A Survey of Discourse Representations for Chinese Discourse Annotation

XIAOMIAN KANG, National Laboratory of Pattern Recognition, Institute of Automation, University of Chinese Academy of Sciences, Chinese Academy of Sciences, China
CHENGQING ZONG, National Laboratory of Pattern Recognition, Institute of Automation, University of Chinese Academy of Sciences, CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, China
NIANWEN XUE, Brandeis University, USA

A key element in computational discourse analysis is the design of a formal representation for the discourse structure of a text. With machine learning being the dominant method, it is important to identify a discourse representation that can be used to perform large-scale annotation. This survey provides a systematic analysis of existing discourse representation theories to evaluate whether they are suitable for annotation of Chinese text. Specifically, the two properties, expressiveness and practicality, are introduced to compare the representations of theories based on rhetorical relations and the representations of theories based on entity relations. The comparison systematically reveals linguistic and computational characteristics of the theories. After that, we conclude that none of the existing theories are quite suitable for scalable Chinese discourse annotation because they are not both expressive and practical. Therefore, a new discourse representation needs to be proposed, which should balance the expressiveness and practicality, and cover rhetorical relations and entity relations. Inspired by the conclusions, this survey discusses some preliminary proposals on how to represent the discourse structure that are worth pursuing.

CCS Concepts: • **General and reference → Surveys and overviews**; • **Computing methodologies → Discourse, dialogue and pragmatics**;

Additional Key Words and Phrases: Discourse representation, discourse theory, discourse structure, discourse analysis

ACM Reference format:
Xiaomian Kang, Chengqing Zong, and Nianwen Xue. 2019. A Survey of Discourse Representations for Chinese Discourse Annotation. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* 18, 3, Article 26 (January 2019), 25 pages.
https://doi.org/10.1145/3293442

The research work described in this article has been supported by the National Natural Science Foundation of China under Grant No. 61333018.

Authors' addresses: X. Kang and C. Zong, National Laboratory of Pattern Recognition, Institute of Automation, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Intelligence Building, No. 95, Zhongguancun East Road, Haidian District, Beijing 100190, China; emails: {xiaomian.kang, cqzong}@nlpr.ia.ac.cn; N. Xue, Brandeis University, Computer Science Department, Brandeis University, 415 South Street, Waltham MA 02453, USA; email: xuen@brandeis.edu.

2375-4699/2019/01-ART26 $15.00
https://doi.org/10.1145/3293442

1 INTRODUCTION

Discourse analysis is a crucial analytic level in natural language processing (NLP), and modern computational discourse analysis cannot proceed without first determining an appropriate formal representation that can be used to annotate a large amount of data. The discourse representation specifies what linguistic phenomena to model, what is the appropriate set of symbols to use, and which mathematical objects can be used to model the linguistic phenomena. Computational tools such as discourse parsers provide crucial information for downstream discourse-level NLP tasks, such as document-level machine translation, text summarization, and machine reading comprehension.

Discourse representation theories have been actively investigated since the 1970s, especially in English. There are numerous theories trying to describe various discourse characteristics from different perspectives. A discourse theory usually represents a text into associations between discourse units. According to the type of the association, we try to divide discourse representation theories into two categories: theories based on ***rhetorical relations*** and ***entity relations***. The former explains the association through rhetorical relations, while the latter represents it by the entity links. For example, to describe the association between discourse unit u_1 and u_2 in Example 1.1, a theory based on rhetorical relations utilizes the rhetorical relation "Cause," while a theory based on entity relations extracts two coreferential entities "John" and "He."

Example 1.1. [u_1 John's leg was injured in the accident.] [u_2 He never stood up again.]

The development of these theories promotes the construction of corpora and the development of parsing algorithms. In the age of machine learning, scalable discourse annotation is essential to train accurate systems. However, not every discourse representation theory is suitable for scalable annotation. There is a gap between theory and technology. Some theories explain discourse from a linguistic point of view but ignore their operability in annotation and system implementation. The situation is more serious for Chinese. Compared with English, discourse representation theories for Chinese are very preliminary. It results in the scarcity of annotated data. Besides, most Chinese discourse theories are transferred from English. However, Chinese has its own properties such as prominent parataxis and ellipsis that should be represented properly. The lack of a Chinese discourse annotation scheme limits the development and application of Chinese discourse analysis.

Therefore, the purpose of this article is to *evaluate existing discourse representation theories and determine which one is more suitable for scalable Chinese discourse annotation in the age of machine learning*.

Given the purpose, two questions naturally arise. (a) What discourse information and characteristics do existing theories represent? (b) Are their representations suitable for annotation and computation? To answer the two questions, we introduce the notions of ***expressiveness*** and ***practicality*** to compare the representations of existing theories based on rhetorical relations and entity relations, respectively. Expressiveness evaluates the capacity to represent discourse information flexibly and comprehensively. And practicality shows the capacity of a representation to be realized and applied to real-world problems. We believe a suitable discourse representation needs to be both "expressive" and "practical" so that sufficient amounts of data can be annotated within a reasonable amount of time to train accurate systems.

Eight typical discourse representation theories, including some specializing in Chinese, are discussed in Section 2. After comparing their expressiveness and practicality (Sections 3 and 4), we argue that none of the existing theories are quite suitable for scalable discourse annotation for Chinese in that they are not both expressive and practical, thus a new representation (or annotation scheme) needs to be proposed to address the shortcomings of existing theories. A reasonable discourse representation should achieve the ideal tradeoff between expressiveness and practicality,

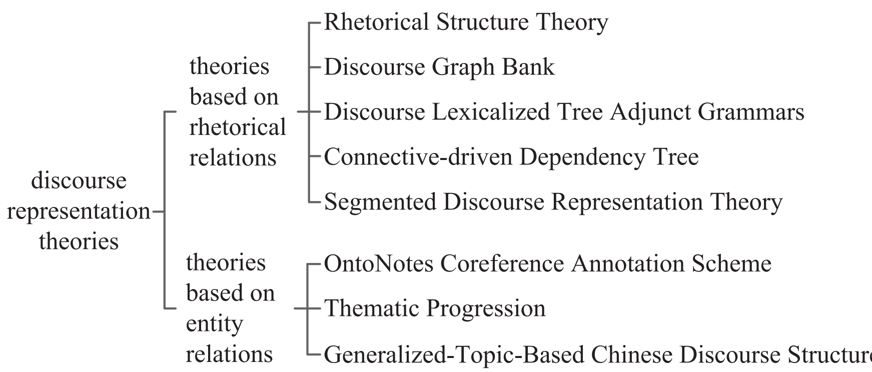

Fig. 1. The existing discourse representation theories.

and represent rhetorical relations and entity relations in a unified framework. Fortunately, comparison sheds some light on the nature of different discourse representations and suggests several desiderata for the content and form of the representation. Based on these desiderata, we outline three future directions for Chinese discourse representation and attempt to give some suggestions (Section 5). In general, we argue that the dependency graph can be used to represent the structure of a text, and the semantic information based on cohesion phenomena and Chinese topics can be annotated.

2 EXISTING DISCOURSE REPRESENTATION THEORIES

Some typical existing discourse theories based on rhetorical relations and entity relations are enumerated in Figure 1.

Before describing them in detail, we would like to lay out a few criteria for the discourse representations that we select for discussion. (a) Given the purpose outlined in the Introduction, some theories without annotation might not be suitable at all. For example, Centering Theory [33] is a theory proposed in the age of rule-based NLP, and while it is very influential, it may not be suitable for our purpose because it has never been used in annotation. (b) We only select some discourse theories representative in the formal representation. So discourse theories that have similar representation forms to the theories discussed in this section and those unable to represent a text into a specific form are excluded. (c) Some theories concentrating on *mention*, *theme*, or *topic* are all classified as theories based on entity relations. Despite the different definitions (that we will introduce in the corresponding theories), these three concepts often appear as entity nouns or phrases and share some common characteristics. (d) This article mainly focuses on the representation rather than machine learning algorithms. For each theory, we will not introduce the implementation algorithms. (e) The theories discussed are limited to two languages: English and Chinese.

2.1 Discourse Theories Based on Rhetorical Relations

2.1.1 Rhetorical Structure Theory (RST). RST [66] explains the structure of a text in terms of a tree-shaped graph, or simply trees. The leaves of the tree correspond to the minimum text fragments, called *elementary discourse units* (*EDUs*). The intermediate nodes of the tree correspond to complex discourse units merged with two or more non-overlapping, adjacent units. In the process of the node combination, the "*nucleus-satellite*" structure is defined where a nucleus indicates an essential unit and a satellite indicates a supporting or background one. This structure is labeled with extensible rhetorical relations, such as Cause, Condition, Sequence, List, and so forth. Figure 2 exhibits the representation of Example 2.1 in RST.

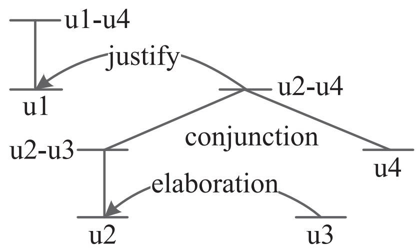

Fig. 2. The RST tree representation of Example 2.1 with four EDUs.

Example 2.1. [u_1 It could have been a great movie.] [u_2 It does have beautiful scenery,] [u_3 some of the best since Lord of the Rings.] [u_4 The acting is well done.]

Carlson et al. [8] build RST-DT, an English corpus annotating 16 classes and 78 types of relations in 385 documents. It contributes to the development of RST-style discourse analysis [28, 29, 37, 43, 46, 56, 86]. An RST-style discourse parser usually contains two modules: EDU segmentation and tree building. For Chinese, Yue [109] annotates an RST-style corpus of 97 Chinese news commentaries. Nevertheless, the lack of corpus restricts the development of Chinese RST-style discourse parsers.

RST suggests a compact, hierarchical discourse representation that mainly embodies two categories of discourse information: the structure and the semantic relations. The structure is built on the foundation of the "nucleus-satellite" relations that characterize the information saliency among discourse units, and the semantic relations are expressed through the rhetorical relations.

