Spoken dialog systems as an application of planning under uncertainty

Jason D. Williams

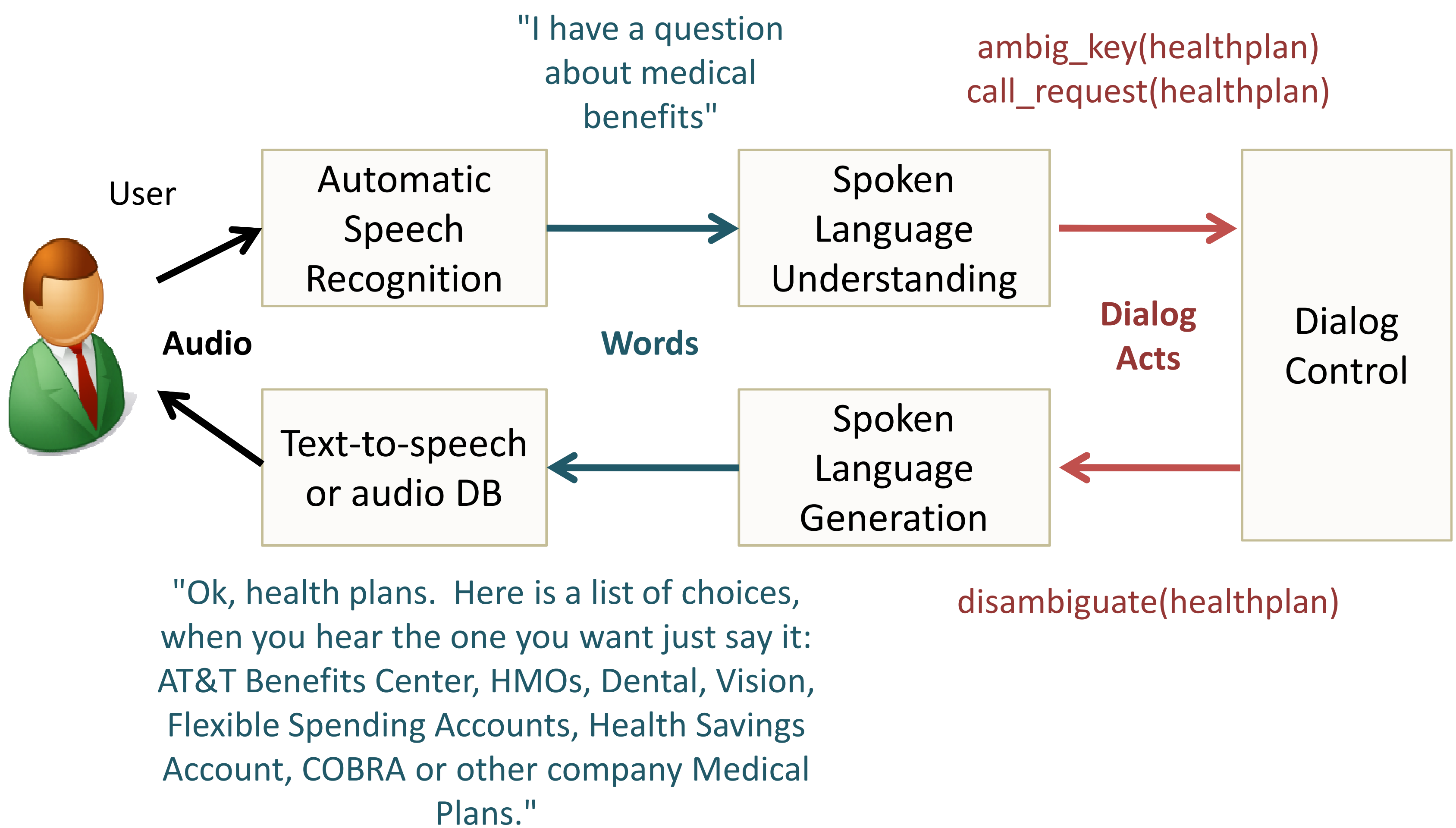
at&t

ICAPS – June 2011

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

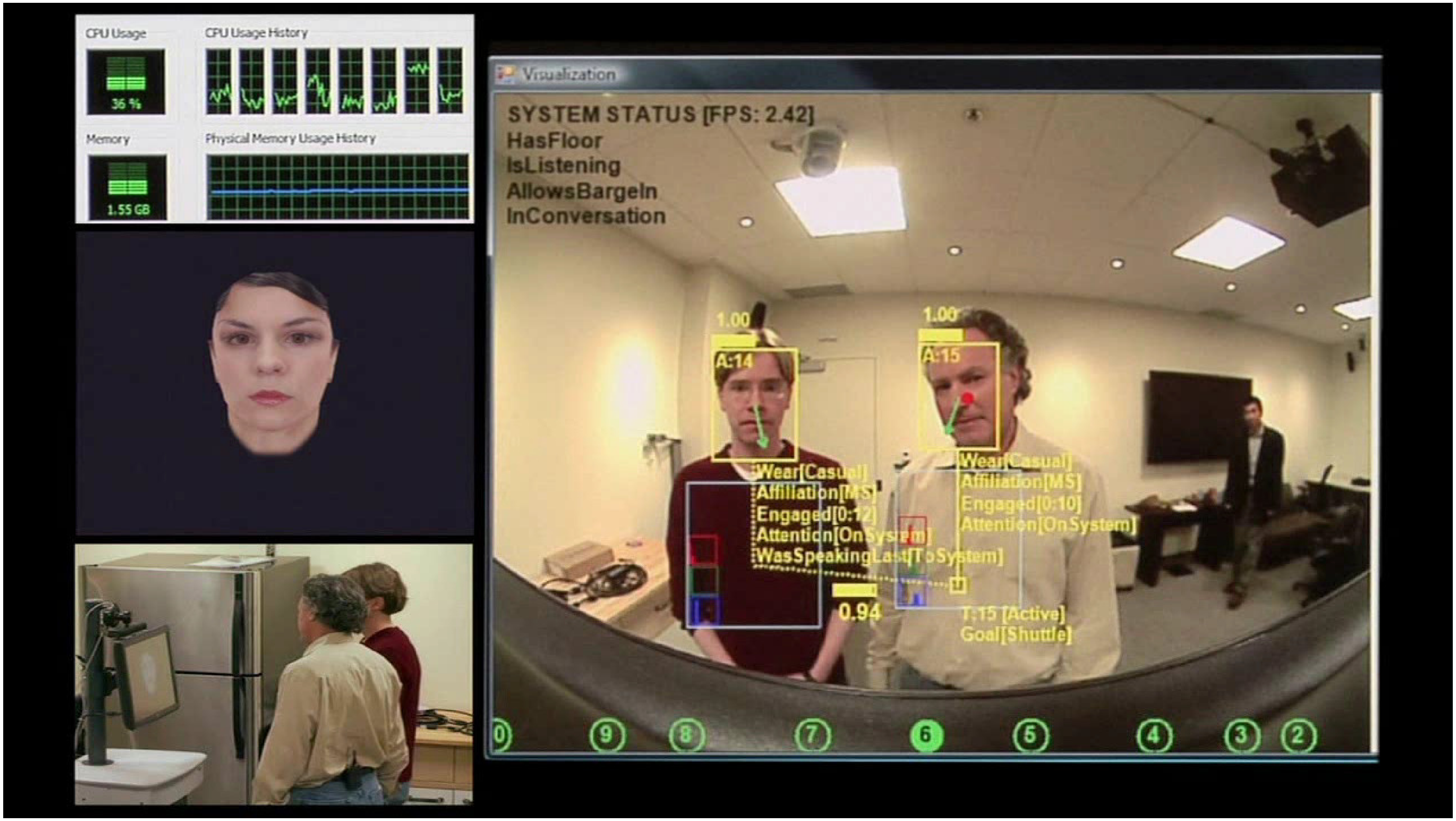
Input	Output	Example
Speech	Speech	Telephone technical support [1]

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

In-car spoken dialogue system

Source: IBM

Automated receptionist

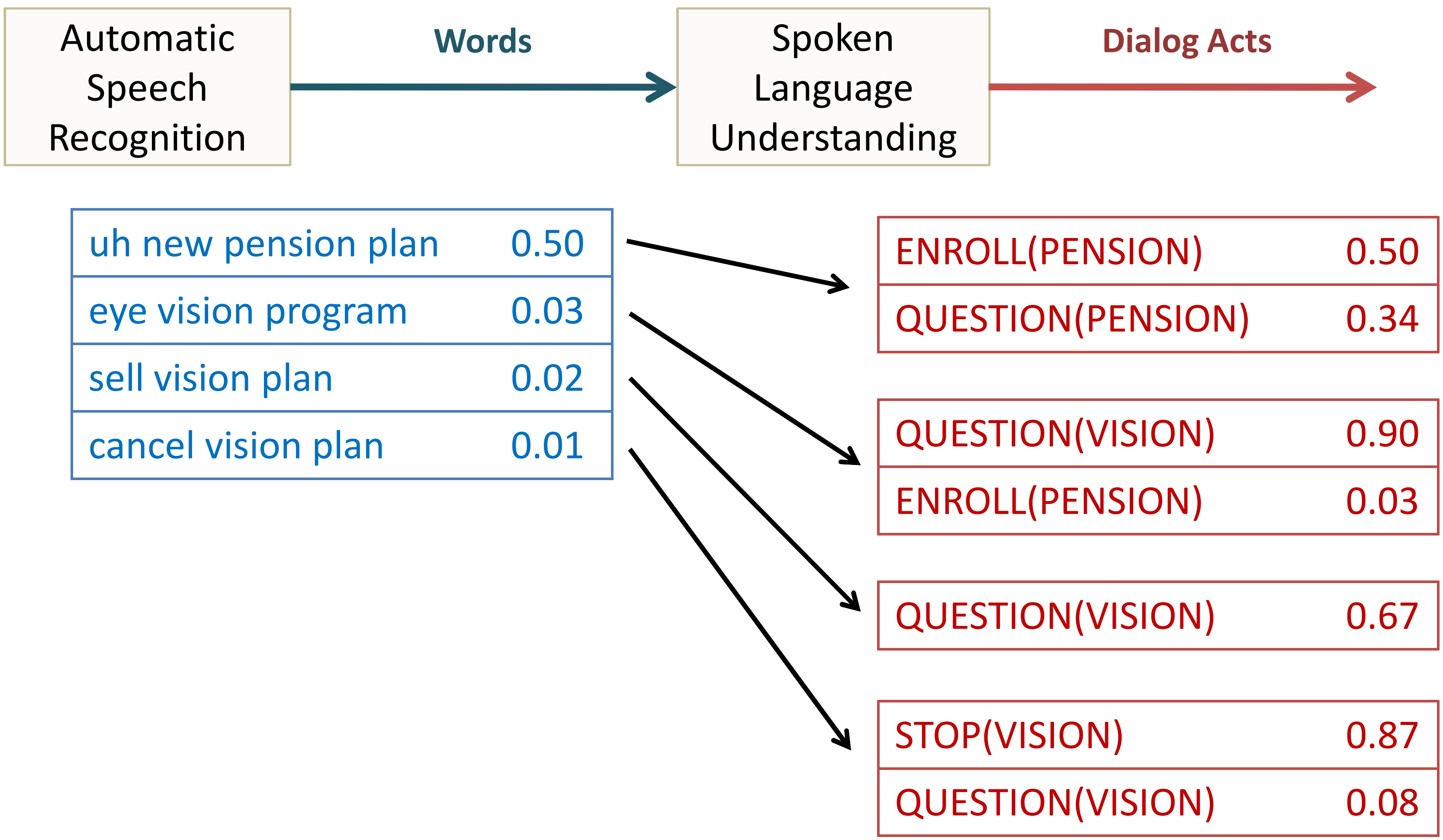


Bohus, D., Horvitz, E. (2009). Models for Multiparty Engagement in Open-World Dialog, in Proceedings of SIGdial'09, London, UK

Outline

- Two key challenges for building dialogue systems
- How dialogue systems are built today
- Dialogue system as planning under uncertainty
- "Growing up" to real-world systems
- Prospects for commercial use
- Thoughts about the future

Speech recognition and spoken language understanding



Why speech recognition is hard: some examples

- Pauses in speech lead to end-pointing problems
- Spontaneous speech contains self-corrections
- "Robot" language (hence examples, "speak naturally")
- Users react to errors, form a "theory-of-mind"
  - > "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>

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%

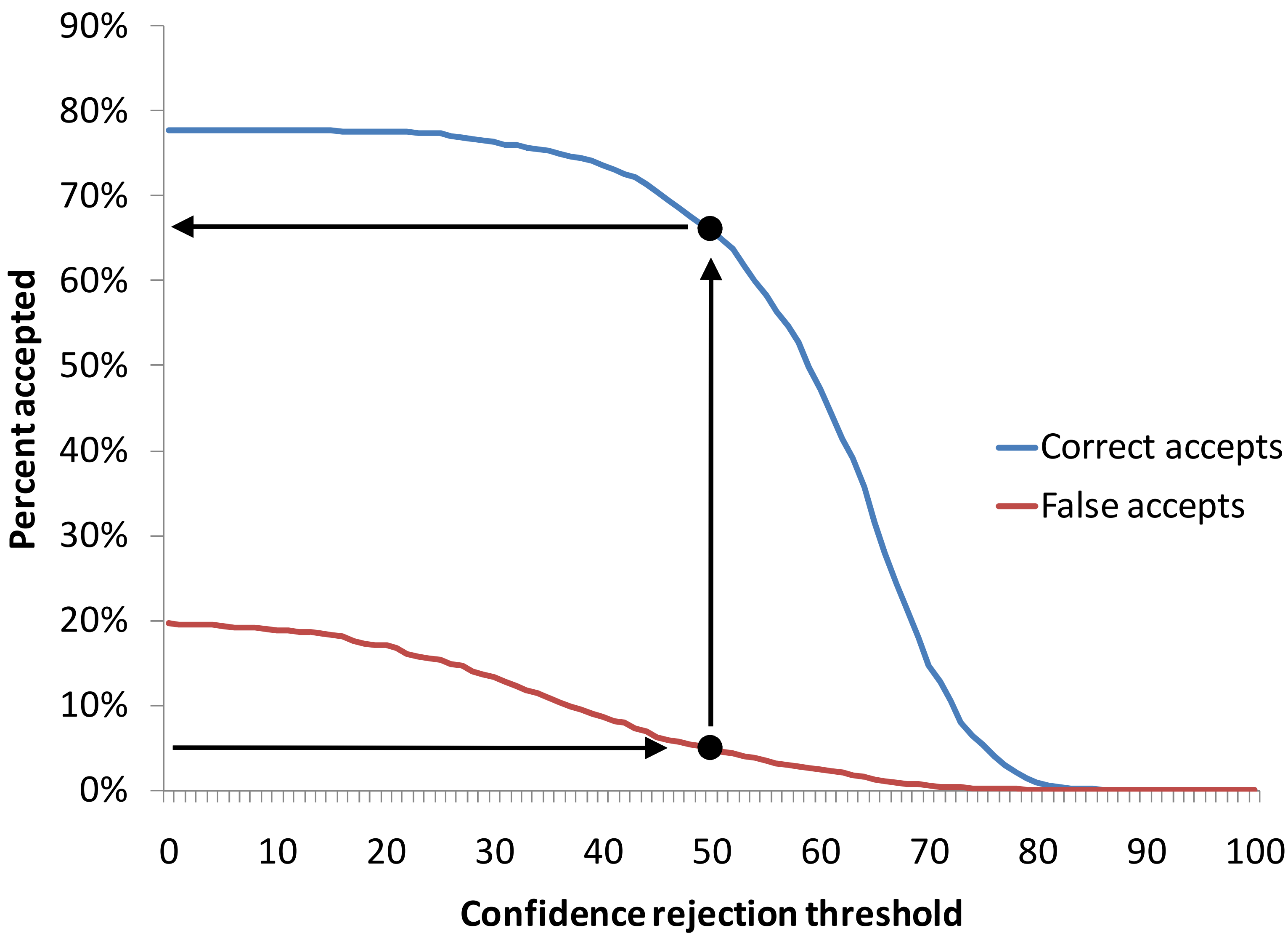
Source: Two different deployed commercial applications running two different speech recognizers

