Spoken dialog systems as an application of POMDPs

Jason D. Williams

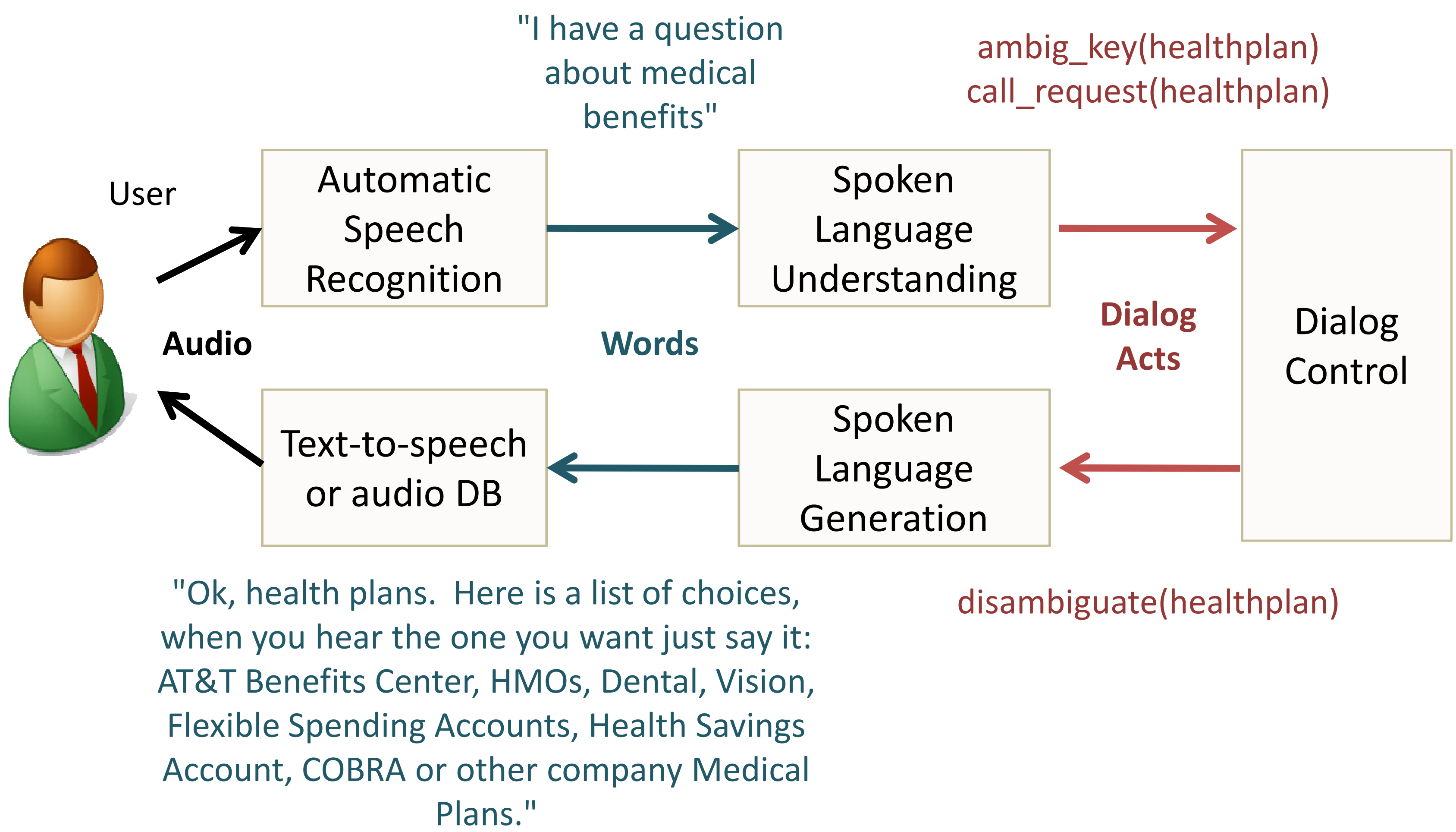
at&t

ICAPS Workshop – May 2010

What is a spoken dialog system?

A spoken dialogue system is
a computer agent that interacts with
people by understanding
spoken language.

Speech recognition and spoken language understanding



Spoken dialogue systems come in many flavours

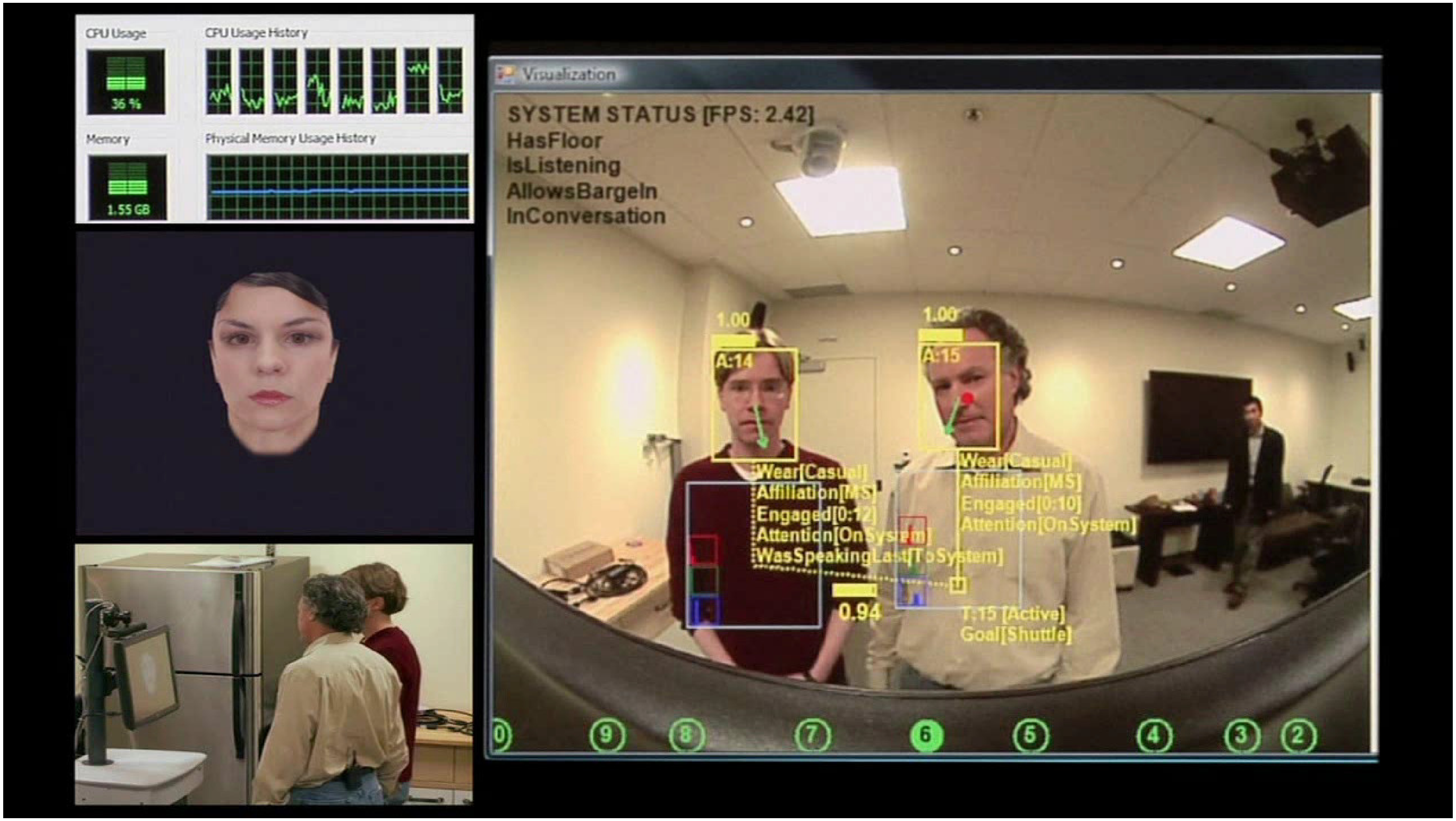
Input	Output	Example
Speech	Speech	Telephone technical support [1]
Speech + ?GUI	Speech + ?GUI	In-car music control, navigation
Speech + GUI	Speech + GUI	Tutoring
Speech + GUI	Speech + GUI	Language learning
Speech + GUI	?Speech + GUI	TV program guide
Speech + GUI	?Speech + GUI	Mobile search interface
Speech + vision	Speech + robot/agent	Eldercare
Speech + vision	Speech + robot/agent	Automated receptionist

[1] Recording of a deployed dialog system, AT&T

In-car spoken dialogue system

Source: IBM

Automated receptionist



Bohus, D., Horvitz, E. (2009). Models for Multiparty Engagement in Open-World Dialog, in Proceedings of SIGdial'09, London, UK

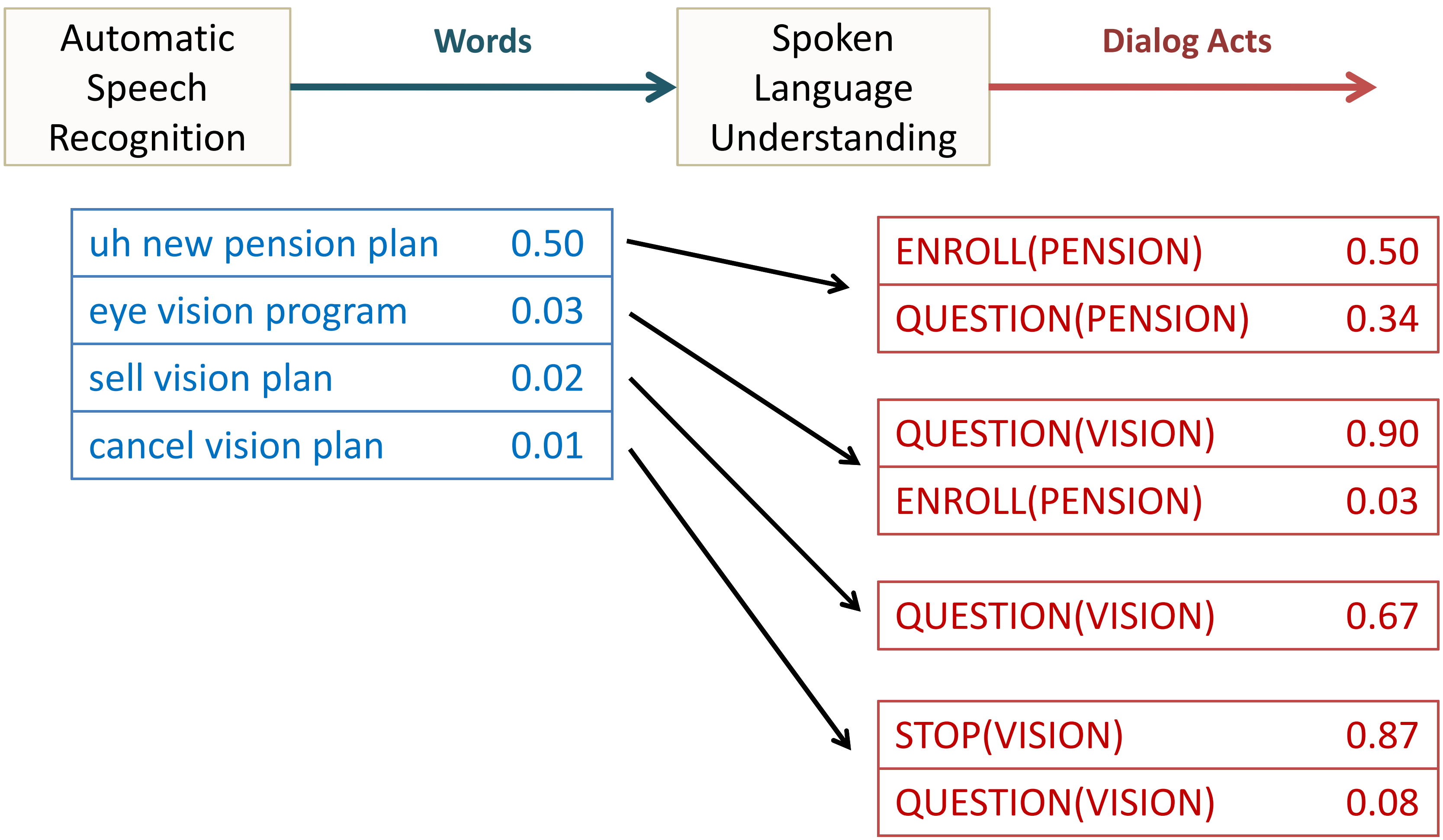
Outline

- Key challenges for building dialogue systems
- How dialogue systems are built today
- Casting a dialogue system as a POMDP
- "Growing up" to real-world systems
- Thoughts about the future

Challenges (among others)

1. Channel errors (ASR, SLU, turn-taking)
2. Curse of history
3. Theory of mind problem

Speech recognition and spoken language understanding



ASR/SLU errors are common

Grammar	Yes/no	City & state	How may I help you?

