

1 緒論

ソフトリアルタイムシステムはデッドラインを過ぎると価値が減少する。価値の減少率は各タスクによって定まるため複数のタスクを処理するためには、これらを考慮した最適なスケジューリングを得る必要がある。Chan と Wanham はスーパーバイザ制御を用いたスケジューリングを提案している [1]。また近年言語測度と呼ばれる形式言語に対する符号付きの実測度の概念の導入により [2]，離散事象システムの定量的な評価がなされている。本報告では周期タスクのソフトリアルタイム処理を時間付きオートマトンでモデル化し，それらを合成する。こうすることで状態に評価値や事象にコストを与えることで価値の減少を表すことができ，言語測度による定量的な評価が可能になる。このモデルに言語測度に基づく最適スーパーバイザの強化学習を適用することで最適なスケジューリングを得る手法を提案する。

2 タスクのモデル

周期タスクのソフトリアルタイム処理を時間付きオートマトンでモデル化する。周期内に処理を完了できる場合のみデッドラインミスを許容する。図 1 にタスク処理のオートマトンの例を示す。W,I,E,D はそれぞれ待ち状態，アイドル状態，実行状態，禁止状態を表す。デッドラインを過ぎてモタスクの処理が完了できていない状態をデッドラインミス状態と呼び図 1 では色付きの状態である。周期内にタスクが処理できない場合は禁止状態へ遷移する。α_i, β_i, γ_i (i = 1, 2) はタスクの処理開始，処理終了，呼び出しを表し，tick は時間の進行を表す事象である。tick はデッドラインミス状態へ遷移する事象 tick₁ とそうでない事象 tick₀ に細分できる。

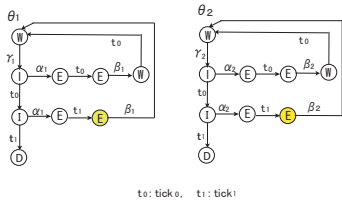


図 1 タスク処理のオートマトン

3 オートマトンの合成

本報告ではマルチタスクシステムを対象とためシステムの振る舞いを表すため，タスク処理のオートマトンを合成し複数タスクの処理を表すオートマトンを構成する。tick はすべてのタスクで同期して生起し，それ以外の事象は各タスクの状態で遷移が定義されていれば合成オートマトンでも遷移が定義される。なお合成オートマトンの tick はタスクの個数分の 0, 1 変数を持ち，生起した tick によってタスクがデッドラインミス状態へ遷移する時はタスクに対応する変数を 1，しない時は 0 とする。目標状態は禁止状態，デッドラインミス状態，タスクの待ち状態，タスクの処理終了直後の状態とする。図 2 に図 1 のオートマトンを合成したものを示す。

4 学習のアルゴリズム

文献 [3] のアルゴリズムを拡張し，Q-learning による G の言語測度に基づく最適スーパーバイザの学習アルゴリズムを示す。

1. 各状態で報酬の期待値と事象の生起確率に初期値を与える。
2. 各状態での Q 値の初期値を計算する。
3. 以下を各エピソードに対して繰り返す。
 - (a) q_i を合成オートマトン G の初期状態にセットする。
 - (b) q_i が終端状態になるまで以下のステップを繰り返す。
 - i. スーパーバイザは生起強制確率，生起禁止確率に基づいて制御パターンを選択し，G に提示する。

- ii. スーパーバイザは G での事象 σ の生起とそれとともなう q_i から q_k への状態遷移を観測する。
- iii. 報酬 (q_i の特性関数から q_i での σ の事象コストを引いたもの) を得る。
- iv. q_i での報酬の期待値と事象の生起確率を更新する。
- v. q_i での他のスーパーバイザの制御パターンに対して，間接的に Q 値を更新する。
- vi. 最適制御パターンを計算し，生起強制確率，生起禁止確率を更新する。
- vii. 現在の状態 q_k を q_i とする。

tick の事象コストは 0, 1 変数を参照し，デッドラインミスの数や各タスクのデッドラインミス後の価値の減少率から決定される。他の事象にコストはかからない。特性関数については禁止状態とデッドラインミス状態に負の値を，タスクの待ち状態の集合，タスクの処理終了直後の状態には正の値を与える。本報告では 1 個のプロセッサでの処理を想定しているため，複数のタスクの実行状態には最小の特性関数を与え生起を禁止する。

5 シミュレーション

図 2 の合成オートマトンに学習アルゴリズムを適用して，Q 値を最大にするスーパーバイザを求めた。エピソードは 1000 回繰り返した。これにより制御されるオートマトンを図 3 で表す。破線は生起が禁止されていることを示す。事象コスト各状態での tick(1, 0) には 0.1, tick(0, 1) は 0.05, tick(1, 1) は 0.15 とし，tick(0, 0) は 10⁻⁹ とした。特性関数は状態 0 は 1.0, 状態 1, 2, 8, 13, 22, 23, 24, 33, 34 は 10⁻⁸, 状態 12, 14, 26, 30, 31 と状態 11, 18, 29, 32 は -1.0, 状態 16 は -0.1, 状態 20 は -0.05, 状態 32 は -0.35 を与えた。なお，事象コストを遷移後受け取るため dummy state を加えた。図 3 より，デッドラインミス状態への時間の進行を表す事象のコストが大きく，デッドラインミス状態の特性関数値が小さいタスク 1 が先に実行されていることがわかる。

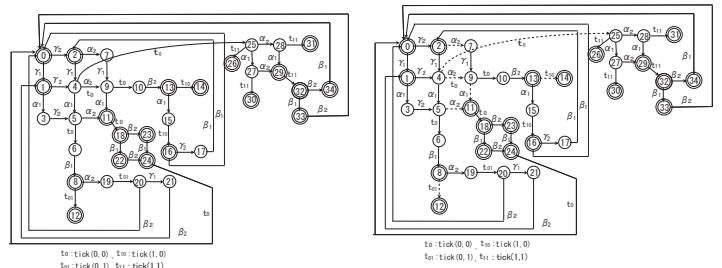


図 2 合成オートマトン

図 3 学習後のオートマトン

6 結論

周期タスクのソフトリアルタイム処理を時間付き離散事象オートマトンでモデル化し，それらを合成した。合成して得られたモデルに言語測度に基づくスーパーバイザの最適制御を強化学習によって設計し，最適なスケジューリングが得られることを確かめた。今後の課題としては各タスクが複数のサブタスクから構成されている場合への拡張や，プロセッサの数を増やした場合への拡張などが考えられる。

参考文献

- [1] P.C.Y.Chan and W.M.Wonham; *Real-Time Systems*, vol.23, no.3, pp.183-208 (2002)
- [2] X.Wang and A.Ray; *Proc. of 2002 American Control Conference*, Anchorage, pp.3937-3942 (2002)
- [3] 谷口 和隆, 山崎 達志, 潮 俊光; システム制御情報学会論文誌, vol.18, No.12, pp.433-439 (2005)