The interplay of linguistic & contextual inferences

Christopher Potts

Stanford Linguistics

SUBTLE meeting, Penn, October 16, 2009
Overview

The original task for pragmatics

- Investigate the way language shapes, and is shaped by, the Questions Under Discussion (QUDs).
- Integrate the results with those of syntax and semantics, machine learning, and robotics.
Overview

The original task for pragmatics

- Investigate the way language shapes, and is shaped by, the Questions Under Discussion (QUDs).
- Integrate the results with those of syntax and semantics, machine learning, and robotics.

Progress

- Ongoing corpus collection efforts provide the empirical foundations.
- We implement pragmatic theories as sets of weighted logical statements, which are interpreted in graphical models.
- The theories provide inferences about both directions of influence: language to context, context to language.
- They also support large-scale lexical pragmatics.
Outline

1. QUDs and robust utterance understanding
2. Pragmatics in Markov Logic
3. Lexical pragmatics
4. Connections with other team members
5. Prospects
Questions Under Discussion

Which flights have seats available?

- Does the 7:00 flight have seats available?
- Does the 7:30 flight have seats?
- The 8:00 pm flight?

- Window seats?
- Aisle seats?
Questions Under Discussion

Figure 1: Part of the QUD structure for a discourse about airplane bookings, with the arrows connecting super-questions to sub-questions.

3.2.3 Question–answer alignment

If we assume that relevance is not binary, but rather organized by the size of the Ans sets defined in definition 8, then we articulate a view on relevance.

Definition 16 (Relevance; G&S, van Rooy 2003):

\[ p \succ Q q \iff \text{Ans}(p, Q) \subset \text{Ans}(q, Q) \text{ or } \text{Ans}(p, Q) = \text{Ans}(q, Q) \text{ and } q \subset p \]

(9) Here, relevance is maximized in the above set-theoretic sense, by resolving she to Lisa and filling out the ellipsis with (the meaning of) tired.

(10) A Where are you from?

BP a l o A l t o / C a l i f o r n i a / T h e U.S. / E a r t h . . .

Where \{city\} \sqsubset Where \{state\} \sqsubset Where \{country\}...

Palo Alto \succ California \succ U.S.

In general, if the Ans set is not a singleton, then something interesting happens — the original question is partly rejected and/or an epistemic/relevance/politeness implicature is generated by the answer and/or . . .

Example

Traveler: Does the 7:30 flight have seats available?

Agent: No, but the 7:00 flight has some window seats.

Traveler: Are they in an exit row?
Questions Under Discussion

Which flights have seats available?

- Does the 7:00 flight have seats available?
- Does the 7:30 flight have seats?
- The 8:00 pm flight?

Window seats? Aisle seats?

Example

Traveler: Are there aisle seats available on the 7:30 flight?
Agent: There are seats available.
Partition semantics and lattice structures

Definition (Declaratives)

$$\lbrack \varphi \rbrack^g = \{ \langle w, w' \rangle \mid \forall d \in D^n : \lbrack \varphi \rbrack^g[\vec{x} \mapsto d](w) = \lbrack \varphi \rbrack^g[\vec{x} \mapsto d](w') = T \}$$

Definition (Interrogatives)

$$\lbrack ?\vec{x}^n \varphi \rbrack^g = \{ \langle w, w' \rangle \mid \forall d \in D^n : \lbrack \varphi \rbrack^g[\vec{x} \mapsto d](w) = \lbrack \varphi \rbrack^g[\vec{x} \mapsto d](w') \}$$

Definition (Entailment)

$Q$ entails $Q'$ iff $Q \subseteq Q'$

Definition (Granularity)

$Q \subseteq Q'$ iff $\forall q \in Q \exists q' \in Q' \; q \subseteq q'$
Implicit QUDs and quantificational domains

Fast and cheap corpus collection

5,000 question–answer pairs concerning the simple context at right.

1. Are the triangles green?
   “Yes”: 10%
   “No”: 16%
   Other: 24%

2. I am curious about the things above shape C. Are the triangles green?
   “Yes”: 36%
   “No”: 4%
   Other: 9%
Task-like QUDs and referential underspecification

Player 2: Look for 2.
Player 1: and the 3?
Task-like QUDs and referential underspecification

**Task:** Find six consecutive cards of the same suit.

Player 2: Look for 2.
Player 1: and the 3?
Task-like QUDs and referential underspecification

**Task**: Find six consecutive cards of the same suit.

**Context**: The players are holding \{4H,KH\}

Player 2: Look for 2.

