Assertions in AFA Set Theory

Anton Benz*

January 10, 2007

Abstract

In this paper we consider an application of the theory of non–well-founded sets to the modelling of epistemic updates triggered by assertions in dialogue. We concentrate mainly on what we might call *defective* dialogue, i.e. dialogue in epistemic contexts where the speaker has either erroneous beliefs or is not sincere in his language use. We develop our theory within a special possible worlds framework which uses AFA set theory for representing information states of dialogue participants. We first characterise the use of assertions in *ideal* dialogue situations and then show how pragmatic principles connected to perspectives and speaker's intentions define extended uses in defective situations. On the formal level, the pragmatic considerations lead to the introduction of operators which can be used systematically to extend the domain of assertions.

1 Introduction

In a possible worlds approach the knowledge of an agent is identified with the set of all worlds where his beliefs hold true (Hintikka, 1962). In a model of communication, these worlds must contain information about the beliefs about each other. Due to the foundation axiom in set theory it is not possible to model these worlds as simple structured objects. Instead, it is necessary to use a relational model in the form of a Kripke structure (Fagin et al., 1995). Working in Aczel's (1988) set theory with anti-foundation axiom allows a direct modelling of circular and self-referential structures in terms of set membership. It is well known that this can be exploited for defining more elegant update operations which model learning effects in communication (Gerbrandy & Groeneveld, 1997). In this paper we show how this approach can be used for analysing learning effects in defective communication, i.e. communication where basic epistemic constraints, as e.g. truthfulness or sincerity, are violated.

1.1 Assertions

How can we characterise the dialogue situations where a sentence can be asserted by the speaker and interpreted by the hearer, and what do the participants

*This paper is based on work done in the DFG project LA 633/5–1 on *Dialogue Semantics* located at the Humboldt–Universität Berlin. Furthermore, I have to thank the IFKI at the University of Southern Denmark and the Centre for General Linguistics, ZAS, in Berlin. The paper benefited from discussions at the Conference on Alternative Set Theories - Alternatives to Set Theory, IFKI, Kolding, 2006.

learn from the fact that it has been asserted? We assume that the speaker can reasonably assert a sentence iff the epistemic update triggered by this assertion leads to a desired outcome. The hearer can interpret an assertion as long as the fact that the assertion has been made does not contradict his beliefs.

We may start with the following simple characterisation: If we identify the information which is conveyed by an assertion with the pure semantics of the uttered sentence, and if ψ represents the semantic meaning, then ψ can be asserted iff ψ is true. The interlocutors not only learn that the asserted sentence is true but also that the other participant learns that it is true, and that both know that the other one learns that it is true, etc. I.e. it will become *mutually known* that the sentence is true. This must be represented by the update potential of assertions, and we will do this by using mutual update operations on the belief states of interlocutors.

(1) Helga calls up her son Stephan who lives in a small town in the Alps and asks him whether he wants to visit her in Munich. Stephan answers: "It is snowing in the mountains. I don't like to drive then."

Here, we consider the sentence ψ: *"It is snowing in the mountains."* Its utterance results in a situation where both speaker and hearer know that it is snowing in the mountains, and that this is mutual knowledge. But this would capture only part of the conveyed information. Helga and Stephan learn also e.g. that *Stephan knows that* ψ, and that he assumes that *Helga does not already know that* ψ. Furthermore, it should hold that *Helga believes that Stephan can know whether* ψ, and that *Stephan does not want to mislead her* etc. We call a dialogue situation where it is common knowledge that all participants have only true beliefs and that the speaker does not want to mislead the hearer an *ideal* dialogue situation. And we call an ideal situation a *basic* situation *for* ψ if all *semantic* and *pragmatic* constraints hold which are necessary for a reasonable assertion and successful interpretation.

Now let us assume that we have characterised the conditions and the update effects of assertions in ideal situations. What about non–ideal situations, i.e. defective situations where beliefs are not true, or where speakers are insincere? Here the perspectives of the participants play an important role, i.e. the limited information of dialogue participants and the fact that participants can exploit the limited information of other participants. In the following example of a non–basic situation both interlocutors believe that all the above mentioned pragmatic and semantic conditions hold but, in fact, the uttered sentence is not true:

(2) Helga calls up her son Stephan and asks him whether he wants to visit her in Munich. Stephan answers: "It is snowing in the mountains. I don't like to drive then." But he has not checked the weather for some time, and it is now raining and the streets are clear.

