Finite Circular Definitions

Anil Gupta
University of Pittsburgh

A finite circular definition is a particularly simple kind of circular definition. The content of a circular definition is given by a revision process that the definition generates. *Finite* circular definitions are those that have finite revision processes: the process of revision is complete, in the sense made precise below, in finitely many stages. Finite circular definitions prove to be useful in at least two ways. First, as André Chapuis has shown, they can be used to construct theories of rational choice for certain kinds of games. The resulting theories improve on earlier ones in some important respects.[1] Second, in view of their relatively simple behavior, finite circular definitions provide a testing ground for alternative treatments of circularity and interdependency. Some treatments that are otherwise attractive become distinctly less plausible when they are applied to finite circular definitions.[2]

My aim in this essay is to record some properties of finite circular definitions. I show that finiteness of a definition is not, in general, effectively decidable (Theorem 5, below). However, first-order finite circular definitions are effectively enumerable (Theorem 15). I show also that finite definitions do not add to the expressive power of the language (Theorem 14). Maricarmen Martinez has shown that if a first-order circular definition is built using at most identity, names, and one-place predicates then it is bound to be finite. I outline a proof of this claim (Theorem 18); a detailed proof may be found in Martinez 2001.

Let us begin with some basic concepts and terminology. Let L be a first-order language, with or without identity. I shall assume that L is equipped with an effective enumeration of its non-logical constants. Let R by an n-ary predicate that does not occur in L. Let L^+ be the language that results when L is extended with R. $\mathcal{D}$ is a *definition* (*possibly circular*) *of* R *in* L iff $\mathcal{D}$ is of the form,

$$R(x_1, \ldots, x_n) =_{Df} A(x_1, \ldots, x_n, R),$$

where $A(x_1, \ldots, x_n, R)$ is a formula of L^+ and has no variables free in it other than $x_1, \ldots, x_n$. R is the *definiendum of* $\mathcal{D}$, and $A(x_1, \ldots, x_n, R)$ is its *definiens*. A definition whose definiendum is a one-place (two-place, n-place) predicate will be called a *monadic* (*dyadic*, *n-adic*) definition.[3] A

---

[1]See Chapuis 2000; see also Gupta 2000.

[2]I think this holds, for example, of the supervaluation treatment modeled on Kripke 1975 and of the revision theory $\mathbf{S}^*$. For $\mathbf{S}^*$, see §5D of Nuel Belnap's and my book *The Revision Theory of Truth* (henceforth cited as *RTT*).

[3]In this essay I shall be concerned only with *circular* definitions of *predicates*. Indeed, most of the essay will be devoted to monadic definitions. The principal results stated below hold

structure or model M (= <D, I>), where D is the domain of M and I provides interpretations for the non-logical constants of L, will sometimes be called a *ground model of* L. L itself will sometimes be called a *ground* language. Subsets h of D^n will be called *hypotheses* (for the interpretation of R in M).[4] Given a ground model M and a hypothesis h, M + h is the structure (of L^+) just like M except that R is assigned the interpretation h. Given a ground model M, $\mathcal{D}$ yields a rule of revision $\delta_{\mathcal{D},M}$ that is defined thus: $\delta_{\mathcal{D},M}$ is a function from D^n into D^n such that, for all h $\subseteq D^n$,

$$<d_1, \ldots, d_n> \in \delta_{\mathcal{D},M}(h) \text{ iff } <d_1, \ldots, d_n> \text{ satisfies } A(x_1, \ldots, x_n, R) \text{ in } M + h.$$

Sometimes, when the context allows it, I shall drop the subscript $\mathcal{D}$ and write δ_M in place of $\delta_{\mathcal{D},M}$. A revision rule yields revision sequences: $<h_i>_{0 \le i}$ is a *revision sequence for* $\mathcal{D}$ *in* M [or *for* $\delta_{\mathcal{D},M}$] iff, for all i ≥ 0, $h_{i+1} = \delta_{\mathcal{D},M}(h_i)$; similarly for j-long sequences $<h_i>_{0 \le i < j}$.[5] $\delta_{\mathcal{D},M}{}^n(h)$ is the hypothesis that results after n successive applications of $\delta_{\mathcal{D},M}$ to h. So, $\delta_{\mathcal{D},M}{}^0(h) = h$, and $\delta_{\mathcal{D},M}{}^{n+1}(h) = \delta_{\mathcal{D},M}(\delta_{\mathcal{D},M}{}^n(h))$. A hypothesis h is p-*reflexive for* $\mathcal{D}$ *in* M [or *for* $\delta_{\mathcal{D},M}$] iff $h = \delta_{\mathcal{D},M}{}^p(h)$; h is *reflexive for* $\mathcal{D}$ *in* M [or *for* $\delta_{\mathcal{D},M}$] iff, for some p > 0, h is p-reflexive for $\delta_{\mathcal{D},M}$. Observe that if $\delta_{\mathcal{D},M}{}^n(h)$ is reflexive then, for all m $\geq$ n, $\delta_{\mathcal{D},M}{}^m(h)$ is reflexive also. We can now define the central concept:

1. Definition. Let $\mathcal{D}$ be a definition in a language L. Then, $\mathcal{D}$ is *finite in* L iff, for all ground models M of L there is a number n such that for all hypotheses h, $\delta_{\mathcal{D},M}{}^n(h)$ is reflexive.

