Ask Not What Semantics Can Do For Dialogue

Ask What Dialogue Can Do For Semantics

Matthew Purver
(and many others)
SemDial 2014
Meaning from Observation
# Meaning from Observation

<table>
<thead>
<tr>
<th>Ronnie B</th>
<th>/fɔːkændəʊʒ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ronnie C</td>
<td>🎃 🎃 🎃 🎃</td>
</tr>
<tr>
<td></td>
<td>Four candles</td>
</tr>
<tr>
<td>Ronnie B</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>‘Andles for forks</td>
</tr>
<tr>
<td>@gaskarthlrh</td>
<td>finally got my 5sos follow back</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>@sleepykidlrh</td>
<td>@gaskarthlrh you mean the band?</td>
</tr>
<tr>
<td>@gaskarthlrh</td>
<td>@sleepykidlrh yeahh</td>
</tr>
</tbody>
</table>
Five Seconds Of Summer

© Dan Jones / The Sun
Dialogue: a Semantic Observatory

• We can tell what things mean by seeing how people respond to them
  – (particularly when they’re trying to repair)

• (Perhaps “what things mean” is “how people respond to them”?)

• Studying dialogue help us study semantics
PROBING SEMANTIC THEORIES
NP Semantics

• The simplistic view

  “John” is of type e:  \( john' \)
  \[
  \text{VP(NP)} \rightarrow \lambda x. \text{snore}(x)(john') \rightarrow \text{snore}(john')
  \]

• The traditional GQ view

  “John” is of type \((e>t)>t\):  \( \lambda P. P(john') \)
  \[
  \text{NP(VP)} \rightarrow \lambda P. P(john')(\lambda x. \text{snore}(x)) \rightarrow \text{snore}(john')
  \]
  “Every man”:  \( \lambda P. P( \forall x. \text{man}(x) \land P(x)) \)
  \[
  \text{NP(VP)} \rightarrow \forall x. \text{man}(x) \land \text{snore}(x)
  \]
Clarification Requests

Ann: I saw John yesterday.
Bob: John??
Ann: Yes, John. Dr Smith. The one with the pipe & monocle. Him.

<points>
Men, Englishmen, old Etonians, people who have climbed Everest in striped pyjamas, ...
Smoking, being shortsighted, being upper class, climbing Everest in striped pyjamas, ...
George: You want to tell them, bring the tourist around show them the spot
Sam: The spot?
George: where you spilled your blood

Unknown: What are you making?
Anon 1: Erm, it’s a do- it’s a log.
Unknown: A log?

Anon 1: It had twenty rooms in it.
Anon 2: Twenty rooms?
Anon 1: Yes.
Clarifying NP Semantics

• Sometimes quantifier, sometimes CN property, sometimes referent set …
  – … but always lower-order: never sets of sets/properties directly
  – (Purver & Ginzburg, 2004)
• Even with logical quantifiers:
  Richard: No I’ll commute every day
  Anon 6: Every day?
  Richard: as if, er Saturday and Sunday
  Anon 6: And all holidays?
  Richard: Yeah
• Denotation of NPs as witness sets of type e
  – “John”: \{john’\}
  – “Every man”: \{x / man(x)\}
  – (or pairs of reference & complement sets)
The GQ Strikes Back

• Cooper (2013): the problem’s not with GQs
  – rather, with standard GQ-compatible NPs
  – proposes a friendly amendment

\[ q\text{-params} : [ w : \text{all(} \text{man} \text{)} ] \]
\[ content \ : \lambda P. [ c=w : \text{all(} \text{man}, P \text{)} ] \]

– explains possible CR readings
  • and why some impossible (GQ scope i.e. VP content)
  • (although possibly not all …)

• Can we tell which should be preferred?
  – perhaps not yet, but we’re better off than we were
Dialogue provides constraints on semantics

(so it can help us work out what things mean – or don’t mean)
GETTING MORE EMPIRICAL
Clarifying Lexical Semantics

• NPs aren’t the only thing we clarify …
• … but they’re by far the most common thing.
• Excluding whole sentences etc:
  – NP/Pro/PN/CN: 76%
  – Adj/Adv/Mod: 12%
  – Det: 4% (mostly numbers)
  – VPs: 4%
  – Verbs: 1%
  – Prep/Conj: <0.5%
Clarifying Lexical Semantics

• NPs aren’t the only thing we clarify ...
• ... but they’re by far the most common thing.
• Excluding whole sentences etc:
  – NP/Pro/PN/CN: 76%
  – Adj/Adv/Mod: 12%
  – Det: 4% (mostly numbers)
  – VPs: 4%
  – Verbs: 1%
  – Prep/Conj: <0.5%
Clarifying Lexical Semantics