Source: Two different deployed commercial applications running two different speech recognizers

ASR/SLU errors are common

Grammar	Yes/no	City & state	How may I help you?
In-grammar/ in-domain accuracy	99.8%	85.1%	89.5%

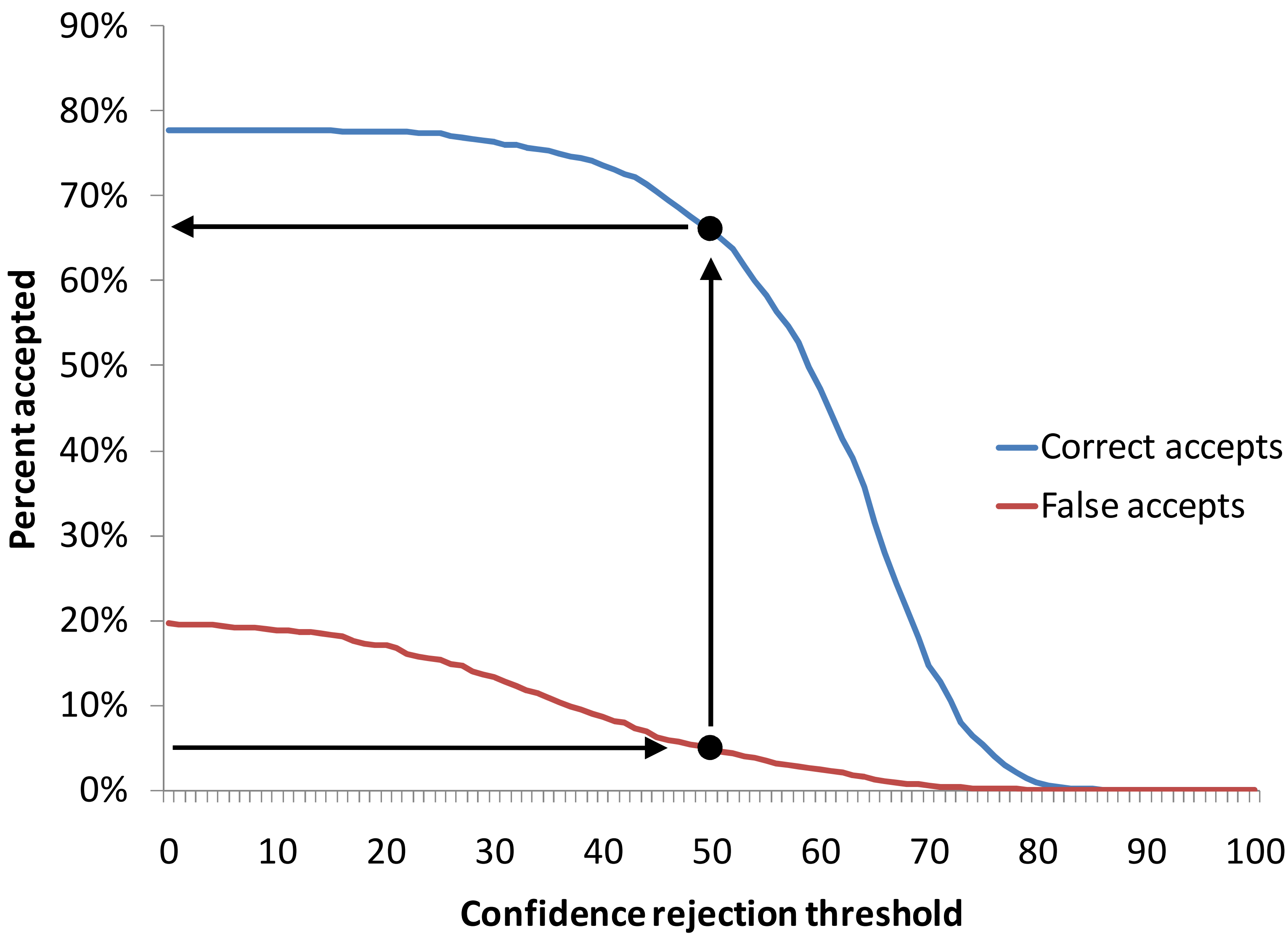
Source: Two different deployed commercial applications running two different speech recognizers

ASR/SLU errors are common

Grammar	Yes/no	City & state	How may I help you?
In-grammar/ in-domain accuracy	99.8%	85.1%	89.5%
% in-grammar/ in-domain	92.3%	91.0%	86.8%
Overall accuracy	92.1%	77.6%	77.7%

Source: Two different deployed commercial applications running two different speech recognizers

ASR errors are hard to detect



ASR/SLU errors are common

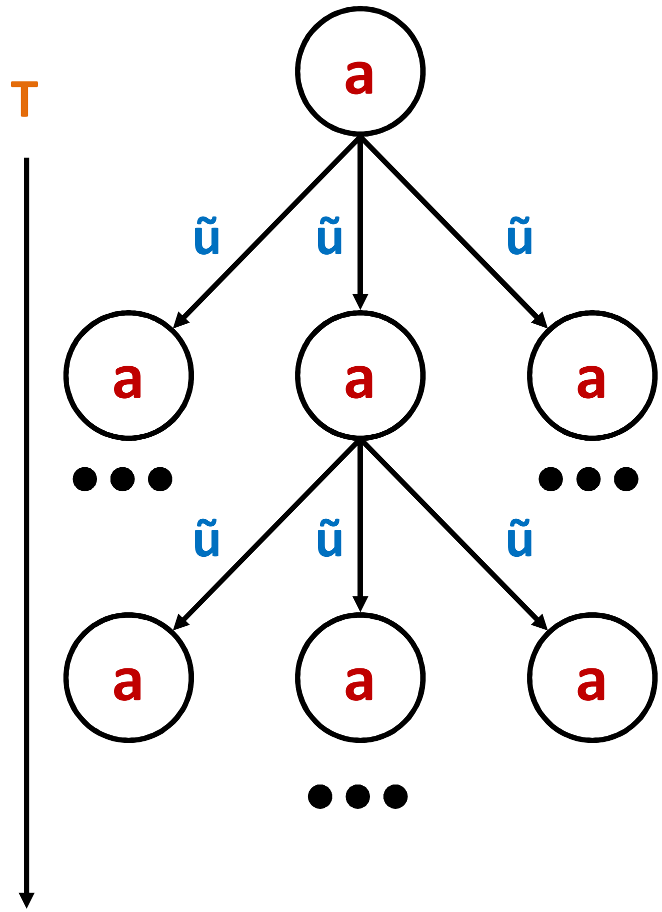
Grammar	Yes/no	City & state	How may I help you?
In-grammar/ in-domain accuracy	99.8%	85.1%	89.5%
% in-grammar/ in-domain	92.3%	91.0%	86.8%
Overall accuracy	92.1%	77.6%	77.7%
Accepted utts (False accepts)	89.6% (1.8%)	60.3% (4.9%)	73.3% (8.3%)

Source: Two different deployed commercial applications running two different speech recognizers

Curse of history

A = {ask(first-name), confirm(last-name=williams), ...}

Ũ = {YES, JASON, WILLIAMS, ... }



~ $A^{\tilde{U}^T}$ possible assignments

Typical system:

$A = 10^{10}$

$\tilde{U} = 10^{10}$

$T = 10$

Curse of history

$$F(\tilde{u}_0,a_1,\tilde{u}_1,a_2,\tilde{u}_2,a_3,\tilde{u}_3,\ldots,a_t,\tilde{u}_t) = a_{t+1}$$

Often it's more convenient to separate the *tracking* problem from the *action selection* problem:

Dialog state	$s_t \approx (\tilde{u}_0,a_1,\tilde{u}_1,a_2,\tilde{u}_2,a_3,\tilde{u}_3,\ldots,a_t,\tilde{u}_t)$
State tracking	$s_{t+1} = G(s_t,a_t,\tilde{u}_n)$
Action selection	$F(s_{t+1}) = a_{t+1}$

Now the problem is what to track in the dialog state *s,* and how to make use of it when choosing actions

The "theory of mind" problem

A real human

What can she/he/it understand? Anything I can explain

How do I know what it can understand? I'm used to speaking to people

Users must think simultaneously about what language the system can understand, and what the system can do – they must form a "theory of mind" about the dialog system

Responses to "How may I help you?"

- Silences and hesitations while users think
  - Leads to end-pointing problems
  - Leads to users confusing themselves
- "Robot" language (hence examples, "speak naturally")
  - Example 1
  - Example 2
- Recognition errors confused with competences
  - > "i need to sign up for a get off benefit" *[no parse]*
  - > "i would like to enroll in a get one" *[no parse]*
  - > "i would like to get help with my dental insurance" <HELP>
  - > "dental insurance" <INSURANCE>

Source: Live calls, human resources dialog system, AT&T

How dialog systems are built today

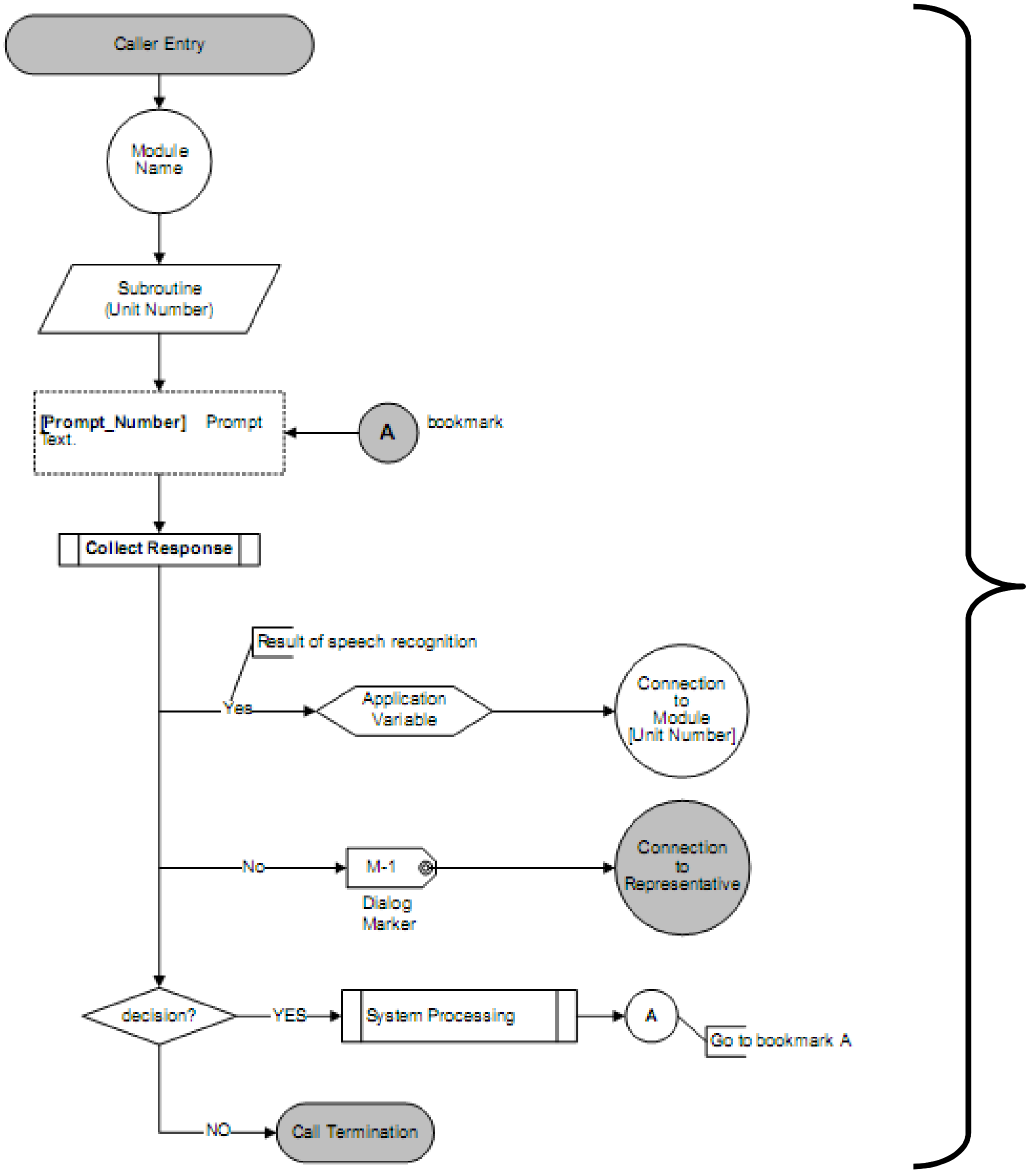
Spoken dialog systems as an application of POMDPs

How dialog systems are built today

s =

reco[1] :	Jason Williams
conf[1]:	0.43
reco[2]:	Jay Wilpon
conf[2]:	0.05
reco[3]:	Jim Wilson
conf[3]:	0.01
name-tries:	2
confirmed-stat:	No
confirmed-tries:	0
confirmed-ID:	{}
match-count[1]:	1
match[1][1]:	jw4796
location[1][1]:	Florham Park
phone-types[1]:	{office, mobile}
phone-types[2]:	{office}
phone-types[3]:	{mobile}
caller-location:	New York
last-call:	Jay Wilpon

How dialog systems are built today



Typical commercial spoken dialog system contains ~100 pages of flowchart

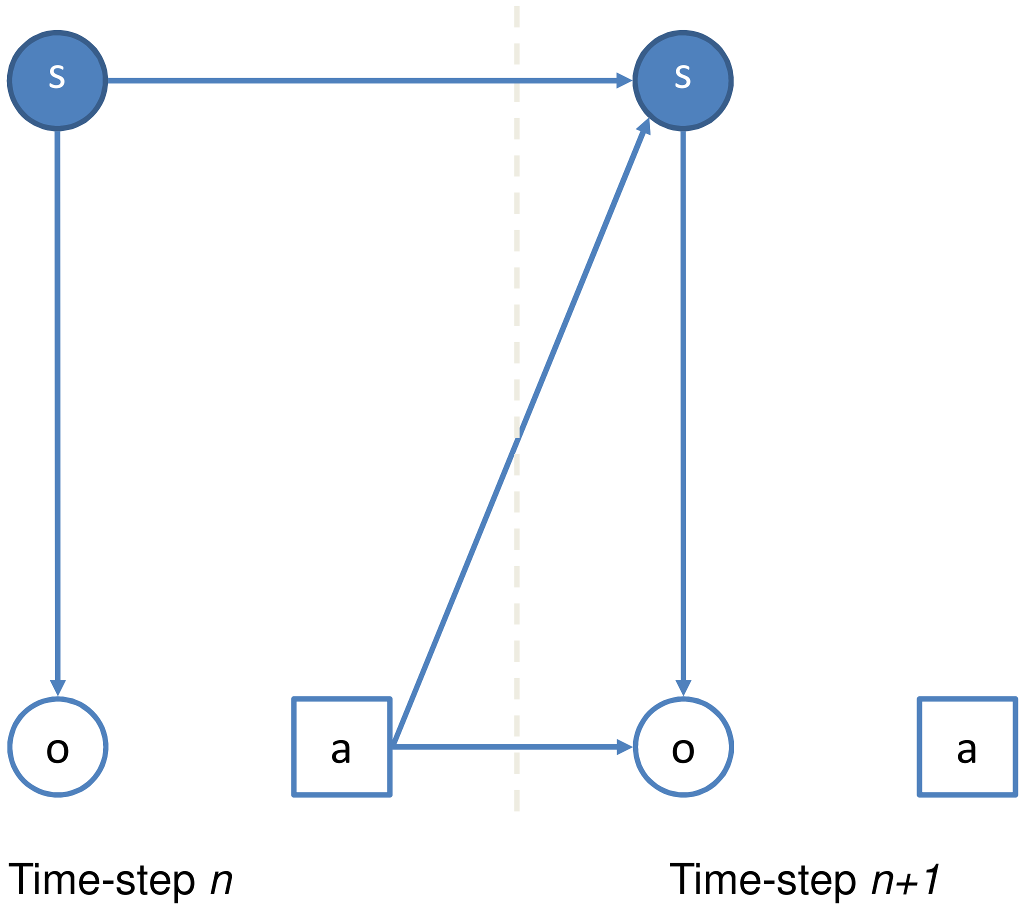
Problems

1. No principled way to encode uncertainty in the dialog
2. No good way to incorporate models of user behavior and ASR errors
3. Actions are chosen locally based on intuition, not globally based on an optimization criteria
4. Good information (N-Best list) is discarded
5. May interact with millions of users, yet will never learn/improve from that experience