Note the order of quantifiers in this definition: ∀M∃n∀h. We shall see below that, for first-order languages, the order of the quantifiers can be switched (Corollary 11). That is, the following three conditions are equivalent:

$\forall M \exists n \forall h[\delta_{\mathcal{D},M}{}^n(h)$ is reflexive],
$\exists n \forall M \forall h[\delta_{\mathcal{D},M}{}^n(h)$ is reflexive],
$\forall M \forall h \exists n[\delta_{\mathcal{D},M}{}^n(h)$ is reflexive].

We shall also see that, for a finite definition $\mathcal{D}$, a bound p can be placed on the cyclicity exhibited by reflexive hypotheses: any reflexive hypothesis for $\mathcal{D}$ in any model is bound to be q-reflexive for some q such that 0 < q $\leq$ p. The revision process for a finite definition is complete, then, after finitely many stages: only finitely many stages are needed to isolate all the reflexive hypotheses. These hypotheses play a central role in the semantics of finite definitions.

---

for systems of interdependent definitions, provided that their definienda are finite in number. Further, circular and interdependent definitions can be given for names and function symbols also. The theory of these definitions parallels that of predicates.

[4]I shall often suppress the parenthetical clause.

[5]The general theory of definition requires consideration of transfinite revision sequences. In this paper, we shall need to consider only revision sequences that are at most ω-long.

2. Definitions. Let $\mathcal{D}$ be a finite definition in L, M (= <D, I>) a ground model of L, Z ⊆ D, B a sentence of L^+, and C(x) a formula of L^+.

(2.1) B is *valid in* M *relative to* $\mathcal{D}$ iff, for all hypotheses h reflexive for $\mathcal{D}$ in M, B is true in M + h; B is *valid* (*simpliciter*) *relative to* $\mathcal{D}$ iff B is valid relative to $\mathcal{D}$ in all ground models.[6]

(2.2) C(x) *weakly defines* Z *in* M *relative* to $\mathcal{D}$ iff C(x) has exactly one free variable x and, for all d ∈ D, d ∈ Z iff, for all reflexive hypotheses h for $\mathcal{D}$ in M, C(x) is true of d in M + h. C(x) *strongly defines* Z *in* M *relative to* $\mathcal{D}$ iff C(x) and ¬C(x) weakly define, respectively, Z and D \ Z in M relative to $\mathcal{D}$. Z is *weakly definable in* M *relative to* $\mathcal{D}$ iff there a formula of L^+ that weakly defines Z in M relative to $\mathcal{D}$; similarly for strong definability.

Remark. The above explanation of validity is cogent only for finite definitions, not for definitions in general. If the explanation were extended to all definitions, it would result in violations of the conservativeness requirement on definitions (see *RTT*, §5A).

3. Examples.

(3.1) Every non-circular definition – i.e., definition whose definiens contains no occurrences of the definiendum – is finite.

(3.2) An example of a circular definition that is finite is this:

$$Gx =_{Df} [x = a \wedge (Ga \vee Gb)] \vee [x = b \wedge \neg Ga \wedge \neg Gb].^{7}$$

Let a and b denote respectively the objects **a** and **b** in the domain of a ground model. The revision rule for the above definition revises any hypothesis that contains either **a** or **b** to {**a**}, and it revises all other hypotheses to {**b**}. The definition is plainly finite: by the second stage of revision we are bound to have the 1-reflexive hypothesis {**a**}. {**a**} is thus the unique reflexive hypothesis, and Ga is valid. Observe that ¬Gb is valid in those ground model in which **a** and **b** are distinct, but not in those in which they are identical. Thus ¬Gb is not valid *simpliciter*. Observe also that weak definability coincides with first-order definability in the ground model.

(3.3) Consider the definition

$$Rxy =_{Df} \exists z(Rxz \wedge Rzy).$$

This definition is *pure* in the sense that it uses no non-logical vocabulary from the ground

---

[6]The notion of validity defined here is equivalent to validity under the system **S**[#] of *RTT*, except that **S**[#] treats all definitions, finite and non-finite. In this paper, our topic is essentially **S**[#] restricted to finite definitions.

[7]This is Example 5A.11 from *RTT*.

language. Consider any ground model with the set N of natural numbers as it domain. The revision sequence $<h_i>_{0 \leq i}$, where h_0 is the successor relation $\{<z, z+1>: z \in N\}$, is plainly non-repeating. Hence, no hypothesis in the sequence is reflexive. So our definition is not finite.

