Introduction to dialogue modelling

Milica Gašić

Dialogue Systems and Machine Learning Group,
Heinrich Heine University Düsseldorf

Course

- 13 1.5h lectures
- 12 45min practical sessions
- Relevant information regarding assessment you will find on ILIAS

In this course…

- ▶ Introduction to dialogue modelling
- ▶ The space of meanings
- ▶ Spoken language understanding
- ▶ Dialogue state tracking and dialogue belief tracking
- ▶ Dialogue policy optimisation
- ▶ Natural language generation
- ▶ End-to-end dialogue models
- ▶ Further topics in dialogue modelling

In this lecture...

Properties of human-computer dialogue

Basic dialogue modelling concepts

Traditional approaches to dialogue modelling

Machine learning approaches to dialogue modelling

Architecture of a spoken dialogue system

What is a dialogue system?

- A dialogue system is a computer system that enables human computer interaction where primary input is natural language.
- A spoken dialogue system is a computer system that enables human computer interaction where primary input is speech.
- Speech does not need to be the only input. We can interact with machines also using touch, gesture or facial expressions and these are multi-modal dialogue systems.

Examples from popular culture

Examples from every-day life

Personal assistants

- Most commonly used dialogue systems are personal assistants such as Siri, Cortana, Google Assistant and Alexa
- These are server-based accessed via a range of devices: smart-phones, tablets, laptops, watches and specialist devices such as Amazon Echo (Alexa).

What kind of dialogues there are?

- Chat (▸ Eliza)
- Social dialogue
- Goal-oriented dialogue - there is a goal or several goals that must be fulfilled during conversation (▸ Medical Bayesian Kiosk)
- Tutorial dialogue

Properties of human-computer dialogue

Properties

What constitutes a spoken dialogue system?

- ▶ Being able to understand the user
- ▶ Responding sensibly and intelligently
- ▶ Handling a variety of topics, large and expanding domains
- ▶ Understanding emotions
- ▶ Being human-like
- ▶ Having variability in behaviour
- ▶ Providing personalised outputs
- ▶ ...

Core Properties

- Being able to understand the user
- Being able to decide what to say back
- Being able to conduct a conversation beyond simple voice commands or question answering

Basic dialogue modelling concepts

Scope or domain of a dialogue systems

In goal-oriented dialogues we typically assume that the conversation belongs to a particular domain. The domain is characterised by

ontology a database that defines properties of entities that a dialogue system can talk about

Ontology can be more complex than that. In the most extreme case it is a knowledge graph consisting of millions of entities and billions of relationships between the entities. In this case we consider the dialogue system to be **open domain**, otherwise it is **limited domain**.

Initiative in dialogue

system-initiative vs user-initiative who takes the initiative in the dialogue:

- System: *Hello. Please tell me your date of birth using the six digit format.*
- System: *Hello, how may I help you?*

Turn-taking in dialogue – Who speaks when?

Dialogue can be described in terms of system and user turns

- ► System: *How may I help you?*
- ► User: *I'm looking for a restaurant*
- ► System: *What kind of food would you like?*
- ► ...

Turn taking can be more complex and characterised by

barge-ins System: *How may I...* User: *I'm looking for a restaurant*

back channels User: *I'm looking for a restaurant* [System: *uhuh*] *in the centre of town*

Dialogue acts

One simple dialogue act formalism would consist of

dialogue act type - encodes the system or the user intention in a (part of) dialogue turn

semantic slots and values - further describe entities from the ontology that a dialogue turn refers to

Is there um maybe a cheap place in the centre of town please?

inform (price = cheap, area = centre)

dialogue act type   **semantics slots and values**

Dialogue acts

Dialogue act formalism describes meaning encoded in each dialogue turn [Traum, 2000].

- ▶ Relation to ontology
- ▶ Encode intention of the speaker
- ▶ Relation to logic
- ▶ Context
- ▶ Partial information from speech recognition (primitive dialogue acts)

Grounding in dialogue

- Dialogue is an interactive speech and language processing problem
- Two-parties in conversation
- Grounding is the notion of how much they share the common understanding [Clark and Brennan, 1991]
- Think of a dialogue between flight control and pilot

Implicature in dialogue

- A: *Which days do you need to be in the office?*
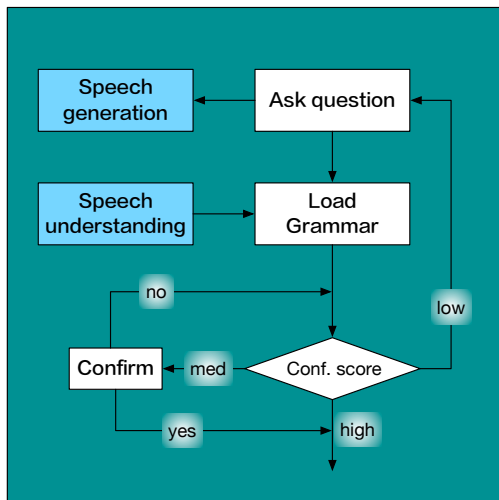- B: *There is an in-person group meeting every Friday.*

