The DialogBank

Harry Bunt*, Volha Petukhova, Andrei Malchanau**, Alex Fang*** and Kars Wijnhoven***
*Tilburg Center for Cognition and Communication, Tilburg University, The Netherlands
**Spoken Language Systems, Saarland University, Germany
***Dialogue Systems Group, City University of Hong Kong, Hong Kong SAR
harry.bunt,kars.wijnhoven@uvt.nl; v.petukhova, a.malchanau@lsv.uni-saarland.de, alex.fang@cityu.edu.hk

Abstract

This paper presents the DialogBank, a new language resource consisting of dialogues with gold standard annotations according to the ISO 24617-2 standard. Some of these dialogues have been taken from existing corpora and have been re-annotated according to the ISO standard; others have been annotated directly according to the standard. The ISO 24617-2 annotations have been designed according to the ISO principles for semantic annotation, as formulated in ISO 24617-6. The DialogBank makes use of three alternative representation formats, which are shown to be interoperable.

Keywords: dialogue annotation, ISO standard, interoperability

1. Introduction

This DialogBank is a new language resource, developed at Tilburg University, which contains dialogues of various kind with gold standard dialogue act annotations according to the ISO 24617-2 standard. This standard builds on previously designed annotation schemes such as DAMSL, DIT^{++}, MRDA, HCRC Map Task, Verbmobil, SWBD-DAMSL, and DIT.[1] Most of these schemes have been used to construct annotated corpora, such as the Switchboard, HCRC Map Task, ICSI-MRDA, and DIAMOND corpora.
For nearly all of these annotation schemes, dialogue act annotation consists of segmenting a dialogue into certain grammatical units and marking up each unit with one or more communicative function labels. ISO 24617-2 (like DIT^{++}) supports semantically more complete annotation by additionally annotating the following aspects:

1. 'Dimension', or category of semantic content: the annotation scheme supports multidimensional annotation, i.e. multiple communicative functions may be assigned to dialogue segments; different from DAMSL and other multidimensional schemes, an explicitly defined notion of 'dimension' is used that corresponds to a certain category of semantic content. The ISO scheme distinguishes nine dimensions on empirical and theoretical grounds.

   'Multidimensional segmentation' is used: a dialogue is segmented in multiple ways, with functional segments in each relevant dimension. 'Functional segments', defined as minimal stretches of behaviour that have a communicative function, are the units that dialogue act annotations attach to. Functional segments are mostly shorter than turns, may be discontinuous, may overlap, and may contain parts contributed by different speakers. A segment carrying a feedback function may for instance overlap with a segment that carries a task-related function.

2. 'Qualifiers' may be added for expressing that a dialogue act is performed conditionally, with uncertainty, or with a particular sentiment.

3. Dependence relations are defined for expressing semantic relations between dialogue acts, e.g. for indicating which question is answered by a certain answer act, or which utterance a feedback act. responds to.

4. Rhetorical relations may be annotated to indicate e.g. that one dialogue act explains the performance of another dialogue act.

Most of the dialogues in the DialogBank have been taken from existing corpora and have been re-segmented and re-annotated; some of these also have their original annotations for comparison. Some of the dialogues are (also) annotated according to the DIT^{++} annotation scheme, which has been a major source of inspiration for the ISO 24617-2 scheme. The DialogBank presently contains (re-)annotated dialogues from four English-language corpora: HCRC Map Task, Switchboard, TRAINS (Allen et al., 1994) and DBOX (Petukhova et al., 2014); and from four Dutch-language corpora: DIAMOND (Geertzen et al., 2004), Schiphol (Prüst et al., 1984), OVIS (`www.let.rug.nl/vannoord/Ovis`), and Dutch Map Task (`http://doc.ukdataservice.ac.uk/doc/4632/mrdoc/pdf/4632userguide.pdf`). Dialogues from other corpora, such as the multi-party AMI corpus are planned to be added in the near future.

2. Interoperable Annotation

2.1. Annotations and Their Representation

The main motivation for designing annotation standards is to promote the interoperability of annotated corpora. Interoperability of annotations is partly a matter of interchangeable representation formats, such as XML, but more importantly of the underlying concepts. Different annotations can be interpreted across platforms and frameworks only if they encode the same information, or information that can be interpreted through a well-defined mapping. Interoperability at conceptual and semantic levels is of more fundamental importance than interoperability at the level of representation formats. In the design of ISO 24617-2, the focus is

[1] See Allen & Core (1997); Bunt (2007); Shriberg et al. (2004); Anderson et al. (1991); Alexandersson et al. (1998); Jurafsky et al. (1997); and Bunt (1994; 2000), respectively.

therefore on the identification of empirically and theoretically well-motivated concepts and precise definitions.