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

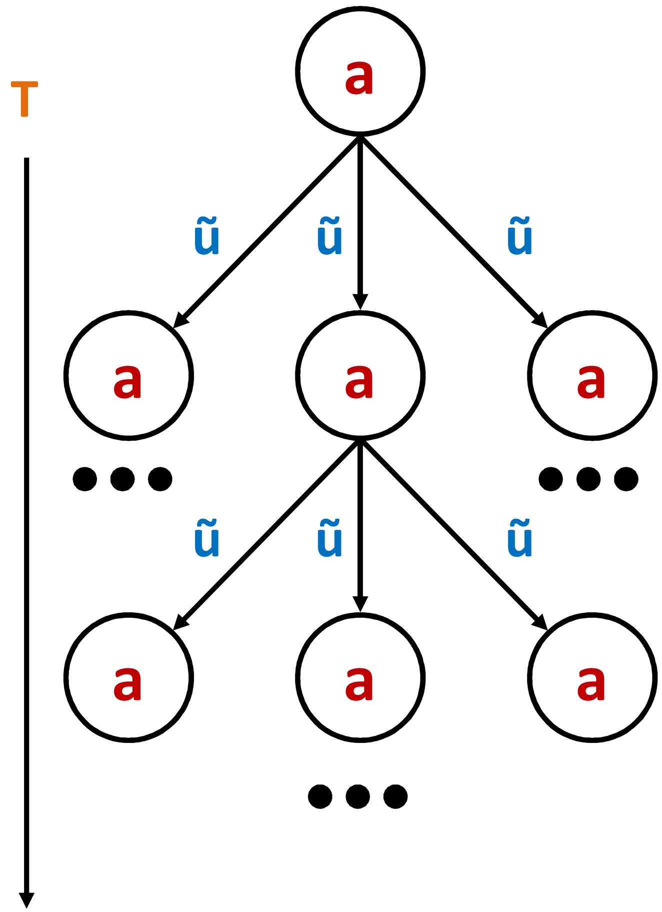
Grammar	Yes/no	City & state	How may I help you?
In-grammar/ in-domain accuracy	99.8%	85.1%	89.5%
% in-grammar/ in-domain	92.3%	91.0%	86.8%
Overall accuracy	92.1%	77.6%	77.7%
Accepted utts (False accepts)	89.6% (1.8%)	60.3% (4.9%)	73.3% (8.3%)

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, ... }



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

Typical system:

$A = 10^{10}$

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

$T = 10$

Curse of history

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

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

Dialog state	$s_t \approx (\tilde{u}_0,a_1,\tilde{u}_1,a_2,\tilde{u}_2,a_3,\tilde{u}_3,...,a_t,\tilde{u}_t)$
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

How dialog systems are built today

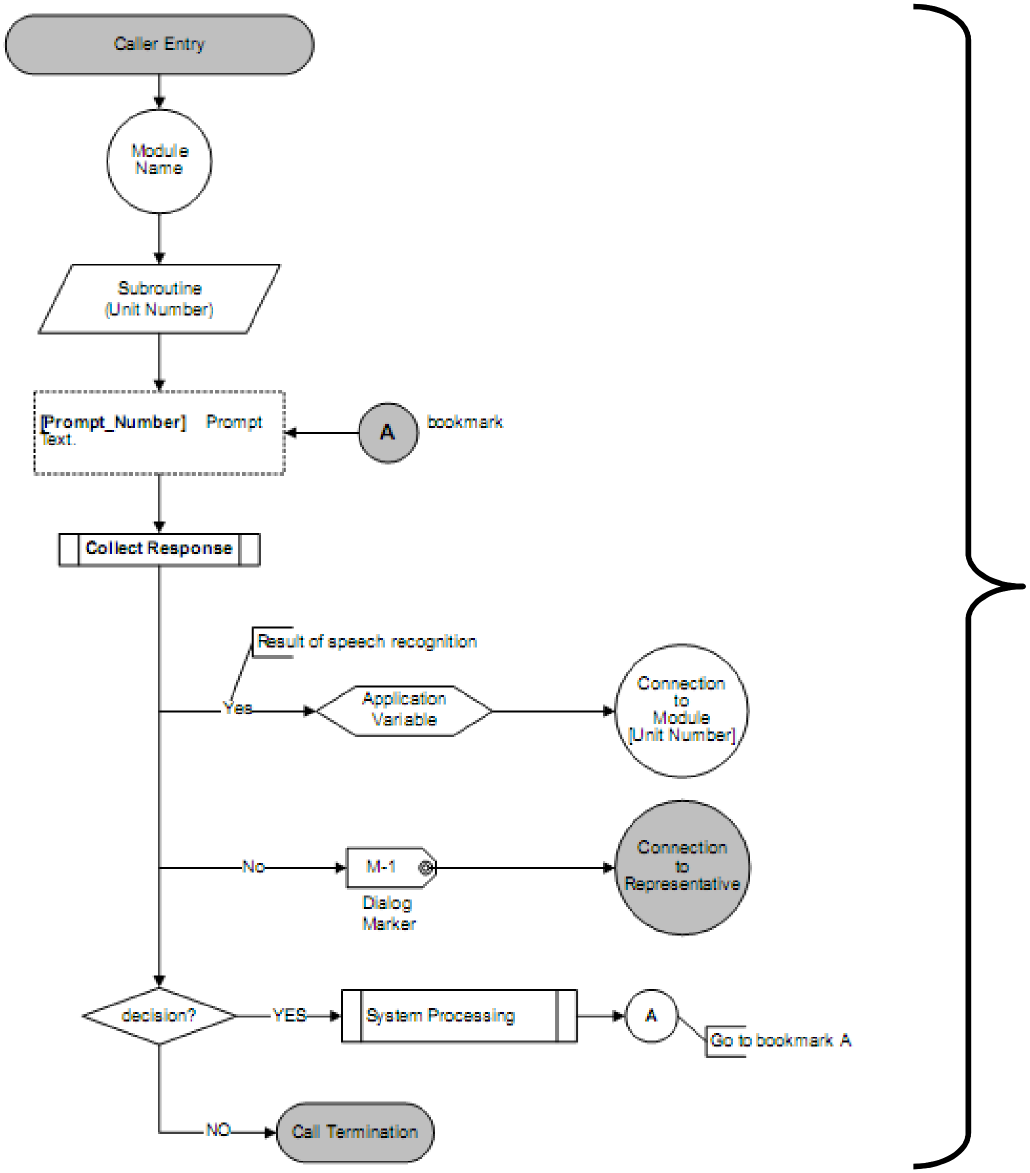
Spoken dialog systems as an application of planning under uncertainty

How dialog systems are built today

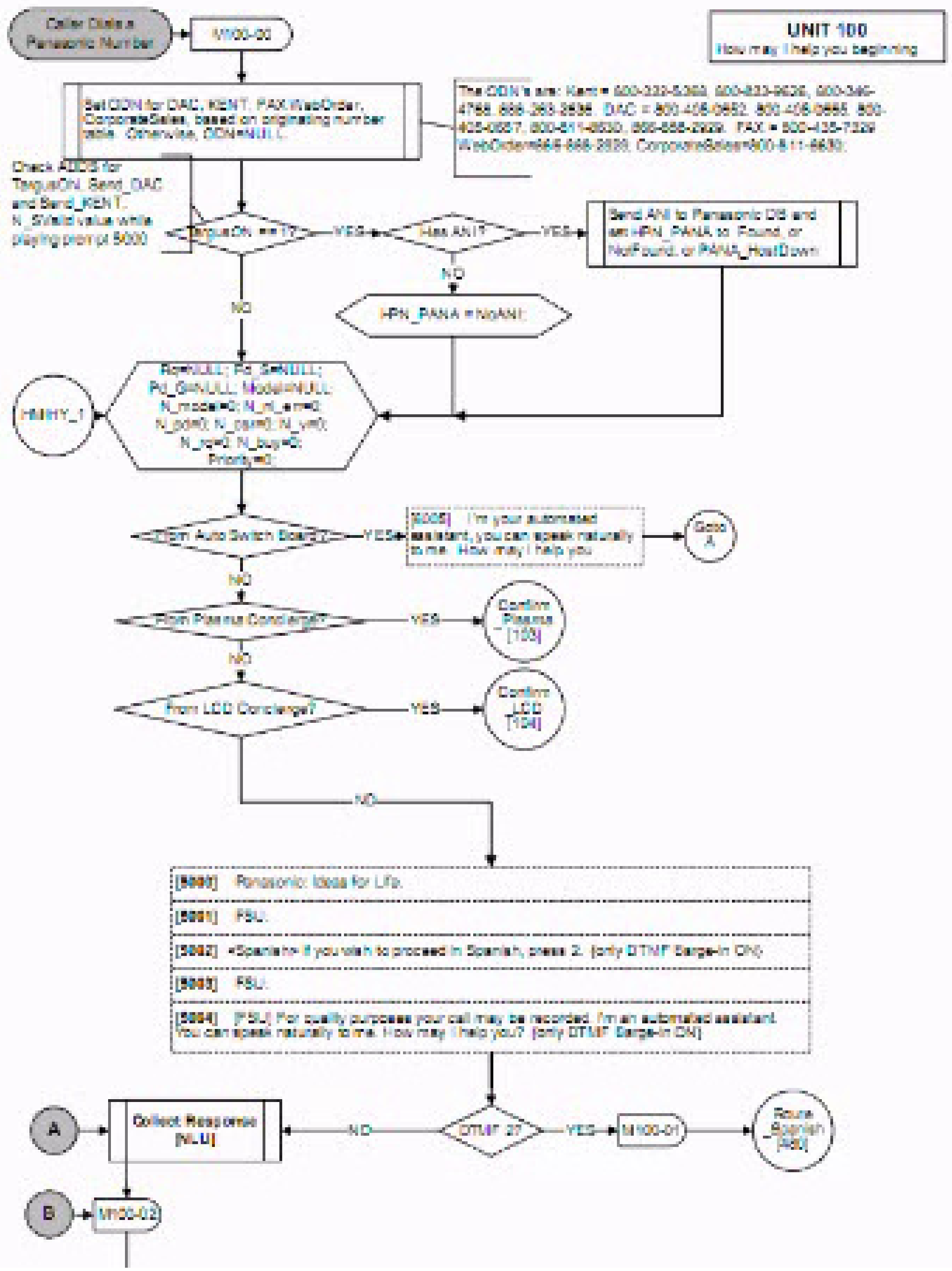
s =

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


UNIT 100
How may I help you beginning
Caller Dials a Panasonic Number
Set DDN for DAC, KENT, FAX WebOrder, CorporateSales, based on originating number table. Otherwise, DDN=NULL.
Check ADDS for TargusON, Send_DAC and Send_KENT, N_SValid value while playing prompt 5000
TargusON == 1?
YES
Has ANI?
YES
Send ANI to Panasonic DB and set HPN_PANA to Found, or NotFound, or PANA_HostDown
NO
HPN_PANA = NoANI
NO
Priority=0;
From Auto Switch Board?
YES
I'm your automated assistant, you can speak naturally to me. How may I help you
Goto A
NO
From Plasma Concierge?
YES
Confirm Plasma [103]
NO
From LCD Concierge?
YES
Confirm LCD [104]
NO
Panasonic: Ideas for Life.
FSU.
<Spanish> If you wish to proceed in Spanish, press 2. (only DTMF Barge-In ON)
FSU.
[FSU] For quality purposes your call may be recorded. I'm an automated assistant. You can speak naturally to me. How may I help you? (only DTMF Barge-In ON)
A
Collect Response [NLU]
NO
DTMF 2?
YES
Route Spanish
B

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 (eg N-Best list) is discarded
5. May interact with millions of users, yet will never learn/improve from that experience

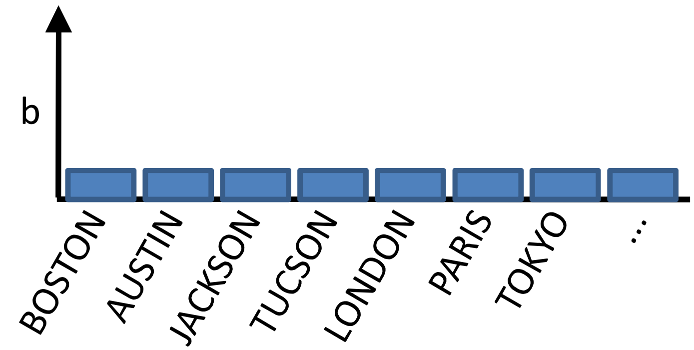
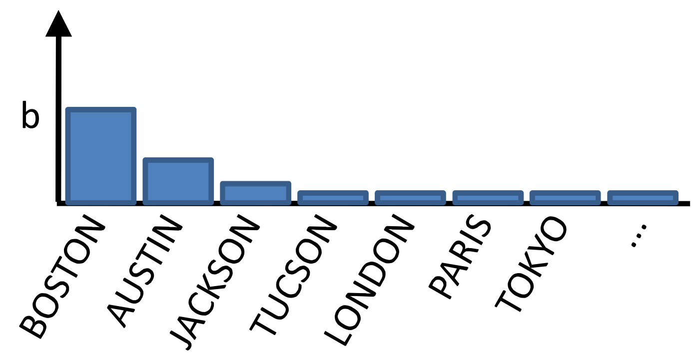
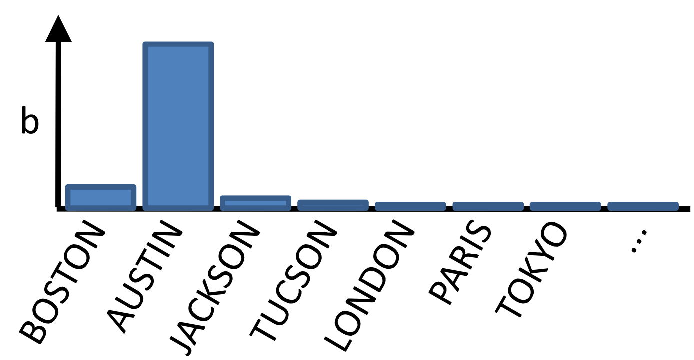
Treating dialogue systems as planning under uncertainty

