

## 過去問題をリソースとした知識ベースの活用可能性の検証

福坂 祥基<sup>†</sup> 高木正則<sup>†</sup> 山田 敬三<sup>†</sup> 佐々木 淳<sup>†</sup>

岩手県立大学 ソフトウェア情報学部<sup>†</sup>

### 1.はじめに

近年、TOEIC や情報処理技術者試験など様々な検定試験が行われている。これらの検定試験では、出題する問題の作成に多くの負担がかかっている。一方で、教育現場で必要となる多様な演習問題が不足していることに対し、演習問題を自動生成する研究が多数存在する<sup>(1)(2)</sup>。しかし、問題を生成するためのリソース情報に XML のタグを付与することや、出題する知識のカテゴリに対応した語尾を定義するなど、問題生成の事前準備に何らかの作業が必要であり、作問者にかかる負担は未だ大きい。これに対し古舘ら<sup>(3)</sup>は、問題のリソースとなる知識ベースを定義し、この知識ベースを自動構築する手法を提案している。これにより、問題の自動生成に要する負担軽減が期待される。しかし、この知識ベースを活用して生成される問題の活用可能性は明らかにされていない。そこで、本研究では、先行研究で提案した知識ベースの構造の妥当性や活用可能性の検証を目的として、手動で知識ベースを構築したうえで、問題自動生成機能を開発し、検定試験の作問委員会に利用してもらった。本稿では、本機能の利用結果から知識ベースと本機能の有効性を評価する。

### 2.生成する問題の形式

本研究では、問題文と解答を自動生成して一問一答形式の問題を生成する。多肢選択形式の問題を生成する場合は、誤答選択肢を自動生成するシステム<sup>(4)</sup>と連携して多肢選択形式に変換する。また、本研究では共同研究を行っている盛岡商工会議所が主催するご当地検定の盛岡もの識り検定（以下、もりけん）で出題される盛岡市の地理・歴史・人物などの固有名詞の知識を問う問題を対象とする。

### 3.作問支援システム

#### 3.1. システムの概要

我々は入力されたキーワードに関連する問題を自動生成し、これを修正または参考にして作問者が新規問題を作成できる作問支援システムを開発した。本機能はもりけんの過去問題をリソースとした知識ベースと、出題テンプレートを用いて問題を自動生成する。

#### 3.2. 知識ベースの概要

もりけんの 2000 問の過去問題をリソースに知識ベースを手動で構築した。表 1 に過去問題の例を示し、図 1 に知識ベースの一部を示す。本研究で扱う知識ベースは、カテゴリの子ノード以降のノードを

Application of Knowledge Base Using Exam Questions as Resources

<sup>†</sup>Shoki FUKUSAKA, Masanori TAKAGI, Keizo YAMADA, Jun SASAKI, Faculty of Software and Information Science, Iwate Prefectural University

表 1 過去問題の例

問題文	解答	対象知識
新渡戸稲造が設立時に事務次長をつとめた国際組織の名称は何ですか。	国際連盟	新渡戸稲造
新渡戸稲造の後に、国際連盟事務次長になった、盛岡出身の外交官は誰ですか。	杉村陽太郎	杉村陽太郎
今年、盛岡駅前滝の広場で胸像が除幕された先人は誰ですか。	新渡戸稲造	新渡戸稲造
平成 24 年、生誕 150 年を迎える盛岡ゆかりの人物は誰ですか。	新渡戸稲造	新渡戸稲造

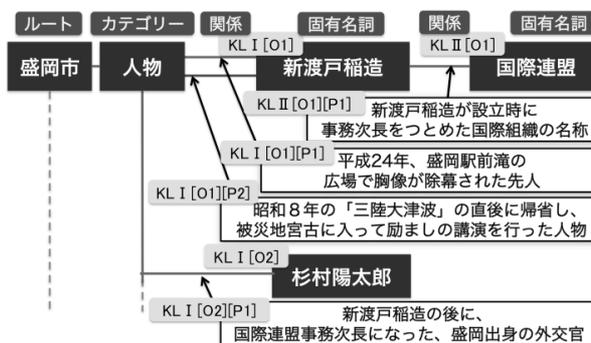


図 1 知識ベースの一部

オブジェクト、ノード間に紐付けられている関係をプロパティと定義する。以下に知識ベースの構築手順を示す。

- (1) 対象となる検定試験で扱う知識の最上位の概念をルートに置く。今回は盛岡市とする。
- (2) 各問題で問われている知識を対象知識（表 1）として定義し、この知識が属するカテゴリを決定する。表 1 の問題はすべて「人物」カテゴリとなるが、この他 7 つのカテゴリがある。
- (3) カテゴリの子ノードに対象知識を配置する。また、対象知識と解答が同じ問題の場合は、このノード間に問題文を紐付ける。対象知識と解答が異なる問題の場合は、対象知識の子ノードに解答を配置し、このノード間に問題文を紐付ける。

#### 3.3. 出題テンプレートの概要

出題テンプレートを作成するために、もりけんの過去問題 250 問の出題形式や知識の出現箇所を分析した。その結果、画像を利用する問題や、穴埋め問題などの 7 つの形式の問題が存在することが分かった。これらの中で、多くの割合を占めていた 2 つの形式（問題文の条件に適合する単語を選ぶ形式と、問題文に示された知識に適した説明を選ぶ形式）を出題テンプレートのモデルとした。出題テンプレートの一覧を表 2 に示す。