ISO 24617-2 includes a comprehensive, application-independent annotation scheme with well-defined concepts and the markup language DiAML (Dialogue Act Markup Language), designed in accordance with the ISO Linguistic Annotation Framework (LAF)[2] and the ISO Principles of Semantic Annotation ('SemAF Principles').[3] LAF makes a fundamental distinction between *annotation* and *representation*: 'annotation' refers to the linguistic information that is added to segments of language data, independent of format; 'representation' refers to the format in which an annotation is rendered. Following SemAF Principles, this distinction is implemented in the DiAML definition in the form of an *abstract syntax* that specifies a class of abstract *annotation structures*, which are set-theoretical constructs like pairs and triples, and a *concrete syntax* that specifies a rendering of these annotation structures in a reference format using XML. This reference format is called DiAML-XML. It uses abbreviated XML-expressions, is complete and unambiguous relative to the abstract syntax: (1) the concrete syntax defines a representation for every structure defined by the abstract syntax; and (2) every expression defined by the concrete syntax represents one and only one structure defined by the abstract syntax. A representation format with these properties is called *ideal*. Any ideal representation format can be converted through a meaning-preserving mapping to any other ideal representation format (see Bunt, 2010 for formal definitions and proofs).

The dialogues in the DialogBank have all been (re-)annotated using the DIAML markup language; some of them are represented in the DiAML-XML format; others are cast in one of two alternative tabular representation formats, defined in such a way that they are demonstrably ideal (complete and unambiguous) and that they are more convenient for human readers than DiAML-XML representations.

2.2. DiAML Representations in XML

The HRCR Map Task dialogues in the DialogBank are re-segmented and re-annnotated using the ANVIL tool (Kipp, 2001), starting from raw speech. These annotations are represented in DiAML-XML, which makes use of two XML elements, one to represent dialogue acts and one to represent a semantic or pragmatic ('rhetorical') relation between dialogue acts. A <dialogueAct> element has attributes whose values represent (1) the speaker; (2) the addressee(s); (3) (optionally) possible other participants; (4) the communicative function; (5) the dimension; (6) qualifiers (if any); and (7) dependence relations. Example (1b) shows the representation of the annotation of the dialogue fragment in (1a), which contains a rhetorical relation (Elaboration) between the dialogue acts in 1 and 3, and a feedback dependence between the dialogue acts in 3 and 4.

(1) a. 1. G: go south and you'll pass some cliffs on your right
2. F: uhm...
3. G: and some adobe huts on your left
4. F: oh okay

b.
```
<diaml xmlns="http://www.iso.org/diaml">
<dialogueAct xml:id="da1" target="#fs1" sender="#g"
  addressee="#f" dimension="task"
  communicativeFunction="instruct" />
<dialogueAct xml:id="da2" target="#fs2" sender="#f"
  addressee="#f" dimension="turnManagement" />
  communicativeFunction="turnTake" />
<dialogueAct xml:id="da3" target="#fs2" sender="#f"
  addressee="#g" dimension="timeManagement"
  communicativeFunction="stalling" />
<dialogueAct xml:id="da4" target="#fs3" sender="#g"
  addressee="#f" dimension="task"
  communicativeFunction="inform" />
<rhetoricalLink dact="#da4" rhetoAntecedent="#da1"
  rhetoRel="elaborate" />
<dialogueAct xml:id="da5" target="#fs4" sender="#f"
  addressee="#g" dimension="autoFeedback"
  communicativeFunction="autoPositive"
  feedbackDependence="#da1" "#da4" />
< /diaml>
```

2.3. Representation in a Tabular Format

The re-annotation of the dialogues that were included in the DialogBank started in some cases from raw primary data or transcriptions, and in some cases from previous annotations in various formats, which mostly used the form of a table in which rows correspond to the segmentation; one column contains the transcribed speech; and the other column(s) contain the annotation. This is illustrated in Figure 1 for a fragment of a Switchboard dialogue, originally annotated according to the SWBD-DAMSL scheme and represented in a 3-column format, and in Figure 2 for the multidimensional annotation of a short dialogue fragment from the TRAINS corpus, using the DIT annotation scheme and the DitAT annotation tool (Geertzen, 20008).

The formats used in Figures 1 and 2 look rather different, and even more different from the XML format used in (1), yet they all contain essentially the same information. For example, the numbers in the first column in Fig. 1 as well as in Fig. 2 can be interpreted as identifiers of functional segments; the strings in the third column in both cases as containing the transcriptions of these segments; the second column in Fig. 2 as indicating the speaker of a segment (and, by implication for a two-person dialogue, the addressee), which in Fig. 1 is part of the information in the third column; and the other columns as representing the dialogue act annotations. The row numbered 1 in Fig. 2 thus corresponds to the following XML expression:

(2)
```
<dialogueAct xml:id="da1" target="#fs1" sender="#s"
  addressee="#u" dimension="turnManagement"
  communicativeFunction="turnTake" />
<dialogueAct xml:id="da2" target="#fs1" sender="#s"
  addressee="#u" dimension="contactManagement"
  communicativeFunction="contactIndication" />
<dialogueAct xml:id="da3" target="#fs1" sender="#s"
```

[2] ISO 24612:2012; see also Ide & Romary (2004).

