Metalanguage dynamics.

Tillmann Pross
Institute of Philosophy, University of Stuttgart
Stuttgart, Germany
tillmann.pross@philo.uni-stuttgart.de

Abstract

This paper outlines the basics of Grounded Discourse Representation Theory, a conservative extension of Discourse Representation Theory to the specific needs of goal-directed joint interactions between humans and robots. Grounded Discourse Representation Theory defines a general framework for the processing of verbal and non-verbal interaction in dialogue that combines multi-agent based control mechanisms with the apparatus of formal natural language semantics.

1 Introduction and Motivation

Central to the formalism of Grounded Discourse Representation Theory (GDRT, [Pross, 2010]) which is outlined in this paper is an agent-based theory of dynamic interpretation of semantic representations (in the sense of Discourse Representation Structures (DRS) in Discourse Representation Theory (DRT), [Kamp et al., 2010]) that makes it possible to capture the dynamics of interaction in dialogue as reciprocal influence between the object language of DRSs and the metalanguage of set-theoretic model theory against which DRSs are evaluated. In this sense, GDRT counters the objection that the dynamics of DRT-like interpretation processes "resides solely in the incremental build-up of the representations, and not in the interpretation of the representations themselves." [Groenendijk and Stokhof, 1999, p. 10]. The approach to the dynamic interpretation of representations proposed by GDRT takes a step behind the scenes of the processing of information and interaction in dialogue in that it specifies *how* an *agent* deals with the dynamics of information and interaction in dialogue. This dynamics manifests itself on the one hand in the form of the semantic representations that are constructed, maintained and used during a dialogue but also - and this is the focus of GDRT - in the model-theoretic semantics for these representations. The central idea that this discussion amounts is that the *modelling* of the dynamics of information and interactions in their application to dialogue must be *itself* dynamic. The context of human-robot interaction from which GDRT draws its basic motivation places special demands on a formalism that is to provide a robot with the necessary means and information structures. Consider the following example (1), uttered by Fred, a human, to robot Clara. Fred and Clara are situated at a table. On the table is a cube, a slat and a screw[1].

(1) Give me the cube.

With respect to the meaning of (1), traditional formal semantics would seek to derive the truth conditions of (1) with the help of which interpreters (such as Clara) of (1) could in turn evaluate (1) against a set-theoretic model theory provided to them. A suitable model theory for the analysis of this sentence should consider that (1)

[1] GDRT has been developed as a part of a project on joint action of humans and robots, where the task domain consisted of a Baufix construction kit, consisting of wooden cubes, screws, nuts, slats with holes and other parts.

is a request for the addressee to choose from her future possibilities for action the course of action which renders true that the speaker has been given a cube. A formal structure that captures future possibilities models time as a tree branching towards the future and in which links connecting successive models (representing momentary situations in various possible future developments of the present) represent basic causal transitions (=actions) from the first situation to the second situation. GDRT integrates this conception of time and action with the multi-agent interpretation of CTL* put forward in [Singh, 1994][2]. This allows us to think of (1) being true iff there exists a sequence of actions in Clara's model of reality starting right after the utterance time n and leading to a timepoint $t_1 > n$ where Fred has the cube[3].

But the *semantic* interpretation of (1) is only half the story. Uttering (1) additionally puts a pragmatic - normative - pressure on Clara to execute the actions leading to t_1 by adding this sequence of actions to her intentions. An agent interpreting (1) must translate (1) into future actions whose *realization* constitutes the proper reaction to (1) by selecting that part of future action in which those actions are performed and in which their final result is therefore true as well, thereby verifying (1). The problem here is that the successful interpretation of (1) is to be determined with respect to possible future developments of reality that we - as designers of robot Clara's processing formalism - cannot foresee in advance and which we consequently cannot capture in terms of models that can be specified *before* the utterance time. Instead, the *dynamics* induced by the interpretation of the information conveyed with (1) pertains to the extraction of instructions how the model of reality against which (1) is to be evaluated must be shaped by future action in order that the truth conditions expressed by (1) are rendered true.

Let me explicate the point with another example, where Fred announces her future plans to Clara.

(2) I am going to build a bike.

Put a simple way, (2) is true iff the information that Fred is going to build a bike is contained in the model against which (2) is evaluated by Clara. However, if we - as designers - would have captured this *information* before the utterance time and provided it to Clara, the utterance would have a very different status, as the information conveyed by (2) would be already available to Clara. Instead, the interpretation of (2) requires Clara to adapt her *modelling* of the future in such a way that it captures the choice and commitment of Fred to those future courses of action that bring about a bike. That is, the interpretation of (2) as informative utterance requires Clara to alter both the universe of her model of reality (adding an individual 'bike') and the interpretation function (adding a (set of) ordered pairs $\langle a, b \rangle$ to the extension of 'build', where a is an agent and b a bike) of the model against which she interprets (2).

What we see here - and saw in the interpretation of (1) - is another dimension of interpretation dynamics than the dynamics which is and can be captured by Standard DRT. It is a dynamics that lies 'outside' of the scope of semantic representations themselves - the object-language - but pertains to the metalanguage of model theory. The way in which GDRT accounts for the metalanguage dynamics that is induced by such interpretation strategies for utterances that concern the future is closely parallel to the notion of context change potential in Dynamics Semantics. GDRT considers meaning, one might say, as model change potential. Two types of model-changing actions have to be distinguished: agent-internal actions that directly affect an agent's models (as planning) and agent-external actions that indirectly affect an

[2]The combination of DRT and CTL* has, besides [Singh and Asher, 1993], not received the attention it probably deserves.

