Uptake and Joint Action

Joris Hulstijn *

Vrije Universiteit, Amsterdam

and

Nicolas Maudet **

LAMSADE, Université Paris-Dauphine

Abstract

In natural language dialogue, the way a responder 'takes up' the initiative of a participant, largely influences the further course of the dialogue. This uptake mechanism can be understood as a negotiation at a meta level: an initiative counts as a bid of a dialogue game; an appropriate response counts as an acceptance of the bid. We propose to extend this account of uptake to other conventional joint activities besides dialogue. We show that for an uptake mechanism to be effective, a joint activity must be characterisable in terms of initiatives and responses, with projection rules that indicate what initiatives count as a bid for a joint action, and appropriateness rules that indicate what responses count as appropriate.

Key words: joint action, dialogue

* Department of Economics and Business Administration, Vrije Universiteit, De Boelelaan 1105, 1081 HV, Amsterdam, The Netherlands, jhulstijn@feweb.vu.nl
**LAMSADE, Université Paris-Dauphine, 75775 Paris Cedex 16, France, maudet@lamsade.dauphine.fr

Preprint submitted to Elsevier Science

1 Introduction

There has been a lot of interest in the nature of joint activities. Tuomela has given a detailed account of what a joint action is, and how it can be understood in terms of mutual beliefs and joint intentions (Tuomela, 2000, 2004). Clark (1996) has been advocating a view in which communication itself is seen as number of coordinated joint actions by speaker and hearer, at different levels of analysis. Joint action has been studied extensively in the area of collaborative problem solving and multi-agent systems, see e.g. Grosz and Kraus (1996). Although these models are quite detailed in specifying minimal conditions for a joint action, it is unclear how joint activities are decided on in practice. How are the joint intentions and mutual beliefs that are needed for joint action, in fact established?

We give part of an answer to this question, by analogy to natural language dialogue. We suggest that, at least for certain types of conventional joint activities, participants engage in a joint action by means of a mechanism called *uptake* (Austin, 1962; Clark, 1996). In natural language dialogue, one can distinguish conventional dialogue games, that determine the possible ways the dialogue may proceed. Which dialogue game will be played, is determined in a joint process of offering and accepting moves. There is usually one participant who takes the initiative; other participants respond to the initiative. The way in which an initiative is 'taken up', allows the responder to modify the course of a dialogue.

In previous work we have given a preliminary analysis of uptake in terms of conditional obligations (Hulstijn and Maudet, 2003). Following Kreutel and Mann (2003) and Reed (1998), we propose that uptake can be understood as a negotiation of the type of dialogue game at a meta level. Each initiative can be interpreted as a 'bid' for a certain game. By convention, an appropriate response is taken as an acceptance of that bid. Because initiatives are open as to what game they start, the responder has a large influence on the way

an interaction will develop. To allow for this process, we predict that each dialogue game must be expressed using two sets of rules. *Projection rules* help participants predict the subsequent dialogue game, on the basis of an initial sequence of moves. *Appropriateness rules* determine what moves are considered an appropriate response. There are essentially two kinds of conditional obligations. First, when a dialogue game bid is made, to either accept or reject that bid, and second, once a dialogue game has started, to give responses that are considered appropriate.

In this paper we extend this uptake approach to joint action in general. We show that for the type of conventional joint activities that we are interested in, projection rules and appropriateness constraints can indeed be found. Under the right circumstances, to be detailed below, certain actions count as an initiative, or bid, to start a joint activity of a certain type. This is specified in the projection rules. Usually, an appropriate reaction – specified by appropriateness constraints – will count as an acceptance of the initiative. We show that uptake can indeed be generalised, and be seen as a mechanism to establish all kinds of joint actions on the basis of conventions.

This paper is not primarily cognitive in its stance (the perspective being initially rather descriptive, and logical). However we reckon these views to be complementarity. By elaborating such models, we definitely intend to pave the way for such mutual research efforts. In particular, we are keen to seeing whether the assumptions underlying our approach could be backed by cognitive evaluation.

The rest of the paper is organised as follows. In section 2 we review the definitions of joint action according to Clark and to Tuomela, and consider some typical examples of joint activities. In section 3 we present our account of the uptake mechanism for dialogue. In section 4 we then consider whether a mechanism like uptake can also be found for other conventionalised joint activities, and under what conditions projection and appropriateness rules can be found.

2 Joint Action

In this section we review the following question: what properties make a number of simultaneous actions by different agents *joint*? We start from Clark's account of joint action (Clark, 1996). After that, we extract some characteristics from Tuomela's work, and discuss a number of typical examples of joint action.

2.1 Requirements for Joint Action

Typical examples of joint action are playing a duet or carrying a piano. Carrying a piano assumes coordination between the carriers: lifting and lowering must be simultaneous, otherwise someone's back might get hurt. Similarly, voices in a duet need to be coordinated, otherwise the resulting chords are off key. So coordination is a necessary condition for joint actions. But it is not sufficient. This is illustrated by the following example. Suppose Ann and Ben are students at the music academy who live in the same student house. Ann studies the flute and Ben plays the piano. Both have to audition at the end of the week, on the same piece, a duet. Since they are rehearsing the whole week, it would not be strange if by accident, they start playing their pieces simultaneously. If both students use a metronome set at the same tempo, to a bystander who walks along the opened windows, it will sound like Ann and Ben are playing a duet. Yet, we would not say that a duet – a joint action – is played.

(1) a. Ann and Ben are playing a duet.

Ann is playing the flute as part of (1a).

Ben is playing the piano as part of (1a).

b. It appears to a bystander that Ann and Ben are playing a duet.

Ann is playing the flute, adagio

Ben is playing the piano, adagio.

Coordinated actions form a joint action, when they are carried out with the intention of them constituting a joint action. This only works if the participants have a common plan that indicates what the separate actions to be achieved are, and which agent should do what. In this example, both the plan and the role-assignment are indicated by the musical score. So this suggests a first requirement for joint action: agents must have a commonly known plan that assigns actions to agents.

Joint action seems paradoxical. People's individual actions are motivated by intentions. You can only intend your own actions; you can certainly not intend other people's actions. So it seems that even though a group can carry out a joint action, it cannot intend it! The paradox is solved because individual actions can be part of a joint action. Such individual actions are special: they are not autonomous but coordinated. Therefore they are called *participatory actions* (Clark, 1996). Participatory actions are distinguished from individual actions by the intentions behind them. This suggests a second requirement for joint action: each participatory action must be intended to be part of the joint action.

But why do we need to reduce group intentions to individual intentions? Why not allow group intention as a separate notion? For example Searle (1990) proposes so called *we-intentions*. One of the problems with group intention has to do with the intimate relationship between intention, accountability

and punishment. People are only accountable for the consequences of actions they commit intentionally. A crime suspect who was acting as part of a group, is only accountable for his share in the group action and for the fact that he did nothing to stop the others. In other words, being an accessory to a crime is a participatory action. This indicates a third requirement for joint action: participants should actively support the other participants in their efforts.

If group intention were recognised, we would have to punish the whole group, as a group. And this does happen, for instance when a whole team gets banned, because one of the athletes used unlawful stimulants. Note that a team is recognised by the sports association as a unity. This suggests that only when there is an institutionalised group structure it makes sense to treat a group as a single agent. Think of a team, a political party, an association or a married couple. In most cases there is no such group structure prior to the joint action; the group is formed in acting jointly.

The discussion suggests the following necessary requirements for joint action.

(i) A joint action requires common knowledge of a plan that defines the participatory actions, a coordination mechanism, and an allocation of participants to participatory actions, that, when executed, will produce the joint action.
(ii) Each of the participants intends to achieve the participatory action, as part of the joint action.
(iii) Each of the participants should actively support the other participants in their participatory actions.

That means that we have to indicate how common knowledge of a plan and a coordination mechanism is established, how the individual and joint intentions are established, and how the obligation for mutual support comes about.

Tuomela made several studies of cooperativity and joint action (Tuomela, 2000). The following list of research questions is taken from a recent version of his ideas (Tuomela, 2004, p.3).

1. The type of the joint performance must be the right kind of thing for a joint action, such as the carrying of a table. This is established by convention.
2. Given the joint action type, its division into parts and their allocation to participants must be accounted for. It needs also to be shown that coordinated performance of the parts will indeed generate the results of the joint action.
3. Intentional joint action requires that it be performed in accordance with and in part because of the agents' relevant joint intention.
4. The participants must have joint control over the performance of the joint action and hence, at least to some extent, over the results.
5. Each of the participants depends on the others to succeed. Success of the participatory actions, both causally and conceptually depends on the coordinated performance of other participatory actions. The nature of such dependencies among participants, and among the part performances must be known.
6. An account of joint actions of different 'conceptual strength' is to be given. Tuomela distinguishes between the strong *we-mode joint action*, namely joint acting as a group and weak or *I-mode joint action*. Such actions are dependent, but do not require joint intention. For example, I can use your slides in a course, although you may have never intended the slides to be used by me.
7. There are many different contexts in which participants act as a group. For example, there can be a prior agreement and thus a social obligation to act jointly, or there can only be a joint plan that is formed on the occasion. Thus, we need to establish exactly the kind of context in which

joint action occurs.

We believe that Tuomela's characteristics 1–7 are complementary to requirements (i)–(iii) requirements established earlier. Note that these may have been influenced by earlier work of Tuomela and others.

2.3 Case Studies

In the discussion above we referred to the example of a duet. Here we add some more illustrative cases.

(2) Consider Ann and Ben now rehearsing the duet together. After some discussion of what went wrong in the previous try, Ann starts again at the fifth bar of the 3rd movement. Ben will join as soon as he has recognised the melody. The initiative is here to start playing again, and where to start. The uptake is signalled by joining at the right point.

(3) Consider the well-known case of lifting a table. The participants and the task allocation are apparent from the table itself. Every agent takes the corner nearest to him or her. After mutual eye glances, it is clear that everybody is ready. Now, as soon as one agent starts lifting, this is meant as an initiative, which is 'taken on' by the others by lifting too, or rejected by someone shouting "Wait!".

