

# Reinforcement Learning for Adaptive Dialogue Systems

## PART IV: Tools and Future Directions

Oliver Lemon    Verena Rieser

School of Informatics  
University of Edinburgh

For updated course materials see:

<http://sites.google.com/site/olemon/eac109>



EACL tutorial, March 2009

# Outline

Tools

Research directions and open problems

Summary

# Outline

Tools

Research directions and open problems

Summary

# Tools for data collection

- ▶ Wizard-of-Oz experiments [Fraser and Gilbert, 1991].
- ▶ WAMI toolkit (Web-Accessible Multimodal Interfaces)  
<http://wami.csail.mit.edu/>.
- ▶ DUDE rapid prototyping (Dialogue and Understanding Development Environment) [Lemon and Liu, 2006].

# PARADISE: logistic regression for reward modelling I

Any toolkit which allows statistical data analysis like linear regression, curve fitting, . . . , e.g.

- ▶ SPSS
- ▶ Matlab
- ▶ R
- ▶ GNUplot
- ▶ etc.

# RL packages/code I

## Matlab tollboxes:

- ▶ **Perseus for Matlab: randomized point-based approximate value iteration algorithm for Partially Observable Markov Decision Processes (POMDPs).** <http://staff.science.uva.nl/~mtjspan/software/approx/>
- ▶ **MDP Toolbox for Matlab (INRIA)** <http://www.inra.fr/internet/Departements/MIA/T//MDPtoolbox/>
- ▶ **C, Lisp, and Matlab code from [Sutton and Barto, 1998]** <http://www.cs.ualberta.ca/~sutton/book/code/code.html>, [http://waxworksmath.com/Authors/N\\_Z/Sutton/sutton.html](http://waxworksmath.com/Authors/N_Z/Sutton/sutton.html)

# RL packages/code II

## Java, C++:

- ▶ Reinforcement Learning Toolbox 2.0, C++ toolbox (Uni Graz), <http://www.igi.tugraz.at/ril-toolbox/general/overview.html>
- ▶ Java code (Q learning and SARSA), <http://www.cse.unsw.edu.au/~cs9417ml/RL1/sourcecode.html>
- ▶ PIQLE platform in Java, <http://sourceforge.net/projects/piqle/>
- ▶ ...

# Example: REALL toolkit, Edinburgh-Stanford-Link

The screenshot displays the REALL Editor interface for a file named 'DudeDialogue\_3Slots.reall'. The interface includes a menu bar (File, Output, Options), a toolbar with buttons for Setup, Go, Pause, Step, Save Learning, and Terminate, and a status bar.

The main workspace is divided into three panels:

- REALL Plan:** A hierarchical tree view showing the structure of the REALL plan. The root is 'REALL\_DUDE\_Dialogue', which contains a 'REALLScope' with various actions like 'NewStart', 'ResetTheEpisode', 'ConductTheDialogue', 'GreetTheUser', 'UserRequestsHelp', 'UserRequestsEnd', 'UserRequestStartOver', 'AskAQuestion', and 'CloseTheDialogue'. Each action has associated 'Features' and 'Means' (e.g., 'SkillRef').
- Plots (specified from State Tree):** A graph titled 'Global' showing 'Reward' on the y-axis (ranging from -103 to 96) and 'Pens' on the x-axis. A red horizontal line is drawn at a reward value of approximately 96.
- Text Output (from 'Step' or 'Go'):** A window displaying the output of the REALL process, listing features and their values:

```
Feature: lastSystemMove = implConfAskASlot
Feature: prv_no = 0
Feature: prv_slot1 = 1
Feature: prv_slot2 = 0
Feature: prv_slot3 = 1
Feature: prv_yes = 0
Feature: toCloseDialogue = 0
Feature: userGreeted = 1
```

At the bottom, a 'State Tree (from 'Setup' of REALL Plan)' panel shows the current state, which is 'Global'.

to Daniel Shapio and Carl Tollander.

# Dialogue toolkits I

## Finite state:

- ▶ CSLU toolkit, AT&T FSM library
- ▶ Voice XML, BeVocal Cafe (Nuance):  
<http://cafe.bevocal.com/>

## Information state update (ISU):

- ▶ DIPPER: [www.ltg.ed.ac.uk/dipper/](http://www.ltg.ed.ac.uk/dipper/)
- ▶ TRINDIKIT:  
[www.ling.gu.se/projekt/trindi/trindikit/](http://www.ling.gu.se/projekt/trindi/trindikit/)

...also see [McTear, 2004].