(3.4) Consider the ordered field of real numbers $\mathfrak{R}$ and its associated language $L_{\mathfrak{R}}$, which contains the constants 0, 1, +, ×, ≤. Let *Closure* abbreviate the formula

$$G0 \wedge \forall x(Gx \rightarrow Gx{+}1).$$

Set $\mathcal{D}$ to be the following monadic definition:

$$Gx =_{Df} [\neg \mathit{Closure} \wedge x = x] \vee [\mathit{Closure} \wedge Gx].$$

Observe that if an hypothetical interpretation of G fails to satisfy *Closure* then the revision rule revises it to the entire domain; otherwise, the revision rule leaves the initial hypothesis unchanged. Therefore $\mathcal{D}$ is finite. Further, it has in $\mathfrak{R}$ uncountably many reflexive hypotheses: any hypothesis that satisfies closure is 1-reflexive. Neither the sentence

$$\exists z(z \neq 0 \wedge z \neq 1 \wedge 0 \leq z \leq 1 \wedge Gz)$$

nor its negation is, therefore, valid in $\mathfrak{R}$. But the sentences G0, G1, $\forall x(Gx \rightarrow Gx{+}1)$ are valid in $\mathfrak{R}$ and, indeed, in all ground models. The set N of natural numbers is weakly definable in $\mathfrak{R}$. But N is not first-order definable in $\mathfrak{R}$. For, if it were so definable then the first-order theory of $\mathfrak{R}$ would end up being undecidable, contradicting a theorem of Alfred Tarski. We shall see below that any set that is strongly definable in a structure by a finite definition is definable already in the ground language. It follows that N is not strongly definable.

There is a natural calculus, $\mathbf{C}_0$, for reasoning with finite definitions. $\mathbf{C}_0$ is a natural deduction system and has the classical introduction and elimination rules for connectives and quantifiers. Further, $\mathbf{C}_0$ has two special rules for definitions: Definiendum Introduction (DfI_r) and Definiendum Elimination (DfE_r).[8] These rules are weaker than the standard rules for the definitions, in the following way. Each step in a derivation in $\mathbf{C}_0$ is assigned an integer as its index. Intuitively, this index may be viewed as representing a stage in the revision process. An applications of DfI_r or DfE_r results in a shift in this index. Suppose, for example, that we have the definition

$$Gx =_{Df} A(x, G).$$

Then given A(t, G) at a step with an index i, DfI_r allows us to infer Gt with the index i+1.[9] DfE_r is

[8]The subscript 'r' indicates that the rules are designed for *r*evision semantics and that they are *r*estricted versions of the classical rules.

[9]Note that A(t, G) is the result of replacing free occurrences of x in A(x, G) by the term t, with the proviso that bound variables in A(x, G) are so changed that the substituted occurrences

the converse of this: from Gt with the index i+1 it allows us to infer A(t, G) with the index i. There is one special rule governing indices in $\mathbf{C}_0$, the rule of Index Shift: the indices of formulas without any occurrences of the definienda can be arbitrarily shifted. Aside from these three rules – DfE_r, DfI_r, and Index Shift –, no other rules of $\mathbf{C}_0$ involve shifting of indices. This is so, in particular, for the introduction and elimination rule for connectives and quantifiers. The rule for Conjunction Introduction, for example, is this: given formulas B and C, both with the index i, we may infer (B & C) with the index i. A sentence A is *derivable from* a definition $\mathcal{D}$ in $\mathbf{C}_0$ iff for some integer i there is a derivation in $\mathbf{C}_0$ of A with the index i from $\mathcal{D}$. For a more detailed presentation of $\mathbf{C}_0$ and for sample derivations, see *RTT*, §5B.

4. Theorem. (Soundness and completeness) A sentence B of L^+ is valid relative to a finite definition $\mathcal{D}$ iff B is derivable in $\mathbf{C}_0$ from $\mathcal{D}$.

Proof. This is an immediate consequence of Theorem 5B.1 of *RTT* : Validity, as defined above, is equivalent, over finite definitions, to validity in the system $\mathbf{S}_0$ of *RTT*. And $\mathbf{C}_0$ is sound and complete with respect to $\mathbf{S}_0$. □

Finite definitions have, then, not only a simple semantics but also a simple calculus. It is this that makes them a useful tool in theorizing about phenomena involving circularity and interdependency (e.g., truth and rational choice). I know of no other semantic treatment of finite definitions that simultaneously preserves classical logic, avoids hierarchies, and at the same time comes with a natural calculus.[10]

In view of their model behavior, it is a natural question whether finite definitions can be characterized in a non-semantical way. More specifically, can we somehow separate out those definientia A(x, G) that yield finite definitions from those that do not?

5. Theorem. There is no effective method for deciding whether a monadic definition in the language of arithmetic $L_{\mathbb{N}}$ is finite.

Proof. Consider the theory **Q** of Robinson's Arithmetic. As is well known, all recursive functions and relations are representable in **Q**. Consider a recursive binary relation R(x, y) such that $\exists$yR(x, y) is *not* recursive. For instance, R(x, y) may be the relation that y is the Gödel number of a proof in **Q** of a formula with Gödel number x. Let A(u, v) be a formula of $L_{\mathbb{N}}$ that represents R(x, y). For all natural numbers p and q, we have therefore the following:

---

of x are free for t.

[10]The calculus $\mathbf{C}_0$ is not sound if finite definitions are evaluated using the supervaluation scheme in the manner of Kripke 1975. A counterexample is provided by the definition given in Example 3.2: sentences Ga and ¬Gb are not grounded truths (i.e., they are not true in the least fixed point supervaluation model), but they are provable in $\mathbf{C}_0$.