(4) Consider a school camp where someone starts to sing a song; the others will join as soon as they have recognised the tune. The initiative here is to sing, and which song it is. Uptake is signalled by joining in the right song.

(5) Consider the football tactic called 'double pass' (figure 1). Suppose player A is in possession. A passes the ball to B and continues to run towards the goal. A defender or goal keeper K will now focus on B. B passes back to A, who has by now passed K and can shoot at will. Here the initiative by A is to start this tactic, and the direction in which to go. Acceptance

Fig. 1. Double Pass

Fig. 2. Collision avoidance

is signalled by B, by indeed passing back to the place where A will be after running. Note that anticipation of B on A's behaviour alone can not explain this behaviour. We need the fact that both players know that they both know the tactic.

(6) Consider sailing on a lake (Figure 2). In case there is a collision danger, a responsible skipper will bear away. The skipper will often temporarily exaggerate the change of course to let the other skipper know which avoiding strategy was chosen. This counts as an initiative. The other skipper can 'take up' the initiative by bearing away in the opposite direction. This mechanism applies to a whole range of collision avoidance tactics.

What can we learn from these examples? Can we deduce from these examples the characteristics of the class of conventional joint activities for which uptake makes sense?

First, before engaging in a joint activity, potential participants are already paying *attention* to each other. They must recognise each other as having a *shared background*, for example based on location, appearance, or on some other more general joint activity.

Second, for recognised potential participants with a shared background, specific circumstances (like rehearsing or being on a school camp) trigger a set of *conventional behaviour patterns* to become salient. Such behaviour patterns have been studied in cognitive science under the name of *schemata, scripts,*

or frames (Schank and Abelson, 1977). Think of the well known example of the restaurant script, that prescribes the different stages of finding a table, ordering, eating and paying in a restaurant. Similar ideas can be found in Goffman's (1974) frame analysis. Frames are basic cognitive structures which guide the perception and representation of (social) reality. In a way, frames structure which parts of reality become noticed. For example, a group of persons lined up in an orderly fashion at the side of a road might evoke the frame of a 'bus queue'. This notion of script or frame is crucial to our approach. Note that even though people get engaged in joint activities voluntarily, and out of their own interests, important elements of a script or frame that regulates a joint activity can be obligatory. For example, after having eaten a dinner in a restaurant, the script prescribes that one must pay. And when joining a queue, one must join at the end.

Third, to engage in a joint activity, one participant needs to take the initiative, by taking a first suggestive action, that must be recognisable as the beginning of a salient coordination pattern or frame. This first action may itself also be called an initiative. Note that the circumstances only make a behaviour pattern salient; not obligatory. There is no norm that forces pupils to sing at a school camp.

Fourth, by taking on the initiative, signalled by an appropriate reaction, participants express a commitment to the joint activity that was initiated. This commitment then forces them to respect the norms that apply to that particular joint activity. For example, while singing a song you are obliged to stay in tune and keep the rhythm. And importantly, when leaving a joint activity if at all, you must make it noticed that you are signing off, thereby indicating that your part is completed, or will not be completed by you [1] .

Fifth, after having engaged in a joint action, the resulting behaviour of the participants is more complex than can be explained on the basis of mere reactive

[1] More about the necessity of obligations can be found in section 5.

coordination, or even on the basis of anticipatory coordination (Castelfranchi, 1998). It requires that other agents are induced to adapt their actions too. In our collision avoidance example, a reactive coordination action would correspond to a skipper avoiding a floating log as soon as the log is noticed. An anticipatory action would correspond to a skipper who uses sailing knowledge to anticipate the other vessel's course, and bear away to avoid a collision. Such behaviour is quite common. But it would not explain the temporarily exaggerated change of course, nor the subsequent change of course by the other skipper. These aspects are typical of joint collision avoidance.

So on the basis of the case studies, we have made five observations that together roughly characterise the type of conventional joint activities to which uptake may be applicable. (1) Potential participants must pay attention and have a shared background. (2) The circumstances suggest a number of possible frames or scripts. (3) Some participant must take the initiative to engage in a joint action. (4) By accepting the initiative, participants engage in a joint action, and become bound by the norms that are part of the script. (5) Uptake accounts for behaviour that could not be explained by reactive or anticipatory coordination.

We do realise that the total class of joint actions is much larger. For instance, not all joint actions need prior knowledge of a plan. To take the football example, the first pair of players who performed a 'double pass' did not need common knowledge of the tactic. By football genius, common knowledge of the relative positions and speed, together with the rules of the game, was enough to coordinate a real 'joint action'. However, it would not classify as an instance of 'uptake'. Uptake requires a conventional frame or script. In particular under time pressure, having a predefined coordination pattern helps.

3 Uptake in Dialogue

By interacting participants execute some social activity. Often the verbal realisation of a social activity is conventionalised, and turned into a genre like information exchange or negotiation. The conventions of a genre may be expressed as a *dialogue game*. This approach has been applied in linguistics, argumentation theory, and multi-agent communication (Carletta et al., 1997; Walton and Krabbe, 1995; McBurney and Parsons, 2002). Typically, a dialogue game consists of rules to express the entry conditions, the moves that participants are allowed to make in a dialogue context, the way participants should update their apparent information states upon uttering or receiving a move, the order in which moves should follow one another and the termination conditions. The dialogue context contains general information about the setting, the participants and their roles in the social activity, as well as a record of the previous moves. How do participants determine under which dialogue game rules to conduct an interaction?

Here we consider dialogue games that contain exchanges consisting of an initiative, followed by a response and possibly an evaluation remark. For example, a question is an initiative which expects relevant information as a response; an inform act is an initiative, that expects an acknowledgement. A proposal expects a counterproposal, an acceptance or a rejection, etc. Such basic exchanges can be combined, e.g. by embedding, merging or by sequential composition. To form a coherent dialogue, there are two constraints. Within an exchange, a response to an initiative must be appropriate in the current dialogue context. And each initiative to start an exchange, must contribute to the apparent overall purpose of the dialogue. Dialogue game rules have been proposed for several genres, such as information exchange, inquiry, persuasion and negotiation (Walton and Krabbe, 1995). We assume the existence of a library with definitions of the relevant dialogue game rules.

Dialogue games are both descriptive and prescriptive. They explain the reoc-

curring patterns found in naturally occurring dialogue (Carletta et al., 1997). Often, such patterns take the shape of so called adjacency pairs, e.g. question-answer, promise-accept, greeting-greeting, etc. Therefore, an initiative may lead to expectations about the continuation of the dialogue. Initiative-response pairs also play an important role in the theory of grounding (Clark and Brennan, 1991; Traum and Allen, 1994). By making an appropriate response, the addressee indicates that he or she received and understood the initiative correctly, and does not challenge the felicity of the initiative in the dialogue context.

Initiatives not only generate expectations, they also constitute a norm to respond (Mann, 1988; Traum and Allen, 1994). In other words, not responding to an initiative, would be rude or odd. It would jeopardise the continuation of the dialogue. Nevertheless, participants are autonomous, and can ignore or violate the norm if they have a good reason. In general, when setting up a normative system, the norm itself should be fairly general. By contrast, the details of the individual cases of applying the norm should be stored and assessed separately. Therefore we suggest that specific linguistic information is encoded separately, in two kinds of rules: *appropriateness rules*, that indicate for each game what moves count as an appropriate response to some initiative, and *projection rules*, that specify which dialogue games are likely to follow a sequence of dialogue moves. Projection makes use of different aspects of the context, either static (e.g. roles of participants) or dynamic (e.g. current dialogue context).

3.1 Uptake Mechanism

Following Kreutel and Mann (2003) and Reed (1998) we suggest the following analysis of the uptake mechanism. Apart from the regular meaning, at a meta level, an initiative can be seen as a *bid* to open a dialogue game of a particular type. For example, a question initiates an information seeking game. By

responding in an appropriate way the responder also indicates to accept the bid. Once committed to a game, the participants are obliged to play by its rules.

Because an initiative is often ambivalent as to which game it starts, the responder has a large influence on the way the dialogue develops. The way the responder 'takes up' the initiative, reflects a possible interpretation that further constrains the dialogue context. For example, a request to shut the window, as in (7), may be taken as a request, permission or command (Austin, 1962), revealing the social relation between A and B.

(7) A: Shut it.

B: Sure (grant request)

B: Thanks. (acknowledge permission)

B: OK. (accept command)

Indirect speech acts are similarly ambivalent. Sometimes the responder will instead postpone the acceptation of a bid in order to 'prepare' another dialogue game, thus using the compositional aspect of dialogue games. Imagine a second hand market with buyer B and seller S. A question by B about the price of a product initiates a plain information seeking game (8). But the responder may choose to postpone the acceptance of this bid, because he is willing, say, to enter a negotiation to buy the product (9).

(8) B: How much for that jacket?

S: 25 euros. (accept information seeking).

(9) B: How much for that jacket?

S: How much would you offer? (propose negotiation)

B: 10 euros. (accept negotiation)

S: You must be kidding! I propose 20 euros.

The examples above are constructed, but examples can be found in naturally occurring dialogues too. Because each dialogue will continue in one particular direction, examples of genuine ambivalence are difficult to find.

The following example shows part of dialogue DIS150JU130 from the MICASE corpus (Simpson et al., 2002). The dialogue takes place at a visit of physics undergraduates to the planetarium. $S1$ is a graduate student, who is the guide on this tour.

(10)

S1: okay um, this is the planetarium and (i'm trying to get this to work here...) this is the evening sky in Ann Arbor as it will look tonight at about, seven P-M. [...] so this is actual north if you've got a good sense of, geography if you go outside the building and look in that direction that's north. *this point, directly above my head, what do i call that if i'm an astronomer?*

Ss: the zenith

S1: the zenith. okay? and we've got a line that runs, from north to south, passing through the zenith what do we call that?

Ss: meridian

S1: the meridian, and ...

