Actions, Beliefs and Intentions in Rationale Clauses and Means Clauses*

Cecile T. Balkanski
Aiken Computation Lab, Harvard University
Cambridge, MA 02138
cecile@das.harvard.edu

Abstract
Utterances that include rationale clauses and means clauses display a variety of features that affect their interpretation, as well as the subsequent discourse. Of particular importance is the information that is conveyed about agents' beliefs and intentions with respect to the actions they talk about or perform. Hence, for a language interpretation system to handle these utterances, it must identify the relevant features of each construction and draw appropriate inferences about the agents' mental states with respect to the actions and action relations that are involved. This paper describes an interpretation model that satisfies this need by providing a set of interpretation rules and showing how these rules allow for the derivation of the appropriate set of beliefs and intentions associated with each construction.

Introduction
Utterances that describe or refer to multiple actions display a variety of features that affect their interpretation, as well as the subsequent discourse. Of particular importance is the information that is conveyed about agents' beliefs and intentions with respect to the actions they talk about or perform. For example, although the following utterances are about the same two actions, some of the beliefs and intentions that are communicated may be diametrically opposed:

(1) a. John dirtied the carpet by walking across the room.
   b. John walked across the room to dirty the carpet.

The speaker of (1a) communicates her belief that John's walking across the room resulted in John's dirtying the carpet, but she may or may not be conveying a belief that John actually intended to walk across the room as a way of dirtying the carpet. The speaker of (1b), on the other hand, does express her belief that John intended to walk across the room as a way of dirtying the carpet, but not necessarily that John's walking across the room actually resulted in his dirtying the carpet.

Utterance (1a) includes a means clause and utterance (1b), a rationale clause. As their names suggest, means clauses express the means by which an action is performed, while rationale clauses express the purpose, or rationale, of the main clause action. The above examples show that these types of utterances express similar relations between actions, but that the speaker and performing agent (i.e., the agent whose actions are being described in the utterance) may have different attitudes with respect to these actions. These differences are critical because of their effect on subsequent discourse; for example, continuation (a) is felicitous in (2), below, but not in (3), and continuation (b) is felicitous in (3) but not in (2). They are also important because of their effect on planning and plan recognition. Among those actions that fail, for example, only those that are intended will constrain replanning (Bratman 1987).

(2) John dirtied the carpet by walking across the room, ...
   a. but he didn't even realize it.
      b. ?but that didn't work because his shoes were clean.

(3) John walked across the room to dirty the carpet, ...
   a. ?but he didn't even realize it.
      b. but that didn't work because his shoes were clean.

For a language interpretation system to handle utterances about actions appropriately, it must therefore be able to determine the relationships between actions expressed in multi-action utterances and derive the beliefs and intentions of the speaker and performing agent with respect to these actions and action relationships. This paper describes an interpretation model that addresses this need. This model comprises an axiomatization of basic principles of belief and intention and

---

*This research has been supported by a contract from U S WEST' Advanced Technologies, by the Air Force Office of Scientific Research under Contract No AFSOR-89-0273, and by an IBM Graduate Fellowship.

1Rationale clauses are to be distinguished from purpose clauses, e.g., "Mary bought a suit to wear at the meeting", and infinitival relative clauses, e.g.: "John found the book to give to his sister". One difference among these constructions is that only rationale clauses can be paraphrased using the words "in order to" (Huettner & al 1987).

296 Natural Language: Interpretation
a set of interpretation rules for deriving the action relations expressed in these utterances. It thereby accounts for both the overlap and the difference in meaning between utterances with rationale clauses and utterances with means clauses, as illustrated in (1), (2) and (3). While utterances with rationale clauses may be about sequential or simultaneous actions (Balkanski 1992), this paper focuses on utterances concerning simultaneous actions.

