Towards the Assessment of Free Learner's Utterances in CALL

Svetla Boytcheva, Irena Vitanova

Sofia University "St. Kliment Ohridski"

5 J. Bauchier Blvd., Sofia, Bulgaria, 1164

svetla@fmi.uni-sofia.bg, itv@gmx.co.uk

Albena Strupchanska, Milena Yankova, Galia Angelova

Linguistic Modelling Dept., IPP, BAS

25A Acad. G. Bonchev Str., Sofia, Bulgaria, 1113

{albena, myankova, galia}@lml.bas.bg

ABSTRACT

Computer-Aided Language Learning (CALL) should play an important role in the modern training process because it provides easy accessible, adaptive and flexible ways of learning. This paper addresses the scenario of tutor-learner question answering and attempts to automate the free answers evaluation using the advantages of Natural Language Processing (NLP). Our current approach integrates shallow parsing for analysing the answers and allows the learners to enter various utterances to express themselves. However this variety does not impede the assessment of the student's answer as we check the utterances against the automatically generated scope of the correct answers. The usage of a "set of answers" instead of one predefined correct answer enables feedback elaboration that helps learners to understand better their knowledge gaps. Briefly, in this paper we show how the combination of shallow and deep semantic NLP techniques can improve the effectiveness of eLearning systems which support communication in free natural language and can make them more satisfactory and pleasant for their users.

