INTERNATIONAL STANDARD

ISO
24617-2

Language resource management — Semantic annotation framework (SemAF) —

Part 2:
Dialogue acts

Gestion des ressources langagières — Cadre d'annotation sémantique (SemAF) —

Partie 2: Actes de dialogue

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 24617-2:2012
https://standards.iteh.ai/catalog/standards/sist/a92d0e8e-c304-421f-b4d8-
c2c1a252f181/iso-24617-2-2012

COPYRIGHT PROTECTED DOCUMENT

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 24617-2 was prepared by Technical Committee ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 4, *Language resource management*.

ISO 24617 consists of the following parts, under the general title: *Language resource management — Semantic annotation framework*:

— *Part 1: Time and events (SemAF-Time, ISO-TimeML)*

— *Part 2: Dialogue acts*

The following parts are under preparation:

— *Part 3: Named entities (SemAF-NE)*

— *Part 4: Semantic roles (SemAF-SRL)*

— *Part 5: Discourse structure (SemAF-DS)*

— *Part 6: Principles of semantic annotation (SemAF-Basics)*

— *Part 7: Spatial information (ISO-Space)*

— *Part 8: Semantic relations in discourse (SemAF-DRel)*

Language resource management — Semantic annotation framework (SemAF) —

Part 2:
Dialogue acts

1   Scope

This part of ISO 24617 provides a set of empirically and theoretically well-motivated concepts for dialogue annotation, a formal language for expressing dialogue annotations — the dialogue act markup language (DiAML) — and a method for segmenting a dialogue into semantic units. This allows the manual or automatic annotation of dialogue segments with information about the communicative actions which the participants perform by their contributions to the dialogue. It supports multidimensional annotation, in which units in dialogue are viewed as having multiple communicative functions. The DiAML language has an XML-based representation format and a formal semantics which makes it possible to apply inference to DiAML representations.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

This part of ISO 24617 specifies data categories for reference sets of communicative functions and dimensions of dialogue analysis and provides principles and guidelines for extending these sets or selecting coherent subsets of them. Additionally, it provides guidelines for annotators and annotated examples. It is applicable to spoken, written and multimodal dialogues involving two or more participants.

2   Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 12620:2009, *Terminology and other language resources — Specification of data categories and management of a Data Category Registry for language resources*

ISO 24610-1:2006, *Language resource management — Feature structures — Part 1: Feature structure representation*

ISO 24612:2011, *Language resource management — Linguistic annotation framework*

3   Terms and definitions

For the purposes of this document, the following terms and definitions apply.[1]

---

1)   In this document, "he", "him" and "his" are used in a generic sense, without implying any gender-related distinctions.

3.1
addressee
dialogue (3.5) *participant* (3.13) oriented to by the *sender* (3.18) in a manner to suggest that his *utterances* (3.22) are particularly intended for this participant and that some response is therefore anticipated from this participant, more so than from the other participants

Note to entry: This definition is a *de facto* standard in the linguistics literature. It has been slightly modified here, in replacing "speaker" by "sender" and avoiding the use of ambiguous pronouns. Goffman's original definition says: "dialogue participant oriented to by the speaker in a manner to suggest that his utterances are particularly intended for him and that some response is therefore anticipated from him/her, more so than from the other participants".

[SOURCE: Goffman (1981).]

3.2
allo-feedback act
feedback act (3.8) where the *sender* (3.18) elicits information about the *addressee*'s (3.1) processing of an *utterance* (3.22) that the sender contributed to the *dialogue* (3.5) or where the sender provides information about his perceived processing by the addressee of an utterance that the sender contributed to the dialogue before

EXAMPLE   A: Now move up.

              B: Slightly northeast you mean?

              A: Slightly yeah.

A performs an allo-feedback act signalling that he thinks B understood his first utterance correctly.

3.3
auto-feedback act
feedback act (3.8) where the *sender* (3.18) provides information about his own processing of an utterance (3.22) contributed to the *dialogue* (3.5) by another *participant* (3.13)

EXAMPLE   B's utterance in the example dialogue fragment in (3.2) signals that he is uncertain whether he understood the previous utterance correctly.

