On-line Learning of Wide-Domain Statistical Spoken Dialogue Systems

(Talking to Machines)

Steve Young

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

Objectives

To develop spoken dialogue systems which:

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

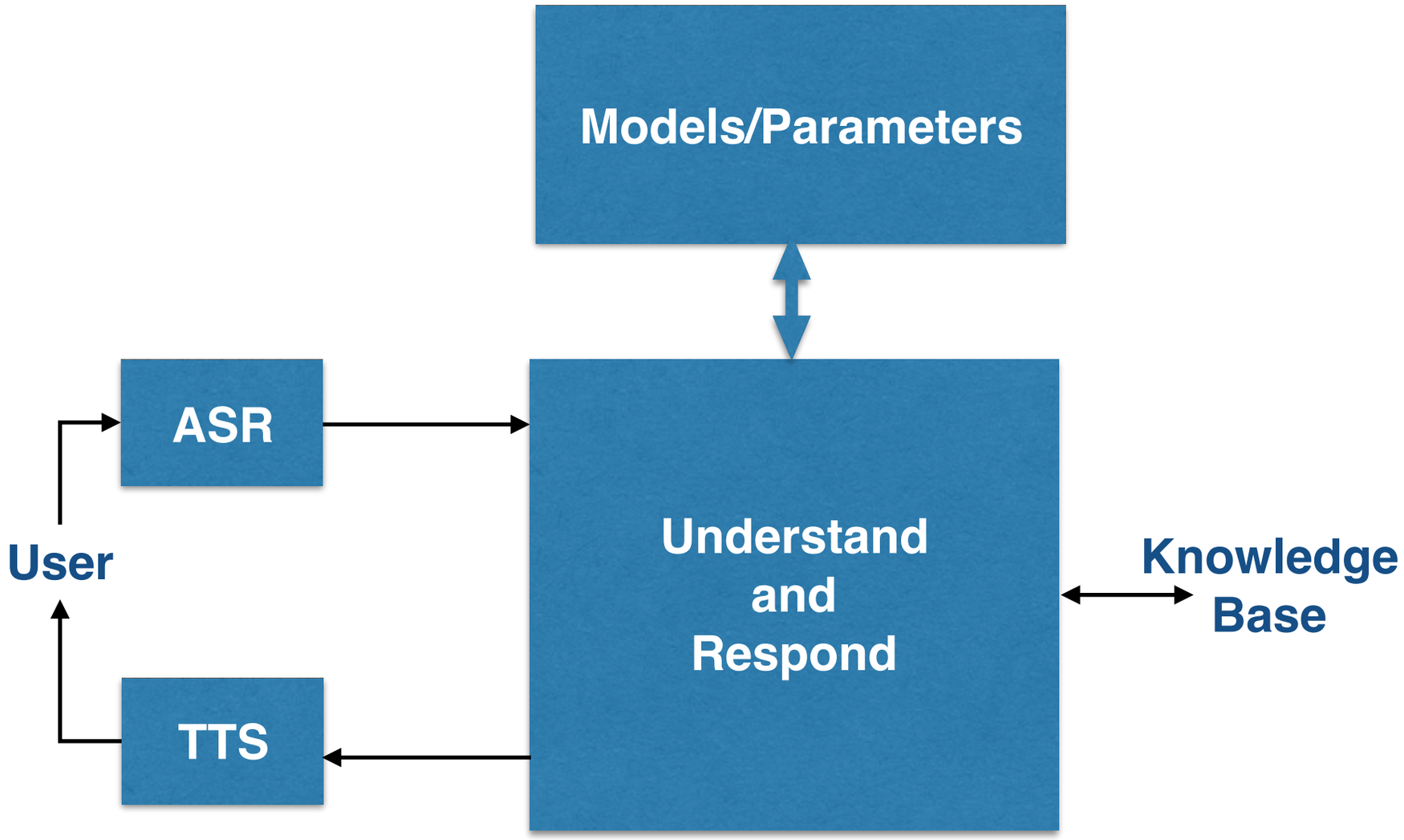
2. support natural conversations even in rarely visited domains

