On "Counterfactual Attitudes and Multi-Centered Worlds" (Ninan 2012)

Will Starr, Cornell University

will.starr@cornell.edu | http://williamstarr.net

September 29, 2012

··

"… I think the generality of my title ["Attitudes *De Dicto* and *De Se*"] is well justified. I'm not sure anything is left out – perhaps some ill-understood attitudes of **imagining**, conceiving, contemplating, or entertaining a thought." (Lewis 1979: 529, my emphasis)

1 The Lewisian View and Ninan's Challenge

- Intuitive basis of *centered worlds*:

  - A god, looking down on the world, can say everything there is to say about that world by specifying the states of affairs, laws, modal properties, etc.
  - This model is inappropriate to capture the rational behavior of agents located *in* a world
  - With an agent located within that world, once they know what the world is like needs to know which of the individuals in that world they are

- **Lewisian'** *De Se* **Semantics**:

  x believes, at w, t, that she (herself) is a spy:

  - $\{\langle w', t', x' \rangle \mid x' \text{ has at } w', t' \text{ all properties } x \text{ self-ascribes in } w, t\}$
    $$\subseteq \{\langle w', t', x' \rangle \mid x' \text{ is a spy at } w', t'\}$$

- Why isn't x the center of her beliefs?

  - *Essential properties problem*: x may not believe she came from a sperm and egg (Lewis 1983: 15)
    - ▸ But, any world where x exists, x came from a sperm and egg
    - ▸ So the set of centered worlds where x has all self-ascribed properties is identical to set of centered worlds where x has all self-ascribed properties and essential/necessary ones too

- **Question 1**: consider an agent x which self-ascribes being identical to Bush, and also fails to believe he came from a sperm and egg.

  - In order for x' to have the first property in w', they have to be Bush, but the second property is incompatible with this.

- **Lewisian** *De Re* **Semantics**:

  Ralph believes, at w, t, of Ortcutt, that he is a spy:

  - There is an acquaintance relation R:
    - (a) Ralph bears R uniquely to Ortcutt at w, t
    - (b) $\{\langle w', t', x' \rangle \mid x' \text{ has at } w', t' \text{ all properties Ralph self-ascribes in } w, t\}$
      $\subseteq \{\langle w', t', x' \rangle \mid \text{the } y \text{ to which } x' \text{ bears } R \text{ in } w', t' \text{ is a spy at } w', t'\}$
  - R is: seeing-sneak-around-on-the-dock

- Relation between *de re* and *de se*:

  - "That suggests that a *de se* belief is also a *de re* belief about oneself. This suggestion is vindicated by Lewis's theory if we count the relation of identity as a relation of acquaintance, and take a *de se* belief to be a *de re* belief about oneself relative to the relation of identity." (Ninan 2012: 13, on Lewis' theory)

- **Question 2**: how does this deal with standard cases of *de re* beliefs about oneself that fail to be *de se* beliefs?

  - Kaplan sees someone in the window with his pants on fire, not realizing it is his reflection
  - Perry, following a sugar trail he is unknowingly creating in the grocery store, thinks that this guy is making a mess
  - Further, if an acquaintance relation is just any relation by which information is being reliably transmitted (at that world-time), doesn't the identity relation always count?

Ninan's challenge:

- What's the content of the following attitude attribution: Ralph imagined that he did not see Ortcutt sneaking around on the dock?

- R is not instantiated in this counterfactual scenario, so clause (b) in the Lewisian semantics fails

  - But R is the only relation satisfying (a)!

- *The gist*: which acquaintance relations hold are contingent and so vary in counterfactual scenarios, but Lewis' theory makes those relations a necessary condition for the truth of *de re* attributions

---

2   Multi-Centered Worlds Solution

- Ninan (2012) proposes to distinguish possibilities more finely than centered worlds: multi-centered worlds

Old	$\langle w', t', x' \rangle$
New	$\langle w', t', \{ \langle \langle \text{Ralph}, Q \rangle, x' \rangle, \langle \langle \text{Ortcutt}, R \rangle, y \rangle, \dots \} \rangle$

(underbrace: f)

  - Centers are replaced with *tagging functions* f
  - Intuitively, f:
    - *Takes*: things x is acquainted with in w and a way x is acquainted with it (an *acquaintance pair*)
    - *Returns*: that thing's 'representative' in w'
  - In practice, the domain the various f's is generally taken to be the acquaintances/relations of the agent whose attitudes we are modeling – so they stay constant across multi-centered worlds
    - In practice, the domain of all tagging functions is the *acquaintance set* of the agent being modeled
    - What varies: individual an acquaintance pair is mapped to

- **Multi-Centered *De Re* Semantics** (informal version):

  x imagines, at w, t, of Orcutt that he (relative to R) is not dock-sneeking:

  - $Im_{x,w,t} \subseteq \{ \langle w', t', f' \rangle \mid f'(\text{Ortcutt}, R) \text{ is not dock-sneeking}_{w', t'} \}$
    - $Im_{x,w,t}$: x's imagination worlds in w, t
    - $Im_{x,w,t} = \{ \langle w', t', f' \rangle \mid \text{dom } f' = \{ \langle a_0, A_0 \rangle, \dots, \langle a_n, A_n \rangle \}$
    $\& \ f'(\langle a_0, A_0 \rangle), \dots, f'(\langle a_n, A_n \rangle) \text{ are related}_{w', t'}$
    $\text{as } x \text{ imagines}_{w, t} \}$

- Unlike Lewis' semantics: R needn't hold between x and Ortcutt at w, t

  - Though we could add that in to analyze *believes*

- **Question 2** (again): since it is seems so easy for the identity acquaintance relation to hold, how does it fail in cases of *de re* beliefs about oneself that aren't *de se*?

  - If identity acquaintance relations aren't so easy, what more is there to them than reliable information transmission?

- This finer conception of content addresses Ninan's challenge

- Let's assume Ninan's challenge is to be addressed by refining content and re-tuning the semantics of attitude verbs

- **My Main Question**: are tagging functions the best tool for doing this?

3   Worries about Multi-Centered Content

- **Worry 1**: tagging functions do not play the theoretical role of *indices*, like w and t, so they don't belong with them in a theory of *content*

  - One role of indices is to capture the ways in which our information can be *partial*
    - w: we can fail to know which total state the world is in
    - t: we can know that total state, but not what stage we're at
    - x: we can know which world-stage we're in, but not which individual we are (Lewis)
    - f: we can know which world-stage we're in, but not (i) what we're acquainted with and how, or if so, (ii) who's playing the role of that acquaintance in this world...
      - ▷ Variation in the *domain* of the tagging functions is not exploited, so type (i) partiality is not useful
      - ▷ Type (ii) is the same as Lewis' model
  - Another role: an essential component for determining truth
    - Ninan (2012: 42-3) relativizes truth to **assignment functions**
      - ▷ The tagging functions implicated in an agent's attitudes are used to set the values of the assignment function
      - ▷ Tagging functions not essential for determining truth!
  - Indices are possibilities communicated content distinguishes btwn
    - But multi-centered contents are often incommunicable, as Ninan (2012: 38n5) admits
    - When I believe I am hungry I believe:
    $\{ \langle w, t, f \rangle \mid f(\text{WS}, \text{identity}) \text{ is hungry}_{w, t} \}$
    No one else can believe a proposition containing this f

- **Worry 2**: what does it mean to have a world (partially) centered on an individual that doesn't exist there?

- **Worry 3**: a certain subset of the multi-centered worlds intuitively model an agent's perspective on a world, but most don't. Does this suggest that this conception of content is too general?

  - In practice, all individuals used as centers are individuals from the world of evaluation
  - All are acquaintances of an agent named in the attribution
  - Tagging functions capture a particular agent's *perspective*
  - We're in w, t, and are acquainted only with ourselves; a doesn't exist here. Where in logical space, with respect to us, is:
  $\langle w', t', \{ \langle \langle a, Q \rangle, b \rangle, \dots \} \rangle$

- **Worry 5**: can we make sense of plural attributions like *We (all) believe we are here in NJ*?
  - ○ Desiderata: analysis should entail I believe *de se* that I'm in NJ, that Ernie believes *de se* that he's in NJ, etc.
    - ▸ Can't say this if we know Ernie thinks he is at a concert in NYC but that the conference magnate he heard about is in NJ
  - ○ *Problem*: treating *we* as *I+you*, multi-centered view only predicts *de se* reading for speaker

4 Another Solution without the Worries?

- Don't let these worries rain on a productive research program, unless there's another party?

- Where's it? Nowhere, completely: Kamp (2011), Cumming (2008, to appeara), Bittner (2012), Kaplan (1968)
  - ○ Not actually sure it's happening, but worth looking into

- Ninan (2012) already makes essential use of variable assignments in his formal semantics:
  - ○ Names and pronouns are treated as variables, including *I* (§5.2)

- Basic picture:
  - ○ **Content**: sets of world-assignment pairs ('discourse content') (Dekker 1993; Groenendijk *et al.* 1996)
    - ▸ Assignments characterize a 'reference relation' between symbols and referents
      - ▹ $\langle w, g \rangle$: w is how the non-linguistic world is, g is how the symbols refer to things in that world (Cumming 2008)
  - ○ Attitudes: an agent can learn about the non-linguistic world (eliminate all possibilities like $\langle w, \ldots \rangle$) or the reference relation (eliminate all possibilities like $\langle \ldots, g \rangle$)
  - ○ *De Dicto*: reference relations compatible with subject's beliefs are used to interpret rigid designators in scope of *believes*
    - ▸ That is, attitude verbs are *assignment-shifters*
  - ○ *De Re*: projecting the reference relation in use in the context, onto the worlds from agent's 'belief set'
  - ○ *De Se*: assignment function has a special variable I which can only be assigned to speaker of context (Kamp 2011; Bittner 2012)
    - ▸ Some languages contain attitude verbs that shift even this variable (e.g. Schlenker 2003; Bittner 2012)

