Downward Compatible Revision of Dialogue Annotation

Harry Bunt[1], Emer Gilmartin[2], Simon Keizer, Catherine Pelachaud, Volha Petukhova, Laurent Prévot and Mariët Theune

[1]Tilburg University, harry.bunt@uvt.nl
[2]Trinity College Dublin, egil@tdc.ei
[3]Vrije Universiteit Brussel, keizer.simon@gmail.com
[4]Université Paris VIII, catherine.pelachaud@upmc.fr
[5]Saarland University, v.v.petukhova@gmail.com
[6]Aix-Marseille Université, laurent.prevot@lpl-aix.fr
[7]University of Twente, Enschede, m.theune@utwente.nl

Abstract

This paper discusses some aspects of revising the ISO standard for dialogue act annotation (ISO 24617-2). The revision is aimed at making annotations using the ISO scheme more accurate and at providing more powerful tools for building natural language based dialogue systems, without invalidating the annotated resources that have been built, with the current version of the standard. In support of the revision of the standard, an analysis is provided of the downward compatibility of a revised annotation scheme with the original scheme at the levels of abstract syntax, concrete syntax, and semantics of annotations.

1 Introduction

ISO standards are examined every five years for the need to be brought up to date or to be improved. The ISO standard for dialogue act annotation, ISO 24617-2,[1], was published in September 2012 and is thus up for revision, if deemed necessary,[2]

When a revised annotation scheme is used to annotate corpus data, the resulting annotations will be in some respects differ from those according to the original version. An important issue concerning the usefulness of a revision is the compatibility between annotations according to the two versions. In particular, it is desirable that old annotations are still valid in the revised version, and do not require to be re-annotated (or converted). In other words, the revised standard should preferably be *downward compatible* with the original version. Downward compatibility is a well-known design property of computer hardware and software, and can be applied also to annotation schemes. This is discussed in Section 3, where the compatibility of annotation schemes is analysed and related to the properties of extensibility, optionality, and restrictability.

First, Section 2 briefly summarizes the ISO 24617-2 standard. Section 3 introduces the notion of *downward compatibility* for the revision of an annotation scheme, and relates it to different forms of optionality. Section 4 discusses some inaccuracies, and outlines possible solutions to be implemented in its second edition. Section 5 briefly considers four different use cases of the standard, and what kind of extensions would be relevant for which use case. Section 6 discusses some inconvenient limitations of the current version, and corresponding extensions that respect the requirement of downward compatibility. Section 7 ends the paper with conclusions and perspectives for revising the standard.

2 The ISO 24617-2 Standard

The ISO 24617-2 annotation standard consists of two main components: (a) a comprehensive, domain-independent set of concepts that may be used in dialogue act annotation, meticulously defined in the form of ISO data categories, and (b) the markup language DiAML (Dialogue Act Markup Language). In its stock of annotation concepts, in particular its taxonomy of communicative functions, ISO 24617-2 builds on previously designed annotation schemes such as DIT^{++}, DAMSL, MRDA, HCRC Map

[1]ISO 24617-2, Language Resources Management, Semantic Annotation Framework, part 2: Dialogue acts.

[2]This issue was discussed at the ISO-13 workshop in September 2017, where it was felt to be desirable to improve and extend the existing standard in some respects. The present paper is partly based on recommendations for revising the standard that were reached at a two-day workshop in April 2018.

Task, Verbmobil, SWBD-DAMSL, and DIT.[3] The ISO 24617-2 scheme supports semantically richer annotations than most of its predecessors in including the following aspects:

Dimension: The ISO scheme supports multidimensional annotation, i.e. the assignment of multiple communicative functions to dialogue segments; following DIT^{++}, an explicitly defined notion of 'dimension' is used that corresponds to a certain category of semantic content. Nine orthogonal dimensions are defined: (1) *Task:* dialogue acts that move forward the task or activity which motivates the dialogue; (2-3) *Feedback*, divided into *Auto-* and *Allo-Feedback*: acts providing or eliciting information about the processing of previous utterances by the current speaker or by the current addressee, respectively; (4) *Turn Management:* activities for obtaining, keeping, releasing, or assigning the right to speak; (5) *Time Management:* acts for managing the use of time in the interaction; (6) *Discourse Structuring:* dialogue acts dealing with topic management or otherwise structuring the dialogue; (7-8) *Own-* and *Partner Communication Management:* actions by the speaker to edit his current contribution or a contribution of another speaker; (9) *Social Obligations Management:* dialogue acts for dealing with social conventions such as greeting, apologizing, and thanking.

Qualifiers for expressing that a dialogue act is performed conditionally, with uncertainty, or with a particular sentiment.

Dependence relations for semantic dependences between dialogue acts, e.g. question-answer (functional dependence), or for relating a feedback act to the utterance(s) that the act reacts to (feedback dependence).

Rhetorical relations , for example for indicating that the performance of one dialogue act explains that of another dialogue act.

The ISO schema defines 56 communicative functions, which are listed in Appendix A. Some of these are specific for a particular dimension; for instance *Turn Take* is specific for Turn Management; *Stalling* for Time Management, and *Self-Correction* for Own Communication Management. Other functions can be applied in any dimension; for example, *You misunderstood me* is an *Inform* in the Allo-Feedback dimension. All types of question, statement, and answer can be used in any dimension, and the same is true for commissive and directive functions, such as *Offer, Suggest,* and *Request*. Such functions are called *general-purpose* functions; the former *dimension-specific* functions.

ISO 24617-2 annotations assume that a dialogue act has seven components: a sender, a set of one or more addressees, zero or more other participants, a dimension, a communicative function, possibly one or more functional or feedback dependence relations (depending on the type of dialogue act), possibly one or more qualifiers, and possibly one or more rhetorical relations to other dialogue acts.

