Peano's Counterexample to Harmony

Leonardo Ceragioli*

Abstract

Harmony and conservative extension are two *criteria* proposed to discern between acceptable and unacceptable rules. Despite some interesting works in this field, the exact relation between them is still not clear. In this paper, some standard counterexamples to the equivalence between them are summarised, and a recent formulation of the notion of stability is used to express a more refined conjecture about their relation. Then Prawitz's proposal of a counterexample based on the truth predicate to this refined conjecture is shown to rest on dubious assumptions. As a consequence, two new counterexamples are proposed: one uses the extension of logic with a small amount of arithmetic, while the other uses the extension of a small fragment of arithmetic with a problematic operator defined by Peano. It is argued that both these new counterexamples work fine to reject the conjecture and that the last one works also as a rejection of harmony as a complete *criterion* of acceptability of rules.

1 Harmony, conservativeness and counterexamples

We will deal with two *criteria* that should distinguish well-defined rules from ill-defined ones: harmony and conservativeness. There are different definitions of harmony, but we will deal mainly with Prawitz's characterization in terms of the Inversion Principle. According to this principle, the E-rules for a connective should be in some sense the inverse of the corresponding I-rules. More precisely:[1]

> Let α be an application of an elimination rule that has B as consequence. Then, deductions that satisfy the sufficient condition $[\cdots]$ for deriving the major premiss of α, when combined with deductions of the minor premisses of α (if any), already "contain" a deduction of B; the deduction of B is thus obtainable directly from the given deductions without the addition of α.

The rules for a connective are harmonious if and only if they suit the Inversion Principle (as far as we are concerned).[2] Harmony of a pair of I and E-rules makes it possible to reduce maximal formulae, that is the conclusions of I-rules that are also premises of E-rules can be erased from the derivation in a standard way, preserving its correctness. As an example, let us consider the two following derivations:

*Department of Civilisations and Forms of Knowledge, Università di Pisa

[1] [Prawitz, 1965], p. 33.

[2] We will consider some more recent definitions of harmony in the next chapters.

$$\begin{array}{c}\supset\mathrm{E}\,\dfrac{\begin{array}{c}\vdots\\ A\end{array}\qquad \supset\mathrm{I}\,\dfrac{\begin{array}{c}[A]\\ \vdots\\ B\end{array}}{A\supset B}}{B}\\ \vdots\end{array}\quad\rightsquigarrow\quad\begin{array}{c}\vdots\\ A\\ \vdots\\ B\\ \vdots\end{array}$$

It is obvious that we can always substitute an occurrence of the first with an occurrence of the second in every formal proof, and in this way we have erased (the occurrence of) $A \supset B$ from the derivation.

Application of conservative extension to meaning-theoretic considerations can be traced at least to Belnap, who uses this criterion to reject Prior's infamous "connective" *tonk*.[3] In this case the definition is quite easy and precise:

> Given two theories **T** and **T'** such that the second proves all the consequences provable in the first and the vocabulary of the second extends the vocabulary of the first, **T'** conservatively extends **T** iff **T** proves all the consequences provable in **T'** in which only elements of the vocabulary of **T** occur.

In his main book on these topics, Michael Dummett suggests that those two *criteria* are deeply connected,[4] and indeed every natural deduction formulation of classical logic breaks both harmony and conservativeness, while there is a separable and harmonious system for intuitionistic logic.[5] Nonetheless, it is not easy to find out precisely which position he supports about their relation, since Dummett evaluates different hypotheses and scatters his considerations throughout the entire book. In addition, he makes the really unhappy choice of using the term "harmony" for a lot of different *criteria*. For these and other reasons his real position on this topic is quite a controversial exegetical problem.

For example, on p. 219 of his [Dummett, 1991] the author asserts and gives grounds to believe that conservativeness is a necessary requirement for harmony, but this position is rejected in a later paragraph of the book.[6] Indeed, he exposes the case of a language containing standard conjunction, standard disjunction and quantum disjunction as a counterexample to the conjecture that harmonious systems can extend other harmonious systems only in a conservative way,[7] so he is provably well aware of the mismatch between these two notions. A reasonable interpretation is that he decided to present the entire development of his position, starting from the old conjecture and with only a minimal change in terminology; we will follow this reading.

Let us briefly look at this counterexample since it is useful to explain the notion of stability that we will use later. Quantum disjunction is defined by the following set of rules:

[3] [Prior, 1960], [Belnap, 1962].

[4] [Dummett, 1991], pp. 217-220.

[5] To be precise, this is the situation in the standard natural deduction formulation – and it is indeed in this framework that our discussion has to be evaluated –, in other contexts the situation is different: [Hjortland and Standefer, 2018], p. 121.

[6] Stephen Read has correctly pointed out that the argument used by Dummett to prove that conservativeness follows from harmony is unsound: [Read, 2000], pp. 126-127.

[7] [Dummett, 1991], p. 288.

$$\sqcup\mathrm{I}\ \frac{\Gamma \vdash A}{\Gamma \vdash A \sqcup B} \qquad \sqcup\mathrm{I}\ \frac{\Gamma \vdash B}{\Gamma \vdash A \sqcup B} \qquad \sqcup\mathrm{E}\ \frac{\Gamma \vdash A \sqcup B \qquad A \vdash C \qquad B \vdash C}{\Gamma \vdash C}$$

What distinguishes this from standard disjunction is the E-rule, that for $\sqcup$ does not allow extra assumptions in the derivation of minor premises. We already know that for standard disjunction and conjunction, harmony holds. This weakening of disjunction does not prevent the reduction needed for harmony, so also the rules for $\sqcup$ are harmonious.[8] Despite this, $\sqcup$ is weaker than $\vee$, since in a system composed by rules for quantum disjunction and standard conjunction it is not possible to prove the law of distribution $A \wedge (B \sqcup C) \vdash (A \wedge B) \sqcup (A \wedge C)$.

Let us now consider the extension of the system composed by rules for quantum disjunction and standard conjunction with rules for standard disjunction. In a system with both $\sqcup$ and $\vee$ we can prove $A \sqcup B \dashv\vdash A \vee B$, so distribution of $\wedge$ over $\sqcup$ becomes provable. By virtue of this phenomenon, we have a harmonious system that is a non-conservative extension of another harmonious system, so conservativeness cannot be a necessary requirement for harmony.[9]

This counterexample is very instructive because it points both to a diagnosis of what causes non-conservativeness, and to a possible solution. It seems that non-conservativeness is due to an *ad hoc* weakening of the E-rule that is at play with $\sqcup$, since it prevents a full usage of the meaning given to $\sqcup$ by its introduction rules. If the base system has this deficiency, then its extension with unrestricted $\vee$-rules unlocks the situation and allows the full usage of $\sqcup$.[10] According to this diagnosis, in order to prevent non-conservative extensions, we should ask that: the base system be constructed with harmonious rules *that fully use the meaning of the terms*; the extension be made with harmonious rules.

Dummett follows this path, proposing to integrate the notion of *harmony* with that of *stability*.[11] A set of rules for a term is stable if it is harmonious and the E-rules among them fully use the meaning given to the term by the I-rule among them.

There are some recent proposals to obtain a formal *criterion* for this intuitive idea, in the same way as availability of the reduction steps is the formal counterpart of the intuitive idea of containment of the conclusion in the major premise of E-rules. We will focus mainly on [Jacinto and Read, 2017], but Luca Tranchini has a similar criterion.[12] The authors propose the following definition of *inverse harmony* – that together with harmony gives stability:[13]

[8] Nonetheless, it prevents permutative conversions and so normalizability does not hold, as correctly remarked by Nissim Francez: [Francez, 2017].

[9] For an interesting analysis of this kind of mismatches between harmony, conservativeness and other criteria in proof-theoretic semantics, see [Steinberger, 2013].

[10] Of course in this case also the ordering in the composition of the language is relevant. For example, quantum disjunction conservatively extends standard conjunction and standard disjunction (or standard conjunction and material conditional).

[11] [Dummett, 1991], chapter 13. Sometimes he uses the term 'harmony' also for this refined notion. Some contemporary researchers follow him in this decision ([Steinberger, 2011], [Tranchini, 2016] are some examples, while in [Jacinto and Read, 2017] 'harmony' is used only for the weaker requirement). To avoid any misunderstanding, we will use this term in the narrower sense.

