Uphill Battles in Task Modeling and Grounding for Dialog

Amanda Stent

In the Beginning were Problems

E: First you have to remove the flywheel.
A: How do I remove the flywheel?
E: First, loosen the two allen head setscrews holding it to the shaft, then pull it off.
A: OK.
I can only find one screw. Where's the other one?
E: On the hub of the flywheel.
A: that's the one I found. Where's the other one?
E: About ninety degrees around the hub from the first one.
A: I don't understand. I can only find one. Oh wait, yes
I think I was on the wrong wheel
…
A: The two screws are loose, but I'm having trouble getting the wheel off.
E: Use the wheelpuller. Do you know how to use it?
A: No.
E: Do you know what it looks like?
A: Yes.
E: Show it to me please.
A: OK
E: Good. Loosen the screw in the center and place the jaws around the hub of the wheel…

Grosz (1978), *Proc. Theoretical Issues in NLP - 2*

Then Problem Definitions…

Technique used	Example task	Task complexity	Dialogue phenomena handled
Finite-state script	Call transfer		User answers questions
Frame based	Getting flight info		User asks questions, simple clarifications by system
Sets of contexts	Booking travel		Shifts between predetermined topics
Plan-based models	Kitchen design consultant		Dynamically generated topic structures, collaborative negotiation subdialogues
Agent-based models	Disaster relief management		Different modalities (e.g. actual and planned worlds)

Adapted from Allen et al (2001), *AI Magazine*

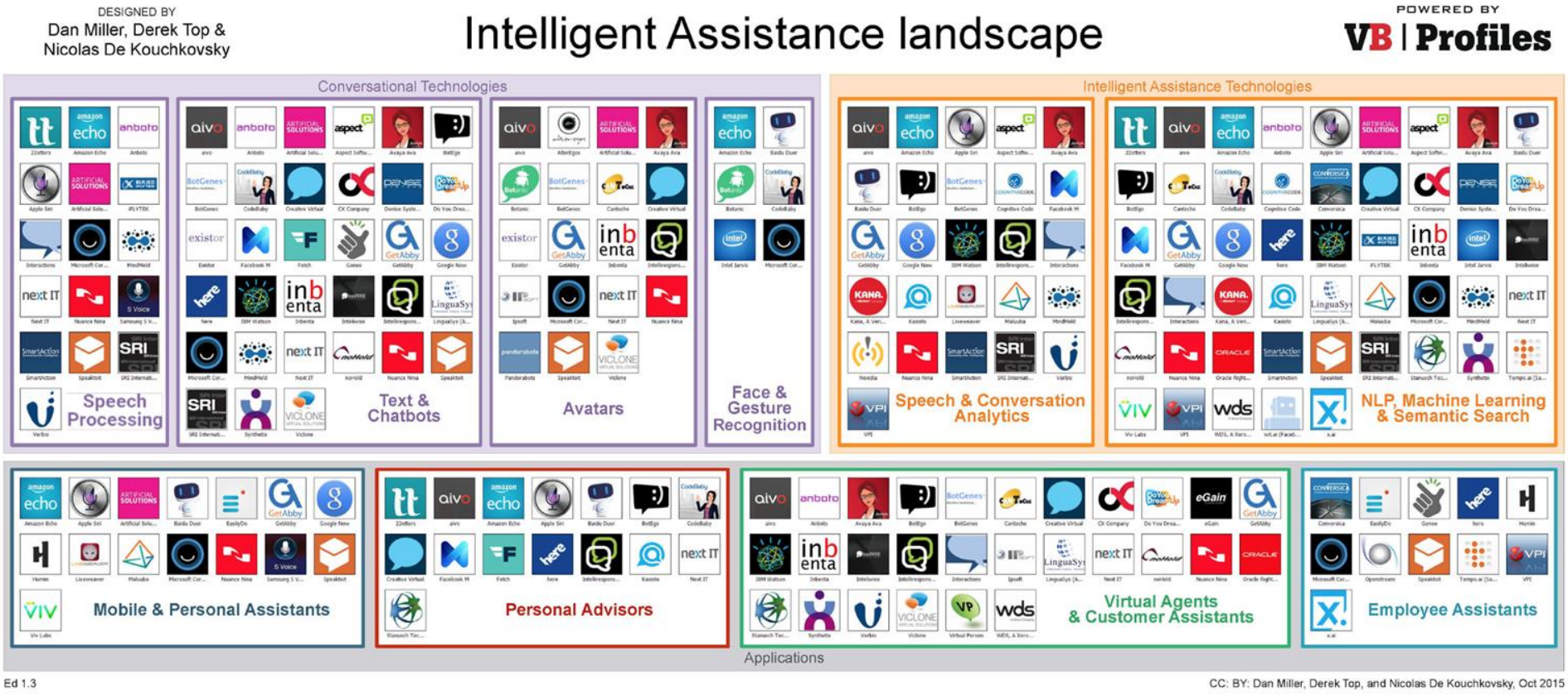
Evaluation Metrics and Parametrizable Methods, Data and Time…