Both will update their beliefs in the same way as in the basic situation. Hence, if we have an update operation for the basic case, we can simply extend it to this wider class and get a correct description of the utterance effects. What happens if only from the perspective of the speaker the basic conditions hold?

(3) Helga calls up her son Stephan and asks him whether he wants to visit her in Munich. Stephan answers: "It is snowing in the mountains." Helga has just talked to her daughter, who lives next to Stephan, and knows that the weather has changed, and that the streets are clear.

S still feels justified to make his utterance, and the hearer may notice this. The update of the speaker remains the same as in the basic situation but the hearer should update only his representation of the speaker's beliefs. Hence we can again derive the update operation for this case from the basic one. We can think of a situation where the speaker thinks that he can mislead the hearer, and where the hearer is aware of the speaker's attempt.

(4) Helga calls up her son Stephan and asks him whether he wants to visit her in Munich. Stephan answers: "It is snowing in the mountains." Helga knows from her daughter that it does not snow at all, and she has also heard that Stephan has a new girl–friend and prefers to stay at home this weekend.

This is a situation where the utterance of *"It is snowing in the mountains"* is reasonable, and it will lead to a new update of the information states of the participants. But it is important that the hearer thinks that the speaker must believe that it is a successful lie. We can see this in contrast to **(5)**:

(5) Helga visits her son Stephan and they take a walk through the town. It is a sunny summer day, and she asks him whether he wants to take her to Munich this day. Stephan answers: "It is snowing in the mountains. I don't like to drive then."

This answer must be absolutely unintelligible for H. It can't be a false assertion, nor a lie. It can't result in a well–defined update of their belief states.

Our central assumption is: If the speaker believes that he can successfully assert some sentence, and the hearer is able to make sense of his utterance, then this is sufficient for the assertion to lead to a well–defined update of their mutual beliefs. We mean by "*being able to make sense*" of an utterance that the fact that an assertion has beed made does not *contradict* the hearer's beliefs.

The dependence on the perspectives explains why it is possible to extend the use of an assertion to new situations. On the other hand, it also leads to stronger conditions for the basic situations. Hence, there are *two* issues which we will address. The first one concerns the extension of a use to new situations, the other one the additional *pragmatic* restrictions in the basic cases. In addition to perspectives we consider speaker's goals. We need them in order to introduce a *sincerity* condition for the base case, for example, to explain the difference between **(2)** and **(3)**.[1]

1.2 Non–well–founded sets

In Section 2 we introduce our general framework. We represent dialogue situations and the beliefs of dialogue participants within a special *possible worlds* framework (Sec. 2). We will build up upon an the approach developed by *Gerbrandy* and *Groeneveld* (1997). Possible worlds w, or *possibilities*, are triples $\langle s_w, w(S), w(H)\rangle$. The first component represents the situation talked about, the second and third the belief sets of speaker and hearer. The sets $w(S)$ and $w(H)$ are again sets of possibilities. As they are proper parts of w, standard set

[1] In **(3)** Helga can interpret Stephan's utterance as a lie only because she knows that he does not want to visit her, hence the result of a successful lie would fit his goals. In Example **(2)** she can infer that his utterance is due to an error because she assumes that he is sincere.

theory would not allow that w is itself an element of $w(S)$ or $w(H)$. But clearly, if the belief sets collect all possibilities that an agent believes to be possible, then a belief set $w(X)$ can only represent a *true* belief iff $w \in w(X)$. In standard possible worlds frameworks (Hintikka, 1962; Fagin et al., 1995), this problem is solved by using relational models, i.e. an agent believes in w that a world v is possible iff w stands in a certain *accessibility relation* R to v. A disadvantage of this approach is that the belief sets, here the set of all worlds accessible to an agent, do no more enter into the identity conditions of worlds. Using AFA set theory (Aczel, 1988; Barwise & Moss, 1996) we can solve both problems, i.e. possibilities w can have belief sets which contain w, and possibilities are identical if, and only if their situations talked about and their belief sets are identical. In AFA set theory, knowledge updates can be reduced to simple set theoretic manipulations of possibilities which leads to elegant and perspicuous mathematical representations (Gerbrandy & Groeneveld, 1997).

