Some closure properties of finite definitions

Maricarmen Martinez[*]
Indiana University

February 14, 2001

Abstract

There is no known syntactic characterization of the class of finite definitions in terms of a set of basic definitions and a set of basic operators under which the class is closed. Furthermore, it is known that the basic propositional operators do not preserve finiteness. In this paper I survey these problems and explore operators that do preserve finiteness. I also show that every definition that uses only unary predicate symbols and equality is bound to be finite.

The Theory of Finite Definitions (henceforth TFD) is a special case of the General Theory of Circular Definitions (henceforth GTCD) as presented by Anil Gupta and Nuel Belnap in their book "The Revision Theory of Truth" [GB 93]. GTCD is concerned with systems $\mathfrak{D}$ of the form

$$\begin{array}{lcl} G_1x_1x_2\dots x_n & =_{df} & A_1(x_1,x_2,\dots,x_n,G_1,\dots,G_k) \\ G_2x_1x_2\dots x_n & =_{df} & A_2(x_1,x_2,\dots,x_n,G_1,\dots,G_k) \\ & \vdots & \\ G_kx_1x_2\dots x_n & =_{df} & A_k(x_1,x_2,\dots,x_n,G_1,\dots,G_k), \\ & \vdots & \end{array}$$

in which the *definiendum* (the left hand side - what is being defined) can appear in the *definiens* (the formula in the right hand side of the definition). Although GTCD works for arbitrary systems, in this paper we will be interested in those which contain only finitely many definitions. Also, for simplicity, most claims here are formulated in terms of unary definitions. It should be understood that at the cost of a more complicated notation, all of those claims could be carried over without difficulty to the realm of definitions with several free variables, unless otherwise stated.

We work here with first order languages $\mathfrak{L}$ based on constant names among $a, b, c\dots$, predicate symbols among $P, Q, R, \dots$, variables among $x, y, z\dots$ and

---

[*] I want to thank Prof. Anil Gupta for teaching me (during a seminar in the Spring 2000) the theory of circular definitions and the theory of truth derived from it, and also for all of his helpful comments on this paper.

the logical connectives $\wedge, \vee, \neg, \rightarrow, \leftrightarrow, \exists, \forall$ and $=$. From such a *ground language* $\mathfrak{L}$ one can obtain *extended languages* $\mathfrak{L}^+$ by adding to $\mathfrak{L}$ defined symbols $G, H, \ldots$ (like the symbols G_i in the definition above).

Correspondingly, in the semantic side there are ground and extended models. A *ground model* M is just a classical model for $\mathfrak{L}$, consisting of a domain D and an interpretation function I which maps n-ary predicate symbols to subsets of D^n $(n \geq 0)$. If $\mathfrak{L}^+ = \mathfrak{L} \cup \{G\}$, then for each ground model M $(= \langle D, I \rangle)$ and set $h \subseteq D$ there is a natural *extended model* $M + h = \langle D, I^+ \rangle$, where I^+ is obtained from I by adding the pair $\langle G, h \rangle$. Therefore, when we add a (unary) symbol G to $\mathfrak{L}$ by means of a definition and we fix a ground model M, there are as many possible extensions of M as there are subsets of G.

Now, when we work in a strictly classical setting in which we ban circular definitions, we can always assign a definite extension to each predicate (either primitive or defined), once we have fixed a ground model M. For if we have the definition $Gx =_{df} Ax$, where G does not appear in the definiens, then the extension of G must be $h = \{c \in M \mid c \text{ satisfies } Ax \text{ in } M\}$. In other words, based on the definition of G we are able to isolate a unique extended model $M + h$ that assigns definite extensions to all predicates in the new extended language. However, if purely circular definitions $Gx =_{df} A(x, G)$ are permitted (so that G appears in the definiens), it is no longer clear how to assign a definite extension to G. In principle, we cannot use the set $\{c \in M \mid c \text{ satisfies } A(x, G) \text{ in } M\}$, since its restricting condition presupposes that we know the extension of G. If we are given a hypothetical extension h' for G though, it makes sense to say that according to the definition the extension of G is $h = \{c \in M \mid c \text{ satisfies } A(x, G) \text{ in } M + h'\}$. The problem is: what if h and h' turn out to be different? Our hope, when looking for a definite extension for G based on the ground model M, is that after "trying" all possible hypotheses h', there is one and only one of them such that h' and h (the revised extension) coincide.

Good or not, that is not always the case. Still, we can claim that we do use circular concepts, some of which cannot be assigned a definite extension. Consider, for example, the following definition.

$$T(x) =_{df} (x = \ulcorner A_1 \urcorner \wedge A_1) \vee (x = \ulcorner A_2 \urcorner \wedge A_2) \vee (x = \ulcorner A_3 \urcorner \wedge A_3) \vee \ldots$$

Here we assume that $A_1, A_2, \ldots$ is an enumeration of all the sentences in $\mathfrak{L}^+$, and that $\ulcorner A_i \urcorner$ is a code (in $\mathfrak{L}$) for A_i. The idea behind this definition is to capture Tarski's intuition about the T-biconditionals as partial definitions of the concept of truth [1]. Suppose now that our language $\mathfrak{L}^+$ allows us to express

---

[1] To be exact, we mean "the *logical* concept of truth", as explained in [GB 93].

If the language $\mathfrak{L}^+$ has only finitely many sentences, then the definiens can be given in a finitary language. In general, though, an infinite disjunction is needed. We use this definition here only as a motivation leading to the concepts to be introduced below. An equivalent way to present the definitions consists on breaking it into many partial definitions:

$$\begin{array}{lcl} T(\ulcorner A_1 \urcorner) & =_{df} & A_1 \\ T(\ulcorner A_2 \urcorner) & =_{df} & A_2 \\ & \vdots & \end{array}$$

the well known sentence

$$\text{(The Liar)} \qquad\qquad \text{The Liar is not true.}$$

A person who is asked to decide whether The Liar is true or false would reason as follows. If The Liar is true, then (this is just what the Liar says) The Liar is not true. Similarly, if The Liar is not true, then it means that The Liar is true. In both cases, after choosing a hypothetical truth status for the sentence we get to the conclusion that the original hypothesis has to be revised, if we want to be coherent with what we mean by "being true".

Let's fix a ground model M now, to see how the paradox is reflected in terms of extended models and our definition above. Notice that if $\ulcorner\text{The Liar}\urcorner$ codes The Liar, then The Liar itself is the sentence

$$\neg T(\ulcorner\text{The Liar}\urcorner).$$

Hence, for this language the definition of truth is circular, because The Liar, being a sentence in our language, has to appear as one of the A_i's.

$$T(x) =_{df} (x = \ulcorner A_1 \urcorner \wedge A_1) \vee \ldots \vee (x = \ulcorner\text{The Liar}\urcorner \wedge \neg T(\ulcorner\text{The Liar}\urcorner)) \vee \ldots$$

Is The Liar true according to this definition of truth? That is, does $\ulcorner\text{The Liar}\urcorner$ satisfy the definiens? First of all, since T appears in the right hand side of the equality, the definiens makes sense only with respect to an extended model $M + h$. Therefore, we are forced to work with an initial hypothesis h about the extension of T. Secondly, it is clear that the definiens is satisfied by $\ulcorner\text{The Liar}\urcorner$ in $M + h$ if and only if $\neg T(\ulcorner\text{The Liar}\urcorner)$ is true on $M + h$. But this is just to say that $\ulcorner\text{The Liar}\urcorner$ is true according to our definition if and only if its interpretation does not belong to h, the hypothetical extension of T. This leads us to discard h as the extension of T. But observe that our hypothesis h here is arbitrary. Therefore, it is impossible to assign a definite extension to truth, because it is impossible to find an extended model $M + h$ in which h, the interpretation of T, coincides with the set of objects that satisfy the definiens. In fact, if we begin to "revise" hypothesis, we observe that we end up with a chain of extended models $M + h_0, M + h_1, M + h_3, M + h_4 \ldots$ in which the corresponding truth status of The Liar is $T, F, T, F, T, \ldots$ or $F, T, F, T, \ldots$, depending on whether the interpretation of $\ulcorner\text{The Liar}\urcorner$ belongs to the initial hypothesis. So, although it is impossible to assign an extension to truth, it is possible to identify, for some sentences, a simple pattern of behaviour in the revision process. In the case of The Liar, the pattern is an oscilation of period two, so to speak.

Behind GTCD there is the belief that circular concepts are interesting and important and that they should not be avoided. GTCD supports that belief by exhibiting sound and complete logical systems in which circular definitions are allowed and where the semantic association assign revision rules - instead of extensions - to defined predicates. Informally speaking, a revision rule describes the process of revision over hypothetical extensions of a defined predicate. A nice feature of GTCD is that it allows us to treat problematic and unproblematic definitions homogeneously. The revision rule yields a definite extension in the case of unproblematic definitions, and provides a description of the pathological behavior of problematic ones, as illustrated by the examples given in the remainder of this section.

