A Polyphonic Model and System for Inter-Animation Analysis in Chat Conversations with Multiple Participants

Stefan Trausan-Matu[1,2], Traian Rebedea[1]

[1] "Politehnica" University of Bucharest, Department of Computer Science and Engineering,
313 Splaiul Independetei, Bucharest, Romania
[2] Research Institute for Artificial Intelligence of the Romanian Academy,
13 Calea 13 Septembrie, Bucharest, Romania
stefan.trausan@cs.pub.ro, traian.rebedea@cs.pub.ro

Abstract. Discourse in instant messenger conversations (chats) with multiple participants is often composed of several intertwining threads. Some chat environments for Computer-Supported Collaborative Learning (CSCL) support and encourage the existence of parallel threads by providing explicit referencing facilities. The paper proposes a discourse model for such chats, based on Mikhail Bakhtin's dialogic theory. It considers that multiple voices (which do not limit to the participants) inter-animate, sometimes in a polyphonic, counterpointal way. An implemented system is also presented, which analyzes such chat logs for detecting additional, implicit links among utterances and threads and, more important for CSCL, for detecting the involvement (inter-animation) of the participants in problem solving. The system begins with a NLP pipe and concludes with inter-animation identification in order to generate feedback and to propose grades for the learners.

Keywords: Discourse analysis, conversation, chat, dialogism, polyphony, Computer-Supported Collaborative Learning

1 Introduction

The goals followed by our approach and, meanwhile, the contributions of this paper are both theoretical and practical. First, we propose polyphony as theoretical model of a particular kind of online conversations: instant messenger (chat) conversations with multiple participants. The practical goal was to implement a system for analyzing such chats and providing feedback in order to encourage the appearance of multiple voices (or positions, in an extended sense [1, 2]), of parallel and intertwining threads of discussions. These aims may be obtained by catalyzing debates and the inter-animation of the participants, which are premises for supporting understanding, studying and creative thinking of virtual teams of learners or researchers. The implemented system was developed as a module in the EU FP7-IST project

'Language Technologies for Lifelong Learning' (LTfLL) and it is now in the evaluation phase [3].

The implemented analysis method integrates results from NLP (content and discourse analysis), Social Networks Analysis (SNA) [4] and, a novel idea, the identification of polyphonic threading in chats [2]. The system was used for Computer Supported Collaborative Learning (CSCL) [5] in assignments for computer science and engineering students. As preparation for these assignments, the tutors group students in small teams of 4-7 participants, each of them being assigned a topic to study and then to support it in chat debates. They read some materials about that topic in order to understand the subject in detail. During the discussions, they present their points of view, they debate and inter-animate (arguing on their assigned topics), all of these improving their own and the others' understanding of the domain. After concluding a chat session, they can launch several widgets from the system, which provide graphical and textual feedback and preliminary scores both for each student and for the group as a whole. The tutors also use the system for providing them insights for writing a detailed feedback and grading the students.

The paper continues with a section introducing some basic theoretical ideas used in the system. The third section presents the implemented system.

2 Polyphony and Inter-animation

For discourse analysis in NLP two different situations are usually considered: monologue and dialogue. In monologues, an unidirectional model of communication is considered, from a speaker to a listener [6]. One of the main ways of analyzing discourse is the detection of local relations and measuring coherence, like in the Rhetorical Schema Theory (RST) [7], which considers a hierarchical decomposition of a text, like Centering Theory [8], or in other co-reference resolution systems [6].

In dialogues usually a phone-like (or face-to-face) type of conversation is considered. Typically, speech acts, dialog acts or adjacency pairs [6] are the units of analysis. Even if there are attempts to analyze conversations with multiple participants using transacts [10], this approach is also based on a two interlocutors' model. For chats, TF-IDF [11, 12], Latent Semantic Analysis [12, 13, 14], Naïve Bayes [15], Social Network Analysis [13], WordNet (wordnet.princeton.edu) [11, 13], Support Vector Machines and Collin's perceptron [10], and the TagHelper environment [16] are used for detection of topics and links [11], dialog acts [15], lexical chains [13] or other complex relations [16].

In phone and face-to-face-like dialogs only one person usually speaks at a given moment in time, determining a single thread of discussions. However, some chat environments, like the one used in the Virtual Math Teams (VMT) project [17] offer explicit referencing facilities, which means that users may indicate to which previous utterance(s) they refer to. This facility is extremely important in chat conversations with more than two participants because it allows the existence of several discussion threads or voices, in parallel. The co-occurrence of several voices gives birth to inter-animation and polyphony, phenomena identified in any text by Mikhail Bakhtin [18].

