

## 集合知を用いた自己管理学習システムの検討

松嶋 和気子, 黒瀬 浩, 中沢 実, 服部 進実

金沢工業大学大学院 工学研究科

eラーニングシステムによる学習は、手軽に取り組めるが、記述式の解答に対してコンピュータ側で処理することが容易でない。本研究では、形態素解析を用いて名詞句を抽出し解答と優良解答群を比較することで、解答のポイントを決定する方法を検討する。解答の精度と学習意識を高めるために、学習者による投票を組み合わせる方法を検討し、参加者が記述解答を改善し知識が自律的に成長できるシステムとして有用性を確認した。

## A Study for Self Administrative e-Learning System With Collective Intelligence

Wakiko Matsushima, Hiroshi Kurose, Minoru Nakazawa, Shimmi Hattori

Graduate Program in Systems for Intellectual Creation,  
Kanazawa Institute of Technology

Study by e-learning system can work easily. However, it is not easy to process it to the description answer on the computer side. In this research, the method of deciding the point of the answer was examined by extracting the noun phrase by using the morphological analysis, and comparing the answer with the excellent answer group. To improve "Accuracy of the answer" and "Study consideration", the method of combining votes by the learner was examined. Consequently, the learner improved the description answer, and utility was confirmed as a system to which knowledge was able to grow up autonomous.

### 1 はじめに

Web 環境での CBT(Computer Based Training) が盛んになっているが、学習者の思考能力の強化に向かない面もある。学習コンテンツは事前に用意された解答候補から選択するケースが多い。学習者が問題を十分に検討する前に繰り返し実施することで解答パターンを覚えてしまう場合がある。解答を選択する学習は受動的である。

一方、記述解答方式では解答のバリエーションが多く、客観的な自動採点が困難である。学習者へのフィードバックは、事前に出題者が用意した

模範解答となる場合が多い。CBTによる学習は、個人がそれぞれ独自に取り組むため、学習者への刺激が少なく学習を継続できなくなる場合もある。

インターネットを通じて、同じテーマに関心を持っている学習者が相互に刺激しあい解答を改善していくことができれば、学習し始めた者には、既存の解答例から理解が進み、スキルの高い者にとっても、新たな気づきが得られるので有用である。

既知の取り組みとしては、以下がある。学習者、教員の交流を通じて学習意識を高めることができる [1][5]。記述解答をコンピュータで処理する方

法が検討されている [6][7]. 知識を共有し, 改善する手段としてウィキペディア [2] は有名である. 名詞句に着目した検索も行なわれている [8].

本稿では, 資格試験などで採用されている数十文字程度で解答する記述問題を題材として, 学習者主体で解答の成長を可能とする e ラーニングシステムを検討する. 記述解答に含まれる名詞に着目して自動採点し, 他の学習者からの投票により精度を高める方法を提案する. 参加者および解答が増えることにより集合知の効果が期待できる.

## 2 提案システム

学習者はブラウザから Web サイトにアクセスし学習を行なう. 学習者は問題に取り組む前にスキルレベルを自己申告する. この方法は SNS(Social Networking Service), BBS(Bulletin Board System) において, 参照者が記述内容の精度を判断するための目安として用いられている. 本システムでは, 申告されたスキルに対応したポイントで解答ポイントの算定に使用する.

解答された答えは, サーバ側でポイント进行計算する. 解答に含まれるキーワードを形態素解析を用いて抽出する. キーワードは, 名詞属性または記号属性を持つ語で, 優良解答群と比較することで採点を行なう. 優良解答群は, 問題ごとに解答ポイントの上位  $n$  個までの集合である. 解答に含まれるキーワードが優良解答群に存在すれば, ポイントの加算を行なう.

学習者は, 解答後に他者の解答を参照できる. 解答群はポイントの高い順に表示するため, 最適な解答例の参照を容易とする. 自分の解答との比較から新たな知識の獲得が行なえる. 学習者は, 他者の解答を参照した後に, 他者の解答に投票することができる. 投票された解答には, 投票者の申告したスキルに応じた投票ポイントを加算する.

