SimpleDS: A Simple Deep Reinforcement Learning Dialogue System

Anton Karazeev

Department of Innovation and High Technologies

MIPT, 2017

Source
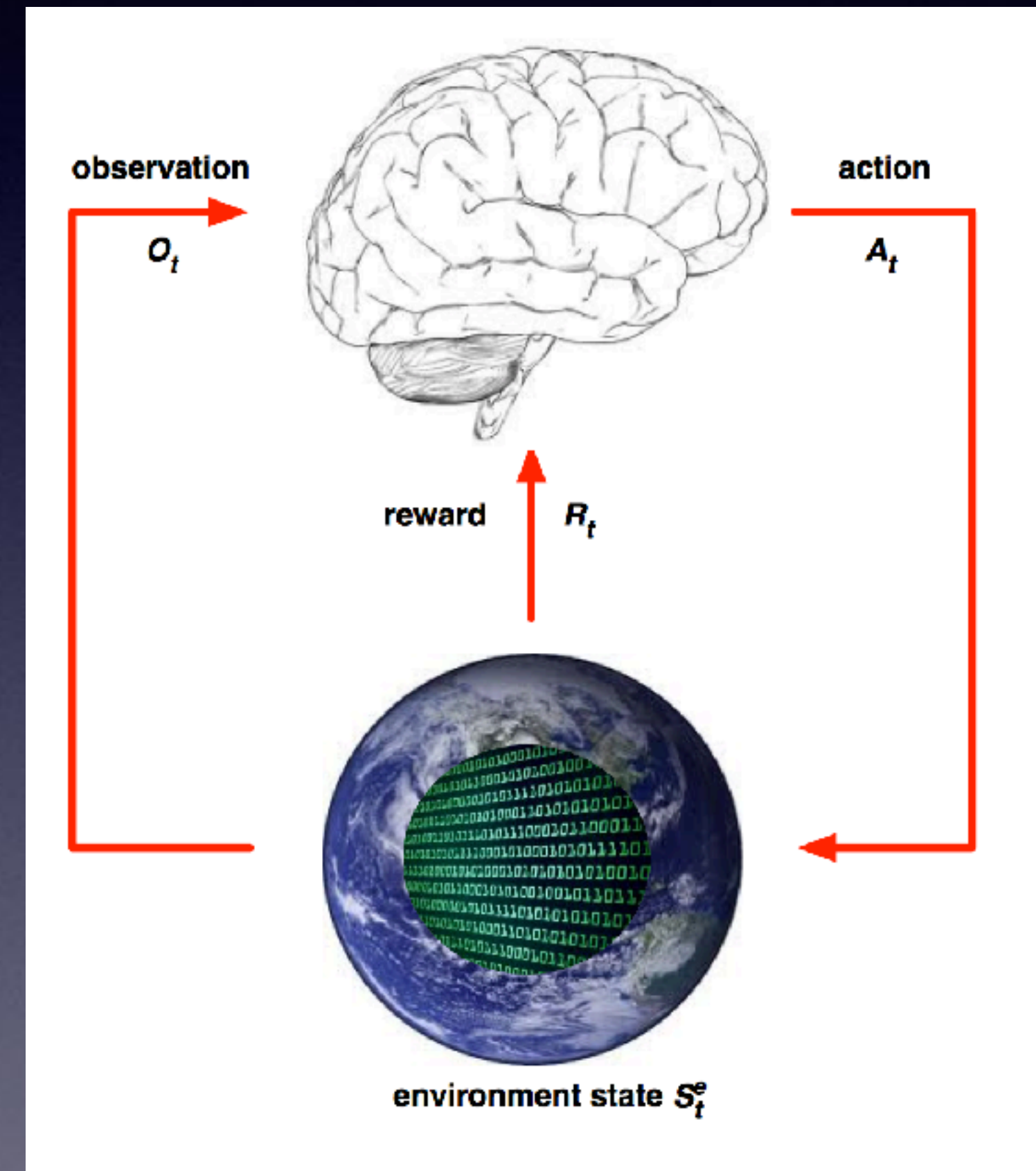
- H. Cuayáhuitl. SimpleDS: A Simple Deep Reinforcement Learning Dialogue System. International Workshop on Spoken Dialogue Systems (IWSDS), 2016. [https://arxiv.org/abs/1601.04574]

