Cooperative Response Generation in Dialogue

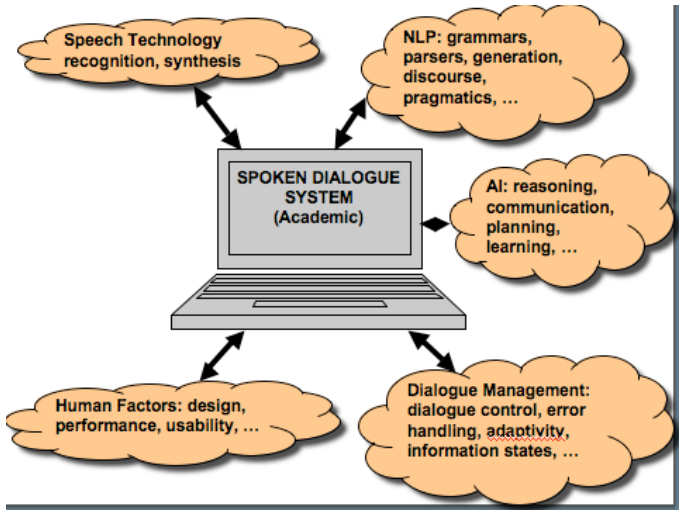
Introduction

Ivana Kruijff-Korbayova

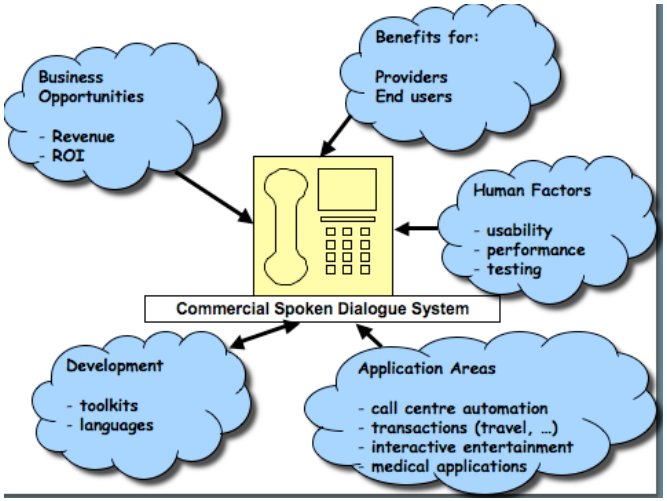
Dialogue Systems

- Emerging interdisciplinary area since the early 1990s
  - integration of speech technology, natural language processing, AI, dialogue / communication theory, human factors, …
  - scientific / academic – based research
  - commercially driven R&D
  - achievements and challenges