3.4
communicative function
property of certain stretches of communicative behaviour, describing how the behaviour changes the *information state* (3.12) of an understander of the behaviour

Note to entry: A communicative function may be "qualified", i.e. one or more **qualifiers** (3.14) may be associated with it. For example, an answer may be qualified as "uncertain" and the acceptance of a request may be "conditional". See 10.3 for explanation and examples.

3.5
dialogue
exchange of *utterances* (3.22) between two or more persons or artificial conversational systems

3.6
dialogue act
communicative activity of a *dialogue* (3.5) *participant* (3.13), interpreted as having a certain *communicative function* (3.4) and *semantic content* (3.16)

Note to entry: A dialogue act may also have certain *functional dependence relations* (3.10), *rhetorical relations* (3.15) and *feedback dependence relations* (3.9) with other units in a *dialogue* (3.5).

3.7
dimension
class of *dialogue acts* (3.6) that are concerned with a particular aspect of communication, corresponding to a particular category of semantic content

EXAMPLE   Dialogue acts advancing the task or activity that motivates the dialogue (the Task dimension), dialogue acts providing and eliciting feedback (the Auto- and Allo-Feedback dimensions) and dialogue acts for allocating the speaker role (the Turn Management dimension).

Note to entry: See Clauses 5, 7 and 9 for discussion and more examples.

3.8
feedback act
dialogue act (3.6) which provides or elicits information about the *sender*'s (3.18) or the *addressee*'s (3.1) processing of something that was uttered in the dialogue

Note to entry: Two classes of feedback are distinguished in this part of ISO 24617: *allo-feedback acts* (3.2) and *auto-feedback acts* (3.3).

3.9
feedback dependence relation
relation between a *feedback act* (3.8) and the stretch of communicative behaviour whose processing the act provides or elicits information about

EXAMPLE   In the example that accompanies definition 3.2, both the allo-feedback act expressed by utterance 3 and the auto-feedback act expressed by utterance 2 have a feedback dependence relation to utterance 1.

3.10
functional dependence relation
relation between a given *dialogue act* (3.6) and a preceding dialogue act on which the semantic content of the given dialogue act depends due to its *communicative function* (3.4)

EXAMPLE   The relation between an answer and the corresponding question, such as between utterance 3 and utterance 2 in the example accompanying definition 3.2, or the relation between the acceptance of an offer and the corresponding offer.

Note to entry: A dialogue act, A2, may also depend on another dialogue act, A1, occurring earlier in a dialogue because of relations between their semantic contents, e.g. because A2 contains a reference to an element occurring in A1. This is not a functional dependence relation, since it is not due to A2's communicative function.

3.11
functional segment
minimal stretch of communicative behaviour that has one or more *communicative functions* (3.4)

EXAMPLE   The functional segment corresponding to the answer given by S in the following dialogue fragment does not include the parts *"Just a moment please"* and *".... let me see..."* but only the parts "*the first train to the airport on Sunday morning is*" and "*at 5:45*":

1. U: What time is the first train to the airport on Sunday morning please?

2. S: Just a moment please... the first train to the airport on Sunday morning is .... let me see... at 5:45.

Note 1 to entry: A consequence of this definition is that functional segments may be discontinuous, may overlap or be embedded and may contain parts contributed by different participants.

Note 2 to entry: The condition of being "minimal" ensures that functional segments do not include material that does not contribute to the expression of a communicative function that identifies the segment.

3.12
information state
context
totality of a *dialogue* (3.5) *participant*'s (3.13) beliefs, assumptions, expectations, goals, preferences, hopes and other attitudes that may influence the participant's interpretation and generation of communicative behaviour

3.13
participant
person or artificial agent involved in the exchange of *utterance*s (3.22)

3.14
qualifier
predicate that can be associated with a *communicative function* (3.4)

EXAMPLE   A: Would you like to have some coffee?

              B: Only if you have it ready.

B's utterance accepts A's offer under a certain condition; this can be described by qualifying the communicative function Accept Offer with the predicate "conditional".

