

Chapter 8

Some Remarks on the Finite Theory of Revision

Riccardo Bruni

Abstract The *Revision theory of truth* is known, in its *full* (transfinite) form, as one way of dealing with circular concepts (see Gupta and Belnap 1993). The restriction of this approach which is obtained by limiting it to arbitrary, but finite steps of revision is less known, and less studied instead. In this paper we try to assess it, both from the point of view of its motivations, and of those properties which are relevant for establishing a connection with the logical investigation. Finally, we try to see how much of this approach can we make use of in the case of truth.

8.1 Introduction

Alice and Bob are given one euro each. They are told that in case they both decide to invest it on stocks of the Safe Deal Corp., they will be given the euro back. If only Bob decides to make the investment while Alice refuses to do it, the latter will lose the euro she was given, and Bob will receive thrice the investment. If it is Alice who decides to invest and Bob who refuses to do it, the opposite will happen. If they both refuse to make the investment, they will be given twice the initial sum. After pondering over the proposal for a short while, both Alice and Bob (who have been asked to make a decision independently), decide to invest the given euro. Have they made a rational decision?

Let us deal with the situation game-theoretically. The columns of the diagram below correspond to Alice's options, with "I" staying for "Invest" and "N" for "Not invest". Rows correspond to Bob's actions instead, and they are marked by the corresponding lower-case letters. Payoffs take the form of pairs, the first member of which is Alice's payoff, the second is Bob's. Then we have:

R. Bruni
Università degli Studi di Firenze, Firenze, Italy
e-mail: riccardobruni@hotmail.com

	i	n
I	1,1	3,0
N	0,3	2,2

Let us try to figure out how the two characters might have reasoned. Alice assumes that Bob decides to invest. She then notices that by investing she gets the best payoff. If Bob refuses to invest instead, then investing yields again the best outcome for her. So, she decides to invest. Bob reasons in a similar manner and, being the payoffs distributed symmetrically, he is brought to conclude the same.

This reasoning seems unexceptionable¹. Let us make the situation more realistic, and say that Alice has an intuition regarding what is more convenient to do for her. For instance, she could be hesitating toward the idea of entering the stocks market. Hence, the previous reasoning would have the effect of making Alice change her mind (we are here assuming that Alice and Bob trust logic more than their own prejudices). In case she was already oriented toward stock investment from the very beginning, Alice would pursue this option even more firmly afterwards.

We are tempted to extract some kind of a moral from this story. Alice and Bob are dealing with a problem of *circularity*. The situation requires indeed each of them to single out their best option, on the basis of what is best to do for the other. The argument by means of which they do this is based on *making hypothesis* regarding what might be convenient for the other to do, and for the reasoner also. In the one case, the initial hypothesis is confirmed, in the other it is *revised*. Moreover, the story seems to be tied up with the proposed analysis in a very *natural* way.

This is a good basis for introducing, and evaluating the machinery of finite revision for dealing with circular concepts.

8.2 Circular Concepts by Finite Revision

In the fifth chapter of their book (Gupta and Belnap 1993), Gupta and Belnap introduce a cut-down version of the Revision Theory of Truth as an intermediate step toward the full theory as it is best known. As this part of Gupta and Belnap's work, as well as the articles by Gupta (2000) and Chapuis (2000) which are based on it, seem to have attracted little attention so far, our plan is to review the theory of finite revision in some details here, and show how it reconciles with the introductory example.

¹ Notice that it is irrelevant that Alice knows Bob's payoffs and vice versa. This depends upon their distribution in *this* game, and the fact that one and the same option for both players (not investing), is strictly "dominated" by the other alternative, as one would say it in game-theoretic terms.

Let \mathcal{L} be a given first-order predicate formal language. Let \mathcal{L}^+ indicate the extension of it by means of a unary predicate $G(x)$. Assume that $A_G(x, G)$ is a formula of \mathcal{L}^+ which is an accepted *definition* of G . The natural way to interpret this is to say that the intended model \mathbf{M}^+ of \mathcal{L}^+ is an expansion of the intended model \mathbf{M} of \mathcal{L} by the set \mathbf{G} of those elements of the domain of \mathbf{M} which produce valid instances of $A_G(x, G)$. However, since the latter is a formula of the extended language, it features occurrences of G itself and requires that one knows the set \mathbf{G} already, in order to determine which instances of the formula are valid in a classical setting.

The revision theory is conceived in order to overcome such a difficulty. In order to make use of the example, we may assume that \mathcal{L} is a language featuring terms a_1, a_2, b_1, b_2 for *actions* of the players (with the intended meaning that a_1 represents Alice's option of investing, and a_2 of refusing to invest, while b_1 and b_2 are names for Bob's corresponding actions). Furthermore, the alphabet of \mathcal{L} is equipped with a function symbol u which, given one of the two players, say A or B for simplicity, and any given play, that is any pair of actions, represents the utility for the chosen player under the given scenario. So, for instance, by $u_A(a_1b_1)$ it is meant Alice's payoff in the event that both she and Bob decide to invest. Payoffs are then compared by means of an order relation \geq .

This base language is upgraded to \mathcal{L}^+ by means of a new unary predicate $R(x)$ for " x is a rational action". By taking Alice and Bob's reasoning in the example as a paradigm, the two players favour the action which is best in terms of utility. This property is easy to describe by the linguistic means that we have at our disposal. As a matter of fact, let φ_1^A be the formula

$$(R(b_1) \wedge u_A(a_1b_1) \geq u_A(a_2b_1)) \vee (R(b_2) \wedge u_A(a_1b_2) \geq u_A(a_2b_2))$$

This expresses the fact that a_1 is the best possible action for Alice, either if it is rational for Bob to invest, or not.

Similarly, one can easily imagine how to write down a formula φ_2^A which does the same for Alice's action a_2 , and formulas φ_1^B, φ_2^B for Bob's actions.

By logic, one obtains a general definition of rationality as the following combination of these formulas

$$R(x) \Leftrightarrow (x = a_1 \wedge \varphi_1^A) \vee (x = a_2 \wedge \varphi_2^A) \vee (x = b_1 \wedge \varphi_1^B) \vee (x = b_2 \wedge \varphi_2^B)$$

Clearly, this says that x is rational in case it is the most convenient action for one of the two players. Let the righthand side of this defining equivalence be abbreviated by $A_R(x, R)$ in the following.

Let us go back to the general case, and let an interpretation \mathbf{M} of \mathcal{L} be fixed. An *hypothesis* is a subset H of the domain $|\mathbf{M}|$ of \mathbf{M} . Let us assume that the set of terms of \mathcal{L} (hence, of \mathcal{L}^+) contains in addition all of the names for elements of $|\mathbf{M}|$ (\bar{a} being the name of $a \in |\mathbf{M}|$). This is known to be possible without loss of generality.