Let's proceed to summarize and formalize the main notions. The subsets h used to form the extended models $M + h$ are called *hypotheses*, because when we work in $M + h$, we use h as the hypothetical extension of G when evaluating formulas $A(x, G)$ that appear in the definiens. The intuition is that the set h' of elements $d \in D$ satisfying $A(x, G)$ in $M + h$ is an improvement of the initial hypothesis h. We can repeat this procedure in $M + h'$ to get a new revised hypothesis h'', and so on. In summary, a definition $\mathfrak{D}$ generates a revision process which is formally defined in terms of sequences, as follows.

Definition 0.1 1. *The revision rule generated by the definition* $\mathfrak{D}$ *in the model* M*, and denoted by* $\delta_{\mathfrak{D},M}$*, is a function from subsets of* D *to subsets of* D *given by*

$$\delta_{\mathfrak{D},M}(h) = \{d \in D \mid d \text{ satisfies the definiens } A(x, G) \text{ in } M{+}h\}.$$

2. *Let* $\delta_{\mathfrak{D},M}$ *be a revision rule and* h *a hypothesis. The revision sequence induced by* h *is the sequence* $\langle {\delta_{\mathfrak{D},M}}^n(h)\rangle_{n\geq 0}$*, where*

$$\begin{array}{lcl} {\delta_{\mathfrak{D},M}}^0(h) & = & h \\ {\delta_{\mathfrak{D},M}}^{n+1}(h) & = & \delta_{\mathfrak{D},M}({\delta_{\mathfrak{D},M}}^n(h)) \end{array}$$

The subscripts M and $\mathfrak{D}$ will be ommited here when the context is clear. Let's see some simple examples. In all of them, $M = \langle D, I\rangle$ is an arbitrary ground model.

Example 0.1 *Ordinary definitions.*

Suppose that $\mathfrak{D}$ is the definition $Gx =_{df} A(x)$, where A is a formula in the ground language $\mathfrak{L}$. Let $h \subseteq D$ be a hypothesis and let

$$h' = \{d \in D \mid d \text{ satisfies } A(x) \text{ in the ground model } M\}$$

Then it should be clear that the revision sequence induced by h is $\langle h, h', h', \ldots\rangle$. Thus, the revision process that corresponds to $\mathfrak{D}$ converges to h' (which coincides with the classical extension), no matter what the initial hypothesis h is.

Example 0.2 *Definitions with relative extensions.*

Consider now the definition $\mathfrak{D}$ given by $Gx =_{df} Gx$. For each hypothesis h, the induced revision sequence is the constant sequence $\langle h, h, h, \ldots\rangle$. So in this case the revision sequences are all convergent, but the sequences corresponding to two different initial hypothesis h_1, h_2 converge to different sets (namely, h_1 and h_2 respectively). Hence, we could say that the revision process for definitions like this $\mathfrak{D}$ determine relative extensions, depending on the initial hypothesis.

Example 0.3 *Definitions without relative nor absolute extensions.*

The previous example showed one way in which the behavior of general circular definitions can differ from the behavior of ordinary ones. Here is an example that shows another possibility. Let $\mathfrak{D}$ be the definition $Gx =_{df} \neg Gx$. The revision sequence induced by $h \subseteq D$ is $\langle h,\ D-h,\ h,\ D-h\ldots\rangle$. Here the revision sequence does not converge, so we cannot even get relative extensions. Observe that the behavior of the revision rule for G here is basically the same as in our discussion about The Liar, in the sense that in both situations any revision sequence oscillates between two "opposite" hypothesis. In this example, the revision sequences oscillate between a set and its complement. In the case of The Liar, the revision sequences oscillate between hypothesis containing the interpretation of The Liar and hypothesis that exclude it.

Based on the same principles and following the same strategies that GTCD, The Theory of *Finite* Circular Definitions (TFD) centers its attention in the study of simple circular definitions, like those used in the examples given so far (except the example on truth). These definitions are simple in the sense that, even when a revision sequence does not converge, it is true that after a finite number of steps we can identify a simple pattern, a finite subsequence of hypotheses that repeats over and over again. In other words, there are only finitely many hypotheses that appear in each revision sequence.

Definition 0.2 *Let $\mathfrak{D}$ be the definition given by $Gx =_{df} A(x, G)$.*

1. *Let $M (= \langle M, I\rangle)$ be a ground model, and let $n > 0$. A hypothesis $h \subseteq D$ is n-reflexive for $\mathfrak{D}$ in M if ${\delta_{\mathfrak{D},M}}^n(h) = h$. Moreover, h is reflexive for $\mathfrak{D}$ in M if h is m-reflexive for $\mathfrak{D}$ in M for some $m > 0$.*

2. *The definition $\mathfrak{D}$ is finite if for each ground model M and hypothesis h, there is $n > 0$ such that ${\delta_{\mathfrak{D},M}}^n(h)$ is reflexive for $\mathfrak{D}$ in M.*

Here is an example of a definition which is *not* finite.

Example 0.4 *A definition which is not finite.*

Assume that our ground language $\mathfrak{L}$ has a symbol of binary predicate F and let $\mathfrak{D}$ be the definition $Gx =_{df} \exists y(F(y, x) \wedge Gy)$. Now, consider the ground model M in which D is the set of natural numbers and F is interpreted as the succesor relation. That is, $(a, b) \in I(F) \iff b = a + 1$. Let $h = \{0\}$. Then the revision sequence induced by h is

$$\langle\{0\},\ \{1\},\ \{2\},\ \{3\}\ldots\rangle,$$

which does not converge, but moreover has no duplicate elements, so none of the $\delta^n(h)$ is reflexive for $\mathfrak{D}$ in M.

The notion of finite definitions and the general ideas underlying the corresponding proof system were introduced in [CG 00]. Just as GTCD provides us with sound and complete deduction systems in which circular definitions are allowed, there is also in TFD a (simpler) proof system which is both sound and complete with respect to the class of finite definitions. TFD can be used as a tool to think about truth in simple settings, for example when the language $\mathfrak{L}^+ = \mathfrak{L} \cup \{T\}$ has only finitely many sentences. When it is the case, the definiens in the definition of truth given above is a finite disjunction, and it is easy to see that any revision hypothesis is just a set of sentences [2]. Since there are only finitely many of those sets, the definition of truth is bound to be finite. A less evident application of TFD is motivated by Andre Chapuis' studies on the similarities between the behaviour of rational choice and the behaviour of truth [3]. Chapuis and Gupta claim that the concept of rational choice in the setting of games in strict normal form [4] is not only circular, but also finite.

For an example of a game in strict normal form for which the definition of rational choice is both finite and circular [5], look at the following payoff table. There are two players A and B, a and $\overline{a}$ are the possible actions for player A, and b and $\overline{b}$ are the possible actions for player B. The left number in each box in the table indicates the payoff for player A.

	b	$\overline{b}$
a	2,3	3,2
$\overline{a}$	3,2	2,3

We want to give a definition for the predicate $Rat(x)$, where x is an action [6]. We use a ground language that include symbols of constants $a, \overline{a}, b, \overline{b}$ to represent actions and $\langle ab \rangle, \langle a\overline{b} \rangle, \langle \overline{a}b \rangle, \langle \overline{a}\overline{b} \rangle$ to represent outcomes. To describe the table we also use a binary predicate $<$ and symbols u_A and u_B of unary functions, which

---

[2] ...or a set of interpretations of sentences, if we want to be very formal.

[3] The author does not intend to defend any particular theory about concepts like truth or rational choice here, but to show how revision processes associated with circular definitions have been applied in concrete cases.

[4] In a game in strict normal form there are finitely many players. Each player has a set of finitely many possible actions, which are exhaustive and mutually exclusive. The players choose simultaneously and independently, and they can make their choice based on the payoff table and on the assumption that the other players are rational. A rational agent chooses the action that yields the highest payoff for him (or her). Common knowledge of the game and of rationality of the players is also assumed. Finally, the strictness requirement says that for any player and any possible combination of actions taken by the other players, there is a unique action that yields the maximum payoff.

[5] There is actually a definition of rationality for each particular game. Such a particular definition can be seen as a partial definition of the general concept of Rational Choice, very much like the T-biconditionals are so frequently taken as partial definitions of Truth. For a more detailed explanation about this and other examples, see [CG 00]

[6] The concept of Rational Choice is essentially a three-placed predicate: action x is rational for player y in the setting z. However, by having the sets of actions of different players be disjoint and by fixing the setting (the payoff table, basically) we can work with a unary predicate. Circularities observed in the behaviour of this unary predicate imply circularities in the general definition of Rational Choice. Here I ignore some details which are formally nedded in the partial definition, for simplicity.

intend to describe the utility functions for players A and B. For example, in modelling the table above we would have that $u_A(\langle ab\rangle) < u_B(\langle ab\rangle)$, because the "ab" entry in the table says that the payoffs in for this outcome are 2 and 3 for A and B, respectively. The payoff table above is basically a pictorial description of a ground model. Now, how can we define $Rat(x)$? The following definition captures the intuitions about Rational Choice for games with 2×2-payoff tables:

$$Rat(x) = (x = a \wedge \phi_a) \vee (x = \overline{a} \wedge \neg\phi_a) \vee (x = b \wedge \phi_b) \vee (x = \overline{b} \wedge \neg\phi_b)$$

where ϕ_a is a sentence which says that doesn't matter what is the rational choice for player B, it is always better for A to choose action a. Similarly, ϕ_b says that doesn't matter what is the rational choice for a, b gives the highest payoff to B. For example, ϕ_b can be

$$[Rat(a) \wedge u_B(\langle ab\rangle) > u_B(\langle a\overline{b}\rangle) \vee [\neg Rat(a) \wedge u_B(\langle \overline{a}b\rangle) > u_B(\langle \overline{a}\overline{b}\rangle)].$$

In the particular case of our game, with its particular payoff values, we see that an equivalent definition of Rat is

$$Rat(x) = (x = a \wedge \neg Rat(b)) \vee (x = \overline{a} \wedge Rat(b)) \vee (x = b \wedge Rat(a)) \vee (x = \overline{b} \wedge \neg Rat(a))$$

For example, to get the first disjoint, observe that a yields a highest payoff than $\overline{a}$ only if B chooses $\overline{b}$, that is, only if B does not choose b. Now, this definition that we have just get is clearly circular, it cannot be reduced to an equivalent noncircular definition. Moreover, the definition is finite, simply because at each stage of the revision process the hypothesis for Rat is a set of actions, and there are only finitely many actions in our game.

