

コミュニケーションに基づく協調学習評価支援ツールの開発 A tool for collaborative learning based on communication

池宮 直十

Nao Ikemiya

花川 典子十

Noriko Hanakawa

1. はじめに

大学教育における学習方法に協調学習という有効な方法がある。グループで作業分担を決めてその目標に向かって調査や制作、発表する学習方法のことであり、技術習得する科目では机上学習より効果的な学習が期待できる。さらに、協調学習について酒井は「協調学習とは学習者がグループ活動の中で互いの学習を助け合い、一人一人の学習に対する責任を果たすことで、グループとしての目標を達成していく、協調的な相互依存学習である」[1]と述べており、協調学習で学生を動機付けすることもできる。

しかし、協調学習には問題点もある。学生個別の評価が困難ということであり、協調作業の成果物だけで評価すると正確な評価が得られない。たとえば、優秀な学生一人だけで全作業を実施した場合でも、グループの学生の評価がすべて同じになるという問題である。その対策として、学習中の学生を観察する方法があるが、少人数授業では可能だが、大人数になるとすべての学生を観察できず、実質不可能となる。

そこで、本研究では協調学習の評価を支援するツールを作成する。これによって効果的な教育の実施と同時に、より正確な評価を実現することが可能となる。

2. 提案する評価方法

2.1 基本的な考え方

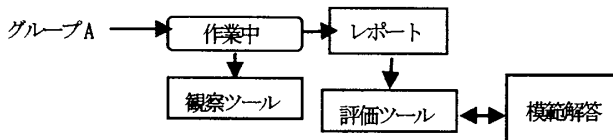


図1. 評価ツールの概念図

まず、基本的な考え方として、学生の作業を観察する観察ツールと作業の成果物を評価する評価ツールの2つで協調作業の評価を支援する。これによって、従来の成果物だけによる評価に加えて、自動的に作業の状況を観察し評価することが可能となる。

評価支援の特徴は作業中のコミュニケーションからグループ全体の状況を把握できることである。まず、作業中の会話のログを記録し、観察ツールを用いてグループの作業状況を定量的に計算する。この結果の値が大きいならば、グループ内でのコミュニケーションが十分になされておらず、話し合いもしないまま、もしくは不必要な会話をしながら非効率に作業がなされている。もしくは、一人だけが作業をしている状況が考えられる。逆に計算した数値が低いならば、作業内容の割り振りについてきちんと話し合いがもたれており、問題解決が効率的になされていることがわかる。

レポート評価ツールで得られたレポートの質を計算した結果に観察ツールの結果を加味し、協調作業全体の評価とすることができる。作業結果だけでなく、その途中経過を見ることにより、より正しい評価を与えることができる。

2.2 協調作業の観察方法

観察ツールは大蔵らの作成したものを使用する[2]。基本アイデアは協調作業が順調であるならばチーム内のコミュニケーションも良いという考えである。つまり、チームで作業する限り

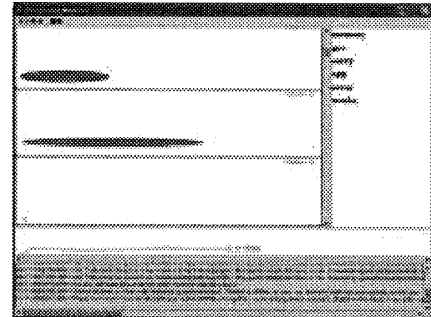


図2. 協調作業観察ツール

コミュニケーションは必ず発生し、良い協調作業が実施されているならば少なくとも良いコミュニケーションが実施されているはずである。さらに、良いコミュニケーションとは「特定の話題について集中的に議論され問題が解決される。以後はその話題について話題が発生しない」状態とみなす。つまり、協調作業において発生する様々な問題を適切にメンバー間で共有し、さらに問題解決のために集中的に議論し、的確に問題解決する状態である。この状態をコミュニケーションであらわすならば、「特定の話題についてのコミュニケーション発生時間が特定の時間に集中しており、長時間継続しない。」状況と考えられる。反対に悪い状態とは「メンバー間で問題を共有しないので極端にコミュニケーション話題数が少なく、かつ特定の話題がなかなか解決できないのでひとつの話題が長期間継続する」状況と考える。そこで、この考えをモデル化すると以下の式となる。

$$Cap1 = \frac{1}{\frac{1}{N} \sum_{i=1}^N SD_i} \times \frac{1}{N} \sum_{i=1}^N Kurt_i \times 100 \quad \dots(1)$$

$$SD_i = \sqrt{\frac{1}{m} \sum_{j=1}^m (x_{i,j} - \bar{x}_i)^2}, \quad Kurt_i = \frac{1}{m} \sum_{j=1}^m \frac{(x_{i,j} - \bar{x}_i)^4}{SD_i^4}$$

