

機械学習活用のための要求分析手法の研究 AI版酒屋倉庫問題のケース

岡野道太郎^{1,a)} 中谷 多哉子^{2,b)}

概要：機械学習を活用したシステムを構築する場合、機械学習部分の開発工程とそれ以外の部分の開発工程では相違点が見られる。本稿ではこれら相違点を述べた上で、機械学習部分とそれ以外の部分の両方が存在するシステムの要求抽出における課題を挙げる。その上で開発工程中、特に要求分析に焦点をあて、課題の解決に向けた要求分析手法を提案し、その手法を「機械学習を活用した酒屋倉庫問題」に適用し、考察する。さらに今後の研究課題について議論する。

Towards a method of the requirements analysis of integrated machine learning into the system A case study of the intelligent liquor store warehouse

1. はじめに

1.1 背景

システム開発には要求分析、設計、実装、テストの各開発工程があり、要求分析において要求が抽出される。そして要求を抽出する一手法として、ゴール指向要求分析のKAOS[1]がある。KAOSでは、ゴールを詳細化して要求を抽出する。つまり、システム開発の全体目標を、最終的には1人あるいは1システム(1エージェントと以下記す)が達成すべき状態まで分解することにより要求を抽出する。ゴールの詳細化手法としては、後述するマイルストーン分解等いくつかの手法をLamsweerdeが挙げている[1]。しかし、各ゴールに対して、どのような詳細化の手法があり、その詳細化の手法はどの程度有効か等については明らかにされていない。

筆者らは、全体目標が達成した状態をQとし、現在の状態をPとしたとき、「現在の状態Pから全体目標が達成した状態Qへ、いつかは状態遷移する」というゴールを設定し([1]ではこのゴールを「達成ゴール」と呼び、 $P \rightarrow \diamond Q$

で表現される)、このゴールをマイルストーン分解と後述する要素分解という2種類の詳細化手法のみを用いて、要求抽出が可能であるかどうかを議論している[2]。その要求抽出手法は、具体的には以下のとおりである。

- (1) 「開発するシステムが達成すべき状態Q」を目標として挙げ、「初期状態Pから目標が達成された状態Qにまで状態遷移すること」(これを $P \rightarrow \diamond Q$ と記す)をトップゴールとして挙げる。
- (2) $P \rightarrow \diamond Q$ を詳細化する。詳細化の手法としては、2種類ある。1つは状態遷移(\rightarrow)に着目し、 $P \rightarrow \diamond Q$ へ状態が遷移する間に、ある状態Mを経るのであれば($P \rightarrow M \rightarrow Q$)、 $P \rightarrow \diamond M$ と $M \rightarrow \diamond Q$ にゴール分解するマイルストーン分解[1]である。もう一つは、目的が達成した状態Qが、Qの構成要素 $\{Q_1, Q_2 \dots Q_n\}$ に分解できる場合、ゴールを「 $P \rightarrow \diamond Q_1$ 」, 「 $P \rightarrow \diamond Q_2$ 」... 「 $P \rightarrow \diamond Q_n$ 」, 「 $\{Q_1 \dots Q_n\} \rightarrow \diamond Q$ 」に分解する要素分解[2]である。
- (3) 詳細化された各ゴールに対しても、詳細化を行う。このゴールに対する詳細化を繰り返す。
- (4) 最終的に1エージェントが達成可能なゴールにまで分解されたとき、末端のゴールを要求とする。

コンピューターシステムに対する要求の場合、1エージェントとは、処理を行うコンピューターになる。よって、ゴールがコンピューターで処理可能なタスク、たとえ

¹ 筑波大学大学院ビジネス科学研究科
3-29-1 Otsuka, Bunkyo, Tokyo 112-0012, Japan

² 放送大学 情報コース
2-11 wakaba, Mihama, Chiba, Chiba 261-8586, Japan

^{a)} okano@gssm.otsuka.tsukuba.ac.jp

^{b)} tinakatani@ouj.ac.jp

ば「ある帳票が出力されている」、「ある画面からデータが入力されている」等にまで分解されたとき、そのゴールは要求となる。機械学習では「予測」、「分類」、「判別」等のタスクが行えるため、「ある値を予測できている」、「分類できている」、「判別できている」等のゴールに分解された場合、要求となる。KAOSによってゴールを詳細化した例を図1に掲げる。