ARCHITECTURE
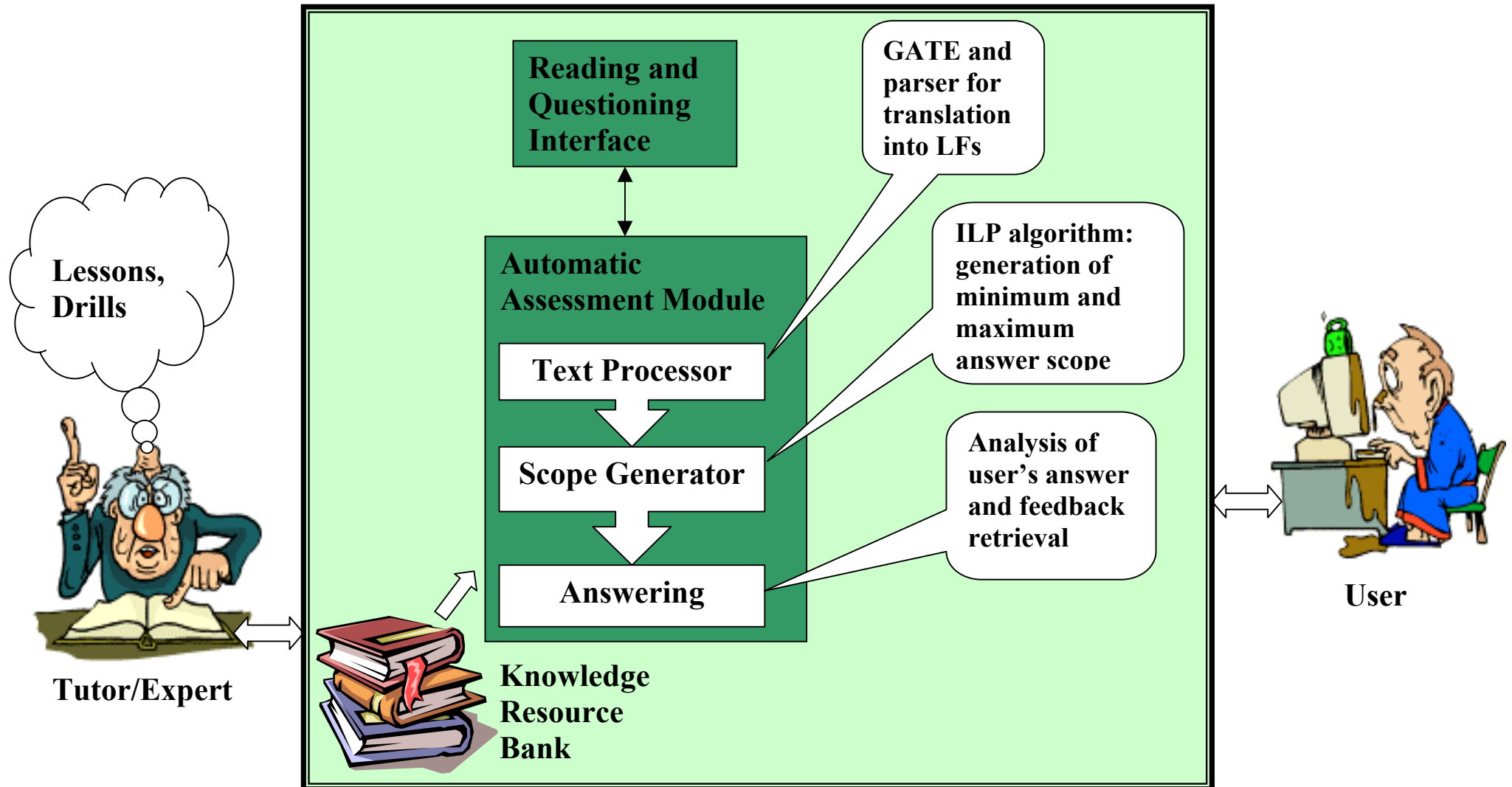
TEXT PROCESSOR

- GATE: lexical analysis and POS tagging
- Parser: left-recursive, top-down depth-first parser (SICSTUS Prolog)

The parser translates lessons's sentences into logical forms (LFs).

- all words are represented as predicates
- thematic roles are represented by predicate symbol "theta" and three arguments

Sentence:
The principal function of the primary market is to raise financial capital.

Logical form:
logical_form(is(_187,'VBZ')&theta(_187,agnt,_688)&function(_688,'NN')&
theta(_688,char,_833)&principal(_833,'JJ')&theta(_688,poss,_1084)&
market(_1084,'NN')&theta(_1084,char,_1217)&primary(_1217,'JJ')&
theta(_187,to,_1720)&raise(_1720,'VB')&theta(_1720,obj,_1878)&
capital(_1878,'NN')&theta(_1878,char,_2003)&financial(_2003,'JJ').

TEXT PROCESSOR

The parser:
- uses grammar rules and rules for translation into LFs;
- is tailored to the financial discourse i.e. it recognises important domain concepts and verbs;
- permits incomplete and syntactically incorrect user's answer;

A knowledge engineer checks and corrects the LFs if necessary, in order to receive one to one correspondences between the lesson sentences and the generated LFs.

Notes:
LFs are suitable representation of the domain knowledge.
LFs of sentences are necessary for performing effective inference.

TEXT PROCESSOR

A pedagogical expert provides both the lessons and the questions relevant to them. The text processor translates the questions into LFs.

Example

Question:

What is the function of primary market?

LF:

function(X) & be(Y) & theta(Y, ptnt, X) &
theta(Y, obj, Z) & univ(Z) & primary_market(A) &
theta (A, poss, X)

What $\longleftrightarrow$ univ(Z)

SCOPE GENERATOR

Problem: question $\Rightarrow$ answer in NL (many correct answers so language expression varies considerably)
Solution: Usage of a "*set of answers*" instead of *one* predefined correct answer:

- minimum (kernel) set i.e. the least generalisation under implication of all correct answers
- maximum (cover) set i.e. the greatest specialisation under implication of all correct answers

An algorithm uses the set of LFs produced from the lessons, the LF of question and some domain knowledge to constructs the sets.
The algorithm applies Inductive Logic Programming techniques.

EXAMPLE

Lesson:

...
Financial markets are the heart of the financial system.
They can be divided into different segments according to
the characteristics of the financial instruments traded
and the needs of the different participants. Primary
markets trade newly issued loans and securities. The
function of the primary market is to raise capital to
supports new investments.
The capital market finances long-term investment by
businesses, governments and households. The construction
of factories, highways, schools and homes relies mainly on
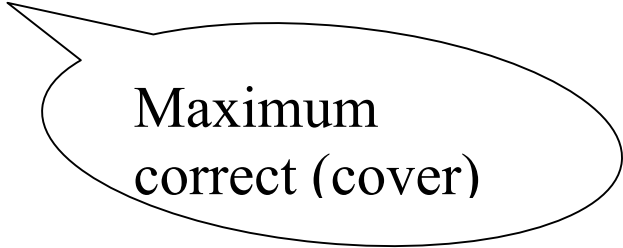the trading of funds on the capital market.
...

EXAMPLE

Question:

What is primary market?