After a long explanation, $S1$ changes the dialogue game. She starts to test the knowledge of the students. As is clear from the answers, the students 'take up' this change remarkably well. Note that $S1$ is not much older, and has no formal teaching position. Nevertheless, she adopts a teaching role and the

students accept this.

A last example is taken from Sbisà (2002). The dialogue takes place between two women: a hostess (A) and a guest (B). B is going to make a recording of an interview with A. The recording has already begun, but not the interview itself. A and B are seated in A's kitchen but A, as it turns out, prefers to have B sit more comfortably in the living room. The dialogue has been translated from Italian by Sbisà.

(11)

1 A: sorry B
2 B: it doesn't matter
3 A: wait, I'll just (–) you'll get a stiff neck (a) eh
4 B: why eh
5 A: but (-) no, if you'd wanted we could have gone into the living room we'd have been much more comfortable there
6 B: it's all the same (-) let's bring the stuff through.(-)
7 A: yes, OK. (-) I'll carry this for you.
8 B: take hold of the handle

This piece of dialogue is analysed in great detail in (Sbisà, 2002). We just mention here the most significant aspect for our study. The second part of 5A is open: it might be either an offer, a commitment to do something desirable for the hearer, or a simple proposal, i.e., allowing the hearer to do something. Because 6B subsequently denies any interest in moving, the latter reading is eventually chosen, by uptake. Note that the second part of 6B is indeed appropriate as a response to a proposal. Sbisà concludes that "turn 6 can be viewed as operating a selection among the forces that 5B might have, that of an offer (which is rejected) and that of a proposal (which is accepted, by accepting the proposal itself)."

These examples show an open and negotiable dialogue context, for example with respect to the social setting or intentions of the participants. Such in-

determinism does not have to be a bad thing. Selecting a particular response further constrains the dialogue context. In this respect, the uptake mechanism is similar to presupposition accommodation (Stalnaker, 1974).

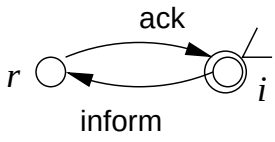
Before we continue to formalise the approach, we first need to evaluate it against the original. Austin identifies three different kinds of effects with an utterance: (i) the securing of uptake, (ii) the production of a conventional effect, and (iii) the invitation of a response or sequel (Austin 1962, p.116-117; Sbisà 2000). How can we account for these three effects, using the analysis of uptake introduced above?

Effect (i), the securing of uptake, means that without being taken up, a speech act is often not really completed. So a promise is nothing without being accepted. In our approach, uptake is secured by the fact that an initiative counts as a 'bid' for some dialogue game, combined with the general norm to respond to such bids. Note that effect (i) only makes sense for initiatives. Effect (ii), the production of a conventional effect, covers the regular meaning of a speech act, and its propositional content. Effect (iii), inviting a response or sequel of a particular kind, is taken care of by the appropriateness rules.

3.2 Towards Formalisation

Because dialogue games describe simple patterns, the order in which moves must be made can be given a straightforward logical representation (Endriss et al., 2003). Figure 3 shows an example of a very simple protocol, with rules distributed among an initiator i and a responder r. The arrow '$\longrightarrow$' represents a kind of production rule, that is indexed by time. The symbol '$\vee$' represents a choice. So $\alpha(t) \multimap \beta_1(t{+}1) \vee \beta_2(t{+}1)$ means that after act or event α at time t, either β_1 or β_2 or both must take place at time $t{+}1$.

However, dialogue rules modelled as hard constraints can not deal with conflicts between applicable rules, and prevents contrary-to-duty reasoning. We



$$i\ \mathsf{start}(t) \longrightarrow \mathsf{inform}(i, r, \varphi)(t{+}1)$$

$$\mathsf{ack}(r, i, \varphi)(t) \longrightarrow \mathsf{start}(t{+}1) \vee \mathsf{stop}(t{+}1)$$

$$r\ \mathsf{inform}(i, r, \varphi)(t) \longrightarrow \mathsf{ack}(r, i, \varphi)(t{+}1)$$

Fig. 3. Continuous update protocol [5]

need an explicit way to represent obligations, that can distinguish between violations and inconsistencies, and can resolve conflicts by means of a priority order. Systems of obligations have been the subject of much research in deontic logic, see e.g. Wieringa and Meyer (1993) for an overview of applied approaches. However, in our case the obligations interact with general principles for assessing a situation and for reasoning about other agents on the basis of expectations. A combination of obligations with other mental attitudes, like goals and beliefs, would be appropriate. One such combination is provided by the BOID cognitive agent architecture (Broersen et al., 2002). A more advanced alternative would be the Normative Multiagent Systems (NMAS) approach developed by Boella and Van der Torre (2004; forthcoming). These approaches are suitable, because they place a heavy emphasis on the fact that obligation rules are contextually dependent. Other declarative formalisms for specifying normative rules may also be used. Our analysis does not depend on the exact choice of formalism.

The BOID defines the goal generation process of an intelligent agent in terms of components for belief B, obligation O, intention I and desire D (Broersen et al., 2002). Components are sets of prioritised default rules of the form: $\varphi -_B\!\!\rightarrow \psi$, $\varphi -_O\!\!\rightarrow \psi$, $\varphi -_I\!\!\rightarrow \psi$ or $\varphi -_D\!\!\rightarrow \psi$, where φ, ψ are formulas constructed by '$\wedge$' (and) and '$\vee$' (or) from so called literals: basic propositions or their negations ($\neg$). Please note that we use a notation convention for expressing dialogue acts that makes the initiator, responder, content and type of dialogue act explicit. Rule schemata contain placeholders, as for example in figure 4. Since for any actual set of rules these placeholders have been instantiated with constants, no real quantification is needed; technically these expressions

remain propositional.

Rules are applied iteratively, to form an extension: a maximally consistent set of literals. Rules are only applicable, in case the antecedent is contained in the extension and the consequent is consistent with it, and is not already contained in it. An extension is interpreted as the goal set for an agent, which may serve as input to a planning procedure to realise the goals. A priority order is used to resolve conflicts: if several rules are applicable, the rule with the highest priority is selected. If the priority order respects the boundaries between components, it may characterise an *agent type*. For example, selfish agents generally give priority to desires over obligations; social agents value obligations over desires and realistic agents only generate goals which are compatible with their belief rules. The BOID architecture has been implemented, and has been used to analyse benchmark examples of conflicts between mental attitudes [2] .

Dialogue game rules are represented as a protocol (Endriss et al., 2003), but now using conditional obligations $-O\rightarrow$ from the BOID. Given some protocol definition, $\alpha(t)$ is an expression that denotes a dialogue move received at time t. Each $\beta_i(t+1)$ is a legal continuation at time $t+1$. The obligation means that at least one of these continuations must be selected. Derived meanings β' of a move β can be inferred by defeasible B-rules. Note that several possible interpretations of moves can exist in parallel.

(12) $\alpha(t) -O\rightarrow \bigvee_i \beta_i(t+1)$

(13) $\beta(t) -B\rightarrow \beta'(t)$

We formalise the examples by the general uptake rule schemata in figure 4, which are shared among initiator i and responder r because they may switch roles. The notation is used as follows. Parameter g can be any game and m, n can be any move. A sequence of moves is indicated by ϵ for the empty sequence,

[2] See `http://boid.info/`.

1. $\mathsf{init}(i,r,m)(t) \wedge \mathsf{project}(m,g) \ -\!\!B\!\!\rightarrow \ \mathsf{bid}(i,r,g)(t)$

2. $\mathsf{bid}(i,r,g)(t) \ -\!\!O\!\!\rightarrow \ \mathsf{accept}(r,i,g)(t{+}1) \vee \mathsf{reject}(r,i,g)(t{+}1)$

3. $\mathsf{accept}(r,i,g)(t) \ -\!\!B\!\!\rightarrow \mathsf{resp}(r,i,n)(t) \wedge \mathsf{appr}(m,n,g) \wedge \mathsf{project}(m;n,g)$

4. $\mathsf{reject}(r,i,g)(t) \ -\!\!B\!\!\rightarrow \neg(\mathsf{resp}(r,i,n)(t) \wedge \mathsf{appr}(m,n,g)) \wedge \mathsf{project}(m,g)$

3′. $\mathsf{resp}(r,i,n)(t) \wedge \mathsf{appr}(m,n,g) \wedge \mathsf{project}(m;n,g) \ -\!\!B\!\!\rightarrow \mathsf{accept}(r,i,g)(t)$

4′. $\neg(\mathsf{resp}(r,i,n)(t) \wedge \mathsf{appr}(m,n,g)) \wedge \mathsf{project}(m,g) \ -\!\!B\!\!\rightarrow \mathsf{reject}(r,i,g)(t)$

5. $\mathsf{comm}(r,i,g)(t) \wedge \mathsf{init}(i,r,m)(t) \ -\!\!O\!\!\rightarrow \mathsf{resp}(r,i,n)(t{+}1) \wedge \mathsf{appr}(m,n,g)$

6. $\mathsf{comm}(i,r,g)(t) \wedge \mathsf{resp}(r,i,m)(t) \ -\!\!O\!\!\rightarrow \mathsf{resp}(i,r,n)(t{+}1) \wedge \mathsf{appr}(m,n,g)$

7. $\mathsf{accept}(r,i,g)(t) \quad \rightarrow \quad \mathsf{comm}(r,i,g)(t'),\ t \leq t'$

8. $\mathsf{bid}(i,r,g)(t) \wedge \mathsf{comm}(r,i,g)(t') \quad \rightarrow \quad \mathsf{comm}(i,r,g)(t''),\ t \leq t' \leq t''$

Fig. 4. Rules for uptake

or by '$m;s$', where m can be any move and s can be any sequence. The predicate '$\mathsf{project}(s,g)$' is true, whenever a sequence of moves s projects that dialogue game g is one of the possible well-formed continuations. The predicate '$\mathsf{appr}(m,n,g)$' indicates that a response n is appropriate to an initiative m, in the context of a given game g. Predicate '$\mathsf{comm}(x,y,g)$' means that an agent x is committed towards another agent y, to continue with a game g. The fact that commitments cannot be overridden, is indicated by the use of a regular material implication.