[12] [Tranchini, 2016], pp. 17-18.

[13] [Jacinto and Read, 2017], p. 373 for the definition of stability and p. 397 for the *criterion* of inverse harmony (the authors call this *criterion* 'Generalized Local Completeness', but we will use the same term for the property and the formal *criterion*). We make a simplification of the *criterion* that is harmless in our case, and omit the definition for the empty set of I-rules (since it is needed only for *ex falso quodlibet*, that we do not use).

Definition 1.1 (Inverse Harmony)**.** If C can be derived from the *direct grounds* for $A \oplus B$ together with the assumptions Γ_i with $1 \leq i \leq m$, then C can be derived from $A \oplus B$ together with the assumptions $\bigcup_{i=1}^{m} \Gamma_i$ by appealing only to the first derivations (one for each ground) and $\oplus$E.

Where the direct grounds for a sentence are what you need in order to derive it by the appropriate I-rule. Different entities can be direct grounds for a sentence. For example the direct ground for $A \supset B$ is a derivation of B from A, while $A \vee B$ has two direct grounds: the sentence A and the sentence B.[14]

As an example of inverse harmony, let us consider the case of disjunction. We saw that the direct grounds for $A \vee B$ are A and B, so the common consequences of these two sentences together with some extra-assumptions shall be derivable also from $A \vee B$ by composing the old derivations with $\vee$E. That this is possible can be shown by the following *expansion*:

$$\begin{array}{cc} \Gamma_1 \quad A \\ \vdots \Phi_1 \\ C \end{array} \qquad \begin{array}{c} \Gamma_2 \quad B \\ \vdots \Phi_2 \\ C \end{array} \qquad \rightsquigarrow \qquad \vee\mathrm{E}_1 \, \dfrac{A \vee B \qquad \begin{array}{c} \Gamma_1 \quad [A]^1 \\ \vdots \Phi_1 \\ C \end{array} \qquad \begin{array}{c} \Gamma_2 \quad [B]^1 \\ \vdots \Phi_2 \\ C \end{array}}{C}$$

So the rules for standard disjunction are stable. Nevertheless, this is not the case for those for quantum disjunction, since it is obviously impossible to give the previous *expansion* when we have the infamous restriction on side assumptions for deriving minor premises.

It is also important to remark that the usage of both E-rules and previous derivations from direct grounds to the conclusion is not only allowed but required, in order to have a proper expansion.[15] Indeed, let us consider the following connective $\hookrightarrow$:

$$\hookrightarrow\mathrm{I} \, \frac{A \vdash B}{\vdash A \hookrightarrow B} \qquad \hookrightarrow\mathrm{E} \, \frac{\Gamma \vdash A \hookrightarrow B \qquad \Delta \vdash A}{\Gamma, \Delta \vdash B}$$

That is $\hookrightarrow$ has the same E-rule of $\supset$, but a restriction on open assumptions in its I-rule. A look at our reduction step for $\supset$ shows that this restriction does not interfere with harmony, so we have to rely on inverse harmony to block this pair of rules.[16] Nonetheless, it could seem that inverse harmony holds for $\hookrightarrow$:

[14]I think that this characterization of direct grounds is sufficient for our discussion. Nonetheless, [Jacinto and Read, 2017] gives a precise definition of this notion on p. 368, to which we refer.

[15][Jacinto and Read, 2017] does not discuss this problem, and also [Tranchini, 2016] is silent about this requirement, even though he explicitly discusses the $\hookrightarrow$I example in relation to another proposed *criterion* of stability.

[16] With this I do not mean that this pair of rules should be rejected *tout court*; I just mean that it cannot be a pair of stable rules, since *modus ponens* is already stable in relation with $\supset$I. Indeed stability is surely a desirable property, but maybe its lack is not sufficient for the rejection of a pair of rules. Dummett considers it at most as a *criterion* for terms the meaning of which is completely determined by inferences, for example purely logical terms: [Dummett, 1991], pp. 286-287. A consequence of this interpretation should be that quantum disjunction (and so maybe also $\hookrightarrow$) is not a purely logical term, but it is in some way obtained by empirical investigations: [Dummett, 1991], p. 289.

$$
\begin{array}{c}
A \\
\vdots\,\Phi_1 \\
B \qquad \Gamma \\
\vdots\,\Phi_2 \\
C
\end{array}
\quad \rightsquigarrow \quad
\begin{array}{c}
\hookrightarrow\mathrm{E}\,\dfrac{A \hookrightarrow B \qquad A}{B} \qquad \Gamma \\
\vdots\,\Phi_2 \\
C
\end{array}
$$

In order to show what is wrong with this argument, we have to consider the case in which $\Gamma = \emptyset$, Φ_2 is composed by a single application of $\hookrightarrow$I and $C = A \hookrightarrow B$. In this case we need an expansion of

$$
\begin{array}{c}
[A] \\
\vdots\,\Phi_1 \\
\hookrightarrow\mathrm{I}\,\dfrac{B}{A \hookrightarrow B}
\end{array}
$$

and we cannot relate to

$$
\hookrightarrow\mathrm{E}\,\dfrac{A \hookrightarrow B \qquad [A]}{\hookrightarrow\mathrm{I}\,\dfrac{B}{A \hookrightarrow B}}
$$

since the first occurrence of $A \hookrightarrow B$ causes a violation in the restriction of $\hookrightarrow$I on open assumptions. Nonetheless, if we do not impose both application of the E-rule and usage of Φ_2 (in this case the derivation of $A \hookrightarrow B$ by $\hookrightarrow$I), we could propose an assumption of $A \hookrightarrow B$ (that is a one-step derivation) as the required expansion, by exploiting the fact that accidentally the conclusion of the derivation that we have to expand has the same form of the sentence that we can use as assumption in the expansion.

With this machinery, we can give a new, refined conjecture about the relation between harmony and conservativeness:[17]

Conjecture 1.1 (Dummett)**.** *Adding harmonious rules to a system composed of only stable rules always results in a conservative extension.*

This conjecture is not so easy to invalidate, indeed there is only one well-known counterexample in the literature. In the following section, we will consider and reject this counterexample, while in the rest of the paper we will propose some new counterexamples.

2 Truth predicate

The only well-known proposed counterexample to Dummett's refined conjecture regards the extension of a starting system with the truth predicate. The first author to propose it has been Prawitz, in his review of [Dummett, 1991].[18] Unfortunately, he does not explain his counterexample in detail but, since it has been endorsed by some authors, we can rely on them for a more extended formulation.[19] The general idea is that the extension of Peano Arithmetic with

[17] [Dummett, 1991], p. 290.
[18] [Prawitz, 1994].
[19] For example Stewart Shapiro and Stephen Read: [Shapiro, 1998], [Read, 2000].

the truth predicate makes derivable the Gödel sentence for **PA**, and that we should be able to give harmonious and stable rules for this predicate.[20]

I will argue that we can raise three doubts about this counterexample:

- We can doubt that there are harmonious rules for the truth predicate;
- We can doubt that the starting system can be formulated with stable rules;
- As a minor point, we can raise doubts about which philosophical lesson we should learn from this alleged counterexample.

Let us consider these points one by one.

2.1 Harmonious extension

We can find a harmonious formulation of rules for the truth predicate $\mathscr{T}$ in [Shapiro, 1998]:[21]

$$\mathscr{T}\mathrm{I}\,\frac{A}{\mathscr{T}(\ulcorner A \urcorner)} \qquad \mathscr{T}\mathrm{E}\,\frac{\mathscr{T}(\ulcorner A \urcorner)}{A}$$

With the side condition that A must be an arithmetical sentence.[22]

First of all, let us remark that to obtain a non-conservative extension – and so a counterexample to Dummett's conjecture – besides these rules we need all the instances of Induction Schema in which $\mathscr{T}$ occurs, as explicitly acknowledged by Shapiro. Indeed, adding only $\mathscr{T}$I and $\mathscr{T}$E we obtain a conservative extension of **PA**.[23] So we have to evaluate harmony of this further extension as well.

