A Framework of Discourse Analysis and Modeling

Feng-Jen Yang, *Member, IAENG*

***Abstract*—A discourse is a span of utterances that are spoken by one or more speakers to convey an action, an event, a concept, or a concession. In human dialogue, the utterances in a discourse are organized in a comprehensible and coherent manner. It is quite natural that people talking on the same subject tend to coordinate their dialogue into a discourse that can be mutually understood and followed as a protocol by all dialogue participants. While a speaker uses a discourse to generate coherent utterances, a listener tries to reconstruct the discourse to understand the speakers intention or information. Formulating a computational discourse model, thus, is a fundamental effort to make a computer system talks like human in the intended application domain.**

Although there are many discourse analysis and modeling approaches stated in the literature of natural language processing, most of them are domain dependent. This paper is an attempt of discovering a generalized practice that can be followed and applied to various application domains. By comparing and concluding some discourse analysis and modeling approaches, I have derived a domain independent framework of this effort.

***Index Terms*—discourse analysis, discourse modeling, discourse segmentation.**

I. Introduction

ONE of the goals that most of the dialogue-based systems are pursuing is to conduct human-like and natural-sounding conversations within their application domains. This concern brings up tow ramifications of efforts. One is making the system understand a users language input. The other is making the system generate fluent and natural sounding language in response to the users input. Both of these two efforts are rooted in the discourse analysis and modeling. By nature, the discourse of a human dialogue can be analyzed into segments in which each segment is a chunk of utterances that a speaker uses to show some intention or to convey some information.

To be considered as a discourse segment, a span of utterances must have a recognizable purpose and can be understood according the informational or intentional relationships among these utterances. In general a discourse is a span of utterances that are spoken by one or more speakers to convey an action, an event, a concept, or a concession. Usually, the utterances in a discourse are organized in a comprehensible and coherent manner. It is quite natural that people talking on the same subject tend to coordinate their dialogue into a discourse that can be mutually understood and followed as a protocol by all dialogue participants. While a speaker uses a discourse to generate coherent utterances, a listener tries to reconstruct the discourse to understand the speakers intention or information [1]. Formulating a computational discourse model, thus, is a fundamental effort to make a computer system talks like human in an application domain.

Manuscript received July 9, 2012; revised August 6, 2012.

F. Yang is with the Department of Mathematics and Information Sciences, University of North Texas at Dallas, TX, 75241, USA, e-mail: feng-jen.yang@unt.edu.

As an attempt of understanding a users language input, the system has to build an internal discourse model according to the users utterances and then use this model as a temporary knowledgebase to figure out what is in the users mind. On the other hand, while generating language output, the system has to use the same discourse knowledge to organize its output utterances in a coherent manner and stay at the current discourse focus. This systematic interaction is a good of practice of conducting a fluent human-machine dialogue and maintaining a coherent discourse.

This paper is a general study of discourse analysis and modeling. By comparing and concluding some discourse analysis and modeling strategies of dialogue systems, I have derived a generalized practice of this effort.

II. The Analysis and Modeling

The modeling of dialogue discourse is a series of works that continuously divide an overall discourse into hierarchical segments in which each segment is correspondent to a substructure in the hierarchy. Many successful researches have identified various factors to be considered in segmenting a discourse, such as attention, intention, initiative, rhetorical structure, story trees, and turn-taking behavior [2]. For dialogue systems per se, a discourse segment is a chunk of utterances that the speaker uses to show some intention or convey some information. The intention or information to be conveyed in a segment is understood by the inference and aggregation of this chunk of utterances.

A. *The Top-Down Analysis and Modeling*

Although it is not particularly pointed out, a computational discourse is usually analyzed and modeled into a hierarchical structure in a top-down manner. Within a top-down model, the macro-structures are considered as the largest functional subparts of the discourse, then each macro-structure is divided into several levels of meso-structures, and finally each meso-structure is divided into several levels of micro-structures. The lower is the level, the finer is the modeling.

The top-down approach works well while the structure of discourse is predictable and can be represented by predictive patterns, but it is less suitable to applications in which the discourse structure is highly diversified and hard to be predicted in advance.

1) A Case Study of the Top-Down Approach: The SHERLOCK system is a typical top-down application of discourse analysis and modeling. The SHERLOCK system is a computer aided fault diagnosis system that can be used to determine the location of a power distribution fault by analyzing some specified symptoms of faults in a ring network [3], [4], [5].

In the SHERLOCK domain, the machine tutor conducts a turn-taking dialogue with a student user. As a typical example of top-down discourse analysis and modeling, the tutorial

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

explanations of SHERLOCK are modeled into a hierarchical discourse structure [6], [7].

