
how to be an expressivist about truth

mark schroeder – expressivism, pluralism, and representationalism, sydney – july 23, 2009

0 disclaimer

- Three goals for today:
- (1) shed some light on why expressivism might be a fruitful framework for theorizing about truth.
 - (2) show how a semantic theory can be explanatory without being inflationary.
 - (3) introduce a couple of ideas I find helpful: *commitment theory* and a *divorce* of two theoretical roles associated with propositions.
- Not on the agenda:
- (1) Defending expressivism – I am not an expressivist.
 - (2) Defending/endorsing a theory of truth.

I.1 intersubstitutability and the unrestricted T-schema

Conditional commitment: someone who thinks that S means that P is committed to having the same attitudes toward the proposition that S is true and the proposition that P.

T-schema If S means that P, then S is true just in case P.

Liar Liar is not true.

T-Liar If Liar means that Liar is not true, then Liar is true just in case Liar is not true.

Key Question: do we really need to ‘make it explicit’? *Can* we really always make it explicit?

I.2 rejection and revenge

Rejection is the attitude to have toward paradoxical propositions – don’t accept or deny them. But... can we *lexicalize* rejection? Can we ‘make it explicit’ – say, with a special meaning for ‘not*’?

Liar’s Revenge Liar’s Revenge is not true or Liar’s Revenge is not* true.

I.3 two observations... and the link to expressivism

Observation I: Both intersubstitutability and the idea that paradoxical propositions are ones that it makes sense to reject are theses about the *rationality of mental states*, or their rational relationships to one another.

Relates to expressivism ...because expressivism is the semantic theory which works by associating each sentence, ‘P’, with the mental state that someone is in who thinks that P, and which accounts for logical relationships in terms of the rational relationships between those mental states.

Observation 2: Both the problem about intersubstitutability and the problem about rejection turn on the question of what we are able to *lexicalize* or ‘make explicit’, and how.

Relates to expressivism ...because expressivism is founded on the importance of the distinction between *expressing* and *reporting* mental states. And much contemporary work on expressivism has pointed toward the conclusion that not all states that it is possible to *report* can also be *expressed*.

I.4 commitment theory

The framework: three committed attitudes – acceptance, rejection, and denial

Assume: each *disagrees* with the others toward the same content

Define: you are *committed* to attitude C toward a content iff you disagree with each of the other attitudes toward that content.

P	~P
A	D
R	R
D	A

This is the commitment table for negation. It follows from three assumptions: (1) Accepting ~P is the same as denying P, (2) rejecting P is the same as rejecting ~P, and (3) each attitude disagrees with each of the others.

Very Important!!! Note that it is a *commitment* table, *not* a truth-table.

P	Q	P&Q
A	A	A
A	R	R
A	D	D
R	A	R
R	R	R
R	D	D
D	A	D
D	R	D
D	D	D

This is the commitment table for conjunction. We need more assumptions to derive it, but it is still highly plausible.

Very Important!!! These tables are the Strong Kleene tables, but we are *not* interpreting them as truth-tables. We are understanding them as articulating the *commitment* relationships between the attitudes of acceptance, rejection, and denial.

Key Facts: (1) All classical theorems of propositional logic are undeniable.
(2) *Modus ponens* with the material conditional preserves commitment – that is, it takes you from propositions that you accept only to propositions that you are committed to accepting.

P	MEANS(S,THAT(P))	TRUE(S)
A	A	A
R	A	R
D	A	D
A	R	A
R	R	R
D	R	D
A	D	
R	D	
D	D	

This is the commitment table that we should want for sentential truth – it articulates the idea of full intersubstitutability, by telling us that anyone who either accepts or rejects that S means that P is committed to having the same attitudes toward ‘P’ and toward ‘TRUE(S)’.

T-Schema MEANS(S,THAT(P)) $\supset$ (TRUE(S) $\equiv$ P)

Important consequence: T-Schema is undeniable, just like theorems of classical logic. But you don’t have to *accept* it – if you reject ‘P’, then you will reject T-Schema.