This report is about TFD, and its main motivation is the question: is there a way to characterize the class of finite definitions on syntactic grounds? Gupta has shown that the set of finite definitions is recursively enumerable but is not recursive, so there is no effective method for deciding whether a definition is finite. Knowing that we cannot look for such a method, a good thing to look at would be a characterization of the class of finite definitions in terms of closure conditions. The goal then would be to identify a set $\mathfrak{B}$ of "basic" finite definitions and a set $\mathfrak{O}$ of "basic" operations on finite definitions such that the class of all finite definitions is the smallest set of definitions that contains $\mathfrak{B}$ and is closed under the operations in $\mathfrak{O}$. The question has not shown to be an easy one to answer, and I do not do it here. Instead, I show some related examples, ruminations and partial results. In Section 1, I survey some easy examples of operations which preserve and which do not preserve finiteness. In section 2 there is a discussion focused on the particular case of the language of arithmetic. The last section is devoted to the proof that any definition that only uses symbols of unary predication and equality is bound to be finite. Based in this result, a bigger class of finite definitions (some of which may include n-ary predicates) is identified.

1 Operators on definitions

In this section I survey some examples that show (the known fact) that finiteness is not closed under negation, conjunction, disjunction nor modus ponens. Those are the most natural operators to start with, but unfortunately none of them works. Later, a couple of operators that do preserve finiteness are introduced; these are used here mainly to illuminate the discussion and results of the following two sections.

Example 1.1 *(Finiteness is not preserved under negation)*

We need a definiens $A(x, G)$ such that the definition $\mathfrak{D}$ given by $Gx =_{df} A(x, G)$ is finite but the definition $\neg\mathfrak{D}$ given by $Gx =_{df} \neg A(x, G)$ is not finite. Take

$$A(x, G) = Gx \wedge \neg\exists y(Gy \wedge F(y, x)).$$

Let $M = \langle D, I\rangle$ be a ground model and let h be a hypothesis in M. Then it is not difficult to verify that the revision sequence induced by h is $\langle h,\ h',\ h' \ldots\rangle$, where $h' = \{d \in h \mid (h \times \{d\}) \cap I(F) = \emptyset\}$. This follows from the fact that for all $d \in h'$, $(h \times \{d\}) \cap I(F) = \emptyset$ by the definition of h', $h' \subseteq h$ and hence $(h' \times \{d\}) \cap I(F) = \emptyset$ for all $d \in h'$. That is to say, $\delta_{\mathfrak{D},M}(h') = h'$. Now we need to find a ground model M and a hypothesis h such that h is not reflexive for

$$Gx = \neg Gx \vee \exists y(Gy \wedge F(y, x))$$

in M. Let M be a ground model where D is the set of integers and F is interpreted as the usual succesor relation. Now, if we take $h = D - \{0\}$, the induced revision sequence that we get is

$$\langle D - \{0\},\ D - \{1\},\ D - \{2\},\ D - \{3\}, \ldots\rangle$$

and so $\neg\mathfrak{D}$ is not a finite definition.

Example 1.2 *(Finiteness is not preserved under conjunction)*

This example shows that if we have finite definitions $\mathfrak{D}$ given by $Gx =_{df} A(x, G)$ and $\mathfrak{D}$' given by $Gx =_{df} B(x, G)$, the definition $\mathfrak{D}\wedge\mathfrak{D}'$ given by $Gx =_{df} A(x, G) \wedge B(x, G)$ does not need to be finite. Suppose that $\mathfrak{L}$ contains a symbol of function f. Take

$$A(x, G) = \exists! y Gy \wedge \forall y(f(y) \neq y) \wedge [\exists y(Gy \wedge x = f(y)) \vee Gx]$$

and

$$B(x, G) = \exists! y Gy \wedge \forall y(f(y) \neq y) \wedge \exists y[Gy \wedge (x = f(y) \vee x = f^2(y)].$$

Now, $\mathfrak{D}$ and $\mathfrak{D}$' are both finite, because given a ground model M and a hypothesis h, we have to cases:

Case 1: $\exists! yGy \wedge \forall y(f(y) \neq y)$ is not true in $M + h$. Then no $x \in D$ satisfies either $A(x, G)$ or $B(x, G)$, and so $\langle h,\ \emptyset,\ \emptyset, \ldots\rangle$ is the revision sequence induced by h (with respect to both $\mathfrak{D}$ and $\mathfrak{D}$') in M.

Case 2: $\exists! yGy \wedge \forall y(f(y) \neq y)$ is true in $M + h$. Then $h = \{d\}$ for some $d \in D$. Moreover, $\delta_{\mathfrak{D},M}(h) = \{d, f(d)\}$ and $\delta_{\mathfrak{D}',M}(h) = \{f(d), f^2(d)\}$. Each one of these sets has exactly two elements, for we have that $\forall y(f(y) \neq y)$ is true in M. But then both $M + \delta_{\mathfrak{D},M}(h)$ and $M + \delta_{\mathfrak{D}',M}(h)$ lie in Case 1, and hence the sequence induced by h in M for $\mathfrak{D}$ is

$$\langle \{d\},\ \{d, f(d)\},\ \emptyset,\ \emptyset, \ldots\rangle$$

and the sequence induced by h in M for $\mathfrak{D}$' is

$$\langle \{d\},\ \{f(d), f^2(d)\},\ \emptyset,\ \emptyset, \ldots\rangle.$$

From Cases 1 and 2 it follows that both $\mathfrak{D}$ and $\mathfrak{D}$' are finite. However, observe that $\mathfrak{D} \wedge \mathfrak{D}'$ is not finite, because we can choose M to be a model such that D is the set of natural numbers and f is the succesor function. If we take $h = \{0\}$, the revision sequence induced by h in M (with respect to $\mathfrak{D} \wedge \mathfrak{D}'$) is

$$\langle \{0\},\ \{1\},\ \{2\},\ \{3\}, \ldots\rangle.$$

Example 1.3 *(Finiteness is not preserved under Modus Ponens)*

Suppose that we have finite definitions $\mathfrak{D} \rightarrow \mathfrak{D}'$ given by $Gx =_{df} A(x, G) \rightarrow B(x, G)$ and $\mathfrak{D}$ given by $Gx =_{df} A(x, G)$. Do we have that the definition $\mathfrak{D}$' given by $Gx =_{df} B(x, G)$ is finite? The answer is that it is not true in general. In particular, if we assume that the ground language $\mathfrak{L}$ has a symbol of binary predicate F, we can take $A(x, G) = \neg\exists yF(y, x)$ and $B(x, G) = \exists y(Gy \wedge F(y, x))$. It is immediate that $\mathfrak{D}$ is finite, because $A(x, G)$ does not really use G, and so it is just an ordinary definition. The definition $\mathfrak{D} \rightarrow \mathfrak{D}'$ is given by

$$Gx =_{df} \neg\exists yF(y, x) \rightarrow \exists y(Gy \wedge F(y, x))$$

and it is finite as well, because by propositional logic it is clear that the definition above is equivalent to

$$Gx =_{df} \exists yF(y, x) \vee \exists y(Gy \wedge F(y, x))$$

which is in turn equivalent (by first order logic) to

$$Gx =_{df} \exists y[F(y, x) \vee (Gy \wedge F(y, x)]$$

and this is just

$$Gx =_{df} \exists y[F(y, x)].$$

In other words, in this case $\mathfrak{D} \rightarrow \mathfrak{D}'$ is equivalent to $\mathfrak{D}$, which is finite. However, it was shown in Example 0.4 that $\mathfrak{D}$' is not a finite definition.

