A machine that knows its own code

Samuel A. Alexander*†
Department of Mathematics, the Ohio State University

May 24, 2013

Abstract

We construct a machine that knows its own code, at the price of not knowing its own factivity.

1 Introduction

It is well known that a suitably idealized mechanical "knowing agent" capable of logic, arithmetic, and self-reflection, cannot know the index of a Turing machine that represents its own knowledge. See Lucas [7], Benacerraf [2], Reinhardt [10], Penrose [8], Carlson [4], and Putnam [9]. However, the proofs always involve (in various guises) the machine knowing its own factivity: that the machine satisfies $K(K\phi \to \phi)$. We will relax this requirement and explicitly construct a machine that knows its own code. The construction resembles that of [4] and [6].

Our result should be compared with that of Carlson [4], who showed that a truthful knowing agent can know its own truth and know that it has *some* code, without knowing which code. A machine can know its own factivity as well as that it has some code (without knowing which), or it can know its own code exactly but not know its own factivity (despite actually being factive). This dichotomy in machine knowledge was first presented in the author's dissertation [1].

In Section 2, we discuss preliminaries.

In Section 3, we construct a machine and prove that it knows its own code.

2 Preliminaries

We will work in the language $\mathscr{L}$ of Epistemic Arithmetic of S. Shapiro [11]. This is the language of Peano arithmetic (with variables $x, y, z, \ldots$, constant symbol 0, unary function symbol S for successor, and binary function symbols $+$ and $\cdot$ for addition and multiplication), extended by a modal operator K for knowledge. The well-formed formulas of $\mathscr{L}$ (and their free variables $\phi \mapsto FV(\phi)$) are defined in the usual way; a formula of the form $K(\phi)$ is called *purely modal*, and will be written $K\phi$ if no confusion results. Formulas without free variables are *sentences*. Terms, substitutability, and the result $\phi(x|t)$ of substituting term t for variable x in ϕ, are defined in the obvious ways.

We borrow the following semantics from T.J. Carlson [4] (pp. 54–55). We have reworded the definition in an equivalent form (except that Carlson allowed for multiple operators while we need only one). The intuition is that purely modal formulas should be treated as much like propositional atoms as possible.

Definition 1. (The Base Logic)

1. If U is some set, an *assignment into* U is a function that maps variables of $\mathscr{L}$ into U.
2. If s is an assignment into U, x is a variable, and $u \in U$, $s(x|u)$ shall mean the assignment into U that agrees with s except that it maps x to u.

*Email: alexander@math.ohio-state.edu
†2010 Mathematics Subject Classification: 03D80

3. An *$\mathscr{L}$-structure* $\mathcal{M}$ consists of a first-order structure $\mathcal{M}_0$ for the first-order part of $\mathscr{L}$, together with a function that takes one assignment s (into the universe of $\mathcal{M}_0$) and one purely modal formula $K\phi$, and outputs either True or False—in which case we write $\mathcal{M} \models K\phi[s]$ or $\mathcal{M} \not\models K\phi[s]$, respectively—satisfying the following three constraints:
   (a) Whether or not $\mathcal{M} \models K\phi[s]$ does not depend on $s(x)$ if x is not a free variable of ϕ.
   (b) If ψ is an *alphabetic variant* of ϕ (meaning that ψ is obtained from ϕ by renaming bound variables so as to respect the binding of the quantifiers) then, for any assignment s, $\mathcal{M} \models K\phi[s]$ if and only if $\mathcal{M} \models K\psi[s]$.
   (c) (Weak Substitution)[1] If x and y are variables, $K\phi$ is a modal formula, y is substitutable for x in ϕ, and s is an assignment, then $\mathcal{M} \models K\phi(x|y)[s]$ if and only if $\mathcal{M} \models K\phi[s(x|s(y))]$.
4. From this, for any formula ϕ, $\mathcal{M} \models \phi[s]$ and $\mathcal{M} \not\models \phi[s]$ are defined in the usual inductive way. We say $\mathcal{M} \models \phi$ if $\mathcal{M} \models \phi[s]$ for every assignment s.
5. If Σ is a set of $\mathscr{L}$-sentences and ϕ is an $\mathscr{L}$-formula, we write $\Sigma \models \phi$ to indicate that for every $\mathscr{L}$-structure $\mathcal{M}$, if $\mathcal{M} \models \Sigma$ (meaning $\mathcal{M} \models \sigma$ for every $\sigma \in \Sigma$) then $\mathcal{M} \models \phi$.
6. An $\mathscr{L}$-formula ϕ is *valid* if $\emptyset \models \phi$.