[3]The account of pragmatic utterance meaning in GDRT shares its basic assumptions with plan-based theories of speech acts [Cohen and Perrault, 1986], where GDRT employs the account of [Singh, 1998].

agent's models via the feedback of the results of perceived action (e.g. the building of a bike or passing over an object). Then the crucial question is how to capture the reciprocal dependency of internal and external action and the interpretation of utterances against dynamic models. The answer of GDRT has three components. At first, we - as designers - need to specify how utterances relate to actions. To this end, GDRT introduces temporal anchors to the object language of DRT and defines their formal semantics in terms of the metalanguage of branching time structures. Temporal anchors render possible to combine semantic - truth-conditional - and pragmatic - success-based approaches of natural language meaning. Second, we must elaborate how the anchor-based interpretation of an utterance specifies courses of action that constitute sequences of future actions as appropriate reaction to the utterance. To this end, GDRT employs a formal theory of planning couched in the metalanguage. Third, we need to explicate how an agent like Clara can *employ* these mechanisms to interpret utterances. For that purpose, GDRT employs a BDI-interpreter[4] to control the execution of intentions on the basis of beliefs, goals and plans, where the interpretation of DRSs is integrated into the overall framework of BDI-based control.

2 Overview of GDRT

In the following, for the analysis of example (1) in the framework of GDRT, I introduce only those fragments of GDRT necessary for the analysis of (1), a full account of GDRT can be found in [Pross, 2010]. Given an agent x at time t_i, I call the agent's configuration of her semantic representations (DRS) and models (External Presentation Structure, EPS) at t_i the 'cognitive structure' CS of x at t_i, $CS(x)(t_i) = \langle DRS(x,t_i), EPS(x,t_i)\rangle$. Figure 1 pictures the cognitive structure of Clara right after she has constructed a DRS K_1 for (1), where her EPS presents initial information delivered by her object recognition about the objects on the table in front of her. As Clara is assumed to be cooperative, she has invoked an attempt to interpret DRS K_1, presented as an internal action $\mathsf{int\text{-}a?}K_1$ in the EPS.

The main DRS K_1 in figure 1 represents (1) with the help of two related events e_1 and e_2. e_1 represents the intentional nature of the action in question, namely that the plan corresponding to giving the cube is or should be amongst the active intentions of Clara. The relation between the propositional attitude INT and the content of the attitude DO is represented by the nested use of temporal anchors for e_1 (denoting choice and commitment wrt. e_1) and e_2 (denoting the plan for giving a cube)[5]. Temporal anchoring in GDRT is built upon the idea that on the one hand, humans structure their temporal perception along "goal relationships and causal structures" [Zacks and Tversky, 2001] and on the other hand, human beings structure their future experiences in terms of the causal relations that they perceive between them and their plan-goal structures or intentions. Formally, this dual role of temporal entities is captured by the introduction of temporal anchors (extending the theory of anchors put forward by [Asher, 1986, Kamp, 1990]), two-place relations between discourse referents for eventualities (which I call floaters) and representations of causal, plan-goal and intentional structures (which I call sources). The temporal structure imposed by the additional constraints on e_1 and e_2 represents the present tense of (1), where e_2 starts after e_1 and e_1 starts after n, Clara's present now defined by the utterance time t_0. DRS K_3 represents the goal of the plan for giving a cube. The goal of the plan denoted by e_2 is represented with the state s_1. Besides the anchors for e_1, e_2, s_1 and n, also the

[4] I do not introduce the functioning of a BDI-interpreter here but refer the reader to e.g. the simple version given in [Singh et al., 1999]

[5] Intentions are analyzed as events as they denote dynamic control over the execution of plans by the BDI-interpreter.

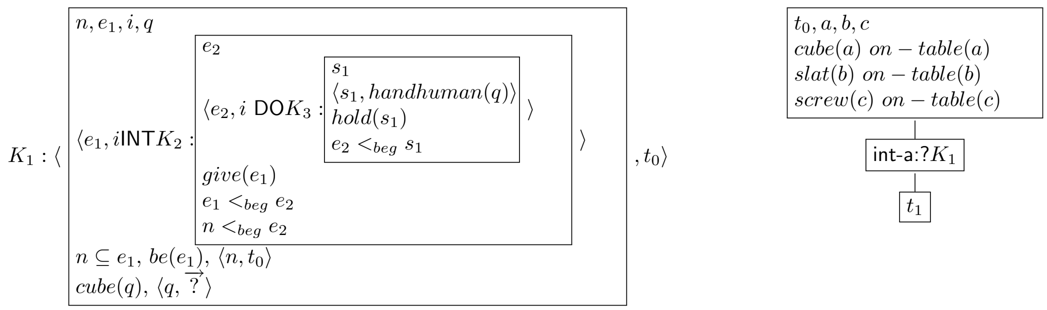

Figure 1: The cognitive state of Clara at t_1, $CS(i)(t_1)$: Representation K_1 of utterance (1) "Give me the cube." (left) making up $DRS(i)(t_1)$ and presentation of Clara's EPS at t_1 (right), $EPS(i)(t_1)$. Note that from an epistemic point of view, Clara's EPS presents a complete picture of the model-theoretic information available to her but further steps of interpretation will extend Clara's EPS.

discourse referent for 'the cube' has an anchor. The initial representation of (1) in GDRT has an *unresolved* anchor for 'the cube', i.e. the referential source of the floater x is unspecified in that it lacks a specification of the object of reference, represented as a variable anchor source ?. The fact that 'the cube' has been identified by means of definite description is captured by the arrow over the variable source ?. This arrow constrains interpretations of K_1 by requiring an unique anchor source for q. The interpretation process for a DRS such as K_1 spelled out in the following seeks to *resolve* both temporal and non-temporal anchors to metalanguage entities (objects for things, temporal structures for eventualities) of the model theory EPS against which the representation is evaluated. In the following, I call discourse referents for eventualities time-individuals and other discourse referents thing-individuals. Unresolved anchors are called variable anchors.