Example 1.4 *(Finiteness is not preserved under disjunction)*

This example is very similar to Example 1.2. Suppose that the ground language $\mathfrak{L}$ contains a symbol $+$ of binary function and a symbol $'$ of unary function, and consider the definition $\mathfrak{D}$ given by

$$Gx =_{df} \exists y \exists z[y \neq z \wedge Gy \wedge Gz \wedge \forall u(Gu \rightarrow (u = y \;\vee\; u = z)) \wedge x = y + z]$$

and the definition $\mathfrak{D}$' given by

$$Gx =_{df} \exists y \exists z[y \neq z \wedge Gy \wedge Gz \wedge \forall u(Gu \rightarrow (u = y \;\vee\; u = z)) \wedge x = (y + z)'].$$

To show that these two definitions are finite, consider an arbitrary ground model M and a hypothesis h. We have again two cases.

Case 1: $\exists y \exists z[y \neq z \wedge Gy \wedge Gz \wedge \forall u(Gu \rightarrow (u = y \;\vee\; u = z))]$ is not true in $M + h$. Then the revision sequence induced by both $\mathfrak{D}$ and $\mathfrak{D}$' is $\langle h,\ \emptyset,\ \emptyset, \ldots\rangle$.

Case 2: $\exists y \exists z[y \neq z \wedge Gy \wedge Gz \wedge \forall u(Gu \rightarrow (u = y \;\vee\; u = z))]$ is true in $M + h$. Then $h = \{d_1, d_2\}$, where $d_1 \neq d_2$. The revision sequence induced by $\mathfrak{D}$ in this case is

$$\langle \{d_1, d_2\},\ \{d_1 + d_2\},\ \emptyset,\ \emptyset, \ldots\rangle$$

and the revision sequence induced by $\mathfrak{D}$' is

$$\langle \{d_1, d_2\},\ \{(d_1 + d_2)'\},\ \emptyset,\ \emptyset, \ldots\rangle.$$

From Cases 1 and 2 it follows that both $\mathfrak{D}$ and $\mathfrak{D}$' are finite. However, $\mathfrak{D} \vee \mathfrak{D}'$ is not finite, because we can take M to be a model such that D is the set of natural numbers and the symbols $+$ and $'$ are interpreted with the usual addition and succesor functions. If we take $h = \{0, 1\}$, the revision sequence induced by h in M (with respect to $\mathfrak{D} \vee \mathfrak{D}'$) is

$$\langle \{0,1\},\ \{1,2\},\ \{3,4\},\ \{7,8\}, \ldots\rangle.$$

Examples 1.1 -1.4 show that the most natural operators on definitions fail to preserve finiteness. There are, however, other operators which do preserve it. Let's close this section by introducing three of them.

Example 1.5 *Finiteness is preserved by self-composition.*

Let $\mathfrak{D}$ be the definition $Gx =_{Df} A(x, G)$. Then we can obtain a new finite definition $\mathfrak{D}^n$ given by

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

where $A^0(x, G) = Gx$ and $A^{n+1}(x, G) = A^n(x, G)[A(t, G)/Gt]$. That is, to obtain A^{n+1} from A^n, we substitute $A(t, G)$ for each occurrence of Gt in $A^n(x, G)$, where t is a term. Let M be a ground model and let δ be the revision rule for $\mathfrak{D}$. We proof by induction on n that the revision rule for $\mathfrak{D}^n$ is $\rho = \delta^n$. This is all we need, because in that case if the set $\{\delta^k(h) \mid k \leq 0\}$ is finite, then the set

$\{\rho^k(h) \mid k \leq 0\}$ is bound to be finite as well, being a subset of the former set. So let us prove the induction.

Case $n = 0$: The definition $\mathfrak{D}^0$ is given by $Gx =_{df} Gx$, with revision rule $\rho(h) = h$. That is, the revision rule is the identity function, which is exactly δ^0.

Case $n = m+1$: Assume that we have that the revision rule for the definition $Gx = A^m(x, G)$ is δ^m. We want to prove that the revision rule for the definition $\mathfrak{D}^{m+1}$ given by $Gx =_{df} A^{m+1}(x, G)$ is δ^{m+1}. Let ρ be the revision rule for $\mathfrak{D}^{m+1}$. Then for every hypothesis h,

$$\begin{array}{lcl} d \in \rho(h) & \Longleftrightarrow & d \text{ satisfies } A^{m+1}(x,G) \text{ in } M+h \\ & \Longleftrightarrow & d \text{ satisfies } A^m(x,G)[A(t,G)/Gt] \text{ in } M+h \\ & \Longleftrightarrow & d \text{ satisfies } A^m(x,G) \text{ in } M + \{d \mid d \text{ satisfies } A(x,G) \text{ in } M+h\} \\ & \Longleftrightarrow & d \text{ satisfies } A^m(x,G) \text{ in } M + \delta(h) \\ & \Longleftrightarrow & d \in \delta^m(\delta(h)) \\ & \Longleftrightarrow & d \in \delta^{m+1}(h). \end{array}$$

Example 1.6 *Finiteness is preserved by classifying ground models.*

A formula ϕ in the ground language $\mathfrak{L}$ can be seen as a type that allows us to classify ground models into those that model ϕ and those that do not. This idea can be used to combine finite definitions in a way that their revision rules do not interfere with each other. Suppose that we have finite definitions

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

and

$$Hx =_{Df} B(x, H),$$

and suppose that ϕ is a sentence in the language of first order logic (no occurrences of G or F in ϕ). Then

$$Fx =_{Df} [\phi \wedge A(x, F)] \vee [\neg\phi \wedge B(x, F)]$$

is a finite definition as well. This is easy to see, because given a ground model M, exactly one of ϕ, $\neg\phi$ is true in M. If $M \models \phi$, then the revision rule for F in M is the revision rule for G. Otherwise, the revision rule for H is used. In any case, the the revision process is finite.

There are at least two questions that we can make about the examples in this section. First, notice that all of the examples of definitions which are not finite involve predicate symbols of arity greater than one[7]. What happens if we

---

[7] ...or function symbols of arity grater than one. But a n-ary function can always be modelled by a $n+1$-ary predicate and the addition of a sentence that states that the predicate is a function. Because of this and to simplify the theoretical reslts and statements, we assume that our languages include only predicate symbols.

restrict ourselves to definitions based on unary predicates only? Section 3 shows that those definitions are bound to be finite. An immediate observation is that the class of (finite) definitions based on unary predicate symbols is closed under the operator introduced in Example 1.5. In contrast, the clase is not closed under the operator introduced in Example 1.6, since one can use any sentence ϕ in the ground language in order to classify models, and that sentence could have predicates of several variables.

The second family of question, addressed in the following section, concerns the use of boolean operators. Why is finiteness not preserved in general under those operators? Is there a way to define a different kind of "conjunction" (or "disjunction") of finite definitions which does preserve finiteness? Section 2 explains a way in which it can be done when our ground language is the language of arithmetic.

2 Operators on finite systems of definitions

Consider two finite definitions $\mathfrak{D}$ and $\mathfrak{D}$' given by $G_x =_{df} A(x,G)$ and $G_x =_{df} B(x,G)$ and with revision rules δ and δ'. Then the revision sequences associated with them (and induced by h) are

$$\mathcal{S} = \langle h, \delta(h),\ \delta^2(h),\ \delta^3(h), \ldots\rangle \text{ and } \mathcal{S}' = \langle h, \delta'(h),\ \delta'^2(h),\ \delta'^3(h), \ldots\rangle.$$

Examples 1.2-1.4 tell us that if $\odot$ is one of the standard boolean operators, then the revision sequence (induced by h) of the definition $\mathfrak{D} \odot \mathfrak{D}'$ given by $Gx =_{df} A(x,G) \odot B(x,G)$ is

$$Dot(\mathcal{S}, \mathcal{S}') = \langle h, \delta(h) \odot \delta'(h),\ \delta(\delta(h) \odot \delta'(h)) \odot \delta'(\delta(h) \odot \delta'(h)),\ \ldots\rangle$$

and that it may include infinitely many different hypothesis. We can imagine then a binary operator Dot on sequences, which takes two sequences $\mathcal{S}$ and $\mathcal{S}$' and returns the sequence $Dot(\mathcal{S}, \mathcal{S}')$ as above. The lesson from Examples 1-2 - 1.4 is that the class of sequences including only finitely many hypothesis is not closed under these "Dot" operators.

However, we would still like to discover some notion of "conjunction", "disjunction", etc. of finite definitions that is guaranteed to preserve finiteness. We explore one approach in this direction here. The idea is the following. Instead of having an operator Dot on sequences which is induced by the plain application of $\odot$ to a pair of finite definitions, we go the other way around. That is, in the concrete case of disjunction, we begin with a simple notion of disjunction on sequences and from this notion we define a "disjunction" operator on finite definitions. Now, the natural counterpart of disjunction in terms of sets is the union operator, and the straightforward way to define an binary operation of sequences based on the union operator is to apply the operator by components. So it is natural to define a disjunction operator on sequences such that if $\mathcal{S}$ and $\mathcal{S}$' are revision sequences as above, then

$$\mathcal{S} \vee \mathcal{S}' = \langle h \cup h, \delta(h) \cup \delta'(h), \delta^2(h) \cup \delta'\ ^2(h), \delta^3(h) \cup \delta'\ ^3(h), \ldots\rangle.$$

It is quite clear that if both sets $\{\delta^i(h) \mid i \leq 0\}$ and $\{\delta'^{\,i}(h) \mid i \leq 0\}$ are finite, then so is $\{\delta^i(h) \cup \delta'^{\,i}(h) \mid i \leq 0\}$. The question is: is there a method to find a finite definition with revision sequence $\mathcal{S} \vee \mathcal{S}'$ out of the two original definitions $\mathfrak{D}$ and $\mathfrak{D}$'? If we can, then we have found the notion of "conjunction" of finite definitions that we were looking for, because by construction this operator would preserve finiteness.