[3] ISO 24617-6; see also Bunt (2015).

id	function	transcript
sw01-0105-0001-A001-01	qw	A.1 utt1: Jimmy, {D so } how do you get most of your news? /
sw01-0105-0002-B002-01	sd	B.1 utt1: {D Well, [I kind of, + {F uh, } I] watch the {F uh, } national news every day, for one /
sw01-0105-0003-B002-02	sd	B.2 utt1 I also read one or two papers a day /
sw01-0105-0004-B002-03	sd	B.3 utt1: {C and } [I'm a, + I'm pretty much a] news junkie /
sw01-0105-0005-B002-04	sd	B.4 utt1: {C and } I tune in to CNN a lot. /
sw01-0105-0006-A003-01	ba	A.3 utt1: {F Oh, } wow. /

Figure 1: Annotation in tabular form of Switchboard (SWBD-DA) dialogue fragment

id	sp	transcript	Task	Auto-Feedback	Allo-Feedb.	Turn Man.	Time Man.	Discourse Structuring	Contact Man.	OC M	PC M	SOM
1	s	hello							Contact indication			Initial greeting
2		can I help you						Offer				
3	u	uhm,				Turn Take	Stalling					
4	u	yes hello, maybe		Evaluation positive				Accept Offer				
5		I'd like to take a tanker ..	Inform					Topic introduct.				

Figure 2: Representation in tabular form of DIT annotations produced with the DitAT tool

addressee="#u" dimension="socialObligationsManagement" communicativeFunction="initGreeting" />

The two tabular formats shown here are less expressive than the DiAML-XML format in that, firstly, the information assigned to dialogue segments is limited to communicative functions only (Fig. 1) or to communicative functions and dimensions (Fig. 2); and secondly only contiguous, non-overlapping dialogue segments can be handled. The former limitation can be overcome by extending the information about a dialogue act in a cell of the table by adding qualifiers, dependences, and rhetorical relations. To overcome the latter limitation, and make the tabular representations compatible with the stand-off requirement of the ISO LAF, we will describe some further adjustments in Section 3. The resulting adaptations of the formats illustrated in Figures 1 and 2 are called DiAML-TabSW and DiAML-MultiTab, respectively, and will be shown to be ideal – complete and unambiguous.

3. Interoperability of Representation Formats

3.1. Abstract Syntax and Alternative Representations

The introduction in the ISO standard of an abstract syntax, besides a concrete representation format, was to allow precise determination of the interoperability of alternative representations. Figure 3 displays the relations between an abstract syntax, one or more alternative ideal representation formats, and the semantics of a markup language according to SemAF-Principles.

Since the DiAML-XML format is defined in such a way that it is both complete and unambiguous, a function F_{XML} can be defined that maps annotation structures as defined by the abstract syntax to an XML expression; due to the unambiguity, this function has an inverse F_{XML}^{-1} which maps any DiAML-XML expression to the annotation structure that it represents.

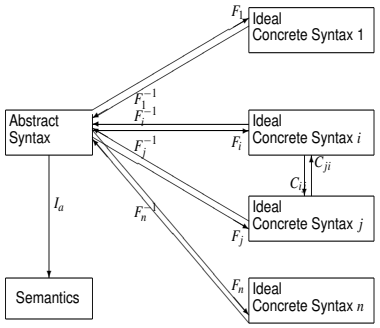

Figure 3: Abstract and concrete syntax, and semantics

A tabular representation can be formally defined as a matrix of which each row is an n-tuple of elements corresponding to the contents of its cells. Using this formalization, the functions $F_{MultiTab}$ and F_{TabSW} are specified below, proving that the modified tabular formats MultiTab and TabSW are complete. Similarly, defining their inverses $F_{MultiTab}^{-1}$ and F_{TabSW}^{-1}, shows their unambiguity. As a result, the composition of functions such as

(3) $C_{MultiTab \rightarrow XML} = F_{XML}\ o\ F_{MultiTab}^{-1}$

defines a conversion from annotations, represented in the DiAML-MultiTab format to representations in the DiAML-XML format. The inter-convertability of the three formats will be exploited in the DialogBank by allowing the user to view the annotations in the form that is most convenient to them, as well as by converting the tabular formats to the XML format for automatic processing.

3.2. Abstract Syntax

The abstract syntax of DiAML reflects the conceptual analysis of dialogue acts according to Dynamic Interpretation Theory (DIT), which underlies the DIT^{++} and ISO 24617-2 annotation schemes. On this analysis, a dialogue act is characterized by seven elements: (1) a sender; (2) one or more addressees; (3) zero or more other participants, such as an audience; (4) a communicative function; (5) a dimension; (6) zero or more qualifiers; (7) zero or more dependence relations.