Lemma 1. (Completeness and compactness)

1. The set of valid $\mathscr{L}$-formulas is r.e.
2. For any r.e. set Σ of $\mathscr{L}$-sentences, $\{\phi : \Sigma \models \phi\}$ is r.e.
3. There is an effective procedure that, given (a Gödel number of) an r.e. set Σ of $\mathscr{L}$-sentences, outputs (a Gödel number of) $\{\phi : \Sigma \models \phi\}$.
4. If Σ is a set of $\mathscr{L}$-sentences and $\Sigma \models \phi$, there is a finite set $\sigma_1, \ldots, \sigma_n \in \Sigma$ such that[2] $\sigma_1 \to \cdots \to \sigma_n \to \phi$ is valid.

Proof. Straightforward. □

Definition 2. The *axioms of Peano arithmetic for $\mathscr{L}$* consist of the axioms of Peano arithmetic, with the induction schema extended to $\mathscr{L}$. To be precise, the axioms of Peano arithmetic for $\mathscr{L}$ are as follows.

1. $\forall x(S(x) \neq 0)$.
2. $\forall x \forall y(S(x) = S(y) \to x = y)$.
3. $\forall x(x + 0 = x)$.
4. $\forall x \forall y(x + S(y) = S(x + y))$.
5. $\forall x(x \cdot 0 = 0)$.
6. $\forall x \forall y(x \cdot S(y) = x \cdot y + x)$.
7. The universal closure of $\phi(x|0) \to (\forall x(\phi \to \phi(x|S(x)))) \to \forall x\phi$ for any $\mathscr{L}$-formula ϕ.

Definition 3.

- The *pre-closure axioms of knowledge* are given by the following schemata.
  - $E1$: The universal closure of $K\phi$ whenever ϕ is valid.
  - $E2$: The universal closure of $K(\phi \to \psi) \to K\phi \to K\psi$.
  - $E3$: The universal closure of $K\phi \to \phi$.

[1] The full Substitution Lemma, where variable y is replaced by an arbitrary term t, is not generally valid in modal logic.
[2] Throughout the paper, $A \to B \to C$ is shorthand for $A \to (B \to C)$, and similar for longer implication chains.

– $E4$: The universal closure of $K\phi \to KK\phi$.

- The *axioms of knowledge* consist of the pre-closure axioms of knowledge along with $K\phi$ whenever ϕ is a pre-closure axiom of knowledge.
- The axioms of *epistemic arithmetic* consist of the pre-closure axioms of knowledge along with $K\phi$ whenever ϕ is a pre-closure axiom of knowledge or ϕ is an axiom of Peano arithmetic for $\mathscr{L}$.
- The *axioms of knowledge mod factivity* consist of the pre-closure axioms of knowledge along with $K\phi$ whenever ϕ is an instance of $E1$, $E2$, or $E4$.
- The axioms of *epistemic arithmetic mod factivity* consist of the pre-closure axioms of knowledge along with $K\phi$ whenever ϕ is an instance of $E1$, $E2$, $E4$, or an axiom of Peano arithmetic for $\mathscr{L}$.

Definition 4. By *Reinhardt's schema* we mean the following schema ([10], p. 327)

- $\exists e K \forall x (K\phi \leftrightarrow x \in W_e)$, whenever $FV(\phi) \subseteq \{x\}$.