Lesson's statements related to the question:

1) Primary markets trade newly issued loans and securities.
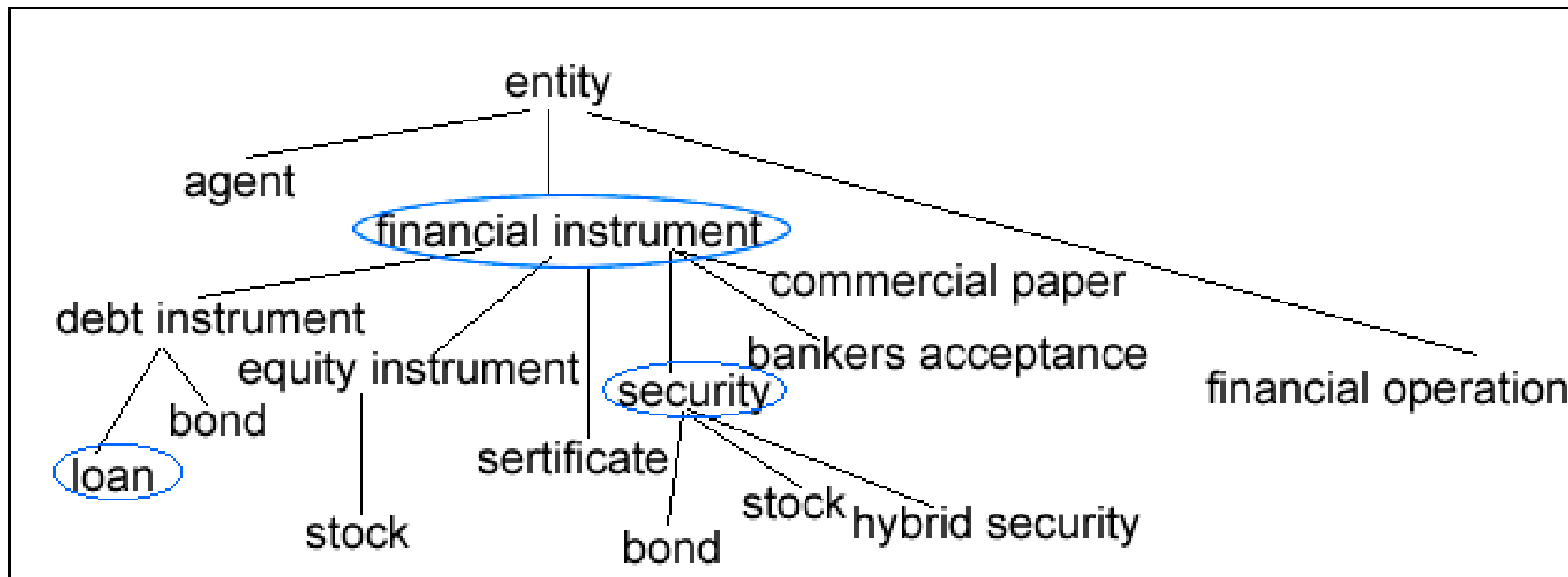2) The function of primary market is to raise capital to supports new investments.

Maximum correct (cover)

EXAMPLE

Minimum correct (kernel)

Primary market trades newly issued loans and securities.

ANSWERING MODULE

- Analysis of the learner's answer

Text processor converts learner's utterance into LFs. The usage of shallow parser permits users to use incomplete and syntactically incorrect sentences in their utterances.

- Check of learner's answer correctness

The module compares the logical form of the learners' utterance to the logical forms of the expected minimum and maximum answers.

- Feedback retrieval

Attempts to perform semantic analysis of errors when they are caused by:
  - Correct question understanding but absent knowledge of the correct term (implies paraphrases and generalisation);
  - Correct question understanding but absent domain knowledge (implies specialisations as well as partially correct, incomplete and wrong answers).

