UiO : University of Oslo

Dialogue Management with Probabilistic Rules

Pierre Lison
Language Technology Group
University of Oslo

KTH, December 2015

Some background

- Started working on human-robot interaction and spoken dialogue systems at the DFKI (Saarbrücken)
- Moved to Oslo in 2011, where I continued my work on dialogue management (PhD defended in 2014)
- Since last year, I also have a postdoctoral project on dialogue modelling for statistical machine translation
- Going to focus on my *dialogue management* work for this talk
    - But if you are interested to know more on my machine translation project, we can talk about this later!

Outline for this talk

- The dialogue management task
- A hybrid logical/probabilistic approach
  - Probabilistic rules
  - Parameter estimation
  - Experiments
- Three open research questions

Outline for this talk

- The dialogue management task
- A hybrid logical/probabilistic approach
  - Probabilistic rules
  - Parameter estimation
  - Experiments
- Three open research questions

What is dialogue management?

- A component in (spoken) dialogue systems
- In charge of "managing" the interaction
  - Maintain a representation of the current *state* of the dialogue
  - Select the next system *actions* based on this state
  - *Predict* how the interaction is going to unfold
- Two intertwined challenges:
  - Dialogue is *complex* (many contextual factors to capture)
  - Dialogue is *uncertain* (ambiguities, unexpected events, etc.)

Existing DM techniques

	Logical approaches	**Statistical approaches**
+	Fine-grained control of conversation	Robust, data-driven models of dialogue
-	Limited account for uncertainties	Need large quantities of training data

A new **hybrid** modelling framework based on *probabilistic rules*

Outline for this talk

- What is dialogue management?
- **A hybrid logical/probabilistic approach**
  - **Probabilistic rules**
  - Parameter estimation
  - Experiments
- Three open research questions

The key idea

- We start with the usual ideas of probabilistic dialogue modelling:
  - Dialogue state encoded as a **Bayesian Network**
  - Each variable captures a *relevant aspect of the interaction* (dialogue history, user intentions, context, etc.)
  - The dialogue state is regularly updated with *new observations* (spoken inputs, new events), according to domain-specific probabilistic models
  - ... and used to determine the *next actions* to execute, according to domain-specific utility models

The key idea

But:

- instead of expressing the domain models using traditional formats (e.g. probability tables)...
- ... we adopt a high-level representation based on **probabilistic rules.**
- The probabilistic rules provide an *abstraction layer* on top of probabilistic (graphical) models

Less parameters to estimate (=easier to learn from small amounts of data)

Can express expert knowledge in human-readable form

Two types of rules

	Probability rules	**Utility rules**
What they encode:	*Conditional probability distributions* between state variables	*Utility functions* for system actions given state variables
General structure:	**if** (condition$_1$) **then** P(effect$_1$)= θ_1, P(effect$_2$)= θ_2, ... **else if** (condition$_2$) **then** P(effect$_3$) = θ_3, ...	**if** (condition$_1$) **then** U(action$_1$)= θ_1, U(action$_2$)= θ_2, ... **else if** (condition$_2$) **then** U(action$_3$) = θ_3,... ...

Examples of probabilistic rules

$\forall$ x,
if (*last-user-act* = x $\wedge$ *last-system-act* = AskRepeat) **then**
P(*next-user-act* = x) = 0.9

"If the system asks the user to repeat his last dialogue act *x*, the user is predicted to comply and repeat *x* with probability 0.9"

$\forall$ x,
if (*last-user-act*=Request(x) $\wedge$ x $\in$ *perceived-objects*) **then**
U(*system-act*=PickUp(x)) = +5

"If the user asks the system to pick up a given object *x* and *x* is perceived by the system, then the utility of picking up *x* is 5"

