Open Domain
Statistical Spoken Dialogue Systems

Steve Young

Dialogue Systems Group
Machine Intelligence Laboratory
Cambridge University Engineering Department
Cambridge, UK

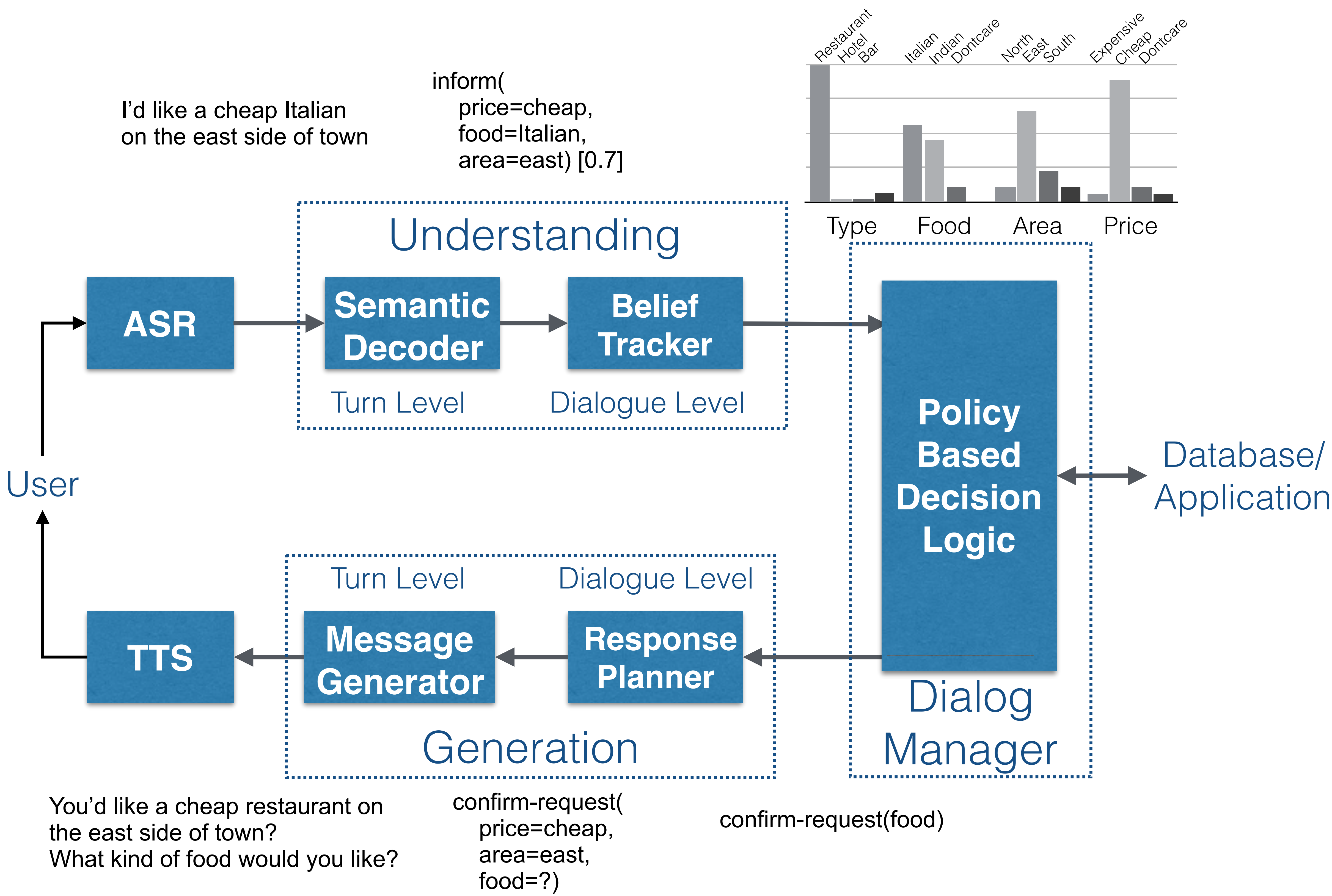
Contents

Statistical Spoken Dialogue

To enable fully automatic on-line learning, all components must be trainable from data.

“Deploy, Collect Data, Improve”

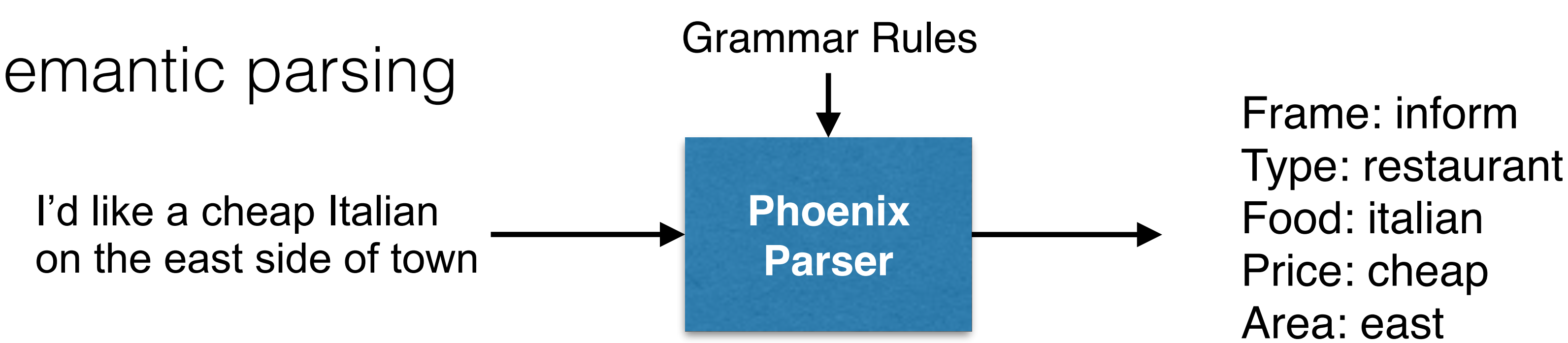
Statistical Spoken Dialog System



Spoken Language Understanding (SLU)

Various decoding strategies

a) Semantic parsing



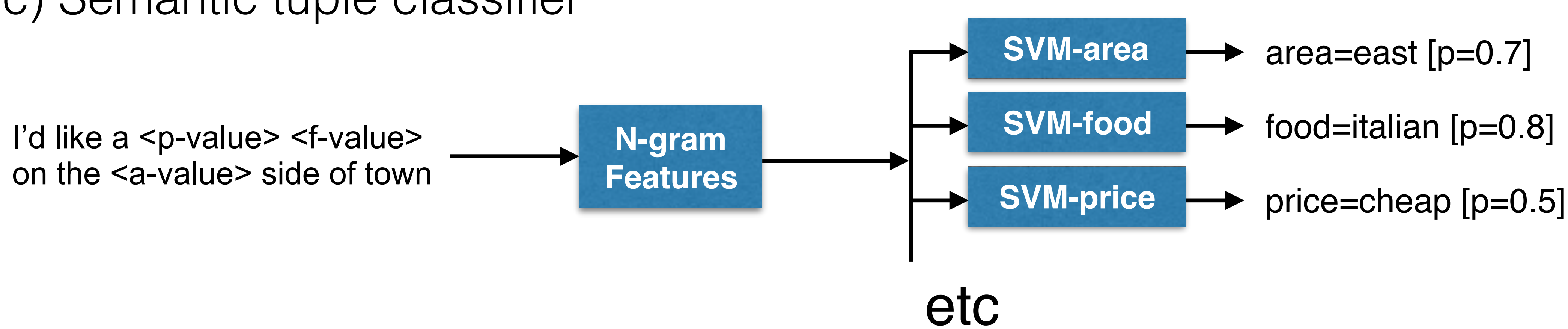
b) Semantic tagging

$$\hat{Y} = \arg\max_{Y} P(Y \mid X)$$ eg. HMM, CRF

X =	I'd	like	a	cheap	Italian	on	the	east	side	of	town
Y =	B-inform	I-inform	o	B-price	B-food	o	o	B-area	I-area	I-area	I-area

inform price=cheap food=italian area=east

c) Semantic tuple classifier



SLU Performance

Cambridge Restaurant System:

- Noisy in-car data, various conditions, 37% average word error rate (WER)
- 10571 training utterances, 4882 test utterances

Features	Trained On	F-Score	Item Cross Entropy
Phoenix	—	0.69	2.78
CRF	ASR 1-best	0.67	2.75
N-grams	ASR 1-best	0.69	1.79
N-grams	ASR 2-best	0.70	1.72
Weighted N-grams	ASR 10-best	0.71	1.76
Weighted N-grams	Confusion Network	0.73	1.68
Weighted N-grams + Context	Confusion Network	0.77	1.43

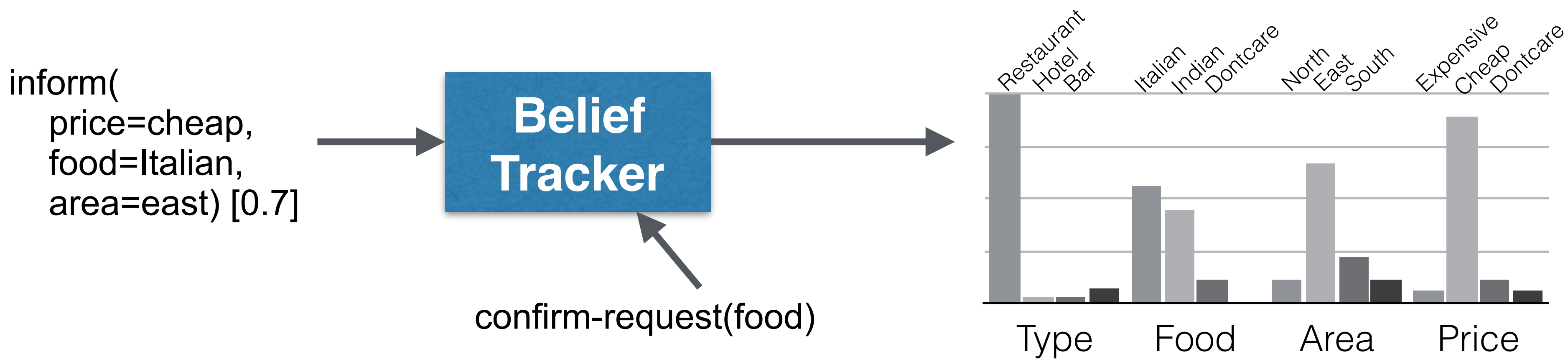
Semantic tuple classifier

choice of classifier not so important

1-best incurs significant information loss!

M. Henderson, et al (2012). "Discriminative Spoken Language Understanding Using Word Confusion Networks." IEEE SLT 2012, Miami, FL

Belief Tracking

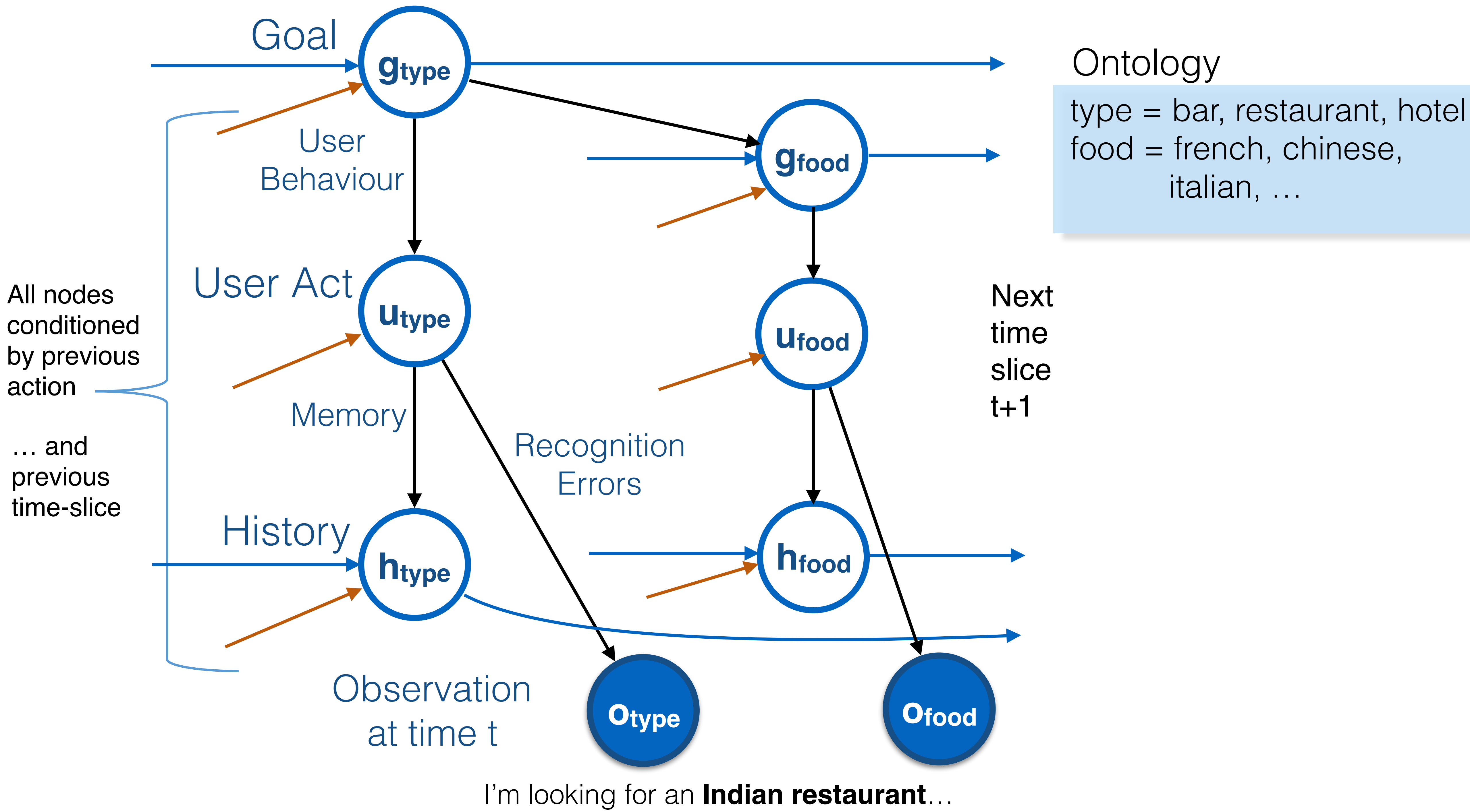


Aim: to maintain a distribution over all dialogue state variables using SLU output at each turn as evidence

3 principal approaches:

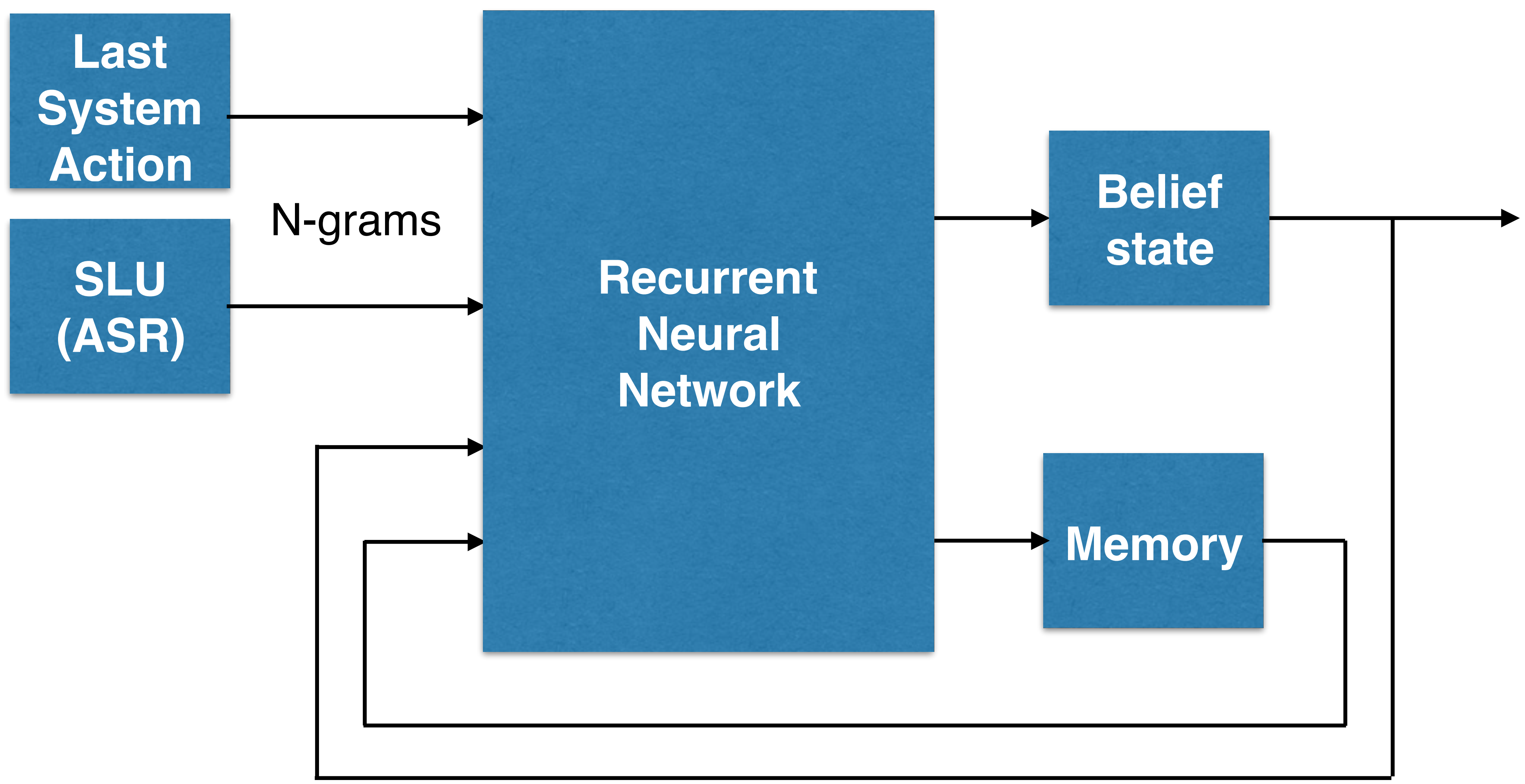
- rule-based
- dynamic Bayesian network
- discriminative model (eg RNN)

Dynamic Bayesian Networks (DBNs)



B. Thomson and S. Young (2010). "Bayesian update of dialogue state: A POMDP framework for spoken dialogue systems." Computer Speech and Language,**24**(4): 562-588. [CSL 2015 Best paper Award]

Recurrent Neural Net Belief Tracking



Word-Based Dialog State Tracking with Recurrent Neural Networks
M. Henderson, B. Thomson and S. Young, SigDial 2014, Philadelphia, PA

Belief Tracking Performance

Cambridge Restaurant System (Dialog State Tracking Challenge 2):

- Telephone data, various conditions, 20% to 40% average word error rate (WER)
- 1612 training dialogs, 1117 test dialogs
- Joint Slot Accuracy (fraction of turns in which all goal labels are correct)
- Joint L2 (L2 norm between tracker output distribution and reference)

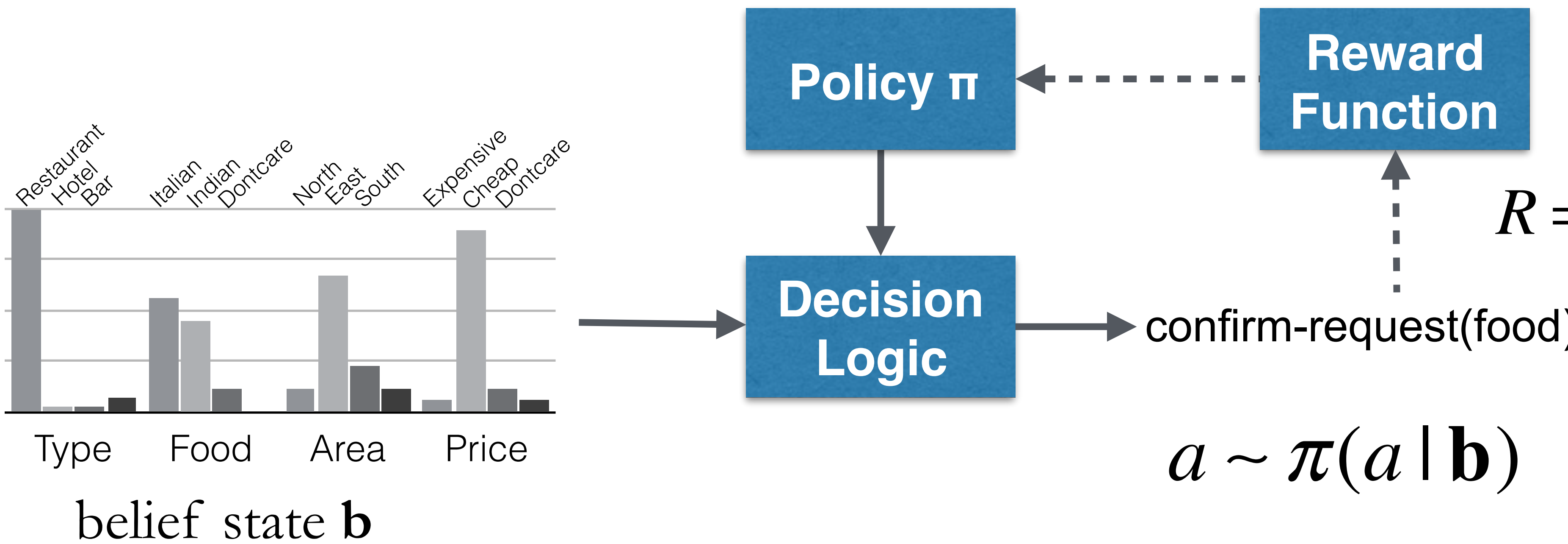
System	Features	Accuracy	L2
Baseline	SLU	61.6%	0.74
Bayes Net	SLU	67.5%	0.55
Delex RNN	SLU	73.7%	0.41
Full RNN	SLU	74.2%	0.39
Delex RNN	ASR	74.6%	0.38
Full RNN	ASR	76.8%	0.35

discriminative tracker significantly better than generative tracker

intermediate semantic representation incurs more information loss!