• NPs aren’t the only thing we clarify …
• … but they’re by far the most common thing.
• Excluding whole sentences etc:
  – NP/Pro/PN/CN: 76%
  – Adj/Adv/Mod: 12%
  – Det: 4% (mostly numbers)
  – VPs: 4%
  – Verbs: 1%
  – Prep/Conj: <0.5%
Clarifying Lexical Semantics

• NPs aren’t the only thing we clarify ...
• ... but they’re by far the most common thing.
• Excluding whole sentences etc:
  – NP/Pro/PN/CN: 76%
  – Adj/Adv/Mod: 12%
  – Det: 4% (mostly numbers)
  – VPs: 4%
  – Verbs: 1%
  – Prep/Conj: <0.5%
Content vs Function Words

• In some cases, this makes sense ...
• Function word clarification very rare
• Function words more familiar:
  – Low type-token ratio (i.e. less rare)
• Function words less contentful:
  – Low variance across genres
  – Low information content (surprisal)
• Perhaps clarification just doesn’t make sense?
  – It would be nice if we could test this ...
Experimenting with Dialogue

- GodFather: wots happening bro!
- GodFather: sat sri akal.
- Gagz: long time no chat
- Gagz: kiddha
- GodFather: yea man
- GodFather: where u bin

- Gagz is typing a line...

- GodFather: wots happening bro!
- GodFather: sat sri akal.
- Gagz: long time no chat
- Gagz: kiddha
- GodFather: yea man
- GodFather: where u bin

- i been hicing

• Insert fake clarifications:
  – Repeat words from previous turns
  – Wait for response

• Content words: 45% responded to
  – The vast majority as direct CRs (92%)

• Function words: only 15% response
  – And none of those as direct CRs

Laura: Can I have some toast please?
Jan: Some?
Laura: Toast

• So maybe we understand content vs. function
What about Verbs?

• But in other cases it seems plain weird!
• Verb clarification is vanishingly rare. Why?
    • no examples found for action-reference class
    • 51% of examples were NP or deictic reference

A: You see this thing did you buy this separately or did it come in the Walkman?
B: We were lent them.
A: Lent them?
B: Yeah.
What about Verbs?

• But in other cases it seems plain weird!
• Verb clarification is vanishingly rare. Why?
    • no examples found for action-reference class
    • 51% of examples were NP or deictic reference
• Verbs are no less contentful than nouns
  – Similar (high) type-token ratio, variance
  – Similar (high) information content
• Verb clarifications are easy to interpret
  – Just as likely to get a response
  – And get responded to in parallel ways
Perhaps Verbs are Not Nouns

• Do verbs & nouns have different semantic (cognitive?) status?
  – Conventionally both $e>t$:
    \[
    \lambda x.\text{snore}(x) \quad \lambda x.\text{woman}(x)
    \]

• Perhaps verbs are structured around arguments
  – ... which are mostly NPs ...
  – ... and then we tend to clarify those NPs?

• Frame semantics:
  – SELL[ buyer, seller, goods, money, ... ]
Dialogue poses questions about semantics

(about what things mean, what things don’t mean, and what differences must be accounted for)
WHAT ABOUT PROCESSING?


## Processing Issues

- We can clarify before the end of a sentence

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>And er they X-rayed me, and took a urine sample, took a blood sample. Er, the doctor</td>
</tr>
<tr>
<td>B</td>
<td>Chorlton?</td>
</tr>
<tr>
<td>A</td>
<td>Chorlton, mhm, he examined me, erm, he, he said now they were on about a slide (unclear) on my heart. Mhm, he couldn’t find it.</td>
</tr>
</tbody>
</table>

- This tells us a lot about semantic processing
  - In interpretation
  - In generation
Processing Issues

• We can clarify before the end of a sentence

<table>
<thead>
<tr>
<th></th>
<th>And er they X-rayed me, and took a urine sample, took a blood sample. Er, the doctor</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Chorlton?</td>
</tr>
<tr>
<td>A</td>
<td>Chorlton, mhm, he examined me [...]</td>
</tr>
</tbody>
</table>

• At this point, both A & B must know:
  – That it’s a constituent
  – That it’s potentially referential to an individual
  – What a possible world/dialogue reference might be
Compound Contributions

• Not just clarification ...
• Completions of incomplete antecedents:

   D: Yeah I mean if you’re looking at quantitative things it’s really you know how much actual-­‐ How much variation happens whereas qualitative is ⟨pause⟩ you know what the actual variations
   U: entails

• Expansions of “complete” antecedents:

   T: It’ll be an E sharp.
   G: Which will of course just be played as an F.
Incrementality

• Incremental processing
• Incremental semantic representation
• Incremental semantic interpretation
• Incremental reference
• Incremental context
• Incremental extensibility
• Incremental reversibility (parsing/generation)
DynDial & RISER

