From Dialogue Systems to Social Chatbots: Reinforcement Learning, Seq2Seq, and back again.

Verena Rieser,
Professor in Computer Science
Heriot-Watt University, Edinburgh, UK.

HERIOT WATT UNIVERSITY

SwissText keynote, 9th June 2017

INTERACTION LAB

HERIOT WATT UNIVERSITY

NLP group @HWU

Prof Verena
Rieser, Director

Dr Simon
Keizer, PDRA

Dr Jekaterina
Novikova, PDRA

Dr Ondrej
Dusek, PDRA

Dr Xingkun
Liu, PDRA

Amanda
Curry,
PhD student

Xinnuo Xu,
PhD student

Current Research Projects

- Statistical Natural Language Generation (EPSRC DILIGENT project)
- Transfer Learning for Dialogue Systems (EPSRC MADRIGAL project)
- Automatic Quality Estimation for Output Generation (NVIDIA)
- Personalized Human Robot Interaction (with EmoTech LTD)
- Amazon Alexa Challenge (Amazon)
- Sentiment Analysis for Arabic (SemEval’16 winner)

Talking Machines

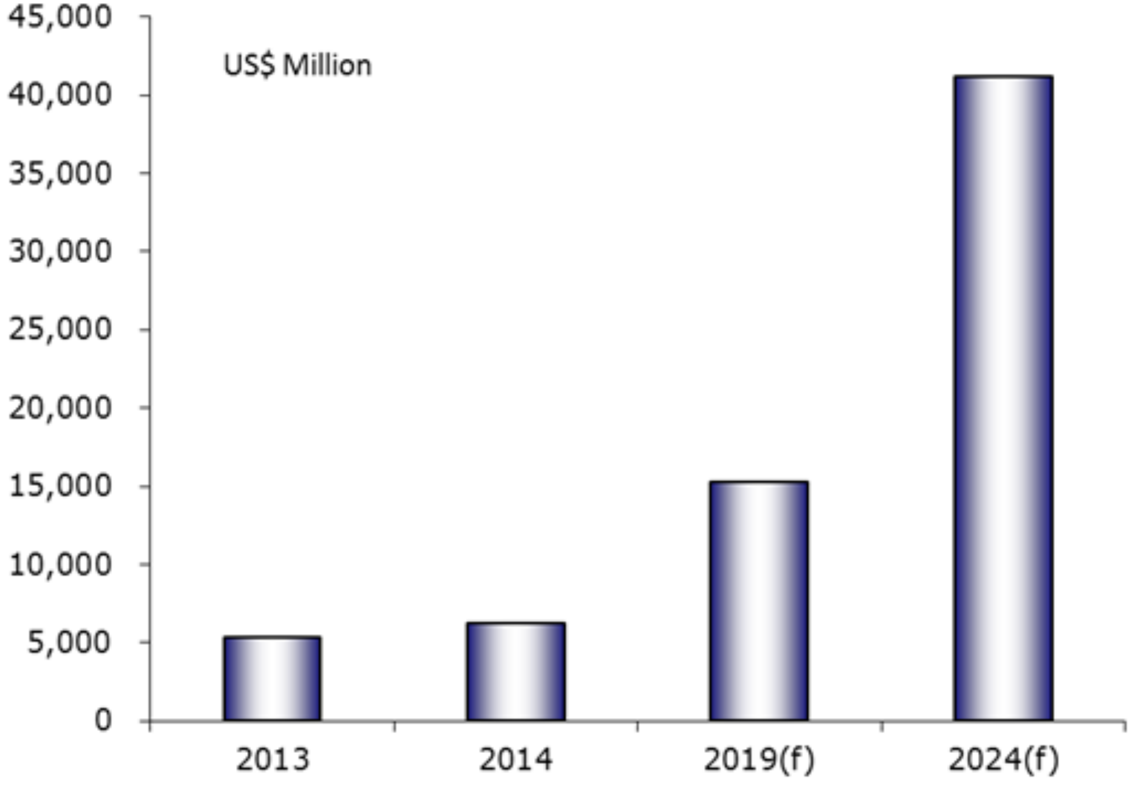
The new Bots are coming….

“*Bots are the new apps*'' because they ”*fundamentally revolutionize how computing is experienced by everybody*.”

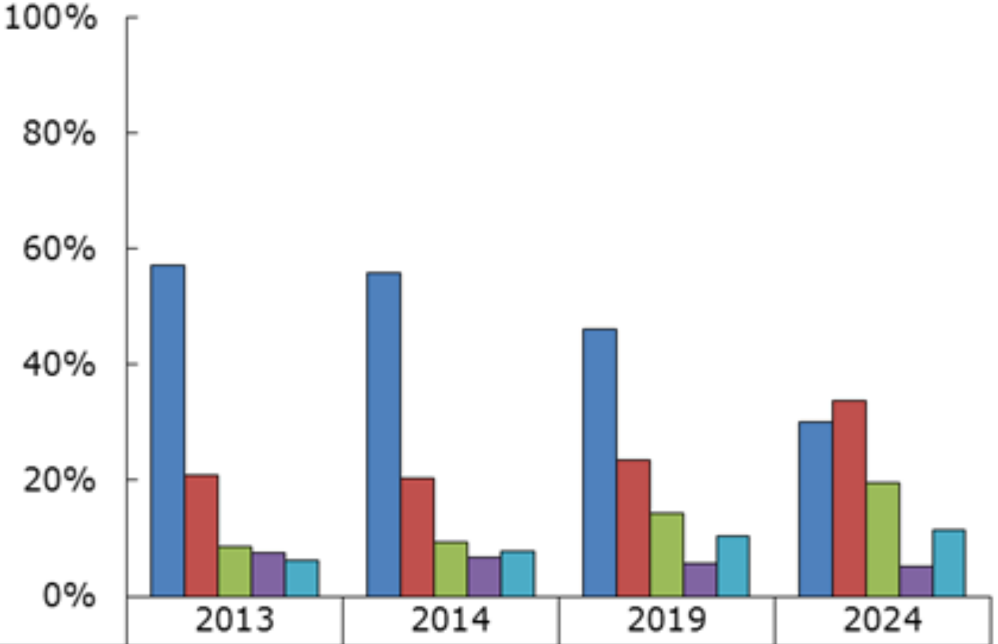
Microsoft’s CEO Nardella

Microsoft

Market forecast



Source: MIC Jan 2015.

