

# Dependency Dialogue Acts — Annotation Scheme and Case Study

Jon Z. Cai, Brendan King, Margaret Perkoff, Shiran Dudy, Jie Cao, Marie Grace, Natalia Wojarnik, Ananya Ganesh, James H. Martin, Martha Palmer, Marilyn Walker and Jeffrey Flanigan

**Abstract** In this paper, we introduce Dependency Dialogue Acts (DDA), a novel framework for capturing the structure of speaker-intentions in multi-party dialogues. DDA combines and adapts features from existing dialogue annotation frameworks, and emphasizes the multi-relational response structure of dialogues in addition to the dialogue acts and rhetorical relations. It represents the functional, discourse, and response structure in multi-party multi-threaded conversations. A few key features distinguish DDA from existing dialogue annotation frameworks such as SWBD-DAMSL and the ISO 24617-2 standard. First, DDA prioritizes the relational structure of the dialogue units and the dialog context, annotating both dialog acts and rhetorical relations as response relations to particular utterances. Second, DDA embraces overloading in dialogues, encouraging annotators to specify multiple response relations and dialog acts for each dialog unit. Lastly, DDA places an emphasis on adequately capturing how a speaker is using the full dialog context to plan and organize their speech. With these features, DDA is highly expressive and recall-oriented with regard to conversation dynamics between multiple speakers. In what follows, we present the DDA annotation framework and case studies annotating DDA structures in multi-party, multi-threaded conversations.

## 1 Introduction

Discourse analysis has become an increasingly popular problem in natural language processing. Broadly, discourse analysis for dialog involves observing a conversation

---

Jon Z. Cai

University of Colorado Boulder, 1134 CO-93, Boulder, CO 80302, e-mail: jon.z.cai@colorado.edu

Brendan King, Jeffrey Flanigan

University of California Santa Cruz, 1156 High St., Santa Cruz, CA 95064, e-mail: bking2@ucsc.edu, jmflanig@ucsc.edu

between two or more individuals and understanding the information that is being exchanged, both explicitly and implicitly. One of the goals of dialogue analysis systems is to be able to understand the intents of the parties involved. This problem becomes more difficult when analyzing conversations between more than two people in an open environment. In these settings, side conversations and non-discourse events can interrupt an ongoing conversation. These multi-party multi-threaded scenarios are ones that we encounter on a daily basis. Additionally, these complex conversations contain abundant information that indicates the relationship between the speakers, their moods, their likes and dislikes, as well as their intentions. Presently, we are seeing an influx of conversational agents that are attempting to mimic our ability to not only interpret this information correctly but to generate an appropriate response to it as well.

Previous research in discourse analysis has led to a variety of annotation schemes that attempt to capture different aspects of conversations. One of the foundational schemes in this space is the Switchboard DAMSL [29] which annotates conversations at the utterance level based on their corresponding dialog acts. Dialog acts are used to represent the intention of the speaker, such as asking a *Question* or expressing a *Statement-Opinion*. Other alternative schemas include Rhetorical Structure Theory (RST) and shallow discourse relation frameworks such as Penn Discourse Tree Bank (PDTB) which are frequently used to analyze text structure and coherence. These schemes have proven extremely valuable in analyzing dialogue but can encounter unique challenges in multi-party, open-environment settings such as our domain of interest – classroom conversations, where conversation threads interweave and are interrupted by events outside the discourse.

With this in mind, we set out to design a discourse analysis scheme that is able to track the intentions of multiple speakers while preserving the relational information from one turn of the dialogue to the next. Furthermore, we want our scheme to be sufficiently useful for a conversational agent to generate an appropriate response in a multiparty conversation, which aligns with our goal of creating more explainable and controllable dialogue generation agents. Prior work has shown rhetorical structures and dialog acts can improve controllability and explainability in response generation in dialog agents [37, 10, 32].

Our proposed Dependency Dialogue Act (DDA) annotation scheme builds upon previous work on discourse annotation by merging different features from existing schemes into a single system that captures a large amount of conversational context while minimizing annotator effort. One of the primary goals is the ability to preserve both rhetorical and response relations between different turns in the utterance.

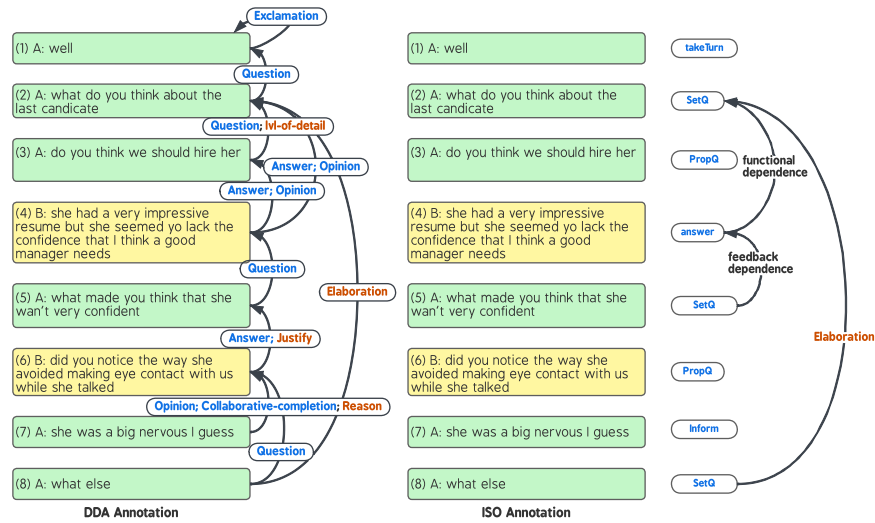
Additionally, we want to embrace the inherent overloading nature of conversations by enabling annotators to select multiple labels per utterance where appropriate. Finally, the DDA scheme anchors the speaker’s intention with context.