Rule instantiation

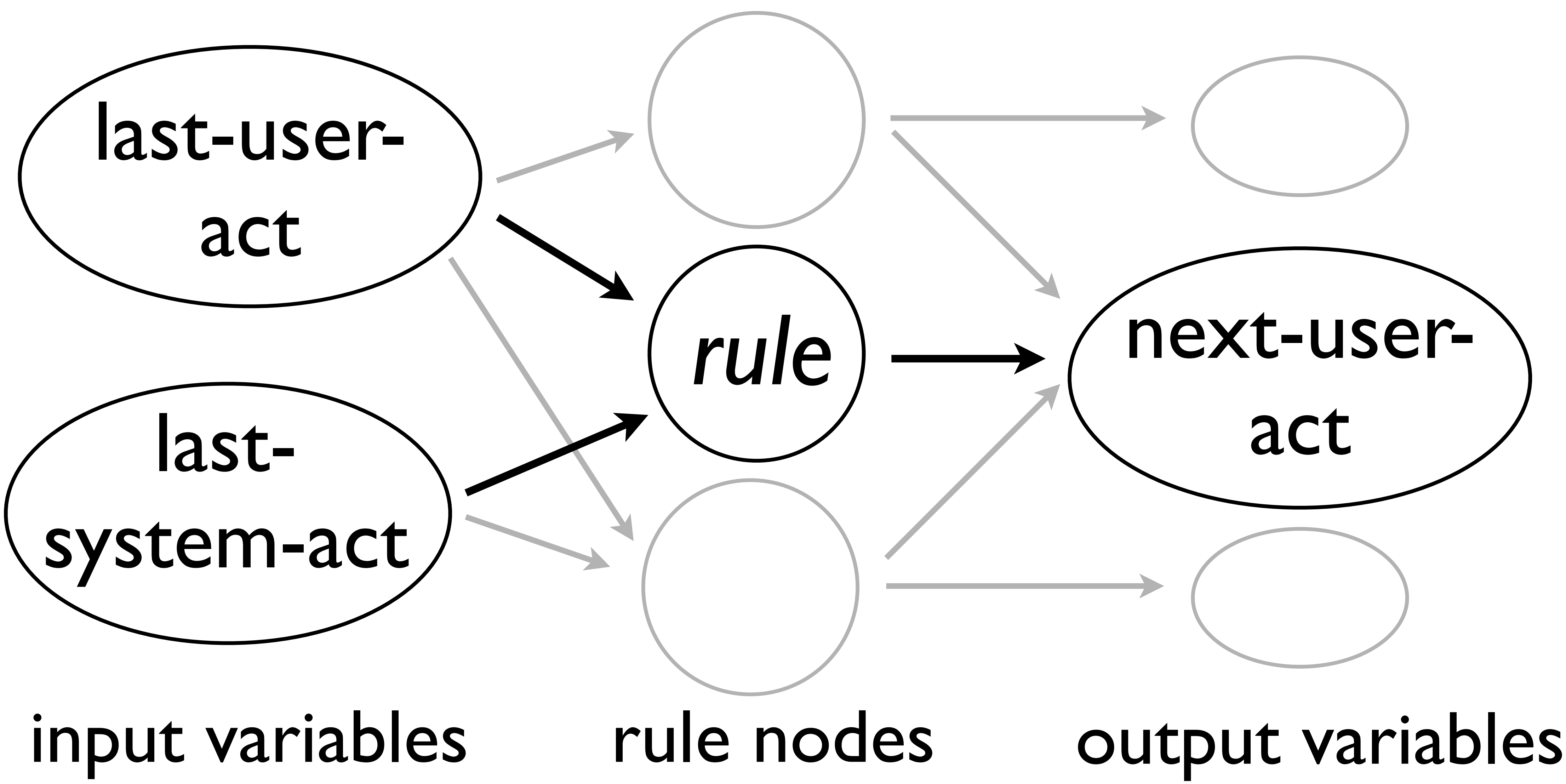
- At runtime, the rules are "executed" by instantiating them in the dialogue state:
  - The rules can be seen as "high-level templates" for the generation of a classical probabilistic model
  - Inference (for state update and action selection) is then performed on this grounded representation
- The use of *logical abstractions* allows us to capture complex relations between variables in a compact, human-readable form