This paper extends a companion paper (Balkanski 1992) in addressing the mental states of the speaker and performing agent and in considering utterances about present and future, as well as past actions. Surprisingly little research has addressed the interpretation of means clauses and rationale clauses. There is a large body of linguistics research on purpose clauses, but it focuses on syntactic aspects of the construction and issues of control (e.g., Bach (1982), Jones (1991)). The planning literature provides very useful theories of action and action relations, which will be referred to later in this paper. In the computational linguistics literature, Huettner's work on the generation of adjunct clauses (Huettner & al 1987) and Di Eugenio's analysis of instructions (Webber & Di Eugenio 1990; Di Eugenio 1992) both examine purpose constructions, but from different perspectives than that of this paper. Huettner focuses on decision making in the generation process, e.g., determining which argument to gap (i.e., delete) when. Di Eugenio uses action relations similar to those presented here, but does not address issues regarding differences in the mental states of the speaker and performing agent, temporal aspects of action occurrences, or the role of contextual conditions, all of which are central to this paper.

**Beliefs and intentions expressed in utterances with rationale and means clauses**

This section examines in more detail the characteristic beliefs and intentions expressed in utterances with rationale clauses and means clauses. These properties, illustrated in (1), (2) and (3), emerged from a detailed analysis of multi-action utterances in task-oriented dialogues (Balkanski 1990) and a subsequent examination of selections from Associated Press news stories.

Utterances with rationale clauses or means clauses are about two actions, the occurrence of one of them possibly generating (Goldman 1970), i.e., resulting in, the simultaneous performance of the other. The generation relation, independently motivated by work in plan recognition (Pollack 1986; Balkanski 1990; Lochbaum 1991), holds of two actions A and B when (a) their agents and times are the same, (b) there is a set of contextual conditions, called generation-enabling conditions (e.g., John's shoes being dirty) that hold during performance time and (c) there is a conditional generation relation between the act-type of A, the act-type of B and these conditions. A conditional generation relation holds among act-types α and β and conditions C if whenever an action of type α occurs while these conditions hold, an action of type β occurs at the same time. In this paper, the generation relation will be notated as GEN(A,B) where A and B are actions, comprising an act-type, agent and time (Balkanski 1990). An important property of this relation is that if it holds of two actions A and B, and A occurs, then so does B. This follows from the generation relation requiring the relevant generation-enabling conditions to hold during performance time.

Utterances with means clauses and rationale clauses, therefore, refer to both a generating and a generated action, performed by the same agent at the same time. In utterances with means clauses (MC), the generating action is expressed in the adjunct clause, whereas in utterances with rationale clauses (RC), it is expressed in the matrix clause. Let A represent the generating action (e.g., John's walking across the room during some time interval) and B the generated action (e.g., John's dirtying the carpet during that time interval). Let S be the speaker, G the performing agent, T, the time of speech, and T_a the time of action (past, present or future). The beliefs and intentions expressed in these two types of utterances are the following:

- **[a]** In both types of utterances, S believes that A occurred (respectively, is occurring, will occur).
- **[b]** In utterances with MCs, but not necessarily those with RCs, S believes that B occurred (is occurring, will occur).
- **[c]** In utterances with MCs, but not necessarily those with RCs, S believes that GEN(A,B).
- **[d]** In the MC case, therefore, but not necessarily in the RC case, S believes there is a conditional generation relation between the act-types of A and B and some set of generation-enabling conditions, and that these conditions held (are holding, will hold).
- **[e]** In utterances with RCs, but not necessarily those with MCs, S believes that A was intended (will be intended) on the part of G.
- **[f]** In utterances with RCs, but not necessarily those with MCs, S believes that B was intended (will be intended) on the part of G.
- **[g]** In utterances with RCs, but not necessarily those with MCs, S believes that A was intended (will be intended) on the part of G as a way of generating B. In the RC case, therefore, but not necessarily in the MC case, S believes that G expects (at the start of action time) the relevant generation-enabling conditions to hold.

The belief in [a] is readily associated with these ut-
terances. Those in [b],[c], [e] and [f] were illustrated in (1), (2) and (3). Although the belief in [d] is likely to hold of utterances with means clauses like (2) (i.e., the walking action was probably intended on the part of John), this need not be the case, as illustrated in (4a) below. Utterance (4b) shows that the use of a rationale clause forces an interpretation in which the speaker believes that the generating action (here the slipping action) is intended on the part of its agent.

(4) a. John broke his arm by slipping on the ice.
   b. John slipped on the ice to break his arm.

The beliefs and intentions listed above are formalized in (5); their truth values with respect to utterances with rationale clauses and means clauses are given in Table 1, where a “1” indicates that the corresponding belief is true, and a “?” that it can be true or false, depending on the context.

