LEARNING SYSTEM SPHINX

Thomas Leopold, Gabriele Kern-Isberner, Gabriele Peters

Combining Reinforcement Learning and Belief Revision – A Learning System for Active Vision

Belief Revision:

- based on logic
- no heuristic approach; has a sophisticated theoretical background (Epistemic State Revision postulates; AGM theory)
- literal: represents an elementary aspect of the world
- ω : model, complete conjunction of all literals
- Ω : set of all models
- OCF $\kappa : \Omega \rightarrow \mathbb{N} \cup \infty$ with $\kappa(0)^{-1} \neq \{\}$
- $\kappa(\omega_1) < \kappa(\omega_2) \Rightarrow \omega_1$ is more plausible than ω_2
- $\kappa(A \vee B) = min\{\kappa(A), \kappa(B)\}$
- conditional: $(B|A)$
- $\kappa \models (B|A) \Leftrightarrow \kappa(AB) < \kappa(A\overline{B})$
- new information: A resp. $(B|A)$
- revise κ so that $(\kappa * A) \models A$ resp. $(\kappa * (B|A)) \models (B|A)$

Example

0	ABC
1	$\overline{A}BC,\ A\overline{B}C,\ AB\overline{C}$
2	$\overline{A}\overline{B}C,\ \overline{A}B\overline{C},\ A\overline{B}\overline{C}$
3	$\overline{A}\overline{B}\overline{C}$

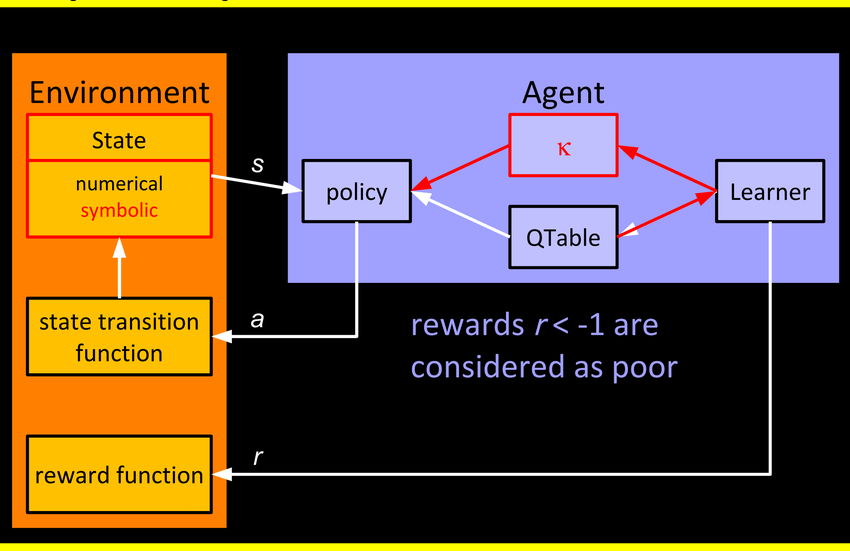
$\xrightarrow{*\overline{B}}$

0	$A\overline{B}C$
1	$\overline{A}\overline{B}C,\ A\overline{B}\overline{C},\ ABC$
2	$\overline{A}\overline{B}\overline{C},\ \overline{A}BC,\ AB\overline{C}$
3	$\overline{A}B\overline{C}$

Hybrid Learning Method Sphinx:

- similar to human learning (two levels)
- top level:
  - explicit learning
  - declarative knowledge (aware of it while applying it)
  - can easily be formulated, e. g. by if-then-rules
  - example: mathematical calculations
- bottom level:
  - implicit learning
  - procedural knowledge (not aware of it while applying it)
  - can hardly be put into words
  - example: riding a bike
- levels are interlinked: bottom-up vs. top-down learning