Reinhardt demonstrated that a formalization of "I am a Turing machine and I know which one" cannot be consistent with epistemic arithmetic. To do this, he found a particular instance of Reinhardt's schema that was inconsistent with epistemic arithmetic. A truthful mechanical knowing agent that knows its own code necessarily knows all instances of Reinhardt's schema (for example, suppose ϕ is the formula "the xth Turing machine runs forever"; if I know my own code, I can deduce a code for the set of those $n \in \mathbb{N}$ such that I know the nth Turing machine runs forever).

We will show that Reinhardt's schema is consistent (in fact, ω-consistent, by which we mean it has a structure with universe $\mathbb{N}$ where the symbols of Peano arithmetic are given the usual interpretations) with epistemic arithmetic mod factivity.

This result should be compared with the main result of [4]. Along with the above schema, Reinhardt introduced ([10], p. 320) a weaker schema, Reinhardt's *strong mechanistic thesis*, $K\exists e \forall x (K\phi \leftrightarrow x \in W_e)$.[3] Reinhardt conjectured, and Carlson proved[4], that the strong mechanistic thesis is consistent with epistemic arithmetic. Thus we have a dichotomy: a truthful knowing machine can know it is *some* machine (but not which one), and also know itself to be truthful; alternatively, a truthful knowing machine can know precisely which machine it is, but not know itself to be truthful.

3 The Construction

Definition 5. Suppose ϕ is an $\mathscr{L}$-sentence and s is an assignment into $\mathbb{N}$. We define ϕ^s to be the sentence

$$\phi^s = \phi(x|\overline{s(x)})(y|\overline{s(y)})\cdots$$

obtained by replacing each free variable in ϕ by a numeral for the natural number it is assigned to.

For example, if $s(x) = 0$ and $s(y) = 2$, then $(x = y)^s$ is the sentence $(0 = S(S(0)))$.

The machine we construct will have the following form for a certain well-chosen set Σ.

Definition 6. If Σ is a set of $\mathscr{L}$-sentences, let $\mathcal{M}_\Sigma$ be the $\mathscr{L}$-structure with universe $\mathbb{N}$, in which symbols of Peano arithmetic are interpreted in the usual way, and in which knowledge is interpreted so that for all $\mathscr{L}$-formulas ϕ and assignments s into $\mathbb{N}$,

$$\mathcal{M}_\Sigma \models K\phi[s] \text{ iff } \Sigma \models \phi^s.$$

Lemma 2. For any Σ as in Definition 6, $\mathcal{M}_\Sigma$ really is an $\mathscr{L}$-structure.

[3] Reinhardt lists only $\exists e \forall x (K\phi \leftrightarrow x \in W_e)$, and $K\exists e \forall x (K\phi \leftrightarrow x \in W_e)$ follows by the rule of necessitation. Clearly the latter formula is what is important. Reinhardt originally referred to this as the Post-Turing thesis, and later decided on the name *strong mechanistic thesis* (see [4] p. 54).

[4] This was accomplished using deep structural theorems on the ordinal numbers [3], later organized into patterns of resemblance [5].

Proof. We must verify the conditions on $\mathcal{M}_\Sigma \models K\phi[s]$ from Definition 1. Let s be an assignment into $\mathbb{N}$.

- (a) If x is not free in ϕ, then ϕ^s does not depend on $s(x)$, so neither does $\Sigma \models \phi^s$, so neither does $\mathcal{M}_\Sigma \models K\phi[s]$.
- (b) An easy inductive argument shows that any time ψ is an alphabetic variant of ϕ, for any assignment s into $\mathbb{N}$, ψ^s is an alphabetic variant of ϕ^s. Another easy induction shows that whenever ψ is an alphabetic variant of ϕ, $\psi \leftrightarrow \phi$ is valid, so certainly $\Sigma \models \phi \leftrightarrow \psi$. It follows that (when ψ is an alphabetic variant of ϕ) $\mathcal{M}_\Sigma \models K\phi[s]$ if and only if $\mathcal{M}_\Sigma \models K\psi[s]$.
- (c) (Weak Substitution) Let x and y be variables. An easy inductive argument shows that for all assignments t into $\mathbb{N}$ and all formulas ϕ such that y is substitutable for x in ϕ, $\phi(x|y)^t \equiv \phi^{t(x|t(y))}$. By definition $\mathcal{M}_\Sigma \models K\phi(x|y)[s]$ if and only if $\Sigma \models \phi(x|y)^s$, which holds if and only if $\Sigma \models \phi^{s(x|s(y))}$, which is true if and only if $\mathcal{M}_\Sigma \models K\phi[s(x|s(y))]$.

