
A Computational Approach to the Interpretation of Indirect Speech Acts

Robbert-Jan Beun, Rogier M. van Eijk, John-Jules Meyer, Nieske Vergunst
Utrecht University, Department of Information and Computing Science
P.O. Box 80089, NL-3508 TB Utrecht, The Netherlands
Email: {rj,rogier,jj,nieske}@cs.uu.nl

An Indirect Speech Act (ISA) is an utterance that conveys a message that is different from its literal meaning, often for reasons of politeness or subtlety. The DenK-system provides us with a non-compositional way to look at Indirect Speech Acts that contain modal verbs. We can extract the non-literal meaning from these Indirect Speech Acts without having to consider the meaning of the modal verbs themselves. The system uses a special ‘verb-construction’ to model the composition of the ISA w.r.t. modal verbs, subject of the sentence, attitude or communicative verbs and the sentence-type. It then makes use of a collection of speech act assignment tables to look up the intended meaning of the ISA.

Keywords: natural language processing, indirect speech acts, modal verbs, human-computer dialogue.

1 INTRODUCTION

The relation between the linguistic form of an utterance and its illocutionary force is a complex matter and the subject of a large body of literature (e.g., [1], [4], [5]). Although surface features such as WH-words, prosodics and the grammatical form of the sentence may considerably contribute to the determination of the speech act, the relation is also determined by contextual cues, such as the participants' physical abilities. So, for instance, the sentence:

(1) ‘Can you switch on the computer?’

can be, depending on the circumstances, interpreted in a direct sense as a *question* about someone's physical abilities or in an indirect sense as a *request* to actually switch on the computer. This depends largely on the context of the utterance.

The aim of this paper is to describe a computational framework for the interpretation of the illocutionary act of utterances that contain modal verbs, such as ‘can’, ‘would’ and ‘must’. The presented framework was developed in the context of the DenK-system [3], a dialogue system that contains information about an electron microscope domain. The architecture of the system reflects a natural dialogue situation where a user has direct access to the domain of discourse (by pointing or direct manipulation) or indirectly by symbolic means (speech acts). The system presents itself as a cooperative dialogue partner and is able to communicate in natural language.

In this paper we will first briefly introduce the DenK-system and its components. Subsequently, we will present the inner workings of the system by discussing the verb-construction, which is a framework for representing indirect speech acts, and the collection of tables the system uses to look up the intended illocutionary act of the utterance. Finally, a conclusion and some future research topics will be presented.

2 THE DENK-SYSTEM

The DenK-system consists of three entities: the *application domain* (a software simulation of an electron microscope), a *digital cooperative assistant* and a single *user*. The user has direct access to the domain, but can also instruct the assistant to manipulate the electron microscope. Both the user and the assistant can observe and manipulate the microscope; the user and the assistant communicate with each other in typed natural language. Communication modalities can be combined, as the user is able, for example, to point to something in the application domain while using a deictic expression (e.g., ‘this’, ‘here’) in an utterance directed to the assistant, like: (*while pointing to a button*) ‘What happens when I push *this* button?’ In principle, the assistant considers itself an expert on the electron microscope domain. This implies that the user cannot add new domain information and, consequently, all declaratives about the subject domain are considered as questions.

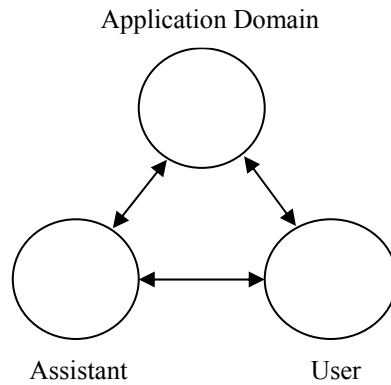


Fig. 1. The triangle metaphor.

The DenK-system incorporates a module that can interpret Indirect Speech Acts (ISA's) in a pragmatic way. If an utterance by the user contains a modal verb or the verb 'want', the Assistant will use a special parsing system and a collection of speech act assignment tables to look up the desired interpretation. The interpretation mechanism of the assistant detects particular characteristics of the utterance that contributes to the interpretation of the speech act, such as sentence type, modal verb or the actor (assistant or user). Based on a combination of these characteristics, the assistant decides about the illocutionary force of the utterance. So, for instance, example 1 contains the modal verb 'can', has a declarative form and the assistant is the actor of the action that follows the modal ('switch on the microscope') (see [2]). The literal meaning is thus abandoned and the system will look up the meaning of the utterance in the table for the modal verb 'can'. The utterance will be interpreted as a request for action, in this case the action of switching on the computer. In the next sections, we will treat in more detail how the system reaches this conclusion.