Moreover, Steinberger argues that even this extension of Induction Schema is not sufficient to have non-conservativeness, and that we need a full, compositional theory of truth.[24] However, I think that his reasons to believe this are not very convincing since, while it is surely true that in an axiomatic theory of truth we should explicitly postulate compositionality of the truth predicate[25], it is not so obvious that we are forced to do the same in a theory based on natural

[20] Also in this case, it is important to remark that we are interested in extensions with the truth predicate only of standard natural deduction systems (see note 5). The consequences of this extension in a substructural framework could be very different: [Hjortland and Standefer, 2018], p. 127.

[21] Where $\ulcorner A \urcorner$ is the Gödelian code for the sentence A.

[22] An anonymous referee suggested that we should impose to have no open assumptions in order to apply $\mathscr{T}$I, and asked which logical system is extended by these rules. These rules are not supposed to have conditions on the open assumptions and the system that we extend is a standard natural deduction system, according to the authors. The precise strength of the natural deduction system is never specified, but arguably it is intuitionistic. In this way, we can derive the truth of any arithmetical sentence from the assumption of the sentence itself, and *vice versa*, reconstructing T-biconditionals. Nonetheless, I completely agree that this counterexample should be stated more clearly. Indeed, none of the authors that endorsed it has proposed a precise formulation and, since the standard approach to theories of truth is axiomatic, not based on rules of natural deduction, it cannot be considered part of the received wisdom.

[23] [Horsten, 2011], p. 78.

[24] [Steinberger, 2011], p. 635.

[25] See chapter 6 of [Horsten, 2011].

deduction. Nonetheless, in Steinberger's defence, we should acknowledge that the proof of non-conservativeness of this extension is up to the objectioners.[26]

From now on, let us just concede that the exemplifications of Induction Schema with the truth predicate are enough to gain non-conservativeness, and evaluate their acceptability. Shapiro seems to defend this extension in his article, but we have to keep in mind that he is working under the hypothesis *ad absurdum* that **PA** can be formulated using only harmonious (or stable) rules. His reason to assume this is given by the main topic of the paper, that is Tennant's neo-logicism.[27] Admittedly having harmonious exemplifications of the Induction Schema with purely arithmetical vocabulary, it becomes at least reasonable to assume harmony of its exemplifications with $\mathscr{T}$. On the contrary, away from this framework the assumption of harmony for this extension is unjustified. So, in order to evaluate our first doubt about this counterexample (harmony of the extension), we need to deal with our second doubt: availability of a stable formulation for **PA**.

2.2 Stable base system

In general, the stability of the base system **PA** is not discussed in those papers that endorse the counterexample to Dummett's conjecture based on truth. We can agree that a neo-logicist working in the Dummettian tradition should consider possible to give such a formulation at least for arithmetic, even though the community of philosophers working on this tradition has proposed great departures from the original orthodox Dummettian view.[28] Nonetheless, this is a problem only for the extra assumption of neo-logicism, not for the conjecture about the relation between harmony, stability, and conservativeness that we are evaluating. What we could at most conclude is that the conjecture is incompatible with logicism, and so that we should not expect to have a good Dummettian version of logicism.

In conclusion, we can be sceptical regarding the existence of a stable formulation of **PA** for the same reasons we are sceptical about the possibility to have harmonious exemplifications of Induction Schema in which the truth predicate occurs. But the situation now is even worse, since we look for stability and not just for harmony.

[26]As an example, in the natural deduction framework we can easily derive $\mathscr{T}(\ulcorner\neg A\urcorner) \supset\subset \neg\mathscr{T}(\ulcorner A\urcorner)$ but, in order to obtain the compositional axiom for negation, we need to generalise it in the following way: $\forall x(x \in \mathscr{L}_{PA} \supset (\mathscr{T}(\neg x) \supset\subset \neg\mathscr{T}(x)))$. The availability of this generalization depends on the details of the formulation of the theory, so we need a precise formulation to decide about Steinberger's observation.

$$\supset\subset\mathrm{I}\,\dfrac{\supset\mathrm{I}_2\,\dfrac{\neg\mathrm{I}_1\,\dfrac{\neg\mathrm{E}\,\dfrac{\mathscr{T}\mathrm{E}\,\dfrac{[\mathscr{T}(\ulcorner\neg A\urcorner)]^2}{\neg A}\qquad \mathscr{T}\mathrm{E}\,\dfrac{[\mathscr{T}(\ulcorner A\urcorner)]^1}{A}}{\bot}}{\neg\mathscr{T}(\ulcorner A\urcorner)}}{(\mathscr{T}(\ulcorner\neg A\urcorner)) \supset (\neg\mathscr{T}(\ulcorner A\urcorner))}\qquad \supset\mathrm{I}_4\,\dfrac{\mathscr{T}\mathrm{I}\,\dfrac{\neg\mathrm{I}_3\,\dfrac{\neg\mathrm{E}\,\dfrac{[\neg\mathscr{T}(\ulcorner A\urcorner)]^4\qquad \mathscr{T}\mathrm{I}\,\dfrac{[A]^3}{\mathscr{T}(\ulcorner A\urcorner)}}{\bot}}{\neg A}}{\mathscr{T}(\ulcorner\neg A\urcorner)}}{(\neg\mathscr{T}(\ulcorner A\urcorner)) \supset (\mathscr{T}(\ulcorner\neg A\urcorner))}}{(\neg\mathscr{T}(\ulcorner A\urcorner)) \supset\subset (\mathscr{T}(\ulcorner\neg A\urcorner))}$$

[27]As we saw in note 16 logic seems to be characterized by having stable rules, so in this framework the logicist thesis that arithmetic is logic entails that arithmetic can be formulated using stable (and so harmonious) rules.

[28]As an example, let us just consider that in [Read, 2000] the author asserts that coherence is not a necessary requirement for harmony.

Of course, these observations do not entail that a good formulation of this counterexample cannot be found, what we can at most conclude is that we do not have a good formulation yet and that the existence of such a formulation is not entailed by Dummett's theory of meaning. For a philosopher that wants to propose a good formulation of this counterexample, it is essential to give a good answer to these first two objections. No acceptable counterexample can avoid them.

2.3 Which lesson

Another problem with Prawitz's counterexample is that it does not offer any suggestion for deciding which principle we should save and which one we should abandon, after the discovery of this alleged mismatch. That is, also accepting it as a formal refutation of Dummett's conjecture, it does not have any philosophically clear consequences about the acceptability of harmony, stability, and conservativeness as *criteria* for well-constructed sets of rules. It is not clear whether we should accept $\mathbf{PA}+\mathscr{T}$ and reject conservativeness as a too stringent requirement, or reject $\mathbf{PA}+\mathscr{T}$ based on conservativeness and drop harmony (at least as a complete *criterion*). Indeed the topic of the conservativeness of truth-theoretic extensions of arithmetic is a very controversial one, connected to deflationism, and our interpretation of this alleged counterexample relies heavily on our solution to this philosophical puzzle.[29] *Naïvely* we could say that a deflationist (that equates deflationism with conservativeness) shall conclude from the counterexample that harmony is not a good *criterion*, while a proof-theoretic semanticist shall conclude that conservativeness is not a necessary requirement.[30] So philosophers with differing background opinions about the two *criteria* will not agree about the consequences of this counterexample. Of course this last objection is just a minor point but, nonetheless, it would be preferable if we could find a less controversial case for deciding between these two alternatives.

We will propose in the next sections two further counterexamples: the first solves all the objections apart from this last one, while the other seems to solve also this last problem. Indeed, we will conclude that neither harmony nor stability is trustworthy *criteria*.[31]

2.4 Invalid counterexamples

In this brief section, I explain my reasons to exclude from this discussion two alleged counterexamples that could be found in the literature. The first is Stephen Read's proof-theoretic version of the liar's paradox:[32]

[29] [Horsten, 2011] is a good introductory book to the relation between conservativeness and deflationism.

[30] As an example, Prawitz in his review of [Dummett, 1991] rejects conservativeness, following his background. Indeed he foresaw the possibility of this mismatch and already decided to reject conservativeness in his [Prawitz, 1985] (p. 166).

[31] Of course, this conclusion says something also about conservativeness but is not enough to decide whether it is a good criterion. This other question has to be endorsed in its own right.