We will use not only single definitions as we have done so far, but also finite systems of definitions. Although the discussion here is centered in the disjunction operator, it should be noticed that everything applies without change to any other boolean operator.

Definition 2.1 *Let $\mathfrak{L}$ be a ground language and let $\{G_1, \ldots, G_n\}$ be a set of fresh predicate symbols. A finite system of definitions $\mathfrak{D}$ is a set*

$$\left\{ \begin{array}{c} G_1 x =_{df} A_1(x, G_1, \ldots, G_n) \\ \vdots \\ G_n x =_{df} A_n(x, G_1, \ldots, G_n) \end{array} \right\}$$

where $A_i(x, G_1, \ldots, G_n)$ is a formula in $\mathfrak{L}^+$, for $1 \leq i \leq n$.

The hypotheses of a ground model $M = \langle D, I \rangle$ are now tuples $h = \langle h_1, \ldots, h_n \rangle$, where $h_i \subseteq D$ $(1 \leq i \leq n)$ and the model $M + h = \langle D, I^+ \rangle$ is a model where I^+ extends I in such a way that $I(G_i) = h_i$. Also, Definitions 0.1 and 0.2 are naturally generalized in order to deal with systems of definitions.

Consider now the problem with disjunction. Single definitions can be seen as systems of definitions with only one element. Suppose that we have systems of definitions

$$\mathfrak{D} = \{Gx =_{df} A(x, G)\} \text{ and } \mathfrak{D}' = \{Gx =_{df} B(x, G)\}$$

Then we can define a new system

$$(*) \qquad \mathfrak{D} \vee \mathfrak{D}' = \left\{ \begin{array}{c} Gx =_{df} A(x, H_1) \vee B(x, H_2) \\ H_1 x =_{df} A(x, H_1) \\ H_2 x =_{df} B(x, H_2) \end{array} \right\}$$

and we have that given a ground model M and an arbitrary hypothesis h, the revision sequence induced by $h = \langle g, h_1, h_2 \rangle$ is the following (each triple in the sequence corresponds to a column in the table):

G	h	$\delta(h_1) \cup \rho(h_2)$	$\delta^2(h_1) \cup \rho^2(h_2)$	$\delta^3(h_1) \cup \rho^3(h_2) \ldots$
H_1	h_1	$\delta(h_1)$	$\delta^2(h_1)$	$\delta^3(h_1) \ldots$
H_2	h_2	$\rho(h_2)$	$\rho^2(h_2)$	$\rho^3(h_2) \ldots$

where

$\delta(h) = \{d \mid d$ satisfies $A(x, H_1)$ in $M\}$ and $\rho(h) = \{d \mid d$ satisfies $B(x, H_2)$ in $M\}$.

Observe that we have come up with a definition such that the sequence of first coordinates in the revision sequence (first row in the table) looks exactly as we wanted (that is like $\mathcal{S} \vee \mathcal{S}'$ in our previous discussion). Two more observations about the definition:

1. The first component g of an arbitrary hypothesis $h = \langle g, h_1, h_2 \rangle$ is not really used during the revision process. However, it seems that when we define $\mathfrak{D} \vee \mathfrak{D}'$ as above, where the intention is to be as close as possible to the problematic $Gx =_{df} A(x, G) \vee B(x, G)$, we only want to revise "homogeneous" initial hypotheses of the form $h = \langle g, g, g \rangle$.
2. It is immediate that if both $\mathfrak{D}$ and $\mathfrak{D}$' are finite, then $\mathfrak{D} \vee \mathfrak{D}'$ is bound to be finite as well.

Unfortunately, what we have done so far is not enough. The problem is that we are not getting a new definition in the original language $\mathfrak{L} \cup \{G\}$, since we have added the two new symbols H_1 and H_2. So we don't have really a binary operator on the set of finite definitions in the language $\mathfrak{L} \cup \{G\}$. We could solve this problem if we were able to "reduce" the system of definitions $\mathfrak{D} \vee \mathfrak{D}'$ (which are given in the language $\mathfrak{L} \cup \{G, H_1, H_2\}$) to an "equivalent" single definition in $\mathfrak{L} \cup \{G\}$. Of course, it is not clear at all how to do that in an arbitrary language $\mathfrak{L}$, as it is not clear what we mean by "equivalent". However, we will see that, when $\mathfrak{L}$ is the language of arithmetic, we can reduce certain systems of definitions to single definitions which, in a precise way, are "equivalent" to the original systems.

First, let us make it clear what we will mean by a single definition being equivalent to a finite system, and then discuss how the reduction from a system to such a definition can be performed. Let $\mathfrak{D}$ be a (single) definition in the language $\mathfrak{L} \cup \{G\}$ and consider the system $\mathfrak{D}$'

$$\left\{ \begin{array}{c} G_1 x =_{df} A_1(x, G_1, \ldots, G_n) \\ \vdots \\ G_n x =_{df} A_n(x, G_1, \ldots, G_n) \end{array} \right\}$$

in the language $\mathfrak{L} \cup \{G_1 \ldots G_n\}$. We will consider $\mathfrak{D}$ as equivalent to $\mathfrak{D}$' if the following holds: if the revision sequences induced by a hypothesis h for $\mathfrak{D}$' are

G_1	h	h_1^1	h_2^1	$h_3^1 \ldots$
G_2	h	h_1^2	h_2^2	$h_3^2 \ldots$
$\vdots$	$\vdots$			
G_n	h	h_1^n	h_2^n	$h_3^n \ldots$

and the revision sequence induced by h for $\mathfrak{D}$ is $\langle h, h_1, h_2, h_3, \ldots \rangle$, then for every $j > 0$ and $0 \leq i < n$,

$$h_j \cap Mod_i = \{nk + i \mid k \in h_j^i\},$$

where $Mod_i = \{d \in D \mid d \text{ satisfies } (x \equiv i \ mod \ n) \text{ in } M\}$ for $0 \leq i \leq n$. In plain words, h_j encodes the j-the row in the table above.

Let us look now at the process of reduction. It is inspired on [K 93], where Phillip Kremer works out a particular example (towards a different objective than ours here) in which $\mathfrak{L}$ is the language of arithmetic and $\mathfrak{D}$ is the system consisting of the two following definitions

$$Gx =_{df} [\exists z \forall y (y < z \leftrightarrow Gy) \wedge \forall y (y < x \rightarrow Gy)]$$

and

$$\begin{array}{ll} Hx =_{df} & [\forall y (Gy \rightarrow Gy') \wedge \forall y \forall z ((Gz \wedge y < z) \rightarrow Gy) \wedge \exists y Gy \wedge Gx] \vee \\ & [[\exists y (Gy \wedge \neg Gy') \vee \exists y \exists z (Gz \wedge y < z \wedge \neg Gy) \vee \neg \exists y Gy] \wedge Hx] \end{array}$$

Again, the motivation for Kremer's definition does not have to do with what our question here [8], but the reduction he makes is interesting in our context. In his paper, Kremer "collapses" the two definitions above into one single definition $\mathfrak{D}$' given by

$$\begin{array}{ll} Fx =_{df} & [(x \text{ is odd }) \wedge \exists z \forall y (y < z \wedge (y \text{ is odd }) \leftrightarrow Fy) \\ & \quad \wedge \forall y (y < x \wedge (y \text{ is odd }) \rightarrow Fy)] \\ & \vee \ [(x \text{ is even }) \wedge \forall y (Fy \wedge (y \text{ is odd }) \rightarrow Fy'') \\ & \quad \wedge \ \forall y \forall z ((Fz \wedge y < z \wedge (z \text{ is odd }) \wedge (y \text{ is odd })) \rightarrow Fy') \\ & \quad \wedge \exists y ((y \text{ is odd }) \wedge Gy) \wedge Gx] \\ & \vee \ [(x \text{ is even }) \wedge [\exists y ((y \text{ is odd }) \wedge Fy \wedge \neg Fy'') \\ & \qquad \qquad \vee \exists y \exists z ((y \text{ is odd }) \wedge (z \text{ is odd }) \wedge Fz \wedge y < z \wedge \neg Fy) \\ & \qquad \qquad \vee \neg \exists y ((y \text{ is odd }) \wedge Fy)] \wedge Fx]. \end{array}$$

