Computational Semantics and Pragmatics

Autumn 2014



Raquel Fernández
Institute for Logic, Language & Computation
University of Amsterdam

Outline

Today:

- Dialogue acts

Next week:

- Grounding

Some key units of analysis

- Turns: stretches of speech by one speaker bounded by that speaker's silence – that is, bounded either by a pause in the dialogue or by speech by someone else.
- Utterances: units of speech delimited by prosodic boundaries (such as boundary tones or pauses) that form *intentional units* – that is, that can be analysed as an action performed with the intention of achieving something.
- Dialogue acts: intuitively, conversations are made up of sequences of actions such as *questioning, acknowledging*,... a notion rooted in *speech act theory*.

Speech Act Theory

Initiated by Austin and developed by Searle in the 60s-70s within philosophy of language.

Speech act theory grows out of the following observations:

- Typically, the meaning of a sentence is taken to be its truth value.
- There are utterances for which it doesn't makes sense to say whether they are true or false, e.g., (2)-(5):

(1) The director bought a new car this year.
(2) I apologize for being late.
(3) I promise to come to your talk tomorrow afternoon.
(4) Put the car in the garage, please.
(5) Is she a vegetarian?

- These (and generally all) utterances serve to perform actions.
- This is an aspect of meaning that cannot be captured in terms of truth-conditional semantics.

Types of Acts

What are exactly the actions that are preformed by utterances?
Austin identifies three types of acts that are performed simultaneously:

- locutionary act: basic act of speaking, of uttering a linguistic expression with a particular phonetics/phonology, morphology, syntax, and semantics.
- illocutionary act: the kind of action the speaker intends to accomplish, e.g. *blaming, asking, thanking, joking,...*
  - these functions are commonly referred to as the illocutionary force of an utterance $\rightsquigarrow$ its speech act.
- perlocutionary act: the act(s) that derive from the locution and illocution of an utterance (effects produced on the audience).

Relations between Acts

Locutionary vs. illocutionary acts:

- The same locutionary act can have different illocutionary forces in different contexts:

> The gun is loaded $\rightsquigarrow$ *threatening? warning? explaining?*

- Conversely, the same illocutionary act can be realised by different locutionary acts:

> Three different ways of carrying out the speech act of requesting:
>
> (6) A day return ticket to Utrecht, please.
>
> (7) Can I have a day return ticket to Utrecht, please?
>
> (8) I'd like a day return ticket to Utrecht.

Illocutionary vs. Perlocutionary acts:

- Illocutionary acts are intended by the speaker and are under the speaker's full control.
- Perlocutionary acts are not always intended and are not under the speaker's control.

Types of Illocutionary Acts

Searle distinguished between five basic types of speech acts:

- Representatives: the speaker is committed to the truth of the expressed proposition (assert, inform)
- Directives: the speaker intends to ellicit a particular action from the hearer (request, order, advice)
- Commissives: the speaker is committed to some future action (promise, oaths, vows)
- Expressives: the speaker expresses an attitude or emotion towards the proposition (congratulations, excuses, thanks)
- Declarations: the speaker changes the reality in accord with the proposition of the declaration (provided certain conventions hold), e.g. baptisms, pronouncing someone guilty.

Felicity Conditions

Speech acts are characterised in terms of felicity conditions (rather than truth conditions): conditions under which utterances can be used to properly perform actions (specifications of appropriate use).

Searle identifies four types of felicity conditions (Speaker, Hearer):

Conditions	REQUESTING	PROMISING
propositional content	S intends future act A by H	S intends future act A by S
preparatory	a) S believes H can do A b) It isn't obvious that H would do A without being asked	a) S believes H wants S doing A b) It isn't obvious that S would do A in the normal course of events
sincerity	S wants H to do A	S intends to do A
essential	The utterance counts as an attempt to get H to do A	The utterance counts as an undertaking to do A

These conditions can be seen as dimensions on which a speech act can go wrong, but also as constitutive of particular speech acts.