Clara's interpretation attempt of K_1 proceeds along the following general procedure. Suppose that $\langle K, t_j\rangle$ is a DRS belonging to $DRS(x, t_i)$ - normally t_j would be the instant 'now' and that is what I will assume here - and that x attempts to interpret K at t_i. As a first step x must find anchor sources for all those thing anchors of the anchor set *Anch* of K where anchor sources are variable. Each such variable anchor source is of one of two sorts: (1) direct[6] or (2) anaphoric, in addition, each anchor may be constrained with a definiteness constraint. If s is a variable direct anchor source, then a (non-variable) anchor source that can replace it must be an object from $EPS(x, t_i)$; when the variable anchor source is anaphoric, then a non-variable anchor source replacing it must be a thing-individual from $DRS(x, t_i)$. If for any variable thing-anchor no suitable sources can be found, then the interpretation of K stops (but can be continued, see section 5 below). Suppose that it is possible to find a suitable non-variable anchor source to replace each of the variable thing anchor sources occurring in anchors of K. Then there will be a nonempty set G of functions g each of which is defined on the set of 'floaters' of anchors in K and maps each floater onto a suitable non-variable anchor source. In the next step of interpretation, each function g in G can be used to identify the time-individuals from $\langle K, t_j\rangle$ by checking whether their branching-time semantics can be embedded into $EPS(x, t_i)$ at t_j. If this step succeeds, the respective g is stored in a set $F \subseteq G$ to check in a final step whether K as a whole has at least one successful anchoring

[6] I exclude the distinction between external and internal anchors that is drawn in the full version of GDRT, as this would involve additional elaborations on symbol grounding and object recognition.

$h \subseteq F$ that enables identification of all conditions $c_1, \ldots, c_n \in K$ with respect to $EPS(x, t_i)$ at t_j[7]. If there exists such a successful anchoring for K, I say that K has a successful 'plain' interpretation. 'Plain' because no manipulations of $EPS(x, t_i)$ were necessary.

If the plain interpretation of K as described in the last section fails, i.e. no successful anchoring of K could be established, 'reactive' interpretation comes into play - this what is required in the case of interpreting (1). It could be the (pragmatic) meaning of K that $EPS(x, t_i)$ has to be changed by the interpreter of K to $EPS(x, t_k)$ with $i < k$ in order to render possible a successful anchoring of K in $EPS(x, t_k)$[8]. The appropriate reaction is guided in particular by the time-individuals contained in K. These time-individuals specify a course of action (via their branching-time semantics) which is to be executed in order to bring about the conditions that render a successful plain interpretation of K possible. That is, in response to a failed plain interpretation of K with respect to $EPS(x, t_i)$ at t_j, the interpreter of K should perform some actions which result in $EPS(x, t_i)$ being transformed into a model structure $EPS(x, t_k)$ which allows for a successful plain interpretation of K at t_k. Technically, this is achieved with the formulation of a semantic (object-language) *and* a pragmatic (metalanguage) identification of time-individuals, specifying the conditions that identify time-individuals in plain interpretation mode (corresponding to classical truth-conditional semantics) and in addition the actions which are to be undertaken in order to make a given time-individual 'true' via an execution of reactive interpretation (corresponding to the unfolding of the pragmatic impact of the utterance)[9].

3 An application of GDRT

The procedure outlined in the last section is applied to (1) in figure 2, where formal details are provided in the next section. For the analysis of (1), the required information for the resolution of the variable anchor source for 'the cube' is given with the following simplified semantic-pragmatic concept, where the [SEM] part specifies the DRS representation and the [PRG]-part the identification conditions in metalanguage associated with the discourse reference marker x for 'cube', where I call 'cube' the *handle* of x.

Sem-Prag-Concept 1 *cube*

$$SEM \; \boxed{\begin{array}{l} x \\ cube(x),\ \langle x, a \rangle \end{array}}$$

$$PRG \; \boxed{\begin{array}{l} a \\ cube(a) \end{array}}$$

A rudimentary semantic-pragmatic concept for a time-individual adapted to the use of 'give' in (1) can be stated as follows, where the [SEM] part specifies the semantic contribution of 'give' and the [PRG] part a metalanguage specification of the pragmatic profile of 'give', a plan[10].

Sem-Prag-Concept 2 *give*

[7] Complex conditions are not discussed here, but are analyzed in [Pross, 2010] in accordance to the dynamic semantics of DRT.

[8] The decision in which cases reactive interpretation is allowed is not discussed here. However, there are some basic cases in which reactive interpretation is not allowed, e.g. if the reaction concerns the manipulation of agent's own history. That is, a question such as "Did you build a bike?" should not receive an reactive interpretation.