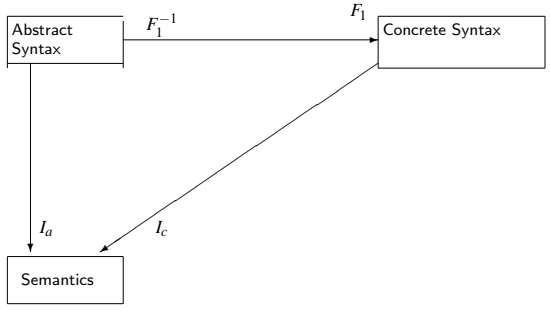

Figure 1: Abstract and concrete syntax, and semantics

The DiAML markup language was designed in accordance with the ISO Linguistic Annotation Framework (LAF)[4] and the ISO Principles of Semantic Annotation (ISO 24617-6).[5] LAF distinguishes between *annotations* and *representations*: 'annotation' refers to the linguistic information that is added to segments of language data, independent of format; 'representation' refers to the rendering of annotations in a particular format.

Following the ISO Principles, this distinction is implemented in the DiAML definition by distinguishing an *abstract syntax* that specifies a class of *annotation structures* as set-theoretical

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

[4] ISO 24612:2010; see also Ide & Romary (2004).

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

constructs, like pairs and triples of concepts, and a *concrete syntax* that specifies a rendering of these annotation structures in a particular format. A representation format is defined called DiAML-XML, which uses abbreviated XML-expressions. The annotations have a semantics which is defined for the abstract syntax (see Fig. 1), thus allowing alternative representation formats to share the same semantics.

According to ISO 24617-2, dialogue acts are expressed by 'functional segments', defined as *minimal stretches of communicative behaviour that have a communicative function and a semantic content*, 'minimal' in the sense of excluding material that does not contribute to the expression of the dialogue act. Functional segments may be discontinuous, may overlap, and may contain parts contributed by different speakers.

Example (1) shows a DiAML-XML annotation representation. It illustrates among other things the annotation of relations between dialogue acts: a rhetorical relation (Elaboration) between the dialogue acts in utterances 1 and 3, a functional dependence relation between the question in 2 and the answer in 3, and a feedback dependence relation between the dialogue acts in utterances 3 and 4.

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

Functional segments:

fs1 = o south and you'll pass some cliffs on your right

fs2 = uhm...

fs3 = straight south?

fs4 = yes

fs5 = passing some adobe huts on your left

fs6 = oh okay

```
<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="#g"
   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="#f" addressee="#g"
   dimension="autoFeedback" communicativeFunction="question"/>
<dialogueAct xml:id="da5" target="#fs4" sender="#g" addressee="#f"
   dimension="alloFeedback" communicativeFunction="answer"
   functionalDependence="#da4"/>
<dialogueAct xml:id="da6" target="#fs5" sender="#g" addressee="#f"
   dimension="task" communicativeFunction="inform"/>
<rhetoricalLink dact="#da6" rhetoAnteceden="#da1" rhetoRel="elaboration"
<dialogueAct xml:id="da7" target="#fs6" sender="#f" addressee="#g"
   dimension="autoFeedback" communicativeFunction="autoPositive"
   feedbackDependence="#da1 #da6"/>
</diaml>
```

3 Formal Properties of Schema Revision

3.1 Compatibility, Optionality, Extensibility, and Restrictibility

Designing a revised version of the ISO 24617-2 standard in a downward compatible way is greatly facilitated by the *extensibility* of the original version, which means that it allows its stock of concepts to be extended with additional concepts. ISO 24617-2 is extensible in four respects:

- **Dimensions**: Due to the orthogonality of the set of dimensions, additional dimensions may be introduced as long as they are orthogonal to the already existing dimensions and to each other.

- **Communicative functions**: The taxonomy of communicative functions defined in the standard expresses the semantic relations between functions: dominance relations express different degrees of specialization; and sister relations express mutually exclusivity of functions. Communicative functions may be added to the taxonomy as long as they respect these relations.
- **Qualifiers**: Like dimensions, due to the orthogonality of the qualifier attributes and their values.
- **Rhetorical relations**: The ISO standard does not specify a particular set of relations, but allows any such set to be plugged in.

The extensibility of ISO 24617-2 is in turn facilitated by the *optionality* of some of its components. Following the ISO Principles of semantic annotation, three types of optionality can be distinguished:

Type I (semantic optionality): a component that a certain type of annotation structure may contain, but does not have to. If it does contain that component then this provides extra information, compared to the case where it does not. Example: the specification of a set of 'other participants' for a dialogue act.

Type II (syntactic optionality): a component may be but does not need to be specified in annotation representations, since it has a default value in the abstract syntax, which is assumed in the encoded annotation structure if it is not specified. Example: the polarity in the annotation of an event by means of an <event> element in ISO-TimeML.

Type III (uninterpreted optionality): a component may be specified in annotation representations but does not encode anything in the abstract syntax, and thus has no semantic interpretation (but the component may be useful for an annotation process or for other purposes). Example: the indication of the part of speech of an event description by means of an <event> element in ISO-TimeML.