Two types of dependence relations are distinguished. A 'functional' dependence occurs when a dialogue act is semantically dependent on one or more previous dialogue acts due to having a communicative function of a responsive character. This is for example the case for answers, whose meaning is partly determined by the question which is being answered, but also for the acceptance or rejection of offers, suggestions, requests, and the acceptance of apologies and thankings. The subset *RSP* of the set of communicative functions, defined as part of the conceptual vocabulary, contains these communicative functions.

Feedback acts provide or elicit information about the processing of something that was said earlier in the dialogue, such as its perception or its interpretation, and their meaning often depends on that. Positive feedback utterances like *"OK"* and *"Yes"*, and negative ones like *"What?"* and *"Excuse me?"* illustrate this phenomenon. While positive feedback acts are typically about the processing of previous dialogue acts, negative feedback acts are often about a problem in understanding something, and may thus refer to a segment of speech rather than to its interpretation as a dialogue act. ISO 24617-2 therefore allows feedback dependence relations to have both dialogue acts and markables as antecedent arguments.

Note that a dialogue act can either have a functional dependence relation or a feedback dependence relation to one or more other dialogue acts (or markables), but not both. Only dialogue acts in one of the two feedback dimensions can have a feedback dependence relation, and only dialogue acts with a responsive communicative function ($f \in RSP$) have a functional dependence relation; such a function can be either a general-purpose one, like Answer, Agreement, or Correction, or a dimension-specific one, such as Return Greeting, Accept Apology, Self-Correction and Completion. A dialogue act in one of the feedback dimensions that has a responsive general-purpose function has a functional dependence relation; all other acts in a feedback dimension have a feedback dependence relation. The following specification lists the possible dependences of a dialogue act:

(4) Dependence relations:

1. Every dialogue act with a responsive communicative function has a functional dependence relation.
2. Every dialogue act in the Auto-Feedback or the Allo-Feedback dimension that does not have a responsive communicative function has a feedback dependence relation.
3. No other dialogue acts have functional or feedback dependence relations.

Since a dialogue act can have a functional or a feedback dependence relation but not both, the component in a dialogue act annotation structure (the component Δ in (7)) can be simply be the set of antecedents that the dialogue act depends on.[4] Since responsive dialogue acts and feedback acts are semantically incomplete without the specification of functional and feedback dependences, these are part of the structures that are used to annotate such acts.

A dialogue act may, finally, also be related to other dialogue acts through rhetorical relations, as in (5).

(5)
1. A: it ties you on in terms of the technology and the complexity that you want
2. A: like for example voice recognition
3. A: because you might need to power a microphone and other things

In this example[5] we see three functional segments, where the second is related to the first through an *Exemplification* relation, and the third to the first through an *Explanation* relation.

Different from functional and feedback dependence relations, rhetorical relations are not part of the meaning of a dialogue act, but add information to the way a fully defined dialogue act is used to establish a semantic relation to other dialogue acts (or how their semantic contents are related – see Petukhova et al., 2011). They are therefore not part of an entity strucutre that describes a dialogue act, but they occur in link structures that relate dialogue acts.

An abstract syntax consists in general of: (a) a specification of the elements from which annotation structures are built up, called a 'conceptual inventory', and (b) a specification of the possible ways of constructing annotation structures using these elements. The DiAML abstract syntax is defined by the following specification:

Specification 1. DiAML abstract syntax.
a. Conceptual inventory
The DiAML conceptual inventory consists of six sets:

1. a set of dimensions (ten in the case of DIT^{++}; nine in ISO 24617-2; these include the dimensions *Task*, *Auto-Feedback* and *Allo-Feedback*);
2. a set of communicative functions, partitioned into 'general-purpose' functions, which can be used in any dimension, and sets of 'dimension-specific' functions for each dimension except *Task*. A subset *RSP* of the set of communicative functions is specified as the 'responsive' communicative functions (including both general-purpose and dimension-specific functions);

[4] This is a small technical improvement over the definition of the DiAML abstract syntax in ISO 24617-2: 2012.

[5] From the AMI corpus, see `http://corpus.amiproject.org`

3. a set of qualifiers that can be associated with dialogue acts, partitioned into subsets for certainty, conditionality, and sentiment;

4. a set of rhetorical relations that can hold between dialogue acts (or their semantic contents);

5. a set of dialogue participants, including possible side-participants or audiences, besides actively participating speakers and addressees;

6. a set of functional segments of primary data.