Voices may be considered as particular positions, which may be taken by one or more persons when they emit an utterance, which may have both explicit (like those

provided by the VMT chat environment [17]) and implicit links (for example, lexical chains, co-references or argumentation links) and influence other voices. Each utterance is filled with 'overtones' of other utterances [1]. Moreover, by the simple fact that they co-occur, voices are permanently *inter-animating*, entering in competition, generating multivocality in any conversation and even in any text (in Bakhtin's dialogic theory everything is a dialog [18]) or, as Bakhtin calls it, a "*heteroglossia*, which grows as long as language is alive" [1].

In order to detect overtones and inter-animation in chats, in our system we start from the explicit and implicit links among utterances. Thus, a graph is constructed connecting utterances and, in some cases, words. In this graph, threads may be identified. Each thread may be considered as a voice which becomes less or more powerful than the others. Among chat voices, both sequential and counterpointal, transversal relations similar to polyphonic music may be identified [18, 2]. From these data, several measures of contributing to the conversation may be computed, for each participant and for the group as a whole.

3 Automatic Analysis of Chats with Multiple Participants

The input of the system for analysis and giving feedback is a chat log similar to the one presented in Figure 1. An XML schema was designed for encoding chat conversations and discussion forums. Each utterance has an unique identifier, ('genid') and the existing explicit references ('ref') to previous utterances, which were specified by the participants using the facility provided by the VMT environment. In addition to annotating the elements of a chat, the schema also includes at the end data generated by the system.

The input data may be in different formats besides the above XML schema. A preprocessing module transforms these formats to respect the XML schema. The supported formats are: saved chats from Yahoo Messenger in text format, other text format chats, VMT format.

Figure 2 presents an overview of the architecture of the system and specifies the communication between the modules. Some of these modules – the ones that are heavily based on NLP technologies – are presented in detail in the following subsections. Others, like the one used for Social Network Analysis were presented in other papers [4].

3.1 The NLP Pipe

The processing starts with a NLP pipe containing the following: spelling correction, stemmer, tokenizer, Named Entity Recognizer, POS tagger and parser, and NP-chunker. The components of the NLP pipe are mainly those provided by the Stanford NLP software group (http://nlp.stanford.edu/software), with the exception of the spell checker (which uses Jazzy, http://www.ibm.com/developerworks/java/library/j-jazzy/ and http://jazzy.sourceforge.net/). Two alternative NLP pipes are under development, integrating modules from GATE (http://gate.ac.uk) and LingPipe (http://alias-i.com/lingpipe/).

```
<?xml version="1.0" encoding="UTF-8" ?>
 <Dialog time="2005-01-11 09:26:11" description="this is an assignment for the NLP
       course" file="chat_input_1.xml" id="Social networks13_6_200610_57_10"
       language="en|fr|ro" name="chat-12-A" subject="about pragmatics" team="12">
 <Participants>
  <Person nickname="Alex" realname="Bibi Ionescu" />
  <Person nickname="vvalcea" realname="" />
  <Person nickname="Adrian" />
 </Participants>
 <Topics>
  <Itemset description="NLP - pragmatics">
   <Item>speech act</Item>
   <Item description="cnf. Grice's theory">implicature</Item>
 </Itemset>
   <Itemset>.....</Itemset>
 </Topics>
 <Body>
  <Turn nickname="Alex">
   <Utterance genid="1" ref="0" time="2005-01-11 09:26:03"> hello all </Utterance>
  </Turn>
  <Turn nickname="Adrian">
   <Utterance genid="2" ref="0" time="2005-01-11 09:27:18">hi</Utterance>
  </Turn>
   <Turn nickname="vvalcea">
    <Utterance genid="3" ref="1" time="2005-01-11-09:29:29"> Hello Alex </Utterance>
  </Turn>
.....
  </Body>
</Dialog>
```