Player 1: and the 3?
Task-like QUDs and referential underspecification

**Task**: Find six consecutive cards of the same suit.

**Context**: The players are holding \{4H,KH\}

Player 2: Look for 2.

Player 1: and the 3?

---

**Basic method**

For any nominal, pick the referent(s) that maximize relevance with respect to the task.
**Task-like QUDs and referential underspecification**

**Task:** Find six consecutive cards of the same suit.

**Context:** The players are holding \( \{4H,KH\} \)

- Player 2: Look for 2.
- Player 1: and the 3?

---

**Basic method**

For any nominal, pick the referent(s) that maximize relevance with respect to the task.

---

**Results (Potts 2009)**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct inference</td>
<td>164</td>
</tr>
<tr>
<td>Incorrect inference</td>
<td>8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>172</strong></td>
</tr>
</tbody>
</table>

(95%)
Intonation and implicit QUD structure

Bart passed  \( \text{BART}_F \) passed.  Bart \( \text{PASSED}_F \)
Intonation and implicit QUD structure

Did Bart pass?  Who passed?  What did Bart do?

Bart passed  BART\textsubscript{F} passed.  Bart \text{PASSED}\textsubscript{F}
Intonation and implicit QUD structure

Did Bart pass?  
↑  
Bart passed

Who passed?  
↑  
BART\textsubscript{F} passed.

What did Bart do?  
↑  
Bart PASSED\textsubscript{F}

Who wore what?  
↑  
The ROCKSTARS wore LEATHER.
Pragmatics in Markov Logic

Alchemy

- Software developed by Pedro Domingos and colleagues to conduct experiments with probabilistic weighted logics.
- Performs a variety of tasks in structured learning.
- Efficiently searches large spaces of possible worlds.
- Delivers formulae with probabilities attached.
Pragmatics in Markov Logic

Alchemy

- Software developed by Pedro Domingos and colleagues to conduct experiments with probabilistic weighted logics.
- Performs a variety of tasks in structured learning.
- Efficiently searches large spaces of possible worlds.
- Delivers formulae with probabilities attached.

Advantages

- Quick for running experiments.
- Can grapple with large numbers of predicates (assuming they are not too interconnected).
- Intelligible; theories take the form of sets of logical formulae.
- Permits an analytic division between general pragmatic pressures and context-specific information.
Model

Given a set of constants, the probability of a given context \( c \) is defined in terms of the logical statements \( j \) and their associated weights \( w_j \):

\[
P(X = c) = \frac{1}{Z} \exp \left( \sum_j w_j f_j(c) \right)
\]
A simple example

Switchboard example (de Marneffe, Grimm & Potts 2009)

A: Are they [your kids] little?
B: I have a seven-year-old and a ten-year-old.

Learning weights

Little(x)
Age(x, y) ^ Little(x)
Age(x, u) ^ Age(y, v) ^ (v >= u) ^ Little(y) => Little(x)

⇒

Stanford typed dependencies

⇒

Age(X21, 10) Little(X21)
Age(X22, 7) Little(X22)
Age(X23, 8) Little(X23)
Age(X35, 10) !Little(X35)
Age(X36, 18) !Little(X36)
Age(X37, 14) !Little(X37)
General pragmatic pressures

- Domain independent, but subject to reweighting in different contexts.
- Delimit the set of admissible contexts, by governing inferences based on shared assumptions about cooperative linguistic behavior rather than basic world knowledge.
- Center around the QUDs.
- Govern hearer inferences and, in turn, speaker utterances.
- Weights should be estimated in human-subjects experiments.
General pragmatic pressures: Some examples

// Quality
AssertTrue(p) => True(p).
True(p) => AssertTrue(p).

// Relevance
AssertTrue(p) => (Qud(q) ^ About(p, q)).
10 (Qud(q) ^ About(p, q)) => AssertTrue(p)

// Quantity
AssertTrue(p) => !AddresseeBelieve(p).
10 !AssertTrue(p) => AddresseeBelieve(p)
General pragmatic pressures: Some examples

// Addressee correction (quality).
10 AssertFalse(p) => (AddresseeBelieve(p) ^ !True(p))
10 (AddresseeBelieve(p) ^ !True(p)) => AssertFalse(p)

// Addressee correction (relevance).
10 AssertFalse(p) =>
   (AddresseeBelieve(p) ^ !True(p) ^ Qud(q) ^ About(p, q))
10 (AddresseeBelieve(p) ^ !True(p) ^ Qud(q) ^ About(p, q)) => AssertFalse(p)
Pragmatics is not first-order

Pragmatics involves reasoning about higher-order objects, in particular, QUDs, but Alchemy is not equipped for this. Over finite domains, the difficulties are surmountable but cumbersome.
Pragmatics is not first-order

Pragmatics involves reasoning about higher-order objects, in particular, QUDs, but Alchemy is not equipped for this. Over finite domains, the difficulties are surmountable but cumbersome.