3. learn automatically on-line through interaction with user

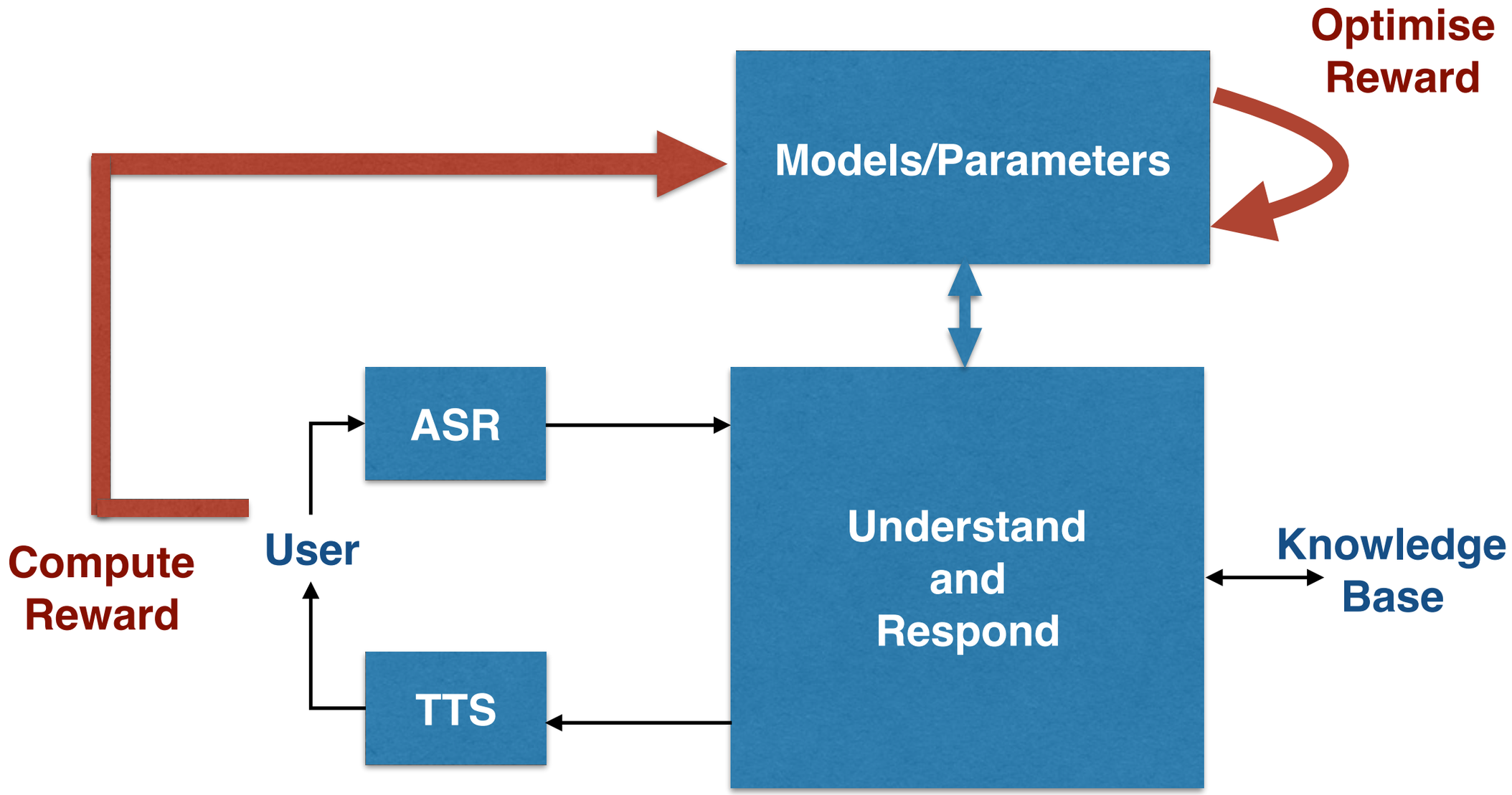
“Deploy, Collect Data, Improve”

