Not so stable

Florian Steinberger

Analysis, 64: 4 (2009)
Penultimate draft

1 Harmony as equilibrium

According to Michael Dummett we may think of the meaning of an expression as given by the principles governing the use we make of it. The principles regulating our linguistic practices can then be grouped into two broad categories (Dummett 1973, p. 396, 1991, p. 211). We might state them as follows:

- **I-principles**: state the circumstances under which an assertion of a sentence containing the expression in question is warranted.

- **E-principles**: state the consequences of asserting a sentence containing the expression.

In the case of the logical constants, we may associate a constant's I-principles with the set of its introduction rules and its E-principles with the set of its elimination rules (Dummett 1973, p. 454).[1]

To ensure that these principles confer coherent meanings on the logical expressions of our language, introduction rules and elimination rules cannot be fixed wholly independently of one another. There must be a certain 'consonance' between the two features of use; I-principles and E-principles must be in *harmony*. This amounts to the proviso that pairs of I- and E-principles must be chosen in such a way that E-principles allow us to infer *no more* and *no less* than what we could have deduced directly from the premises of the corresponding I-principles. Put another way, disharmony can take one of two forms:

- **E-strong disharmony**: the E-principles are unduly permissive;

[1] A logical constant may of course be governed by more than one introduction or elimination rule or both. For simplicity I will nevertheless speak loosely of 'pairs of rules' where strictly speaking I should write 'pairs of sets of rules'.

- **E-weak disharmony**: the E-principles are unduly prohibitive.

Whereas the regular principle of harmony understood as a form of conservativeness constraint[2] or reduction procedure[3] requires only that E-principles not outstrip I-principles—that is, it only bars E-strong disharmony—this stronger logic-specific brand of harmony, which Dummett calls *stability* (Dummett 1991, p. 287), promises to ward off both types of disharmony. Disappointingly, Dummett never goes on to propose a rigourous formulation of it. It is far from clear how the details of his rather sketchy account are to be filled in.[4]

The account to examine, rather—the most detailed and careful account of stability—is due to Neil Tennant.[5] The remainder of this paper is devoted to this account. I begin by expounding Tennant's rendering of the principle of stability. I then present a counterexample to it, showing that the account gets the rules for the existential quantifier wrong: instead of picking out our standard quantifier rules, it selects a patently disharmonious strengthened version of these rules.

2 Tennant's notion of harmony

We had said that for a stable balance to obtain an elimination rule must exploit, by way of the inferential consequences it licenses, all and only the content conferred on its major premise by the corresponding introduction rule. In other words, we should be able to derive from a statement with the constant in question in a dominant position no more than what we are entitled to in virtue of being in a position to assert it properly; nor should we be able to derive from it any less than is warranted on the basis of the premises of the introduction rule. To the end of giving this intuitive idea a precise form, Tennant introduces the notions of the logically *strongest* and *weakest* propositions with a certain property.

- A is the *strongest* proposition with property P if, for any proposition B with the same property, A entails B.

- A is the *weakest* proposition with property P if A is entailed by any proposition B with the same property.

It should be noted that I am here following Tennant's non-standard use of 'proposition': a proposition A is the logical equivalence class of which A is a member, i.e. the class of statements logically equivalent to A.

[2]See e.g. Belnap 1962, Dummett 1991, p. 217.

[3]See e.g. Dummett 1991, p. 250.

[4]In addition, his discussion is slightly obscured by the fact that he juggles several issues at once. See Dummett 1991, ch. 13.

[5]Although Tennant does not himself use the term 'stability', I will retain it.

With the aid of these tools, Tennant then proceeds to define the principle of harmony (the lower case ‘h’ is crucial here). An (arbitrary binary logical) operator $ is harmonious if the following two conditions are met:

(S) $(A, B) is the strongest conclusion possible under the conditions described by $-I. Moreover, in order to show this,

1. one needs to exploit all the conditions described by $-I;
2. one needs to make full use [of] $-E; but
3. one may not make any use of $-I.

(W) $(A, B) is the weakest major premise possible under the conditions described by $-E. Moreover, in order to show this,

1. one needs to exploit all the conditions described by $-E;
2. one needs to make full use [of] $-I; but
3. one may not make any use of $-E. (Tennant forthcoming, p. 25, with some minor notational adjustments)

When harmony obtains between $-I and $-E we may write *h*($-I, $-E). Tennant states the motivating idea behind his formulation as follows.

> Introduction and elimination rules for a logical operator $ must be formulated so that a sentence with $ dominant expresses the strongest proposition which can be inferred from the stated premises when the conditions for $-introduction are satisfied; while it expresses the weakest proposition possible under the conditions described by $-elimination (ibid., p. 19, again with notational adjustments).