[9] The two options of interpretation (reactive/plain) come close to the distinction between declarative and imperative semantics [Gabbay, 1987] or from another point of view, conceptual and procedural meaning [Sperber and Wilson, 1993]

[10] Serious attempts of modelling 'give' must of course be more fine-grained and specify invocation, context and feedback conditions for [PRG] and a more detailed connection between [SEM] and [PRG]. I also exclude the contribution of tense and aspect which is spelled out in detail in [Pross, 2010] as well as the integration into a syntactic framework such as the lambda-calculus

$$SEM \quad \boxed{\begin{array}{l} x, e \\ give(e) \\ \langle e, xDO \boxed{\begin{array}{l} y, s \\ \langle s, handhuman(y)\rangle \\ \langle y, a\rangle \end{array}} \rangle \end{array}}$$

PRG t_0 - ext-a:grasp(a) - t_1 - ext-a:present(a) - t_2

4 Formal Definitions

This section sketches the basic formal ideas underlying the account of semantics and pragmatics in GDRT as discussed with examples (1) and (2). In order to capture the dynamic nature of the EPS structure and consequently of the model theory against which DRSs are evaluated, we first need a specification of EPSs and the branching structure of EPSs.

Definition 1 *EPS vocabulary*

- *A set T_R of EPS reference markers for things: $\{a_1, \ldots, a_n, \ldots\}$*
- *For each $n > 0$ a set Rel^n of n-place predicate constants for handles $\{C_1, \ldots, C_m, \ldots\}$*
- *A set **Times** of EPS times $\{t_0, \ldots, t_n, \ldots\}$*[11]

Definition 2 *Syntax of EPSs and EPS conditions*

1. *If $U \subseteq T_R \bigcup$ **Times**, Con a (possibly empty) set of conditions then $\langle U, Con\rangle$ is an EPS*
2. *If $R_1 \in Rel^n$ and $a_1, \ldots, a_n, \ldots \in T_R$ then $R_1(a_1, \ldots a_n)$ is an EPS-condition*
3. *A time-indexed EPS is a tuple $\langle t, \langle U, Con\rangle\rangle$.*

The branching structure of time-indexed EPSs - the EPS structure - can be formally described in terms of a modal model structure (cf. [Singh, 1994], [Emerson, 1990]).

[11] The numerical subscripts are used only to clarify the design of the EPS structure.

Definition 3 *EPS Structure*
An EPS structure is a tuple $E = \{\boldsymbol{T}, I, Actions\}$ of an agent x at time t, where

- *$\boldsymbol{T} = \langle <, \boldsymbol{Times}_A\rangle$ is a time structure of an agent x at time t, where $\boldsymbol{Times}_A \subseteq \boldsymbol{Times}$ and $\boldsymbol{T}$ is a labeled directed graph with node set $\boldsymbol{Times}_A$, arc set Actions and node labels given by I. In addition, we require the graph of $\boldsymbol{T}$ to be a tree.*
- *I associates times $t \in \boldsymbol{Times}_A$ with EPSs, i.e. I is a function from $\boldsymbol{Times}_A$ to EPSs according to definition 2*[12]*.*
- *Actions is a function from pairs $\langle t, t'\rangle$ of adjacent members of $\boldsymbol{Times}_A$ to the set of internal and external actions available to an agent.*

For the use of EPS structures as models for the interpretation of the language of DRSs, it is useful to convert the 'raw form' of the EPS structure into the logically more manageable form of sets and assignment functions. Here we make use of the function I from EPS times to EPSs (definition 3.) That is, with a given EPS structure $E = \{\mathbf{T}, I, Actions\}$ of an agent x at time t we are provided with the following sets:

Definition 4 *EPS sets of an agent x stored in her EPS structure at t*

- *The set of EPS times $\boldsymbol{Times}_A = \{t_0, .., t_n, \ldots\}$ (occurring in $Dom(I)$*[13]*)*
- *The set of EPS things **Things**$= \{a, b, c, \ldots\}$ (occurring in the universes of EPSs in $Ran(I)$)*

[12] That is, the interpretation I of an EPS-time $t \in \mathbf{Times}_A$ is a function from time indices t to EPSs as defined by a set of time-indexed EPSs $\langle t_1, \langle U_1, Con_1\rangle\rangle, \ldots \langle t_n, \langle U_n, Con_n\rangle\rangle, ...$

[13] I write $Dom(F)$ for the domain and $Ran(F)$ for the range of a function F.