優良解答群は, 初期は精度が低いものでも構わない. または, 存在しなくてもよい. 参加人数が増えると, 解答および投票によりポイントが動的に変動し, より正しい解答が優良解答群に含まれ,

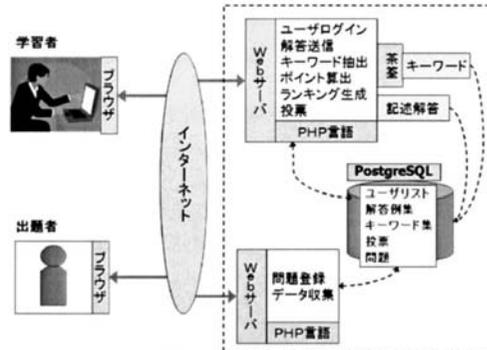


図1 集合知を用いた学習システム

採点精度の向上が期待できる.

図1にシステム概要を示す. サーバー側のコンポーネントは, Web サーバ, Web 言語, SQL DBM と形態素解析から成る.

### 2.1 関係者

本システムの関係者の役割とポイント公開について述べる. 学習者は, ポイントについて意識する必要は無い. 本システムでは, 得点格差が参加者の増加を妨げる要因と考え, 得点は公表しない. 学習者は自分の順位, 他学習者の解答および学習進捗を確認できる. 出題者, 学習者は, 利用に先立ちユーザー登録行なう.

**出題者** 問題を登録する

スキル項目と投票ポイントを設定する

**学習者** 解答と他者へ投票を行なう

**参照者** 問題, 解答, 学習状況の参照のみ可

### 2.2 学習手順

問題はテーマごとに複数, 登録できる. 1 題ごとの採点時に他者の解答を参照することで, 学習者の印象に残ることが期待できる. 以下に学習の手順を示す.

1. 学習するテーマを選定する
2. スキルレベルを自己申請する
3. 解答する
4. 採点ボタンを押し, 結果を確認する
5. 他者の解答との比較を行なう
6. 他者の解答に投票を行なう

7. 次の問題に取り組む(手順の3へ)

表1 スキル要素  $P_s$  の例

2.3 ポイントの算出方法

解答は問題・学習者ごとにスカラー値のポイントを持つ。ポイントは、解答時と投票時に計算する。解答時のポイント計算で参照する優良解答群は、その時点でのポイントの上位  $n$  個までを採用する。ポイントは、固定的ではなく動的に変化する。解答のポイント  $P$  は、キーワード一致ポイント  $P_k$  と投票ポイント  $P_v$  から成る。

$$P = P_k + P_v \quad (1)$$

2.3.1 キーワード一致ポイント  $P_k$

解答が、優良解答群に含まれるキーワードを有していれば、一致数に応じてポイントを加算する。キーワードは名詞または記号で、解答判定に有効な語である。

解答の文字列を  $A$ 、解答が含むキーワードの文字列リストを  $Keys$ 、優良解答群を示す添字を  $b$ 、文字列リスト  $string$  から名詞属性または、記号属性の語句を取り出す関数を  $getNoun(string)$ 、語  $word$  が  $string$  に含まれている個数を返す関数を  $numMatch(word, string)$ 、 $word$  が判定に使用できる場合に真を返す関数を  $isValid(word)$ 、優良解答群に含まれる名詞属性または、記号属性の語が解答に存在したときに加算するポイントを  $P_x$  とすると、キーワード一致ポイント  $P_k$  はリスト 1 のように求めることができる。  $isMatch()$  で の文字列の一致方法は、完全一致を用いる。

リスト1 キーワード一致ポイント  $P_k$  の計算

- 1:  $P_k = 0$  ;
- 2:  $Keys = getNoun(A)$  ;
- 3:  $Keys_b = getNoun(A_b)$  ;
- 4: forall  $word (Keys)$ {
- 5:      $num = numMatch(word, Keys_b)$  ;
- 6:     if(  $num \wedge isValid(word)$  ){
- 7:          $P_k = P_k + P_x * num$  ;
- 8:     }
- 9: }

