An Agent that Helps Children to Author Rhetorically-Structured Digital Puppet Presentations

Paola Rizzo[1], Erin Shaw[2] and W. Lewis Johnson[2]

[1] Department of Computer Science
University of Rome "La Sapienza"
Via Salaria 113, 00189 Rome, Italy
rizzo@dsi.uniroma1.it
http://www.dsi.uniroma1.it/~rizzo

[2] Center for Advanced Research in Technology for Education
Information Sciences Institute, University of Southern California
4676 Admiralty Way, Marina del Rey, CA 90292-6695 USA
{shaw,johnson}@isi.edu
http://www.isi.edu/~shaw/
http://www.isi.edu/isd/johnson.html

Abstract. This paper describes a pedagogical agent that helps children to learn to author structured presentations about explanations of concepts. Using a Rhetorical Structure Theory analysis of a source Web page, the agent performs pedagogical tasks to support the user's understanding of rhetorical relations, stimulates reflection about the relations between the structure of the original text and the structure of the presentations, and suggests ways to improve the user's performance. Upon completion of the authoring, the presentations are organized into coherent structures that can be performed by animated characters, or Digital Puppets, in a learning-by-teaching classroom context.

1. Introduction

When properly designed, multimedia presentations result in deeper learning, compared to equivalent textual presentations [1]. Likewise, authoring multimedia presentations forces students to organize their thoughts and clarifies their understanding of the subject material, encouraging the development of important procedural and metacognitive skills while achieving mastery of the subject matter.

To support both multimedia presentation and multimedia authoring, we are developing a system aimed at young children for authoring and generating Digital Puppet presentations. Digital Puppets (DPs), like animated pedagogical agents (APAs) [2], are animated characters that help learners understand a subject. Like APAs, DPs use text-to-speech software and a variety of nonverbal gestures to provide voiceover narration and personalized commentary, which have been shown to be particularly effective multimedia presentation methods [3]. Like APAs, we expect DPs to evoke a positive affective response in the viewer, referred to as the *persona effect*, which is produced even by lifelike characters that do not perform autonomous

pedagogical behaviors [4]. Unlike APAs, however, DPs are not necessarily intelligent or autonomous; hence, we call them puppets.

How can we assist students in structuring their knowledge and explanations about a subject, so as to create coherent and effective presentations? Agents such as Adele, STEVE, and Herman [5, 6, 7] are designed to generate explanations based upon knowledge structures that support explanation. But young children have relatively poor understanding of what counts as causation, evidence, etc. As a consequence, agent authoring tools such as Diligent [8] and VIVIDS [6] are too complex to be used by children. In contrast, off-the-shelf Web page and presentation building tools provide little structure, and no guidance as to how to think about the knowledge students must present.

As a basis for structuring knowledge and explanations we have adopted Rhetorical Structure Theory (RST) by Mann & Thompson [9]. According to the RST, a coherent natural language text is structured in terms of functional relations that hold between parts of it. The relations, such as elaboration, motivation, and evidence, represent different pragmatic goals within the authoring tool. The goals provide a high-level structure for presentation authoring.

We focus in this paper on a pedagogical agent that assists with presentation authoring by monitoring the user's performance, and by intervening to assist him. The tutor is embedded in the Digital Puppet System's authoring tool. Using an RST analysis of the source text, the agent performs pedagogical tasks to support the user's understanding of rhetorical relations, stimulates reflection about the relationships between the structure of the original text and the structure of the presentations, and suggests ways to improve the user's performance. Upon completion of the authoring, the authored paragraphs are organized into coherent structures that can be presented by DPs in a learning-by-teaching classroom context.

2. Related work

The Digital Puppet system is, at its core, a tool for building knowledge representations, and is similar to systems like Belvedere [10] and ConvinceMe [11]. These hypothesis-oriented tools address high school and undergraduate level scientific inquiry into broad-based problems in a collaborative setting. In contrast, we present an RST-based tool for young children whose goal is to explain the relationships of local artifacts (paragraphs of text) in a learning-by-teaching context. Tools that address rhetorical issues, such as argument construction, e.g., SenseMaker [12], and explanation construction, e.g. Explanation Constructor [13], are fundamentally similar, though are not RST- or agent-based.

