九州大学学術情報リポジトリ
Kyushu University Institutional Repository

タグ付きテキストからの問題自動生成システム

菅沼，明
九州大学システム情報科学研究院知能システム学部門

峯，恒憲
九州大学システム情報科学研究院知能システム学部門

正代，隆義
九州大学システム情報科学研究院知能システム学部門

https://hdl.handle.net/2324/5587

出版情報：情報処理学会研究報告. DD, [デジタル・ドキュメント]. 2000 (28), pp.17-24, 2000-03. 情報処理学会
バージョン：
権利関係：ここに掲載した著作物の利用に関する注意 本著作物の著作権は（社）情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。

Information Processing Society of Japan

デジタル・ドキュメント 22－3
（2000．3．17）

タグ付きテキストからの問題自動生成システム

菅沼 明　峯 恒憲　正代 隆義
九州大学 大学院システム情報科学研究科

教育機関へのコンピュータやLANの導入により，教材をWebのデータとして作成した講義も行われている．しかし，教材を電子化するだけではなく，コンピュータ世界の変化に対応した教育方法，教育支援方法が必要とされている．教材に沿った練習問題があり，ブラウザ上で解答できる環境が整っていれば，学生は容易に練習問題にチャレンジすることができる．これにより，学生の理解を深めることが可能になる．本論文では，教材として用意しているタグ付きテキストを利用して，練習問題を生成するシステムAEGIS(Automatic Exercise Generator based on the Intelligence of Student)について述べる．AEGISは学生の理解力に応じた練習問題の出題を可能としている．また，練習問題の出題形式を制限することで，自動採点も可能にしている．

Automatic Exercise Generator with Tagged Documents

Akira SUGANUMA, Tsunenori MINE, and Takayoshi SHOUDAI
Graduate School of Information Science and Electrical Engineering
Kyushu University
e-mail : {suga,mine}@is.kyushu-u.ac.jp, shoudai@i.kyushu-u.ac.jp

Popularization of computers and the Internet enable us to hold lectures using Web contents as a teaching material. We should, however, make better use of the new computer and the Internet technologies, and have to develop a new lecture method using them. If we arrange some exercises which follow a lecture note and prepare an answering mechanism for the exercises through the Internet, every student can try the exercises any time. This system aims to help the students understand the lecture. In this paper, we propose AEGIS (Automatic Exercise Generator based on the Intelligence of Student) system which generates some kinds of exercises automatically based on the tagged documents with considering both the ability of a student and the abstrusity to solve the exercises. It realizes a feedback mechanism by restricting the exercise generation to the well-known three types.

NII-Electronic Library Service

Information Processing Society of Japan

1. はじめに

インターネットの普及とそれに接続するコンピュータの低価格化によって，大学などの教育機関でもコンピュータの導入やLANの構築が行われている．大学において，学生が一人一台のコンピュータを使用できる環境で受ける講義を開講することも可能になってきている．また，個人所有のコンピュータを大学のネットワークコンセントなどに接続し，公開されている講義資料を講義時間以外に参照する学生も増えてきている．このような教育を取り巻く環境の変化に対応した教育方法，教育支援方法が必要とされている．

我々は，Webを使用した講義支援に関する研究を行っている[1][2]．コンピュータリテラシー教育においてプログラムの実行を理解するためのツール[3]に関する研究も行ってきた．これらは，九州大学の学部1，2年生を対象にして開講されるコンピュータリテラシーの科目で使用され，講義の効率を高めている．

その教育方法のひとつとして，教材をWebのデータとして用意する方法が多用されている．用意した教材を教師のブラウザで表示し，教師のディスプレイをスクリーンに投影して講義を行う．また，教師が参照しているWebページを学生のディスプレイ上のブラウザに同期表示するようなツール[4]を使用して講義を行う．講義以外の時間に，学生の予習・復習のためにWebのデータを使用する．このような形で講義を進めることが普通に可能になってきている．そのため，Webのデータとして用意された教材も増えている．