SECOND

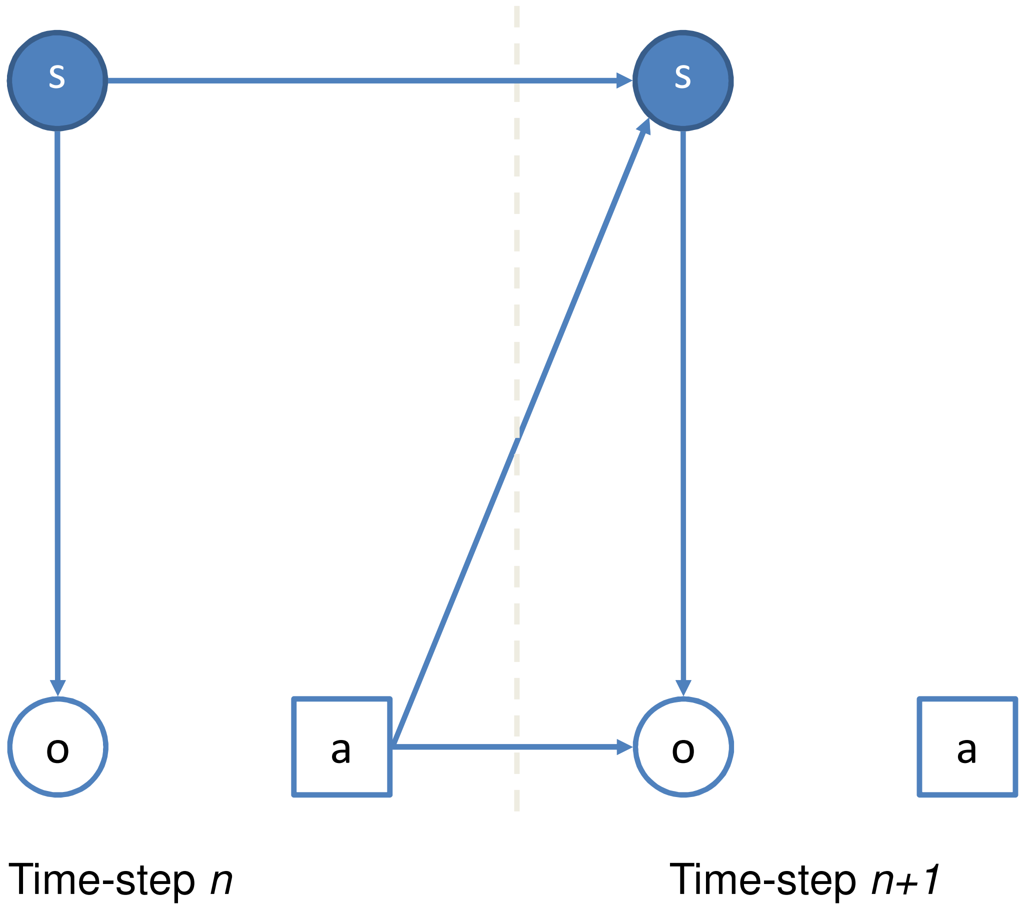
FIRST

Spoken dialog systems as an application of planning under uncertainty

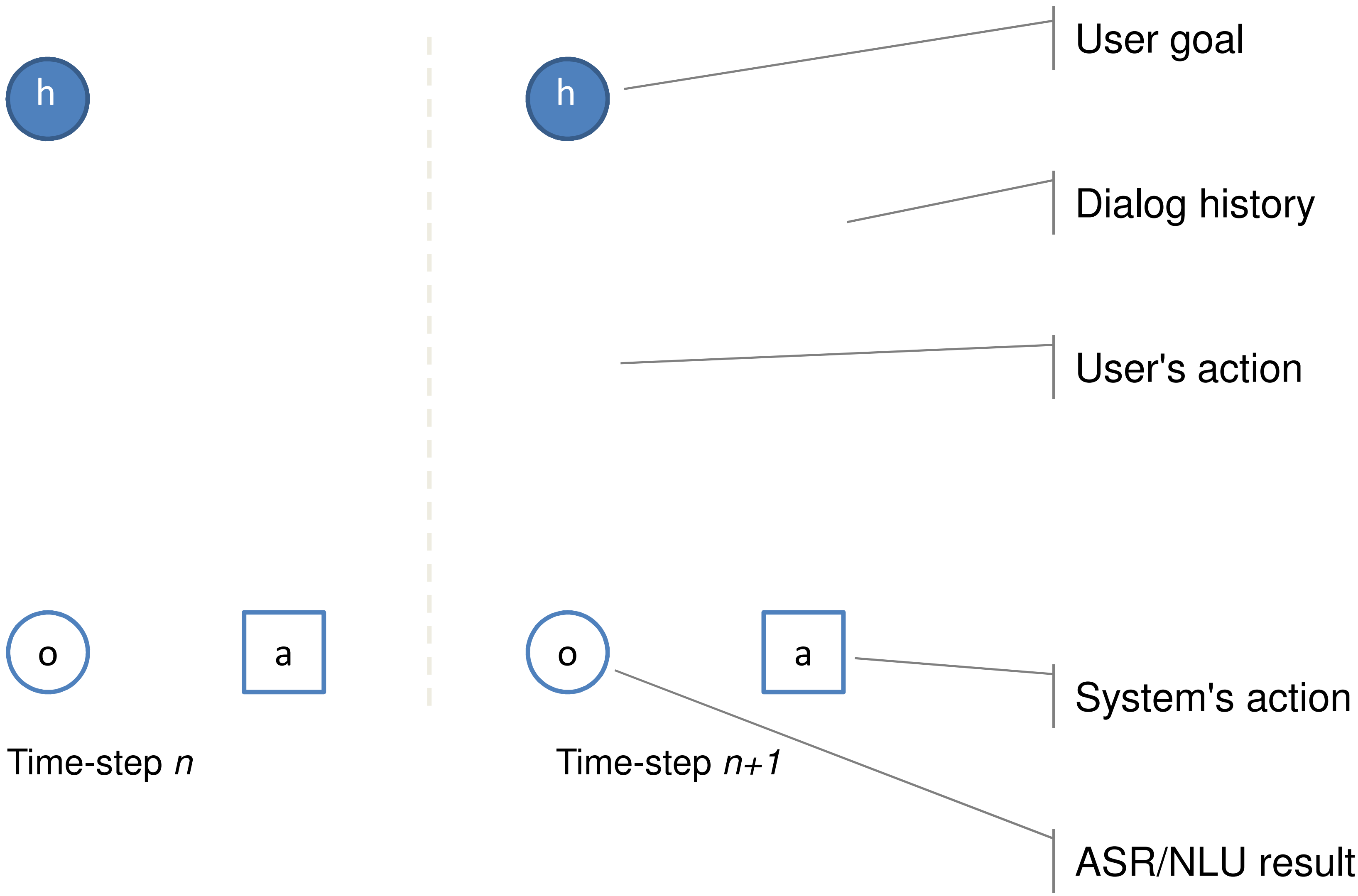
Idea: track a distribution over dialogue states

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	

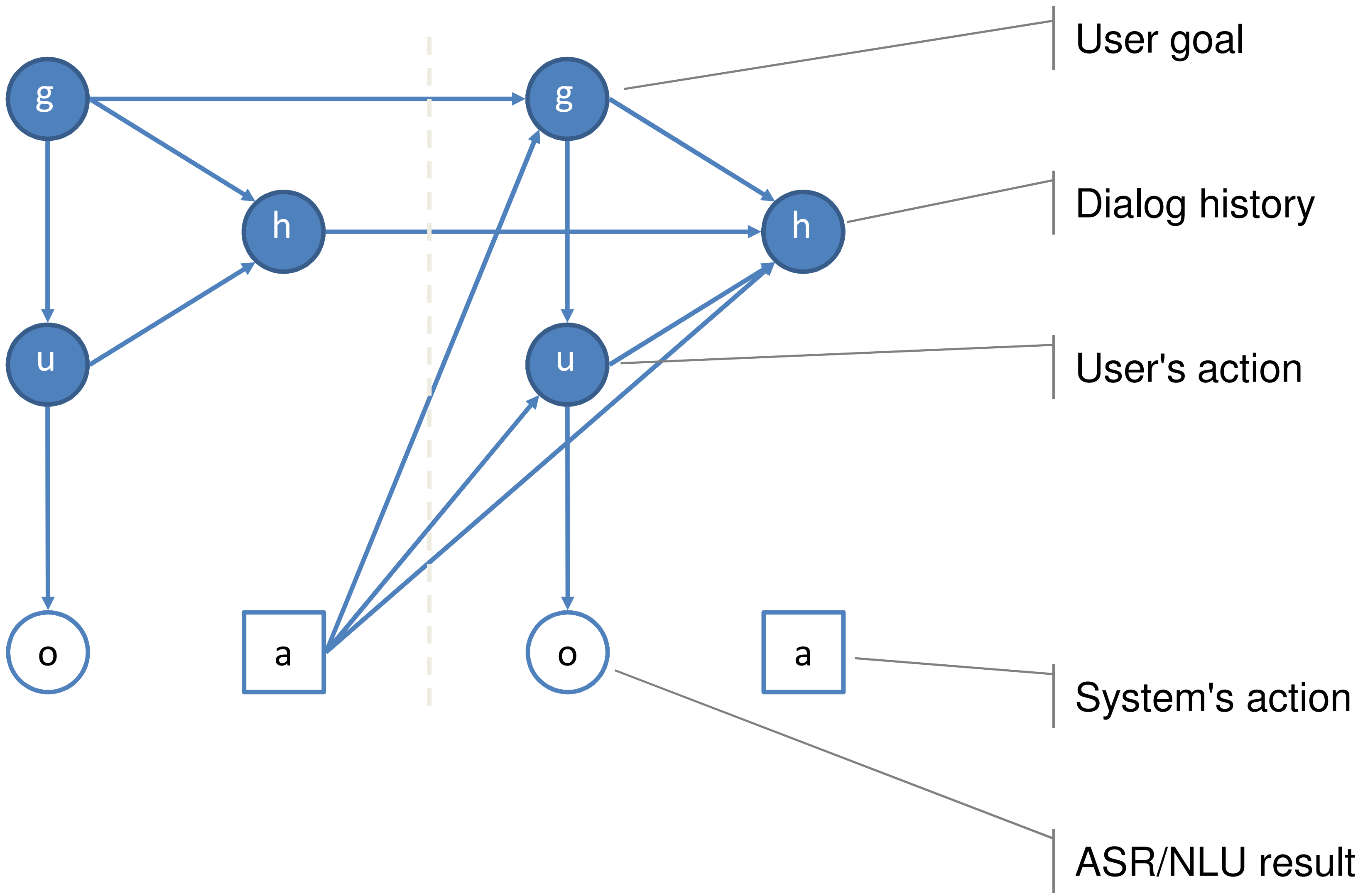
Tracking a distribution over dialogue states



Tracking a distribution over dialogue states



Tracking a distribution over dialogue states



Distribution update equation

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

Problem: Updating belief in real-time

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

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

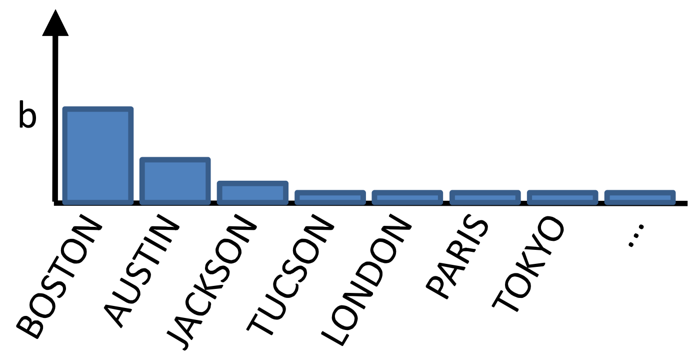
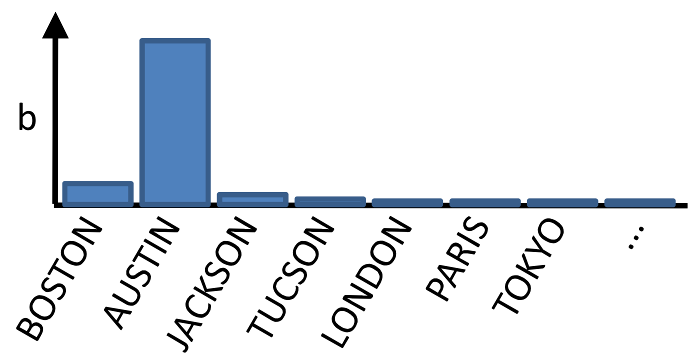
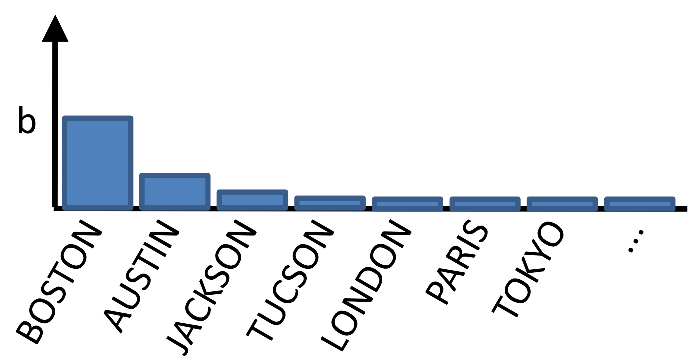
$O(10^{24})$ impossible in real time !