At the macro-level, each tutor explanation or answer is considered as macro-structure called a segment. A segment is a meta-structure consists of at least a core and possibly some contributors as described follows:

1) Segment: Typically each tutorial explanation is a segment with a discourse purpose to answer the student's immediately previous question.

2) Core: In order to maintain a discourse focus, a segment must have at least one constituent that directly expresses the purpose of the segment. This constituent is called the core of the segment.

3) Contributor: In addition to the core, a segment may also have some contributors, which help to achieve the purpose expressed by the core of the segment.

4) Minimal unit: Segments that have only cores but no contributors are recognized as minimal units for the reason of not having further intentional structure, but they may have further informational structures.

Based on the relationship between a core and its contributors, each segment is further divided into the following meso-structures:

1) Intentional relation: An intentional relation between a contributor and its core describes what the speaker is trying to accomplish by providing the contributor in addition to the core.

2) Informational relation: An informational relation describes how the content of the contributor and its core are related in the domain.

At the micro-level, each meso-structure is yet analyzed and modeled to catch some finer discourse properties. Hierarchically, an intentional relation is further divided into convince relation, enable relation, concede relation, join relation, and indeterminate relation.

Similarly, for the sake of hierarchical modeling, the micro-structures within an informational relation is further analyzed and modeled according to the causality, similarity, elaboration and temporal between a contributor and its core.

A special feature of minimum unit is that its components are all defined functionally instead of syntactically. For the hierarchical sake, a minimum unit is further divided into domain unit, matrix and relation cluster.

B. The Bottom-Up Analysis and Modeling

When the discourse structures are unpredictable and variable in a wide range, it is hard for a system to predict the lower level of discourse structures in advance. In such domains, using a top-down approach may end up with trying too many errors. Instead, using the bottom-up approach is more specific and more efficient.

1) A Case Study of the Bottom-Up Approach: The TRAINS system is a typical application of bottom-up discourse analysis and modeling that is built to discuss the efficient routes for trains in the Northeastern United States [8], [9].

In the TRAINS domain, the dialogue discourse is analyzed and modeled into a hierarchy of three levels [2]. While most of the dialogue systems are modeling their discourses in a top-down manner, the analysis and modeling of the TRAINS system dialogue discourse is in a bottom-up manner because it allows users to initiate dialogues. This particular feature makes it hard for the system to predict and plan a discourse contents in advance and, thus, a top-down analysis and modeling may end up with too much try and error.

In the TRAINS domain, the micro-level of discourse consists of tokens. At this level, the dialogues are split into utterance-tokens based on prosody and grammar. Intuitively, a token is correspondent to a single phrase of intonation or a single grammatical clause. The meso-level of discourse consists of Common Ground Units (CGUs) in which a CGU clusters distinct tokens to achieve a mutual understanding. The clustering is more concentrated on modeling what is being said at the level of information exchange than other discourse properties.

At the macro-level, each discourse consists of several I-Units (IUs)in which the I stands for either informational or intentional. The relationships among CGUs, can be used to group CGUs into a hierarchical topic structure or planning based structure called IU trees.

C. Some Observations

The emphasis of discourse analysis and modeling emerged in the mid 1980s. As more and more dialogue systems are being developed, more and more domain-oriented approaches are practiced. Through the past three decades, the theory of discourse analysis and modeling keeps evolving, but two historical approaches have been standing well and surviving the test of time. Both of these two approaches are top-down. The first approach is well known as the Rhetorical Structure Theory (RST) proposed by Mann and Thompson in 1987 that can be followed as guidelines to model text spans hierarchically. Each span is either the nucleus (also known as the central) or a satellite (also known as a support) within a discourse relation [10], [11]. The second approach is known as Hierarchical Schemata proposed by McKeown in 1985 which models discourse into a set of hierarchical schemata [12].

Schemata approach has been proved to be a good way of guaranteeing discourse coherence and selecting discourse content for text generation. Also, the core and contributor in SHERLOCK is correspondent to the nucleus and satellite in RST.

In TRAINS domain, the modeling of CGUs is a good way of getting the level of commonality between participants in dialogue. Also, the modeling of IUs provides a good way of identifying the hierarchical purpose of within discourse structures. Overall, the discourse analysis and modeling of TRAINS is good at marking the mixed-initiative interaction between a user and a dialogue system. This is especially worthy of attention when the task is planning discourse for dialogue systems allowing user initiatives.