“User in the loop” enables on-line reinforcement learning

Reinforcement Learning



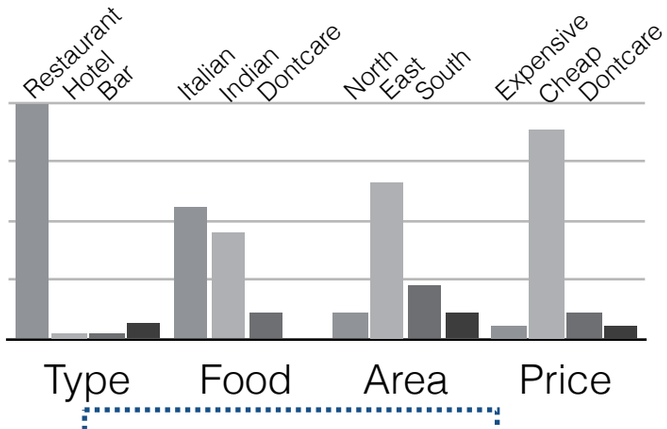
Reinforcement Learning



Single Limited Domain Statistical SDS

I'd like a cheap Italian
on the east side of town

inform(
price=cheap,
food=Italian,
area=east) [0.7]



Understanding

ASR

Semantic Decoder

Belief Tracker

Turn Level

Dialogue Level

Policy

Policy Based Decision Logic

Knowledge Base

User

Turn Level

Dialogue Level

TTS

Message Generator

Response Planner

Generation

Dialog Manager

You'd like a cheap restaurant on
the east side of town?
What kind of food would you like?

confirm-request(
price=cheap,
area=east,
food=?)

confirm-request(food)

Extending to Wide Domains

Two key problems:

1. How to expand coverage from a single limited domain to wide or even unlimited domains
2. How to measure success (and hence a reward signal)

Multi-domain SDS

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?