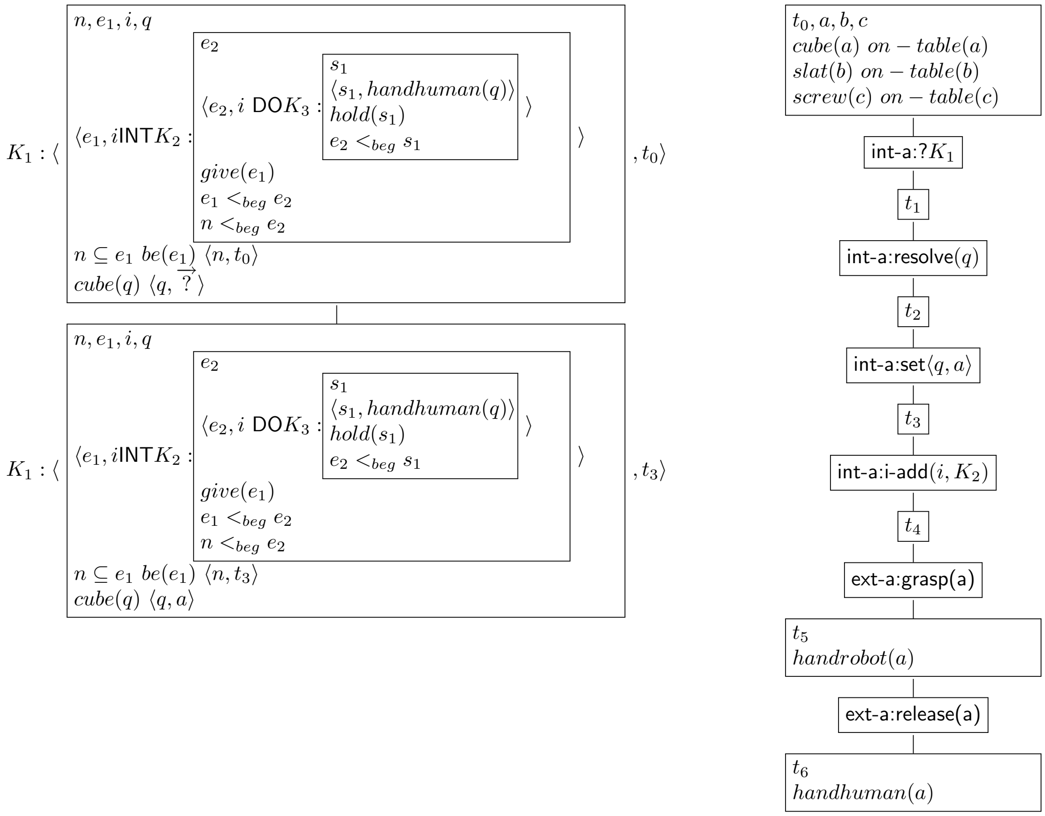

Figure 2: Processing of example (1) "Give me the cube" by Clara in the framework of GDRT. The figure shows her cognitive structure $CS(i)(t_5)$. Earlier stages of processing are recorded in $CS(i)(t_5)$, representing a discourse history. Clara's interpretation attempt int-a:?(K_1) at t_1 invokes a plan that pushes K_1 to the list of DRSs to be interpreted. Next, it is checked whether K_1 contains variable anchor sources. The variable anchor source ? in K_1 at t_1 triggers a plan int-a:resolve(q) for the resolution of this source. The anchor source for q is resolved at t_3 by a successful unification of the semantic-pragmatic concept 1 under consideration of the definiteness constraint on q with $EPS(i)(t_1)$. Once the variable anchor source for q has been resolved to a, K_1 is passed over to the main interpretation process. As a plain embedding of K_1 fails at t_3 - up to t_3 there is no temporal structure in which the [PRG] of 'give' could be embedded, reactive interpretation of K_1 is executed. This results in an extension of Clara's EPS at t_3 with the pragmatic part [PRG] of the semantic-pragmatic concept for 'give'. As the EPS-path $t_4 - t_5 - t_6$ is added to Clara's intentions by the command i-add, Clara's BDI-interpreter executes this intention. Finally, with t_5, K_1 can be embedded into $EPS(i)(t_6)$ and Clara realizes a successful interpretation of (1), where Fred has the cube in his hands. Note that the DRS with unresolved anchors at t_1 differs from the DRS with resolved anchors at t_3 in its now-anchor. The information in the EPS presents only updates to the EPS, i.e. newer information replaces old information if there exist incompatibilities between an existing and a new EPS condition concerning an already registered thing (e.g. a move to a different location). Similarly, new information (e.g. a thing appearing for the first time in the agent's area of vision) is presented in the EPS.

- *The set of EPS properties* ***Properties***= $\{p_1, \ldots, p_n, \ldots\}$ *(occurring in EPSs in* $Ran(I)$*)*
- *The set of EPS atomic actions* ***Actions***= $\{\mathsf{a}_1, \ldots, \mathsf{a}_n, \ldots\}$ *(occurring in* $Ran(Actions)$*)*

Next, we define a set of functions that assigns sets of (tuples of) agents and/or things and times to subsets of I as specified in definition 4.

Definition 5 *EPS assignment functions*

- *A function* $\boldsymbol{T}$ *that assigns EPS structures to an agent* τ *at* t*, i.e. the time structure of an agent at* t*:* $\boldsymbol{T}(\tau)(t)$
- *A function* S *that assigns Scenarios to an agent* τ *at* t*:* $S(\tau)(t)$.

A scenario is an EPS structure $\{\boldsymbol{T}, I, Actions\}$ *such that* $<$ *is a linear ordering. Let* $R = \{\boldsymbol{T}, I, Actions\}$ *be an Model structure.* $S = \{\boldsymbol{T'}, I', Actions'\}$ *is a scenario of* R *iff* S *is a scenario,* $I' \subseteq I$ *and* $\boldsymbol{T'}$ *is a substructure of* $\boldsymbol{T}$*. If* S *is a scenario of* R*, then there will be* $t, t' \in Dom(I)$ *so that* $\boldsymbol{T'}$ *is the segment* (t, t') *of* $\boldsymbol{T}$*.* t *is called the starting point of* S *in* R*. Of particular interest are those scenarios* S *of* R *in which* t' *is a leaf of* $\boldsymbol{T}$*. When* t *is the starting point of* S *then we write '*$S(t)$*'.* $\boldsymbol{S}(t) \subseteq \boldsymbol{T}$ *denotes the set of all scenarios of* $\boldsymbol{T}$ *at* t*. The notation* $[S; t, t']$ *denotes an inclusive interval on a scenario* $\boldsymbol{S}$ *from* t *to* t' *with* $t, t' \in S$ *and* $t \leq t'$*.*

- *A function* P *that assigns Plans to an agent* τ *at* t*:* $P(\tau)(t)$.