ANSWERING MODULE

- Maximum correct (cover) - MCC
- Minimum correct (kernel) - MCK
- Logical model of user's answer
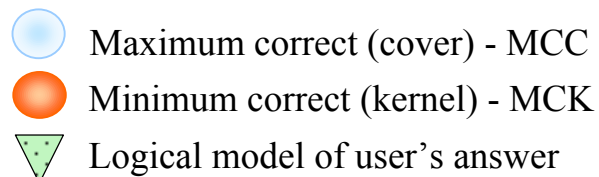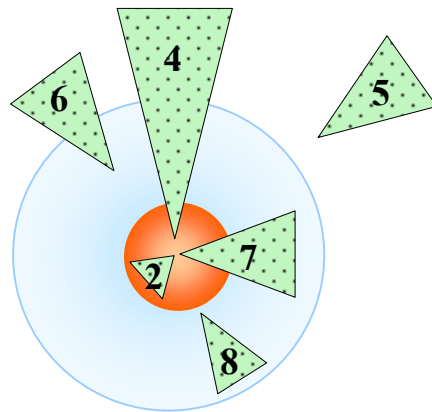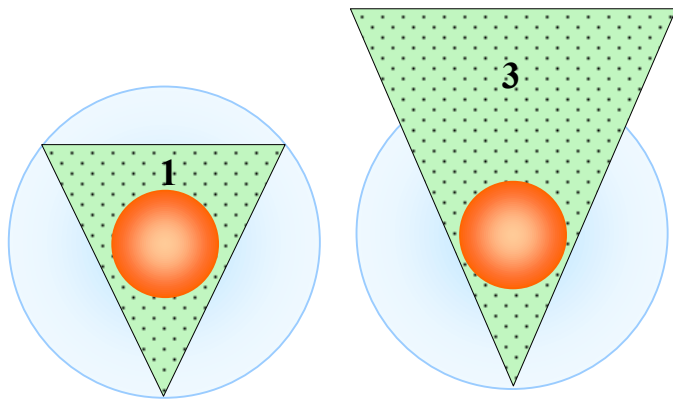
- **correct** - 1
- **wrong**
  - 5: completely wrong
  - 6: contains true and wrong statements (missing statements from MCK)
- **incomplete** (contains no wrong statements)
  - 2: contains parts of MCK, missing MCC
  - 7: contains parts of MCK and MCC
  - 8: missing MCK, contains parts of MCC
- *more specific*
  - 2: contains concepts that are specialisation of the expected ones
- *paraphrase* (usage of concept definition instead of the proper term) – 7, 8
- *partially correct* (contains true and wrong statements)
  - 3: contains MCK and wrong statements
  - 4: contains parts of MCK and MCC but also wrong statements
- *more general*
  - 3: contains MCK and concepts that are generalizations of the expected ones

ANSWERS AND FEEDBACK

(A1) *Primary market trades newly issued loans and securities.*
Primary market supports new investments.

(F1) `correct`

(A2) *Primary market* is an ***organisation in which total worth is divided into commercial papers***.

(F2) `wrong`
Wrong: ***organisation in which total worth is divided into commercial papers***

ANSWERS AND FEEDBACK

(A3) *Primary market trades* **financial instruments**.

(F3) `partially correct, more genral`
Missing: *newly issued*
Generalisation: **financial instruments** are generalisation of *loans and securities*
(see slide N.10 where the extract of the hierarchy of terms is shown)

(A4) *The principal function of the primary market is to raise financial capital to support new investment* **in buildings, equipment and inventories**.

(F4) `partially correct`
Missing: *Primary market trades newly issued loans and securities.*
Wrong: **buildings, equipment and inventories**

KNOWLEDGE RESOURCES

- Pedagogical resources: provided by the tutor

- Domain knowledge:

  - hierarchy of all important domain concepts;
  - synonym sets for the concepts and for some of the domain related verbs;
  - knowledge base with assertions and definitions of domain relevant terms as well as definitions of relations.

Note: We reuse the domain knowledge, which was developed in the project LARFLAST. The assertions and definitions of domain relevant terms were presented as Conceptual Graphs (CGs) so we have automatically converted the CGs into LFs.

EVALUATION

An implemented a NLP module that performs deep semantic analysis in LARFLAST project was restrictive to users as it forces them to answer with complete and syntactically correct sentences. So we tried to reduce these limitations by applying shallow and partial NLP techniques in Information Extraction style. The students like that they can answer by sentence phrases only in a relatively liberal style and that the word order is irrelevant (as the logical forms are conjunctive terms). At the same time the prototype is not over-permissive as (roughly) each unnecessary word leads to a wrong answer.

We admit however that the semantic analysis as such is a rather expensive task. Defining the domain knowledge and testing the inference and the diagnostics procedures take man-years (in our case this was done in a previous project). Writing the lessons and the related questions in the corresponding way takes time too, exactly as the checks of the logical forms of the lessons and the questions. These efforts make sense only if they are multiplied and the resources are reused in a larger context.