Beyond Speech Acts

Speech act theory was developed by philosophers of lanauge (Austin 1962, Searle 1975) $\rightsquigarrow$ their methodology forgoes looking at actual dialogues.

Empirical traditions that have also shaped current dialogue research:

- Conversation Analysis (sociology): Sachs, Schegloff, Jefferson
- Joint Action models (cognitive psychology): Clark, Brennan, ...

Speech act theory focusses on the intentions of the speaker. But a dialogue is not simply a sequence of actions each performed by individual speakers.

- Dialogue is a joint action that requires coordination amongst participants (like playing a duet, dancing a waltz)
  - many actions in dialogue serve to manage the interaction itself
  - they are overlooked by speech act theory
- There are regular patterns of actions that co-occur together

Adjecency Pairs

Certain patterns of dialogue acts are recurrent across conversations

question – answer
proposal – accetance / rejection / counterproposal
greeting – greeting

Adjacency pairs (term from Conversation Analysis)

- pairs of dialogue act types uttered by different speakers that frequently co-occur in a particular order
- the key idea is not strict adjacency but *expectation*.
  - given the first part of a pair, the second part is immediately relevant and expected
  - any intervening material is perceived as an *insertion sequence* or a *sub-dialogue*

```
Waitress:  What'll ya have girls?
Customer:  What's the soup of the day?
Waitress:  Clam chowder.
Customer:  I'll have a bowl of clam chowder and a salad.
```

The Joint Action Model

Also called collaborative model or grounding model.

- Clark & Schaefer (1989) put forward a model of dialogue interaction that sees conversation as a joint process, requiring actions by speakers and addressees.
- Conversation is a continuos process of establishing common ground between speaker and addressee $\Rightarrow$ grounding
- Speakers and addressees have mutual responsibility in managing the grounding process and making communication successful.
- This structures the dialogue into contributions:
  - each contribution to dialogue is made up of a presentation phase and an acceptance phase.
- More on Clark's model and grounding next week.

Clark & Schaefer (1989) Contributing to discourse. *Cognitive Science*, 13:259–294.

Clark (1996) *Using Language*. Cambridge University Press.

A trascript fragment from the Switchboard corpus:

```
B.52 utt1:  Yeah, /
B.52 utt2:  [it's,+ it's] fun getting together with immediate family.  /
B.52 utt3:  A lot of my cousins are real close /
B.52 utt4:  {C and} we always get together during holidays and
             weddings and stuff like that, /
A.53 utt1:  {F Uh, } those are the ones that are in Texas?  /
B.54 utt1:  # {F Uh, } no, # /
A.55 utt1:  # {C Or } you # go to Indiana on that?  /
B.56 utt1:  the ones in Indiana, /
B.56 utt2:  uh-huh.  /
A.57 utt1:  Uh-huh, /
A.57 utt2:  where in Indiana?  /
B.58 utt1:  Lafayette.  /
A.59 utt1:  Lafayette, I don't know where, /
A.59 utt2:  I used to live in Indianapolis.  /
B.60 utt1:  Yeah, /
B.60 utt2:  it's a little north of Indianapolis, about an hour.  /
```

From Speech Acts to Dialogue Acts

The concept of dialogue act (DA) extends the notion of speech act to incorporate ideas from conversation analysis and joint action models of dialogue.

It is the term favoured within computational linguistics to refer to the function or the role of an utterance within a dialogue.

- Taxonomies of DAs aim to cover a broader range of utterance functions than traditional speech act types
  - importantly, they include grounding-related DAs (meta-communicative).
- They aim to be effective as tagsets for annotating dialogue corpora.

Dialogue Act Taxonomies: DAMSL

One of the most influential DA taxonomies is the DAMSL schema (Dialogue Act Markup in Several Layers) by Core & Allen (1997).

- Communicative Status
- Information Level
- Forward-looking Function
- Backward-looking Function

Explore the annotation manual:

http://www.cs.rochester.edu/research/speech/damsl/RevisedManual/RevisedManual.html

Utterances can perform several functions at once: possibly one tag per layer.