Figure 1. A fragment of a chat log encoding

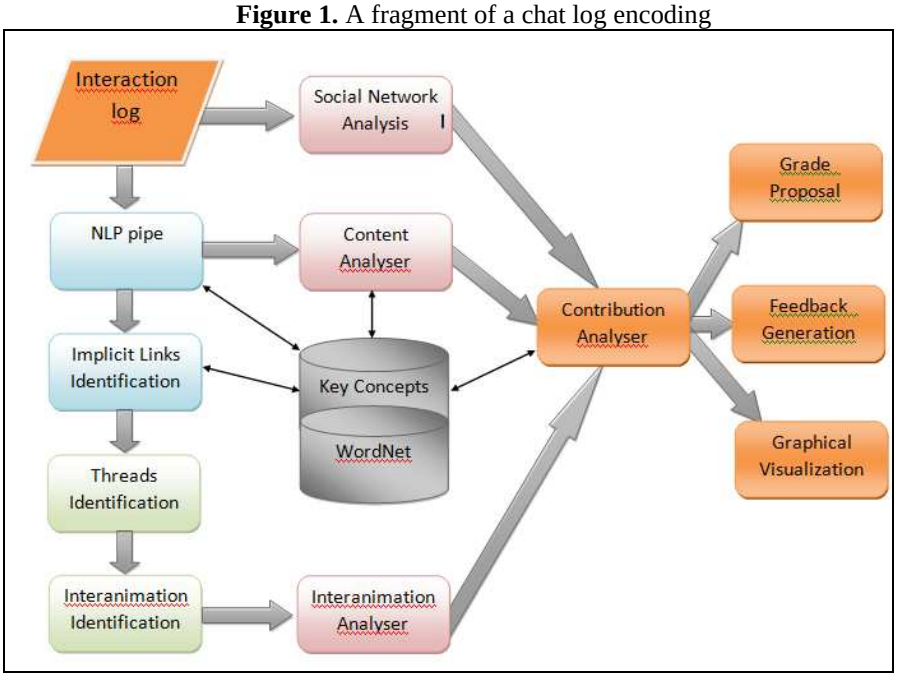

Figure 2. Main modules of the analysis and feedback system

3.2 Pattern Language

Because important parts of the processing in the system are based on patterns identified by cue phrases, a module, called 'PatternSearch' was implemented for searching occurrences that match expressions specified by the user in a log of a chat or a forum. In addition to a simple regular expression search, the module allows considering not only words, but also synonyms, hypernyms and hyponyms via WordNet, words' stems and their part of speech (POS). Another novel facility is the consideration of utterances as a search unit, for example, specifying that a word should be searched in the previous *n* utterances and that two expressions should be in two utterances.

For example, the expression `<S "convergence"> #[*] cube` searches pairs of utterances that have a synonym of "convergence" in the first utterance and "cube" in the second. One result from a particular chat is the pair of utterances 1103 and 1107:

```
1103 # 1107.  overlap # cube [that would stil have to acount for the
overlap that way] # [an idea: Each cube is assigned to 3 edges. Then add
the edges on the diagonalish face.]
```

The search is made at utterance level - the program checks the utterances one by one (and if there is a match between a part of the utterance and the searched expression, both the utterance and the specific text that matched are indicated).

PatternSearch is used in several other modules: cue-phrases identification, implicit links identification and adjacency pairs identification.

3.3 Content Analysis

The content analysis identifies the main concepts of the chat or forum using the NLP pipe, cue-phrases and graph algorithms [2]. It also identifies speech acts (a set derived from DAMSL [19]) and argumentation types in utterances (as in Toulmin's theory [20]: Warrant, Concession, Rebuttal and Qualifiers). Concepts and their synonyms are searched in the lexical database WordNet (wordnet.priceton.edu) and in a collection of key concepts and their inter-relations for the subject, provided by the teacher.

Advanced NLP and specific discourse analysis identify various types of implicit links:

- Repetitions (of ordinary words or of Named Entities), which were identified by Tannen as very important for detecting the involvement of the participants in a conversation [21];
- Lexical chains, which identify relations among the words in the same post / utterance or in different ones, by using semantic similarity measures based on WordNet;
- Adjacency pairs [6] – pairs of specific speech acts – e.g. answers to a single question in a limited window of time (in which the echo of the "voice" of the question remains), greeting-greeting;
- Co-references (the BART system [22] is used – see also http://bart-coref.org/)

3.4 Words, Key Concepts, Voices, and Threads

In the implementation of our analysis tool, we start from the key concepts and associated features that have to be discussed and that are provided by the teacher. Each participant is assigned to support a position which corresponds to a key concept. That corresponds to a kind of implicit voice emitting that concept and the associated features. We may identify other, additional voices in the conversation by detecting recurrent themes, new concepts. Therefore, a first, simple perspective is to have a word-based approach on voices: We consider that a repeated word (that is a noun, verb, adjective or adverb) becomes a voice [21, 2]. The number of repetitions and some additional factors (e.g. presence in some specific patterns) may be used to compute the strength of that voice (word).