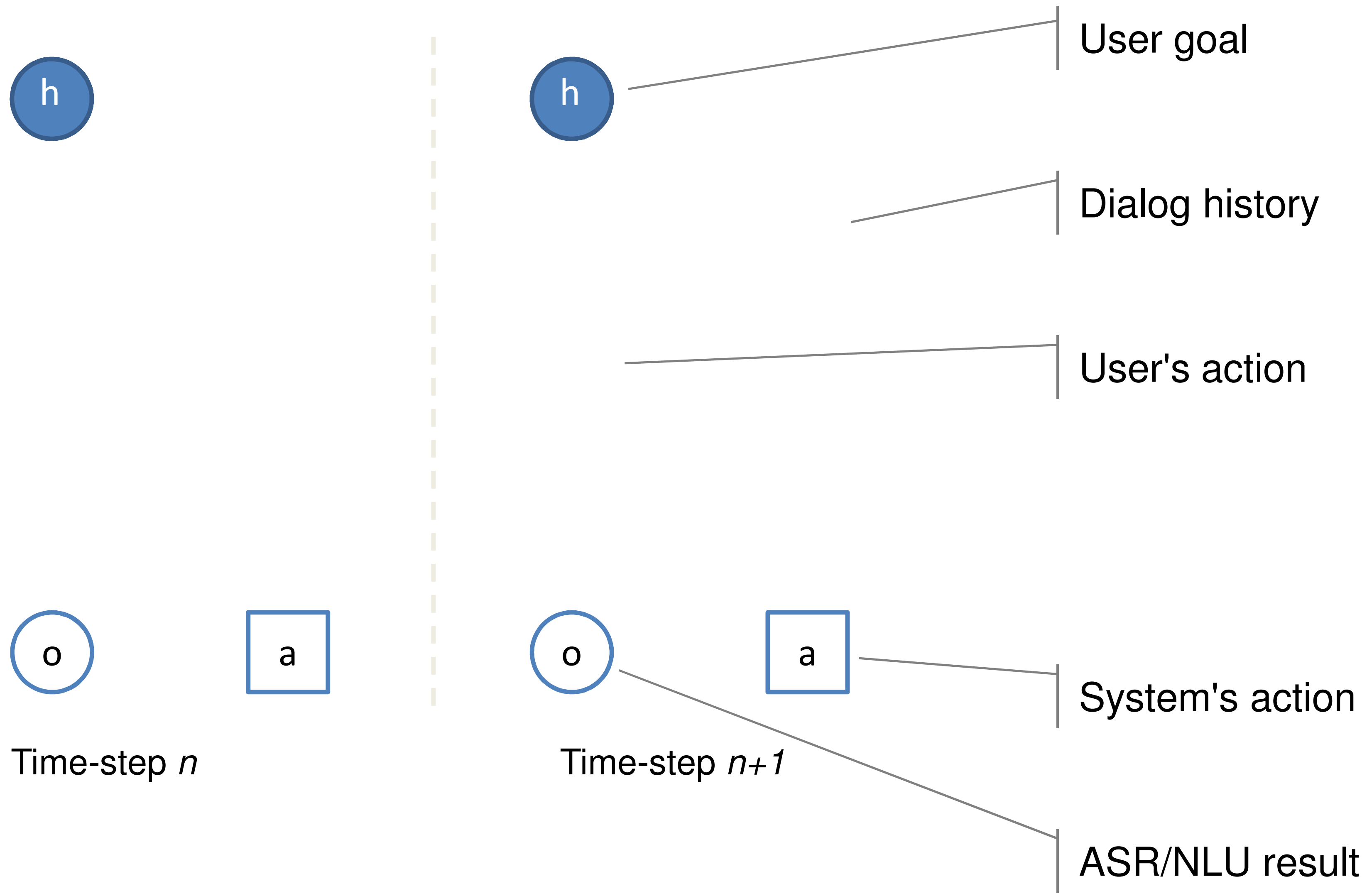
Casting dialog systems as POMDPs

Spoken dialog systems as an application of POMDPs

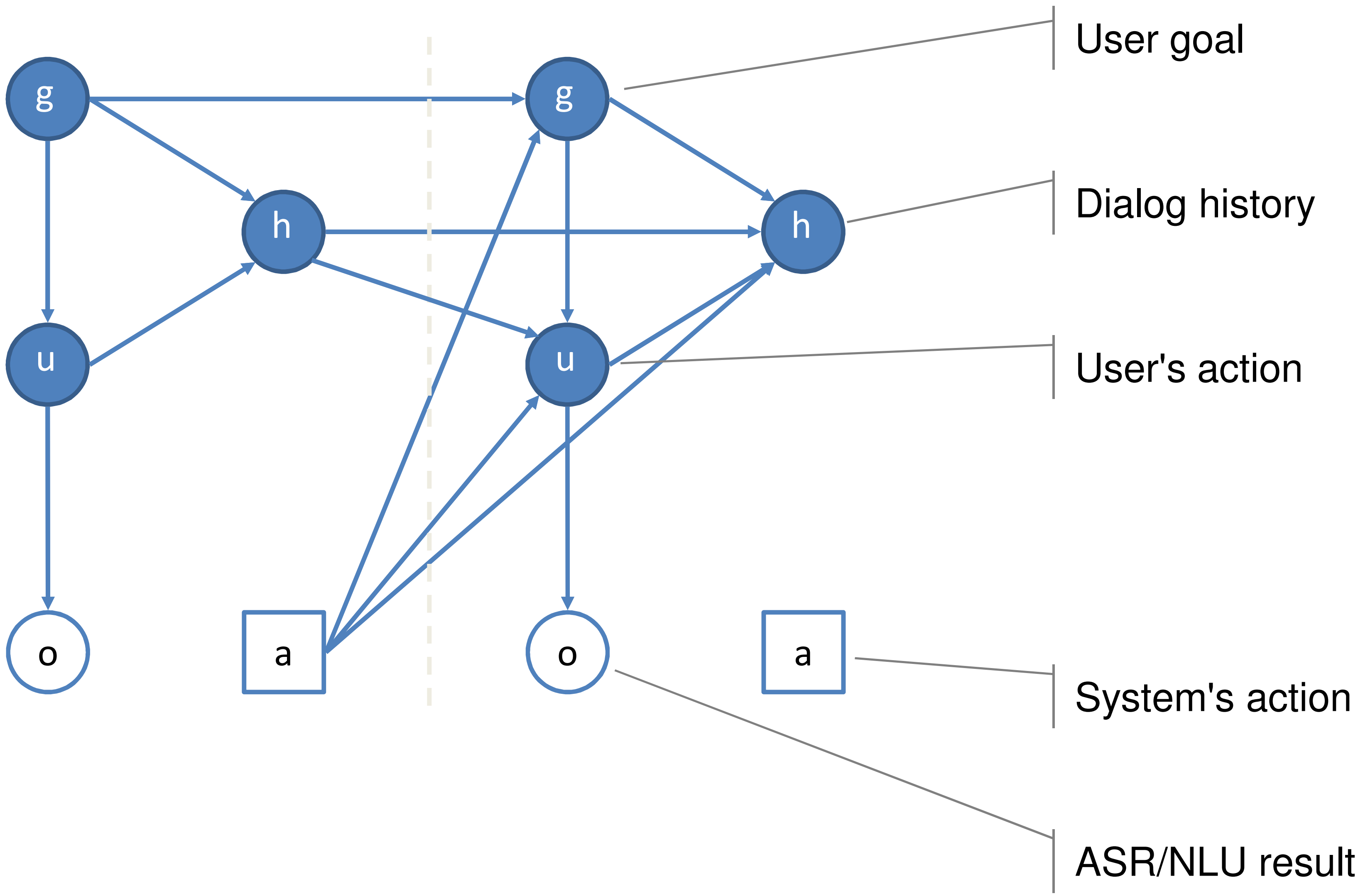
Casting dialog as a POMDP



Casting dialog as a POMDP



Casting dialog as a POMDP

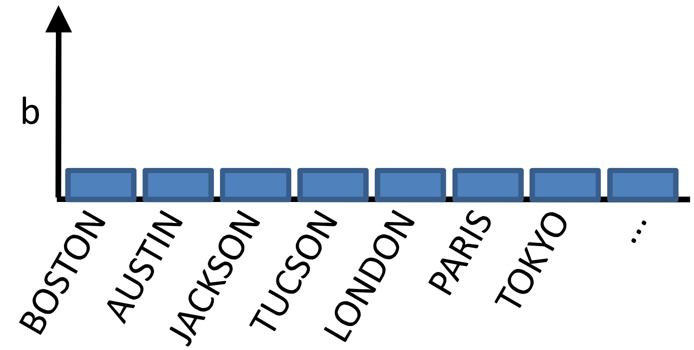
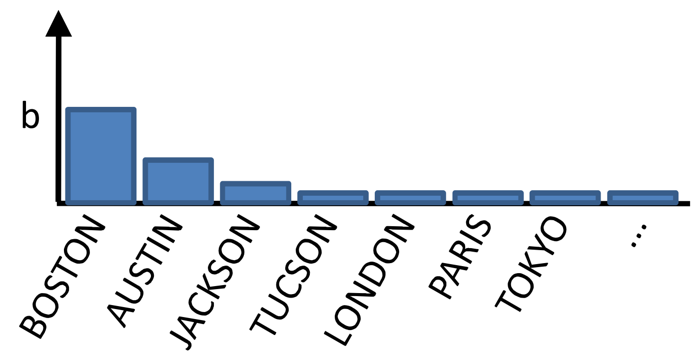
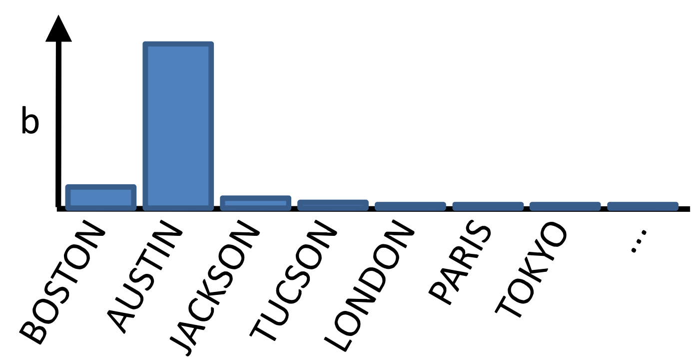


SDS-POMDP update equation

$$\underbrace{b'(g',h')}_{\text{new belief state}} = \underbrace{\eta}_{\text{normalizing constant}} \cdot \sum_{u'} \underbrace{P(o' \mid u',a)}_{\text{ASR model}} \sum_{h} \underbrace{P(u' \mid g',h,a)}_{\text{user action model}} \underbrace{P(h' \mid u',g',h,a)}_{\text{dialog history model}} \sum_{g} \underbrace{P(g' \mid g,a)}_{\text{user goal model}} \underbrace{b(g,h)}_{\text{old belief state}}$$

- Assert a reward function R(s,a)
- Choose actions to maximize expected sum of rewards over the whole dialog

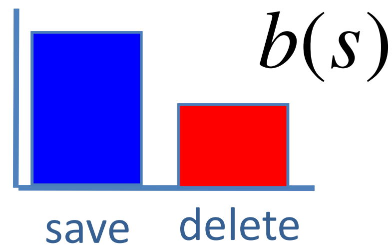
Illustration: synthesizing across N-Best lists

System action	*N-Best list*	*User goal belief*
[prior to start of dialog]		
Which city?	BOSTON ~0.50 AUSTIN ~0.20 JACKSON ~0.10	
Sorry which city?	TUCSON ~0.20 AUSTIN ~0.10	