これらの教材に沿った練習問題があり，それをWebブラウザ上に表示し，解答できる環境が整っていれば，学生は容易に練習問題にチャレンジすることができ，講義内容についての理解を深めることが可能になる．また，学生の自習にも使用できる．練習問題を作成するときには，解答する学生の理解度によって問題の難易度を変えられると，学習効果はさらに高まる．しかし，難易度毎に個別に練習問題を作成するのは教師に対する負荷が大きくなる．そのため，教材として用意しているHTMLデータを利用して，練習問題を生成するシステムAEGIS(Automatic Exercise Generator based on the Intelligence of Student)を設計し試作した．このシステムは，過去に学生が答えた履歴から学生の理解度を推定し，練習問題の難易度を変更して出題するものである．

練習問題を出題できても，学生への採点後の提示がなければ，学生も継続してチャレンジしなくなる．そのため，教師は学生の解答を採点して学生へ情報のフィードバックを行わなければならない．この作業は学生の数に比例して大きな負荷となる．そのため，練習問題を提示するシステムには，自動的に採点する機能が欲しくなる．AEGISでは，練習問題の出題形式を制限することで，自動採点を可能にしている．

2. 練習問題の出題形式

講義などで習った分野に関して問題を作成する場合には，論述式の問題からマークシート方式の問題まで様々な出題形式が考えられる．しかし，コンピュータによる自動採点を可能にしようとすると，出題形式は限られてくる．例えば，論述式の問題であれば，学生が書いた文章を読んで内容を理解しなければ採点できない．文章の内容理解に関しては，自然言語処理の研究でも正確に行えていない分野である．そのた

NII-Electronic Library Service

Information Processing Society of Japan

選択問題	空白に当てはまる適当なものを下から選べ. Data structures need to be studied [ ] order to understand the algorithms. (1) an (2) in (3) on (4) at (5) by
穴埋め問題	下の空白を埋めよ. Data structures need to be studied [ ] order to understand the algorithms.
誤り訂正問題	誤りの箇所を指摘して，正しく訂正せよ. Data structures need to be studied an order to understand the algorithms.

図 1: 3 つの出題形式

め，このような問題を出題形式に選ぶのは妥当ではない．そのため，本研究では，下に示す3つの出題形式を対象として問題の自動作成を考えている．

- **選択問題：** 文章中に空白が挿入されていて，その部分に最も当てはまるものを選択肢から選ぶ問題である．
- **穴埋め問題：** 文章中に空白が挿入されていて，その部分に最も当てはまるものを記述する問題である．
- **誤り訂正問題：** 文章中のどこかに誤った表現が挿入(もしくは置換)されていて，誤りを捜し出すとともに正しく修正する問題である．

これら3種類の出題形式の問題は，正解の文章の一部を空白に置き換える，もしくは誤った表現に置き換えることによって作成できる．説明の都合上，以下では，空白に置き換える箇所もしくは誤った表現に置き換える箇所のことを『出題箇所』と呼ぶことにする．出題箇所を同じにした場合の各出題形式の例を図1に示す．これらの例では，文中に現れる慣用句「in order to」の「in」の部分を出題箇所としている．

問題文中の出題箇所を同じにしても上に示した3種類の出題形式で難易度が変化する．これは，問題用紙上に載っている情報の量を考えることではっきりする．各形式の難易度について，問題を解く場合と問題を作る場合とに分けて以下で考察する．

問題を解く場合を考えると，選択問題は，選択肢に必ず正解が含まれている．学生は，その選択肢の中から正解を選べば良いので，比較的容易な問題となる．選択問題は，正解の情報が問題用紙にあることを意味する．次に，穴埋め問題は，空白に挿入すべき正解を学生自身が考え出して埋めなければならない．これは，正解の情報が全くないことを意味する．この点で，穴埋め問題は選択問題より難易度の高い問題と判断できる．誤り訂正問題は，問題文中にある誤りの箇所を探さなければならない．さらに，誤りと判断した箇所に対して正解の記述を学生が考えて訂正を行う．つまり，正解の情報が全くないばかりか，出題箇所自身も隠されている．そのため，誤り訂正問題が最も難易度が高いといえる．

次に，問題を作成する場合を考える．出題箇所は出題者が指定するので，どの出題形式でも出題者は同じ作業を行うことになる．出題箇所

NII-Electronic Library Service

Information Processing Society of Japan

が決まった後の問題作成は，各出題形式で以下のように異なる．穴埋め問題は，出題箇所さえ決まればそれだけで問題が作成できる．穴埋め問題の出題形式では，問題用紙に書き込む正解の情報は全くないことになる．そのため，3種類の出題形式のなかでは最も作成しやすいと考えられる．次に，誤り訂正問題は，正解とは異なる記述を1つ用意して，出題箇所の表現と入れ換えることで作成可能である．さらに，選択問題では，正解と異なる記述を複数用意して，それを選択肢にしなければならない．このため，選択問題の方が誤り訂正問題よりも問題用紙に書き込む正解の情報が多いことになる．これらより，問題の出題に関しては，穴埋め問題，誤り訂正問題，選択問題の順に作成が難しくなる．