The Second Dialog State Tracking Challenge
M. Henderson, B. Thomson and J. Williams, SigDial 2014, Philadelphia, PA

Dialog Management



$$R = \sum_{\tau} \gamma^{\tau-1} r(a_\tau, \mathbf{b}_\tau)$$

$$a \sim \pi(a \mid \mathbf{b})$$

Partially Observable Markov Decision Process

- action at each turn is function of belief state **b**
- policy optimised by maximising expected cumulative reward R
- trained on corpora, user simulator or on-line

Exact solutions intractable, but wide range of approximations:

- gradient ascent directly on policy π (NAC)
- maximise GP approximation of Q-function (GP-SARSA)

S. Young, M. Gasic, B. Thomson and J. Williams (2013). "POMDP-based Statistical Spoken Dialogue Systems: a Review." Proc IEEE, **101**(5):1160-1179

Natural Actor-Critic

$$\pi(a \mid \mathbf{b}, \theta) = \frac{e^{\theta . \phi_a(\mathbf{b})}}{\sum_{a'} e^{\theta . \phi_{a'}(\mathbf{b})}}$$

Action specific features $\phi_a(\mathbf{b})$ defined on **b**

Policy defined directly on softmax $\theta . \phi_a(\mathbf{b})$

$$J(\theta) = \mathbb{E}\left[\frac{1}{T} \sum_t r(\mathbf{b}_t, a_t) \mid \pi_\theta \right]$$

Cost function is sum over observed per turn rewards

Optimise using natural gradient ascent

$$\tilde{\nabla} J(\theta) = F_\theta^{-1} \nabla J(\theta)$$

Gradient is estimated by sampling dialogues so Fisher Information Matrix does not need to be explicitly computed.

F. Jurcicek, B. Thomson and S. Young (2011). "Natural Actor and Belief Critic: Reinforcement algorithm for learning parameters of dialogue systems modelled as POMDPs." ACM Transactions on Speech and Language Processing, 7(3)

GP-SARSA

$$Q_0^\pi(\mathbf{b},a) \sim GP(0, k((\mathbf{b},a),(\mathbf{b},a)))$$

$Q^\pi(\mathbf{b},a) = E_\pi(R)$ is expected total reward R following policy π from point (**b**,a)

Given trajectory $\mathbf{B}_t = (\mathbf{b}_1, a_1), \ldots, (\mathbf{b}_t, a_t)$ and rewards $\mathbf{r}_t = r_1, \ldots, r_t$

posterior is $$Q_t^\pi(\mathbf{b},a) \mid \mathbf{r}_t, \mathbf{B}_t \sim N(\bar{Q}(\mathbf{b},a), \mathrm{cov}((\mathbf{b},a),(\mathbf{b},a)))$$

GP-SARSA Reinforcement Learning

- Choose: $a_{t+1} \sim Q_t^\pi(\mathbf{b}_t, a_t)$
- Update: $\mathbf{b}_t \rightarrow \mathbf{b}_{t+1}$
- Observe: r_{t+1}
- Update: $Q_t^\pi \rightarrow Q_{t+1}^\pi$

Gaussian processes for POMDP-based dialogue manager optimization - M. Gasic and S. Young (2014). IEEE Trans. Audio, Speech and Language Processing, **22**(1):28-40.

Dialog Manager Performance

Cambridge Restaurant System:

- Reward = +20 for success -1 per turn
- User simulator-based training, 100k dialogs
- Telephone-based on-line training, 1200 dialogs
- Telephone-based real-user testing, 500 dialogs
- Telephone speech recognition, 20% average word error rate (WER)

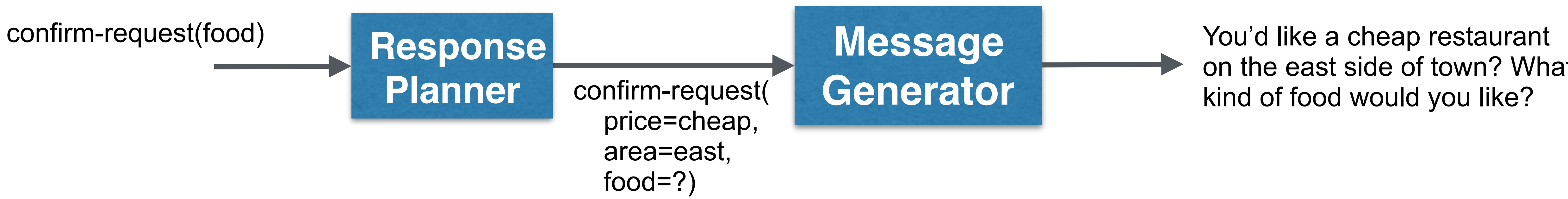
Method	Training	Reward	Success Rate	#Turns
NAC	Simulator	11.9	91.8%	6.5
GP-Sarsa	Simulator	11.6	91.2%	6.6
GP-Sarsa	On-line	13.4	96.8%	6.0

similar performance but GP must faster

Learning from real interactions makes significant difference

S. Young, et al (2014). "Evaluation of Statistical POMDP-based Dialogue Systems in Noisy Environments." International Workshop Spoken Dialogue Systems (IWSDS 2014), Napa, CA