An additional necessary (and strong) assumption is that we are looking only at models of $PA^- \wedge AX$ (models of Peano arithmetic minus the inductive scheme, which are strict linear orders with 0 as minimum and where $'$ stands for the immediate succesor operation). The idea behind the definition of F is that the behaviour of F in the odd numbers mimics the behaviour of G and the behaviour of F in the even numbers mimics the behaviour of H. In other words, if the revision sequence induced by a hypothesis $h = \langle f_0, f_0 \rangle$ for $\mathfrak{D}$ is

$$\begin{array}{c|lllll} G & f_0 & g_1 & g_2 & g_3 & \ldots \\ \hline \\ H & f_0 & h_1 & h_2 & h_3 & \ldots \end{array}$$

and the revision sequence induced by $f_0 \subseteq D$ for $\mathfrak{D}$' is $\langle f_0, f_1, f_2, \ldots \rangle$, then if $Even, Odd$ are the sets of even and odd numbers in D, we have that for all $i > 0$,

$$f_i \cap Odd = \{2k + 1 \mid k \in g_i\}$$

---

[8] Kremer uses the system of definitions $\mathfrak{D}$ to show that the Gupta-Belnap systems $\mathbf{S}^{\#}$ and $\mathbf{S}^{*}$ are not axiomatisable, and that the complexity of each one of them is at least Δ^1_1. The reduction of $\mathfrak{D}$ to the single definition $\mathfrak{D}$' allows him to show a stronger complexity result. Namely, the complexity of the $\mathbf{S}^{\#}$ and $\mathbf{S}^{*}$ theories of a single definition is also at least Δ^1_1. The systems $\mathbf{S}^{\#}$ and $\mathbf{S}^{*}$ belong to the General Theory of Circular Definitions.

and similarly

$$f_i \cap Even = \{2k \mid k \in h_i\}.$$

This is just to say that the new definition is equivalent to the original system, according to our definition of "being equivalent".

It is clear that the very same idea can be used without major difficulties to encode our binary operators, given that we keep the assumptions: that our ground language is the language of arithmetic, that we look only at models of $PA^- \wedge AX$ and that the only predicate symbols from the ground language $\mathfrak{L}$ that appear in our definitions are $<$ for the order and $'$ for the succesor. So for our example with disjunction, we would start with two finite single definitions $\mathfrak{D}$ and $\mathfrak{D}$', then we would build a finite system $\mathfrak{D} \vee \mathfrak{D}'$ as in the beginning of this section, and then we would reduce the system to an equivalent single definition in the original extended language. This will always give us a new finite definition that can be thought of as the "disjunction" of $\mathfrak{D}$ and $\mathfrak{D}$', in a certain sense[9]. The basic difference with Kremer's reduction, in the case of our binary operators, is that we need to use the formulas $(x \equiv 0 \ mod\ 3)$, $(x \equiv 1 \ mod\ 3)$ and $(x \equiv 2 \ mod\ 3)$ instead of $(x\ is\ odd)$ and $(x\ is\ even)$, because we want to encode now a table with three rows (the system $\mathfrak{D} \vee \mathfrak{D}'$ has three definitions). Evidently, there is nothing especial about disjunction here, and this procedure applies to any other boolean operator on finite definitions.

Final remark. The procedure described in this section can also be applied to n-ary boolean operators. When doing so, we end up coding tables with $n+1$ rows.

3 Languages and operators

In this section we forget about the language of arithmetic and come back to the discussion on arbitrary languages. The main result is the following.

Theorem 3.1 *Let $\mathfrak{D}$ be the definition*

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

Assume that every predicate symbol that appears in $\mathfrak{D}$ is either unary or the equality predicate. Then $\mathfrak{D}$ is finite.

The proof of Theorem 3.1. is involved when writing it in detail, but the main ideas can be explained briefly before giving the formal proof. Let $\mathfrak{D}$ be a definition as in the statement of Theorem 3.1. Say that the unary predicate symbols that appear in $\mathfrak{D}$ are among $G, F_0, \ldots F_n$. We know that if any other predicate symbol appears in $\mathfrak{D}$, it has to be the equality symbol. Now, fix an extended model $M + h = \langle D, I^+\rangle$ and consider the family

$$\mathfrak{B}(h) = \{\mathfrak{B}^h_n \mid n > 0\},$$

---

[9] The new operator on finite definitions does not behave exactly as the usual disjunction operators. For example, it is not associative.

where $\mathfrak{B}_n^h$ is the collection of subsets of D^n which are definable by using only $G, F_1, \ldots, F_n$, equality and the boolean connectives. The following facts are true:

1. Each $\mathfrak{B}_n^h$ is a finite algebra, that is, each $\mathfrak{B}_n^h$ is closed under boolean operations.
2. The family is closed under projections and cartesian products. For example, if $A \in \mathfrak{B}_{n+1}^h$, then the projection of A on the first $n+1$ coordinates belongs to $\mathfrak{B}_n^h$.
3. If $g \subseteq \mathfrak{B}_1^h$, then $\mathfrak{B}^g$ is a subfamily of $\mathfrak{B}^h$, in the sense that for all n, $\mathfrak{B}_n^g \subseteq \mathfrak{B}_n^h$.

The details of the proofs of facts 2-3 take most part of this section. Basically, facts (1) and (2) together tell us that any set which is definable by using only the predicates $G, F_1, \ldots F_n$, equality and the logical constants belongs to $\mathfrak{B}_n^h$ for some n. In particular, since the set

$$\delta(h) = \{d \mid d \text{ satisfies } A(x, G) \text{ in } M + h\}$$

is definable, we have that $\delta(h) \in \mathfrak{B}_1^h$. Also, in virtue of fact 3, $\mathfrak{B}^{\delta(h)}$ is a subfamily of $\mathfrak{B}^h$. If we continue the revision process, we obtain a sequence of hypothesis $\langle h, \delta(h), \delta^2(h), \ldots\rangle$ such that $\delta^k(h) \in \mathfrak{B}_1^h$ for all k. Since $\mathfrak{B}_1^h$ is finite, and h was chosen arbitrarily, we conclude that $\mathfrak{D}$ is a finite definition.

Now we can proceed to formalize the proof.

Definition 3.1 *1. Let D be a set and $\mathcal{C}$ a collection of subsets of D. The algebra generated by $\mathcal{C}$ on D is the smallest collection of subsets of D that contains $\mathcal{C} \cup \{\emptyset\}$ and is closed under finite unions and complementation (with respect to D).*

2. Let $A(x, F_0)$ be a formula in $\mathfrak{L}^+ = \mathfrak{L} \cup \{F_0\}$ such that every predicate symbol that appear in A is either the equality or one of $F_0, \ldots, F_N$, where F_i unary for all i. Fix a ground model M and let $M + h = \langle D, I^+\rangle$ be an extended model. For each $n > 0$, we define $\mathfrak{B}_n(A, M, h)$ as the algebra on D^n generated by the sets

$$\begin{aligned} B_{n,i,j}(A, M, h) &= \{\langle d_1, \ldots, d_n\rangle \mid d_i \in f_j\} \\ &= \{\langle d_1, \ldots, d_n\rangle \mid d_i \textit{ satisfies } F_j(x) \textit{ in } M\} \end{aligned}$$

and

$$B_{n,i,k,=}(A, M, h) = \{\langle d_1, \ldots, d_n\rangle \mid d_i = d_k\}$$

where $1 \leq i, k \leq n$ and $0 \leq j \leq N$.

3. With A and M as above, the collection $\mathfrak{B}(A, M, h)$ of algebras generated by A is

$$\mathfrak{B}(A, M, h) = \{\mathfrak{B}_n \mid n > 0\}.$$

We will write $\mathfrak{B}(h)$ instead of $\mathfrak{B}(A, M, h)$ when A and M are known by context. The same consideration applies to the algebras $B_{n,i,j}(A, M, h)$ and $B_{n,i,k,=}(A, M, h)$.

We need two lemmas before attempting to prove the main result. The first one introduces some properties of the collection of algebras $\mathfrak{B}(A, h)$ of a given ground model M. Just a sketch of the proof is included here, since a very detailed proof would involve too much decoration and indexes on names, making it difficult to grasp the main ideas, which are not complicated.

Lemma 3.1 [10] *Let $A(x, F_0)$ be a formula in $\mathfrak{L}^+ = \mathfrak{L} \cup \{F_0\}$ and let $M + h = \langle D, I^+ \rangle$ be an extended model. Suppose that the unique non-unary predicate that appears in A (if it appears at all) is "$=$". Let $F_0, F_1, \ldots, F_N$ be the unary predicates that appear in A and let $f_0, f_1, \ldots, f_N$ be the interpretations corresponding to them (so $f_0 = h$). Then*

1. *Let $n > 0$. Each set in $\mathfrak{B}_n$ is the union of finitely many sets of the form $h_1 \cap h_2 \cap \ldots \cap h_{k_n}$ where for each $1 \leq i, k \leq n$ and $0 \leq j \leq N$, exactly one of $B_{n,i,j}$ or $\overline{B_{n,i,j}}$ and exactly one of $B_{n,i,k,=}$, $\overline{B_{n,i,k,=}}$ appears as a term in the intersection. [Remark on notation: if $B \subseteq M$, then $\overline{B}$ stands for the set $M \setminus B$.]*

2. *The collection $\mathfrak{B} = \{\mathfrak{B}_n \mid n > 0\}$ is closed under projections. That is, for all $n > 0$, if $C \in \mathfrak{B}_{n+1}$, then the set*

   $$\pi_k(C) = \{\langle d_1, \ldots, d_{k-1}, d_{k+1}, d_{n+1} \rangle \mid \exists d(\langle d_1, \ldots, d_{k-1}, d, d_{k+1}, d_{n+1} \rangle \in C)\}$$

   belongs to $\mathfrak{B}_n$.