Both top-down and bottom-up approaches are commonly practiced in the current academia. Nevertheless, top-down approach has more followers than that of bottom-up approach. The reason for choosing SHERLOCK and TRAINS as study examples in this section is that both of these two domains are naming discourse structures by general terms instead of domain dependent terms which can be adopted in the formulation of a domain-independent approach.

```
BEGIN
    IF (allow user initiatives)
        Follow the top-down analysis and
        modeling
    ELSE
        Follow the bottom-up analysis and
        modeling
    ENDIF
END
```

Fig. 1. The Choice between Top-Down and Bottom-Up Approaches

III. A Framework

This framework starts from choosing either top-down or bottom-up approach. This critical checking is depending on the nature of discourse structures in an application domain. When the higher level structures dominate the lower level structures, the top-down approach is a good choice. This is usually working well when a system is conducting Socratic dialogue and does not allow user initiatives. On the other hand, when the lower level structures dominate the higher level structures, the bottom-up approach is a better choice. This is usually the case when a dialogue system allows user initiatives. This process is represented as the algorithm shown in Figure 1.

A. The Top-Down Analysis and Modeling

The process of top-down analysis and modeling is following the sequence of macro-level analysis and modeling first, then meso-level analysis and modeling and finally micro-level analysis and modeling.

At the macro-level, the analysis and modeling focuse on the division of largest functional subparts within the overall discourse which are used for the purpose of either conveying information or showing intention. As a preparation for meso-level of analysis and modeling, some meta-constituents should also be identified within macro-structures, such as cores and contributors. The core is the constituent that most directly expresses the purpose of a macro-structure. A macro-structure may also have contributors to help achieve the purpose of the core. Some alternative naming for a core as suggested in RST is a nucleus or a central. Similarly manner, a contributor is also known as a satellite or a support.

At the meso-level of analysis and modeling, each macro-structure is further divided according to the relationship between a contributor and its core and comes up with either an informational meso-structure or an intentional meso-structure. If the relationship between a contributor and its core is not clear or hard to be categorized, it is necessary to reanalysis and remodel the macro-structures by using different factors.

The micro-level of analysis and modeling tend to be domain dependent. In order to model these finer discourse properties, an informational meso-structure may be further divided according what kind of information is conveyed such as a concept or an emphasis. Each kind of information leads to a special kind of micro-structure such as a conceptual micro-structure or an emphatic micro-structure. Similarly, an intentional meso-structure may be further divided according what kind of purpose is intended such as eliciting a truth, or asking a question. Each kind of intention leads to a special kind of micro-structure such as an eliciting micro-structure or an asking micro-structure. If the subdivision within a meso-structure is not clear or hard to be categorized, it is necessary to reanalysis and remodel the meso-structures or even the macro-structures by using other dimensions.

```
BEGIN
    implementable = false
    redo-macro-structures = true
    redo-meso-structures = ture

    WHILE (NOT implementable)
        WHILE (redo-macro-structures)
            Divide the discourse into major
            functional sub-parts to form
            macro-structures and identify
            the core and contributors
            within each macro-structure

            IF (clearly divided)
                redo-macro-structures = false
            ENDIF
        ENDWHILE

        WHILE (redo-meso-structures)
            Divide each macro-structure into
            informational meso-structures or
            intentional meso-structure
            according to the relationship
            between a contributors and its
            core

            IF (clearly divided)
                redo-meso-structures = false
            ENDIF
        ENDWHILE

        Divide each informational
        meso-structures or intentional
        meso-structure into semantic
        micro-structures according to the
        semantic of each informational
        or intentional purpose

        IF (implementable)
            implementable = true
        ENDIF
    EDNWHILE
END
```

Fig. 2. The Top-Down Analysis and Modeling

Eventually, all of the modeling comes to the implementation stage. After all, if the top-down discourse structures are not implementable, it is necessary to reanalysis and remodel the macro-structures, the meso-structures, and the micro-structures. This process is represented as the algorithm shown in Figure 2.

B. The Bottom-Up Analysis and Modeling

The process of bottom-up analysis and modeling is following the sequence of micro-level of analysis and modeling first, then meso-level of analysis and modeling and finally macro-level of analysis and modeling.

At the micro-level of analysis and modeling, each utterance consists of some prosodic or grammatical phrases. As a preparation for meso-level of analysis and modeling, the semantic of each phrase should also be identified and then

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

```
BEGIN
    implementable = false
    redo-micro-structures = true
    redo-meso-structures = ture

    WHILE (NOT implementable)
        WHILE (redo-micro-structures)
            Look for the semantic
            micro-structure for each
            prosodic or grammatical phrase
            within an utterance

            IF (clearly modeled)
                redo-micro-structures = false
            ENDIF
        ENDWHILE

        WHILE (redo-meso-structures)
            Aggregate related semantic
            micro-structures that achieve a
            mutual understanding to form a
            mutual meso-structure

            IF (clearly divided)
                redo-meso-structures = false
            ENDIF
        ENDWHILE

        Use the relationships among mutual
        meso-structures to form either
        an informational macro-structure
        or an intentional macro-structure

        IF (implementable)
            implementable = true
        ENDIF
    ENDWHILE
END
```

Fig. 3. The Bottom-Up Analysis and Modeling

each kind of semantic leads to a special kind of semantic micro-structure.