Question: Is there a natural sound and complete calculus for finite definitions under the supervaluation semantics?

(1) If $R(p, q)$ holds then $\vdash_{\mathbf{Q}} A(\mathbf{p}, \mathbf{q})$,

and

(2) If $R(p, q)$ fails then $\vdash_{\mathbf{Q}} \neg A(\mathbf{p}, \mathbf{q})$,

where bold **p** and **q** indicate numerals in $L_{\mathbb{N}}$ for p and q respectively.

We prove the theorem by constructing effectively, for each natural number n, a definition $\mathcal{D}_n$ such that

(*) $\exists y R(n, y)$ holds iff $\mathcal{D}_n$ is finite.

Since the construction of $\mathcal{D}_n$ is effective, (*) suffices to establish the theorem: if there were an effective procedure for deciding whether a definition is finite, there would be an effective procedure for deciding whether a number satisfied $\exists y R(x, y)$. But this latter is, by hypothesis, impossible.

We construct the definitions $\mathcal{D}_n$ as follows. Let Q abbreviate the conjunction of the axioms of Robinson's Arithmetic and let B abbreviate the formula

(B) $\exists z \exists y(Gz \wedge y \leq z \wedge A(\mathbf{n}, y))$,

where $\leq$ is defined thus:

$$y \leq z \leftrightarrow \exists u(u + y = z).$$

Then $\mathcal{D}_n$ is:

($\mathcal{D}_n$) $Gx =_{\mathrm{Df}} Q \wedge \neg B \wedge [x = \mathbf{0} \vee \exists y(Gy \wedge x = y')]$.

Let the rule of revision generated by $\mathcal{D}_n$ in an interpretation M be δ_M. It is easy to see that, for arbitrary interpretation M and arbitrary hypothesis h,

(3) If Q is false in M then $\delta_M(h) = \varnothing$.

(4) If B is true in M + h, then $\delta_M(h) = \varnothing$.

Further, for the standard model of arithmetic $\mathbb{N}$ and for arbitrary hypothesis h,

(5) If B is false in $\mathbb{N}$+h then $\delta_{\mathbb{N}}(h) = \{0\} \cup \{p' \mid p \in h\}$.

We shall describe the behavior of δ_M in nonstandard models of Q below. Let us now verify that (*) holds.

($\Leftarrow$ part) Suppose R(n, y) holds for no natural number y. By (2), A(**n**, **m**) is false in $\mathbb{N}$ for all numbers m. Hence, for any arbitrary hypothesis h, B is bound to be false in $\mathbb{N}$ + h. So (5) applies, and we have

$$\delta_{\mathbb{N}}(h) = \{0\} \cup \{p' \mid p \in h\}.$$

The following is therefore a revision sequence of $\delta_{\mathbb{N}}$:

$$\varnothing, \{0\}, \{0, 1\}, \{0, 1, 2\}, \ldots .$$

None of the hypotheses in this sequence is reflexive. Hence, $\mathcal{D}_n$ fails to be finite.

($\Rightarrow$ part) Suppose R(n, y) holds for some natural number y. Let m be the least such number. We claim that for all M and h,

(#) $\quad\delta_M^{m+1}(h)$ is reflexive.

This suffices to show that $\mathcal{D}_n$ is finite.

Now if Q is false in M then, by (3), $\delta_M(h) = \varnothing$ for all h. It follows that $\varnothing$ is reflexive and that (#) holds. So we need only verify (#) for those M in which Q is true. Note that such an M may be a nonstandard model of Q and may contain in its domain objects other than (copies of) the natural numbers. We know, however, that theorems of **Q** are true in M. Note in particular the following:

(Fact A) $\quad\vdash_{\mathbf{Q}} \mathbf{p} \neq \mathbf{q}$, if $p \neq q$,

(Fact B) $\quad\vdash_{\mathbf{Q}} \forall x(x \leq \mathbf{p} \leftrightarrow x = \mathbf{0} \vee \ldots \vee x = \mathbf{p})$, for all p,

(Fact C) $\quad\vdash_{\mathbf{Q}} \forall x(x \leq \mathbf{p} \vee \mathbf{p} \leq x)$, for all p.

Fact A tells us that the denotations in M of **0**, **0'**, **0''**, . . . , **p**, . . . are all distinct from each other. Call these denotations the *standard numbers of* M and call all the other objects in the domain of M the *nonstandard numbers of* M. Let $\underline{p}$ be the denotation of **p** in M. Let us now consider how the objects in M are ordered by the interpretation of $\leq$ in M. Call this interpretation $\leq^*$. Fact B tells us that

$$\underline{p} \leq^* \underline{q} \text{ iff } p \leq q.$$

That is, the standard numbers are ordered in the standard way in M. What about the nonstandard numbers? By Fact C, every nonstandard number is either above or below any given standard number $\underline{p}$. It cannot be below, for then (by Fact B) it would be identical to one of the standard numbers, which is impossible. Hence, the nonstandard numbers are greater than all the standard numbers. We thus have the following picture:

$$\underline{0} \leq^* \underline{1} \leq^* \underline{2} \leq^* \underline{3} \leq^* \ldots \leq^* \text{ the nonstandard numbers.}$$

Now, by our assumption, m is the least number for which R(n, y) holds. Hence, by (1) and (2), we have

$\vdash_{\mathbf{Q}} A(\mathbf{n}, \mathbf{m})$ and, for all $p < m$, $\vdash_{\mathbf{Q}} \neg A(\mathbf{n}, \mathbf{p})$.

