

令和元年5月17日現在

機関番号：17102

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K00304

研究課題名(和文) 推論・学習機能を備えた実験候補推薦システムの研究開発

研究課題名(英文) Research and Development of Experiment Candidate Recommendation System with Reasoning and Learning

研究代表者

越村 三幸 (Koshimura, Miyuki)

九州大学・システム情報科学研究所・助教

研究者番号：30274492

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：機械学習技術を用いた「尿の顕微鏡画像の解析」と「電流データに基づく異常診断」に取り組んだ。前者については、病院で採取された100人の尿の顕微鏡写真から約14000個の細胞の画像を切り抜きその分類を行い、正答率98.30%のモデルの作成に成功した。また、SATによる最適化問題の解法の事例として「提携構造形成問題」と「ILP(帰納論理プログラミング)のMaxSAT符号化」に取り組んだ。前者については従来の手法に比べ一桁以上の性能向上を達成した。

研究成果の学術的意義や社会的意義

本研究は、目標性能達成のため、これまでの実験データを基に次に試行すべき実験候補を推薦するシステムの開発を目指した。統計的手法に論理的手法を組み合わせることにより、探索や説明機能を強化する。ILPをSAT符号化して高速化を図る点も研究の特色である。また、多くの問題に現れる推移律の効率的なSAT符号化法を考案しその効果を実証した点は、組合せ最適化解法の適用分野を広げる成果である。

研究成果の概要(英文)：We engaged in "analyzing microscopic image of urine formed element" and "abnormal diagnosis based on current data" with techniques of machine learning. We cut out images of approximately 14,000 cells from photomicrographs of 100 urine samples collected at a hospital and classified them into red blood cells, white blood cells, and others. We succeeded in acquiring a model with a correct answer rate of 98.30%. We also worked on MaxSAT encoding of coalition structure generation (CSG) problem and ILP (inductive logic programming) as case studies of problem solving with SAT technology. And we succeeded in achieving an order of magnitude performance improvement than a previous method for CSG.

研究分野：知能情報学

キーワード：機械学習応用 組合せ最適化 SAT

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

資源およびエネルギー消費を抑制する技術は世界中に普及しつつあるが、新興国需要の増加等を背景に世界の資源およびエネルギー消費は一貫して増加し深刻な問題になってきている。そのため、新素材や新システムの開発においては、少ない実験回数で効率的かつ短期間に開発するための手法の開発と導入が切望されている。しかしながら、研究開発は過去の実験データベースを基に研究員の経験と勘により行われているのが現状である。以前から実験計画法、多変量解析等の導入も試みられているが、これらは化学技術者にとって馴染みが薄く、有効に活用されていない。

上記課題を解決するシーズ候補は、自動推論と学習の二つである。我々は、MGTP と称する独自のモデル生成型定理証明システムを核として、自動推論に関する様々な技術を研究開発してきた。中でも命題論理の充足性可能性判定 (SAT)、準命題論理の定理証明 (EPR)、帰納論理プログラミング (ILP) に関して得た成果については、多くの分野で効果的な応用が期待できる。また、特に産業界で重要とされる組合せ最適化問題の解法に関して、我々は種々の進化型アルゴリズムを効果的に組み合わせた独自のハイブリッド探索手法 HPTS を開発した。

目指す実験候補推薦システムを実現するには、(1) 予算および開発時間が限られた開発現場では、小規模データしか入手できない、(2) 回帰分析等の結果の解釈・読み方は化学技術者には難しい、といった問題を解決しなければならない。少ない標本数で最適解を予測するには、大規模データを前提とする従来の統計的学習手法では十分な精度が得られず、実験者の求める解から乖離する可能性が高い。そこで、小規模データを補間してデータ数を増し学習の精度を高める手法を導入すると共に、統計的学習手法の代替として、記号論理的学習手法である ILP の適用を図ることとした。ILP は、規則の形で表現した結果を返すので、直観的であり化学技術者にとっても分かり易い。しかし、予備実験の結果、(3) パラメータ数の増大に伴い ILP の処理時間も膨大となる、ことが判明した。この解決法として、ILP で行われる探索を SAT 符号化して SAT ソルバーで解かせるアプローチをとる。

我々は、帰納推論の技法が、新素材の開発実験においても有効に適用できることを、既に幾つかの事例でプロトタイプを試作して確認した。その成果は、現場の実験技術者からも高い評価を受けている。今後は、より精密な予測や分析を行うために探索や学習に関する機能強化と推論性能の向上を図り、真に使い勝手のよい実験候補推薦システムを構築することが望まれる。

