

言語処理を取り入れた英語学習者向けディベート支援環境の構築

岡本 紘幸 肥塚 真輔 曾我 真也 西岡 伸紘 滝本 湖 小原 京子 斎藤 博昭
慶應義塾大学 理工学部

Email: {motch,shizuka,soga,nishioka,lake}@nak.ics.keio.ac.jp,
ohara@hc.cc.keio.ac.jp,
hxs@ics.keio.ac.jp

1 はじめに

自然言語処理を利用した e ラーニングでは、e ラーニング特有の利点である、

- Web をはじめとするさまざまなコンテンツの活用
- 学習者間での情報の共有
- 個人に応じた学習状況の把握やカリキュラムの設定

といった機能をより効果的に提供できることが期待されている。昨年 of 年次大会で併設ワークショップ「e-Learning における自然言語処理」が開催されたことでもわかるように、こうした期待を受けて e ラーニングに自然言語処理技術を適用する研究が盛んに行われるようになってきている。たとえば、安藤らは補習のための大規模なシステムを構築している (安藤ハル他 2004)。そこでは個人に特化した教材を生成するために、生徒がデジタルペンで書く単語を同定するという処理が見られる。ただ、このような初歩的な言語処理の適用に留まっている大きな原因は、この種の学習アプリケーションには処理の誤りが許されないためである。実際、現在の言語処理技術の精度でそのまま教育に適用することはできないとの指摘もなされている (Segond and Parmentier 2004)。確かに現状の言語処理を積極的に e ラーニングに使っていかうという動きもある (Smrž 2004) もの、意味解析技術が未発達な現状では、適用範囲は狭くならざるをえない。

一方で、自然言語処理技術を適用した e ラーニング環境を実用的なレベルで提供するには、意味処理を含む高度な言語処理が要求されていることも事実である。この要求に対しては、談話解析といった未成熟 of 自然言語処理技術をそのまま使用するのではなく、Tane らの研究 (Tane et al. 2004) のようにコンテンツ間の関係などをあらかじめ意味的に構造化して管理するという解決方法が考えられる。しかしこの方法は、コンテンツ作成者に大きな負担を与えるだけでなく、学習者の回答といった動的なコンテンツに対しては適用でき

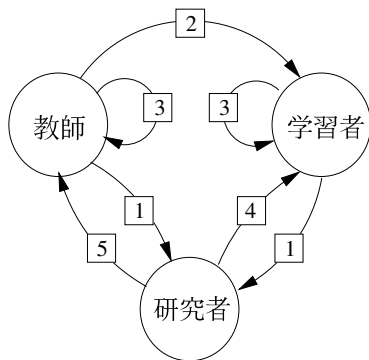
ない。そこでわれわれは、動的なコンテンツに対する構造化の作業を使用者に明示的・暗黙的にしてもらい、アノテーションとして XML 形式でコンテンツに埋め込み、それを解析することにした。同様の方法で動的なコンテンツにアノテーションを行う研究はいくつかある (掛川淳一他 2004; 山根木浩平他 2005) が、いずれもコンテンツ構造化の作業が使用者に大きな負担となっているという問題があるため、特に学習者に対してアノテーションを意識させないような枠組が必要となる。

そこで本稿では、学習過程における学習者や教師の操作をコンテンツに対する暗黙的なアノテーションと見なすことにより、学習者にアノテーションを意識させることなくコンテンツを構造化し、蓄積されたコンテンツを利用した高度な言語処理アプリケーションが容易に組み込めるような e ラーニング支援環境の枠組を提案する。また、提案した枠組に基づいて構築した、英語学習者向けディベート支援を題材としたシステムを紹介する。このシステムは大学の英語の授業において試験的に運用されている段階である。

なお、本稿 of 提案のように独自の XML を設定するという意味では、Keogh らが、教材を Perl や POS tagger にか、XML のタグ付きコーパスに変換し、それから練習問題を生成する試みをしている (Keogh et al. 2004)。また、われわれと同様のアプリケーションである議論支援という点では、太田らが遠隔授業での使用を想定したシステムを作成している (太田政宏他 2004)。ただ、本稿のように自然言語処理を取り込むことは想定していない。

2 概要

本稿で提案する e ラーニング支援環境 of 枠組は、環境が提供する学習支援によって教師や学習者が受ける利益と、環境が蓄積するコンテンツによって自然言語処理技術が受ける利益を両立する。枠組が提供する利益の相関を図 1 に示す。