Instantiation of probability rules

∀ x,

if (*last-user-act* = *x* ∧ *last-system-act* = AskRepeat) **then**

P(*next-user-act* = *x*) = 0.9

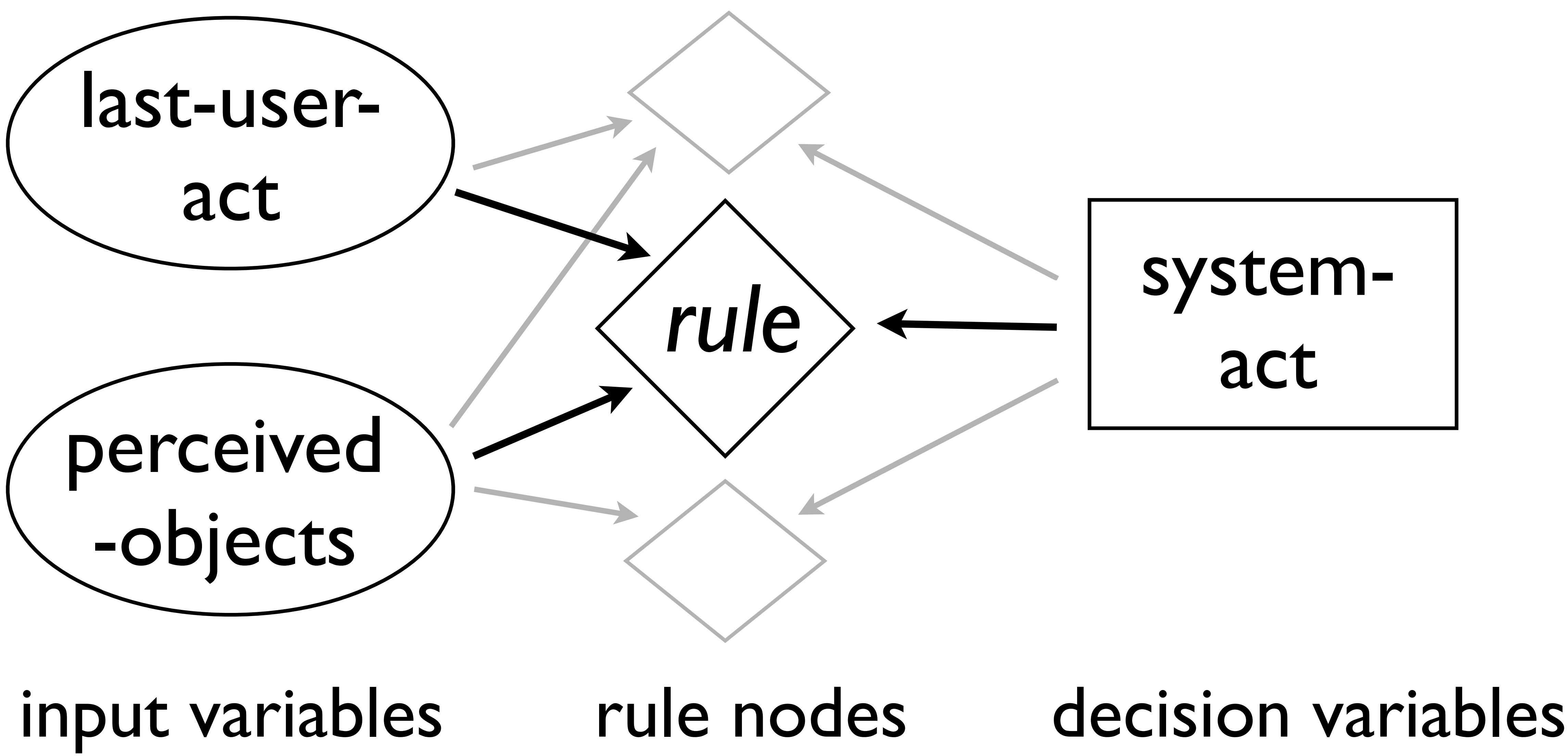


Instantiation of utility rules

∀ x,

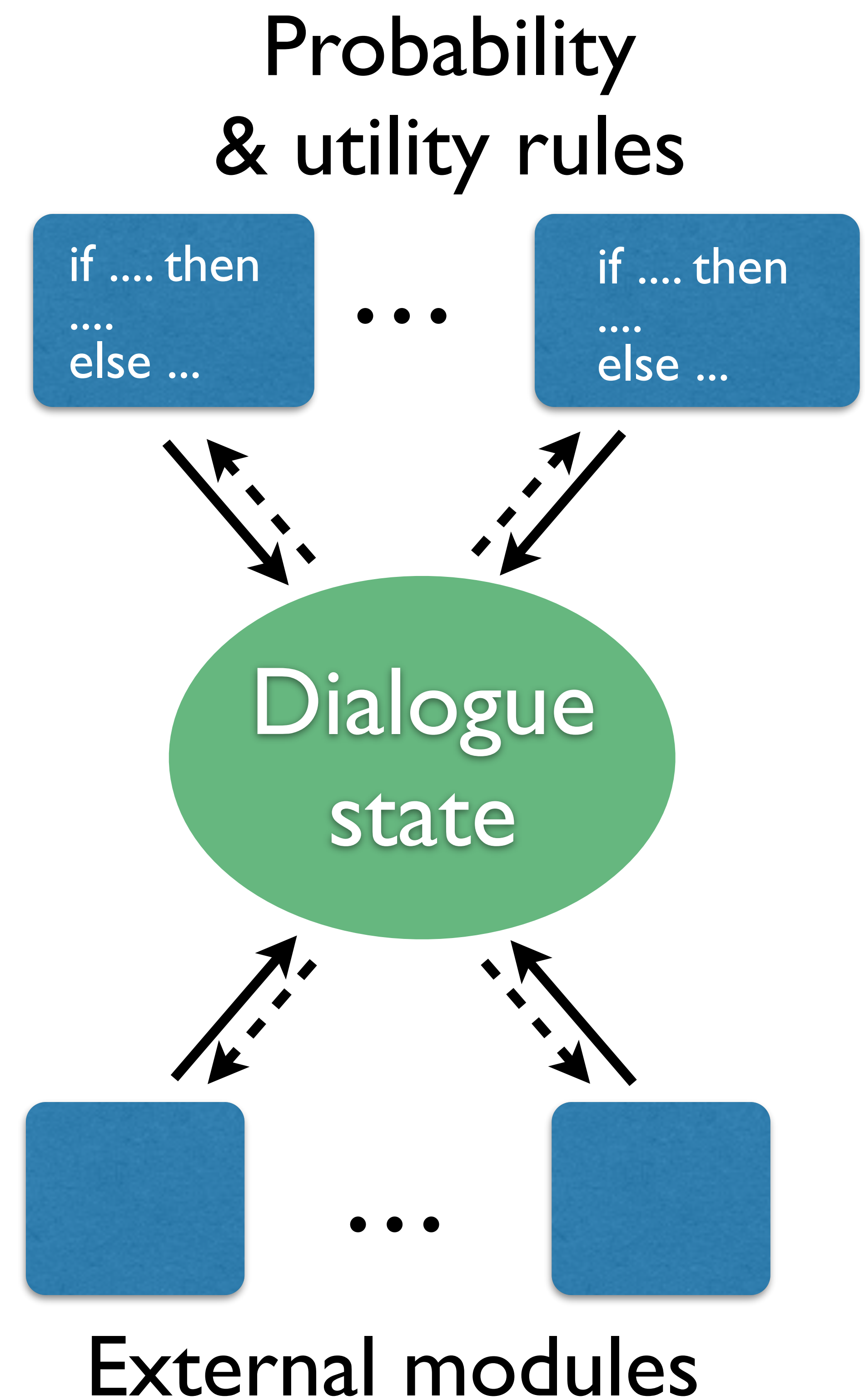
if (*last-user-act*=Request(x) ∧ x ∈ *perceived-objects*) **then**

U(*system-act*=PickUp(x)) = +5



Processing workflow

- *Information state architecture*, with the dialogue state expressed as a Bayesian Network
- External modules add new observations
- Probability rules employed to update the dialogue state (following the new observations)
- Utility rules employed to determine the system actions
- Implementation: OpenDial toolkit [http://www.opendial-toolkit.net]



Outline for this talk

- What is dialogue management?
- **A hybrid logical/probabilistic approach**
  - Probabilistic rules
  - **Parameter estimation**
  - Experiments
- Three open research questions

Parameter estimation

- Probabilistic rules may include parameters (unknown probabilities or utilities)
- **Bayesian learning** approach:
  - Start with initial prior over possible parameter values
  - Refine the distribution given the observed data $\mathcal{D}$

$$P(\boldsymbol{\theta} \mid \mathcal{D}) = \eta \, P(\mathcal{D}\,;\boldsymbol{\theta}) \, P(\boldsymbol{\theta})$$