Cap1 : 協調作業状況値
SD_i : i 番目の話題発生時間の標準偏差。
Kurt_i : i 番目の話題発生時間の尖度。
x_{i,j} : i 番目の話題の j 番目のコミュニケーション

つまり、同一話題のコミュニケーション発生時間のばらつきが大きく、かつ特定時間に集中していない場合は協調作業状況値が小さい。つまり特定問題が集中して議論されず、その解決も迅速に行われていない。また適切に情報が共有されていないのでコミュニケーションがまとまって発生していない状況も示す。

図2に観察ツールを示す。まず協調作業中のコミュニケーションを記録する。コミュニケーションの内容を TF-IDF 法[5]と、ベクトル空間法[3]と、クラスタリングアルゴリズム[4]を用いて分析し話題に分類する。話題ごとのコミュニケーション発生時間の分散を示したのが図2の観察ツールに示される赤い丸である。すなわち、赤丸が大ききときはその話題でとりあげられている問題がなかなか解決できない状況を示し、赤丸が小さい時は問題が早期に解決したことを示す。また、極端な話題(コミュニケーション)が少ない時も協調作業状況値は下がる。

2.3 成果物の評価方法

成果物を自動評価するためにレポート評価ツールを使用する。自然言語処理技術を使って模範解答との文書類似度を計算し、模

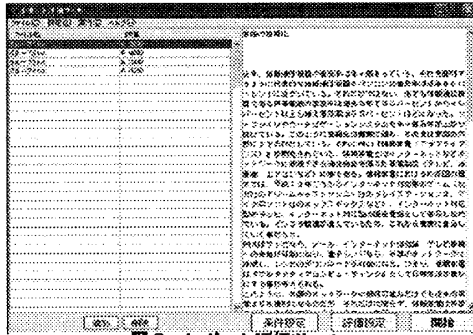


図3. レポート評価ツール

範解答に近いほど評価が高くなる。日本語の解析には自然言語処理の技術である形態素解析を用いる。これを実現するため形態素解析器Sen[6]を用いる。図3がレポート評価システムの画面である。画面の左側に評価したいレポートをドラッグ&ドロップするか、追加ボタンによって指定する。右側はクリックされたレポートをプレビューするエリアである。そして、条件設定ボタンから模範解答を指定し、類似度計算のためのキーワードを抽出する。キーワードは手動追加することもでき、詳細な設定が可能である。評価設定ボタンでは、優良可 (A,B,C) とする類似度を指定する。これらの設定をした上で開始ボタンを押すとレポートの評価が開始され、各レポートの評価結果が表示される。

3. ツールの評価実験

まず、2つのツールの単体の評価を行う。観察ツールの評価実験結果は文献[2]にて実施し、バグ数による品質と計測した協調作業状況値の相関係数は 0.62 であった。また、レポート評価ツールはソフトウェア工学科目のレポートを評価した結果、教員が実際に評価した結果とツールによる評価結果の相関係数は 0.79 となった。

次に二つのツールを組み合わせた協調作業全体の評価のために実験を行う。4グループ計8人が課題について各グループで協力しながらレポートを書き、その協力過程の会話のログを取る。課題は「家庭の情報化について」であり、ログの取得には MSN Messenger[7]を使用する。結果を表1に示す。レポート評価ツールによって得られた「評価値」と教員がマニュアルで評価した「Manual」値 (100 点満点とする)、さらに観察ツールで計算した「状況値」と教員がマニュアルで作業実施状況を評価した「Manual」値 (10 点を普通とする) を示す。結果は、レポート評価が高い順にグループ2, 4, 1, 3、状況値はグループ3, 2, 1, 4の順に良い結果を残している。二つのツールで総合判定した結果が表1の「総合評価」の欄である。ツールの計算ではグループ2の成績が最も良くグループ4が最も悪い。教員のマニュアル評価ではグループ1が最もよく、グループ3, 4が共に最も悪い結果となった。この実験では2つのツールの評価とマニュアルの評価で差が出た。この差について考察する。