- 1: 自然言語処理ツールに有用なコンテンツの動的生成
- 2: 教師からのフィードバック
- 3: コンテンツの効果的な提示
- 4: 学習効果の向上
- 5: 教師の負担削減

図 1: 利益相関図

本節では、図1に示したそれぞれの利益を実現するための枠組を、データ記述言語仕様とインタフェースの面から詳述する。

2.1 データ記述言語仕様

本環境では、XML で定義された Keio e-learning Markup Language(KeML) を用い、コンテンツを蓄積する。KeML は橋田の提案する GDA (Global Document Annotation)(橋田浩一 1999) を包含していることから、GDA のコンテンツを対象とした検索 (Miyata and Hashida 2003) などの既存の各種言語処理ツールをそのまま利用できるため、蓄積されるコンテンツは自然言語処理ツールに有用なコンテンツとなる (図 1 利益 1)。また、それらのツールを用いて学習効果を促進したり、教師の評価等にかかる負担を削減することが可能となる (図 1 利益 4, 5)。

さらに、GDA が意味的・構文的な言語構造を記述できることから、次節で詳説する出力インタフェースとの組合せにより、学習効果の向上と教師の負担削減が期待できる (図 1 利益 3)。

KeML では GDA で定義された意味的な構造の他に、教師による修正やコメント等の学習に有用な情報を含むことができるため、蓄積されるコンテンツからこうした情報を学習者に提示できるという利点もある (図 1 利益 2)。

2.2 インタフェース

環境の利用者にとって、e ラーニング支援環境を利用する目的である学習以外の操作を明示的に行うことは負担であり、研究者はそれを期待できない。また、確立された操作モデルが存在する場合には、その操作

モデルとあまりかけ離れた操作を要求することも利用者の負担につながる。そこで本環境では、既存のアプリケーションと同様の操作モデルにおいて、学習に必要な操作から得られる情報のみをアノテーションとして蓄積するものとする。これに従うことで、利用者に負担を強いることのない入力インタフェースを提供することができる (図 1 利益 1)。

一方、学習者の学習効果の向上や教師の評価等の負担削減のためには、コンテンツの効果的な表示が必要となる。例えば、発想支援効果のあるセマンティックオーサリングの考え方を適用し、構造化されたコンテンツをグラフ形式で表示する出力インタフェースを提供することで、コンテンツに対する学習者の理解を容易にすることができる (図 1 利益 3)。

3 ディベート支援システム

本節では、前節で提案した枠組に基づいて構築した、英語学習者向けディベート支援システムについて詳説する。

3.1 システム概要

本稿が提案するディベート支援システムは、サーバ・クライアント型システムである。サーバはコンテンツの管理や自然言語処理の機能を、クライアントはユーザインタフェースを提供する (図 2)。

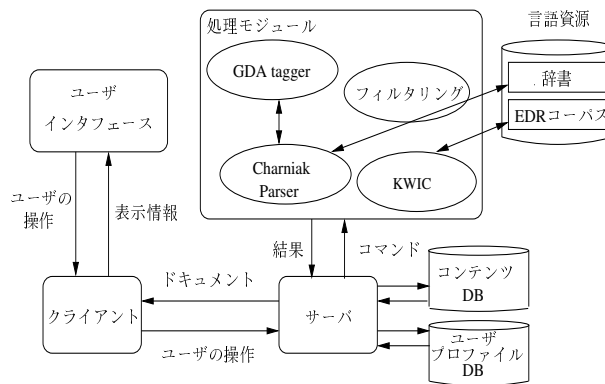


図 2: システム構成

ユーザの発言は、Charniak Parser(Charniak 1999) を用いて構文解析され、GDA 形式に変換されてデータベースに蓄積される。

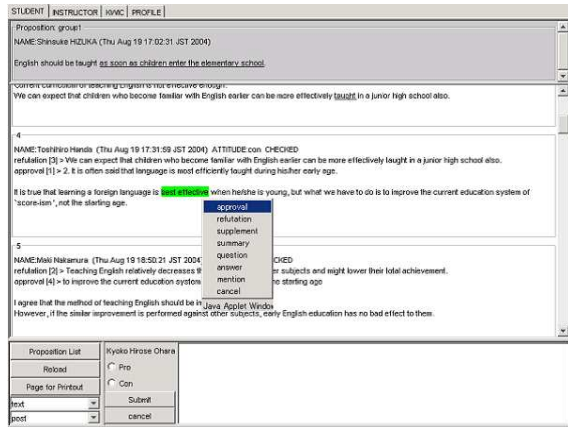


図 3: ユーザインタフェース (テキスト形式)