A Simple Two State Example

"Save"
"Delete"
mumble

o



a

ask
doSave
doDelete

Observation Probability

$$P(o' \mid s', a)$$

eg

"Save"	0.7
"Delete"	0.1
mumble	0.2

$$P(o' \mid save, ask)$$

Transition Probability

$$P(s' \mid s, a)$$

	save	delete
save	1.0	0.0
delete	0.0	1.0

$$P(s' \mid s, ask)$$

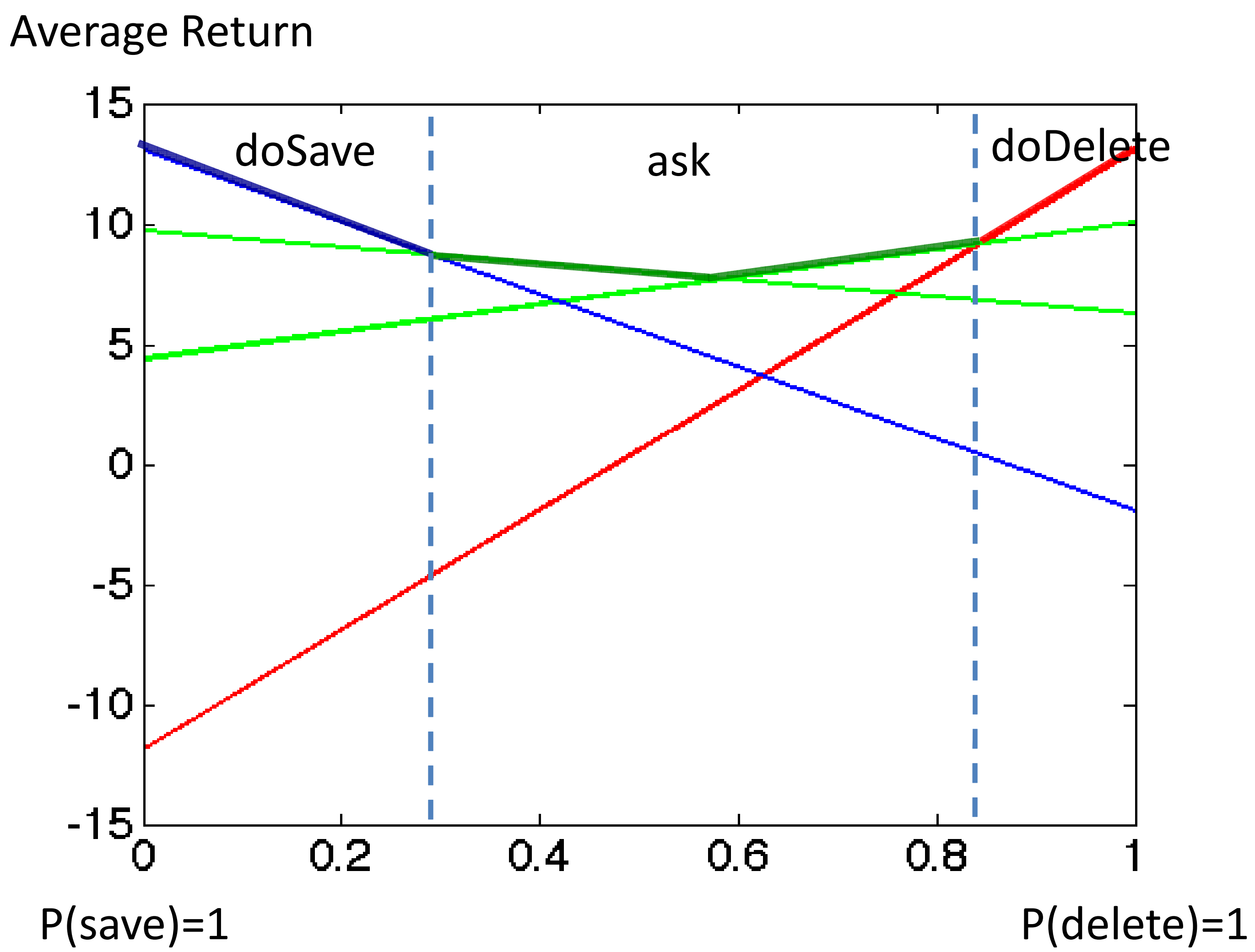
Reward Function

$$R(s, a)$$

	save	delete
ask	-1	-1
doSave	+5	-10
doDelete	-20	+5

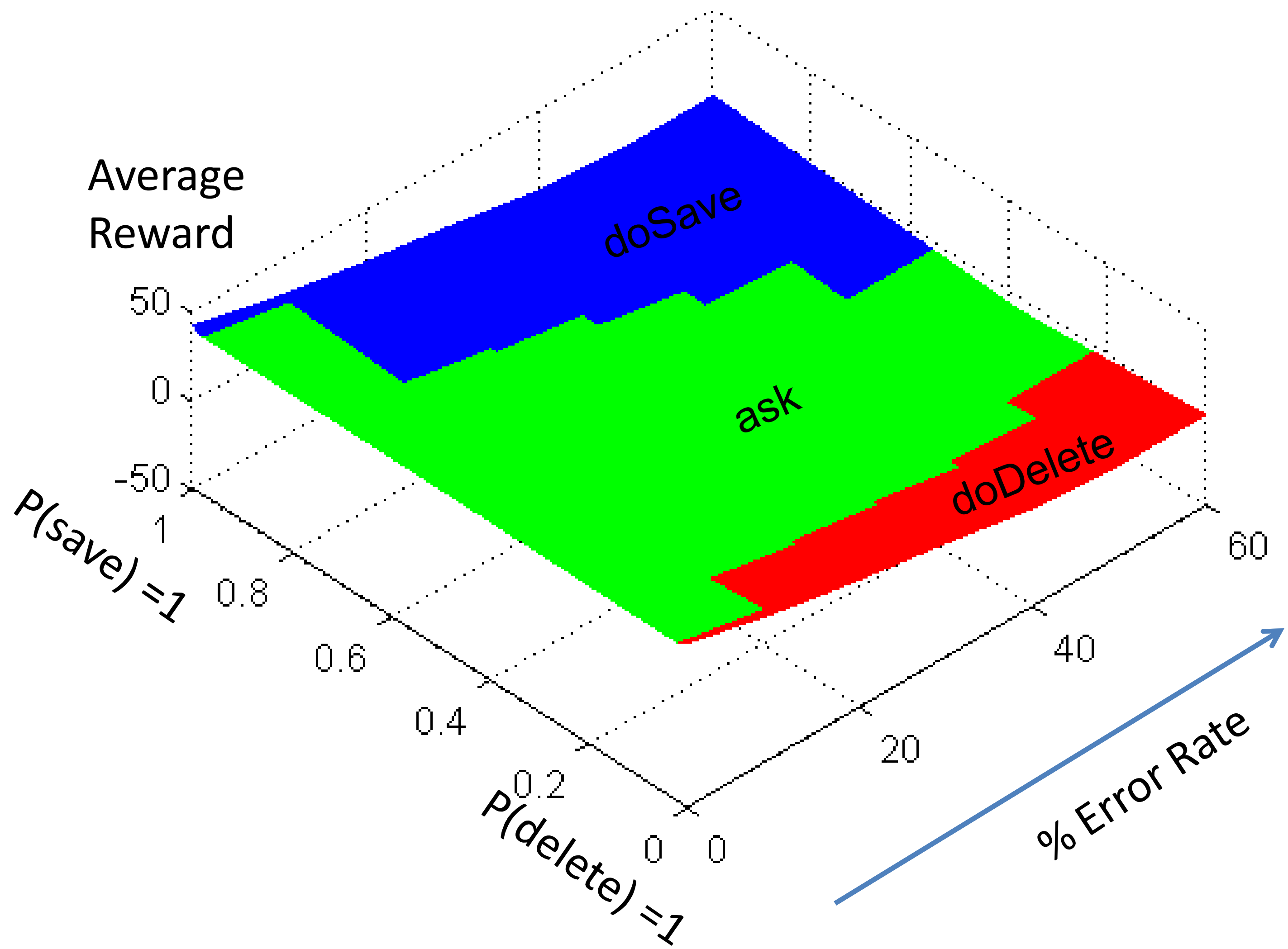
Williams, Young, and Thomson. Statistical approaches to dialogue systems. Tutorial at INTERSPEECH 2009. Brighton, UK.

Policy Value Function at 30% Error Rate



Williams, Young, and Thomson. Statistical approaches to dialogue systems. Tutorial at INTERSPEECH 2009. Brighton, UK.

Policy Value Function vs Error Rate



Williams, Young, and Thomson. Statistical approaches to dialogue systems. Tutorial at INTERSPEECH 2009. Brighton, UK.

Growing up to real-world systems

Spoken dialog systems as an application of POMDPs

Main issues

1) Belief update : must run in real-time

2) Planning

- Scalability
- More sophisticated simulated users
- Expert knowledge & business rules

Although the dialog problem is well-stated as a POMDP, growing to real-world problems has moved away from traditional POMDP solution algorithms

Problem: Updating belief in real-time

$$\underbrace{b'(g',h')}_{\text{new belief state}} = \underbrace{\eta}_{\text{normalizing constant}} \cdot \sum_{u'} \underbrace{P(o' \mid u',a)}_{\text{ASR model}} \sum_{h} \underbrace{P(u' \mid g',h,a)}_{\text{user action model}} \underbrace{P(h' \mid u',g',h,a)}_{\text{dialog history model}} \sum_{g} \underbrace{P(g' \mid g,a)}_{\text{user goal model}} \underbrace{b(g,h)}_{\text{old belief state}}$$

from	1000 values
to	1000 values
time	1000 values
date	1000 values

$|G| = 1000^4$
$= 10^{12}$

Update is $O(|G|^2) = 10^{24}$

We need a response in < 1 s

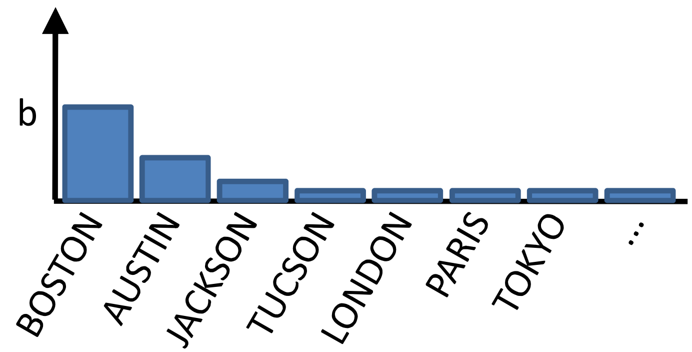
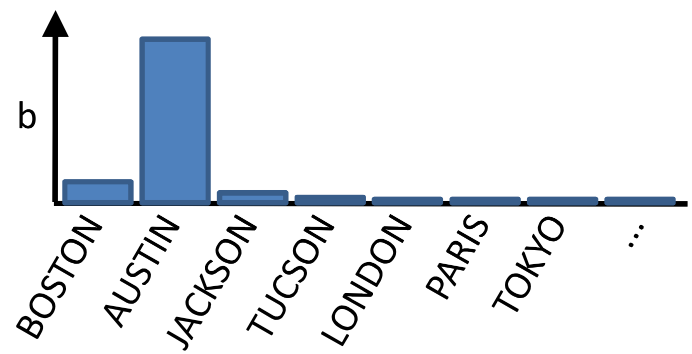
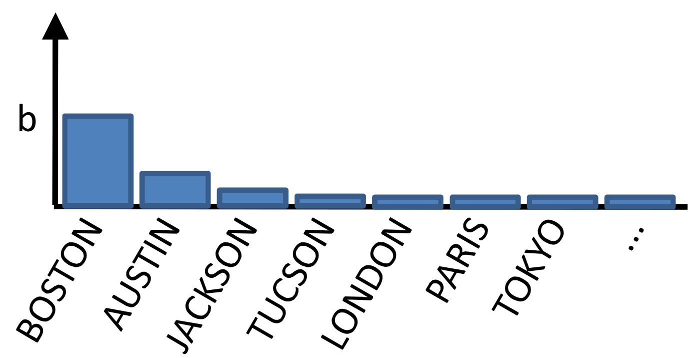
$O(10^{24})$ impossible in real time !

2 methods for efficient belief monitoring

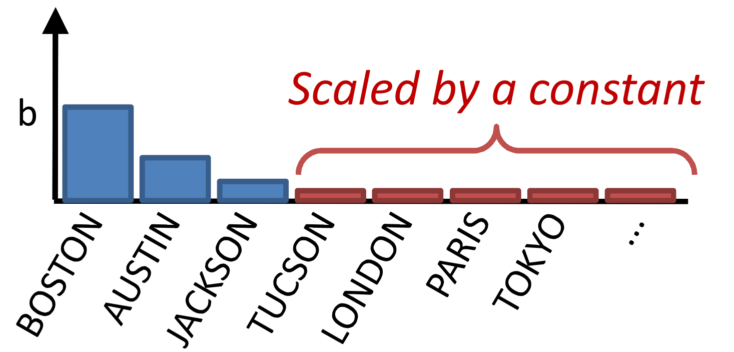
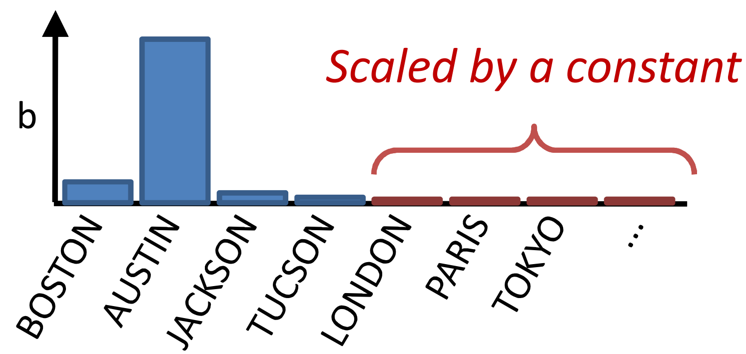
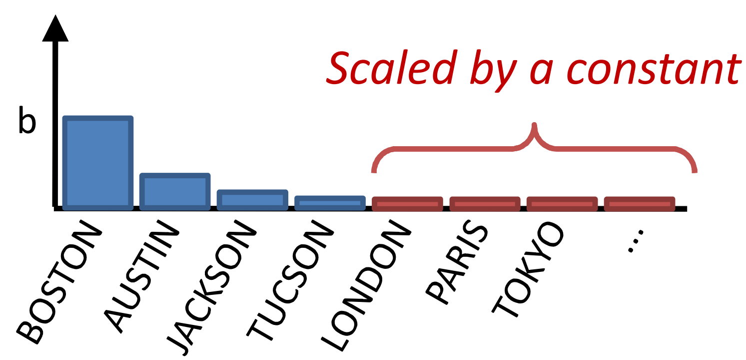
1. **M-Best**: Constrain aspects of the model such that un-observed goals can be tracked en-masse

2. **Factorization**: Decompose the network as much as possible; apply approximate inference techniques from the Bayesian network literature

M-Best partitions: Intuition

System action	*N-Best list*	*User goal belief*
Which city?	BOSTON ~0.50 AUSTIN ~0.20 JACKSON ~0.10	
Sorry, which city?	TUCSON ~0.20 AUSTIN ~0.10	
Was that Austin?	NO ~0.99 YES ~0.01	

M-Best partitions: Intuition

System action	N-Best list	User goal belief
Which city?	BOSTON ~0.50 AUSTIN ~0.20 JACKSON ~0.10	
Sorry, which city?	TUCSON ~0.20 AUSTIN ~0.10	
Was that Austin?	NO ~0.99 YES ~0.01	

Partition update example (maxPartitions = 3)

"First name?"