- The *de re* idea in a little more detail, for x believes of Ortcutt that he is a spy:
  - ○ Basic idea: if x referred as we do with Ortcutt then the proposition expressed by Ortcutt is a spy would be true in all of x's belief world-assignment pairs

- **De Re Semantics**:

  $[\![$x believes of Ortcutt that he is a spy$]\!]$ =
  $\{\langle w, g \rangle \mid \langle w', g' \rangle \in B_{g(x),w}$ & $g' \in g^{\frown w'}$ & $\langle w', g' \rangle \in [\![$Ort is a spy$]\!]$
  & $g(\text{Ort}) \in D_{w'}\}$

  - ○ $g^{\frown w'}$ is the set of projections of g onto w':
    - ▸ Every projection agrees with g when $g(\text{v})$ exists in w'
    - ▸ Otherwise, a projection has a suitable ersatz for $g(\text{v})$
      - ▹ Lots of options to explore for doing this
      - ▹ Acquaintance relations, counterparts, contextually salient property maximization, etc.
  - ○ If we didn't care about the essential properties problem, we could just alter x's g' to match g on Ort

- Getting around essential properties problem:
  - ○ In worlds where Ortcutt doesn't exist, g' can assign Ortcutt to something else and still, technically, be rigid
  - ○ One worry: do Ralph's anti-science belief worlds contain actual Ortcutt, who came from a sperm-and-egg, as well as miraculous-Ortcutt who was assembled in the womb from fairy-wings?

- *De re* imagination requires no special treatment, it's just a case where projection will very likely be necessary

- Getting around the incommunicability of *de se* contents:
  - ○ For this approach to work to Frege's problem, the symbols in the reference relation need to be an agent's *mental symbols*
  - ○ Since no two agent's share mental symbols they don't share discourse content
  - ○ However, one can tell a story about two speakers who share a public language can come into alignment such that we can say they *carry the same information* (Cumming to appearb)
  - ○ For names, roughly:
    - ▸ They have aligned signaling strategies: if x encodes mental symbol s_x with a and y construes a with their mental symbol s_y, then y uses a to encode s_y and x construes it with s_x
  - ○ For indexicals, the inverse!
  - ○ So communicating *de se* contents is the same as communicating *de dicto* contents

References

BITTNER, M (2012). 'Perspectival Discourse Referents for Indexicals.' In H GREEN (ed.), *Proceedings from SULA 7*. URL http://www.rci.rutgers.edu/~mbittner/bitt12_sula7.pdf.

CUMMING, S (2008). 'Variablism.' *Philosophical Review*, **117(4)**: 525-554.

CUMMING, S (to appeara). 'Coordinating in the Dark: from private symbols to Fregean content.' *Philosopher's Imprint*. URL https://dl.dropbox.com/u/31660624/kpuzzle-8-12.pdf.

CUMMING, S (to appearb). 'Creatures of Darkness.' *Analytic Philosophy*. URL http://dl.dropbox.com/u/31660624/creatures-7-11.pdf.

DEKKER, P (1993). *Transsentential Meditations*. Ph.D. thesis, University of Amsterdam, Amsterdam.

GROENENDIJK, J, STOKHOF, M & VELTMAN, F (1996). 'Coreference and Modality.' In S LAPPIN (ed.), *The Handbook of Contemporary Semantic Theory*, 179-213. Oxford: Blackwell Publishers.

KAMP, H (2011). 'Representing *De Se* Thoughts and Their Reports.' Ms. University of Stuttgart, URL http://nasslli2012.com/files/kamp_2011.pdf.

KAPLAN, D (1968). 'Quantifying In.' *Synthese*, **19(1/2)**: 178-214.

LEWIS, DK (1979). 'Attitudes *de Dicto* and *de Se*.' *The Philosophical Review*, **88**: 513-543.

LEWIS, DK (1983). 'Individuation by Acquaintance and by Stipulation.' *The Philosophical Review*, **92(1)**: 3-32. URL http://www.jstor.org/stable/2184519.

NINAN, D (2012). 'Counterfactual attitudes and multi-centered worlds.' *Semantics and Pragmatics*, **5(5)**: 1-57.

SCHLENKER, P (2003). 'A Plea for Monsters.' *Linguistics and Philosophy*, **26(1)**: 29-120.