

FIRST ORDER DECISION DIAGRAMS FOR RELATIONAL MDPs

SAKET JOSHI, RONI KHARDON, CHENGGANG WANG, TUFTS UNIVERSITY

LOGISTICS DOMAIN		DOMAIN DYNAMICS CAPTURED BY TRUTH VALUE DIAGRAM		ACTION CHOICE PROBABILITY FOR $unloadS(b^*, t^*)$	REWARD FUNCTION V^0
$Bin(B,C)$	$unload(b^*, t^*)$	Generic [left]			$Bin(b, Paris)$
$On(B,T)$	$load(b^*, t^*, c^*)$	$unloadS(b^*, t^*)$ [middle]			10
$Tin(T,C)$	$drive(t^*, c^*)$	$unloadF(b^*, t^*)$ [right]			0

FODD SEMANTICS

Multiple Path semantics more suitable for Value Iteration

$D = \{1, 2\}$

$I = \{p(1), q(2), h(2)\}$

SINGLE PATH SEMANTICS

$\exists x, p(x)$ Bloekel & DeRaedt 97

$\wedge \neg \exists z [p(z) \wedge q(z)] \Rightarrow 0$

$MAP_B(I) = 0$

MULTIPLE PATH SEMANTICS

Groote & Tv 03

$\zeta \equiv \{x/1, y/1\} \rightarrow MAP_B(I, \zeta) = 0$

$\zeta \equiv \{x/1, y/2\} \rightarrow MAP_B(I, \zeta) = 0$

$\zeta \equiv \{x/2, y/1\} \rightarrow MAP_B(I, \zeta) = 0$

$\zeta \equiv \{x/2, y/2\} \rightarrow MAP_B(I, \zeta) = 1$

$MAP_B(I) = \text{MAX}_{\zeta} [MAP_B(I, \zeta)]$

$MAP_B(I) = 1$

VALUE ITERATION ALGORITHM

$$T_{n+1}^{A(x)}(V_n) = \gamma [\oplus_j (\text{prob}(A_j(x)) \otimes \text{Regr}(V_n, A_j(x)))]$$

$$Q_{n+1}^A = R \oplus \text{obj-max}(T_{n+1}^{A(x)}(V_n))$$

$$V_{n+1} = \max_A (Q_{n+1}^A)$$

● Operation implemented by: Standardize apart – combine - reduce

[1] Regression over deterministic alternatives of stochastic actions by block replacement.

[2] Conditional probability for $unloadS(b^*, t^*)$ combined with [1].

[2'] Diagram corresponding to [2] under $unloadF(b^*, t^*)$.

[3] Expected abstract value function over stochastic action $unload(b^*, t^*)$ obtained by the combination of [2] and [2']. The diagram gets complicated but it can be reduced using reduction operators. The parts inside the colored contours can be removed. Finally the diagram is reduced to [4] prior to object maximization.

[4] Object maximization is performed on this reduced diagram. The equality node will be eliminated (R6) and its true branch will survive since b^* is now a variable and its value can be chosen to make the equality true.

[5] Maximization over all actions gives V^1 .

[6] Block replacement for the next iteration. Implemented by block combination using $[B \times B_i] + [(1-B) \times B_j]$