JASON ~0.6

first: *(any)*
last: *(any)*
1.0000 1.0000

Belief **Prior**

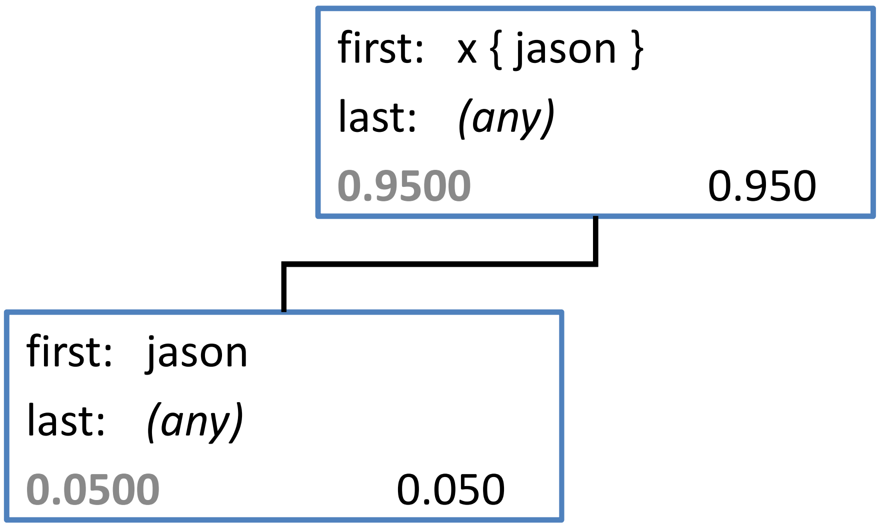
Database of priors

First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010

Partition update example (maxPartitions = 3)

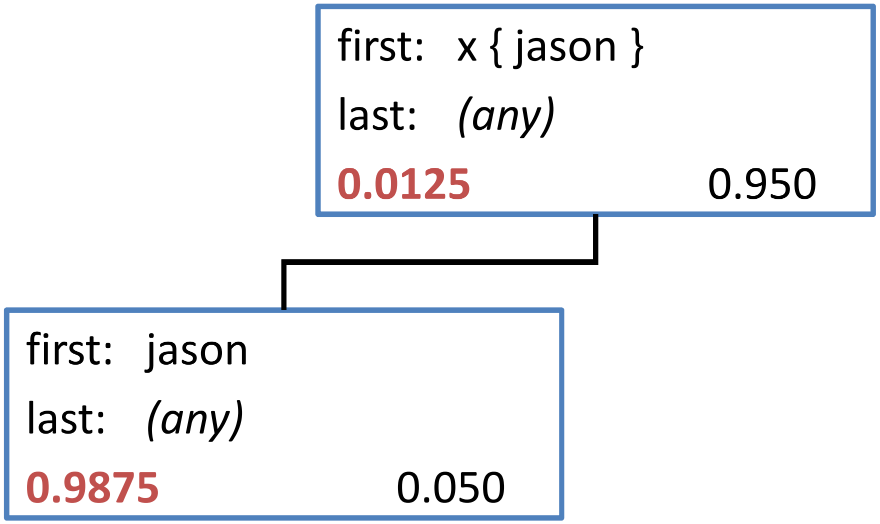
"First name?"

JASON ~0.6



First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010

Partition update example (maxPartitions = 3)

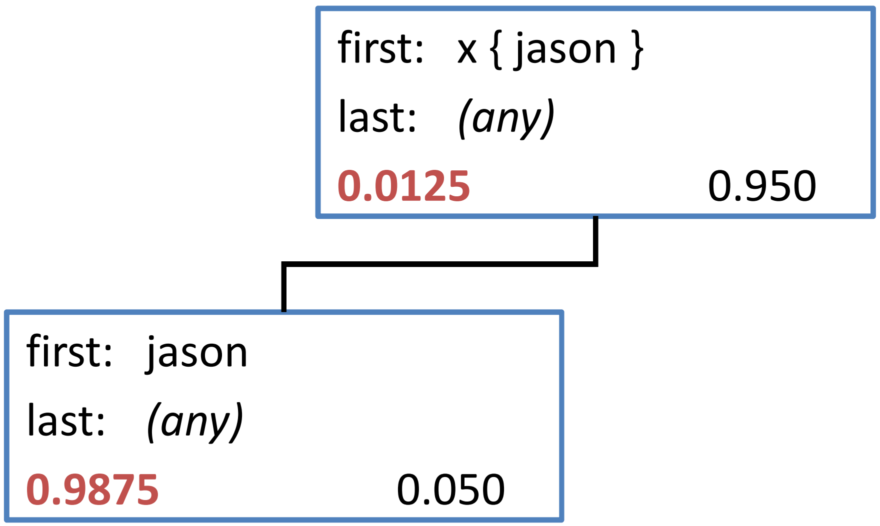


"First name?"

JASON ~0.6

First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010

Partition update example (maxPartitions = 3)



"First name?"

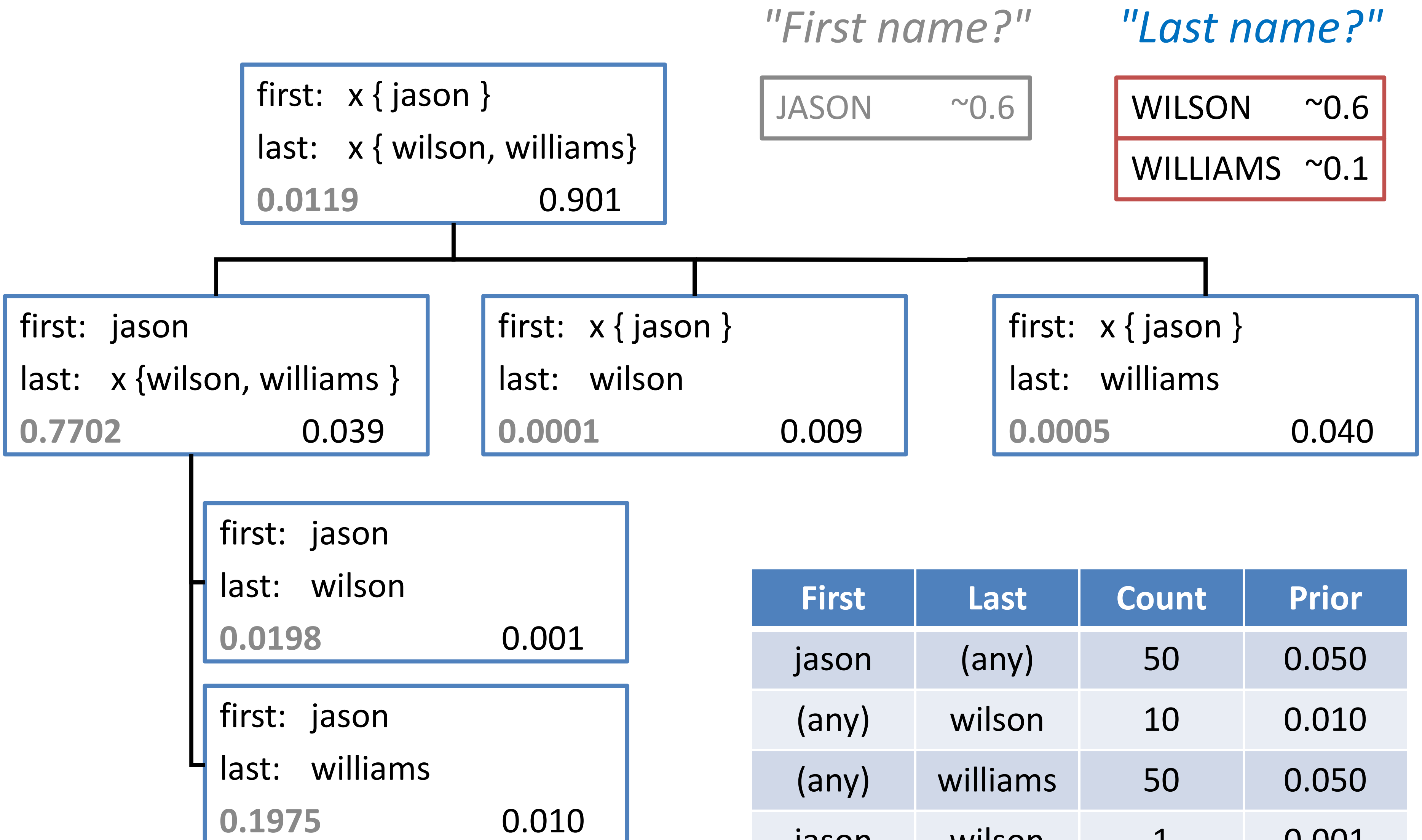
JASON ~0.6

"Last name?"

WILSON ~0.6
WILLIAMS ~0.1

First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010

Partition update example (maxPartitions = 3)



First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010

Partition update example (maxPartitions = 3)

"First name?"

JASON	~0.6

"Last name?"

WILSON	~0.6
WILLIAMS	~0.1

first: x { jason }
last: x { wilson, williams}
0.0001 0.901

first: jason
last: x {wilson, williams }
0.0072 0.039

first: x { jason }
last: wilson
0.0022 0.009

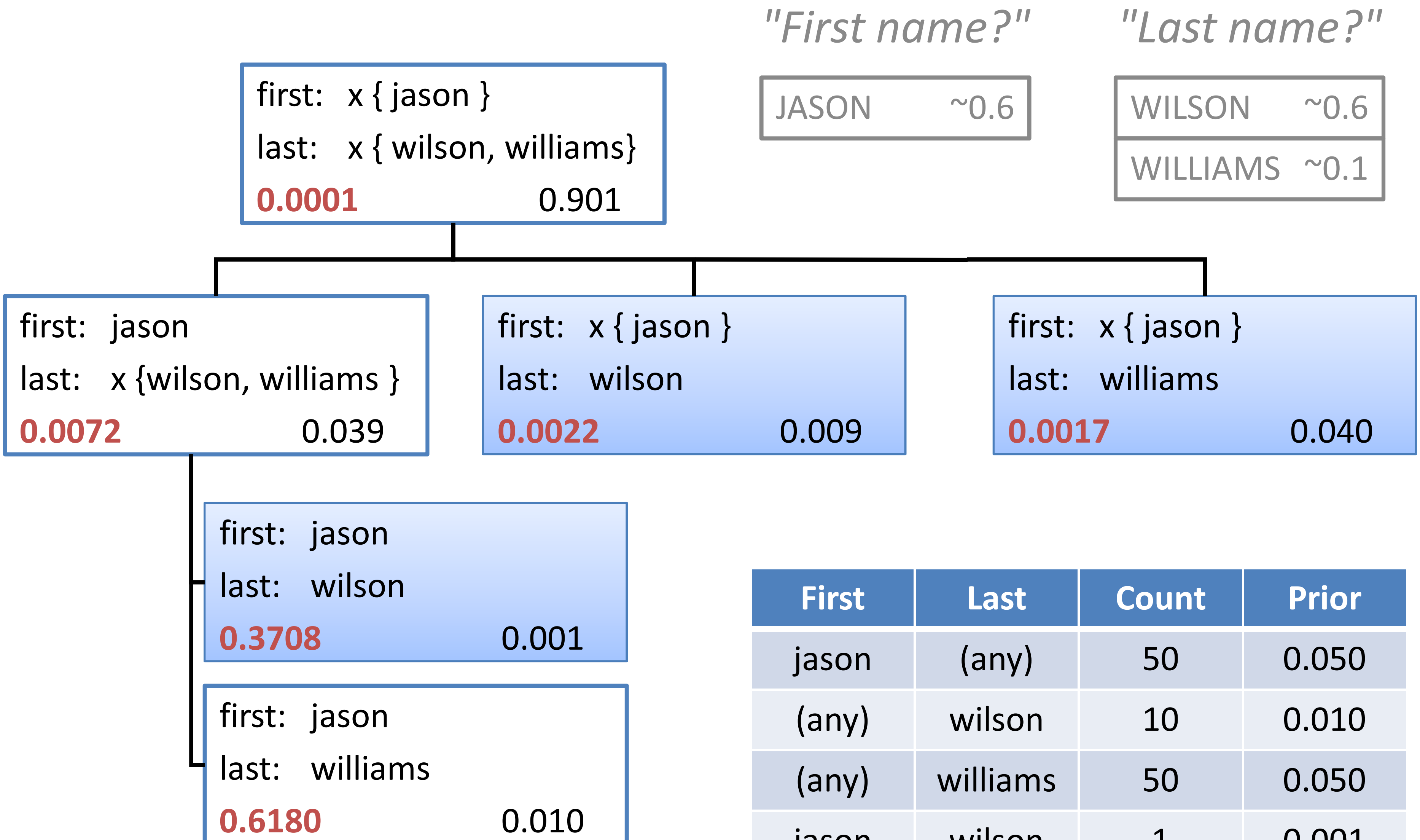
first: x { jason }
last: williams
0.0017 0.040

first: jason
last: wilson
0.3708 0.001

first: jason
last: williams
0.6180 0.010

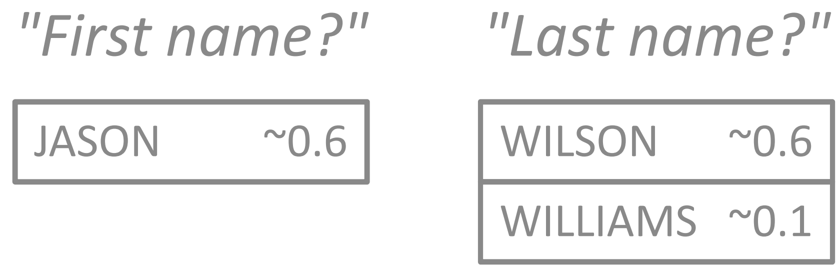
First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010