Posterior distribution

Normalisation factor

Likelihood of the data

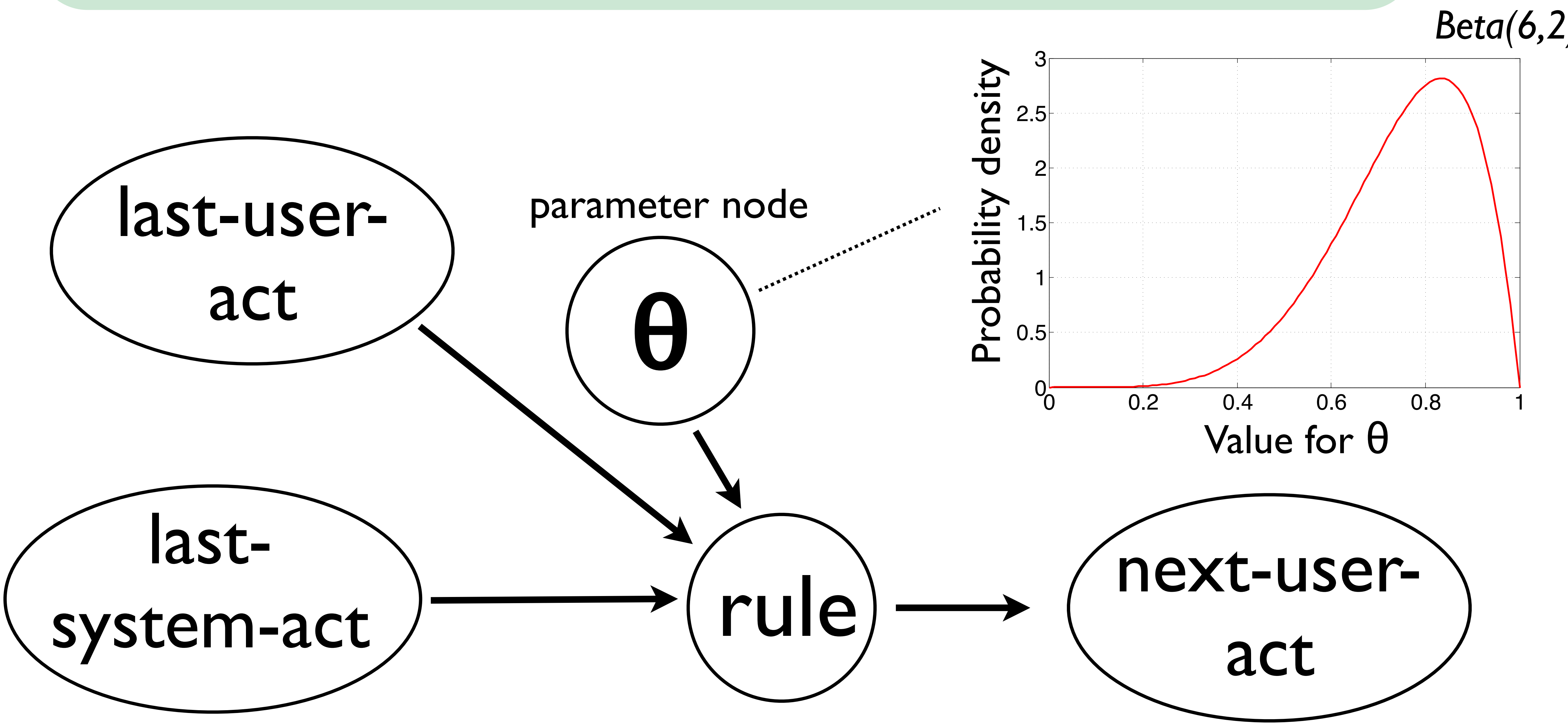
Prior distribution

Parameter estimation

∀ x,

if (*last-user-act* = x ∧ *last-system-act* = AskRepeat) **then**

P(*next-user-act* = x) = θ



Learning paradigms

- Different types of training data:
  - *Supervised learning*: Wizard-of-Oz interactions

Goal: find the parameter values that best "imitate" the Wizard's conversational behaviour

  - *Reinforcement learning*: real or simulated interactions

Goal: find the parameter values that provide the best fit for the collected observations

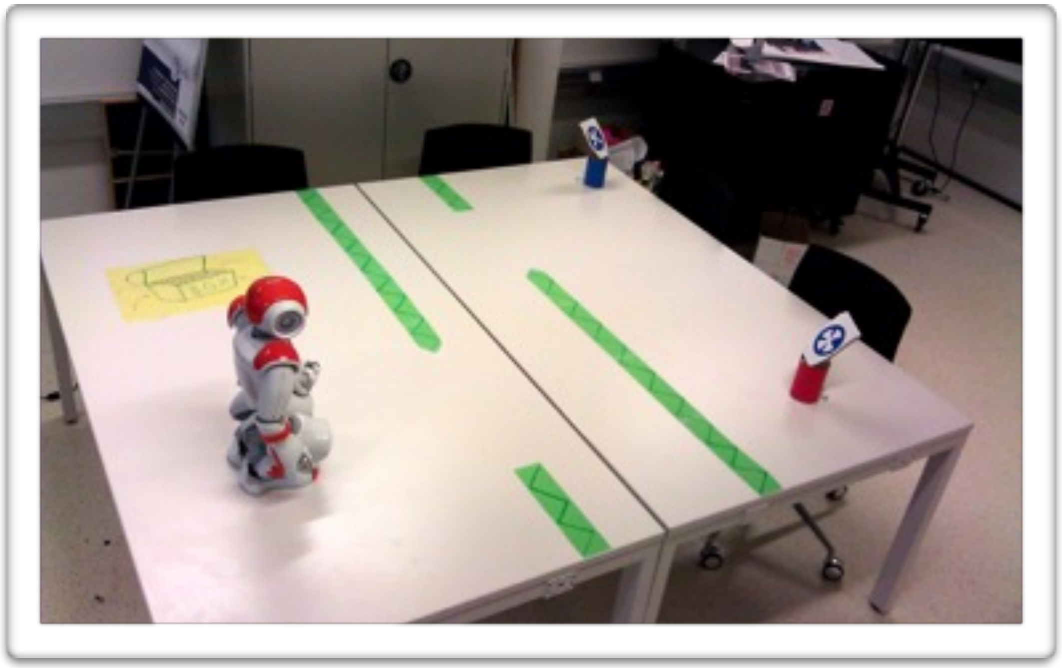
[P. Lison. A hybrid approach to dialogue management based on probabilistic rules. *Computer Speech & Language*, 2015]
[P. Lison. Model-based Bayesian Reinforcement Learning for Dialogue Management (Interspeech 2013)]