- Walker, M. A., Litman, D. J., Kamm, C. A., & Abella, A. (1997). PARADISE: A framework for evaluating spoken dialogue agents. In *Proceedings of the EACL.*
- Levin, E., Pieraccini, R., & Eckert, W. (2000). A stochastic model of human-machine interaction for learning dialog strategies. *IEEE Transactions on speech and audio processing*, *8*(1), 11-23.

- 2000-2001 – Communicator
- 2010 – Spoken Dialog Challenge
- 2011-2014 – DSTC 1-4

Microsoft

Facebook



Amazon

Apple

Google

But We Still Didn't Solve Our Problems

E: **First** you have to remove *the flywheel*.
A: How do I remove *the flywheel*?
E: **First**, loosen *the two allen head setscrews* holding *it* to the shaft, then pull *it* off.
A: OK.
I can only find one screw. Where's the other one?
E: On the hub of the flywheel.
A: that's the one I found. Where's the other one?
E: About ninety degrees around the hub from the first one.
A: I don't understand. I can only find one. Oh wait, yes
I think I was *on the wrong wheel*
…
A: The two screws are loose, **but** I'm having trouble getting the wheel off.
E: Use the wheelpuller. Do you know how to use it?
A: No.
E: Do you know what it looks like?
A: Yes.
E: Show it to me please.
A: OK
E: Good. Loosen the screw in the center and place the jaws around the hub of the wheel…

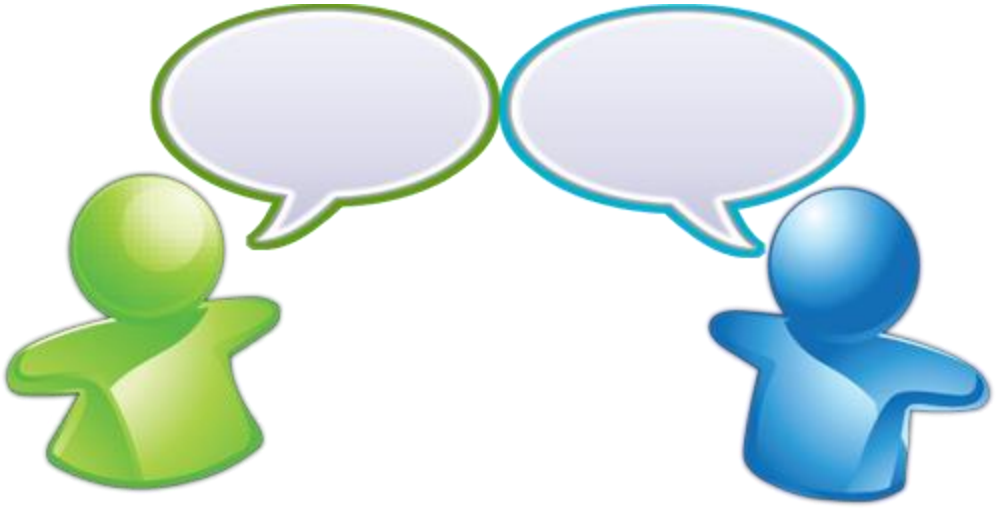
Grosz (1978), *Proc. Theoretical Issues in NLP - 2*

“The formulation of the problem is often more essential than its solution, which may be merely a matter of mathematical or experimental skill.”
— **Albert Einstein**

What is Dialog?

3 or more (language-rich) turns

Coherent



A:
B:
A:

What is a Dialog System?