(5) a. \(\text{BEL}(S, T_1, \text{OCCUR}(A))\)
   b. \(\text{BEL}(S, T_1, \text{OCCUR}(B))\)
   c. \(\text{BEL}(S, T_1, \text{GEN}(A,B))\)
   d. \(\text{BEL}(S, T_1, \text{INT}(G,\text{start}(T_a),\text{act-type}(A),T_a))\)
   e. \(\text{BEL}(S, T_1, \text{INT}(G,\text{start}(T_a),\text{act-type}(B),T_a))\)
   f. \(\text{BEL}(S, T_1, \text{INT}(G,\text{start}(T_a),\text{act-type}(A) & \text{GEN}(A,B),T_2))\)

<table>
<thead>
<tr>
<th>Beliefs:</th>
<th>[a]</th>
<th>[b]</th>
<th>[c]</th>
<th>[d]</th>
<th>[e]</th>
<th>[f]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utterances with RCs:</td>
<td>1</td>
<td>?</td>
<td>?</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Utterances with MCs:</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

Table 1: Beliefs and intentions expressed in utterances with MCs and RCs

The functions “start” and “act-type” return, respectively, the starting point of a time interval and the act-type of an action (e.g., the act-type “dirty the carpet” for the action of John’s dirtying the carpet during \(T\)). The predicate \(\text{BEL}\) holds of an agent, a time (interval) and a proposition if the agent believes the proposition during that time. The predicate \(\text{OCCUR}\) holds of an action if that action occurs. Because actions have associated times, the \(\text{OCCUR}\) predicate inherits its time from its argument. When the time of action \(A\) is past, then \(\text{OCCUR}(A)\) is true if \(A\) occurred in the past; when the time of \(A\) is present, then \(\text{OCCUR}(A)\) is true if \(A\) is an action currently being performed; finally, when the time of \(A\) is future, then \(\text{OCCUR}(A)\) is true if \(A\) is an action that will necessarily occur in the future. Although it is impossible to determine whether an action will necessarily occur in the future, when used in belief contexts (as in this paper), this predicate makes claims about an agent’s beliefs about the past, present or future occurrence of actions. The predicate \(\text{INT}\) holds of an agent, a time (point) \(T_1\), an act-type, and a time (interval) \(T_2\) if the agent intends at \(T_1\) to perform an action of that type during \(T_2\). Because \(T_1=\text{start}(T_2)\) in [d], [e] and [f], these beliefs are about present-directed intentions (Bratman 1987).

Whether or not \(G\) also had, at an earlier time, a future-directed intention to perform an action of type \(A\) (or \(B\)) is not of concern here. The \(\text{GEN}\) relation in [f] is used as a modifier indicating the way in which an action is performed. That is, \(\text{INT}(G, T_1, \text{act-type}(A) & \text{GEN}(A,B), T_2)\) means that \(G\) intends at \(T_1\) to perform an action of the type of \(A\) during \(T_2\) as a way of generating \(B\) (Lochbaum & al 1990).

The beliefs represented in Table 1 hold for past, present and future actions, as indicated by the verb tenses used in the English descriptions. The examples in the Introduction were about past actions, but the same acceptability judgments are obtained with present and future actions. For example, in uttering (6a), below, \(S\) may have reason to believe that Mary’s keeping Sue up is unintentional, or, in uttering (6b), that John will fail to reset the printer.

(6) a. Mary is keeping Sue up by playing the piano.
   b. John will press the red button to reset the printer.

These beliefs also hold for utterances in which the speaker and the agent are the same person. For example, the speaker of (7a), below, believes that she intended to extend her arm out of the window [d], intended to signal for a left turn [e], and intended to extend her arm as a way of signaling for a turn [f]. She also believes that she extended her arm [a], but not necessarily that in doing so, she signaled for a turn [b,c] (e.g., maybe she knows that the driver behind her could not see her arm). Similar results obtain for present and future actions. For example, the speaker of (7b) may have doubts about her success in signaling for a turn. In utterances with means clauses, e.g. (7c), the speaker may or may not intend to dirty the carpet even if she believes that she is doing so (or will do so).