As far as we know, no computer-based learning environment is aimed at teaching students how to organize presentations of a target text according to rhetorical principles. Most Intelligent Language Tutoring Systems (ILTS) focus on vocabulary and grammar; some systems (e.g. [14]) help learners plan the essay by dividing it into functional units like introduction, body, and conclusion, but the text to be written is self-contained and does not constitute a presentation of some other text.

The Rhetorical Structure Theory is used in the area of natural language processing for both analyzing and generating texts in terms of rhetorical relations [15, 16]. Burstein et al. [17] use the RST to automatically summarize GMAT essays produced by students. Their work differs from ours in that the GMAT essay is a self-contained text, and the summary provided to the student does not outline the rhetorical structure of the essay nor of the summary. André et. al [18] use an RST-based planning system for controlling an agent that displays animated presentations of Web pages or multimedia material. Our system is focused on helping the user to author the text of a presentation, rather than on automatically building multimedia presentations, and the text authored by the children is organized into a coherent structure using templates, rather than by means of a planning algorithm. Other works on ITSs use natural language methods, such as performatives designed from speech acts, for producing coherent and effective dialogues between system and student [19]. Here, we use the RST to guide students through the process of building coherent presentations.

3. The Digital Puppet System

Digital Puppet presentations are authored in the Digital Puppet System (DPS), which was created for authoring a simplified version of Adele (Agent for Distance Education), a Web-based animated pedagogical agent technology [5]. Whereas Adele is designed to support simulation-based learning, DPS serves a complementary functions, to augment Web-based presentations. There are three types of users of the DPS: designer, author, and viewer. The designer annotates the text of a Web page using the RST relations; the author, a fourth grader, creates a puppet-enhanced presentation about a Web page; and the viewer, a classmate, teacher, or even the author herself, plays the resulting presentation. The DPS, which includes the pedagogical agent and RST annotation tool [20, 21], is illustrated in Figure 1.

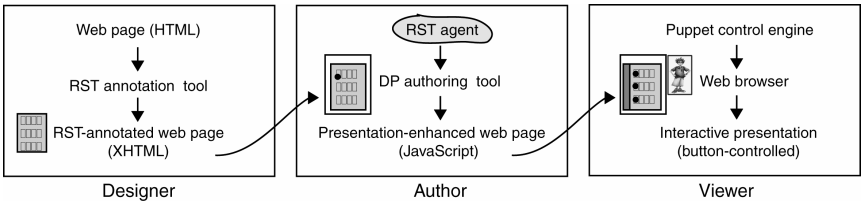

Fig. 1. The DP system, including the pedagogical agent and RST annotation tool

The DPS consists of three main parts: (1) The designer's tools for annotating the text, (2) the authoring tool for creating the presentation, which this paper is focused on, and (3) the browser environment for playing the presentations. The pedagogical agent is embedded in the authoring tool, where it monitors and assists the author. The operation of the system can be summarized in the following steps:

Designer

1. A Web page on the topic of study is either selected or created by a teacher or instructional designer, and is input to an RST-based manual annotation tool.

2. The designer, by means of the tool, generates a paragraph-level RST analysis of the page and produces hierarchical markup tags that represent the tree of rhetorical relations, that are inserted into the Web page using XHTML.

Author

3. The RST-annotated Web page is read in and displayed by the DP authoring tool.
4. The RST agent interacts with the author, explaining the meaning of the RST relations, highlighting the structure of the page, and suggesting ways to structure the presentations.
5. The DP authoring tool is a Java application that enables the learner to further annotate the Web page for the purpose of adding introductory and explanatory presentation text. RST relation boxes are presented as a means of outlining, or structuring, the presentation. The authored text and associated interactive buttons are inserted into the page with Java Script.

Viewer

6. At anytime, the presentation-enhanced Web page can be displayed and tested in a Web browser. The authoring tool creates a display Web page with two frames. In one frame, it inserts a JavaScript function called the Puppet Control Engine, which controls the Digital Puppet's animated behavior and communicates with the client-based text-to-speech engine. In the other frame, it displays the annotated Web page that includes a synthesis of the authored text for each explanation. The tool then calls a browser to display the results.
7. The viewer activates the Digital Puppet presentation by clicking on the interactive buttons that have been placed before relevant paragraphs, i.e. annotated paragraphs of the Web page by the DPA tool. The puppet then presents the material to the viewer. In our initial version of the system, the interaction between the puppet and viewer is minimal; the puppet may ask questions but cannot respond.