The sets of functional segments and dialogue participants are specific for a particular annotation task; the other concepts are tasl-independent.

b. Annotation structures

An annotation structure is a set $\{\varepsilon_1, \ldots, \varepsilon_k, L_1, \ldots, L_m\}$ consisting of the entity structures $\{\varepsilon_1, \ldots, \varepsilon_k, \}$, with $k \geq 1$, and the link structures $\{L_1, \ldots, L_m\}$ (with $m \geq 0$). Entity structures contain semantic information about a functional segment; link structures describe semantic relations between functional segments. An entity structure in DiAML is a pair

(6) $\varepsilon = \langle m, \alpha \rangle$

consisting of a functional segment m (a 'markable') and the characterization of a dialogue act, which is either a 6-tuple (7a) or a 7-tuple (7b), where S is the sender of the dialogue act; A is a non-empty set of addressees; H is a (possibly empty) set of other dialogue participants; d is a dimension; f is a communicative function; Q is a (possibly empty) set of qualifiers, and Δ is a set of other dialogue acts that the dialogue act in focus depends on.

(7) a. $\alpha = \langle S, A, H, d, f, Q \rangle$
b. $\alpha = \langle S, A, H, d, f, Q, \Delta \rangle$

Case (7a) occurs when the communicative function f is not a responsive one and the dialogue act does not belong to the auto- or the allo-feedback dimension.

A link structure is a triple $\langle \varepsilon, E, \rho \rangle$ consisting of an entity structure ε, a non-empty set E of entity structures, and a rhetorical relation ρ, which relates the dialogue act α in ε to the entity structures in E.

3.3. DiAML Representations

3.3.1. Anchoring Annotations in Primary Data

Representation in DIAML relies on a three-level architecture: (1) a primary source, which may correspond to a speech recording, textual transcription or any further low-level annotation thereof; (2) the marking of functional segments in the primary source; (3) the actual dialogue act information associated with a functional segment. DiAML annotation is concerned with level (3) and follows the stand-off annotation approach: annotations refer to segments of the primary data specified at level (2), and the primary data are kept separate. This is quite clear in DiAML-XML representations, such as (1), where functional segments appear as the values of the 'target' attribute, which