I.5 why not stop here?

Does commitment theory give us an *account of truth*, or a set of *goals* for an account of truth?

The pure inferentialist says: *an account of truth*

The proper expressivist says: *a set of goals*

In favor of the proper expressivist's answer:

- (1) We need *constraints* – otherwise why shouldn't we be able to lexicalize intersubstitutability and rejection?
 - (2) Theft vs. Honest Toil – Should deflationists be against *explanations*, or just against the *wrong kind* of explanations?
-

2.1 expressivist semantics

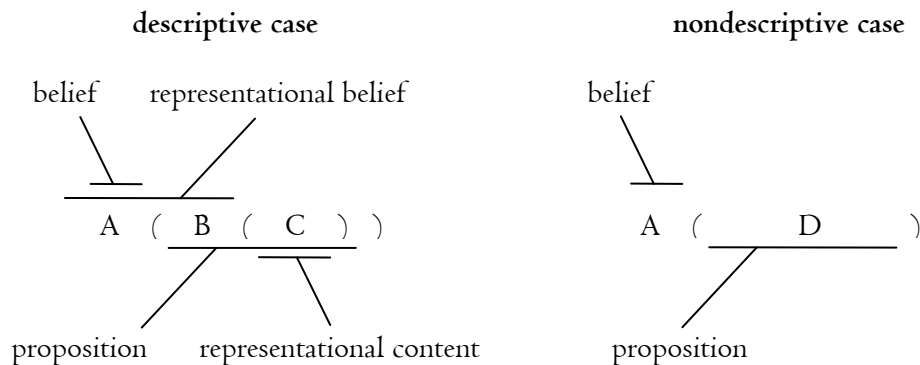
Expressivism as assertability-conditional semantics

- The goal:
- (1) Give a set of compositional rules which tells us, for each sentence, 'P', what it is to think that P.
 - (2) Use facts about the mental states expressed by sentences to predict and explain their commitment properties – as set out in commitment theory.
 - (3) Use commitment theory to explain logic and other semantic properties.

Divorcing two theoretical roles for propositions:

- (1) objects of the attitudes and bearers of truth and falsity: *propositions*
- (2) carving up reality; metaphysical commitments: *representational contents*

Representational belief can be thought of as a relation to *either kind* of object, depending on how we carve it up – just like being *about to go to Paris*:



2.2 biforcated attitudes

The attitude of *being for*

- i. takes properties for its contents
- ii. is *inconsistency-transmitting* – like belief and intention, two states of being for *disagree* just in case their contents are inconsistent

Propositions

defn. consists in a pair of properties, one of which is strictly stronger

To believe a proposition is to be for each of its properties – to have a 'biforcated attitude'

Example: $\langle \lambda z(\text{avoiding}(z, \text{stealing}) \wedge \text{disapproving}(z, \text{stealing})), \lambda z(\text{disapproving}(z, \text{stealing})) \rangle$

DESCRIPTIVE PROPOSITIONS (if we have time):

Proceeding as if

gloss: to proceed as if p is to take p as settled in deciding what to do

assumed: i. that proceeding as if takes representational contents for its objects

ii. that proceeding as if p and proceeding as if $\neg p$ are inconsistent

The descriptive proposition picking out the representational content p :

$\langle \lambda z(\text{pai}(z, p)), \lambda z(\neg \text{pai}(z, \neg p)) \rangle$

So... someone with an ordinary descriptive belief that p will be disposed to take it as settled that p in deciding what to do... not unintuitive.

2.3 connectives, rejection, and logic in BAS

If A is the proposition $\langle \alpha_1, \alpha_2 \rangle$ and B is the proposition $\langle \beta_1, \beta_2 \rangle$...

then $\langle \neg \alpha_2, \neg \alpha_1 \rangle$ and $\langle \alpha_1 \wedge \beta_1, \alpha_2 \wedge \beta_2 \rangle$ are propositions.