これらのことから，出題箇所を同じにした問題でも出題形式が異なれば，問題を解く学生にとって難易度が異なる問題となる．さらに，問題を作成する場合にも作成の難易度が異なることが分かる．

3. 問題の自動生成

3.1 出題に必要な情報

教材から練習問題を作成する手順は，以下のようになるであろう．

(1) **出題箇所の決定：** まず，何を練習問題として学生に出題するのかを決めなければならない．出題箇所の決定には，「この教材の内容を学生が理解しているかを判定するには，何を問えばいいのか」を考えなければならない．この様に，出題箇所の決定は出題意図に大きく関わっている．

(2) **出題文の決定：** これは，教材のどの部分を問題として学生に示すかを表している．以下では，問題として学生に示す部分のことを出題文と呼ぶ．出題箇所が決まっても，出題箇所を含む前後の文脈がなければ，正解が一意に定まらなくなることがある．そのため，出題文は練習問題を作成する上で重要である．

(3) **選択肢の決定：** 選択問題では，選択肢を作るために，出題箇所の表現に代わる表現を複数用意しなければならない．選択肢を選定する場合には，出題箇所に入れても不自然でない表現を揃える．また，誤りに陥りやすい表現，ちょっと考えるとすぐに誤りと分かるような表現も含めるであろう．このように，選択肢にどのような表現を揃えるかは，出題意図に大きく左右される．

上記3つの項目は，全て出題意図に関係している．この様な情報を教材から自動的に抽出するのは困難である．そのため，AEGIS を設計する際には，出題者が教材にこれらの情報を書き込むことにした．次の節で述べるタグを使用して出題者が問題作成の情報を埋め込み，埋め込まれた情報を元に AEGIS が学生の理解度に合わせた出題を行う．このような作業分担を考えて設計し実装した．

3.2 問題自動生成のためのタグ

練習問題を自動的に作成する場合，上で述べた3つの情報を教材に埋め込む必要がある．AEGIS では新たにタグを定義し，それを用いて教材から練習問題を自動生成する．

- **出題文を示すタグ：** 教材のどの部分を出題文とするかを示すタグとして，QUESTION タグを用意した．このタグの開始タグと終

NII-Electronic Library Service

Information Processing Society of Japan

```
<QUESTION>
Data structures need to be studied
<DEL CAND="an,on,at,by"> in </DEL>
order to understand the algorithms.
</QUESTION>
```

図 2: 問題を生成するためのタグ付きデータ

了タグに囲まれた内容が練習問題の出題文になる.

- **出題箇所を示すタグ:** 出題箇所を指定するためにDEL タグを用意した. このタグの開始タグと終了タグで教材中の出題箇所を囲んで出題箇所を指定する. 囲まれた部分を空白にすれば, 穴埋め問題は自動的に作成可能である.

- **選択肢を示す属性:** 前節で述べたように選択問題を作成する場合が最も情報を必要とする. そのため, 選択問題を作成できるだけの情報をタグ付き教材は含んでないといけない. 選択肢は出題箇所に依存するので, 別のタグを用意して記述するのではなく, 出題箇所を示すタグの属性として記述できるようにするのが自然である. AEGIS のタグでは, DEL タグに属性CAND を用意してそれに対処した. 出題者が誤った表現を属性CAND の値として記述する.

ここまでで定義した2つのタグを使用すると, 図1の問題を得るためのタグ付きデータは図2の様になる. DEL タグの内容「in」を空白に置き換えると, 図1の空白問題になる. また, DEL タグのCAND 属性の値を使って選択肢を作れば, 図1の選択問題になる. CAND 属性の値の1つをDEL タグの内容と置き換えると, 図1の誤り訂正問題になる.

