Lecture 9: Educational Dialogue
(including material from Porayska-Pomsta)

Contents

1. What is dialogue?
2. Why NL dialogue?
3. Tutors' expertise
4. Example Systems

1. What is dialogue?

Most commonly used to mean

- Natural Language dialogue

Different forms of dialogue

- Communication between at least two parties
- Mixed or single initiative:
  - Unrestricted more difficult to handle, because of NLU
- Through natural language:
  - Text or
  - Speech
- Through graphical user interface
  - Buttons, e.g. hint, help, give answer, give definition (Wallis)
  - Direct manipulation interface, e.g. BEETLE, Bob the Builder pipes (www.bobthebuilder.com) etc.
- Multimodal:
  - Text
  - Speech
  - Deictic devices (detected and tracked through, e.g. mouse and keyboard actions)
  - Embodiment (e.g. ECAs)

2. Why NL Dialogue?
(slides based on Moore, 2004)

Effective Learning

Learning occurs when students:

- engage in active generation of knowledge
  - solve problems independently
- encounter obstacles and work around them
- explain to themselves
  - what worked and what did not
  - how new information fits in with what they already know (generalisation)

(Chi et al. 1989, 1994, 2001; Ohlsson & Rees 1991; Van Lehn 1990)

Effective Tutoring

Appropriate guidance with intervention to ensure that errors are detected and repaired

Human tutors maintain a delicate balance:

students do as much of the work as possible

tutors provide just enough guidance to keep students from floundering

⇒ **students maintain a feeling of control and sense of achievement; motivational benefits**

(e.g., Fox 1993; Lepper & Chabay 1988; Merrill et al. 1992; Graesser & Person 1994)

Dialogue-based learning

- Human 1-2-1 tutoring is the most effective form of instruction
- Much of the success of human tutors hinges on their ability to engage students in dialogue
- Natural language dialogue offers indirect techniques for:
  - signalling disagreement or uncertainty, suggesting solutions, etc.
  - switching topic
  - taking or relinquishing initiative
  - eliciting knowledge construction, via techniques such as co-construction of explanations, and directed lines of reasoning

Dialogue-based learning

- Intelligent tutoring systems lead to learning gains that are half that of human tutoring

So…

- Can intelligent tutoring systems be more effective if they engage students in a dialogue?
  - What constitutes effective tutoring?
  - Should we look at human tutors?
    - Not all successful strategies of human tutors may be suitable for computer-based tutoring
    - Not all information available to tutors will be equally available to computer tutors and vice versa.
  - What kind of architectures can support good tutorial dialogue?

Tutor's expertise: the KEY!

- **Diagnosis**: the ability to diagnose the student in a specific context
- **Planning**: the ability to plan their actions based on the diagnoses
- **Action**: the ability to act upon their diagnosis and plans

⇒ **Adaptivity**

Effective tutoring relies on adaptation to context

Inexperienced human tutors often:

- ignore signs of students' confusion
- follow inappropriate plans
- give long-winded, didactic explanations

(Chi et al., 2001)

Effective Tutoring is planned behaviour

Tutorial techniques (e.g. in Algebra) are often ***complex***, i.e., require a series of actions (McArthur et al. 1990)

But…strategy may fail

Repeated incorrect answers or "I don't know", other signs of confusion

Students may:

- skip steps, jumping to right answer
- change topic
- introduce new tutorial goals in response

Tutorial Dialogue can be effective

Through appropriate use of

- tutorial strategies said to improve learning, e.g.
  - Graesser's 5-step strategy (Graesser et al. 1994):
    - Step1: T asks question
    - Step2: S answers question
    - Step3: T gives short feedback
    - Step4: T and S collaborate on improving the answer
    - Step5: T assesses S' understanding of the answer
  - McArthur's "micro-plans" (McArthur et al. 1990)
  - Chi's prompting strategies (Chi et al. 2001)
- shared linguistic conventions, e.g.
  - Politeness strategies (Brown and Levinson 1987; Porayska-Pomsta 2003)

Architectural requirements: Flexibility

In order to:

- Support multiple complex strategies
- Adapt to:
  - Failure to recognise student's response
  - Failure of strategy
  - Interruptions (clarifications)
  - Student skips ahead
  - Student changes topic
  - Student provides more information than expected

Architectural requirements: Re-Usability

Clean separation of knowledge sources

- share knowledge among the system components
  - domain reasoner
  - student modeller
  - dialogue manager
  - understanding and generation
- minimise re-representation
- maximise domain independence, ease of maintenance, and re-usability

3. Modelling Tutors' Feedback *(Porayska-Pomsta, 2004)*

Tutors' Feedback

What language do tutors' produce in corrective situations?

