# Assumptions about admissible models and the semantics

Igor Yanovich

Universität Tübingen

Models in Formal Semantics and Pragmatics @ ESSLLI 2014

August 19, 2014

# The "standard" task of the semantic theory

semantic competence

$\Uparrow$ explains

semantic theory

- We assume there is one real-world system to account for.
- Therefore we want our semantic theory to also be unique, at least in the limit.
- Seemingly little gain from explicit use of models, compared to disquotational semantics. So models may seem superfluous or even harmful.
  - Disquotational: 'Ernie is happy' is true iff Ernie is happy.
  - Model-theoretic: 'Ernie is happy' is true iff '$\text{Ernie}'^{\mathcal{M}} \in \text{'happy}'^{\mathcal{M}}$

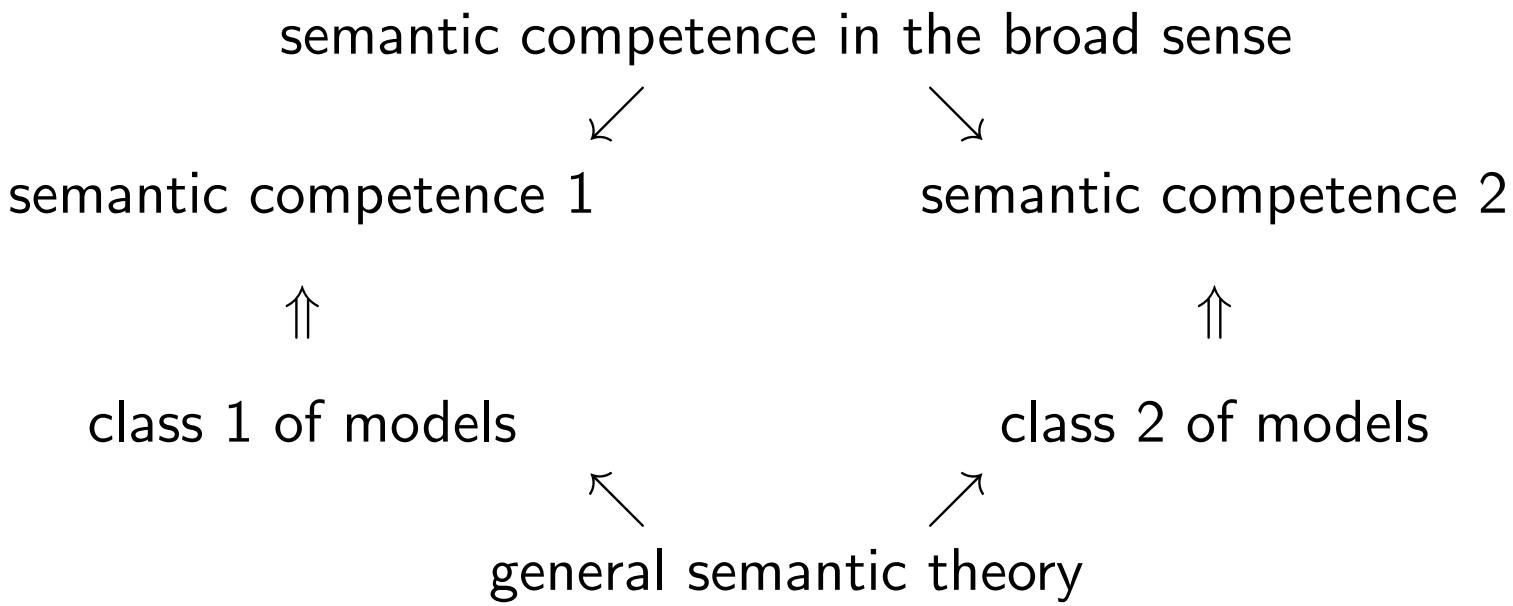
# What I hope to show sometimes happens


Why would we think there can be 2 semantic competences?

- When same expressions have arguably the same "general sense", but may be used with very different inference patterns. Those patterns may have different domains of applicability, but are not ordered by how "good" they are.

# So what about models?

In a model-theoretic setup, there is a straightforward account:




Two alternatives: 1) disquotational; 2) "naive" model-theoretic that does not really employ model classes. But neither would do the job.

# The sorites paradox

(1) a. **First polar claim:** if you have $100M, you are rich.

b. **Second polar claim:** if you have $0, you are not rich.

c. **Neighborhood claim:** if you have $5 less than some rich person, you are also rich.