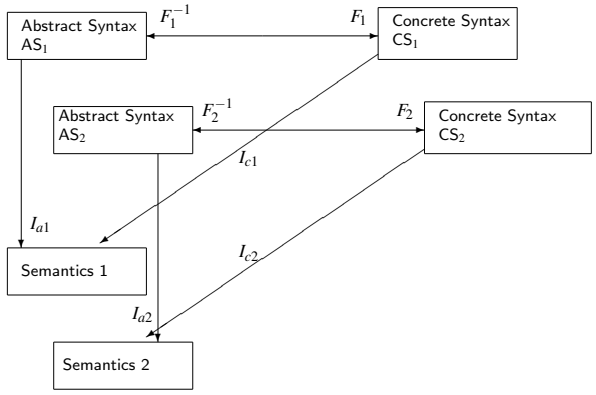

Figure 2: Optionality in abstract and concrete syntax, and semantics

These distinctions can be made precise in terms of the abstract and concrete syntax of annotations and their semantics, as shown in Figure 2, where two versions of an annotation scheme are considered, with abstract syntax specifications AS_1 and AS_2, two semantic specifications by means of the interpretation functions I_{a1} and I_{a2}, and two concrete syntax specifications CS_1 and CS_2. The encoding and decoding functions F_1, F_1^{-1}, F_2, and F_2^{-1} relate the structures generated by the two abstract and concrete syntax specifications, respectively, and define the semantics of concrete representations by means of the composite functions $I_{c1} = I_{a1} \circ F_1^{-1}$ and $I_{c2} = I_a \circ F_2^{-1}$.

Let α be an annotation structure generated by AS_1, with encoding $F_1(\alpha) = \beta$. Let δ_a be an optional addition to α according to the abstract syntax AS_2, forming the annotation structure designated by $\alpha+\delta_a$, and let δ_c be the corresponding element in the concrete syntax CS_2, forming an annotation representation designated by $\beta + \delta_c$.

Semantic optionality (Type I) can now be defined formally as the case where δ_c represents additional semantic information:

(2) $F^{-1}(\beta + \delta_c) = \alpha + \delta_a$
$I_c(\beta + \delta_c) = I_a(F^{-1}(\beta + \delta_c)) = I_a(\alpha + \delta_a)$

Syntactic optionality (Type II) is the case that an optional addition δ_c in a representation $\beta + \delta_c$ (such as polarity="positive") indicates that the abstract annotation structure $\alpha[\delta_a]$ that it encodes, includes its default value δ_{ad}:

(3) $F^{-1}(\beta + \delta_c) = \alpha[\delta_{ad}]$
$I_c(\beta + \delta_c) = I_a(F^{-1}(\beta)) = I_a(\alpha[\delta_{ad}])$

Finally, uninterpreted optionality (Type III) is the case where the representation with the optional element encodes the same semantic information as the structure without the optional element, not requiring a default value in the abstract annotation structure:

(4) $F^{-1}(\beta + \delta_c) = F^{-1}(\beta)$
$I_c(\beta + \delta_c) = I_a(F^{-1}(\beta + \delta_c)) = I_a(F^{-1}(\beta)) = I_a(\alpha)$

The following elements of ISO 24617-2 are optional in one of these three senses:

- **Qualifiers:** The qualifier attributes *Certainty* and *Conditionality* have default values ('certain' and 'unconditional', respectively), hence they form a Type II optionality. The attribute *Sentiment* has no values defined; in this respect the annotation scheme is extensible: any set of values may be used. If this set contains a default value; then the specification of that value is an optionality of Type II; for all other values it is of Type I, since the semantics is defined (in terms of predicates used in information state updates, see Bunt, 2014).

- **Rhetorical relations:** If specified, these add semantic information about relations between dialogue acts or their semantic content. There is no 'default' rhetorical relation, hence this is a Type I optionality.

ISO 24617-2 currently has no cases of Type III optionality, but its revision is expected to have some.

Annotation schemes are usually considered only at the level of concrete syntax, and have no abstract syntax or semantics. Notions such as extensibility are thus typically considered only at that level, in terms of adding attributes and/or values to XML elements. In the 3-layer architecture of DiAML, extensibility must be considered at all three levels; extending the representations defined by the concrete syntax is only semantically significant if the corresponding extensions are introduced in the abstract syntax, and their semantic interpretation is defined. Since this is technically nontrivial, user-defined extensions are typically Type III optional, and are disregarded by software that interprets the annotations.

The converse of extensibilty is the *'restrictability'* of an annotation scheme: the possibility to not use the entire stock of concepts offered by the scheme, but only a subset. ISO 24617-2 is restrictable in its set of dimensions and its set of communiative functions; as the official description of the standard in the ISO 24617-2:2012 document stipulates:

- *"A dimension and the corresponding set of dimension-specific communicative functions may be left out; by virtue of the orthogonality of the set of core dimensions, this has no influence on the remaining dimensions."*
- *"Communicative functions may be left out for which there is a less specific function in the taxonomy"*

In order to ensure that desirable extensions of ISO 24617-2 are well-defined at all three levels, it seems attractive to define such extensions in ISO 24617-2 Edition 2 while insisting on its restrictability, thus supporting the use of additional dimensions and communicative functions with a well-defined semantics without making their use obligatory.

3.2 Constraints on Revisions

Figure 3 shows the three levels of an Edition 1 annotation scheme and a revised version, Edition 2 with the functions A_{12}, S_{12}, and C_{12} which describe the revision at each level, i.e. if α_1 is an Edition 1 annotation structure, then $A_{12}(\alpha_1)$ is the revised annotation structure, and similarly at the other levels.

