Comments on Peter Vranas's New Foundations for Imperative Logic II: Pure Imperative Inference

Reza Mahmoodshahi

June 4, 2007

Peter's principal result is an equivalence theorem. In his words, "it is a consequence of this theorem that the validity of a pure imperative argument, which I defined in terms of the intuitively appealing but initially nebulous relation of support between reasons and prescriptions, can be captured in terms of the clear and precise relation of entailment between propositions."

I want to speak to what this result does and does not show. It does show that one can *test for* "strong" and "weak" pure imperative validity by checking whether or not certain propositions stand in a relation of entailment to one another. It does not, however, succeed as a *reduction* of "strong" and "weak" pure imperative validity to a relation of entailment between propositions - it does not succeed, because the proof of the theorem appeals (in the direction from strong or weak validity to the corresponding entailment conditions on propositions) to the irreflexivity and transitivity of a so-called *relation of comparative support* between "reasons" and prescriptions.

That there is a need to appeal to the properties of such a relation in the course of the proof is not surprising given the paper's definitions. A pure imperative argument is defined as *strongly valid* just in case, necessarily, every reason that strongly supports the conjunction of the premises of the argument also strongly supports the conclusion of the argument. By definition, a reason *strongly supports* a prescription exactly if (i) [dominance condition] it favors every proposition which entails the satisfaction proposition of the prescription over every different proposition which entails the violation proposition of the prescription and (ii) [satisfaction indifference condition] it does not favor any proposition which entails the satisfaction proposition of the prescription over any other such possible proposition (10).

Strong support is thus a relation that holds between "reasons" and propositions. On just what a "reason" is, following Broome, a reason is identified with something that is a consideration for something else. That is, a reason is something that *counts in favor of* something else. So it's a two-place relation. As to the first component, following Raz, the considerations in question are taken to be facts rather than e.g. propositions, although it's claimed that "nothing of substance...hangs on this choice." As to the second component, the "something else" that's favored is assumed to be a proposition. So, in short, a relation of comparative support is a relation that reasons qua "facts" bear to propositions.

I come now to my first point. Given any pair $< R, X >$ where R is a reason and X is a proposition, I submit that the question of whether R counts in favor of (or against) X is not a question of logic. Rather, I take it to be a question of agency and as such a question of preference. Take the example from the paper connecting the fact that I have promised to help you with the proposition that I help you. The former fact is a reason that counts in favor of the latter only if "relative to the fact that I have promised to help you it is better if I do anything which entails that I dont break my promise than if I do anything else which entails that I break it." This assumption, that one state-of-affairs is *better* than another, encodes a preference, viz., a general preference, on the part of the promiser, to fulfill rather than break promises. A preference to break rather than fulfill promises, though perverse, does not violate any principle of logic. This illustrates, quite clearly I think, that the so-called "relation of comparative support" is not a logical one.

This is not a point that necessarily spells trouble for the account of imperative inference under consideration. In fact, the enterprise of adducing a valid pure imperative argument is motivated on grounds of "convincing people that they should *act* according to its conclusion." That is, to a first-approximation, "if a pure imperative argument is valid and one should act according to its premises, then one should (also) act according to its conclusion." While matters are somewhat more complicated, the motivation is on-target insofar as it highlights the fact that an imperative "logic" is inextricably bound up with matters of rational or prudential agency.

The question thus arises: given that the so-called "relation of comparative support" between reasons and propositions is not a purely logical one, how might it be developed so that valid pure imperative arguments transmit something akin to prudential warrant? I believe the most natural development is to have it encode utility judgement. This should not be confused with judgments of expected utility.

To this end, I will need to interpret facts as propositions. While this is a departure from the paper, I believe it's one that's amenable to the author. For my purposes, then, a reason is a proposition that counts in favor of another proposition.

Let R be a proposition *qua* reason and U some agent's utility function. Let's us say that R *counts in favor of X* just in case

$$U(R \wedge X) > U(R \wedge \neg X)$$

Let's further say R *counts against X* just in case

$$U(R \wedge X) < U(R \wedge \neg X)$$

Finally, for logically distinct X and Y, we say R *comparatively favors X over Y* just in case

$$U(R \wedge X) > U(R \wedge Y)$$

More generally, let's say that X is in the *comparison class* of R iff either R counts in favor or against X. So, if X is not in the comparison class of R, it must be the case that $U(R \wedge X) = U(R \wedge \neg X)$. The plausibility of these definitions is transparent. The basic idea is to make the relation of comparative support analogous to the relationship between evidence and hypothesis in confirmation theory. We say e confirms h just in case $Pr(h|e) > Pr(h)$, e disconfirms h just in case $Pr(h|e) < Pr(h)$, and e is independent of h whenever $Pr(h|e) = Pr(h)$. I further make the crucial assumption that utilities are *not* additively separable. For if they were, the conditions above would be trivialized.[1] Moreover, it's worth noting that these conditions collapse the author's distinction between comparative and non-comparative support. On this account, all support is comparative support.