So beliefs [c] and [f] can be true or false, as desired, despite the speaker’s having beliefs [b] and [e].

(7) a. I extended my arm out of the car window to signal for a left turn.
   b. I am extending (or will extend) my arm out of the window to signal for a turn.
   c. I am getting (or will get) the carpet dirty by walking across the room.

The goal of the analysis presented in the remainder of this paper is to derive the beliefs and intentions given in Table 1. These beliefs and intentions will be derived on the basis of the logical form of an utterance, a set of axioms about belief and intention, and interpretation rules defining the meaning of the relevant linguistic constructions.

---

2We are also not concerned with the performing agent’s intentions during action time, since when an agent is actually performing an action \(A\), we no longer say that he intends to \(A\) (Bratman 1987).

3This distinction corresponds to the difference between doing \(A\) intentionally and intending to do \(A\) (Bratman 1987).
Processing framework

This section presents the logical forms that are input to the interpretation model, and a set of axioms about belief and intention that are necessary for the interpretation process.

Logical forms

Logical forms represent the literal meaning of an utterance and are derived compositionally by semantic interpretation rules based on the syntactic structure of the utterance. We present here the main aspects of our representations; a detailed description is given in the companion paper (Balkanski 1992). Our logical forms reify actions (Davidson 1967) and are represented as existentially quantified sentences of first-order logic, with predicates that include an additional argument position for the action being described. Multi-action utterances are handled by introducing multiple action variables and by representing connectives like "by" and "to" using two-place predicates ranging over these action variables, as illustrated in (8)\(^6\).

(8) John will dirty the carpet by walking across the room.

LF: \(\exists x_1, x_2, \text{dirty}(x_2) \land \text{agt}(x_2, \text{John}) \land \text{obj}(x_2, \text{Carpet}) \land \text{future}(x_2) \land \text{walk}(x_1) \land \text{loc}(x_1, \text{Room}) \land \text{by}(x_2, x_1)\)

b. John will walk across the room to dirty the carpet.

LF: \(\exists x_1, x_2, \text{walk}(x_1) \land \text{agt}(x_1, \text{John}) \land \text{loc}(x_1, \text{Room}) \land \text{future}(x_1) \land \text{dirty}(x_2) \land \text{obj}(x_2, \text{Carpet}) \land \text{to}(x_1, x_2)\)

Every action argument is represented as a separate predicate to avoid having either to introduce existential variables for missing arguments or to determine how these arguments are recovered during the process of building the logical form. As will be shown later in this paper, the missing agent and time arguments of the adjunct clause actions in the logical forms in (8) will be recovered through the interpretation of the action relations.

Rather than denoting "real" action tokens, action variables in our formalism correspond to action entities in the discourse model that may represent either actual (i.e., "real") actions, namely, actions that have occurred, or unrealized actions, that is, actions that have not yet occurred or might never occur (Balkanski 1991). Quantification is therefore over a universe that contains everything that can be talked about, with no commitment to existence in the real world (as in, e.g., Hobbs (1985), Schubert & Hwang (1990)). Since the mapping from action variables to action entities is not an issue relevant to the main topic of this paper, we assume from here forward that action variables have been replaced by their corresponding action entities.

Axioms

Agents' beliefs are taken to be closed under logical consequence and distributive over conjunctions:

\[\text{Closure: } \text{BEL}(g, t, p) \land \text{BEL}(g, t, p \rightarrow q) \rightarrow \text{BEL}(g, t, q)\]

\[\text{Distributivity: } \text{BEL}(g, t, p_1 \land \ldots \land p_n) \rightarrow \text{BEL}(g, t, p_1) \land \ldots \land \text{BEL}(g, t, p_n)\]

Agents uttering a declarative sentence are assumed to believe the propositional content of that utterance\(^6\):

\[\text{Declarative rule: } \text{UTTER}(s, t, "p") \rightarrow \text{BEL}(s, t, p)\]

where "\(p\)" is a declarative utterance and \(p\) is its LF.