CONCLUSION

Although the task of automatic processing of free NL in CALL is very hard and, similarly to NLU, cannot be solved completely, we believe that the combination of shallow and deep NLP techni-ques is an attempt to improve the up-to-date CALL solutions especially when the expected learner utterances are relatively short and well-focused. At the same time the communication in NL is more effective and more attractive for the student, so we expect further projects and new attempts to approach the semantic analysis in CALL.

REFERENCES

- Weinberg, J. Garman, J. Martin and P. Merlo 1995. A Principle-Based Parser for foreign Language Tutoring in German and Arabic. *In Intelligent Language Tutors: Theory Shaping Technology*, Lawrence Erlbaum Ass. UK, 23-44.

- Dorr, J. Hendler, S. Blanksteen and B. Migdaloff. 1995. On Beyond Syntax: Use of Lexical Conceptual Structure for Intelligent Tutoring. *In Intelligent Language Tutors: Theory Shaping Technology*, Lawrence Erlbaum Associates, UK, 289-310.

- Rose, P. Jordan, M. Ringenberg, S. Siler, K. VanLehn and A. Weinstein. 2001. Interactive Conceptual Tutoring in Atlas-Andes. *AI in Education*, J.D. Moore, et al (Eds), IOS Press.

- Rose. 2000. A Framework for Robust Semantic Interpretation, Proc. of the *1st Annual Conf. of the North American Chapter of the Association for Computational Linguistics (*NAACL'00), 1129-1135.

- C. Rose et al. 2003. Overcoming the Knowledge Engineering Bottleneck for Understanding Student Language Input, Proc. of *AI in Ed.*

- G. Angelova, S. Boytcheva, O. Kalaydjiev, S. Trausan-Matu, P. Nakov and A. Strupchanska. 2002. Adaptivity in Web-Based CALL, *Proc. of ECAI - 2002*, 445-449.

- GATE 2003. See http://gate.ac.uk/

- K. VanLehn et al. 2002. The architecture of Why2-Atlas: A coach for qualitative physics essay writing, Proc. *Intelligent Tutoring Systems Conference, volume LNCS 2363*, 158—167.

- L. Michaud, K. McCoy and L. Stark. 2001. Modeling the Acquisition of English: an Intelligent CALL Approach, *Proc. of the 8th Int. Conf. on User Modeling*, July 2001, 14-23.

REFERENCES

- M. Glass. 2000. Processing Language Input in the CIRCSIM-Tutor Intelligent System. *AAAI 2000 Fall Symp. on Building Dialogue Systems for Tutorial Applications*, http://www.csam.iit.edu/ ~circsim/index.html

- M. Sams. 1995. Advanced Technologies for Language Learning: the BRIDGE Project within the ARI Language Tutor Program. *In Intelligent Language Tutors: Theory Shaping Technology*, Lawrence Erlbaum Associates, UK, 17-21.

- N. Nagata. 2002. BANZAI: An application of na-tural language processing to web based language learning, *CALICO Journal*, 19(3): 583-599.

- R. Freedman. 1999. Atlas: A Plan Manager for Mixed-Initiative, Multimodal Dialogue. *AAAI-99 Workshop on Mixed-Initiative Intelligence*, Orlando, FL.

- R. Freedman et al. 2000. ITS Tools for natural language dialogue: A domain-independent parser and planner, *Proc. Intelligent Tutoring Systems: 5th Int. Conf., ITS 2000, G. Gauthier, et al (Eds)*, Springer: Berlin, 433-442.

- RECALL, a Telematics Language Engineering project (1997), see http://iserve1. infj.ulst.ac.uk/~recall. (CASTLE was developed at the IBM Centre in Heidelberg, Germany).

- Sv. Boytcheva. 2002. ILP Techniques for Free-text Input Processing, *In Proc. of AIMSA, 10th Int. Conf. on AI: Methodology, Systems and Applications, Springer, LNAI 2443*, 101-110.

- V. Aleven, O. Popescu, and K. Koedinger. 2001. Pedagogical content knowledge in a tutorial dialogue system to support self-explanation. *In Papers of the AIED-2001 Workshop on Tutorial Dialogue Systems*, 59-70.