- Dialogues analysis

What drives the selection of tutors' responses?

- Contextual factors relevant to tutors' decisions
- Politeness considerations

How can the process of selecting tutors' responses be modelled formally?

- Outline of the model of tutors selecting corrective responses
- Model's implementation and evaluation

Example of feedback variation

Tutor's question:
What is needed to light a light bulb?
Student's answer:
Heat. (incorrect)
Tutor's possible feedback:

1. No, that's incorrect.
2. Try again.
3. Well, why don't you try again?
4. Are you sure about that?
5. Well, if you put the light bulb in the oven it will certainly get a lot of heat, but is it likely to light up?
6. Is it the heat or the source that are needed to light a light bulb?
7. Why?

What language do tutors produce in corrective situations?

Approx. 50% of all dialogue moves produced by tutors are questions.

A distinction between communicatively "straight" acts, testing acts and corrective acts, e.g.:

What do you mean by this? (a straight act)
vs.
What are the main components needed to light a light bulb? (a testing act)
vs.
Well, if you put the light bulb in the oven it would get heat, but would it light up? (a corrective act)

Taxonomy of Corrective Acts

Negative Assertives (53.5%)
Direct Negatives (72.6%): *"That's not right"*
Hidden Negative Assertives (27.3%): *"It's the source that is needed"*

Hidden Negative Instructions (7%)
Direct instructions (92.8%): *"Try again"*
Indirect instructions (7.1%): *"Why don't you try again?"*

Hidden Negative Questions (39.3%)
Positive polarity (6.4%): *"Is heat needed to light a light bulb?"*
Negative Polarity (6.4%): *"Isn't heat needed to light a light bulb?"*
IF-THEN questions (24.3%): *"If you put the light bulb in the oven, it will certainly get a lot of heat, but will it light up?"*
WH-content seeking (41%): *"What formula do you have for calculating power?"*
WH-explanation seeking (21.8%): *"Why would you say that?"*

What is the difference?

Indirectness:

Illocutionary specificity:
the degree to which the teacher hides the rejection of the student's answer.

Content Specificity:
the degree to which the teacher gives the relevant content away.

Why do the possible responses differ?

Because they allow the tutor to achieve slightly different communicative and educational goals to various degrees,
e.g.

- ***tell the student*** his answer was problematic
- ***prompt/guide the student*** to make further attempts at finding a solution.
- ***boost the student'***s confidence and curiosity

What determines which form of response is best: Face

Face is a person's self-image which needs to be maintained, respected and approved of by self and others (Brown and Levinson, 1987)

Autonomy: a dimension of a student's Face which refers to his need to be allowed the freedom of initiative to discover knowledge by himself

Approval: a dimension of a student's Face which refers to the student's need for his positive self-image to be maintained.

A student's **positive self-image** relies on his motivational states being maintained or boosted.

What determines which form of response is best: Context

Based on the tutor's awareness of contextual factors, e.g.,

- student's characteristics,
- the characteristics of the material taught,
- time and place of teaching, etc.

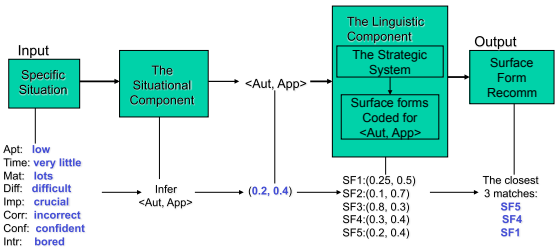
(e.g., Lepper and Chabay 1988; Graesser 1995; Person 1995; deVicente 2003; etc.)

