Article:
Elstein, D (2007) Against Sonderholm: still committed to expressivism. Proceedings of the Aristotelian Society, 107 (1). 111 - 116. ISSN: 0066-7374

https://doi.org/10.1111/j.1467-9264.2007.00213.x

AGAINST SONDERHOLM: STILL COMMITTED TO EXPRESSIVISM

by DANIEL ELSTEIN

This is the accepted version of the following article: Elstein, D. (2007), Against Sonderholm: Still Committed to Expressivism. Proceedings of the Aristotelian Society (Hardback), 107: 111–116, which has been published in final form at http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9264.2007.00213.x/abstract

ABSTRACT: Jorn Sonderholm (2005) has argued that Simon Blackburn's commitment semantics for evaluative discourse is unable to explain the validity of simple inferences involving disjunction. This is true insofar as the basic rules which Blackburn suggests are not strong enough, but it is relatively simple to augment those rules so as to meet Sonderholm's challenge, whilst respecting the spirit of commitment semantics. One way of doing this is to add a reduction rule such that if accepting p commits one to inconsistent commitments, one is committed to accepting $\neg p$. Thus Sonderholm has not provided any reason to doubt the adequacy of commitment semantics to explain validity in evaluative discourse.

Jorn Sonderholm (2005) presents a difficulty for Simon Blackburn's commitment semantics for evaluative discourse. Blackburn aims to provide an explanation of our acceptance of various inference patterns that does not require validity to be defined in terms of truth-preservation. Such an account appears necessary for defending expressivism, the view that moral utterances express attitudes. Accordingly he characterises valid inferences as ones where accepting the premises and rejecting the conclusion results in inconsistent commitments. Sonderholm points out that as it stands Blackburn's account is unable to validate certain simple arguments involving disjunction, the most basic of which is the inference:

$$(p \vee p) \vdash p$$

Blackburn interprets the disjunction p v q as two conditional commitments: if ¬p is accepted, accept q, and if ¬q is accepted, accept p. This is symbolised as $[A(\neg p) \rightarrow A(q)] \,\&\, [A(\neg q) \rightarrow A(p)]$. Sonderholm explains that Blackburn cannot deduce an inconsistent set of commitments from accepting the premise of the inference in question whilst not accepting its conclusion. An attempted proof will proceed as follows, stopping before an inconsistency is reached:

1	(1) $[A(\neg p) \rightarrow A(p)] \,\&\, [A(\neg p) \rightarrow A(p)]$	1 ASS
2	(2) $\neg A(p)$	2 ASS
1	(3) $A(\neg p) \rightarrow A(p)$	1 &E1
1,2	(4) $\neg A(\neg p)$	2,3 MTT

Sonderholm rightly says that this argument could be completed if we were allowed a rule (K) such that $\neg A(\neg A) \Rightarrow A(A)$, and that this rule is unacceptable: refusing to reject p is not the same as accepting p. There is, however, a different rule that would allow us to complete the argument: a reduction rule which says that if one is conditionally committed to absurdity on accepting A, then one is committed to accepting $\neg A$. We can symbolise this as $[A(A) \vdash A(\bot)] \Rightarrow A(\neg A)$. We want to generalise this rule to cases where it is only on certain assumptions that accepting A commits one to accepting absurdity, in which case it is only given those assumptions that one is committed to accepting $\neg A$. Thus the full reduction rule is:

(R) $\phi, [\phi, A(A) \vdash A(\bot)] \Rightarrow A(\neg A)$[1]

This licenses running the argument as follows, now deducing A(p) directly, rather than showing that the premiss is inconsistent with $\neg A(p)$:

1	(1) $[A(\neg p) \rightarrow A(p)] \,\&\, [A(\neg p) \rightarrow A(p)]$	1 ASS
1	(2) $A(\neg p) \rightarrow A(p)$	1 &E1
3	(3) $A(\neg p)$	SUPP
1,3	(4) $A(p)$	2,3 MP
1,3	(5) $A(\bot)$	3,4 ⊥

[1] Here and throughout ‘ϕ’ is a schematic letter for any wff, whereas ‘A’ is a schematic letter for any string (not containing any occurrences of the A(…) acceptance operator) such that ‘A(A)’ is a wff.