The information has been applied to several NLP tasks [88]. For summarization, some researchers extract the discourse units by considering their depth on the tree and the nuclearity [38, 67, 95]. Gerani et al. [32] utilize the depth and the number of EDUs to select aspects for abstractive summarization of product reviews. Radev [80] defines the relations between units across the text for multi-document summarization. For sentiment analysis, the "nucleus-satellite" labels and a few special rhetorical relation types are usually utilized as features or clues to help to determine the polarity [16, 36, 111, 115]. Bhatia et al. [6] try to embody the RST tree through recursive neural network and train weight matrices to capture rhetorical relations. For machine translation, Guzmán et al. [34] introduce RST structure to improve evaluation metrics. Tu et al. [91] extract tree-to-string rules according to the relations, then the decoding procedure is bottom-to-up on the source RST tree. For question answering, the rhetorical relations between sentences are considered as features for answer selection or re-ranking [42, 72, 93].

2.1.2 Discourse Graph Bank (DGB). Despite the extensive application, RST trees cannot cover all the discourse relations naturally occurring in texts. It does not support crossed dependencies and nodes with multiple parents. This has been demonstrated by the statistics from Discourse Graph Bank, a corpus of 135 news texts [99]. The proportion of the arcs participating in crossed dependencies is more than 12.5% and 41% of all nodes have multiple parents. Both phenomena can be presented in DGB which uses an unconstrained graph structure.

DGB represents a text in terms of a graph. The adjacent discourse units topically related are grouped together. Each previously unconnected discourse unit or group is tested to see if it has rhetorical relations with any of the (groups of) discourse units in the partially built discourse representation. It is worth mentioning that connectives are enumerated to help delimit discourse units and determine the relations, despite their lack of a formal role in the graph. Figure 3 shows the DGB representation of Example 2.2, where the nonadjacent connection between u_2 and u_4 is acceptable while it is not allowed in RST.

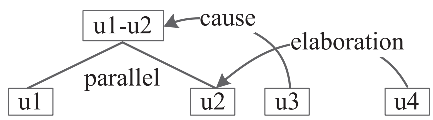

Fig. 3. The representation of Example 2.2 in DGB. There is a crossed dependency between $(u_1 - u_2, u_3)$ and (u_2, u_4).

Type: Implicit
Connective: NULL
Arg1: It's a problem that clearly has to be resolved
Arg2: The agency has already spent roughly $19 billion selling 34 insolvent S&Ls, and it is likely to sell or merge 600 by the time the bailout concludes
Sense: Contingency.Cause.Reason

Type: Explicit
Connective: and
Arg1: The agency has already spent roughly $19 billion selling 34 insolvent S&Ls
Arg2: it is likely to sell or merge 600 by the time the bailout concludes
Sense: Expansion.Conjunction

Fig. 4. The representation of Example 2.3 in PDTB. The relation set contains an implicit relation and an explicit relation. The two arguments of a connective are labeled as *Arg*1 and *Arg*2.

Example 2.2. [u_1 Susan wanted to buy some tomatoes] [u_2 and she also tried to find some basil] [u_3 because her recipe asked for these.] [u_4 The basil would probably be quite expensive.]

DGB can flexibly describe rhetorical relations between almost any discourse units. It constructs a graph with little if any constraints. But the complex structure makes it difficult to apply in large-scale annotation efforts.

2.1.3 Discourse Lexicalized Tree Adjunct Grammars (DLTAG). DLTAG [31, 96] provides theoretical support for *Penn Discourse Treebank* (*PDTB*), currently the largest annotated discourse corpus. The theory extends the sentence-level *Lexicalized TAG* [45] to discourse, where a connective is treated as a discourse predicate that requires two clausal arguments. The PDTB annotation scheme represents a text in terms of a relation set. A relation corresponds to a predicate-argument structure consisting of a connective, two arguments, and their rhetorical relations. The implicit connectives are identified between adjacent sentences. A legitimate argument may be a single clause, a single sentence, or a sequence of clauses and/or sentences. PDTB defines a hierarchical ontology of rhetorical relations containing 4 classes, 16 types, and 23 subtypes. Example 2.3 in PDTB is represented in Figure 4.

Example 2.3. "It's a problem that clearly has to be resolved," said David Cooke, executive director of the RTC. The agency has already spent roughly $19 billion selling 34 insolvent S&Ls, and it is likely to sell or merge 600 by the time the bailout concludes.

PDTB annotates 2,159 English documents and 40,600 relations [81]. The corpus and two CoNLL Shared Tasks [105, 106] make the PDTB-style discourse analysis popular [61]. Researchers are especially attracted by the subtask of implicit relation classification [14, 44, 55, 63, 83, 110]. For

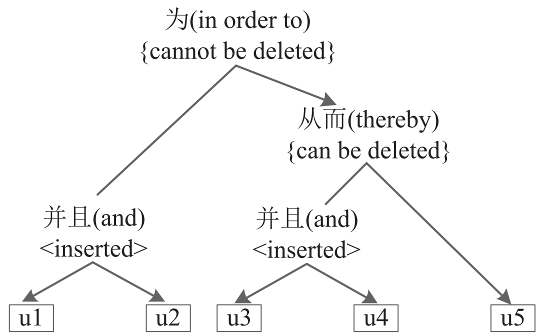

Fig. 5. The representation of Example 2.4 in CDT.

Chinese, Zhou and Xue [113] annotate *Chinese Discourse Treebank* (*CDTB*), a PDTB-style corpus including 500 documents. The dataset is utilized in CoNLL-2016 Shared Task.

DLTAG (PDTB) emphasizes the role of connectives in rhetorical relations. The practice of annotating individual discourse relations simplifies the discourse annotation task. One consequence the PDTB did not foresee is that the arguments for top-level relations become really large. For example, in Figure 4, the *Arg*2 of the implicit relation contains two sentences. This is worse in real data because the arguments there can be even larger. The boundaries of large arguments are hard to predict and the unanalyzed structures in the arguments would not be useful in NLP applications.

PDTB mainly offers the connectives and relation information for NLP tasks. For summarization, Lin [114] extracts relation types and argument tags as features to score sentences. For sentiment analysis, the sentiment transitions between sentences are constrained by the connectives and their rhetorical relations [89, 107]. Mishra and Jain [69] select the more important arguments according to relation types to determine sentiment polarity. For machine translation, Meyer [68] investigates the disambiguation of connectives to correct some translation errors. Li et al. [58] show that translation quality is influenced by discourse connectives and their relation types. For text quality assessment, Lin [114] utilizes the relation transitions in the discourse role matrix filled with relation types and argument tags to evaluate the text coherence.

2.1.4 Connective-Driven Dependency Tree (CDT). Motivated by PDTB and RST, a CDT scheme [60] represents the structure of a text as a tree, with EDUs as leaf nodes and connectives as non-leaf nodes. This theory is designed for Chinese. The EDUs are limited to Chinese clauses, defined in syntactic, functional, and morphological terms. The "nucleus-satellite" structure in RST is retained through the arrows on the tree. The intermediate nodes of the tree are the connectives rather than the complex discourse units in RST. For an implicit relation, an appropriate connective is inserted. The properties of connectives (e.g., whether can be deleted, the language sense) also are labeled. Figure 5 shows the representation of Example 2.4 in CDT.

Example 2.4. [u_1 为规范建筑行为,] [u_2 防止出现无序现象,] [u_3 新区管委会根据国家和上海市的有关规定,] [u_4 结合浦东开发实际,] [u_5 及时出台了一系列规范建设市场的文件。]

"In order to regulate the building behavior and prevent disorderly phenomena, the New District Administrative Committee issued a series of documents to regulate the building market according to the relevant regulations of the state and Shanghai and the actual development of Pudong."

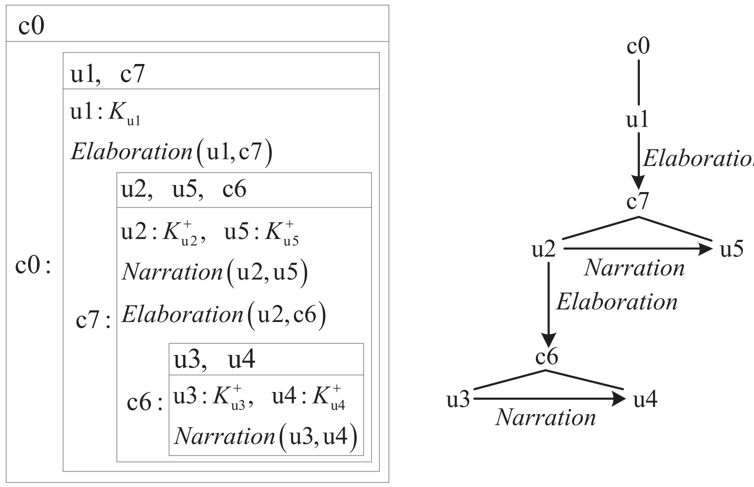

Fig. 6. Two kinds of representation of Example 2.5 in SDRT. The left is the SDRS where c is the complex unit and K is the semantic expression. The right is the graph form.

Li et al. [60] annotate 500 Xinhua newswire documents from the Chinese Treebank. The corpus contains 7,310 relations and 282 connectives. A CDT-style parser is trained based on the corpus [51].

CDT organically integrates the overall tree structure and the connectives that directly indicate rhetorical relations. But the annotation of the rich information is costly. Besides, the RST-similar tree raises the same issue that it cannot cover some discourse structures.

CDT has been applied to zero pronoun (ZP) resolution [15], where some EDU-level features are extracted to generate ZP candidates and some features about the rhetorical relations are employed to improve the resolution.

2.1.5 Segmented Discourse Representation Theory (SDRT). SDRT [3, 5] is a semantic representation theory that extends *Discourse Representation Theory* (*DRT*) [47]. DRT describes the semantics of a text using *Discourse Representation Structure* (*DRS*), a "box-style" logical notation which contains a set of discourse referents and a set of DRS conditions. SDRT introduces rhetorical relations to represent the semantics of a text. Similar to DRS, SDRT can describe a text by its boxed representation, called *Segmented Discourse Representation Structure* (*SDRS*).