Outline for this talk

- What is dialogue management?
- **A hybrid logical/probabilistic approach**
  - Probabilistic rules
  - Parameter estimation
  - **Experiments**
- Demonstration of the OpenDial toolkit

User evaluation

- **Task**: instruct the robot to move across the table, pick one cylinder and release it on the landmark



- Comparison of three modelling approaches:

1. A handcrafted finite-state automaton
2. A factored statistical model
3. A model structured with probabilistic rules

Experimental procedure

- **Step 1:** collect Wizard-of-Oz interaction data
- **Step 2:** Estimate the internal parameters for the 3 models with the collected data
- **Step 3:** Conduct user trials for the 3 approaches
- **Step 4:** Compare them on dialogue quality metrics

Dialogue domain:

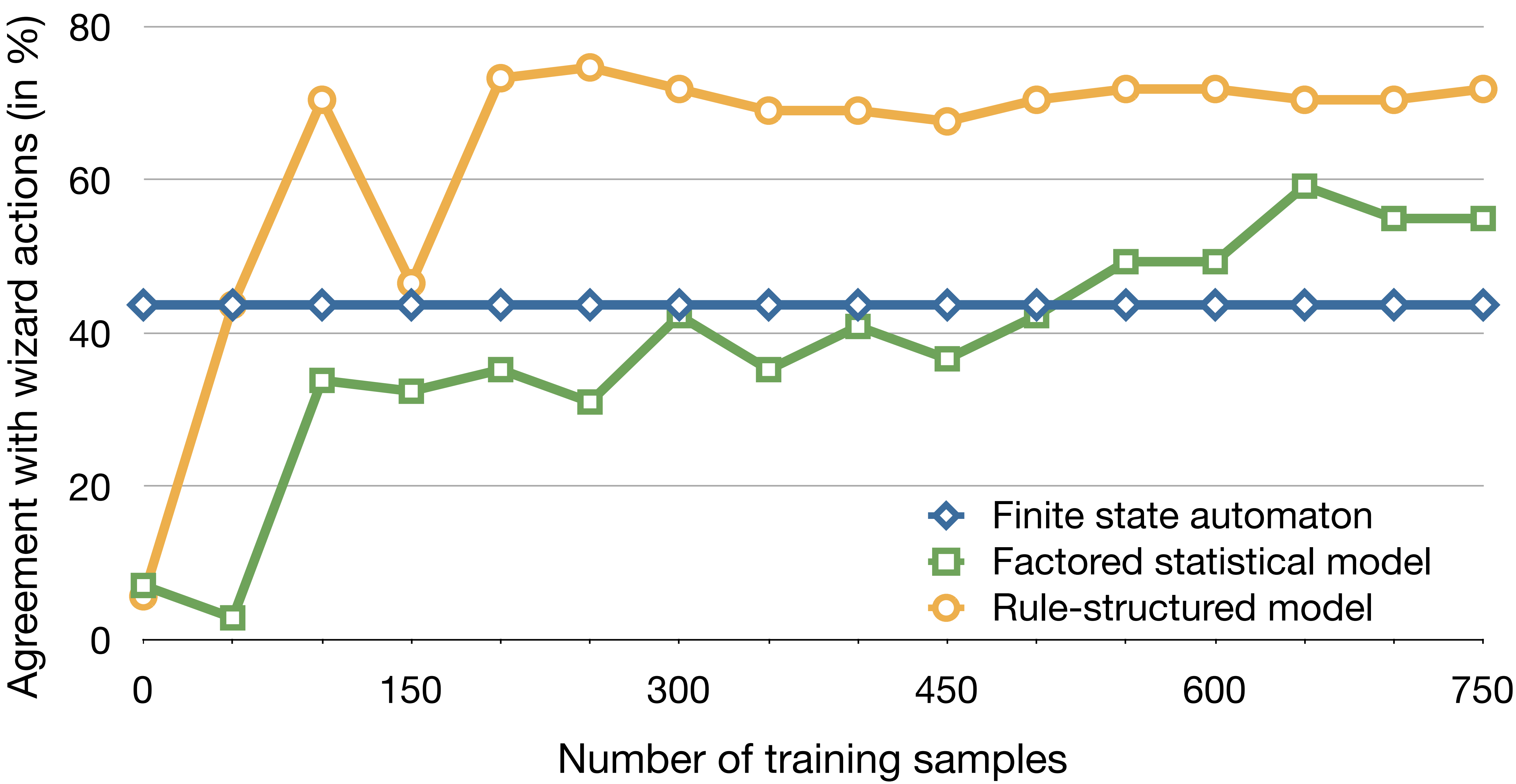
- 26 user actions
- 41 system actions
- State size: 35×10^6 (10 variables)