Plan

- Reinforcement Learning for Dialogue Systems

- Introduction to Reinforcement Learning

- Structure of *SimpleDS*

- Results

Reinforcement Learning for Dialogue Systems

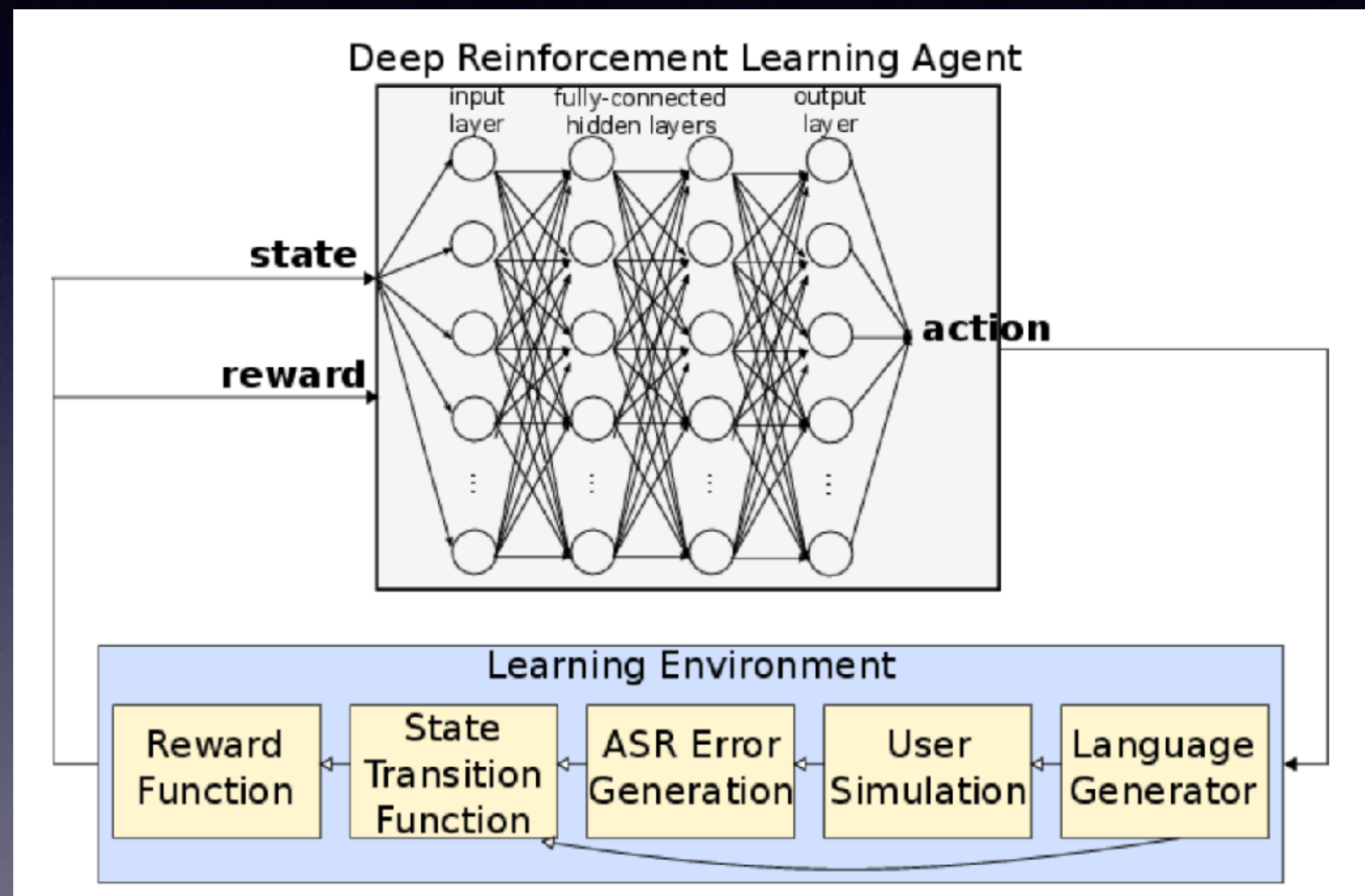| Environment State	Action	Verbalisation
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0	Salutation (greeting)	Hello!
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0	Request (food,price,area)	What type of food, price range, and area are you looking for? [reasonably priced mexican ...food in the east of town]
0,0,0,1,0,1,1,0,0,0,0,0,0,0,0,0,0.53,0,0,0,0,0,0.63,1,0,0,0,0,0,0.83,0,0,0,0,0,0,0,0,1,0,0, 0,0.78,0,0,0,0,0,0.31,0,0,0,1,0.36,1,0.25,0,0,0,0,0,0,0,0,0.57,0,0,0.82,1,0,0,0,0,1,0,0,0,0,0,1	ImpConfirm (food,price,area)	Okay, reasonably priced mexican food in the east.
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0, 0,0,0,1,0,0,0,0,0,1,0,0,0,1,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0	Retrieve (info)	Let me see.
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0, 0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0	Provide (known)	Restaurant X is an excellent choice. It is located in Y.
1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0, 0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,0,1,0,0,0,0,1,0,0,0,0,0	AskFor (more)	Anything else? [no]
0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0.72,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0	Salutation (closing)	Okay, talk to you soon. Bye!
0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0, 0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1		