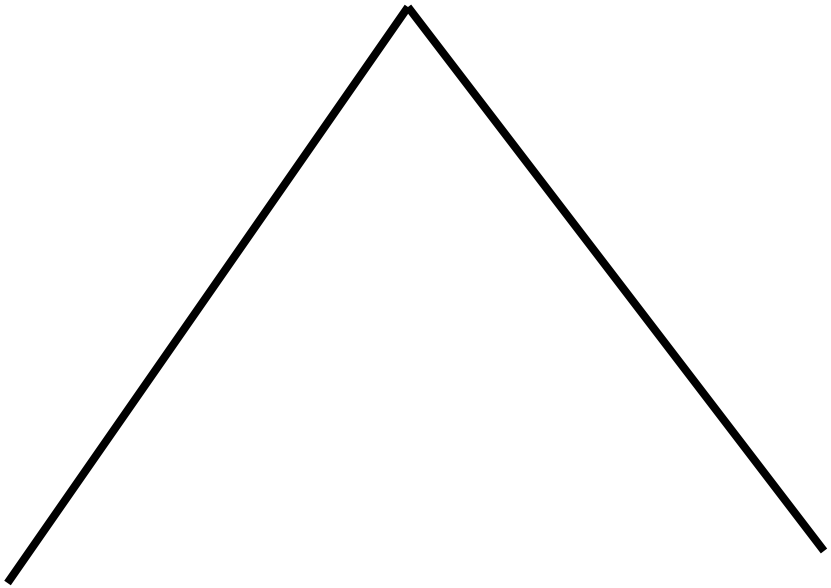
A dialog system is a computer system that can engage in a dialog *to do something*

aAa bBb aAa bBb aAa bBb *bBb* bBb aAa bBb aAa bBb bBb

Dialog Systems Research – Some Variables

- Task-related: complex and hierarchical tasks
  - How many objects? And for each, how many attributes?
    - *Dozens (Let's Go, DSTC) to millions (directory lookup)*
  - How many verbs?
    - *One*
  - How fixed/flexible?
    - *Fixed: info retrieval*
  - What evaluation metrics?
    - *Info presented + minimal time*
- Language-related: shared representations of common ground
  - Anaphor resolution
    - *Very little*
  - Entrainment
    - *Almost impossible due to lack of info sharing at the language level*
  - Response generation
    - *Limited by lack of rich task representation and lack of common ground modeling*
  - Duration of modeling
    - *Too short! An adjacency pair is not a dialog!!*

Why Are You Building Your Dialog System?

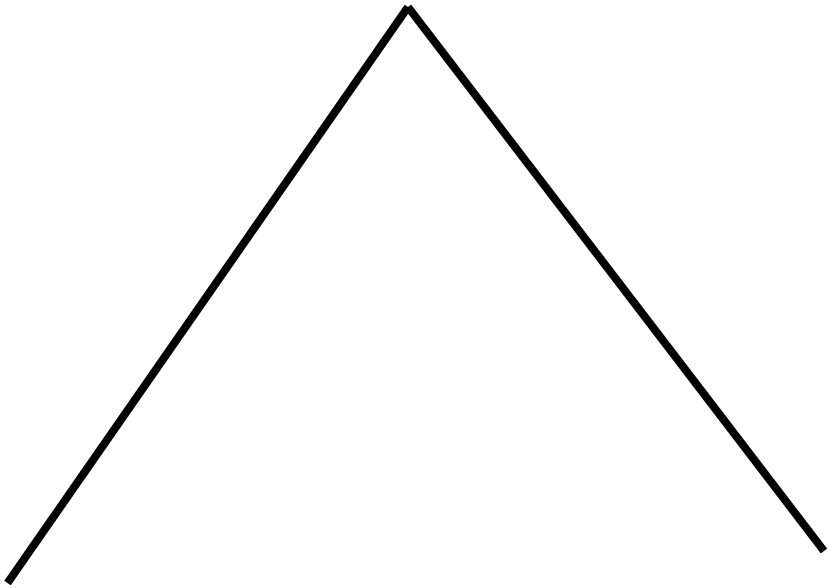


Engineering:
what optimization function?
~ Efficiency of interaction

Modeling:
what phenomena?
~ ASR error handling

Adapted from Brennan (2016), SIGDIAL

Why Are You Building Your Dialog System?