For a clearer presentation of the discourse structure, an SDRS can be transferred to a two-dimensional graph. SDRT takes into account a text in terms of a directed acyclic graph [71]. The nodes are basic or complex units. There are two types of edges: *lines* for scope relations, which connect a complex unit and its parts, and *arrows* for rhetorical relations, which connect two units and label relations between them. The construction of the graph is dynamic. The context of the current unit is represented by processed previous units. A new unit attaches its previous units by rhetorical relations. *Subordinating* relations like Elaboration extend the vertical dimension of the graph and *coordinating* relations like Narration expand the structure horizontally [2]. Through this operation, the existing graph structure and relations are updated. For the current unit, the *Right Frontier Constraint* (*RFC*) restricts its attachment points on an existing graph, in which only the last introduced node and its dominators are available. The dominators of a node contain transitive closure over the arrows given by subordinating relations and those holding between a complex unit and its parts. Figure 6 shows the SDRS and its equivalent graph form of Example 2.5.

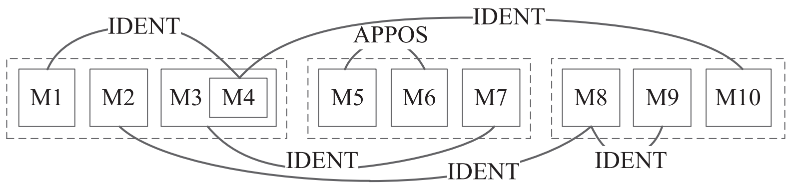

Fig. 7. The coreference chains in Example 2.6. There are four chains: "$M1, M4, M10$," "$M2, M8, M9$," "$M3, M7$," and "$M5, M6$." Their coreference types are labeled on the edges.

Example 2.5. [u_1 John had a great evening last night.] [u_2 He had a great meal.] [u_3 He ate salmon.] [u_4 He devoured lots of cheese.] [u_5 He won a dancing competition.]

Guided by SDRT, Afantenos et al. [1] perform SDRT-style annotation on English 1,091 dialogues from 36 multi-party games and present the first discourse parser for multi-party chat dialogues.

The subordinating and coordinating relations in SDRT construct the hierarchy that is represented by the "nucleus-satellite" in RST. But the structure in SDRT is the graph instead of the tree so that some nonadjacent units can be connected. Meanwhile, the construction of the graph is more constrained than DGB.

The structure and semantic information of SDRT is utilized to set rules for several NLP tasks. For summarization, a rule-based system is built to extract EDUs according to rhetorical relations [49]. For sentiment analysis, Asher et al. [4] design rules based on five relation types to convert the SDRT graph into a representation with polarity.

2.2 Discourse Theories Based on Entity Relations

2.2.1 OntoNotes Coreference Annotation Scheme (OCAS). OCAS annotates a multilingual (English, Chinese, and Arabic) corpus with coreference, a phenomenon where two or more specific referring expressions (i.e., *mentions*) in a text refer to the same object or set of objects in the world (i.e., entity) [40, 79, 98]. Mentions are limited to noun phrases (NPs), pronouns, possessives (e.g., "Fred's" in "Fred's wife"), proper noun premodifiers (e.g., "Army Corps" in "the Army Corps of Engineers"), and verbs (e.g., "grew" in "Sales grew 22%."). The dropped subjects and objects in Chinese and Arabic are also considered. OCAS distinguishes two types of coreference: identity (*IDENT*) that links the same entities, and appositive (*APPOS*) that links a mention with its appositives. In Example 2.6, the mentions are labeled by "$[mention]_M$." This example illustrates the fact that the mentions can be nested (e.g., $M3$ and $M4$). The coreferential mentions can be connected to a chain, so a text contains one or multiple chains, as shown in Figure 7.

Example 2.6. [The U.S.]$_{M1}$ removed [Taiwan]$_{M2}$ from [a list of countries [it]$_{M4}$ is watching for failing to honor U.S. copyrights]$_{M3}$. However, [three other countries]$_{M5}$—[China, Thailand and India]$_{M6}$—will remain on [that watch list]$_{M7}$. [Taiwan]$_{M8}$ has improved [its]$_{M9}$ standing with [the U.S.]$_{M10}$ by initialing a bilateral copyright agreement.

The coreference chains in OCAS reflect the continuity of mentions in a text. Sentences sharing same mentions are connected. Two other influential coreference annotation schemes, MUC [25, 39] and ACE [26], have different definitions of mentions. Mentions are nominal in MUC, and only contain seven types in ACE. In addition, they do not distinguish between identical and appositive types.

OntoNotes annotates 35,143 chains in 2,384 English documents and 28,257 chains in 1,729 Chinese documents. The corpus spans multiple genres and promotes research on coreference

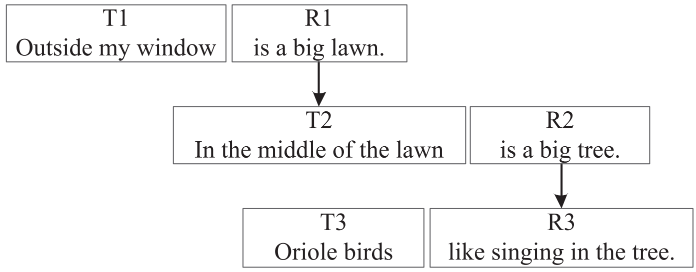

Fig. 8. The representation of Example 2.7 in TP. Each unit is represented by a theme-rheme structure where T and R represent its theme and rheme, respectively. u_1 and u_2 connect through $R1$ and $T2$. u_2 and u_3 connect through $R2$ and $R3$.

resolution [19–21, 53, 73, 77, 78]. Some studies focus on the resolution of particular mentions, such as zero pronouns [11, 12, 50, 62, 108].

Coreference resolution is a basic task for some downstream NLP applications. For summarization, coreference chains are utilized to introduce context information [9, 30] or improve the coherence of output [76, 87]. Bing et al. [7] select coreferential NPs and their VPs to generate abstractive summarization. For question answering, Chen et al. [13] show that the preprocessing of coreference resolution is important. For sentiment analysis, coreferential sentences enhance the context representation [74, 107]. For machine translation, many researchers show interest in the pronoun translation that mainly uses the result of pronoun resolution on the source side. For text coherence assessment, Roth et al. [82] extract the features about coreference chains to model local coherence.

2.2.2 Thematic Progression (TP). TP [22, 27] describes the organization of discourse information on the basis of "*theme-rheme*" structure. Theme-rheme structure is a well-known notion in functional grammar [35]. It regards a clause as a message that conveys the writer's intention. The *theme* is the element which serves as the point of departure of the message. It is the first component of a clause, generally a noun phrase, adverbial group, or prepositional phrase. The remainder of the message is called the *rheme*, which develops the theme. Therefore, a clause is a theme-rheme structure where the theme and rheme usually carry the given and new information, respectively.

TP extends the theme from a single clause to a text. It claims that the discourse information is expressed through the progression and transition of themes and rhemes. Thus, a clause is usually linked to the nearest previous clause connecting with its theme or rheme. These connections represent a text as an information stream progressing clause by clause. Figure 8 explains how the information in Example 2.7 progresses.

Example 2.7. [u_1 Outside my window is a big lawn.] [u_2 In the middle of the lawn is a big tree.] [u_3 Oriole birds like singing in the tree.]

For Chinese, Xi et al. [102] explore a discourse micro-topic scheme based on the theme-rheme structure. Eight types of *micro-topic links*, such as *anaphora*, *omission*, *repetition*, and so forth, are defined to describe the connection of themes or rhemes. The researchers define four thematic progression types for two adjacent clauses. Guided by this scheme, they annotate *Chinese Discourse Topic Corpus* (*CDTC*) containing 500 documents [104], and also built a baseline system [101].

TP proposes a framework to represent discourse information flow. A theme is generally an entity. However, different from limited concentration on the notion about entity in OCAS, TP introduces the rheme to convey the completion information and thus explains wider connections between units. A sentence in OCAS may contain multiple mentions, but usually has only one

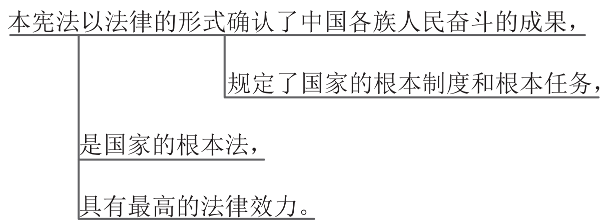

Fig. 9. The representation of Example 2.8 in GCDS. The adverbial "以法律的形式 (in legal form)" in the first clause is a generalized topic. The section after it in the first clause and the second clause are its two comments, answering what is done "in legal form." This adverbial topic is a sub-topic of the topic "本宪法 (this Constitution)," so this topic-comments structure and the last two clauses are all comments on the topic "本宪法 (this Constitution)" in the first clause, answering what "this Constitution" is about.

theme in TP. TP patterns have been applied to coreference resolution, where a possible candidate antecedent is likely to be in the units it links [103].

2.2.3 Generalized-Topic-Based Chinese Discourse Structure (GCDS). A widely accepted opinion in comparative linguistics is that Chinese is topic-prominent while English is subject-prominent [10, 54]. Chu [17, 18] identifies a *topic* in a Chinese clause by three properties, where two basic properties are that it is nominal and it connects with other clauses. Similar to the theme in TP, the notion of a topic is characterized as the focus of attention [90], and syntactically it is usually located at the beginning of a clause or before a verb.

Influenced by theoretical research along this line, GCDS extends the definition of the topic into a *generalized topic* and reveals the properties of Chinese discourse structure based on it [85]. Besides the traditional notion of the topic that is usually nominal and serves as the subject, object, or attributive, a generalized topic in a clause can also be verbal, adverbial, or prepositional.[1] A *comment* describes what or how about the topic. It can be the other parts of the clause except for the topic, or an adjacent complete clause.