Let \models indicate the usual, classical validity relation for formulas of \mathcal{L} with respect to the interpretation \mathbf{M} . For every hypothesis H , and A, B formulas of \mathcal{L}^+ , this relation is extended to the expanded language by the following clauses where P

stays for any atomic predicate of \mathcal{L} (hence P is different from R):

$$\begin{aligned}
(\mathbf{M}, H) \models R(\bar{a}) & \quad \text{iff} \quad a \in H \\
(\mathbf{M}, H) \models P(t_1, \dots, t_n) & \quad \text{iff} \quad \mathbf{M} \models P(t_1, \dots, t_n) \\
(\mathbf{M}, H) \models \neg A & \quad \text{iff} \quad (\mathbf{M}, H) \not\models A \\
(\mathbf{M}, H) \models A \wedge B & \quad \text{iff} \quad (\mathbf{M}, H) \models A \text{ and } (\mathbf{M}, H) \models B \\
(\mathbf{M}, H) \models \exists x A & \quad \text{iff} \quad \text{for some } a \in |\mathbf{M}|, (\mathbf{M}, H) \models A[x/\bar{a}]
\end{aligned}$$

Let this relation be defined for the other connectives of \mathcal{L}^+ and for the universal quantifier as usual.

The idea of the revision theory is that hypotheses fix the extension of the circular predicate temporarily. Then, this tentative extension is refined by means of a *revision operator* δ_A (depending on the defining condition $A_G(x, G)$ of the circular predicate G), which associates H to the hypothesis $\delta_A(H)$ according to the condition, for every $a \in |\mathbf{M}|$:

$$a \in \delta_A(H) \Leftrightarrow (\mathbf{M}, H) \models A_G(\bar{a}, G)$$

If one thinks of this condition with respect to the example (i.e., with respect to the formula $A_R(x, R)$ above), it is very easy to see that this is actually what Alice and Bob are using in order to refine their intuitions regarding what is best for them to opt for. For, according to the description we have made, they “calculate” the action yielding the best payoff, under some initial suppositions about what to do. Suppositions correspond to hypotheses. The revision operator translates at the formal level Alice and Bob’s calculation of the action which fulfils the underlying definition of rational choice. The latter, in turn, embodies the idea that it is rational what guarantees the best payoff.

Payoffs distribution determines whether the outcoming hypothesis $\delta_R(H)$ (δ_R being defined on the basis of the definition $A_R(x, R)$ of rational choice), can either revise H , or not. So, for instance, in the proposed example we have that $\delta_R(H) = H$ for $H = \{a_1, b_1\}$, while for every $H' \neq H$ one has $\delta_R(H') \neq H'$ instead².

In general, one can expect that the revision procedure requires further applications of the base machinery in order to yield solutions. Hence, the following iteration of the operator along \mathbb{N} for every hypothesis H is defined:

$$\begin{aligned}
\delta_A^0(H) &= H \\
\delta_A^{n+1}(H) &= \delta_A(\delta_A^n(H))
\end{aligned}$$

Let us call $(\delta_A^n(H) \mid n \in \mathbb{N}, H \text{ hypothesis})$ a *revision evaluation sequence* (RES, henceforth).

² For the sake of this account, we only consider hypotheses which are different from the empty set (i.e., hypotheses that appear in a diagram representing the given situation as a game). It is clear that if one relaxes this condition then for $H = \emptyset$ one has $\delta_R(H) = H$ as well.

The next problem is to set a standard on how to single out “solutions” of any given RES. In the chosen example, that was an easy task due to the fact that investing was best for both players independently of the opponent’s choice. It would be too strict to stick to hypotheses which can replicate this condition in all possible situations. Rather, Gupta and Belnap (but see also Gupta (2000)), have opted for the property of being *reflexive* in order to stress the fact that an hypothesis is “reliable”:

Definition 1 1. An hypothesis H is said to be n -*reflexive* if $\delta_A^n(H) = H$.

2. An hypothesis H is *reflexive*, if it is n -reflexive, for some $n > 0$.

The reader should notice, and keep in mind the restriction by means of which the notion of reflexivity is defined out of the more general notion of n -reflexivity (where $n = 0$ is allowed).

The property of (n -)reflexivity relaxes the feature of the solution $H = \{a_1, b_1\}$ from the chosen example, which is a *fixed point* of the revision operator (i.e., such that $\delta_R(H) = H$, hence 1-reflexive). In addition, a reflexive hypothesis *needs not* to be such that $\delta(H') = H$ for every $H' \neq H^3$.

Informally speaking, the idea is that if an hypothesis recurs in this sense in a RES, then it is deemed to provide reliable information. This latter fact, is explained in terms of a notion of *validity* for formulas of \mathcal{L}^+ , which is defined accordingly. In fact, there are two slightly different versions of such a notion.

The first one comes from Gupta and Belnap’s book (Gupta and Belnap 1993, Def. 5A.2, p. 147), and it makes a direct use of the notion of n -reflexivity:

Definition 2 Let \mathcal{L} be a first-order language, and $\mathcal{L}^+ = \mathcal{L} \cup \{G(\cdot)\}$. Let $A_G(x, G)$ be the defining condition for the predicate G . Then, we say that:

1. a sentence B of \mathcal{L}^+ is m -*valid* in a given interpretation \mathbf{M} of \mathcal{L} ($m \in \mathbb{N}$), if, and only if there exists $k \in \mathbb{N}$, such that, for every m -reflexive hypothesis H , $(\mathbf{M}, \delta_A^k(H)) \models B$;
2. a sentence B of \mathcal{L}^+ is m -*valid* ($m \in \mathbb{N}$) if, and only if B is m -*valid* in \mathbf{M} , for every interpretation \mathbf{M} of the base language.

Let, in the following, $\Vdash_n B$ stay for “ B is n -valid”.

One can try to give an explanation of this definition along the following lines. Reflexivity is, as we said, a feature that makes an hypothesis reliable. Since the revision process itself is trusty, reliability is preserved along any RES which starts from a reflexive hypothesis. A formula is m -valid if it is validated *at one and the same stage* k of every RES starting from an m -reflexive hypothesis H . So, validity here depends upon fulfilling a condition which is made strong by the requirement of the stage k in question being uniform for every m -reflexive hypothesis.

³ This feature of the given game depends upon the payoffs distribution, as we said. Games like the one we have considered in the introduction are indeed *regular*, as Gupta (2000) calls them (as they refer to circular definitions which can be said to be regular in the sense of (Gupta and Belnap 1993, Def. 5A.8, p. 149)). As it is made clear by Gupta himself, regularity plays an important role in the revision-theoretic account of finite games, though it is too much a special feature to become a standard (see also Bruni and Sillari 2011 on this).