1.3 Assertions in AFA set theory

In Section 3 we present our basic considerations concerning the role of perspectives (Sec. 3.1) and speaker's goals (Sec. 3.2) in dialogue. There we introduce operators linked to the perspectives of speaker and hearer which allow us to *derive* extended classes of situations where the speaker believes to be justified in making an assertion and where the hearer can make sense of an assertion. If M is a class where the use of an assertion is well–defined, then we can apply the operators for speaker and hearer to M and get the classes where from the perspective of speaker or hearer the assertion is justified. One of the major problems is to characterise the belief–structures of the dialogue participants in situations which are possible candidates for an extended use, i.e. to characterise the class of dialogue situations where it makes sense to ask whether an extended use of an assertion is possible. In this regard, we introduce a class of possibilities with *internal hierarchical* structure, $\mathcal{H}(\mathcal{T})$, in Section 3.3. The motivation for this structure is closely related to the role of perspectives in deriving extended uses. This defines the class of intended applications of our theory.

In Sections 4 and 5 we show in detail how to derive the basic cases and extended uses. We start with the mutual update operations which are defined by the pure semantics of an assertion in ideal dialogue situations. We consider two groups of pragmatic principles: a) principles linked to perspectives, and b) rationality principles linked to speaker's goals. We will divide the main body of our investigation (Sec. 4/5) into two parts. In the first part we consider only principles linked to perspectives and their interaction with the pure semantics of an assertion (Sec. 4). This part again divides into a sub–part (Sec. 4.1) where we define the pragmatically basic situations where a sentence can be asserted, and a second sub–part (Sec. 4.2) where we show how perspectives give rise to extended uses. The principles linked to perspectives provide for the justification of these extended uses and allow us to define the update operations for a derived use of an assertion in terms of the update operation which represents the underlying use. In the second part we discuss effects of rationality principles connected to speaker's goals (Sec. 5). This section again divides into a sub–part (Sec. 5.1) where we consider the base case problem, and a sub–part (Sec. 5.2) where we consider derived uses of assertions. The underlying circularity of structures and conditions makes the whole enterprise a non–trivial one. In Section 6 we show

how the examples are handled using our theory.

2 Possibilities in AFA set theory

The possibility approach is essentially a possible worlds approach, i.e. it identifies the beliefs of an individual with the set of all worlds which are *possible* according to those beliefs. We denote the set of participants by $\mathrm{DP} = \{S, H\}$, where S denotes the *speaker*, H the *hearer*. A *possibility* consists of a model for the *outer* situation and *information states* for each participant, where those states are again sets of possibilities. The *outer* situation describes the *non–modal* part of the dialogue context. In case of e.g. assertions, this outer situation will be identified with the situation *talked about*. The possibility approach was first developed by *J. Gerbrandy* and *W. Groeneveld* in (1997). It is based on an extension of classical set theory, the theory of *Non–Well–Founded Sets* developed by *P. Aczel* (1988).[2] The original problem motivating the development of the possibility approach was to define suitable *update operations* for dialogue. Here the approach proved to be especially useful.[3] Before we introduce possibilities, we first state some of the fundamental theorems of AFA set theory.

AFA set theory

For Cantor or Frege, a *set* was an unrestricted collection of objects that satisfy a certain predicate. Understood thus, a set can be seen as the semantic extension associated to a predicate. In an effort to avoid the paradoxes generated by unrestricted comprehension, the cumulative conception of the set universe emerged.[4] Intuitively, we can think of the set universe as generated from base sets by iterated application of a power set operation. The set universe is organised in levels where each higher level comprises all subsets of the lower levels. In AFA set theory the emphasis shifts away from the cumulative conception. Here, the aspect that is relevant to us is the conception of a set as a structured object where the identity of objects follows from the identity of its constituents, i.e. elements.