図1は「予測する」という機械学習部分を含んでいるので、前述の要求抽出手法を、機械学習を含むシステムに適用可能なように見える。しかし、前述の手法は、達成ゴール $P \rightarrow \diamond Q$ を詳細化している。達成ゴールの場合、目標とする状態 Q がいつかは起こることとしている。しかし、機械学習を用いた場合、目標とする状態 Q が必ず起こるとは言い切れない。例えば、予測を行う場合、過去のデータを学習して、未来を予測できるとは必ずしも言い切れず、 $P \rightarrow \diamond Q$ の Q (例では「予測できている」) は永遠に成立しないかもしれない。このような場合が考えられるため、機械学習を含む場合は、前述の手法をそのままでは適用できない。そして図1は、適用できないことが指摘できていない。本稿では、適用できないことを指摘可能な図の記法を提案し、前述の要件抽出手法を、機械学習を含むシステムにも適用可能になるように拡張することを目指す。そのためにも、機械学習の開発工程を述べ、機械学習を行うシステムと行わないシステムの相違点について述べる。なお本稿において機械学習は教師あり学習の場合のみを取り上げる。

1.2 機械学習システムの開発方法

機械学習を行うシステムを開発する場合、要求分析段階で、機械学習を行って求める変数（以下「目的変数」と記す）を決定する。また目的変数の値を決定するのに必要な変数（以下「説明変数」と述べる）の候補も挙げる。その後、機械学習を行うプログラムを用意し、データを収集・加工して機械学習を実施する。機械学習を実施した結果、期待した精度が得られた場合、その学習結果（以下「学習済みモデル」と記す）を用いて実運用する。すなわち、運用時に収集したデータを学習済みモデルに適用し、結果を取得する（以下、このことを「推論する」と記す）。機械学習を行っても期待した精度が得られない場合は、機械学習の手法（以下「モデル」と記す）や説明変数を変更し、期待する精度になるまで試行を繰り返す。ただし、収集したデータをもとに、期待する精度に達した学習モデルが得られる保証はない。また、環境変化が生じ、推論した結果が求める精度に達しなくなった場合等は、新たなデータを用いて再度機械学習するという開発工程を経る。

この開発工程の場合、要求分析段階で目的変数項目は決定するが、学習済みモデルができるまでは、説明変数項目やモデルは確定しない。さらに、収集データを用いて十分

な精度の学習済みモデルが得られる保証はない。つまり学習モデルを用いて目的変数の値を求められるかどうかは不確かである。機械学習を用いない場合は、この不確かさはなく、要求分析で、ある変数の値を求めるように要求された場合、設計段階でその変数の値を求める手順が確定し、実装可能になるという相違点がある。そして図1で、従来手法が適用できないことを指摘できなかったのは、従来手法は項目に基づいて作成していただけで、不確かさについての配慮がなかったからである。

1.3 本稿での課題

実際の機械学習を含むシステムには、予測等の機械学習を活用する部分（以下「ML活用部分」と記す）と、帳票出力や画面表示等、機械学習を行わない部分（以下「非ML部分」と記述する）がある。これらの部分は連携している（連携箇所を以下「連携部分」と記す）。ここで、ML活用部分のゴールを $P \rightarrow \diamond M$ 、その M を利用する非ML部分のゴールを $M \rightarrow \diamond Q$ とすると、連携部分は状態 M となる。

「1.2 機械学習システムの開発方法」の議論から、機械学習を含むシステムで問題となるのは、 P が確定しないこと、ML活用部分の M が必ずしも成立しないこと、 M が成立しないと非ML部分の $M \rightarrow \diamond Q$ が実行できないことが挙げられる。そして、このことを考察するためには、ML活用部分 $P \rightarrow \diamond M$ の M と非ML部分 $M \rightarrow \diamond Q$ の M が別々に記述でき、この間の関係を議論できる記法が必要である。このことと、「1.2 機械学習システムの開発方法」で議論した結果から「適用できないことを指摘可能な図の記法」は、少なくとも以下のことが表現可能でなければならない。

- (1) (ML活用：入出力) ML活用部分では、ゴール達成時の出力である目的変数は示すが、そのゴールを達成する為に必要な入力情報である説明変数は必ずしも明示しないでよいという記述ができなければならない。
- (2) (ML活用：状態遷移) ML活用部分では、アルゴリズム(モデル)は必ずしも明示しなくてもよい記述ができなければならない(ただし、詳細化は可能にする)。
- (3) (非ML：入出力) 非ML部分は、入出力を表現でき、設計者等の第三者に明示できなければならない。
- (4) (非ML：状態遷移) 非ML部分は、KAOSと同様、詳細化を段階的に表現できなければならない。
- (5) (連携) ML活用部分と非ML部分を連携させるため、ML活用部分の出力と非ML部分の入力は、別々に明示できなければならない。

本稿ではまず、これらの条件を満たす記法を提案することを課題とする。そのうえで、実際にゴール分解を行い、不確かさを図で表現するための考察や、従来手法を拡張して、不確かさに対応させる方法について考察する。