GCDS interprets the nature of the frequent subject ellipsis in Chinese as the one-to-many correspondence between one generalized topic and its multiple comments. Based on this opinion, a text is represented as a set of "topic-comments" structures. The structure can be nested, where some sub-topics and their comments can serve as comments of the main topic. For the convenience of representation, each clause in a text is placed in a different line and indented after its topic. This way of expression is called *indented new-line representation.* Figure 9 shows the indented new-line representation of Example 2.8.

Example 2.8. [u_1 本宪法以法律的形式确认了中国各族人民奋斗的成果，] [u_2 规定了国家的根本制度和根本任务，] [u_3 是国家的根本法，] [u_4 具有最高的法律效力。]

"This Constitution, in legal form, affirms the achievements of the struggles of the Chinese people of all nationalities, and defines the basic system and basic tasks of the state. It is the fundamental law of the state, and has supreme legal authority."

Supplementing the generalized topic to each clause without a subject, GCDS offers a group of complete clauses for a text. It is valuable to the topic segmentation and information complement for discourse understanding. The topic-comments structure is utilized to derive the translation unit in English-Chinese machine translation [84].

[1] A related Chinese discourse annotation framework is Topic-chain-based Coherence Annotation [112], which is also based on the notion of topic, but it is narrower in scope in that it only considers nominal topics.

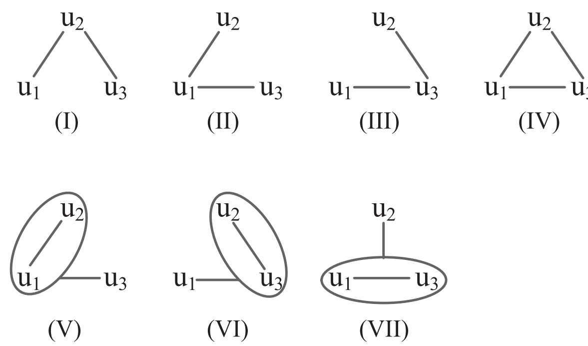

Fig. 10. The possible structures in a text consisting of three units. The line between two units means they have an association, no matter rhetorical relations or entity relations. The ellipse represents a complex discourse unit.

3 THE EXPRESSIVENESS OF DISCOURSE REPRESENTATIONS

In this section, we introduce the **expressiveness** to describe the ability of a discourse theory to represent the discourse information flexibly and comprehensively. We compare aspects of existing discourse theories, especially in terms of their formal representations. The comparison of expressiveness is important in helping us assess which formal representation may be appropriate for a reasonable theory. Because theories based on rhetorical and entity relations model different linguistic phenomena, they are compared separately. While the discourse theories are based on rhetorical relations model connectives (or discourse markers), entity-based discourse theories model anaphoric devices such as pronouns and ellipsis.

3.1 Expressiveness

For the convenience of analysis and exposition, we stipulate some constraints and explain the expressiveness from the following four aspects. The first three are qualitative descriptions, while the last one is quantitative.

Formal Representation. It characterizes the topological structure and main content represented by each discourse theory.

Connection Constraint. Stipulated explicitly or implicitly in a discourse theory, the constraint restricts the permissible range of connections to generate the corresponding formal representation.

Text Coverage. It evaluates whether a discourse representation can cover the text completely, which has also been discussed in [97]. A representation may provide a *full* or *partial* coverage of a text. For example, an RST tree covers every clause in the text, while OCAS only represents the mentions in clauses.

Structural Coverage. It describes the capacity of a theory's representation to cover possible discourse structures, inspired by Danlos [24] who introduces the weak generative capacity to compare RST and SDRT. In order to quantify the capacity, without loss of generality, we discuss structures in a simple text including three basic discourse units, denoted by $d = u_1u_2u_3$. There is an assumption that the text forms one coherent unit, which means isolated units are not allowed. Hence, there are seven possible structures in all, as shown in Figure 10. (V) represents the situation where a complex unit produced by u_1 and u_2 has a relationship with u_3. (VI) and (VII) are similar cases. The structural coverage of a representation is defined as the number of possible structures it can represent.

Table 1. Aspects of the Expressiveness of Representations Based on Rhetorical Relations

Theory	Formal Representation	Connection Constraint	Text Coverage	Structural Coverage
RST	A tree with the EDUs as leaves, the combined spans as intermediate nodes, and the nuclearity and rhetorical relations as edges	Adjacent non-overlapping units	full	2 (V, VI)
DGB	An unconstrained graph with (group of) units as nodes, and the nuclearity and rhetorical relations as edges	No constraint	full	7 (I, II, III, IV, V, VI, VII)
DLTAG (PDTB)	A set of predicate-argument structures anchored by connectives	No constraint	partial	0 (no overall structure)
CDT	A tree with the EDUs as leaves, the connectives as intermediate nodes, and the nuclearity and rhetorical relations as edges	Adjacent non-overlapping units	full	2 (V, VI)
SDRT	A constrained directed acyclic graph with (complex) units as nodes and the scope relations and rhetorical relations as edges	Right Frontier Constraint on the graph	full	4 (I, II, V, VI)

3.2 The Expressiveness of Representations Based on Rhetorical Relations

The details of each aspect of expressiveness of discourse representation theories based on rhetorical relations are listed in Table 1. We will analyze and compare their expressiveness in terms of the four criteria we outlined above.

Given the formal representation, all theories based on rhetorical relations are aimed at representing the overall discourse structure of the entire text except for DLTAG (PDTB). RST and CDT choose a tree as the topological structure, while DGB and SDRT favor a directed graph. The construction of the RST tree highlights the hierarchical nature of a text. A CDT tree is similar to an RST tree other than the fact that the intermediate nodes in the latter are connectives instead of complex units. RST and CDT trees are most naturally viewed as constructed in a bottom-up process. In contrast, SDRT graphs are constructed incrementally as new clauses are sequentially processed. New nodes and scope lines are introduced into the graph to represent the complex units, which also focuses on the hierarchical nature of a text. DGB only organizes some units into groups as the local hierarchical structures. DLTAG (PDTB) does not attempt to represent the overall structure. A connective and its two arguments compose a discourse relation instance that is independent of other discourse relations in the relation set.

The representation of a discourse theory is restricted by its connection constraint. For RST and CDT, only the connections between adjacent discourse units are allowed, and the units cannot overlap. Thus, a tree that obeys the constraint prevents crossing dependencies and reentrancies. For SDRT, the Right Frontier Constraint (RFC) is set to select the attachment point set of a new unit. The nonadjacent units dominating the last unit are allowed to connect, so a directed acyclic graph is constructed. As a comparison, the directed graph of DGB is more unconstrained because no constraint is assumed when a DGB graph is constructed.

In terms of the text coverage, representations of RST, DGB, CDT, and SDRT can provide a full coverage of a text owing to their overall structures. By contrast, DLTAG (PDTB) covers partial text because an argument only includes the clauses (or phrases) that are needed to interpret a rhetorical relation [81], and any unnecessary clauses are omitted. For example, in Example 2.3, the clause "said David Cooke, executive director of the RTC" is omitted in the representation of PDTB (Figure 4). While in the RST tree, it connects with the clause "It's a problem that clearly has to be resolved" with the relation "Attribution."

Table 2. Aspects of the Expressiveness of Representations Based on Entity Relations

Theory	Formal Representation	Text Coverage	Structural Coverage
OCAS	One or multiple chains with the coreferential mentions as nodes	partial	4 (I, II, III, IV)
TP (CDTC)	A chain with the themes or rhemes as nodes	full	2 (I, II)
GCDS	A set with the generalized topics and their comments as elements, which can be represented in the term of indented new-line	full	3 (I, II, III)

To compare the possible structures different discourse representations can cover, we analyze the structural coverage. The tree forms in RST and CDT cannot represent long-distance relations (u_1 and u_3 in VII). In contrast, the graph forms in SDRT and DGB have more flexible structures. The Right Frontier Constraint in SDRT requires that u_3 connects with one unit in the partially built representation (i.e., u_1 or u_2 or the complex unit u_1-u_2) and u_3 cannot construct a complex unit with u_1. So SDRT has four structures. For DGB, a graph with little constraint can represent all seven structures. In general, a graph has larger structural coverage than a tree.

3.3 The Expressiveness of Representations Based on Entity Relations

Table 2 presents a comparison of the expressiveness of discourse representation theories based on entity relations.[2]

Unlike the formal representation of theories based on rhetorical relations, there are almost no strong assumptions about the topological shape in theories based on entity relations. The links between mentions (or themes, topics) build the core structure. Therefore, for OCAS and TP (CDTC), the discourse structure is regarded as one or multiple chains. Even though GCDS looks different in formal representation, to a certain extent, its structure can be regarded as multiple chains where a generalized topic corresponds to a chain with the topic as nodes. The distinctions between these formal representations are mainly from the number of chains and the content of nodes (i.e., mentions, themes/rhemes, or generalized topics). A text usually contains multiple different coreferential mentions or topics that can constitute multiple independent chains in a representation. But the representation of TP (CDTC) is an exception, which usually contains only one chain because all clauses are sequentially connected to represent a text as a continuous information stream.

The text coverage of a discourse representation based on entity relations is mainly determined by the content of nodes. Text information represented by OCAS is incomplete. The representation of OCAS does not cover the words or phrases that are not in coreferential mentions. By contrast, TP (CDTC) provides a full coverage of a text. Each word in a clause belongs to a theme or rheme, and each theme or rheme is represented. For GCDS, each word in a clause belongs to a generalized topic or comment, therefore its representation also fully covers a text.

All theories based on entity relations cannot cover the structure (V), (VI), and (VII) because they do not attempt to represent the hierarchical structures. Their structural coverage is influenced by the number of nodes in a clause. There are usually multiple mentions in a clause, so any two units can be connected by coreferential mentions in OCAS, which makes it possible to represent the structures (I), (II), (III), and (IV). However, TP (CDTC) and GCDS cannot represent the ring

[2]The connection constraint is not compared because there is no constraint to restrict the entity links in all representations based on entity relations. Long-distance connections are frequent in OCAS. For TP (CDTC) and GCDS, although the connections between theme-rhemes or generalized topics are usually in adjacent units, it is legitimate to connect units in long distance.

structure (IV) where each unit is connected to the other two units because a clause usually has only one topic or theme (rheme) as the connected node.

