Dialogue and Conversational Agents

Speaker: Sunjing Ji
C₁: ...I need to travel in May.
A₁: And, what day in May did you want to travel?
C₂: OK uh I need to be there for a meeting that’s from the 12th to the 15th.
A₂: And you’re flying into what city?
C₃: Seattle.
A₃: And what time would you like to leave Pittsburgh?
C₄: Uh hmm I don’t think there’s many options for non-stop.
A₄: Right. There’s three non-stops today.
C₅: What are they?
A₅: The first one departs PGH at 10:00am arrives Seattle at 12:05 their time. The second flight departs PGH at 5:55pm, arrives Seattle at 8pm. And the last flight departs PGH at 8:15pm arrives Seattle at 10:28pm.
C₆: OK I’ll take the 5ish flight on the night before on the 11th.
C₇: OK.

Figure 24.4  Part of a conversation between a travel agent (A) and client (C).
How is conversation different from Monologue?

- Turn-taking
- adjacency pairs
- significant silence

- Utterance
- shorter
- single clauses
- pronoun as subjects

- Grounding
- continuer
-backchannel
- acknowledgement token (such as Mm hmm)

- repetition
- continued attention
- request for repair
How to construct a machine system that can understand and talk in such conversation?
Plan-Inferential Interpretation of Dialogue Acts

- BDI (Belief, Desire and Intention) models: Allen, Cohen, Perrault etc.

- Predicate Calculus
  1. $B(S,P)$—”S believes the proposition P”
  2. $B(A,P) \land B(A,Q) \Rightarrow B(A, P \land Q)$
  3. $\text{KNOW}(S,P)$—”S truly belief the proposition P”
  4. $W(S,P)$—”S wants P to be true”
  5. $W(S,ACT(H))$—”S wants H to do ACT”

- BDI: action schemas
- a set of parameters with constraints about the type of each variable
Agent A booking Flight F1 for Client C

**BOOK-FLIGHT(A,C,F):**

- **Constraints:** Agent(A) ∧ Flight(F) ∧ Client(C)
- **Precondition:** Know(A, depart-date(F)) ∧ Know(A, depart-time(F)) ∧ Know(A, origin(F)) ∧ Know(A, flight-type(F)) ∧ Know(A, destination(F)) ∧ Has-Seats(F) ∧ W(C, BOOK(A,C,F)) ∧ ...
- **Effect:** Flight-Booked(A,C,F)
- **Body:** Make-Reservation(A,F,C)
Encoding Dialogue Act

The speaker wants the hearer to know something

\textbf{INFORM}(S,H,P):

- **Constraints:** Speaker(S) \land Hearer(H) \land Proposition(P)
- **Precondition:** Know(S,P) \land W(S, INFORM(S, H, P))
- **Effect:** Know(H,P)
- **Body:** B(H,W(S,Know(H,P)))

The speaker requests the hearer to perform some

\textbf{REQUEST}(S,H,ACT):

- **Constraints:** Speaker(S) \land Hearer(H) \land ACT(A) \land H is agent of ACT
- **Precondition:** W(S,ACT(H))
- **Effect:** W(H,ACT(H))
- **Body:** B(H,W(S,ACT(H)))
An example

C₁: I need to travel in May.
A₁: And, what day in May did you want to travel?

• How does a plan-based agent know to ask A₁?
• From Book-Flight Plan
• REQUEST-INFO to represent a plan to ask for the rest information
Cue-based Interpretation of Dialogue Acts

• What types of cues can be used?

1. Words and collocations
   “please” or “would you”— cue for REQUEST;
   “are you”— cue for YES-NO-QUESTIONs

2. Prosody
   “rising pitch”— cue for YES-NO-QUESTIONs
   “loudness of yeah”— cue for AGREEMENT (not BACKCHANNEL)

3. Conversational Structure
   “yeah following a proposal is an AGREEMENT”
   “yeah following an INFORM is probably a BACKCHANNEL”
A simple N-gram Approach

• These cues can be captured by training a separate word-N-gram grammar for each dialogue act
• Given an input $u$ consisting of a sequence of words $W$, the systems then choose the dialogue act $d$ whose N-gram grammar assigns the highest likelihood to $W$

$$d^* = \arg\max P( d|W) = \arg\max P(d) P(W|d)$$

d: dialogue act
$W$: sequence of words