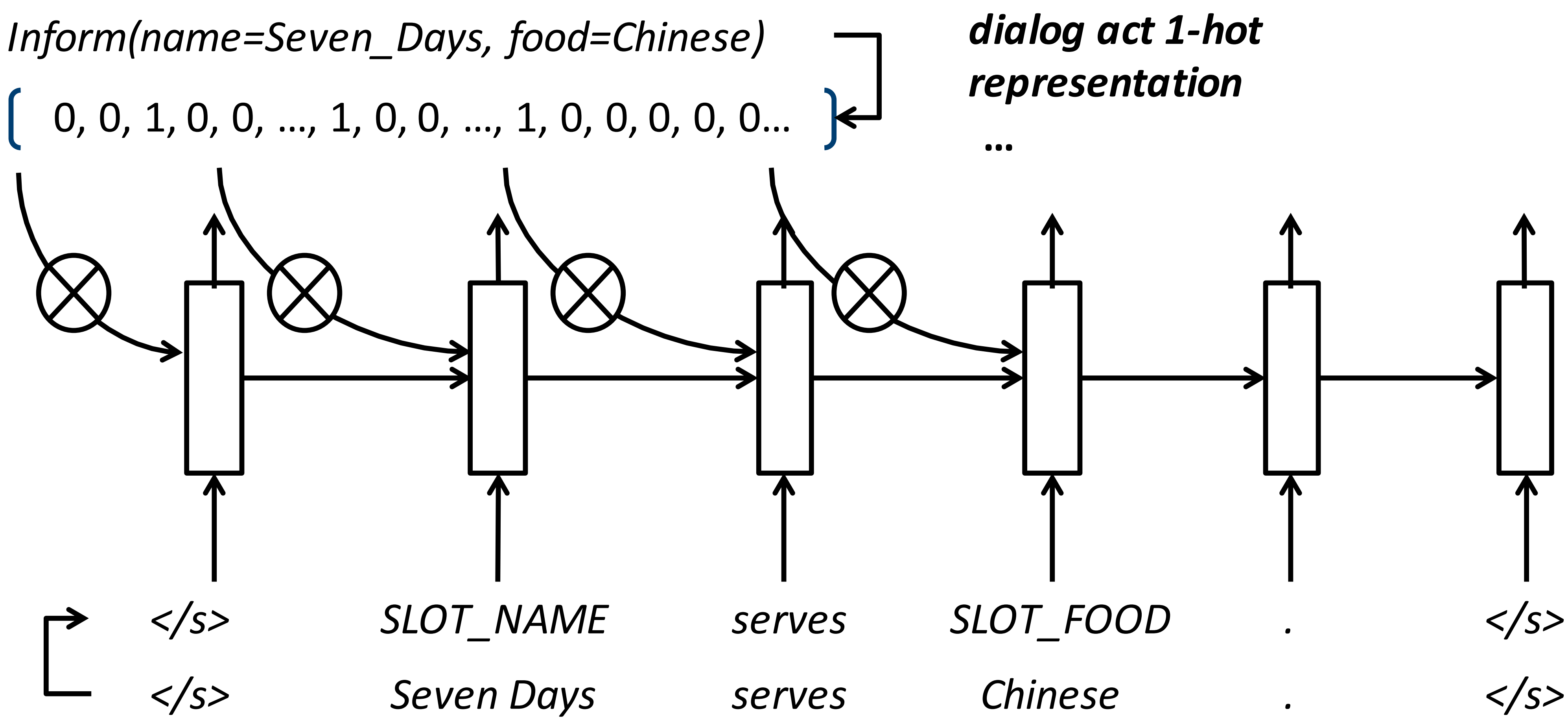
Natural Language Generation



3 principal approaches:

- hand-crafting with parameterised templates
- generative linguistic rules
- data driven using "over-generate and filter" approach

Constrained RNN Generation

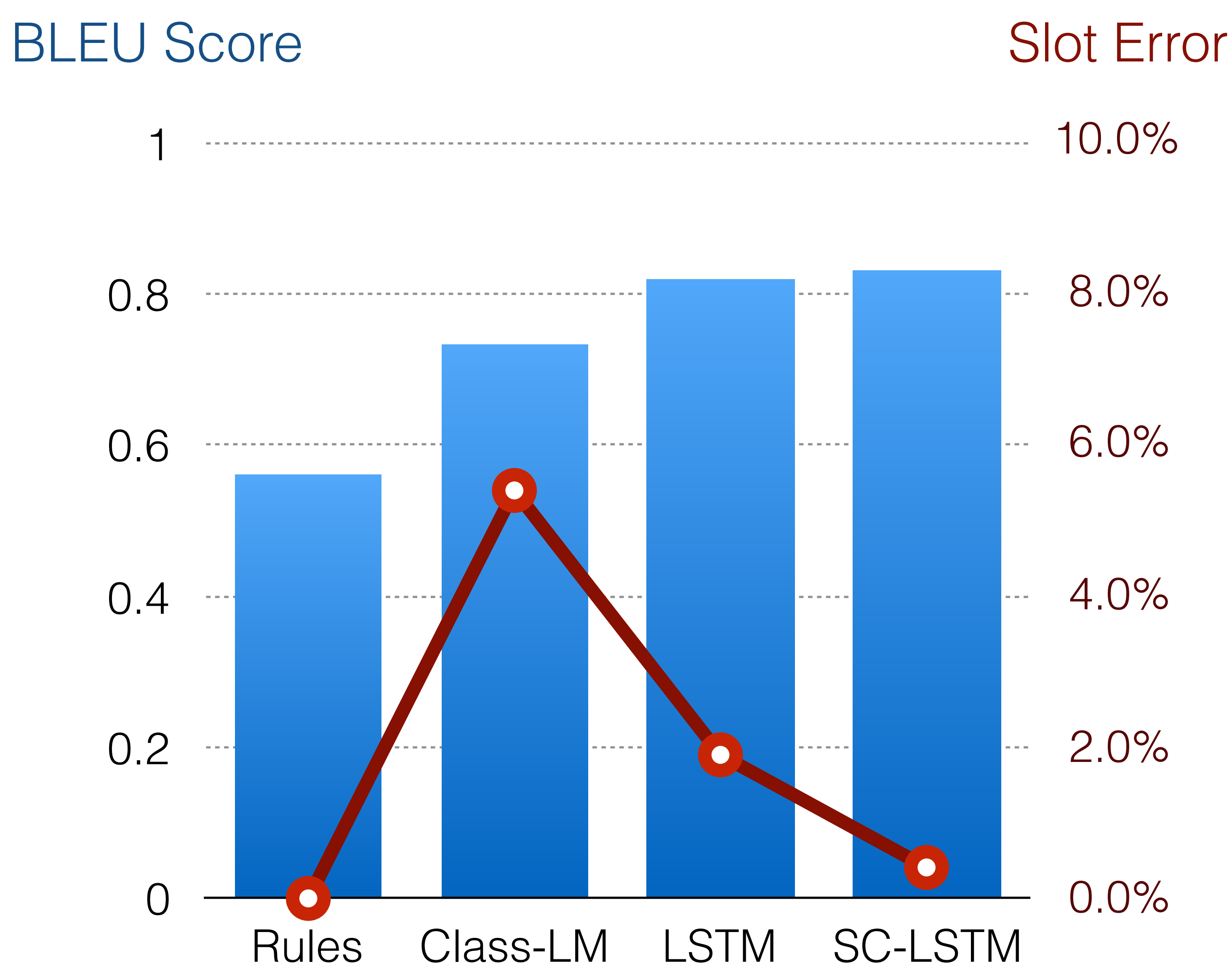


RNN trained on data pairs consisting of

a) 1-hot representation of system dialog act
b) corresponding delexicalised output utterance

T-H. Wen et al (2015). "Stochastic Language Generation in Dialogue using Recurrent Neural Networks with Convolutional Sentence Reranking." Sigdial 2015, Prague, Cz. [Best paper award]

Generation Performance



T-H. Wen et al (2015). "Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems." EMNLP 2015, Lisbon, Portugal. [Best paper award]

Open-Domain Statistical Dialogue

Objectives

To develop spoken dialogue systems which:

1. allow users to reference multiple domains within a single conversation

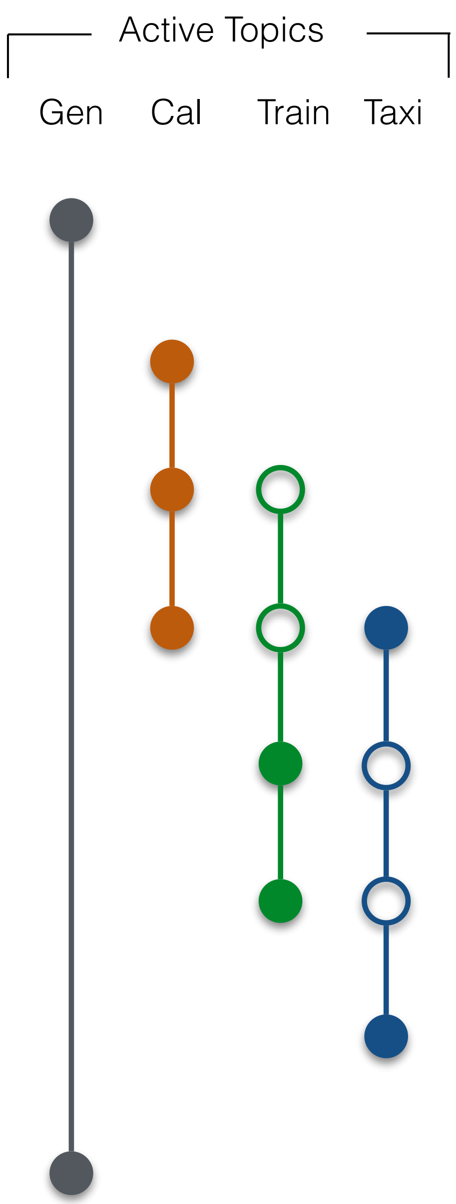
2. supports natural conversation even in rarely visited domains

3. can learn automatically on-line through interaction with user

“Deploy, Collect Data, Improve”

Note that user in the loop enables on-line reinforcement learning

Example Dialogue



Hello, how can I help you?