A Step in Sphinx



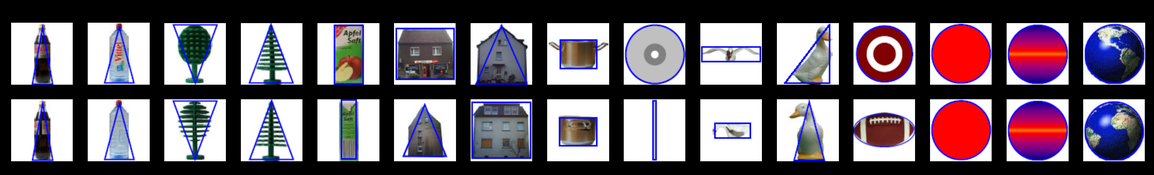
Models, Chosing Actions, and Belief Revision in Sphinx:

- *multi*-valued literals $d_i = v_{i.j}$ represent elementary aspects of current state of the world
- $s_{sym} = (d_1 = v_{1.j(1)}) \wedge \ldots \wedge (d_n = v_{n.j(n)})$
- dual representation: $s = (s_{num}, s_{sym})$:
  - s_{num} for RL part
  - s_{sym} for BR part
- literal $action = a_i$ represents an action
- models have the form: $s_{sym} \wedge (action = a_i)$

- consider models with $\omega \models s_{sym}$ only
- determine the most plausible models among these models (i. e. smallest κ-value)
- extract the actions from these models: $A_\kappa(s)$
- among these actions, find the ones with greatest $Q(s_{num}, \cdot)$-value: $A_{best}(s)$
- perform a random action from $A_{best}(s)$

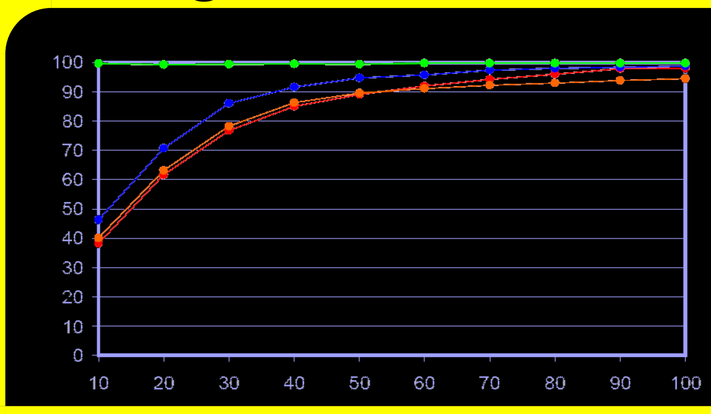
- performed at the end of each step
- read the new $Q(s_{num}, \cdot)$-values and determine the new best actions for s
- pattern p: conjunction of *some* literals
- revision with up to 4 types of conditionals
  - $(\overline{action = a} \mid s_{sym})$
  - $(\overline{action = a} \mid p)$, where p is a pattern satisfying s_{sym}
  - $(\bigvee_i action = a_i \mid s_{sym})$, where all a_i are best actions for s
  - $(\bigvee_i action = a_i \mid p)$, where each a_i is a best action in at least *one* state covered by the pattern p
- background knowledge: revision before starting to learn

Results:



- background knowledge: rules that enabled the agent to classify each object correctly, if it has perceived all views; and rules instructing it to look at the object from all views at first
- goal in this case: minimize the number of necessary views
- some rules that have been found:
  - If *FrontViewShape* = *TriangleUp* and *FrontViewSize* = *Tall*, then *Action* = *RecognizeBottle*
  - If *FrontViewShape* = *Circle* and *SideViewShape* = *Unknown* and *Texture* = *Simple*, then *Action* = *RotateLeft*
  - If *Texture* = *Simple*, then $\overline{Action = RecognizeBottle}$
- Sphinx learns faster than Q-learning even without background knowledge

Recognition Rates after *n* Episodes



- Q(0)
- Q(0.5)
- Sphinx without background knowledge
- Sphinx with background knowledge

University of Technology Dortmund, and Dortmund University of Applied Sciences and Arts

contact info: thleopold@hotmail.com

BMVC 2008