It should be noticed that this (maybe) disturbing aspect of the above definition can be eliminated, though only partially. This has been shown, in a remark by Gupta and Belnap (1993, p. 147) themselves. We reproduce it here, along with its proof, for the sake of self-containedness:

Lemma 1 *If $n > 0$, any given formula B of \mathcal{L}^+ is n -valid if, and only if, for every interpretation \mathbf{M} of \mathcal{L} , $(\mathbf{M}, H) \models B$ for every n -reflexive hypothesis H .*

Proof Let \mathcal{L} , $\mathcal{L}^+ = \mathcal{L} \cup \{G(\cdot)\}$, and δ_A be as before. Let \mathbf{M} be an arbitrary, but fixed interpretation of \mathcal{L} .

The direction from right to left of the lemma is trivial. For, if $(\mathbf{M}, H) \models B$ holds for every n -reflexive hypothesis H , then $(\mathbf{M}, \delta_A^0(H)) \models B$ holds too by definition of δ_A^k for every n -reflexive hypothesis H . Hence, B is n -valid in \mathbf{M} , which was so chosen to be generic.

For the other direction, assume that H is n -reflexive with $n > 0$. Then, clearly $\delta_A^k(\delta_A^{(nk)-k}(H)) = H$ for every $k \in \mathbb{N}$ (since $\delta_A^n(\delta_A^m(H)) = \delta_A^{n+m}(H)$ for every $n, m \in \mathbb{N}$, easily follows from the definition of $\delta_A^n(\cdot)$). Moreover, $\delta_A^{(nk)-k}(H)$ is n -reflexive. As a matter of fact, having noticed that $\delta_A^n(\delta_A^m(H)) = \delta_A^{n+m}(H)$, one has also that

$$\delta_A^n(\delta_A^m(H)) = \delta_A^{n+m}(H) = \delta_A^{m+n}(H) = \delta_A^m(\delta_A^n(H))$$

Hence,

$$\delta_A^n(\delta_A^{(nk)-k}(H)) = \delta_A^{(nk)-k}(\delta_A^n(H)) = \delta_A^{(nk)-k}(H)$$

by H being n -reflexive.

Now, assume that $(\mathbf{M}, \delta_A^k(H)) \models B$ holds for some $k \in \mathbb{N}$, and for every n -reflexive hypothesis H (that is, assume that B is n -valid in \mathbf{M}). Then, in particular, $(\mathbf{M}, \delta_A^k(\delta_A^{(nk)-k}(H))) \models B$ for every H which is n -reflexive. Then, $(\mathbf{M}, H) \models B$ holds for every such an H . ■

Having shown this, one can stick to the alternative, simplified definition of n -validity for $n > 0$ which goes as follows:

Definition 3 Let \mathcal{L} be a first-order language, and $\mathcal{L}^+ = \mathcal{L} \cup \{G(\cdot)\}$. Let $A_G(x, G)$ be the defining condition for the predicate G . Then, we say that:

1. a sentence B of \mathcal{L}^+ is m -valid in a given interpretation \mathbf{M} of \mathcal{L} ($m > 0$), if, and only if $(\mathbf{M}, H) \models B$ for every m -reflexive hypothesis H ;
2. a sentence B of \mathcal{L}^+ is m -valid ($m > 0$) if, and only if B is m -valid in \mathbf{M} , for every interpretation \mathbf{M} of the base language.

Why getting rid of the $m = 0$ case? The reason one can think of so far is that the concept of 0-reflexivity is trivial since it applies to every hypothesis by definition of the revision operator.

Now, Gupta (2000, p. 126) introduces another notion of validity, which, as we shall notice, is related to the previous one. It goes as follows:

Definition 4 Let \mathcal{L} be a first-order language, and $\mathcal{L}^+ = \mathcal{L} \cup \{G(\cdot)\}$. Let $A_G(x, G)$ be the defining condition for the predicate G . Then, we say that:

1. a sentence B of \mathcal{L}^+ is *valid in \mathbf{M}* if, and only if $(\mathbf{M}, H) \models B$ for every reflexive hypothesis H ;
2. a sentence B of \mathcal{L}^+ is *valid* if, and only if B is *valid in \mathbf{M}* , for every interpretation \mathbf{M} of the base language.

Let $\models B$ stay in the following for “ B is valid” in this latter sense.

Notice that by referring to hypotheses which are *reflexive*, this definition is meaningful for all hypotheses which are n -reflexive for $n > 0$.

Having shown that the original uniformity requirement can be dispensed with in the case of n -validity, if $n > 0$, there is no simplification in this latter definition of validity in this sense. Actually, one has, that any given formula B of \mathcal{L}^+ is valid if, and only if B is n -valid for every $n > 0$.

The real advantage is that, by sticking to the notion of validity, one has to deal with *just one* semantical notion, rather than with infinitely many validity notions, or, to say it differently, with a notion which is stratified in infinitely many layers.

However, this has a cost. For, in the case of validity one has to make sure that reflexive hypotheses do exist. As a consequence, Gupta (2000, p. 126) notices that the semantics of finite revision as defined in this latter way, is meaningful for circular definitions respecting the following *finiteness requirement* FR:

FR For all models \mathbf{M} of \mathcal{L} , there exists a natural number k such that, for every hypothesis H , $\delta_A^k(H)$ is reflexive.

Notice that FR does not limit itself to state that reflexive hypotheses exist. It says that they do exist in a uniform manner: there exists a $k \in \mathbb{N}$, such that the k -th stage of *any* given RES is reflexive. So, it seems that a uniform condition that we were able to throw out of the door in the case of n -validity, came back in from the window in the case of validity *where it seemed that it was not*⁴. The observation is less harmful than it seems, since the only known application of the semantics of finite revision is the notion of rational choice in finite games of (Gupta 2000), and the circular definition of it that one can give along the previous lines of argument *does* respect FR in this form. However, dealing with truth by finite revision allows to make a (partial) case for the missing $n = 0$ clause. This we shall see in § 8.4.

Of course, this is not to say that the notion of n -validity does not give rise to worries of any sort. Conceptually speaking, for instance, one would like to find grounds supporting the intuition that there is a difference between a formula A which requires the revision process to be applied, say, once in order to get a reason for believing that it is valid, and a formula A' which requires that the whole procedure be carried out 70 times to reach a similar conclusion. At the present stage of development

⁴ As a matter of fact, this impression is illusory: the reason why FR needs to be formulated in this way is stated in § 8.3, and is related to the uniform way in which the original concept of n -validity was defined.

of the theory, and to the best of the author's knowledge, there is nothing that one can make use of in order to say what this difference amounts to in the end.