3.3 理解力に応じた出題のための情報

問題を出題するだけであれば, 上記の2つのタグで可能である. しかし, 学生の理解度に応じた練習問題を出題するには, 問題の難易度に関する情報が必要となる. AEGIS のタグでは, DEL タグの属性としてLEVEL, GROUP, REF の3つを用意して, 出題箇所の難易度を表すようにした. また, 出題箇所と関係する表現に印を付けるためのタグLABEL を追加した.

LEVEL 属性には出題箇所自体の難易度を指定する. この属性は, 1〜10 の数値の対になっており, 出題する学生の理解度の範囲を表す. 学生の理解度とLEVEL 属性の値を比較して, その DEL タグの箇所を出題するか否かを判断する.

GROUP 属性には出題箇所間の関連を指定する. 出題箇所には, 他の出題箇所が空白になることによって正解に到達しにくくなるものがある. これを出題箇所の依存関係と考える. 複数の出題箇所に同じGROUP 名を付けることで, 依存関係があるとことを明示する. 難易度を高くする場合には, 同じGROUP 名が付いている出題箇所をすべて出題する. 難易度を低くする場合には, 一部だけを出題し, 出題しなかった箇所を一種のヒントとすることもできる.

REF 属性には出題箇所以外の表現との関係を指定する. GROUP 属性は出題箇所同士の依存関係を明示するために導入したが, 出題箇所によっては, 出題箇所以外の表現と依存するものもあるであろう. この属性は, 次に述べるLABEL タグと対をなして, 教材中のどの部分と依存関係があるかを明示するために使用する.

練習問題の出題文の記述には, 出題箇所と関

NII-Electronic Library Service

Information Processing Society of Japan

```
<QUESTION> 出題文 </QUESTION>

<DEL CAND="CANDIDATE"
     LEVEL="PAIR"
     GROUP="ID"
     REF="ID"          > 解答 </DEL>
CANDIDATE ::= W_S | W_S,CANDIDATE
W_S       ::= 語，または記号
              記号がカンマ(,)か2重引用符(")，
              バックスラッシュ(\)を含む場合は，
              文字の前にバックスラッシュをつける．
PAIR      ::= LOW,HIGH
LOW       ::= 1～10までの数値
HIGH      ::= 1～10までの数値
ID        ::= 名前

<LABEL NAME="ID"> 参照内容 </LABEL>
ID        ::= 名前
```

図 3: 出題用のタグの定義(BNF 風)

係のある記述が存在しているであろう．ある記述がある出題箇所の正解を得るためのヒントとなる場合もある．そのような箇所をタグで示しておくと，誤った解答をした場合にヒントとして学生に示すことも可能である．AEGIS のタグでは，LABEL タグを用意している．LABEL タグにはNAME 属性があり，どの出題箇所と関係があるのかを指定する．

本研究で定義した3つのタグの定義を図3に示す．この図ではBNF 形式風に記述している．

4. AEGIS の概要

AEGIS は Perl を用いて作成され，Web ブラウザから CGI (Common Gateway Interface) として起動される．AEGIS は問題生成と採点の2つのコンポーネントから成っている．既存の教材に問題作成用のタグを埋め込んだ「タグ付き教材」を出題者が用意し，AEGIS はそれを利用する．また，学生ごとに作成するユーザプロファイルも利用する．ユーザプロファイルには，どの出題箇所をどの出題形式で出題したかという情報と採点した結果の情報とを採点の際に記録する．これを基にして AEGIS は学生の理解度を判断する．AEGIS の処理の流れを図4に示す．

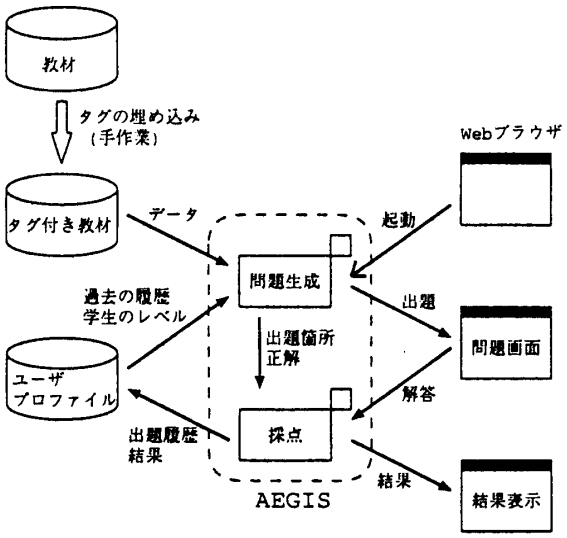


図 4: AEGIS による処理の流れ

AEGIS は，学生の要求によって Web ブラウザから起動されると，タグ付き教材とユーザプロファイルを読み込む．ユーザプロファイル中の過去の出題履歴をもとに学生の理解度を計算する．求めた理解度とDEL タグのLEVEL 属性の値とを比較して，その出題箇所を出題するか否かを判断する．また，過去の出題記録も参照して出題箇所を選択する．選択したDEL タグを含んでいるQUESTION タグを教材テキスト中から検索し，出題文を作成する．