# Outline

Tools

Research directions and open problems

Summary

# Overview: Research directions and open issues

- ▶ Tractability, and dimensionality reduction methods, e.g. Summary POMDPs [Williams and Young, 2005]
- ▶ Available corpora, e.g [Rieser and Lemon, 2008b]
- ▶ What makes a good user simulation?, e.g. [Ai and Litman, 2008]
- ▶ Is RL suitable for commercial dialogue strategy development? [Paek and Pieraccini, 2008]
- ▶ RL for Natural Language Generation [Lemon, 2008, Janarthanam and Lemon, 2008, Rieser and Lemon, 2009]
- ▶ See [Lemon and Pietquin, 2007] for further discussion.

# Announcement

**Interspeech 2009**, Brighton, special session on:

“Machine Learning for Adaptivity in Spoken Dialogue Systems”,

<http://www.interspeech2009.org/conference/specialsessions.php>

# Outline

Tools

Research directions and open problems

Summary

# Summary: PART I+II (Oliver Lemon)

## **PART I: Introduction**

1. Introduction to Dialogue Strategy Development
2. Introduction to Reinforcement Learning
3. RL-based dialogue system development

## **Part II: Literature review on RL for DM**

# Summary: PART III +IV (Verena Rieser)

## **PART III: Simulation-based Dialogue Strategy Optimisation**

1. Simulation-based RL
2. Data collection and Corpus requirements
3. Simulated environments for dialogue optimisation
  - ▶ State-action space
  - ▶ Noise model
  - ▶ User simulation
  - ▶ Data-driven reward modelling
4. Policy training and evaluation

## **Part IV: Tools and Future directions**

## Summary: Take home messages

- ▶ DM is a hard problem for human designers.
- ▶ DM has a large decision space, long term effects of actions, stochastic/non-deterministic environment,...
- ▶ RL is better than manually setting thresholds, hand coding, e.g. [Rieser and Lemon, 2008a].
- ▶ Simulation-based training and testing allows effective “system-in-the-loop” development.
- ▶ RL promises advances in *adaptive* and *robust* HCI.
- ▶ Principled mathematical framework for dialogue management.



Ai, H. and Litman, D. (2008).

Assessing dialog system user simulation evaluation measures using human judges.

*In Proc. of the 21st International Conference on Computational Linguistics and 46th Annual Meeting of the Association for Computational Linguistics (ACL/HLT).*



Fraser, N. M. and Gilbert, G. N. (1991).

Simulating speech systems.

*Computer Speech and Language*, 5:81–99.



Janarthanam, S. and Lemon, O. (2008).

User simulations for online adaptation and knowledge-alignment in troubleshooting dialogue systems.

*In Proc. of the 12th SEMdial Workshop on on the Semantics and Pragmatics of Dialogues.*



Lemon, O. (2008).

Adaptive natural language generation in dialogue using Reinforcement Learning.

*In Proc. of the 12th SEMdial Workshop on on the Semantics and Pragmatics of Dialogues.*



Lemon, O. and Liu, X. (2006).

DUDE: a dialogue and understanding development environment, mapping business process models to Information State Update dialogue systems.

*In Proc. of the Conference of the European Chapter of the ACL (EACL).*



Lemon, O. and Pietquin, O. (2007).

Machine Learning for spoken dialogue systems.

*In Proc. of the International Conference of Spoken Language Processing (Interspeech/ICSLP).*



McTear, M. F. (2004).  
*Towards the Conversational User Interface.*  
Springer Verlag.



Paek, T. and Pieraccini, R. (2008).  
Automating spoken dialogue management design using machine learning: An industry perspective.  
*Speech Communication, Special Issue on Evaluating New Methods and Models for Advanced Speech-Based Interactive Systems*, 50(8-9):716–729.



Rieser, V. and Lemon, O. (2008a).  
Does this list contain what you were searching for? Learning adaptive dialogue strategies for Interactive Question Answering.  
*J. Natural Language Engineering*, 15(1).



Rieser, V. and Lemon, O. (2008b).  
Learning effective multimodal dialogue strategies from Wizard-of-Oz data: Bootstrapping and evaluation.  
*In Proc. of the 21st International Conference on Computational Linguistics and 46th Annual Meeting of the Association for Computational Linguistics (ACL/HLT).*



Rieser, V. and Lemon, O. (2009).  
Natural Language Generation as Planning Under Uncertainty for Spoken Dialogue Systems.  
*In Proc. of the Conference of the European Chapter of the ACL (EACL).*



Sutton, R. and Barto, A. (1998).  
*Reinforcement Learning.*  
MIT Press.



Williams, J. and Young, S. (2005).

Scaling up POMDPs for dialogue management: The “Summary POMDP” method.

In *Proc. of the IEEE workshop on Automatic Speech Recognition and Understanding (ASRU)*.