The goal of this paper is to define the Dependency Dialogue Act (DDA) annotation scheme for discourse analysis and investigate its effectiveness in the context of multi-threaded multi-party conversations. In Section 3, we define the response structure of DDA and present the tagset composed of two class types: Dialog Acts and Rhetorical Relations. We demonstrate the usefulness of the DDA scheme with examples from

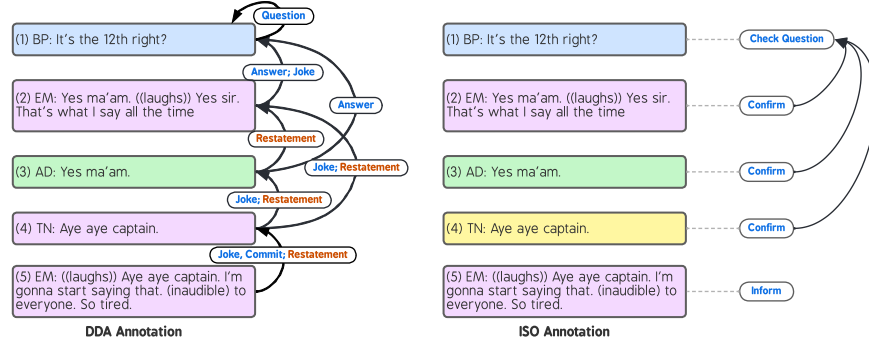
diverse conversation settings throughout. We discuss applications in the dialog analysis space in Section 4. We briefly review prior work on discourse annotation schemes and highlight key features that each of them captures in Section 5. The ability to adequately interpret multi-party multi-threaded conversations has significant implications for conversational technology across many domains; we hope that the DDA scheme is a step towards capturing more of the critical information present in these settings.

## 2 Motivation

We motivate our Dependency Dialogue Act (DDA) annotation scheme with three examples, shown in Figs. 1-3. DDA aims to capture as much information about the interrelationships between utterances as possible while also representing the multiple dialog acts and rhetorical relations that a single utterance can have. DDA captures the response structure, dialog acts, and rhetorical relations in one integrated graph structure. Compared to ISO 24617-2, DDA has more dialog acts for each utterance,



**Fig. 1** An example from the DialogBank corpus as originally annotated with ISO 24617-2 (right) and with Dependency Dialogue Act (DDA) (left). Dialog act labels are in blue, and rhetorical relations are in red. The ISO annotation contains a functional dependency (4 → 2), a feedback dependency (5 → 4) and a rhetorical relation (8 → 2), giving context for units 4,5,&8 respectively. DDA annotations are context-oriented, explicitly marking context with response dependencies for all units in a dialogue. This broader view of dialogue structure leads to a fully connected dialog thread that can be disentangled from others in the multi-party setting, by design. DDA annotations are also recall-oriented, encouraging the use of multiple labels for multi-function conversation units.



**Fig. 2** An example from our classroom setting, annotated with DDA (left) and ISO (right). DDA annotation embraces overloading of multi-purpose units such as in 2, ...5, yielding a comprehensive, recall-oriented structure.<sup>1</sup>

and more relations (Figs. 1 and 2).<sup>1</sup> DDA’s edges represent response relations, with conversation threads forming connected components, similar to reply-structure graphs in the Ubuntu-IRC corpus (Fig. 3) [31].

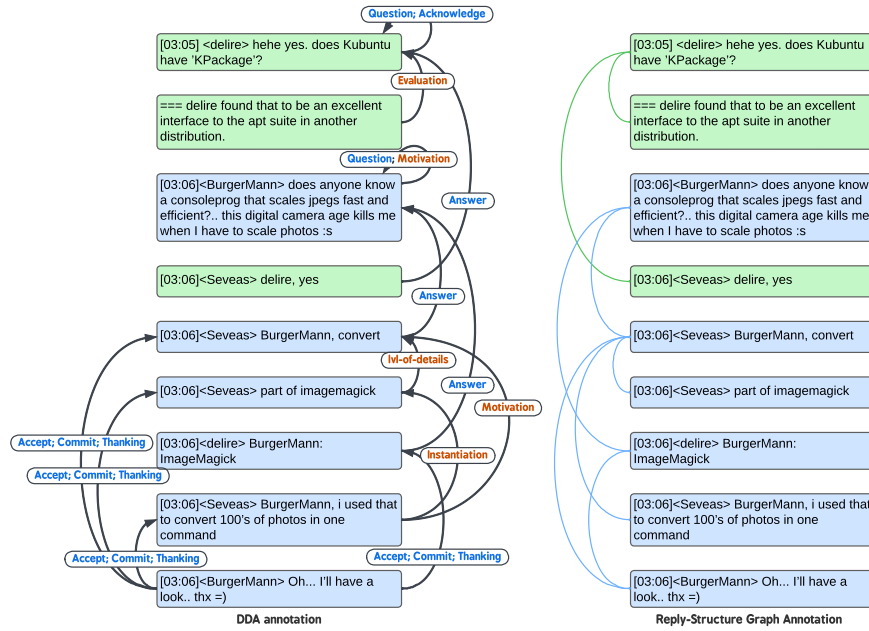
### 3 Dependency Dialog Act

We propose the Dependency Dialog Act (DDA) annotation scheme to capture a broad range of speaker intentions and their relationships to the dialogue context in the multi-party setting. We emphasize the following key design philosophies:

1. DDA is context oriented and recommends annotators think from a relational perspective. This is reflected in DDA by annotating dialog acts and relations on response edges to the surrounding context, rather than on dialog units (see end of Section 3.1).
2. DDA is recall-oriented and encourages annotators to put all response relations that fit for the given context. It embraces overloading as an important feature of the framework (see end of Section 3.2).
3. DDA pays attention to speaker intentions, trying to capture both the purpose of speech and “how” a speaker plans and arranges their speech conditioned on the context. This philosophy is reflected in the design decisions of DDA.

