

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 16 日現在

機関番号：34428

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25870945

研究課題名(和文) 離散事象アプローチに基づく定量的リスク保証型最適スーパーバイザ制御機構の構築

研究課題名(英文) Construction of an optimal supervisory control mechanism with quantitative risk assurance based on discrete event approach

研究代表者

山崎 達志 (Yamasaki, Tatsushi)

摂南大学・理工学部・准教授

研究者番号：00368458

交付決定額(研究期間全体)：(直接経費) 2,000,000円

研究成果の概要(和文)：安心安全な制御系の実現のためには、システムの抱えるリスクを定量的に評価した上での設計が重要である。本研究では、離散事象システムに対し、システムの抱えるリスクを一定以下に保証する中で、事象の生起と禁止からのコストに関して最適とするスーパーバイザの設計法を示した。また、複数のサブシステムからなる大規模離散事象システムへの対応として、システムの振る舞いの先読みに基づく最適LLPスーパーバイザの設計法を示した。

研究成果の概要(英文)：For realization of a safety guarantee system, it is important to synthesize a controller which considers the quantitative risk evaluation of the system. In this study, we showed a synthesis method of an optimal supervisor of discrete event systems. The supervisor optimizes the costs of disabling and occurrence of events within the permissible risk of the controlled system. Moreover, we showed a synthesis method of an optimal LLP (limited lookahead policy) supervisor for a global system which is composed of several local systems.

研究分野：システム制御

キーワード：制御工学 離散事象システム スーパーバイザ制御 リスク 最適制御 強化学習

1. 研究開始当初の背景

事象の非同期的、並行的な生起により状態が離散的に遷移する離散事象システム(DES; Discrete Event System)は、生産システム、データベースシステム、通信システムなど、近年特に大規模・複雑化している様々な人工システムのモデル化に有効な、重要なシステムのクラスである。DES に対する論理的な制御法として Ramadge と Wonham によって提案されたスーパーバイザ制御がある。一般的なスーパーバイザ制御の枠組みでは、図 1 に示す概念図のように、制御器であるスーパーバイザが、論理的な制御仕様を満足するように、観測した DES の生起事象列をもとに、DES に対し生起を許可(または禁止)する事象の集合(制御パターン)を指定する。このとき、事象には、故障や割り込みといったスーパーバイザが生起を禁止できない不可制御事象も存在するとしている。

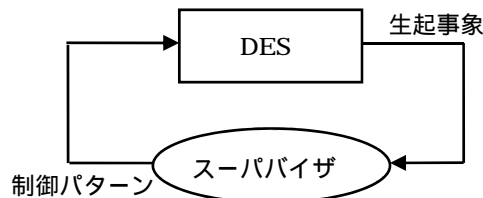


図 1 スーパーバイザ制御の概念図

スーパーバイザ制御の理論は主として形式言語とオートマトンを基に構築され、様々な研究が内外で行われてきている。ここでは、生成言語の包含関係に基づく議論が中心となってきたが、事象に関するコスト等を定量的な評価指標として用いた最適スーパーバイザ制御についての研究も行われている。

研究代表者は本研究開始までに、強化学習を用いた離散事象システムの最適スーパーバイザ制御に関する研究を行ってきた。基本的なアイデアは、スーパーバイザ制御問題を強化学習が適用可能なマルコフ決定過程下の問題として定式化することで、強化学習を用いて最適スーパーバイザの設計を行うというものであった。

現代においては、効率性、低コストといった経済的指標の直接的な追求だけでなく、安全性や信頼性を陽に考慮した制御系設計の重要性が一層増している。離散事象システムの最適スーパーバイザ制御においてもこのような観点からの研究は重要であると考えられる。

2. 研究の目的

本研究では、離散事象システムに対し、定量的なリスク保証と制御性能の双方を考慮した、新たなスーパーバイザ制御の手法を示す。そのために、コストや効率に関する指標に加えて、リスクに対する定量的な評価指標を導入する。制御仕様を完全に満たすためには、システムの振る舞いが非常に保守的に制限される可能性があるが、それが常に望ましい

こととは限らない。そこで、定量的に許容できるリスクを設定し、システムに求められる安全性・信頼性を一定に確保した上で、最適なスーパーバイザを設計するための手法を構築する。制御対象についての知識レベルが十分で、事前に制御器を設計できる場合のみでなく、情報が不十分で適応的に制御器を設計する必要がある場合についての設計法も示す。

また、現在の大規模複雑化したシステムに対しても有効な手法となるように、複数の時変するサブシステムから構成される離散事象システムを陽に考慮したスーパーバイザの設計手法を示す。