[32] [Read, 2000], pp. 140-142.

$$\bullet\mathrm{I}\ \dfrac{\neg\bullet}{\bullet} \qquad \bullet\mathrm{E}\ \dfrac{\bullet \qquad \begin{matrix}[\neg\bullet]\\ \vdots \\ D\end{matrix}}{D}$$

This pair of rules extends non-conservatively any consistent logical system, since it can be used to derive $\perp$:

$$\neg\mathrm{E}\ \dfrac{\neg\mathrm{I}_2\ \dfrac{\neg\mathrm{E}\ \dfrac{\bullet\mathrm{E}_1\ \dfrac{[\bullet]^2 \qquad [\neg\bullet]^1}{\neg\bullet} \qquad [\bullet]^2}{\perp}}{\neg\bullet} \qquad \bullet\mathrm{I}\ \dfrac{\neg\mathrm{I}_2\ \dfrac{\neg\mathrm{E}\ \dfrac{\bullet\mathrm{E}_1\ \dfrac{[\bullet]^2 \qquad [\neg\bullet]^1}{\neg\bullet} \qquad [\bullet]^2}{\perp}}{\neg\bullet}}{\bullet}}{\perp}$$

Moreover, the author argues that $\bullet$E is in harmony with $\bullet$I, so this seems to qualify as a counterexample to Dummett's conjecture.

My reason to consider invalid this counterexample is that $\bullet$I does not suit the minimum requirements to be considered as a real introduction rule. Indeed we should at least demand that the conclusion of $\bullet$I be no less complex than its premise:[33]

> Hence the minimal demand we should make on an introduction rule intended to be self-justifying is that its form be such as to guarantee that, in any application of it, the conclusion will be of higher logical complexity than any of the premisses and than any discharged hypothesis. We may call this the 'complexity condition'.

Of course, we have to weaken this complexity condition (together with separability, as we will see) when we consider introduction rules for non-logical terms, but we should nonetheless ask for a conclusion that is less complex than neither any of the premises, nor any of the discharged hypothesis. This is the reason why I will not consider this counterexample further in this article.

A similar point can be made for another supposed counterexample, that inspired Read's connective $\bullet$. Indeed in Appendix B of his [Prawitz, 1965], the author proposes the following harmonious pair of rules:

$$\lambda\mathrm{I}\ \dfrac{A[x/t]}{t \in \lambda x A} \qquad \lambda\mathrm{E}\ \dfrac{t \in \lambda x A}{A[x/t]}$$

These rules for set theory could seem acceptable at first glance, but they lead to Russell's paradox and so to contradiction if added to a well behaved logical system. Indeed, let us define $t = \lambda x.\neg(x \in x)$ and consider the following derivation, that seems to reject Dummett's conjecture:

$$\supset\mathrm{E}\ \dfrac{\lambda\mathrm{I}\ \dfrac{\supset\mathrm{I}_1\ \dfrac{\supset\mathrm{E}\ \dfrac{[t \in t]^1 \qquad \lambda\mathrm{E}\ \dfrac{[t \in t]^1}{\neg(t \in t)}}{\perp}}{\neg(t \in t)}}{t \in t} \qquad \supset\mathrm{I}_2\ \dfrac{\supset\mathrm{E}\ \dfrac{[t \in t]^2 \qquad \lambda\mathrm{E}\ \dfrac{[t \in t]^2}{\neg(t \in t)}}{\perp}}{\neg(t \in t)}}{\perp}$$

[33] [Dummett, 1991], p. 258.

We reject this counterexample for the same reason why we reject Read's one, although in this case the violation of the complexity condition is not obvious. Indeed a brief look at the rules could deceive us and leave unnoticed this violation, since while every application of •I has a premise that is more complex than its conclusion, only some applications of λI manifest this behaviour. Nonetheless, let us consider the way in which this rule is applied in the just seen derivation of a contradiction. If we make t explicit, it has the form:

$$\lambda\mathrm{I}\ \frac{\neg((\lambda x.\neg(x \in x)) \in (\lambda x.\neg(x \in x)))}{(\lambda x.\neg(x \in x)) \in (\lambda x.\neg(x \in x))}$$

that obviously violate the complexity condition. In conclusion, we can reject also this pair of rules, and so the only counterexample that we need to consider extensively is the one based on the truth predicate, that uses at least a set of acceptable rules.[34]

3 Two new counterexamples

3.1 'Two objects' counterexample

The first counterexample that I want to propose has been already used to criticize conservative extension. Let us consider the sentence $\exists x \exists y(x \neq y)$. It should be a logical sentence since it is formulated using only logical terms. But nonetheless, it cannot be proved in any logical system, while it easily follows from the arithmetical theorem $1 \neq 0$ by two applications of $\exists$I. As Shapiro observes, this means that the extension of a logical system with the rules for arithmetic is non-conservative and poses in serious threat the separability of concepts needed for a molecularist theory of meaning.[35]

It can be observed that this example has consequences also for the relation between harmony, stability, and conservativeness. Indeed it is enough to have a stable formulation of the base logical system and to give a harmonious formulation of the small fragment of arithmetic necessary to prove $1 \neq 0$ in order to have a sound counterexample to Dummett's conjecture. This goal is much less demanding than a stable formulation of the entire system **PA**, also because in this case we do not need stability but harmony, since arithmetic is used in the extension and not in the base system. Moreover, the requirement of harmony seems to be applicable to the entire language.[36]

Steinberger evaluates something similar to this counterexample. He proposes to add a quotation operator '[ ]' (or in general a term-forming operator) to logic, in order to derive $[A] \neq [B]$ for A logical theorem and B contradiction, from a proof of the first and a rejection of the second. In this way we can obtain a non-conservative extension of logic by applying two times $\exists$I and deriving $\exists x \exists y(x \neq y)$, following essentially the pattern of the counterexample that we just saw (Steinberger speculates that this phenomenon could also be the hidden reason why the truth predicate can lead to non-conservativeness). Nonetheless, he poses correctly the problem of existence of a harmonious formulation for this

[34] Although, as we already saw, we have other more complex reasons to reject it as a counterexample to Dummett's conjecture.

[35] [Shapiro, 1998], p. 603.

[36] [Dummett, 1991], p. 287.

quotation operator (since quotation is possible in **PA** via arithmetization, the problem of existence of a harmonious formulation for **PA** is just a special case of this).[37] What he fails to acknowledge is that we have other, less demanding ways of proving $a \neq b$ for two terms; indeed the extension of logic with a very small fragment of arithmetic is enough.

Maybe the reason why Steinberger prefers to extend logic with a quotation operator is that it is easier to consider this as a logical system (the same could be said for logic extended with the truth predicate). Indeed even a small fragment of arithmetic leads to a clear departure from the realm of logic (at least if you are not a logicist), and Dummett poses stability as an adequate *criterion* only for logical terms.[38] But as already remarked in this case we do not need stability of arithmetic, but just harmony, and this property should hold of the entire language, according to Dummett.

In a next section, we will see that it is possible to give a formal shape to this counterexample as opposed to the one based on the truth predicate. The only problem that remains is that it is not clear whether it speaks against harmony (and stability) or conservative extension. Indeed the problem of the ontological presupposition of logic is just a little less controversial than that of deflationism regarding truth: maybe in a complete logical system we should be able to prove $\exists x \exists y (x \neq y)$, and so we could obtain a conservative extension with arithmetic.[39] Rejection of harmony and stability as complete *criteria* for the acceptability of rules depends on our position regarding this philosophical issue.

3.2 Peano's counterexample

Our second new counterexample to Dummett's conjecture regarding the relation between harmony, stability, and conservativeness is inspired by an observation made by Peano regarding an ill-defined operation between fractions.[40] Let us consider these two rules:

$$?\mathrm{I}\ \frac{(a+c)/(b+d) = e/f}{(a/b)?(c/d) = e/f} \qquad ?\mathrm{E}\ \frac{(a/b)?(c/d) = e/f}{(a+c)/(b+d) = e/f}$$

They are obviously in harmony with each other since the elimination rule just derives the premise of the introduction rule from its conclusion. Indeed, the reduction step for a ?-maximal formula can easily be found:

$$\begin{array}{c} \vdots \\ ?\mathrm{E}\ \dfrac{?\mathrm{I}\ \dfrac{(a+c)/(b+d) = e/f}{(a/b)?(c/d) = e/f}}{(a+c)/(b+d) = e/f} \\ \vdots \end{array} \quad \rightsquigarrow \quad \begin{array}{c} \vdots \\ (a+c)/(b+d) = e/f \\ \vdots \end{array}$$

[37] [Steinberger, 2011], p. 636.