2 methods for efficient belief monitoring

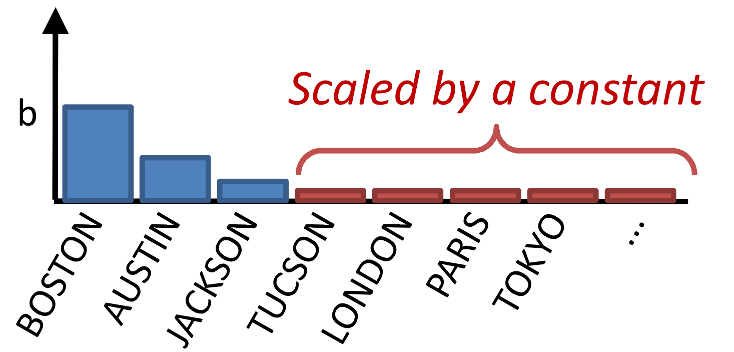
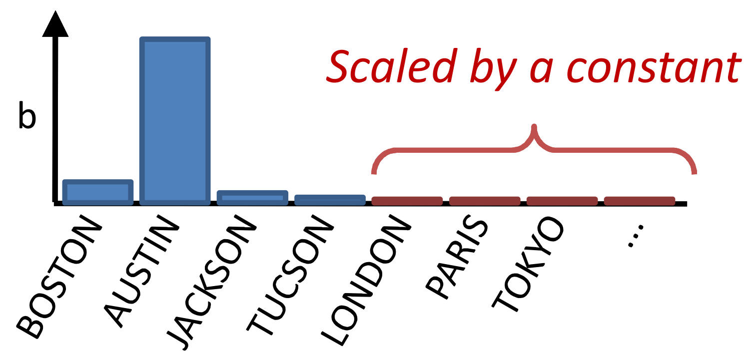
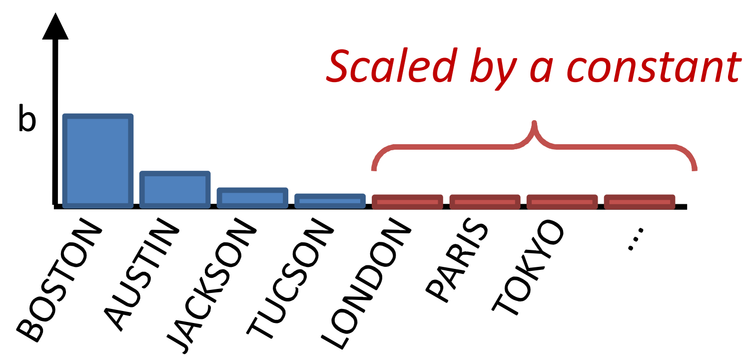
1. **M-Best**: Constrain aspects of the model such that un-observed goals can be tracked en-masse

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 equation

$$b'(g',h') = \eta \cdot \sum_{u'} P(o' \mid u',a) \sum_{h} P(u' \mid g',h,a) P(h' \mid u',g',h,a) \underbrace{\sum_{g} P(g' \mid g,a)}_{\text{goal is fixed; goal model drops out}} b(g,h)$$

$$\underbrace{b(q',h')}_{\text{new belief state}} = \underset{\text{normalizing constant}}{\eta} \cdot \sum_{u'} \underbrace{P(o' \mid u',a)}_{\text{ASR model}} \sum_{h} \underbrace{P(u' \mid q',h,a)}_{\text{user action model}} \underbrace{P(h' \mid u',q',h,a)}_{\text{dialog history model}} \underbrace{\frac{b_0(q')}{b_0(q)}}_{\text{belief refinement}} \underbrace{b(q,h)}_{\text{old belief state}}$$

Partition update example (maxPartitions = 3)

"First name?"

JASON	~0.6

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

Belief **Prior**

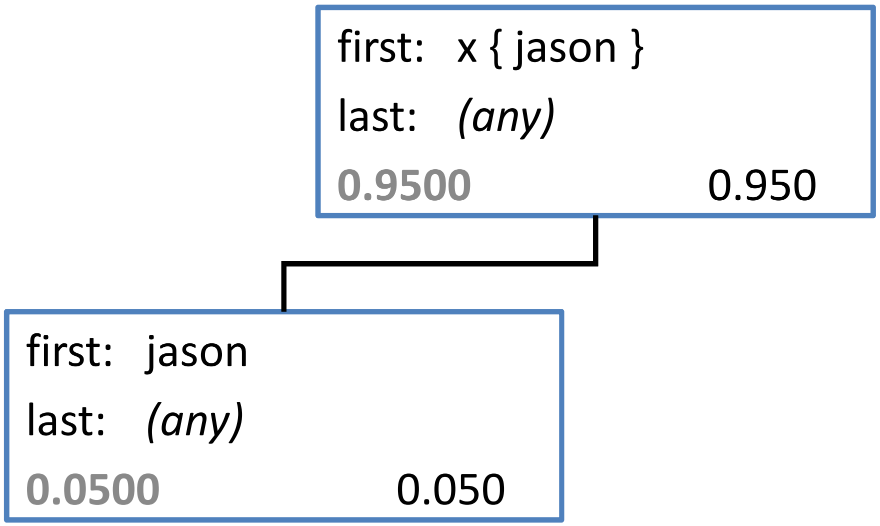
Database of priors

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

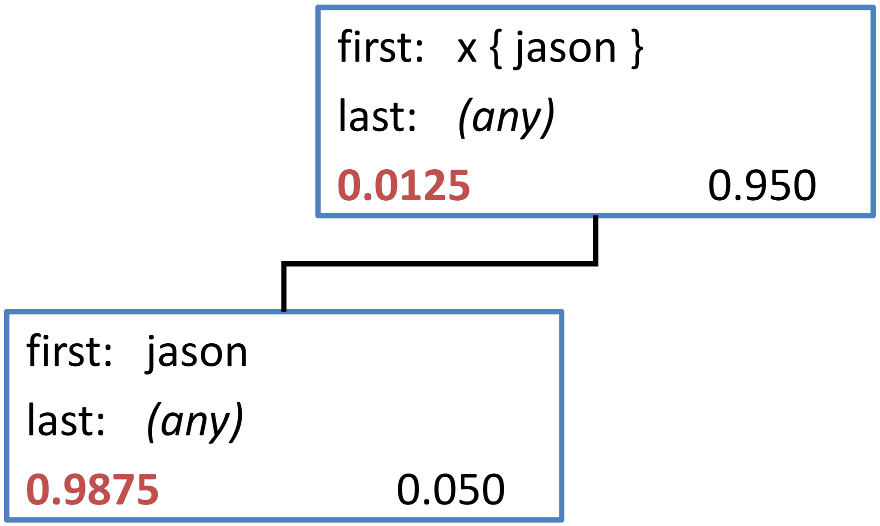


First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010

Partition update example (maxPartitions = 3)

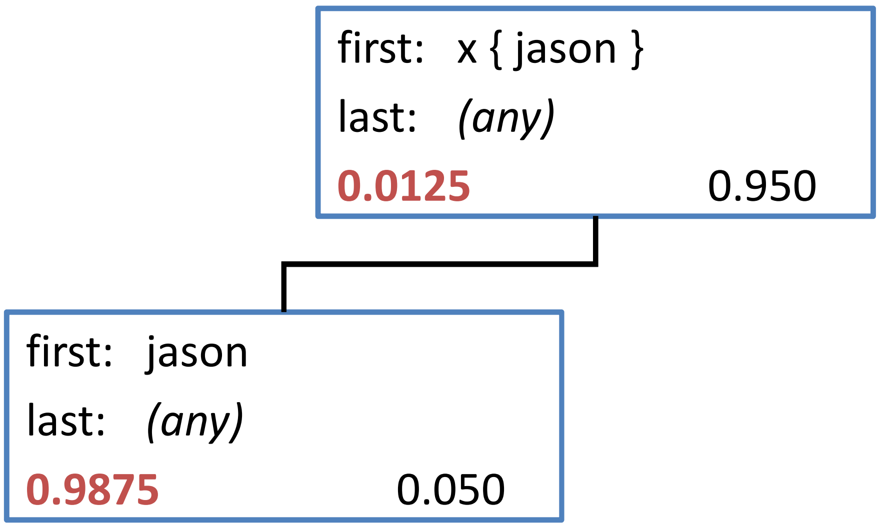
"First name?"

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

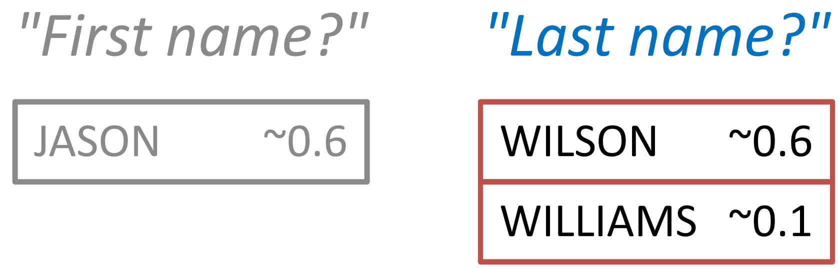
"Last name?"

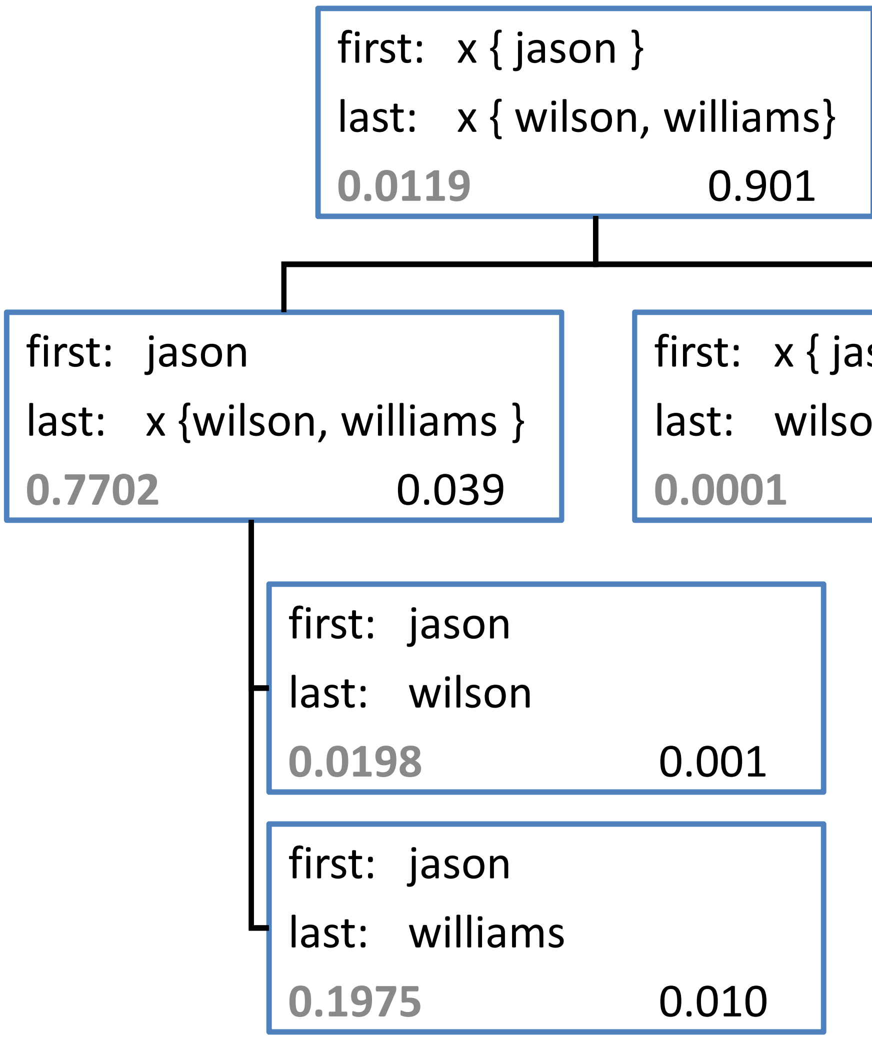
WILSON ~0.6

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

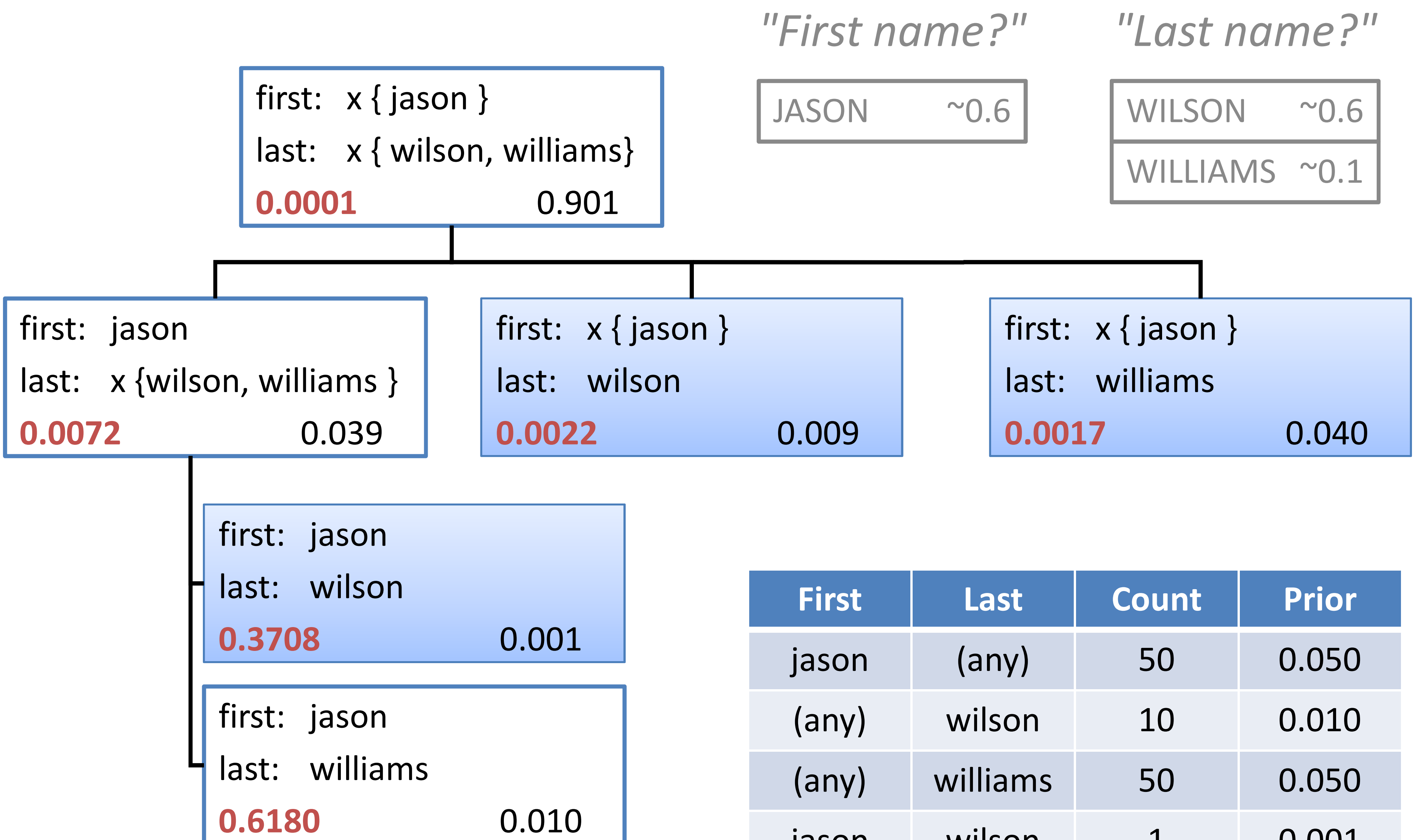
first: x { jason }
last: williams
0.0017 0.040