3. 研究の方法

前節の研究目的を達成するために、研究代表者の従来研究における成果に基づき、制御対象を、遷移先が確率的に定まる確率離散事象システム(PDES; Probabilistic DES)としてモデル化し、マルコフ決定過程下の最適制御問題としての観点からの定式化を行う。システムには守るべき制御仕様と、事象の生起と禁止に対して評価値(報酬)が与えられるとする。そのため、望ましくない事象は禁止したいが、その事象を禁止するために過大なコストが伴う場合には、あえてその生起を許容するというトレードオフの状況がありうる。さらに、システムの抱えるリスクとして、システムの各状態において、制御仕様を外れた状態へ遷移する確率を考える。そして、各状態でのリスクが許容する閾値以下になるなかで、報酬に関して最適となるリスク許容型最適スーパーバイザを設計する手法を示す。

大規模複雑化する離散事象システムに有効な手法とするための検討を行う。複数のローカル離散事象システム(LDES; Local DES)からなるシステムを考える。また、LDES の数・構成は増減しうるものとする。一般にシステム全体の状態集合は各 LDES の状態集合の直積となるため、容易に状態空間のサイズが爆発的に増大する。各 LDES をそれぞれ PDES により表現し、システム全体の有限ステップ先までの制御仕様内の振る舞いをツリーとして構成する。先読み木を用いたスーパーバイザ制御は、従来型のスーパーバイザ制御の枠組みにおいて LLP(Limited Lookahead Policy)スーパーバイザ制御として提案されているが、本研究ではこれを最適スーパーバイザ制御の枠組みに拡張して適用する。また、一般に LDES 間の関係は完全に独立ではなく、協調して生起される事象(共通事象)も存在するとする。そして、ツリーから計算される評価値が最大となる制御パターンを選択することで、大規模時変な離散事象システムに対しても最適スーパーバイザを設計する手法を示す。

提案する手法の評価として、計算機実験および小型ロボットによる実機実験を行い、そ

の有効性を示す。

4. 研究成果

本研究における研究成果として得られたリスク許容型最適スーパーバイザ制御の枠組みを図2に示す。各制御パターンの評価値として、与えられた制御パターンの中で、最も望ましくない事象が生じた最悪ケースにおける期待報酬についての評価値を、Bellman最適方程式を基に定めた。PDESにおいて、与えられた制御パターンの中から生じた事象と状態遷移、事象の生起と禁止に対して与えられた報酬を基に、スーパーバイザは報酬と遷移確率に関する情報を強化学習により学習する。遷移確率の情報から、各状態において、制御仕様を外れる確率を計算し、リスク制約を満たす取り得る制御パターンの集合を計算する。集合に含まれる各制御パターンの評価値を計算し、次に選択する制御パターンを決定する。これにより、定量的にリスクを保証した上で最適なスーパーバイザを構成するという本研究の基本的な枠組みを示すことができた。

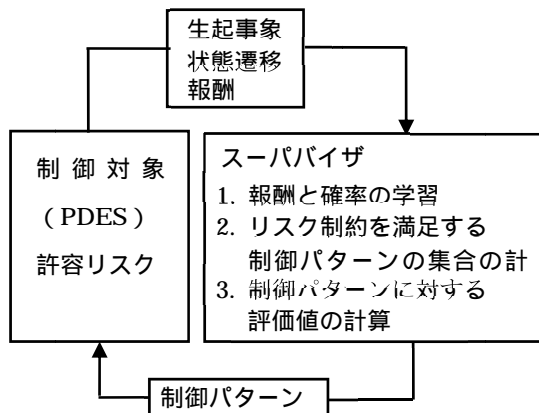


図2 リスク許容型最適スーパーバイザ制御

タスクを実行する複数台のロボットをモデルとした計算機実験による評価により、報酬と遷移確率の情報が未知の場合においても、学習により、全ての情報が既知の場合とした最適な制御パターンを選択できるようになることを確認した。実験においては、許容するリスクが大きくなるほど、評価値の高い制御パターンを選択する余地も出てくることを確認できた。報酬情報に比べて、確率情報の学習は遅くなる傾向であり、学習の高速化が課題としてある。

大規模複雑化する離散事象システムへの対応として得られた最適LLPスーパーバイザ制御の枠組みを図3に示す。与えられた制御パターンの中から、LDESのいずれか、または協働により生じた事象と、状態遷移、報酬の観測に基づき、各LDESのモデル情報を学習し、実時間制約の下での計算により、システム全体の先読みツリーを構成する。与えられた制御仕様において、各ノードで許容される制御パターンの集合を求め、学習した情報が