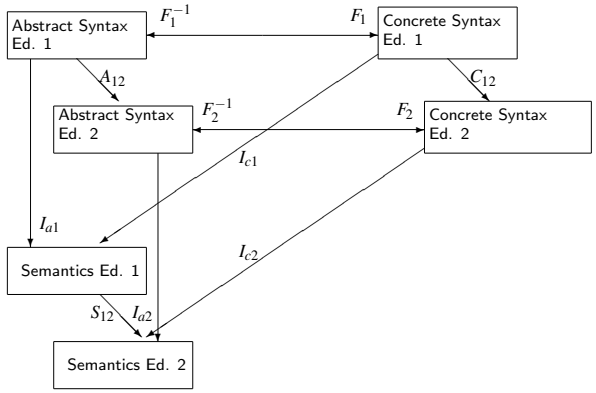

Figure 3: Annotation schema revision in abstract and concrete syntax, and semantics

Note that the revised representation of an Edition 1 annotation structure α_1 can be computed in two ways: (1) by applying the Edition 2 encoding function F_2 to the revised annotation structure $A_{12}(\alpha_1)$, and (2) by applying the representation revision function C_{12} to the Edition 1 representation $F_1(\alpha_1)$. The result should in both cases of course be the same:

(5) $F_2(A_{12}(\alpha_1)) = C_{12}(F_1(\alpha_1))$.

Since this is true for any Edition 1 annotation structure α_1, a requirement on consistent revision is that the function compositions $F_2 \circ A_{12}$ and $C_{12} \circ F_1$ are identical:

(6) $F_2 \circ A_{12} = C_{12} \circ F_1$.

Similarly, two ways of computing the Edition 2 meaning of an Edition 1 annotation structure are: (1) computing its Edition 1 meaning $I_{a1}(\alpha_1)$ according to Edition 1 and applying the semantic revision function S_{12}, and (2) determining the revised annotation structure $A_{12}(\alpha_1)$ and computing its Edition 2 meaning by applying the interpretation function I_{a2}. Again, the result should in both cases be the same:

(7) $S_{12}(I_{a1}(\alpha_1)) = I_{a2}(A_{12}(\alpha_1))$

Since this is true for any Edition 1 annotation structure α_1, a second consistency requirement on annotation schema revision is:

(8) $S_{12} \circ I_{a1} = I_{a2} \circ A_{12}$

3.3 Downward Compatible Revision

Whether an annotation according to the original standard ('Edition 1') is valid according to its revised version ('Edition 2'), should be considered at all three levels of the definitions: abstract syntax, concrete syntax, and semantics. An Edition 1 annotation structure α_1 is valid according to Edition 2 if and only if (1) it belongs to the set of annotation structures defined by the Edition 2 abstract syntax and (2) it has the same meaning as in Edition 1. In other words, for Edition 2 to be downward compatible with Edition 1 the functions A_{12} and S_{12} are the identity function, and the interpretation functions I_{a1} and I_{a2} assign the same meanings to Edition 1 annotation structures and their Edition 2 versions, respectively (thus respecting constraint (8)). The Edition 2 set of annotation structures is thus a superset of the Edition 1 set of annotation structures, whose meanings are not changed. (Additional, in particular 'richer' meanings, are assigned to the Edition 2 annotation structures that are not also Edition 1 annotation structures.)

The Edition 2 annotation representations are defined by the Edition 2 concrete syntax, and in order to be downward compatible also at the level of concrete representations, this representation is preferably the same as the Edition 1 representation, but there is room for variation here: according to constraint (6) with A_{12} being the identity function, the representation conversion function C_{12} and the Edition 2 encoding function F_2 may be defined in such a way that, applied to an annotation structure that is also an Edition 1 annotation structure:

(9) $F_2(\alpha) = (C_{12} \circ F_1)(\alpha)$

(For those Edition 2 annotation structures that are not also Edition 1 annotation structures there are no consistency constraints on the definition of the encoding function F_2.) The effect of this is that, while the revision leaves the annotation structures of the Edition 1 abstract syntax unchanged, a conversion procedure implementing the function C_{12} may change their representations into a new form to become Edition 2 representations.

Note that, if a revision of an annotation scheme consists of extensions, optional elements, and/or refinements (more fine-grained annotations, or annotations with greater accuracy), then the revised version may indeed be downward compatible in the sense described here; if, by contrast, the revision includes corrections of errors in the earlier version, then the revised edition is not entirely downward compatible. The revisions of ISO 24617-2 recommended in this paper can all be viewed as extensions, including new optional elements, and refinements, leading to a downward compatible Edition 2.

4 Accuracy of Annotations

4.1 Dependence Relations

4.1.1 Dependence Relations for Feedback

ISO 24617-2 defines a feedback act as a *"dialogue act which provides or elicits information about the sender's or the addressee's processing of something that was uttered in the dialogue"*. A feedback act is thus a dialogue act in either the auto-feedback or the allo-feedback dimension. Moreover, it defines the feedback dependence relation as the *"relation between a feedback act and the stretch of communicative behaviour whose processing the act provides or elicits information about"*. The feedback dependence relation serves to identify this "something that was uttered in the dialogue". This is illustrated in (10), where the segment "The first train to the airport on Sunday" in S's utterance repeats material from C's question, which can be interpreted as a positive auto-feedback act by which S indicates to have understood which train C wants to know the departure time of.

(10) C: Do you know what time the first train to the airport leaves on Sunday?
S: The first train to the airport on Sunday is at 6:15.

The annotation of S's utterance thus considers this segment as a functional segment, with the communicative function *autoPositive*, and with a feedback dependence relation to what C said. However, ISO 24617-2 does not consider segments other than *functional* segments, so rather than a dependence relation to the corresponding (discontinuous) segment in C's utterance, the feedback dependence relation uses the smallest functional segment that includes the repeated material - in this case C's entire utterance. This is rather inaccurate. It is therefore recommended that ISO 24617-2 Edition 2 should include the possibility to refer to non-functional segments, whose relevance comes from the fact that they are referred to by feedback acts – "reference segments".