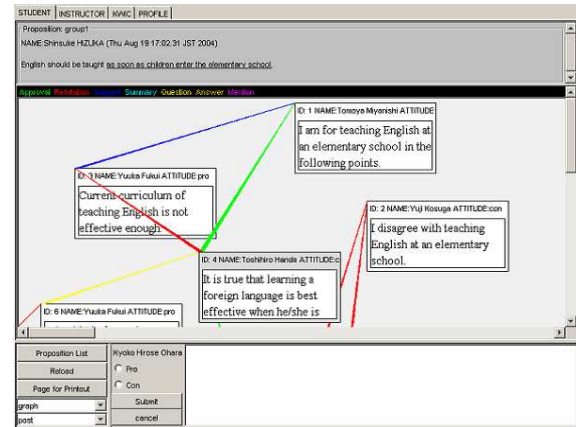


図 4: ユーザインタフェース (グラフ形式)

3.2 データ記述言語仕様

本稿で紹介するディベート支援システムではその目的のため、データ記述仕様 KeML をディベート支援に特化させた。ディベート支援向け KeML では、

- ディベートのテーマ
- ディベートのテーマに対する発言
- ディベートのテーマに対する発言の姿勢 (pro: 賛成 / con: 反対)
- 既存の発言に対する新しい発言の位置づけ
- 教師による、発言におけるスペルミスやシンタクスエラーの修正
- 教師による、発言に対するコメント

といった情報が記述できる仕様となっている。なお、発言の位置づけとしては以下の 7 種が定義されている。

approval (同意)	refutation (反論)
supplement (補足)	summary (要約)
question (質問)	answer (回答)
mention (言及)	

3.3 インタフェース

本稿で紹介するディベート支援システムのユーザインタフェースを図 3 に示す。本システムでは、上部に蓄積されたコンテンツが表示され、下部で利用者が発言等を行う。

発言に際し、学習者は既存の発言の全体ないし一部を指定することによって、既存の発言に対する新たな発言の位置づけを指定することができる。この操作はディベート学習に不可欠な操作であり、学習者の陽のアノテーション操作ではないが、これにより動的にコ

ンテンツが生成、蓄積される。また、本システムのインタフェースは、学習者の発言に対する教師の修正やコメントなどの機能を提供するため、より学習効果の高いコンテンツが生成、蓄積可能である。

加えて本システムは、構造化されたコンテンツをグラフ形式で表示する出力インタフェースを提供している(図 4)。これにより、ディベートの進行状況や多数の発言の関係を視覚的に把握することができ、より高度な学習効果が期待できる。

3.4 その他の機能

本稿で紹介するシステムは、前節までで詳説したもののほかにも、ディベート支援のための様々な機能を提供する。

- [フィルタリング機能]

コンテンツに記述された各学習者の姿勢 (pro/con) や発言同士の関係に基づくフィルタリング機能。紆余曲折を経るディベートにおいては大局的にディベートを総括することが困難となるため、全体の中での発言の位置づけを把握しやすくする。

例: 「テーマに対する賛成意見のうち、ある発言に対する反論を 3 発言まで遡って表示する」

- [KWIC 機能]

例文表示により文生成を支援する機能。学習者が英文を作成する際に、句動詞や連語の用法を実英文で確認できるように、指定された単語が含まれているコーパス中の文を KWIC(Key Word in Context) 形式で表示する。コーパスには、EDR 英語コーパス*を利用した。このコーパス中の各々

*EDR ホームページ: <http://www2.nict.go.jp/kk/e416/EDR/>

の単語には概念情報が定義されており、例文は概念ごとに分類出力される。

4 評価・考察

本稿で紹介するディベート支援システムの評価のため、慶應義塾大学理工学部の英語の授業において本システムを利用した英語のディベートを行った。以下に、利用後に集計したアンケートの結果とその考察について記述する。