3. *The collection $\{\mathfrak{B}_n \mid n > 0\}$ is closed under cartesian products. That is, if $C_1 \in \mathfrak{B}_n$ and $C_2 \in \mathfrak{B}_m$, then the set $C_1 \times C_2$ belongs to $\mathfrak{B}_{n+m}$.*

Proof of lemma (sketch). The proofs of parts (1) and (3) are not difficult, although they can be messy. The key part is (2). Let's show a toy case with $N = 1$. Because of (1) and the fact that the projection of a union is the union of the projections, it suffices to show that $\pi_k(C)$ belongs to $\mathfrak{B}_n$ whenever $C \in \mathfrak{B}_{n+1}$ and $C = h_1 \cap h_2 \cap \ldots \cap h_k$ where for each $0 \leq i, k \leq n$ and $0 \leq j \leq N$, exactly one of $B_{n,i,j}$ or $\overline{B_{n,i,j}}$ and exactly one of $B_{n,i,k,=}$, $\overline{B_{n,i,k,=}}$ appears as a term in the intersection. Again to simplify the illustration of the argument, let us assume that $n = 2$. Then there is a formula

$$\psi(x, y) = \phi_{0,x}(x) \wedge \phi_{1,x}(x) \wedge \phi_{0,y}(y) \wedge \phi_{1,y}(y) \wedge \phi^{=}$$

in the language $\mathfrak{L}^+$ such that

$$\begin{array}{ll} \phi_{0,x} \in \{F_0, \neg F_0\}, & \phi_{1,x} \in \{F_1, \neg F_1\}, \\ \phi_{0,y} \in \{F_0, \neg F_0\}, & \phi_{1,y} \in \{F_1, \neg F_1\}, \\ \phi^{=} \in \{x = y, x \neq y\}. & \end{array}$$

---

[10] This lemma says that our family of algebras is a *structure*, in the terminology of Lou van den Dries [VdD 98]

and

$$C = \{\langle d_1, d_2\rangle \mid \langle d_1, d_2\rangle \text{ satisfies } \psi(x,y) \text{ in } M + f_0\ \}.$$

Now, remember that $f_0 = h$. It follows from the semantics of the existential quantifier that

$$\begin{aligned} \pi_1(C) &= \{d_2 \mid \langle d_1, d_2\rangle \text{ satisfies } \psi(x,y) \text{ in } M + f_0 \text{ for some } d_1\} \\ &= \{d_2 \mid d_2 \text{ satisfies } \exists x \psi(x,y) \text{ in } M + f_0\ \}. \end{aligned}$$

But it is a theorem of classical logic that

$$\exists x(\theta_1 \wedge \theta_2) \leftrightarrow \theta_1 \wedge \exists x \theta_2 \text{ when } x \text{ does not occur free in } \theta_1,$$

and therefore

$$\pi_1(C) = \{d_2 \mid d_2 \text{ satisfies } \phi_{0,y}(y) \wedge \phi_{1,y}(y) \wedge \exists x(\phi_{0,x}(x) \wedge \phi_{1,x}(x) \wedge \phi^{=}) \text{ in } M + f_0\ \}.$$

Now we can consider several cases, and in each of them we'll find that either $\pi_1(C) = \emptyset$ or

$$\pi_1(C) = \{d_2 \mid d_2 \text{ satisfies } \phi_{0,y}(y) \wedge \phi_{1,y}(y) \text{ in } M + f_0\ \} = h'_0 \cap h'_1.$$

where h'_0 is either $\{d \mid d \text{ satisfies } F_{0,y}(y) \text{ in } M + f_0\ \}$ or its complement and h'_1 is either $\{d \mid d \text{ satisfies } F_{1,y}(y) \text{ in } M + f_0\ \}$ or its complement, so that $\pi_2(C)$ belongs to $\mathfrak{B}_1$. As we said, we need to check several cases.

CASE 1: Suppose that $\psi^{=}$ is $x = y$. If $\pi_1(C) = \emptyset$, we have nothing to prove. Suppose $\pi_1(C) \neq \emptyset$. Then we must have $\phi_{0,y} = \phi_{0,x}$ and $\phi_{1,y} = \phi_{1,x}$, so

$$\begin{aligned} \pi_1(C) &= \{d \mid d \text{ satisfies } \phi_{0,y}(y) \wedge \phi_{1,y}(y) \wedge \exists y(\phi_{0,y}(y) \wedge \phi_{1,y}(y)) \text{ in } M + f_0\ \} \\ &= \{d \mid d \text{ satisfies } \phi_{0,y}(y) \wedge \phi_{1,y}(y) \text{ in } M + f_0\ \}. \end{aligned}$$

CASE 2: Suppose that $\psi_{=}$ is $x \neq y$, that $\psi_{0,y}$ and $\psi_{0,x}$ are both atomic or both negated atomic formulas and that $\psi_{1,y}$ and $\psi_{1,x}$ are both atomic or both negated atomic formulas. Then

$$\pi_1(C) = \{d \mid d \text{ satisfies } \psi_{0,y}(y) \wedge \psi_{1,y}(y) \wedge \exists x(\psi_{0,x}(x) \wedge \psi_{1,x}(x) \wedge x \neq y) \text{ in } M + f_0\ \}$$

and therefore $\pi_1(C) = \emptyset$ if the set

$$\{d \mid d \text{ satisfies } \psi_{0,y}(y) \wedge \psi_{1,y}(y) \text{ in } M + f_0\ \}$$

has less than two elements, and otherwise

$$\pi_1(C) = \{d \mid d \text{ satisfies } \psi_{0,y}(y) \wedge \psi_{1,y}(y) \text{ in } M + f_0\ \}.$$

CASE 3: Suppose that $\psi_{=}$ is $x \neq y$, and we are not in CASE 2. Then without loss of generality, we can assume that $\psi_{0,x} = F_0$ while $\psi_{0,y} = \neg F_0$. Then

$$\pi_1(C) = \{d \mid d \text{ satisfies } \neg F_0(y) \wedge \psi_{1,y}(y) \wedge \exists x(F_0(x) \wedge \psi_{1,x}(x)) \text{ in } M + f_0\ \}$$

(since $x \neq y$ is redundant in this case). Now, if the set

$$S = \{d \mid d \text{ satisfies } \exists x(F_0(x) \wedge \psi_{1,x}(x)) \text{ in } M + f_0\ \}$$

is empty, then $\pi_1(C) = \emptyset$. If S is not empty, then it should be clear that

$$\pi_1(C) = \{d \mid d \text{ satisfies } \psi_{0,y}(y) \wedge \psi_{1,y}(y) \text{ in } M + f_0\ \}.$$

And this finishes the proof of the lemma.

Lemma 3.2 *Let $A(x, F_0)$ be a formula in $\mathfrak{L}^+ = \mathfrak{L} \cup \{F_0\}$ with only one free variable and such that all the predicate symbols that appear in A are unary or equality. Let $M + f_0$ be an extended model and let $\{\mathfrak{B}_k \mid k > 0\}$ be the collection of algebras induced by $A(x, F_0)$. Then*

$$\{d \mid d \textit{ satisfies } A(x, G) \textit{ in } M + f_0\}$$

belongs to $\mathfrak{B}_1$.

Proof. Without loss of generality, we can assume that

$$A(x, G) = Q_1 x_1 \ldots Q_{n-1} x_{n-1} \ldots Q_n x_n (\psi(x, x_1, \ldots, x_n))$$

where $Q_j \in \{\forall, \exists\}$ for all $1 \leq j \leq n$ and where no quantifiers appear in $\psi(x, x_1, \ldots, x_n)$. We prove by induction on i that the set

$$\begin{aligned} h_i = \{\langle d, d_1, \ldots d_{n-i}\rangle \mid\ & \langle d, d_1, \ldots d_{n-i}\rangle \text{ satisfies} \\ & Q_{n-i+1} \ldots Q_n x_n (\psi(x, x_1, \ldots, x_n)) \text{ in } M + f_0\} \end{aligned}$$

belongs to $\mathfrak{B}_{n-i+1}$ for all $0 \leq i \leq n$. The case $i = 0$ says that

$$h_i = \{\langle d, d_1, \ldots d_n\rangle \mid \langle d, d_1, \ldots d_n\rangle \text{ satisfies } \psi(x, x_1, \ldots, x_n) \text{ in } M + f_0\}$$

belongs to $\mathfrak{B}_{n+1}$. But this is true, because $\psi(x, x_1, \ldots, x_n)$ is a boolean combination of atomic formulas in the $n+1$ variables $x, x_1, \ldots, x_n$, since no quantifiers appear in $\psi(x, x_1, \ldots, x_n)$.