スキル	ポイント
何も知識がない	0.1
セキュリティ学習中	0.2
資格保有 情報処理 (SU)	0.4
資格保有 情報処理 (SV)	0.4
資格保有 ISMS審査員補	0.4
資格保有 ISMS審査員	0.6
セキュリティ業務に従事	0.6
セキュリティ指導に従事	0.6
その他	0.2

2.3.2 投票ポイント  $P_v$

投票ポイント  $P_v$  は、スキル要素  $P_s$  とインセンティブ要素  $P_i$  から成り、学習者が問題に取り組む際に入力フォームから選択する。要素には、出題者の考慮不足を補うために「その他」を指定でき、項目を学習者が必要に応じて入力することができる。出題者は、記載事項を確認し、必要に応じて項目と投票ポイントを追加することができる。

$$P_v = P_s + P_i \quad (2)$$

情報セキュリティアドミニストレータ試験 [4] の学習において、スキル要素  $P_s$  の例を表 1 に、インセンティブ要素  $P_i$  の例を表 2 に示す。これらは出題者が問題登録時に設定する。

表 1, 表 2 では、学習サイトの参画意識向上と、他者に影響を与えられるように、スキルが低くてもポイントを与えている。同様に、その他の配点も 0.2 と大きくしている。

学習者は、不適切な解答に対して、投票時に否定投票ができる。否定投票が一定数以上である解答は、優良解答群から除外する処理を行ない、キーワード一致ポイント算出の欠点を補うとともに、他者の解答を参照した場合に、参照されにくくする。

2.3.3 キーワード抽出

記述解答からポイント加算に用いるキーワードリストを抽出する方法を示す。

表 2 インセンティブ要素  $P_i$  の例

目的	ポイント
資格受験のための学習	0.05
定期試験・レベルを確認	0.01
その他	0.2

1. 記述解答の文字列の半角を全角に変換する
2. chasen[3] により形態素解析を行なう
3. 記号属性を持つ語を名詞と同様に扱う
4. 名詞属性を持つ語をリストとして抽出する
5. フィルタ処理
6. キーワードのリストに追加する

半角文字を全角に変換する理由は、記述解答では、簡潔に要旨を述べるために文字数制限が指定されており、全半角を同様に扱うことと、半角で形態素解析を行なった場合に、記号が1文字ごとに分割されることを防ぐためである。記号属性を持つ語をキーワードとして扱うことで、アルファベット、数値をキーワード一致処理に利用する。

本システムでは、記述解答の文法解析、重要語抽出、要約処理または、類義語の対応を行っていない。数十字程度の記述解答での主体は、設問で明示されている場合が多い。また、現象の原因、行動の理由や対応を求める場合も多い。日本語の否定や関係の記述を特定する方法は、短文では難しい。以上の理由から、名詞属性を持つ単語を比較する方式を採用している。類義語・同義語については、投票から、よりふさわしい語が優位となることから検討範囲外としている。

解答に「A社に事前の同意を得ずに再委託することを禁止する」が入力された場合の形態素解析結果を図2に示す。

フィルタ処理は、名詞属性でも「こと」のようにキーワードに向かない語があるので除外する処理を実施する。chasenの結果で「名詞句-非自立-一般」属性を持つ語は判定では無効としている。結果として、{A, 社, 事前, 同意, 委託, 禁止}が解答のキーワードリストとなる。

形態素	読み	品詞	活用型	活用形
A	エイ	記号-アルファベット		
社	シャ	名詞-接尾-一般		
に	ニ	助詞-格助詞-一般		
事前	ジゼン	名詞-一般		
の	ノ	助詞-連体化		
同意	ドウイ	名詞-サ変接続		
を	ヲ	助詞-格助詞-一般		
得	エ	動詞-自立	一段	未然形
ず	ズ	助動詞	特殊・ヌ	連用二接続
に	ニ	助詞-格助詞-一般		
再	サイ	接頭詞-名詞接続		
委託	イタク	名詞-サ変接続		
する	スル	動詞-自立	サ変・スル	基本形