We use voices to keep track of the position that each participant has to support, in order to identify divergences and conjunctions. This position is, as mentioned above, an implicit voice. For a given small period of time, the last utterances are echo-like voices. For example, answers may be associated to questions that are present in a given time window.

Voices continue and influence each other through explicit or implicit links. In this perspective, voices correspond to chains or threads of utterances. They may be a reasoning or argumentation chain [20], a chain of rhetorical schemas, chains of co-references, lexical chains and even only chains of repeated words, in the idea of Tannen [21]. The identification of argumentation chains, rhetorical schemas or co-references in texts and conversations are very difficult tasks for Natural Language Processing. Chains of repeated words, however, are very easy to detect, the sole problem being the elimination of irrelevant repeated words. Lexical chains can also be detected, but their construction is more difficult and the resulted lexical chains are greatly influenced by the choice of the ontology and similarity measures.

3.5 Polyphony, Inter-animation and Collaboration

In polyphony, the most advanced kind of music compositions, a number of melodic lines (or "voices," in an extended, non-acoustical perspective) jointly construct a harmonious musical piece, generating variations on one or several themes. Dissonances should be resolved, even if several themes (melodies) or theme variations are played simultaneously, and even if sometimes the voices situate themselves in opposing positions.

Voices in polyphonic music have two dimensions, the sequential threading of utterances or words and the transversal one implicitly generated by the coincidence of multiple voices. In addition, another dichotomy, the unity-difference (or centrifugal-centripetal [1]) opposition may also be observed.

The evaluation of the contributions of each learner considers several features like the coverage of the expected concepts, readability measures, the degree to which they have influenced the conversation or contributed to the inter-animation. In terms of our polyphonic model, we evaluate to what degree they have emitted *sound* and *strong* utterances that influenced the following discussion, or, in other words, to what degree the utterance became a strong voice [2].

The automatic analysis considers the inter-animation patterns in the chat [2]. It uses several criteria such as the presence in the chat of questions, agreement, disagreement or explicit and implicit referencing. In addition, the strength of a voice (of an utterance) depends on the strength of the utterances that refer to it. If an utterance is referenced by other utterances that are considered important, obviously that utterance also becomes important [2].

By using this method of computing their importance, the utterances that have started an important conversation within the chat, as well as those that began new topics or marked the passage between topics, are more easily emphasized. If the explicit relationships were always used and the implicit ones could be correctly determined in as high a number as possible, then this method of calculating the contribution of a participant would be considered [2, 4].

The implemented system supports the analysis of collaboration among learners: It produces different kinds of information about discussions in chat and forum discussions, both quantitative and qualitative, such as various metrics, statistics and content analysis results such as the coverage of the key concepts related to executing a task and the understanding of the course topics or the inter-threaded structure of the discussion [2, 23]. In addition, the system provides feedback about the involvement of each learner, generates a preliminary assessment and visualizes the interactions and the social participation. Finally, the system identifies the most important chat utterances or forum posts (that express different opinions, missing topics/concepts, misleading posts, misconceptions or wrong relations between concepts).

The results of the contribution analyzer are annotated in the XML file of the chat or forum. The annotations are on utterances:

```
<UtteranceFeedback genid="53">
        <Grade type="overall">8.15</Grade>
        <SpeechAct>Continuation</SpeechAct>
        <SpeechAct>Info Request</SpeechAct>
        <SpeechAct>Statement</SpeechAct>
        <Argumentation>Claim</Argumentation>
</UtteranceFeedback>
```

and on participants:

```
<GeneralGrade nickname="AlexI">
      <Grade type="diction">20.07</Grade>
      <Grade type="spelling">13.16</Grade>
      <Grade type="fluency">25.63</Grade>
      <Grade type="pageRanking">20.44</Grade>
      <Grade type="utteranceStructure">22.89</Grade>
      <Grade type="nbrWordsProc">27.37</Grade>
      <Grade type="nbrDiffWordsProc">24.98</Grade>
      <Grade type="nbrUtterancesProc">25.06</Grade>
      <Grade type="nbrUtterancesProc">23.19</Grade>
      <Grade type="meanUtteranceWords">13.21</Grade>
      <Grade type="correctWordsProc">13.16</Grade>
      <Grade type="flesch">55.18</Grade>
      <Grade type="kincaid">8.23</Grade>
      <Grade type="fog">10.35</Grade>
      <Grade type="inDegree">25.51</Grade>
      <Grade type="outDegree">23.19</Grade>
```

```
        <Grade type="rank">22.09</Grade>
        <Grade type="eigen">100.01</Grade>
        <Grade type="closeness">16.79</Grade>
        <Grade type="centrality">17.46</Grade>
</GeneralGrade>
```