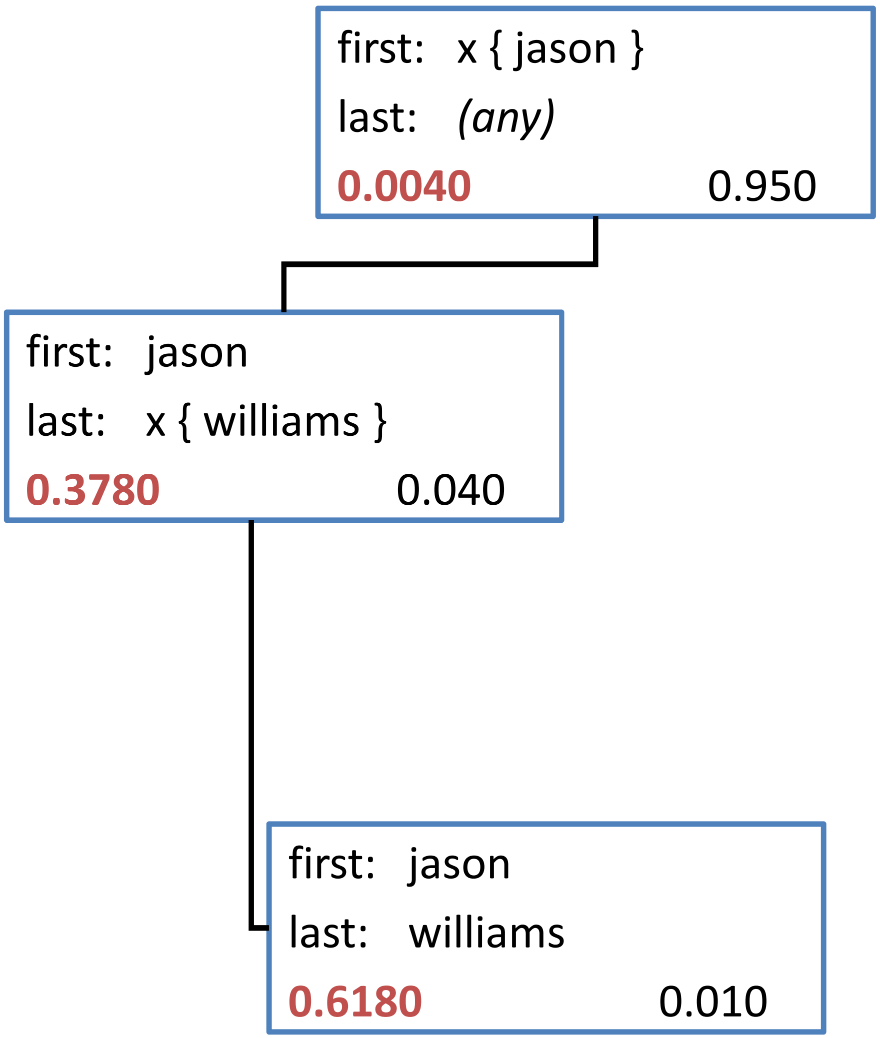
Partition update example (maxPartitions = 3)



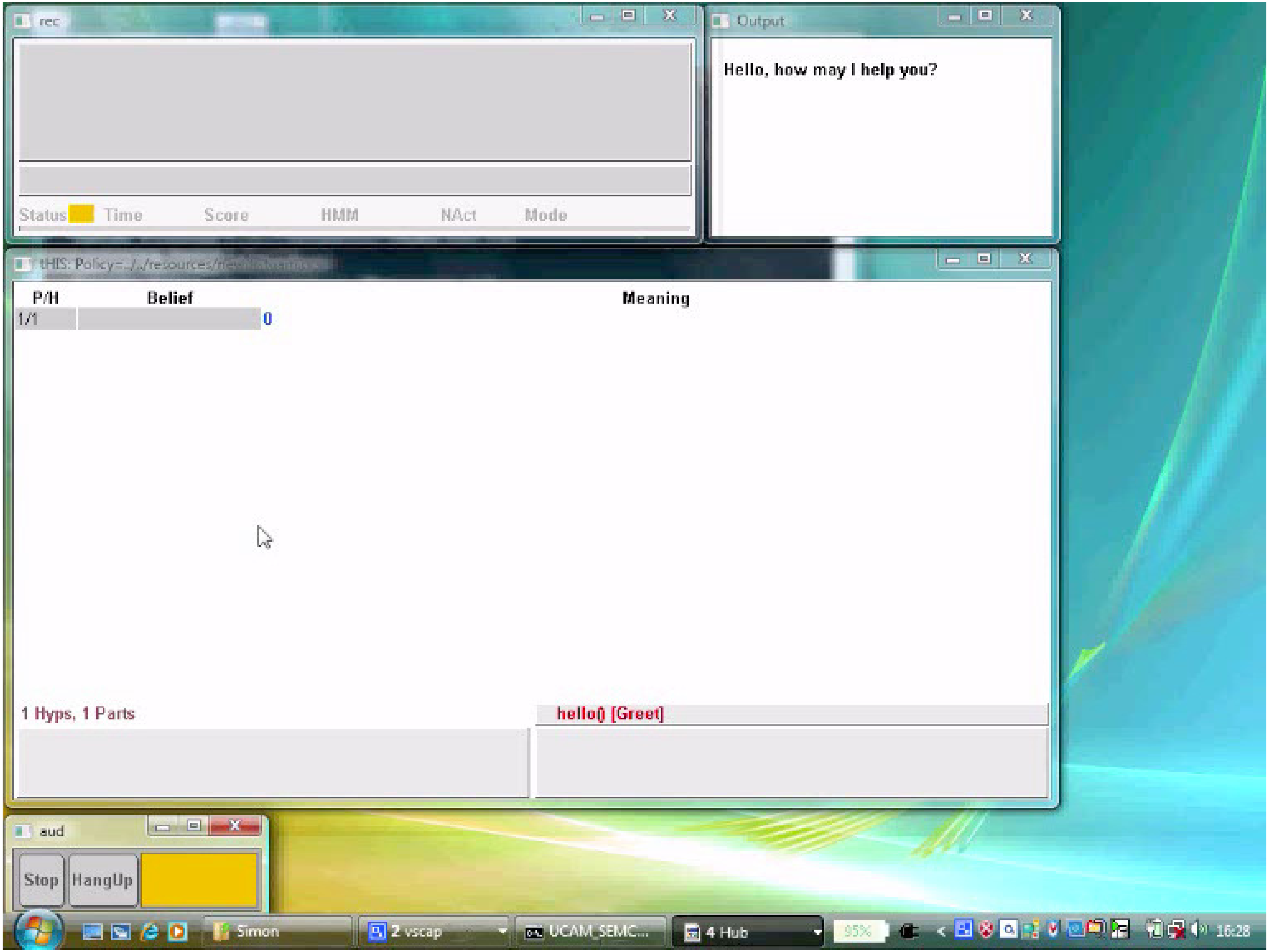
First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010

Partition update example (maxPartitions = 3)

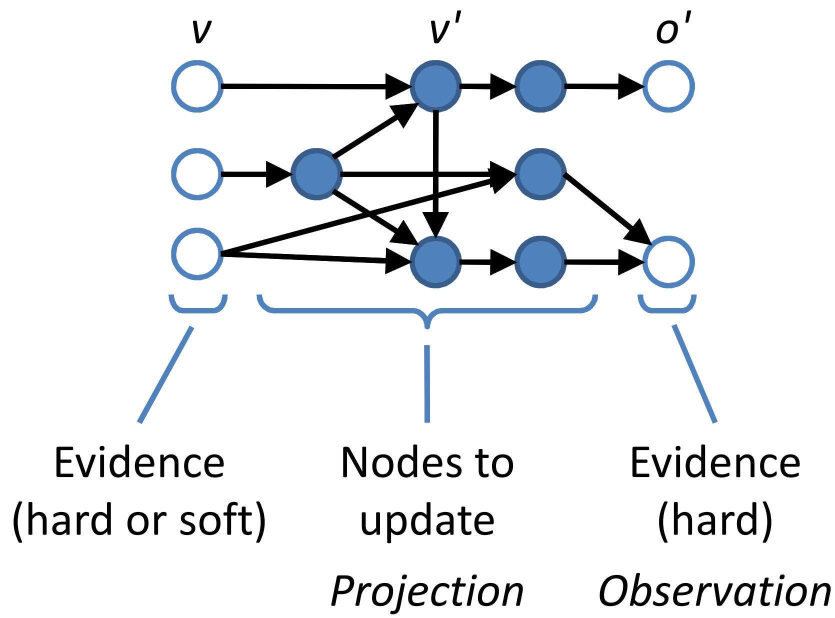




First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010


rec
Status
Time
Score
HMM
NAct
Mode
Output
Hello, how may I help you?
P/H
Belief
Meaning
1/1
0
1 Hyps, 1 Parts
hello() [Greet]
aud
Stop
HangUp
Simon
2 vscap
UCAM_SEMC...
4 Hub
16:28

Network-based approaches



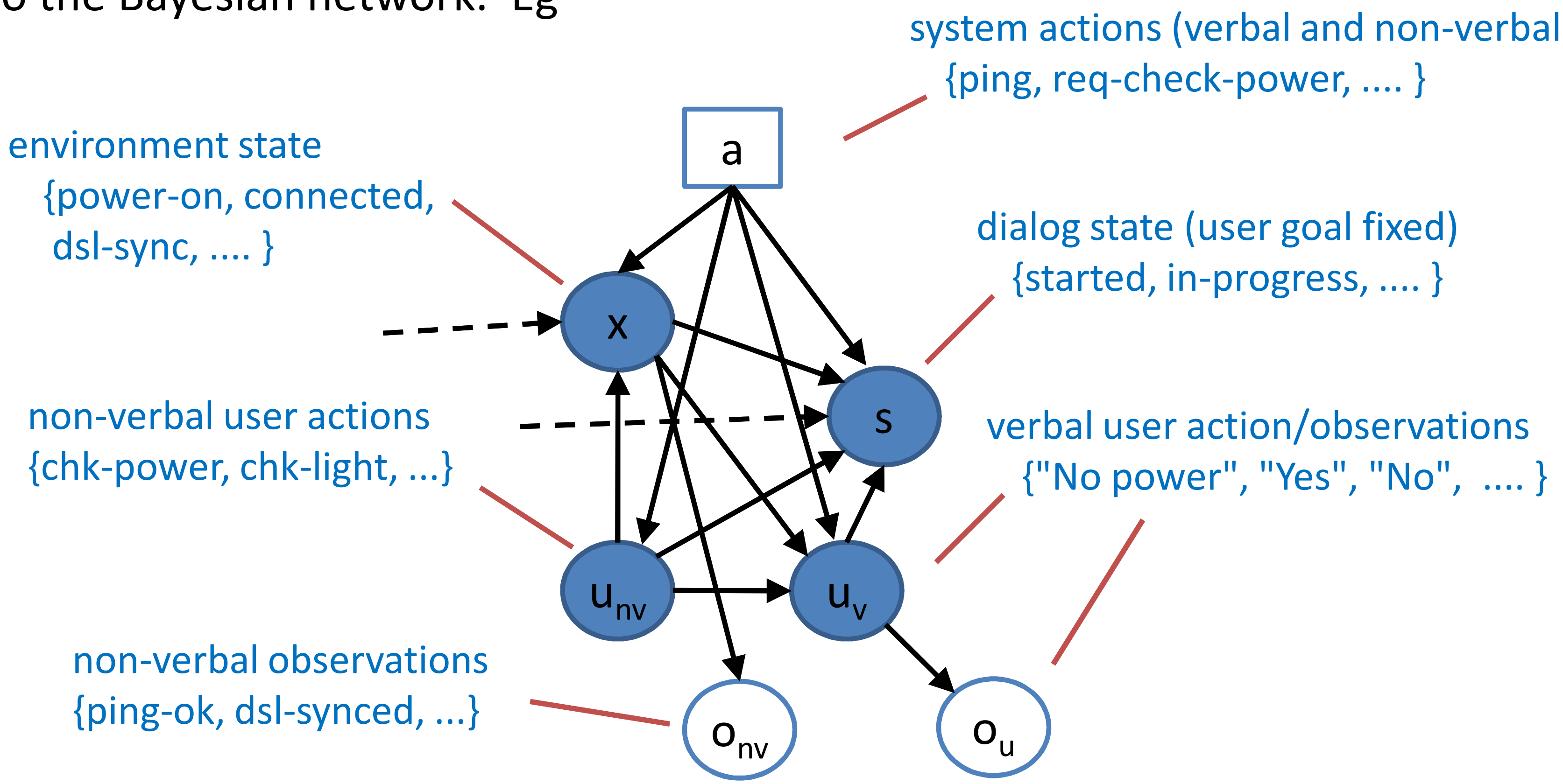
Idea: Apply general purpose Bayes Network inference techniques

Approximate inference can be much faster than exact

Examples: loopy belief propagation and particle filters

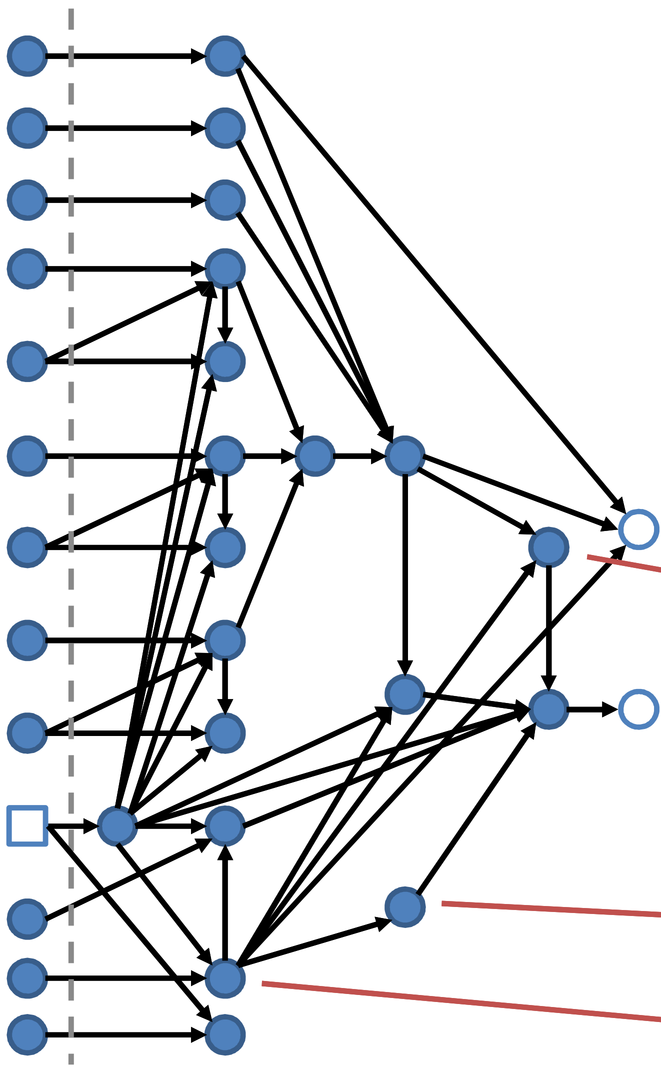
AT&T's Trouble-shooting System

In some applications such as help-lines for DSL modem faults, there are additional sources of uncertainty. These can be easily incorporated into the Bayesian network. Eg