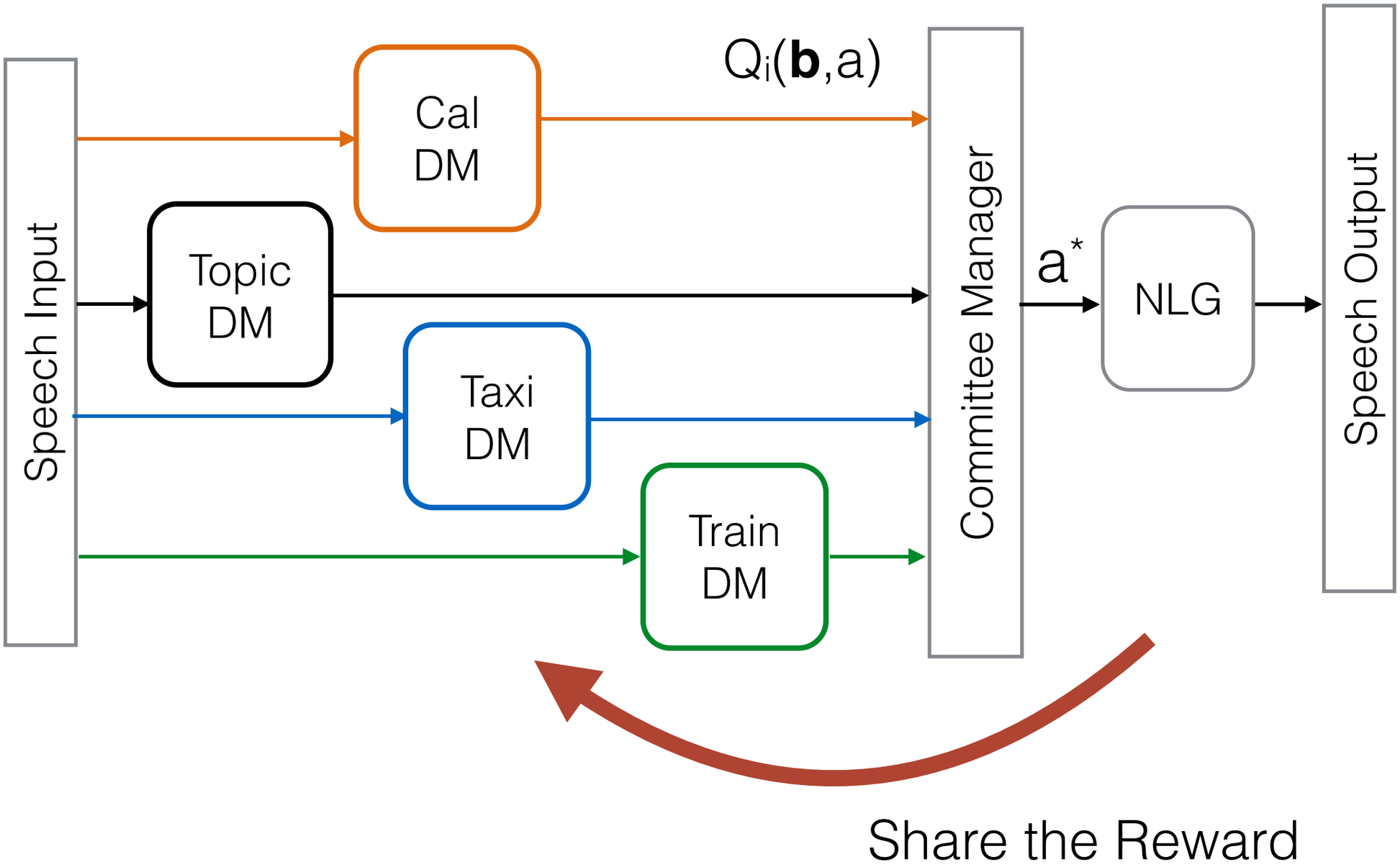
Yes, thats fine. I also need a taxi to the station.

What time do you need the taxi?

When does the train depart to London?

The 9.15am gets in at 10.06.

….



Bayesian Committee Machine

- 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:

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

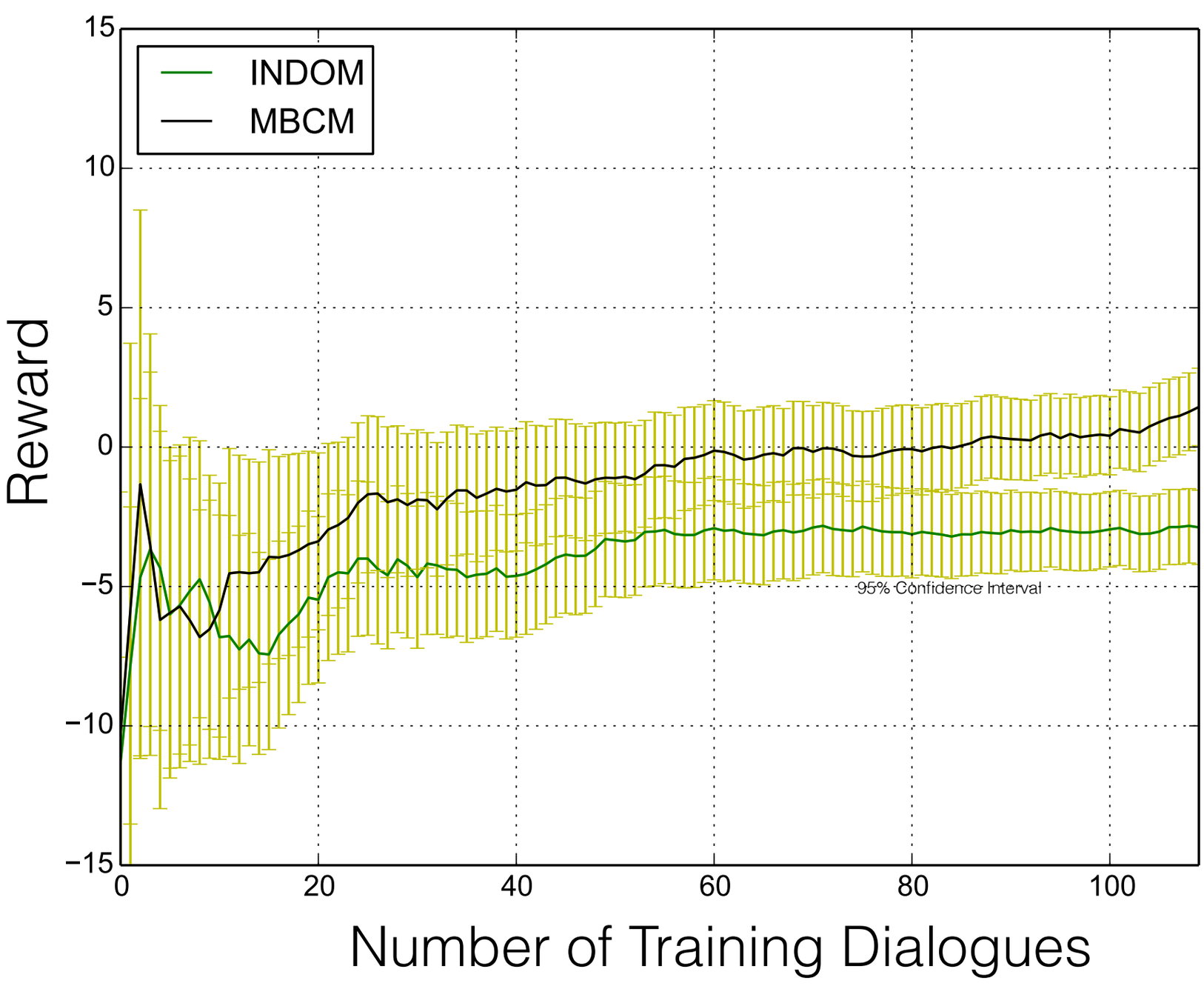
$$\Sigma^{Q}(\mathbf{b},a)^{-1} = -(M-1)*k((\mathbf{b},a),(\mathbf{b},a))^{-1} + \sum_{i=1}^{M}\Sigma_i^{Q}(\mathbf{b},a)^{-1}$$

- 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.

