A Denotational Semantics for Deliberation Dialogues

Peter McBurney[1] and Simon Parsons[2]

[1] Department of Computer Science,
University of Liverpool,
Liverpool L69 7ZF, UK
`p.j.mcburney@csc.liv.ac.uk`
[2] Department of Computer and Information Science,
Brooklyn College,
City University of New York,
Brooklyn NY 11210, USA
`parsons@sci.brooklyn.cuny.edu`

Abstract. We present a denotational semantics for agent deliberation dialogues, i.e., dialogues over proposed actions, conducted under a broad class of interaction protocols. The semantics uses category-theoretic entities to represent deals proposed by agents and the preferences they articulate between these. The semantics is constructed jointly and incrementally by the participating agents in the course of the dialogue, and evolves with the dialogue. We consider properties of the semantics relating to deals and dialogue termination.

1 Introduction

Over the last two decades, considerable attention has been given to the design of agent communications languages and interaction protocols, and their semantics. Most of this attention has focused on the semantics of utterances in agent dialogues, rather than on the semantics of dialogues or the semantics of dialogue protocols. Speech act theory, for example, has been used to provide a semantics for individual utterances in the FIPA Agent Communications Language, FIPA ACL [10]. However, such fixed, pre-defined utterance-level semantics does not allow for the meaning of utterances to change with the context of utterance, or for the meaning of utterances to be created by the participants in the course of dialogue together. Both of these are features of human dialogues [20]. While it is possible that the semantics of dialogues and dialogue protocols are compositional, it is not obvious that this is a property of every type of dialogue or protocol.

The contribution of this paper is to present the first formal, denotational semantics for a particular class of dialogues, namely deliberations. We call this semantics a *trace semantics*. In the influential typology of human dialogues proposed by Walton and Krabbe [29], deliberation dialogues involve two or more participants seeking to agree upon an action or a course of action, actions which may or may not be undertaken by the participants. Negotiation dialogues, in the Walton and Krabbe typology, are a special case of deliberations, when the action(s) under discussion involve(s) the division of some scarce resource. Both deliberations and negotiations are distinguished from dialogues over beliefs, such as Information-Seeking dialogues and Mutual Inquiries.

A deliberation dialogue arises with a need for action in some circumstance. In general human discourse, this need may be initially expressed in governing questions which are quite open-ended, as in, *What shall we do for dinner this evening?* or *How should we respond to the prospect of global warming?* Proposals for actions to address the expressed need may only arise late in a dialogue, after discussion of the governing question, and discussion of what considerations are relevant to its resolution. When possible courses of action are proposed, they may be evaluated on a large number of attributes, including: their direct or indirect costs and benefits; their opportunity cost; their consequences; their practical feasibility; their ethical, moral or legal implications; their resourcing implications; their likelihood of realization or success; their conformance with other goals or strategies; their timing or duration; etc.

Given such complexity and multi-dimensionality, it would be possible to develop quite complex models for deliberation dialogues, such as those in [11, 15]. Our approach will be simpler than these. We will assume that the parties to the dialogue are willing participants, and that resolution of the dialogue requires all parties to agree to a proposed course of action. We further assume that the participants co-operate sufficiently to commence a dialogue together to achieve this joint agreement, although they may have mutually-incompatible objectives for the content of the agreement. Each agent may also withdraw at any time. We will then define (in Section 2) two broad classes of protocols for deliberation dialogues; our results will apply to any dialogue conducted under any protocol in the respective class. As will be seen, these results cover many deliberation and negotiation interactions.

Following the definition of the classes of deliberation protocols, we give in Section 3 some examples of them. Section 4 then presents a denotational semantics for these protocols. In the theory of programming semantics (e.g., [12]), a denotational semantics for a programming language assigns an object in a mathematical space to each well-formed statement in the language syntax. For example, the well-known possible-worlds (or Kripke) semantics defines a class of relational structures for logical languages containing modal operators. Because mathematics provides us with tools to reason about mathematical objects, such an assignment can enable us to reason about programming languages, to study the properties of languages, and to compare one language with another. In this paper, we define a denotational semantics for deliberation dialogues using the mathematics of category theory. Our formalism attempts to make precise some intuitions about agent interactions presented graphically and informally in recent work on agent negotiations, for example, [4, 16]. Section 5 will follow the semantics with an exploration of deal properties, and the paper concludes with a discussion in Section 6.

Why use category theory? Our long-term objective is a formal theory of interaction protocols which incorporates the protocols and languages studied in the agent communications community, e.g., [3], and the interaction mechanisms studied in mathematical economics, e.g., [14]. Existing semantic frameworks do not provide this single theory of all types of deliberations. For example, as mentioned above, speech act semantics provides a semantic understanding of individual utterances, but not necessarily of dialogues or protocols. The real-valued mathematical spaces studied in mathematical economics,

on the other hand, do not apply to negotiation or deliberation interactions over more general domains, or where the consequences of outcomes can not be readily quantified. Because category theory is an abstraction of mathematics itself [21], it is a plausible candidate to provide the basis for a single, unified framework for these various forms of deliberation interaction. Such a unified framework would aid understanding of the differences between protocols and potentially permit the generalization of results about specific protocols in both agent communications and mathematical economics.