For the values describing the activity of the participants, Social Network Analysis, Latent Semantic Analyses and other techniques were used [4, 23]. These values are used for generating textual feedback, which include, besides the above numerical values:

- the list of most important (used, discussed) concepts in a chat / forum;
- the coverage of the important concepts specified by the tutor;
- the most important utterances of each participant (the ones with the largest scores)
- the score for an utterance (which uses a complex formula that takes into account the concepts used, dialog acts, the links between utterances and SNA factors [2, 23]);
- a score for each participant in the conversation;
- areas of the conversations with important collaboration (inter-animation, argumentation, convergence and divergence);
- other indicators and statistics that are going to be added with the development of the system.

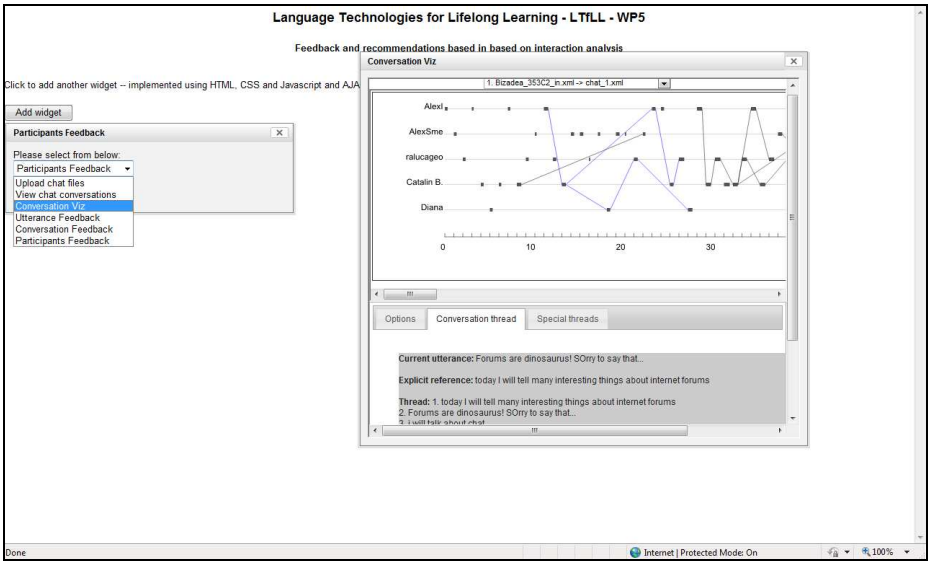

Figure 3. A screenshot illustrating the graphical feedback and analysis system

As graphical feedback, the service provides interactive visualization and analysis of the conversations graph with filtering enabled. The graphical representation of chats was designed to facilitate an analysis based on the polyphony theory of Bakhtin and to permit the best visualization of the conversation. For each participant in the chat,

there is a separate horizontal line in the representation and each utterance is placed in the line corresponding to the issuer of that utterance, taking into account its positioning in the original chat file – using the timeline as an horizontal axis (see Figure 3). Each utterance is represented as a rectangular node having a horizontal length proportional with the textual length of the utterance. The distance between two different utterances is proportional to the time between the utterances [2].

An image of the facilities of graphical and textual visualization is presented in Figure 3.

4 Conclusions

A new theory, inspired from Bakhtin's ideas was proposed for explaining and evaluating collaboration and inter-animation in chats. Its main idea is the consideration of intertwining of discussion threads similarly with counterpoint in polyphonic music. Graphical visualization and various metrics are computed using a wide range of NLP techniques for the lexical, semantic and discourse analysis levels.

The first experiments with the implemented system showed that the polyphony model eases the development of algorithms and implementation of a system that analyses and gives feedback to participants in chats.