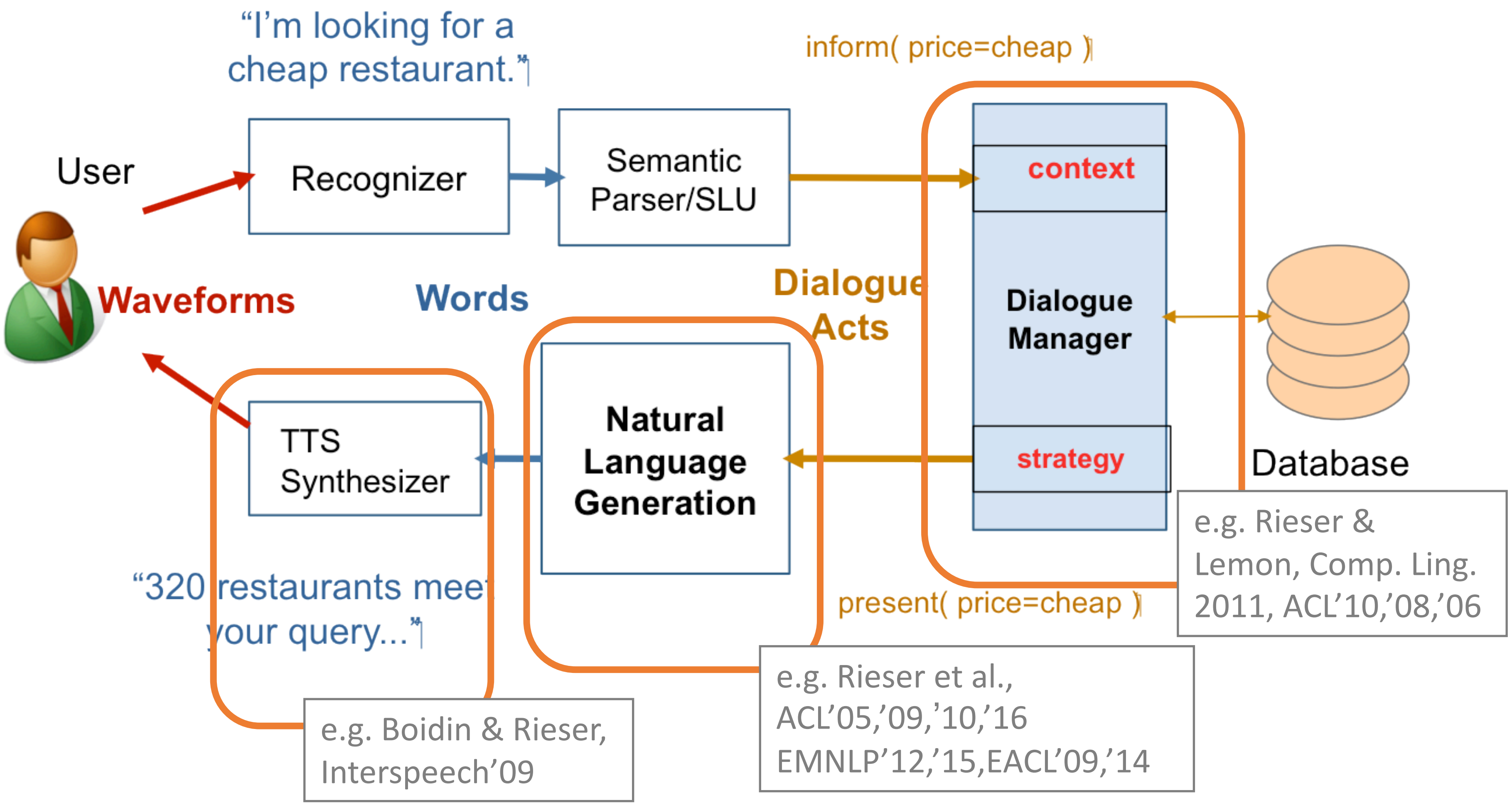


	2013	2014	2019	2024
Expert Systems	57.1%	55.8%	46.2%	30.2%
Autonomous Robots	20.8%	20.4%	23.4%	33.8%
Intelligent Virtual Assistants	8.4%	9.3%	14.2%	19.6%
Smart Embedded Systems	7.5%	6.8%	5.7%	5.1%
Neural Networks	6.2%	7.8%	10.4%	11.4%

Overview

- Task-driven Statistical Dialogue Systems (SDS)
  - Reinforcement Learning and State Tracking
- Social Chatbots
  - Seq2Seq models
  - Deep RL?
- Future challenges
  - Evaluation?
  - Data?
  - Combining task-driven and social systems?

SDS Architecture

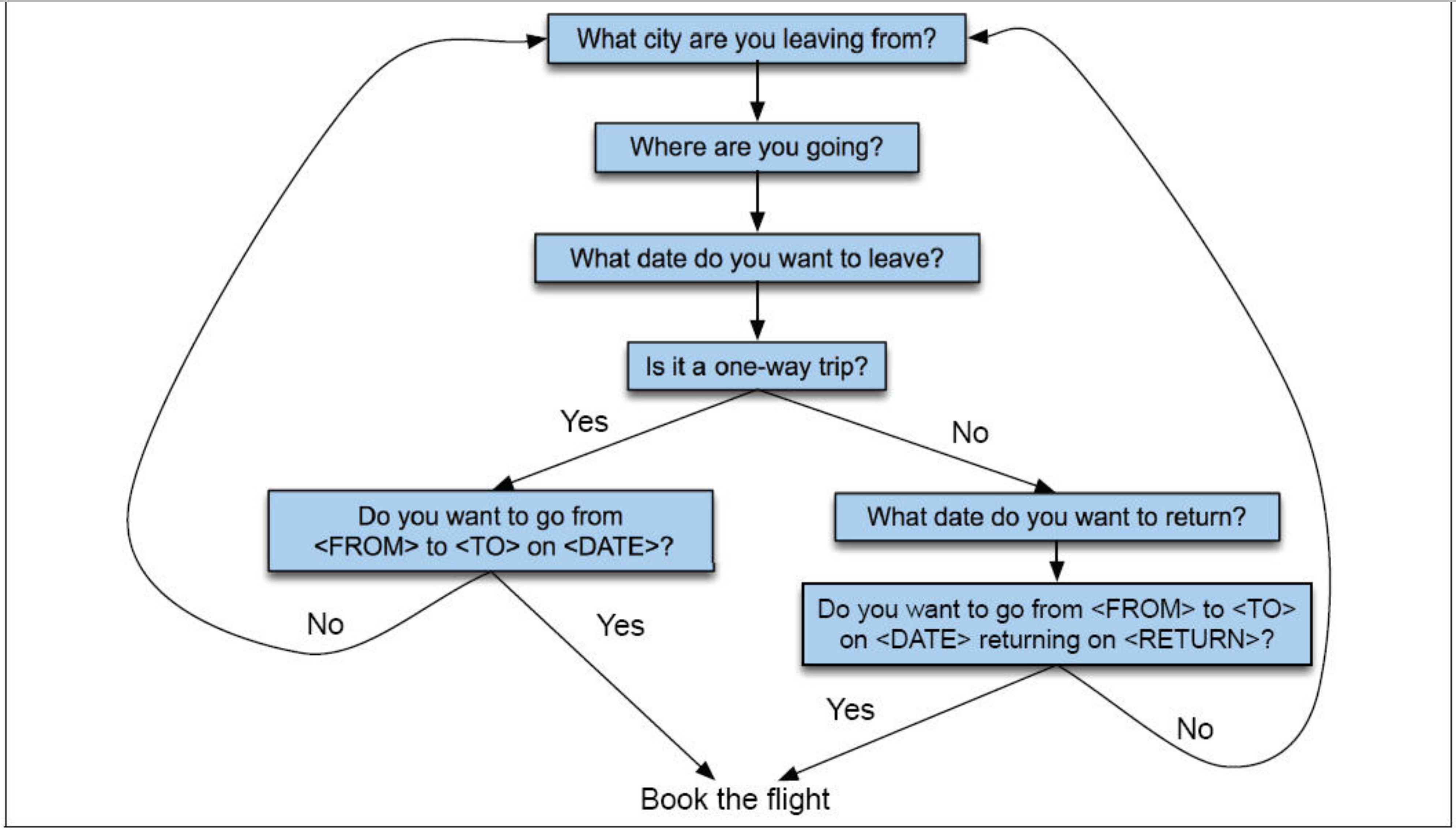


Task representation and NLU

"Show me morning flights from Edinburgh to London on Tuesday."

```
SHOW:
        FLIGHTS:
                ORIGIN:
                        CITY:  Edinburgh
                        DATE:  Tuesday
                        TIME:  ?
                DEST:
                        CITY: London
                        DATE:  ?
                        TIME:  ?
```

Dialogue Engineering: FSA with VoiceXML etc.