What appointments do I have tomorrow?

You have a meeting at 10am with John and a teleconf at noon with Bill.

I need to go to London first thing, can you reschedule the meeting with John?

John is free tomorrow at 3pm, is that ok?

Yes, thats fine. I also need a taxi to the station.

Meeting with John at 15.00 is confirmed. What time do you need the taxi?

When does the train depart to London?

The 9.15am gets in at 10.06.

When is the one before that?

The train before that leaves at 8.45am and arrives at 9.40.

Ok I will take that, book the taxi for 8.15am from my house.

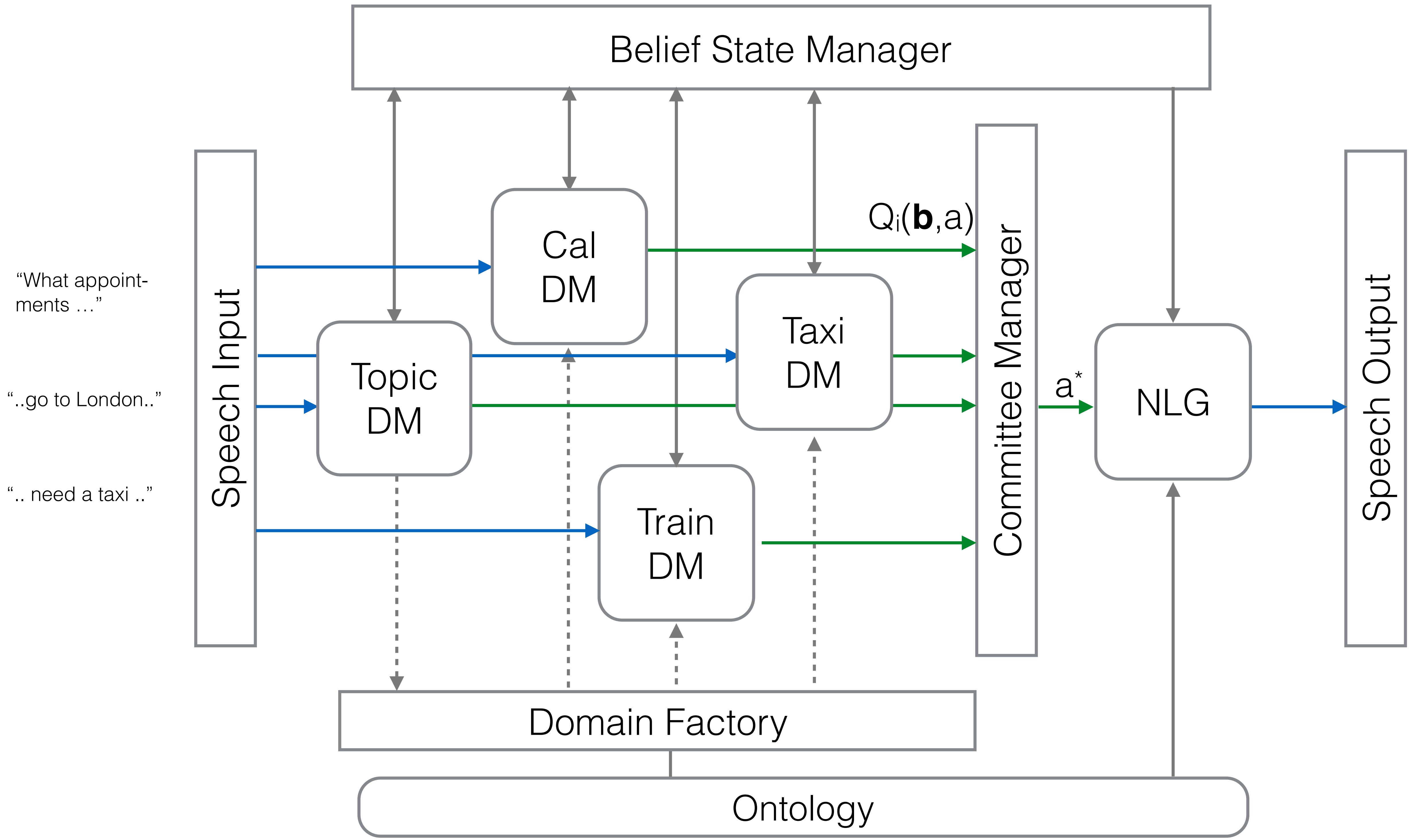
Ok, I will book the taxi for 8.15am, is that correct?

Yes that's right.

Ok. Do you need anything else?

Not for now thanks.

Run-time Architecture



Distributed Dialog Management

- Each DM operates independently, receives speech, tracks its own beliefs and proposes system actions

- DM's operate as a Bayesian Committee Machine, each machine's Q-value has a confidence attached to it:

$$\bar{Q}(\mathbf{b},a) = \Sigma^Q(\mathbf{b},a) \sum_{i=1}^{M} \Sigma_i^Q(\mathbf{b},a)^{-1} \bar{Q}_i(\mathbf{b},a)$$

$$\Sigma^Q(\mathbf{b},a)^{-1} = -(M-1) * k((\mathbf{b},a),(\mathbf{b},a))^{-1} + \sum_{i=1}^{M} \Sigma_i^Q(\mathbf{b},a)^{-1}$$