3.4 Conclusions of the Expressiveness

From the comparison of the representation characteristics on four aspects, we draw the following conclusions about the expressiveness of existing discourse theories.

- Different topological forms have diverse structural expressiveness. To sum up, there are four forms: set, chain, tree, and graph. The structural expressiveness of the set in DLTAG (PDTB) is the weakest. The tree in RST and CDT can represent the hierarchical nature of a text but cannot represent the nonadjacent long-distance connections. In contrast, the graph in SDRT and DGB has stronger structural expressiveness. It can cover all possible structures if no constraint is imposed. The chain in most representations based on entity relations is very flexible to connect long-distance units, but it cannot represent the hierarchical structure.
- For the connection constraints in theories based on rhetorical relations, RST and CDT are the most strict and DGB is the most unconstrained. SDRT stands in between.
- For the representations based on entity relations, TP and GCDS have larger text coverage than OCAS. Although entity relations hold between coreferential words or phrases, other elements in the clauses can be defined (e.g., rhemes in TP and comments in GCDS) to describe something related to the entities, which can provide more complete information.

4 THE PRACTICALITY OF DISCOURSE REPRESENTATIONS

Although the sufficient expressiveness in different aspects is necessary, the discourse representation theory should not be discussed in a vacuum. Actually, when we consider NLP, the **practicality** is more important for a reasonable discourse representation. We believe a practical representation should be suitable for large-scale annotation, learnable through machine learning algorithms, and useful for downstream tasks.

4.1 Practicality

It is not easy to assess whether a representation is practical. Given that each discourse theory has its own set of assumptions, it is impossible to compare them directly. Hence, three aspects are defined to compare their practicality.

Scalability. It evaluates the difficulty of annotation. The construction of a large-scale corpus benefits from a scalable annotation scheme. Intuitively, the simpler the annotation scheme, the more scalable it will be. So the scale of existing corpora for each language and each theory are compared.

Learnability. It evaluates the complexity of system implementation. The performance of a discourse parser trained with machine learning algorithms should be good enough to provide reliable discourse information for downstream NLP tasks. We present state-of-the-art performance for discourse parsers that we are aware of trained on data annotated with each representation.

Usability. It evaluates the range of applications using data annotated with each representation. The existing applications of each discourse theory have been simply described in Section 2. Now they are systematically compared and these include *summarization*, *machine translation*, *sentiment analysis*, *question answering*, *text quality* (*coherence*) *assessment*, and *coreference* (*pronoun*) *resolution*.

Table 3. Aspects of the Practicality of Representations Based on Rhetorical Relations

Theory	Scalability	Learnability	Usability
RST	385 English news; 97 Chinese new comments	61.6% F1 in English	Sentiment analysis, machine translation, question answering, and summarization
DGB	135 English news	None	None
DLTAG (PDTB)	2,159 English news; 500 Chinese news	27.8% F1 in English; 26.9% F1 in Chinese	Sentiment analysis, machine translation, text quality assessment, and summarization
CDT	500 Chinese news	20.0% F1 in Chinese	Zero pronoun resolution
SDRT	1,091 English multi-party game dialogues	51.6% F1 in English	Sentiment analysis, and summarization

The "None" in "Learnability" and "Usability" represents no end-to-end parser and no application utilizing the discourse theory, respectively. There is the same meaning in Table 4.

4.2 The Practicality of Representations Based on Rhetorical Relations

The details of each aspect for practicality of discourse representation theories based on rhetorical relations are summarized in Table 3.

Data availability partly reflects the scalability of a theory. PDTB, the largest corpus based on rhetorical relations, annotates individual relations instead of the overall structures to simplify annotation. Most disagreements in its annotation come from the determination of the argument boundaries and the recognition of implicit relations. The annotated datasets in RST, DGB, CDT, and SDRT are relatively small. Their annotations are more difficult, because annotators disagree on the overall discourse structure of a text. In addition, just presenting a text in its entirety in an annotation tool is quite a challenge. Among them, DGB faces the most serious challenge because it does not impose any constraint on which pairs of discourse units can be rhetorically related.

In terms of practical discourse parsers, end-to-end RST and PDTB parsers have been widely studied. The state-of-the-art performance of RST parser is 61.6% F1 [43]. As a comparison, in PDTB, they are only 27.8% F1 [75] and 26.9% F1 [48] in English and Chinese, respectively. Despite the fact that there is a larger training set annotated with PDTB discourse relations, its end-to-end performance is low. There are two main reasons for this. The first one is that there is a long pipeline from detecting discourse connectives to predicting argument boundaries to predicting discourse relations between the two arguments. Errors propagate in a long-pipeline system. The second reason is that some arguments are very long and accurately predicting their boundaries is quite a challenge. The difficulty of implicit relation recognition also has a great impact on CDT parser, whose state-of-the-art performance is 20.0% F1 [51]. Afantenos et al. [1] build their dialogue parser with the 51.6% F1-value guided by SDRT.

Given the usability of representations based on rhetorical relations, RST and DLTAG (PDTB) are two popular representations. Both of them are applied to summarization, sentiment analysis, and machine translation. In contrast, the attention paid to the application of SDRT and CDT is relatively little. DGB has never been tried in NLP tasks. On one hand, the application of a theory is influenced by the performance of a system. On the other hand, it is also influenced by the needs of NLP tasks.

Different tasks require different discourse information. For summarization where the important messages are retained and unnecessary messages are omitted, the hierarchical structure in RST is naturally applicable [67]. Louis et al. [64] have proven that the structure information is the most robust indicator on content selection, and the rhetorical relation is also valuable to sentence extraction. Thus, the rhetorical relations in PDTB and SDRT are utilized [49, 114]. For sentiment

Table 4. Aspects of the Practicality of Representations Based on Entity Relations

Theory	Scalability	Learnability	Usability
OCAS	2,384 English documents; 1,729 Chinese documents	68.8% Avg. F1 in English; 63.9% Avg. F1 in Chinese	Coreference resolution, sentiment analysis, question answering, machine translation, text coherence assessment, and summarization
TP (CDTC)	500 Chinese news	52.6% Acc. in Chinese	Coreference resolution
GCDS	Chinese documents with 400,000 words	None	Machine translation

analysis, some researches show that the sentiment polarity has significant association with the relation type [41, 52, 57, 70, 94]. Therefore, the relations in RST, PDTB, and SDRT are widely utilized [4, 6, 89]. For machine translation, the translation of complex sentences needs the guidance of discourse structures provided by RST [91, 92]. The translation of connectives in PDTB has also received considerable attention [68]. For question answering, there exists the association between the answering sentences and a sequence of questions. RST is considered to capture the rhetorical association [42, 72]. For text quality (coherence) assessment, a common hypothesis is that a high-quality text has better coherence which can be evaluated through the transitions of relations in PDTB [114]. For zero pronoun resolution, EDUs in CDT are fine units to extract features [15].

4.3 The Practicality of Representations Based on Entity Relations

Now we discuss the practicality of discourse representations based on entity relations. Table 4 lists each aspect.

OCAS is the largest multilingual coreference corpus, while TP (CDTC) and GCDS only annotate Chinese data. The core effort of the annotation in representations based on entity relations is the determination of mentions, theme-rhemes, or generalized topics. OCAS does not annotate additional information and the entity relations have been established in the process of annotation. The same is true for GCDS. The annotation of TP (CDTC) is more difficult because of the cost on the annotation of lexical cohesion.

Given the learnability, there are many coreference resolution systems trained on the dataset of OCAS. Lee et al. [53] achieve state-of-the-art end-to-end English coreference resolution performance with 68.8% average F1 (Avg. F1), and the performance of a Chinese system is 63.9% average F1 [20]. Xi [101] builds a Chinese discourse topic analysis system guided by TP (CDTC), with 52.6% accuracy (Acc.). The error is mainly derived from the recognition of themes. The parsers for GCDS have not been developed yet.

OCAS has a wide range of the applications. On one hand, coreference links are regarded as important indicators to assess the text coherence [76] or ensure a coherent output in summarization [82]. On the other hand, discourse units with coreferential links are grouped together as the context information to improve the sentence-level sentiment analysis [107]. In addition, question answering and machine translation require the resolution of special coreference forms like pronoun [13, 65]. By contrast, other theories based on entity relations have not been widely used yet.

4.4 Conclusions of the Practicality

A comparison of the practicality of existing discourse representation theories leads to the following conclusions:

- In theories based on rhetorical relations, the performance of RST or SDRT parser is far higher than DLTAG (PDTB), despite the different corpus and framework. To a certain extent, this indicates that RST and SDRT are more suitable for automatic analysis. A simpler annotation does not necessarily lead to easier implementation. One reason why this is the case is that RST parsers do not have a long pipeline like PDTB parsers do. Another reason is that RST and SDRT are represented with well-defined mathematical objects like trees and graphs that are easier to manipulate computationally. This suggests one way forward to design a representation that is easier to annotate than RST and SDRT and more amenable computation than PDTB.
- Different discourse representations are adept in different applications. The representations of RST, PDTB, and OCAS have a wider range of applications than others. In general, the representations based on rhetorical relations are more appropriate for the tasks that rely on the structure or semantic information, such as summarization and sentiment analysis. The representations based on entity relations mainly concentrate on the interpretation of cohesion phenomena, such as coreference (pronoun) resolution. This suggests both types of discourse relations are complementary with each other and both are needed.

4.5 Further Discussion

The conclusions in Sections 3.4 and 4.4 summarize some important characteristics of existing discourse representation theories. Now, we take the expressiveness and practicality into account together to evaluate the representations of existing theories and obtain the following conclusions that reveal some basic principles and perspectives to explore a discourse representation. We firmly believe that they are enlightening for scalable Chinese discourse annotation.