Note to entry: See 10.3 for more examples.

3.15
rhetorical relation
relation between two *dialogue acts* (3.6), indicating a pragmatic connection between the two or between their *semantic content*s (3.16)

EXAMPLE 1   The statement in the second utterance which follows provides a *motivation* for the question in the first utterance:

A: Can you tell me what flights there are to Sydney on Saturday? I'd like to attend my mother's 80th birthday.

EXAMPLE 2   A rhetorical relation between the semantic contents of two dialogue act occurs in the following, where the content of B's statement mentions a *cause* for the content of A's statement:

A: I can never find these stupid remote controls

B: That's because they don't have a fixed location

Note to entry: Relations such as *elaboration, explanation, justification, cause* and *concession* have been studied extensively in the analysis of (monologue) text, where they are often called "rhetorical" relations or "discourse relations" and are mostly viewed either as relations between text segments or as relations between events or propositions, described in text segments. See, for example, Hovy and Maier, 1992, Lascarides & Asher, 2007 or Mann & Thompson, 1988. Many of these relations also occur in dialogue, either as relations between dialogue acts or between the semantic contents of dialogue acts.

3.16
semantic content
information, situation, action, event or objects that a stretch of communicative behaviour refers to

3.17
semantic content category
semantic content type
kind of information, situation, action, event or objects that form the *semantic content* (3.16) of a *dialogue act* (3.6)

EXAMPLE   The various *dimensions* (3.7) defined in this part of ISO 24617 correspond to categories of semantic content. In particular, the Task dimension corresponds to the category of task-specific actions and information; the Allo- and Auto-Feedback dimensions correspond to the categories of information about the processing by the current speaker or by the addressee, respectively, of something that was said before; the Turn Management dimension corresponds to the category of information about the allocation of the speaker role and so forth.

3.18
sender
dialogue (3.5) *participant* (3.13) who produces a *dialogue act* (3.6)

3.19
speaker
sender (3.18) of a *dialogue act* (3.6) in the form of speech, possibly combined with nonverbal communicative behaviour

Note to entry: A dialogue participant may say something while another participant occupies the *speaker role* (3.20), therefore the term "speaker" is not synonymous with "participant who occupies the speaker role".

3.20
speaker role
role occupied by a *dialogue* (3.5) *participant* (3.13) who has temporary control of the *dialogue* and speaks for some period of time

[SOURCE: DAMSL Revised Manual.]

3.21
turn unit
stretch of communicative activity produced by one *participant* (3.13) who occupies the *speaker role* (3.20), bounded by periods where another participant occupies the speaker role

3.22
utterance
anything said, written, keyed, gesticulated or otherwise expressed

Note to entry: An utterance is mostly a part of what a sender contributes in a turn unit.

4 Purpose and justification

The notion of a dialogue act plays a key role in the analysis of spoken and multimodal dialogue, as well as in the design of spoken dialogue systems and embodied conversational agents. These activities all depend on the availability of dialogue corpora, annotated with dialogue act information.

Over the years a variety of dialogue act annotation schemes have been developed, such as those of the TRAINS human-computer dialogue project (Allen et al., 1994), the Map Task studies of human-human dialogue (Carletta et al., 1996) and of the Verbmobil speech translation project (Alexandersson et al., 1998). These schemes were developed for specific purposes and application domains. They contain overlapping sets of concepts and make use of often mutually inconsistent terminology, sometimes employing different terms for the same concept or the same term for different concepts.

The multidimensional DIT scheme (Bunt, 1984) was developed for information-seeking dialogues without depending on a particular domain. The DAMSL scheme (Dialogue Act Markup using Several Layers, Allen and Core,1997; Core et al., 1998) constitutes an application-independent multidimensional annotation scheme. The DIT^{++} scheme (Bunt, 2006; 2009) combines the DIT scheme with concepts from DAMSL and other more recent schemes into a comprehensive general-purpose annotation scheme.