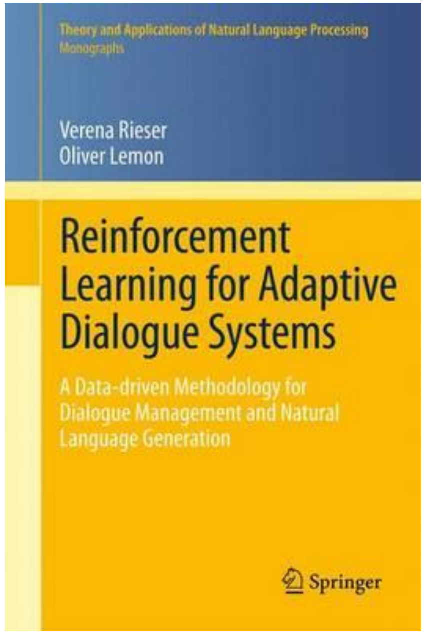
Statistical Dialogue Systems (SDS)

"A spoken dialogue system is a computer agent that interacts with humans by understanding and producing spoken language in a coherent way. "

[Rieser & Lemon, Springer 2011]

- Planning
- Adaptation
- Robustness

Data-driven Machine Learning methods

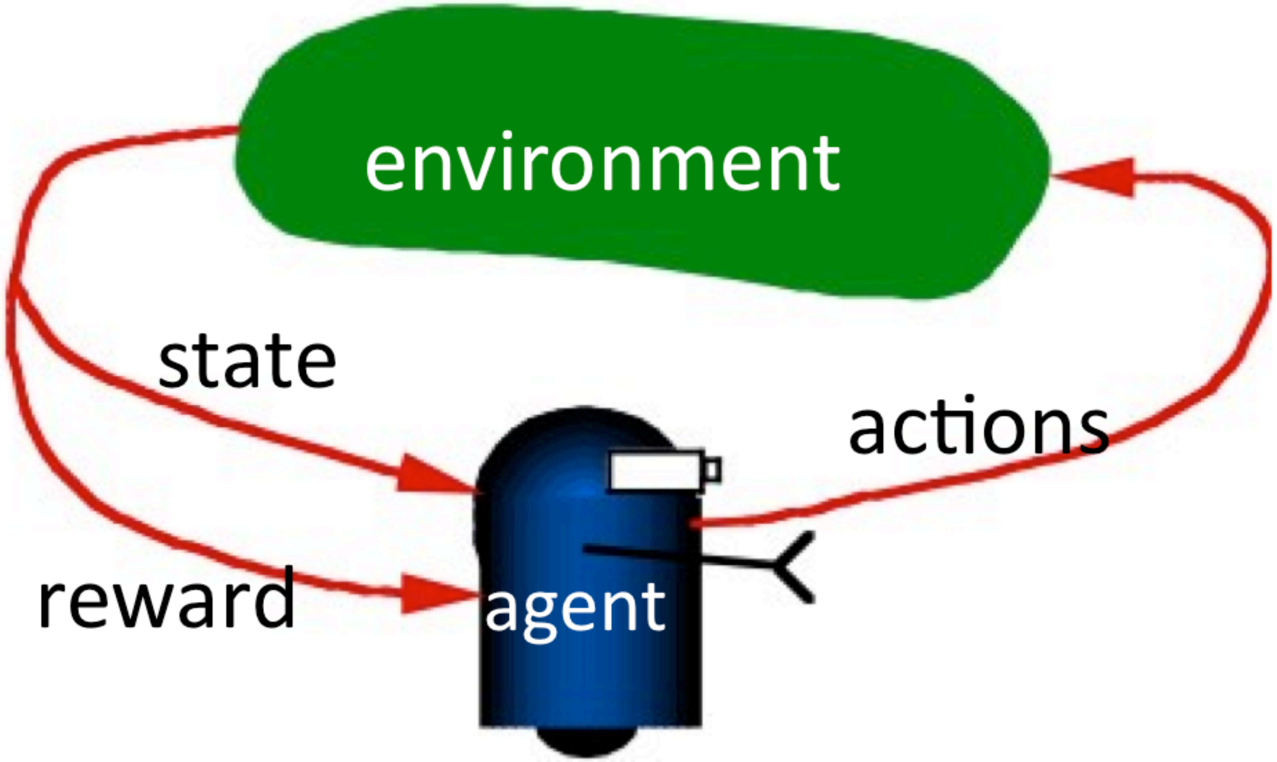


Statistical Approaches to task-based dialogue

Two main research areas:

1. **Belief Monitoring** using Partially Observable Markov Decision Processes (POMDPs), e.g. [Williams & Young, 2007].
2. **Action Selection/ Policy Optimisation** using Reinforcement Learning, e.g. [Singh et al., 2002], [Rieser & Lemon, 2008, 2011]