first: jason
last: wilson
0.3708 0.001

first: jason
last: williams
0.6180 0.010

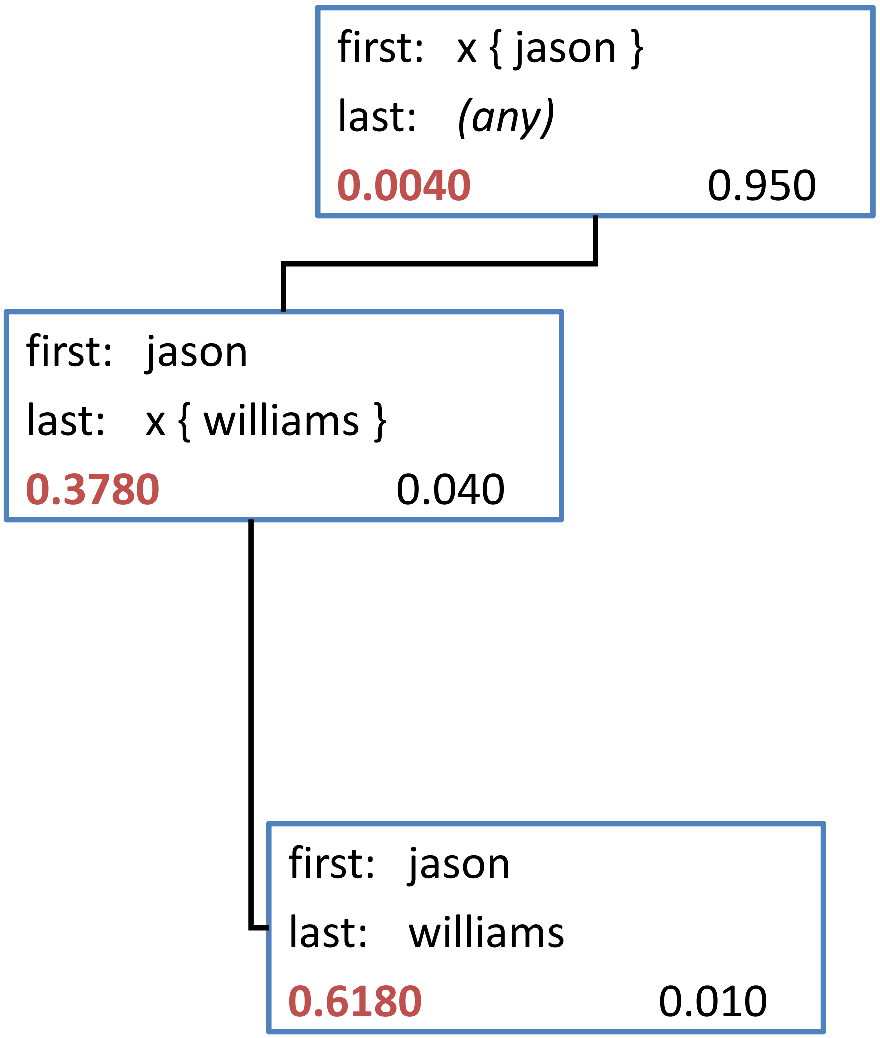
First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010

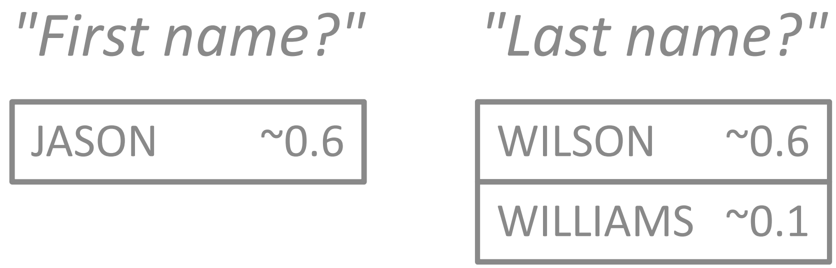
Partition update example (maxPartitions = 3)



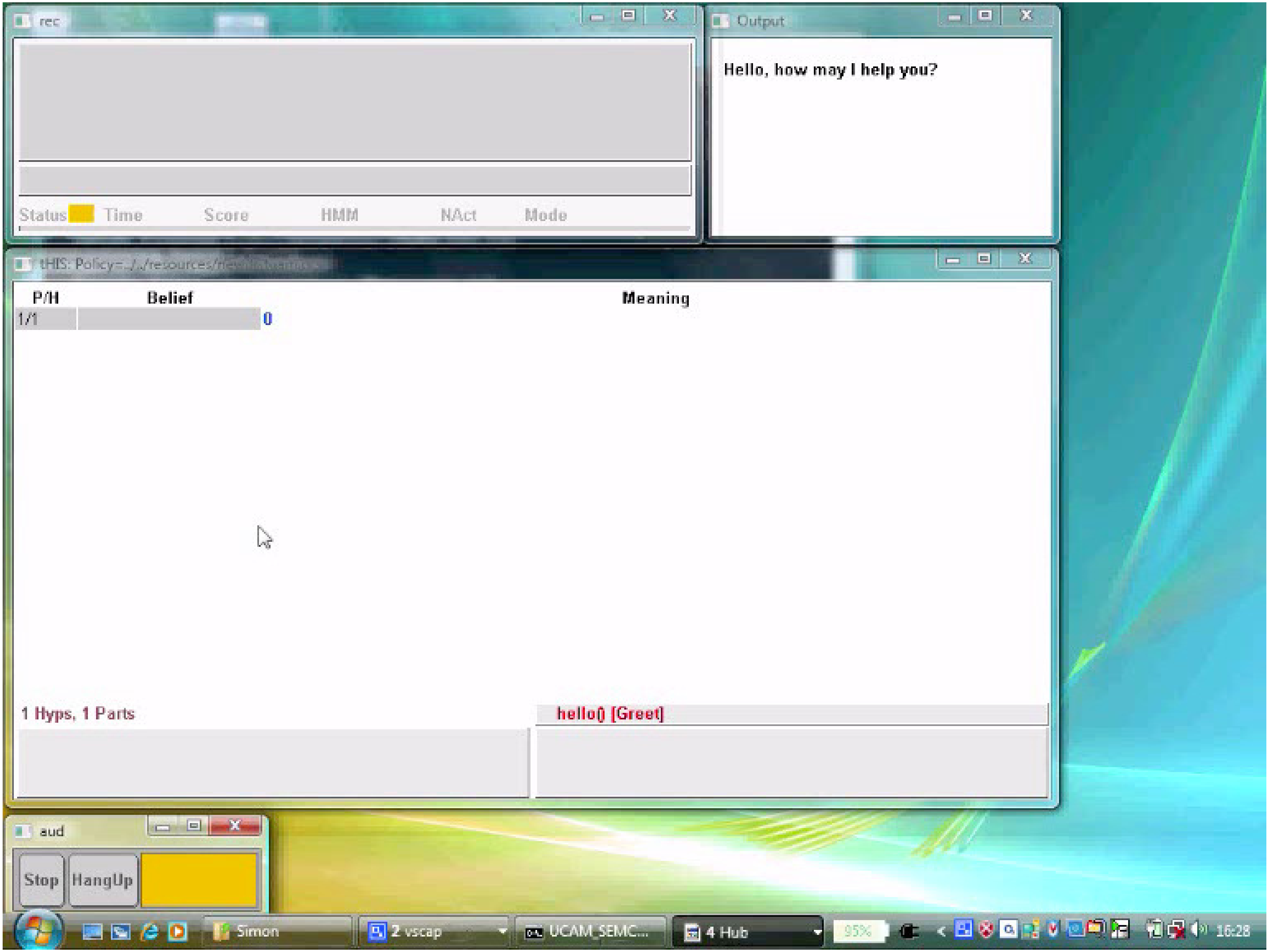
First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010

Partition update example (maxPartitions = 3)

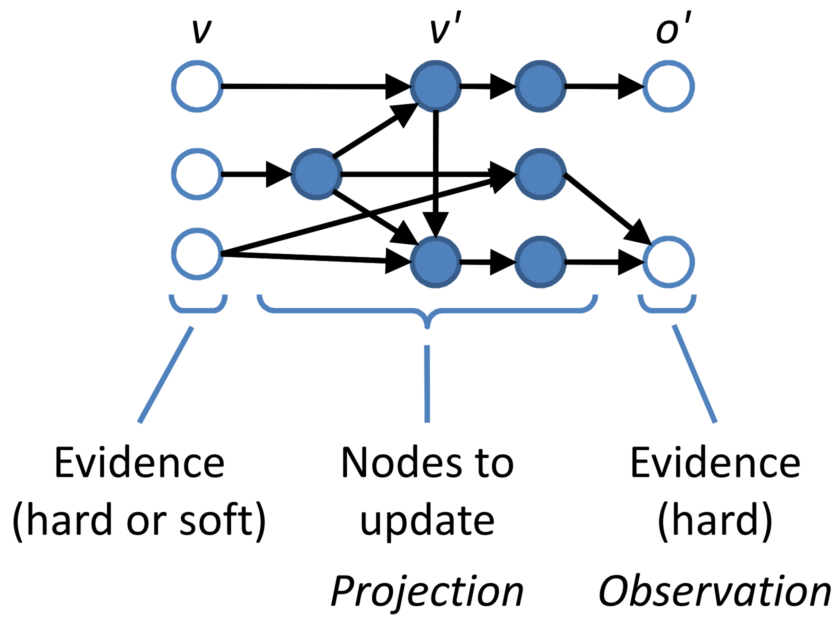




First	Last	Count	Prior
jason	(any)	50	0.050
(any)	wilson	10	0.010
(any)	williams	50	0.050
jason	wilson	1	0.001
jason	williams	10	0.010


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

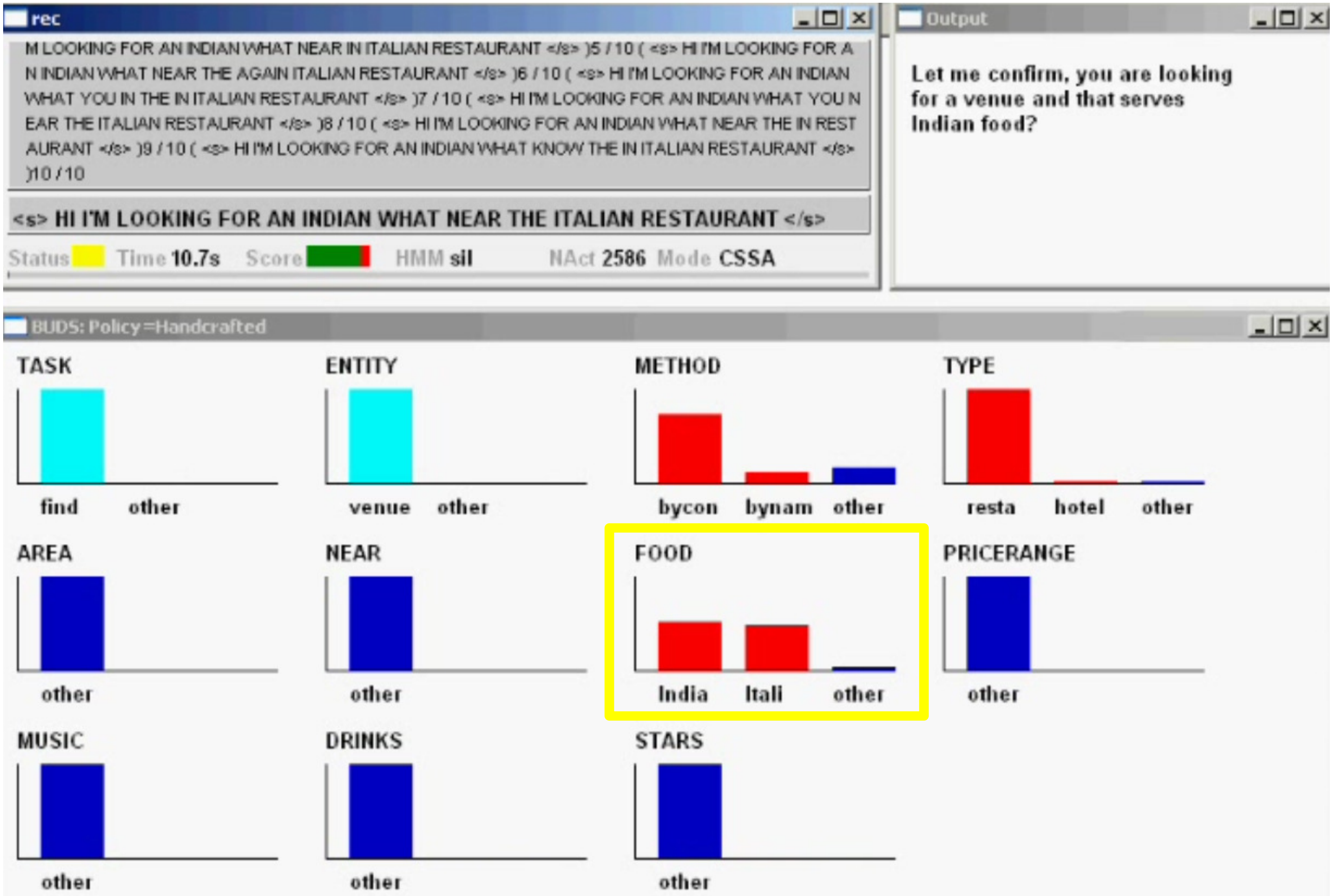
Network-based approaches



Idea: Apply general purpose Bayes Network inference techniques

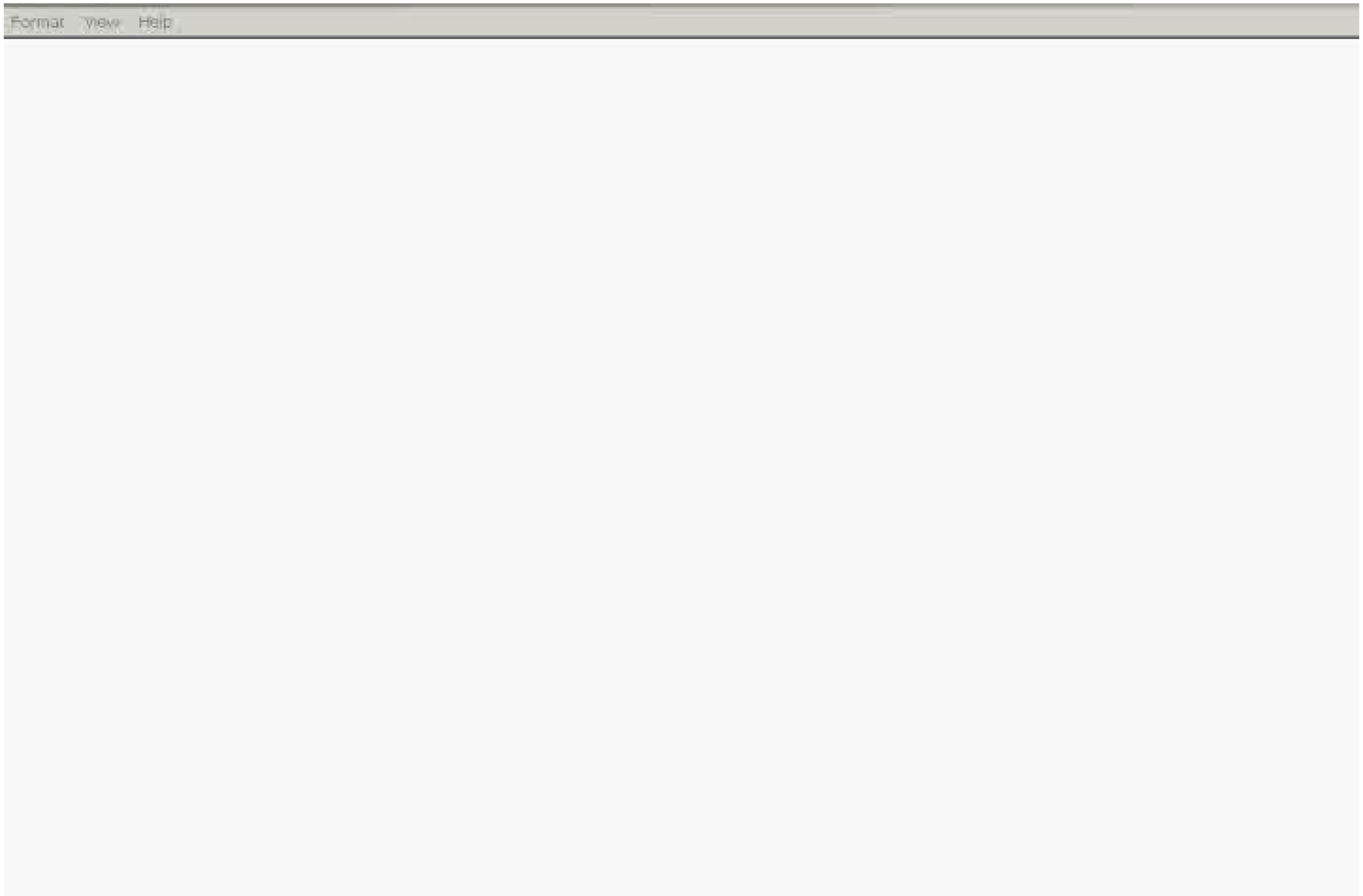
Approximate inference can be much faster than exact

Examples: loopy belief propagation and particle filters



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

Tracking multiple dialogue states: example



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

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, 24(4): 562-588.