□

Lemma 3. For any Σ as in Definition 6, any $\mathscr{L}$-formula ϕ, and any assignment s, $\mathcal{M}_\Sigma \models \phi[s]$ if and only if $\mathcal{M}_\Sigma \models \phi^s$.

Proof. By induction on formula complexity of ϕ. The most interesting case is when ϕ is $K\phi_0$ for some formula ϕ_0. Suppose $\mathcal{M}_\Sigma \models K\phi_0[s]$, so $\Sigma \models \phi_0^s$. If we let t be an arbitrary assignment, since ϕ_0^s is a sentence, $\phi_0^s \equiv (\phi_0^s)^t$ and thus $\Sigma \models (\phi_0^s)^t$. By definition this means $\mathcal{M}_\Sigma \models K\phi_0^s[t]$. By arbitrariness of t, $\mathcal{M}_\Sigma \models K\phi_0^s$. The converse is similar. □

Lemma 4. For any Σ as in Definition 6, $\mathcal{M}_\Sigma$ satisfies all instances of $E2$.

Proof. Let s be an assignment and suppose $\mathcal{M}_\Sigma \models K(\phi \to \psi)[s]$ and $\mathcal{M}_\Sigma \models K\phi[s]$. This means $\Sigma \models (\phi \to \psi)^s$ and $\Sigma \models \phi^s$. Clearly $(\phi \to \psi)^s \equiv \phi^s \to \psi^s$, so by modus ponens, $\Sigma \models \psi^s$, so $\mathcal{M}_\Sigma \models \psi[s]$. □

Lemma 5. For any Σ as in Definition 6, $\mathcal{M}_\Sigma$ satisfies the axioms of Peano arithmetic for $\mathscr{L}$.

Proof. Let $\mathcal{M} = \mathcal{M}_\Sigma$. Let ψ be an axiom of Peano arithmetic. If ψ is any other axiom besides an instance of induction, $\mathcal{M} \models \psi$ because $\mathcal{M}$ has universe $\mathbb{N}$ and interprets the symbols of Peano arithmetic in the intended ways. But suppose ψ is a universal closure of

$$\phi(x|0) \to (\forall x(\phi \to \phi(x|S(x)))) \to \forall x\phi.$$

Let s be an assignment and assume $\mathcal{M} \models \phi(x|0)[s]$ and $\mathcal{M} \models \forall x(\phi \to \phi(x|S(x)))[s]$. We must show $\mathcal{M} \models \forall x\phi[s]$.

Since $\mathcal{M} \models \phi(x|0)[s]$, Lemma 3 says $\mathcal{M} \models \phi(x|0)^s$. Clearly $\phi(x|0)^s \equiv \phi^{s(x|0)}$, so $\mathcal{M} \models \phi^{s(x|0)}$.

For each $m \in \mathbb{N}$, since $\mathcal{M} \models \forall x(\phi \to \phi(x|S(x)))[s]$, in particular $\mathcal{M} \models \phi \to \phi(x|S(x))[s(x|m)]$. And thus, *if* $\mathcal{M} \models \phi[s(x|m)]$, then $\mathcal{M} \models \phi(x|S(x))[s(x|m)]$. By Lemma 3, that last sentence can be rephrased: *if* $\mathcal{M} \models \phi^{s(x|m)}$, then $\mathcal{M} \models \phi(x|S(x))^{s(x|m)}$; but clearly $\phi(x|S(x))^{s(x|m)} \equiv \phi^{s(x|m+1)}$, so in summary:

- $\mathcal{M} \models \phi^{s(x|0)}$.
- For each $m \in \mathbb{N}$, if $\mathcal{M} \models \phi^{s(x|m)}$, then $\mathcal{M} \models \phi^{s(x|m+1)}$.

