Stalnaker's theory of belief ascriptions

PHIL 93515
Jeff Speaks

January 16, 2007

Possible worlds semantics identifies propositions with functions from worlds to truth-values or, equivalently, with sets of possible worlds. Since the proposition expressed by a sentence is the set of possible worlds with respect to which it is true, sentences true with respect to just the same possible worlds must express the same proposition.

To see the problems that this causes with propositional attitiude ascriptions, note that any proposition is necessarily equivalent to the conjunction of itself and any of its necessary consequences. Hence, if Q is among the necessary consequences of P, it follows that

$$\Box\ (a \text{ believes } P \equiv a \text{ believes } P \ \&\ Q)$$

from which it follows, given the distribution of belief over conjunction, that

$$\Box\ (a \text{ believes } P \rightarrow a \text{ believes } Q)$$

So belief is closed under necessary consequence: if one believes p, then one also believes all of p's necessary consequences. From this two particularly damaging consequences follow: (a) No one believes any necessary falsehoods since, all propositions being necessary consequences of a necessary falsehood, if one believed a necessary falsehood one would thereby believe every proposition; and no one believes every proposition. (b) Everyone who has any beliefs at all believes every necessarily true proposition, since all necessary propositions are necessary consequences of every other proposition. From (a) it follows that, for example, no one has ever held a false mathematical belief or believed that water is not H_2O; from (b) it follows that every creature with any beliefs believes that arithmetic is incomplete, and that water is H_2O. These conclusions seem clearly to be incorrect.

Stalnaker's central response to this problem is to deny the naive relational theory of attitude ascriptions: the view that an ascription $\ulcorner \alpha$ believes that $\sigma \urcorner$ is true just in case the referent of the value of 'α' bears the belief relation to the semantic content of the value of 'σ' (in the context of the ascription). Instead, Stalnaker thinks, such ascriptions sometimes report a relation to a meta-linguistic proposition about the truth of the sentence in the complement clause of the ascription. Because this proposition will always be contingent, and the possible worlds account of the objects of belief runs into trouble precisely with necessarily true and necessarily false propositions, this meta-linguistic reinterpretation promises to deliver a more intuitive assignment of truth-conditions to attitude ascriptions than the unmodified possible worlds theory.[1]

[1] Note that Stalnaker does not deny that, for example, anyone who has any beliefs at all bears the belief relation to the (one and only) necessary proposition, expressed by, among many other sentences, "Arithmetic is incomplete"; what he denies is that, in all such cases, an ascription $\ulcorner \alpha$ believes that arithmetic is incomplete$\urcorner$ will be true.

The main problem with this analysis is not so much that it is implausible as that it does very little to palliate the counter-intuitive consequences of Stalnaker's theory. Consider the sentence, "No whole number raised to a power greater than two is equal to the sum of two other whole numbers, each raised to that power." This is an example of a sentence which poses problems for the view of the objects of belief as sets of possible worlds, because (i) since it expresses a necessary proposition, it follows from the closure of belief under necessary consequence that any agent who has any beliefs at all believes what it says, and yet (ii) there is no difficulty in finding an example of an agent A such that the sentence

[1] A believes that no whole number raised to a power greater than two is equal to the sum of two other whole numbers, each raised to that power.

seems clearly false. Intuitively, many agents have beliefs without believing Fermat's last theorem. The meta-linguistic strategy is designed to block our having to treat [1] as true in these cases by interpreting it as attributing to A, not belief in the necessary proposition expressed by

[2] No whole number raised to a power greater than two is equal to the sum of two other whole numbers, each raised to that power.

but rather belief in the contingent meta-linguistic proposition expressed by the sentence

[3] "No whole number raised to a power greater than two is equal to the sum of two other whole numbers, each raised to that power" is true.

Since the proposition expressed by [3] is contingent, closure under necessary consequence doesn't entail that A believes it; hence Stalnaker's semantics for belief ascriptions seems to make room for the wanted result that [1] is not true.

A problem with this strategy of systematically reinterpreting attitude ascriptions is that agents often in the problematic cases have beliefs about the meaning of the relevant sentence. For example, we can suppose that the agent of [1] believes the proposition expressed by [4]:

[4] "No whole number raised to a power greater than two is equal to the sum of two other whole numbers, each raised to that power" means that no whole number raised to a power greater than two is equal to the sum of two other whole numbers, each raised to that power.

The problem is that, given that Fermat's Last Theorem is a necessary truth, [4] has [3] as a necessary consequence. So, by the closure of belief under necessary consequence, if Stalnaker grants that the agent believes what is expressed by [4], he must grant that the agent believes what is expressed by [3]. But, if the agent believes what is expressed by [3], [1] is true on the meta-linguistic reinterpretation of the ascription — which is the result that the reinterpretation was introduced to avoid.