Although the linguistic expressions uttered by the user may take an enormous variety of surface forms, the system distinguishes only a limited number of illocutionary acts: RI (request for information), RA (request for action) and RI_POS (request for possibility).

3 SPECIAL VERBS AND THE VERB-CONSTRUCTION

In order to determine the illocutionary force of the utterance, the system uses special verbs that pertain to relations between the dialogue partners and the domain information. We distinguish between *modal auxiliaries* (can, will, etc.), *attitude verbs* (believe, know, desire, etc.) and *communicative verbs* (tell, explain, describe, ask, etc.). Moreover, attitude verbs are subcategorised in so-called *information attitudes* (know) and *pro-attitudes* (want). Other verbs, which pertain to the domain, are called *domain verbs*.

Utterances by the user are transformed into so-called verb-constructions, represented in BNF-notation:

```

VERB_CONS(partner, verb1, polarity, partner, verb2, sentence_type, domain_info)
partner      ::= A | U | nil
verb1       ::= modal | want | nil
modal       ::= can | could | may | must | might | shall | should | will | would
polarity    ::= pos | neg
verb2       ::= tell | ask | know | nil
sentence_type ::= dec | int
domain_info ::= method | prop
  
```

The *partner* (A for Assistant, U for User) occurs twice in this framework: the first time as the subject of *verb1*, the second time as the (implicit) subject of *verb2*. *Verb1* is the modal auxiliary or the verb want. If no modal verb or want occurs in the sentence, this field will be nil. Furthermore, it should be noted that only in the case that *verb1* is want there can be a difference between both occurrences of *partner* in the verb-construction ('I want to...' vs. 'I want you to...'). The values pos and neg in the field *polarity* mean that the sentence is, respectively, positive or negative.

Verb2 is the second verb in the sentence; we use tell, ask and know as prototypical verbs to represent the verb categories mentioned above, but they can be replaced with others; e.g., 'believe' = know.

If no such verb occurs in the sentence, *verb2* is either non-existent or a domain verb and represented as ‘nil’ in the verb-construction. The field *sentence_type* represents the sentence type: declarative (*dec*) or interrogative (*int*). Imperative sentences are not included in this system, because they do not occur in combination with modal verbs (e.g., sentences like ‘Can switch the microscope on!’ are grammatically incorrect). Finally, *domain_info* represents information about the electron microscope in a type theoretical formalism. We make a distinction between methods (e.g., ‘switch on the microscope’) and (*open*) propositions *prop* (e.g., ‘which state the microscope is in’). The content of the type theoretical formula itself is not relevant for the speech act interpretation of the sentence; we will not treat it further in this article.

Because the verb-construction was developed specifically for the DenK-system, we assume that it only parses utterances that are directed to the system, in a single-user situation. Therefore, the word ‘you’ in the utterance to be parsed is always the assistant (A) and ‘I’ is always the user (U). The verb-construction enables us to represent Example 1 as `VERB_CONS(A, can, pos, nil, nil, int, method)`.

Other examples of sentences with their representation in the verb-construction are:

- (2) ‘You can tell me how to switch on the microscope.’ → `VERB_CONS(A, can, pos, A, tell, dec, method)`
- (3) ‘May I ask you in which state the microscope is in?’ → `VERB_CONS(U, may, pos, U, ask, int, prop)`
- (4) ‘I want to know which button controls the contrast.’ → `VERB_CONS(U, want, pos, U, know, dec, prop)`
- (5) ‘I know which button controls the contrast.’ → `VERB_CONS(nil, nil, pos, U, know, dec, prop)`

Note that in case of a communicative verb, only the actor of this verb is represented (see Examples 2 and 3, where the actors are, respectively, Assistant and User). Finally, these features and their values have to be converted into the three speech act types discussed in the previous section. In order to do this, a number of tables have been developed in which the interpreter can look up the intended speech act of an utterance, depending on the parameters of the verb-construction.