Engineering: model building
feature engineering ☹
generalizability ☹

Modeling: hypothesis testing
variable selection ☺
explanatory power ☺

Adapted from Brennan (2016), SIGDIAL

Challenges

- Move beyond info-seeking dialog
  - Dialog systems move from cost centers to profit centers
  - New evaluation metrics
  - Must look at additional interaction parameters
- Go back to our roots
  - AI ~ (queryable, studyable) models of human intelligences
    - Entrainment
    - Anaphor resolution and grounding
    - Contextually appropriate, mixed-initiative generation
    - Non-cooperative and semi-cooperative dialog
- Consider other types of AI
  - Moving beyond session-based dialog
    - Dialog systems remember better than we do
    - Dialog systems that can handle many parallel threads
    - Dialog systems that can help us manage interruption

Interesting Current Directions

- Complex and flexible tasks, evaluation beyond efficiency:
  - Sun, M., Chen, Y. N., & Rudnicky, A. I. (2016). An intelligent assistant for high-level task understanding. In *Proceedings of IUI.*
  - Lee, S., & Stent, A. (2016). Task Lineages: Dialog State Tracking for Flexible Interaction. In *Proceedings of SIGDIAL.*
  - Mohseni-Kabir, A., Rich, C., Chernova, S., Sidner, C. L., & Miller, D. (2015). Interactive hierarchical task learning from a single demonstration. In *Proceedings of HRI.*
- Anaphora and entrainment, shared language representations:
  - Misu, T., Raux, A., Gupta, R., & Lane, I. (2015). Situated language understanding for a spoken dialog system within vehicles. *Computer Speech & Language*, *34*(1), 186-200.
  - Kennington, C., & Schlangen, D. (2017). A simple generative model of incremental reference resolution for situated dialogue. *Computer Speech & Language*, *41*, 43-67.
  - Hu, Z., Halberg, G., Jimenez, C. R., & Walker, M. A. (2016). Entrainment in Pedestrian Direction Giving: How many kinds of entrainment?. In *Situated Dialog in Speech-Based Human-Computer Interaction* (pp. 151-164). Springer International Publishing.
- Non-cooperative (situated) dialog:
  - Efstathiou, I., & Lemon, O. (2016). Learning Better Trading Dialogue Policies by Inferring Opponent Preferences. In *Proceedings of AAMAS.*
  - Andrist, S., Bohus, D., Yu, Z., & Horvitz, E. (2016). Are You Messing with Me?: Querying about the Sincerity of Interactions in the Open World. In *Proceedings of HRI*.

https://dialrc.org/dialport/

Some Additional References Everyone Should Read

- Austin, J. L. (1975). *How to do things with words*. Oxford University Press.
- Bohus, D., & Rudnicky, A. I. (2003). RavenClaw: Dialog management using hierarchical task decomposition and an expectation agenda. In *Proceedings of SIGDIAL.*
- Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation*, *53*, 301-344.
- Chu-Carroll, J., & Carberry, S. (1998). Collaborative response generation in planning dialogues. *Computational Linguistics*, *24*(3), 355-400.
- Clark, H. H. (1996). Using language. *Cambridge University Press.*
- Danlos, L. (1984). Conceptual and linguistic decisions in generation. In *Proceedings of COLING-ACL.*
- Grice, H. P. (1969). Utterer's meaning and intention. *The Philosophical Review*, 147-177.
- Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational linguistics*, *12*(3), 175-204.
- Guinn, C. I. (1996, June). Mechanisms for mixed-initiative human-computer collaborative discourse. In *Proceedings of the ACL.*
- Jameson, A., et al. (1994). Cooperating to be noncooperative: The dialog system PRACMA. In *Proceedings of the Annual Conference on Artificial Intelligence*.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.
- Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of CHI.*
- Power, R. (1979). The organisation of purposeful dialogues. *Linguistics*, *17*(1-2), 107-152.
- Schlangen, D., & Skantze, G. (2009). A general, abstract model of incremental dialogue processing. In *Proceedings of EACL.*
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language* (Vol. 626). Cambridge University Press.
- Sidner, C. L. (1983). Focusing and discourse∗. *Discourse Processes*, *6*(2), 107-130.
- Traum, D., & Hinkelman, E. (1992). Conversation acts in task-oriented spoken dialogue. *Computatinal Intelligence, 8, 575-599.*
- Williams, J., Raux, A., & Henderson, M. (2016). The dialog state tracking challenge series: A review. *Dialogue & Discourse*, *7*(3), 4-33.