Semantic graph clustering for POMDP-based spoken dialog systems
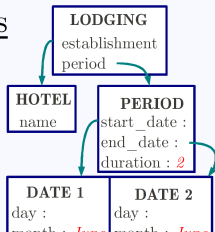
Florian Pinault and Fabrice Lefèvre, University of Avignon, France
{florian.pinault,fabrice.lefevre}@univ-avignon.fr

FLEXIBLE SEMANTICS

Ontology of frames

FrameNet inspired
Rich representation
  semantic graphs
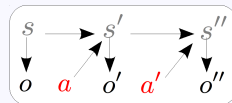**Data-driven**
  flexible structure

Ex: « *I need an hotel for two days in June* »

+ REINFORCEMENT LEARNING

POMDP
(Partially Observable
Markov Decision Process)

a action: System action
s state: Exact user turns (unobserved)
o observation: Noisy *n-best* list of user turns

Belief tracking: $b(s)$ updated from o and a
Policy depends on the belief: $a = \pi(b)$

= INTRACTABLE

SUMMARY POMDP

Mapping master into summary POMDP

$$\tilde{s} = M_s(s) \qquad \tilde{o} = M_o(b)$$

HandCrafted Summaries: Logical rules, linear function, etc.
Automatic Clustering Summaries: Need distances
  $d(s_1, s_2)$ distance between graphs
  $d(b_1, b_2)$ distance between *n-best* lists of graphs
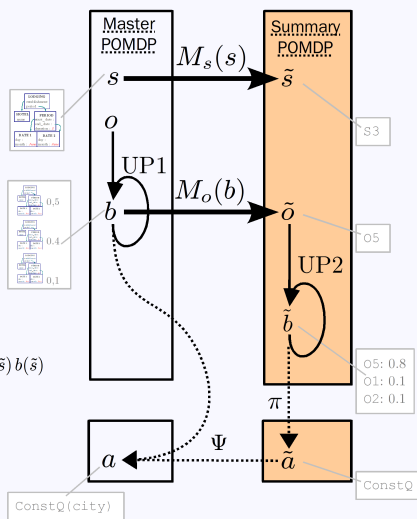
Two belief tracking: master AND summary

UP1: *in master space with n-best scores*
UP2: *in summary space with full model* $b'(\tilde{s}') = k.P(\tilde{o}'|\tilde{s}') \sum_{\tilde{s} \in \tilde{\mathbb{S}}} P(\tilde{s}'|\tilde{a}, \tilde{s}) b(\tilde{s})$
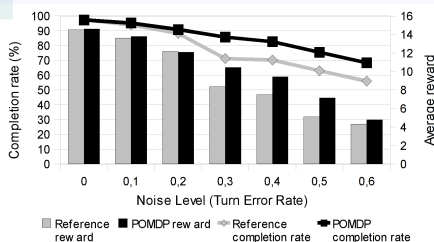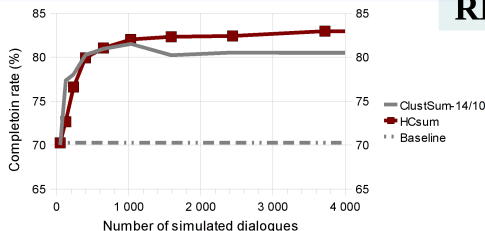
True POMDP policy optimization
  with full model algorithm (PBVI-Perseus).

Back to master action *with a heuristic:* $a = \Psi(\tilde{a}, b)$

RESULTS

Training improves the system policy
with *handcrafted* summary (HCSum)
as well as
with *automatic clustering* summary (ClustSum)

Robust performance improvement
Better cumulated reward
Better competion rate