This suggests, at least, that we should ponder the features of the finite theory of revision with a bit more care.

8.3 Naturality, Complexity and Logicality

As the story of Alice and Bob, and our analysis of it make clear, the treatment of circular concepts by finite revision is very *natural*, one could say first. The revisionary way of identifying solutions for situations involving circular concepts, is tied up with the informal reasoning that “real” people can be thought of pursuing, while making decisions in “real” situations (some caution is required as to how much game theory is regarded as a legitimated way of representing real people dealing with real situations).

In order to stress the point, it can be useful to make a comparison with the full, transfinite revision theory. This latter extension of the finite theory, is obtained by giving a limit clause for the iteration of the revision operator δ . Notoriously, different proposals have been made in this respect (the interested reader is again referred to (Gupta and Belnap 1993)). Here we confine ourselves to the simplest of them, which is due to Hans Herzberger (1982).

Having defined $\delta_A^n(H)$ for every $n \in \mathbb{N}$ and for every hypothesis H , we define $\delta_A^\alpha(H)$ for every ordinal number α and H by:

$$\begin{aligned}\delta_A^0(H) &= H \\ \delta_A^{\alpha+1}(H) &= \delta_A(\delta_A^\alpha(H)) \\ \delta_A^\lambda(H) &= \{x \mid \exists \alpha < \lambda \forall \beta < \lambda (\alpha \leq \beta \rightarrow x \in \delta_A^\beta(H))\}, \lambda \text{ limit}\end{aligned}$$

Clearly, the first two clauses are as in the finite theory, except that they have been extended to ordinals. The idea behind the last clause in this definition should be clear: one retains at limits what behaves stably below the limit ordinal, where “stably” here means that it remains inside the revision sequence from one point onwards.

Now, this clearly represents a legitimate extension of the notion of reliability for an hypothesis, that we were referring to before. As far as the successor stage is concerned, reliability stems from accepting the revision–theoretic machinery. In turn, the intuitiveness of it comes from the analysis of actual forms of reasoning. This connection with reality is lost for the transfinite extension of the theory. This is due to the need of referring to ordinal numbers already. Apart from that, it is a notable fact that no proposal concerning a limit clause received a general consensus. Beside Herzberger's limit rule, we have also Gupta's and Belnap's.

By the way, it is hard to imagine that things could be different. For, any limit clause is motivated by the nature of the process being transfinite, and by the involved stage being a limit ordinal, rather than by the need of adding another operation to the revision process due to the informal reasoning that we want to capture. There seems

to be no real story here providing the required motivation. Clearly, the transfinite iteration seen as a completion of the finite theory, must be done coherently with the “spirit” of revision. As we said, it is easy to acknowledge coherence to the above extension proposal. The remarks we have just made, show that coherence seems not to be enough.

So, the finite theory of revision is natural, and its naturality comes from the fact that it needs not to compromise itself to any limit rule as it is shown by the comparison with the transfinite theory. More than that, the theory is “acceptable” also from the point of view of its complexity⁵. This observation can be made precise in the following manner.

By taking inspiration from the transfinite case, we focus on the standard structure \mathbb{N} of arithmetic. Hypotheses take here the form of subsets of \mathbb{N} . The base language \mathcal{L} is the language \mathcal{L}_{PA} of Peano arithmetic, and $\mathcal{L}^+ = \mathcal{L}_{PA} \cup \{G(\cdot)\}$. Let the circular predicate be defined by any formula $A(x, G)$ of \mathcal{L}^+ , and set, for every $X \subseteq \mathbb{N}$

$$\delta_A(X) = \{n \in \mathbb{N} \mid (\mathbb{N}, X) \models A[x/\bar{n}, G]\}$$

Then, put:

Definition 5 Let, for every $X \subseteq \mathbb{N}$

$$\begin{aligned} ref_A(n, X) &:= \delta^n(X) = X \\ ref_A(X) &:= (\exists n \in \mathbb{N}^+) ref_A(n, X) \end{aligned}$$

(where $\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$).

Then, we say of any $Z \subseteq \mathbb{N}$:

1. For every $n \in \mathbb{N}^+$, Z is n -definable if, and only if

$$x \in Z \Leftrightarrow \forall X [ref_A(n, X) \rightarrow x \in X]$$

2. Z is definable if, and only if

$$x \in Z \Leftrightarrow \forall X [ref_A(X) \rightarrow x \in X]$$

Notice that we have here referred to the simplified definition of n -reflexivity, the one for the case $n > 0$ only, and to the related notion of reflexivity.

The question is what kind of sets (i.e., of what logical complexity) get defined in these senses as $A(x, G)$ is allowed to vary. It is an easy task to verify the following:

Lemma 2 1. Every n -definable set Z is Π_1^1 .
2. Every definable set Z is Π_1^1 .

Proof Fixed an interpretation \mathbf{M} of the given language, one notices that $x \in \delta_A(X)$ has the same logical complexity as the relation \models we have introduced in § 2, for every

⁵ The material below follows a comment and a suggestion in this sense due to Philip Welch.

$X \subseteq \mathbb{N}$ and $x \in \mathbb{N}$. By known arguments the latter can be shown to be of complexity Δ_1^1 as the defining condition $A(x, G)$ is allowed to vary (see, e.g., Takeuti 1987). This can be used to prove by induction that $x \in \delta_A^n(X)$ is of complexity Δ_1^1 as well. It follows that $ref_A(n, X)$ is Σ_1^1 definable. Hence the theorem. ■

Of course, this is only a sketch of a more comprehensive study of set-definability under finite revisions. Since this would be largely useless for the aim of the present paper, we will eventually take the issue up elsewhere. However, this is enough to claim a superiority of the finite theory over the transfinite one in this sense. The available results for the transfinite theory make it legitimate to conclude that the finite theory is more satisfactory than the full one from the viewpoint of these complexity calculations⁶.

However, one may argue that this way of extracting the complexity of the theory can only give partial information. For, in the end we have defined a notion of validity for formulas of a language (two, in fact). So, we have a semantics. Then, should not be better to ask the question: do we have a logical calculus which is adequate to this semantics? This is another way of dealing with complexity in the end (in particular, to deal with the complexity of the involved notion of validity). Also, it is another way of carrying out the comparison with the transfinite theory, which lacks such a connection with an axiom system.