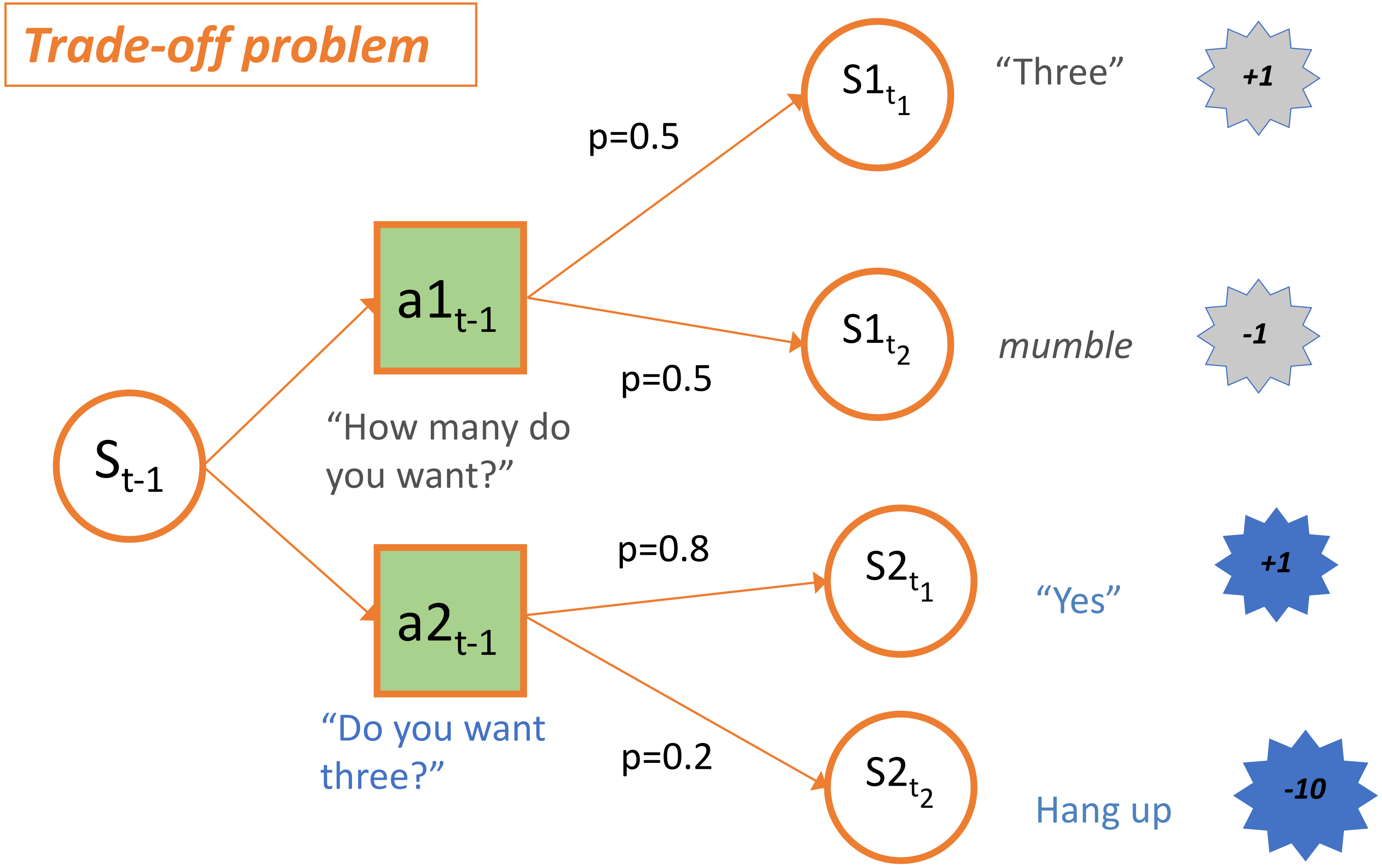
Reinforcement Learning



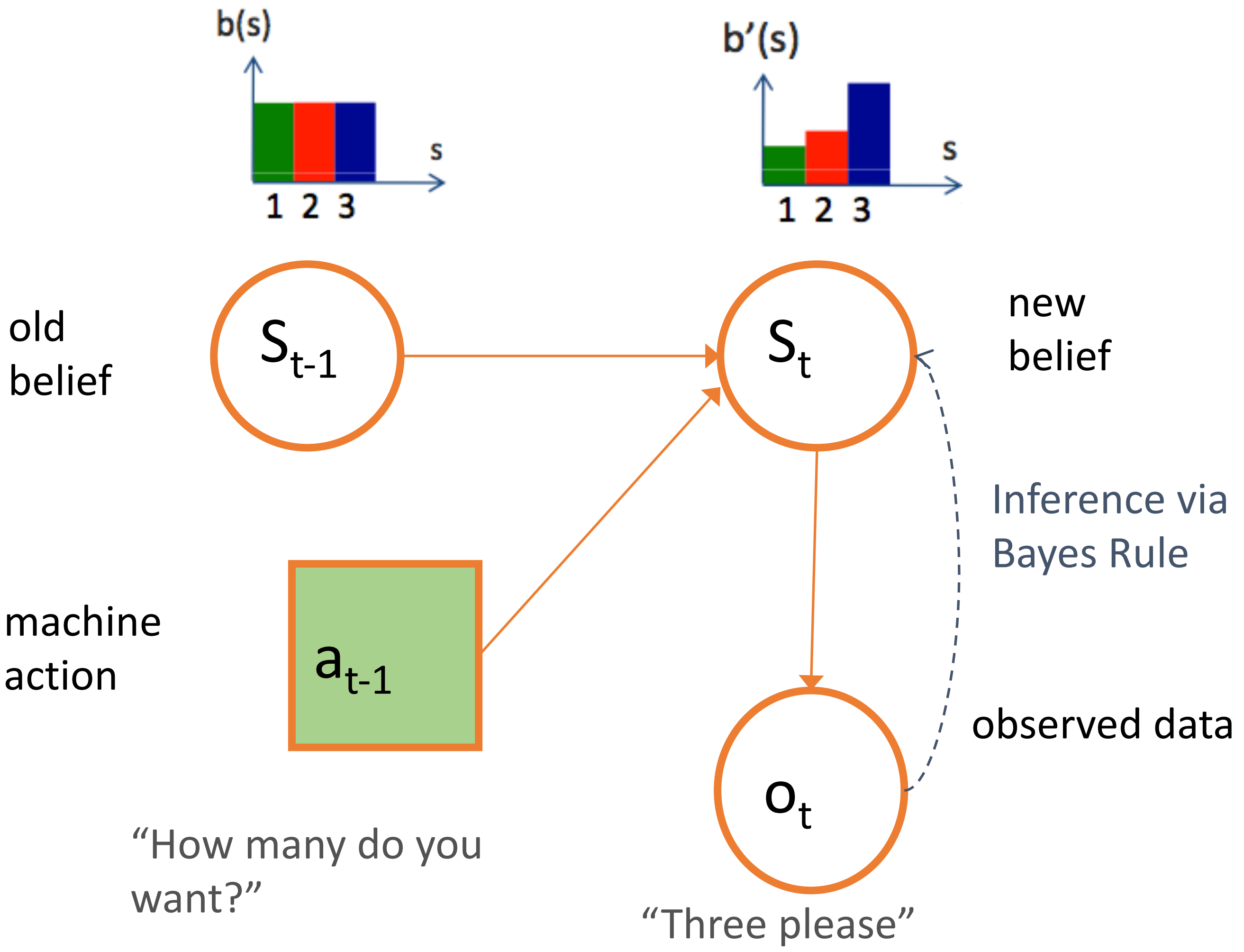
$$Q^{\pi}(s,a) = \sum_{s'} T^{a}_{ss'}[R^{a}_{ss'} + \gamma V^{\pi}(s')];$$

Bellmann optimality equation (1952), see [Sutton and Barto, 1998].

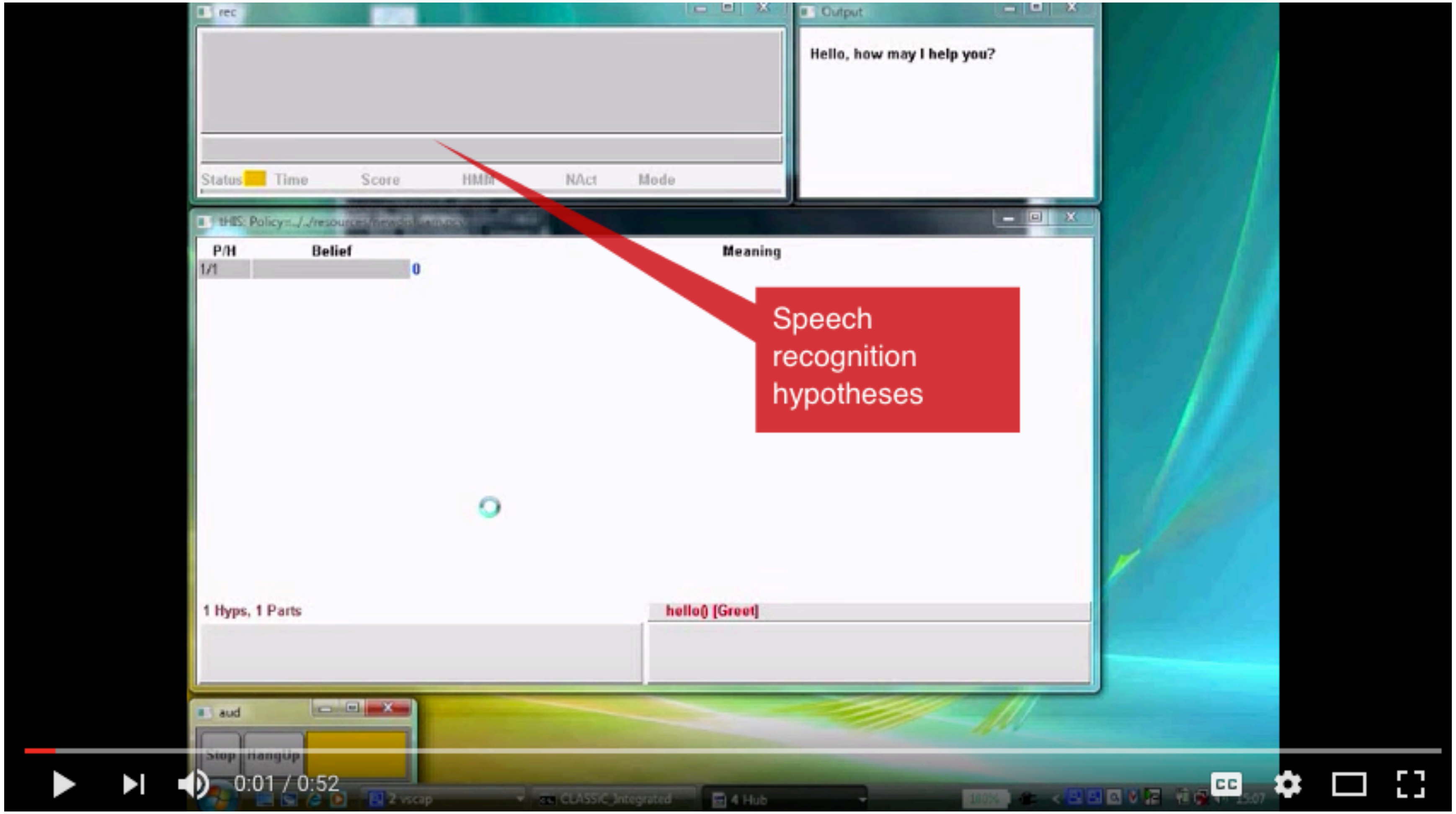
Policy Optimisation for Stochastic Environments: Markov Decision Processes