Parameter estimation:

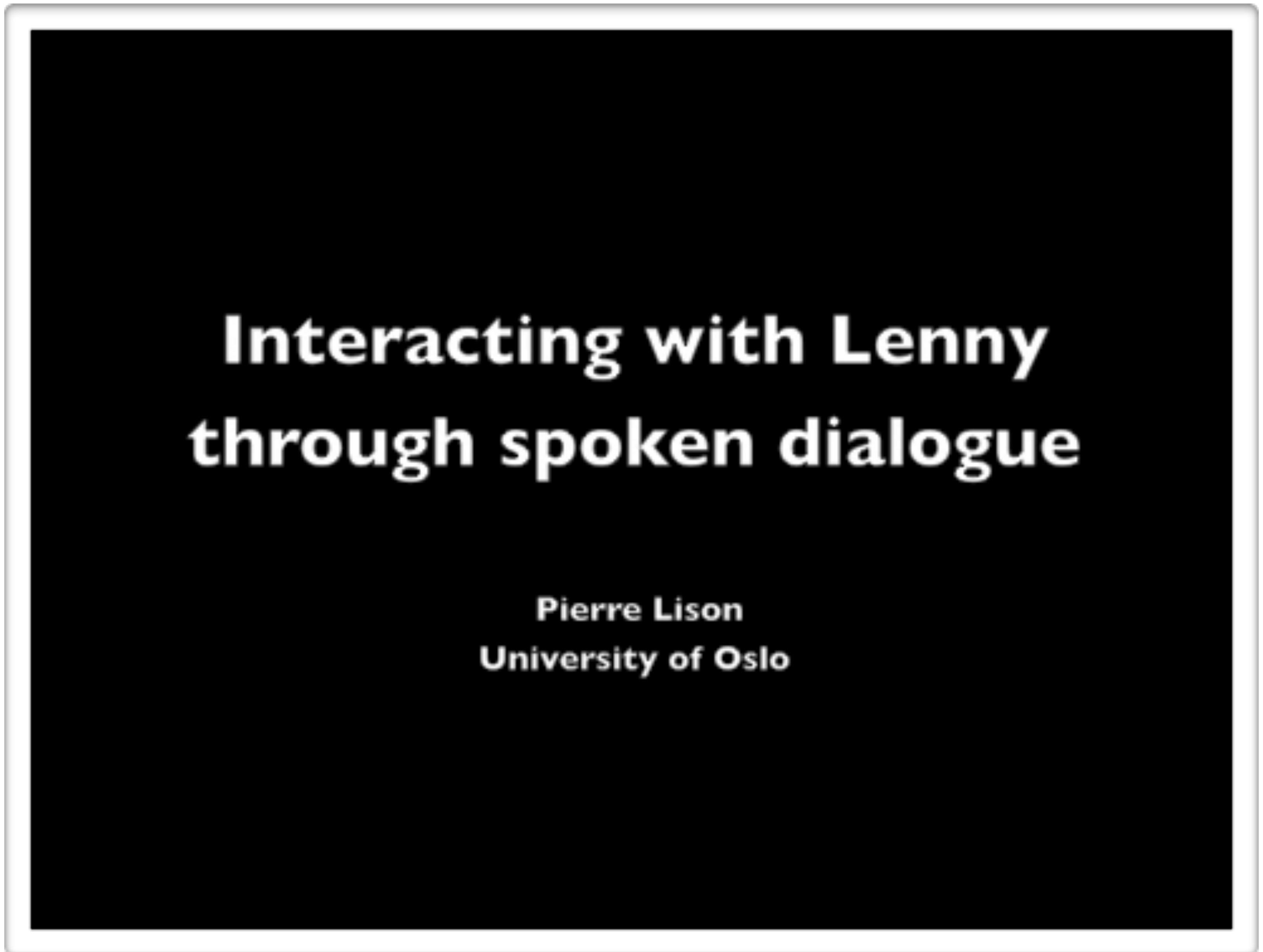
- 10 recorded WoZ interactions
- 3 parameters in handcrafted automaton (thresholds)
- 433 parameters in factored statistical model
- 28 parameters in model encoded with probabilistic rules