Applying POMDPs to dialog systems in the troubleshooting domain, Williams (Proc W'Shop Bridging the Gap, ACL, 2007)

DSL troubleshooting SDS as a Bayesian network

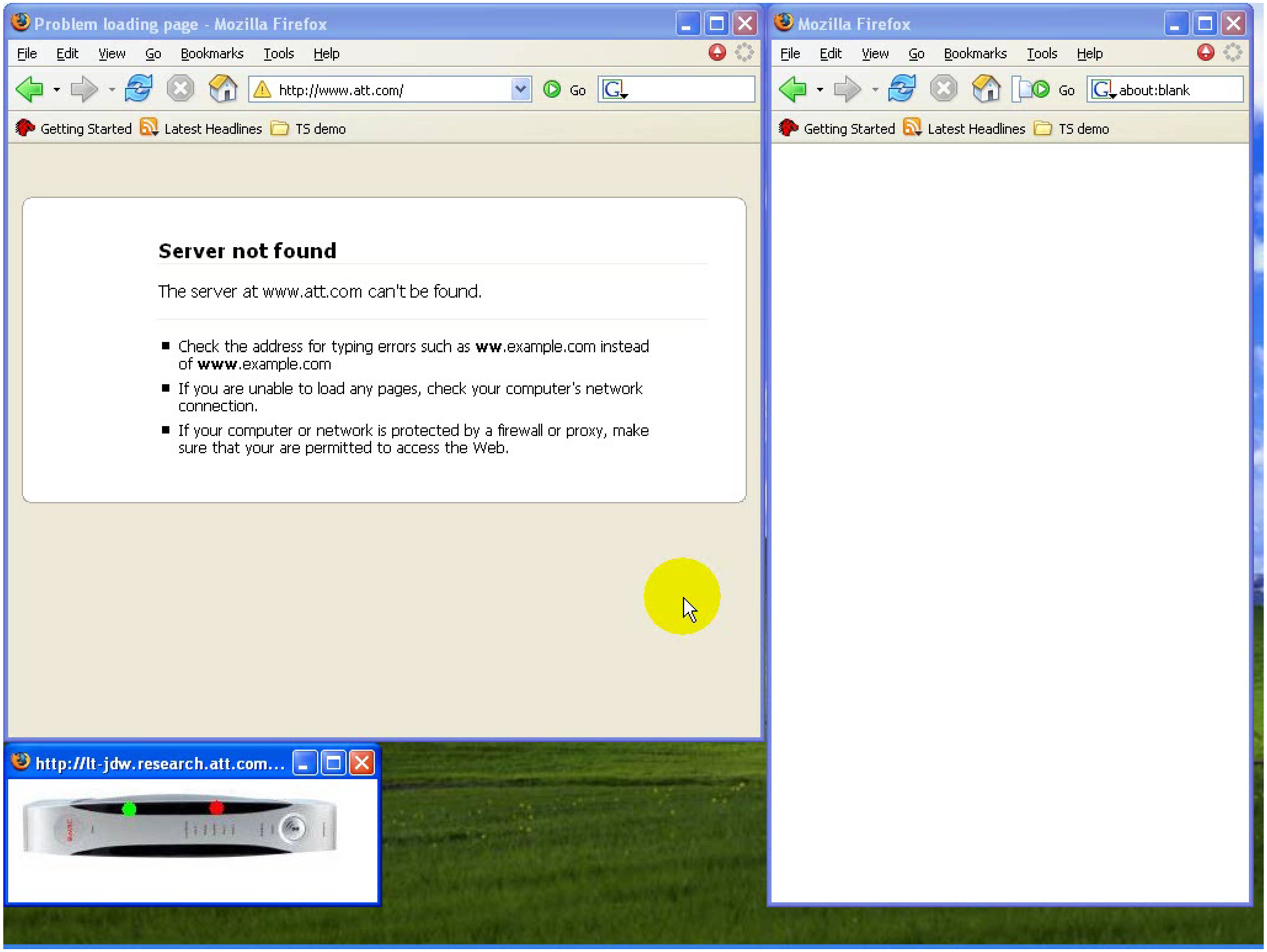


Hidden state components

Service outage
Upstream network failure
Unknown, unfixable problem
Correct username in browser
Correct username on modem
Correct password in browser
Correct password on modem
Correct service type in browser
State of modem network light
Correct service type on modem
Config screen visible in browser
Modem configuration is correct
DSL connection is working
State of modem power light
User opened a webpage
State of DSL modem

Applying POMDPs to dialog systems in the troubleshooting domain, Williams (Proc W'Shop Bridging the Gap, ACL, 2007)

Demonstration of AT&T troubleshooting system



Tracking multiple dialogue states: results

	Domain	Task completion rates: Single state	Multiple states
[1] Higashinaka et al	Room reservation	88%	91%
[2] Henderson & Lemon	Tourist info	67%	73%
[3] Young et al	Tourist info	66%	79%
[4] Thomson & Young	Tourist info	65%	84%

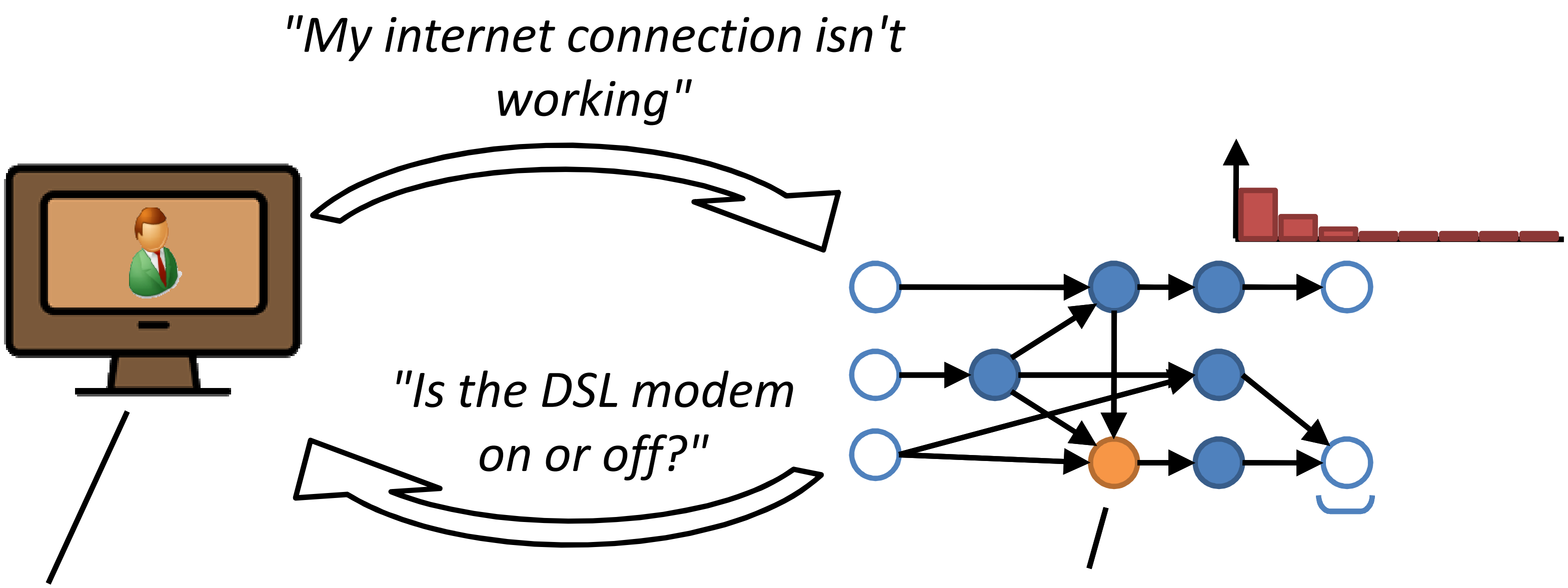
[1] Ryuichiro Higashinaka, Mikio Nakano, Kiyoaki Aikawa, "Corpus-based Discourse Understanding in Spoken Dialogue Systems", ACL, pp240-247, 2003

[2] James Henderson and Oliver Lemon, "Mixture Model POMDPs for Efficient Handling of Uncertainty in Dialogue Management", ACL 2008

[3] S. Young, M. Gasic, S. Keizer, F. Mairesse, J. Schatzmann, B. Thomson and K. Yu (2009). "The Hidden Information State Model: a practical framework for POMDP-based spoken dialogue management." Computer Speech and Language, 24(2): 150-174.

[4] B. Thomson and S. Young (2009). "Bayesian update of dialogue state: A POMDP framework for spoken dialogue systems." Computer Speech and Language, To appear.

Simulated users – two places



External simulated user

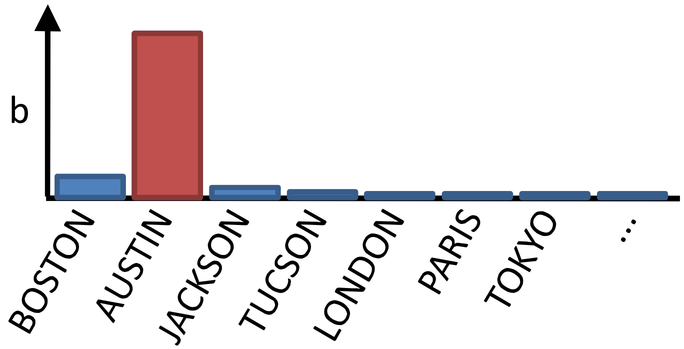
- Generates observations
- High-fidelity
- Programmatic: only generative

Internal simulated user

- Updates belief state
- Low-fidelity
- Tabular: usable for inference

To take advantage of high-fidelity user simulation, an external simulation needs to be "in the learning loop"

Planning : what are the difficult decisions?



All possible actions:

ask

confirm(boston)

confirm(austin)

confirm(jackson)

...

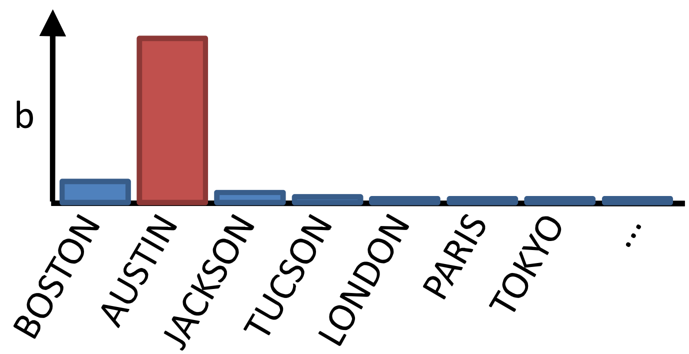
read-weather(boston)

read-weather(austin)

read-weather(jackson)

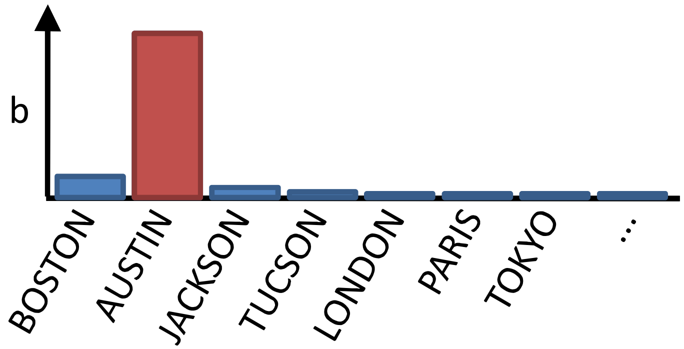
...

Planning : what are the difficult decisions?



All possible actions:	**Useful actions:**
ask	ask
confirm(boston)	confirm(boston)
confirm(austin)	confirm(austin)
confirm(jackson)	confirm(jackson)
...	...
read-weather(boston)	read-weather(boston)
read-weather(austin)	read-weather(austin)
read-weather(jackson)	read-weather(jackson)
...	...

Planning : what are the difficult decisions?



All possible actions:

ask
confirm(boston)
confirm(austin)
confirm(jackson)
...
read-weather(boston)
read-weather(austin)
read-weather(jackson)
...

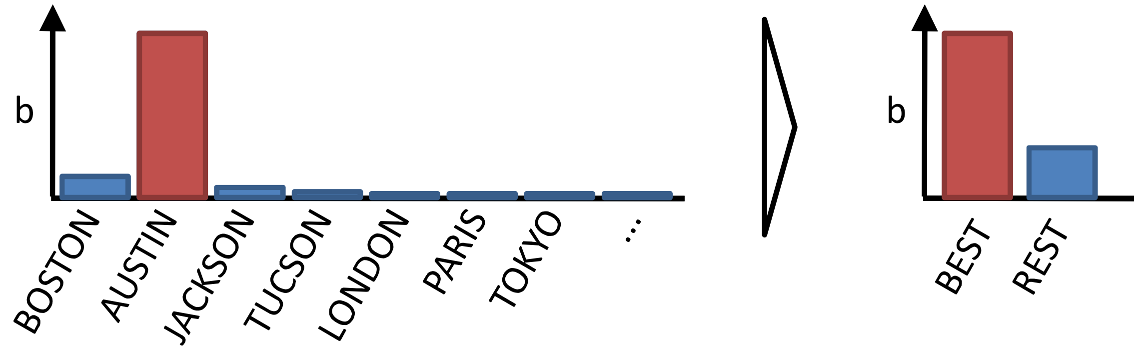
Useful actions:

ask
confirm(boston)
confirm(austin)
confirm(jackson)
...
read-weather(boston)
read-weather(austin)
read-weather(jackson)
...