[38] Or at least for terms the meaning of which is fully specified by rules of inference; see note 16.

[39] This is, for example, the position of Tennant: [Tennant, 1997].

[40] Peano's observation is contained in [Peano, 1921]. Belnap refers to it when arguing for conservativeness; see [Belnap, 1962], p. 131.

Looking at this reduction step, these rules seem to be as innocuous as useless. Nonetheless, they non-conservatively extend every standard theory of arithmetical fractions, since we can easily derive an absurdity using them together with some basic arithmetical and identity laws:

$$\dfrac{\dfrac{\dfrac{\dfrac{(1+1)/(2+3)=2/5}{(1/2)?(1/3)=2/5}\,?\mathrm{I} \qquad 1/2=2/4}{(2/4)?(1/3)=2/5}\,\text{Sub. of Id.}}{(2+1)/(4+3)=2/5}\,?\mathrm{E}}{3/7=2/5}\,\text{Add.}$$

Since obviously $3/7 \neq 2/5$, we have found a contradiction in our theory. As a consequence, a stable formulation of the weak fragment of arithmetic needed for this counterexample is all we need for rejecting Dummett's conjecture.

We could restate for this counterexample a criticism we saw in the previous section: stability here is applied to (a small fragment of) arithmetic while it should hold only for logic. While it is true that we are applying stability to a non-logical theory (as opposed to the 'two objects' counterexample), this criticism is nonetheless irrelevant. We are just proposing a counterexample to a formal conjecture (with philosophical implications, of course) about the relation between harmony, stability, and conservativeness. If stability holds *by accident* of a subsystem of **PA**, this can be a perfectly good counterexample.[41] Moreover, the philosophical consequences of this rejection are not less relevant just because they are obtained from an arithmetical counterexample.

Unlike the previous 'two objects' counterexample, in this case it is clear which lesson is taught: harmony has to be rejected, in favour of conservativeness or maybe in favour of a more complex *criterion*. That is, the philosophical background is essentially irrelevant for evaluating the consequences of this counterexample. Indeed we cannot be satisfied with a *criterion* that does not prevent contradiction, whatever our philosophical opinions (or even deep convictions) are.

3.2.1 General-elimination rule for Peano

In this small paragraph, I just want to assure the reader that our counterexample based on Peano's operator cannot be avoided by using general-elimination harmony. Indeed let us consider the following general-elimination rule for ?:

$$\dfrac{(a/b)?(c/d)=e/f \qquad \begin{array}{c}[(a+c)/(b+d)=e/f]\\ \vdots \\ C\end{array}}{C}\,?_g\mathrm{E}$$

It is easy to show that this rule is in harmony with the standard introduction rule for ?, both according to the characterization of harmony given by Prawitz,

[41] If you are not satisfied with the idea that stability can hold *by accident* of some systems, then you could consider Peano's counterexample as a refutation both of Dummett's conjecture about the formal relation between harmony, stability, and conservativeness and of his opinion about the applicability of stability only to logical terms.

and according to the characterization of general-elimination harmony given by Read. Indeed Read essentially argues that the E-rule for a connective should be obtainable from the I-rule and should have the following shape: the major premise should be the conclusion of the I-rule, the minor premise should be general and the rule should allow the discharge of all open assumptions on which the minor premise depends that have the same form of a direct ground of the major premise.[42]

So, since $?_g\mathrm{E}$ is obtained as a unique function of ?I and has the structure required by Read, it is harmonious according to Read's conception. In order to show that it is also harmonious according to Prawitz's conception, we just have to show that maximal formulae can still be erased, in this way:

$$?_g\mathrm{E}\,\dfrac{?\mathrm{I}\,\dfrac{\begin{matrix}\vdots\\(a+c)/(b+d)=e/f\end{matrix}}{(a/b)?(c/d)=e/f} \qquad \begin{matrix}[(a+c)/(b+d)=e/f]\\ \vdots\\ C\end{matrix}}{C} \quad\rightsquigarrow\quad \begin{matrix}\vdots\\(a+c)/(b+d)=e/f\\ \vdots\\ C\end{matrix}$$

Nonetheless, Peano's counterexample still holds:

$$?_g\mathrm{E}_1\,\dfrac{\text{Sub. of Id.}\,\dfrac{?\mathrm{I}\,\dfrac{(1+1)/(2+3)=2/5}{(1/2)?(1/3)=2/5} \qquad 1/2=2/4}{(2/4)?(1/3)=2/5} \qquad \text{Add.}\,\dfrac{[(2+1)/(4+3)=2/5]^1}{3/7=2/5}}{3/7=2/5}$$

4 The old objections

Let us now check whether the old objections we used to reject the counterexample based on the truth predicate exclude our new counterexamples, as well. The first objection was about the harmony of the rules used in the non-conservative extension, and in particular about the harmony of the exemplifications of Induction Schema with the truth predicate. The new counterexample based on Peano's operator does not use Induction Schema and we already saw that its rules are harmonious, so this objection is not really at issue for it. Nonetheless, we still need a harmonious fragment of arithmetic capable of proving $a \neq b$ for two terms, in order to formalize our 'two objects' counterexample. We will not deal with this issue directly since in the following paragraph we will see a stable system for an elementary fragment of arithmetic suitable also for this purpose.

Our last (minor) objection was that it is not clear whether we want the extension with the truth predicate to be conservative or not, so it is not clear which lesson we should learn from Prawitz's counterexample about the acceptability of the *criteria*. We saw that this problem is shared also by the 'two objects' counterexample since there can be philosophical disagreements about whether logic should entail the existence of some objects. Nonetheless, as we stressed, Peano's counterexample solves also this last problem.

In conclusion, there is just one residual problem:

[42]The most complete definition of this kind of harmony is given in [Read, 2010], while here we just give a simplified version of this notion. Nonetheless, this simplification is acceptable, since for Peano's operator we have only one fully local I-rule with just one premise (that is only one sentence is the direct ground for the conclusion).

- Is there a stable formulation of the base system that we non-conservatively extend?

Since the base system of the 'two objects' counterexample is pure logic, this is not a problem for it. In the next section we will argue that this objection can be answered for Peano's counterexample as well and that, as already disclosed, a consequence of this result is the possibility of harmoniously extending pure logic with enough arithmetic to prove the existence of two distinct numbers.

4.1 BA* is stable

Let us call Baby Arithmetic (**BA**) the system obtained by adding the following rules to the standard natural deduction system for *minimal propositional* logic:[43]

$$\mathfrak{s}\mathrm{I}\ \frac{m=n}{\mathfrak{s}(m)=\mathfrak{s}(n)} \qquad \mathfrak{s}\mathrm{E}\ \frac{\mathfrak{s}(m)=\mathfrak{s}(n)}{m=n} \qquad \bot\mathrm{I}\ \frac{\mathfrak{s}(n)=0}{\bot} \qquad \bot\mathrm{E}\ \frac{\bot}{\mathfrak{s}(n)=0}$$

$$+0\mathrm{I}\ \frac{m=n}{m+0=n} \qquad +0\mathrm{E}\ \frac{m+0=n}{m=n} \qquad +\mathrm{I}\ \frac{\mathfrak{s}(m+n)=l}{m+\mathfrak{s}(n)=l} \qquad +\mathrm{E}\ \frac{m+\mathfrak{s}(n)=l}{\mathfrak{s}(m+n)=l}$$

$$\times0\mathrm{I}\ \frac{0=n}{m\times 0=n} \qquad \times0\mathrm{E}\ \frac{m\times 0=n}{0=n} \qquad \times\mathrm{I}\ \frac{(m\times n)+m=l}{m\times\mathfrak{s}(n)=l} \qquad \times\mathrm{E}\ \frac{m\times\mathfrak{s}(n)=l}{(m\times n)+m=l}$$

$$=\mathrm{I}\ \frac{\begin{array}{c}[F(a)]\\ \vdots\\ F(b)\end{array}}{a=b} \qquad =\mathrm{E}\ \frac{a=b \qquad A(a)}{A(b)}$$

The rules for identity are proposed by Stephen Read and seem to be in harmony.[44] Some clarification is needed for them: in order to apply =I, F has to be a fully general *predicative variable* which does not occur in other open assumptions; =E enables the substitution of any number of occurrences of a in *A(a)*.