2   Deliberation Protocols

We begin by defining a general class of protocols for deliberation dialogues. We assume that time is continuous, and isomorphic to the positive real numbers, but that utterances occur only at integer values, with precisely one utterance made at each integer time-point. We further assume that these protocols are specified as dialogue games, in accordance with current research in agent communications protocols, e.g., [23, 25]. In this approach, the syntax of legal utterances comprises two layers, with the lower, content layer being wrapped in a higher, speech-act locution.[1] We denote participating agents by P_i, for $i \in \mathcal{I}$ a positive integer for some finite set $\mathcal{I}$, and locution contents by lower-case letters of the Greek alphabet. We let $\mathcal{L} = \{\alpha, \beta, \ldots\}$ denote this collection of locution contents, and each of these represents an action or plan of action to be undertaken following agreement by the dialogue participants.[2] Although not strictly necessary, for ease of presentation, we assume the first field in the content of utterances is the integer time t of the utterance, and the second field in the content is an identifier P_i of the agent uttering the locution.

Definition 1: Class $\mathcal{D}$: General Deliberation Dialogue Protocols
An agent interaction protocol is a member of the class of General Deliberation Dialogue Protocols (denoted $\mathcal{D}$) if it satisfies these five conditions:

Condition 1: General Locutions
The protocol contains locutions for participants to initiate, enter and withdraw from the protocol, such as those defined in other recent dialogue game protocols, e.g., [22]. We assume the syntax of the withdrawal illocution is *WITHDRAW(t, P_i)*.

Condition 2: Specific Locutions
The protocol contains three locutions of the following form:

2.1 *PROPOSE(t, P_i, α)*, which enables the speaker, agent P_i, to propose the deal α. We further assume that utterance of *PROPOSE(t, P_i, α)* by a speaker expresses a willingness of the speaker P_i to accept the proposal α at the time t of utterance.
2.2 *ACCEPT(t, P_i, α)*, which indicates to the hearer that the speaker, agent P_i, wishes to accept the deal α, which has been the subject of a prior *PROPOSE(s, P_j, α)* locution by some agent P_j (possibly P_i), and with $s < t$.

---

[1] The FIPA ACL uses the same two-layer syntax [10].
[2] For example, the contents in $\mathcal{L}$ may represent commitments, as in [27].

2.3 *PREFER(t, P_i, α, β)*, which indicates to any hearers that the speaker, agent P_i, prefers proposal β to proposal α.[3]

Condition 3: Combination Rules
The three locutions listed in Condition 2 are subject to the following combination rules:

3.1: The instantiated locution *ACCEPT(t, P_i, α)* can only be legally uttered if there has been a prior utterance of *PROPOSE(s, P_j, α)* by some agent P_j at some time $s < t$.

3.2: The instantiated locution *PREFER(t, P_i, α, β)* may only be legally uttered if there have been prior utterances of *PROPOSE(s_1, P_j, α)* and *PROPOSE(s_2, P_k, α)* by some agents P_j and P_k at some times $s_1, s_2 < t$.

3.3 The protocol has a voting rule indicating when an agreement is reached on an action, and this results in the termination of the dialogue and execution of the action, called *the deal*. For example, for unanimous agreement, the rule could be as follows: If there is a proposal α such that all participants P_i have uttered either *PROPOSE(t, P_i, α)* or *ACCEPT(t, P_i, α)*, then the dialogue ends immediately, with the participants agreeing to execute the action or action plan represented by the deal α.

Condition 4: Transitivity of Preferences
Expressed participant preferences are transitive, i.e. utterance of the following two locutions at any times t and $t + k$ in a dialogue

> *PREFER(t, P_i, α, β)*
> *PREFER(t + k, P_i, β, γ)*

entitles a hearer to infer the following relationship:

> *PREFER(t + k, P_i, α, γ)*.

Condition 5: Reflexivity of Preferences
Participant preferences are reflexive, i.e. for every deal α, every speaker P_i is able to utter:

> *PREFER(t, P_i, α, α)*. □

In the remainder of this paper, we will assume that unanimous agreement (Condition 3.3) is required for a deal. Conditions 4 and 5 are required for the resulting mathematical structure to be a category. Note that we do not assume that every participant is able to express a preference between any two proposals. At any given time, a participant may prefer one proposal to a second, or may prefer the second to the first, or may be indifferent between them, or may not yet have determined its preference between them.

Definition 2: Class $\mathcal{D}_M$: Monotonic Deliberation Dialogue Protocols
We also define a sub-class of class $\mathcal{D}$, called Class $\mathcal{D}_M$, Monotonic Deliberation Dialogue Protocols, which satisfy all five conditions above, in addition to:

---

[3] Note that preference is not the same as private welfare: an agent may prefer one outcome to another even though the first outcome makes the agent personally worse off. In other words, an agent's preferences may incorporate social aspects of its utility.

Condition 6: Monotonicity of Proposals

Assume $\alpha \neq \beta$ are two non-identical proposals. If participant P_i utters the locution *PROPOSE(s, P_i, α)* in a dialogue, and, later in the same dialogue, utters the locution *PROPOSE(t, P_i, β)*, hearers are entitled to infer that participant P_i prefers proposal α to proposal β. In other words, for integers $s < t$, the sequence

> *PROPOSE(s, P_i, α)*
>
> $\vdots$
>
> *PROPOSE(t, P_i, β)*

is equivalent to the sequence:

> *PROPOSE(s, P_i, α)*
>
> $\vdots$
>
> *PROPOSE(t, P_i, β)*
> *PREFER(t + 1, P_i, β, α).*   □

Dialogues undertaken using protocols from Class $\mathcal{D}_M$ require that agents utter new proposals that are less preferred by themselves than any of their own previous proposals.