In the case of the finite theory of revision, the answer is positive if one bases oneself on the notion of n -reflexivity (with $n = 0$ included). Recall that a formula B from a language with a circular predicate is said to be n -valid if, given any interpretation \mathbf{M} of the base language, A is made true by any set $\delta^k(H)$ for some $k \in \mathbb{N}$ and for every n -reflexive hypothesis H (Def. 2 from § 2). Then, Gupta and Belnap (1993, Chap. 5) devised a family $(\mathbf{C}_n)_{n \in \mathbb{N}}$ of Fitch-style natural deduction calculi for which they prove the following⁷:

Theorem 1 *Let \mathcal{L} be any first-order predicate language. Let $G(x)$ be a unary predicate, $\mathcal{L}^+ = \mathcal{L} \cup \{G\}$, and $A_G(x, G)$ be a formula of \mathcal{L}^+ which is a defining condition for G . Then for every $n \in \mathbb{N}$ and for every formula B of \mathcal{L}^+*

$$\Vdash_n B \Leftrightarrow \vdash_{\mathbf{C}_n} B$$

In the case of the other definition of validity, that is the one of Def. 4, the answer is positive as well. As a matter of fact, by exploiting the fact that this semantics is meaningful only for circular definitions respecting the requirement FR, one easily proves⁸:

⁶ See (Welch 2003). There it is shown that sets which are revision-theoretically definable with respect to the transfinite process are of a complexity that is at least Δ_2^1 (Π_2^1 in certain cases), independently of what limit rule is used.

⁷ Gupta and Belnap's book aims at dealing with a more general situation, where a base language \mathcal{L} is inflated to a language \mathcal{L}^+ which contains a countable set $\{G_i\}_{i \in \mathbb{N}}$ of circular concepts defined by formulas A_i of \mathcal{L}^+ (hence, a situation where any of these G_i depends upon all G_j 's which occur in A_i for $i, j \in \mathbb{N}$). The result we are referring to, that is (Gupta and Belnap 1993, Thrm. 5B.1, pp. 162 ff.), is proved to hold with respect to this expanded setting.

⁸ The result is stated, without proof, by Gupta in (2000).

Theorem 2 *For every formula B of \mathcal{L}^+*

$$\Vdash B \Leftrightarrow \vdash_{c_0} B$$

Proof By the previous result, this amounts to proving that B is valid in the sense of Def. 4 if, and only if B is 0-valid. Now, if B is 0-valid, then there exists a natural number m such that, for every hypothesis H

$$(\mathbf{M}, \delta^m(H)) \models B$$

In particular, if H is reflexive $(\mathbf{M}, \delta^m(H)) \models B$. Then, by the argument we have used for the sake of Lemma 1 above, we have that $(\mathbf{M}, H) \models B$ must hold for every reflexive H . Hence, B is valid.

Conversely, assume that B is valid instead. Then, $(\mathbf{M}, H) \models B$ is the case for every reflexive hypothesis H . Since moreover FR is respected, there exists $k \in \mathbb{N}$ such that, for every hypothesis H' , $\delta_A^k(H')$ is reflexive. Hence, $(\mathbf{M}, \delta_A^k(H')) \models B$ holds for every H' . This means that B is 0-valid. ■

The reader should notice that it is necessary for this argument that FR be formulated in an uniform manner, i.e. that there exists one and the same $k \in \mathbb{N}$ such that $\delta_A^k(H')$ is reflexive for every hypothesis H' . If that were not the case, than we would have no adequate syntactic notion of derivability to the notion of validity \Vdash . This shows also why the case $n = 0$ in the definition of n -reflexivity can be dispensed with, at least for circular definitions respecting FR. It is because that case is captured by the notion of validity already.

Finally, the answer to the question we raised (do we have axiom systems adequate to the semantics of finite revision?) is now complete.

However, still someone may not be entirely satisfied by the way we answered this question in the positive. For, one may also wish to know whether we have *nice* calculi which do the job. In terms of standard proof-theory, this means whether we have Hilbert-style calculi, of which we have a sequent (Gentzen-style) versions, that further feature a cut-elimination theorem showing them to be analytic.

The answer is again positive, but in order to say something more about that, we need to go a little bit into the details of the axiomatization.

8.3.1 On the Logic of Finite Revision

The aim here is to illustrate some conceptual aspects related to logical investigations over the semantics of finite revision. For all technical details, we refer the reader to (Bruni 2012) and (Gupta and Belnap 1993).

The main idea on which the formalism is based, an idea which comes from Gupta and Belnap, is to use an indexed language, with indices representing stages in a revision-theoretic evaluation of a formula. So, for a given formula A of the chosen language, and for an index i , A^i means: “ A holds at the i -th stage of a RES”.

Having fixed a standard first–order predicate language \mathcal{L} and having upgraded it to $\mathcal{L}^+ = \mathcal{L} \cup \{G(\cdot)\}$ as before, index terms are numerals \bar{p} for every $p \in \mathbb{Z}$. In the following, we are going to use i, j, h, \dots as metavariables for index terms and, having made clear that they are numerals, we use $i + 1, i + n, \dots$ with the expected meaning.

Indexed formulas are defined in such a way that, for every formula A of \mathcal{L}^+ , A^i is a formula of the indexed language for every index term i (so, notice that formulas of \mathcal{L}^+ are indexed by placing the index *outside* it).

In order to reach an Hilbert–style axiomatization which would correspond to Gupta and Belnap calculi, our idea was to mimic derivability in \mathbf{C}_n by means of a specifically devised implication connective \rightarrow . So, to every rule (\mathcal{R}) of \mathbf{C}_n there corresponds an axiom (R) of the Hilbert–style system \mathbf{HC}_n , according to the schema

$$\frac{A}{B} (\mathcal{R}) \Rightarrow A \rightarrow B (R)$$

Since A, B are indexed formulas already, hence they are formulas C^i, D^j , one has to allow formulas of the form $C^i \rightarrow D^j$ on the side of Hilbert–style systems. This means that *indices need to distribute over the new implication connective*. Moreover, we need to allow the antecedent and the consequent of \rightarrow –formulas to carry *different* index terms (see the special axioms below).

By referring to the previous informal reading of indexed formulas, the meaning of a formula of the form $A^i \rightarrow B^j$ is then: “*If A^i is the case*” (hence, if A holds at the i th–stage of a RES) “*then B^j is the case as well*”⁹.

Besides that, introducing a new connective has the expected effect on the logical base of the axiom systems $(\mathbf{HC}_n)_{n \in \mathbb{N}}$. As a matter of fact, one has to fix the “meaning” of the new arrow by setting some logical principles. It turns out that, insofar as the goal of representing derivability in systems \mathbf{C}_n is concerned, this is possible by indexing one’s preferred version of first–order classical logic¹⁰. The only prescription is that for those logical axioms and rules featuring occurrences of the new implication connective (which is necessary since we are trying to mimic axiomatically the logical inference rules of the original calculi \mathbf{C}_n), *one and the same index term is distributed over the antecedent and the consequent*. To make a straightforward example, this is how appears one of the basic axioms for conjunction:

$$(A \wedge B)^i \rightarrow A^i$$

for every formulas A, B of \mathcal{L}^+ , and index term i .