• Dynamic Syntax
• Type Theory with Records
  – (IWCS 2011, TTNLS 2014)

I want to go to Paris

\[
\begin{aligned}
e &= \text{now} & : & e_s \\
e_1 &= \text{future} & : & e_s \\
x_1 &= \text{Paris} & : & e \\
p_2 &= \text{to}(e_1, x_1) & : & t \\
x &= \text{speaker} & : & e \\
p_1 &= \text{go}(e_1, x) & : & t \\
p &= \text{want}(e, x, p_1) & : & t
\end{aligned}
\]
Self-repair

• Incrementality & monotonicity:
  
  The interview was – it was alright
  I went swimming with Susan – or rather, surfing
  – Maintain semantic context, but with ...
  – incremental parsing & choice mechanisms (Hough, 2012-14)
Dialogue provides constraints on semantic processing

(about when & how we understand and produce meaning and components of meaning)
GETTING EVEN MORE EMPIRICAL
Learning

- It’s all very well testing our existing theories
  – (my armchair is very comfy, actually)
- But can we **learn** a good framework?
- If dialogue gives us evidence for semantics, we should be able to learn that semantics
- Of course, we’d need a lot of data with people talking to each other about stuff …
Distant Supervision

• A common technique for sentiment detection

  Best day in ages! #Happy :)  
  just because people are celebs they dont reply to your tweets! NOT FAIR :(  

^_^*
Distant Supervision

• A common technique for sentiment detection

Best day in ages!
just because people are celebs they dont reply to your tweets! NOT FAIR
Distant Supervision

- A common technique for sentiment detection

  Best day in ages!

  just because people are celebs they don't reply to your tweets! NOT FAIR

再做个梦如果明天我中奖了该怎么支配呢每次想这个问题都很美**^_^**

离队倒计时，期待奇迹的发生 (T_T)
Distant Supervision

• A common technique for sentiment detection

Best day in ages!
just because people are celebs they dont reply to your tweets! NOT FAIR
再做个梦如果明天我中奖了该怎么支配呢每次想这个问题都很美
离队倒计时,期待奇迹的发生

• Go et al (2009): works well if you have a reliable but (semi-)independent label to hand
Distant Supervision

• Often independent labels aren’t reliable
  
  : - O  : - @  : - $  : - P

• Often reliable labels aren’t independent

Vodafone signal #fail
Gets so #angry when tutors don’t email back
Distant Supervision

• Often independent labels aren’t reliable
  :-O     :-@     :-$     :-P

• Often reliable labels aren’t independent

  Vodafone signal
  Gets so     when tutors don’t email back

• Poor results for many emotions
  – (Purver & Battersby, 2012)
Responses as Distant Supervision

• But what if someone responds?

_AggieGirl16: @captain_lizard lol yeaaaah. I'm pretty lucky! Haha!
captain_lizard: @_AggieGirl16 I'm glad you're happy, Monica! :)
Responses as Distant Supervision

• What do these have in common?

MattHDGamer: EA Servers down again?!
OrFIFAProdigy: you're surprised?
mattryanharris: Another school shooting? What the actual fuck.
BasedGoDEnigma: You seem like you're surprised?
danni_13_ONLY: HES GAY?!?! What the hell!
BeastyyLove: you're surprised? ! Lol

• Build classifiers better or same with much less data
Responses as Distant Supervision

• Does OK for simple distinctions (*happy vs not*)
  – better than hashtags, worse than emoticons
  – with a dataset half the size

• Similar on 6-way emotions, with 10% of the data
  – 77% accuracy
  – similar per-class f-scores

• Better at subtle distinctions e.g. *angry vs surprised*
  – 75% accuracy with <1000 training examples
  – (an emoticon-based *angry* classifier achieves 76% “accuracy” on *surprised* data!)
Questions as Distant Supervision

• Q8. In what city is the maracana stadium located?
  #Nairabet #Mightygeorgegiveaway
  – Brasilia

• Thanks. What city is the 24-hour fitness?
  – Oxnard

• what city is the quarry in?
  – Monroe. Exit 11 off 75. Lol.
Clarification as Distant Supervision

- @xxsylviaaxx Midnight Red is actually really good
  - you mean the band?
  - haha yeah the boyband

- @gaskarthlrh finally got my 5sos follow back
  - you mean the band?
  - yeahh

- @Joe_rauchet Any girl that likes the red hot chili peppers immediately becomes 100x more attractive
  - oh you mean the band

- @sbrezenoff Why is it so hard to find a youth-size Boston T-shirt with the guitar spaceships?!?
  - you mean the band & not city, don’t you? #notwherebrainwent
  - the band the band!!
Dialogue provides a basis for learning meaning