4.1.2 Dependence Relations for Own and Partner Communication Management

Reference segments are also needed for the accurate annotation of Own Communication Management acts and Partner Communication Management acts. For example, the accurate annotation of a self-correction (in the OCM dimension) or a partner correction (in the PCM dimension) requires the specification of the dialogue segment that is corrected, which may very well be a single word or morpheme.

4.1.3 Types of Dependence Relations

ISO 24617-2 defines the functional dependence relation as the *"relation between a given dialogue act and a preceding dialogue act on which the semantic content of the given dialogue act depends due to its communicative function."* Examples of such dialogue acts are the inherently responsive acts such as answers, (dis-)confirmations, (dis-)agreements, corrections and the acceptance or rejection of requests, offers, and suggestions.

Auto- and allo-feedback acts, which in a different sense are also responsive, come in two varieties: those whose communicative function is specific for these dimensions (*AutoPositive, AutoNegative, AlloPositive, AlloNegative, FeedbackElicitation*) and those whose communicative function is a general purpose function, such as *Question* (for clarification), *CheckQuestion*, or *Confirm*. The two varieties aare illustrated by the examples in 11):

(11) a. G: the turn left just above the adobe huts
F: okay [AutoPositive]

b. C: Best before nine on Monday, or else on Tuesday
S: Monday before nine you said? [Auto-Feedback, CheckQuestion]
C: That's right. [Allo-Feedback, Confirm]

The specification in ISO 24617-2 could be understood as saying that for the latter type of feedback act, if it has a responsive communicative function, like the *Confirm* act in (11b), then it should be annotated as having both a functional and a feedback dependence relation. This was not intended, however. In such cases the functional dependence relation, required for interpreting the responsive act, identifies the material that the feedback is about, so the use of both would be redundant. The same applies to dialogue acts in the OCM and PCM dimensions with a responsive communicative function. It is therefore recommended that the assignment of functional and feedback dependence relations should be specified more accurately than in ISO 24617-2 Edition 1, as follows:

1. For all dialogue acts in the Auto-Feedback, Allo-Feedback, OCM or PCM dimension:
   (a) if the communicative function is a responsive one, then assign a *functional* dependence relation to the dialogue act that is responded to;
   (b) if the communicative function is a general-purpose function but not a responsive one, or is dimension-specific for Auto-Feedback, Allo-Feedback, OCM or PCM, then assign a *feedback* dependence relation to the material that is reacted to.
2. In all other cases do not assign a dependence relation.

Note that, according to this specification, a feedback dependence relation is assigned to a feedback question like the CheckQuestion in (11b).

4.2 Rhetorical Relations

The dialogue acts that make up a dialogue are often rhetorically related. ISO 24617-2 supports the marking up of rhetorical relations (also know as *discourse relations*) as an optional addition to dialogue act annotation, but does not specify any particular set of relations to be used; it only specifies *how* a rhetorical relation may be marked up as relating two dialogue acts. The experience in dialogue act annotation is that rhetorical relations tend to be very important for a good understanding of the interaction. Users of the ISO scheme have often added these annotations, using a variant of the set of relations defined in ISO standard 24617-8, a.k.a. 'DR-Core'. This is a set of 18 'core' relations that are shared by many annotation schemes. It has been used in most of the dialogues in the DialogBank. Two problems were noted when annotating discourse relations in ISO 24617-2.

First, many rhetorical relations have two arguments that play different roles, for example, a Cause relation has a "Reason" and a "Result" argument. DiAML currently has no provision for indicating the roles in a rhetorical relation between dialogue acts. The DR-Core annotation scheme does have attributes and values for this purpose, so the annotation of rhetorical relations in dialogue could be made more accurate by importing some of the elements from DR-Core into DiAML.

Second, rhetorical relations may occur either between two dialogue acts, or between their semantic contents, or between one dialogue act and the semantic content of another. This phenomenon is known in the literature as the 'semantic-pragmatic' distinction. Example (12) illustrates this.

(12) a. *'Semantic Cause':*
A: Have you seen Pete today?
B: He didn't come in. He has the flu.

b. *'Pragmatic Cause':*
A: Have you seen Pete today?
B: He didn't come in. He sent me a message saying that he has the flu.

This distinction can only be made in DiAML if it is extended with the possibility to say something about the semantic content of a dialogue act. This is taken up in Section 6.4.

5 Use cases

The concepts and mechanisms defined in ISO 24617-2 can be used in at least four different situations:

U1 manual annotation of corpus data;
U2 automatic annotation of corpus data;
U3 online recognition of dialogue acts by interactive systems;
U4 dialogue management and dialogue act generation by a dialogue system.

These different use cases present different desiderata and requirements, in particular concerning the granularity of the available communicative functions. Concerning use cases U1 and U2, a trained manual annotator may bring richer background and context information to bear in the annotation process than an automatic system, and may therefore benefit from the availability of fine-grained, context-dependent communicative functions. Manual annotators with little training or experience may, on the other hand, benefit more from the use of more coarse-grained functions in order to produce consistent results.

Concerning use cases U3 and U4, for example, Malchanau et al. (2017) have shown the usefulness of DiAML as an interface language between the modules of a multimodal dialogue system, and Keizer et al. (2011) have shown the use of the DIT^{++} taxonomy of communicative functions, which underlies the ISO standard, in a multidimensional Dialogue Manager. In both cases issues of granularity of the communicative functions arise, in particular in the generation of feedback acts, where the Dialogue Manager typically has detailed information about the level of processing that it would be appropriate to provide feedback about. The DIT^{++} taxonomy of communicative functions distinguishes between feedback acts at five different levels of processing: (1) attention; (2) perception; (3) understanding; (4) evaluation; and (5) execution. For use cases U3 and U4 such a fine-grained set of feedback functions would be useful.