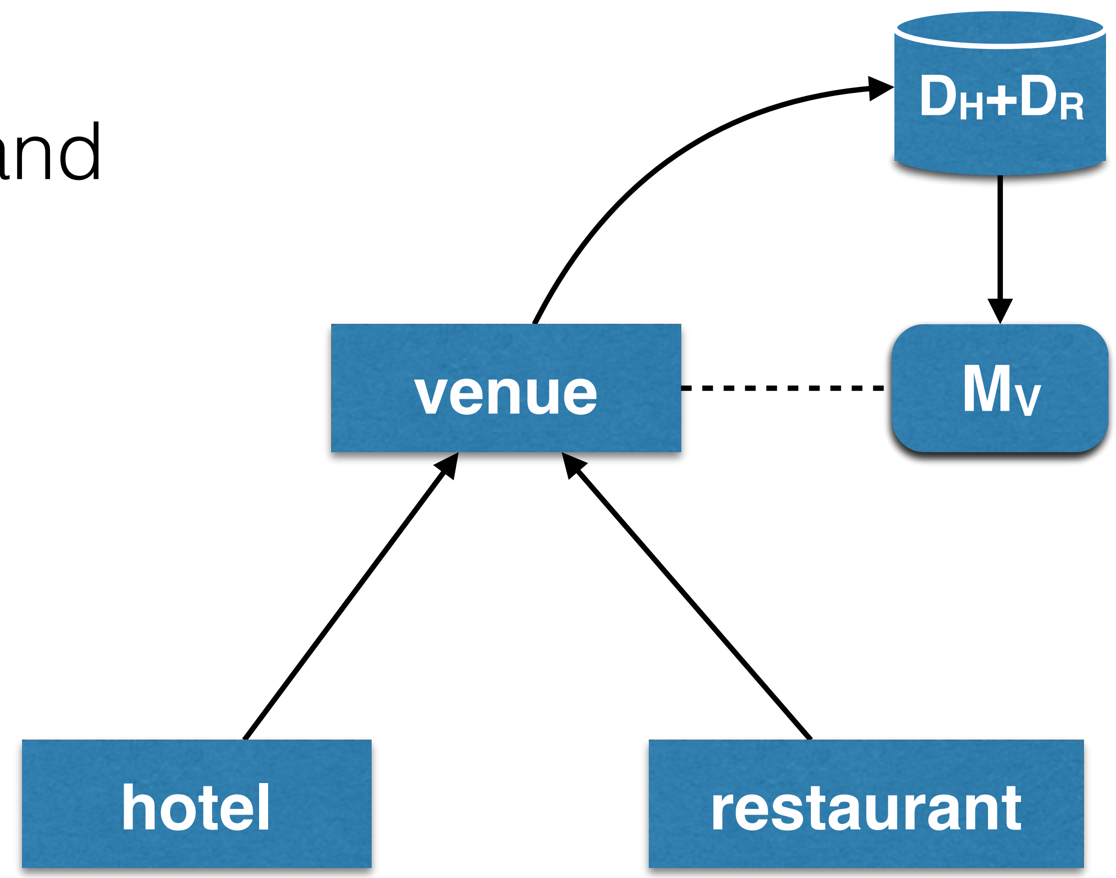
- Reinforcement learning operates on the group, distributing rewards at each turn according to previous action selection.

Modular, flexible, incremental, trainable on-line, …

M. Gasic et al (2015). "Policy Committee for Adaptation in Multi-Domain Spoken Dialogue Systems." IEEE ASRU 15, Scotsdale, AZ.

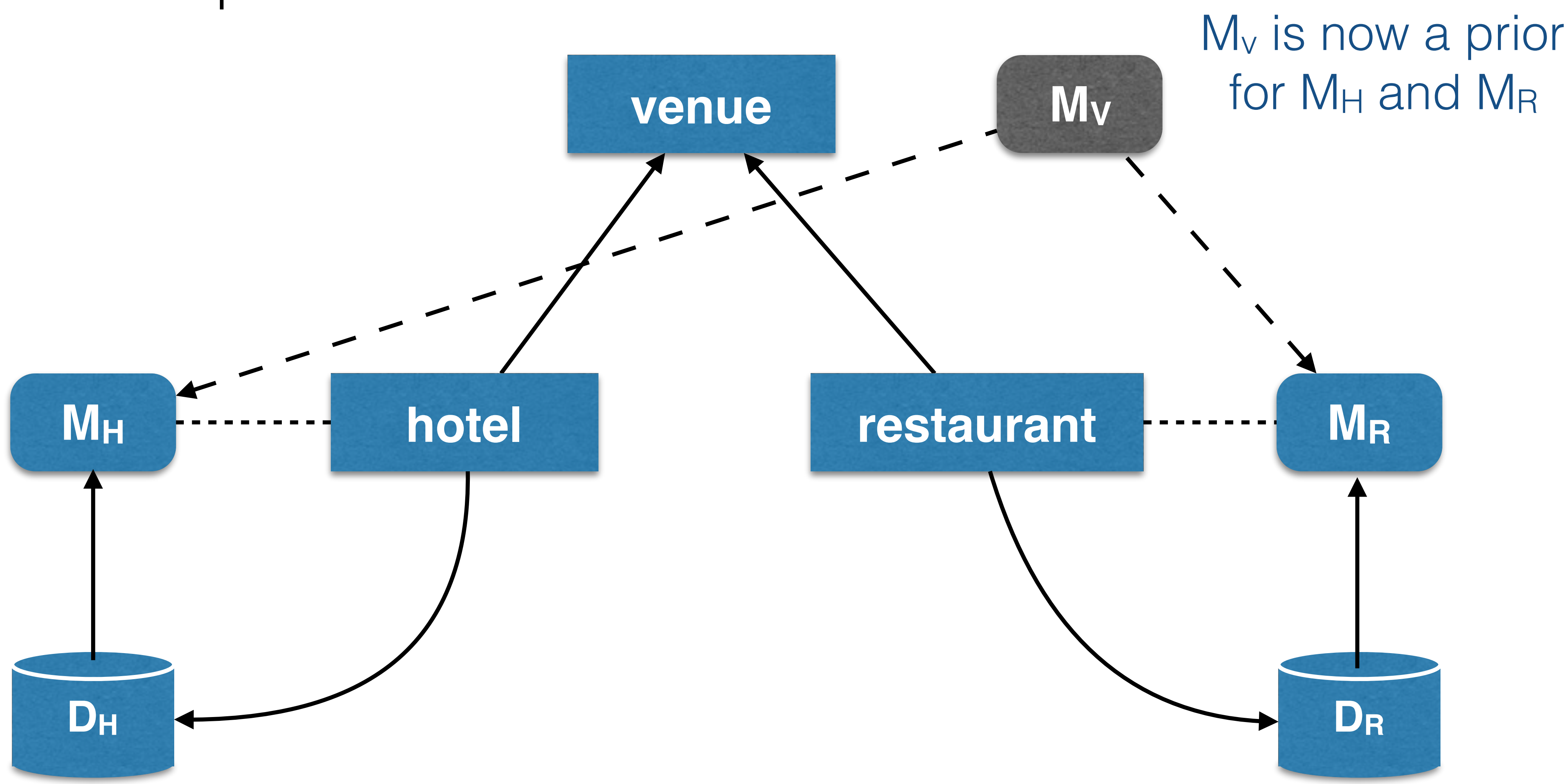
Incremental Domain Learning

Initially pool all available data and learn generic models



Incremental Domain Learning

Refine with more data using generic models as priors



M. Gasic et al (2015). "Distributed Dialogue Policies for Multi-Domain Statistical Dialogue Management." IEEE ICASSP 15, Brisbane, Sydney.