Identifying the situational variables relevant to tutors' corrective response selection

1. **Temporal factors:** (from observation of the dialogues)
   amount of time available
   amount of material left
2. **Characteristics of the material taught:** (Lepper and Chabbay 1988; Person et al. 1995; Chi 2001)
   difficulty of the material
   importance of the material
3. **Characteristics of the student:** (Lepper and Chabbay 1988; Person et al. 1995; Chi 2001; deVicente 2003)
   student's ability
   correctness of student's answer
   student's confidence
   student's interest

Validated through empirical study with teachers

Model of tutors selecting corrective responses



The situational component

1. **Motivation oriented Factors (MOFs)**
   Student's confidence
   Student's interest
2. **Lesson oriented Factors (LOFs)**
   Time oriented factors
   Amount of time
   Amount of material
   Content oriented factors
   Difficulty of material
   Importance of material
3. **Performance oriented Factors (POFs)**
   Correctness of student's answer
   Student's Aptitude

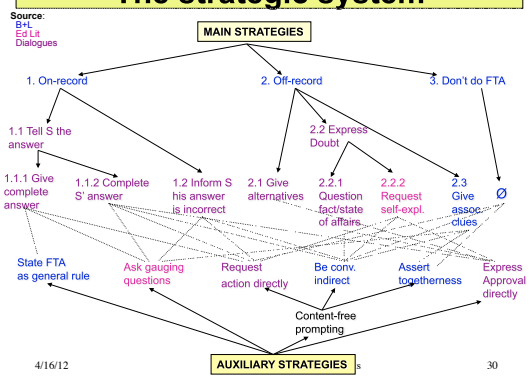
The Linguistic Component: The strategic system

Brown and Levinson's strategies:
On-record, Off-record and Don't do FTA are adopted

Strategies are adapted to suit the educational context in which corrective responses are made (sources: dialogue analysis, educational literature)

Split between main strategies and auxiliary strategies.

The strategic system



Method for assigning <Aut, App> values to strategies and surface forms

Strategies are examined in general terms with respect to the level of Autonomy and Approval that they seem to express

Range of values which intuitively describes those levels the best is assigned to each strategy (fuzzy descriptions (plenty of guidance), fuzzy values (low –medium), numerical values (0—0.45)

The strategies are compared with one another to ensure coherence in the way these values are actually assigned

Method for assigning <Aut,App> values to surface forms

Surface forms which combine the qualities of several strategies are assigned <Aut,App> values based on the values of all the strategies involved (a weighted sum calculation)

Example Form:
"Let us try again. Say the black lead is connected to tab 4. Which tab positions would be included?":

Strategy:

GIVE ASSOCIATION CLUES (Aut = 0.3 and App = 0.7) + IMPLY TOGETHERNESS (Aut = 0.4 and App = 0.7) .

<Aut, App> values:

Aut = (0.3 * 1 + 0.4 * 0.5)/ 1.5 = 0.33333

Aut = (0.7 * 1 + 0.7 * 1) /2 = 0.7

Autonomy, Approval and Linguistic Choice

"No, that's not right." (Aut: 1.0, App: 0.1)

"Are you sure that this is the right way to de-energize the circuit?" (Aut: 0.8, App: 0.4)

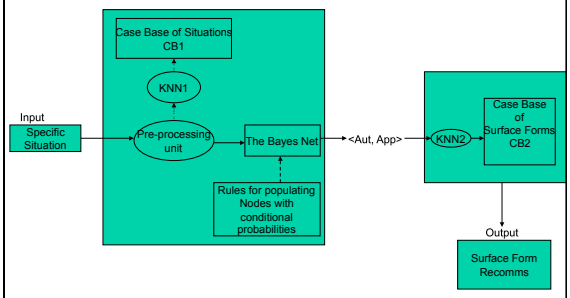
"Not quite, why don't you try again?" (Aut: 0.6, App: 0.4)

"Removing the wire does not de-energize the circuit." (Aut 0.4, App: 0.1)

"If you remove the wire, then this will break the circuit but does it de-energize it?" (Aut: 0.3, App: 0.5)

"Isn't this breaking the circuit rather than de-energizing it?" (Aut: 0.2, App: 0.3)

Implementation of the model



Comparing model- and teacher-generated surface forms

Participants: 4 very experienced tutors in the domain of basic electricity and electronics

Materials and Procedure: For each dialogue/situation:

1. **Teacher produced response**
2. **Model's preferred response**
3. **Model's less preferred response**