In the EU-funded project LIRICS (Linguistic Infrastructure for Interoperable Resources and Systems, Romary et al., 2007) a reference set of dialogue acts, taken from the DIT^{++} taxonomy, was defined in the form of data categories, following ISO 12620. This set of concepts has been tested for its usability and coverage a) in the manual annotation of spoken dialogues in English, Dutch and Italian and b) in the automatic annotation of spoken and multimodal dialogue in English and forms a significant part of the background of this part of ISO 24617.

The main purpose of this part of ISO 24617 is to define a reference set of domain-independent basic concepts for dialogue act annotation, plus a formal language, based on XML, for representing such annotations. Guidelines are provided for how to use the defined concepts and the annotation language, supported by extended examples. This formal language, the Dialogue act markup language (DiAML) has a formal semantics, which makes it possible to apply techniques for automatic reasoning to DiAML annotations.

Guidelines and principles are also provided for extending the set of concepts defined in this part of ISO 24617, for example, with domain-specific concepts, as well as for selecting coherent subsets.

5   Basic concepts and metamodel

The term "dialogue act" is often used rather loosely in the sense of a speech act used in dialogue. Indeed, the idea of interpreting communicative behaviour in terms of actions, such as questions, promises and requests, goes back to speech act theory (Austin, 1962; Searle, 1969). But where speech act theory is primarily an action-based approach to meaning within the philosophy of language, dialogue act theory is an empirically-based approach to the computational modelling of linguistic and nonverbal communicative behaviour in dialogue.

Dialogue acts offer a way of characterizing the meaning of communicative behaviour in terms of update operations, to be applied to the information states of participants in the dialogue; this approach is commonly known as the "information-state update" or "context-change" approach — see e.g. Bunt (1989; 2000a); Traum and Larsson (2003). For instance, when an addressee understands the utterance "*Do you know what time it is?*" as a question about the time, then the addressee's information state is updated to contain (among other things) the information that the speaker does not know what time it is and would like to know that. If, by contrast, it is understood that the speaker is reproaching the addressee for being late, then the addressee's information state is updated to include (among other things) the information that the speaker *does* know what time it is. Distinctions such as that between a question and a reproach concern the *communicative function* of a dialogue act, which is one of its two main components. The other main component is its *semantic content*, which describes the objects, properties, relations, situations, actions or events that the dialogue act is about. The communicative function of a dialogue act specifies how an addressee should update his information state with the information expressed in the semantic content when he understands the dialogue act.

A dialogue act as defined in this part of ISO 24617 (3.6) is a semantic unit of communicative behaviour. Dialogue act annotation is the marking up of stretches of dialogue with information about the dialogue acts performed in these stretches and is often limited to assigning communicative function tags. A dialogue act being a semantic unit in communicative behaviour, the question arises as to which stretches of communicative behaviour are considered as corresponding to dialogue acts. Spoken dialogues are traditionally segmented into *turns*, defined as stretches of communicative behaviour produced by one speaker, bounded by periods of inactivity of that speaker. Turns in this sense can be quite long and complex and are therefore not very useful units of behaviour for assigning communicative functions to. Communicative functions can be assigned more accurately to smaller units, which are called *functional segments* and which are defined as the minimal stretches of communicative behaviour that are functionally relevant. See Clause 8 for more details about dialogue segmentation.

Inherent to the notion of a dialogue act is that there is an agent who produces the dialogue act, called the "sender" and one or more agents who are addressed, called "addressees". Dialogue studies often focus on two-person dialogues, in which case the dialogue acts have only one addressee. Besides sender and addressee(s), there may be various types of side-participants who are present but do not or only marginally participate (see Clark, 1996).

Dialogue act annotation is often limited to assigning communicative functions to dialogue segments, which corresponds intuitively to indicating the type of communicative action that is performed. A semantically more complete characterization also provides information about the *type of semantic content.* The DAMSL annotation scheme distinguishes three categories of semantic content: task, task management and communication, which indicate whether the semantic content of the dialogue act is concerned with performing the task which underlies the dialogue or with discussing how to perform the task or with the communication. The DIT++ scheme distinguishes a number of subcategories of communication-related information, such as feedback information, turn allocation information and topic progression information. The various categories of semantic content are also called "dimensions" and are discussed in more detail in Clause 7.