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
<body / >
<div><head>The dialogue turns, segmented into words (TEI-compliant)</head>
   <u>
   <w xml:id="w1">right</w>
   <w xml:id="w2">go</w>
   <w xml:id="w3">south</w>
   <w xml:id="w4">and</w>
   <w xml:id="w5">you'll</w>
   <w xml:id="w6">pass</w>
   <w xml:id="w7">some</w>
   <w xml:id="w8">cliffs</w>
   <w xml:id="w9">on</w>
   <w xml:id="w10">your</w>
   <w xml:id="w11">right</w>
   </u>
</div>
<div><head>Identification of functional segments</head>
<spanGrp xml:id="ves1" type="functionalVerbalSegment">
  <span xml:id="ts1" type="textStretch" from="w1" to="w1"/>
</spanGrp>
<fs type="functionalSegment" xml:id="fs1"/>
  <f name="verbalComponent" fVal="#ves1"/ ><fs/ >
<spanGrp xml:id="ves2" type="functionalVerbalSegment">
  <span xml:id="ts2" type="textStretch" from="w2" to="w11"/>
</spanGrp>
<fs type="functionalSegment" xml:id="fs2">
<f name="verbalComponent" fVal="#ves2"/></fs>
</div>
</body>
</TEI>
```

Figure 4: TEI-compliant segmentation of primary data.

are assumed to be given as markables. Figure 4 shows how these markables can be defined at level 2 in a TEI-compliant way.

To make the tabular representation formats shown in Figures 1 and 2 fit into this 3-level architecture, these formats were modified as described in the next subsections.

3.3.2. DiAML-TabSW

First, it may be noted that the third column in Fig. 1 represents three things: the speaker, the slash units into which a turn may be subdivided ('utt1' etc.), and a transcript of what the speaker said (with in-line markups, mostly related to disfluencies). These ingredients were separated by introducing a column to represent the speaker, and replacing the 'utt1' marks by functional segment identifiers in the left-most column, where they replace the contents shown in Fig. 1; in fact, the replacing identifiers are references to specifications of stretches of the primary data in a separate file, for instance as a sequence of word tokens or as a stretch of speech with a given start- and end point. This file corresponds to level (2) in the 3-level architecture, and forms an implementation of stand-off annotation in tabular form that remedies the limitations of the Fig. 1 representation of being unable to deal with discontinuous or overlapping functional segments. For example, the discontinuous functional segment sw01-0105.fs3 in Fig. 5 is specified in the file sw01-0105-fs as consisting of the word tokens w12, w13, w14, and w16 (I, kind, of, I). For the sake of readability, the

markables	ID	Dialogue acts	Sp	FS text	Turn transcript
sw01-0105-fs.1	da1	Ta:setQuestion	A	Jimmy, so how do you get most of your news?	Jimmy, {D so } how do you get most of your news? /
			B		{D Well, } [I kind of, + {F uh, } I] watch the, national news every day, for one. / I also read one or two papers a day / {C and } [I'm a, + I'm pretty much a] news junkie / {C and } I tune in to CNN a lot. /
sw01-0105-fs.2	da2 da3	TiM:stalling TuM:turnTake	B:	Well,	
sw01-0105-fs.3 w16	da4	OCM:selfCorrection	B B	I kind of, I	
sw01-0105-fs.4	da5	TiM;stalling	B	uh	
sw01-0105-fs.5	da6	Ta:answer (da2)	B	I watch the national news every day, for one	
sw01-0105-fs.6	da7	TiM:stalling	B	uh	
sw01-0105-fs.7	da8	Ta:answer (da2) (Expansion foregr da7)	B	I also read on or two papers a day	
sw01-0105-fs.8	da9	TuM:turnKeep	B	and	
sw01-0105-fs.9	da10	Ta:inform	B	I'm pretty much a news junkie	
sw01-0105-fs.10	da11	OCM:selfCorrection	B	I'm a, I'm pretty much a	
sw01-0105-fs.11	da12	TuM:turnKeep	B	and	
sw01-0105-fs.12	da13	Ta:answer (da2) (Expansion foregr da7, d9]	B	I tune in to CNN a lot	
sw01-0105-fs.13	da14	AuF:autoPositive	A	Oh, wow.	Oh, wow.

Figure 5: ISO 24617-2 annotation of dialogue fragment in Fig. 1, represented in DiAML-TabSW format.

text of a functional segment is represented in an extra column (column 5 in Fig. 5); the transcripts of speaker turns were retained as in Fig. 1, allowing one to see immediately where a functional segment occurs. The textual information in the columns 4 and 5 in Fig. 5 is strictly speaking redundant, and has no formal status in DiAML, but makes the annotations more readable.[6] Contentwise, the most significant modification is the replacement in the second column of the SWBD-DAMSL tags by the DiAML representation of (1) communicative function; (2) dimension; (3) qualifiers (if any); and (4) dependences (if any) - and this for all dialogue acts expressed by the functional segment of that row. Figure 5 shows the resulting DiAML-TabSW format applied to the same dialogue fragment as Fig. 1.

3.3.3. DiAML-MultiTab

The tabular representation format produced by the DitAT tool for DIT annotations, shown in Fig. 2, was likewise modified in order to be fully ISO-compliant. The identifiers of functional segments in the leftmost column in Fig. 2 were replaced by references to the functional segment specifications. The 'transcript' column in Fig. 2 was split into a column containing functional segment texts and one containing turn transcripts.

The Contact Management column in Fig. 2 was deleted, since this dimension has not been adopted in ISO 24617-2. Like in DiAML-TabSW, the contents of the cells in the dimension columns were enriched to contain complete dialogue act information according to ISO 24617-2. The resulting format is shown in Fig. 6.

3.4. Encodings and Mappings