We also adopt an intention axiom, \(\Pi\), assumed to be believed by all agents; that states necessary conditions on an agent's intention to perform an action as a way of generating another. According to \(\Pi\), if \(G\) intends at \(t_1\) to do \(\alpha\) at \(t_2\) as a way of generating \(b\), then he must intend to do \(\alpha\) at \(t_2\) and to do \(\beta\) at \(t_2\), and he must believe that the two actions are related by generation (Pollack 1986).

\[\Pi: \text{INT}(g, t_1, \alpha \land \text{GEN}(a, b), t_2) \rightarrow\]

\[\text{INT}(g, t_1, \alpha, t_2) \land \text{INT}(g, t_1, \beta, t_2) \land \text{BEL}(g, t_1, \text{GEN}(a, b))\]

where \(\alpha = \text{act-type}(a)\) and \(\beta = \text{act-type}(b)\)

Interpretation

The process of interpreting a logical form consists in applying interpretation rules to the various predicates of that logical form. This section presents these rules and describes the way in which the appropriate beliefs and intentions expressed in utterances with rationale and means clauses are derived on the basis of them. To emphasize the fact that reasoning takes place in belief contexts, interpretation rules are embedded in belief predicates.

Interpretation rules

The inference rules defining the meaning of means clauses and rationale clauses are given below. They formalize those given in the companion paper (Balkanski 1992) and extend them in a number of ways, as discussed below. The "by" and "to" predicates in the left hand side of the rules are the logical form predicates representing the NL connectives introducing means clauses and rationale clauses. Note that the order of the arguments in the "by" and GEN predicates is reversed.

---

\(^6\) A more complete treatment of declaratives would require a complex theory of speech acts (e.g., Cohen & Levesque (1990), Perrault (1990)); these topics, however, lie beyond the scope of this paper.

---

Balkanski 299
The LF1 rule maps the LF representation of means clauses to the generation relation, capturing the fact that the speaker of an utterance with a means clause believes that the occurrence of the main clause action follows from that of the adjunct clause action. The second conjunct of this rule specifies the speaker’s belief of the occurrence of the generating action. This occurrence is indeed a feature of the means clause construction and not of the form of the verb, namely a gerund; as explained below, the occurrence of an action may be derived from certain tensed verbs.

The LF2 rule maps the LF representation of these constructions to the performing agent’s intention to perform the main clause action as a way of generating the adjunct clause action. By virtue of intention axiom 11, it then follows that the speaker of an utterance with a rationale clause believes that the agent believes (believed, or will believe) that these actions are related by generation (in addition to believing that the agent intended (or will intend) to perform both actions). Given that the speaker may believe that the agent’s beliefs are incorrect, the generation relation in these utterances is, in a sense, only potential. This is as desired since, as discussed earlier, the speaker does not necessarily believe that a generation relation actually holds between the two actions that are referred to in the utterance.

In the companion paper (Balkanski 1992), we discussed the potentiality of the generation relation in utterances with rationale clauses, but did not propose a way of capturing it. In this paper, casting the analysis in the context of the beliefs and intentions of the speaker and performing agent allows for a very simple and elegant treatment of this aspect of rationale clauses.

The form of a verb may sometimes provide information about the occurrence of the associated action. For example, a past tense (action) verb indicates the speaker’s belief that the corresponding action occurred. Similarly, a present or future tense verb indicates the speaker’s belief of the present or future occurrence of the associated action. Tensed verbs are represented in the logical form by the predicates past, present or future, as illustrated in (8). The following inference rule captures the fact that these predicates express assertions about action occurrences:

**LF3**: $\text{BEL}(g, t, \text{past}(a) \lor \text{present}(a) \lor \text{future}(a) \rightarrow \text{OCCUR}(a))$

**Applying the rules**

This section illustrates the interpretation process by applying the interpretation rules and axioms presented in the preceding sections to the sample utterances and logical forms in (8), thereby showing how the beliefs and intentions recapitulated in Table 1, are derived. As mentioned earlier, action variables are assumed to have been replaced by action constants (let $x_1$ and $x_2$ be replaced by A and B respectively). The logical forms under consideration are therefore conjunctions of ground literals.