Performance of Generic Policies

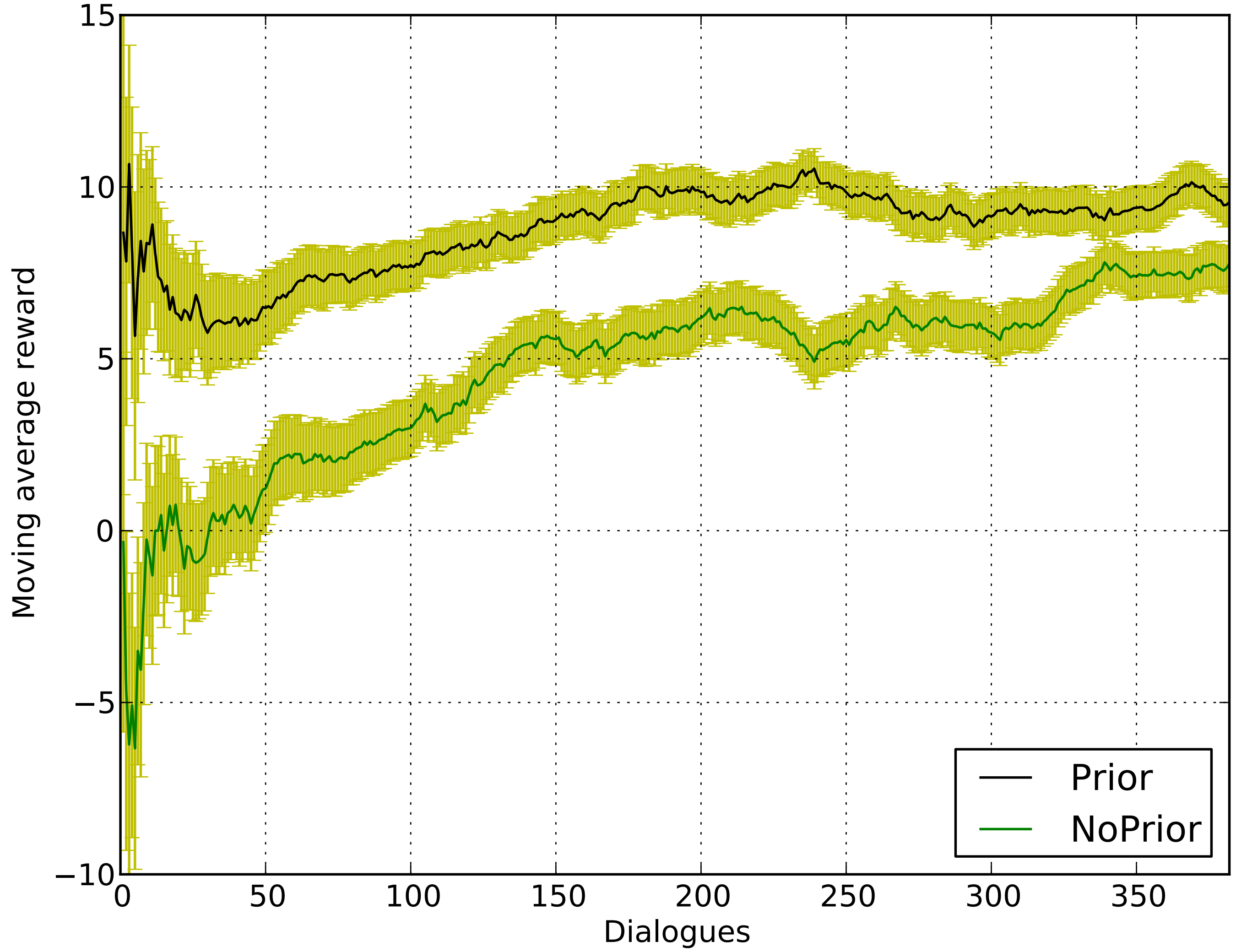
Strategy	#Dialogs	Restaurant	Hotel
in-domain	250	62.5%	64.3%
in-domain	500	67.5%	70.1%
generic	500	73.0%	76.2%
in-domain	2500	83.9%	85.9%
in-domain	5000	86.4%	86.9%
generic	5000	86.5%	87.1%

i.e. 250 from each domain

Success rates averaged over 10 policies and 1000 dialogues per condition

Distributed Dialogue Policies for Multi-Domain Statistical Dialogue Management
M. Gasic, D. Kim, P. Tsiakoulis and S. Young, Proc ICASSP, Brisbane, 2015

On-line Adaptation with Real Users



San Francisco Restaurant Domain

a) with generic prior
b) no prior

Performance is acceptable after only 50 dialogues in the new domain.

Conclusions

- End-to-end statistical dialogue is feasible, and can match or exceed hand-crafted systems in limited domains

- User-in-loop makes on-line learning feasible, even for previously unseen domains

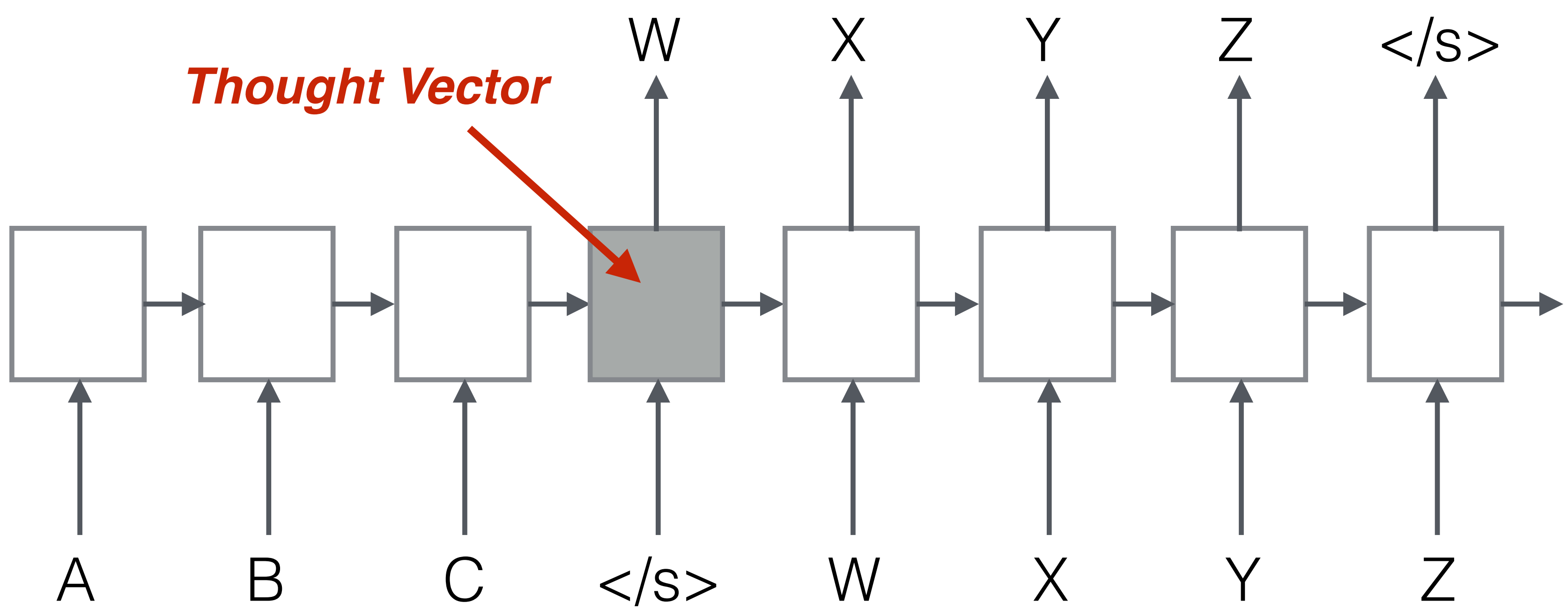
- Distributed hierarchical models, with generic parameters and "committees of experts" enable systems to learn to expand coverage whilst avoiding unacceptable user experience.

- Focus today has been on expanding dialogue management. Current work suggests that similar ideas extend to SLU and NLG.

CUED Dialogue Systems Group

Current	Past
Steve Young	Blaise Thomson, Apple
Milica Gasic	Dongho Kim, Apple
David Vandyke	Matt Henderson, Google
Lina Rojas-Barahona	Prof Kai Yu, SJTU
Nikola Mrksic	Jason Williams, Microsoft
Eddy Su	Pirros Tsiakoulis, Innoetics Ltd
Shawn Wen	Francois Mairesse, Amazon
*Stefan Ultes**	Catherine Breslin, Amazon
**starting Jan 2016*	Prof Filip Jurcicek, Charles U.

Deep Learning - Seq2Seq Models



Key strengths:

- automatic feature extraction
- ability to compactly encode sequence information

But hard to build a practical system without pulling out and explicit action set and without individually trainable modules.