

U: Hello

R: Good morning. How can I help you?

U: where can I find frank

R: Frank Jones is in 4001 Wean Hall. Would you like directions?

U: Sure

R: Go to the stairs on your left [. . .] Room 4001 will be on your right.

▶ U: Weather?

R: The weather tonight in Pittsburgh, Pennsylvania will be cloudy and 27 degrees.

:

▶ U: I was looking for another frank

R: The closest match I have in my directory is Frank Jones in 4001 Wean Hall. Would you like directions?

U: any frank in smith hall

R: I didn't understand that. It might help if you rephrase it.

U: Come on.

R: Whatever

Incorporating a User Model to Improve Detection of Unhelpful Robot Answers

Maxim Makatchev, Reid Simmons

Carnegie Mellon University
`maxim.makatchev@cs.cmu.edu`

Acknowledgements:
Greg Armstrong, Brett Browning,
Min Kyung Lee, Sanako Mitsugi, and Antonio Roque

November 6, 2009

Problem: Detecting errors in understanding

- Errors, problems, mis-understandings, non-understandings.
 - Prosodic cues [Krahmer et al., 2001]
 - Content cues: repetitions, corrections [Hirasawa et al., 2000, Swerts et al., 2000]
- **Unhelpful answers** (this study)
 - Intention level: e.g. lexical ambiguity, semantic interpretation:

U: i am looking for beer

R: John Beer is at Wean 4002.

- Conversation levels (uncooperativeness) [Paek, 2003]:

U: Where can I eat?

R: So go and get some food.

Unhelpful answers:

24% of all questions are answered unhelpfully.

Unhelpful answer detection can be used. . .

- . . . for dialogue management (e.g. apply recovery policies [Bohus et al., 2006]).
- . . . as automated annotations for data analysis/learning.

Social dialogue acts

Adherence to social conventions: Greeting, Thanks, Goodbye
Linguistic equivalent of the “social media” phenomena
[Reeves & Nass, 1996].

Social dialogue acts

Can we use social dialogue acts to detect Unhelpfulness?

Yes, we can!

Methodology

Binary variables:

U:thanks \in {Did not thank, thanked }

U:greeting \in {Did not greet, Greeted }

R:answer \in {Unhelpful, Helpful }

Significant associations U:thanks, U:greeting and R:answer.

Classify R:answer using U:thanks and U:greeting as features and do significantly better than random.

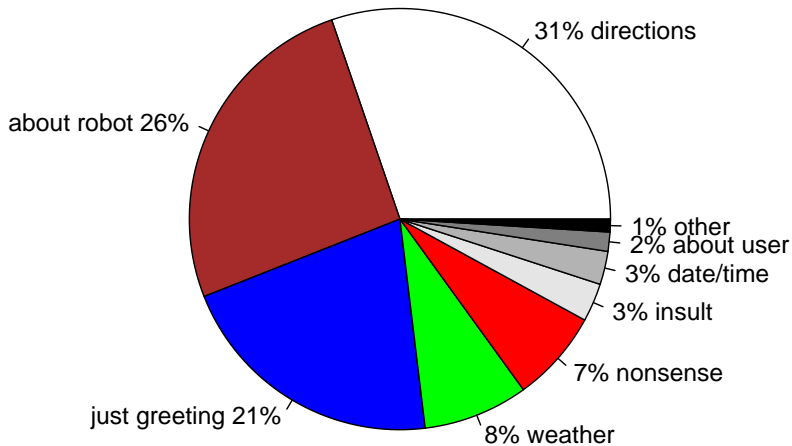
Roboceptionist



- Actuated head display engages passers-by
- Input: typed
- Output: spoken & captions
- Since 2004
- Does not authenticate users (generally)
- Has a back-story

Dialogue topics

2 months, 1676 dialogues, 11024 turns



[Lee & Makatchev, 2009]

Example dialogue

U: hello

R: Welcome to Newell-Simon. How can I help you?

U: how are you today /

R: Dr. Spitolá suggested that a self defense seminar might help to reassemble my self esteem. [...]

U: okay

R: <nod>

:

U: i need to go to wean hall

R: Wean Hall? Go up the stairs in front of you. Turn left and walk to the end of the hall way, and you will be in the computer science building, Wean Hall.