- None of the existing representations are quite suitable for scalable Chinese discourse annotation in the age of machine learning. As discussed above, different representations have their own advantages and disadvantages. For representations based on rhetorical relations, the representation of DLTAG (PDTB) is easy to annotate but suffers from weak structural expressiveness and low system performance. The trees in RST and CDT can represent the hierarchical structures but cannot represent long-distance relations. Both cases can be represented by the graphs of DGB and SDRT. The trees and graphs provide a full coverage of a text. Despite their strong expressiveness, the construction of large-scale corpora is relatively difficult for them. Representations based on entity relations can represent long-distance relations. It is easy to annotate entity relations and OCAS has been proved to have good learnability and usability. Unfortunately, they cannot represent the hierarchical structures, and OCAS does not cover a text completely.

  Generally speaking, all representations are not both expressive and practical. In this case, a new discourse representation (or annotation scheme) needs to be proposed to take full consideration of the expressiveness and practicality.
- There is a tradeoff between the expressiveness and practicality. More powerful representation in terms of expressiveness provides valuable information for discourse analysis. However, an overly complex representation is more expensive to annotate and harder to compute. This inspires us to set appropriate connection constraints to simplify the representation of structure, and define the content of annotation with existing lower-level tasks (e.g., semantic role labeling) and algorithms as a reference.
- Rhetorical relations and entity relations coexist in a text and both are needed to represent discourse information properly. Rhetorical relations express the semantic association between discourse units, while entity relations can capture the entities in a text and

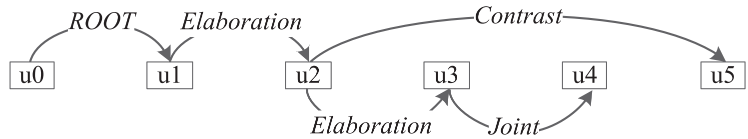

Fig. 11. The dependency representation of Example 5.1. u_0 is introduced as the root node. The arrow points from the head to the dependent and semantic relations are labeled.

group units describing the same entity. The two types of relations provide complementary information for downstream NLP tasks. Therefore, it is better to cover them in a unified annotation scheme.

5 FUTURE DIRECTIONS

The comparison and analysis of the existing discourse representations reveals at least three major challenges ahead. Meanwhile, inspired by the conclusions, some preliminary solutions are suggested.

(1) **The representation of discourse structure.** Discourse structure has been represented into various forms but existing representations have not achieved the ideal tradeoff between expressiveness and practicality. Specifically, we think there are two main issues that need to be further explored.

The first is what kind of topological forms a discourse structure should represent. As discussed in Section 3.4, DLTAG (PDTB) does not represent the overall structure and the graphs in SDRT and DGB have stronger structural expressiveness than the trees in RST and CDT. So the graph form can be considered first. Unfortunately, the complexity of the graphs in SDRT and DGB creates difficulties in annotation and parsing. We argue that the dependency graph may strike the right balance between expressiveness and practicality.

A dependency graph connects discourse units with a set of dependency relations. A dependency relation holds between a subordinate unit called *dependent* and another unit on which it depends, called the *head*. The semantic relations can be labeled on the head-dependent links. For example, Figure 11 represents a possible dependency structure in Example 5.1.

Example 5.1. [u_1 John has two very different brothers.] [u_2 One is a scientist] [u_3 who is friendly] [u_4 and has a great achievement in his field,] [u_5 while the other is an irascible murderer.]

We believe that the dependency graph has three advantages. (a) The head-dependent links can flexibly represent the hierarchy and nonadjacent connections. (b) It is easier to annotate and compute because of some simplifications. The category of labels is simplified. Unlike RST that labels "nucleus-satellite" relations and rhetorical relations separately, a dependency representation only labels the head-dependent relations with the "nucleus-satellite" relations represented structurally. Another simplification is the elimination of "non-terminal" nodes that are used to represent composite discourse units in RST. All nodes in a dependency graph are elementary discourse units. The nodes of complex units are replaced by the relations among elementary units. This results in simpler structures for discourse parsing. This also simplifies the annotation task. Annotators only need to worry about relations between head-dependent pairs rather than the large structures in RST, CDT, and SDRT. A simple and intuitive representation is essential to building large-scale resources. (c) Existing corpora and techniques can be used to train dependency discourse parsers. Danlos [23, 24] has shown that SDRT graph and RST tree can be converted into dependency graphs. Hence, corpora annotated with these representations can be converted to dependency representations. In addition, some algorithms for syntactic dependency analysis are adaptable to training discourse dependency parsers.

In fact, there has been some related research on transforming other discourse representations into the dependency form [38, 59, 71]. However, there is almost no comprehensive dependency-based discourse theories and annotated corpora. As far as we know, Wu et al. [100] annotate the first discourse dependency corpus on Chinese and English texts with 26 relation types. Nevertheless, the corpus is small in size and the annotation is preliminary. There are a lot of details to be defined and perfected, such as how to segment the discourse units, how to determine the connections, and so on.

The second issue about the representation of discourse structure is how to constrain the size of the candidate connection set. A discourse dependency graph without any constraints would be computationally intractable. A natural constraint would be that there can only be one head discourse unit per dependency unit. This will result in a dependency tree, which will be computationally tractable.

(2) **The understanding of semantic relations.** The semantic relations (including rhetorical relations and dependency relations) between discourse units play a significant role in discourse comprehension and reasoning. However, semantics is an abstract concept. Its understanding is often dependent on knowledge or common sense without any support of cohesion phenomena except for connectives, which causes the arbitrariness and ambiguity of the relation annotation and recognition. To better understand the semantic relations, a discourse representation theory should pay attention to the correlation between semantic relations and cohesion phenomena.

Different semantic relation types may be statistically correlated with certain cohesion phenomena. Existing theories like CDT and GCDS have proved that cohesion phenomena such as connectives and zero topics, can help to organize the discourse structure and explain semantic relations. Furthermore, the cohesion phenomena are rich and obvious in text, which will offer clues to reduce the subjectivity of the relation type determination. Therefore, we think it may be meaningful to annotate some cohesion phenomena for the understanding of deep semantic relations.

(3) **The analysis of Chinese discourse topic.** Topic is an important concept in Chinese linguistics and the topics in a text contribute to its coherence alongside the rhetorical relations that hold a text together. For example, some units sharing a topic chain can be grouped so a text is segmented into some groups with different topics. Hence, it is important to take into account the role of topics in a Chinese discourse representation theory.

TP (CDTC) and GCDS have focused on topics. However, a lot of problems have not been solved satisfactorily, such as the definition of topic and its representation, the reference of zero topics, the extension of topic chains from complex sentences to entire document, and so forth. These basic problems encourage us to find answers from topic-based discourse analysis. In addition, when exploring a topic-based discourse representation, the boundary of topic should be easy to annotate and compute.

Existing discourse theories about the Chinese topic only connect topics with entity relations, but they do not annotate the semantic relations between a topic and its comments. Although TP (CDTC) and GCDS introduce rhemes (for themes) and comments (for generalized topics) to ensure a full text coverage, they do not annotate the roles of a comment.

For example, a comment can be the explanation, evaluation, background, or result of a topic. These semantic relations may be helpful for some Chinese NLP tasks. The representation and annotation of Chinese topic needs to be further studied.

REFERENCES

[1] Stergos Afantenos, Eric Kow, Nicholas Asher, and Jérémy Perret. 2015. Discourse parsing for multi-party chat dialogues. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 928–937.

[2] Stergos D. Afantenos and Nicholas Asher. 2010. Testing SDRT's right frontier. In *Proceedings of the 23rd International Conference on Computational Linguistics*. Association for Computational Linguistics, 1–9.

[3] Nicholas Asher. 2012. *Reference to Abstract Objects in Discourse*. Vol. 50. Springer Science & Business Media.

[4] Nicholas Asher, Farah Benamara, and Yvette Yannick Mathieu. 2008. Distilling opinion in discourse: A preliminary study. *Coling 2008: Companion volume: Posters* (2008), 7–10.

[5] Nicholas Asher and Alex Lascarides. 2003. *Logics of Conversation*. Cambridge University Press.

[6] Parminder Bhatia, Yangfeng Ji, and Jacob Eisenstein. 2015. Better document-level sentiment analysis from RST discourse parsing. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2212–2218.

[7] Lidong Bing, Piji Li, Yi Liao, Wai Lam, Weiwei Guo, and Rebecca J. Passonneau. 2015. Abstractive multi-document summarization via phrase selection and merging. *Computational Linguistics* 31, 4 (2015), 505–530.

[8] Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski. 2001. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of the 2nd SIGdial Workshop on Discourse and Dialogue*. 1–10.

[9] Yllias Chali, Moin Tanvee, and Mir Tafseer Nayeem. 2017. Toward abstractive multi-document summarization using submodular function-based framework, sentence compression and merging. In *Proceedings of the 8th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*. Asian Federation of Natural Language Processing, 418–424.

[10] Yuen Ren Chao. 1965. *A Grammar of Spoken Chinese*. University of California Press.

[11] Chen Chen and Vincent Ng. 2013. Chinese zero pronoun resolution: Some recent advances. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 1360–1365.

[12] Chen Chen and Vincent Ng. 2016. Chinese zero pronoun resolution with deep neural networks. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 778–788.

[13] Danqi Chen, Jason Bolton, and Christopher D. Manning. 2016. A thorough examination of the cnn/daily mail reading comprehension task. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, Vol. 1. Association for Computational Linguistics, 2358–2367.

[14] Jifan Chen, Qi Zhang, Pengfei Liu, Xipeng Qiu, and Xuanjing Huang. 2016. Implicit discourse relation detection via a deep architecture with gated relevance network. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, Vol. 1. Cambridge University Press, 1726–1735.

[15] Sheng Cheng, Kong Fang, and Zhou Guodong. 2017. Towards better Chinese zero pronoun resolution from discourse perspective. In *National CCF Conference on Natural Language Processing and Chinese Computing*. Springer, 406–418.

[16] Jose M. Chenlo, Alexander Hogenboom, and David E. Losada. 2013. Sentiment-based ranking of blog posts using rhetorical structure theory. In *International Conference on Application of Natural Language to Information Systems*. Springer, 13–24.

[17] Chauncey C. Chu. 1993. The prototypicality of Ttpic in Mandarin Chinese. *Journal of the Chinese Language Teachers Association* 28, 1 (1993), 25–48.