(helping build systems that can understand or produce meaning)
GETTING MORE EMPIRICAL AND MORE SEMANTIC ...
Learning Semantic Grammars

- We can learn lexical entries/grammars from sentential LFs:

“you read the book”
\textit{read(you, the(x, book(x)))}

LF constrains meaning (as situations do?)
Learning a Dialogue Grammar

- We can learn one that’s suited for dialogue:
  - i.e. incremental in all the necessary ways
  - Eshghi et al (2013): 92% coverage, 85% F-score on CHILDES
Learning Without Grammar

- We might even be able to learn one without a grammar ...
  – (ConCreTe project: Wiggins, Forth, Griffiths et al)

Semantic context

- Speech recognition

Semantic association

- Recognise → Speech → Language

Word

- To

- Wreck → Nice → Beach

Morpheme

- Tek

- Rek → Keg → Niviz → Spi:tf

Phoneme

- T → E → R → E → K → A → N → L → S → B → I → T → J

Viewpoint Model

- D1 → D2 → D3 → D4 → D5 → D6 → D7 → D8 → D9 → D10

- Chromatic Pitch
- Scale Degree
- Mode
- Duration Ratio
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duration
- Duratio
Learning from Dialogue?

• We can even learn from questions & answers:
  – Liang et al, 2011

    \[
    \begin{array}{cc}
    \text{column 1} & \text{column 2} \\
    \text{(OK)} & \text{(TX,2.7e5)} \\
    \text{(NM)} & \text{(TX,2.7e5)} \\
    \text{(NV)} & \text{(CA,1.6e5)} \\
    \ldots & \ldots \\
    \end{array}
    \]

  – But these are database query results, not utterances ...

• See Eshghi & Lemon (this afternoon)
• And Moradlou & Ginzburg (the day after tomorrow)

• Dialogue utterances provide similar (less specific) information:
  – Responses restrict the space of antecedent meanings
  – So could we learn semantic grammars from dialogue?
Learning from Dialogue?

• Can we learn semantics from dialogue alone?

• General problem:
  – Learn to construct representations which match the distribution of responses
  – (cf: learn to construct queries which match the distribution of answers)
  – A very unconstrained space
  – Large number of latent variables

• Distributional semantics
  – Vector space representations of meaning
  – Geometric modelling of distributions & relations
  – (cf: learn to construct representations which match distributions of lexical context)
Distributional Semantics

- Vector space representations of words
  - Co-occurrence-based or learned (Mikolov et al, 2013)
  - *apple* close to *orange*, far from *pavement*
  - *(king – queen)*
    - \( \approx (man – woman) \)
    - \( \approx (uncle – aunt) \)

- Compositional approaches:
  - Learned e.g. neural net-based (Socher et al, 2012)
  - Tensor-based (Coecke et al, 2010)
    \[
    \sum_{ijk} C_{ijk} \langle \text{dogs} | \overrightarrow{n_i} \rangle s_j \langle \overrightarrow{n_k} | \text{cats} \rangle
    \]
VSMs for Dialogue Act Tagging

- Kalchbrenner & Blunsom (2013)
  - learn word representations & context update functions jointly

\[ p_i = \text{softmax}(O_i h_i + b_o) \]

\[ h_i = (I x_{i-1} + H_i h_{i-1} + S s_i + b_h) \]

\[ p_i = P(x_i | x_{<i}, s_i, a_i) \]
Distributional Pragmatics?

• Can we produce a *compositional* version?
  – Compositional distributional models help DA tagging
    • (Milajevs & Purver, 2014; Milajevs et al, 2014)

• Sentences as vectors plus “dialogue act” tensors
  – e.g. (Paperno et al, 2014) vector + tensor model
    \[
    \left\langle \vec{x}, \hat{1}^1, \ldots, \hat{n}^n \right\rangle
    \]
  – estimate contextual tensors directly

• Estimate tensors and lexical vectors jointly:
  – Learn lexical (& phrasal) semantics & pragmatics directly!
  – i.e. learn what things mean from how people respond
In Summary

Hooray for dialogue!

(even if you don’t care about dialogue)
IWCS 2015 in London

13th-16th(ish) April 2015
Queen Mary University of London
in London’s vibrant and fashionable East End™

I hope there will be dialogue!
Thanks!

• To you and:
  – Julian Hough
  – Arash Eshghi
  – Christine Howes
  – Dmitrijs Milajevs
  – Mehrnoosh Sadrzadeh
  – Dimitri Kartsaklis
  – Stephen McGregor
  – Zheng Yuan
  – Stuart Battersby
  – Geraint Wiggins
  – Pat Healey
  – Ruth Kempson
  – Jonathan Ginzburg