Given the restrictability that would be required from the second edition in order to be downward compatible, it follows that it is recommended to add more fine-grained concepts to the standard, and to provide use-case dependent guidelines for how to optimally make use of the concepts that the standard makes available.

6 ISO 24617-2 Extensions

6.1 Dimensions

Users of ISO 24617-2 have mentioned two dimensions that they missed, namely Task Management, known from DAMSL, and Contact Management, known from DIT^{++}. Task Management acts discuss or explain a certain task or activity that is pursued through the dialogue (as opposed to performing that task/activity). They occur for example in TV debates and in interactive games (see e.g. Petukhova et al., 2014).

Contact Management acts serve to establish and manage contact and attention. Casual conversations are known to contain a rich variety of greetings and leavetaking acts (Gilmartin et al., 2017), which often have such a function (see also the next subsection).

Since one of the attractive features of the ISO scheme is that its dimensions are 'orthogonal', Task Management and Contact Management can be added as optional additions without interfering with the existing 9-dimensional system, keeping the extended system downward compatible with the existing system, and are available in a given use case when needed.

6.2 Communicative Functions

The taxonomy of communicative functions in ISO 24617-2 makes it possible to add fine-grained communicative functions without making existing annotations incompatible with the standard. Experience in applying the ISO standard has given rise to the desire to have more fine-grained communicative functions for Social Obligations Management, Discourse Structuring, and Auto- and Allo-Feedback.

ISO 24617-2 was intended to be domain-independent, applicable to a wide range of tasks and domains, and consequently does not have domain-specific communicative functions. This has been felt

as a limitation when using the concepts of the standard for online dialogue act recognition or generation. It is recommended that the documentation of the standard discusses (informatively) two ways of defining domain-specific communicative functions: (a) as a way of specifying the semantic content of a general-purpose function (as illustrated by communicative functions for negotiation in the MIB corpus (Petukhova et al., 2016) ; and (b) as a dimension-specific communicative function for the Task domain, in which case the information-state update semantics of dialogue acts with that communicative function has to be defined.

Note that there can be no objection to the introduction of some examples of task-specific communicative functions in view of the restrictability of the standard in the use of the communicative functions that it defines.

6.3 Qualifiers

The available qualifiers for optional representation of certainty (default: certain) and conditionality (default: unconditional) seem adequate for their intended purpose. For emotion and sentiment the DiAML concrete syntax has the optional attribute 'sentiment', for which the standard does not specify any set of possible values, let alone a semantics, which makes the use of sentiment qualifiers Type III optional. For specifications of possible sets of emotion and sentiment values, and for more sophisticated annotation of the affective aspects of dialogue behaviour, it is recommended to look to EmotionML.

EmotionML, the W3C standard for annotating emotion (Baggio et al., 2014), does not prescribe the use of any particular set of emotion values, but supports the articulate annotation of emotions using alternative sets of values. Moreover, EmotionML is explicitly aimed at supporting the integration of emotion descriptions with other annotations. It would be attractive to extend the possibility to annotate emotion and sentiment (especially in multimodal dialogue) in DiAML by allowing EmotionML expressions in the concrete syntax of DiAML as optional elements of Type III that represent emotions with reference to dialogue acts.

6.4 Semantic Content

In dialogue act theory, a dialogue act is formally defined as a 8-tuple of which one of the elements is a semantic content (see Bunt, 2014). ISO 24617-2 focuses on the functional meaning of dialogue acts, and therefore annotates dialogue acts in DiAML (in the abstract syntax) as 7-tuples (see Section 2), leaving out the semantic content. For use in dialogue annotation (use cases U1 and U2) and for online recognition of dialogue acts (use case U3) this seems appropriate, but in online use in the dialogue management of a dialogue system (use case U4), there is a need to be able to specify information about the semantic content of dialogue acts. It is therefore recommended to explore the possibilities of extending DiAML with semantic content information. This has for example been done in the Virtual Negotiation Coach (Petukhova et al., 2017), where semantic content is specified by a set of attribute-value pairs that represent the state of a negotiation.

It may be noted that the semantics of dialogue act annotations is defined in a way that expects the specification of a semantic content as the argument of an update function, defined by the 7-tuples used in DiAML, namely as a mechanism for updating the dialogue participants' information states with that content. From a semantic point of view, it is thus fairly straightforward to extend DiAML with the semantic content of dialogue acts. Moreover, when DiAML is used in a dialogue system, the way in which semantic content is specified can be customized for the system's application domain.

The marking up of semantic content would mean in the concrete syntax the introduction of a <semanticContent> element which can be used e.g. for the improved annotation of rhetorical relations as follows (annotating B's utterance in example (12a)) :

(13)
```
<dialogueAct xml:id="da2" target="#s2" sender="#b" addressee="#a"
   dimension="task" communicativeFunction="answer"
   functionalDependence="#da1">
  <dialogueAct xml:id="da3" target="#s3" sender="#b" addressee="#a"
   dimension="task" communicativeFunction="inform" >
  <event xml:id="e3" target="#s3" type="ill" />
  <semanticContent dialogAct="#da3" content="#e3"/>
  <DRLink rel="cause" reason="#e3" result="#da2" />
```

The <event> element introduced in (13) for specifying information about the semantic content of a dialogue act could be the same as, or a simplified version of, the element with the same name that is used in the ISO standards for time and events (ISO 24617-2, see also Pustejovsky et al., 2010), for annotating semantic roles (ISO 24617-4, see also Bunt & Palmer, 2013), and for spatial information (ISO 24617-7, see also Pustejovsky et al., 2013), and that has also been proposed for the annotation of modality (Lapina & Petukhova, 2017) and quantification (Bunt et al., 2017). This suggests that the introduction of <semanticContent> and <event> elements, with their underlying abstract syntax and semantics, may open the possibility to specify quite detailed information about the semantic content of dialogue acts.