次に，3つの出題形式のうちどの形式を選択するかを決定する．同じ出題箇所でも出題形式を変化させることで難易度が変化する．学生の理解度がLEVEL 属性に与えた範囲の下側に近い場合には選択問題として出題する確率を高くし，

NII-Electronic Library Service

Information Processing Society of Japan

前節では，2つの整数の和を計算し表示するプログラムを学んだ．算数には基本演算として4つの演算がある．これらに付いてはパスカルプログラムではどのように記述されるかをこの節では見ていく．

```
<QUESTION>
下のプログラムは，2つの整数を入力すると，積と商を
計算し表示するプログラムである．
program enzan;
uses wincrt;
var x, y : integer;
    seki, shou : integer;
begin
  write('Input two integers : ');
  readln(x, y);
  seki := <DEL CAND="x,xy,x X y,x mul y"
              LEVEL="1,5"> x * y </DEL>;
  shou := <DEL CAND="x / y,x ÷ y,xdivy,x mod y"
              LEVEL="1,5"> x div y </DEL>;
  writeln('積 : ', seki);
  writeln('商 : ', shou);
end.
</QUESTION>
```

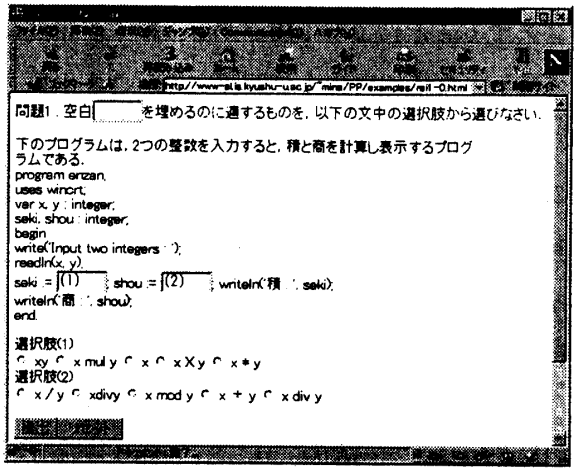
8行目で積を計算し，9行目で商を計算している．このことから，積と商の演算は，算数での記述とは異なる記述をすることを確認して欲しい．

図 5: タグ付き教材の例

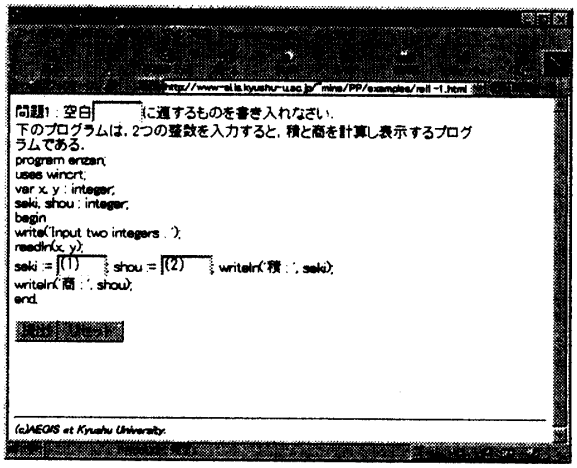
学生の理解度がLEVEL属性に与えた範囲の上側に近い場合には誤り訂正問題として出題する確率を高くする．この様にして，同じ出題箇所でも出題形式を変化させることで，学生の理解度に応じた出題を可能としている．

DELタグで囲まれた内容は，その出題箇所の正解になる．また，出題者によって選択肢の候補(誤りの選択肢)がCAND属性に記述してある．これらの両方を選択肢として学生に提示することで選択問題を生成する．また，誤り訂正問題では，DELタグで囲まれた内容をCAND属性に記述してある表現の1つで置き換える．これにより，問題文中に誤りを挿入する．