AFA set theory has no axiom of foundation. The following so-called *Solution Lemma* can either be proved as a theorem in AFA set theory (Aczel, 1988) or be introduced as an axiom (Barwise & Moss, 1996). In the latter case, it is the axiom which replaces the axiom of foundation:

Theorem 2.1 (Solution Lemma) *Let* $\mathbf{V}[X]$ *be the universe of sets over a class* X *of* urelements. *A* system of equations over X *is a function* $e : X \longrightarrow \mathbf{V}[X] \setminus X$. *Then, every system of equations has a unique solution; i.e. there exists a unique function* $s : \mathbf{V}[X] \longrightarrow \mathbf{V}$ *such that*

$$\begin{aligned} s(m) &= \{s(n) \mid n \in m\} \quad \textit{for sets } m \\ s(x) &= s(e(x)) \quad \textit{for urelements } x \in X \end{aligned}$$

[2]For more information about (AFA) set theory we can refer to (Barwise & Moss, 1996), and for the possibility approach to the thesis of Gerbrandy (1998).

[3]There is a larger literature concerning the proper definition of updates in communication: (Jaspars, 1994), (Groeneveld, 1995), (Gerbrandy & Groeneveld, 1997), (Zeevat, 1997), (Gerbrandy, 1998), (Baltag, Moss, Solecki, 1999), (Baltag, 1999), (Baltag et al., 2006) and more.

[4]For the history of this development see (Fraenkel & Bar-Hillel, 1958).

To see the significance of this lemma, let us consider an arbitrary set m. We can associate to m the set of equations $n = \{k|k \in n\}$, where n is in the transitive hull of m. In a second step we can replace each set n by a distinct urelement u_n, which leads to the system of equations $u_n = \{u_k \mid k \in n\} =: e(u_n)$. A solution to this system of equations is a function s that assigns to each urelement u_n a set $s(u_n)$ such that $s(u_n) = \{s(u_k) \mid k \in n\} = \{s(e(u_k)) \mid k \in n\}$. We can think of the system of equations as a description of the internal structure of a set. In standard set theory every set has such a structural description and if a system of equations has a solution, then this solution is unique. AFA set theory generalises this property by assuming that *every* system of equations has a unique solution. Two sets are identical if they can be described by the same system of equations, i.e if they have the same structural description. The following identity condition follows from the Solution Lemma:

Theorem 2.2 *A relation $R \subseteq \mathcal{W} \times \mathcal{W}$ is a* bisimulation*, iff vRw implies that $s_v = s_w$ and for all $X \in \mathrm{DP}$*

$$\forall v' \in v(X)\, \exists w' \in w(X)\, v'Rw' \text{ and } \forall w' \in w(X)\, \exists v' \in v(X)\, v'Rw'.$$

If R is a bisimulation, then vRw implies $v = w$.

Possibilities

In order to define certain subclasses of non–well–founded sets we need the following special fixed–point lemma:

Lemma 2.3 *For each class M there exists a unique maximal class C which satisfies:*

$$C = \{\langle m, x, y\rangle \mid m \in M \,\&\, x, y \subseteq C\} \quad (*)$$

C can be constructed as a fixed-point of the operator $\Gamma X := \{\langle m, x, y\rangle \in X | m \in M \,\&\, x, y \subseteq X\}$. Let $C^0 := \mathbf{V}$, $C^{n+1} := \Gamma\, C^n, C := \bigcap_n C^n$, then C is a fixed–point of Γ and the largest class that satisfies $(*)$.[5]

Let $\mathcal{M}$ be a class of models for the possible outer situations. We define *possibilities* and *information states* in the following way:

- A *possibility* w is a triple $\langle s_w, w(S), w(H)\rangle$ where $s_w \in \mathcal{M}$ and $w(S)$ and $w(H)$ are information states.
- An *information state* σ is a set of possibilities.

s_w describes an outer situation, $w(S)$ and $w(H)$ the set of worlds S and H believe to be possible. We denote the class of all possibilities with $\mathcal{W}$. The theory of non–well–founded sets allows for sets containing themselves, hence it is possible that there exist possibilities w with $w \in w(X)$, $X \in \mathrm{DP}$.