- **Fact 1**: people find each of the three claims quite OK in isolation.
- **Fact 2**: together, they seem contradictory. 20M times $5 is $100M!

- Other versions:
  - bald: a person with 1 hair vs. with 100000 hairs
  - heap: one grain vs. enough grain to make a heap
  - yellow: 1 drop yellow+10K drops orange vs. 10K drops yellow + 1 drop orange

# Sorites: how to derive a contradiction

(1) a. **First polar claim:** if you have $100M, you are rich.
b. **Second polar claim:** if you have $0, you are not rich.
c. **Neighborhood claim:** if you have $5 less than some rich person, you are also rich.

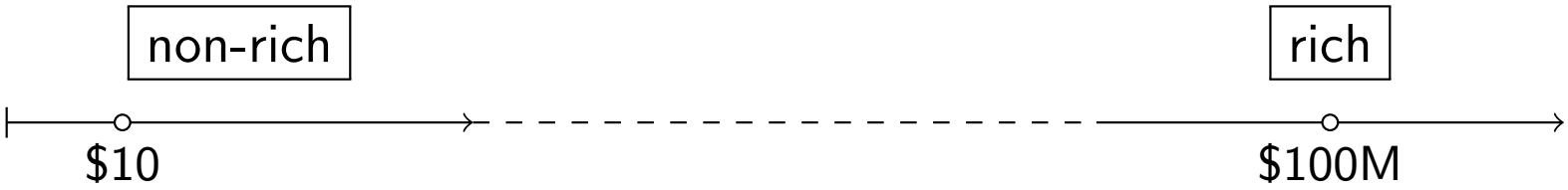
- **Built-in assumption about the model theory of** *rich*: you can turn a non-rich person into a rich one by a finite number of small increments on their fortune.
- The assumption follows if we make richness parasitic on the scale of money which is isomorphic to real numbers.

# Sorites: how to *not* derive a contradiction

- **Alternative assumption about the model theory of** *rich*: you can *not* turn a non-rich person into a rich one by small increments on their fortune.
- Quite reasonable by practical standards:
  - If A's fortune increases by \$5 every minute, it will take 38.03 years for A to get to \$100M.
  - If A's fortune increases by \$5 every second, it will still take 231.5 days.
  - And, just how likely is it for A with \$0 to have such steady increases?
- Interestingly, there is a well-defined and very natural mathematical theory of richness coming with the alternative assumption. It is inspired by non-standard models of arithmetic.

# 'Non-standard' account for the sorites

- **A real-number structure for richness**:




You can always get from one region to another in a finite # of steps.

- **A 'non-standard' structure for richness**:


No finite # of finite steps will take you from one region to the other.

# Non-standard model theory of richness

- Predicates: *rich*, *richer*, *almost.same.fortune*
- Postulates:
  - $\forall x : rich(x) \vee \neg rich(x)$
  - $\forall x, y : \neg rich(x) \wedge rich(y) \rightarrow richer(y, x)$
  - $\forall x, y : rich(x) \wedge richer(y, x) \rightarrow rich(y)$
  - $\forall x, y : \neg rich(x) \wedge richer(x, y) \rightarrow \neg rich(y)$
  - $\forall x, y : almost.same.fortune(x, y) \rightarrow (rich(x) \leftrightarrow rich(y))$

- We still derive that rich people are richer than poor people, and that everyone is either rich or poor.
- But we do not require that you could get rich by adding a finite number of small amounts of money.
  $\Rightarrow$ Richness is not parasitic on real-number amounts of money.
  $\Rightarrow$ *almost.same.fortune* is a primitive.

# Non-standard richness and the sorites

(2) a. **First polar claim:** if you have \$100M, you are rich.
$\Rightarrow$ true

b. **Second polar claim:** if you have \$0, you are not rich.
$\Rightarrow$ true

c. **Neighborhood claim:** if you have \$5 less than some rich person, you are also rich.
$\Rightarrow$ true

d. **Deriving contradiction:** if you are poor, but get a \$5 increment of your fortune 20M times, you become rich.
$\Rightarrow$ false

# 'Traditional' model theory of richness

- Indirect option:
  - pred. *fortune* maps people to real amounts (representing their money)
  - small *fortune* is $\neg$*rich*, big *fortune* is *rich*.
  - same *fortune* $\rightarrow$ both rich or both poor

- As this exploits the structure of real numbers, we have several consequences:
  - Adding enough small (but not infinitely small) increments can turn a poor person rich.
  - There must be a point where $\neg$*rich* ends and *rich* starts.
  - For most pairs of individuals with almost the same fortune, both would be *rich* or both would be $\neg$*rich*. But for a small number of pairs, it won't be so.
    So our earlier postulate will not hold:
    $$\forall x, y : \mathit{almost.same.fortune}(x, y) \rightarrow (\mathit{rich}(x) \leftrightarrow \mathit{rich}(y))$$

# 'Traditional' model theory of richness

- The direct option is trickier. If we can't exploit the structure of real values, we have to do special things to force it that poor people may be turned rich by finite increments.