The rules are applied in derivation (14). The example specific notation is meant as follows. A dialogue game is of the form $game(i,r,topic)$, where $game$ can be 'nego' or 'info' and i and r alternate between seller s and buyer b. The expression $topic$ refers to the object or issue the game is about. Notation $?x.P(x)$ expresses an issue like "What value is P?". So for example, since $\mathsf{price}(j,25)$ represents "the price of jacket j is 25 euros", we can turn it into a

question $?x.\mathsf{price}(j,x)$, meaning "What is the price of jacket j?". The following derivation will produce the dialogue in example (8).

(14) 1. $\mathsf{init}(b,s,?x.\mathsf{price}(j,x))(t_1)$,

$\mathsf{project}(?x.\mathsf{price}(j,x),\mathsf{info}(b,s,j))$

2. $\mathsf{bid}(b,s,\mathsf{info}(b,s,j))(t_1)$ (B-rule 1)

3. $\mathsf{accept}(s,b,\mathsf{info}(b,s,j))(t_2) \vee \mathsf{reject}(s,b,\mathsf{info}(b,s,j))(t_2)$ (O-rule 2)

4. $\mathsf{accept}(s,b,\mathsf{info}(b,s,j))(t_2)$ (cooperative)

5. $\mathsf{resp}(s,b,\mathsf{price}(j,25))(t_2)$,

$\mathsf{appr}(?x.\mathsf{price}(j,x),\mathsf{price}(j,25),\mathsf{info}(b,s,j))$ (B-rule 3)

In step 1 and 2 the question about the price of the jacket counts as a bid for an information seeking game about jacket j. In step 3, the general obligation to address a bid means that the seller must either accept or reject it. In step 4, we suppose the seller s is cooperative, which triggers a goal to accept the information seeking bid. This goal may for example be caused by a desire, typically modelled by a D-rule in the BOID, but this is not shown in the example. In step 5, the seller shows his acceptance by uttering an appropriate response, namely a price for the jacket.

A similar derivation can be constructed for example (9). Step 1-3 are just like in (14). But now the seller is asking $?y.\mathsf{offer}(b,y,j)$: what price y is b prepared to offer for the jacket? This initiates an embedded information seeking game, about the value the buyer attaches to the jacket: $\mathsf{info}(s,b,\mathsf{val}(b,j))$. Because of the entry conditions of an information seeking game – that the answer is not yet known – this only makes sense if the seller has not fixed a price yet, which further suggests that negotiation is possible (step 4). So an alternative projection is $\mathsf{nego}(s,b,\mathsf{val}(j))$. Both these projections trigger a bid (step 5), and thus an obligation to respond (step 6). Now suppose the buyer is coop-

erative, and also competitive, i.e. wants the best price. That means it will accept the negotiation bid. The buyer's reply, "10 euros", is an appropriate response, both with respect to the information seeking and to the negotiation bids (step 7, 8). So until now both participants can live with the underlying ambivalence of the moves. However, the seller's rejection of the buyer's offer and the counterproposal are only appropriate in the context of a negotiation (step 10), which solves the ambivalence. Note furthermore that the counter-proposal can also be regarded as a postponed response to the buyer's initial question (step 11). That means that in this case the negotiation subsumes two information exchanges: one about what each of the participants would offer.

(15) 1. $\mathsf{init}(b, s, ?x.\mathsf{price}(j, x))(t_1), \mathsf{project}(?x.\mathsf{price}(j, x), \mathsf{info}(b, s, j))$

2. $\mathsf{bid}(b, s, \mathsf{info}(b, s, j))(t_1)$ (B-1)

3. $\mathsf{accept}(s, b, \mathsf{info}(b, s, j))(t_2) \vee \mathsf{reject}(s, b, \mathsf{info}(b, s, j))(t_2)$ (as in (14))

4. $\mathsf{init}(s, b, ?y.\mathsf{offer}(b, y, j))(t_2)$,

   $\mathsf{project}(?y.\mathsf{offer}(b, y, j), \mathsf{info}(s, b, \mathsf{val}(b, j)))$,

   $\mathsf{project}(?y.\mathsf{offer}(b, y, j), \mathsf{nego}(s, b, \mathsf{val}(j)))$

5. $\mathsf{bid}(s, b, \mathsf{info}(s, b, \mathsf{val}(b, j)))(t_2)$,

   $\mathsf{bid}(s, b, \mathsf{nego}(s, b, \mathsf{val}(j)))(t_2)$ (B-1)

6. $\mathsf{accept}(b, s, \mathsf{info}(s, b, \mathsf{val}(b, j))(t_3) \vee \mathsf{reject}(b, s, \mathsf{info}(s, b, \mathsf{val}(b, j))(t_3)$,

   $\mathsf{accept}(b, s, \mathsf{nego}(s, b, \mathsf{val}(j))(t_3) \vee \mathsf{reject}(b, s, \mathsf{nego}(s, b, \mathsf{val}(j))(t_3)$ (O-2)

7. $\mathsf{accept}(b, s, \mathsf{info}(s, b, \mathsf{val}(b, j))(t_3)$ (coop.)

   $\mathsf{resp}(b, s, \mathsf{offer}(b, 10, j))(t_3)$,

   $\mathsf{appr}(?y.\mathsf{offer}(b, y, j), \mathsf{offer}(b, j, 10), \mathsf{info}(s, b, \mathsf{val}(b, j)))$ (B-3)

8. $\mathsf{accept}(b, s, \mathsf{nego}(s, b, \mathsf{val}(j))(t_3)$ (comp.)

   $\mathsf{resp}(b, s, \mathsf{offer}(b, 10, j))(t_3)$,

   $\mathsf{appr}(?y.\mathsf{offer}(b, y, j), \mathsf{offer}(b, 10, j), \mathsf{nego}(s, b, \mathsf{val}(j)))$ (B-3)

10. $\mathsf{resp}(s, b, \mathsf{offer}(s, 20, j))(t_4)$,

    $\mathsf{appr}(\mathsf{offer}(b, j, 10), \mathsf{offer}(s, 20, j), \mathsf{nego}(s, b, \mathsf{val}(j)))$ (7,O-6)

11. $\mathsf{resp}(s, b, \mathsf{offer}(s, 20, j))(t_4)$,

    $\mathsf{appr}(?x.\mathsf{val}(j, x), \mathsf{offer}(s, 20, j), \mathsf{info}(b, s, j))$ (7,O-5)

How can we generalise from this example? The general idea behind the reasoning is summarised in figure 5. We start in the lower left corner. Suppose at time t participant i initiates a move m. Because m can be projected as the beginning of a dialogue game g, we can say that $\mathsf{init}(i, r, m)$ counts as a bid to start game g. This could be an instance of the 'counts as' relation studied by Searle (1997) [3] . The institution that guarantees the meaning of the rule, is the group of participating language users with their conventions. In our account, the 'counts as' relation is represented by a B-rule in the BOID. The general norm is that, once recognised, all bids must be addressed, i.e. be either accepted or rejected. That is represented by the O-rule on top of the figure. Ambiguity about the underlying bid of an initiative can be located in the 'counts as' rule. Therefore it is crucial that B-rules in the BOID are defeasible. By contrast, the obligation to respond is not defeasible. It could also be modelled by a conditional obligation in a more standard deontic logic.

If a bid is not appropriately responded to, the corresponding game is ended. In step 10, the fact that a counterproposal is not appropriate in the context of information seeking ends this possible continuation. This process is similar to presupposition accommodation (Stalnaker, 1974), and is used to weed out the multitude of possible continuations. But if the last possible continuation is weeded out, the participants have a problem. Such deadlock situations are to be avoided.

Suppose r does want to accept the bid, for example because he or she is cooperative, as we assumed in the example. How can r express this acceptance? To make acceptance explicit would require additional resources. Therefore the typical way to signal acceptance of a dialogue game bid is implicit, by some appropriate response in the context of the dialogue game. If responses to several alternative games would be appropriate, as in the example, the

[3] But see the discussion in section 5.4 below.

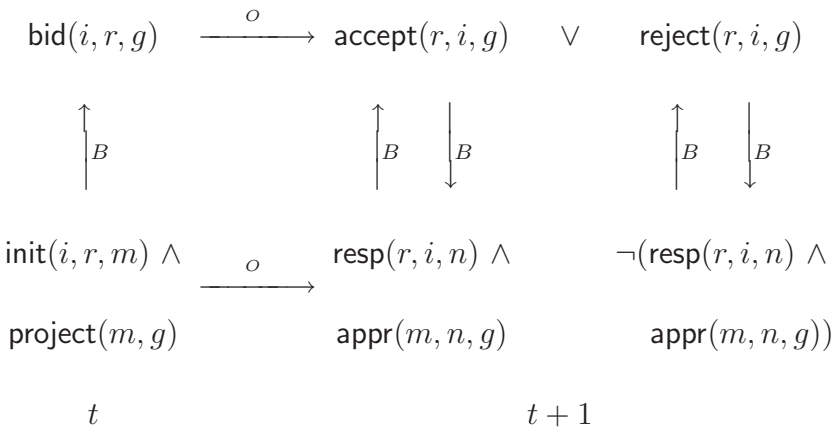


Fig. 5. Reasoning Process

responder's choice will influence the subsequent dialogue game. Two kinds of reasoning are involved. The responder needs a so called practical reasoning rule to select the best response, also represented as a B-rule in the BOID. Examples are rule 3 and 4 in figure 4. The rules can be paraphrased as: if a responder would want to accept the bid for game g, then he or she must respond by a move n which is appropriate in g after m. Practical reasoning rules are directed from top to bottom in figure 5. The rule works, because from the point of view of the hearer of the response, i, an appropriate response counts as an acceptance. Similarly, an inappropriate response or no response at all, would count as a rejection. In figure 4 rule $3'$ and $4'$ represent this kind of reasoning. These 'count as' rules are directed from bottom to top in figure 5.