4.1 評価結果に関して

ここでは、アンケートで指摘された以下の点について考察する。

- [修正機能]
教師による修正機能は良い評価を受けた。実際にスペルミスや文法ミスなどをする学習者が多く、この機能が役に立ったのだと思われる。しかし、この機能は現在、自動修正ではなく教師による修正であるため、教師に負担がかかるのが難点である。
- [フィルタリング機能]
フィルタリング機能は、多くの意見が交わされているディベートにおける意見の集計に有用であった。この機能によって、さらにグラフ表示によるコンテンツの可視化の有用性が高まっている。
- [グラフ表示での発言の位置づけの指定]
本システムでは現在、発言に際してグラフ上で入力することができない。コンテンツの構造化、可視化の利点を活かすためには、今後この機能を追加することは不可欠であると考えられる。
- [アノテーションなしのグラフ表示]
既存の発言に対する位置づけが示されていない場合、グラフ表示機能はテキスト表示に比較して一覽性に欠ける。前述の通り、グラフ表示での発言の位置づけ指定ができないことが、アノテーションのついていないコンテンツを生成する要因となっている。

4.2 アノテーションに関して

学習者によるアノテーションとして、実際に多く付与された発言間の位置づけは、ある発言に対する質問、そしてそれに対する答えという関係を示すことが多かった。しかし、生成されたコンテンツには想定していたほど多くの位置づけの付与がなされていなかった。こ

れは、本システムの評価を英語の授業で行ったため、学習者同士が意見を交わし合うディベート本来の形を環境として整えることができなかったためであると考えられる。

5 おわりに

本稿では、学習者にアノテーションを意識させることなくコンテンツを構造化し、蓄積されたコンテンツを利用した高度な言語処理アプリケーションが容易に組み込めるようなeラーニング支援環境の枠組を提案した。更に、この枠組に基づいて、英語学習者向けディベート支援環境を構築し、評価と考察を行った。

GDA を利用した言語処理ツールを含め、利用可能なツールを組み込むことにより、より効果的な学習支援を行うことや、暗黙のアノテーションがより促進されるようにインタフェースを洗練させることが今後の課題である。

参考文献

- 安藤ハル, 関本信博, 広池敦 (2004). “個人状況適応型映像コンテンツ生成システムの開発.” 言語処理学会年次大会併設ワークショップ “e-Learning における自然言語処理”, pp. 5-8.
- Charniak, E. (1999). “A Maximum-Entropy-Inspired Parser.” Tech. rep. 12, Brown University.
- 橋田浩一 (1999). “GDA 意味的修飾に基づく多用途の知的コンテンツ.” 人工知能学会, **13** (4), 528-535. (<http://www.i-content.org/gda/>).
- 掛川淳一, 石川賢太郎, 海野俊介, 藤井雅弘, 伊丹誠, 伊藤紘二 (2004). “比較による表現獲得を支援する日本語読解・作文学習環境.” 言語処理学会年次大会併設ワークショップ “e-Learning における自然言語処理”, pp. 29-32.
- Keogh, K., Koller, T., Ward, M., Dhonnchadha, E. U., and van Genabith, J. (2004). “CL for CALL in the Primary School.” In *COLING2004 Workshop “eLearning for Computational Linguistics and Computational Linguistics for eLearning”*, pp. 79-85.
- Miyata, T. and Hashida, K. (2003). “Information Retrieval Based on Linguistic Structure.” In *the Japanese-German Workshop on Natural Language Processing*.
- 太田政宏, 横山輝明, 衛藤将史, 門林雄基, 山口英 (2004). “遠隔授業における共同レポート作成のための議論支援システムの実装と評価.” 電子情報通信学会研究報告. IA2003, 38 号, pp. 13-18.
- Segond, F. and Parmentier, T. (2004). “NLP Serving the Cause of Language Learning.” In *COLING2004 Workshop “eLearning for Computational Linguistics and Computational Linguistics for eLearning”*, pp. 11-17.
- Smrž, P. (2004). “Integrating Natural Language Processing into E-Learning — A Case of Czech.” In *COLING2004 Workshop “eLearning for Computational Linguistics and Computational Linguistics for eLearning”*, pp. 1-10.
- Tane, J., Schmitz, C., and Stumme, G. (2004). “Semantic Resource Management for the Web: An E-Learning Application.” In *WWW Conference 2004*.
- 山根木浩平, 伊藤一成, 斎藤博昭 (2005). “汎用アノテーションシステム (MAML System) の e-Learning への適用.” In *DEWS2005*. (to appear).