Learning curve

Training: 9 Wizard-of-Oz interactions (770 system turns)
Testing: 1 Wizard-of-Oz interaction (71 system turns)



User trials



- 37 participants (16 M / 21 F)
- Average age : 30.6
- Average duration: 5:06 mins
- All captured on videos

User trials

- Each participant in the trial repeated the task **three** times
  - One interaction for each modelling approach (in randomised order)
- Evaluation metrics:
  - **Objective metrics**: list of 9 measures extracted from the interaction logs
  - **Subjective metrics** : survey of 6 questions filled by the participants after each interaction

Empirical results

	Metrics	Finite-state automaton	Factored statistical model	Rule-structured model
Objective	Average number of repetition requests	18.68	12.24	**0***
	Average number of confirmation requests	9.16	10.32	**5.78***
	Average number of repeated instructions	3.73	7.97	**2.78**
	Average number of user rejections	**2.16**	2.59	2.59
	Average number of physical movements	**26.68**	29.89	27.08
	Average number of turns between moves	3.63	3.1	**2.54***
	Average number of user turns	78.95	77.3	**69.14**
	Average number of system turns	57.27	54.59	**35.11***
	Average duration (in minutes)	6:18	7:13	**5:24***
Subjective	*"Did you feel that…*			
	… the robot correctly understood what you said?"	3.32	2.92	**3.68**
	… the robot reacted appropriately to your instructions?"	3.70	3.32	**3.86**
	… the robot asked you to repeat/confirm your instructions?"	2.16	2.19	**3.3***
	… the robot sometimes ignored when you were speaking?"	3.24	2.76	**3.43**
	… the robot thought you were talking when you were not?"	3.43	3.14	**4.41***
	… the interaction flowed in a pleasant and natural manner?"	2.97	2.46	**3.32**

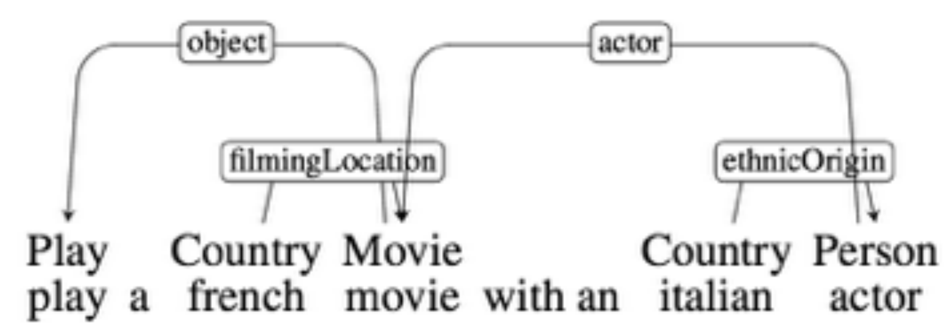
Scale from 1 (worse) to 5 (best)

Outline for this talk

- The dialogue management task
- A hybrid logical/probabilistic approach
  - Probabilistic rules
  - Parameter estimation
  - Experiments
- Three open research questions

Open research questions (1)

- The probabilistic rules allow us to capture complex relations between state variables
- **But** the underlying state representation remains propositional (slot-value pairs)
- Many variables are better viewed as *relational structures*



  - Semantic content, user intentions, task structures, etc.
- Need to extend the probabilistic rules to be able to operate on such types of state variables

[D. Ramachandran and A. Ratnaparkhi. "Belief Tracking with Stacked Relational Trees" (SIGDIAL 2015)]

Open research questions (2)

- Optimising dialogue policies from social signals?
  - Users spontaneously produce a variety of *multimodal feedback signals* (emotional cues, grounding actions, etc.)
  - Can we optimise the model parameters against these signals ?
- Distinct from traditional reinforcement learning:
  - Detecting these multimodal signals and determining their "feedback value" is difficult and prone to errors
  - No one-to-one mapping between signals and system actions (*credit assignment problem*)

Open research questions (3)

- The information-state architecture of OpenDial works well for "high-level" reasoning tasks
  - Tracking the user intention(s), planning system actions
  - One central information hub: the dialogue state
- But it is less appropriate for lower-level tasks
  - Turn-taking, (high-throughput) perception processes, etc.
- How to reconcile the "high-level" and "lower-level" aspects of dialogue processing in a principled manner?
  - In other words: can we combine OpenDial and IrisTK?

Conclusion

- Development of a new modelling framework for dialogue management, based on **probabilistic rules**
  - *Hybrid* approach at the crossroads between logical and statistical methods
  - Rule parameters can be learned from data
- HRI experiments demonstrate the benefits of the approach
- Concrete implementation in the OpenDial software toolkit