson [1].

Before getting under way, it will be useful to say a few words about the relationship between the Knower and two nearby topics whose relation to the paradox has not (at least to the best of our knowledge) been systematically explored. The first of these is the long-standing debate as to whether our everyday conception of knowledge admits to analysis into other putatively simpler or more fundamental terms. For instance, the traditional view holds that knowledge can be properly analyzed as justified, true belief. Needless to say, however, this view is now widely thought to be either false or in need of substantial modification. Partly in response to this, Williamson [51] has argued that while knowledge is correctly regarded as a factive propositional attitude, it is primitive in the sense that it cannot be "factored" into a combination of other mental states and non-mental conditions. Such a view which has, in turn, at least an affinity to the schematic understanding of $K(x)$ mentioned above.

Another topic which abuts the literature on the Knower is work on axiomatic theories of truth. The origins of this subject can be traced to some of the same papers (e.g. [36,34,14]) which also bear a close historical connection to the discovery of the Knower. The setting of the two subjects is also similar – i.e. a primitive predicate is adjoined to an arithmetical base theory along with principles which are presumed to

[1] The use of a predicate rather than an operator to formalize knowledge notwithstanding, the adoption of such a perspective is similar to that often adopted in epistemic logic (e.g. in the tradition of Hintikka [25]) wherein it is conventional to forego attempts to provide an analysis of knowledge in favor of studying its properties axiomatically.