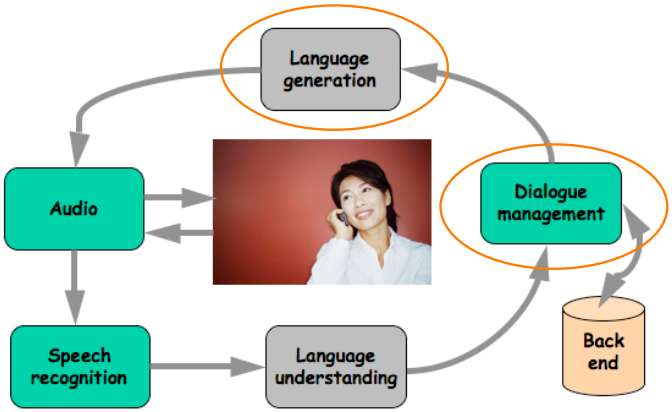
Dialogues System Research



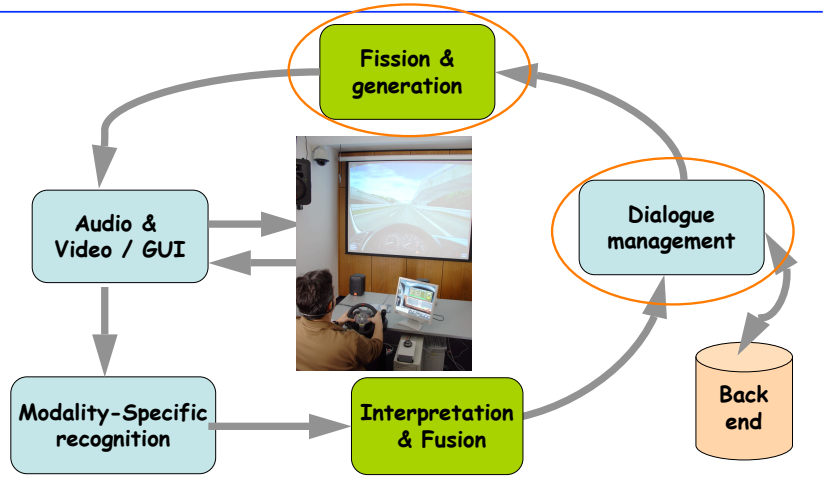
Dialogue System Industry



Typical Pipeline Architecture

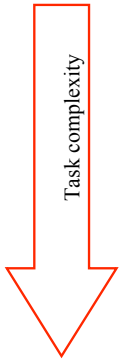


Typical Pipeline Architecture (Multimodal)



Dialogue Management

- Finite state systems
  - Sequence of predefined steps (dialogue script)
- Frame-based systems (form-filling)
  - Task represented as a set of slots to fill (frame, template)
- Agent-based systems
  - Joint problem solving by collaborating agents



Output Generation

- Canned text
- Template-based
- Concept-to-text/speech ("deep generation")
  - Content selection
  - Utterance planning
  - Surface realization
  - Speech synthesis

Key Issues

- Collaboration
  - Gricean maxims
- Initiative
- Grounding and error recovery

Collaboration

- Communication is a joint activity: agents collaborate to establish and achieve their goals
- Cooperative Principle (Grice)
  - Make your contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged
    - Maxims of Conversation
      - Maxim of quality
      - Maxim of quantity
      - Maxim of relevance
      - Maxim of manner
- Neither agent can accomplish the task alone
  --> joint goals, mixed initiative
- Need mutual understanding
  --> grounding

Initiative

- Who is in control of the dialogue progression?
  - Being the one who's talking does not necessarily mean being in control, e.g., just answering a question
  - Dialogue initiative vs. task initiative
- Basically, two models:
  - Fixed initiative model (one participant in control)
    - System-initiative: can drive dialogue as wanted by prompting user, but may be unnatural and inconvenient for user
    - User initiative: can do what wants when wants, but difficult for system, because it doesn't know what is coming
  - Mixed initiative model (either participant can assume initiative, depending on knowledge, skills, situation, etc.)
    - Typical in human-human conversation
    - System needs to decide when/whether to take initiative, e.g.
      - "overanswering"
      - avoiding or recovering from dead-ends

Grounding

(Clark 1996)

- Principle of (joint) closure: Agents performing a (joint) action require (CG) evidence, sufficient for current purposes, that they have succeeded in performing it
- Levels of interpretation: channel, signal, proposition, intention
- The optimal evidence isn't usually the strongest, most economical and most timely evidence possible, for that may be too costly.
- Positive grounding feedback:
  - Continued attention
  - Relevant next contribution
  - Acknowledgement (nod or "continuer", e.g., uh-huh, yeah; or assessment, e.g., that's great)
  - Demonstration (by paraphrasing, reformulating or cooperatively completing)
  - Display (verbatim repetition)

Verification Strategies in Systems

- Immediate explicit feedback (and verification request)
  - S: Where do you want to go?
  - U: Hamburg.
  - S: Traveling to Hamburg. (OK?)
  - U: Yes.
  - S: When do you want to go?
- Immediate "implicit" feedback by incorporating material to be grounded in the next system turn (see if user accepts or protests)
  - S: Where do you want to go?
  - U: Hamburg.
  - S: And when do you want to go to Hamburg?
- Delayed explicit feedback by summarizing at task end
  - …
  - S: So. Traveling from Saarbrücken to Hamburg on Monday June 6
  - …

More Key Issues

- Global and local structure, sub-dialogues
- Style, Personality
- Influence of context
  - Intonation
  - Referring expressions
  - Indirectness
- Alignment
- Adaptivity (tailoring)