However, as it stands the account is incomplete. harmony is no more than a necessary condition for stability; given a permissible introduction rule it fails to fix a uniquely matching harmonious elimination rule and vice versa.[6] The requisite *addendum* takes the form of the requirement of Harmony (the capital ‘H’ here is significant). Tennant states it as follows:

> Given $-E we determine $-I as the *strongest* introduction rule $-i such that *h*($-i, $-E). Given $-I we determine $-E as the *strongest* elimination rule $-e such that *h*($-I, $-e) (Tennant 2007, p. 22, with slight notational adjustments).

Thus to establish the Harmony of a pair of rules $-I and $-E we begin by holding a particular rule $-I fixed and then run through all of the $-e rules such that *h*($-I, $-e). Among these we choose the strongest, $-E. We now repeat the same

[6] Tennant mentions that this was pointed out to him by Peter Schroeder-Heister (1987, p. 94).

procedure, this time taking \$-E as our point of departure and arrive at \$-I′. If \$-I′ and the original \$-I are identical we have established that $H(\$\text{-I}, \$\text{-E})$. If \$-I′ is stronger we repeat the process until we reach a new equilibrium.[7]

3 A counterexample to Tennant's account

So far, so good. Tennant's notion of Harmony appears to be a good candidate for rendering the intuitive notion of stability precise. As desired, Harmony seeks to determine a unique ideal choice, 'a kind of Nash equilibrium between introduction and elimination rules' (Tennant forthcoming, p. 22). However, as I now show, it fails to account for our quantifier rules. The trouble is that Tennant's demand that we choose the *strongest* possible harmonious counterpart for any given rule obliterates crucial restrictions on our quantifier rules. Consider the introduction rule for ∃:

$$\exists\text{-I}\;\dfrac{\begin{matrix}\Gamma\\ \vdots\\ A[t/x]\end{matrix}}{\exists x A(x)}$$

As is standard, $A[t/x]$ is the result of substituting t for all free occurrences of x in A, and we require x to be freely substitutable for t in $A[t/x]$. The corresponding elimination rule is:

$$\exists\text{-E},\, i\;\dfrac{\begin{matrix}\Gamma\\ \vdots\\ \exists x A(x)\end{matrix}\qquad\begin{matrix}\Gamma', [A[a/x]]^i\\ \vdots\\ C\end{matrix}}{C}$$

where the parametric a may not occur in Γ', $\exists x A(x)$ or C.

Let us begin by showing that these rules are harmonious in Tennant's sense, i.e. $h(\exists\text{-I}, \exists\text{-E})$. First we show that (S) is satisfied. To this end, suppose there is a proposition X which is, for any term t, entailed by $A[t/x]$. In other words, suppose the following rule (∘) holds:

[7]Tennant's notion of Harmony presupposes that our grasp of what is to count as a legitimate candidate to be an inference rule of the type required (an introduction or an elimination rule) is sufficiently clear for the talk of 'running through' the set of such rules to make sense. For present purposes let us grant that the collection is sufficiently determinate (or at least that it can be made so).

$$\circ\,\frac{\begin{array}{c}\Gamma\\ \vdots\\ A[t/x]\end{array}}{X}$$

We now want to show that $\exists x A(x) \vdash X$, making full use of the elimination rule $\exists$-E and none of the introduction rule $\exists$-I. *Ex hypothesi*, X is implied by $A[t/x]$ *for any* t via $\circ$, therefore in particular for any t satisfying the constraints imposed by $\exists$-E. Consequently, X follows from any arbitrary parametric a.[8] We thus obtain the desired result:

$$\exists\text{-E, }1\,\frac{\exists x A(x) \qquad \circ\,\dfrac{[A[a/x]]^1}{X}}{X}$$

Turning now to (W), assume that X features as the major premise of $\exists$-E and thus entails every proposition that follows from $A[a/x]$ where a is again parametric and the usual conditions apply. Our assumption amounts to the following rule ($\bullet$):

$$\bullet,\,i\,\frac{X \qquad \begin{array}{c}\Gamma, [A[a/x]]^i\\ \vdots\\ C\end{array}}{C}$$

With its help we obtain $X \vdash \exists x A(x)$:

$$\bullet,\,1\,\frac{X \qquad \exists\text{-I}\,\dfrac{[A[a/x]]^1}{\exists x A(x)}}{\exists x A(x)}$$

This establishes that our standard rules for the existential quantifier are indeed $\underline{\text{h}}$armonious.

