

Human-like dialogue: Key challenges for AI

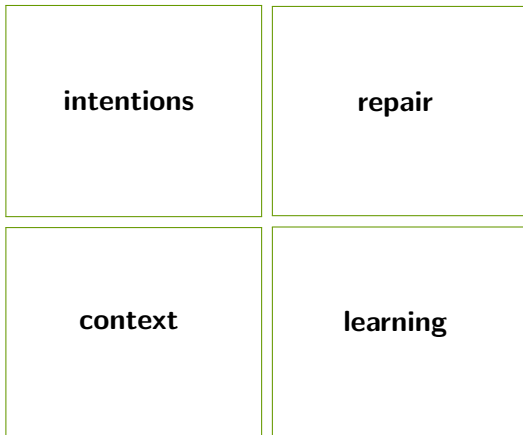
Raquel Fernández

Institute for Logic, Language & Computation
University of Amsterdam



MAIN @ NIPS 2016

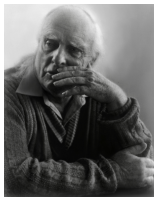
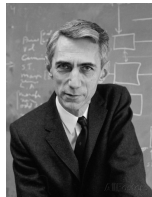
4 key features of human dialogue



intentions

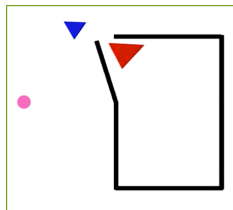
Two views of communication:

Shannon (1948): communication as information encoded by a sender and decoded by a recipient.



Grice (1957): human communication is characterised by the process of **intention recognition**

intentions

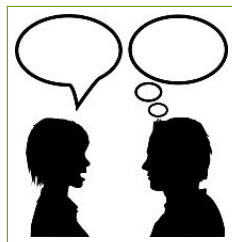


Heider & Simmel (1944): An experimental study in apparent behavior

- tendency to ascribe **goals** and **intentions**
- theory of mind / attribution of causation

Linguistic actions also trigger the attribution of mental attitudes and goals

- recognising type of **dialogue game**



↪ *From pattern recognition to intention recognition*

repair

Human conversations are not always flawless ...

non-native and native speakers

A: and they have the *chwach* there

B: the *what*?

A: the *chwach* – I know someone that

B: what does it mean?

A: like um like American people they
always go there every Sunday

B: oh to church – I see

maptask dialogue

G: see where the fenced meadow is
besi– on the left-hand side of the *r*–
uh the left-hand side of your page

F: no m– mine's

G: about halfway up

F: no that's not there on mine

G: is it not?

... but we are equipped with **flexible and robust repair mechanisms**

- self- and other- monitoring and repair
- fine-grained and targeted
- coordinated trade-offs to reduce joint effort

⇒ *Monitoring and repair system in production and comprehension*

Coherence in dialogue hinges on intricate linguistic phenomena:

- the discourse context is complex and structured
- not only discourse referents: also *questions under discussion*, *discourse commitments* ...
- long distance dependencies
- reasoning with discourse context and previous assumptions

British national corpus

J: I remember once I roasted our piece.
I roasted it and we couldn't eat it on the Sunday and

A: Could not? Why could you not eat it?

J: That was bull beef.

A: Oh right.

H: our second class beef, you see.

J: Then I, I put it in a saucepan and I stewed it the next day

↪ *Keeping track of the dialogue gameboard (context / state)*

learning

Widely accepted characteristics of human language learning:

- humans do receive a lot of input (no need for innate universal grammar)
- but not in a training / testing setting
- by and large, we learn through interaction, from few exposures

Two more points:

(1) learning is not (fully) unsupervised – implicit supervision

CHI: you're good to sharing.
MOT: I'm good at sharing?

CHI: I'm read it.
DAD: you read it to mummy.

↪ *Ability to provide, recognise, and exploit feedback is critical*

↪ *Evidence of communicative success or lack thereof*

learning

(2) learning is 'context-dependent'

- reciprocity: interacting agents learn to communicate *with each other*
- different degrees of stability of what is 'learned'
 - ▶ specific dialogue, interlocutor, sub-community, ...

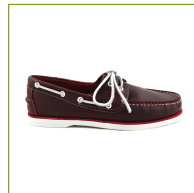
conceptual pact (Brennan & Clark, 1996)

A: A docksider.

B: A what? Is that a kind of dog?

A: No, it's a kind of um leather shoe,
kinda pennyloafer.

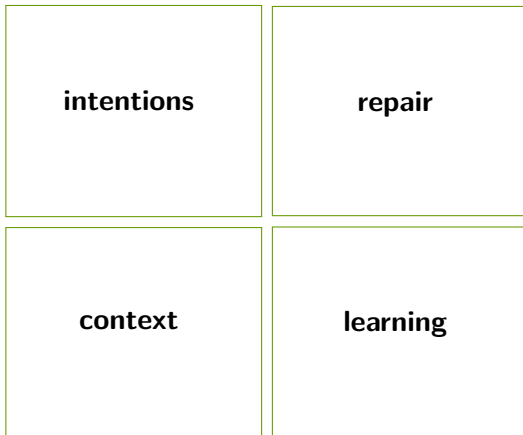
B: Okay, okay, got it.



↪ *Our ways of communicating are determined by the common ground we share with our interlocutors*

↪ *Learning is not independent from these processes*

summing up



summing up: 4 key features of human dialogue

- Beyond pattern recognition: recognising **intentions**
- Flexible, precise, and robust **repair** mechanisms
- Keeping track of dialogue **context** to capture coherence
- Online incremental **learning**:
 - ▶ via supervision / feedback, at least to some extent
 - ▶ indexed by common ground with interlocutors