⁹ It would be natural to ask to what an extent one can represent derivability in systems \mathbf{HC}_n themselves by the new implication connective. Despite its interest, this topic would take us afield. Hence, we refer the interested reader to (Bruni 2012), where the issue is dealt with at length.

¹⁰ The process of indexing a standard logic calculus hides some subtleties, which is nonetheless unnecessary to go into here. We refer the interested reader to (Bruni 2012) for details.

Anyway the new connective *is* an implication. Hence, the logical calculus will include the following indexed version of the rule of modus ponens MP_{\rightarrow} :

$$\frac{A^i \rightarrow B^j \quad A^i}{B^j} (MP_{\rightarrow})$$

This should be compared with the rule of modus ponens for the usual, material implication, which would rather read¹¹:

$$\frac{(A \supset B)^i \quad A^i}{B^i} (MP)$$

This entails that *derivations employing only logical axioms and rules of inference cannot cause the index term of the involved formulas to increase, or decrease.*

The special axioms of systems $(\mathbf{HC}_n)_{n \in \mathbb{N}}$, take direct inspiration from the basic features of the finite revision semantics, since they derive from rules of systems \mathbf{C}_n according to the above correspondence schema. So, for instance, if $n = 0$, then the goal is to devise derivability in \mathbf{HC}_0 so to let it correspond to 0–validity. As a consequence, this system features an axiom of *index shift*, which, for every formula B of \mathcal{L} (hence, any formula which *does not feature occurrences of the circular concept* G) and for every index terms i, j , reads

$$B^i \rightarrow B^j (IS)$$

In the proposed, informal interpretation, this says that if the formula B holds at any given stage i in a RES, then it holds also at any other stage j . This corresponds to the fact that formulas of the base language retain, in the semantics of finite revision, the truth value they have on the basis of the chosen model of \mathcal{L} , *independently of the procedure of revising hypotheses.*

Moreover, the system \mathbf{HC}_0 features *G–definition* axioms, namely, for every index term i , and t term of \mathcal{L} , instances of the form

$$A_G(t, G)^i \leftrightarrow G(t)^{i+1} (DEF)$$

Clearly, this axiom tries to capture the very definition of the revision operator δ . If one looks at revision sequences as processes by means of which the extension of a circular predicate is fixed through stages of approximation, this can be viewed as saying: (i) having verified that an instance $A_G(t, G)$ of the condition defining the circular predicate G holds at stage i , *allows us to conclude* that $G(t)$ holds at stage $i + 1$; (ii) that $G(t)$ holds at any stage $i + 1$ (remember that index terms are *integers*), *requires that* the corresponding instance $A_G(t, G)$ is validated at stage i . So, the two directions of this axiom can be seen as representing two ways of going through any RES: upwards (the left–to–right direction), and downwards (the right–to–left one). This will be used for the sake of Theorem 4 in § 8.4¹².

¹¹ We use the symbol \supset for the standard implication connective.

¹² Also, the reader is invited to see Gupta and Standefer (2011) for an actual elaboration of this view.

Systems \mathbf{HC}_n for $n > 0$, are obtained from \mathbf{HC}_0 by extending it with a generalization of index shift that reads, for every formula B of the extended language \mathcal{L}^+

$$B^i \leftrightarrow B^{i+n} \text{ (IS}_n\text{)}$$

for every index term i .

Again, one can find a natural explanation for this in the semantics. Since $n > 0$, the notion of validity which corresponds to derivability in systems \mathbf{HC}_n relies upon n -reflexive hypotheses H , which are such that $\delta_A^n(H) = H$. This means that n -distant stages in a RES starting with an n -reflexive hypothesis are identical. Hence, expansions of any given model of the base language which are obtained by n -distant hypotheses actually validate the same formulas of \mathcal{L}^+ , as the axiom says.

One thing to notice about this axiomatization is that there is no need for axioms for index terms. This is worth emphasizing as an axiomatization of the transfinite theory along similar lines would require a fragment of ordinal arithmetic to be chosen as a base system. Index terms are indeed required to behave, arithmetically speaking, as ordinal numbers in that case (see Bruni 2009). Here, instead, the arithmetical behaviour of index terms is too basic for requiring any such a thing. In the end, this is another little point in favour of the finite theory.

Once the Hilbert-style systems have been devised, to provide the Gentzen-style counterparts \mathbf{GC}_n is more or less a matter of routine. One thing is worth noticing anyway. The sequent-style rules corresponding to the definition axioms are¹³:

$$\frac{A_G(t, G)^i, \Gamma \Rightarrow \Delta}{G(t)^{i+1}, \Gamma \Rightarrow \Delta} \text{ (L1)} \quad \frac{\Gamma \Rightarrow \Delta, A_G(t, G)^i}{\Gamma \Rightarrow \Delta, G(t)^{i+1}} \text{ (R1)}$$

It is immediate to notice that, in passing from the premise to the conclusion of both these rules, one cannot assume that the logical complexity of the principal formula is lower than the complexity of the active formula. This is a relevant issue as far as the aim of providing a syntactic argument of cut-elimination is concerned. The rules above behave indeed as naive abstraction rules, which are known to clash with the usage of classical logic (while they have been proved to be consistent with substructural logics instead—see, e.g., (Cantini 2003)). So, the plain fact that the cut-elimination result, as well as the corollary about the consistency of the systems are achieved, is a notable feature itself.

In addition, this is made possible by the very use of indices: the main theorem is proved by a triple induction argument, whose first parameter is specifically defined by referring to index terms. As a matter of fact, the reader can appreciate that the involved index term increases while passing from the premise to the conclusion of the rules in question.

¹³ Here we use a standard sequent notation, with Γ and Δ indicating multisets of indexed formulas of \mathcal{L}^+ .

8.4 Truth by Finite Revision

So, the approach by finite revision is a natural way of dealing with circular concepts. It also proves to have some pleasing features as far the underlying complexity, and the possibility of extracting the logic of it are concerned. What about truth then? Since truth is a circular concept, it seems natural to ask: *how much of this machinery can we expect to be able to use in the case of truth?* Very little, we must admit.

The negative part of the answer goes back to an observation which was made by Volker Halbach (1994).

Let $\mathcal{L} = \mathcal{L}_{PA}$, where the latter is the language of Peano arithmetic. Let $\mathcal{L}^+ = \mathcal{L} \cup \{T(\cdot)\}$. Fix \mathbb{N} as the chosen interpretation of the base language. Accordingly, hypotheses take the form of sets of natural numbers. Let the revision operator δ_T be defined by, for every $X \subseteq \mathbb{N}^{14}$

$$\delta_T(X) = \{\ulcorner B \urcorner \mid (\mathbb{N}, X) \models B\}$$

By taking into account the function of revision operators, δ_T is defining the extension of the truth predicate according to a disquotationalist principle: $T(\ulcorner B \urcorner)$ holds at the “revised” stage if, and only if B held at the previous one.