But is $\exists$-E also the *strongest* elimination rule $\underline{\text{h}}$armonious with $\exists$-I? That is, is $\exists$-E $\underline{\text{H}}$armonious with $\exists$-I? Herein lies the problem. For consider the elimination rule $\exists$-E$'$, identical to the regular elimination rule for the existential quantifier with the exception that it lacks some or all of the restrictions on the parameter a. An inspection of the above demonstration of $\underline{\text{h}}$armony reveals that the same

[8] It is not quite clear how the restriction that a must not occur in X is to be understood given that X is a proposition in Tennant's sense, i.e. a logical equivalence class of sentences which, as such, has no syntactic form. Where such restrictions are operative we may restrict our attention to the subclass of sentences in the equivalence class that satisfy them. We know that this subclass is non-empty in this case because '$\exists x A(x)$' is a member of the class.

demonstration goes through for any of the more permissive ∃-E′. Hence h(∃-I, ∃-E′) holds also. But now let ∃-E′ be the elimination rule devoid of any restrictions. Surely *that* rule is the strongest such elimination rule as it enables us to 'extract' the most from sentences containing ∃ in a dominant position. In particular, of course, it is 'stronger' than the ordinary ∃-elimination rule. It follows that it is the wholly unrestricted elimination rule that is in Harmony with the introduction rule given above, i.e. we have H(∃-I, ∃-E′).

But do we really get H(∃-I, ∃-E′)? Or must we in turn revise our introduction rule? Only if we can also show that ∃-I is the strongest introduction rule harmonious with ∃-E′ will H(∃-I, ∃-E′) obtain and our conclusion pass muster. Again we might be tempted to try out a modified introduction rule devoid of any constraints. However, the restrictions in the standard elimination rule play a rather different role from the restrictions we impose on introductions of the existential quantifier. In the former case, the restrictions—the term a may not occur in any of the undischarged hypotheses, nor in the major premise, nor in the conclusion—are specifically tailored to capture the idea that a is arbitrary. This constraint is beholden ultimately only to the intended meaning of ∃. Or, to put it in a way more germane to the spirit of the inferentialist project, the restrictions are partially determinative of the meaning of the existential quantifier. By contrast, the restriction we impose on the standard ∃-introduction rule carries no such semantic weight; it serves sole the purpose of ensuring that every application of the rule will result in a well-formed formula, i.e. that non-well-formed formulas like the following are prevented.

$$\exists\text{-I}\ \frac{\forall x R(x,a)}{\exists x \forall x R(x,x)}$$

Therefore, it seems that ∃-I really *is* the strongest introduction rule that matches ∃-E′. And we get H(∃-I, ∃-E′) after all.[9]

Something has gone very wrong here. Clearly the quantifier whose meaning we have so fixed—call it 'Ⅎ'—bears little resemblance to the existential quantifier that we have come to know and love. Moreover, it is clearly incoherent as the following derivation makes plain.[10]

$$Ⅎ\text{-E, 1}\ \frac{Ⅎ\text{-I}\ \dfrac{F(a)}{Ⅎ x F(x)} \qquad [F(b)]^1}{F(b)}$$

[9] An analogous demonstration of the insufficiency of Tennant's account could be given for the universal quantifier.

[10] The following proof violates the standard restriction imposed upon the ∃-E rule that the parameter a ought not to occur in the conclusion of the minor premise.

The introduction of ⅃ induces non-conservativeness that spills over into the non-logical regions of language: it makes available new information about the world—indeed a lot of new information! For by appeal to ⅃, we can justify the assertion of an atomic sentence '$F(b)$' for any individual b so long as the predicate 'F' is correctly assertible of at least one individual. Yet it is ⅃ that comes out Harmonious. We thus find ourselves with a clear-cut case of E-strong disharmony, showing that our intuitive principle of stability has been violated.[11] Tennant's account, at least in its current form, thus fails to capture the notion of stability.[12]

[11] Clearly, therefore, the resulting system is not normalizable either. The detour resulting from the introduction and immediate subsequent elimination of ⅃ is ineliminable since there cannot be any direct deductive path from '$F(a)$' to '$F(b)$'.

[12] In writing this paper I benefited from helpful comments from Arif Ahmed, Yoon Choi and Michael Potter. Needless to say, only I am to blame for remaining errors. Also I want to thank Neil Tennant for generously helping me to understand his position better. Research for this paper was funded by the AHRC and the Gates Cambridge Trust whose support I gratefully acknowledge.

References

[1] Belnap, N. (1962) 'Tonk, plonk and plink', *Analysis*, 23:130–134.

[2] Dummett, M. A. E. (1973) *Frege: philosophy of language*, Harvard University Press, Cambridge MA.

[3] Dummett, M. A. E. (1991) *The logical basis of metaphysics*, Harvard University Press, Cambridge MA.

[4] Tennant, N. (1978) *Natural logic*, Edinburgh University Press, Edinburgh.

[5] Tennant, N. (1987) *Anti-realism and logic*, Oxford University Press, Oxford.

[6] Tennant, N. (1997) *The taming of the true*, Oxford University Press, Oxford.

[7] Tennant, N. (forthcoming) 'Inferentialism, logicism, harmony, and a counterpoint', to appear in Miller, A. (ed.) *Essays for Crispin Wright: language, logic and mathematics*, in preparation at Oxford University Press.