DDA aims at capturing *speaker intention* as a key feature. “Intention” is a widely studied concept in philosophy [5, 41], theory of action, and logic [28]. We follow

<sup>1</sup> In Fig. 2, (1) poses a question. (2) can be considered an “Answer” to (1). Similarly, the units that follow restate this joke while answering the question posed in (1), each having a different functional response dependency to (1) and rhetorical response dependency on (2), (3), and/or (4). Despite the appearance of an answer to (1), the intent of the speaker in (5) is more likely to be participation in the joke, as the question has already been answered by the same speaker.



**Fig. 3** Annotation of DDA (left) as compared to a reply-structure graph (right) from the Ubuntu-IRC corpus [31]. DDA contains all the edges in Ubuntu-IRC’s reply structure, with labeled dialog acts and rhetorical relations. DDA can contain additional edges for rhetorical relations and dialog acts. Both structures yield separable conversation threads as connected components.

the functionalist philosophy [20], defining **intentions** as operational plans either in our mind or can be entailed by current actions. For example, when a speaker provides a “action-directive” utterance, the speaker’s entailed plan is to have certain actions performed. Then, when providing further “elaboration,” the speaker’s plan is to make the existing statement more convincing or clear. DDA uses an enhanced dialog acts set from SWBD-DAMSL as the basis to describe actions performed, and enhanced discourse relations as the basis to describe discourse purpose plans. It embeds intentional information and context into the labeled dependency edges.

We introduce the Dependency Dialog Act annotation scheme in two sections: In Section 3.1, we define our response dependency relations between units of dialog (see Slash Units, below). In Section 3.2, we describe the adaptation of existing tag schema for dialogue acts and rhetorical relations to form the basis of DDA’s intention space.

**Dialog Units of Annotation: Slash Units** - Similar to *functional segments* in ISO 24617-2 standard and *elementary discourse units* (EDUs) in RST, we assume that a dialogue is broken up into units for annotation. Following the SWBD-DAMSL annotation scheme, we term these **slash units**.

### 3.1 DDA Edges: Response Dependencies

The edges in DDA indicate response relationships between slash units. Specifically, for a slash unit of interest, a **response dependency** is a directed edge from the unit of interest to the slash unit it depends on or originated from conversationally. When a slash unit  $u_i$  has no unit to relate to in the prior context, we use a **self-pointing dependency**  $u_i \rightarrow u_i$  to specify the start of a new thread of conversation.

DDA takes an expansive view of response relations between slash units which encompasses the functional, rhetorical, and reply relationships in other frameworks:

- **Functional dependency:**<sup>2</sup> the meaning of a dialogue act for a local slash unit depends crucially on a particular slash-unit in the dialogue context, such as how an Answer depends on a Question [15].
- **Rhetorical relations:** the coherent organization of two slash units, for example labeling units which elaborate on or contrast with previous units [33]. Also known as discourse relations.
- **Response or continuation dependency** represents continuation of a conversation thread but no explicit functional or rhetorical dependencies between two slash units.

In DDA, conversation threads form separate connected components in the DDA annotation graph. Consider Fig. 1, which includes DDA annotations for a snippet from the DialogBank corpus as originally annotated with the ISO 24617-2 standard [17]. While the ISO annotation includes multiple relation types, some dialogue units in the conversation thread remain disconnected from the structure. While two-party dialogues like this one often follow a single linear thread, this is not always the case[23]. For example, coherent threads can overlap and might require disentangling for further analysis, as seen in Figs. 3, 5 and 6.

We annotate dialog acts on response edges rather than on slash units. This is in contrast with most previous annotation schemes for dialog acts such as SWBD-DAMSL and ISO. The benefit of our approach is that it explicitly labels the context for each dialog act. For example, in Fig. 5, utterance 32 contains a question, asking “who wants to go first?”. In the DDA annotation, the context is explicitly marked by the response dependency, such that going first can be understood as a leading discussion of the first question in their packet. In the ISO annotation, this context would need to be inferred from the dialogue history, which may be difficult as many of the nearest slash units belong to a different conversation thread.

---

<sup>2</sup> In addition to functional dependence, the ISO 24617-2 standard defines feedback dependencies for particular feedback acts. Feedback acts largely correspond to backward-communicative-function dialogue acts in SWBD-DAMSL, which we adapt for use in DDA. Thus, we consider feedback dependencies as similar to functional ones, where the interpretation of the slash-unit and label heavily rely on the dependent unit(s).

	DDA	ISO	Ubuntu-IRC	STAC	SWBD-DAMSL
Dialog acts	yes	yes	no	yes (limited)	yes
Discourse relations	PDTB+	subset of PDTB	no	task specific	no
Reply structure	yes	partial	yes	no	no
Functional dependence	yes	yes	no	no	partial

**Table 1** Annotation Scheme Features Comparison. **Dialog act** and **PDTB** rows represent whether a scheme uses this set. **reply structure, continuation and functional dependence** denote three types of dependence structure defined in Section 3.1; *partial* means only a subset of the feature can be annotated under a scheme. STAC refers to the annotation scheme [7].

### 3.2 DDA Tagset: Dialog Act and Discourse Relation Classes

Many dialogue annotation frameworks label conversation units from one of two perspectives. First, there are frameworks for labeling the function or “act” of a dialog unit of interest, including DAMSL, SWBD-DAMSL, and the ISO 24617-2 standard. Other frameworks aim to model the discourse relations between units, drawing from Rhetorical Structure Theory [34, 40] or Segmented Discourse Representation Theory (SDRT) [8]. Since we want to capture speaker intentions, we aim to capture both categories of these phenomena in multi-party dialogue in a single annotation scheme, by adapting dialog acts from the SWBD-DAMSL scheme [29] and discourse relations from the Penn Discourse Tree Bank 3.0 scheme [36].