Bayesian Committee Machine Training Performance



Laptop domain trained in parallel with Hotels and Restaurants

Laptop domain trained in isolation

Training with Real Users

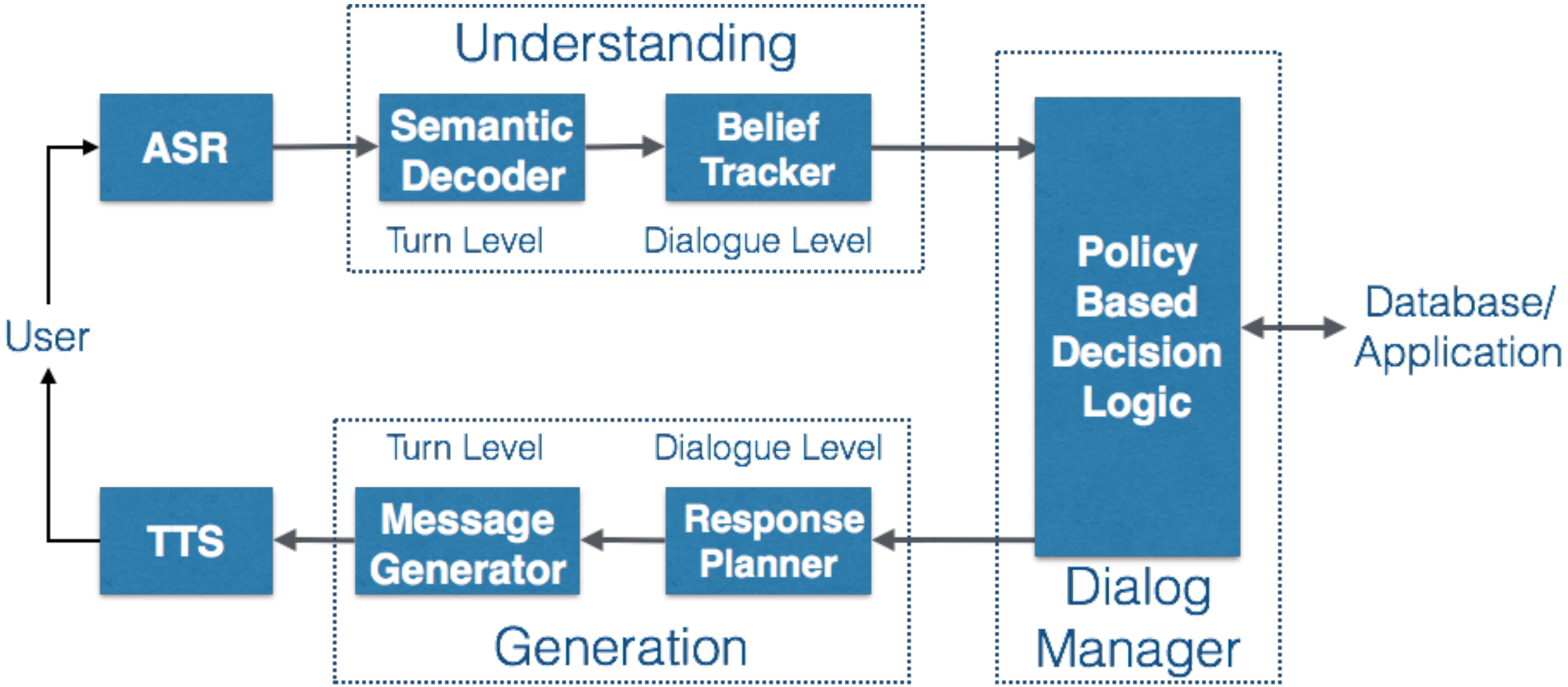
Most research results are obtained using paid subjects given prescribed tasks. Moving to real systems presents problems:

- Reward depends on task success which is very hard to measure
- Explicit user feedback is costly and unreliable