Consequently, $A(\mathbf{n}, \mathbf{m})$ is true in M and, for all $p < m$, $A(\mathbf{n}, \mathbf{p})$ is false in M. This, in conjunction with the properties we have noted of the models of **Q**, yields that

(6) B is false in M + h iff $h \subseteq \{\underline{q} \mid q < m\}$.

Hence, by definition $\mathcal{D}_n$,

(7) If $h \subseteq \{\underline{q} \mid q < m\}$ then $\delta_M(h) = \{\underline{0}\} \cup \{\underline{p'} \mid \underline{p} \in h\}$.

Observe that, by (4), (6), and (7),

(8) $\delta_M(h) = \varnothing$ iff $h \not\subseteq \{\underline{q} \mid q < m\}$.

Now consider the revision sequence generated by δ_M when the initial hypothesis is $\varnothing$. By (7) and (8), the sequence looks like this:

$\varnothing, \{\underline{0}\}, \{\underline{0}, \underline{1}\}, \{\underline{0}, \underline{1}, \underline{2}\}, \ldots\ldots, \{\underline{0}, \underline{1}, \ldots \underline{m}\}, \varnothing$.

So $\varnothing$ is reflexive. And (#) holds for M when h is $\varnothing$. What if h is nonempty? Now, if $h \not\subseteq \{\underline{q} \mid q < m\}$ then, by (8), $\delta_M(h) = \varnothing$. And again (#) holds. Suppose, on the other hand, that $h \subseteq \{\underline{q} \mid q < m\}$. Now all members of h are standard numbers of M, and there must be one that is the greatest (under $\leq^*$ ordering). Let this be $\underline{p}$. By (7), $\underline{p+1} \in \delta_M(h)$. More generally, it follows by induction that, for all $q \leq m - p$,

$\underline{p+q} \in \delta_M{}^{q}(h)$.

In particular, $\underline{m} \in \delta_M{}^{m-p}(h)$. That is,

$\delta_M{}^{m-p}(h) \not\subseteq \{\underline{q} \mid q < m\}$.

So, by (8), $\delta_M{}^{(m-p)+1}(h) = \varnothing$. Hence $\delta_M{}^{(m-p)+1}(h)$ is reflexive. Since $(m - p) + 1 \leq m + 1$, claim (#) holds again. □

Remarks. (i) A small modification of the above argument yields a stronger result: the undecidability of finite monadic definitions in any first-order language L with identity and a binary predicate. Treat the binary predicate as set-theoretic membership and conduct the above argument not with **Q** but with the finitely many axioms of **ZFC** needed to derive the translations of the axioms of **Q**. The problem of deciding whether a number belongs to an r.e. set can be reduced, as above, to the problem of deciding whether a monadic definition in L is finite.

(ii) José Martinez has observed that the claim holds also for first-order languages L' with just one binary predicate R. Conduct the argument given in (i) with R as set-theoretic membership and with indiscernibility in place of identity, where indiscernibility ($\approx$) is defined thus:

$$x \approx y =_{Df} \forall z(zRx \leftrightarrow zRy) \wedge \forall z(xRz \leftrightarrow yRz).$$

(iii) A simple argument establishes the undecidability of finite dyadic definitions in the language L' of (ii). Let S be a binary predicate distinct from R and let A be a sentence of L'. Then the definition

$$Sxy \ =_{Df} \neg A \wedge \exists z(Sxz \wedge Szy)$$

is finite if and only if A is a logical truth of L'. Since logical truth is not effectively decidable, the same holds for finite dyadic definitions.

6. Problem. Is there an effective method for deciding whether a *pure* dyadic definition in a first-order language (with or without identity) is finite?

We shall show that monadic finite definitions in a first-order language are recursively enumerable. As a step in this argument, we show that for these definitions there is a finite bound by which revision reaches reflexive hypotheses irrespective of the *ground model* and the initial hypothesis. Furthermore, there is similar bound on the length of the cycle that reflexive hypotheses exhibit. The following terminology will prove useful.

7. Definitions. A natural number m is an *initial number of $\mathcal{D}$* iff m is the least number such that, for all ground models M and all hypotheses h, $\delta_{\mathcal{D},M}{}^{m}(h)$ is reflexive. A natural number n is a *cyclic number of $\mathcal{D}$* iff n is the least number such that, for all ground models M and all reflexive hypothesis h of $\delta_{\mathcal{D},M}$, h is p-reflexive for some p such that $0 < p \leq n$. A pair <m, n> is a *revision index* of $\mathcal{D}$ iff m is the initial number and n the cyclic number of $\mathcal{D}$.

We show that every finite definition has a revision index. But first a preliminary definition and a lemma.