Functional and feedback dependence relations give rise to nested structures in the abstract syntax. For example, an answer by participant A to a question by participant B about the task domain takes the form of an entity structure with the following schematic form (for simplicity omitting empty sets of 'other participants' and qualifiers):

$$\langle m, \langle A, B, Task, answer, \langle m, \langle B, A, Task, question \rangle \rangle \rangle \rangle$$

'Flat' representation of annotation structures is made possible by the introduction of identifiers for dialogue act representations and using these to refer from one dialogue act to another, as in Fig. 5 for dialogue act da6.

In a tabular format, rhetorical relations are are most conveniently rendered as a property of the second argument of a relation, since this is typically where the existence of a rhetorical relation becomes apparent; see e.g. dialogue act da8 in Fig. 5.

DiAML annotation structures can be represented in MultiTab through the following procedure, which defines the encoding function $F_{MultiTab}$.

Specification 2. Encoding DiAML annotation structures in DiAML-MultiTab representation format.
For a given annotation structure $\{\varepsilon_1, \ldots, \varepsilon_k, L_1, \ldots, L_m\}$:

1. Step 1: introduction of identifiers for entity structures. Sort the entity structures $\varepsilon_1, \ldots, \varepsilon_k$ according to their markables. Sort entity structures with the same mark-

[6] See ISO Principles or Bunt (2015) for the use of elements in a concrete representation that have no correspondence to elements in the underlying abstract syntax.

markables	sp	fs text	turn transcript	Task	Auto-Feedback	Turn Man.	Time Man.	Discourse Structuring	Soc.Obl. Man.
			hello, can I						
			help you						
TR1-fs.1	s	hello							da1:Initial Greeting
TR1-fs.2	s	can I help you						da2:Offer	
			uhm, yes hello, maybe, I'd like to take a tanker with orange juice from...						
TR1-fs.3	u	uhm				da3:Turn Take	da4:Stalling		
TR1-fs.4	u	yes hello			da5: Pos. da1)				
TR1-fs.5	u	yes maybe						da6: Accept Offer(da2) [uncertain]	
TR1-fs.6	u	I'like to take a tanker..		da7: Inform					

Figure 6: ISO 24617-2 annotation of TRAINS dialogue fragment represented in DiAML-MultiTab format

able according to their dimension, with Task = 1, Auto-Feedback = 2, Allo-Feedback = 3, Turn Management = 4, Time Management = 5, Own Communication Management = 6, Partner Communication Management = 7, Discourse Structuring = 8, Social Obligation Management = 9. Assign to each entity structure an index corresponding to its position in the resulting ordering.
Output of this step is a set $E = \{\langle \varepsilon_i, i\rangle, \ldots, \langle \varepsilon_n, n\rangle\}$ of entity structures with indices.

2. Step 2: extraction of elements for cells in DiAML-MultiTab cells from indexed entity structures:
$T_e(\langle\langle m, \langle S, A, H, d, f, Q, \Delta\rangle\rangle, i\rangle) = \langle\langle m, \langle S, A, H, \langle d, \langle i, f, Q, \Delta\rangle\rangle\rangle\rangle\rangle$

3. Step 3: restructuring the information in link structures in entity-like form. If $L = \langle \varepsilon_1, \{\varepsilon_2, ..., \varepsilon_k\}, \rho\rangle$, with $\varepsilon_1 = \langle m_1, \langle S_1, A_1, H_1, d_1, f_1, Q_1, \Delta_1\rangle\rangle$, then
$T_L(L) = \langle m_1, \langle S_1, A_1, H_1, \langle d_1, \langle i_1, f_1, Q_1, \Delta_1 \langle \rho, \rho_1, \{i_2, .., i_k\}\rangle\rangle\rangle\rangle\rangle$

   where ρ_1 is the argument role of the related dialogue act(s)[7]; $i_2, .., i_k$ are the indices of the structures $\{\langle \varepsilon_2, i_2\rangle, ..., \langle \varepsilon_k, i_k\rangle\}$, built in step 1.
   The structures buit in this step are copies of structures built in step 2, extended with information from rhetorical links. The next step will eliminate duplicated information by merging the structures with and without rhetorical link information.

4. Step 4: merge of structures built in the previous two steps with and without rhetorical link information. This merge operation succeeds only if both arguments are identical except that one of them has additional rhetorical link information specified.

5. Step 5 finally forms combinations of all the structures constructed so far that have the same markable. This operation is defined as: $\langle m, \langle S, \langle d_1, \alpha\rangle\rangle\rangle \oplus \langle m, \langle S, \langle d_2, \beta\rangle\rangle\rangle = \langle m, \langle S, \{\langle d_1, \alpha\rangle, \langle d_2, \beta\rangle\}\rangle\rangle$

These steps produce a set of structures of the form $\langle m, S, A, H, \{\langle d_1, \alpha_1\rangle, ...\langle d_k, \alpha_k\rangle\}\rangle$ where α_j is maximally a quintuple $\langle i, f, Q, \Delta, R_{rh}\rangle$ (each of the last three elements may be absent) and corresponds to the content of a cell in of one of the nine dimension-related columns in DiAML-MultiTab representations.

A convenient layout of such a table uses the first 4 columns for representing markable (m), speaker (S), addressee (A) and 'other participants' (H), and the remaining 9 columns as corresponding to the 9 dimensions of ISO 24617-2, representing the dialogue acts within a particular dimension in the corresponding column. For the sake of readability, the addressee column may be suppressed for two-party dialogues, and the 'other participants' column may be suppressed if there are no such participants; instead, as we did in Fig. 6, columns representing the textual content of functional segments and turns greatly increases the readability.