A plan P *of some EPS structure* R *has a starting point* t*. Its time structure is a subtree of* $\boldsymbol{T}$ *with* t *as root. When* t *is the starting point of* P*, then we write '*$P(t)$*'.* $P(t) \subseteq \boldsymbol{T}$ *denotes the set of plans at* t*.* $[P; t, t_1]$ *denotes a plan starting at* t *with its goal located at* t_1*.*

- *A function* $\boldsymbol{P}$ *that assigns* ***Properties*** *to (tuples of) EPS-things* $\langle a_1, \ldots, a_n \rangle$ *at* t*:* $\boldsymbol{P}(a_1, \ldots, a_n)(t)$

The information structures of the BDI-interpreter of an agent x *provides us with sets of plans (the knowledge base) and intentions (the current configuration of the intention stack), sow we can define*

- *A function* ***Attitudes*** *that assigns attitudes of a certain type* ϕ *(*DO *or* INT*) to an agent* x *at* t*:* $\boldsymbol{Attitudes}(\phi)(x)(t)$

The syntactic definition of DRSs in GDRT follows the standards of DRT [Kamp et al., 2010], so I skip this step and directly move on to the main point, the use of EPS structures as models for DRS interpretation. In defining the models for DRS interpretation we have to consider an important point that distinguishes the models for DRS interpretation in GDRT from the models in Standard DRT. As the models an agent can employ for DRS interpretation are derived from the agent's *current* EPS, those models only present the agent's current information about the state of affairs. The 'indexed' nature of the models for DRS interpretation is captured by recording the agent from whose EPS the model was derived and the time at which this was done. Mindful of this consideration, a model M for the semantic definition of the language of DRSs could be defined as follows.

Definition 6 *A model* M *at a time* t *of an agent* x *is a tuple*
$M(x)(t) = \langle P, S, \boldsymbol{T}, \boldsymbol{P}, \boldsymbol{PRG}, \boldsymbol{Things}, \boldsymbol{Attitudes} \rangle$

M differs from traditional models (e.g. those for the interpretation of Standard DRSs) in that there is no 'interpretation function' included, i.e. a *pre*defined function that maps predicates and individual constants to their model-theoretic counterparts. Instead, the concept of an 'interpretation function' is replaced by two components. First, this is the **PRG** function, which contains the pragmatic profiles associated with the handles of discourse reference markers in the respective semantic-pragmatic concepts. The identification conditions provided by the $[PRG]$ part of an individual guide the transformation of a DRS with variable thing anchor sources to a DRS with a set of possible thing individual anchor source resolutions. The other component is the identification procedure for time-individuals, which determines at runtime the 'extension' of time-individuals in that it identifies sets of thing individual anchor sources among the set of possible thing anchor resolutions that satisfy the identification conditions $[PRG]$ of the time-individuals with respect to $\uparrow M$. In addition, we have to consider the fact that adding a DRS to $DRS(x)(t)$ may result in a revised set of referents, conditions and anchors. So a successful anchoring must be defined with respect to the existing anchors of $DRS(x)(t)$. That is, suppose a sequence of DRSs in $DRS(x)(t)$ consisting of a set of reference markers $U \subseteq Ref$, a set of conditions $Con = \{C_1, \ldots, C_n\}$ and a set of anchors $G = \{A_1, \ldots, A_m\}$ is given. An update to $DRS(x)(t)$ with K will result in the sets U_K^{update}, Con_K^{update}, G^{update}, where G^{update} is a successful anchoring of K with respect to M iff G^{update} extends G in a way that it identifies the floaters in U_K^{update} with EPS entities given the pragmatic identification conditions $[PRG]$ associated with the floaters. Formally, a successful anchoring is defined as follows:

Definition 7 *Successful anchoring*[14] *of a DRS K*
Given sets of possible anchorings G, H, a model M and a DRS K

- $\langle\!\| G, H \|\!\rangle \models_M K$ *iff* $G \subset_D H$ *and for all* $\gamma \in Con_K : H \models_M K$, *where* $A \subset_D B$ *reads as: 'the domain of A is a subset of the domain of B'*
- $G \models_M K$ *reads as: G successfully anchors K in M and*
- $\langle\!\| G, H \|\!\rangle \models_M K$ *reads as: H extends G to a successful anchoring of K in M.*

[14] Successful anchoring mirrors the notion of a verifying embedding in DRT.

Definition 8 *Successful interpretation of a DRS K*

- *A DRS K has a successful interpretation in a model M iff there exists a successful anchoring G for K in M that extends the empty anchoring ξ.*
- *I write $\models_M K$ iff there exists a successful anchoring G such that $\langle\!\| \xi, G \|\!\rangle \models_M K$.*
- *When $G \models_M \gamma$, where γ is a DRS-condition, I say that G identifies γ in M.*

In addition, interpretations can be determined with respect to a time, a scenario, a plan and a model, which will be written as $\models_{M,S,P,t} K$. In the EPS (e.g. as part of a plan), an interpretation attempt of a DRS K can be triggered by the command int-a $:?K$. *The EPS constituents which were identified as a successful interpretation of a DRS K with respect to a model and a time are denoted by $[K]_{M,t}$. If the respective EPS constituents have not been identified yet, $[K]_{M,t}$ triggers an interpretation attempt of K,* int-a $:?K$.

Thing-individuals are identified as follows.