- hard-copy questionnaires
- 20 different situations each as dialogue between a student and a tutor
- scale from 1-5 for rating each of the 3 responses according to how appropriate they seem for a dialogue
- basic electricity and electronics domain

Results of the Evaluation

Significant difference between **human** and **system's less preferred** responses ($t_2(19) = 4.40$, $p<0.001$)

Significant difference between **system's preferred** and **system's less preferred** responses ($t_2(19) = 2.72$, $p = 0.013$)

No significant difference between **system's preferred** and **human** responses ($t_2(19) = 1.99$, $p = 0.061$)

Future Work

Hypertext lessons and multiple choice questions cover the topics of current, voltage, resistance, and power.
Dialogue capabilities support discussion centred around the lab exercise, measuring current.

Future work:
- **extending the system's coverage** to other topics of discussion
- generally **enriching its knowledge sources** (e.g. tutorial strategies, natural language understanding and generation resources).

In addition, the architecture was ported to the domain of calculus.

Issues

- Mixed Initiative dialogues
  - If unrestricted – difficult to handle because of NLU
- Multi-modality
  - Pointing
  - Speaking
  - Typing
  - Embodied conversational Agents (Talking Heads)
- Not all successful strategies of human tutors may be suitable for computer-based tutoring
- Not all information available to tutors will be equally available to computer tutors and vice versa.

4. References and other sources

References

Porayska-Pomsta, K. (2004). Influence of Situational Context in Language Production: *Modelling Teachers' Corrective Responses*. Ph.D. thesis, University of Edinburgh.

De Vicente, A. & Pain, H. (2002). Informing the detection of the students' motivatonal state: An empirical study. In S.A. Cerri, G. Gouardères, & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems,* 933-943. Berlin: Springer.

See http://www.cogsci.ed.ac.uk/~jmoore/tutoring/papers.html for various Beetle papers.

Other relevant papers

Bloom, B.S. (ed.) (1956), Taxonomy of educational objectives: The classification of educational goals. Handbook I, cognitive domain. London: Longman.
Brown, P. & Levinson, S.C. (1987). Politeness: Some universals in language use. New York: Cambridge University Press.
Lepper, M.R., Woolverton, M., Mumme, D., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S.P. Lajoie and S.J. Derry (Eds.), Computers as cognitive tools, 75-105. Hillsdale, NJ: Lawrence Erlbaum Associates.
Johnson, W.L. (2003). Interaction tactics for socially intelligent pedagogical agents. Int'l Conf. on Intelligent User Interfaces, 251-253. New York: ACM Press.
Johnson, W.L., Rickel, J.W., and Lester, J.C. (2000): Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. International Journal of Artificial Intelligence in Education 11, (2000) 47-78
Pilkington, R.M. (1999). Analysing educational discourse: The DISCOUNT scheme. Technical report 99/2, Computer-Based Learning Unit, University of Leeds.

Other useful references

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P. and Glaser, R. (1989). Self-Explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2): 145-182.
Chi, M. T. H., de Leeuw, N., Chiu, M-H. and Lavancher, C. (1994), Eliciting self-explanations improves understanding. *Cognitive Science,* 18(3): 439-477.
Chi, M. T. H., Siler, S.A., Jeong, H., Yamauchi, T. and Hausmann, R.G. (2001), Learning from Human Tutoring, *Cognitive Science*, 25: 471-533
Fox. B.A. (1993), *The Human Tutorial Dialogue Project: Issues in the design of instructional systems.* Lawrence Erlbaum Associates, Hillsdale, NJ.
Graesser, A. C. and Person, N. K. (1994) Question asking during tutoring. *American Educational Research Journal*, 31 (1): 104-137.
Lepper, M. R. and Chabay, R. W. (1988) Socializing the intelligent tutor: Bringing empathy to computer tutors. In H. Mandl and A. Lesgold, editors, *Learning Issues for Intelligent Tutoring Systems*, pp 114-137. Springer-Verlag, New York.
Merrill, D. C., Reiser, B. J., Ranney, M. and Trafton, J. G. (1992), Effective tutoring techniques: Comparison of human tutors and intelligent tutoring systems. *Journal of the Learning Sciences*, 2 (3): 277-305.
Ohlsson, S. and Rees, E. (1991), The function of conceptual understanding in the learning of arithmetic procedures. *Cognition and Instruction*, 8, 103-179.