Though relatively few schemes attempt to unify these approaches, ours is not the first. In particular, the ISO 24617-2 standard includes dialog acts as well as an additional dimension for rhetorical relations, most commonly annotated with the DR-CORE<sup>3</sup> relation set [15, 18]. While one could annotate multi-party dialogue with the ISO standard by using the finer-grain PDTB relations in the rhetorical dimension, we found this to not fit our distinct approach to the structures described previously, which departs significantly from the ISO annotation guidelines.

**Dialog Act Set:** DDA’s dialog act set covers 40 out of the 42 most frequently used Dialog Act (DA) classes from the SWBD-DAMSL scheme. 26 out of the 40 classes are kept with the original definition and class name, while the remaining 15 are collapsed into coarser classes. This leads us to 31 DA classes. The most noticeable merger of SWBD-DAMSL DA classes is from the “question” and “answer” DAs. We replaced 5 classes of “answer” type from SWBD-DAMSL with a single “answer” tag and 8 “question” DAs from SWBD-DAMSL with 3 coarser “question” classes. This is because most of the sub-type “question” and “answer” tags can be resolved from the lexical level analysis. Additionally, we add “joke” as a new DA to cover the social acts in our domain of interest. In this regard, the taxonomy of DDA’s dialogue acts

<sup>3</sup> While corpora annotated with ISO typically use the DR-CORE rhetorical relations, the guideline itself does not actually specify this, and any set which relates dialog unit pairs may be used.

labels still follow SWBD-DAMSL’s hierarchy with 6 top-level categories, bolded below. We list DDA’s dialogue acts set with this hierarchy as follows:

- **Statements:** Statement, Opinion
- **Communicative Status:** Self-talk, Abandoned
- **Backward-Communicative Functions:** Answer, Stalling, Accept, Reject, Collaborative Completion, Appreciation, Downplayer, Sympathy, Acknowledge, signal-non-understanding
- **Forward-Communicative Function:** Task-Management, Offer, Action-Directive, Commit, Question/Info-request, Open-Question, Rhetorical-Question, Apology, Thanking, Exclamation, Explicit-performative, Welcome
- **Information Level:** Greeting, Correction, Conventional-closing
- **Other:** Hedge, Joke

**Rhetorical Relation Set:** DDA uses discourse relations from PDTB expanded with some extra relations. Aside from dialog acts, discourse relations are very useful for describing speaker intentions, especially speech organizational intentions that dialog acts do not cover. We use the discourse relations set from PDTB 3.0, but extended it with some finer-grained relations for the “Contingency” and “Expansion” types. For the “Contingency” class, we add 4 more asymmetric sub-types for Cause (“Justify”, “Motivation”, “Enablement” and “Evaluation”). Similarly, we extended the “Expansion” class with 3 more relations (“Process-step”, “Object-attribute” and “List”), which are inspired by Amanda Stent’s work on RST in Dialog [40]. We added these relations because we found them to be useful distinctions in our conversational datasets, in which students discuss, collaborate and negotiate with each other. DDA still leverages the benefits of PDTB’s taxonomy hierarchy with this extension.

For comparison, we list the discourse relations adapted in DDA as well as discourse relations from other common frameworks in Table 2.

DDA edges always point backward in the conversation (from a slash unit to another slash unit that it is responding to) or are self-edges. In order to support this directionality without losing expressive power, we make use of the dual discourse relations introduced in PDTB 3.0, such that any asymmetric relationship is annotated from the context of a reply without changing the meaning of the response dependency structure. For instance, if a unit  $A$  is a “Reason” for a future unit  $B$ , this  $A \rightarrow B$  can be equivalently annotated as  $A \leftarrow B$  such that  $B$  is a “Result” of  $A$ . As for asymmetric relations that can be verbified, DDA uses the active or passive voice of the verb to encode the directionality. For example,  $C \xrightarrow{\text{Enabling}} D$  is equivalent to  $C \xleftarrow{\text{Enabled}} D$ . This can be read naturally with English:  $C \xrightarrow{\text{Enabling}} D$  is read “C is enabling D” and  $C \xleftarrow{\text{Enabled}} D$  is read “D is enabled by C”.

**Overloading and Multi-functionality:** As part of DDA’s recall-oriented annotation philosophy, we embrace multi-edges to encode the overloading of responsiveness. For example, in Fig. 2, utterance (2) can be considered a “reply” given the only context utterance in this example is a question from utterance (1), such that the reply-to edge is also a response dependency edge. As the conversation proceeds,