Once a dialogue game is established, and participants are committed to the game, either because they accepted, or because they initiated it and it was accepted, much stricter rules apply. In particular, we have the usual dialogue game rules, which can be seen as obligations to respond appropriately. Rule 5 and 6 of figure 4 are examples of such rules, shown at the bottom of the figure. The outcome of the obligation to respond appropriately is the same as the outcome of the acceptance reasoning above. The same utterance does double duty.

Note that we do not make any claim about the cognitive validity of the model. Researchers have convincingly argued against separate modules to process primary and secondary meanings in human language processing, see Clark (1996). The typical example is "Could you tell me the time?". Empirical evidence does not support the hypothesis of separate derivation steps in such cases. So when we say that an initiative (primary) counts as a bid at a meta level (secondary), the terms should be taken in a theoretical sense. Initiative-response patterns are examples of reactive patterns of human behaviour, that are often simply triggered by the circumstances. The model explains why, in case of a misunderstanding for example, these conventional patterns turn out to have such complex meanings with regard to dialogue obligations. In other words, having a meaning at a meta level, does not preclude fast processing.

4 Uptake for Joint Activities

Can we extend this approach to joint activities that are, unlike dialogue games, not expressed verbally? The hypothesis would be 'yes', since the verbal aspect does not seem to be crucial to the analysis. What matters is that dialogue games describe conventional interaction patterns, which produce well-formed sequences of action. The work is done by the projection rules and the appropriateness constraints. That means, that if we can find projection rules and appropriateness constraints that make sense for the non-verbal examples of joint activities in section 2.3, we can apply the uptake mechanisms to explain how such activities are established. The question then remains whether the uptake mechanism really leads to a state that characterises a joint action, displaying the properties discussed in section 2. That question is discussed in section 6 below. In what follows we describe how the uptake mechanism can be used to represent crucial aspects of joint activities, using the BOID representation language. But first, we review a number of essential aspects of the notion of a recipe.

As we established in section 2.3, uptake only succeeds in a context in which potential participants with a shared background are paying attention, and a set of conventional coordination patterns – frames or scripts – has become salient. How can we analyse such scripts? We will use the technical concept of a partial plan, or *recipe* (Pollack, 1990; Grosz and Kraus, 1996), to account for some of the properties of such a coordination pattern. Below we discuss some additional constraints on recipes and on participatory actions.

A joint action progresses in phases: periods in which participatory actions are carried out simultaneously. If these phases are mutually known, participants only have to synchronise on the so called *entry* and *exit points* of each phase. Synchronisation of phases becomes easier when there are jointly salient reference points (Clark, 1996). In a periodic process, like digging or marching, reference points are provided by a rhythm. Rhythm plays an important part in our examples (1), (2) and (4), namely the duet and the singing. Often a kind of anticipatory phase precedes the actual joint action. Consider the conductor of an orchestra that raises his baton, just before the start of a symphony. Our example (3), lifting the table, shows this kind of anticipatory phase, that makes it easier to synchronise on the entry of the actual joint action.

In general, planning consists of task decomposition and scheduling. Both aspects are subject to constraints. See Pollack (1990) for more about planning, and Grosz and Kraus (1996) for more about joint planning. Regarding individual actions, there are *preconditions*, that must be true at the start of an action in order for it to be successfully completed, and *effects*, that will be true after a successful completion. So when the preconditions are true, the effects can be guaranteed. The effect provides a declarative specification of the state of affairs that is to be achieved by the action. Composite actions can be composed of basic actions, by means of operators like ';', for sequential composition, '|' for choice and '||' for parallel composition. Scheduling constraints

can be reduced to three basic types: *precedence*, *dominance* and *alignment*.

Precedence puts a constraint on the order in which actions are carried out. We have $\mathsf{exit}(\alpha_1) < \mathsf{entry}(\alpha_2)$, whenever α_1 needs to be completed before α_2 can be started [4]. Precedence can derive from preconditions and effects: if $\mathsf{effect}(\alpha_1) \rightarrow \mathsf{pre}(\alpha_2)$, then $\mathsf{exit}(\alpha_1) < \mathsf{entry}(\alpha_2)$. Precedence can also derive from the use of a common resource. If both α_1 and α_2 require a resource r, we have either $\mathsf{exit}(\alpha_1) < \mathsf{entry}(\alpha_2)$ or $\mathsf{exit}(\alpha_2) < \mathsf{entry}(\alpha_1)$.

Dominance has to do with composition. When a complex action α consists of various sub-action $\alpha_1, ..., \alpha_n$, i.e., we say that α *dominates* α_i, for $1 \leq i \leq n$. Composite actions are often depicted as an and/or tree, with the branches being dominated by the root. Preconditions and effects aggregate to composite actions. For example, the precondition of the composite action consists of the preconditions of each of the sub-actions, except for those that have been accomplished by previous sub-actions: $\mathsf{pre}(\alpha_1; \alpha_2) = \mathsf{pre}(\alpha_1) \cup (\mathsf{pre}(\alpha_2) \setminus \mathsf{effect}(\alpha_1))$ [5]. Using such equations, an algebra of composite actions and their preconditions and effects can be constructed.

Finally, *alignment* accounts for synchronisation. For example, the *chaining strategy*, in which sub-actions $\alpha_1; ...; \alpha_n$ are carried out in sequence, is expressed by the constraint $\mathsf{entry}(\alpha_{i+1}) = \mathsf{exit}(\alpha_i)$, for $1 \leq i < n$. If the activity is periodic, we further get that $\mathsf{exit}(\alpha) = \mathsf{entry}(\alpha) + d$, where d represents the duration of the period of the activity.

We are now in a position to clarify what we understand by a recipe. A *recipe* is a partial specification of a joint action. In other words, a recipe is nothing but a mutually consistent set of constraints on a set of actions. Think of a recipe as

[4] We write $\mathsf{entry}(\alpha)$ and $\mathsf{exit}(\alpha)$ for the entry and exit points of action α, expressed in some proper temporal formalism. Event calculus would be a good candidate (Dean and Boddy, 1988).

[5] We use $\mathsf{pre}(\alpha)$ and $\mathsf{effect}(\alpha)$ to denote the preconditions and effects of α, represented as consistent sets of literals, just like extensions in the BOID.

a complicated formula, with many variables still open. The more constraints there are in the recipe, the less possible instantiations. As we argued above, the constraints in a recipe are typically defined in terms of an over all precondition and effect, a set of basic actions with preconditions and effects, a set of roles with constraints regarding capabilities or skill of potential participants, a set of possible resources, again with constraints on their required properties, along with a set of dominance constraints, a set of precedence constraints and a set of alignment constraints.

Because a recipe is a partial specification of a plan, common knowledge of the recipe does not require that all participants have complete knowledge of each others task. The participants can assume that the others will achieve the effects of their parts, and rely on this, but they can delegate how this is achieved.

In the BOID formalism, some constraints must be represented as strict '$\rightarrow$' rules, such as for example the mutual exclusion of resources. Others, such as constraints on the skills of the participants, are better expressed as defeasible '$-_B\!\!\rightarrow$' rules.

4.2 Towards Formalisation

We use the following notation. We use A as a variable ranging over joint actions. The participants in a joint action A form an indexed list $X = \langle x_1, ..., x_n \rangle$, of which initiator $i \in X$ has a special role. Unlike in section 3.2 we no longer distinguish initiatives and responses, but we just have actions being executed. Expression $\mathsf{exec}(x, X, \alpha)(t)$ means that agent x executes action α towards other participants X. Instead of discrete time points $t + 1$, we now use t' to represent a new time such that $t \leq t'$, expressed in some appropriate temporal representation.

The first thing that we need to clarify are the *projection rules*: what actions

'count as' a bid for a joint action? In all the examples of section 2.3, the context was such that potential participants already where paying attention, and could expect a joint action to be initiated. With regard to initiating play, lifting the table, starting a song, starting a double-pass, or signalling avoidance behaviour, the initial act itself projects the joint action as a conventional continuation, and therefore counts as a bid to engage in the joint action. Again, we also have the obligation to either accept or reject such a bid, once recognised.

1. $\mathsf{exec}(i, X, \alpha)(t) \wedge \mathsf{project}(\alpha, A) \ -_B\!\rightarrow \ \mathsf{bid}(i, X, A)(t)$
2. $\mathsf{bid}(i, X, A)(t) \ -_O\!\rightarrow \ \mathsf{accept}(x, X, A)(t') \vee \mathsf{reject}(x, X, A)(t'),$

   $x \in X, t \leq t'$

The second thing that we need to clarify is the role of the *appropriateness constraints*. Here lies one of the main differences between dialogue games and joint actions in general. Whereas dialogue games can be defined as simple patterns allowing a straightforward sequential representation as initiative-response structures, we can no longer make such assumptions for general joint actions: they will typically involve more than two agents, with actions being coordinated in different manners, e.g., synchronously if you take the example of the agents lifting the table. In this context, the appropriateness constraints must therefore determine a task allocation: a mapping from agents and time-points to actions. Task allocations essentially specify who should do what and when. We assume that these task allocations meet the usual properties, e.g., that actions are not conflicting with each other. We will make use of a 5-ary predicate $\mathsf{appr}(\alpha, A, x, X, t)$ which captures that it is indeed appropriate for agent x to undertake action α at time t as part of the joint action A involving agents X.

Just like in the dialogue case, appropriateness constraints are used as part of the practical reasoning rules, for agents who already decided to accept or reject a bid.

3. $\mathsf{accept}(x, X, A)(t) \multimap_B \mathsf{exec}(x, X, \alpha)(t) \wedge \mathsf{appr}(\alpha, A, x, X, t)$

4. $\mathsf{reject}(x, X, A)(t) \multimap_B \neg(\mathsf{exec}(x, X, \alpha)(t) \wedge \mathsf{appr}(\alpha, A, x, X, t))$

And just like in the dialogue case, an appropriate action is used to signal acceptance. Therefore we add the following 'counts as' rules.