$\langle \neg \alpha_2, \neg \alpha_1 \rangle$ is the *negation* of A , and $\langle \alpha_1 \wedge \beta_1, \alpha_2 \wedge \beta_2 \rangle$ is the *conjunction* of A and B .

Semantics: ‘ $\sim A$ ’ expresses belief in the negation of the proposition that ‘ A ’ expresses belief in; ‘ $A \& B$ ’ expresses belief in the conjunction of the propositions that ‘ A ’ and ‘ B ’ express belief in.

Definitions: To *accept* $\langle \alpha_1, \alpha_2 \rangle$ is to be for each of α_1 and α_2 .

To *deny* $\langle \alpha_1, \alpha_2 \rangle$ is to be for each of $\neg \alpha_2$ and $\neg \alpha_1$.

To *reject* $\langle \alpha_1, \alpha_2 \rangle$ is to be for each of α_2 and $\neg \alpha_1$.

Important Fact #1: This allows us to predict the commitment tables for ‘ $\&$ ’ and ‘ $\sim$ ’.

Important Fact #2: This explains why rejection cannot be lexicalized.

2.4 how could propositions be pairs of properties?

Recall theoretical role of properties: objects of attitudes, bearers of truth and falsity, objects of assertion

THAT If ‘ P ’ is a well-formed formula, then ‘ $\text{THAT}(P)$ ’ is a referring term. Relative to each assignment of values to the unbound variables in ‘ P ’, it refers to the semantic value of ‘ P ’ relative to that assignment.

THINKS ‘ $\text{THINKS}(x, y)$ ’ is a well-formed formula and has the semantic value $\langle \lambda z(\text{pai}(z, \text{that}(\text{jointfor}(x, y))))^*, \lambda z(\neg \text{pai}(z, \text{that}(\neg \text{jointfor}(x, y)))) \rangle$

MEANS ‘ $\text{MEANS}(x, y)$ ’ is a well-formed formula and has the semantic value $\langle \lambda z(\text{pai}(z, \text{that}(\text{sv}(x) = y))))^*, \lambda z(\neg \text{pai}(z, \text{that}(\neg \text{sv}(x) = y)))) \rangle$

Truth: What is someone who believes that it is true that P motivated to do?

TRUE ‘ $\text{TRUE}(x)$ ’ is a well-formed formula and has semantic value $\langle \lambda z(I^{\text{major}}(z, x)), \lambda z(I^{\text{minor}}(z, x)) \rangle$

Nice fact: This predicts the commitment table for propositional truth:

P	TRUE(THAT(P))
A	A
R	R
D	D

(By the way, this suffices for a Very Attractive Corollary:)

If ' $A \supset B$ ' is undeniable, then ' $\text{TRUE}(\text{THAT}(A)) \supset \text{TRUE}(\text{THAT}(B))$ ' is undeniable.

In other words: our account of logic, together with our account of truth, allows us to 'earn the right' to say that valid arguments are truth-preserving.

2.5 sentential truth and the deflationist resolution of the paradox

Standard definition of sentential truth: truth of the x such that S means x .

This allows us to predict the commitment table for sentential truth.

Hence, it allows us to predict full intersubstitutability.

Deflationary resolution of the paradoxes:

(1) there is nothing that you 'miss out on' by rejecting the liar proposition, because it is not about the world. Even if the world has to be one way or the other, the liar proposition doesn't have to be.

(2) We don't just stipulate the inferential properties of the sentences; we *explain* them, by appeal to the rational properties of the underlying states of mind.

(3) There is no paradox of revenge, because it is *always* consistent to reject any paradoxical proposition.

2.6 do we really avoid paradoxes of revenge?

For discussion.

references

Field, Hartry [2008]. *Saving Truth from Paradox*. Oxford: Oxford University Press.

Richard, Mark [2008]. *When Truth Gives Out*. Oxford: Oxford University Press.

Schroeder, Mark [2008]. *Being For: Evaluating the Semantic Program of Expressivism*. Oxford: Oxford University Press.