Now suppose that $0 < i+1 \leq n$ and suppose that we know that $h_i \in \mathfrak{B}_{n-i+1}$. Then we want to see that $h_{i+1} \in \mathfrak{B}_{n-i}$. If $Q_{n-i} = \exists$, we have that

$$\begin{aligned} h_{i+1} = \{\langle d, d_1, \ldots d_{n-i-1}\rangle \mid\ & \langle d, d_1, \ldots d_{n-i-1}\rangle \text{ satisfies} \\ & \exists x_{n-i} \ldots Q_n x_n (\psi(x, x_1, \ldots, x_n)) \text{ in } M + f_0\} \end{aligned}$$

Thus, $h_{i+1} = \pi_{x_{n-i}}(h_i)$, and so $h_{i+1} \in \mathfrak{B}_{n-i}$ because our collection of algebras is closed under projection. Finally, if $Q_{n-i} = \forall$, then

$$\begin{aligned} h_{i+1} = \{\ & \langle d, d_1, \ldots d_{n-i-1}\rangle \mid \langle d, d_1, \ldots d_{n-i-1}\rangle \text{ satisfies} \\ & \neg\exists x_{n-i} \neg Q_{n-i+1} x_{n-i+1} \ldots Q_n x_n (\psi(x, x_1, \ldots, x_n)) \text{ in } M + f_0\}, \end{aligned}$$

so h_{i+1} is the complement of $\pi_{x_{n-i}}(\bar{h_i})$. Since $h_i \in \mathfrak{B}_{n-i+1}$ and $\mathfrak{B}_{n-i+1}$ is an algebra, we also have $\bar{h_i} \in \mathfrak{B}_{n-i+1}$. Since our collection of algebras is closed under projection, we have $\pi_{x_{n-i}}(\bar{h_i}) \in \mathfrak{B}_{n-i}$, and since $\mathfrak{B}_{n-i}$ is an algebra, we have that $h_{i+1} \in \mathfrak{B}_{n-i}$. This finishes the induction. Hence, for $i = n$, we have

$$\{d \mid d \text{ satisfies } A(x, G) \text{ in } M + f_0\}$$

belongs to $\mathfrak{B}_1$.

Proof of Theorem 3.1. Let $\mathfrak{D}$ be the definition

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

and assume that every predicate symbol that appears in A is either unary or the equality symbol. Let $M = \langle D, I\rangle$ be a ground model. Recall that for each hypothesis $h \subseteq D$ we have defined the collection $\{\mathfrak{B}_j(h) \mid j > 0\}$ of algebras associated with A in the extended model $M + h$. By Lemma 3.2, we have that $\delta_{\mathfrak{D},M}(h) \in \mathfrak{B}_1(h)$. If $G, F_0, \ldots F_N$ are all the predicate symbols that appear in A, then $\mathfrak{B}_1(h)$ is the collection of sets obtained by boolean combinations of sets in $\{h\} \cup \{B_i \mid 1 \leq i \leq N\}$, where

$$B_i = \{d \mid d \text{ satisfies } F_i(x) \text{ in } M + f_i\}$$

for $1 \leq i \leq N$. Similarly, we have that $\mathfrak{B}_1(\delta_{\mathfrak{D},M}(h))$ is the collection of boolean combinations of sets in $\{\delta_{\mathfrak{D},M}(h)\} \cup \{B_i \mid 1 \leq i \leq N\}$. Therefore, since $\delta_{\mathfrak{D},M}(h) \in \mathfrak{B}_1(h)$, we have that $\mathfrak{B}_1(\delta_{\mathfrak{D},M}(h)) \subseteq \mathfrak{B}_1(h)$. We can repeat this argument to prove inductively that for all $k \geq 0$,

$$(*) \qquad \mathfrak{B}_1({\delta_{\mathfrak{D},M}}^{k+1}(h)) \subseteq \mathfrak{B}_1(h).$$

Consequently, ${\delta_{\mathfrak{D},M}}^{k}(h) \in \mathfrak{B}_1(h)$ for all $k \geq 0$. But $\mathfrak{B}_1(h)$ is finite, so there must be indexes k, l with $k \neq l$ and ${\delta_{\mathfrak{D},M}}^{k}(h) = {\delta_{\mathfrak{D},M}}^{l}(h)$. Thus, $\mathfrak{D}$ is finite. This finishes the proof of the theorem.

Corollary 3.1 *If $\mathfrak{L}$ is a ground language all of whose predicate symbols are unary, then every definition $\mathfrak{D}$ given by $Gx =_{df} A(x, G)$ is finite.*

The Examples 1.1-1.3 show that the proof of Theorem 3.1 does not apply to definitions that involve binary predicates, because Lemma 3.1 can be falsified. However, there are at least one generalization of the result, which allows the use of arbitrary predicates, in a restricted way. The idea is to add to the generators of the algebras the projections of the predicates which appear in the definition.

We say that a formula $\phi(x)$ is a *basic projection on* F (where F is a n-ary predicate symbol) if it has the form

$$\exists x_1 \exists x_2 \ldots \exists x_k \exists x_{k+1} \exists x_{k+2} \ldots \exists x_n F(x_1, \ldots, x_n)$$

.

Theorem 3.2 *Let $\mathfrak{D}$ be the definition*

$$Gx_1, \ldots x_n =_{df} A(x_1, \ldots, x_n, G).$$

Assume that $A(x_1, \ldots, x_n, G)$ has the form

$$Q_1 x_1 \ldots Q_{m-1} x_{m-1} \ldots Q_n x_m (\psi(x, x_1, \ldots, x_m))$$

where $Q_j \in \{\forall, \exists\}$ for all $1 \leq j \leq m$ and where $\psi(x, x_1, \ldots, x_n)$ is a boolean combination of basic projections. Then $\mathfrak{D}$ is finite.

Sketch of Proof. Say that $G, F_1, \ldots, F_N$ (and perhaps equality) are all the predicate symbols appearing in $\mathfrak{D}$. Assume that F_i has arity a_i and identify F_0 with G. Repeat the proof of the Lemmas and of Theorem 3.1, but this time we redefine the set $\mathfrak{B}_n(A, M, h)$ as the algebra of sets on D^n generated by [11]:

1. Embeddings of predicates into higher dimensional spaces.

   $B_{n,\overline{s},j}(A, M, h) =$

   $\{\langle d_1, \ldots, d_n \rangle \mid \langle d_{s(1)}, \ldots, d_{s(a_j))} \rangle$ satisfies $F_j(x_1, \ldots, x_{a_j})$ in $M\}$

   where $1 \leq n$, $0 \leq j \leq N$ and $\overline{s}$ varies over the set of increasing functions from $\{1, \ldots a_j\}$ to $\{1, \ldots n\}$.

2. Diagonals:
   $B_{n,i,k,=}(A, M, h) = \{\langle d_1, \ldots, d_n \rangle \mid d_i = d_k\}$,
   where $1 \leq i, k \leq n$ and $0 \leq j \leq N$.

3. Projections:
   $B_{n,\overline{s},\overline{t},j}(A, M, h) =$
   $\{\langle d_1, \ldots, d_n \rangle \mid \begin{array}{l} \langle d_{s(1)}, \ldots, d_{s(a_j - q))} \rangle \text{ satisfies} \\ \exists x_{t(1)}, \exists x_{t(2)}, \ldots \exists x_{t(q)} F_j(x_1, \ldots x_{a_j}) \text{ in } M \end{array} \}$

   where $1 \leq n$, $0 \leq j \leq N$, $q \leq a_j$, $\overline{s}$ varies over the set of increasing functions from $\{1, \ldots, a_j - q\}$ to $\{1, \ldots n\}$, and $\overline{t}$ varies over the set of increasing functions from $\{1, \ldots q\}$ to $\{1, \ldots a_j\}$.

Notice that the requirement about basic projections, in the statement of Theorem 3.2, is necessary. If we drop it, we can find definitions like

$$Rxy =_{df} Rxy \wedge \exists z Rzx$$

which are infinite, even when the ground language is monadic with equality. To see why this definition is infinite, consider a ground model with universe

---

[11] Again, the notation here is heavy, but the idea behind the algebras and the proof has not changed.

the set of natural numbers, and let h be the usual succesor relation, that is, $\langle x, y\rangle \in h \iff y$ is the succesor of x. Then the revision sequence induced by h is

$$\langle h, h - \{\langle 0,1\rangle\}, h - \{\langle 0,1\rangle, \langle 1,2\rangle\}, h - \{\langle 0,1\rangle, \langle 1,2\rangle, \langle 2,3\rangle\}, \ldots\rangle,$$

which of course includes infinitely many different hypotheses.

References

[GB 93] Gupta, A. and N. Belnap, "The Revision Theory of Truth", MIT Press, 1993.

[CG 00] "On Circular Concepts", In: *"Circularity, Definition and Truth"*, Chapuis, A. and Gupta, A (eds.),Indian Council of Philosophical Research, 2000.

[K 93] Kremer, Philip, "The Gupta-Belnap Systems S^* and $S^\#$ are not Axiomatisable". *Notre Dame Journal of Formal Logic* 34 (4), Fall 1993; 583-596.

[VdD 98] van den Dries, L. "Tame Topology and O-minimal Structures". Cambridge University Press, 1998.