Therefore, by mathematical induction, $\mathcal{M} \models \phi^{s(x|m)}$ for every $m \in \mathbb{N}$. By Lemma 3, for all $m \in \mathbb{N}$, $\mathcal{M} \models \phi[s(x|m)]$. So $\mathcal{M} \models \forall x\phi[s]$, as desired. □

Lemma 6. Suppose Σ (as in Definition 6) is *closed under* K, by which we mean that for every $\phi \in \Sigma$, $K\phi \in \Sigma$. Furthermore, assume Σ contains all instances of $E1$ and $E2$ from Definition 3. Then $\mathcal{M}_\Sigma$ satisfies all instances of $E4$.

Proof. Assume $\mathcal{M}_\Sigma \models K\phi[s]$. This means $\Sigma \models \phi^s$. By Lemma 1 there are finitely many $\sigma_1, \ldots, \sigma_n \in \Sigma$ such that $\sigma_1 \to \cdots \to \sigma_n \to \phi^s$ is valid. Thus, the universal closure of $K(\sigma_1 \to \cdots \to \sigma_n \to \phi^s)$ is an instance of $E1$, hence in Σ. By repeated instances of $E2$ in Σ, Σ implies the universal closure of

$$K(\sigma_1 \to \cdots \to \sigma_n \to \phi^s) \to K\sigma_1 \to \cdots \to K\sigma_n \to K\phi^s.$$

It follows that $\Sigma \models K\phi^s$, so $\mathcal{M}_\Sigma \models KK\phi[s]$. □

Definition 7. By *assigned validity* we mean the following schemata of $\mathscr{L}$-sentences:

- ϕ^s, whenever ϕ is valid and s is any assignment.

Lemma 7. For any Σ as in Definition 6, if Σ contains all instances of assigned validity, then $\mathcal{M}_\Sigma$ satisfies all instances of $E1$.

Proof. Suppose ϕ is valid and s is any assignment, we will show $\mathcal{M}_\Sigma \models K\phi[s]$. Since ϕ is valid, ϕ^s is an instance of assigned validity, so $\Sigma \models \phi^s$ by assumption. Thus $\mathcal{M}_\Sigma \models K\phi[s]$. □

Definition 8. For every $n \in \mathbb{N}$, let $\Sigma(n)$ be the family of axioms consisting of the following $\mathscr{L}$-schemata.

1. $E1$, $E2$, and $E4$.
2. The axioms of Peano arithmetic for $\mathscr{L}$.
3. $\forall x(K\phi \leftrightarrow \langle x, \overline{\ulcorner \phi \urcorner}\rangle \in W_{\overline{n}})$, ϕ any $\mathscr{L}$-formula with $FV(\phi) \subseteq \{x\}$.
   - Here $\ulcorner \bullet \urcorner$ denotes canonical Gödel number, $\overline{\bullet}$ denotes numeral, and $\langle \bullet, \bullet \rangle$ abbreviates a definition (in Peano arithmetic) of a canonical computable bijection $\mathbb{N}^2 \to \mathbb{N}$.
4. Assigned validity.
5. $K\phi$, whenever ϕ is an instance of any of lines 1–4 or (recursively) 5.

Lemma 8. For every $n \in \mathbb{N}$ and every $\phi \in \Sigma(n)$, ϕ is a sentence.

Proof. By inspection. □

Lemma 9. There is a total computable function $f : \mathbb{N} \to \mathbb{N}$ such that for every n,

$$W_{f(n)} = \{\langle m, \ulcorner \phi \urcorner\rangle \in \mathbb{N} : \phi \text{ is a formula with } FV(\phi) \subseteq \{x\} \text{ and } \Sigma(n) \models \phi(x|\overline{m})\}.$$

Proof. Follows from Lemma 1 and the Church-Turing Thesis. □

Corollary 10. There is an $n \in \mathbb{N}$ such that

$$W_n = \{\langle m, \ulcorner \phi \urcorner\rangle \in \mathbb{N} : \phi \text{ is a formula with } FV(\phi) \subseteq \{x\} \text{ and } \Sigma(n) \models \phi(x|\overline{m})\}.$$

Proof. By Kleene's Recursion Theorem and Lemma 9. □

Proposition 11. Let n be as in Corollary 10. Then $\mathcal{M}_{\Sigma(n)} \models \Sigma(n)$.