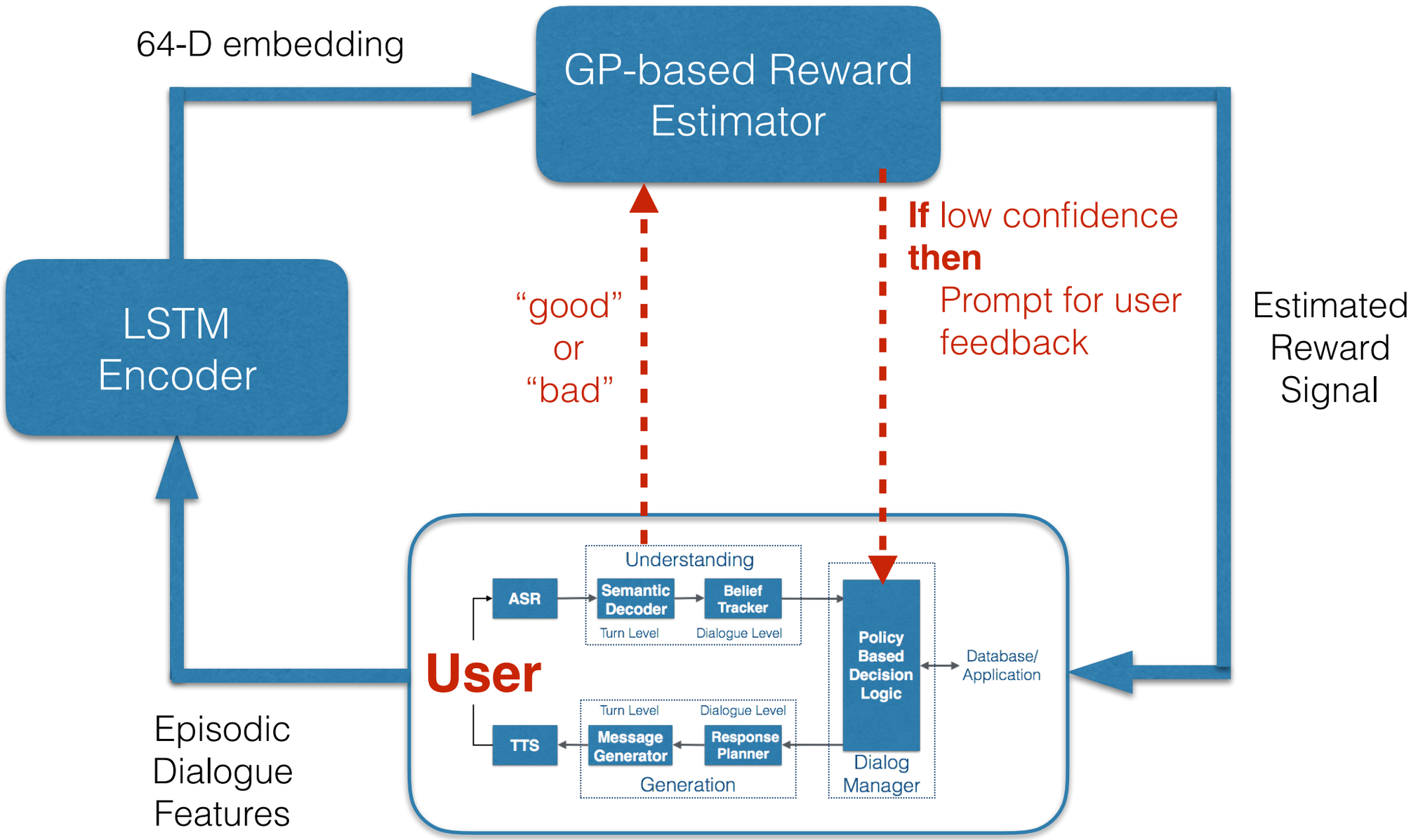
Solution:

- Learn an embedding function for dialogues (using a Bi-LSTM)
- Train a Gaussian Process based classifier to estimate reward success
- Use GP uncertainty estimate to limit use of explicit user feedback
- Use GP noise model to compensate for unreliable user feedback