The DiAML-TabSW encoding of annotation structures (the function F_{TabSW}) may be defined in a similar way, the only difference being that all dialogue acts expressed by a functional segment are represented in a single column with an indication of their dimension, as illustrated in Fig. 5.
The definition of the encoding functions $F_{MultiTab}$) and F_{TabSW} demonstrates the completeness of the two tabular representation formats; their unambiguity can likewise be demonstrated by defining the reverse functions. Together

[7]The abstract syntax of ISO standard 24617-8:2016 for annotating rhetorical relations assigns to each relation two argument roles, such as the roles *reason* and *result* to the relation *Cause*.

Origin	Lang	Original representation	Original annotation	DiAML representation
HCRC Map Task	EN	NITE XML	HCRC Map Task communicative functions	DiAML-XML
Switchboard	EN	3-column tabular	SWBD-DAMSL communicative functions	DiAML-TabSW
TRAINS	EN	13-column tabular	DAMSL communicative functions	DiAML-MutiTab
DBOX	EN	DiAML-XML	ISO 24617-2 annotations	DiAML-XML
Dutch Map Task	NL	plain text transcript	no dialogue act annotation	DiAML-MultiTab
DIAMOND	NL	13-column tabular	DIT^{++} communicative functions and dimensions	DiAML-MultiTab
OVIS	NL	plain text transcript	no dialogue act annotation	DiAML-MultiTab
Schiphol Airport	NL	plain text transcript	no dialogue act annotation	DiAML-XML

Table 1: Current contents of the DialogBank corpus.

with the corresponding functions for the DiAML-XML format, this demonstrates the interoperability of the three representation formats.

4. Conclusions and Future Work

The DialogBank is presently in beta version and had its first public release in December 2015. It contains at the time of writing annotated dialogues with the properties shown in Table 1. The dialogues from the HCRC Map Task and TRAINS corpora were re-segmented and re-annotated according to ISO 24617-2. The annotations of the Switchboard dialogues were converted semi-automatically by Fang et al. (2012) from SWBD-DAMSL tags to ISO 24617-2 function tags; this has been used as an intermediate step to the construction of full-blown ISO 24617-2 annotations. The DBOX corpus was collected at the University of Saarland and annotated using the ISO 24617-2 standard. The annotations of the dialogues from the DIAMOND corpus were obtained by adapting DIT^{++}. The dialogues from the Dutch Map Task, OVIS, and Schiphol corpora had not been annotated before.

In the near future, more annotated dialogues from these corpora will be added. as well as multi-party dialogues from the AMI corpus.

Bibliography

Allen, J. and Core, M. (1997) DAMSL: Dialogue Act Markup in Several Layers. Technical Report, Multiparty Discourse Group.

Allen, J. et al. (1994) The TRAINS Project: a case study in building a conversational planning agent. *Technical Note 94-3*. University of Rochester.

Alexandersson, J., et al. (1998) Dialogue acts in Verbmobil-2. Second edition. *Report 226*. DFKI Saarbrücken.

Andersson, A., et al. (1991) The HCRC Map Task Corpus. *Language and Speech* 34, 351-366.

Bunt, H. (1994) Context and dialogue control. *Think Quarterly 3(1)*, 19-31.

Bunt, H. and Black, W. (2000) Dialogue pragmatics and context specification. In Harry Bunt and William Black (eds.) *Abduction, Belief and Context in Dialogue*, Benjamin, Amsterdam, pp. 81-150.

Bunt, H. (2009) The DIT^{++} taxonomy for functional dialogue markup. *AAMAS Workshop 'Towards a Standard Markup Language for Embodied Dialogue Acts*, 13-23.

Bunt, H. (2010) A methodology for designing semantic annotation languages exploiting syntactic-semantic isomorphisms. In *Proceedings of ICGL 2010, Second International Conference on Global Interoperability for Language Resources*, Hong Kong, pp. 29-45.

Bunt, H. (2015) On the principles of semantic annotation. In *Proceedings of ISA-11*, London, pp. 1-13.

Bunt, H., Alexandersson, J., Carletta, J., Choe, J.-W., Fang, A., Hasida, K., Lee, K., Petukhova, V., Popescu-Belis, A., Romary, L., Soria, C. and Traum, D. (2010) Towards an ISO standard for dialogue act annotation. In *Proceedings of LREC 2010*.

Fang, A., Cao, J., Bunt, H., and Liu, X. (2012) The annotation of the Switchboard corpus with the new ISO standard for Dialogue Act Analysis. In *Proceedings of ISA-8*, Pisa, pp. 13-18.

Geertzen, J., et al. (2004) The DIAMOND project. In *Proceedings of SIGSEM 2004 (CATALOG)*, Barcelona.

Geertzen, J. (2007) DitAT: a flexible tool to support web-based dialogue annotation. *Proceedings of IWCS 2007*, Tilburg, pp. 320-323

Ide, N. and Romary, L. (2004) International Standard for a Linguistic Annotation Framework. *Natural Language Engineering* 10: 221-225.

Jurafsky, D., Schriberg, E., and Biasca, D. (1997) *Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation: Coders Manual*, University of Colorado.

Kipp, M. (2001) Anvil - A Generic Annotation Tool for Multimodal Dialogue. *Eurospeech 2001*, 1367-1370.

Petukhova, V. and Prévot, L. and Bunt, H. 2011 Discourse Relations in Dialogue. In: *Proc. ISA-6, Oxford)*, 18-27.

Petukhova, V. et al.(2014) The DBOX Corpus Collection of Spoken Human-Human and Human-Machine Dialogues, In *Proceedings of LREC 2014*, Reykjavik.

Prüst, H., Minnen, G., and Beun, R.-J. (1984) *Transcriptie dialooogesperiment juni/juli 1984*, IPO Rapport 481. Eindhoven University of Technology.

Shriberg, E. et al. (2004) The ICSI meeting recorder dialogue act (MRDA) corpus. *Proc. SIGDIAL 2004*, 97-100.