Definition 9 *Identification of thing-individuals.*

- $\langle x, a \rangle \models_{M,t}$ *handle*(x) *iff* $PRG_{handle}(x) \in \boldsymbol{P}(a)(t)$

Constraints imposed by definite descriptions are captured by the following clause.

- $\langle x, \overrightarrow{source} \rangle \models_{M,t}$ *handle*(x) *iff there is exactly one source with which* $PRG_{handle}(x)$ *can be identified in M at t.*

Time-individuals in present tense are resolved via the following clauses, where the reactive interpretation of a time-individual is formulated in terms of metalanguage actions, i.e. the addition of beliefs (b-add), goals (g-add) or intentions (i-add).

Definition 10 *Identification of time-individuals in present tense.*

- $\langle s, R(x_1, \ldots, x_n)\rangle \vDash_{M,t}$ *handle*(s)
plain: *iff* $\exists G = \{\langle x_1, a_1\rangle, \ldots, \langle x_n, a_n\rangle\}$ *sth.* $PRG_{handle}(a_1, \ldots, a_n) \in \boldsymbol{P}(a_1, \ldots, a_n)(t)$*;*
reactive: $\textsf{b-add}(x, t, PRG_{handle}(a_1, \ldots, a_n))$

- $\langle e, x\textsf{DOK}\rangle \vDash_{M,S,P,t}$ *handle*(e)
plain: *iff* $\exists[S; t, n] \in \boldsymbol{S}(x)(t)$ *and* $\exists[P; n, t_1] \in \boldsymbol{T}(x)(n)$ *sth.* $(S \cup P) \in PRG_{handle}(e)$ *and* $\vDash_{M,t_1} K$ *and* $[K]_{M,t} \in \boldsymbol{Attitude}(Do, x, t)$*;*
reactive: $\textsf{g-add}(x, PRG_{handle}(e))$

- $\langle e, x\textsf{INT}K\rangle \vDash_{M,S,P,t}$ *handle*(e)
plain: *iff* $\exists[S; t, n] \in \boldsymbol{S}(x)(t)$ *and* $\exists[P; n, t_1] \in \boldsymbol{T}(x)(n)$ *sth.* $(S \cup P) \in PRG_{handle}(e)$ *and* $[K]_{M,t} \in \boldsymbol{Attitude}(Int, x, t)$*;*
reactive: $\textsf{i-add}(x, PRG_{handle}(e))$

5 The logic behind GDRT

With this picture of dynamic partial models in mind, consider another utterance of Fred to Clara.

(3) Show me all cubes.

In the intended application scenario of GDRT, an agent x will always find herself in a specific situation in which she is supposed to evaluate her semantic representations by default[15]. We - as designers - do not want that the agent interprets (3) as involving quantification over an infinite set of possibly existing cubes but as pertaining to the possible anchors for cubes provided by the current model of reality. However, the staged design of interpretation in GDRT supports the implementation of 'switches' between this situation-bounded interpretation and non-situation-bounded interpretation. At each level of interpretation via the resolution of anchors, the agent can adopt different logical attitudes towards the interpretation of DRSs. That is, depending on the situation, different logics (classical or non-classical) can be employed by an agent for the semantic interpretation of DRSs. GDRT models can be incomplete in several ways. First, the extensions (the referents) of DRS thing-individuals may be unknown to an agent and thus missing in the agent's modeling of reality. Second, the extensions of DRS time-individuals may be unknown to an agent and thus are not contained in the agent's models. Third, the models against which an agent evaluates DRSs involving quantification over thing-individuals have finite domains. At first sight, these limitations seem to be in conflict with two fundamental assumption of bivalent formal semantics. First, that models are *complete* in that they include all information that is relevant with respect to evaluation; with respect to complete models, a sentence evaluates to either true or false. Second, that models provide *infinite* quantification domains. The account of this problem in GDRT makes use of the possibility to intervene into the interpretation process of semantic representations. That is, in principle GDRT allows an agent to take different logical attitudes towards the interpretation of DRSs depending on the context of interpretation such as supervaluation semantics [van Fraassen, 1966] or Kleene Logic [Kleene, 1952] to deal with incomplete information due to unresolved anchors. However, the practically motivated default account of incomplete information implemented in GDRT forces an agent to put on hold the interpretation of DRSs until the given models have been extended with the information necessary for the resolution of anchors of a DRS. Extensions of models may result from asking for more information (concerning e.g. the reference of thing-individuals) or bringing about certain state of affairs (concerning e.g. the reference of time-individuals). With respect to the finite quantification domains GDRT models provide, it is

[15] In its semantic conception of evaluation in specific situations, GDRT is reminiscent of the information limitation proposed by situation semantics [Barwise and Perry, 1983]

possible to drop the restriction of GDRT quantifiers to the closed world of the robot's finite model structures by switching from finite to infinite quantification domains with the definition of an 'inflated' model $\uparrow M$ obtained from a finite model M by adding a finite or infinite set of objects Unk (for Unknown) to the set **Things**, resulting in the set $\uparrow$ **Things** replacing **Things** in M[16]. Finally, given the limitation of the intended application scope of GDRT and its primarily practical motivation, issues of e.g. decidability may not be as relevant from that specific point of view as they are from the more general metaphysical point of view.