$3'$. $\mathsf{exec}(x, X, \alpha)(t) \wedge \mathsf{appr}(\alpha, A, x, X, t) \multimap_B \mathsf{accept}(x, X, A)(t)$

$4'$. $\neg(\mathsf{exec}(x, X, \alpha)(t) \wedge \mathsf{appr}(\alpha, A, x, X, t)) \multimap_B \mathsf{reject}(x, X, A)(t)$

Once committed to a joint action, each participant is bound to undertake its participatory actions as required by the current task allocation. This task is specified by the appropriateness constraints. The predicate 'comm' is still used for bookkeeping which agents are committed to what joint actions.

5. $\mathsf{comm}(x, X, A)(t) \wedge \mathsf{appr}(\alpha, A, x, X, t) \multimap_O \mathsf{exec}(x, X, \alpha)(t')$,

6. $\mathsf{accept}(x, X, A)(t) \rightarrow \mathsf{comm}(x, X, A)(t')$,

7. $\mathsf{bid}(i, X, A)(t) \wedge \mathsf{comm}(x, X, A)(t') \rightarrow \mathsf{comm}(i, X, A)(t'')$,

   $x \in X,\ t \leq t' \leq t''$

This structure roughly captures the establishment of the joint intentions and the common knowledge of the task allocation. The third required element, mutual support, is more difficult to handle. The most natural way to model helping is by a separate joint activity, initiated by the person who needs help. This works analogously to a bid for a regular dialogue game, although a request for help need not to be expressed verbally.

We need to ensure that, once agents are committed to a joint action, it cannot be an option to reject a request for help about some action α which is appropriate in the context of joint action A, unless that would jeopardise their own part of the plan. Assuming the existence of a dialogue game $\mathsf{help}(i, r, \alpha)$, this can be captured by the following rule:

8. $\mathsf{bid}(i, r, \mathsf{help}(i, r, \alpha))(t) \wedge \mathsf{comm}(r, X, A)(t) \wedge \mathsf{appr}(\alpha, A, i, X, t) \wedge$

$\mathsf{appr}(\beta, A, r, X, t{+}1) \wedge \neg\mathsf{conflict}(\beta, \mathsf{help}(i, r, \alpha))$

$-_O\!\!\rightarrow \ \mathsf{accept}(r, i, \mathsf{help}(i, r, \alpha))(t{+}1)$

In many cases however, bids for help will not be explicit. It is then necessary to define what circumstances 'count as' such bids. One example we may cite is the case of failed actions: when perceiving that another agent involved in the same joint action fails to achieve its participatory action, a partner should consider that as a bid for help. In other words, the projection rules must be adapted such that, when an action has failed, they project a *help* dialogue. Because of rule 1, a failed action will then count as a bid for a help dialogue.

9. $\mathsf{fail}(i, X, \alpha)(t) \wedge \mathsf{appr}(\alpha, A, i, X, t) \ -_B\!\!\rightarrow \ \mathsf{project}(\alpha, \mathsf{help}(r, i, \alpha)),$

$r \in X$

This would be the case for example when agents involved in the 'lifting table' joint action notice that one of them is in trouble. Usually, 'recovery actions' are then invoked to somehow fix the joint action, being it by interrupting the action, or by supporting the agent.

To illustrate that the rules also work for joint actions, we give an example based on the collision avoidance mechanism discussed in Section 2. To keep the case simple, we assume that the joint action involves a couple of 'sharp starboard' actions to be executed, by each agent, in sequence (precedence relation).

(16) 1.$\mathsf{exec}(a, \langle a, b\rangle, \mathsf{sharp\_starboard})(t_1), \mathsf{project}(\mathsf{sharp\_starboard}, \mathsf{avoidance})$

2.$\mathsf{bid}(a, \langle a, b\rangle, \mathsf{avoidance})(t_1)$ (B-1)

3.$\mathsf{accept}(b, \langle a, b\rangle, \mathsf{avoidance})(t_2) \vee \mathsf{reject}(b, \langle a, b\rangle, \mathsf{avoidance})(t_2)$ (O-2)

4.$\mathsf{accept}(b, \langle a, b\rangle, \mathsf{avoidance})(t_3)$ (coop.)

$\mathsf{exec}(b, \langle a, b\rangle, \mathsf{sharp\_starboard})(t_3)$

$\mathsf{appr}(\mathsf{sharp\_starboard}, \mathsf{avoidance})(t_3)$ (B-3)

As a further illustration, here is a version of the table lifting example. Now suppose that agent b is not ready to lift the table, and wants to reject the bid.

(17) 1. $\mathsf{exec}(a, \langle a, b, c, d\rangle, \mathsf{glance; lift})(t_0), \mathsf{project}(\mathsf{glance; lift},\ \mathsf{lifting})$

2. $\mathsf{bid}(a, \langle a, b, c, d\rangle, \mathsf{lifting})(t_1) \quad (t_0 \leq t_1)$ (B-1)

3. $\mathsf{accept}(x, \langle a, b, c, d\rangle, \mathsf{lifting})(t_2) \vee \mathsf{reject}(x, \langle a, b, c, d\rangle, \mathsf{lifting})(t_2)$($O$-2)

$(x \in \{b, c, d\}, t_1 \leq t_2)$

4. $\mathsf{reject}(b, \langle a, b, c, d\rangle, \mathsf{lifting})(t_3) \quad (t_2 \leq t_3)$ (not ready)

$\mathsf{exec}(b, \langle a, b, c, d\rangle, \mathsf{warning})(t_3)$

$\neg\mathsf{appr}(\mathsf{warning}, \mathsf{lifting})(t_3)$ (B-3)

Here we illustrate the mutual support constraint. Steps 1-3 are just like above. But now everybody accepts, and thereby becomes committed. Now later, when agent b fails, e.g. stumbles, the others will help by for example putting the table down temporarily.

(18) 4. $\mathsf{accept}(x, \langle a, b, c, d\rangle, \mathsf{lifting})(t_3)$ $(x \in \{b, c, d\})$ (coop)

$\mathsf{exec}(x, \langle a, b, c, d\rangle, \mathsf{lift})(t_3)$

$\mathsf{appr}(\mathsf{lift}, \mathsf{lifting})(t_3)$ (B-3)

5. $\mathsf{comm}(x, \langle a, b, c, d\rangle, \mathsf{lifting})(t_4)$ $(t_3 \leq t_4, x \in \{a, b, c, d\})(6, 7)$

6. $\mathsf{fail}(b, \langle a, b, c, d\rangle, \mathsf{lift})(t_5)$ (external event)

7. $\mathsf{project}(\mathsf{lift}, \mathsf{help}(x, b, \mathsf{lift})$ (B-9)

So by extending the techniques introduced in the context of dialogue games, we captured that each agent involved in a joint action, is obliged to undertake its participatory actions, and is also obliged to support its fellow partners. One way to guarantee that these obligations are not violated, is to assume that the agent is of the *social* agent type, as defined in (Broersen et al., 2002). By contrast to a selfish agent, a social agent lets its obligations prevail over individual desires.

5 Discussion and Related Research

We have tried to demonstrate that an uptake mechanism can be used to explain how participants engage in a joint action. Partly because of the constraints of formalisation, the mechanism has been rather crude and simplified. We do realise that there are difficult issues underlying the mechanism. Some of these issues will now be discussed in more depth, giving pointers to related literature on the subject.

At various places throughout the paper we have argued that once dialogue participants are engaged in a joint action, they are bound by the rules of the recipe or script. This includes the obligation of mutual support, and the obligation to respond. Our formalisation in terms of $\multimap\!\!o \rightarrow$ rules has made these norms even more prominent. But, one might argue, is the notion of obligation really necessary here? After all, would not co-interest be sufficient to explain mutual support? Obligation is a rather strong notion. If indeed all agents have a commitment towards the success of the joint action, would that not be enough to secure success? To see why, we refer to the convoy example invoked by Cohen and Levesque (1991). Imagine that Bob wants to go home, but unfortunately doesn't know his way. Now he knows that Alice is going near his place, and also that she knows her way. They really need to work together, otherwise Alice could, for instance, drive too quickly for Bob to keep up. Now maybe at some point Bob might come to believe that Alice does not know her way, after all. Certainly, it would be in his interest to leave the convoy, and try to get home some other way. But instead he will typically try to communicate that the joint activity is over, before splitting up.

Obligations prevent agents from deviating from the script or frame. This in turn provides stability, in the sense that agents can have solid expectations as to what actions will be performed by partners, regardless of what mere interest-driven behaviour might favour.

In the context of natural language dialogue, Traum and Allen (1994), make the following observation.

> Consider a stranger approaching an agent and asking, Do you have the time? It is unlikely that there is a joint intention or shared plan, as they have never met before. From a purely strategic point of view, the agent may have no interest in whether the strangers goals are met. Yet, typically

agents will still respond in such situations.

Moreover, the very same rules that govern dialogue in cooperative circumstances, can be used against someone's interests, in non-cooperative circumstances (Gabbay and Woods, 2001). Consider a police interrogation. The suspect will be trying to give away as little information as possible, while still technically answering the question. This shows that some of these conventional interaction patterns have a normative power. The typical sanction upon violating such an obligation, is that the interaction as a whole breaks down. Some researchers have argued that in addition to sanctions, obligations may work because of an internal motivation: people feel obliged to help others, and they feel bad when they do not (Castelfranchi, 1998).

5.2 Rules

The previous section has stressed the prescriptive aspect of dialogue game rules, or of recipes for joint action. But rules also have an important descriptive aspect. Dialogue game rules, or recipes, describe common patterns of interaction, that have been conventionalised. In the end, projection and appropriateness rules are not much more than statistically relevant dependencies between occurrences of dialogue acts, or participatory actions. This descriptive perspective on dialogue game rules is apparent in the empirical work on natural language dialogue, such as that by Carletta et al (1997). Historically one might trace this view back to the ethnomethodological approach to dialogue, which stresses the importance of observing human interaction without too much interpretation. This observational aspect remains relevant today. For instance, if one would like to build an artificial dialogue agent that could learn to adapt to different interaction protocols, one might try to make it learn the projection and the appropriateness relations, by observation. Initial experiments in this direction have proved encouraging (Hiel, 2005). In such a learning setting, it makes sense to have a basic architecture for the agent,

that contains among other things the obligation to respond, and the obligation to act appropriately. But the details of what to respond, or how to act appropriately can be learned from observation of other agents interacting.