7 Conclusions and Perspectives

In this paper we have considered the requirements for a revision of the ISO standard for dialogue act annotation. One of the requirements is that, where possible, a second edition should be downward compatible with the original (current) version of the standard. The notion of compatibility between annotation schemes was analysed and related to the properties of extensibility, restrictability, and optionality.

Applying the ISO 24617-2 scheme in various use cases, such as the creation of the DBOX corpus (Petukhova et al., 2014) and the ADELE corpus (Gilmartin et al., 2017), and the design of the Virtual Debate Coach (Malchanau et al., 2017) show that it would be convenient to add Task Management and Contact Management to the ISO dimensions, as well as certain communicative functions for more fine-grained annotation of feedback, social obligations management, and discourse structuring.

Limitations of ISO 24617-2 were brought to light by the development of ISO standard 24617-6 for discourse relation annotation, of which rhetorical relations between dialogue acts or their semantic contents are a special case. The possibility was discussed to import elements from DR-Core into the annotation scheme for dialogue acts and to optionally add provisions for indicating the semantic content of a dialogue act. Doing so could be a step towards a more general merging of elements from annotation schemes for different semantic information, such as time and events, spatial information, semantic roles and quantification.

Acknowledgement

Thanks are due to the participants in a two-day meeting in April 2018 where ideas for a possible revision of ISO 24617-2 was discussed, including (besides the authors of this paper) also Pierre Albert, Shammur Chowdhury, Andrei Malchanau, and Kars Wijnhoven.

References

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

Alexandersson, J., B. Buschbek-Wolf, T. Fujinami, E. Maier, N. Reithinger, B. Schmitz, and M. Siegel 1998 Dialogue acts in Verbmobil-2. Second edition. *Report 226*. DFKI Saarbrücken.

Andersson, A., M. Bader, E. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, Miller, F. Sotillo, H. Thompson, and R. Weinert (1991) The HCRC Map Task Corpus. *Language and Speech* 34, 351-366.

Baggio, P. , Pelachaud, C., Peter, C. and Zovato, E. (2014) Emotion Markup Language (EmotionML) 1.0. W3C Recommendation, 22 May 2014, edited by F. Burkhardt and M. Schröder. http://www.w3.org/TR/emotionml/

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

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

Bunt H. (2009) The DIT^{++} taxonomy for functional dialogue markup. In: D. Heylen, C. Pelachaud, and D. Traum (eds) *Proceedings of EDAML/AAMAS Workshop "Towards a Standard Markup Language for Embodied Dialogue Acts"*, Budapest, pp. 13–24. Available (with updates) at http://dit.uvt.nl

Bunt, H. (2010) A methodology for designing semantic annotation languages exploiting syntactic-semantic iso-morphisms. In A.Fang, N. Ide, and J. Webster (eds.) *Proceedings of ICGL 2010, Second International Conference on Global Interoperability for Language Resources*, Hong Kong, pp. 29-45.

Bunt, H. (2011) The Semantics of Dialogue Acts. In *Proceedings* 9^{th}*International Conference on Computational Semantics (IWCS 2011)*, Oxford, pp. 1-14.

Bunt, H. (2014) A Context-change Semantics for Dialogue Acts. In Harry Bunt, Johan Bos and Stephen Pulman (eds) *Computing Meaning, Vol. 4*, Berlin: Springer, pp. 177-201.

Bunt, H. (2015) On the principles of semantic annotation. In *Proceedings 11th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-11)*, London, pp. 1-13.

Bunt, H. (2016) The DialogBank. In *Proceedings 10th International Conference on Language Resources and Evaluation (LREC 2016)*, Portoroz, Slovenia,

Bunt, H. (2017) Towards Interoperable Annotation of Quantification. In *Proceedings 13th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-13)*, Montpellier.

Bunt, H. and M. Palmer (2013) Conceptual and Representational Choices in Defining an ISO Standard for Semantic Role Annotation. In *Proceedings 9th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-9)*, Potsdam, pp. 45-54.

Bunt, H. and R. Prasad (2016) ISO DR-Core (ISO 24617-8): Core concepts for the annotation of discourse relations. In *Proceedings 12th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-12)*, Portoroz, Slovenia, pp. 45-54.

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* 8^{th} *International Conference on Language Resources and Evaluation (LREC 2010)*, Malta, pp. 2548-2558.

Bunt, H, J. Alexandersson, J.-W. Choe, A.Fang, K. Hasida, K. Lee, V. Petukhova, A. Popescu-Belis, L. Romary, and D. Traum (2012) A semantically-based standard for dialogue annotation. In: *Proc.* 8^{th} *Intern. Conference on Language Resources and Evaluation (LREC 2012)*, Istanbul. ELRA, Paris.

Bunt, H., V. Petukhova, D. Traum, J. Alexandersson (2016) Dialogue Act Annotation with the ISO 24617-2 Standard. In: D. Dahl (ed.) *Multimodal Interaction with W3C Standards*, Springer, Berlin, pp. 109-135.

Carletta, J., A. Isard, S. Isard, J.Kowtko and G. Doherty-Sneddon (1996) HCRC dialogue structure coding manual. *Technical Report HCRC/TR-82*, University of Edinburgh.

Dhillon, R., S. Bhagat, H. Carvey, and E. Shriberg (2004) Meeting recorder project: dialogue labelling guide, ICSI Technical Report TR-04-002.