By the Declarative rule, the speaker, S, believes these logical forms and by the Distributivity rule, S believes each conjunct of those propositions, and in particular the one expressing the action relation, namely:

(9) **MC case**: $\text{BEL}(S, T, \text{by}(B,A))$

**RC case**: $\text{BEL}(S, T, \text{to}(A,B))$

The beliefs in (10) then follow from (9) and the interpretation rules LF1 and LF2 (along with the Closure and Distributivity axioms):

(10) **MC case**: $\text{BEL}(S, T, \text{GEN}(A,B)) \land \text{BEL}(S, T, \text{OCCUR}(A))$

i.e., beliefs [c] and [a] from Table 1.

**RC case**: $\text{BEL}(S, T, \text{INT}(\text{agt}(A), \text{start}(\text{time}(A)), \text{act-type}(A) \land \text{GEN}(A,B), \text{time}(A)))$

i.e., belief [f] from Table 1.

In the RC case, it follows from S’s belief of intention axiom 11 that S also believes that the agent of A believed (believes, or will believe) that A and B are related by generation:

(11) **RC case**: $\text{BEL}(S, T, \text{BEL}(\text{agt}(A), \text{start}(\text{time}(A)), \text{GEN}(A,B)))$

Because S only believes that the performing agent believed (believes or will believe) a GEN relation between A and B, S herself may or may not believe that relation, depending on whether or not S believes the agent’s beliefs are correct. This is as desired, given the “?” for belief [c] in Table 1.

By the Declarative and Distributivity rules, S also believes the conjuncts in the logical forms that express the tense of the matrix clause verbs, namely:

(12) **MC case**: $\text{BEL}(S, T, \text{future}(B))$

**RC case**: $\text{BEL}(S, T, \text{future}(A))$

It then follows from these beliefs and the interpretation rule LF3 that S believes the corresponding action occurred:

(13) **MC case**: $\text{BEL}(S, T, \text{OCCUR}(B))$

i.e., belief [b] from Table 1.

**RC case**: $\text{BEL}(S, T, \text{OCCUR}(A))$

i.e., belief [a] from Table 1.

In the MC case, S’s belief about the occurrence of B, the generated action, can also be derived from beliefs [a] and [c] (see (10)), on the basis of the definition of generation. In the RC case, however, this belief cannot be derived from sentential information alone: neither from the form of the verb (the logical form does not
include future(B)), nor from the interpretation of the construction (i.e., LF2). This is as desired, given the “?” in Table 1 for belief [b].

Looking now at intentions, nothing can be derived in the MC case about the agent’s intention to perform either A or B, independently, or one as a result of the other, as desired given the “?” for beliefs [d], [e] and [f] in Table 1. In the RC case, given that S believes that the agent of A intends to perform A as a way of generating B (see (10)), it follows from S’s belief of intention axiom II that S also believes that the agent of A intends both to perform A and to perform B.

We have now shown how each belief in Table 1, along with its associated truth value, is derived. In addition, the present analysis also accounts for the missing agent and time arguments of the adjunct clause actions in (8a) and (8b). Because the generation relation requires the agents and times of the generating and generated actions to be identical, these missing arguments can be recovered on the basis of the agent and time information associated with the actions expressed in the main clause of the utterances.

Conclusion

This paper presented an interpretation model that accounts for the particular features of utterances with means clauses and rationale clauses. The model achieves this goal by deriving the appropriate set of beliefs and intentions of the speaker and performing agent regarding the actions and action relations expressed in these utterances.

This model is being refined and extended in a number of directions. We are analyzing utterances with rationale clauses and means clauses embedded in negated contexts and modal contexts. Initial analysis of these utterances shows that the interpretation rules can account for them as well as they are applied within the embedding context. For example, given the utterance “John wanted to dirty the carpet by walking across the room”, the model should derive Want(John, GEN(A,B) \land OCCUR(A)) from LF1 and Want(John, OCCUR(B)) from LF3, where A is the action of John’s walking across the room (during some time interval) and B that of John’s dirtying the carpet (during the same time interval). We are also examining the meaning of the “by” and “to” connectives in contexts other than means clauses and rationale clauses in order to integrate the interpretation rules presented here into an interpretation model of wider scope. Finally, we are investigating further additions to our theory of intention and working on a Prolog implementation of our interpretation model.

Acknowledgments.

I would like to thank Barbara Grosz, Karen Lochbaum and Stuart Shieber for many helpful discussions and comments regarding this paper.

References