  One version: a postulate schema that says that for any amount of money, you can get it down to 0 by subtracting \$1 $n$ times.

  Another version: $\forall x, y : \exists n_{finite} : fortune.difference(x, y) < (n * \$1)$. But this requires us to be able to define what a finite number is.

- Interestingly, the 'traditional' model theory for *rich* is theoretically more complex than the 'non-standard' model theory!

  Just as you can't define true arithmetic without going second-order...

# What two model theories derive

(3) a. **First polar claim:** if you have \$100M, you are rich.

non-standard: true traditional: true

b. **Second polar claim:** if you have \$0, you are not rich.

non-standard: true traditional: true

c. **Neighborhood claim:** if you have \$5 less than some rich person, you are also rich.

non-standard: true traditional: false

d. **Deriving contradiction:** if you are poor, but get a \$5 increment of your fortune 20M times, you become rich.

non-standard: false traditional: true

# Two semantic competences

- **Non-standard models of** *rich*: very easy inference, but work poorly with a large number of increments
- **Traditional models of** *rich*: work well with a large number of increments, but more complex, and entail a border between non-rich and rich persons.

- When people make their assumptions about the intended class of models explicit, they have to become committed to a particular set of truth values for the paradox's claims.

- But importantly, oftentimes we don't even have to choose.

  The perspective of logical comparison games (cf. Goranko's class at this ESSLLI): given a model $\mathcal{M}$, it may take very many rounds for us to determine which of the two classes it belongs to.

# The structure of our explanation

- The common semantic component explains the properties of *rich* common to both classes of models.
- The two classes of models **each** explain a particular pattern of human behavior associated with the use of *rich*.
- Both types of behavior are justified, though there are circumstances where only one of them would make sense. So we consider both classes of models to describe the full array of semantic behavior.
- A parallel: does the Sun orbit the Earth, or vice versa?
  - We know that strictly speaking, the trajectories are relative to a coordinate system, so neither is true.
  - When reasoning physically about the internal workings of the Solar system, it makes sense to choose the Sun as our 0, and then the Earth orbits the Sun.
  - When reasoning practically about the times of day, it makes sense to treat the Sun as if it orbited our location on the Earth.

# A designated class of models for temporal operators

(4) Condoravdi's (2002) designated class of models:

- for each $\langle w, t \rangle$, all future-accessible $\langle w', t' \rangle$ are such that $w$ and $w'$ agree on everything up to time $t$ (=are $t$-historical alternatives)

- In this class of models, the past and present are always already determined.

  Therefore, metaphysical modality about the present and the past loses sense: there's always only one way things can be among the metaphysical alternatives.

$\Rightarrow$ division of labor similar to the one I suggested for the sorites!

# A designated class of models for temporal operators

- Temporal *still* presupposes the presence of genuine metaphysical alternatives, and hence is a diagnostic for metaphysical readings. (Cf. "concessive" *still*.) In ordinary discourse, such *still* is incompatible with modal claims about the present:

  (5) She still may [get the flu tomorrow].
  $^{OK}$ 'There's still an objective, metaphysical possibility that she catches the virus tomorrow — and another one, that she doesn't.'

  (6) She still may [have the flu].
  # 'There's still an objective, metaphysical possibility that she has the flu — and another one, that she doesn't.'
  $^{OK}$ 'Well, there's also the epistemic possibility that she's ill. She's one or the other, but we don't yet know.'

- Semantic entries + the choice of a designated class of models together account for the speakers' semantic behavior.

# Alternative classes of models: multiple pasts

- But we can use our language with different intended classes of models.
- Suppose we are in a world where each moment may have multiple pasts, and the pasts available now (we can call them "the current pasts") may become unavailable later.