Gilmartin, E., B. Spillane, M. O'Reilly, C. Saam, K. Su, B.R. Cowan, K. Levacher, A.Calvo Devesa, L. Cerrato, N. Campbell, and V. Wade (2017) Annotation of Greeting, Introduction, and Leavetaking in Text Dialogues. In *Proc. ISA-13*, Montpellier.

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

ISO 24610 (2006) *Language Resource Management: Feature structures*. International Standard. International Organisation for Standardisation ISO, Geneva.

ISO 12620 (2009) *Terminology and other language and content resources – Specification of data cateories and management of a Data Category Registry for language resoures*. International Standard. International Organisation for Standardisation ISO, Geneva.

ISO 24612 (2010) *ISO 24612: Language resource management: Linguistic annotation framework (LAF)*. International Organisation for Standardisation ISO, Geneva.

ISO 24617-1 (2012) *ISO 24617-1: Language resource management – Semantic annotation framework – Part 1: Time and events.* International Organisation for Standardisation ISO, Geneva.

ISO 24617-2 (2012) *ISO 24617-2: Language resource management – Semantic annotation framework – Part 2: Dialogue acts.* International Organisation for Standardisation ISO, Geneva.

ISO 24617-4 (2014) *ISO 24617-4: Language resource management – Semantic annotation framework – Part 4: Semantic roles.* International Organisation for Standardisation ISO, Geneva.

ISO 24617-6 (2016) *ISO 24617-6: Language resource management – Semantic annotation framework – Part 6: Principles of semantic annotation.* International Standard. International Organisation for Standardisation ISO, Geneva.

ISO 24617-7 (2015) *ISO 24617-7: Language resource management – Semantic annotation framework – Part 7: Spatial information (ISOspace).* International Organisation for Standardisation ISO, Geneva.

Jurafsky, D., E. Schriberg, and D. Biasca (1997) *Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation: Coders Manual*, Draft 13. University of Colorado, Boulder.

Keizer, S., H.Bunt and V. Petukhova (2011) Multidimensional Dialogue Management. In A. van den Bosch and G. Bouma (eds.) *Interactive Multi-modal Question Answering*, Springer, Berlin, pp. 57-86.

Lapina, V. and V. Petukhova (2017) Classification of Modal Meaning in Negotiation Dialogues. In *Proceedings 13th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-13)*, Montpellier.

Malchanau, A., V. Petukhova, and H. Bunt (2017) 'Virtual Debate Coach design: Assessing multimodal argumentation performance.' In *Proceedings International Conference om Multimodal Interaction ICMI 2017*.

Petukhova, V., C. A. Stevens, H. de Weerd, N. Taatgen, F. Cnossen, A. Malchanau (2016) Modelling Multi-issue Bargaining Dialogues: Data Collection, Annotation Design and Corpus. *LREC 2016.*

Petukhova, V. and A. Malchanau and H. Bunt (2014) Interoperability of Dialogue Corpora through ISO 24617-2-based Querying. In *Proceedings 9^{th} International Conference on Language Resources and Evaluation (LREC 2014*, Reykjavik, pp. 4407-4414.

Petukhova, V., Gropp, M., Klakow, D., Eigner, G., Topf, M., Srb, S., Moticek, P., Potard, B., Dines, J., Deroo, O., Egeler, R., Meinz, U., Liersch, S., Schmidt, A. (2014) The DBOX Corpus Collection of Spoken Human-Human and Human-Machine Dialogues. In *Proceedings 9^{th} International Conference on Language Resources and Evaluation (LREC 2014*, Reykjavik.

Petukhova, V. and L. Prévot, and H. Bunt (2011) Discourse Relations in Dialogue. In: *Proceedings 6th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-6)*, Oxfor, pp. 80-92.

Pustejovsky, J., J. Moszkowicz,, and M. Verhagen (2013) A linguistically grounded annotation language for spatial information. *Traitement Automatique des Langues* 53: 87-113.

Pustejovsky, J., K. Lee, H. Bunt, and L. Romary (2010) ISO-TimeML: An International Standard for Semantic Annotation. In: *Proceedings Seventh International Conference on Language Resources and Evaluation (LREC 2010)*, Malta. ELRA, Paris, pp. 394–397.

Wijnhoven, K. (2016) Annotation Representations and the Construction of the DialogBank. MA Thesis, Tilburg University.

Appendix A Dimensions and Communicative Functions in ISO 24617-2:2012

The table below lists the 56 communicative functions defined in ISO 24617-2.

Table 1: ISO 24617-2 communicative functions

General-Purpose	Dimension-Specific Communicative Functions	
Communicative Functions	Function	Dimension
Inform	AutoPositive	Auto-Feedback
- Agreement	AutoNegative	
- Disagreement	AlloPositive	Allo-Feedback
- - Correction	AlloNegative	
- Answer	FeedbackElicitation	
- - Confirm	Stalling	Time Management
- - Disconfirm	Pausing	
Question	Turn Take	Turn Management
- Set-Question	Turn Grab	
- Propositional Question	Turn Accept	
- - Check-Question	Turn Keep	
- Choice-Question	Turn Give	
Request	Turn Release	
- Instruct	Self-Error	Own Communication Man.
- - Address Offer	- Retraction	
- - - Accept Offer	- - Self-Correction	
- - - Decline Offer	Completion	Partner Communication Man.
Suggest	Correct Misspeaking	
Address Suggest	Interaction Structuring	Discourse Structuring
- Accept Suggest	- Opening	
- Decline Suggest	Init-Greeting	Social Obligations Man.
Offer	Return Greeting	
- Promise	Init-Self-Introduction	
Address Suggest	Return Self-Introduction	
- Accept Suggest	Apology	
- Decline Suggest	Accept Apology	
	Thanking	
	Accept Thanking	
	Init-Goodbye	
	Return Goodbye	