4 SPEECH ACT ASSIGNMENT

In the DenK-system, tables were developed for all modal verbs. In these tables, the columns represent the different types of *VP*, which consists of *verb2* and *domain_info*. Distinctions are made between information-attitude verbs (*know*), communicative verbs (*tell* and *ask*) and domain verbs (*nil*). Resulting interpretations are RI (request for information), RI_POS (request for possibility) or RA (request for action). ‘*’ indicates that the utterance is pragmatically inadequate; utterances of this type are not expected, but if they do occur, the Assistant will respond ‘What do you mean?’.

The desired interpretations of sentences of the form ‘*X can VP*’ (e.g., ‘You can switch on the microscope.’) and ‘*Can X VP?*’ (e.g., ‘Can you tell me which state the microscope is in?’) are summarized in Table 1. In this table, we have indicated, for instance, that a combination of ‘can’ with an information-attitude verb sometimes yields a pragmatically inadequate interpretation of the speech act. Typical examples of this construction are utterances like ‘Can you believe that the microscope is on?’ and ‘I can know that the microscope is in nP-mode.’ These sentences are grammatically correct, but they do not make sense in the context of the DenK-system and are therefore indicated in the tables as pragmatically inadequate. In practice, the Assistant will respond ‘What do you mean?’ to these utterances.

Table 1. Speech act interpretation of `VERB_CONS(partner1, can, pos, partner2, verb2, sentence_type, domain_info)`, where *verb2* is not *nil* and *partner1* equals *partner2*.

CAN	tell		ask		know	
	method	prop	method	prop	dec	int
U	RA	*	RA	RI	*	RI
A	*	RI	*	*	RI	*

Some typical examples of utterances with their speech act assignment:

(6) ‘You can tell me how to switch on the microscope.’
 VERB_CONS(A, can, pos, A, tell, dec, prop) Classified as RI.

(7) ‘Can I ask you to switch on the microscope?’
 VERB_CONS(U, can, pos, U, ask, int, method) Classified as RA.

(8) ‘Can I ask you which state the microscope is in?’
 VERB_CONS(U, can, pos, U, ask, int, prop) Classified as RI.

Note that for the verbs ‘tell’ and ‘ask’, no distinction is made between the sentence types `dec` and `int`, and for the verb `know`, there is no distinction between `method` and `prop` as domain information. This is because there is also no distinction between the speech act interpretations of sentences with this information. There is also no row in the table for `nil` as *partner*, because `nil` (a domain object) cannot occur in a sentence with `tell`, `ask` or `know` as *verb2* (e.g., ‘The microscope can know...’), but only with domain verbs, represented as `nil` in the field *verb2*. There is a separate table for *verb2* = `nil` (not presented in this paper), because in this case there is a distinction between `method` and `proposition` and also between declarative and imperative sentences.

In Table 2 we have presented the speech act assignments of ‘*X should VP*’ and ‘*Should X VP?*’. Again, many combinations are considered as not interpretable (*). The assignment slightly differs from the modal ‘can’, e.g.:

(9) ‘You should tell me to switch on the microscope.’
 VERB_CONS(A, should, pos, A, tell, dec, method) Classified as *.

(10) ‘Should I know which state the microscope is in?’
 VERB_CONS(U, should, pos, U, know, int, prop) Classified as *.

(11) ‘You should know which state the microscope is in.’
 VERB_CONS(A, should, pos, A, know, dec, prop) Classified as RI.

Table 2. Speech act interpretation of VERB_CONS(*partner1*, `should`, `pos`, *partner2*, *verb2*, *sentence_type*, *domain_info*), where *verb2* is not `nil` and *partner1* equals *partner2*.

SHOULD	tell		ask		know	
	method	prop	method	prop	dec	int
U	*	*	RA	RI	RI	*
A	*	RI	*	*	RI	*

Note the differences in Table 2 between the communicative verbs `tell` and `ask`, e.g.:

(12) ‘You should tell me which state the microscope is in.’
 VERB_CONS(A, should, pos, A, tell, dec, prop) Classified as RI.

(13) ‘You should ask me which state the microscope is in.’
 VERB_CONS(A, should, pos, A, ask, dec, prop) Classified as *.

Example 12 is a request for information. Although both examples are grammatically correct, Example 13 is uninterpretable in the context of the DenK-system. The combination of modals and other sentence features in Example 13 could be pragmatically correct in other contextual settings, such as:

(14) An instruction for a (possibly hypothetical) situation in the future: ‘If you need help, *you should ask me.*’

(15) An educative environment: ‘At this point *you should ask me* to push this button while you adjust

- the microscope.’
- (16) A reproach: ‘You insensitive brute! When I come home from work, *you should ask me* how my day was!’