Some types of dialogue acts are inherently dependent for their full meaning on one or more dialogue acts that occurred earlier in the dialogue. This is, for example, the case for answers, whose meaning is partly determined by the question being answered and for the acceptance or rejection of offers, suggestions, requests and apologies. The following example illustrates this, where the meaning of (1.3) clearly depends very much on whether it is an answer to the question (1.1) or to the question (1.2).

EXAMPLE 1

(1.1)   B: Do you know who's coming tonight?

(1.2)   B: Which of the project members d'you think will be there?

(1.3)   A: I'm expecting Jan, Alex, Claudia and David, and maybe Olga and Andrei.

As an answer to (1.1), it says that nobody else is expected to come than the people that are mentioned, but as an answer to (1.2) it leaves open the possibility that other people will come, who are not members of "the project".

For dialogue acts which have such a dependence on other dialogue acts, due to their responsive character, the marking up of the links to these "antecedent" dialogue acts allows the annotation not just to express e.g. that the utterance is an answer, but also to express *to which question* it is an answer. This type of relation between dialogue acts is called a *functional dependence relation*.

Dialogue acts may also be semantically related through other relations, as shown in the following example:

EXAMPLE 2

(2.1)   A: It ties you on in terms of the technology and the complexity that you want

(2.2)   A: like for example voice recognition

(2.3)   A: because you might need to power a microphone and other things

(2.4)   A: so that's one constraint there

In this example[2] we see a sequence of four functional segments contributed by the same participant. Segment (2.2) is related to the initial statement through an *Exemplification* relation and (2.3) through an *Explanation* relation, while (2.4) is related to the preceding three segments through a *Summarization* relation. Such relations are known as *rhetorical relations*. In view of the wide diversity of the sets of rhetorical relations that have been proposed (see, e.g., Mann and Thompson, 1988; Hovy and Maier, 1993; Sanders et al., 1992), this part of ISO 24617 does not propose any specific set of such relations, but only provides a conceptual category for which a particular set of relations may be specified.

Feedback-providing and eliciting acts also relate to what happened earlier in the dialogue, but in a different way. They are concerned with the processing of what was said before — such as its perception or its interpretation:

EXAMPLE 3

(3.1)   A: Is this flight also available on Thursday?

(3.2a)   B: On Thursday you said?

(3.2b)   B: The twelfth you mean?

With utterance (3.2a), B checks whether he heard correctly what A said. This is a response to A's *utterance*, rather than to the dialogue act that the utterance expresses; with utterance (3.2b), by contrast, B checks whether he has correctly interpreted what A said. Both types of dependence are called a *feedback dependence relation*.

Note that nonverbal feedback, for instance in the form of nodding or vocal backchannels like "*uh-huh*", "*um*", "*huh*", "*m-hm*", may have a feedback dependence relation to what is being said at that moment, rather than to what was said before. This is also the case for speech editing acts like self-corrections ("*on Tuesday I mean Thursday*") and completions of what the partner is trying to say.

Example 1 above also illustrates another phenomenon that is frequently found in dialogue, namely that speakers may have incomplete or uncertain information. The use of "*maybe*" in (1.3) expresses that A is uncertain about part of the information that he provides.

---

2)   From the AMI corpus, see http://corpus.amiproject.org.

In addition, speakers may express a certain sentiment about the information or event that is being discussed, as in (4.2) or express a reservation in the form of a condition, as in (4.3), where an offer is conditionally accepted:

EXAMPLE 4

(4.1)   A: Would you like to have some coffee?

(4.2)   B: That would be great, thank you!

(4.3)   B: Only if you have it ready.

For the annotation of conditions, uncertainty and sentiment, this part of ISO 24617 makes use of so-called *function qualifiers*, which can be attached to communicative functions — see 10.3 for more detail.