Let $\mathcal{L}$ be a language of predicate logic for the class $\mathcal{M}$. We assume that $\mathcal{L}$ contains all the predicates the dialogue participants can use to talk about an outer situation. Let $w = \langle s_w, w(S), w(H)\rangle$ be a possibility. We define truth conditions for $\varphi \in \mathcal{L}$:

$$w \models \varphi \text{ iff } s_w \models \varphi, \ \varphi \text{ a sentence in } \mathcal{L}.$$

[5] For the set–theoretic machinery behind fixed–point constructions we refer to (Aczel, 1988; Barwise & Moss, 1996).

For a dialogue participant X a possibility w is *epistemically possible* in v iff $w \in v(X)$. X *believes* that φ in w iff φ holds in all his epistemic alternatives in w. We write $w \models \Box_X \varphi$ *iff* $\forall v \in w(X)\, v \models \varphi$, and $w \models \Diamond_X \varphi$ *iff* $\exists v \in w(X)\, v \models \varphi$. If σ is an information state, then we define $\sigma \models \varphi$ *iff* $\forall w \in \sigma\, w \models \varphi$.

Until now, we did not restrict the properties of possibilities. A subclass $M \subseteq \mathcal{W}$ is called *transitive*, iff $\forall w \in M\, \forall X \in DP\, w(X) \subseteq M$. Let $\mathcal{I} \subseteq \mathcal{W}$ be the largest transitive subclass with

$$\forall w \in \mathcal{I}\, \forall X \in DP\, \forall v \in w(X) :\ w(X) = v(X).$$

This property is called *introspectivity*. It means: (1) If a dialogue participant believes φ, then he knows that he believes it; (2) if he does not believe that φ, then he knows that he does not believe φ; and (3) it means that (1) and (2) are common knowledge. Let $\mathcal{T} \subseteq \mathcal{I}$ be the largest transitive subclass with

$$\forall w \in \mathcal{T}\, \forall X \in \mathrm{DP}\, w \in w(X).$$

If $w \in \mathcal{T}$, then w is for both participants an element of their sets of epistemic alternatives. Hence, if a participant believes that φ, then φ must in fact hold. Therefore, $\mathcal{T}$ denotes the class of possibilities where (1) the dialogue participants can only have *true beliefs*, i.e. *knowledge*, and (2) where this fact is common knowledge. The *Anti–Foundation–Axiom* (AFA) of the underlying set theory guaranties that $\mathcal{T}$ is not empty.

We are only interested in non–contradicting information states of participants. This means that the set containing all their epistemic possibilities should contain at least one element. Let $\dot{\mathcal{I}}$ denote the largest transitive subclass of $\mathcal{I}$ with $w \in \dot{\mathcal{I}} \Rightarrow w(S) \neq \emptyset \neq w(H)$.

If $D(a) \subseteq \mathcal{W}$ and $a :\ D(a) \longrightarrow \mathcal{W}$, then we call a a *normal mutual update* if

(**NMU**) $a(w) = \langle s_w, a(w(S)), a(w(H))\rangle$ for $w \in D(a)$,
$a(\sigma) = \{a(w) \mid w \in \sigma \cap D(a)\}$ for information states σ.

It is clear that a normal mutual update is uniquely determined by its domain $D(a)$. If $[\varphi] := \{w \in \mathcal{T} \mid w \models \varphi\}$, then the normal mutual update determined by $[\varphi]$ describes the effect of '*mutually learning that* φ' in $\mathcal{T}$. We can think of this mutual learning as a step by step process of eliminating possibilities. First both participants eliminate all possibilities where φ is not true from their information states, and then they take the remaining possibilities and again eliminate all worlds from both information states where φ is not true. This they repeat again and again. Such definitions of update operations for possibilities have been introduced by (Gerbrandy & Groeneveld, 1997). From the set theoretic point of view, (**NMU**) defines a system of equations, i.e. a structural description of the possibility that results from updating.

