Augmentation of tLDGBAs and Synthesis Method

We introduce an automaton augmented with binary vectors. The automaton can explicitly represent whether transitions in each accepting set occur at least once, and ensure transitions in each accepting set occur infinitely often.

Let $V = \{(v_1, \ldots, v_n)^T \,;\, v_i \in \{0,1\},\ i \in \{1, \ldots, n\}\}$ be a set of respectively. In order to augment a tLDBA B_φ, we introduce three functions $visitf : \delta \to V$, $reset : V \to V$, and $Max : V \times V \to V$ as follows. For any $e \in \delta$, $visitf(e) = (v_1, \ldots, v_n)^T$, where v$_i$ = $\{1 if\ e \in F_i, 0 otherwise.$ For any $v \in V$, reset(v) = $\{0 if\ v = 1, v otherwise.$ For any $v, u \in V$, $Max(v, u) = (l_1, \ldots, l_n)^T$, where $l_i = max\{v_i, u_i\}$ for any $i \in \{1, \ldots, n\}$.

Each vector v is called a memory vector and represe-

$\mathrm{P}^{\otimes}(s^{\otimes\prime}|s^{\otimes},a) = \left\{ 1 if\ (s{=}s')\wedge(v{=}v')\wedge(x,\varepsilon_{x'},x')\in\bar{\delta}, 0 otherwise, \right.$ where $s^{\otimes} = (s,(x,v))$ and $s^{\otimes} = (s',(x',v'))$.

[Reward assignments] The reward function $\mathcal{R} : S^{\otimes} \times A^{\otimes} \times S^{\otimes} \rightarrow R_{\geq 0}$ is defined as $\mathrm{R}(\mathrm{s}^{\otimes}, a, s^{\otimes\prime}) = \left\{ r_p\ if\ \exists i \in \{1,\ldots,n\},\ (s^{\otimes},a,s^{\otimes\prime}) \in \bar{F}_i^{\otimes}, 0\ \ otherwise, \right.$ where r_p is a positive value.

Under the product MDP $M^{\otimes}$ and the reward function $\mathcal{R}$, which is based on the acceptance condition of $M^{\otimes}$, we show that if there exists a positional policy state and the action space is specified with $\mathcal{A}(s) = \{Right,\ Left,\ Up,\ Down\}$ for any state $s \neq s_4$ and $\mathcal{A}(s_4) = \{to\_s_0,\ to\_s_1,\ to\_s_2,\ to\_s_3,\ to\_s_5,\ to\_s_6,\ to\_s_7,\ to\_s_8\}$, where to_s_i means attempting to go to the state s_i for $i \in \{0,\ 1,\ 2,\ 3,\ 5,\ 6,\ 7,\ 8\}$. The robot moves in the intended direction with probability 0.9 and it stays in the same state with probability 0.1 if it is in the state s_4. In the states other