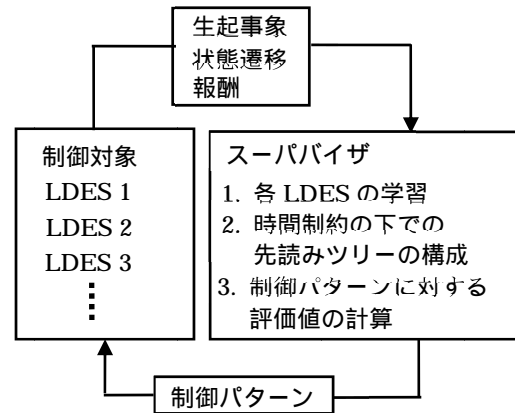


図3 最適LLPスーパーバイザ制御

ら、報酬に基づく各制御パターンの評価値を計算することで、現在の状態において最適な制御パターンを求める。なお、有限ステップの先読みのため、葉ノードの扱いについて、評価値をどう見積もるか、仕様に対して保守的・楽観的な方針をとるかいくつかの方針が考えられる。本手法をリスク許容型最適スーパーバイザ制御の枠組みと併用することで、大規模で時変な離散事象システムにたいしても、リスク制約を満たした上で最適なスーパーバイザを構成することが可能となる。

タスクを遂行する複数のロボット(LDES)が存在し、それらが増減するシステムに対し、計算機実験での評価により、LDESの情報全てが既知であるとして求められる最適な制御パターンを学習により選択できるようになることが確認できた。また、適切な深さまで先読みすることにより、制御仕様を満たすことが可能であることが確認できた。また、実験として、LEGO Mindstormsを用いて図4に示す、複数のサブシステムから構成される生産システムミニチュアを作成し、その最適スーパーバイザ制御を行った。学習により、加工や移動の時間などから得られる報酬から、より評価の高い制御パターンを選択するようになっていった。しかしながら、ハードウェア的な通信の安定性もあり、複雑度の高い場合の評価については今後さらなる実装が必要である。

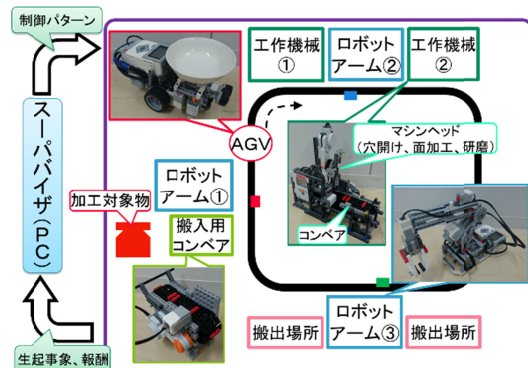


図4 生産システムミニチュアの概要

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 3 件)

梅本聖, 山崎達志, “状態遷移モデルの学習に基づく最適 LLP スーパーバイザの構成,” 電子情報通信学会技術研究報告 (MSS), Vol.114, No.493, pp.7-12, 2015.(査読無)

梅本聖, 山崎達志, “強化学習を用いた離散事象システムの最適 LLP スーパーバイザ制御について,” 電子情報通信学会技術研究報告 (MSS), Vol.114, No.313, pp.135-140, 2014.(査読無)

山崎達志, 入江太志, 松原伸, “離散事象システムのリスク考慮型最適スーパーバイザ制御,” 電子情報通信学会技術研究報告 (MSS), Vol.113, No.466, pp.1-6, 2014.(査読無)

[学会発表](計 4 件)

梅本聖, 山崎達志, “最適 LLP スーパーバイザによる生産システムミニチュアの制御,” 電子情報通信学会 2016 年総合大会, 2016-3-16, 福岡.

梅本聖, 山崎達志, “リスク許容型最適スーパーバイザの強化学習,” 計測自動制御学会システム・制御部門学術講演会 2015, 2015-11-20, 函館.

Hijiri Umemoto and Tatsushi Yamasaki, “Optimal LLP supervisor for discrete event systems based on reinforcement learning,” 2015 IEEE International Conference on Systems, Man, and Cybernetics, 2015-10-10, 香港.

梅本聖, 山崎達志, “離散事象システムのリスク許容型最適スーパーバイザ制御の検討,” 電子情報通信学会 2015 年ソサイエティ大会, 2015-9-9, 仙台.

6. 研究組織

(1)研究代表者

山崎 達志 (YAMASAKI, Tatsushi)
摂南大学・理工学部機械工学科・准教授
研究者番号: 00368458

(2)研究協力者

梅本 聖 (UMEMOTO, Hijiri)