The author's interface to the DPA tool is shown on the left in Figure 2. The Web page is displayed in the main window, to the left. The user writes an introduction to the page in the small window at the top-right. She then selects a paragraph from the original page and writes an explanation about it in a second small window (bottom-right). An annotation button is inserted before the main paragraph so that the authored text can be retrieved for editing, and ultimately, presentation. The authored text is linked to the original text by means of a cause relation that has been selected by the author from a list of possible rhetorical relations. The agent monitors the author's activities and makes suggestions as appropriate, or provides help on demand via the Help button. To activate the presentation, the author clicks on a button on the toolbar.

The viewer's interface to the Web browser and the Digital Puppet is shown on the right in Figure 2. The Web page is augmented with small buttons that, when pressed, activate the animated puppet to present explanations that were authored for a particular paragraph. The puppet uses a client-based text-to-speech synthesizer to narrate the presentation and an XML-based animation engine to perform the presentation. The teen persona of the Digital Puppet is designed as a character with whom younger students can easily identify, and as a role model.

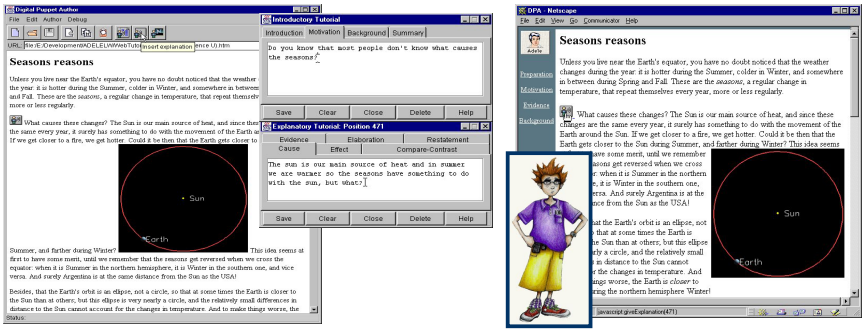

Fig. 2. DP authoring tool interface (left) and the puppet-enhanced Web page (right)

4. Organizing presentations using Rhetorical Structure Theory

According to the RST [9], a text can be analyzed in terms of a tree of relations, each of them holding between two non-overlapping text spans named nucleus and satellite respectively. The nucleus is more important for expressing the writer's intention, and is independent of the satellite. A relation specifies a set of constraints on nucleus and satellite, and a pragmatic effect that the writer intends to produce in the reader. For example, the "evidence" relation has as intended effect that the reader's belief in the nucleus is increased. The following text, taken from [9], p. 10, and regarding a federal income tax program, shows an instance of this relation; the first span is the nucleus, while the second span is a satellite providing evidence for the nucleus: "(1) The program as published for calendar year 1980 really works. (2) In only a few minutes, I entered all the figures from my 1980 tax return and got a result which agreed with my hand calculations to the penny."

Table 1. Rhetorical relations for the Introductory and the Explanatory Tutorials

Introductory Tutorial	Explanatory Tutorial
Introduction Background Motivation	Cause Result Evidence Compare/Contrast Elaboration Restatement

We have selected and partially modified a subset of the relations originally proposed in [9], and we have classified them according to their purpose: 1) structuring an *Introductory tutorial*, i.e. a text that introduces the Web page as a whole, and 2) structuring *Explanatory tutorials*, i.e. texts that are aimed at explaining specific portions of the Web page (see Table 1). For Introductory relations, the whole text of the Web page is considered a nucleus, for which three types of satellite information are to be provided by the author: the introduction, the background, that should

increase the reader's ability to comprehend the nucleus, and the motivation, that should foster the reader's desire to read the original text. As for the Explanatory, the author is requested to identify some important spans within the original text, and to provide one or more satellites for each of them, choosing suitable relations. For example, the author might provide some evidence or restatement for several concepts mentioned in the original text.

The connection between the original text (i.e., the Web page) and the authored text is illustrated in Figure 3. Each piece of text written by the author corresponds to a relation. There is only one introductory presentation, comprising three pieces of text, one per relation. There may be several explanatory presentations, each comprising at least one piece of text about a unique paragraph.