3   Examples

In this section we present some examples of common deliberation interactions expressed in the syntax of Section 2.

Example 1: Open-Cry Dutch Auction. A Dutch auction has a single potential seller of an item interacting with multiple potential buyers. The seller (or an auctioneer, acting on the seller's behalf) shouts successively decreasing selling prices until a buyer indicates a willingness to purchase the item at the most-recently quoted price. Using the illocutions given in Definition 1, a dialogue for a Dutch auction would have the following general form, where each successive proposed price, price-p*, is lower than the one before it,* price-(p-1)*:*

> PROPOSE(1, seller, sell-item-at-price-1)
> PROPOSE(2, seller, sell-item-at-price-2)
>
> $\vdots$
>
> PROPOSE(s, seller, sell-item-at-price-s)
> ACCEPT(s+1, buyer-k, sell-item-at-price-s).

The dialogue then terminates, with buyer-k *executing a transaction with* seller *at* price-s*.*   □

Because proposed prices are descending, this is an example of a monotonic protocol. Provided the other conditions are satisfied (i.e., Conditions 1, 3–5), then the Dutch Auction protocol would be a member of Class $\mathcal{D}_M$. Note that the syntax presented here is similar to the specification given by FIPA for these auctions [9].

Example 2: Open-Cry English Auction. *In an English auction a single potential seller of an item interacts with multiple potential buyers. The seller (or an auctioneer) shouts successively increasing prices, and buyers indicate their willingness to accept these. As prices rise, fewer buyers indicate acceptance. The item is sold to the last-remaining buyer for the most recent price. Using the illocutions given in Definition 1, a dialogue for an English auction would have the following general form, where each successive proposed price,* price-p, *is higher than the one before it,* price-(p-1):

> PROPOSE(1, seller, sell-item-at-price-1)
>> ACCEPT(2, buyer-h, sell-item-at-price-1)
>> ACCEPT(3, buyer-i, sell-item-at-price-1)
>> $\vdots$
>> ACCEPT(n1, buyer-j, sell-item-at-price-1)
> PROPOSE(n1+1, seller, sell-item-at-price-2)
> $\vdots$
> PROPOSE(s, seller, sell-item-at-price-s)
>> ACCEPT(s+1, buyer-k, sell-item-at-price-s).

The dialogue then terminates, with buyer-k *executing a transaction with* seller *at* price-s.   □

The English auction protocol is not monotonic in the sense of Definition 2, but is in class $\mathcal{D}$ if Conditions 1, 3–5 hold.

Example 3: Monotonic Concession Protocol (MCP) *Zeuthen [26, 32] described a negotiation process in which two parties each make successive proposals to one another. At each proposal, the other party can either accept the proposal, or make a counter-proposal, or withdraw. For each participant, every subsequent proposal after its first must concede something to the opponent. Thus, relative to the most recent proposal made by a participant, the next proposal made by that same participant could be no more attractive to that participant and no less attractive to the other participant.*   □

If we assume we can map "attractiveness" onto preferences in the obvious way, then the MCP is an example of a protocol in class $\mathcal{D}_M$, provided Conditions 1, 3–5 hold.

4   Trace Semantics

We now define a denotational semantics, which we call a *trace semantics*, for dialogues conducted using protocols in Class $\mathcal{D}$, using concepts from Category Theory [21]. Assume $G \in \mathcal{D}$ is a deliberation protocol in $\mathcal{D}$. Let $\mathcal{P} = \{P_1, \ldots, P_n\}$ be a finite set of n distinct agents, engaged in a deliberation dialogue conducted in accordance with protocol G, with $\mathcal{L} = \{\alpha, \beta, \ldots\}$ the topics of the dialogues (i.e., the contents of locutions). We let $g_1, g_2, \ldots$ denote dialogues — sequences of instantiated locutions — conducted by $\mathcal{P}$ under protocol G. We denote the agent index set $\{1, \ldots, n\}$ by $\mathcal{I}$. For each agent $P_i, i \in \mathcal{I}$, we assume there exists two sequences of mathematical categories:[4]

---

[4] We use the letter $\mathcal{C}$ for the public stores, since these are inspired by the Commitment Stores of dialogue games [29]; we use $\mathcal{M}$ for the private stores, since these embody mentalistic notions.

- Each $\mathcal{C}_i^t$, with t a non-negative integer, is called the *public proposal store* of agent P_i at time t, and contains objects corresponding to the proposals presented by agent P_i up to and including time t in the dialogue.
- Each $\mathcal{M}_i^t$, with t a non-negative real number, is called the *private proposal store* of agent P_i at time t. Agent P_i is assumed to commence the deliberation dialogue with private proposal store $\mathcal{M}_i^0$, which may be empty. This store contains proposals which agent P_i is considering at time t, but may not yet have revealed to the dialogue.

These categories are constructed by the following trace-semantics rules, linking dialogue statements to objects and arrows in the appropriate categories. In all categories, we label those objects corresponding to proposed deals with lower-case Greek letters, while certain other objects have mnemonic labels; arrows are labelled with lower-case Roman letters. An object labelled θ^k may be understood as the action (or course of action) θ to be agreed and executed at time k. Arrows are used to indicate preferences, with the arrow pointing to the more-preferred object. Time-stamping in this way allows us to model an agent's preferences between the same action agreed at different times. We first list the rules for the public stores:

TS1: Each agent P_i begins the dialogue with a public proposal store $\mathcal{C}_i^0$ which is empty.