Proof. For brevity, write Σ for $\Sigma(n)$ and $\mathcal{M}$ for $\mathcal{M}_{\Sigma(n)}$.

Claim 1 $\mathcal{M}$ satisfies all instances of $E1$. By Lemma 7.

Claim 2 $\mathcal{M}$ satisfies all instances of $E2$. By Lemma 4.

Claim 3 $\mathcal{M}$ satisfies all instances of $E4$. By Lemma 6.

Claim 4 $\mathcal{M}$ satisfies the axioms of Peano arithmetic for $\mathscr{L}$. By Lemma 5.

Claim 5 For any $\mathscr{L}$-formula ϕ with $FV(\phi) \subseteq \{x\}$, $\mathcal{M}$ satisfies $\forall x(K\phi \leftrightarrow \langle x, \overline{\ulcorner \phi \urcorner}\rangle \in W_{\overline{n}})$.

Let s be an arbitrary assignment (say with $s(x) = m$), we must show $\mathcal{M} \models K\phi[s]$ if and only if $\mathcal{M} \models \langle x, \overline{\ulcorner \phi \urcorner}\rangle \in W_{\overline{n}}[s]$. The following are equivalent:

$\mathcal{M} \models K\phi[s]$	
$\Sigma \models \phi^s$	(Definition of $\mathcal{M}$)
$\Sigma \models \phi(x\|\overline{m})$	(Since $FV(\phi) \subseteq \{x\}$)
$\langle m, \ulcorner \phi \urcorner\rangle \in W_n$	(By choice of n (Corollary 10))
$\mathcal{M} \models \langle \overline{m}, \overline{\ulcorner \phi \urcorner}\rangle \in W_{\overline{n}}$	(Since $\mathcal{M}$ has standard first-order part)
$\mathcal{M} \models (\langle x, \overline{\ulcorner \phi \urcorner}\rangle \in W_{\overline{n}})^s$	(Since $s(x) = m$)
$\mathcal{M} \models \langle x, \overline{\ulcorner \phi \urcorner}\rangle \in W_{\overline{n}}[s]$.	(By Lemma 3)

Claim 6 $\mathcal{M}$ satisfies all instances of assigned validity. Suppose ϕ is valid and s is an assignment, we must show $\mathcal{M} \models \phi^s$. By Lemma 3, it suffices to show $\mathcal{M} \models \phi[s]$. But this is immediate, because ϕ is valid.

Claim 7 $\mathcal{M} \models K\phi$ whenever $K\phi$ is an instance of line 5 from Definition 8. For any such $K\phi$, ϕ itself lies in Σ, so $\Sigma \models \phi$. Let s be any assignment. By Lemma 8, ϕ is a sentence, thus $\phi^s = \phi$ and so $\Sigma \models \phi^s$, meaning $\mathcal{M} \models K\phi[s]$. □

Theorem 12. Let n be as in Corollary 10.

1. $\mathcal{M}_{\Sigma(n)}$ satisfies the axioms of epistemic arithmetic mod factivity.
2. $\mathcal{M}_{\Sigma(n)}$ satisfies all instances of Reinhardt's schema, that is,

$$\exists e K \forall x (K\phi \leftrightarrow x \in W_e)$$

   whenever $FV(\phi) \subseteq \{x\}$.
3. Additionally, there is a fixed $m \in \mathbb{N}$ such that $\mathcal{M}_{\Sigma(n)}$ satisfies the schema $K(K\phi \leftrightarrow \ulcorner \phi \urcorner \in W_{\overline{m}})$, where ϕ ranges over $\mathscr{L}$-sentences.