表2 出題テンプレート

No.	問題文	解答
①	KL I [O][Pi] ( $1 \leq i \leq n$ )であり, KL I [O][Pj] ( $1 \leq j \leq n, i \neq j$ )のは何か	KL I [O]
②	KL I [O][Pi] ( $1 \leq i \leq n$ )である KL I [O]+ KL II [Oj][Pk] ( $1 \leq j \leq o, 1 \leq k \leq p$ )のは何か	KL II [Oj] ( $1 \leq j \leq o$ )
③	KL I [Oi][Pj] ( $1 \leq i \leq n, 1 \leq j \leq o$ )である KL I [Oi] ( $1 \leq i \leq n$ )+KL I [O][Pk] ( $1 \leq k \leq p$ ) のは何か	KL I [O]
④	KL I [O]の説明	KL I [O][Pi] ( $1 \leq i \leq n$ )
⑤	として正しいも のはどれか	KL II [Oi][Pj] ( $1 \leq i \leq n, 1 \leq j \leq n$ ) は KL II [Oi] ( $1 \leq i \leq n$ ) である

本テンプレートでは、知識ベースの知識を挿入する箇所や内容が記号で表現されている。KL I と KL II は知識ベースのルートからの知識階層を意味し、これらに付随する [O]はオブジェクトを、[O][P]はそのオブジェクトに紐づくプロパティを意味する。i, j, k はオブジェクトやプロパティの識別番号を示す。n, p, o はオブジェクトやプロパティの総数を示す。作問者が入力したキーワードが KL I [O]のいずれかに一致するとき問題が生成される。生成時は、テンプレート中の記号に該当する知識ベースの情報を1つずつ挿入する。

表2のテンプレート①は、入力されたキーワードに一致した KL I [O]が複数のプロパティに紐付けられている場合に利用する。解答に KL I [O]を挿入し、問題文に KL I [O][Pi]を挿入後、KL I [O][Pi]以外のプロパティ (KL I [O][Pj]) を問題文に挿入する。

テンプレート②は、KL I [O]とその子ノード間に紐付いているプロパティ KL II [Oj][Pk]の中に KL I [O]と同一の文字列が含まれる場合に利用する。KL I [O][Pi]と KL I [O]を「である」で連結した文を KL II [Oj][Pk]文中の KL I [O]と置き換えた文を問題文に挿入する。解答には KL II [Oj]を挿入する。

テンプレート③は、KL I [O][Pk]の中に別の第1層オブジェクト (KL I [Oi]) と同一の文字列が含まれている場合に利用する。KL I [Oi][Pj]と KL I [O]を「である」で連結した文を KL I [O][Pk]に含まれる KL I [Oi]と置き換えた文を問題文に挿入する。解答には KL I [O]を挿入する。

テンプレート④⑤では、どちらも問題文に KL I [O]を挿入する。解答は④が KL I [O][Pi], ⑤が KL II [Oi][Pj]と KL II [Oi]が挿入される。

#### 4. 自動生成された問題の評価

もりけんの作問委員会 (5名) が検定試験の問題を作成する際に、作問支援システムを利用してもらった。システム利用後に実施したアンケート結果やシステムに登録された問題から、システムで生成された問題の妥当性や活用可能性を評価した。

作問支援システムによる作問負担の軽減効果に関するアンケート結果を図2に示す。図2の結果から、本システムで自動生成された問題を参考にすることで、作問作業の負担軽減効果が期待できること

支援システムを利用して、問題作成の負担が軽減されたと思いますか



図2 作問支援システムに関するアンケート結果

表3 修正内容一覧

No.	修正内容	問題数
①	修正されていない	2(11.1%)
②	問題文の問い方のみを修正	1(5.6%)
③	文を追記・削るなどの修正	8(44.4%)
④	問題文や正答間での文・単語の移動	1(5.6%)
⑤	大幅に書換えられている	6(33.3%)

が示唆された。また、本システムを利用して作成された全18問と、自動生成された問題との差分を分析し、作問者が自動生成された問題にどのような修正を加えたかを分析した。分析から確認できた修正内容を表3に示す。表3⑤は生成時と比較して、数個の単語以外、ほぼ全文が異なる情報に書換えられたものを指す。表3から、修正されなかった問題または軽度な修正のみの問題 (表3①~④) は、大幅な書換えをされた問題 (表3⑤) の2倍あったことが分かる。この結果から、軽微な修正を加えれば自動生成された問題を検定試験に活用でき、ある程度妥当性のある問題が生成されていたことが示唆される。また、過去問題をリソースとした知識ベースから生成される問題が新しい問題を作る際に活用できることが示された。

#### 5. まとめ

本稿では、もりけんの過去問題をリソースとした知識ベースを構築し、問題自動生成機能を利用した作問支援システムを開発した。本システムをもりけんの作問現場で利用してもらった結果、作問負担の軽減効果や知識ベースの活用可能性が示唆された。今後は知識ベースの自動生成機能を開発し、作問負担の軽減効果の検証を行う。

#### 謝辞

本研究に協力して頂いた盛岡商工会議所ともりけん作問委員会の皆様に感謝を申し上げます。

#### 参考文献

- 菅沼明, 峯恒憲, 正代隆義: “学生の理解度と問題の難易度を動的に評価する練習問題自動生成システム”, 情報処理学会論文誌, Vol.46, No.7, p.1810-1818 (2005)
- 田村吉宏, 山内崇資, 林佑樹, 中野有紀子: “Wikipediaを用いた質問応答と多肢選択問題による歴史学習”, 人工知能学会全国大会論文集, 29th, ROMBUNNO.1N2-2, 2015
- 古館昌伸, 福坂祥基, 高木正則: “試験問題をリソースとした知識ベース自動構築手法の検討”, 教育システム情報学会第39回全国大会, p.75-76 (2014)
- 菅原達介, 高木正則: “記述式問題の誤回答を用いた誤答選択肢自動生成システムの開発”, 情報処理学会情報教育シンポジウム, p.177-181 (2013)