Using the above definitions of support, it's straightforward to reformulate the author's condition of dominance and satisfaction indifference, which jointly are necessary and sufficient for strong support. R strongly supports the satisfaction proposition of a prescription (and hence the prescription) iff (i) [dominance condition] for every X s.t. X entails S and every Y s.t. Y entails $\neg S$, $U(R \wedge X) > U(R \wedge Y)$ and (ii) [satisfaction indifference] for every X, Z s.t. U(X) > U(Z), X entails S, and Z entails S: $U(R \wedge X) = U(R \wedge Z) + Q$ where Q = U(X) - U(Z). I submit that this framework is expressive enough to handle virtually all judgments of comparative support, but I show next why it will not do for the purposes of a logic of imperative inference.

In *The Logic of Decision* R. Jeffrey, using a result due to Ethan Bolker, shows that if the pair of functions $< Pr, U >$ defined over an algebra of propositions obey the well known axioms of the probability calculus as well as a desirability axiom (see Ch. 5), then there exists an expected utility representation of an agent's preferences s.t. for any propositions X, Y: X$\succ$Y iff EU(X) > EU(Y). The trouble with Jeffrey's representation result is the failure of uniqueness up to positive affine transformations. On Jeffrey's account, if $< Pr, U >$ is a pair

[1] I add that this assumption is relatively innocuous on Jeffrey's decision theory, which is unique up to bilinear transformations. Additive separability is only preserved up to linear transformations.

of functions reflecting the agent's preferences among propositions, then so does any pair $< Pr^*, U^* >$ related to the original functions by what's known as the Bolker transformations:

$$U^*(A) = \frac{U(A)}{cV(A) + 1} \tag{1}$$

$$Pr^*(A) = (cU(A) + 1)Pr(A) \tag{2}$$

where c is a real number s.t. cU(A) > -1 whenever Pr(A) > 0.[2] Where U is bounded above (at max) and below (at min), the possible values of c are those allowed by the inequalities:

$$\frac{-1}{max} < c < \frac{-1}{min} \tag{3}$$

The upper bound on indeterminacy due the failure of uniqueness on Jeffrey's account can now be made precise. John Collins (1996) and Joyce discuss indeterminacy for Pr. I am interested in indeterminacy for U. It can be shown that given any utility function U, U(A) can differ from its Bolker equivalent transformation $U^*(A)$ by as much as

$$\frac{-1}{max} + \frac{1}{min} \tag{4}$$

This makes for a quite a bit of indeterminacy. In fact, the Bolker transformation does not (in general) preserve the truth of statements of the form "U(A) > U(B)" or "U(A)= 3U(B)". E.g. it might be the case that $U(A) > U(B)$ while $U^*(A) < U^*(B)$. It should now be clear why this type of indeterminacy is problematic for the proposed logic of imperative inference. Suppose we discover that an agent's degree of belief and utility function can be represented by the pair $< Pr, U >$. It follows, as a matter of mathematical fact, that this pair can be represented by any Bolker equivalent pair $< Pr^*, U^* >$. But the Bolker equivalent pair does not respect the relation of comparative support. We can therefore conclude the following. Given that an agent's degree of belief and utility function can be represented by the pair $< Pr, U >$, it's possible for the same pure imperative argument to be both valid and invalid in both the strong and weak sense. This type of result, I take it, is a surprising one given the prima facie plausibility of a preference-based development of the relation of comparative support.

What can we conclude from all this? There are, in my view, two important points of disanalogy between the standard logic of declarative inference and the proposed logic of pure imperative inference. The first point of disanalogy is not surprising given Peter's definitions and my development of the relation of comparative support. While declarative validity holds across persons, pure imperative validity does not in general. What's valid for me might not be valid for

[2] I've simplified the transformation functions by fixing the unit and zero point and calibrating preferences s.t. U(Tautology)=0.

you. The second point of disanalogy follows from the result I've argued for and strikes me as far more surprising. While a third party is always in the position to adduce the validity of a declarative argument, the problem of indeterminacy entails that under certain conditions it is impossible, as a matter of principle, for a third party to adduce the validity of a pure imperative argument. Admittedly, this second point of disanalogy is *epistemological*, but I think this new dimension to inference is precisely what makes the enterprise of pure imperative validity novel and worthy of further investigation.