1	(6) A(¬¬p)	2-5 R
1	(7) A(p)	6 N[2]

The strategy also works for Sonderholm's less simple case: the inference [(p&q) v (p&r)] ├ p:

1	(1) [A¬(p&q) → A(p&r)] & [A¬(p&r) → A(p&q)][3]	1 ASS
1	(2) A¬(p&q) → A(p&r)	1 &E1
3	(3) A(¬p)	SUPP
3	(4) A¬(p&q)	3 J[4]
3	(5) A¬(p&r)	3 J
1,3	(6) A(p&r)	2,4 MP
1,3	(7) A(⊥)	5,6 ⊥
1	(8) A(¬¬p)	2-7 R
1	(9) A(p)	8 N

It is worth mentioning four possible concerns. The first is that rule R and rule K are equivalent; if this were so then my reply would be in no better shape than the one which Sonderholm rightly rejects. But the rules are not equivalent: R is weaker than K. There are circumstances where one does not accept ¬p, but accepting ¬p would not involve absurdity. According to rule K one is committed to accepting p, but rule R does not apply.

The next worry is that I play fast and loose with absurdity: there is a difference between the internal contradiction involved in the combination A(p) and A(¬p), and the external contradiction in A(p) and ¬A(p). I concede the point, and that is why I use the expression 'A(⊥)' rather than '⊥' when internal contradiction is in play, but if my

[2] N is the rule that A(¬¬A) ⇒ A(A). We are entitled to assume a commitment to classicism.

[3] Note that there is something dubious about Sonderholm's translation here, because we should be unwilling to allow 'A¬(p&q)' as a wff, since it clearly means the same as 'A(¬p v ¬q)', which Blackburn disallows. It would be better to paraphrase the former in the same way as the latter. In the text we give Sonderholm the benefit of the doubt in framing his alleged counterexample.

[4] J is the rule that A(¬A) ⇒ A¬(A&B). This rule is eliminable, because both A(¬A) ⟶ [A(A) ⟶ A(¬B)] and A(¬A) ⟶ [A(B) ⟶ A(¬A)] are theorems, and so A(¬A) ⟶ {[A(A) ⟶ A(¬B)] & [A(B) ⟶ A(¬A)]} is also a theorem. The latter is equivalent to A(¬A) ⟶ A¬(A&B), so J is a shorthand, rather than a substantive addition to the system.

symbolism offends, feel free to substitute a different symbol. The crucial point is that we have a standing commitment to avoiding contradiction, and this is reflected in our use of reductio reasoning. Anyone who accepts reductio reasoning must accept rule R, since what rule R records is simply our commitment to accepting the results of reductio reasoning. Those points are the ones on which my reply stands, and they are unaffected by quibbles about what 'absurdity' means.

A third concern is that rule R commits us to accepting a contradiction.[5] If we let L be the liar sentence ('This sentence is false'), then both we get both $A(L) \rightarrow A(\bot)$ and $A(\neg L) \rightarrow A(\bot)$ as theorems. But then R allows to deduce both $A(\neg L)$ and $A(L)$, so it turns out that $A(\bot)$ is a theorem! But it is if anything an advantage for commitment semantics that it takes the Liar to be paradoxical. It is not as if standard reduction rules can cope consistently with the Liar; truth-conditional semantics is thus a companion in guilt. So this worry is only worth taking seriously if the expressivist's opponent has her own solution to the Liar, which cannot be adapted to commitment semantics. If commitment semantics has the same problem with the Liar that everyone else does, that counts as a success (albeit an odd one) for expressivists in stealing the clothes of realists.

The most serious worry is that my reply has the advantage of theft over honest toil. After all, the point of Blackburn's project is to explain why we are inclined to accept certain inferences. But I am just taking it for granted that reductio inferences are acceptable. Clearly this would not be acceptable as a general strategy: whenever a valid argument is presented that commitment semantics in its present form cannot deal with, invent a rule licensing the argument. On the other hand, the expressivist must be allowed some materials to work with. Blackburn takes as basic rules that (are tailored to) validate modus ponens and modus tollens. Sonderholm has demonstrated that Blackburn's rules are insufficient for all the arguments we want to validate. The

[5] I thank an anonymous referee for bringing this problem to my attention.

following rules seem sufficient[6]:

(→I)	$\theta, [\theta, \phi \vdash \psi] \Rightarrow [\phi \rightarrow \psi]$
(MP)	$[\phi \rightarrow \psi], \phi \Rightarrow \psi$
(&I)	$\phi, \psi \Rightarrow [\phi \,\&\, \psi]$
(&E1)	$[\phi \,\&\, \psi] \Rightarrow \phi$
(&E2)	$[\phi \,\&\, \psi] \Rightarrow \psi$
(¬I)	$\phi, [\phi, \psi \vdash \bot] \Rightarrow \neg\psi$
(¬E)	$\neg\neg\phi \Rightarrow \phi$[7]
(R)	$\phi, [\phi, A(A) \vdash A(\bot)] \Rightarrow A(\neg A)$
(N)	$A(\neg\neg A) \Rightarrow A(A)$

The need for R and N follows from the obvious point that once there is a distinction between internal and external negation, there need to be rules for internal negation introduction and elimination, and R and N are natural parallels of the classical rules for negation. Ideally we would like a meta-proof of the conservativeness of these rules with respect to classical logic.[8] Here is a sketch of such a proof:

The rules above are conservative just in case we can derive either an external contradiction (A(p) and ¬A(p)) or an internal contradiction[9] (A(p) and A(¬p)) from a set of wffs iff we can derive a contradiction from the correct translation of that set in the propositional calculus. The central difference between commitment semantics and the propositional calculus is that when the latter has a negated wff e.g. '¬p', the correct translation in commitment semantics may be either '¬A(p)' or 'A(¬p)', depending on

[6] Assuming that we do not have internal conjunction – see footnotes 3 and 4.
[7] Note that this rule covers ¬¬A(A) ├ A(A), since the acceptable substitutions for 'ϕ' are a superset of the acceptable substitutions for 'A(A)' (see note 1). If this sub-rule seems controversial, consider that all it amounts to is that non-non-acceptance commits one to acceptance. If we translate sentences of English involving non-non-acceptance by '¬¬A(A)', and ones involving acceptance by 'A(A)', then the sub-rule is evidently sound.
[8] I do not mean that the expressivist is committed to providing such a proof.
[9] If the contradiction is internal then taking the internal negation of a conclusion as a supposition and applying R and N will only allow us to derive conclusions of the form 'A(A)', so there may appear to be a breakdown of *ex falso quodlibet*. But it will be legitimate to assume that the only atomic sentences are of that form too, so in fact *ex falso quodlibet* will hold.

what sentence of natural language '¬p' is a translation of. Whichever translation is correct, it must be applied uniformly to all occurrences of '¬p', since otherwise we are admitting that there is an equivocation in the argument as rendered in the propositional calculus. When external negation is the correct translation, there is no problem, because then R and N can be ignored and the other rules above are a complete classical system when the possibility of internal negation is off the table. When internal negation is used, we take advantage of the fact that, in the propositional calculus, for any wff in which negation has broad scope with respect to some other connective it is possible to give a logically equivalent wff where negation takes narrow scope. This can be done even when we restrict ourselves to the connectives '&', '⊃', and '¬'. And this suffices to show that the rules above excluding ¬I and ¬E are a complete classical system, ignoring the possibility of external negation. Once we put the wffs of the propositional calculus into the form of narrow-scope negation, proofs of contradiction will even proceed isomorphically to proofs of internal contradiction in commitment semantics. So however negation is translated, the conservativeness condition is satisfied. Whilst it is true that commitment semantics allows for wffs with both internal and external negation, this just shows the greater expressive power of commitment semantics, and such cases are irrelevant to whether it is conservative.

The simplicity in the set of rules required should allow my proposal to count as a reply in the spirit of Blackburn's project. That project is to reconstruct the validity of all valid inferences from the acceptance of a set of rules designed to validate a small set of basic inferences. Reductio arguments can legitimately be seen as part of this basic set, which gives expressivists a right to rule R. It would be unfair to Blackburn to say that he is not allowed a rule of internal negation introduction. R is the most obvious candidate for such a rule, and it meets Sonderholm's objection. Thus, *pace* Sonderholm, quasi-realism has not yet broken its promise to make expressivism non-revisionist.[10]

[10] Thanks to Simon Blackburn, Hallvard Lillehammer, Neil Sinclair, Jorn Sonderholm and an anonymous

Reference

Sonderholm, Jorn. 2005. 'Why an Expressivist should not Commit to Commitment-Semantics', *Proceedings of the Aristotelian Society*, CV: 403-409.

referee for helpful criticisms and suggestions.