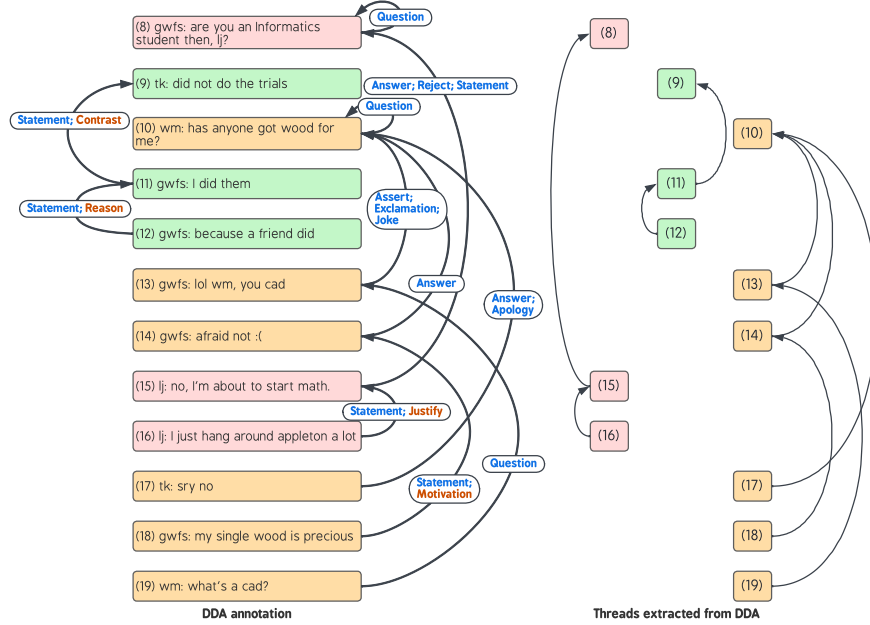
Category	ISO-DR-core	PDTB3.0	SDRT	DDA
Temporal	Async, Sync	Async, Precedence, Succession	Sync, Narration, Precondition	Async, Sync, Before, After
Contingency	Cause, Condition, Neg-Condition, Purpose	Cause, Cause+Belief, Cause+SA, Condition, Neg-Condition, Purpose, Reason, Result	Explanation, Result, Consequence	Cause, Justify, Motivation, Condition, Neg-Condition, Purpose, Enablement, Reason, Result, Evaluation
Comparison	Contrast, Similarity, Concession	Contrast, Similarity, Concession, Concession+SA	Consequence, Explanation, Contrast, Parallel	Contrast, Similarity, Concession
Expansion	Exception, Conjunction, Disjunction, Substitution, Manner, Elaboration, Restatement, Expansion, Exemplification	Instantiation, Level-of-details, Substitution, Equivalence, Disjunction, Exception, Conjunction, Manner	Continuation, Alternation, Elaboration, Background, Commentary, Attribution, Source	Expansion, Instantiation, Level-of-details, Substitution, Restatement, Summary, Disjunction, Exception, Conjunction, Manner, Process-step, Object-attribute

**Table 2** A comparison of discourse relations across frameworks. We choose to replace some of PDTB’s relation name from PDTB2.0 for ease of memory such as “Precedence”=“Before”, “Succession”=“After”, “equivalence”=“restatement”, assigning them identical definition.

the intention rendered in utterance (4) shifted away from being an “Answer” to utterance (1) (the marginal information gain of yet another same answer to the previous question diminishes) and therefore serves more toward rhetorical functions instead of communicative ones.

## 4 Applications

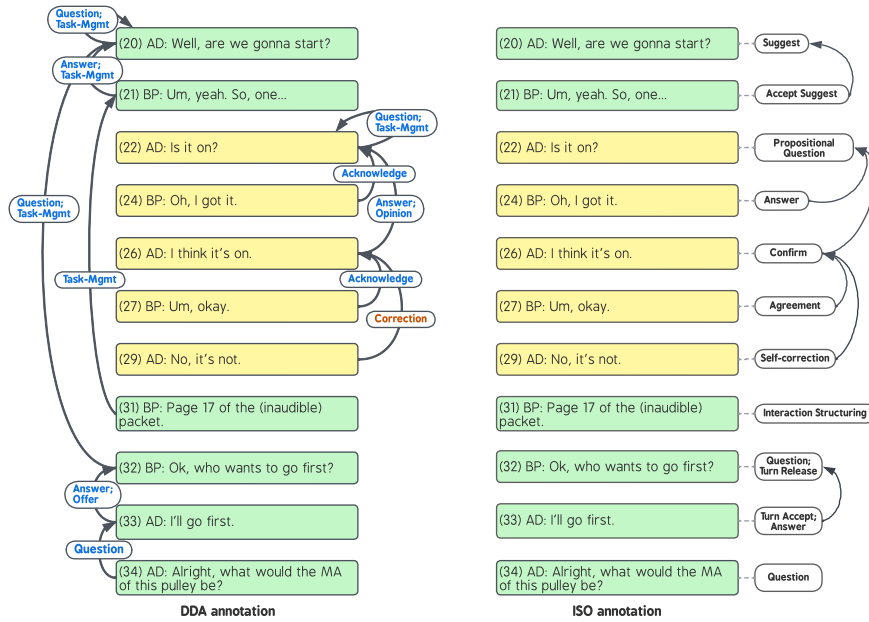
**Conversation Threads:** Despite the fact that dialogue annotations in DDA might include more structural links than in reply structure graphs, they share the same useful property in which separable conversation threads form connected components in the resulting graph. Given a complete dialogue annotation, this allows a simple method for disentangling threads, a processing step that has been shown to improve dialogue understanding methods [23] and is of analytical interest in our classroom multi-party dialogue setting. See Fig. 4 for an example. The dependency chains (8)-(15)-(16),



**Fig. 4** DDA annotation on an example in the STAC corpus [7] showing that DDA connects conversation threads into connected components; On the left, we have DDA annotation for the conversation snippet; On the right, we highlight the extracted threads.

(9)-(11)-(12) and (10)-(13)-(14)-(17)-(18)-(19) can be naturally derived from DDA's response dependency structure. Similarly, in Fig. 5, we show a classroom example, in which the threads of conversation among students are naturally disentangled by following the response dependencies.

**Response Dependencies for Discourse Analysis:** The theoretical benefits of DDA's response dependency structure go beyond the threads disentanglement and annotation simplification. DDA can be potentially used as an analytical tool to identify interpersonal relationships and power dynamics. For example, if DDA dependencies show significantly more connections between certain participants, it may indicate they are having more engaged conversations and forming bonding. Further, if the topological structure of the DDA for a conversation shows balanced connectivity between speakers, it could indicate the power is evenly distributed. Alternatively, if the dependencies are mostly pointing at a single or a few people, it's more likely that they are leading the conversation. We aim to explore these analyses in future work. In Table 1, we compare features among different annotation schemes.