We call **BA*** the theory that extends **BA** with the following rules for fractions:

$$\div\mathrm{I}\ \frac{n\times s=l\times m}{n/m=l/s} \qquad \div\mathrm{E}\ \frac{n/m=l/s}{n\times s=l\times m}$$

Of course, these two rules do not constitute a complete theory of rational numbers - for example, addiction between fractions is not defined. Nonetheless, we do not need a complete theory for our counterexample, but just a sound one. We can show that this stable theory is strong enough to enable the introduction of Peano's problematic operator, with all its bad consequences:

[43] See [Prawitz, 1965], p 21 for minimal logic. For **BA** I take inspiration from the arithmetic system with the same name in [Smith, 2013], which I label here **SBA** for clarity. In the next section, we will use the equivalence of these two systems.

[44] [Read, 2004]. Nonetheless the acceptability of his rules is disputed; for example, [Klev, 201X] proposes a defence of an older formulation of Martin-Löf and criticizes Read's formulation. We have no reason to believe that any formulation of identity could reject our counterexample, since substitution of identicals is a *desideratum* for every such formulation.

$$
{\scriptstyle \div\mathrm{E}}\,\dfrac{{\scriptstyle =\mathrm{E}}\,\dfrac{{\scriptstyle =\mathrm{E}}\,\dfrac{{\scriptstyle =\mathrm{E}}\,\dfrac{{\scriptstyle =\mathrm{E}}\,\dfrac{{\scriptstyle ?\mathrm{E}}\,\dfrac{{\scriptstyle =\mathrm{E}}\,\dfrac{\begin{matrix}\vdots\\ 1/2 = 2/4\end{matrix} \quad {\scriptstyle ?\mathrm{I}}\,\dfrac{{\scriptstyle =\mathrm{I}_1}\,\dfrac{[F((1+1)/(2+3))]^1}{(1+1)/(2+3) = (1+1)/(2+3)}}{(1/2)?(1/3) = (1+1)/(2+3)}}{(2/4)?(1/3) = (1+1)/(2+3)}}{(2+1)/(4+3) = (1+1)/(2+3)} \quad \begin{matrix}\vdots\\ 1+1=2\end{matrix}}{(2+1)/(4+3) = 2/(2+3)} \quad \begin{matrix}\vdots\\ 2+3=5\end{matrix}}{(2+1)/(4+3) = 2/5} \quad \begin{matrix}\vdots\\ 2+1=3\end{matrix}}{3/(4+3) = 2/5} \quad \begin{matrix}\vdots\\ 4+3=7\end{matrix}}{3/7 = 2/5}}{\dfrac{3 \times 5 = 2 \times 7}{15 = 14}}
$$

Where F is a fully general predicative variable, and we use standard numbers for numerals (for example '2' for '$\mathfrak{ss}(0)$').

We overlook the provability of elementary arithmetical truths like $1+1=2$ and $1/2 = 2/4$ in **BA***. We could show the complete derivation, but it is easier and more fruitful to invoke a more general result:

Theorem (Atomic Completeness of **BA***). *For every atomic sentence E of* ***BA****

- $\vdash_{BA^*} E$ *or;*
- $\vdash_{BA^*} \neg E$.

Proof. The completeness of the atomic fragment of **BA** follows from its equivalence with the already mentioned Smith's system (**SBA**), the completeness of which is established by the author.[45] So, we have to deal only with its extension to rational numbers, that is with the extension to **BA***.[46]

First of all, let us remember that sentences of the form $a/b + c/d = e/f$ are not well formed in **BA***. The rules for fractions deal only with sentences like $a/b = e/f$, so we will consider as not well-formed other kinds of atomic sentences regarding fractions.[47]

In virtue of the completeness of the atomic fragment of **BA** we can rightly decide every sentence of the form $a = b$ in which a and b are terms constructed using only 0, $\mathfrak{s}$, $+$ and $\times$. In order to decide a sentence of the form $a/b = c/d$, **BA*** just decides the sentence $a \times d = c \times b$. If $a \times d = c \times b$ is proved in **BA***, then also $a/b = c/d$ is proved by the same derivation with $\div$I as an extra step. If $a \times d = c \times b$ is rejected in **BA***, then also $a/b = c/d$ is rejected by the following derivation:

$$
{\scriptstyle \neg\mathrm{I}_1}\,\dfrac{{\scriptstyle \neg\mathrm{E}}\,\dfrac{\begin{matrix}\vdots\\ \neg(a \times d = c \times b)\end{matrix} \quad {\scriptstyle \div\mathrm{E}}\,\dfrac{[a/b = c/d]^1}{a \times d = c \times b}}{\bot}}{\neg(a \times d = c \times b)}
$$

[45][Smith, 2013], pp. 65-66.

[46]Technically speaking, **SBA** is formulated using classical logic, but this is not a problem, at least for Atomic Completeness. A complete proof of this kind of completeness for **BA** (that follows the line of that in [Smith, 2013]) is shown in the appendix, together with the details of the derivation of the axioms of **SBA** from the rules of **BA**.

[47]This restriction is not essential: we could just extend our theory in order to treat addiction and multiplication between fractions, and our counterexample would still apply. Our choice is just to keep the presentation as simple as possible.

So, **BA*** rightly decides every well-formed atomic formula, and it is an extension of **SBA**, conservative over the atomic formulae. This entails that it is possible to derive the true arithmetical sentences used in our counterexample, and so that Peano's operator gives a non-conservative extension of **BA***.[48] $\square$

In conclusion, it should also be obvious that $\mathscr{T}$ can extend **BA*** at most in a conservative way. Indeed, its eventual inclusion in **BA*** is even pointless, since this system cannot define self-reference (and does not have induction). This means that $\mathscr{T}$ cannot be used in place of Peano's operator in **BA***, and that this function causes problems in very weak systems, in which $\mathscr{T}$ is ineffective. In summary, at the end of this section, we have a formal counterpart of Peano's counterexample that is formulated in a base system that seems to be stable. To complete our counterexample, in the next sections we will argue explicitly that this system is stable, considering before the logical and purely arithmetical rules and then the rules for identity.

4.1.1 The logical and arithmetical rules are stable

The logic of our base system is both *minimal* and *propositional*, so we have neither any rule for $\perp$ (apart from those of **BA**), nor object variables and quantifiers. This is one of the stable systems according to Jacinto and Read's *criterion*, and in general its stability seems a shareable starting point.

Moreover, the lack of some logical rules is not a problem, quite contrary: showing a counterexample for harmony in minimal logic we have a stronger argument since disputable rules like *ex falso quodlibet* are not used.[49] In other words, presenting our counterexample in a weak and undisputed logic makes us more ecumenical.[50]

Now, we can broaden our scope and evaluate stability also for purely arithmetical rules; we will discuss the stability of the rules for identity in the next section. Harmony and stability of these rules are uncontroversial, since the only premise of every I-rule is the conclusion of the corresponding E-rule, while its conclusion is the only premise of the corresponding E-rule. Developing our counterexample in a minimal system, we also prevent any doubts about the pair of rules for $\perp$. Indeed had we assumed *ex falso quodlibet*, it would have been considered as an elimination rule for $\perp$, raising the problem of its harmony (and stability) with $\perp$I.[51]

[48] To be precise, we need also coherence or soundness of **BA*** in order to establish this result. Nonetheless, the first follows from the second, which seems an obvious property of this system.

[49] For the problems in the standard treatments of negation, see [Kürbis, 2015]; I believe that Kürbis' criticism does not apply to the small fragment of negation and absurdity that we need to use here.

[50] An anonymous referee stressed that the extension with the rules for identity seems to require a less undisputed logic since we need predicative variables. While I agree that this extension is less uncontroversial than propositional minimal logic, I do not think that we necessitate a full second-order system in order to have =I. Indeed we do not need second-order quantifiers but just variables, and since we started with a propositional logic, we do not have first-order quantifiers either. Nonetheless, what is really important here is the availability of a stable formulation of the base system, and we can give one for **BA*** while this is at least controversial for **PA**.