[18] Cheng Hsi Chu. 1998. *A Discourse Grammar of Mandarin Chinese*. P. Lang.

[19] Kevin Clark and Christopher D. Manning. 2015. Entity-centric coreference resolution with model stacking. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, Vol. 1. Association for Computational Linguistics, 1405–1415.

[20] Kevin Clark and Christopher D. Manning. 2016. Deep reinforcement learning for mention-ranking coreference models. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2256–2262.

[21] Kevin Clark and Christopher D. Manning. 2016. Improving coreference resolution by learning entity-level distributed representations. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 643–653.

[22] E. Dane. 1974. Functional sentence perspective and the organization of the text.

[23] Laurence Danlos. 2004. Discourse dependency structures as constrained DAGs. In *5th SIGDIAL Workshop on Discourse and Dialogue*.

[24] Laurence Danlos. 2005. Comparing RST and SDRT discourse structures through dependency graphs. *Proceedings of Constraints in Discourse* (2005), 53.

[25] Kees van Deemter and Rodger Kibble. 2000. On coreferring: Coreference in MUC and related annotation schemes. *Computational Linguistics* 26, 4 (2000), 629–637.

[26] George R. Doddington, Alexis Mitchell, Mark A. Przybocki, Lance A. Ramshaw, Stephanie Strassel, and Ralph M. Weischedel. 2004. The automatic content extraction (ACE) program-tasks, data, and evaluation. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*. European Language Resources Association (ELRA).

[27] Angela Downing. 2001. Thematic progression as a functional resource in analysing texts. *CLAC (Circulo de Lingüística Aplicada a la Communicacion)*.

[28] David A. Duverle and Helmut Prendinger. 2009. A novel discourse parser based on support vector machine classification. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*. Association for Computational Linguistics, 665–673.

[29] Vanessa Wei Feng and Graeme Hirst. 2012. Text-level discourse parsing with rich linguistic features. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 60–68.

[30] Rafael Ferreira, Frederico Freitas, Luciano de Souza Cabral, Rafael Dueire Lins, Rinaldo Lima, Gabriel França, Steven J. Simskez, and Luciano Favaro. 2013. A four dimension graph model for automatic text summarization. In *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 01*. IEEE Computer Society, 389–396.

[31] Katherine Forbes, Eleni Miltsakaki, Rashmi Prasad, Anoop Sarkar, Aravind Joshi, and Bonnie Webber. 2003. D-LTAG system: Discourse parsing with a lexicalized tree-adjoining grammar. *Journal of Logic, Language and Information* 12, 3 (2003), 261–279.

[32] Shima Gerani, Yashar Mehdad, Giuseppe Carenini, Raymond T. Ng, and Bita Nejat. 2014. Abstractive summarization of product reviews using discourse structure. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 1602–1613.

[33] Barbara J. Grosz, Scott Weinstein, and Aravind K. Joshi. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21, 2 (1995), 203–225.

[34] Francisco Guzmán, Shafiq Joty, Lluís Màrquez, and Preslav Nakov. 2014. Using discourse structure improves machine translation evaluation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Vol. 1. Association for Computational Linguistics, 687–698.

[35] Michael Halliday, Christian MIM Matthiessen, and Christian Matthiessen. 2014. *An Introduction to Functional Grammar*. Routledge.

[36] Bas Heerschop, Frank Goossen, Alexander Hogenboom, Flavius Frasincar, Uzay Kaymak, and Franciska de Jong. 2011. Polarity analysis of texts using discourse structure. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. ACM, 1061–1070.

[37] Hugo Hernault, Helmut Prendinger, Mitsuru Ishizuka, et al. 2010. HILDA: A discourse parser using support vector machine classification. *Dialogue and Discourse* 1, 3 (2010), 1–33.

[38] Tsutomu Hirao, Yasuhisa Yoshida, Masaaki Nishino, Norihito Yasuda, and Masaaki Nagata. 2013. Single-document summarization as a tree knapsack problem. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 1515–1520.

[39] Lynette Hirschman. 1997. MUC-7 coreference task definition.

[40] Eduard Hovy, Mitchell Marcus, Martha Palmer, Lance Ramshaw, and Ralph Weischedel. 2006. OntoNotes: The 90% solution. In *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*. Association for Computational Linguistics, 57–60.

[41] Hen-Hsen Huang, Chi-Hsin Yu, Tai-Wei Chang, Cong-Kai Lin, and Hsin-Hsi Chen. 2013. Analyses of the association between discourse relation and sentiment polarity with a Chinese human-annotated corpus. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*. Association for Computational Linguistics, 70–78.

[42] Peter Jansen, Mihai Surdeanu, and Peter Clark. 2014. Discourse complements lexical semantics for non-factoid answer reranking. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Vol. 1. Association for Computational Linguistics, 977–986.

[43] Yangfeng Ji and Jacob Eisenstein. 2014. Representation learning for text-level discourse parsing. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 13–24.

[44] Yangfeng Ji and Jacob Eisenstein. 2015. One vector is not enough: Entity-augmented distributional semantics for discourse relations. *Transactions of the Association for Computational Linguistics* 3 (2015), 329–344.

[45] Aravind K. Joshi and Yves Schabes. 1991. Tree-adjoining grammars and lexicalized grammars. *Technical Reports (CIS)* (1991), 445.

[46] Shafiq Joty, Giuseppe Carenini, Raymond Ng, and Yashar Mehdad. 2013. Combining intra- and multi-sentential rhetorical parsing for document-level discourse analysis. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, Vol. 1. 486–496.

[47] Hans Kamp, Josef Van Genabith, and Uwe Reyle. 2011. Discourse representation theory. In *Handbook of Philosophical Logic*. Springer, 125–394.

[48] Xiaomian Kang, Haoran Li, Long Zhou, Jiajun Zhang, and Chengqing Zong. 2016. An end-to-end Chinese discourse parser with adaptation to explicit and non-explicit relation recognition. *Proceedings of the CoNLL-16 Shared Task* (2016), 27–32.

[49] Iskandar Keskes. 2015. *Discourse Analysis of Arabic Documents and Application to Automatic Summarization*. Ph.D. dissertation. Université de Toulouse, Université Toulouse III-Paul Sabatier.

[50] Fang Kong and Hwee Tou Ng. 2013. Exploiting zero pronouns to improve Chinese coreference resolution. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 278–288.

[51] Fang Kong and Guodong Zhou. 2017. A CDT-styled end-to-end Chinese discourse parser. *ACM Transactions on Asian and Low-Resource Language Information Processing* 16, 4 (2017), 26.

[52] Angeliki Lazaridou, Ivan Titov, and Caroline Sporleder. 2013. A Bayesian model for joint unsupervised induction of sentiment, aspect and discourse representations. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 1630–1639.

[53] Kenton Lee, Luheng He, Mike Lewis, and Luke Zettlemoyer. 2017. End-to-end neural coreference resolution. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 188–197.

[54] Charles N. Li. 1976. *Subject and Topic*. Academic Press.

[55] Haoran Li, Jiajun Zhang, and Chengqing Zong. 2017. Implicit discourse relation recognition for English and Chinese with multiview modeling and effective representation learning. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)* 16, 3 (2017), 19.

[56] Jiwei Li, Rumeng Li, and Eduard Hovy. 2014. Recursive deep models for discourse parsing. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2061–2069.

[57] Junjie Li, Yu Zhou, Chunyang Liu, and Lin Pang. 2014. Sentiment classification of Chinese contrast sentences. In *Natural Language Processing and Chinese Computing*. Springer, 205–216.

[58] Junyi Jessy Li, Marine Carpuat, and Ani Nenkova. 2014. Assessing the discourse factors that influence the quality of machine translation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Cambridge University Press, 283–288.

[59] Sujian Li, Liang Wang, Ziqiang Cao, and Wenjie Li. 2014. Text-level discourse dependency parsing. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 25–35.

[60] Yancui Li, Fang Kong, Guodong Zhou, et al. 2014. Building Chinese discourse corpus with connective-driven dependency tree structure. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. Cambridge University Press, 2105–2114.

[61] Ziheng Lin, Hwee Tou Ng, and Min-Yen Kan. 2014. A PDTB-styled end-to-end discourse parser. *Natural Language Engineering* 20, 2 (2014), 151–184.

[62] Ting Liu, Yiming Cui, Qingyu Yin, Weinan Zhang, Shijin Wang, and Guoping Hu. 2017. Generating and exploiting large-scale pseudo training data for zero pronoun resolution. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Vol. 1*. Association for Computational Linguistics, 102–111.

[63] Yang Liu, Jiajun Zhang, and Chengqing Zong. 2017. Memory augmented attention model for Chinese implicit discourse relation recognition. In *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*. Springer, 411–423.

[64] Annie Louis, Aravind Joshi, and Ani Nenkova. 2010. Discourse indicators for content selection in summarization. In *Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. Association for Computational Linguistics, 147–156.

[65] Ngoc Quang Luong and Andrei Popescu-Belis. 2016. Improving pronoun translation by modeling coreference uncertainty. In *Proceedings of the 1st Conference on Machine Translation: Volume 1, Research Papers*. Association for Computational Linguistics, 12–20.