Local Coherence

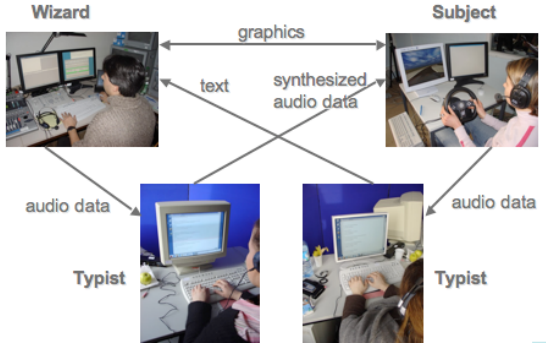
S: What is the patient's sex?
U: Female with severe nipple discharge
S: What is the patient's age?
U: Fifty five
S: Is the discharge bilateral?
U: No

S: What is the patient's sex?
U: Female with severe nipple discharge
S: Is the discharge bilateral?
U: No
S: What is the patient's age?
U: Fifty five

[Milward&Beveridge 2003]

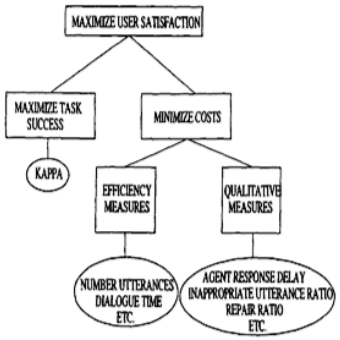
Development Methodology

- Requirement Specification
  - Analysis of human-human dialogues
  - Wizard-of-Oz experiments (simulations) to gather user behavior samples and test design ideas in early stages of development
    - e.g., the TALK project WOZ experiment setup:



Development Methodology

- Usability Evaluation
  - PARADISE framework [Walker et al. 1997]:
    - Maximize user satisfaction through maximizing task success while minimizing dialogue costs
    - User satisfaction (surveys)
    - Objective measures:
      - Task success (in terms of filling a set of slots)
      - Dialogue costs:
        - » Efficiency, e.g., no. of turns and time
        - » qualitative phenomena, e.g., no. of inappropriate utterances or repairs
    - Performance function: relative contribution of objective factors to user satisfaction
  - Questionnaires, questionnaires ….



Deployment Platforms

- PC
  - GoDIS
  - Circuit-Fix-It Shop, TRIPS/TRAINS
  - Autotutor, Why-Atlas, BE&E, PACO …
- Telephone
  - Philips Train Timetable System, Deutsche Bahn info, …
  - It-Spoke weather
- Embedded voice systems
  - HAL (Home Automated Living), D'Homme project
- In-car voice or multimodal systems
  - BMW navigation, TALK project: MP3 player
- PDA, tablet PCs, next generation phones
  - MATCH, SmartKom
- Embodied agents
  - REA, SAM, MRE, …
- Robots
  - WITAS
  - MEL, BIRON, COSY and CogX system, Companions

Applications

- Speech interfaces to devices, e.g., TV, lamps, heating, washer, MP3 player, navigation system, …
- Speech interfaces to databases, e.g., TV, MP3 player, timetable info(train, flight, …), restaurants, movie info, stock-exchange info, soccer results, weather forecast, …
  - Philips, DBahn, ItSpoke Weather, MATCH
- Expert systems / decision support, collaborative agents
  - TRAINS/TRIPS, WITAS
- Educational systems, e.g.,
  - Tutoring language, math, physics, electric circuits, …
    - AutoTutor
  - Communication skills (e.g., story-telling or -listening systems)
    - SAM, LISTEN, MRE
  - Decision skills
    - MRE
- Conversational or entertainment systems
  - MEL, REA, Companions

Key Issues for the Future

- Pervasive systems
  - distributed dialogues: shifts between dialogue situations
  - concurrent dialogues: multitasking (co-ordination, synchronisation, redundancy)
  - interaction model needs to be predominantly event-based (external events, opportunistic)
- Adaptivity:
  - Systems need to be dynamically adaptive in a number of different ways: to the environments in which they are used (modality), to their user's preferences and needs (personalisation), to changes in task and context, to interaction progress.
- Ability to learn:
  - Systems need to be able to learn from interactions with users in order to provide an optimally usable interface that matches the current environment and user.
- Standardization:
  - There is a need for a common set of standards to support re-usability for developers and to support usability for the users of spoken dialogue systems, e.g. constraining vs. open-ended prompts, explicit vs. implicit verification.

Reading

- D. Jurafsky and J. Martin (2000): Speech and Language Processing, Chapters 19 and 20.
- McTear (2002): Spoken Dialogue technology. In ACM Surveys. pp. 1-80.