6 Comparison and Conclusion

While there exist numerous *semantic* approaches to dialogue processing, the in-depth discussion of the semantics-*pragmatics interface* and GDRT's integration of DRT-based formal semantics and pragmatic planning with temporal anchors is new to the literature. Consequently, it is difficult to compare GDRT with other approaches to dialogue processing. Prominent pragmatic accounts based on plan recognition are limited to propositional logic (to name some:[Cohen and Perrault, 1986, Grosz and Sidner, 1986, Pollack, 1990, Singh, 1994] and do not explicitly spell out the connection between propositional planning and complex semantic representations of natural language whereas GDRT integrates planning into the formal semantics of complex DRSs.

On the other hand, approaches to discourse processing that are built on top of DRT such as SDRT [Asher and Lascarides, 2003] tackle the problems presented in this paper on a different level of analysis than GDRT does. As SDRT adopts the formal semantics underlying DRT, the considerations on metalanguage dynamics put forward with GDRT apply to SDRT too. However, SDRT and GDRT can be considered natural companions in DRT-based dialogue processing: GDRT does not spell out how sequences of DRSs constructed and interpreted during a dialogue are (rhetorically) interconnected so it is at this point where the mechanisms of SDRT can be connected to GDRT. In turn, it would be interesting to see how GDRT can flexibilize SDRT's logical system of axioms and inference by providing the possibility to ground rhetorical structures of and pragmatic inferences from DRSs in an action-theory based account of pragmatics. As GDRT is 'backwards'-compatible to DRT, axiom-based reasoning about agents and discourse (e.g. [Asher and Lascarides, 2003]) is put back into the game if an agent x 'freezes' her cognitive structure $CS(x)(t)$ and uses the anchors in $CS(x)(t)$ to reconstruct a static universe and interpretation function that can be employed as a classical model theory.

This paper introduced the basics of GDRT, where the systematic use of anchors allows to combine *normative* (pragmatic) and *descriptive* (semantic) approaches to discourse processing. I call GDRT normative in the sense that its central goal is to derive appropriate *future* options of (re-action) that serve the realization of discourse goals. Theories such as DRT are descriptive in the they *describe* the processes which are supposed to take place in the minds of the discourse participants when they try to make sense of a *given* discourse. The combination of descriptive semantic and normative pragmatic meaning via the concept of dynamic interpretation proposed in this paper probably constitutes the main technical innovation of GDRT with respect to DRT.

[16]One could even think of the option of finite models with extensions spelled out in [Bonevac and Kamp, 1987]

References

N. Asher. Belief in discourse representation theory. *Journal of Philosophical Logic*, 15:127 – 189, 1986.

N. Asher and A. Lascarides. *Logics of Conversation.* Cambridge University Press, Cambridge, 2003.

J. Barwise and J. Perry. *Situations and Attitudes*. MIT Press, Cambridge, 1983.

D. Bonevac and H. Kamp. Quantifiers defined by parametric extensions. Technical report, Center for Cognitive Science GRG 220 The University of Texas at Austin, Austin, Texas 78712, 1987.

P. R. Cohen and C. R. Perrault. Elements of a plan-based theory of speech acts. In B. Grosz, editor, *Readings in natural language processing*, pages 423–440. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1986.

E. A. Emerson. Temporal and modal mogic. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science Vol. B*, Amsterdam, 1990. North-Holland Publishing Company.

D. M. Gabbay. The declarative past and imperative future: Executable temporal logic for interactive systems. In *Temporal Logic in Specification*, pages 409–448, London, 1987. Springer.

J. Groenendijk and M. Stokhof. Meaning in motion. In K. von Heusinge and U. Egli, editors, *Reference and Anaphoric Relations*, pages 47 – 76. Kluwer, Dordrecht, 1999.

B. J. Grosz and C. L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.

H. Kamp. Prolegomena to a structural account of belief and other attitudes. In J. Anderson, C. Owens, editor, *Propositional Attitudes. The Role of Content in Logic, Language and Mind.*, volume 20, pages 27 – 90. CSLI Lecture Notes, Stanford, 1990.

H. Kamp, J. van Genabith, and U. Reyle. Discourse Representation Theory. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, Dordrecht, 2010. Kluwer.

S. C. Kleene. *Introduction to Metamathematics*. North-Holland Publishing Company, 1952.

M. E. Pollack. Plans as complex mental attitudes. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 77–103. MIT Press, Cambridge, 1990.

T. Pross. *Grounded Discourse Representation Theory. Towards a semantics-pragmatics interface for human-machine collaboration.* PhD thesis, Institute for Natural Language Processing, University of Stuttgart, 2010. URL `d-nb.info/1000641007/34`.

M. P. Singh. *Multiagent Systems. A theoretical framework for Intentions, Know-How and Communications*. Springer, New York, 1994.

M. P. Singh. A semantics for speech acts. In M. N. Huhns and M. P. Singh, editors, *Readings in Agents*, pages 458 – 470. Morgan Kaufman, San Francisco, 1998.

M. P. Singh and N. Asher. A logic of intentions and beliefs. *Journal of Philosophical Logic*, 22: 513 – 544, 1993.

M. P. Singh, A. S. Rao, and M. P. Georgeff. Formal methods in DAI: Logic-based representation and reasoning. In G. Weiss, editor, *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, chapter 8, pages 331 – 376. MIT Press, Cambridge, MA, 1999.

D. Sperber and D. Wilson. Linguistic form and relevance. *Lingua*, 90(2):1 – 25, 1993.

B. van Fraassen. Singular terms, truth-value gaps and free logic. *Journal of Philosophy*, 63 (17):481 – 495, 1966.

J. M. Zacks and B. Tversky. Event structure in perception and conception. *Psychological Bulletin*, 127:3 – 21, 2001.