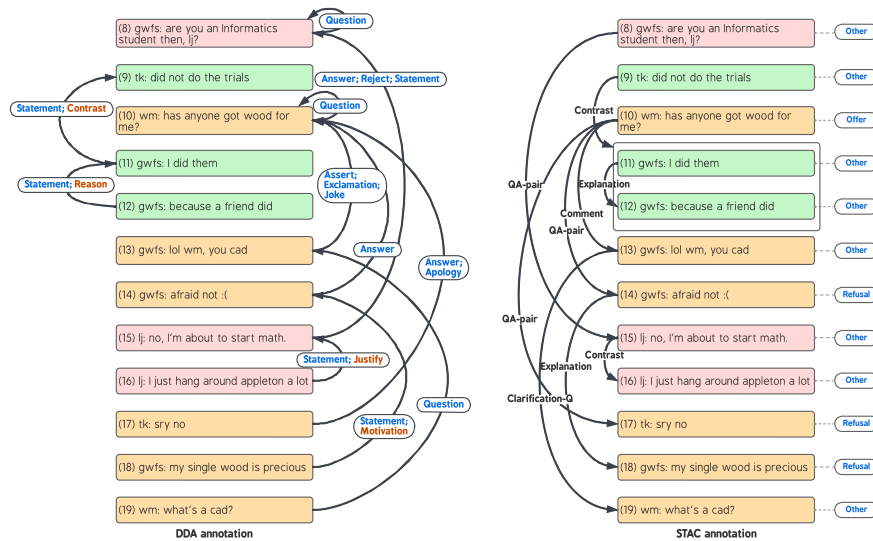


**Fig. 5** A comparison of DDA annotation and corresponding ISO annotation for an example from our classroom setting. Students participate in two interleaved dialog threads: preparing to discuss solutions to a worksheet on pulleys (green) and ensuring their equipment is on correctly (yellow). Response dependencies in DDA connect thread slash units by construction. Functional and similar dependencies in the ISO annotation yield only partial connectivity.

## 5 Related Work

**Dialogue Acts:** There is a long history of analyzing the “actions” of utterances, known as dialog acts [44, 9, 39, 3, 30]. Dialog act annotation schemes developed include DAMSL (Dialogue Act Markup in Several Layers) [2, 21], SWBD-DAMSL [29, 25], DIT (Dynamic Interpretation Theory) [11], and DIT++ schema [12, 13]. The ISO 24617-2 standard proposed a semantically-based standard for dialogue annotation, and includes both dialogue acts and the relations between discourse units [14, 15, 16]. Researchers have long noted that multi-functionality (pragmatic overloading) is hard to capture with a single utterance purpose, especially in multi-party multi-threaded dialogues [4, 20, 26, 22].

In our work, we follow SWBD-DAMSL’s approach by augmenting its flattened DA tag set. Nevertheless, we made two augmentations: first, DDA handles multi-functionality phenomenon with multi-label and multi-dependency; second, DDA resolves the response structure, which not only unveils a deeper discourse structure in conversations but also anchors the dialogue act and discourse relations into context, which is fundamentally different than tagging schemes. In limited experiments, we



**Fig. 6** A comparison between DDA annotation (left) and SDRT as adapted for the STAC corpus (right), on an example from the STAC corpus [7]. The dialog consists of three distinct threads, denoted by color. Due to the high degree of rhetorical organization in this dialog, both annotations form distinct connected threads. While the relation set in the STAC corpora is tailored for strategic dialog, DDA is able to capture more general phenomena with greater specificity, such as the labeling of apologies within answers or of jokes as opposed to comments.

find the efficiency of annotating DDA to be comparable to SWBD-DAMSL.

**Rhetorical Relations in Multi-party Dialogue:** Previous work on structured analysis for multi-party dialogue mainly focused on simple thread disentanglement, rather than analyzing the resulting rhetorical structures [31, 24, 43, 42]. In the current work, we mainly focus on the rhetorical structures in multi-party dialogue. Four of the most influential frameworks have been used in dialogue analysis: Rhetorical Structure Theory (RST) [33], Segmented Discourse Representation Theory (SDRT) [8], Hobbs' theory of discourse [27], and Penn Discourse Treebank (PDTB) framework [35, 36]. In RST, an RST tree is built recursively by connecting the adjacent discourse units, forming a hierarchical structure covering the whole text. RST Bank [19] created a reference corpus for community-wide use, while [40] provides a practical analysis on annotating dialogue with RST. Similar to RST, SDRT also provides a hierarchical structure of text organization with full annotation. For example, DISCOR corpus [38], the ANNODIS corpus [1], and the STAC [7] use directed acyclic graphs that allow for multiple parents, but not for crossing. Based on Hobbs' theory, Discourse Graphbank [45] allows for general graphs that allow multiple parents and crossing. Unlike the above frameworks, PDTB adopts a theory-neutral approach to the annotation, which does not aim at achieving complete annotation of the text but focuses on local discourse relations anchored by structural connectives or discourse

adverbials. This theory neutrality makes no commitments to what kinds of high-level structures may be created from the low-level annotations of relations and their arguments, thus it permits more freedom of investigating complex dependency structures in multi-party multi-threaded dialogue. Furthermore, ISO DR-Core also follows the theory-neutral stance in PDTB, annotating only high-level, coarse-grained discourse relations that can then be annotated further to capture a finer-grained tree or graph structure, depending on one's theoretical preferences.

DDA follows PDTB's discourse relation taxonomy since it has been demonstrated effective in annotation practice to yield good annotator agreement, however, we augmented it with dense response structure annotation instead of partial annotation (as shown in Fig. 1).