Acknowledgements

We would like to thank the students of "Politehnica" University of Bucharest, Department of Computer Science, which participated to our experiments. The research presented in this paper was partially performed under the FP7 EU STREP project LTfLL and the national CNCSIS grant K-Teams.

References

1. Bakhtin, M., The Dialogic Imagination: Four Essays, University of Texas Press, Austin (1981)
2. Trausan-Matu, S., Rebedea, T., Polyphonic Inter-Animation of Voices in VMT, in Stahl.G. (Ed.), Studying Virtual Math Teams, Springer, pp. 451 – 473 (2009)
3. Berlanga, A. J., Rosmalen, P. V., Trausan-Matu, S., Monachesi, P., & Burek, G., The Language Technologies for Lifelong Learning Project. In D. S. I. Aedo., N. Chen, Kinshuk, Proceedings of the 9th IEEE International Conference on Advanced Learning Technologies, Riga, pp. 624-625 (2009)
4. Dascalu, M., Chioasca, E.V., Trausan-Matu, S., ASAP-An Advanced System for Assessing Chat Participants, in D. Dochev, M. Pistore, and P. Traverso (Eds.): AIMSA 2008, LNAI 5253, Springer, Heidelberg, pp. 58–68 (2008)
5. Stahl, G.: Group Cognition: Computer Support for Building Collaborative Knowledge. MIT Press, (2006)
6. Jurafsky, D. and Martin, J.H., Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Second Edition, Pearson Prentice Hall, (2009)

7. Mann, W. C. and Thompson, S. A., Rhetorical structure theory: A theory of text organization. Tech. rep. RS-87-190, Information Sciences Institute, (1987)
8. Grosz, B. J., Joshi, A. K., and Weinstein, S. Centering: A framework for modeling the local coherence of discourse. Computational Linguistics, 21(2), 203–225, (1995)
9. Dysthe, O., 'The Multivoiced Classroom: Interactions of Writing and Classroom Discourse'. Written Communication, 13.3: 385-425, (1996)
10. Joshi, M., Rosé, C. P.: Using Transactivity in Conversation Summarization in Educational Dialog. In Proceedings of the SLaTE Workshop on Speech and Language Technology in Education, (2007)
11. Adams, P. H., Martell, C. H., Topic detection and extraction in chat. In Proceedings of the 2008 IEEE International Conference on Semantic Computing, pp. 581-588, (2008)
12. Schmidt, A.P., Stone, T.K.M.: Detection of topic change in IRC chat logs. http://www.trevorstone.org/school/ircsegmentation.pdf
13. Dong, A.: Concept formation as knowledge accumulation: A computational linguistics study. Artif. Intell. Eng. Des. Anal. Manuf. 20, 1, 35-53 (2006)
14. Manning, C., Schütze, H., Foundations of statistical Natural Language Processing, MIT Press, Cambridge (Mass.), (1999)
15. Kontostathis, A. , Edwards, L., Bayzick, J., McGhee, I., Leatherman, A. and Moore, K.: Comparison of Rule-based to Human Analysis of Chat Logs. In 1st International Workshop on Mining Social Media Programme, Conferencia de la Asociación Española para la Inteligencia Artificial, (2009)
16. Rose, C. P., Wang, Y.C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., Fischer, F.: Analyzing Collaborative Learning Processes Automatically: Exploiting the Advances of Computational Linguistics in Computer-Supported Collaborative Learning , International Journal of Computer Supported Collaborative Learning, (2007)
17. Stahl., G. (Ed.), Studying Virtual Math Teams, Springer, Boston (2009)
18. Bakhtin, M. M.: Problems of Dostoevsky's poetics. University of Minnesota Press, Minneapolis (1993)
19. DAMSL, http://www.cs.rochester.edu/research/cisd/resources/damsl/RevisedManual/, downloaded on 22 November, 2009
20. Toulmin, S., The Uses of Arguments. Cambridge Univ. Press, (1958)
21. Tannen, D.. Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse, Cambridge University Press, (1989)
22. Versley, Y., Ponzetto, S.P., Poesio, M., Eidelman, V., Jern, A., Smith, J., Yang, X., Moschitti, A., BART: A Modular Toolkit for Coreference Resolution. Companion Volume of the Proceedings of the 46th Annual Meeting of the Association for Compuatational Linguistics, (2008)
23. Dascalu, M., Trausan-Matu, S., Dessus, P., Utterances Assessment and Summarization in Chat Conversations, to appear in CICLING'2010 Complementary Proceedings Volumes (2010)