図 2 形態素解析の例

## 3 実装

### 3.1 プログラム

問題は、HTMLのコンテンツとしてフォームから入力する。図表を含む問題は、添付文書として登録する。問題、ポイント値、ユーザ属性、解答、進捗状況およびポイント値はPostgreSQLに格納している。プログラムはPHP言語を用いた。ポイント算出以外にも、他者との情報交換、他者の進捗状況把握(図3)、自己による計画管理機能も加えることにより、学習継続の意識低下を防止する効果を付加している。PHPプログラムファイルは機能別に作成し、30程度であった。構成の概要を示す。

1. ユーザー管理
  - 1.1 ユーザー登録・削除・変更
2. メニュー
  - 2.1 学習案内
  - 2.2 問題と解答入力
  - 2.3 解答ランキング (優良解答群)
    - 2.3.1 投票
  - 2.4 進捗ランキング
  - 2.5 自己進捗管理
  - 2.6 問題入力
    - 2.6.1 ポイント設定
3. 掲示板機能

## 学習進捗ランキング

順位	ユーザ番号	解答した問題数
1	30	16
2	1	10
3	20	4
4	11	4
5	29	3
6	2	3
7	21	2
8	26	2
9	17	2
10	22	2
11	15	2
12	27	2
13	13	1
14	12	1

図3 進捗状況把握

### 3.2 環境

クライアントはフォーム入力ができるブラウザが、サーバ側は PHP, PostgreSQL, chasen が動作すればよい。実装では、Linux 上で、Apache 1.3.27, PHP 4.4.4, PostgreSQL 7.4.14 を用いた。対象とする問題は、情報セキュリティアドミニストレータ試験の記述問題を用いた。学習者は、インターネットを介して、時間・場所の制約なしに学習できる。

投票ポイントは、表 1、表 2 に示した値と同じである。優良解答群から除外する拒否投票数は 5 に、キーワード一致ポイント  $P_x$  は、0.02 に、優良解答数は 10 とした。

## 4 結果

本システムを使用して、セキュリティアドミニストレータ試験試験学習者十数名がインターネット経由で記述解答と投票を実施した。

学習で得られた解答の一部を図 4 に示す。ここでポイント値は各学習者からは参照できないが、評価用として表示している。解答数が少ない状態のため、名詞一致ポイントの割合は少ないが、投票された結果、ポイントに優劣がついている。この解答群では同義語を集約すると「電話して本人であることを確認する」という意味の解答と判断で

きる。同義語の集約について、本システムでは、学習者が投票することにより参画意識の向上に結び付けることができる。解答例を見ると別の手段を提案するもの（解答識別番号 130070005）や手段を記載していないもの（解答識別番号 130070007）があり、気づきの機会となる。掲示板での情報交流や他学習者からの指摘により、能力向上が期待できる。優良解答群から投票する画面例を図 5 に示す。解答例の左の選択リストで、「投票」、「投票しない」または、「却下」を選択できる。ここでも学習者が複数の解答を比較することにより能動的な学習が期待できる。

投票によるポイント値の変化の様子を図 6 に示す。解答者による投票により優良解答の順位が変更されている。

解答識別番号	解答	ポイント
130070001	折り返し電話して本人であることを確かめる	1.42
130070008	P部長の携帯に電話をして本人を確認する	0.10
130070004	折り返し電話をして本人を確かめる	0.06
130070009	P部長にコールバックをして、本人であるか確かめる	0.06
130070002	こちらから電話して本人であることを確かめる	0.04
130070003	本人であることを確認する	0.02
130070005	絶対に電話でパスワードを教えない	0.02
130070006	ホテルに電話をして確かめる	0.02
130070007	本当にP部長本人か確かめる	0.02

図4 解答例

投票	解答例
-	個人所有のパソコンをリモートアクセスに使用し、無断で社内のネットワークに接続することの禁止
-	個人のPCを勝手に社内のネットワークに接続しないこと
-	個人所有のパソコンをリモートアクセスに使用し、無断で社内のネットワークに接続することを禁止
-	パスワード管理に関する規定