3 The Basic Considerations

In this section we describe the basic ideas of our approach and provide a definition of the class of possibilities which represent the intended applications. We have already explained that we make a distinction between assertions in ideal and non–ideal dialogue situations. We represent the class of ideal situations by $\mathcal{T}$, i.e. we assume that all participants have only true beliefs, and that this is common knowledge. The *basic* assertions of a sentence ψ are all uses of ψ in $\mathcal{T}$ where all pragmatic and semantic constraints hold. We will denote this class by $\mathcal{B} = \nabla[\psi] \subseteq \mathcal{T}$. We can represent the effect of mutually learning ψ by a normal mutual update $a : [\psi] \longrightarrow \dot{\mathcal{I}}$, where $[\psi] := \{w \in \mathcal{T} \mid w \models \psi\}$. But if ψ is really asserted, then both participants not only learn that ψ holds but also that all pragmatic conditions hold, i.e. they learn that they are in a situation in $\mathcal{B}$. This means that we can identify an assertion with the normal mutual update $a : \mathcal{B} \longrightarrow \dot{\mathcal{I}}$. We then want to extend this basic use to non–ideal situations. We will see that it is not possible to identify these non–ideal cases with the complement $\dot{\mathcal{I}} \setminus \mathcal{T}$. This forces us to describe the intended applications more closely. In this respect we introduce in Section 3.3 a class $\mathcal{H}(\mathcal{T})$. In the same way as $\mathcal{T}$ represented the class of situations where it was reasonable to ask whether a basic use is possible, $\mathcal{H}(\mathcal{T})$ will represent the class where it is reasonable to ask whether a derived extended use is possible. We extend $\mathcal{B}$ to a class $\Delta(\mathcal{B}) \subseteq \mathcal{H}(\mathcal{T})$ by iterated applications of certain operators which are linked to perspectives. They reflect the way how perspectives restrict and allow use and interpretation of assertions. Later (Sec. 4.1) we will see that we can also derive $\mathcal{B} = \nabla[\psi]$ by an iterated application of a related combination of operators to $[\psi]$. We first (Sec. 3.1) motivate these operators and then characterise the non–ideal dialogue situations where we intend to apply our theory.

3.1 Considerations concerning Perspectives

The Operators

We reconsider the assertion *It is snowing in the mountains* ψ in Example **(1)**. If we define the felicity conditions by pure semantics then asserting ψ would be felicitous just in case the sentence ψ is true. We represent the situation of Example **(1)** by a possibility w with[6]:

$$\begin{aligned} w &= \langle \psi, \{w\}, \{w, v\} \rangle \\ v &= \langle \neg\psi, \{v\}, \{w, v\} \rangle \,. \end{aligned}$$

I.e. w is such that ψ holds, the speaker knows the actual situation, and the hearer can not distinguish between w and v, where the only difference between w and v is that $\neg\psi$ holds in v. Hence the semantic condition for asserting ψ holds. But, of course, the speaker S can rely for what he says only on what he believes to be true. We assume throughout this paper that the speaker is only justified to assert a sentence if he is *convinced* that he can assert it. For the basic case this means that $w(S)$ should be a subset of all situations where ψ holds. This is, of course, in many real life situations too strong a requirement, as agents may well perform an action if they think e.g. that the possible negative results are negligible, or that it is highly unlikely that it has no success. If $M \subseteq \dot{\mathcal{I}}$ is a class which represents some property of possibilities, e.g. a class where some sentence can be asserted successfully, then we explicate the fact that this property obtains *under the perspective* of a dialogue participant X in world w as $w(X) \subseteq M$.

Now we look at the hearer H. We see that one of his epistemic alternatives is that ψ does not hold. It is also essential that the addressee does not already

[6] As we are only interested in the truth or falsity of ψ we denote the outer situation just by ψ or $\neg\psi$.

know that the uttered sentence is true. Hence, if M characterises the pragmatic conditions for a reasonable use of an assertion, then the requirement that he must believe that all his epistemic alternatives are situations in M is too strong. We assume that he *can make sense* out of an assertion if there is at least one epistemic alternative in his set of possibilities where the uttered sentence is reasonable. *Make sense* means here that if the hearer *learns* that a sentence ψ is asserted, then this fact does not contradict with what he believes about the world. The actual situation w *may* have the property represented by a class M *under the perspective of* a participant X iff $w(X) \cap M \neq \emptyset$.