8. Definition. A definition $\mathcal{D}$ *has an* ω*-long* [*n-long*] *non-repeating revision sequence* iff there is a ground model M and a hypothesis h such that, the sequence $\langle\delta_{\mathcal{D},M}{}^{q}(h)\rangle_{0 \leq q}$ [$\langle\delta_{\mathcal{D},M}{}^{q}(h)\rangle_{0 \leq q < n}$] is non-repeating – in other words, for all i, j < ω [i, j < n], if $i \neq j$, then $\delta_{\mathcal{D},M}{}^{i}(h) \neq \delta_{\mathcal{D},M}{}^{j}(h)$.

9. Lemma. If a definition $\mathcal{D}$ has, for each p, a p-long non-repeating revision sequence then $\mathcal{D}$ has an ω-long non-repeating revision sequence.

Proof. Suppose $\mathcal{D}$ is a definition, say,

$$G(x_1, \ldots, x_m) \ =_{Df} A(x_1, \ldots, x_m, G),$$

that has, for each p, a p-long non-repeating revision sequence. Consider a language L* just like the ground language L but with countably many new m-place predicates G_i ($0 \leq i$). Let E_i

designate the following formula of L*:

(E_i) $\forall x_1 \ldots \forall x_m(G_{i+1}(x_1, \ldots, x_m) \leftrightarrow A(x_1, \ldots, x_m, G_i))$

And let C_i designate the formula,

$$(C_i) \quad \neg\forall x_1 \ldots \forall x_m(G_{i+1}(x_1, \ldots, x_m) \leftrightarrow G_0(x_1, \ldots, x_m)) \wedge$$
$$\neg\forall x_1 \ldots \forall x_m(G_{i+1}(x_1, \ldots, x_m) \leftrightarrow G_1(x_1, \ldots, x_m)) \wedge$$
$$\vdots$$
$$\neg\forall x_1 \ldots \forall x_m(G_{i+1}(x_1, \ldots, x_m) \leftrightarrow G_i(x_1, \ldots, x_m)).$$

Recursively define theories Γ_n thus:

$$\Gamma_0 = \varnothing$$
$$\Gamma_{n+1} = \Gamma_n \cup \{E_n, C_n\}.$$

Set $\Gamma = \bigcup_{0 \leq n}\Gamma_n$. We show that each Γ_n is consistent. Γ_0 is obviously consistent; hence we may suppose that n > 0. We know that $\mathcal{D}$ has an (n+1)-long non-repeating revision sequence. So there is a ground model M and a hypothesis h such that $<\delta_{\mathcal{D},M}{}^i(h)>_{0 \leq i \leq n}$ is non-repeating. Let M + $<\delta_{\mathcal{D},M}{}^i(h)>_{0 \leq i}$ (= M*) be the model of L* that agrees with M on the constants in L and that assigns to each G_i the interpretation $\delta_{\mathcal{D},M}{}^i(h)$. Plainly, each E_i is true in M*, and, for each j such that $0 \leq j < n$, so also is C_j. We conclude therefore that each Γ_n is consistent. By the Compactness Theorem, Γ is consistent as well.

Since Γ is consistent, there is a model of L* in which all members of Γ are true. This model can be represented as M + $<h_i>_{i \leq 0}$, where M is a ground model and h_i is the interpretation of G_i. Since each C_i is true in M + $<h_i>_{i \leq 0}$, $<h_i>_{i \leq 0}$ is a non-repeating sequence. Since each E_i is true in M + $<h_i>_{i \leq 0}$, $<h_i>_{i \leq 0}$ is a revision sequence generated by $\delta_{\mathcal{D},M}$. It follows that $\mathcal{D}$ has an ω-long non-repeating revision sequence. □

This lemma yields the desired theorem.

10. Theorem. Every finite definition $\mathcal{D}$ has a revision index.

Proof. Suppose, for reductio, that a finite definition $\mathcal{D}$ fails to have an initial number. Hence, for every natural number p, there is a ground model M and an hypothesis h such that $\delta_{\mathcal{D},M}{}^p(h)$ is not reflexive. It follows that $<\delta_{\mathcal{D},M}{}^q(h)>_{0 \leq q \leq p}$ is a non-repeating revision sequence of $\delta_{\mathcal{D},M}$. Hence, by the previous lemma, $\mathcal{D}$ has an ω-long non-repeating revision sequence. This violates the finiteness of $\mathcal{D}$. A parallel argument shows that $\mathcal{D}$ has a cyclic number. It follows that $\mathcal{D}$ has a revision index. □

11. Corollary. The following three conditions define the same notion of finiteness for first-order languages:

(i) $\exists n\forall M\forall h[\delta_{\mathcal{D},M}{}^{n}(h)$ is reflexive],
(ii) $\forall M\exists n\forall h[\delta_{\mathcal{D},M}{}^{n}(h)$ is reflexive],
(iii) $\forall M\forall h\exists n[\delta_{\mathcal{D},M}{}^{n}(h)$ is reflexive].

Proof. An argument parallel to the one for Theorem 10 shows that (iii) implies (i), and this suffices to establish the corollary. □

Remark. The above equivalences hold because of the compactness property of first-order languages; they fail to hold for second-order languages.