There is a second argument for this type of architecture. Application of interaction patterns or recipes should be easy and fast. Much easier than having to reason about joint intentions. Recall the discussion at the end of section 3.2. Interaction patterns are often 'hardwired' into the cognitive machinery of people. It seems feasible that sets of appropriateness and projection rules can be hardwired in this way. But this does not mean that joint intentions do not exist. When applying joint intentions, people are not necessarily aware of them. Only when the joint action breaks down, do the underlying intentions become apparent.

5.3 Signalling

A large part of our approach is based on the assumption that some state of affairs or action – both practical and dialogue actions – can be interpreted as a kind of message. They signal something. Earlier in the paper we have used terms like 'be understood as', 'signal', 'indicate', 'count as', etc. to refer to such a notion of signalling. Clearly we need a theoretical basis for these terms. In this section we will review some relevant literature.

One way to start, would be to use Schelling's (1960) notion of coordination game. In the typical pay-off matrix of a coordination game, all participants win when they all play the same move, i.e. when they coordinate. Even in the absence of communication, people manage to coordinate in such games, at least better than chance.

> When a man loses his wife in a department store without any prior understanding on where to meet if they get separated, the chances are good that they will find each other. It is likely that each will think of some obvious

> place to meet, so obvious that each will be sure that the other is sure that it is 'obvious' to both of them. One does not simply predict where the other will go, since the other will go where he predicts the first to go, which is wherever the first predicts the second to predict the first to go, and so on ad infinitum. [...] They must 'mutually recognise' some unique signal that coordinates their expectations of each other (Schelling 1960; p. 54).

How does this come about? Based on a shared history, some places are more salient than other places. Such a salient place may act as a kind of focal point. In numerical coordination games, i.e., to try and name the same number as the others, people often select the number '1'. Apparently the number '1' stands out. Interestingly, after a series of such encounters, the salient solution to a coordination game, may slowly turn into a proper convention (Lewis, 1969). So after the third time, man and wife will simply meet at the south entrance, as they did before.

Language creates a special kind of coordination game: a signalling game (Lewis, 1969). What people in a language community try to do, is coordinate on signals and meanings. One tries to develop and maintain a code: a list of tuples $\langle s, m \rangle$ that conventionally assign signals s to meanings m and vice versa. In principle, any signal could be paired with any meaning, but it turns out there are regularities in this process. One of the most well known regularities is called Zipf's Law: expressions that are relatively more frequent, tend to be shorter or abbreviated. There is a cognitive counterpart to this regularity, that may be called the 'principle of least effort' (van Rooy and Jaeger, to appear). The principle works differently for speaker and addressee. For example, speakers tend to choose the most frequent words, because they are easier to find in memory. The effort for the addressee has to do with determining the meaning of the word. The higher the ambiguity of a word, the higher the effort for the hearer. The net effect is that marked expressions, i.e. longer or more elaborate expressions, are typically paired with an unexpected interpretation, whereas unmarked or reduced expressions are paired with an expected meaning. These

phenomena have been analysed in Optimality Theory (OT) and evolutionary game theory (van Rooy and Jaeger, to appear).

In our case, these principles explain why many initiatives or responses do double duty: they have a regular linguistic meaning, but they also serve as a 'bid' for a joint action. Similarly, an appropriate response has its regular meaning, but also signals acceptance of the bid. Only when things go wrong, which is generally unexpected, a marked expression is required: an explicit no, or a warning. In other words, the signal that represents an acceptance, has been reduced, and has become implicit.

A second possible foundation for our approach, especially for the non-verbal examples like double pass and lifting a table, is found in systematic accounts of non-verbal communication and every day gestures (Posner et al., 2003). The idea is that gestures may have developed by ritualisation of iterated body movements (Posner and Serenari, 2003). Thus salient body movements that are repeated in similar circumstances, will slowly acquire a meaning. They will come to stand for those circumstances. Posner and Serenari's observations may help to understand the conventionalisation process of the initiative-response pairs.

Non-verbal communication has received much interest from the field of artificial intelligence. According to Poggi et al. (2003) communication systems can be *codified* or *creative*. In a codified system, such as a set of symbolic gestures or a set of words, the correspondence rules between symbol and meaning are shared and remain fixed. In a creative system, memory contains a set of inference rules, that specify how a meaning can be encoded, or a how a signal can be interpreted. This is analogous to the morphological rules by which new words can be created in a language.

In our examples, the songs in the school camp are codified. The double pass tactic is creative. The tactic can be applied with players in many different positions. Rules about the relative speed of the players, the distance to their

opponents and the angle, determine whether the tactic will be interpretable as such.

A third source of ideas on signalling is the theory developed by Castelfranchi and colleagues on Behavioural Implicit Communication (BIC) (Castelfranchi, 1998; Tummolini et al., 2004). In this context, communication, as distinguished from mere interaction, is defined as a process of information exchange between sender s and receiver r, aimed at informing r. Behavioural Implicit Communication is defined by the following situation: sender s is performing a usual practical action α, but it is part of the goals motivating α that receiver r observes and understands such behaviour, and derives some meaning m from it. Thus the communicative intent is implicit in the behaviour. Clearly, all our examples of uptake classify as instances of BIC.

Tummolini et al. (2004) identify at least three different requirements for such a form of communication. The first requirement has to do with the environment. For example, action α must at least be observable, and receivers must at least pay attention. The second requirement is related to the ability of participants to interpret action α. That means among other things that participants must have a shared background, so that they can derive similar meanings. The third requirement is that the sender should be able to observe the effect that his actions have had on the others, so that he can react appropriately.

5.4 Counts As

In the previous subsections we have tried to understand some of the requirements on the relationship that underlies rule 1 and 3': that an initiative also *signals* a bid for a joint action, and that an appropriate response also signals acceptance. But in section 3.3 we even claim that some of these signalling cases are an instance of Searle's (1997) notion of constitutive rules. In other words, that an initiative *counts as* a bid for a joint action. Constitutive rules generate or constitute the social context, by establishing social or institutional facts.

For example, the rules of chess constitute the game by defining legal moves and positions. A constitutive rule applies only under certain circumstances and is intimately linked to an institution. This institution can be a person or an abstract entity such as a community of language users. Whether a rule applies depends on the jurisdiction of the institution. Thus constitutive rules are of the form "state of affairs or event x *counts as* state of affairs or event y, under circumstances c in institution i". Constitutive rules are often used to explain the changes to a legal system, as for example the transfer of ownership rights in a sale. Now one might ask two questions: can uptake rules be modelled as constitutive rules of this kind, and if so, is that necessary for the analysis?

In the case of uptake rules, we would suggest that the institution is formed by the community of language users, exemplified by the participants present. In case of rules 1 and $3'$, the institutional fact that is created, is the fact that a bid was just made, and that a commitment was just made. Moreover, the initial triggering of potential recipes for joint actions, as reviewed in the context of frame analysis above, already requires many institutional facts. Think of the establishment of a group in the schoolcamp example.

But why would such an approach be necessary? Here we have run out of arguments. As far as we know, an analysis of uptake is also possible without constitutive rules. However, some examples are more institutional than others, and are better understood when analysed in terms of 'counts as' rules. Also note that constitutive rules have both a social and a cognitive effect. In the formalisation, we have only explicitly modelled the cognitive effect, using B-rules and obligations. We have merely assumed the social effect. Possibly a more thorough analysis of uptake in terms of Normative Multiagent Systems (Boella and van der Torre, forthcoming) could demonstrate the necessity for constitutive uptake rules. This remains a topic for further research.

There has not been much work on the notion of uptake as such. Apart from Austin (1962), and Clark (1996), we know of little other research about uptake itself. In the dialogue community, ideas from Bill Mann (1988) have been worked out in more detail recently (Kreutel and Mann, 2003). Like us, Mann perceives the engagement in a joint action as a negotiation at a meta level, with bids and acceptances. Ideas of negotiating the type of protocols have also been voiced in the community that studies agent communication languages (Reed, 1998). One of the main motivations for such a mechanism, is that it may lead to a more flexible way of regulating agent interaction.

The only in-depth work we know on uptake is by Marina Sbisà (2002). Sbisà starts from the inability of classical speech act theory to account for sequential aspects. She puts emphasis on the fact that no dialogue move can produce its conventional aspect apart from its reception, i.e. the way it is taken up. In Sbisà (2002), she explores a "general form of our understanding of sequences of events which contributes to the structuring of speech acts sequences". In the same spirit as our approach, she claims that "the participants' way of reacting to each other's moves is enough for constituting or failing to constitute the required agreement". Of course the technical solution proposed – a move is framed into the so called manipulation/action/sanction-scheme, which has been proposed in narrative semiotics – differs from our recipes for joint action. An in-depth investigation of the connections between the two approaches could prove to be fruitful.

6 Conclusion

To what extent does this analysis conform to Clark's or Tuomela's accounts? To compare our uptake mechanism with the approaches discussed in Section 2, we would need to enrich our agent model with the various mental attitudes

involved in requirements (i)–(iii) and 1–7. Below we will nevertheless give a sketch. A complete logical analysis is beyond the scope of this paper.

Regarding requirements (i)–(iii), we can say the following.

(i) *common knowledge* — common knowledge of the distribution of the joint action into parts and for the allocation of parts to agents is ensured, since the projection rules project a particular recipe, which is often known by convention, and because from the appropriate responses a task allocation can be derived. Thus a combination of a projection and an accepted bid, signalled by an appropriate participatory action, counts as enough 'basis' for common knowledge that a joint action of the right type has indeed been established (Lewis, 1969). In Lewis' theory, common knowledge of various strengths is derived from a common perceptual basis, to which all participants have immediate access. The more immediately accessible the basis is to participants, the stronger the common ground that can be derived. We should stress, however, that conventions are by no means the only way to achieve common knowledge of a recipe. Explicit communication, salience, previous agreement, or even shared reasoning, can lead to the establishment of a common ground.