Treating dialogue systems as planning under uncertainty

SECOND

FIRST

Spoken dialog systems as an application of planning under uncertainty

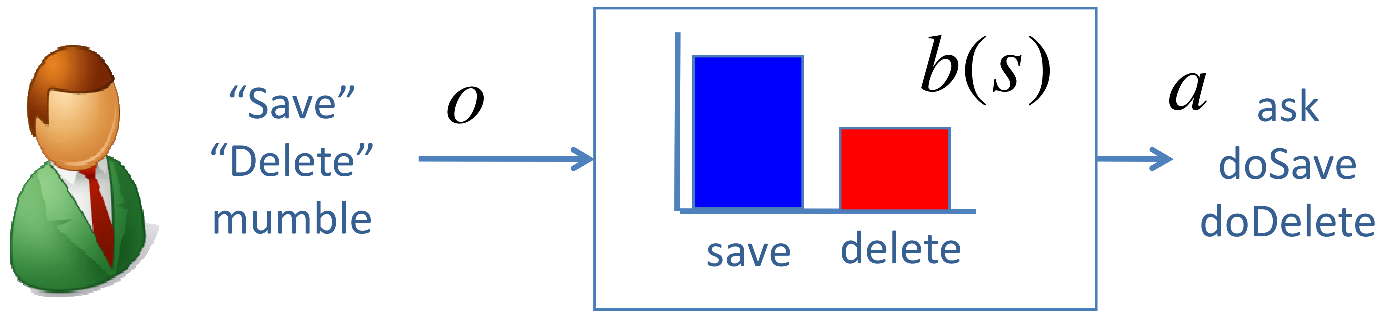
Idea: choose actions via a reward function

Instead of specifying F(s) = a directly, the human designer **specifies a reward function R(s,a)**

An **optimization process** chooses F(s) = a in order to **maximize the sum of expected rewards** until the end of the dialogue

- Allows much more detailed plans to be created
- Actions chosen for global optimality (not local)
- Opens the door to improving through experience

A Simple Two State Example (voicemail)



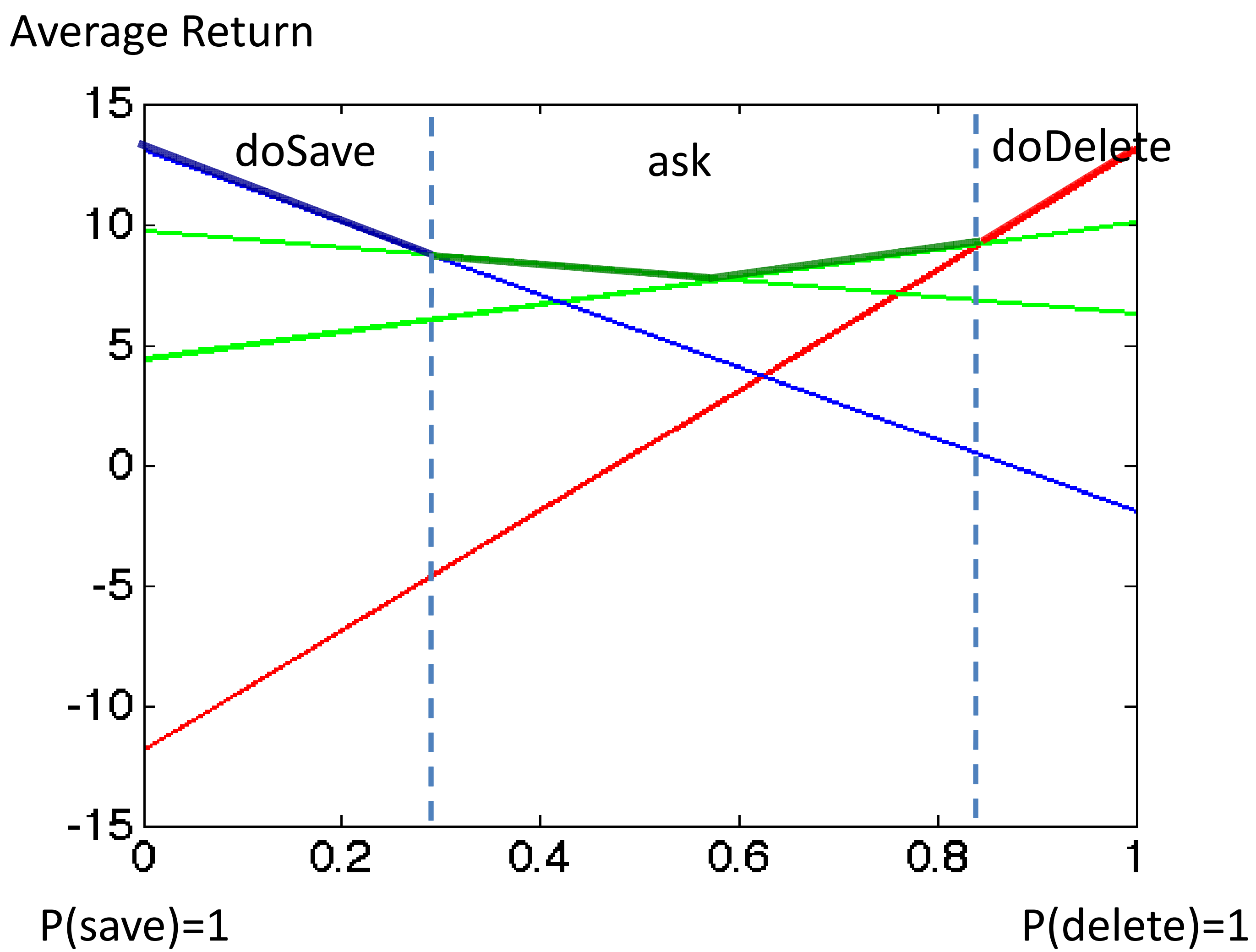
Reward Function

$$R(s,a)$$

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

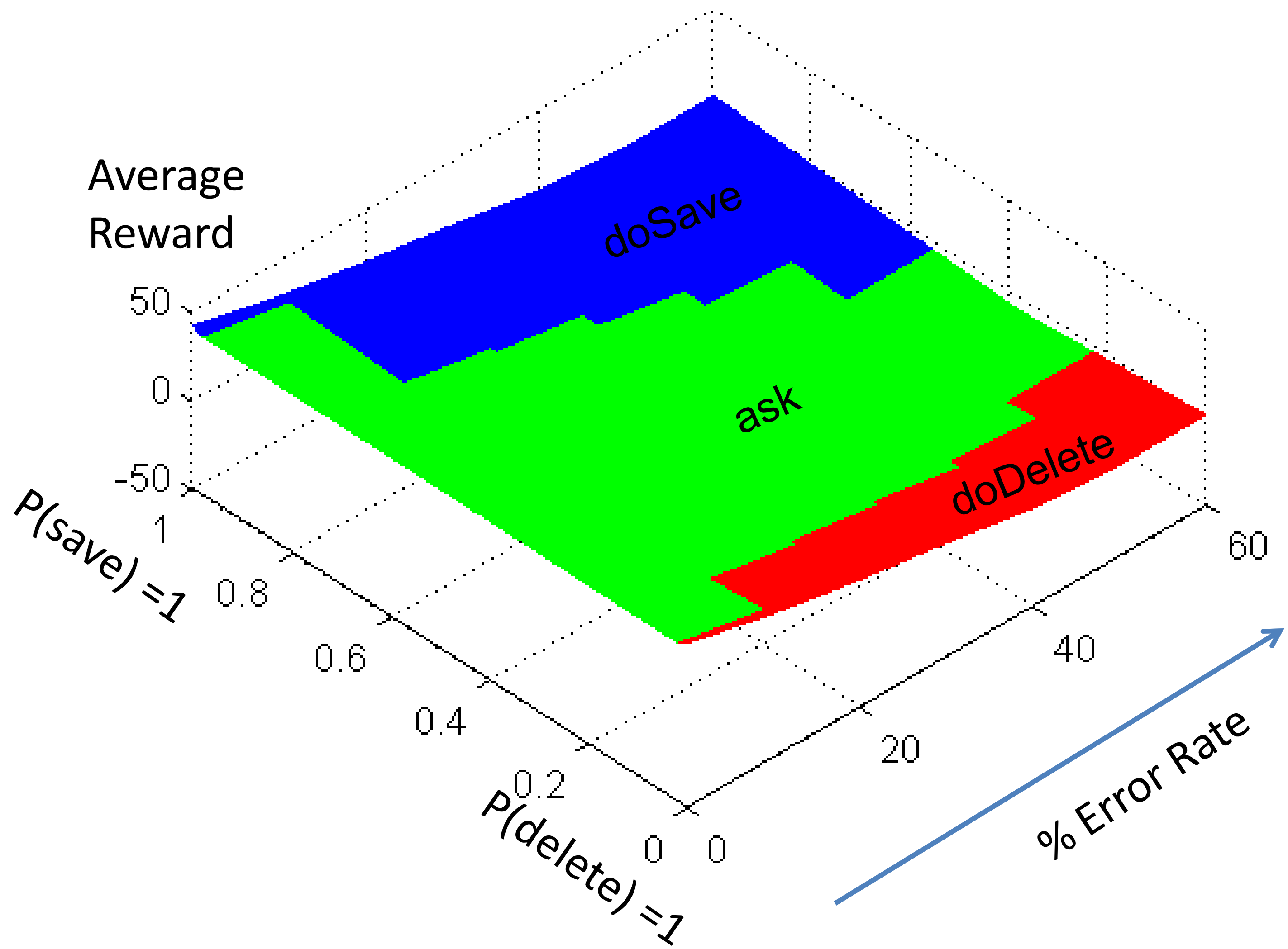
Williams, Young, and Thomson. Statistical approaches to dialogue systems. Tutorial at INTERSPEECH 2009. Brighton, UK.

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

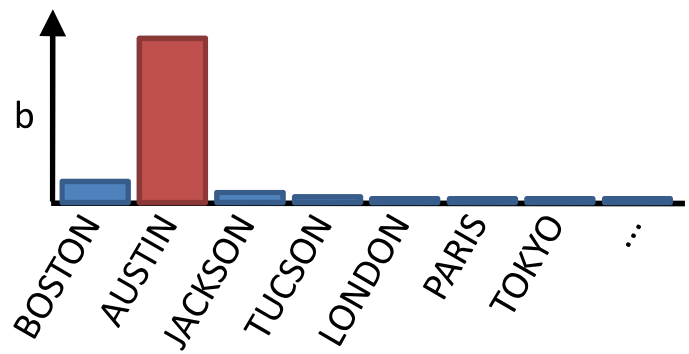


Williams, Young, and Thomson. Statistical approaches to dialogue systems. Tutorial at INTERSPEECH 2009. Brighton, UK.

Planning : issues for real-world use

- Scaling to large state and action spaces
- Expert knowledge & business rules

Scaling up : what are the difficult decisions?



All possible actions:

ask

confirm(boston)

confirm(austin)

confirm(jackson)

...

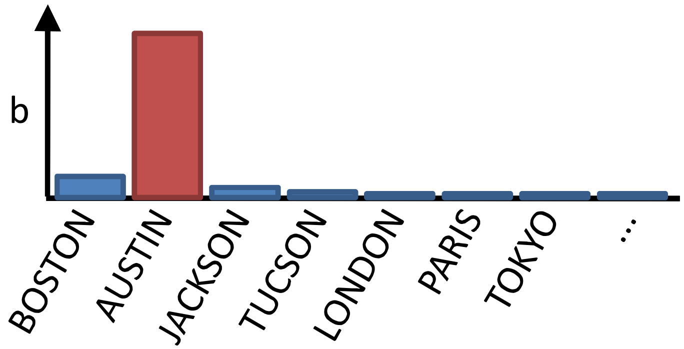
read-weather(boston)

read-weather(ausin)

read-weather(jackson)

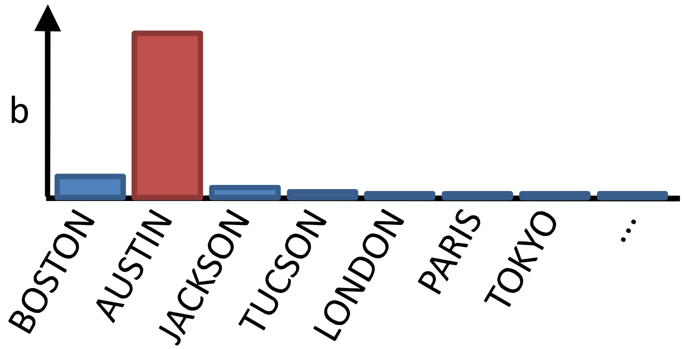
...

Scaling up : 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(ausin)	read-weather(ausin)
read-weather(jackson)	read-weather(jackson)
...	...

Scaling up : what are the difficult decisions?



All possible actions:

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

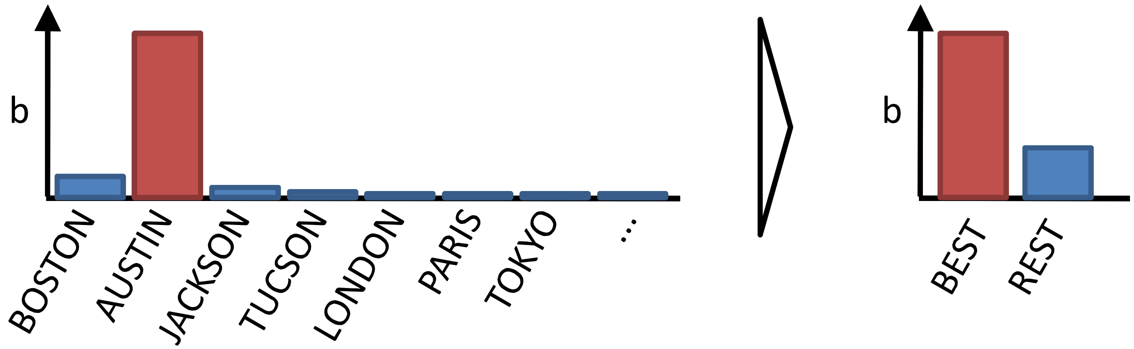
Useful actions:

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

Summary actions:

ask
confirm(best)
read-weather(best)

Scaling up : what are the difficult decisions?