The taxonomy is meant to be general but not totally domain independent $\rightsquigarrow$ it has been adapted to several types of dialogue.

DA Taxonomies: SWBD DAMSL

The SWBD DAMSL schema is a version of DAMSL created to annotated the Switchboard corpus. Here are the 18 most frequent DA in the corpus:

Tag	Example	Count	%
Statement	*Me, I'm in the legal department.*	72,824	36%
Continuer	*Uh-huh.*	37,096	19%
Opinion	*I think it's great*	25,197	13%
Agree/Accept	*That's exactly it.*	10,820	5%
Abandoned/Turn-Exit	*So, -/*	10,569	5%
Appreciation	*I can imagine.*	4,633	2%
Yes-No-Question	*Do you have to have any special training*	4,624	2%
Non-verbal	*<Laughter>,<Throat_clearing>*	3,548	2%
Yes answers	*Yes.*	2,934	1%
Conventional-closing	*Well, it's been nice talking to you.*	2,486	1%
Uninterpretable	*But, uh, yeah*	2,158	1%
Wh-Question	*Well, how old are you?*	1,911	1%
No answers	*No.*	1,340	1%
Response Ack	*Oh, okay.*	1,277	1%
Hedge	*I don't know if I'm making any sense*	1,182	1%
Declarative Question	*So you can afford to get a house?*	1,174	1%
Other	*Well give me a break, you know.*	1,074	1%
Backchannel-Question	*Is that right?*	1,019	1%

The average conversation consists of 144 turns, 271 utterances, and took 28 min. to annotate. The inter-annotator agreement was 84% (κ=.80).

http://www.stanford.edu/~jurafsky/manual.august1.html

DA Recognition

There isn't a one-to-one relation between a locutionary act and its illocutionary force. How can we derive the dialogue act performed by an utterance?

Two computational models of the interpretation of dialogue acts:

- Inferential plan-based models: based on epistemic logic (beliefs, desires, and intentions - BDI); use of logical inference to reason about the speaker's intentions - builds directly on speech act theory.
- Probabilistic cue-based models: the surface form of the sentence is seen as a set of cues to the speaker's intentions; use of probabilistic machine learning models.

Both models use a kind of inference: the hearer infers something that was not contained directly in the semantics of the utterance.

Daniel Jurafsky (2004) Pragmatics and Computational Linguistics. *Handbook of Pragmatics*. Oxford: Blackwell.

Inferential Plan-Based Models

The BDI approach is meant to be a general model of rational action that can be applied to conversation. It proposes an axiomatization of BDIs to account for

- what motivates our actions
- how to understand actions by others

It aims to explain indirect speech acts

(9) Can you pass me the salt?

⇝ Literal speech act [literal force hypothesis]: *yes-no question*
⇝ Indirect speech act after an inference chain: *request* (pass me the salt)

and also, for instance, answers that appear to be overinformative:

(10) Customer: When does the train to Montrteal leave?
Clerk: At 3:15 at gate 7.

⇝ the clerk recognises the plan of the customer and identifies possible obstacles and relevant information to solve them

Can you pass me the salt?

A BDI system can interpret the correct speech act by simulating an inference chain along the following lines, as suggested by Searle:

1. X has asked me a question about whether I have the ability to pass her the salt.
2. I assume that X is being cooperative in the conversation (in the Gricean sense) and that her utterance therefore has some aim.
3. X knows I have the ability to pass her the salt, and there is no alternative reason why X should have a purely theoretical interest in my ability.
4. Therefore X's utterance probably has some ulterior illocutionary point. What can it be?
5. A preparatory condition for a directive is that the hearer have the ability to perform the directed action.
6. Therefore X has asked me a question about my preparedness for the action of passing X the salt.
7. Furthermore, X and I are in a conversational situation in which passing the salt is a common and expected activity.
8. Therefore, in the absence of any other plausible illocutionary act, X is probably requesting me to pass her the salt.