2. 研究の目的

新素材の開発では、実験データを基に最適組成を得ることが求められる。実験にはコストおよび時間がかかる為、目標の性能を満たす素材を生成するパラメータの組合せを、少ない実験回数で予測することが極めて重要である。我々は MGTP や QMaxSAT のような効率的な自動推論システム、並びに帰納論理プログラミング等の学習手法に関する研究蓄積を有する。本研究では、データマイニングで用いられる方法の内、帰納論理プログラミング、決定木、および多目的最適化の手法を用いて、目標条件を満たすパラメータの推定を行い、これに基づいて次に試すべき実験パラメータ候補を提示する、**推論・学習機能を備えた実験候補推薦システム**を実現することを目的とする。

3. 研究の方法

研究開発テーマを、(A) 機械学習技術の適用、(B) 帰納論理プログラミング (ILP) ツール、(C) SAT を用いた最適化問題の解法、の3つに定め、初年度に、機械学習技術を適用する具体的なデータの選定を行い、尿の顕微鏡画像と電気機器を流れる電流の時系列データの2種類のデータの解析を行うことにした。前者は尿中の有形成分 (赤血球や白血球など) の計数、後者は機器の異常診断が目的である。

研究方法として、研究代表者と研究分担者の3名の下に大学院生3名の研究協力者を適宜割当て、研究を遂行した。また、尿の顕微鏡画像は病院の検査で実際に採取されたものを東洋紡株式会社から、電流の時系列データは株式会社スカイディスクから提供を受けた。そして解析にあたっては、それぞれの会社の専門家の協力を得た。

力研究分担者とは数ヶ月に1回程度、九州大学あるいは徳山工業高等専門学校において、成果発表・研究交流・情報交換を行った。

4. 研究成果

(A) 機械学習技術の適用

(ア) 尿の顕微鏡画像の解析: 先ず、分類問題に帰着、回帰問題に帰着、する2種類の手法について、いずれも深層学習の技法を用いて取り組んだ。前者は帰着する際に従来の機

械学習技術の援用が必要になるが、後者はそのようなものは不要である。実際の顕微鏡画像を用いた実験を行なった結果、前者の解析精度は実用的なレベルに近かったが、後者のレベルは低かった。次いで、100人の尿の顕微鏡写真から14033枚の有形成分の画像を切り抜き、深層学習を用いて分類を行なった。ネットワークは、CNN(Convolution Neural Network)であるLeNetを参考に、全層畳み込みネットワークをベースにしたものを用い、正答率97.26%のモデルを得た。さらにハイパーパラメータ探索を行い、正答率98.30%に高めることができた。

- (イ) 電流データに基づく異常診断：サポートベクタマシン(SVM)、決定木学習、ランダムフォレスト、k近傍法の4種類の分類器を用いた実験の結果、分類精度はSVMが最も高かった。なお、実験には異常振動発生装置を用いた。
- (B) ILPツール：ILPのMaxSAT符号化の研究を進めた。データログと呼ばれるILPの部分クラスを符号化の対象とし、ILPシステムで広く利用されている探索法の符号化を目指した。前処理で最弱仮説の生成を行い、候補仮説内探索をMaxSAT符号化した。正例はソフト節、負例はハード節に符号化される。1回のMaxSAT呼び出しで1つのルールが生成され、全ての正例を被覆するまでMaxSAT呼び出しが繰り返される。化学実験データの解析を例にシステムの調整を行い、一般的なILPツールとほぼ同等の性能が得られた。
- (C) SATを用いた最適化問題の解法：
 - (ア) 分割決定木で記述された提携構造形成(CSG)問題をSATを用いて解く手法と分枝限定法を用いる手法の比較を行い、負の効用がある問題ではSAT手法が格段に効率的であることを明らかにした。
 - (イ) MC-netsで記述されたCSG問題をSATで解く際の推移律の符号化の制約数を減らす手法の提案と評価を行い、従来手法と比べて1桁以上の性能向上を定量的に確かめた。
 - (ウ) 擬似ブール制約の新しいSAT符号化の提案と評価を行い、多くのベンチマークを用いて効果を確認した。

5. 主な発表論文等

[雑誌論文](計5件)

- 藤田 博, 越村 三幸. SATとラムゼー数 ~数学の未解決問題への挑戦~, 査読有, 情報処理, Vol.57, No.8, pp.716-719, 2016年
- 越村 三幸, 藤田 博. MaxSAT: SATの最適化問題への拡張 -MaxSATソルバーの活用法, 査読有, 情報処理, Vol.57, No.8, pp.730-733, 2016年
- Xiaojuan Liao, Miyuki Koshimura, Kazuki Nomoto, Suguru Ueda, Yuko Sakurai, Makoto Yokoo. Improved WPM Encoding for Coalition Structure Generation under MC-nets, 査読有, Constraints, Vol.24, Issue 1, pp.25-55, 2019.
- 査 澳龍, 越村 三幸, 櫻井 祐子, 横尾 真. 分割決定木を用いた分割関数ゲームの提携構造形成アルゴリズム, 査読有, 電子情報通信学会論文誌 D, Vol.J102-D, No.4, pp.313-323, 2019年
- Aolong Zha, Miyuki Koshimura, Hiroshi Fujita. N-level Modulo-Based CNF encodings of Pseudo-Boolean constraints for MaxSAT, 査読有, Constraints, Vol.24, Issue 2, pp.133-161, 2019.