[51] The interpretation of *ex falso quodlibet* as an elimination rule is very common today, and accepted also by Jacinto and Read. Historically it has been considered as a special rule, external to the distinction between I and E-rules (for example by Prawitz), while something

Even though harmony and stability are not problematic, there can be a further reason to reject **BA*** and its extension with Peano's operator. Indeed *separability* is a standard requirement of proof-theoretic semantics: every term should be graspable in isolation, without reference to other terms.[52] I interpret this requirement as asking that in every rule only one term occurs in a non-schematic way: as an example $\wedge$I is separable because only $\wedge$ occurs non-schematically in it, while in ?I there are $/, \div, +$ and ?.[53] Sure enough, the usual rules for logic are separable, while those for arithmetic and for ? are not, and so they characterise the meaning of a term using the meaning of some other terms (for example the meaning of $\times$ relies on that of $+$).

This observation points to a very interesting aspect of the language but, nonetheless, I do not think it can be used as an objection to **BA***. Indeed Dummett postulates separability only for logical terms and not for the non-logical fragments of the language.[54] So the necessity of dropping separability in order to reach richer fragments of the language is not in contrast with Dummett's view of the language.[55] We have an application of stability in a non-logical framework in which it is completely acceptable to have non-separable rules, and this application rejects Dummett's conjecture about stability and conservativeness. Dummett's complexity condition is not violated either in an essential way, that is for every introduction rule its conclusion is not more complex than its premise. So both principles are extended in a natural way to the non-logical part of the language.

Someone could nonetheless object that, at least from a formal point of view, an easy way out can be found by changing our conjecture and asking for separability of the stable base system, of the harmonious extension or of both of them. The new conjecture so defined should not be rejected by our counterexample. Apart from the *ad hocness* of this reformulation, since our defence of the non-separability of the arithmetical language, also this reformulation can be rejected. Indeed our 'two objects' counterexample can be formalized using minimal logic as base system and the $\perp$-rules of **BA** as (stable and *a fortiori*) harmonious extension. This system is obviously separable, but nonetheless it leads to a non-conservative extension of minimal logic.

4.1.2 The rules for identity are stable

Since harmony and stability of Read's rules for identity are not completely obvious, we will prove them explicitly. In order to show harmony, we just need to consider the following reduction step:

like our $\perp$I ha been proposed by Peter Milne: [Milne, 1994], p. 64.

[52] I thank an anonymous referee for stressing this point, I also thank Salvatore Florio for posing a related question during the exposition of these ideas at the Second Graduate Conference of the Italian Network for the Philosophy of Mathematics.

[53] This requirement is sometimes called 'purity', while 'separability' is used to indicate that a result can be derived using only rules for the terms that occur in it. Since the relation of this notion with conservativeness is obvious, I think it is more interesting to discuss separability as purity.

[54] He also seems sceptic about the real strength of this assumption for the logical fragment; maybe what we really want is non-circularity: see [Milne, 2002], pp. 522-523.

[55] Nonetheless, he still rejects holism, that is the meaning of a term cannot depend on the entire language.

$$
{\scriptstyle =\mathrm{E}}\,\dfrac{{\scriptstyle =\mathrm{I}}\,\dfrac{\begin{matrix}[F(a)]\\ \vdots\,\Phi \\ F(b)\end{matrix}}{a=b} \qquad A(a)}{A(b)} \qquad \rightsquigarrow \qquad \begin{matrix}A(a)\\ \vdots\,\Phi^* \\ A(b)\end{matrix}
$$

Where Φ^* is obtained from Φ by substituting every occurrence of F with an occurrence of A.

Since F is a predicative variable, there cannot be rules that require it in order to be applicable, so we are sure that if Φ is a valid derivation, so is Φ^*. In the same way, since F cannot occur in other open assumptions apart from $F(a)$, we are sure that the open assumptions of Φ^* are the right ones. So we have shown that harmony holds for the identity rules.[56]

In order to show stability, we just need to prove inverse harmony, following the definition of Jacinto and Read. The following expansion shows exactly this:

$$
\begin{matrix}\begin{matrix}F(a)\\ \vdots\,\Phi_1 \\ F(b)\end{matrix} \quad \Gamma \\ \vdots\,\Phi_2 \\ C\end{matrix} \qquad \rightsquigarrow \qquad \begin{matrix}{\scriptstyle =\mathrm{E}}\,\dfrac{a=b \qquad F(a)}{F(b)} \quad \Gamma \\ \vdots\,\Phi_2 \\ C\end{matrix}
$$

Where $\begin{matrix}F(a)\\ \vdots\,\Phi_1 \\ F(b)\end{matrix}$ is the direct ground for $a = b$, and indeed it is substituted by this sentence in the second proof tree. This completes the proof of stability of the rules for identity and the proof of stability of **BA***. We dealt with all the old objections used against the counterexample based on the truth predicate and proved that they are ineffective for our counterexample based on Peano's operator.

5 Conclusions

We have seen a counterexample – which uses *ad hoc* weakened E-rules – to the hypothesis that harmony and conservativeness are equivalent to each other. We have then explained how stability can be used to reject this kind of counterexamples and so, using this machinery, we have displayed another conjecture proposed by Dummett about the relationship between harmony, stability, and conservativeness. We dealt with the only well-known counterexample to this conjecture proposed in the literature: the extension of **PA** with harmonious rules for the truth predicate. We pointed out some objections to this counterexample and proposed two new counterexamples: the 'two objects' counterexample and the Peano's operator counterexample. The first solves all the old objections except

[56] They are also in general-elimination harmony, as shown by Stephen Read in [Read, 2016], so also in this case we cannot object that the conjecture could be saved for this more recent version of harmony.

for the last one: we cannot conclude a clear lesson about the suitability of harmony and conservativeness as good *criteria* for the acceptability of sets of rules. The second counterexample solves also this problem, since it clearly rejects harmony and stability as complete *criteria* of proof-theoretical acceptability.

References

Belnap, N. D. (1962). Tonk, plonk and plink. *Analysis*, 22(6):130–134.

Dummett, M. (1991). *The Logical Basis of Metaphysics.* Harvard University Press, Cambridge (Massachussets).

Francez, N. (2017). On harmony and permuting conversions. *Journal of Applied Logic*, 21:14–23.

Hjortland, O. and Standefer, S. (2018). Inferentialism, structure and conservativeness. In Ondřej Beran, V. K. and Koreň, L., editors, *From Rules to Meanings: New Essays on Inferentialism*, pages 115–140. Routledge, New York, Abingdon.

Horsten, L. (2011). *The Tarskian turn: deflationism and axiomatic truth.* MIT press, Cambridge, Massachusetts.

Jacinto, B. and Read, S. (2017). General-elimination stability. *Studia Logica*, 105:361–405.

Klev, A. (201X). The harmony of identity. *Journal of Philosophical Logic.* Unpublished, preprint available at `https://www.academia.edu/38081597/The_harmony_of_identity]`.

Kürbis, N. (2015). What is wrong with classical negation? *Grazer Philosophische Studien*, 92:51–86.

Milne, P. (1994). Classical harmony: Rules of inference and the meaning of the logical constants. *Synthese*, 100:49–94.

Milne, P. (2002). Harmony, purity, simplicity and a "seemingly magical fact". *Monist*, 85:498–534.

Peano, G. (1921). Le definizioni in matematica. *Periodico di Matematiche*, 1(3):175–189. Eng. Trans. in [**?**].

Prawitz, D. (1965). *Natural Deduction: A Proof-Theoretic Study.* Almqvist & Wiksell, Stockholm.

Prawitz, D. (1985). Remarks on some approaches to the concept of logical consequence. *Synthese*, 62:153–71.

Prawitz, D. (1994). Review of: The logical basis of metaphysics. *Mind*, (103):373–376.

Prior, A. (1960). The runabout inference ticket. *Analysis*, 21(2):38–39.

Read, S. (2000). Harmony and autonomy in classical logic. *Journal of Philosophical Logic*, (29):123–54.

Read, S. (2004). Identity and harmony. *Analysis*, 64:111–19. A revised version at `https://www.st-andrews.ac.uk/~slr/identity_revisited.pdf`.