U: thank you

R: You are welcome

Dialogue corpus

Coverage

- 7 days in March 2008
- 221 dialogues, 1524 dialogue turns

Annotations

- dialogue acts: greeting, thanks, goodbye, question, answer, non-understanding and insult

Automatically annotated corpus

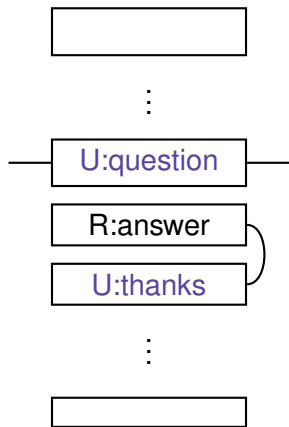
Classifiers

- decision tree binary classifiers, bag-of-words features
- at least 89% accuracy, 10-fold cross validation.

Coverage

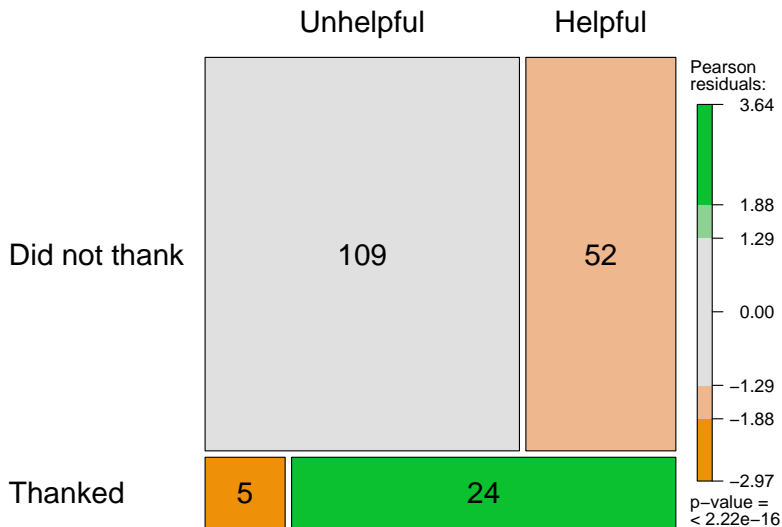
- 2 months: March and April 2008
- 1676 dialogues, 11024 dialogue turns

Associating R:answer and U:thanks



	Unhelpful	Helpful
Did not thank	109	52
Thanked	5	24

Associating R:answer and U:thanks

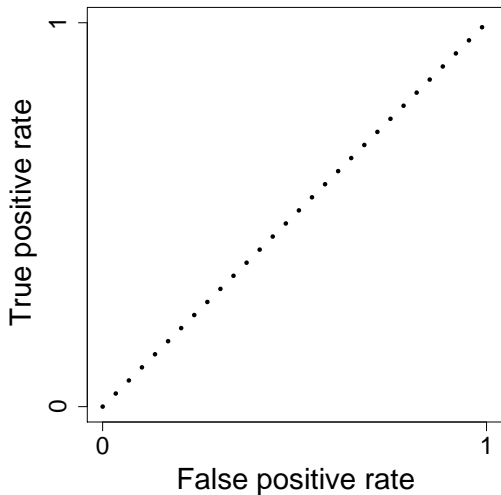


Classifier performance evaluation

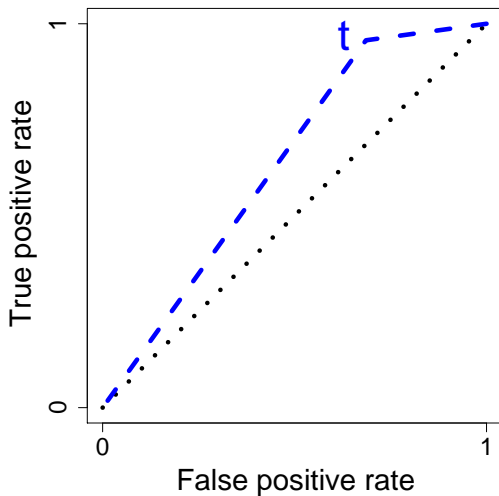
$$\text{True positive rate} = \frac{\text{True positive}}{\text{Positive}}$$

$$\text{False positive rate} = \frac{\text{False positive}}{\text{Negative}}$$

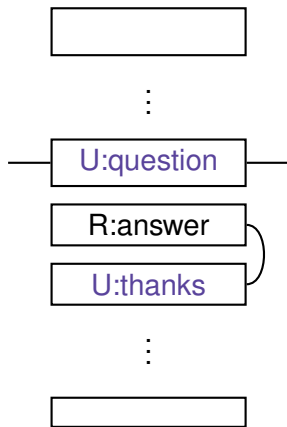
Let's ROC: classifier performance evaluation