Introduction to Reinforcement Learning

- Agent and Environment

- Observation and Action

- Reward

Structure of *SimpleDS*

- Diagram of *SimpleDS*

- Setup of learning

- Code at GitHub: https://github.com/cuayahuitl/SimpleDS

H. Cuayáhuitl (2016) [https://arxiv.org/abs/1601.04574]

Results

H. Cuayáhuitl (2016) [https://arxiv.org/abs/1601.04574]
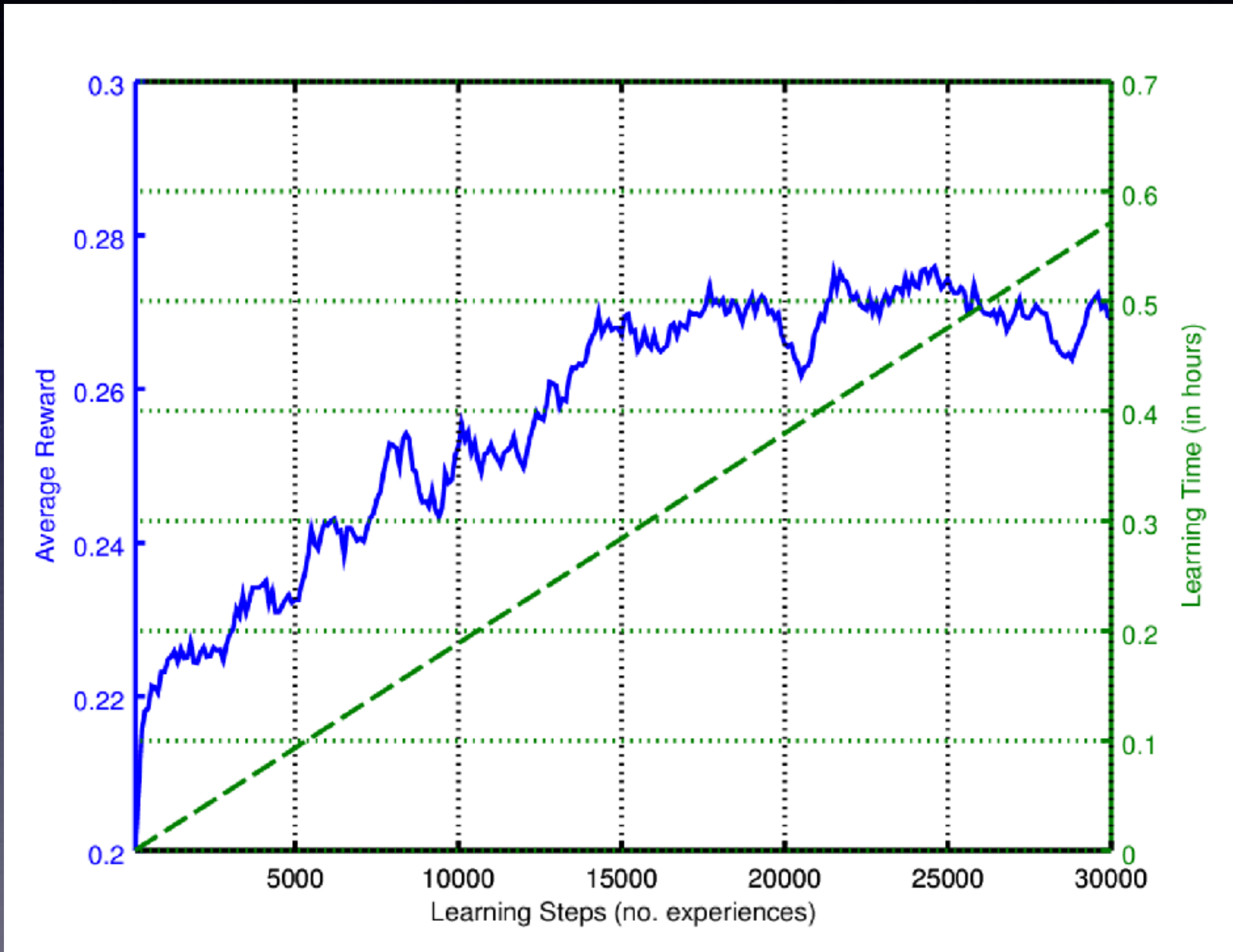
Summary

- Reinforcement Learning for Dialogue Systems

- Introduction to Reinforcement Learning

- Structure of *SimpleDS*

- Results

H. Cuayáhuitl (2016) [https://arxiv.org/abs/1601.04574]

Conclusion

[human]: "What do we want?"

[bot   ]: "Chatbots!"

[human]: "When do we want them?"

[bot   ]: "Sorry, I didn't understand your request."

–ChatbotConf 2016