Notice that B does not directly answer the question of A, but A should be able to infer that B needs to be in the office every Friday. This is **conversational implicature**.
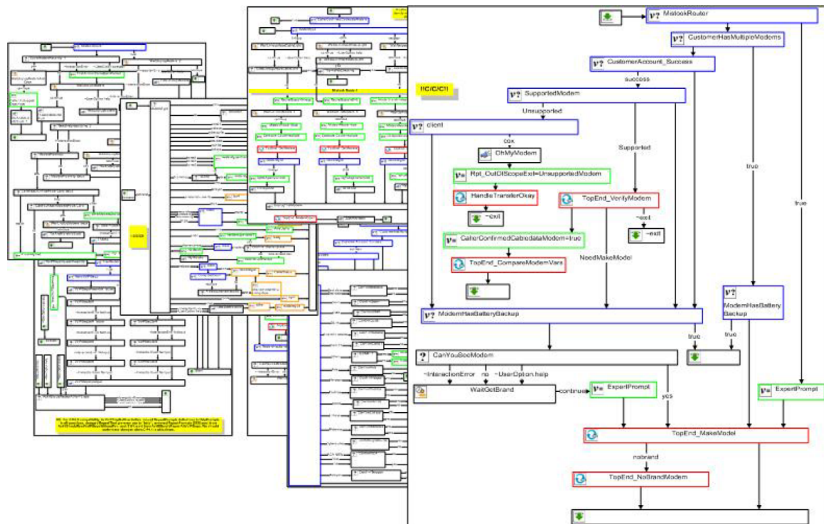
Traditional approaches to dialogue modelling

Traditional approach to dialogue systems – Call flow

What are you looking for a restaurant or a hotel?

- restaurant → What kind of food?
- hotel → How many stars?

What kind of food? — Chinese → What area?

How many stars? — three → What area?

What area? — center →
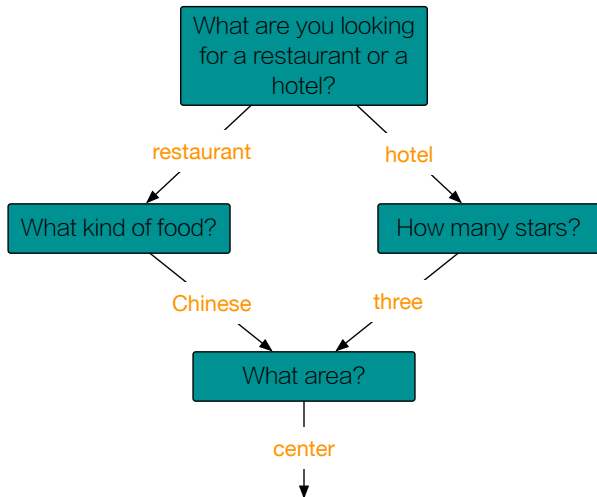
Node processing in a call flow

Part of a deployed call-flow [Paek and Pieraccini, 2008]

Problems

What breaks dialogue systems?

- ▶ Speech recognition errors
- ▶ Not keeping track of what happened previously
- ▶ Need to hand-craft a large number of rules
- ▶ Poor decisions
- ▶ Repetitive behaviour
- ▶ User's request is not supported
- ▶ ...

# Machine learning approaches to dialogue modelling

Dialogue as a core AI problem

- ▶ Turing posed dialogue as a core AI problem [Turing, 1950]
- ▶ Turing test: Are we talking to a machine or a human?
- ▶ Dialogue is hard: infinite possible trajectories of user input and system output
- ▶ We can always think of a dialogue that was never produced before

Dialogue is an AI complete problem

- A problem is considered AI complete if solving it solves a general problem or producing a generally intelligent program [Shapiro, 1992].
- AI complete problems include natural language processing (NLP), speech processing, computer vision, problem solving and search, knowledge representation and reasoning, machine learning, robotics.
- Dialogue encompasses a number of NLP tasks and is therefore also AI complete.