Summary actions:

ask
confirm(best)
read-weather(best)

Planning : what are the difficult decisions?



All possible actions:

ask
confirm(boston)
confirm(austin)
confirm(jackson)
...
read-weather(boston)
read-weather(austin)
read-weather(jackson)
...

Useful actions:

ask
confirm(boston)
confirm(austin)
confirm(jackson)
...
read-weather(boston)
read-weather(austin)
read-weather(jackson)
...

Summary actions:

ask
confirm(best)
read-weather(best)

Domain knowledge & business rules

People know how to build good dialog systems

- The problem is that people can't consider all of the possible situations

Some actions are just silly and shouldn't be explored

- Don't begin the conversation with a confirmation.
- Don't say "Welcome" except at the start
- ...

Guarantees about system performance must be made

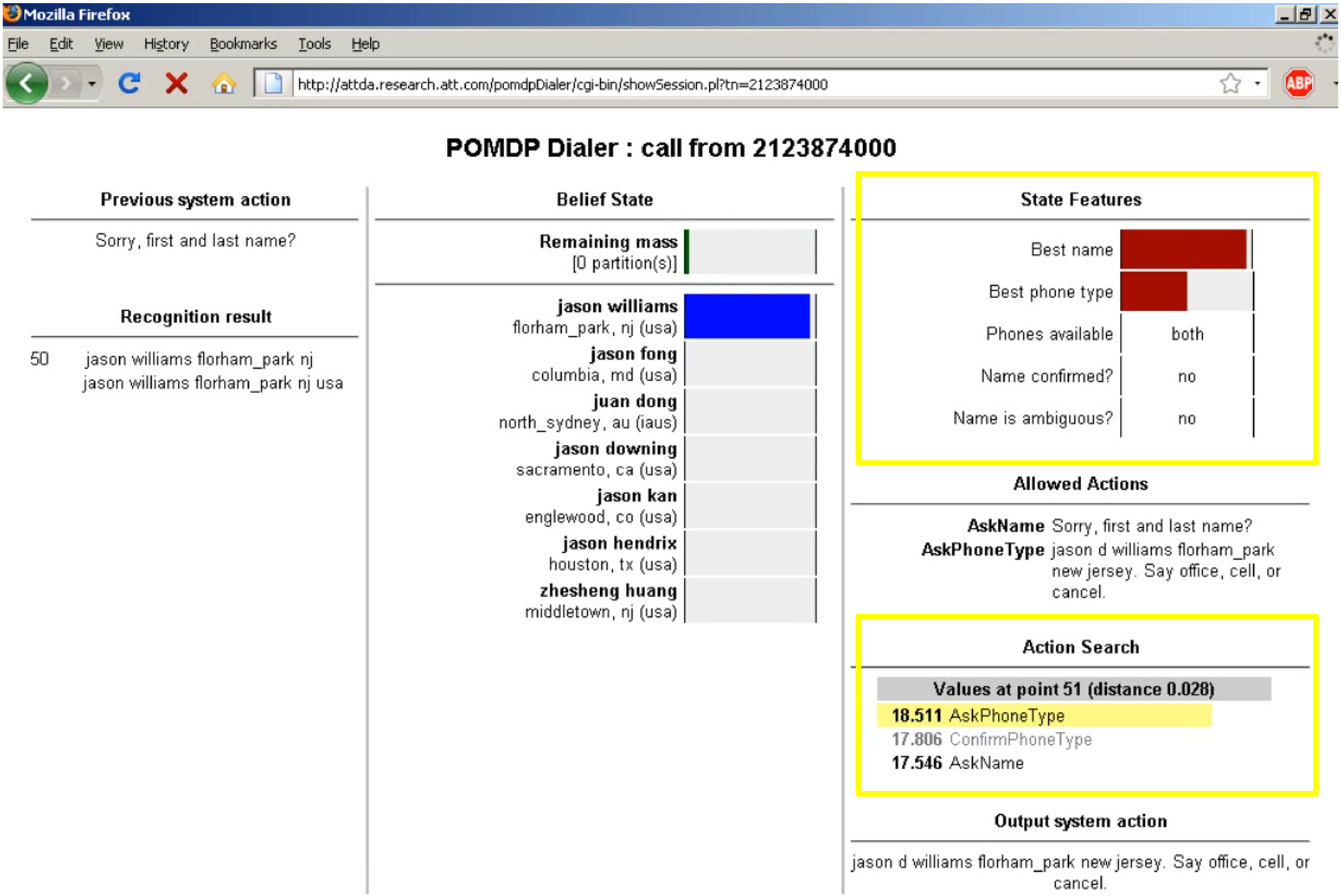
- Only allow funds transfer after password is entered

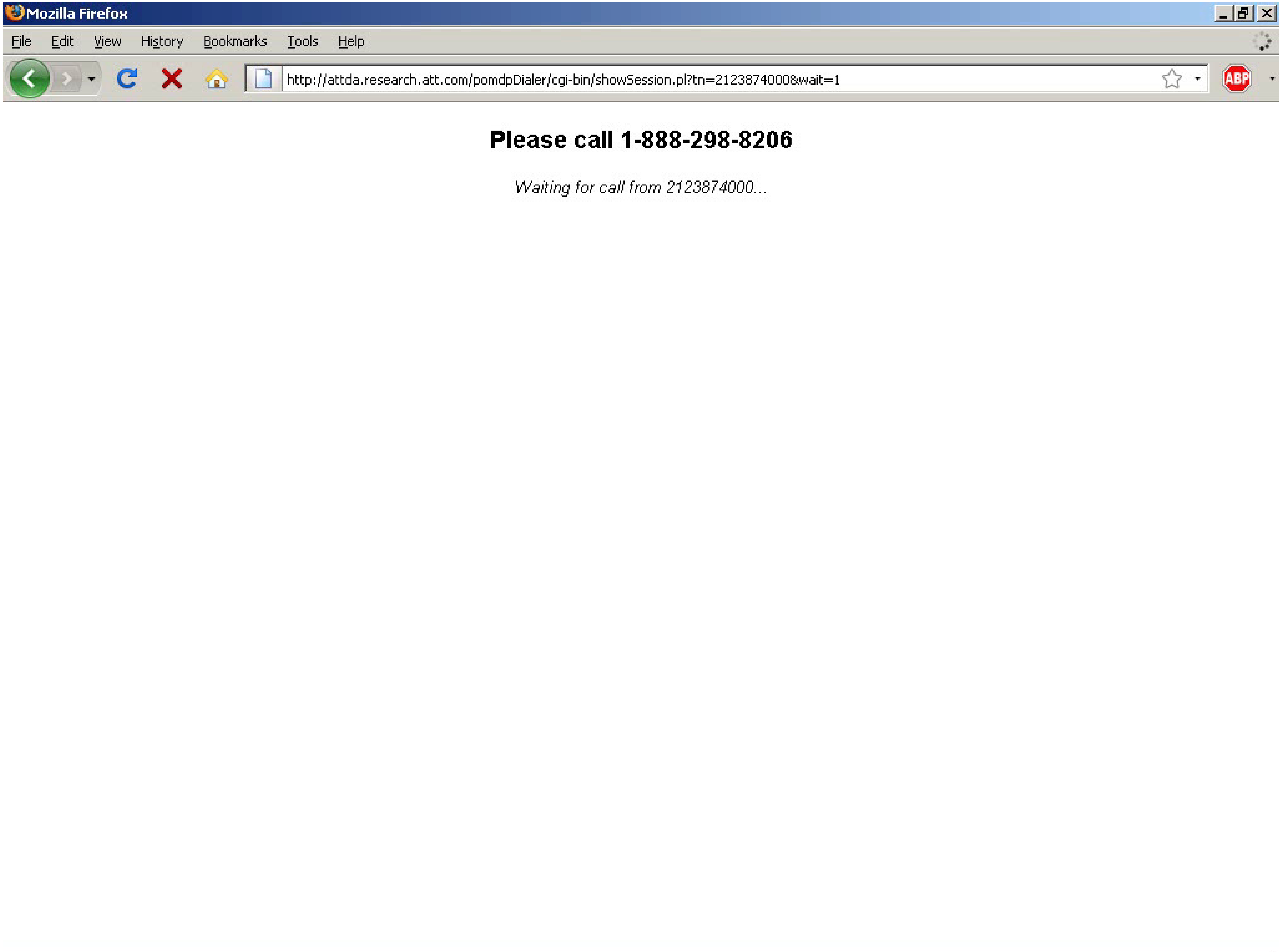
POMDP "Tabula rasa" approach to planning seems inappropriate.
Need a way of incorporating constraints and expert knowledge.

Current approach : Reinforcement learning

- Create a *partial program* which outputs a *set* of one or more acceptable actions
- Extract *features* from the state of the partial program *and* the belief state
- Use RL to choose among the available actions based on the current features

Algorithm	Reference
Natural actor-critic	B. Thomson and S. Young (2010). "Bayesian update of dialogue state: A POMDP framework for spoken dialogue systems." Computer Speech and Language, To appear.
Monte-Carlo sampling	S. Young, M. Gasic, S. Keizer, F. Mairesse, J. Schatzmann, B. Thomson and K. Yu (2009). "The Hidden Information State Model: a practical framework for POMDP-based spoken dialogue management." Computer Speech and Language, 24(2): 150-174. PDF
LSPI w/ feature selection	Lihong Li, Jason D. Williams, and Suhrid Balakrishnan. (2009). Reinforcement Learning for Dialog Management using Least-Squares Policy Iteration and Fast Feature Selection. Proc Interspeech, Brighton, United Kingdom.
SARSA(λ)	J. Henderson, O.Lemon, K.Georgila. (2008). Hybrid reinforcement/supervised learning of dialogue policies from fixed data sets. Computational Linguistics, 34(4):487-511,.
Grid-based value iteration	Jason D. Williams. (2008). Integrating expert knowledge into POMDP optimization for spoken dialog systems. Proc AAAI Workshop on Advancements in POMDP Solvers, Chicago, USA.
Q-MDP	J. Henderson and O.Lemon. (2008). Mixture model POMDPs for efficient handling of uncertainty in dialogue management. In Proc. 46th Annual Meeting of the Association for Computational Linguistics (ACL'08), Columbus, Ohio.


Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://attda.research.att.com/pomdpDialer/cgi-bin/showSession.pl?tn=2123874000
POMDP Dialer : call from 2123874000
Previous system action
Sorry, first and last name?
Recognition result
50 jason williams florham_park nj
jason williams florham_park nj usa
Belief State
Remaining mass
[0 partition(s)]
jason williams
florham_park, nj (usa)
jason fong
columbia, md (usa)
juan dong
north_sydney, au (iaus)
jason downing
sacramento, ca (usa)
jason kan
englewood, co (usa)
jason hendrix
houston, tx (usa)
zhesheng huang
middletown, nj (usa)
State Features
Best name
Best phone type
Phones available both
Name confirmed? no
Name is ambiguous? no
Allowed Actions
AskName Sorry, first and last name?
AskPhoneType jason d williams florham_park new jersey. Say office, cell, or cancel.
Action Search
Values at point 51 (distance 0.028)
18.511 AskPhoneType
17.806 ConfirmPhoneType
17.546 AskName
Output system action
jason d williams florham_park new jersey. Say office, cell, or cancel.


Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://attda.research.att.com/pomdpDialer/cgi-bin/showSession.pl?tn=2123874000&wait=1
ABP
Please call 1-888-298-8206
Waiting for call from 2123874000...

Reinforcement Learning: results

	Domain	Task completion Baseline	RL
[1] Singh et al, 2002	Tourist info	20-64%	88%
[2] Lemon et al, 2006	Tourist info	68%	82%
[3] Frampton & Lemon, 2008	Tourist info	82%	91%
[4] Young et al, 2009	Tourist info	64%	79%
[5] Thomson & Young, 2009	Tourist info	84%	75%
[6] Cuayahuitl et al, 2010	Flight booking	94%	95%

[1] S Singh, DJ Litman, M Kearns, and M Walker, “Optimizing dialogue management with reinforcement learning: Experiments with the NJFun system,” Journal of Artificial Intelligence Research, 2002.

[2] Oliver Lemon, Kallirroi Georgila, James Henderson, "Evaluating Effectiveness and Portability of Reinforcement Learned Dialogue Strategies with real users: the TALK TownInfo Evaluation", IEEE/ACL Spoken Language Technology, 2006.

[3] Matthew Frampton and Oliver Lemon. 2008. Using dialogue acts to learn better repair strategies. Proc ICASSP 2008.

[4] S. Young, M. Gasic, S. Keizer, F. Mairesse, J. Schatzmann, B. Thomson and K. Yu (2009). "The Hidden Information State Model: a practical framework for POMDP-based spoken dialogue management." Computer Speech and Language, 24(2): 150-174.

[5] B. Thomson and S. Young (2009). "Bayesian update of dialogue state: A POMDP framework for spoken dialogue systems." Computer Speech and Language, To appear.

[6] Heriberto Cuayáhuitl, Steve Renals, Oliver Lemon, Hiroshi Shimodaira, "Evaluation of a hierarchical reinforcement learning spoken dialogue system", Computer Speech and Language, (to appear)

Some thoughts on the future

Spoken dialog systems as an application of POMDPs

What happened to POMDPs?

The problem is definitely a POMDP...

... but current solutions do not apply (what we call) "POMDP algorithms"

Are there useful learnings for developers of POMDP algorithms?

A wish-list for POMDP algorithm developers

- **Lifted POMDPs**: Can inference *and planning* be done in a lifted (first-order logic) space?
- **Designer knowledge:** Are there good ways of incorporating designer knowledge into planning?
- **High-fidelity simulations:** Is there a principled way they can be incorporated into optimization?
- **POMDPs and standard programming languages**: Is there a good toolkit based on Python or Java?

If you want to get started...

Some tools are available:

- **AT&T Statistical Dialog Toolkit**
  Efficiently track multiple dialog states
  www.research.att.com/people/Williams_Jason_D
- **AT&T Speech Mash-ups**
  Speech recognition & synthesis "in the cloud"
  https://service.research.att.com/smm

Thanks!

Jason D. Williams

Spoken dialog systems as an application of POMDPs

ICAPS Workshop – May 2010