  (7) Context: Only those are eligible to enroll who attended Harvard in one of their current pasts.
  *Mary*: Hooray! I still may have gone to Harvard, so I am eligible to enroll!

- With a properly defined set of models, we can use the same semantics by Condoravdi for metaphysical modality — and capture Mary's semantic behavior.

# Alternative classes of models: circular time

- Or imagine worlds with circular time. We will still be able to use our English there just fine!
- In the class of models with circular time, the following is a validity:

  (8) If I ever was in London, then I will be there again.

  This validity will follow if we take the usual temporal semantics and apply it to the right class of models.
- So again, we can take our good old semantics for temporal expressions, add the intended (unusual) class of models — and voilà, we get an explanation for interesting semantic behavior.

# Consequences of the two case studies

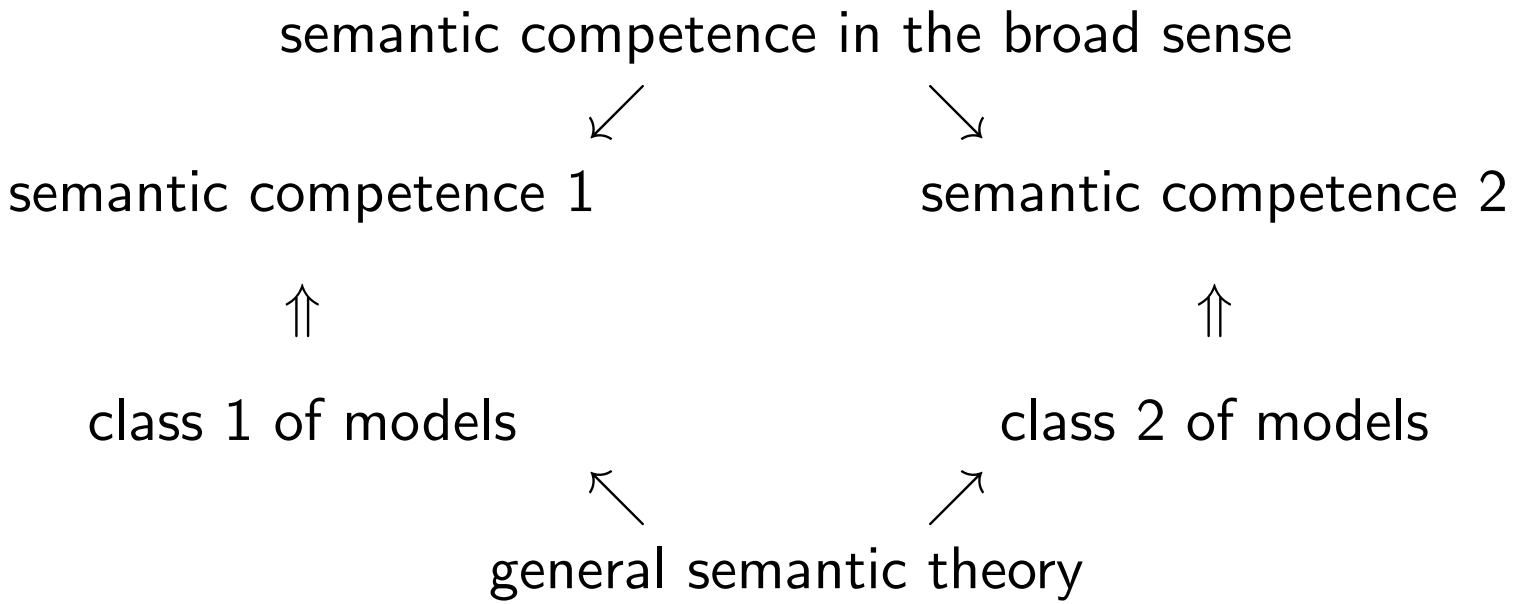
- The same expression may give rise to different inference patterns, in different circumstances.
- The broad semantics of the expression may stay the same.
- The differences may be captured via behavior of the same operator in different classes of models.

- Methodological consequence[1]: we should not equate the broad semantics of an expression with its interpretation in a given class of models.

  "Meta-intensionality": listing the reference of *Op* in all intended models does not capture the broad semantics of *Op*. Just as showing all and only chairs does not necessarily allow us to build the proper concept of a chair.

[1] Formulated thanks to a discussion with Stanley Peters.

# The form of our theories




- If we don't have the layer of models, we have to choose one of the semantic competences as the only "true" one.

  It's as if we asked which *really* orbits which, the Earth or the Sun.