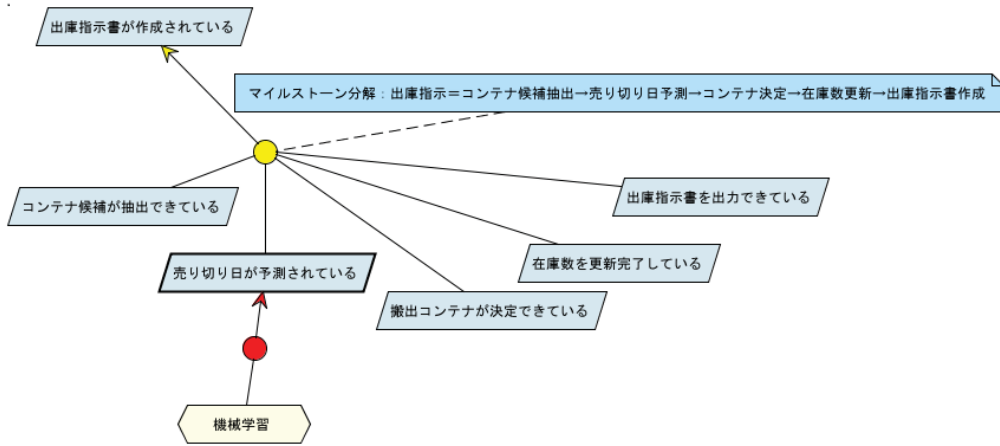


図 1 KAOS の例

Fig. 1 The example of KAOS.

2. 提案手法

2.1 提案へのアプローチ

本稿では課題を解決するために、KAOSで行った詳細化をGSN[3]の記法で表現するというアプローチを採用する。GSNは、保証のための構造化された議論の記述法であり、保証したいことをゴールとしてあげ、そのゴールをサブゴールに分解していく。分解するときの観点が戦略である。また議論する際の環境等の条件を前提として挙げる。サブゴールの分解を繰り返し、末端のゴールをエビデンスにより成立することを保証する。

GSNは図示することができる。図2はGSNの例であり、長方形がゴール、角丸四角形が前提、円がエビデンス、平行四辺形が戦略である。

2.2 提案手法

本稿では、KAOSのゴールを、以下のようにGSNで表現する。図1のKAOSを、以下の表現で書き表したものが図2である。

- 従来手法のKAOSのゴールはGSNのゴールとして表現する
- ゴールを入力をゴールに接続する前提の中に記述する
- ゴールをサブゴールに分解する観点、すなわちマイルストーン分解か要素分解かを、戦略に記述する
- 末端のゴールが達成されるに伴い生成される出力を、エビデンスに記述する。
- 非ML部分の前提(入力)は必須であるが、ML活用部分の前提は必須ではない(書いても良い)

上記提案が、「1.3 本稿の課題」で挙げた条件に適合しているか確認する。KAOSのゴールをGSNのゴールにするため、条件(4)は満たしている。また、従来手法のKAOSでは、ゴールに「目標となる結果の状態」を記述する。よって提案手法のGSNのゴールにも「結果の状態」である出

力がかかれ、そのゴールの前提に入力がかかれるので、条件(3)も満たしている。また前提は必須ではないことから条件(1)も満たしている。ゴールも前提も状態であり、アルゴリズムを示しているわけではないので(2)も満たしている。そして、非ML部分の入力は、非ML部分のゴールの「前提」に示され、ML活用部分の出力はML活用部分のゴールとエビデンスに書かれる。このように別々に明示されているので条件(5)は満たしている。そして、ゴールには不確かさは書かれない。そこで、ゴールと同等のことが書かれるエビデンスに不確かさを記述することを検討する。具体的な書き方は「考察」で議論するが、その議論を行うために、提案手法を用いた事例を示す。

3. 事例

本章では提案手法に対し「酒屋倉庫問題に機械学習を活用した事例」(以下「AI版酒屋倉庫問題」と記す)を示す。まずAI版酒屋倉庫問題について説明する。

3.1 AI版酒屋倉庫問題について

酒屋倉庫問題とは、情報処理[4]に掲載された、ある酒類販売会社(以下酒屋と記す)の受付係の仕事プログラミングする問題である。受付係の仕事とは、「在庫無し連絡」、「出庫指示書作成」、「在庫不足リスト作成」である。出庫指示書を作成するためには、倉庫から酒を出庫しなければならないが、この問題文中には、受付係が、どのような基準(先入先出等)で出庫を行うかは記載されていない。ただし、「倉庫内のコンテナ数はできる限り最小にしたいと考えている」との記載がある。そこで、出庫のゴールは「倉庫内のコンテナ数を出来るだけ最小になるように出庫する」と考えられる。本稿では、このゴールを満たすために、機械学習を採用したシステムを「AI版酒屋倉庫問題」と定義する。