The next lemma and theorem hold for n-adic definitions generally. For notational simplicity, I state them for monadic definitions only. Let $\mathcal{D}$ be the following definition:

$$Gx =_{Df} A(x, G).$$

12. Definition. Set:

$$A^{0}(x, G) = Gx,$$
$$A^{n+1}(x, G) = A^{n}(x, G)[A(t, G)/ Gt],$$

where $A^{n}(x, G)[A(t, G)/ Gt]$ is the result of substituting $A(t, G)$ for all occurrences of Gt in $A^{n}(x, G)$.

13. Lemma. Let M be a ground model, h an arbitrary hypothesis, and d arbitrary member of the domain of M. Then, d satisfies $A^{n}(x, G)$ in M + h iff $d \in \delta_{\mathcal{D},M}{}^{n}(h)$. □

14. Theorem. If $\mathcal{D}$ is finite then all sets that are strongly definable in a ground model M are also first-order definable in M.

Proof. Let C(x, G) be a formula (of L^{+}) that strongly defines a set Z in M relative to $\mathcal{D}$. Let h be an arbitrary reflexive hypothesis for $\mathcal{D}$ in M. Then, for all d in the domain of M,

$d \in Z$ iff C(x, G) is true of d in M + h.

Since, $\mathcal{D}$ is finite, there is a number n such that $\delta_{\mathcal{D},M}{}^{n}(\varnothing)$ is reflexive. Let C*(x, G) be the formula that results by substituting all occurrences of Gt in C(x, G) by $A^{n}(t, G)$. Further, let C**(x) be the formula of L that results from replacing all occurrences of Gt in C*(x, G) by $\perp$ (or any other logically false sentence). We have, in virtue of the substitution theorems of first-order logic, that

$d \in Z$ iff C(x, G) is true of d in $M + \delta_{\mathcal{D},M}{}^{n}(\varnothing)$,
iff C*(x, G) is true of d in $M + \varnothing$,
iff C**(x) is true of d in M.

Thus C**(x) defines Z in M. That is, Z is first-order definable in M. □

Remark. In one sense, then, finite definitions do not add to the expressive power of the ground language.[11] In contrast, non-finite definitions do add to expressive power: they allow sets to be strongly definable that are not first-order definable. This contrast reflects another contrast: the logic of finite definitions is axiomatizable, but that of non-finite definitions, as Philip Kremer (1993) has shown, is not axiomatizable. The failure of axiomatizability is a direct consequence of the expressive power that non-finite definitions bring with them.

15. Theorem. Finite monadic definitions are recursively enumerable.

Proof. For all number n and all numbers m > 0, set $K(\mathcal{D}, n, m)$ to be the following formula:

$$\begin{aligned}
&\forall x(A^n(x, G) \leftrightarrow A^{n+1}(x, G)) \lor \\
&\forall x(A^n(x, G) \leftrightarrow A^{n+2}(x, G)) \lor \\
&\quad\vdots \\
&\forall x(A^n(x, G) \leftrightarrow A^{n+m}(x, G)).
\end{aligned}$$

Lemma 13 implies that, for all ground models M and hypotheses h,

(1) $K(\mathcal{D}, n, m)$ is true in $M + h$ iff $\delta_{\mathcal{D},M}{}^n(h)$ is q-reflexive, for some q such that $0 < q \leq m$.

This yields the following crucial fact:

(2) $K(\mathcal{D}, n, m)$ is a logical truth iff $\mathcal{D}$ is a finite definition with an index $<p, q>$ such that $p \leq n$ and $q \leq m$.

Argument for (2): ($\Rightarrow$ part) Suppose that $K(\mathcal{D}, n, m)$ is a logical truth, and consider an arbitrary ground model M and an arbitrary hypothesis h. By (1), $\delta_{\mathcal{D},M}{}^n(h)$ is q-reflexive, for some q such that $0 < q \leq m$. Further, for every reflexive hypothesis h' there is an hypothesis h'' such that $h' = \delta_{\mathcal{D},M}{}^n(h'')$. It follows that $\mathcal{D}$ has an index $<p, q>$ such that $p \leq n$ and $q \leq m$.

($\Leftarrow$ part) Suppose $\mathcal{D}$ has an index $<p, q>$, $p \leq n$ and $q \leq m$, and consider an arbitrary ground model M and hypothesis h. Now $\delta_{\mathcal{D}, M}{}^n(h)$ is r-reflexive for some r such that $0 < r \leq q$. Hence by (1), $K(\mathcal{D}, n, m)$ is true in $M + h$. Since M and h are arbitrary, $K(\mathcal{D}, n, m)$ is logically true.

Now, the four-place relation "Z is a proof of $K(\mathcal{D}, n, m)$" is recursive. And by fact (2),

$\mathcal{D}$ is finite iff there is a Z, n, and m such that Z is proof of $K(\mathcal{D}, n, m)$.

Since the relation "proof of" is recursive, it follows that finite definitions are recursively enumerable. □

---

[11] In another sense – that captured by weak definability – they do add to expressive power.

Remark. A generalization of the above argument shows that finite n-adic definitions are recursively enumerable also.