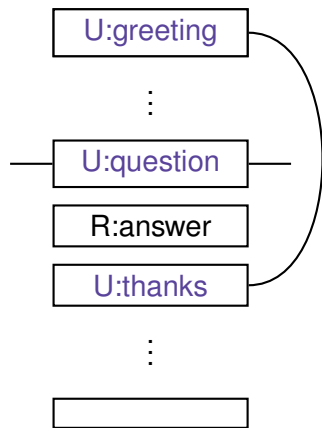
Detecting R:answer(Unhelpful) using U:thanks feature



Recap: U:thanks is associated with R:answer



Other associations?

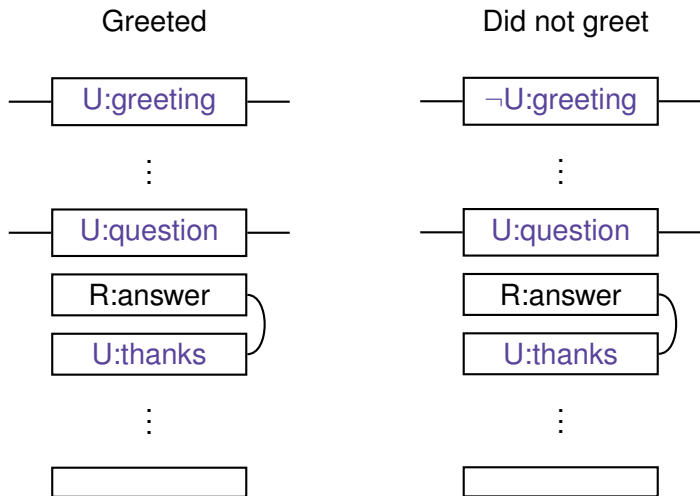


Users that greet the robot exhibit different linguistic behaviors

[Fischer, 2006]

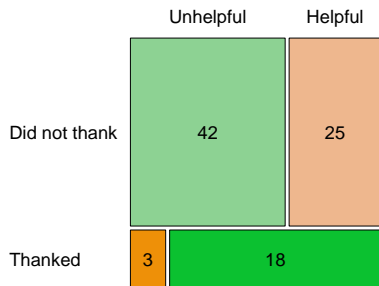
	<i>U:greeting</i>	$\neg U:greeting$
<i>U:thanks</i>	0.25	0.09

Split on whether user Greeted or Did not greet.



Did not greet:
fewer U:thanks
more R:answer(Unhelpful)