5 CONCLUSIONS AND FUTURE RESEARCH

In this article, we have shown that it is possible to extract the intended meaning from indirect speech acts that contain modal verbs without considering the meaning of the modal verbs themselves. A collection of tables have been developed that can be used to assess the meaning of an ISA based on a small number of characteristics of the sentence. This appears to be a successful way to deal with these indirect speech acts. However, it should be noted that these tables are just a rule of thumb, and highly context-dependent. An ISA can still mean different things in different contexts; in some cases it could, for example, have its literal meaning instead of the pragmatic speech act meaning. As stated before, the verb-construction and the tables have been developed specifically for the electron microscope domain of the DenK-project. The system can be adapted to suit other domains, most importantly by taking into account the context of the utterance. The current system uses only information in the utterance itself to determine its pragmatic interpretation. However, the system could be improved and generalized by researching which aspects of the context have substantial influence on the desired interpretation of the speech act, and incorporating this in the system. This is a difficult task, because the context of a discourse is theoretically infinite; it consists, for instance, of the role of the system and user, the dialogue history, the action history and all other aspects of the situation and surroundings. We could also extend the system to incorporate possibilities for a multi-user system and for also parsing utterances that are not directed to the system, but instead to one of the users.

On the other hand, the simplicity of the current system also has its advantages: it is a fast and easy way to extract the pragmatic meaning of speech acts, which returns the correct result in many cases. Also, making mistakes is not fatal, because of the interactive nature of the system; the system’s interpretation of an utterance can be checked and, if necessary, corrected, by means of verification in the dialogue. The system also asks for clarification when a pragmatically incorrect utterance occurs (indicated by ‘*’ in the tables), to avoid a total communication breakdown when the user says something unexpected.

Another future topic of research is to extend the system to support utterances with a reflexive pronoun, like ‘You should ask yourself whether...’, which is usually intended as something along the lines of ‘Are you sure that...?’ Sentences like this cannot be interpreted correctly in the current system, because the reflexive pronoun is not represented in the current verb-construction. The representation of this sentence would be VERB_CONS(S, should, pos, S, ask, dec, domain_info), the same as sentences of the form ‘You should ask me...’ This utterance is classified as pragmatically incorrect (see Table 2). However, a sentence of the form ‘You should ask yourself whether...’ is actually pragmatically correct and mostly used as a warning or scolding. Because of these ‘emotional’ contexts, it would be interesting to add the dimension of emotions to correctly identify and interpret this utterance.

Another important question is why the tables contain the values as we presented them and how the tables can be generalized into other applications in different contextual settings. In the analysis of the modal verbs, we assume that the modal dimensions ‘possibility’, ‘ability’ and ‘permission’ (and their counterparts ‘necessity’, ‘inability’ and ‘obligation’) play a central role. It is unclear, at this stage of the research, how contextual and utterance features influence the interpretation of the modal verbs in terms of these modal dimensions.

ACKNOWLEDGMENT

This work is supported by SenterNovem, Dutch Companion project grant nr: IS053013.

REFERENCES

- [1] Asher, N. & Lascarides, A. 2001. Indirect Speech Acts. *Synthese*, 128, pp. 183-228.
- [2] Beun, R.J. & Piwek, P. 1996. Pragmatische features in DenK: Pragtags. DenK report 97/29, Eindhoven: SOBU.
- [3] Bunt, H.C., Ahn, R.M.C., Beun, R.J., Borghuis, T. & van Overveld, C.W.A.M. 1998. Multimodal cooperation with the DenK system. In: H. Bunt, R. Beun and T. Borghuis (eds). *Multimodal Human-Computer Communication. Lecture Notes in Computer Science*, vol.1374. Heidelberg: Springer-Verlag, pp. 39-67.
- [4] Levinson, S.C. 1983. *Pragmatics*, Cambridge: CUP, pp. 263-283, pp. 356-364.
- [5] Searle, J.R. 1975. Indirect Speech Acts. In: P. Cole and J.L. Morgan (eds). *Syntax and Semantics, Volume 3: Speech Acts*, Academic Press, pp. 59-82.