16. Problem. Is there a natural syntactic characterization of finite monadic definitions?

This problem is made difficult by the fact that finite definitions are not closed under the usual logical operations: $\neg$, $\rightarrow$, $\leftrightarrow$, $\wedge$, $\forall$, etc. This is illustrated by the following examples. (Further examples may be found in Martinez 2001.)

17. Examples.

(17.1) Let A(x, G) be the formula

$$Gb \vee x = b \vee \exists y(Gy \wedge Syx).$$

Then a definition of Gx with A(x, G) as definiens is finite: by the second stage of revision we invariably obtain the entire domain as the revised hypothesis. But a definition with $\neg$A(x, G) as the definiens is not finite. Consider the structure with natural numbers as the domain that interprets b as 0 and S as the successor relation. The revision rule generates the following infinite revision sequence.

$$\varnothing, N\backslash\{0\}, \{1\}, N\backslash\{0, 2\}, \{1, 3\}, N\backslash\{0, 2, 4\}, \{1, 3, 5\} \ldots.$$

Variants of this example show that finite definitions are not closed under $\rightarrow$, $\leftrightarrow$, and $\forall$. Both $x \neq x$ and A(x, G) yield finite definitions, but ($A(x, G) \rightarrow x \neq x$) and ($A(x, G) \leftrightarrow x \neq x$), being equivalent to $\neg$A(x, G), fail to do so. Further, every instantiation of

$$\forall z[Gz \vee x = z \vee \exists y(Gy \wedge Syx)]$$

with a closed term yields a finite definition, but the formula itself does not do so. This shows that finite definitions are not closed under $\forall$. Finally, observe that both $x \neq x$ and ($x \neq x \rightarrow \neg A(x, G)$) yield finite definitions but not $\neg$A(x, G). Hence, closure under modus ponens fails as well.

(17.2) The following two definiens

$$Gb \vee b = c \vee [x \neq c \wedge (x = b \vee \exists y(Gy \wedge Syx)]$$
$$Gc \vee b = c \vee [x \neq b \wedge (x = c \vee \exists y(Gy \wedge Syx)]$$

yield finite definitions but not their conjunction. Similar examples can be constructed to illustrate the failure of closure under $\vee$ and $\exists$.

Maricarmen Martinez (2001) has investigated the closure properties of finite definitions. She has established the following useful theorem.

18. Theorem. (Maricarmen Martinez) Let $\mathcal{D}$ be a monadic definition in a first-order language with identity. If the non-logical resources of $\mathcal{D}$ consist solely of names and one-place predicates, then $\mathcal{D}$ is finite.

Outline of a proof. Let L be any first-order language with identity whose non-logical resources include at most finitely many names and finitely many one-place predicates. The proof of the theorem rests on two facts about L:

(i) Any set or relation definable in a structure by a formula of L is also definable by a quantifier-free formula of L.

(ii) Only finitely many sets are definable in a structure by quantifier-free formulas of L.

Fact (i) can be established by a quantifier elimination argument. Fact (ii) is a consequence of the finiteness of the non-logical resources of L. Now consider an arbitrary monadic definition $\mathcal{D}$,

$$Gx =_{Df} A(x, G),$$

in L, and an arbitrary ground model M and hypothesis h. By Lemma 13, $\delta_{\mathcal{D},M}{}^{n}(h)$ is definable in M + h by $A^n(x, G)$. It follows by fact (i) that $\delta_{\mathcal{D},M}{}^{n}(h)$ is definable in M + h by a quantifier-free formula of L^+ (= L $\cup$ {G}). By fact (ii), only finitely many sets $h_0, \ldots, h_p$ are definable in M + h by quantifier-free formulas of L^+. Hence,

$$\delta_{\mathcal{D},M}{}^{n}(h) \in \{h_0, \ldots, h_p\}.$$

It follows that, for some $q \leq p$, $\delta_{\mathcal{D},M}{}^{q}(h)$ is reflexive. In view of Corollary 11, it follows that $\mathcal{D}$ is finite. □

References

Chapuis, André 2000. Rationality and circularity. In Chapuis and Gupta 2000, pp. 49-78.

Chapuis, André and Anil Gupta 2000. *Circularity, Definition and Truth*. New Delhi: Indian Council of Philosophical Research. Distributed outside India by Ridgeview Publishing Company, Atascadero CA.

Kripke, Saul 1975. Outline of a theory of truth. *Journal of Philosophy* 72, 690-716. Reprinted in Robert L. Martin, ed., *Recent Essays on Truth and the Liar Paradox*, Oxford: Oxford University Press, pp. 53-81.

Kremer, Philip 1993. The Gupta-Belnap systems $\mathbf{S}^{\#}$ and $\mathbf{S}^{*}$ are not axiomatisable. *Notre Dame Journal of Formal Logic* 34, 583-596.

Gupta, Anil 2000. On circular concepts. In Chapuis and Gupta 2000, pp. 123-153.

Gupta, Anil and Nuel Belnap 1993. *The Revision Theory of Truth*. Cambridge MA: The MIT Press.

Martinez, Maricarmen 2001. Some closure properties of finite definitions. *Studia Logica* **68**, 43-68.