(ii) *joint intention* — the joint intention to achieve the purpose of the joint action is a consequence of each of the participants' intentions to achieve the purpose of the joint action, which can in turn be inferred from the offer or acceptance of the bid. In rules 6 and 7 this is expressed by the establishment of a commitment. In general, we take it that an accepted proposal conveys an intention, or a commitment. These meta level bids are no different.

(iii) *mutual support* — this point is discussed at length in section 4.2 above. Mutual support is covered by the additional obligation to accept all bids for help that are related to the joint actions that an agent has committed itself to. The obligation to assist in case of a failure, can be covered by extending the projection rules. In dialogue, this can also be observed in situations where an interlocutor is supported by another conversation participant, when

collaboratively completing a sentence, for instance.

Regarding the requirements from Tuomela, we believe that item 1 – 5 can indeed be established in an uptake mechanism. More precisely, items 2 and 5 are closely related to the aforementioned point (i). Item 3 has a straightforward connection to point (ii).

1. *right type of joint action* — we refer to frame analysis (Goffman, 1974) and accounts of scripts and frames in cognitive science (Schank and Abelson, 1977), to indicate what kinds of activities would conventionally be regarded as candidates for a joint action in a particular situation. In the formalisation, such scripts are represented by the projection and the appropriateness rules.

2. *common knowledge and success of a recipe* — this has been dealt with under (i) above. Given a good recipe, coordinated performance of the parts will secure success of a joint action, but exactly which recipes will be successful in which circumstances is an empirical matter.

3. *relevant joint intention* — this has been dealt with under (ii) above.

4, 5. *joint control* and *mutual dependence* — these points are more delicate. The natural interpretation would be to say that (the results of) joint actions are jeopardised by a failure of the uptake process. For example, when not everybody lifts the table at the same moment, someone might hurt his back. This seems to be related to a point discussed earlier, namely the fact that the behaviour induced by the joint action is more than can be explained on the basis of reactive or even anticipatory coordination. Each participant can influence the outcome, and must therefore be explicitly involved. In the formalism, this aspect is dealt with by the definition of the recipes of a joint action, and in the representation of the dependencies between the participatory actions.

6. *strength of joint commitment* — uptake can only account for one type of joint action, namely largely conventionalised joint activities, with clear scripts or frames. Although they are heavily reliant on convention and context, we

do believe that these activities involve joint intentions. That such intentions exist over and above mere rule following becomes apparent especially when things go wrong. In that case participants can try to resort to a different way of accomplishing their goals.

7. *context* — uptake is one of the ways to explain how conventional joint plans are activated by groups. In general we believe groups are created in acting jointly. So uptake requires quite a lot from the context. An initiative must be identifiable as a 'bid' to the potential participants, and by projection it must be clear which joint action is initiated. No prior agreement is necessary, but there must already be a communication channel in place, and participants must be paying attention. Once engaged in a joint plan, compliance becomes obligatory.

In summary, we have outlined an account of 'uptake' of joint actions as a kind of negotiation at a meta level. If a move which, in context, counts as a bid to initiate a joint action is accepted, we can show how the most crucial aspects of joint action hold: common knowledge of a coordination mechanism and a plan, as well as a joint intention. Along these lines, it becomes possible to set up an empirical study into the fulfillment conditions of particular joint activities.

Acknowledgements We are greatly indebted to the anonymous reviewers, whose fair and challenging comments helped to improve the paper.

References

Austin, J., 1962. How to do things with words. Harvard University Press.

Boella, G., van der Torre, L., 2004. Regulative and constitutive norms in normative multiagent systems. In: Proceedings of the 9th International Conference on the Principles of Knowledge Representation and Reasoning (KR'04). Whistler (CA), pp. 255–265.

Boella, G., van der Torre, L., forthcoming. A game theoretic approach to contracts in multiagent systems. IEEE Transactions on Systems, Man and CyBernetics - Part C.

Broersen, J., Dastani, M., Hulstijn, J., Van der Torre, L., 2002. Goal generation in the BOID architecture. Cognitive Science Quarterly 2 (3-4), 431–450.

Carletta, J., Isard, A., Isard, S., Kowtko, J. C., Doherty-Sneddon, G., Anderson, A. H., 1997. The reliability of a dialogue structure coding scheme. Computational linguistics 23 (1), 13–32.

Castelfranchi, C., 1998. Modelling social actions for AI agents. Artificial Intelligence 103, 157–182.

Clark, H., Brennan, S. A., 1991. Grounding in communication. In: Resnick, L., Levine, J., Teasley, S. (Eds.), Perspecives on socially shared cognition. APA Books, Waskington, D.C., pp. 127–149.

Clark, H. H., 1996. Using Language. Cambridge University Press.

Cohen, P. R., Levesque, H. J., 1991. Teamwork. Nous 25 (4), 487–512, special Issue on Cognitive Science and Artifical Intelligence.

Dean, T., Boddy, M., 1988. Reasoning about partially ordered events. Artificial Intelligence 36, 375–399.

Endriss, U., Maudet, N., Sadri, F., Toni, F., 2003. Protocol conformance for logic-based agents. In: IJCAI'03. Morgan Kaufmann, pp. 679–684.

Gabbay, D. M., Woods, J., 2001. Non-cooperation in dialogue logic. Synthese 127, 161186.

Goffman, E., 1974. Frame Analysis: An Essay on the Organization of Experience. Harper and Row, New York.

Grosz, B. J., Kraus, S., 1996. Collaborative plans for complex group action. Artificial Intelligence 86 (2), 269–357.

Hiel, M., 2005. Learning interaction protocols. Master's thesis, Department of Computer Science, Utrecht University.

Hulstijn, J., Maudet, N., 2003. Uptake by conditional obligations. In: Dekker, P., van Rooy, R. (Eds.), Proceedings of the 14th Amsterdam Colloquium. ILLC, University of Amsterdam, pp. 141 – 146.

Kreutel, J., Mann, W., 2003. Analysing bids in dialogue macrograme the-

ory using discourse obligations. In: Kruijff-Korbayova, I. (Ed.), Proceedings of the 7th Workshop on Formal Semantics and Pragmatics of Dialogue (Diabrück'03). Universität des Saarlandes, pp. 59–66.

Lewis, D., 1969. Convention: A Philosophical Study. Harvard University Press, Cambridge.

Mann, W. C., 1988. Dialogue games: Conventions of human interaction. Argumentation 2, 511–532.

McBurney, P., Parsons, S., 2002. Games that agents play: A formal framework for dialogues between autonomous agents. Journal of Logic, Language and Information 11 (3), 315–334.

Poggi, I., Cirella, F., Zollo, A., Agostini, A., 2003. The communicative system of touch, alphabet, lexicon, and norms of use. In: Gesture Workshop 2003. pp. 77–89.

Pollack, M. E., 1990. Plans as complex mental attitudes. In: Cohen, P. R., Morgan, J., Pollack, M. E. (Eds.), Intentions in Communication. MIT Press, Cambridge, Mass., pp. 77–103.

Posner, R., Hanke, T., Krger, R., Noll, T., Serenari., M., 2003. Berlin dictionary of everyday gestures. Tech. rep., Research Center for Semiotics, Technische Universitt Berlin, http://www.ims.uni-stuttgart.de/projekte/nite/BLAG/index.html.

Posner, R., Serenari, M., 2003. The emergence of gestures from body movements. In: Rector, M., Poggi, I., Trigo, N. (Eds.), Gestures, Meaning and Use. Universidad Fernando Pessoa, Porto.

Reed, C., 1998. Dialogue frames in agent communication. In: Demazeau, Y. (Ed.), Proceedings of the 3rd International Conference on Multi-Agent Systems (ICMAS'98). IEEE Press, pp. 246–253.

Sbisà, M., 2002. Cognition and narrativity in speech act sequences. In: Fetzer, A., Meierkord, C. (Eds.), Rethinking Sequentiality. John Benjamins, pp. 71 – 97.

Schank, R. C., Abelson, R. P., 1977. Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures. Erlbaum, New York.

Schelling, T., 1960. The Strategy of Conflict. Harvard University Press, Cam-

bridge, Mass.

Searle, J. R., 1990. Collective intentions and actions. In: Cohen, P. R., Morgan, J., Pollack, M. E. (Eds.), Intentions in Communication. MIT Press, Cambridge, Mass., pp. 401 – 416.

Searle, J. R., 1997. The Construction of Social Reality. Free Press; (January 1, 1997).

Simpson, R. C., Briggs, S. L., Ovens, J., Swales, J. M., 2002. The Michigan Corpus of Academic Spoken English. University of Michigan.

Stalnaker, R. C., 1974. Pragmatic presupposition. In: Munitz, M., Unger, P. (Eds.), Semantics and Philosophy. New York University Press, New York, pp. 197–213.

Traum, D., Allen, J., 1994. Discourse obligations in dialogue processing. In: ACL'94. Morgan Kaufmann, pp. 1 – 9.

Tummolini, L., Castelfranchi, C., Ricci, A., Viroli, M., Omicini, A., 2004. What I see is what you say: Coordination in a shared environment with behavioral implicit communication. In: Vouros, G. A. (Ed.), Proceedings of the ECAI workshop on Coordination in Emergent Agent Societies (CEAS'04).

Tuomela, R., 2000. Cooperation: A Philosophical Study. Kluwer Academic Pulishers, Dordrecht.

Tuomela, R., 2004. Joint action. In: Proceedings of the Workshop on Holistic Epistemology and Theory of Action. University of Leipzig.

van Rooy, R., Jaeger, G., to appear. Language structure: Psychological and social constraints. Synthese.

Walton, D. N., Krabbe, E. C., 1995. Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning. State University of New York Press.

Wieringa, R., Meyer, J., 1993. Applications of deontic logic in computer science: A concise overview. In: Deontic Logic in Computer Science. John Wiley & Sons, Chichester, England, pp. 17–40.