Defining complex properties

\[ \varphi \iff (\psi_1 \cdots \psi_n) \]

Tractable if \( n \) is small (around 2), but often intractable if \( n \) is around 8. This makes it hard to define complex properties.
Challenges

**Pragmatics is not first-order**

Pragmatics involves reasoning about higher-order objects, in particular, QUDs, but Alchemy is not equipped for this. Over finite domains, the difficulties are surmountable but cumbersome.

**Defining complex properties**

\[ \varphi \iff (\psi_1 \cdots \psi_n) \]

Tractable if \( n \) is small (around 2), but often intractable if \( n \) is around 8. This makes it hard to define complex properties.

**SUBTLE team to the rescue**

FACTORIE is expressively unconstrained and allows the user to write intelligent search functions to deal with constraints like the above.
Lexical pragmatics: Vague predicates

What is likely to count as *little* for kids?

Stanford typed dependencies

(a fragment of ongoing work by Marie-Catherine de Marneffe)
Scalar modifiers

What is the emotive contribution of scalar modifiers?
(Potts & Schwarz 2008; Constant, Davis, Potts et al. 2008)

Customer Reviews

726 Reviews

5 star: 246
4 star: 198
3 star: 146
2 star: 73
1 star: 63

Average Customer Review

🌟🌟🌟🌟 (726 customer reviews)

Utter(S1, Superb) HighlyPositive(Superb)
Utter(S2, Superb) HighlyPositive(Superb)
Utter(S3, Superb) Positive(Superb)
Utter(S4, Superb) HighlyPositive(Superb)
Utter(S5, Superb) Positive(Superb)
Utter(S6, Superb) HighlyPositive(Superb)
Utter(S7, Superb) HighlyNegative(Superb)
Utter(S8, Superb) Positive(Superb)
Utter(S9, Superb) Neutral(S1)
Utter(S10, Somewhat) Neutral(S1)
Utter(S11, Somewhat) Positive(S1)
Utter(S11, Somewhat) Neutral(S1)
...
Epistemic (un)certainty

Will the Swine Flu infect more than one-tenth of one percent of the world’s population in 2009?

6 sqiman

There are an estimated 100,000+ cases in the U.S., 30,000+ in the UK, and today Australia’s top doc said 2 million of that country’s people could be affected. With the way it’s taken off in the Southern Hemisphere in the past few days, this seems totally possible.

posted 19 weeks ago

38 bookie

I note that 6.7 million is more than 3 times as many as the 20 lakh figure. If vaccinations do start soon, it is possible that - unless we see significant spread in Asia or Africa - this pandemic may yet be halted or slowed. Worth a punt on no...

posted 12 weeks ago

Utter(bookie, Possible, m41529) Bet(bookie, 25, m41529)
Utter(sqlman, Possible, m41529) Bet(sqlman, 10000, m41529)
Utter(tuff_sledding, sure, m25339) Bet(sqlman, 500000, m25339)

...
We’ll soon transition from Alchemy to UMass Amherst’s FACTORIE for its expressivity and efficiency.

Ongoing collaboration with UMass Lowell to bring pragmatic inference on board the Lowell ATRV-Jr.

Future data collection using Penn’s Pragbot 2.0, which elicits more realistic expressions concerning location and field of vision (among other things) than Pragbot 1.0 does.
Immediate plans

1. Additional experiments on resolving underspecification, using the richer Pragbot 2.0 environment for data collection.
2. Additional experiments using explicit and implicit prosody to infer what the presumed QUD structure is like.
3. Richer modeling of intermediate QUDs in task oriented dialogue.
4. Fuller integration of general pragmatic pressures with lexical information.
5. Weight setting and evaluation with real users.
References I

Kok, Stanley, Marc Sumner, Matthew Richardson, Parag Singla, Hoifung Poon & Pedro Domingos. 2006. The Alchemy system for statistical relational AI. Tech. rep., Department of Computer Science and Engineering, University of Washington, Seattle, WA.
References II