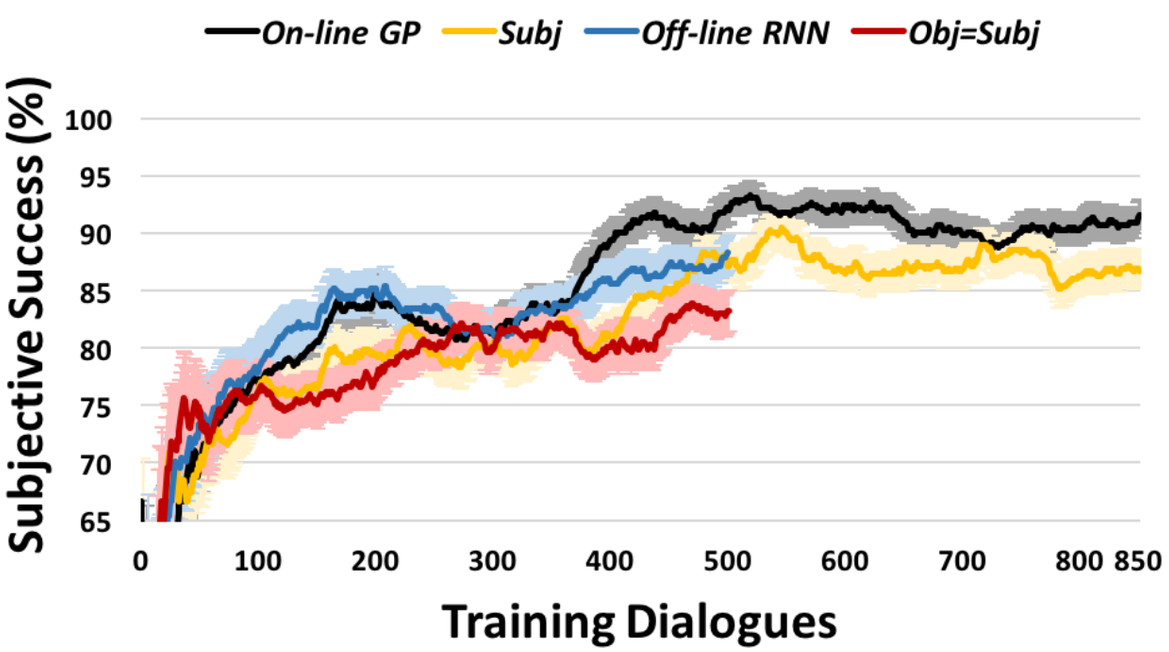
On-line Reward Estimation

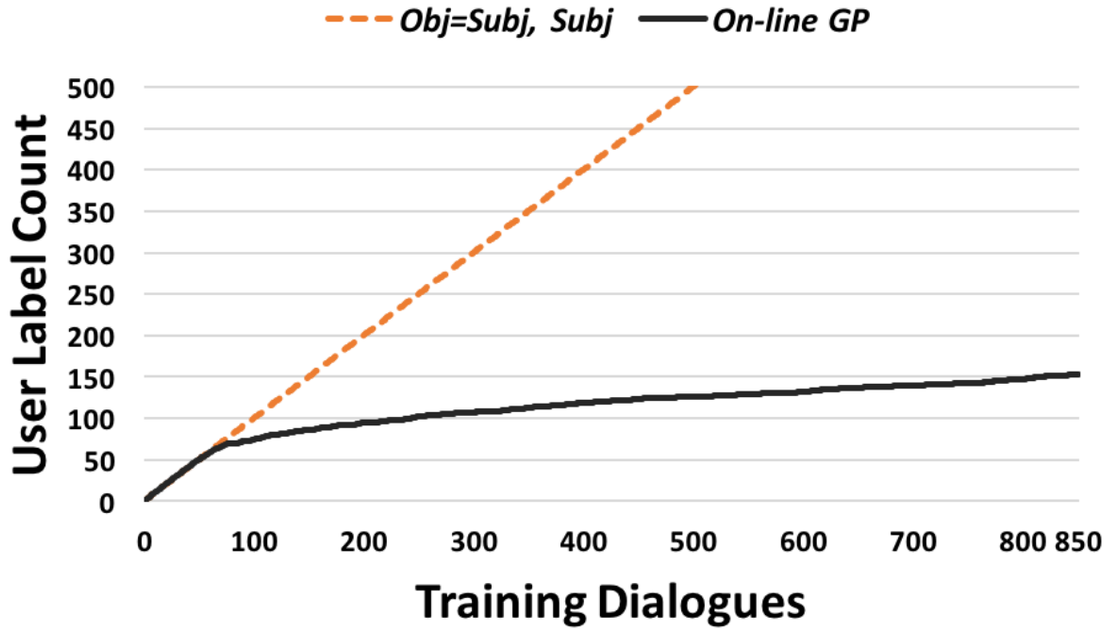


On-line Reward Estimation



On-line Reward and Policy Learning





P-H. Su et al (2016). "On-line Active Reward Learning for Policy Optimisation in Spoken Dialogue Systems." ACL 2016, Berlin.

Conclusions

- Technology components are now in place to build large scale wide-domain spoken dialogue systems
- Capability and user acceptance of Virtual Personal Assistants (VPAs) will increase rapidly
- Key is ability to learn on-line thru interaction with users and sharing data with other VPAs
- VPAs will become autonomous entities, independent of any specific device
- This will raise many issues: ensuring veracity of VPA derived information, personal privacy, consumer protection, …