All possible actions:

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

Useful actions:

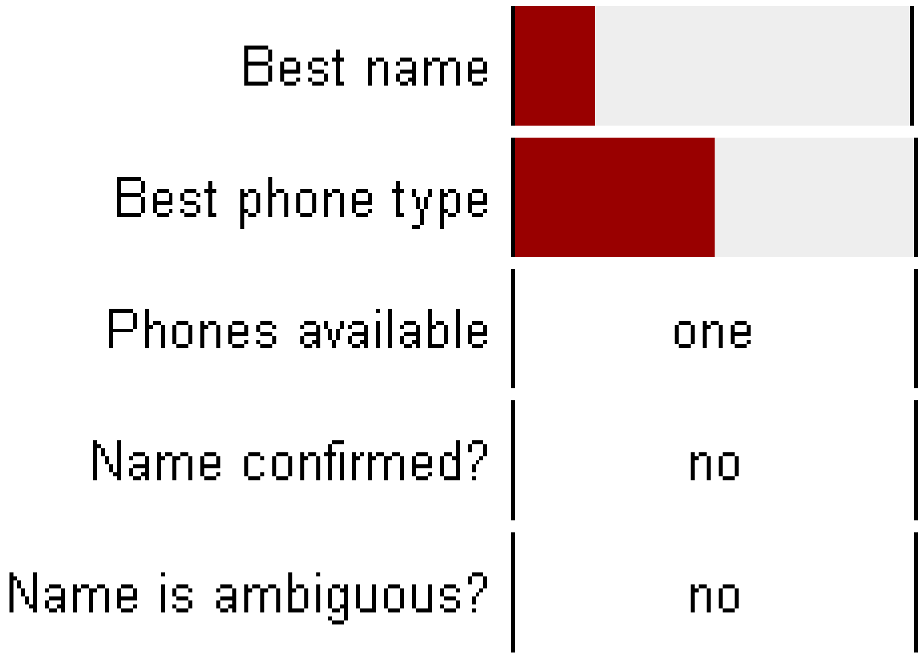
ask
confirm(boston)
confirm(austin)
confirm(jackson)
...
read-weather(boston)
read-weather(ausin)
read-weather(jackson)
...

Summary actions:

ask
confirm(best)
read-weather(best)

Planning : Useful features for difficult decisions

Categorical features are also useful



AskName Sorry, first and last name?
ConfirmName jason wing.

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
- ...

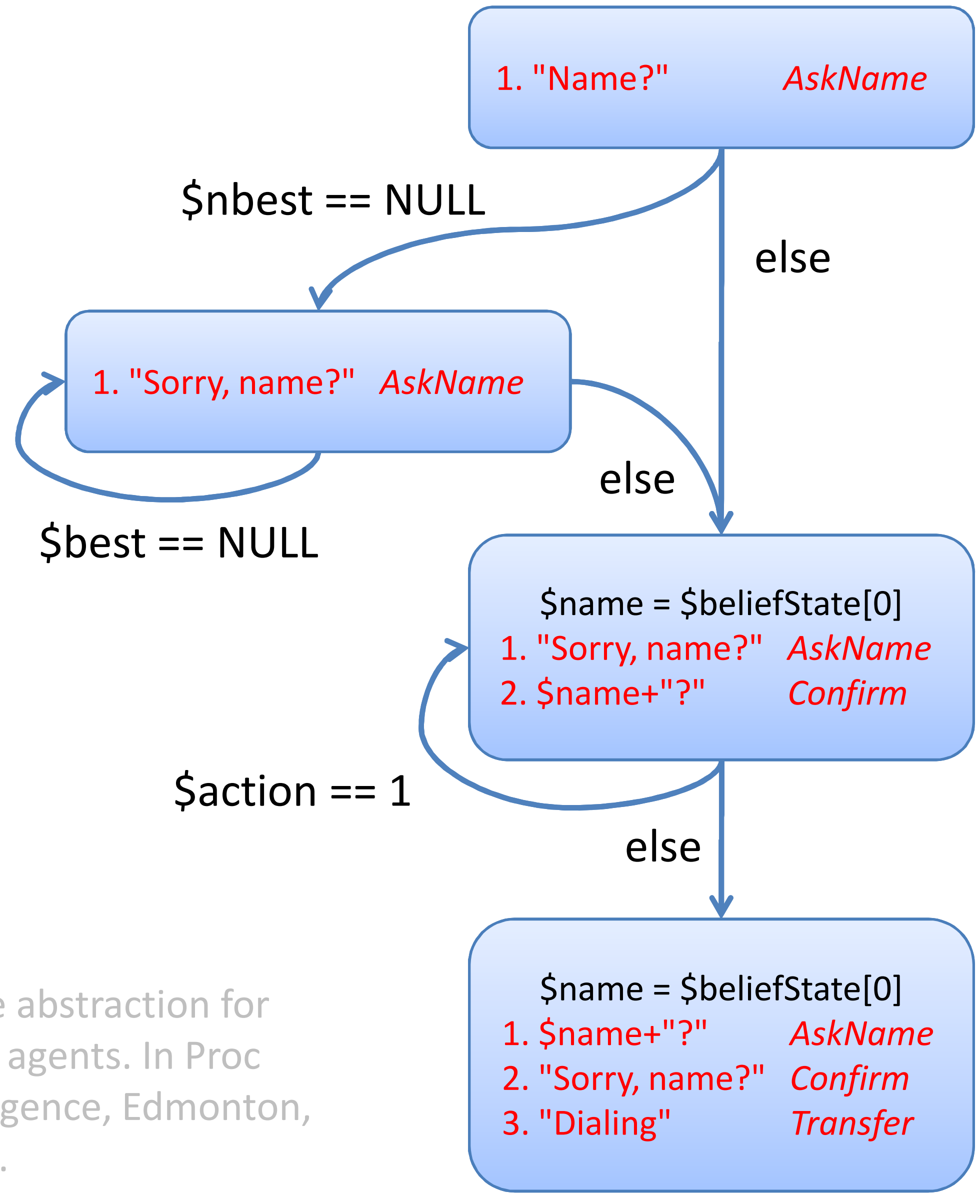
Guarantees about system performance must be made

- Only allow funds transfer after password is entered

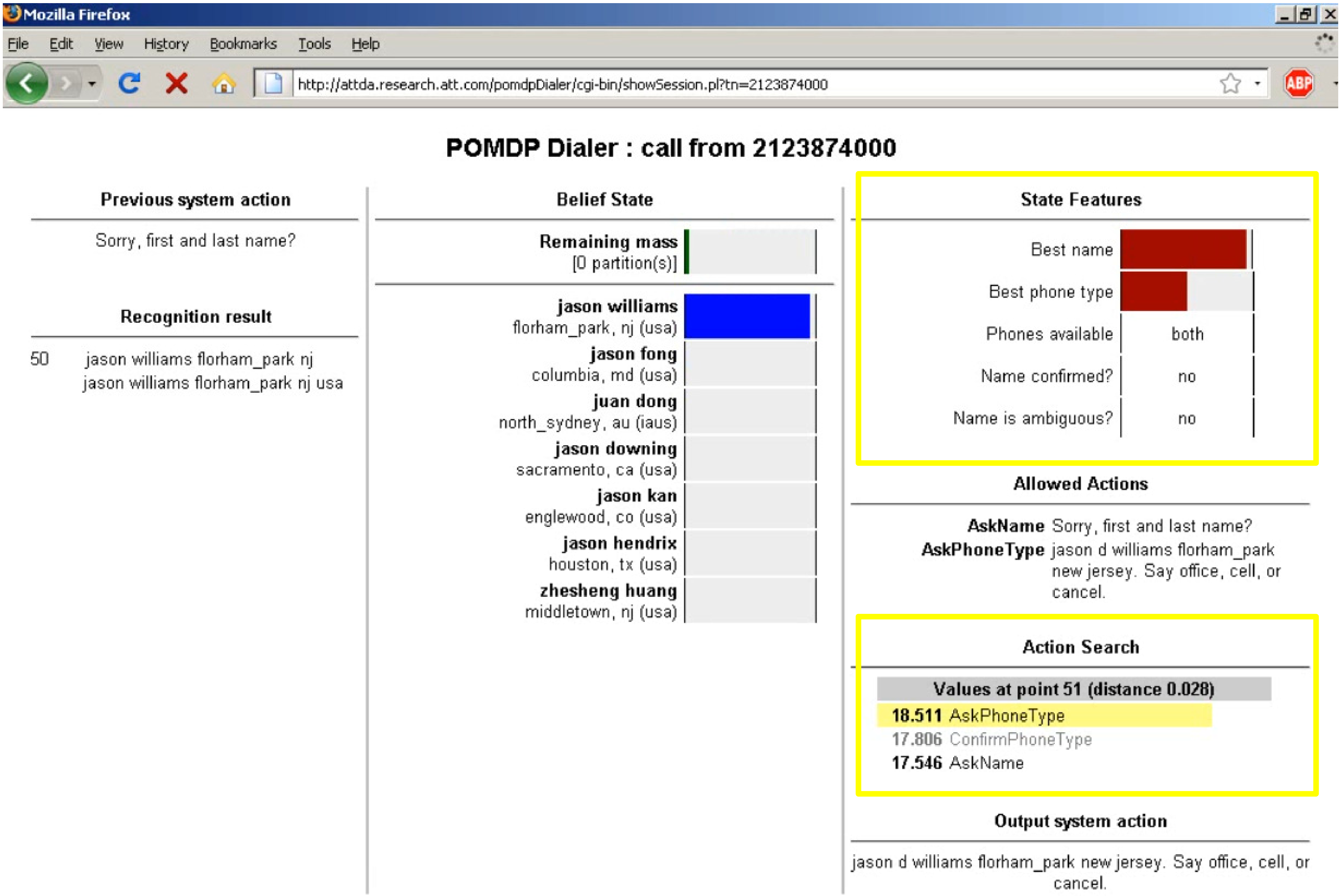
"Tabula rasa" approach to planning seems inappropriate.
Need a way of incorporating constraints and expert knowledge.

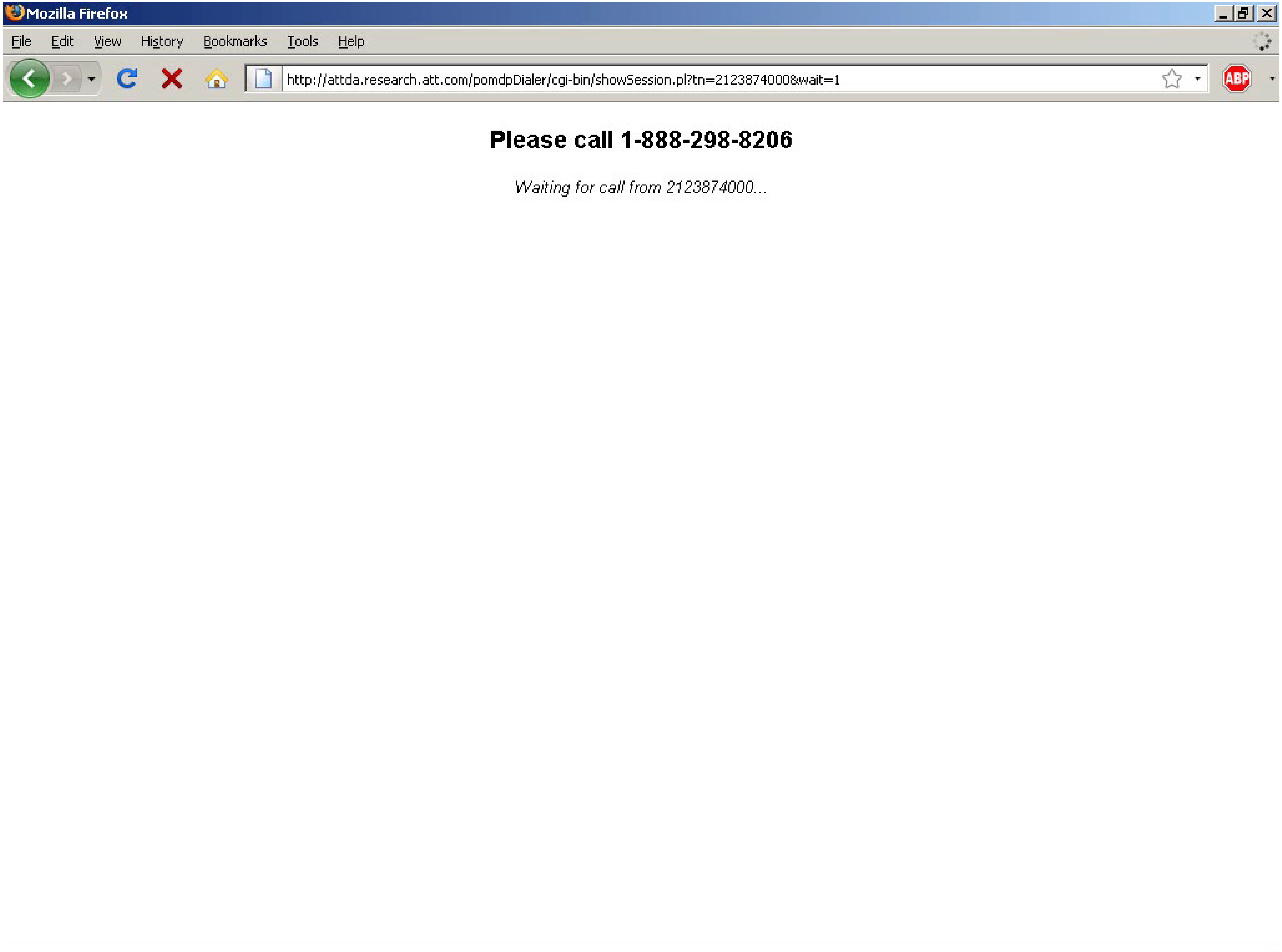
Domain knowledge & business rules

- Create a *partial program* [1] 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



[1] David Andre and Stuart J Russell. State abstraction for programmable reinforcement learning agents. In Proc National conference on Artificial Intelligence, Edmonton, Alberta, Canada, pages 119–125, 2002.


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
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 (2010). "Bayesian update of dialogue state: A POMDP framework for spoken dialogue systems." Computer Speech and Language, 24(4): 562-588.

[6] Heriberto Cuayáhuitl, Steve Renals, Oliver Lemon, Hiroshi Shimodaira (2010). "Evaluation of a hierarchical reinforcement learning spoken dialogue system", Computer Speech and Language, 24(2): 395-429.