The above characterization of the notion of a dialogue act makes use of the following key concepts, which form the backbone of the metamodel for dialogue act annotation in Figure 1:

a)   sender, addressee and participants in other roles (side-participants);

b)   functional segment;

c)   dialogue act, communicative function, communicative function qualifier and semantic content category (or "dimension");

d)   functional dependence relation, rhetorical relation and feedback dependence relation.
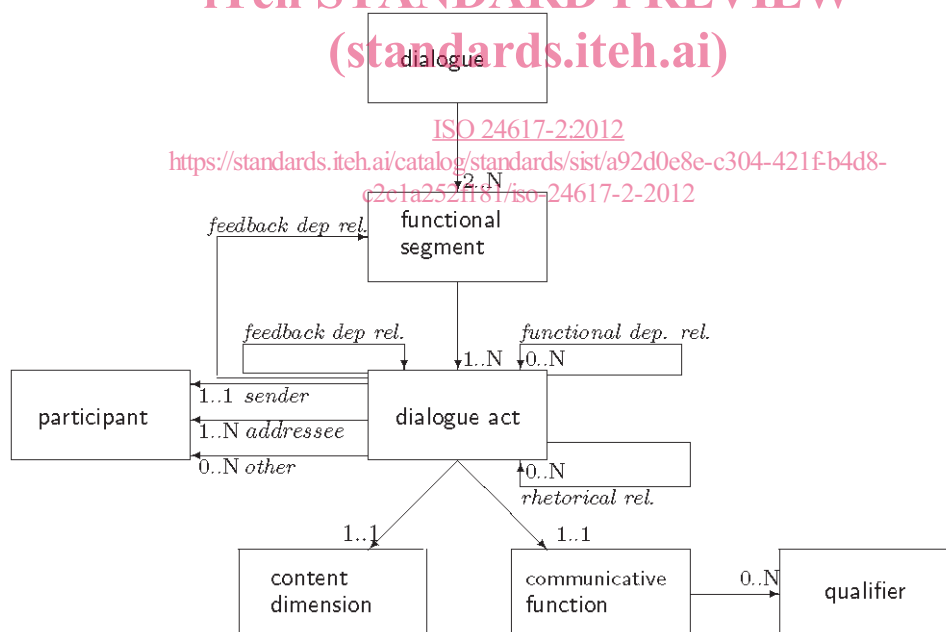
Figure 1 — Metamodel for dialogue act annotation

6   Definition of communicative functions

Existing dialogue act annotation schemes use one of the following two approaches to defining communicative functions or a combination of the two: (1) in terms of the effects on addressees intended by the sender; (2) in terms of properties of the signals that are used. Defining a communicative function by its linguistic form has the advantage that its recognition can be straightforward, but runs into the problem that the same linguistic form can be used to express different functions. For example, the utterance "*Why don't you start?*" has the form of a question and can be intended as such, but can also be used to invite or encourage somebody to

start. Similarly for so-called "declarative questions" (questions in the form of a declarative sentence), like "*You're going home tomorrow*", which are intended as questions although they look like statements.

Form-based definitions also run the risk of being purely descriptive, rather than semantic. For example, when a speaker repeats something that was said before, this behaviour may be characterized as a repetition; however, that would only say something about the *form* of the behaviour, nothing about its communicative function. A repetition for instance often has a feedback function, as in (5.2a) but it can also have other functions, as in (5.3), where it is used as a confirmation in response to a check question:

EXAMPLE 5

(5.1)   S: There are evening flights at seven-fifteen and eight-thirty

(5.2a)   C: Seven-fifteen and eight-thirty

(5.2b)   C: And that's on Sunday too

(5.3)   S: And that's on Sunday too

This part of ISO 24617 follows a strictly semantic approach to the definition of communicative functions. But while linguistic form is taken not to be part of the definition of a communicative function, a requirement for introducing a communicative function is that there are ways in which a sender can indicate that his behaviour should be understood as having that particular function, by shaping his (linguistic and/or nonverbal) behaviour so as to have certain observable features which are indicative for that function in the context in which the behaviour occurs. This requirement puts all communicative functions on an empirical basis.

A particular case where form and function are not related in a straightforward way is that of indirect speech acts, where a speaker uses a linguistic form that is standardly used to express one type of dialogue act, but in context means something else. Questions of the form *Do you know* [*X*] are illustrative: while an utterance of this form would standardly seem to ask an addressee whether he possess the knowledge [*X*], it is more often used to request the addressee to provide the information [*X*], if possible. This makes such a question a conditional request.