TS2: An utterance of the locution *PROPOSE(t, P_i, α)* by an agent P_i at integer time t results in an object labelled α^t, corresponding to the execution of α at time t, being inserted into the public proposal store $\mathcal{C}_i^t$ of P_i.

TS3: An utterance of the locution *ACCEPT(t, P_j, α)* by an agent P_j at integer time t results in an object labelled α^t, corresponding to the execution of α at time t, being inserted in the public proposal store $\mathcal{C}_j^t$ of P_j.

TS4: For each agent P_i and for all times $t \geq 0$, every object θ^k in the public proposal store $\mathcal{C}_i^t$ of P_i has associated to it an identity arrow $id_{\theta^k} : \theta^k \rightarrow \theta^k$. This rule encodes Condition 5.

TS5: An utterance of the locution *PREFER(t, P_i, α, β)* by an agent P_i at integer time t results in an arrow from the object corresponding to α^t to the object corresponding to β^t being inserted into the public proposal store $\mathcal{C}_i^t$ of P_i.

TS6: An utterance of the locution *PREFER(s, P_i, α, β)* by an agent P_i at integer time s following at a later integer time t by an utterance of the locution *PREFER(t, P_i, β, γ)* results in an arrow from the object corresponding to α^s to the object corresponding to γ^t being inserted into the public proposal store $\mathcal{C}_i^t$ of P_i. This rule encodes Condition 4.

TS7: For protocols in class $\mathcal{D}_M$, the utterance by an agent P_i of the two locutions *PROPOSE(s, P_i, α)* and *PROPOSE(t, P_i, β)*, with integer times $s < t$, creates an arrow in the public proposal store $\mathcal{C}_i^t$ of P_i from every object corresponding to β^t to the object corresponding to α^s. This rule encodes Condition 6.

TS8: An object inserted at time s in a public proposal store remains in the store for all times $t \geq s$. An arrow a from object α to object β inserted at time s in a public proposal store remains in the store for all times $t \geq s$ unless and until an arrow b is inserted from object β to object α. The presence of an arrow $a : \alpha \rightarrow \beta$ between two distinct objects α and β in a public proposal store means there is no arrow $b : \beta \rightarrow \alpha$ in that store.

We now list the rules for the private stores:

TS9: Each agent P_i begins the dialogue with a private proposal store $\mathcal{M}_i^0$ (which may be empty).

TS10: An utterance of the locution *PROPOSE(t, P_i, α)* by an agent P_i at integer time t means that there exists $\epsilon > 0$ such that an object corresponding to α^t is in the private proposal store $\mathcal{M}_i^{t-\epsilon}$ of P_i at time $t - \epsilon$.

TS11: An utterance of the locution *PROPOSE(t, P_i, α)* by an agent P_i at integer time t results in an object corresponding to α^t being inserted in the private proposal store $\mathcal{M}_j^t$ of agent P_j, for every $j \neq i$.

TS12: For each agent P_i and each time $t \geq 0$, every object θ^k in the private proposal stores $\mathcal{M}_i^t$ of P_i has associated to it an identity arrow $id_{\theta^k} : \theta^k \to \theta^k$.

TS13: For every agent P_i and every time $t > 0$, the private proposal store $\mathcal{M}_i^t$ has a distinguished object, called ND_i^t, intended to represent *"No Deal"*.

TS14: For every agent P_i and every time $t > 0$, the private proposal store $\mathcal{M}_i^t$ has a distinguished object, called FP_i^t, an abbreviation for *"Future Prospects at t"*, intended to represent the valuation at time t by agent P_i of all possible future deals, allowing for the estimation by the agent of any uncertainty in their achievement.[5]

TS15: An utterance of the locution *PREFER(t, P_i, α, β)* by an agent P_i at integer time t means that there exists $\epsilon > 0$ such that there is an arrow from the object corresponding to α^t to the object corresponding to β^t in the private proposal store $\mathcal{M}_i^{t-\epsilon}$ of P_i at time $t - \epsilon$.

TS16: An utterance of the locution *PREFER(s, P_i, α, β)* by an agent P_i at integer time s following at a later integer time t by an utterance of the locution *PREFER(t, P_i, β, γ)* means that there exists $\epsilon > 0$ such that there is an arrow from the object corresponding to α^t to the object corresponding to γ^t in the private proposal store $\mathcal{M}_i^{t-\epsilon}$ of P_i at time $t - \epsilon$.

TS17: For every agent P_i and every time $t \geq 0$, whenever there are arrows $a : \alpha \to \beta$ and $b : \beta \to \gamma$ in the private proposal stores $\mathcal{M}_i^t$ then there is also an arrow $c : \alpha \to \gamma$ in $\mathcal{M}_i^t$.

TS18: For protocols in class $\mathcal{D}_M$, the utterance by an agent P_i of the two locutions *PROPOSE(s, P_i, α)* and *PROPOSE(t, P_i, β)*, with integer times $s < t$ means that there exists $\epsilon > 0$ such that there is an arrow from the object corresponding to β^t to the object corresponding to α^s in the private proposal store $\mathcal{M}_i^{t-\epsilon}$ of P_i at time $t - \epsilon$.