Thus, the machine that knows the things known by $\mathcal{M}_{\Sigma(n)}$ is a machine that knows its own code.

Proof.

(1). The only axiom schema that remains to be proven is $E3$, the universal closures of formulas of the form $K\phi \to \phi$. Suppose s is any assignment and $\mathcal{M}_{\Sigma(n)} \models K\phi[s]$. This means $\Sigma(n) \models \phi^s$. By Proposition 11, $\mathcal{M}_{\Sigma(n)} \models \Sigma(n)$, therefore $\mathcal{M}_{\Sigma(n)} \models \phi^s$. By Lemma 3, $\mathcal{M}_{\Sigma(n)} \models \phi[s]$, establishing (1).

(2) and (3). By combining lines 3 and 5 of Definition 8, $\Sigma(n)$ contains $K\forall x(K\phi \leftrightarrow \langle x, \overline{\ulcorner \phi \urcorner} \rangle \in W_{\overline{n}})$ whenever ϕ is an $\mathscr{L}$-formula with $FV(\phi) \subseteq \{x\}$. (2) and (3) follow. □

4 Conclusion and Related Work

A knowing machine (implicitly meaning, a knowing machine that knows its own factivity) cannot know its own code. Carlson showed that such a machine can know that it has *some* code (without knowing exactly which). Our result complements Carlson's: it is possible for a machine to know its code quite precisely, at the price of knowing its factivity (despite really being factive).

In our dissertation [1] we explore related issues surrounding multiple interacting machines. Suppose $\prec$ is an r.e. well-founded partial ordering of $\mathbb{N}$.

- There are machines $M_0, M_1, \ldots$ such that each M_i knows precise codes of each M_j, and knows factivity of M_j when $j \prec i$.
- There are machines $M_0, M_1, \ldots$ such that each M_i knows precise codes of M_j when $j \prec i$; factivity of M_j when $j \preceq i$; and each M_i knows that each M_j has some code (without necessarily knowing which).
- There are machines $M_0, M_1, \ldots$ such that each M_i knows precise codes of M_j ($j \prec i$); factivity of M_j ($j \preceq i$); a slight weakening of factivity of M_j (all j); and that M_j has some code ($j \preceq i$).
- But if $\prec$ is ill-founded, there are no such machines as above, provided the machines are also required to know rudimentary facts about computable ordinals.

We are preparing a streamlined paper on these results for journal submission.

References

[1] Alexander, S. (2013). The Theory of Several Knowing Machines. Dissertation, The Ohio State University.

[2] Benacerraf, P. (1967). God, the Devil, and Gödel. *The Monist*, **51**, 9–32.

[3] Carlson, T.J. (1999). Ordinal arithmetic and Σ_1 elementarity. *Archive for Mathematical Logic*, **38**, 449–460.

[4] Carlson, T.J. (2000). Knowledge, machines, and the consistency of Reinhardt's strong mechanistic thesis. *Annals of Pure and Applied Logic*, **105**, 51–82.

[5] Carlson, T.J. (2001). Elementary patterns of resemblance. *Annals of Pure and Applied Logic*, **108**, 19–77.

[6] Carlson, T.J. (2012). Sound Epistemic Theories and Collapsing Knowledge. Slides from the *Workshop on The Limits and Scope of Mathematical Knowledge* at the University of Bristol.

[7] Lucas, J.R. (1961). Minds, machines, and Gödel. *Philosophy*, **36**, 112–127.

[8] Penrose, R. (1989). *The Emperor's new mind: concerning computers, minds, and the laws of physics*. Oxford University Press.

[9] Putnam, H. (2006). After Gödel. *Logic Journal of the IGPL*, **14**, 745–754.

[10] Reinhardt, W. (1985). Absolute versions of incompleteness theorems. *Noûs*, **19**, 317–346.

[11] Shapiro, S. (1985). Epistemic and intuitionistic arithmetic. In Shapiro S. (ed), *Intensional Mathematics*, pp. 11–46. Elsevier.