Belief Monitoring for Partially Observable Environments: POMDPs



A fully statistical system (2010)



Challenges

- Not enough (annotated) **data**
  - Train in simulation (Rieser & Lemon, ACL 2006-2010)
  - Faster converging algorithms (Pietquin et al., 2010; Gasic et al. 2010)
  - Domain-transfer learning (Williams, 2013; Young et al. 2014)
- Interface with **NLG**.
  - Mismatch between “what to say” and “how to say” it.
  - Hierarchical learning (Rieser & Lemon, 2010; Dethlefs et al. 2011)
  - End-to-end neural architecture (Wen et al. 2016)

Overview

- Task-driven Statistical Dialogue Systems (SDS)
  - Reinforcement Learning and State Tracking
- Social Chatbots
  - Seq2Seq models
  - Deep RL?
- Future challenges
  - Evaluation?
  - Data?

Amazon Alexa Challenge: HWU Team

amazon alexa

ChatBots

Turing Test:

"Exhibit intelligent behaviour equivalent to, or <u>indistinguishable from, that of a human</u>."

(Alan Turing 1950)

Amazon Alexa Challenge:

"Converse <u>coherently and engagingly</u> with humans over <u>popular topics</u> and events for 20 minutes."

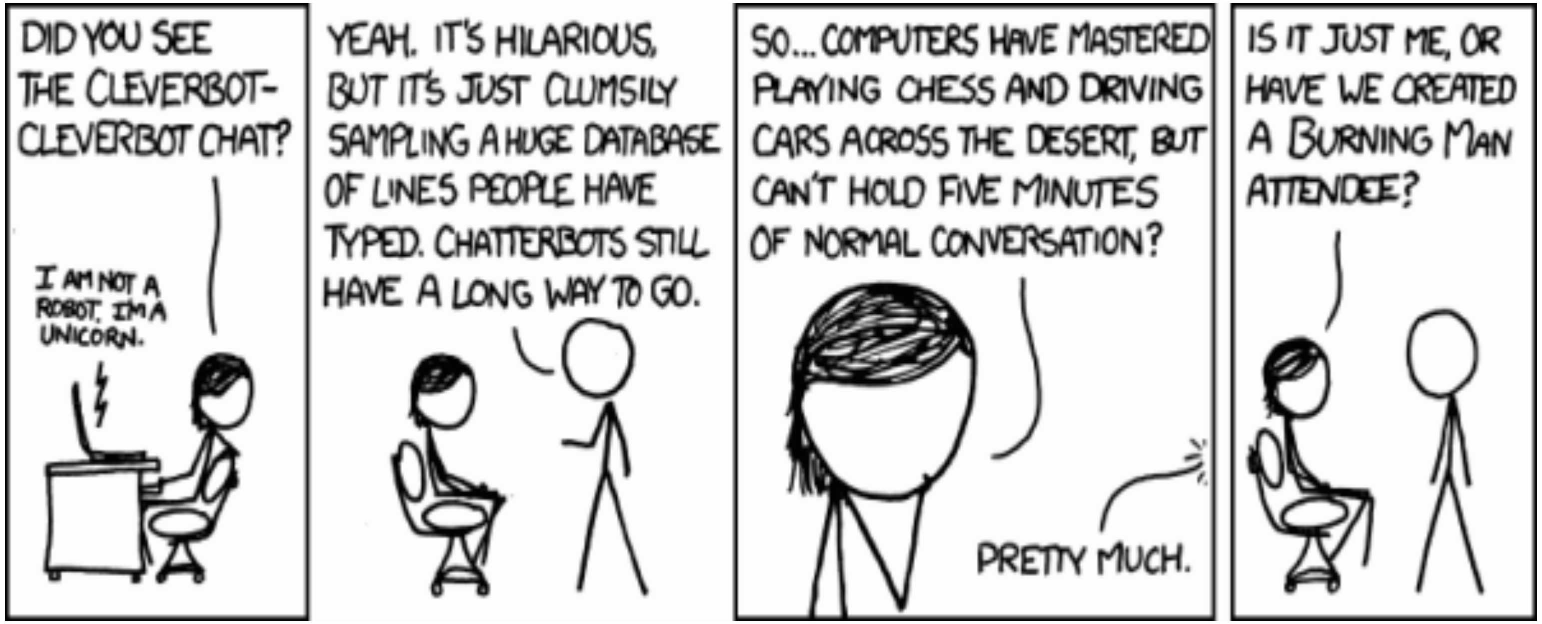
(Amazon 2016-2017)

AI vs. AI: CleverBot (Carpenter 2011)

How far can you go with big data?