TS19: The presence of an arrow $a : \alpha \to \beta$ between two distinct objects α and β in a private proposal store means there is no arrow $b : \beta \to \alpha$ in that store.

The rules for the private stores (TS9–TS19) create a mathematical model of the private states of the participating agents. It is important to note that agents may not necessarily conform to this model in their actual decision processes when engaged in deliberation dialogues. In any case, such conformance would be in general unverifiable [30]. Rule TS17 corresponds to an assumption that the private preferences of each agent are

---

[5] Thus, for an agent engaged in utility-maximizing behavior, FP_i^t would represent its estimated maximum expected utility, evaluated at t, of all future deals believed possible by the agent P_i.

transitive. Note that we make no assumption that an agent's preferences are fixed or pre-determined. Thus, objects may enter and leave the private proposal stores of the participants throughout a dialogue, and arrows likewise may change. In other words, there is no assumed relationship between $\mathcal{M}_i^s$ and $\mathcal{M}_i^t$, for $s \neq t$. We believe this captures nicely the notion that agents may have resource-constraints on their processing powers, and so they may not consider all options at all times throughout an interaction.

Using these rules, we now define a denotational semantics for dialogues conducted under protocols in class $\mathcal{D}$:

Definition 3: Given a finite set of agents $\mathcal{P}$, a collection of locution contents $\mathcal{L}$, and a deliberation dialogue protocol $G \in \mathcal{D}$, we define the *Deliberation Trace Semantics*, or *Trace Semantics*, of a dialogue g undertaken by $\mathcal{P}$ about topics in $\mathcal{L}$ according to protocol G by the pair:
$$\langle \mathcal{C}, \mathcal{M} \rangle$$
where $\mathcal{C} = \{\mathcal{C}_i^t \mid i \in \mathcal{I}, \ t \in \mathbb{Z}^+ \cup \{0\}\}$ is a collection of public proposal stores for each agent in the dialogue, created according to rules TS1–TS8, and $\mathcal{M} = \{\mathcal{M}_i^t \mid i \in \mathcal{I}, \ t \in \mathbb{R}^+ \cup \{0\}\}$ is a collection of private proposal stores for each agent in the dialogue, created according to Rules TS9–TS19. We also call $\langle \mathcal{M}, \mathcal{C} \rangle$ a *deliberation trace* of $\mathcal{P}, \mathcal{L}$ and G, denoted:
$$\langle \mathcal{C}, \mathcal{M} \rangle \models (\mathcal{P}, \mathcal{L}, G). \qquad \square$$

Proposition 1: *Each element of $\mathcal{C}$ and $\mathcal{M}$ is a category.*

Proof. A category contains zero or more objects and zero or more arrows between objects, subject to two conditions: (a) from each object to the same object there is an identity arrow; and (b) if there exists an arrow between objects α and β and between objects β and γ, then there exists an arrow between objects α and γ [21]. These conditions are guaranteed by Rules TS4 and TS6 respectively, in the case of elements of $\mathcal{C}$, and Rules TS12 and TS16 respectively, in the case of elements of $\mathcal{M}$. $\qquad \square$

It is easy matter to demonstrate consistency of the trace semantics with respect to deliberation dialogues in $\mathcal{D}$.

Proposition 2: (Consistency) *For any finite set of agents $\mathcal{P}$, any collection of locutions $\mathcal{L}$ and any dialogue protocol $G \in \mathcal{D}$, there is a trace semantics $\langle \mathcal{C}, \mathcal{M} \rangle$ such that $\langle \mathcal{C}, \mathcal{M} \rangle \models (\mathcal{P}, \mathcal{L}, G)$.*

Proof. This is straightforward from the rules of construction above. $\qquad \square$

We can also demonstrate completeness of the trace semantics with respect to deliberation dialogues in $\mathcal{D}$. For this, we must confine attention to collections of categories satisfying the properties implied by rules TS1–TS19.

Proposition 3: (Completeness) *Suppose the two collections of categories $\langle \mathcal{C}, \mathcal{M} \rangle$, with $\mathcal{C} = \{\mathcal{C}_i^t \mid i \in \mathcal{I}, \ t \in \mathbb{Z}^+ \cup \{0\}\}$ and $\mathcal{M} = \{\mathcal{M}_i^t \mid i \in \mathcal{I}, \ t \in \mathbb{R}^+ \cup \{0\}\}$, have the following properties:*
(a) $\mathcal{I}$ is finite.

(b) $\mathcal{C}_i^0 = \emptyset$, $\forall i \in \mathcal{I}$.

(c) Each $\mathcal{C}_i^t$ is isomorphic to a subcategory of $\mathcal{M}_i^t$, $\forall i \in \mathcal{I}$ and $\forall t \in \mathbb{Z}^+ \cup \{0\}$.

(d) Each category $\mathcal{M}_i^t$ has at most a countable number of objects, $\forall i \in \mathcal{I}$ and $\forall t \in \mathbb{R}^+ \cup \{0\}$.

(e) Every object and arrow of $\mathcal{C}_i^s$ is also an object and arrow of $\mathcal{C}_i^t$, $\forall s \leq t$ integers and $\forall i \in \mathcal{I}$.

(f) There is at most one arrow between any two distinct objects in each category in the two collections $\langle \mathcal{C}, \mathcal{M} \rangle$.

(g) The total combined number of objects and arrows in the union of categories $\bigcup_{\mathcal{I}} \mathcal{C}_i^t$ is at most t, $\forall t \in \mathbb{Z}^+ \cup \{0\}$.