At the meso-level of analysis and modeling, all related semantic micro-structures that achieve a mutual understanding are aggregated together to form a mutual meso-structure. If there is no obvious common ground among semantic micro-structures, it is necessary to reanalysis and remodel the micro-structures by using different factors.

The macro-level of analysis and modeling is based on the relationships among mutual meso-structures. While some mutual meso-structures may form an informational relationship and result in an informational macro-structure, some others may form an intentional relationship and result in an intentional macro-structure.

Finally, if the bottom-up discourse structures are not implementable, it is necessary to reanalysis and remodel the micro-structure, meso-structures, the macro-structures. This process is represented as the algorithm shown in Figure 3.

C. Putting Together

The overall process of this framework starts from the decision of choosing either top-down or bottom-up approach and then gets into the detail of its correspondent process of analysis and modeling. Eventually, a discourse analysis and modeling comes to the final stage of implementation. If the resulted discourse structures can not be implemented efficiently, it is necessary to start over and redo the analysis and modeling, until an implementable result is reached.

IV. Conclusion

Discourse analysis and modeling is an essential work in terms of language understanding and language generation. Since human dialogues are purposeful, machine dialogues should fulfill the same goal. As a result, discourse contents should not be just random aggregations of utterances. Instead, they should be well planned and structured according to preformed computational models that are built based the intended application need. For a dialogue system to really benefit its users, the system must be robust, efficient, coherent and fluent. Any flaw in the analysis and modeling of dialogue discourse may result in unnatural dialogues and frustrate its users. With a delicate analysis and modeling, each level of discourse content can form a coherent and thematic structure that is easy to read and understand.

The research of computational discourse is relatively new and still evolving. In this paper, I surveyed several related literatures to formulate a domain-independent framework that can be followed as a generalized practice while analyzing and modeling discourse structures and segmenting discourse contents in an application domain.

References

[1] E. Schuster, "Establishing the Relationship between Discourse Models and User Models," *Journal of Computational Linguistics*, 1988, vol. 14, no 3, 46 - 50.
[2] C. Nakatani and D. Traum, "Draft: Discourse Structure Coding Manual," *Technical Report of the 3rd Discourse Resource Initiative (DRI) Meeting*, Chiba, Japan, 1998.
[3] K. Wong, C. Tsang, and W. Chan, "SHERLOCK: A System for Diagnosing Power Distribution Ring Network Faults," *Proceedings of the 1st International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems*, Tullahoma, TN, 1988, vol. 1, 109 - 113.
[4] K. Wong, H. Cheung, "Artificial Intelligence Approach to Load Allocation in Distribution Substations," *IEE Proceedings, 1987, Part C, Generation, Transmission and Distribution*, 1987, vol. 134, no. 5, 357-365.
[5] K. Wong, C. Tsang, "A Logic Programming Approach to Fault Diagnosis in Distribution Ring Networks," *Electric Power Systems Research*, 1988, vol. 15 Issue 2 77-87.
[6] J. Moore, B. Lemaire, and J. Rosenblum, "Discourse Generation for Instructional Applications: Identifying and Exploiting Relevant Prior Explanations," *Journal of Learning Sciences*, 1996, vol. 5, no. 1, 49 - 94.
[7] M. Moser, J. Moore, and E. Glendening, "Instructions for coding explanations: Identifying segments, relations and minimal units," *Technical Report 96-17*, University of Pittsburgh, Department of Computer Science, 1996.
[8] J. Allen, L. Schubert, G. Ferguson, P. Heeman, C. Hwang, T. Kato, M. Light, N. Martin, B. Miller, M. Poesio, and D. Traum, "The TRAINS Project: A Case Study in Building a Conversational Planning Agent," *TRAINS Technical Note 94-3*, Computer Science Department, The University of Rochester, 1994.
[9] J. Allen, L. Schubert, G. Ferguson, P. Heeman, C. Hwang, T. Kato, M. Light, N. Martin, B. Miller, M. Poesio, and D. Traum, "The TRAINS Project: A Case Study in Defining a Conversational Planning Agent," *Journal of Experimental and Technical Artificial Intelligence*, 1995, vol. 7, no. 1, 7 - 48.
[10] W. Mann and S. Thompson, "Rhetorical Structure Theory: A Theory of Text Organization," *Technical Report ISI/RS-87-190*, Information Sciences Institute, University of Southern California, 1987.
[11] W. Mann and S. Thompson, "Rhetorical Structure Theory: Toward a Functional Theory of Text Organization," *Text*, 1988, vol. 8, no. 3, 243 - 281.
[12] K. McKeown, "Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text," *Cambridge, UK: Cambridge University Press*, 1985.