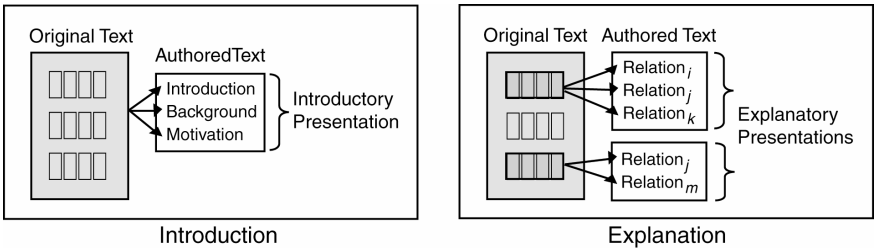

Fig. 3. Relations between the original text and the authored texts

5. Pedagogical tasks and knowledge of the tutor

In order to realize RST-based presentations, especially the explanatory ones, the author should perform three tasks while interacting with the DPA tool: (1) understand the basic structure of the original text; (2) build sensible relations between the original text and the texts to be authored; (3) edit the texts of the presentations. The tutor helps the author perform these tasks by executing several types of behaviors, that can be classified according their "focus", i.e. the issue they concern, and to four pedagogical functions: *stimulating reflection, supporting comprehension, encouraging action,* and *improving performance* (see Table 2 for examples).

Tutor's behaviors focused on the original text.
Thanks to an a priori RST analysis of the original text, the pedagogical agent is able to highlight, comment, and explain the following items in the original text:

1. The paragraph containing the basic nucleus, and the most significant satellite paragraphs. This information helps the author to choose which paragraphs to work on first, or can be used by the agent for giving feedback to the author about his choice of paragraphs to annotate.
2. The rhetorical relations holding between paragraphs. This helps the author understand the structure of the original text, provides him with examples of the relations he wishes to edit, and gives him guidelines about how to structure his own explanatory text.

Table 2. Examples of pedagogical behaviors of the agent

FUNCTIONS	FOCUS			
		Structure of original text	**Preparation of authored text**	**Relation between original and authored text**
	Stimulating reflection	*What do you think is the most significant paragraph?*	*Do you think what you have is long enough?*	*Can you think of other relations you may write wrt this paragraph?*
	Supporting Comprehension	*Look, this paragraph is an example of an evidence*	*For writing some evidence, look at the following example...*	*When explaining a result, it is good to provide some evidence for it*
	Encouraging action	*Try to identify some evidence in the text. If you can't, I can help you.*	*Maybe you could write something more...*	*Why don't you try to elaborate on this comparison?*
	Improving performance	*What you have highlighted is not a generic elaboration, it is a comparison*	*I suggest that you try to rewrite the text in your own words*	*It would be good to elaborate on this evidence*

Tutor's behaviors focused on relations between original and authored texts.
The RST is a descriptive rather than a prescriptive theory; hence, the tutor can only give some general suggestions about the relations to choose for linking the authored paragraphs to paragraphs in the original text, based on preferences shown in Table 3. The preferences are based on the idea that a relation instance in the original text should be linked to the most relevant and specific relation instances in the authored text.

Table 3. Preference-based rules for suggesting relations to author

If relation instance in original text is:	Then agent suggests user author:
Cause	Result, then Evidence, then Elaboration
Result	Cause, then Evidence, then Elaboration
Compare-contrast	Elaboration
Evidence	Compare-contrast, then Elaboration
Elaboration	Restatement
Restatement	[no specific suggestion]

Tutor's behaviors focused on the authored text.
Two kinds of performance variables about the author can be continuously measured by the agent: *global* and *local*. Global variables refer to the relations and relation instances the author creates, regardless of the specific paragraphs they are attached to. Local variables refer to the relations and relation instances concerning specific paragraphs in the original Web page. The variables are listed in Table 4. In the names of the variables, the prefixes "o-" and "a-" stand for "original" and "authored" respectively. Some of the variables are computed on the basis of two constants:

total_rels, the number of all the relations the author can choose from (see Table 1); and *total_o-pars*, the number of paragraphs in the original text.

The agent can also keep track of other variables that refer to the actual actions performed by the user: *tool_actions*, concerning the author's use of the tool in general (for example, how often and about what topic help is requested), and *task_actions*, regarding the user's progress during the authoring task (for example, how often he edits, clears and deletes relations).

By means of these variables the agent can notice, for example, whether the author is (a) over/under-using some relations; (b) adequately explaining the original text or ignoring some parts of it; (c) plagiarizing the original text; (d) authoring text that is too long or short with respect to the original paragraph. The values of the variables are used for inferring the author's needs and for triggering agent's behaviors aimed at, for example, encouraging him to experiment with all the available relations, or better explaining to him the overlooked relations.