図5 投票画面

## 5 今後の課題

本システムは、資格試験合格を目標とし教育機関に通わず自宅で自己学習する場合に、同じ目的を持つ人たちとの知識共有さらには意識共有を実現すべく集合知を活用し、インセンティブの継続意欲を保持する学習システムを考察した。

特に記述式問題では、正解は直ちに公開されな

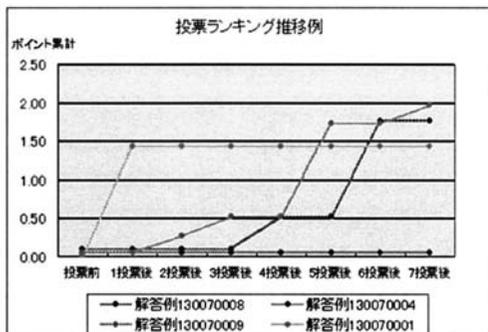


図6 投票によるポイント値の変化

い場合が多いため、教師に相当する機能が期待できない場合に有効である。集合知により記述式問題の解を、かたよりが生成されることなく出題者の労力を極端に少なくまたは、無しの条件で実現するフォークソノミー的アプローチが理想であるが今後の課題である。

本システムで改善可能な点を以下に記す。

スキルが低い学習者が多数参加した場合に、優良解答群の精度が低くなる可能性がある。採点については、解答群と解答のキーワードの完全一致で行っているが、同義語・類義語の対応については、推奨するキーワードを出題者が出題時に入力することで加算ポイントに重みを持たせる方法がある。出題者が意図したキーワードを含む解答が、ランキングの上位に移動しやすくなる。

キーワードの接続の扱いについて改善の余地がある。図2では「再委託」という語句が{再, 委託}として扱われ、計算では、「委託」のみを解答してもポイントの加算となる。語を接続し、最長一致にて処理する方法がある。同様に記号属性を含む「A社」、「10回」なども1語として扱うことが望ましい。

本システムでは、名詞に着目しているために肯定・否定の区別がつかない。記述解答では、体言止めや「～ので」、「～ため」などの理由を答える記述があることから、文法処理を行っていないが、記述解答に向けた最小限の意味理解処理機能の追加が必要と考える。記述解答では、動詞も重要なキーワードになると予想できるが、出題での意図

との対応が必要であり、実装に至っていない。出題文の要約および題意抽出について検討したい。

## 謝辞

有益なご助言とご支援を頂いた株式会社アルファシステムズ 千葉大作氏に感謝致します。

## 参考文献

- [1] 授業支援型 e-Learning システム CEAS 関西大学, <http://www.atl.kansai-u.ac.jp/ceas/>
- [2] ウィキペディア, <http://ja.wikipedia.org/>
- [3] 形態素解析システム 茶筌, <http://chasenaist.jp/hiki/ChaSen/>
- [4] 情報処理技術者試験センター, [http://www.jitec.jp/1\\_04hanni\\_sukiru/index\\_han\\_i\\_sukil.html](http://www.jitec.jp/1_04hanni_sukiru/index_han_i_sukil.html)
- [5] 石川 賢:「情報メディアを用いた授業改善の試み：電子掲示板や Web ページの機能を用いた学習情報の交流について」, 宇都宮大学教育学部教育実践総合センター紀要 27, 20040401(ISSN 13452495) pp. 63-71
- [6] 川場 隆:「自動採点可能なウェブ表示用試験問題を記述するマークアップ言語とその処理系の開発」, 日本教育工学会論文誌, Vol.29, No.Suppl(20060320) pp. 69-72
- [7] 澤井 進, 若木, 池田, 國藤:「暗黙知・形式知を用いた e ラーニング用 Web 検索エンジン」, 教育情報研究 21(2), 20051115(ISSN 09126732) pp. 13-19
- [8] 石井恵, 渡辺一成:「分類体系と名詞句を用いた検索インターフェースの提案と評価」, 情報処理学会研究報告, Vol.2000, No.12(20000128) pp. 1-6