これらの差を検証するためにレポート内容を実際に読み、さらにコミュニケーションログを解析した。その結果各ツールの特徴が明確になった。(1) 箇条書きをするとレポート評価ツールの結果が良くなる、(2) 雑談があると観察ツールの状況値が悪くなる、(3) Web サイトから文章をコピー&ペーストするとレ

表1 実験結果

	レポート評価		観察ツール		総合評価	
	評価値	Manual	状況値	Manual	ツール	Manual
Group1	36	75	109	8	3.3	A
Group2	46	60	9.7	10	4.7	B+
Group3	34	60	9.0	11	3.7	B
Group4	42	65	15.0	12	2.8	B

表2 各ツールの評価可能性

状態	期待する評価	観察ツール	評価ツール
Copy & Paste	Low	High	High
箇条書き	Low		High
雑談	Low	Low	
問題解決できない	Low	Low	
一貫性	High		Not available
良い作業分担	High	High	

表2 各ツールの評価可能性

ポート評価、状況値ともに評価が上がる、という3点である。

(1) はレポートを作成しやすいように調べる項目を実験前に列挙し、それによってレポート作成した場合、グループ2, 3, 4が箇条書きをしている。逆に箇条書きではなく文章で記述したグループ1は評価が低くなった。しかしグループ1は教員が直接評価すると一番内容がよかった。(2) ではグループ4が作業中に雑談をしたので作業全体の能力が低くなった。これは食べ物や遊びの雑談が混在し、またチャット用の顔文字も混在していたからである。(3) ではレポート評価ツールも観察ツールも Web サイトの文章をコピーしそのままチャット上に送信、またはレポートの文章に利用すると評価値が良くなるということがわかった。これはグループ1, 2, 3に当てはまる。この3グループの状況値が同じような数値になったのはこのためだと考えられる。

表2はこれらのツールによっての評価の良し悪しと、通常要求される評価の条件についてまとめた表である。今回作成したツールで表現できたのは「雑談をすると評価が下がる」「問題解決がスムーズなら評価が上がる」「分担が正常なら評価が上がる」の3つである。これらより現在のツールを使った評価には制約が必要であり「Web サイトから文章をコピー&ペーストはしない」「箇条書きすることを禁じる」「文章として全体を書く」などを事前条件として設定しなければいけないことがわかった。さらに大きな傾向としてはインターネットサイトのコピー&ペーストした文章をレポートに貼り付けた際に双方のツールとも高評価になるのだが、不正検出の機能を拡張し、Web サイトからコピーされた文章も検出できるようにすることで改善できると考える。

4. まとめ

様々な学習方法の中から協調学習に着目し、協調学習の評価を支援するツールを作成した。協調作業観察ツールを組み合わせることにより、協調学習の評価を支援することができる。協調学習観察ツールで、学習中の途中経過を評価し、結果はレポート評価ツールで評価する。結果のみで評価するのではなく途中経過も評価することにより、グループ全員が作業せずに、一人だけでやっている場合などを検出できる。より協調学習らしい学習をしているグループに対し評価を上げることができるようになった。

参考文献

[1] <http://www.bun-eido.co.jp/aste/sakai.html>
 [2] 大蔵君治, 花川典子: "アジャイルソフトウェア開発におけるプロジェクト能力計測ツールの開発", 第3回情報科学技術フォーラム講演論文集 (FIT2004), Vol.1 pp183-184, Sep.2004.
 [3] Salton, G and McGill, M. J.: "Introduction to Modern Information Retrieval", McGraw-Hill(1983)
 [4] 鳥協純一郎: 認識工学, コロナ社(1993)
 [5] <http://mikilab.doshisha.ac.jp/dia/research/report/2004/0812/001/report20040812001.html>
 [6] <http://ultima.org/sen/>
 [7] <http://messenger.msn.co.jp/>