Table 4. Global and local variables for measuring the authoring task outcomes

Global Variables			**Local Variables**
About a-relations	**About a-instances**	**About o-paragraphs**	
total_a-rels: number of used relations	*total_a-insts*: number of all authored relation instances	*total_explained_o-pars*: number of annotated original paragraphs	*relative_a-rels*: number of a-relations authored for each o-paragraph
relative_a-rels: percentage of *total_a-rels* to *total_rels*	*a-inst_length*: length of each authored piece of text	*relative_explained_o-pars*: percentage *of total_explained_o-pars* to *total_o-pars*	*ao_length_distance*: difference between length of o-paragraph and length of the a-text
a-rel_insts: number of instances authored for a given relation			*ao_distance*: semantic distance between o-paragraph and a-paragraph

6. Evaluating the RST-based approach

The work presented in this paper is new and ongoing. During the 2001-2002 school year we are scheduled to perform field trials of the authoring tool at an Elementary School in Los Angeles Unified school district with English Language Learners (students whose first language is not English). Our evaluation plan calls for a preliminary study to evaluate the usability of the tool and understanding of the task, and a pilot study to assess its effectiveness. To evaluate how well an arbitrary ten-year-old might create an RST-structured presentation, we asked a fourth grader to prepare a sample presentation. His answers give us an indication of the challenges we will face in our future work, including how to parse awkward sentences, how to assist at both grammatical and semantic levels, and how to synthesize authored components:

Intro: "Hi my name is Dale Lin. I am Doing a report of seasons."
Motivate: "Say if your in the USA and you wanted to go some where cold you would go down to the southern hemisfear because it would be winter there"
Evidence: "The cause of seasons is that it is on one imaginary line called and axis. The earth rotates around the axis and form an angle of 23.4. the the ear start to tilt and a cause a certin amout of area."
Backgrnd: "The axis cause a tilt that mak the north or south hemisfear to receive more sun then the other."

The challenge is then to turn the authored results into a coherent presentation. We may be able to catch some of the grammatical errors before the components are synthesized, and the authoring tool enables the student test out a presentation before it is performed for the peer group. However, we expect for younger children, especially children whose first language is not English, that there will be awkwardness and that the awkwardness will be a catalyst for learning. The following is an example of one synthesis of the components above. Italicized text is the student's verbatim text:

Hi! My name is Oliver, and I will be speaking for *Dale* today. We are *doing a report of seasons*. It's good to know *what causes the seasons*. Why? *Say if your in the USA and you wanted to go some where cold you would go down to the southern hemisfear because it would be winter there*. What causes the seasons? *The cause of the seasons is that it is on one imaginary line called and axis. The earth rotates around the axis and form an angle of 23.4. The ear start to tilt and a cause a certin amout of area. The axis cause a tilt that mak the north or south hemisfear to receive more sun then the other*.

For the pilot study [22], as designed by Mayer, we will work with a class of fourth grade students. We will evaluate both learning performance and outcomes for both a puppet and control group, analyzing the performance based on learning process dimensions such as number of self-explanations, and inquiry episodes generated by students, to determine whether the Digital Puppet learners scored significantly higher on process measures of cognitive activity during learning. To study the learning outcomes, we will analyze the data from a battery of instruments in order to determine whether the Digital Puppet task improved student learning.

Acknowledgements

We would like to thank Kate LaBore for her support and insightful comments, Nancy Chang, for performing the student evaluation, Andrew Marshall, who created the first animation engine for Adele, Karyn Cordova, who is creating the new animation engine, Daniel Marcu and Edward Hovy for insightful discussions, and Stefano Levialdi and Marilena De Marsico for useful comments on a previous version of this paper. This work was supported by an internal R&D grant from the USC Information Sciences Institute, and by a scholarship from NATO and the Italian National Research Council awarded to Paola Rizzo during her visit to the USC.