**Annotating Response Structure Graphs:** Another line of work aims to improve conversational understanding systems by uncovering the response dependencies between utterances in multi-party speech [6] or online chat [24, 31, 23]. Given dialogue segmented into utterances, the task is to connect an utterance of interest with all previous utterances to which it responds. The resulting connected components form dialogue threads that can be understood individually in downstream systems. [31] present one of the largest available corpora annotated with reply structure graphs, consisting of 77,653 messages from the Ubuntu Internet Relay Chat (IRC). Our notion of response dependency is similar to this line of work with three key differences: (1) All of our dependencies are labeled with the dialogue act and/or rhetorical relation initiated in the responding utterance. (2) All non-self DDA edges point towards previous utterances, but with the duality coding mentioned in Section 3.2, the edges still denote the semantic roles of each utterance. (3) An utterance of interest can respond to any number of previous utterances using any number of labels.

## 6 Future Work

In the future, we plan to apply DDA to annotate multiparty conversations, including conversations from K-12 classrooms, where students form small groups to solve problems collaboratively.

**Limitations:** DDA, like all other discourse-level annotation schemes, has its own limitations in terms of scope, generalizability and domain-specific bias. First, DDA assumes sufficient information exists for annotation in the conversation records. If certain references in the context is meant to be resolved from non-verbal communication channels, such as pointing and gestures, DDA may need situated transcripts of the conversation to be properly deployed. Besides, DDA inherits the limitation of the expressive power from PDTB and SWBD-DAMSL and benefits from their scalability in practice. Second, the addressee information is not guaranteed to be reflected in DDA solely.

## References

1. Afantenos, S., Asher, N., Benamara, F., Bras, M., Fabre, C., Ho-Dac, L.M., Le Draoulec, A., Muller, P., Péry-Woodley, M.P., Prévot, L., et al.: An empirical resource for discovering cognitive principles of discourse organisation: the annodis corpus. In: Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12). European Language Resources Association (ELRA) (2012)
2. Allen, J., Core, M.: Draft of damsl: Dialog act markup in several layers (1997)
3. Allen, J.F., Perrault, C.R.: Analyzing intention in utterances. *Artif. Intell.* **15**, 143–178 (1980)
4. Allwood, J.: An activity based approach to pragmatics (1995)
5. Anscombe, G.E.M., Teichman, J., Diamond, C.: Intention and Intentionality Essays for G. E. M. Anscombe (1979)
6. Aoki, P.M., Szymanski, M.H., Plurkowski, L., Thornton, J.D., Woodruff, A., Yi, W.: Where's the "Party" in "Multi-Party"? Analyzing the Structure of Small-Group Sociable Talk p. 10 (2006)
7. Asher, N., Hunter, J., Morey, M., Farah, B., Afantenos, S.: Discourse structure and dialogue acts in multiparty dialogue: the STAC corpus. In: Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), pp. 2721–2727. European Language Resources Association (ELRA), Portorož, Slovenia (2016). URL <https://aclanthology.org/L16-1432>
8. Asher, N., Lascarides, A.: Logics of conversation. Cambridge University Press (2003)
9. Austin, J.L.: How to do things with words. Oxford university press (1975)
10. Balakrishnan, A., Rao, J., Upasani, K., White, M., Subba, R.: Constrained decoding for neural NLG from compositional representations in task-oriented dialogue. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 831–844. Association for Computational Linguistics, Florence, Italy (2019). DOI 10.18653/v1/P19-1080. URL <https://aclanthology.org/P19-1080>
11. Bunt, H.: Dynamic interpretation and dialogue theory. The structure of multimodal dialogue **2**, 139–166 (1999)
12. Bunt, H.: Dimensions in dialogue act annotation. In: Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06) (2006)
13. Bunt, H.: The dit++ taxonomy for functional dialogue markup. In: AAMAS 2009 Workshop, Towards a Standard Markup Language for Embodied Dialogue Acts, pp. 13–24 (2009)
14. Bunt, H., Alexandersson, J., Carletta, J., Choe, J.W., Fang, A.C., Hasida, K., Lee, K., Petukhova, V., Popescu-Belis, A., Romary, L., Soria, C., Traum, D.: Towards an ISO standard for dialogue act annotation. In: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10). European Language Resources Association (ELRA), Valletta, Malta (2010). URL [http://www.lrec-conf.org/proceedings/lrec2010/pdf/560\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2010/pdf/560_Paper.pdf)
15. Bunt, H., Alexandersson, J., Choe, J.W., Fang, A.C., Hasida, K., Petukhova, V., Popescu-Belis, A., Traum, D.: ISO 24617-2: A semantically-based standard for dialogue annotation. In: Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12), pp. 430–437. European Language Resources Association (ELRA), Istanbul, Turkey (2012). URL [http://www.lrec-conf.org/proceedings/lrec2012/pdf/530\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2012/pdf/530_Paper.pdf)
16. Bunt, H., Petukhova, V., Gilmartin, E., Pelachaud, C., Fang, A., Keizer, S., Prévot, L.: The ISO standard for dialogue act annotation, second edition. In: Proceedings of the Twelfth Language Resources and Evaluation Conference, pp. 549–558. European Language Resources Association, Marseille, France (2020). URL <https://aclanthology.org/2020.lrec-1.69>
17. Bunt, H., Petukhova, V., Malchanau, A., Wijnhoven, K., Fang, A.: The DialogBank. In: Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), pp. 3151–3158. European Language Resources Association (ELRA), Portorož, Slovenia (2016). URL <https://aclanthology.org/L16-1503>