Prospects for commercial use

Spoken dialog systems as an application
of planning under uncertainty

Tracking a distribution over dialog states

There is a technique in research which accumulates information from all of the N-Best lists across all recognitions to yield a combined, whole-dialog confidence score. For example, if the same city were recognized twice with low local confidence, after the second recognition it would have a much higher global confidence. This same technique can synthesize together multiple N-Best lists to find the most likely user goals. What is the potential for success of this approach in commercial systems?

6/10 practitioners already aware

Potential	N
High	7
Moderate	3
Low	0

Pros:

- Right now I'm reading up on this and trying to figure out how we can use it in our applications
- We built an application which uses a more basic version of this approach – it “really helped" task completion
- If this could re-rank N-Best lists, it “could improve dialog quite a lot”
- Our apps are context free with no memory; if we could improve on this, I'm sure it would help performance

Source: Anonymous interviews with 10 industry practitioners

Multiple dialog states

Cons:

- The problem is <u>comprehensibility</u> to testers and designers, so the cost and obstacles may not be paid back in the benefit.
- Might make sense for <u>large-scale, centrally managed systems</u>, but not in one-off smaller systems (N=2)

Conclusions for researchers:

- This is a radically different approach to design vs. industry
- Need to figure out how to communicate this approach to practitioners – including concepts, engineering, APIs, etc.
- Start addressing large-scale problems hard recognition tasks (e.g., business search)
  - Scalability
  - Relationship to search/question-answering

Source: Anonymous interviews with 10 industry practitioners

Automatic action selection

There is a technique in research which tries to learn the best action to take in each dialog state automatically. The idea is for a designer to specify, at each dialog state, a small set of possible actions. Then the designer specifies an overall objective function, such as +10 points for successful completion of the dialog, and -1 point for each question asked. Then the machine tries taking each action in each state, and works out which combination is optimal. What is the potential for success of this approach in commercial systems?

5/10 practitioners already aware

Potential	N
High	5
Moderate	4
Low	1

Pros:

- Designers often have to make many guesses about how people will react. This takes out some of the guesswork.
- "Extremely high" potential for success, especially if this could link abandonment to specific prompts and interactions.
- "I could see how this could save some time" because I wouldn't have to define all the arcs in the callflow.

Source: Anonymous interviews with 10 industry practitioners

Automatic action selection

Cons:

- One obstacle is the level of skills required to do the optimization: doing this without a PhD right now is impossible.
- To get a client to "sign off", you need to make it clear what they're signing off on – documenting all the different versions could be very tedious to produce.
- If there are many paths, how do we know that all paths make sense? Would some paths be crazy?
- See also "VUI Completeness" (Paek and Pierracini. 2008. Automating spoken dialogue management design using machine learning: An industry perspective. Speech Communication.)

Conclusions for researchers:

- Need to be able to assure that all possible user experiences are acceptable
- Incorporation of business rules is crucial (cf Williams, 2008)

Source: Anonymous interviews with 10 industry practitioners

Spoken Dialogue Challenge 2010

First deployment to <u>real callers with real needs</u>

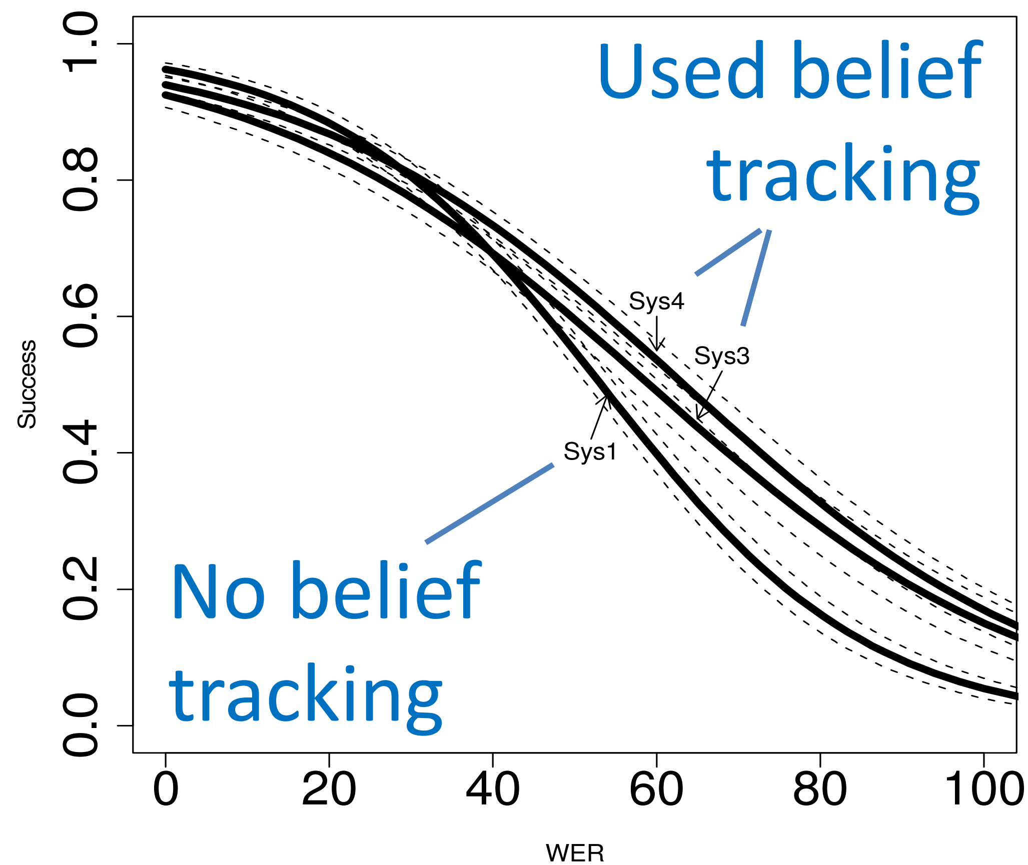
Task: Bus timetable information for Pittsburgh, PA

NSF-sponsored

3 systems deployed

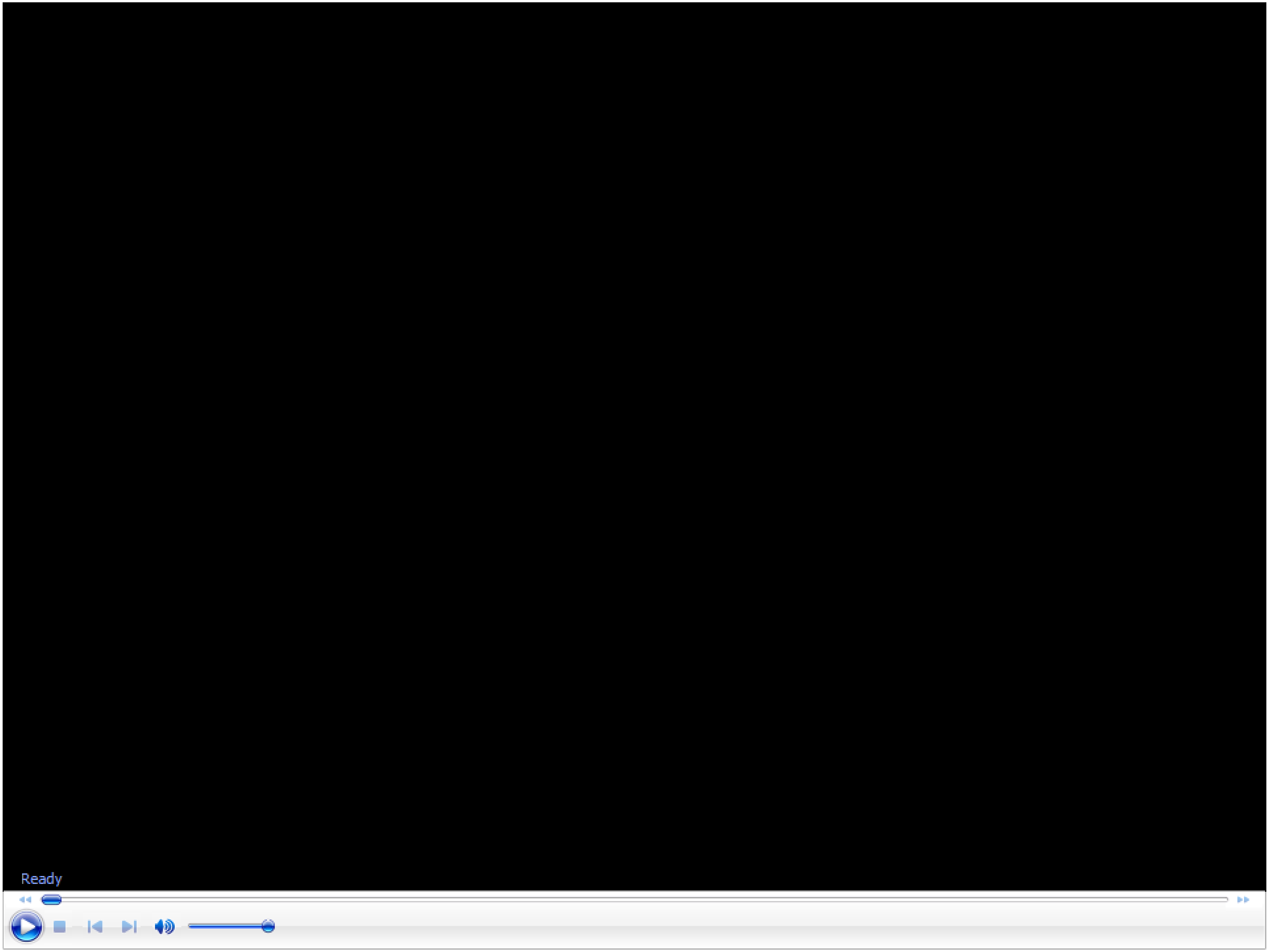
400-700 calls/sys

No RL/planning (yet?)



Alan W Black, Susanne Burger, Alistair Conkie, Helen Hastie, Simon Keizer, Oliver Lemon, Nicolas Merigaud, Gabriel Parent, Gabriel Schubiner, Blaise Thomson, Jason D. Williams, Kai Yu, Steve Young and Maxine Eskenazi. 2011. Spoken Dialog Challenge 2010: Comparison of Live and Control Test Results. Proc SigDial. Portland, Oregon, USA.

.


Ready

Some thoughts on the future

Spoken dialog systems as an application of planning under uncertainty

Which planning algorithms to use?

To date, many RL algorithms have been applied

Which **should** we be using?

Are there special properties we can exploit?

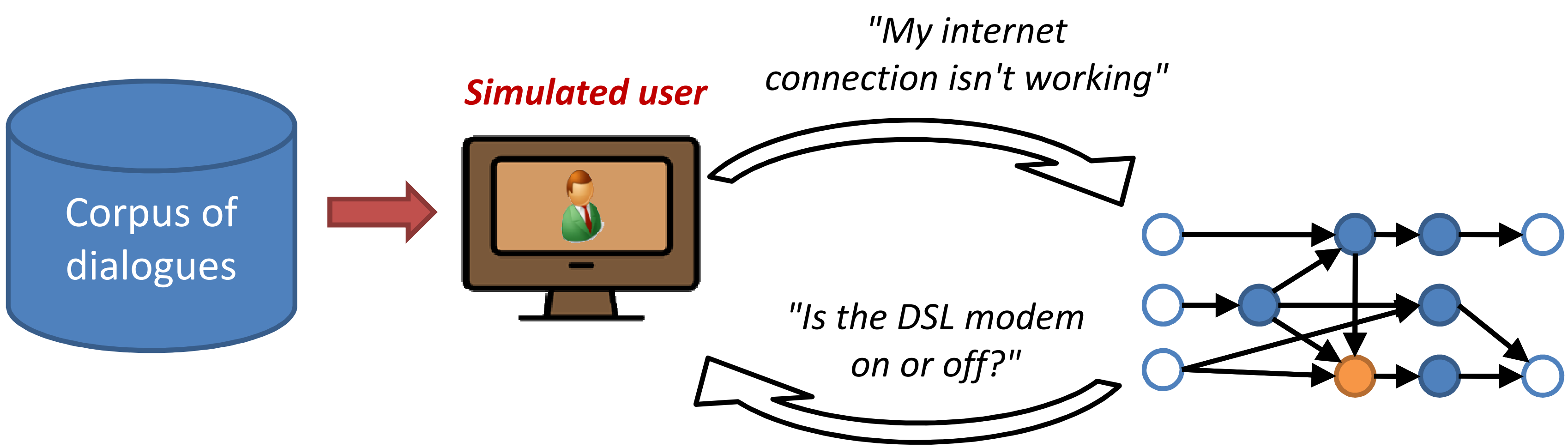
Algorithm	Reference
Natural actor-critic	B. Thomson and S. Young (2010). "Bayesian update of dialogue state: A POMDP framework for spoken dialogue systems." Computer Speech and Language, 24 (4): 562-588.
Monte-Carlo sampling	S. Young, M. Gasic, S. Keizer, F. Mairesse, J. Schatzmann, B. Thomson and K. Yu (2009). "The Hidden Information State Model: a practical framework for POMDP-based spoken dialogue management." Computer Speech and Language, 24(2): 150-174.
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.

Feature selection in RL

- I have lots of **possibly** useful features in the dialog state
- But I don't know which ones are **actually** useful for planning
- Modern machine learning algorithms perform feature selection (eg random forests)
- How can feature selection be incorporated into planning/RL?

What is a good simulated user?

Planning often requires a simulated user.



How to measure the **quality** of a simulated user?

Metrics have been proposed; none yet widely accepted

How is this done in other fields?

What is a good reward function?

How should the reward function be set?

Can this be an interactive process?

We have examples of reasonable dialog systems – can we apply inverse reinforcement learning (IRL) to infer the reward function?

(IRL is hard!)

In conclusion

Building spoken dialog systems is difficult

Cast the problem as planning under uncertainty

- Distribution over states gives robustness to errors
- Automated planning (RL) yields more detailed plans

Scale up distribution by partitioning / factoring

Scale up planning with features / expert knowledge

Good potential for commercial impact; promising first real-world results

Interesting open practical and theoretical questions

If you want to get started...

Reading

- **Book chapter**
  Jason D. Williams. A case study of applying decision theory in the real world: POMDPs and spoken dialog systems. Chapter in *Decision Theory Models for Applications in Artificial Intelligence: Concepts and Solutions*, L.E. Sucar, E. Morales, H. Hoey (Eds.). IGI Global.
  www.research.att.com/people/Williams_Jason_D
- **Bibliography**
  www.research.att.com/~jdw/sdsbib.html

Tools

- **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 to my collaborators

Steve Young
Matt Stuttle
Blaise Thomson
Pascal Poupart
Alan Black
Maxine Eskenazi
Iker Arizmendi

Alistair Conkie
Amanda Stent
Diamantino Caseiro
Jay Wilpon
Vincent Goffin
Suhrid Balakrishnan
Lihong Li
Umar Syed

Thanks!

Jason D. Williams

Spoken dialog systems as an
application of planning under uncertainty

ICAPS – June 2011