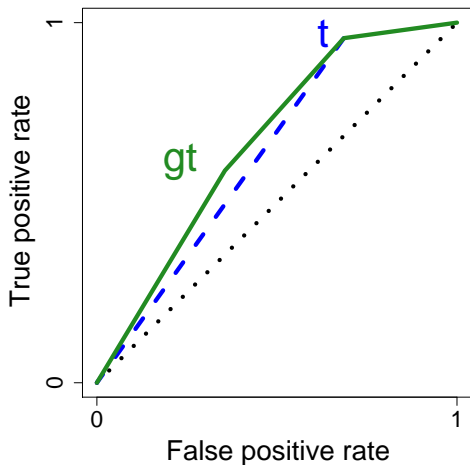
Greeted



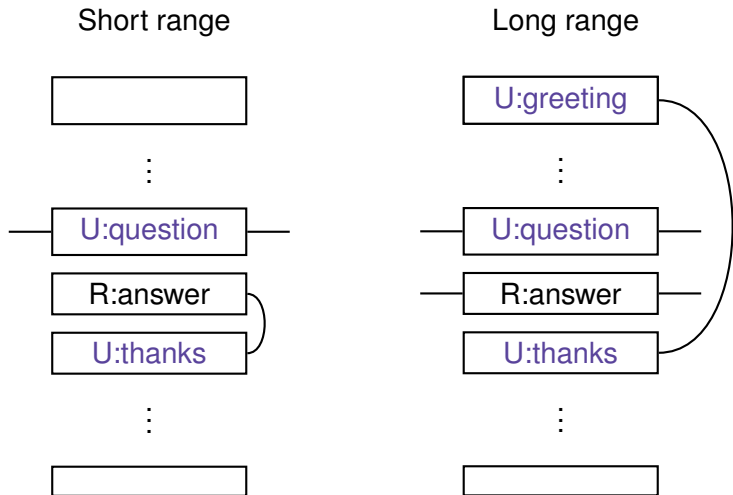
Did not greet



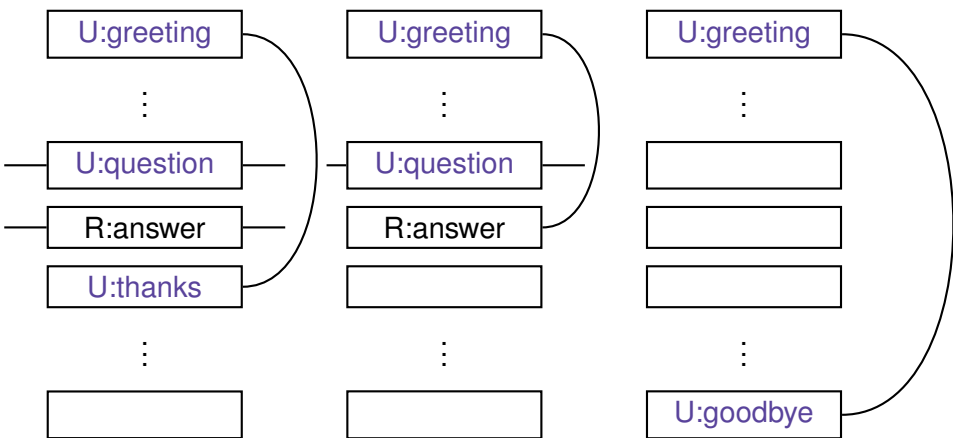
Adding U:greeting as a feature (vs. just U:thanks)



The Message



U:Greeting is associated with U:Thanks, R:Answer and U:Goodbye.



Summary

- Short range association exists between user's thanks and helpfulness of the robot's answer.
- Long range associations exist between social dialogue acts such as between user's greeting, thanks and goodbye.
- Long range association exist between user's greeting and having user's question answered helpfully.
- Long range associations can be interpreted as a user model (trait).
- Combining short and long range associations improves detection of unhelpful robot answers.

Future work

- A practical classifier should also use content words as features.
- Use unhelpful answer detection to guide the dialogue.

U: Hello

R: Good morning. How can I help you?

U: where can I find frank

R: Frank Jones is in 4001 Wean Hall. Would you like directions?

U: Sure

R: Go to the stairs on your left [...] Room 4001 will be on your right.

U: Weather?

R: I guess that wasn't too helpful. Anyways, the weather tonight in Pittsburgh, Pennsylvania will be cloudy and 27 degrees.


U: Woah

 Bohus, D., Langner, B., Raux, A., Black, A., Eskenazi, M. & Rudnicky, A. (2006).

In Spoken Language Technologies, Palm Beach, Aruba.

 Fischer, K. (2006).

In Proc. of the Workshop on 'How People Talk to Computers, Robots, and Other Artificial Communication Partners', (Fischer, K., ed.), pp. 112–130,, Hansewissenschaftskolleg, Delmenhorst.

 Hirasawa, J., Miyazaki, N., Nakano, M. & Aikawa, K. (2000).

In Proc. Int Conf. of Spoken Language Processing (ICSLP) vol. 2, pp. 154–157,, Beijing, China.

 Krahmer, E., Swerts, M., Theune, M. & Weegels, M. (2001).

Int. J. of Speech Technology 4, 19–29.

 Lee, M. K. & Makatchev, M. (2009).

In Proceedings of CHI pp. 3769–3774,.



Paek, T. (2003).

In Proc. ISCA Tutorial and Research Workshop on Error Handling in Spoken Dialogue Systems pp. 53–58,.



Reeves, B. & Nass, C. (1996).

The Media Equation: how people treat computers, television, and new media like real people and places.

Cambridge University Press.



Swerts, M., Litman, D. & Hirschberg, J. (2000).

In Proc. ICSLP-2000 pp. 615–618,.