18. Bunt, H., Prasad, R.: Iso dr-core (iso 24617-8): Core concepts for the annotation of discourse relations. In: ACL 2016 (2016)
19. Carlson, L., Marcu, D., Okurowski, M.E.: Building a discourse-tagged corpus in the framework of rhetorical structure theory. In: Current and new directions in discourse and dialogue, pp. 85–112. Springer (2003)
20. Cohen, P.R., Morgan, J.L., Pollack, M.E.: Intentions in communication. MIT press (1990)
21. Core, M.G., Allen, J.: Coding dialogs with the damsl annotation scheme. In: AAAI fall symposium on communicative action in humans and machines, vol. 56, pp. 28–35. Boston, MA (1997)
22. Di Eugenio, B., Webber, B.L.: Pragmatic overloading in natural language instructions. International Journal of Expert Systems Research and Applications **9**, 53–84 (1996)
23. Du, W., Poupart, P., Xu, W.: Discovering Conversational Dependencies between Messages in Dialogs (2016). URL <http://arxiv.org/abs/1612.02801>. ArXiv:1612.02801 [cs]
24. Elsner, M., Charniak, E.: You Talking to Me? A Corpus and Algorithm for Conversation Disentanglement. In: Proceedings of ACL-08: HLT, pp. 834–842. Association for Computational Linguistics, Columbus, Ohio (2008). URL <https://aclanthology.org/P08-1095>
25. Godfrey, J.J., Holliman, E.C., McDaniel, J.: Switchboard: Telephone speech corpus for research and development. In: Acoustics, Speech, and Signal Processing, IEEE International Conference on, vol. 1, pp. 517–520. IEEE Computer Society (1992)
26. Hancher, M.: The classification of cooperative illocutionary acts I. Language in society **8**(1), 1–14 (1979)
27. Hobbs, J.R.: Literature and cognition. 21. Center for the Study of Language (CSLI) (1990)
28. Jeffrey, R.C.: The Logic of Decision. New York, NY, USA: University of Chicago Press (1965)
29. Jurafsky, D.: Switchboard swbd-damsl shallow-discourse-function annotation coders manual. Institute of Cognitive Science Technical Report (1997)
30. Kautz, H.A.: A formal theory of plan recognition (1987)
31. Kummerfeld, J.K., Gouravajhala, S.R., Peper, J.J., Athreya, V., Gunasekara, C., Ganhotra, J., Patel, S.S., Polymenakos, L.C., Lasecki, W.: A Large-Scale Corpus for Conversation Disentanglement. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 3846–3856. Association for Computational Linguistics, Florence, Italy (2019). DOI 10.18653/v1/P19-1374. URL <https://aclanthology.org/P19-1374>
32. Li, X., Stevens-Guille, S., Maskharashvili, A., White, M.: Self-training for compositional neural NLG in task-oriented dialogue. In: Proceedings of the 14th International Conference on Natural Language Generation, pp. 87–102. Association for Computational Linguistics, Aberdeen, Scotland, UK (2021). URL <https://aclanthology.org/2021.inlg-1.10>
33. Mann, W.C., Thompson, S.A.: Rhetorical structure theory: Toward a functional theory of text organization. Text-interdisciplinary Journal for the Study of Discourse **8**(3), 243–281 (1988)
34. Mann, W.C., Thompson, S.A.: Rhetorical structure theory: Toward a functional theory of text organization. Text & Talk **8**, 243 – 281 (1988)
35. Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., Webber, B.: The Penn Discourse TreeBank 2.0. In: Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC’08), European Language Resources Association (ELRA), Marrakech, Morocco (2008). URL [http://www.lrec-conf.org/proceedings/lrec2008/pdf/754\\_paper.pdf](http://www.lrec-conf.org/proceedings/lrec2008/pdf/754_paper.pdf)
36. Prasad, R., Webber, B., Lee, A., Joshi, A.: Penn Discourse Treebank Version 3.0 (2019). DOI 11272.1/AB2/SUU9CB. URL <https://hdl.handle.net/11272.1/AB2/SUU9CB>
37. Reed, L., Oraby, S., Walker, M.: Can neural generators for dialogue learn sentence planning and discourse structuring? In: Proceedings of the 11th International Conference on Natural Language Generation, pp. 284–295. Association for Computational Linguistics, Tilburg University, The Netherlands (2018). DOI 10.18653/v1/W18-6535. URL <https://aclanthology.org/W18-6535>
38. Reese, B., Hunter, J., Asher, N., Denis, P., Baldridge, J.: Reference manual for the analysis and annotation of rhetorical structure. Ph.D. thesis, University of Texas at Austin (2007)

39. Searle, J.R., Searle, J.R.: *Speech acts: An essay in the philosophy of language*, vol. 626. Cambridge university press (1969)
40. Stent, A.: Rhetorical structure in dialog. In: INLG'2000 Proceedings of the First International Conference on Natural Language Generation, pp. 247–252. Association for Computational Linguistics, Mitzpe Ramon, Israel (2000). DOI 10.3115/1118253.1118288. URL <https://aclanthology.org/W00-1433>
41. Vermazen, B., Hintikka, M.: *Essays on Davidson: Actions and Events*. Clarendon Press (1985). URL [https://books.google.com/books?id=1R\\\_1AAAAIAAJ](https://books.google.com/books?id=1R\_1AAAAIAAJ)
42. Wang, H., Wang, C., Zhai, C., Han, J.: Learning online discussion structures by conditional random fields. In: Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, pp. 435–444 (2011)
43. Wang, Y.C., Rose, C.: Making conversational structure explicit: identification of initiation-response pairs within online discussions. In: *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 673–676 (2010)
44. Wittgenstein, L.: *Philosophical investigations*. John Wiley & Sons (2010)
45. Wolf, F., Gibson, E.: Representing discourse coherence: A corpus-based study. *Computational linguistics* **31**(2), 249–287 (2005)