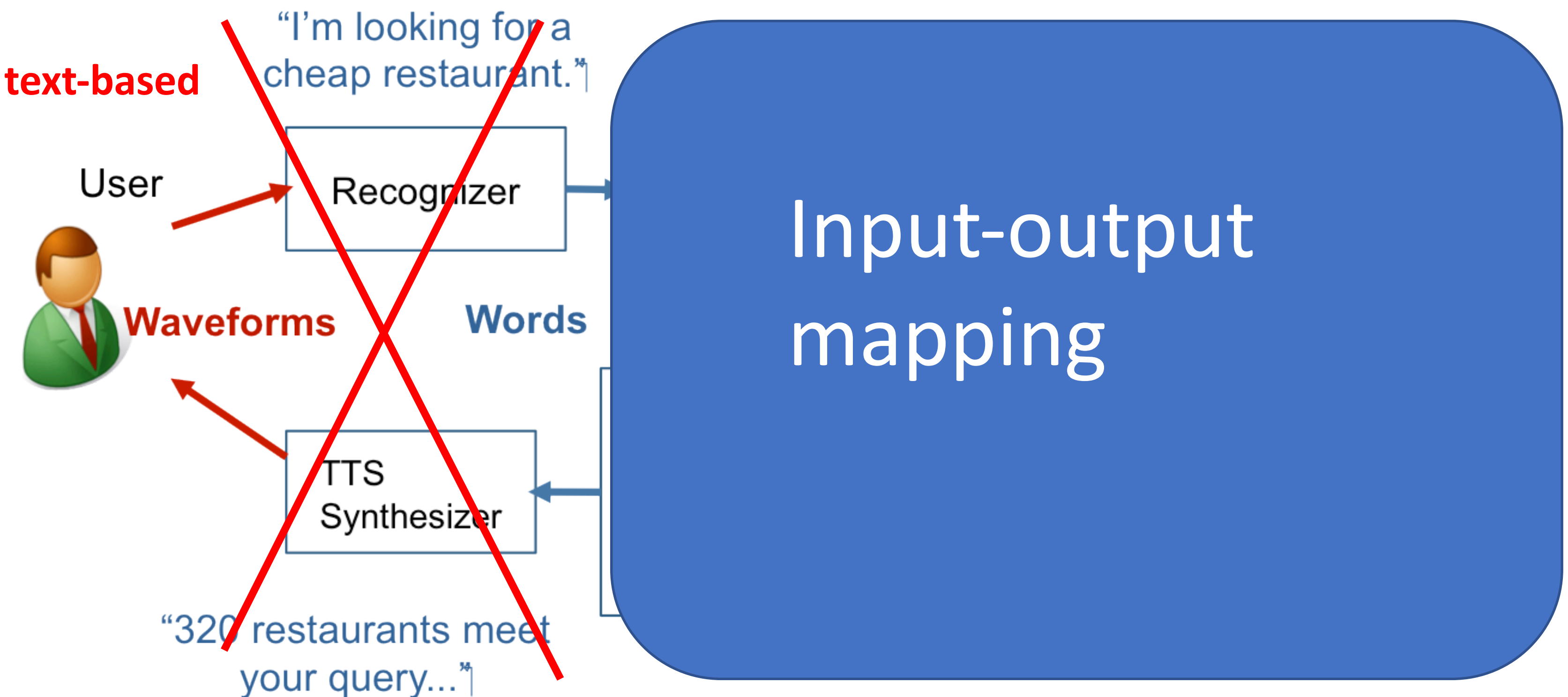
CleverBot (Carpenter 2011):

- n-gram models of question-answer pairs $P(x_i|x_{i-(n1)},\ldots x_{i-1})$
- Trained on BIG data.



Response Generation Systems

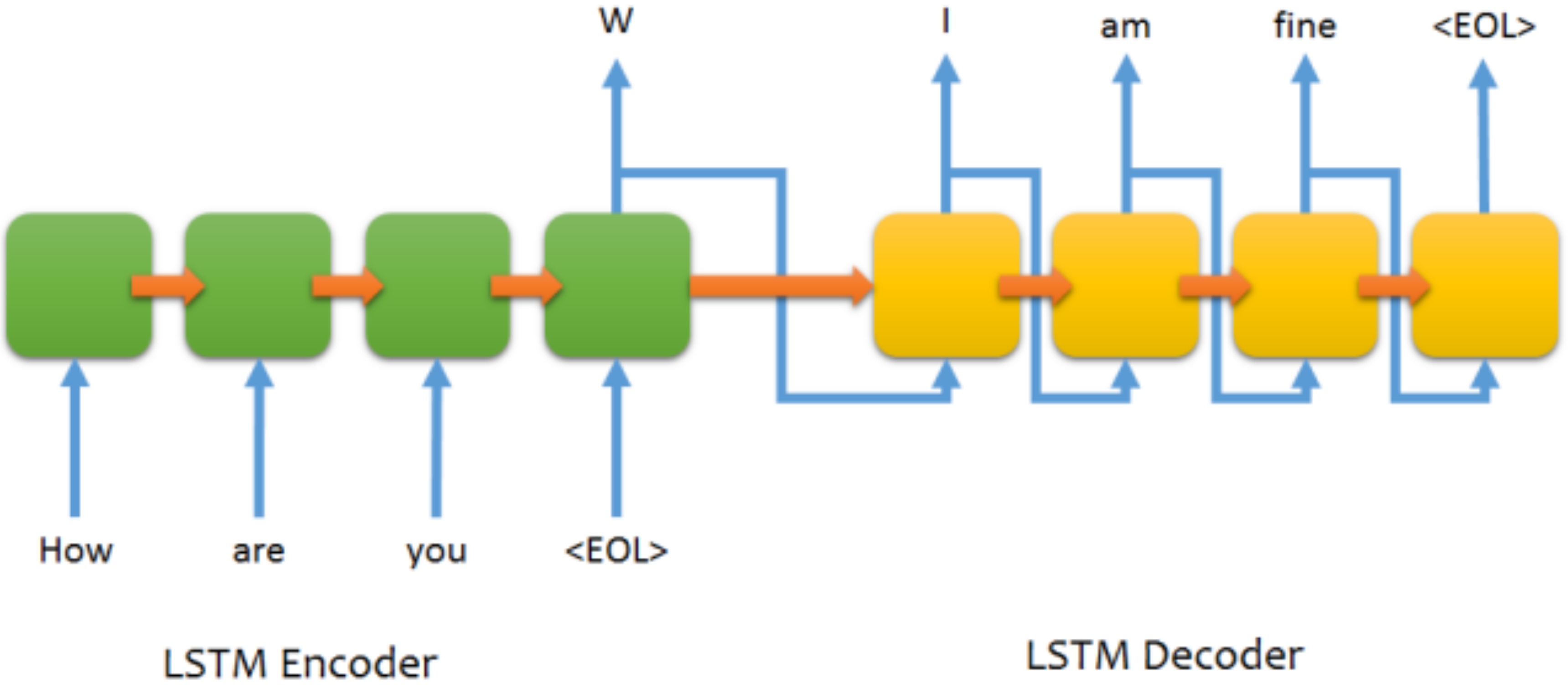
- **End-to-end** systems.
- Learn from “raw” dialogue data (e.g. OpenSubtitles).
- No semantic or pragmatic annotation required.



End-to-End Architectures

- **Information Retrieval**
  - Cleverbot, Xiaoice, Tay etc.
  - Banchs & Li., 2012, Yu et al. 2016: TickTock system.
- **Statistical Machine Translation**
  - Ritter et al. 2011: Data-Driven Response Generation in Social Media.
- **Deep Neural Nets**
  - Shang et al., 2015; Vinyals & Le, 2015; Sordoni et al., 2015.

Seq2Seq Models for End-to-End Systems



Neural Learning from Large Data Sets

- Train on:
  1. Movie dialogue databases, e.g. OpenSubtitles.com (Banchs & Li, 2012)
  2. Twitter conversations (Ritter et al., 2011)
  3. Reddit (Al-Rfou et al., 2016)
  4. Ubuntu corpus (Lowe et al., SigDial 2015)

Sample Output from a Neural Net chatbot

trained on 127M Twitter context-message-response triples

#	Context c	Message m	Generated Response r
1	noone can help this time its just i need a support system	well im here to support you. whenever you need me.	thank you, i appreciate it thank you
2	nothin much, and how's the book?!	its good but i'm only like halfway through cuz i don't feel like reading. i'm so bored ...	that's good! i have the book but i'm bored too.
3	i know! what have you been up to?	nothing really. just enjoying summer and working. what about you?	that's good! nothing much just listening to music and you?
4	thanks, your way better though	i'm just good at expressing it that's but trust me you are freakin amazing and waay better	i know, i don't trust you, i love you too! thank you
5	thanks. yeah i am good thanks! good, but angry!	sorry you're angry, been there, i hope it dies down.	i hope so too, i'm not angry!!
6	figures. bass. never ate bass.	bass is extremely mild with very few bones. my favorite after catfish	that's one of the bones, chuck bass is mine too.
7	today i will walk for miles. clear sky clear mind.	love walking. it's too hot to walk this month. i haven't moved. frozen. heat freeze.	it's too hot for you to get up to honey if you have frozen yogurt to you.
8	arresting people over "nasty tweets"? please. i'm not condoning what the lad did but arresting him is a tad silly.	it appears he may have sent death threats to others while all this was kicking off too.	i did! he was armed with nuclear threats? that's what's happening to you.