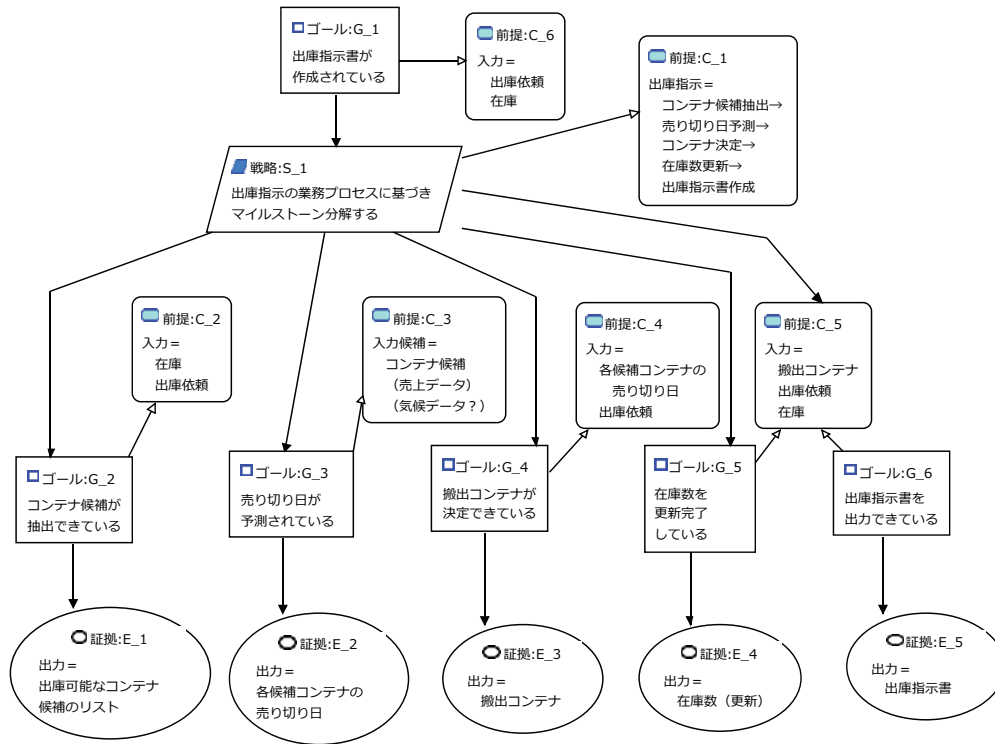


図 2 GSN の例

Fig. 2 The example of GSN.

3.2 提案手法の適用

AI 版酒屋倉庫問題において、機械学習が適用できるのは「出庫指示書作成」に関することのみなので、この部分のみを詳細化する。詳細化した図は、すでに図 2 に記述している。

4. 考察

4.1 不確かさの表現について

不確かさが問題となるのは、ML 活用部分の出力部分である「証拠 E_2」とその出力を入力とする非 ML 部分の前提「前提 C_4」である。ここは「各候補コンテナの売り切り日」という同じ言葉が入っている。非 ML 部分が ML 活用部分の出力精度にかかわらず、常に出力結果を入力とする場合には問題は生じないが、非 ML 部分の入力精度に制約がある場合は、非 ML 部分の前提に、求める精度を記述する。この記述がある場合、精度を満たせるかどうかは不確かとなる。

4.2 不確かさへの対応

上述のように入力精度に制約がある場合、機械学習結果がその制約を満たさないかもしれない。その場合、入力精度を満たすその他の手段があればそれも記述する。機械学習およびその他の手段を実行した結果、入力精度を必ず満たす場合は、ML 活用部分と非 ML 部分は連携できるとみなせる。

4.3 今後の研究

本稿では、AI 版酒屋倉庫問題を取り上げたが、ほかの機械学習についても当てはまるか、機械学習プロセスが 2 か所以上含むシステムの場合どのようなようになるかは今後の研究となる。

また、本稿では酒屋倉庫問題に基づいた予測業務を取り扱ったが、本稿の提案手法は業務を特定していない。ということは、提案手法を「要求分析業務」に適用した場合、機械学習を活用した要求分析支援システムが開発できるはずである。これについては今後の研究となる。

参考文献

- [1] Axel van Lamsweerde, "Requirements Engineering: From System Goals to UML Models to Software Specifications", Wiley, 2009.
- [2] 岡野道太郎, 中谷多哉子: ゴール指向分析 KAOS における依存性を考慮した要求抽出法の考察-酒屋倉庫問題の場合-, ソフトウェア・シンポジウム 2017
- [3] Origin Consulting (ork) Limited, on behalf of the Contributors "GSN COMMUNITY STANDARD VERSION 1" (online), http://www.goalstructuringnotation.info/documents/GSN_Standard.pdf, 2011.(2017/8/2)
- [4] 山崎利治, 共通問題によるプログラム設計技法解説, 情報処理 Vol25 No9(1984).