Machine learning in spoken dialogue systems

Data
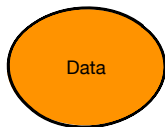
- Dialogues
- Labelled user intents
- Transcribed speech

Model
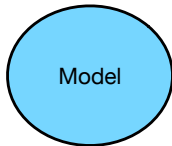
- Regression
- Classification
- Markov decision process
- Neural networks

Predictions

- What the user wants
- What is the best response
- How to formulate the response

Architecture of a spoken dialogue system

Modular approach

- ▶ Divide and conquer
- ▶ Divide the task in sub-tasks
- ▶ Each sub-task has a well defined input and a well-defined output
- ▶ If we have labeled data with inputs and outputs we can apply machine learning to model sub-tasks

Modules

- ▶ Spoken language understanding (SLU)
- ▶ Dialogue management (DM)
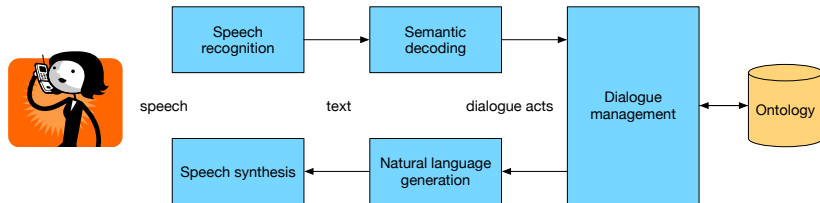- ▶ Spoken language generation (SLG)

Modular approach
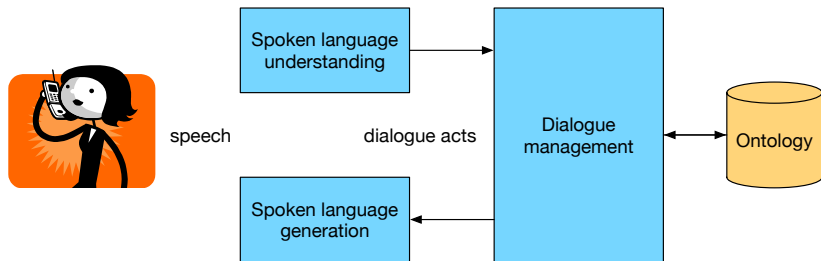
Modules

- Spoken language understanding is further decomposed in
  - Speech recognition
  - Semantic decoding
- Spoken language generation is further decomposed in
  - Natural language generation
  - Speech synthesis

Modular approach

Both semantic decoding and natural language generation depend on the ontology. Dialogue manager in addition needs to query the ontology to be able to provide an adequate response.
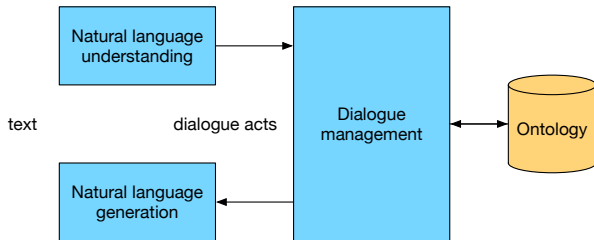
Dialogue system

- If the dialogue input is simply text than we have a dialogue system
- In this case we refer to the first module as Natural language understanding

Modular approach

Components of dialogue management

- Components of dialogue management address core dialogue properties.
- Dialogue state tracker (DST) remembers what has happened in the conversation so far.
- Dialogue policy (DP) is the decision maker, it conducts conversation.

Components of dialogue management

CC by Free Clip Art

Natural language understanding

Dialogue state tracking

Ontology

text

dialogue acts

dialogue states

Natural language generation

Dialogue policy
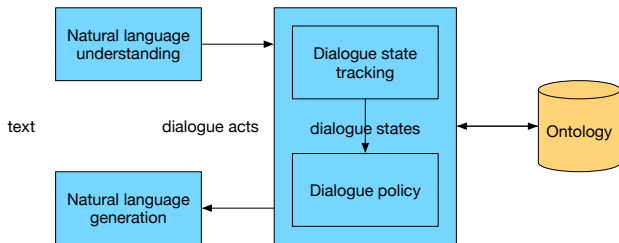
Summary

- Dialogue is one of the most natural means for humans to exchange information.
- Dialogue modelling is hard, it is one of core AI problems.
- Basic dialogue modelling concepts include: scope (domain), initiative, turn, dialogue acts, grounding and implicature.
- Hand-coding dialogue system behaviour is sub-optimal, instead we utilise machine learning for dialogue modelling.
- This typically involves adopting a divide and conquer approach and dividing the task in sub-tasks that can more easily be modelled using machine learning.

Next lecture

- The space of meanings

References I

📄 Clark, H. and Brennan, S. (1991).
Grounding in Communication.
APA.

📄 Paek, T. and Pieraccini, R. (2008).
Automating spoken dialogue management design using machine learning: An industry perspective.
Speech Commun., 50(8-9):716–729.

📄 Shapiro, S. C. (1992).
Encyclopedia of Artificial Intelligence.
John Wiley & Sons, Inc., USA, 2nd edition.

📄 Traum, D. R. (2000).
20 questions on dialogue act taxonomies.
JOURNAL OF SEMANTICS, 17:7–30.

References II

📄 Turing, A. (1950).
Computing machinery and intelligence.
Mind, LIX(236):433–460.