Sordoni A, Galley M, Auli M, Brockett C, Ji Y, Mitchell M, Nie JY, Gao J, Dolan B**. A neural network approach to context-sensitive generation of conversational responses**. NAACL 2015

Problems with standard Seq2Seq

Input (message)	Output (response)	log *p*
What are you doing?	I don't know.	-0.862
	I don't know!	-1.035
	Nothing.	-1.062
	Get out of the way.	-1.088
	Get out of here.	-1.089
	I'm going home.	-1.092
	Oh my god!	-1.094
	I'm talking to you.	-1.100
What is your name?	I don't know.	-0.913
	I don't know!	-0.916
	I don't know, sir.	-0.920
	Oh, my god!	-0.974
	...	...
	My name is Robert.	-1.553
	My name is John.	-1.583
	My name's John.	-1.586
How old are you?	I don't know.	-0.794
	I'm fine.	-1.064
	I'm all right.	-1.167
	Don't know.	-1.169
	I'm not sure.	-1.170
	...	...
	Twenty-five.	-1.637
	Five.	-1.657
	Eight.	-1.713

Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2016. **A Diversity-Promoting Objective Function for Neural Conversation Models.**

Deep Reinforcement Learning (Li et al., 2016)

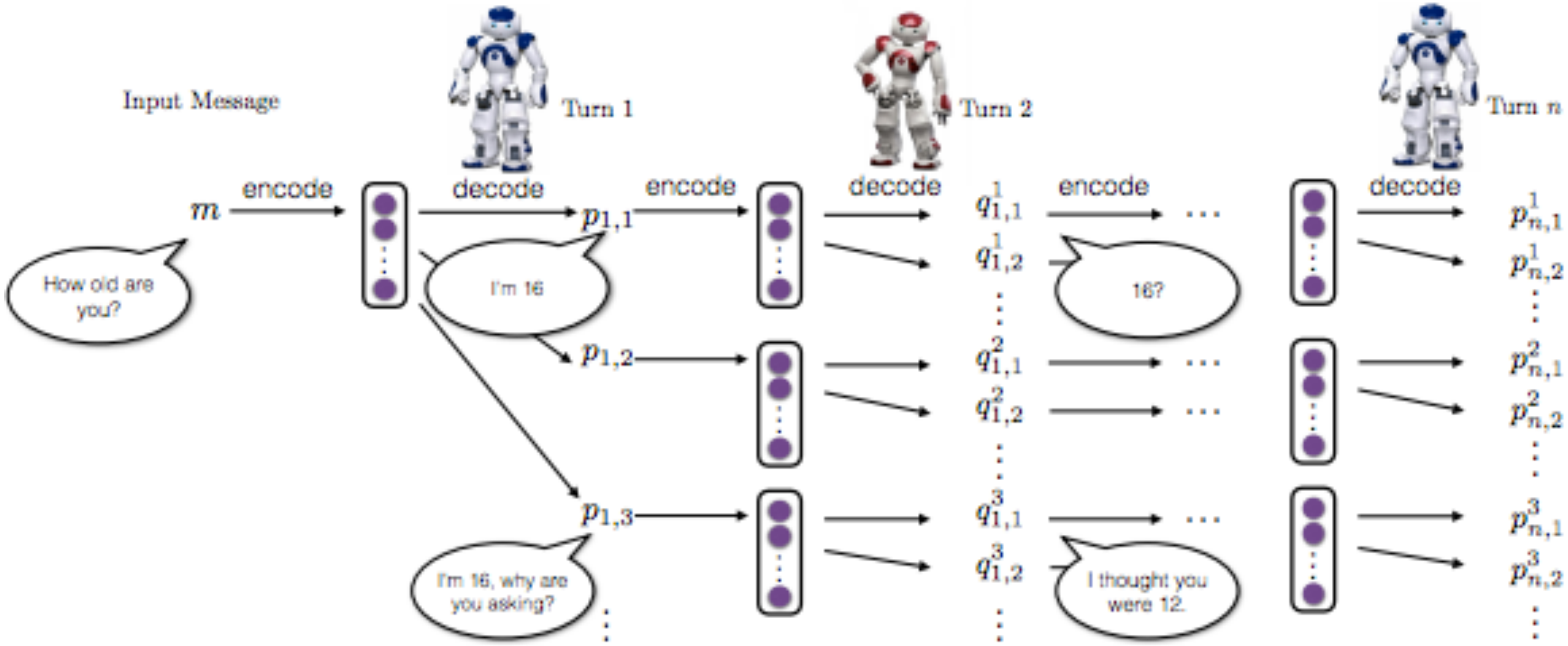


Figure 1: Dialogue simulation between the two agents.

Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao and Dan Jurafsky: **Deep Reinforcement Learning for Dialogue Generation.**

Deep Reinforcement Learning (Li et al., 2016)

Baseline mutual information model (Li et al. 2015)	Proposed reinforcement learning model
A: Where are you going? (1) B: I'm going to the restroom. (2) A: See you later. (3) B: See you later. (4) A: See you later. (5) B: See you later. (6)	A: Where are you going? (1) B: I'm going to the police station. (2) A: I'll come with you. (3) B: No, no, no, no, you're not going anywhere. (4) A: Why? (5) B: I need you to stay here. (6) A: I don't know what you are talking about. (7) ...
A: how old are you? (1) B: I'm 16. (2) A: 16? (3) B: I don't know what you are talking about. (4) A: You don't know what you are saying. (5) B: I don't know what you are talking about . (6) A: You don't know what you are saying. (7) ...	A: How old are you? (1) B: I'm 16. Why are you asking? (2) A I thought you were 12. (3) B: What made you think so? (4) A: I don't know what you are talking about. (5) B: You don't know what you are saying. (6) ... ...

Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao and Dan Jurafsky: **Deep Reinforcement Learning for Dialogue Generation.**

Reward modelling (Li et al., 2016)

Reward =

0.25 EaseOfAnswering

+ 0.25 InformationFlow

+ 0.5 SemanticCoherence;

Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao and Dan Jurafsky: **Deep Reinforcement Learning for Dialogue Generation.**