Main influences of these approaches:

- Austin's and Searle's characterisation of speech acts in terms of felicity conditions that appeal to the mental attitudes of speakers (rationality & cooperativity à la Grice)
- Hintikka's logic of belief (axiomatisation of BDI and action/planning)

BDI Approaches

For more details on the BDI axiomatization and the plan-inference rules see Jurafsky (2004) for a short summary and the original papers by Allen et al.

Jurafsky (2004) Pragmatics and Computational Linguistics. *Handbook of Pragmatics*. Oxford: Blackwell.
Allen & Perrault (1980) Analyzing Intention in Utterances, *Artificial Intelligence* 15(3).
Perrault & Allen (1980) A Plan-based Analysis of Indirect Speech Acts, *Computational Linguistics* 6(3):167-182.

BDI approaches have been used as the basis to implement conversational agents in the TRAINS/TRIPS projects.

- see the project's website for access to a dialogue corpus collected to develop the system, movies of the system in action, and links to publications. http://www.cs.rochester.edu/research/trains/

Allen et al. (2001) Towards Conversational Human-Computer Interaction, *AI Magazine*.
Allen et al. (2001) An architecture for more realistic conversational systems, in *Proc. of Intelligent User Interfaces*.

DA Recognition: Probabilistic Model

The most common approach in computational linguistics is to use a probabilistic cue-based model:

- the listener uses cues in the input to infer a particular interpretation.
- use of several sources of knowledge: lexical, collocational, syntactic, prosodic, conversational-structure (the *micro-grammar* of each DA)

- Lexical and Syntactic Cues: words/phrases that occur more often in particular DAs. presence of particular words, such as *'please'* (requests), word order (questions), tag particle *'right?'* in final position (declarative questions or checks)
- Prosodic Cues: final pitch rise (polar questions and declarative questions); loudness or stress can help distinguish *'yeah'* agreement from backchannel.
- Conversational Structure Cues: *'No it isn't'* is an agreement after *'It isn't raining'* and a disagreement after *'It is raining'*. *'yeah'* is more likely to be an agreement after a proposal. ($\rightsquigarrow$ adjacency pairs)

Cue-based Algorithms

Cue-based models are supervised machine learning algorithms trained on a dialogue corpus hand-labeled with DAs.

Typically, a statistical classifier is trained for each particular type of DA. The classifier learns to recognise the combination of features that suggests the presence of a question, an assessment, an inform...

Example: Given the observed cues c, the goal is to find the DA d^* that has the maximum posterior probability $P(d|c)$ given those cues.

$$d^* = \underset{d}{\operatorname{argmax}}\, P(d|c)$$

$$= \underset{d}{\operatorname{argmax}}\, P(d)P(c|d)$$

We need to choose the DA that maximises the product of two probabilities: the prior probability of a DA $P(d)$ and the likelihood $P(c|d)$ of observing a particular combination of features when a particular DA is present.

Different machine learning algorithms can be used (HMMs are common).

Some References

Shriberg et al. (1998) Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech? *Language and Speech*, 41:439-487.

Stolcke et al. (2000) Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech, *Computational Linguistics*, 26(3).

Keizer et al. (2002) Dialogue act recognition with Bayesian networks for Dutch dialogues. *Proc. SIGdial*

Klüwer et al. (2010) Using Syntactic and Semantic based Relations for Dialogue Act Recognition, *Proc. COLING*

Cuayáhuitl et al. (2013) Impact of ASR N-Best Information on Bayesian Dialogue Act Recognition. *Proc. SIGdial*

Summing Up

- Speech act theory: truth-conditional content falls short of characterising the role utterance play in conversation. Utterances are actions, with certain *felicity conditions*.
- Conversation analysis / joint action models: we should actually look beyond individual speech acts and embrace the fact that conversations involve multiple participants performing *joint actions* (adjacency pairs, contributions: presentation/response)
- The notion *dialogue act* extends the notion of speech act to incorporate ideas from CA and joint action models.
- DA taxonomies provide inventories of dialogue act types that aim to be suitable for dialogue corpora annotation.
- Two main approaches to DA recognition: logic-based models vs. probabilistic models.