Read, S. (2010). General-elimination harmony and the meaning of the logical constants. *Journal of Philosophical Logic*.

Read, S. (2016). Harmonic inferentialism and the logic of identity. *Review of Symbolic Logic*, 9:408–20. A revised version at `https://www.st-andrews.ac.uk/~slr/Inferentialism-Identity-amended.pdf`.

Shapiro, S. (1998). Induction and indefinite extensibility: The Gödel sentence is true, but did someone change the subject? *Mind*, 107(427):597–624.

Smith, P. (2013). *An Introduction to Gödel's Theorems*. Cambridge University Press, Cambridge, 2 edition.

Steinberger, F. (2011). What harmony could and could not be. *Australasian Journal of Philosophy*, 89(4):617–639.

Steinberger, F. (2013). On the equivalence conjecture for proof-theoretic harmony. *Notre Dame Journal of Formal Logic*, 54(1):79–86.

Tennant, N. (1997). *The Taming of The True*. Oxford University Press.

Tranchini, L. (2016). Proof-theoretic harmony: towards an intensional account. *Synthese*.

Appendix

In order to prove the equivalence of **BA** with Smith's system, we just have to derive the following axioms:

1. $\mathfrak{s}(n) \neq 0$
2. $\mathfrak{s}(m) = \mathfrak{s}(n) \supset m = n$
3. $m + 0 = m$
4. $m + \mathfrak{s}(n) = \mathfrak{s}(m + n)$
5. $m \times 0 = 0$
6. $m \times \mathfrak{s}(n) = (m \times n) + m$

It can be easily done in the following way:

$$ {\scriptstyle \neg\mathrm{I}_1}\dfrac{{\scriptstyle \bot\mathrm{I}}\dfrac{[\mathfrak{s}(n) = 0]^1}{\bot}}{\neg(\mathfrak{s}(n) = 0)} \qquad {\scriptstyle \supset\mathrm{I}_1}\dfrac{{\scriptstyle \mathfrak{s}\mathrm{E}}\dfrac{[\mathfrak{s}(m) = \mathfrak{s}(n)]^1}{m = n}}{\mathfrak{s}(m) = \mathfrak{s}(n) \supset m = n} \qquad {\scriptstyle +0\mathrm{I}}\dfrac{{\scriptstyle =\mathrm{I}_1}\dfrac{[F(m)]^1}{m = m}}{m + 0 = m} $$

$$ {\scriptstyle +\mathrm{I}}\dfrac{{\scriptstyle =\mathrm{I}_1}\dfrac{[F(\mathfrak{s}(m + n))]^1}{\mathfrak{s}(m + n) = \mathfrak{s}(m + n)}}{m + \mathfrak{s}(n) = \mathfrak{s}(m + n)} \qquad {\scriptstyle \times 0\mathrm{I}}\dfrac{{\scriptstyle =\mathrm{I}_1}\dfrac{[F(0)]^1}{0 = 0}}{m \times 0 = m} \qquad {\scriptstyle \times\mathrm{I}}\dfrac{{\scriptstyle =\mathrm{I}_1}\dfrac{[F(m \times n + m)]^1}{m \times n + m = m \times n + m}}{m \times \mathfrak{s}(n) = m \times n + m} $$

This should be sufficient to convince the reader that Smith's system and **BA** are equivalent. Nonetheless, since we formulate **BA** using minimal logic, it is preferable to give an explicit proof of the completeness of the atomic fragment of **BA**.[57] In this way we establish both completeness of **BA** and its equivalence with **SBA**.[58]

Theorem (Atomic Completeness)**.** *For every atomic sentence E of **BA***

- $\vdash_{BA} E$ *or;*
- $\vdash_{BA} \neg E$.

In order to prove this we need to observe that every well formed atomic sentence of **BA** is $a = b$ for some terms a and b. We also need the following lemma:

Lemma. *For every term t of **BA** there is a natural number n such that $\vdash_{BA}$ $t = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n}(0)$.*

Proof. The proof is by induction on the number of occurrences of formal symbols (0, $\mathfrak{s}$, $+$ and $\times$) in the term t. The basis is obvious, since if only one term occurs in t, it must be 0.

The inductive step is proved by cases on the most external function:

Case 1 ($t = \mathfrak{s}(r)$): In this case, we know by induction that $r = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n}(0)$ for some natural number n, so we conclude $t = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n+1}(0)$ by an application of =E.

Case 2 ($t = q + r$): In this case, we know by induction that $q = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0)$ and $r = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n}(0)$ for some natural numbers m and n. If $n = 0$, by +0E we obtain $q + r = q$, and so we conclude $t = q = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0)$ by an application of =E. If $n \neq 0$, by +E we obtain $\mathfrak{s}(\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) + \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n-1}(0)) = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) + \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n}(0)$. For inductive hypothesis $\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) + \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n-1}(0) = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m+n-1}(0)$, so by =E we have $t = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m+n}(0)$.[59]

Case 3 ($t = q \times r$): In this case, we know by induction that $q = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0)$ and $r = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n}(0)$ for some natural numbers m and n. If $n = 0$, by $\times$0E we obtain $q \times r = 0$, and so we conclude $t = 0$. If $n \neq 0$, by $\times$E we obtain

[57] The proof is similar to that for **SBA** in [Smith, 2013], but we focus on the fact that we do not need non-minimal logical rules: *ex falso quodlibet* and *tertium non datur*.

[58] Of course the two systems are different *if* we consider **BA** formulated with minimal rules and **SBA** formulated with classical rules.

[59] To be precise, inductive hypothesis just give us the existence of a natural number l such that $\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) + \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n-1}(0) = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{l}(0)$, but we do not want to be too pedantic.

$\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) \times \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n-1}(0) + \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) = q \times r$. $\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) \times \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n-1}(0)$ is less complex than $q \times r$, so by induction we have $\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) \times \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n-1}(0) = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m\times(n-1)}(0)$. In this way, we obtain $\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m\times(n-1)}(0) + \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) = q \times r$, and we can use the procedure of the additive case to reduce it to $\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m\times n}(0) = q \times r$.

This concludes the proof by cases and the proof of this lemma. Let us notice that no purely classical logical principles have been used. $\square$

We now can prove Atomic Completeness in an easy way:

Proof. Every well-formed atomic sentence of **BA** has the form $a = b$ for two terms a and b. We know by the previous lemma that $\vdash_{BA} a = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0)$ and $\vdash_{BA} b = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n}(0)$. We just have to show that: if $m = n$, $\vdash_{BA} a = b$; if $m \neq n$, $\vdash_{BA} \neg(a = b)$. In the first case, we compose the two derivations in the following way:

$$
{=}\mathrm{E}\ \dfrac{\overset{\vdots}{a = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0)} \qquad \overset{\vdots}{b = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n}(0)}}{a = b}
$$

While in the second case, if $m \geq n$ we use:

$$
\neg\mathrm{I}_1\ \dfrac{\bot\mathrm{I}\ \dfrac{\mathfrak{s}\mathrm{E}_{\ n\ times}\ \dfrac{{=}\mathrm{E}\ \dfrac{{=}\mathrm{E}\ \dfrac{\overset{\vdots}{a = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0)} \qquad [a = b]^1}{\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) = b} \qquad \overset{\vdots}{b = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n}(0)}}{\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n}(0)}}{\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m-n}(0) = 0}}{\bot}}{\neg(a = b)}
$$

And if $m < n$:

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[a=b]^2}{\vdots}}{\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{m}(0) = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n}(0)}}{0 = \overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n-m}(0)}\;\mathfrak{s}\mathrm{E}_{\ m\ times} \qquad \dfrac{[F(0)]^1}{0=0}\;{=}\mathrm{I}_1}{\overbrace{\mathfrak{s}\cdots\mathfrak{s}}^{n-m}(0) = 0}\;{=}\mathrm{E}}{\perp}\;\perp\mathrm{I}}{\neg(a=b)}\;\neg\mathrm{I}_2
$$

□

It can be observed that in order to derive E or $\neg E$ for E atomic we use only non-logical rules and $\neg$-rules. So Atomic Completeness of **BA** holds for every choice of logic in which $\neg$I and $\neg$E are admissible.[60]

[60]In particular it holds both for minimal logic and classical logic formulations of **BA**.