Then there is a dialogue g undertaken by a finite set of agents $\mathcal{P}$, about a collection of topics $\mathcal{L}$ according to a dialogue protocol $G \in \mathcal{D}$, for which $\langle \mathcal{C}, \mathcal{M} \rangle$ is the trace semantics of $(\mathcal{P}, \mathcal{L}, G)$.

Proof. [Outline] Assign a distinct agent identifier P_i to each $i \in \mathcal{I}$. Starting with $t = 1$, and then for each successive integer value of t, label the objects and arrows of $\bigcup_{\mathcal{I}} \mathcal{C}_i^t$ as follows: $\alpha(1)_i^t, \alpha(2)_i^t, \ldots$ and $a(1)_i^t, a(2)_i^t, \ldots$ etc. Do this only for objects and arrows on their first appearance in each sequence, i.e., for the smallest value of t in which the object or arrow appears. Thus the objects and arrows are indexed both by a count (in parentheses) and by the category $\mathcal{C}_i^t$ in which they first appear. It is then possible to construct a dialogue between the agents using the illocutions of Definition 1, instantiated with these labels. One can readily show that this dialogue is conducted according to the rules of a protocol which is a member of class $\mathcal{D}$. □

5   Deals

In this section we consider some of the circumstances of deal agreement. Throughout, we are assuming a finite set of agents $\mathcal{P}$, a collection of locution contents $\mathcal{L}$, and a deliberation dialogue protocol $G \in \mathcal{D}$, for the class $\mathcal{D}$ defined earlier. For simplicity, when agents are willing to accept a proposal, we ignore the time taken for each of them to express this acceptance. Since some properties depend on the nature of the participants, we first need to define a class of agents.

Definition 4: A *serious* agent P_i has the following three properties:

S1: P_i utters *WITHDRAW(s, P_i)* iff
$\forall t > s, \forall \beta^t \in \mathcal{M}_i^s$, and $\forall \alpha^s \in \mathcal{M}_i^s$, there exist arrows $\alpha^s \to ND^s$ and $\beta^t \to ND^s$ in $\mathcal{M}_i^s$.

S2: P_i utters *PROPOSE(s, P_i, α)* iff
$\exists \alpha^s \in \mathcal{M}_i^s$, such that
(i) $\mathcal{M}_i^s$ has an arrow $ND^s \to \alpha^s$, OR
(ii) $\exists t > s$ and $\beta^t \in \mathcal{M}_i^s$ with arrows $\alpha^s \to \beta^t$ and $ND^s \to \beta^t$.

S3: P_i utters *ACCEPT(s, P_i, α)* iff
$\exists \alpha^s \in \mathcal{M}_i^s$, such that $\forall t > s$ and $\forall \beta^t \in \mathcal{M}_i^s$ there are arrows $ND^s \to \alpha^s, \beta^t \to \alpha^s$ and $\beta^t \to ND^s$ in $\mathcal{M}_i^s$. □

We intend these conditions to permit agents to be insincere, i.e., to propose deals they do not wish to accept, but not to be capricious or whimsical. Condition S2(ii), for example, permits an agent to propose a deal α^s at time s with the strategic intention of agreeing a more preferred deal β^t at some future time t in the dialogue. This property enables the two following results, whose proofs are straightforward from Definitions 1, 3 and 4.

Proposition 4: *Let α^u be a deal agreed at time u, according to the voting rule of Condition 3.3. Suppose all participating agents are serious. Then $\forall i \in \mathcal{I}, \exists t_i < u$ such that $\alpha^u \in \mathcal{M}_i^{t_i}$ and $\exists s$ with $max_{\mathcal{I}}\{t_i\} \le s \le u$ such that $\forall i \in \mathcal{I} \alpha^u \in \mathcal{C}_i^s$.*   □

Proposition 5: *Let α^u be a deal agreed at time u, according to the voting rule of Condition 3.3. Suppose all participating agents are serious. Then, $\forall i \in \mathcal{I}, \exists s_i \le u$ such that $\forall t \in (s_i, u]$, there is no arrow $\alpha^u \to ND_i^u$ in $\mathcal{M}_i^t$ and, $\forall v > u$ and $\forall \beta^v \in \mathcal{M}_i^t$, for β^v possibly the same as α^v, there is no arrow $\alpha^u \to \beta^v$.*   □

Proposition 4 says that, for serious agents, deals must have been considered prior to proposal or acceptance, and must appear in the public stores of all agents before a deal is reached. Proposition 5 says that, again for serious agents, a proposal cannot become a deal at some time point if an agent prefers no deal to that proposal, or prefers some future proposal to that deal. We now define a notion of Pareto-Optimality in our semantic framework.