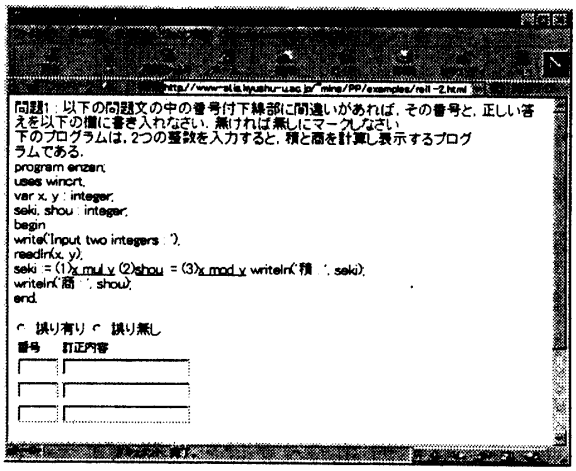
採点に関しては，DELタグで囲まれた内容がその出題箇所の正解であることを利用して，学生の解答を採点する．選択問題と穴埋め問題では，出題時に出題箇所と出題形式，正解を採点



(a) 選択問題での出題



(b) 穴埋め問題での出題



(c) 誤り訂正問題での出題

図 6: AEGISによる出題の表示例

NII-Electronic Library Service

Information Processing Society of Japan

側に伝え，学生が記入した解答と正解をマッチングすることによって採点する．誤り訂正問題では，上の情報に加えて問題文のどこを出題したのかを採点側に伝える．学生の解答の「誤りの場所」と「訂正内容」とが伝えられている情報とマッチングした場合を正解として採点する．採点が終了すると，出題箇所と出題形式，採点結果をユーザプロファイルに追記する．

我々が定義したタグを埋め込んだ教材の例を図5に示す．この例は，Pascal言語のプログラムの初歩の教材で積と商を計算する演算式の理解を目的としている．この教材の目的からして，プログラム中の積を計算する式と商を計算する式を出題箇所として選び，2つのDELタグで囲んでいる．さらにプログラム全体とプログラムの説明文とを出題文にするためにQUESTIONタグで囲んでいる．このタグ付き教材をAEGISで処理した結果が図6である．HTMLのフォームを使用して出題している．この図から，1つのタグ付き教材を使用して3種類の出題形式で出題できることがわかる．

5. おわりに

学生の練習問題挑戦への要求に応え，自動的に出題するシステムAEGISを構築した．教師が出題用のタグを教材に埋め込んでタグ付き教材を作成し，これを使用してAEGISが練習問題を自動生成する．出題に際しては，過去の学生の履歴を参考に学生の理解度を推定し，その理解度に応じた難易度の問題に変化させることが可能である．

現在，AEGIS用のタグは出題者がエディタで入力している．タグの埋め込み作業は労力が必要なので，それを支援するツールを作成する必要があろう．また，DELタグで囲む出題箇所は，解答として曖昧さがないようにタグ付けされなければならない．現在のAEGISでは，曖昧さがないことを出題者が保証することを前提にしている．曖昧さがある場合への対応は今後の課題である．また，AEGISは開発段階であり，講義等での実使用には至っていない．今後，多くのタグ付きデータを準備し，実際の講義の場で使用して評価することが必要である．

謝辞

この研究の一部は，九州大学教育研究プログラム・研究拠点形成プロジェクト(Cタイプ)の助成により行われた．

参考文献

[1] R.Fujimoto, A.Suganuma, and T.Mine, "Development of a Classroom Management System on the Web", International Conference on Computers in Education, Vol.2, pp.756–759 (1999).

[2] 菅沼 明，藤本 竜之介，峯 恒憲，"Webを用いた講義支援システムの開発とその適用"，情報処理教育研究集会，pp.631–634 (1999).

[3] T.Mine, D.Nagano, K.Baba, T.Shoudai, and S.Hirokawa, "On-Web visualizing a mechanism of a single chip computer for computer literacy courses", International Conference on Computers in Education 98, Vol.2, pp.496–499 (1998).

[4] R.Fujimoto, Y.Tsutsumi, and A.Suganuma, "CACCE : Computer Aided Cooperative Classroom Environment", World Conference on Educational Multimedia, Hypermedia & Telecommunications, p.1686 (1999).

NII-Electronic Library Service