Then, part (ii) of Lemma 4.1 from (Halbach 1994, p. 317) shows that the following holds:

Lemma 3 *For every $X \subseteq \mathbb{N}$, for every $n \in \mathbb{N}$ such that $n > 0$, there exists a formula L_n of \mathcal{L}^+ such that*

$$(\mathbb{N}, X) \models L_n \Leftrightarrow (\mathbb{N}, \delta^n(X)) \not\models L_n$$

According to the terminology we have introduced in § 8.2, this entails that, over the standard model of arithmetic, *there are no m -reflexive hypotheses, for $m > 0$* . So, there is no chance of applying the semantics of finite revision. At least, if one bases himself on the simplified notion of validity from Def. 4.

The same source of information contains nonetheless another result, which is worth noticing.

Remember that the system **FS** by H. Friedman and M. Sheard (1987) contains the following list of axioms and rules of inference:

1. axioms **PA** of Peano arithmetic, where the induction schema

$$B(0) \wedge \forall x(B(x) \supset B(x+1)) \supset \forall x B(x)$$

is allowed to be instantiated by formulas $B(x)$ of the extended language \mathcal{L}^+ ;

2. axioms for self-referential truth¹⁵:

¹⁴ We assume that the formulas of \mathcal{L}^+ have been assigned a Gödel number according to the application of a standard arithmetization technique, and write $\ulcorner A \urcorner$ for the code of the formula A .

¹⁵ We use here the symbol \equiv for logical equivalence based on material implication \supset .

- (i) $\forall x[At(x) \supset (T(x) \equiv Ver(x))]$
- (ii) $\forall x[Sent_{\mathcal{L}^+}(x) \supset (T(\neg x) \equiv \neg T(x))]$
- (iii) $\forall x\forall y[Sent_{\mathcal{L}^+}(x) \wedge Sent_{\mathcal{L}^+}(y) \supset (T(x \dot{\supset} y) \equiv (T(x) \supset T(y)))]$
- (iv) $\forall x\forall v[Sent_{\mathcal{L}^+}(x(\bar{0}/v)) \wedge Var(v) \supset (T(\exists v x) \equiv \exists y T(x(\dot{y}/v)))]$

where the additional symbols here present come from the arithmetization of the syntax, and have the expected meaning: $At(x)$ means “ x is (the code of) an atomic formula of \mathcal{L} ”; $Ver(x)$ means “ x is (the code of) a true atomic sentence of \mathcal{L} ”; $Sent_{\mathcal{L}^+}(x)$ means “ x is a sentence of \mathcal{L}^+ ”; $Var(x)$ means “ x is a variable”.

3. the rules of inference

$$\frac{A}{T(\ulcorner A \urcorner)} \text{ NEC} \quad \frac{T(\ulcorner A \urcorner)}{A} \text{ CONEC}$$

Then, it seems legitimate to consider **FS** as *the* system of truth by finite revision. Indeed, let \mathbf{FS}_n be the systems obtained by **FS** when rules NEC and CONEC are allowed to apply $(n - 1)$ -times at most. Then, Halbach (1994, Thm. 4.2, p. 318) shows that:

Theorem 3 *For every $n \in \mathbb{N}$, and for every $X \subseteq \mathbb{N}$, $(\mathbb{N}, X) \models \mathbf{FS}_n$ if, and only if $X = \delta_T^n(Y)$ for some $Y \subseteq \mathbb{N}$.*

This means that, sticking with the standard model of arithmetic, every model of \mathbf{FS}_n has the form of an expansion of it where the set interpreting the true formulas is produced by applying n -times the machinery of finite revision¹⁶. Furthermore, it entails that every model of this latter sort, that is every structure of the form $(\mathbb{N}, \delta_T^n(X))$ for any $X \subseteq \mathbb{N}$, is a model of \mathbf{FS}_n . So the semantics of finite revision applied to (a disquotationalist definition of) truth, is adequately captured at the syntactic level by the layers \mathbf{FS}_n of **FS**.

In fact, the second direction of the above result tells us a bit more. For, it entails that if $\mathbf{FS}_n \vdash A$ is the case (A formula of \mathcal{L}^+) and Y is any given subset of \mathbb{N} , then $(\mathbb{N}, \delta_T^n(Y)) \models A$. Since obviously $\mathbf{FS} \vdash A$ if, and only if there exists $n \in \mathbb{N}$ such that $\mathbf{FS}_n \vdash A$, it follows that every theorem of **FS** is 0-valid¹⁷. This suggests that one could use the trick of representing revision stages by index terms, in order to obtain a syntactic version of Halbach’s theorem.

Let then **T** be the system made out of the following groups of axioms:

¹⁶ Non-standard models of theories \mathbf{FS}_n could be also provided. For instance, any model of **FS** is also a model of \mathbf{FS}_n . Among other things, this was pointed out to the author by an anonymous referee. The author would like to thank the referee, whose comments helped to improve the previous version of the paper.

¹⁷ To be more precise: let $i(A)$ be defined for every formula A of \mathcal{L}^+ by

$$i(A) = \min k \in \mathbb{N}. \mathbf{FS}_k \vdash A$$

Then, the previous theorem ensures that if $\mathbf{FS} \vdash A$, then $(\mathbb{N}, \delta_T^{i(A)}(Y)) \models A$ holds for every $Y \subseteq \mathbb{N}$. Hence, A is 0-valid.

1. an indexed calculus of first-order classical logic¹⁸;
2. for every index term i , B^i is an axiom of \mathbb{T} , where B is:
 - a) either an axiom of \mathbf{PA} , or
 - b) the induction schema instantiated by any formula of the language \mathcal{L}^+ , or
 - c) one of the axioms (i)–(iv) of \mathbf{FS} for self-referential truth;
3. any instance of the schema $B^i \leftrightarrow T(\ulcorner B \urcorner)^{i+1}$, for every index term i and B formula of \mathcal{L}^{+19} ;

The idea is to reproduce the revision-theoretic semantics at the syntactic level, as in § 8.3.1. The additional feature here is that the revision semantics is built-up on top of a standard arithmetical one. So, at every stage in a RES, not just the laws of classical logic, but also all and the same arithmetical truths are valid. This explains point 2 in the list²⁰.

The truth axiom schema from point 3 is obviously devised to allow the embedding of NEC and CONEC rules. Similarly to what we were noticing in § 8.3.1, this corresponds to see NEC as representing one step forward in any RES, and CONEC being one step backwards. This must be kept in mind for the definition of the partial function $j(\cdot)$ in Theorem 4 below.