Definition 5: A proposal α^t is said to be *Pareto-Optimal at time t* iff $\forall \beta^t \in \bigcup_{\mathcal{I}} \mathcal{C}_i^t$, with $\beta^t \ne \alpha^t$, $\exists j \in \mathcal{I}$ such that it is not the case that there is an arrow $\alpha^t \to \beta^t$ in $\mathcal{C}_j^t$.   □

In other words, a proposal is *Pareto-Optimal at time t* precisely when, for every alternative proposal presented by this time, there is at least one participant who has not yet described the alternative proposal as preferred. Thus, the definition only concerns publicly-known proposals, and only those which have been uttered up to the time of consideration. Definition 5 is therefore a constructive definition of Pareto-Optimality. We are able to demonstrate the following result regarding deliberations between two parties using a monotonic deliberation protocol:

Proposition 6: *Let $\mathcal{I} = \{1, 2\}$ index two serious agents engaged in a deliberation dialogue using a monotonic protocol $G \in \mathcal{D}_M$. Suppose that the rules of G require that an agent P_i may only utter $\mathrm{ACCEPT}(t, P_i, \alpha)$ for the most recent proposal of agent $P_j, j \ne i$. Let α be a deal agreed at time t. Then the following are equivalent:*

  – *α is Pareto-Optimal at time t.*
  – *If $\beta \in \mathcal{C}_1^t \bigcup \mathcal{C}_1^t$ is any other proposal, distinct from α, then if $\exists s_1 \le t$ with the arrow $\alpha \to \beta$ contained in $\mathcal{M}_i^{s_1}$, then $\exists s_2 \le t$ with the arrow $\beta \to \alpha$ contained in $\mathcal{M}_j^{s_2}$, for $i \ne j$, and $i, j \in \mathcal{I}$.*

Proof. ($\Longrightarrow$) The result follows, with some care, from Definitions 2, 4 and 5. ($\Longleftarrow$) Straightforward from Definition 5.   □

We may use Proposition 6 to generate a corollary regarding Zeuthen's Monotonic Concession Protocol (Example 3 in Section 3), provided that we can map "attractiveness" onto preferences in the obvious manner.

Proposition 7: *Suppose α is a deal agreed at time t by two serious agents using the MCP. Then, α is Pareto-Optimal at t.*

Proof. Straightforward from Proposition 6 and Definition 4.   □

This proposition generalizes a result of Harsanyi [13] regarding the MCP. Our definitions of agent strategies (i.e., that agents are serious, Definition 4), of the protocol (Definition 2) and its semantics (Definition 3), and of Pareto-Optimality (Definition 5) are all more general than has usually been the case in economics. We conjecture that a version of Proposition 6 also holds with more than two participants; however, we have not yet identified the conditions under which this conjecture is true.

6   Discussion

The research reported here is original. The only previous work relating category theory with argumentation was Ambler's categorical semantics for static, monolectical (one-party) argument over beliefs [2]; in contrast, our work concerns dynamic, dialectical (multi-party) argument over possible actions. Within economics, the study of negotiation mechanisms has a long history; however, mathematical economics, even when undertaken by mathematicians, has not sought to find the most general mathematical representation for these mechanisms, but confined attention to real spaces, e.g. [5, 14]. Even in the one publication known to us where category theory was applied in mathematical economics [28], categorical methods were used to prove a result about real spaces. Our semantics is not confined to real-valued proposals, nor to those denominated in prices. In any case, the problem of defining semantics for interaction mechanisms — a very important problem for computer science — has not been considered in economics.

Within theoretical computer science, category theory has been applied to the development of game semantics for interaction, e.g., [1]. That work views interactions more abstractly than the specific deliberation dialogues of interest to us, and has not treated semantic structures as objects created and manipulated by participants in an interaction. Moreover, it has only considered very simple sets of illocutions, such as questions and answers. Finally, within category theory itself, little attention appears to have been given to sequences of categories indexed by time. The only such structures known to us are the Memory Evolutive Systems of [7], designed to model emergent phenomena in complex adaptive systems, such as ecologies; these structures are monotonic over time, which is not true in our case.

An obvious question in response to this paper is: *Why not Kripke semantics?* Our reason for proposing a categorical rather than Kripkean semantic framework is that our focus in deliberation is on preferences between alternative outcomes, rather than on the outcomes themselves. For example, a rational agent choosing between: (a) accepting a proposal; (b) suggesting an alternative proposal; or (c) withdrawing from the dialogue; would make its decision on the basis of its preferences between these options. Category theory, because it emphasizes arrows not objects, is better suited to a formalization of such

preference relationships.[6] In addition, categories provide greater scope for generalization than do Kripke frames, and have available a richer and more sophisticated mathematical theory. In particular, the category-theoretic treatment of the differential calculus [19] potentially means that a single categorical theory of agent interactions could model both argumentation interactions and economic transactions.

This paper has revealed a garden we believe to be profuse with interesting flora. Much work remains to study and exploit these delights, however. In future work we plan to explore, firstly, categorical definitions of other dialogue properties, such as other types of outcomes [8]. Secondly, we aim to consider the similarity of protocols. Our long-term objective is a formal, semantic classification of protocols to complement the preliminary classifications in [18, 24]. This should help to better understand protocol properties, such as the computational complexity of dialogues under specific protocols [6, 31]. Finally, we plan to re-visit Condition 4, the assumption of transitivity of preferences. It may be possible to do without this assumption if we map non-transitive preferences to one or more arrows representing "illegal" compositions, as in [17].