References

1. Mayer R. E.: Multimedia Learning. Cambridge University Press, Cambridge, UK (2001)
2. Johnson, W.L., Rickel, J.W., Lester, J.C.: Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. International Journal of Artificial Intelligence in Education 11 (2000) 47-78
3. Moreno, R. and Mayer, R. E.: Engaging students in active learning: The case for personalized multimedia messages. Journal of Educational Psychology, 92 (2000) 724-733
4. Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., Bhogal, R. S.: The persona effect: Affective impact of animated pedagogical agents. In: Proceedings of CHI '97 (1997) 359-366

5. Shaw E., Johnson W. L., Ganeshan R.: Pedagogical agents on the Web. In: Proceedings of the Third International Conference on Autonomous Agents (1999) 283-290
6. Johnson, W.L., Rickel, J., Stiles, R., Munro, A.: Integrating Pedagogical Agents into Virtual Environments, Presence: Teleoperators and Virtual Environments 7(6) (1998) 523-546
7. Lester, J. C., Converse, S. A., Stone, B. A., Kahler, S. E., Barlow, S. T.: Animated pedagogical agents and problem-solving effectiveness: A large-scale empirical evaluation. In: Proceedings of the Eighth World Conference on Artificial Intelligence in Education. IOS Press, Amsterdam (1997) 23-30.
8. Angros, R., Johnson, W.L., Rickel, J.: Agents that Learn to Instruct. In: AAAI 1997 Fall Symposium Series: Intelligent Tutoring Systems Authoring Tools, Technical Report FS-97-01, AAAI Press, Menlo Park (CA) (1997)
9. Mann W. C. and Thompson S. A.: Rhetorical structure theory: A theory of text organization. Technical Report ISI/RS-87-190, Information Sciences Institute, University of Southern California, Marina del Rey, CA. (1987)
10. Suthers, D. D.: Representational Support for Collaborative Inquiry. In: Proceedings of the 32nd Hawai'i International Conference on the System Sciences (HICSS-32), Maui, Hawai'i (CD-ROM), Institute of Electrical and Electronics Engineers, Inc. (1999)
11. Schank, P, and Ranney, M.: Improved reasoning with Convince Me. In: Human Factors in Computing Systems CHI '95 Conference Companion, Association for Computing Machinery, New York, NY (1995) 276-277
12. Bell, P. Using argument representations to make thinking visible for individuals and groups. In: Proc. Computer Supported Collaborative Learning '97, University of Toronto (1997) 10-19
13. Sandoval, W. A., and Reiser, B. J.: Evolving explanations in high school biology, Paper presented at the Annual Meeting of the American Educational Research Assn., Chicago, IL (1997)
14. Rowley, K. and Crevoisier, M.: MAESTRO: Guiding students to skillful performance of the writing process. In: Proceedings of the Educational Multimedia and Hypermedia conference, Calgary, CA. (1997)
15. Marcu D.: The Rhetorical Parsing of Unrestricted Texts: A Surface-Based Approach. Computational Linguistics, 26 (3) (2000) 395-448
16. Hovy, E.H.: Automated Discourse Generation Using Discourse Structure Relations. Artificial Intelligence 63(1-2), Special Issue on Natural Language Processing (1993) 341-386.
17. Burstein, J. and Marcu D.: Towards Using Text Summarization for Essay-Based Feedback. In: La 7e Conference Annuelle sur Le Traitement Automatique des Langues Naturelles TALN'2000, Lausanne, Switzerland (2000)
18. André, E., Rist, T., Müller, J.: Integrating Reactive and Scripted Behaviors in a Life-Like Presentation Agent. In: Proc. of the Second Int'l. Conf. on Autonomous Agents (Agents '98), Association for Computing Machinery (1998) 261-268.
19. Penstein Rosé C. and Freedman R.: Building Dialog Systems for Tutorial Applications. In: AAAI Fall Symposium 2000, AAAI Press Technical Report FS-00-01, Menlo Park (CA) (2000)
20. Marcu D., Amorrortu B., Romera M.: Experiments in Constructing a Corpus of Discourse Trees. In: The ACL'99 Workshop on Standards and Tools for Discourse Tagging, Maryland (1999) (RST tool: http://www.isi.edu/licensed-sw/RSTTool/index.html)
21. O'Donnell, M.: RSTTool 2.4 -- A Markup Tool for Rhetorical Structure Theory. In: Proceedings of the Int'l Natural Language Generation Conference (INLG'2000), Mitzpe Ramon, Israel (2000) 253-256
22. Johnson, W.L., Mayer, R., and Shaw, E.: K12 Agents: USC/ISI grant 12-1540-0002 (2001)