[学会発表](計15件)

- 越村 三幸. モデルとUnsat Coreを利用したMaxSATソルバーの試作, 2016年度人工知能学会全国大会(第30回), 2016年
- Miyuki Koshimura, Kazuki Nomoto, Yuko Sakurai, and Makoto Yokoo. MaxSAT Encoding for MC-net/PDT-based Coalition Structure Generation Problem, 10th CSPSAT Seminar, 2016.
- 上村 直輝, 藤田 博, 越村 三幸, 査 澳龍. 混合基数を用いた擬似ブール制約のSAT符号化, 人工知能学会 第103回人工知能基本問題研究, 2017年
- 越村 三幸, 査 澳龍, 野本 一貴, 櫻井 祐子, 横尾 真. 分割決定木で表現された提携構造形成問題のMaxSAT符号化, 情報処理学会第79回全国大会, 2017年
- 越村 三幸, 査 澳龍, 野本 一貴, 櫻井 祐子, 横尾 真. 分割関数ゲームを対象とした提携構造形成問題のMaxSAT符号化, 人工知能学会全国大会(第31回), 2017年
- 越村 三幸, 廖 曉鵬, 野本 一貴, 上田 俊, 櫻井 祐子, 横尾 真. MC-netsを用いた提携構造形成問題のMaxSAT符号化の改良, 日本ソフトウェア科学会 第34回大会, 2017年
- Aolong Zha, Kazuki Nomoto, Suguru Ueda, Miyuki Koshimura, Yuko Sakurai, Makoto Yokoo. Coalition Structure Generation for Partition Function Games Utilizing a Concise Graphical Representation, 査読有, Proc. of PRIMA 2017, pp.143-159, 2017.
- Aolong Zha, Naoki Uemura, Miyuki Koshimura, and Hiroshi Fujita. Mixed Radix Weight Totalizer Encoding for Pseudo-Boolean Constraints, 査読有, Proc. of 29th ICTAI, pp.868-

875, 2017.

Aolong Zha, Miyuki Koshimura, and Hiroshi Fujita. A Hybrid Encoding of Pseudo-Boolean Constraints into CNF, 査読有, Proc. of TAAI, pp. 9-12, 2017.

越村 三幸, 廖 曉鵬, 野本 一貴, 上田 俊, 櫻井 祐子, 横尾 真. MC-nets を用いた提携構造形成問題の MaxSAT 符号化の改良と評価, 人工知能学会 第 106 回人工知能基本問題研究会, 2018 年

満尾 成亮, 安部 竜太, 橋下 司, 越村 三幸, 藤田 博. 機械学習を用いた電流データに基づく異常判定システム, 電子情報通信学会 2018 年総合大会, 2018 年

浅倉 健太, 越村 三幸, 藤田 博. CNN を用いた尿中有形成成分の分類の初期検討, 電子情報通信学会 医用画像研究会, 2018 年

Miyuki Koshimura, Aolong Zha, Kazuki Nomoto, Suguru Ueda, Yuko Sakurai, Makoto Yokoo. Coalition Structure Generation for Partition Function Games Utilizing a Concise Graphical Representation, The 4th Kakenhi Kiban-A&B/NII Collaborate Research Meeting, 2019 年

Xiaojuan Liao, Miyuki Koshimura, Kazuki Nomoto, Suguru Ueda, Yuko Sakurai, Makoto Yokoo. Improved WPM Encoding for Coalition Structure Generation under MC-nets, The 4th Kakenhi Kiban-A&B/NII Collaborate Research Meeting, 2019 年

越村 三幸, 佐藤健. SAT ソルバーを用いた MCS 列挙の実装とその評価, 人工知能学会 第 109 回人工知能基本問題研究会, 2019 年

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

名称:
発明者:
権利者:
種類:
番号:
出願年:
国内外の別:

取得状況(計 0 件)

名称:
発明者:
権利者:
種類:
番号:
取得年:
国内外の別:

〔その他〕

ホームページ等

<https://sites.google.com/site/qmaxsat/>

6. 研究組織

(1) 研究分担者

研究分担者氏名: 藤田 博

ローマ字氏名: Fujita Hiroshi

所属研究機関名: 九州大学

部局名: システム情報科学研究院

職名: 准教授

研究者番号(8桁): 70284552

研究分担者氏名: 力 規晃

ローマ字氏名: Chikara Noriaki

所属研究機関名：徳山工業高等学校
部局名：情報電子工学科
職名：助教
研究者番号(8桁)：50290804

(2)研究協力者

研究協力者氏名：上村 直輝
ローマ字氏名：Uemura Naoki

研究協力者氏名：查 澳龍
ローマ字氏名：Zha Aolong

研究協力者氏名：浅倉 健太
ローマ字氏名：Asakura Kenta

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。