Acknowledgments

References

1. S. Abramsky and R. Jagadeesan. Games and full completeness for multiplicative linear logic. *J. Symbolic Logic*, 59(2):543–574, 1994.
2. S. J. Ambler. A categorical approach to the semantics of argumentation. *Mathematical Structures in Computer Science*, 6:167–188, 1996.
3. L. Amgoud, S. Parsons, and N. Maudet. Arguments, dialogue, and negotiation. In W. Horn, editor, *Proc. ECAI 2000*, pages 338–342, Berlin, 2000. IOS Press.
4. M. Bratu, J. M. Andreoli, O. Boissier, and S. Castellani. A software infrastructure for negotiation within inter-organisational alliances. In J. Padget *et al.*, editor, *Agent-Mediated Electronic Commerce IV*, LNAI 2531, pages 161–179. Springer, Berlin, 2002.
5. W. D. A. Bryant. *Information, Adjustment and the Stability of Equilibrium*. Macquarie Economics Research Paper 6/2000, Macquarie University, Sydney, Australia, 2000.
6. P. E. Dunne and P. McBurney. Optimal utterances in dialogue protocols. In J. S. Rosenschein *et al.*, editor, *Proc. AAMAS 2003*, pages 608–615, New York, 2003. ACM Press.
7. A. C. Ehresmann and J-P. Vanbremeersch. Hierarchical evolutive systems: a mathematical model for complex systems. *Bulletin of Mathematical Biology*, 49(1):13–50, 1987.
8. U. Endriss and N. Maudet. Welfare engineering in multiagent systems. In A. Omicini *et al.*, editor, *Engineering Societies in the Agents World (ESAW-2003)*, LNAI, 2004.
9. FIPA. *Dutch Auction Interaction Protocol Specification*. Experimental Standard XC00032F, FIPA, 2001.
10. FIPA. *Communicative Act Library Specification*. Standard SC00037J, Foundation for Intelligent Physical Agents, 2002.

---

[6] Indeed, it is possible to define categories entirely in terms of arrows, without any objects.

11. K. Greenwood, T. Bench-Capon, and P. McBurney. Towards a computational account of persuasion in law. In *Proc. Ninth Intern. Conf. on AI and Law (ICAIL-03)*, pages 22–31, New York, 2003. ACM Press.
12. C. A. Gunter. *Semantics of Programming Languages: Structures and Techniques*. MIT Press, Cambridge, MA, 1992.
13. J. C. Harsanyi. Approaches to the bargaining problem before and after the theory of games: a critical discussion of Zeuthen's, Hicks' and Nash's theories. *Econometrica*, 24:144–157, 1956.
14. P. J-J. Herings. Universally converging adjustment processes — a unifying approach. *J. Mathematical Economics*, 38:341–370, 2002.
15. D. Hitchcock, P. McBurney, and S. Parsons. A framework for deliberation dialogues. In H. V. Hansen *et al.*, editor, *Proc. Fourth Biennial Conf. Ontario Society for the Study of Argumentation (OSSA 2001)*, Windsor, Ontario, Canada, 2001.
16. N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, M. Wooldridge, and C. Sierra. Automated negotiation: prospects, methods and challenges. *Group Decision and Negotiation*, 10(2):199–215, 2001.
17. M. W. Johnson. *On Pointed Enrichments and Illegal Compositions*. Technical Report ULCS-03-010, Department of Computer Science, University of Liverpool, UK, 2003.
18. M. W. Johnson, P. McBurney, and S. Parsons. When are two protocols the same? In M-P. Huget, editor, *Communication in Multi-Agent Systems*, LNAI 2650, pages 253–268. Springer, Berlin, 2003.
19. A. Kock. *Synthetic Differential Geometry*. Cambridge University Press, Cambridge, UK, 1981.
20. S. C. Levinson. *Pragmatics*. Cambridge University Press, Cambridge, UK, 1983.
21. S. Mac Lane. *Categories for the Working Mathematician*. Springer, New York, NY, USA, 1971.
22. P. McBurney, R. M. van Eijk, S. Parsons, and L. Amgoud. A dialogue-game protocol for agent purchase negotiations. *J. Auton. Agents & Multi-Agent Systems*, 7(3):235–273, 2003.
23. P. McBurney and S. Parsons. Games that agents play: A formal framework for dialogues between autonomous agents. *J. Logic, Language and Information*, 11(3):315–334, 2002.
24. S. Parsons, P. McBurney, and M. J. Wooldridge. The mechanics of some formal inter-agent dialogues. In F. Dignum, editor, *Advances in Agent Communication*, LNAI 2922. Springer, Berlin, 2004.
25. S. Parsons, M. Wooldridge, and L. Amgoud. On the outcomes of formal inter-agent dialogues. In J. S. Rosenschein *et al.*, editor, *Proc. AAMAS 2003*, pages 616–623, New York, 2003. ACM Press.
26. J. S. Rosenschein and G. Zlotkin. *Rules of Encounter: Designing Conventions for Automated Negotiation among Computers*. MIT Press, Cambridge, MA, 1994.
27. M. P. Singh. An ontology for commitments in multiagent systems: toward a unification of normative concepts. *Artificial Intelligence and Law*, 7:97–113, 1999.
28. H. Sonnenschein. An axiomatic characterization of the price mechanism. *Econometrica*, 42(3):425–434, 1974.
29. D. N. Walton and E. C. W. Krabbe. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. SUNY Press, Albany, NY, 1995.
30. M. J. Wooldridge. Semantic issues in the verification of agent communication languages. *J. Autonomous Agents & Multi-Agent Systems*, 3(1):9–31, 2000.
31. M. J. Wooldridge and S. Parsons. Languages for negotiation. In W. Horn, editor, *Proc. 14th ECAI 2000*, pages 393–397, Berlin, Germany, 2000. IOS Press.
32. F. Zeuthen. *Problems of Monopoly and Economic Warfare*. Routledge and Sons, London, UK, 1930.