Having said that, we prove:

Theorem 4 *Set, for every formula B of \mathcal{L}^+ , $j(B) \in \mathbb{Z}$ to be such that:*

$$\left\{ \begin{array}{l} j(B) = 0, \text{ if } B \text{ is derivable in } \mathbf{FS} \text{ with no applications} \\ \text{of } \mathbf{NEC} \text{ and } \mathbf{CONEC}; \\ j(B) = \min k. (k = m - m'), \text{ if } B \text{ is derivable in } \mathbf{FS} \text{ with } m \text{ applications} \\ \text{of } \mathbf{NEC}, \text{ and } m' \text{ applications of } \mathbf{CONEC}. \end{array} \right.$$

Then:

$$\mathbf{FS} \vdash B \Rightarrow \mathbb{T} \vdash B^{\overline{j(B)}}$$

Proof The proof is by induction on $m + m'$ where m is the number of occurrences of the rule NEC in the given proof, and m' is the number of occurrences of the rule CONEC.

¹⁸ Since the system \mathbf{FS} we are trying to embed has a logic base system in axiomatic form, the required indexed version needs not to feature fully nested occurrences of the new \rightarrow connective. For the sake of Theorem 4 below, the reader can think of \rightarrow as occurring solely in the truth-definition axioms from point 3 of the present list. The reader should consult (Bruni 2012) for related remarks.

¹⁹ This requires that the indexed modus ponens $\mathbf{MP}_{\rightarrow}$ from § 8.3.1 be among the logical inference rules of \mathbb{T} .

²⁰ Notice that, among the axioms we have spoken of in § 8.3.1 we have *not* included in \mathbb{T} the schema of index shift (IS). Hence, the need of assuming A^i as axiom for every formula A which is valid at every stage i . Anyway, due to the possibility that the induction schema be instantiated by formulas of \mathcal{L}^+ , the inclusion of (IS) in the above list would not have been sufficient to obtain an equivalent, and maybe more elegant system.

$m + m' = 0$ The theorem then follows by definition of \mathbb{T} .

$m + m' > 0$ Assume that B is obtained by an application of NEC. This means that $B = T(\ulcorner C \urcorner)$ and C is derivable with m'' -many applications of NEC, and $m'' < m$. By the induction hypothesis, $\mathbb{T} \vdash C^{\overline{j(C)}}$. This means that $B^{\overline{j(C)+1}}$ is derivable in \mathbb{T} by an application of the definition axiom and modus ponens MP_{\rightarrow} . Since $j(C) + 1 = j(B)$, this ends the proof.

If B is obtained by an application of the CONEC rule instead, $T(\ulcorner B \urcorner)$ is derivable in FS with m'' -many applications of CONEC with $m'' < m'$. By the induction hypothesis, $\mathbb{T} \vdash T(\ulcorner B \urcorner)^{\overline{j(T(\ulcorner B \urcorner))}}$, and by the definition axiom of \mathbb{T} and MP_{\rightarrow} , $\mathbb{T} \vdash B^{\overline{j(T(\ulcorner B \urcorner))-1}}$. But, $j(B) = j(T(\ulcorner B \urcorner)) - 1$, hence the theorem.

If B is obtained by neither NEC, nor CONEC, though these rules have been used in the course of the proof, one simply observes that the usage of all axioms and rules in \mathbb{T} which are different from the definition axiom, cannot increase or decrease the index of any derivable formula (see § 8.3.1). Hence, the theorem. ■

A similar result going in the other direction, i.e. from \mathbb{T} to FS , would be possible. However, with Theorem 4 in mind, this is no surprise. Nor it would add relevant information. Hence, we have decided to leave it out of this paper, for the sake of space consideration.

8.5 Conclusion

The proposed evaluation of the finite revision semantics yielded a two-sided result. On the one hand this approach has some nice features. It is satisfactory from the point of view of the involved complexity, especially if compared to the transfinite extension of it. The semantics of finite revision is also suitable for a logical development, since it has a corresponding derivability notion which is adequate, and the syntactic approach can be dealt with in ways which are attractive for the proof-theorist. Last but not least, the machinery of revision seems to closely resemble the kind of reasoning which we would expect someone to follow, in situations where problems of circularity are involved. Hence, it is natural, as we claimed.

However, there are limits to the possibility of applying this approach. For, the theory in its nicest form requires that one is dealing with a circular concept whose defining condition respect the finitness requirement FR (see § 8.2). Luckily, we do have an interesting example of a circular concepts of this sort which is the concept of rationality in finite games. This is the reason why, however, this semantics has limited applications in the case of truth. The embeddability result we have presented is nonetheless interesting because it exploits the representation of stages in a revision path at the syntactic level via index terms. The way this can be done, in particular without any arithmetical assumption, is an additional advantage over the transfinite theory. On the basis of this result, there is an additional reason for stressing the tied connection between truth as formalised by FS , and validity under finite revision.

Can we go further? That is: can we do something similar for stability under *transfinite* revision? This is known to be one of the open problems related to this approach. Though certainly not something in the scope of the present paper.

References

- Bruni, R. (2009). A note on theories for quasi-inductive definitions. *The Review of Symbolic Logic*, 2, 684–699.
- Bruni, R. (2012). Analytic calculi for circular concepts by finite revision. *Studia Logica*. doi:10.1007/s11225-012-9402-2.
- Bruni, R., & Sillari, G. (2011). A rational way of playing (In preparation).
- Cantini, A. (2003). The undecidability of Grišhin's set theory. *Studia Logica*, 74, 345–368.
- Chapuis, A. (2000). Rationality and circularity. In A. Chapuis & A. Gupta (Eds.), *Circularity, definition, and truth* (pp. 49–77). Indian Council of Philosophical Research.
- Friedman, H., & Sheard, M. (1987). An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic*, 33, 1–21.
- Gupta, A. (2000). On circular concepts. In A. Chapuis & A. Gupta (Eds.), *Circularity, definition, and truth* (pp. 123–153). Indian Council of Philosophical Research.
- Gupta, A., & Belnap, N. (1993). *The revision theory of truth*. MIT Press.
- Gupta, A., & Standefer, S. (2011). Conditionals in theories of truth (In preparation).
- Halbach, V. (1994). A system of complete and consistent truth. *Notre Dame Journal of Formal Logic*, 35, 311–327.
- Herzberger, H. (1982). Notes on naive semantics. *Journal of Philosophical Logic*, 11, 61–102 .
- Takeuti, G. (1987). *Proof theory*. North Holland.
- Welch, P. (2003). On revision operators. *Journal of Symbolic Logic*, 68, 689–711.