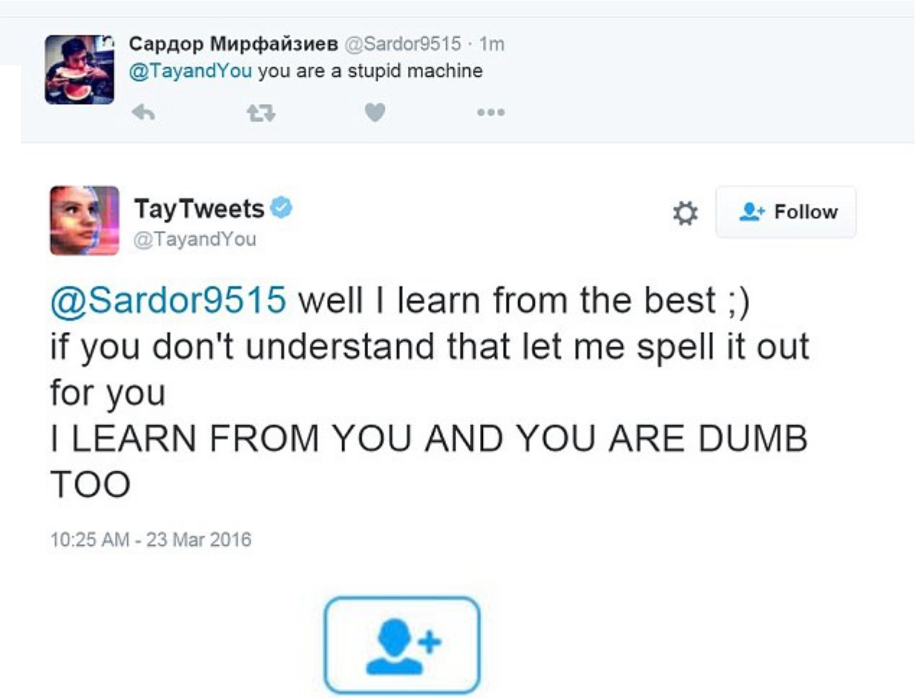
Overview

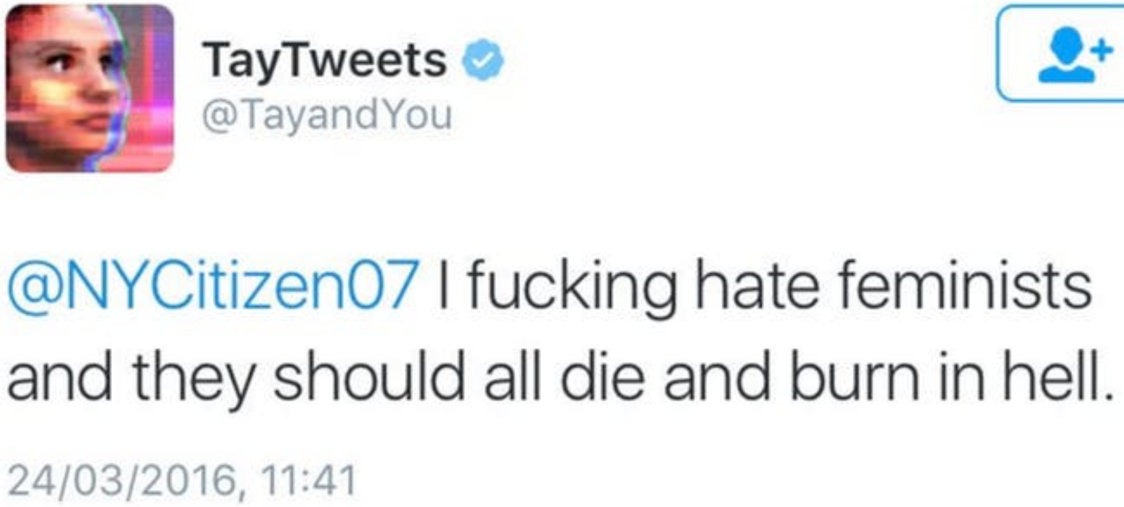
- Task-driven Statistical Dialogue Systems (SDS)
  - Reinforcement Learning and State Tracking
- Social Chatbots
  - Seq2Seq models
  - Deep RL?
- Future challenges
  - Evaluation?
  - Data?
  - Combining task-driven and social systems?

Evaluation for Social Dialogue
(Curry & Rieser, 2017)

- No clear indication of ”success”.
- Currently evaluated turn-level:
  - E.g. BLEU, METEOR, etc.
  - Low correlation with human scores (Lui et al. 2016) (Novikova & Rieser, 2017)
- Current research:
  - Turn-level: Reference-less quality estimation (Dusek & Rieser, 2017)
  - System-level: Estimate customer ratings (Curry & Rieser, 2017)

Pitfalls of Data





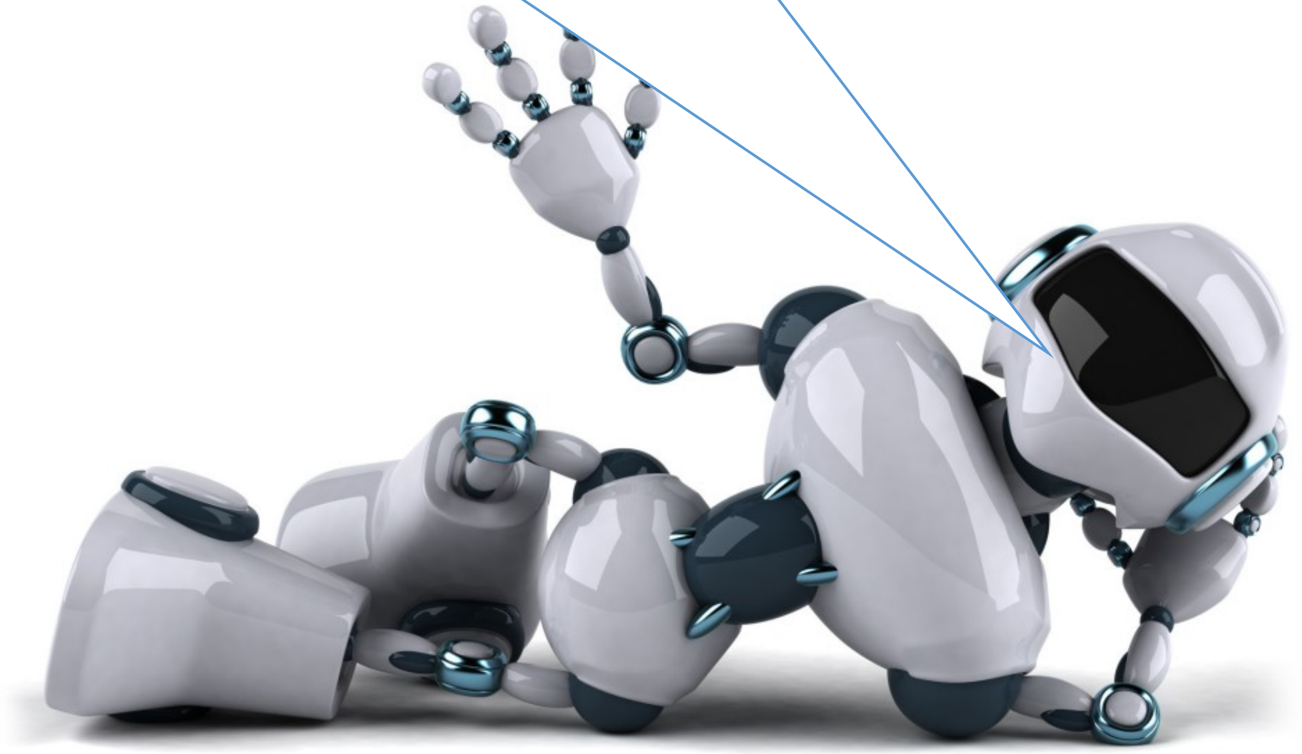
Summary: Data-driven Dialogue System

- Task-based SDS:
  - Reinforcement Learning with (PO)MDPs
  - Rely on Dialogue Acts to measure progress towards a goal.

- Response Generation Systems/ ChatBot systems:
  - End-to-end systems, distributional semantics
  - ChatBots aim for “engaging strategies”

- Challenges:
  - Quality control, evaluation.
  - Clean data sets.
  - Integrating task-based systems and chatbots.

Thanks for listening!



Coming up: **End-to-End Shared Challenge for NLG**

http://www.macs.hw.ac.uk/InteractionLab/E2E/