The full complexity of the phenomenon of indirect speech acts is beyond the scope of this part of ISO 24617, but an important class of indirect speech acts can be covered by qualifying them as conditional — see 10.3.

7   Annotation schemes

7.1 Structure of annotation schemes

Existing dialogue act annotation schemes can be divided into one-dimensional and multidimensional schemes. One-dimensional schemes have a set of mutually exclusive tags and are used for coding stretches of dialogue with a single tag. Multidimensional schemes, on the other hand, are intended for encoding stretches of dialogue with multiple tags. Schemes of the latter kind typically have a relatively large tag set. There are several advantages to the structuring of such a tag set into clusters of communicative functions tags:

— Clustering semantically related tags improves the transparency of the tag set, as each cluster is concerned with a certain kind of information. This also makes the coverage of the tag set clearer, since each cluster typically corresponds to a certain class of dialogue phenomena.

— A structured tag set can be searched more systematically and more "semantically" (i.e. on the basis of semantic differences and similarities) than an unstructured one.

— The tags within a cluster are usually mutually exclusive; this has the advantage that, once a particular tag has been assigned, the rest of the tags within that cluster do not need to considered any further. If a cluster is hierarchically organized, as is the case in this part of ISO 24617, with finer-grained functions being dominated by less fine-grained ones (such as "confirmation" being more fine-grained than

"answer"), then the most sensible use of these tags is to choose the most specific tag for which there is sufficient evidence.

7.2 Multidimensionality and multifunctionality

Participation in a dialogue involves several activities beyond those strictly related to performing the task or activity for which the dialogue is instrumental. In natural conversation, the participants among other things constantly "evaluate whether and how they can (and/or wish to) continue, perceive, understand and react to each other's intentions" (Allwood, 1997). Communication is thus a complex, multi-faceted activity and this is reflected in the multifunctionality that dialogue utterances often exhibit.

Multifunctionality comes in a variety of forms. Allwood (1992) distinguishes between *sequential* and *simultaneous* multifunctionality and provides the following example as an illustration:

EXAMPLE 6   A: Yes! Come tomorrow. Go to the church. Bill will be there. OK?

                B: The church, OK.

Sequential multifunctionality occurs when a turn has several parts which each have a different communicative function. In Example 6 we see A's utterance containing five functional segments, with communicative functions such as *feedback giving, request, request, statement and response elicitation* The occurrence of sequential multifunctionality depends on the way in which a dialogue is segmented (see also Clause 8) and disappears when sufficiently small segments are considered as markables.

Simultaneous multifunctionality, by contrast, persists even when minimal segments are used as markables. The following example illustrates this.

EXAMPLE 7

(7.1)   A: Do you know what date it is?

(7.2)   B: Today is the fifteenth.

(7.3)   A: Thank you.

A's utterance (7.3) has the function of thanking and will mostly be taken to imply that A has understood and accepted the information in (7.2) — i.e. as having a positive feedback function. But *"Thank you"* does not *always* express positive feedback; a participant in an unsuccessful dialogue may just want to terminate the interaction in a polite way. The feedback function of the thanking in (7.3) can be inferred along the following lines: By saying *"Thank you",* A expresses his gratitude to B. This can only be for what B just said; this would constitute a reason for being grateful if A considers B's utterance as relevant and useful, which means that A accepted B's utterance as an answer to his question. The feedback function in such a case can be viewed as a conversational implicature (Grice, 1979), i.e. as a contextually plausible consequence which the addressee is intended to infer.

The implication relation between thanking and positive feedback is different from that between a propositional answer (*"yes"* or *"no"*) and a confirmation, where the relation is one of *entailment*, i.e. an implication which is logically valid. (Every confirmation by its very nature is also an answer.) Entailment relations occur when the definition of one communicative function is a special case of that of another.

It may be argued that such cases should not be considered as instances of multifunctionality, e.g. a speaker who wants to issue a confirmation can hardly have the intention of *additionally* giving an answer, since the recognition of that intention is already part of the recognition of a confirmation.