[66] William C. Mann and Sandra A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text-Interdisciplinary Journal for the Study of Discourse* 8, 3 (1988), 243–281.
[67] Daniel Marcu. 1997. From discourse structures to text summaries. *Intelligent Scalable Text Summarization*
[68] Thomas Meyer. 2011. Disambiguating temporal-contrastive discourse connectives for machine translation. In *Proceedings of the ACL 2011 Student Session*. Cambridge University Press, 46–51.
[69] Amit Mishra and Sanjay Kumar Jain. 2014. An approach for computing sentiment polarity analysis of complex why-type questions on product review sites. *Research in Computing Science* 84 (2014), 65–76.
[70] Subhabrata Mukherjee and Pushpak Bhattacharyya. 2012. Sentiment analysis in Twitter with lightweight discourse analysis. *Proceedings of COLING 2012* (2012), 1847–1864.
[71] Philippe Muller, Stergos Afantenos, Pascal Denis, and Nicholas Asher. 2012. Constrained decoding for text-level discourse parsing. *Proceedings of COLING 2012* (2012), 1883–1900.
[72] Karthik Narasimhan and Regina Barzilay. 2015. Machine comprehension with discourse relations. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, Vol. 1. Association for Computational Linguistics, 1253–1262.
[73] Vincent Ng. 2017. Machine learning for entity coreference resolution: A retrospective look at two decades of research. In *AAAI*. 4877–4884.
[74] Nicolas Nicolov, Franco Salvetti, and Steliana Ivanova. 2008. Sentiment analysis: Does coreference matter. In *AISB 2008 Convention Communication, Interaction and Social Intelligence*, Vol. 1. 37.
[75] Stephan Oepen, Jonathon Read, Tatjana Scheffler, Uladzimir Sidarenka, Manfred Stede, Erik Velldal, and Lilja Øvrelid. 2016. OPT: Oslo–Potsdam–Teesside. Pipelining rules, rankers, and classifier ensembles for shallow discourse parsing. *Proceedings of the CoNLL-16 shared task* (2016), 20–26.
[76] Emily Pitler, Annie Louis, and Ani Nenkova. 2010. Automatic evaluation of linguistic quality in multi-document summarization. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 544–554.
[77] Sameer Pradhan, Alessandro Moschitti, Nianwen Xue, Olga Uryupina, and Yuchen Zhang. 2012. CoNLL-2012 shared task: Modeling multilingual unrestricted coreference in OntoNotes. In *Joint Conference on EMNLP and CoNLL-Shared Task*. Association for Computational Linguistics, 1–40.
[78] Sameer Pradhan, Lance Ramshaw, Mitchell Marcus, Martha Palmer, Ralph Weischedel, and Nianwen Xue. 2011. Conll-2011 shared task: Modeling unrestricted coreference in ontonotes. In *Proceedings of the 15th Conference on Computational Natural Language Learning: Shared Task*. Association for Computational Linguistics, 1–27.
[79] Sameer S. Pradhan, Lance Ramshaw, Ralph Weischedel, Jessica MacBride, and Linnea Micciulla. 2007. Unrestricted coreference: Identifying entities and events in OntoNotes. In *International Conference on Semantic Computing*. IEEE, 446–453.
[80] Dragomir R. Radev. 2000. A common theory of information fusion from multiple text sources step one: Cross-document structure. In *Proceedings of the 1st SIGdial Workshop on Discourse and Dialogue-Volume 10*. Association for Computational Linguistics, 74–83.
[81] Prasad Rashmi, Dinesh Nihkil, Lee Alan, Miltsakaki Eleni, Robaldo Livio, Joshi Aravind, Webber Bonnie, et al. 2008. The Penn discourse Treebank 2.0. In *Lexical Resources and Evaluation Conference*.
[82] Michael Roth and Anette Frank. 2013. Automatically identifying implicit arguments to improve argument linking and coherence modeling. In *2nd Joint Conference on Lexical and Computational Semantics (* SEM), Volume 1: Proceedings of the Main Conference and the Shared Task: Semantic Textual Similarity*, Vol. 1. Association for Computational Linguistics, 306–316.
[83] Attapol Rutherford, Vera Demberg, and Nianwen Xue. 2017. A systematic study of neural discourse models for implicit discourse relation. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*. Cambridge University Press, 281–291.
[84] Rou Song and Shili Ge. 2013. English-Chinese translation unit and translation model for discourse-based machine translation. *Journal of Chinese Information Processing* 29, 15 (2013), 125–135.
[85] Rou Song, Yuru Jiang, and Jingyi Wang. 2010. On generalized-topic-based Chinese discourse structure. In *CIPS-SIGHAN Joint Conference on Chinese Language Processing*.
[86] Radu Soricut and Daniel Marcu. 2003. Sentence level discourse parsing using syntactic and lexical information. In *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 149–156.
[87] Josef Steinberger, Massimo Poesio, Mijail A. Kabadjov, and Karel Ježek. 2007. Two uses of anaphora resolution in summarization. *Information Processing and Management* 43, 6 (2007), 1663–1680.
[88] Maite Taboada and William C. Mann. 2006. Applications of rhetorical structure theory. *Discourse Studies* 8, 4 (2006), 567–588.

[89] Rakshit Trivedi and Jacob Eisenstein. 2013. Discourse connectors for latent subjectivity in sentiment analysis. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Cambridge University Press, 808–813.

[90] Feng-fu Tsao. 1990. Clause and sentence structure in Chinese: A functional perspective. *Taipei: Student Book Co.*.

[91] Mei Tu, Yu Zhou, and Chengqing Zong. 2013. A novel translation framework based on rhetorical structure theory. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, Vol. 2. Association for Computational Linguistics, 370–374.

[92] Mei Tu, Yu Zhou, and Chengqing Zong. 2014. Enhancing grammatical cohesion: Generating transitional expressions for SMT. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Vol. 1. Association for Computational Linguistics, 850–860.

[93] Suzan Verberne, L. W. J. Boves, N. H. J. Oostdijk, and P. A. J. M. Coppen. 2007. Discourse-based answering of why-questions.

[94] Fei Wang, Yunfang Wu, and Likun Qiu. 2012. Exploiting discourse relations for sentiment analysis. *Proceedings of COLING 2012: Posters* (2012), 1311–1320.

[95] Xun Wang, Yasuhisa Yoshida, Tsutomu Hirao, Katsuhito Sudoh, and Masaaki Nagata. 2015. Summarization based on task-oriented discourse parsing. *IEEE Transactions on Audio, Speech, and Language Processing* 23, 8 (2015), 1358–1367.

[96] Bonnie Webber. 2004. D-LTAG: Extending lexicalized TAG to discourse. *Cognitive Science* 28, 5 (2004), 751–779.

[97] Bonnie Webber, Markus Egg, and Valia Kordoni. 2012. Discourse structure and language technology. *Natural Language Engineering* 18, 4 (2012), 437–490.

[98] Ralph Weischedel, Martha Palmer, Mitchell Marcus, Eduard Hovy, Sameer Pradhan, Lance Ramshaw, Nianwen Xue, Ann Taylor, Jeff Kaufman, Michelle Franchini, et al. 2013. Ontonotes release 5.0. *LDC2013T19, Linguistic Data Consortium*.

[99] Florian Wolf and Edward Gibson. 2004. Representing discourse coherence: A corpus-based analysis. In *Proceedings of the 20th International Conference on Computational Linguistics*. Association for Computational Linguistics, 134.

[100] Yongfan Wu, Sujian Li, Mukun Qin, An Yang, and Houfeng Wang. 2017. Exploring Chinese and English discourse dependency treebanks. *Proceedings of Chinese Computational Linguistics 2017*.

[101] Xuefeng Xi. 2017. *Research on Chinese Discourse Topic Structure: Representation, Resource Construction and Its Analysis*. Ph.D. dissertation. Soochow University.

[102] Xuefeng Xi, Xiaoming Chu, Qingying Sun, and Guodong Zhou. 2017. Research and prospect of discourse topic structure analysis for discourse intentionality. *Chinese Journal of Computers* 40 (2017).

[103] Xuefeng Xi and Guodong Zhou. 2016. A micro-topic model for coreference resolution based on theme-rheme structure. In *Natural Language Understanding and Intelligent Applications*. Springer, 648–656.

[104] Xuefeng Xi and Guodong Zhou. 2017. Building a Chinese discourse topic corpus with a micro-topic scheme based on theme-rheme theory. *Big Data Analytics* 2, 1 (2017), 9.

[105] Nianwen Xue, Hwee Tou Ng, Sameer Pradhan, Rashmi Prasad, Christopher Bryant, and Attapol Rutherford. 2015. The conll-2015 shared task on shallow discourse parsing. In *Proceedings of the 19th Conference on Computational Natural Language Learning-Shared Task*. Association for Computational Linguistics, 1–16.

[106] Nianwen Xue, Hwee Tou Ng, Sameer Pradhan, Attapol Rutherford, Bonnie Webber, Chuan Wang, and Hongmin Wang. 2016. Conll 2016 shared task on multilingual shallow discourse parsing. *Proceedings of the CoNLL-16 Shared Task* (2016), 1–19.

[107] Bishan Yang and Claire Cardie. 2014. Context-aware learning for sentence-level sentiment analysis with posterior regularization. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Cambridge University Press, 325–335.

[108] Qingyu Yin, Yu Zhang, Weinan Zhang, and Ting Liu. 2017. Chinese zero pronoun resolution with deep memory network. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 1309–1318.

[109] Ming Yue. 2008. Rhetorical structure annotation of Chinese news commentaries. *Journal of Chinese Information Processing* 22, 4 (2008), 19–23.

[110] Biao Zhang, Jinsong Su, Deyi Xiong, Yaojie Lu, Hong Duan, and Junfeng Yao. 2015. Shallow convolutional neural network for implicit discourse relation recognition. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Cambridge University Press, 2230–2235.

[111] Lanjun Zhou, Binyang Li, Wei Gao, Zhongyu Wei, and Kam-Fai Wong. 2011. Unsupervised discovery of discourse relations for eliminating intra-sentence polarity ambiguities. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 162–171.

[112] Qiang Zhou and Xiaocong Zhou. 2014. Topic-chain-based coherence annotation scheme for Chinese text. *Journal of Chinese Information Processing* 28, 5 (2014), 102–111.

[113] Yuping Zhou and Nianwen Xue. 2012. PDTB-style discourse annotation of Chinese text. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1*. Cambridge University Press, 69–77.

[114] Lin Ziheng. 2012. *Discourse Parsing: Inferring Discourse Structure, Modeling Coherence, and Its Applications*. Ph.D. dissertation. National University of Singapore.

[115] Cäcilia Zirn, Mathias Niepert, Heiner Stuckenschmidt, and Michael Strube. 2011. Fine-grained sentiment analysis with structural features. In *Proceedings of the 5th International Joint Conference on Natural Language Processing*. Asian Federation of Natural Language Processing, 336–344.

Received June 2018; revised November 2018; accepted November 2018