There are also cases of multifunctionality where the different functions do not have any logical relation. This is, for example, the case for turn-initial hesitations, as in the following dialogue fragment:

EXAMPLE 8

(8.1)   A: Is that your opinion too?

(8.2)   B: Uh,.. well,... I guess so.

In (8.1), speaker A asks a question to B and assigns the turn to B. In (8.2) B performs a stalling act in order to buy some time for deciding what to say; the fact that he starts speaking without waiting until he has made up his mind about what to say, indicates that he accepts the turn. So the segment *"Uh,.. well,..."* is multifunctional, having both a stalling function and a turn-accepting function. Note that A's utterance is also multifunctional: it asks a question about B's opinion and it assigns the turn to B (due to its intonation, in combination with A looking at B and raising his eyebrows).

The design of a dialogue act annotation schema can reflect the multifunctional view of utterances in two ways: 1) by structuring the tag set into clusters (see below); 2) by accompanying instructions to annotators for how to apply multiple tags. If the tag set is fairly extended and does not have any structure, it is next to impossible to formulate good instructions for how to apply multiple tags, since there is no easy way to refer to groups of tags. Therefore, the recognition that utterances in dialogue tend to be multifunctional naturally leads to the introduction of dimensions in a dialogue annotation schema.

7.3 Multidimensionality, clustering and dimensions

The clusters of communicative functions that can be found in existing annotation schemes are typically chosen on the basis of a conceptual similarity of certain functions. An early version of the DIT schema, for example, has a cluster of "information-seeking functions" for a range of question types and a cluster of "information-providing" functions for various kinds of informs and answers (Bunt, 1989).

The DAMSL schema (Core and Allen, 1997) is organized into "layers" and "dimensions". Four layers are distinguished: communicative status, information level, and forward looking and backward looking communicative functions (FLF and BLF); the latter two are indeed clusters of communicative functions (the tags in the other layers are concerned with other kinds of information). The FLF cluster is subdivided into five clusters, including the classes of commissive and directive functions, well known from speech act theory. The BLF cluster has four subclasses: Agreement, Understanding, Answer and Information Relation. Core & Allen (1997) refer to these nine subclasses as "dimensions".

Popescu-Belis (2005) mentions six aspects of utterance function as relevant for choosing dimensions: 1) the traditional clustering of illocutionary forces in speech act theory into representatives, commissives, directives, expressives and declarations; 2) turn management; 3) adjacency pairs; 4) topical organization in conversation; 5) politeness functions; 6) rhetorical roles.

Bunt (2005; 2006) proposes to structure a multidimensional tag set by basing the notion of dimension on the observation that participation in a dialogue involves a range of communicative activities other than those for advancing the task. Dialogue participants share information not only about the task that is pursued but also about the processing of each other's messages, about the allocation of turns, about contact and attention, about the use of time and about various other aspects of the interaction. They thus perform communicative activities of various types, such as giving and eliciting feedback, taking turns, stalling for time, establishing contact and showing attention. Each of these types of activity is concerned with a different category of information. This part of ISO 24617 uses the term "dimension" to refer to these various semantic content categories or to the communicative activities concerned with these content categories. This leads to dimensions such as feedback, turn management, time management and contact management in addition to the dimension formed by the task that motivates the dialogue. Clause 9 describes the set of dimensions defined in this part of ISO 24617.

7.4 Dimension- specific and general-purpose functions

Not every grouping of communicative functions qualifies as a dimension in the sense of this part of ISO 24617. For example, the group of information-giving acts (statements, warnings, answers, confirmations and so on) does not form a dimension, since information can be given about any aspect of the dialogue, such as the underlying task, feedback, change of topic or contact. Information-giving acts are thus not specifically related to a particular category of information. The same is true of information-seeking acts (open questions, check questions, menu questions and so on) and of the commissive and directive acts (request, suggest, instruct, offer, promise and so on), which can be about any kind of action. These clusters of functions therefore do not qualify as dimensions; since these functions can be combined with any kind of information or

11