We define two operators on subclasses of possibilities for each dialogue participant X. They are closely related to the modal operators $\Box_X$ and $\Diamond_X$, so we denote them by the same symbols:

$$\begin{aligned} \Box_X M &:= \{w \in \dot{\mathcal{I}} \mid w(X) \subseteq M\} \\ \Diamond_X M &:= \{w \in \dot{\mathcal{I}} \mid w(X) \cap M \neq \emptyset\} \end{aligned}$$

With these operators at hand we can reformulate our observations as: S is convinced that the actual world w belongs to M iff $w \in \Box_S M$; Learning that M does not contradict what H believes iff $w \in \Diamond_H M$. Notice that these operators generate proper classes, even if M is a set. Our aim is to derive the class of extended uses of an assertion by an iterated application of certain combinations of these operators to the base $\mathcal{B}$.

Combinations of Operators for Derived Uses

Let us now look at Example **(2)**. We represent it by

$$u = \langle \neg\psi, \{w\}, \{w, v\} \rangle ,$$

where w and v are the situations from Example **(1)**. We see that $u(S) = w(S)$ and $u(H) = w(H)$. If the assertion leads to an update in w, then it should also do so in u. If we assume that we have proven that w belongs to the basic cases for a successful use, and if we assume that we have characterised these basic uses by a class $\mathcal{B}$, then we see also that $u, w \in \Box_S \mathcal{B} \cap \Diamond_H \mathcal{B}$.

In general we assume that an assertion leads to a successful update if the speaker is convinced that he can use the assertion, and if the fact that it is asserted does not lead to a contradiction with the beliefs of the hearer. Hence, if we have proven that an assertion leads to an update for a class of situations M, then it also leads to an update of the speaker's information state for situations in $\Box_S M$, and to an update of the hearer's information state for situations in $\Diamond_H M$.

If the speaker thinks that the hearer can make sense of an assertion, he can exploit this fact and mislead him. This happens in case of lying. We can see here how the limited perspective of one dialogue participant can give rise to an extension of a dialogue act. If M is a class of possibilities where an assertion leads to an update, then this act can be extended to the class where S is convinced that they are in a situation in M or where H can at least make sense of the assertion. This is the class $\Box_S(M \cup \Diamond_H M)$. If $M \subseteq \Diamond_H M$, the definition of the extension can be simplified to $\Box_S \Diamond_H M$. This will always be the case.

Now we can again turn to the perspective of H. Assume that the speaker S is convinced of the truth of ψ but H knows it to be false. This was the case in Example **(3)** which we can represent by:

$$s = \langle \neg\psi, \{w\}, \{s\} \rangle ,$$

where w is the situation from Example **(1)**. We see that $s \in \Diamond_H \Box_S \mathcal{B}$, where $\mathcal{B}$ denotes the class of basic situations for the assertion ψ (*It is snowing in the mountains*). We assume H can make sense of the utterance if he thinks it is possible that S might believe that he can assert ψ. Let M be again a class of possibilities where we know that the assertion leads to an update. Then the assertion also leads to an update of the hearer's information state if H thinks that it is possible that S is convinced that the actual situation belongs to M. I.e. there exists an update of the hearer's information state for the class $\Diamond_H \Box_S M$.

We find in this way four operations which give us new classes where some sentence can be asserted due to the perspective of *one* participant. Let M be given. Then we classify the perspectivally derived classes as shown in Figure 1.

	direct	indirect
speaker	$\Box_S M$	$\Box_S(M \cup \Diamond_H M)$
hearer	$\Diamond_H M$	$\Diamond_H \Box_S M$

Figure 1: The four basic perspectival operations.

We can simplify $\Box_S(M \cup \Diamond_H M)$ to $\Box_S \Diamond_H M$ for the indirect operation for the speaker if $M \subseteq \Diamond_H M$. We will see later that this is the case in all intended applications.

It follows by *introspectivity* that we can not get new possible extensions if we apply the direct operation twice for the same participant. I.e. for operators $P, Q \in \{\Box, \Diamond\}$, $M \subseteq \dot{\mathcal{I}}$ and $X \in \mathrm{DP}$ we find that $P_X Q_X M = Q_X M$.

To get a real extension it is necessary that S is convinced that he can perform the act, and H must be able to make sense of this. It is not sufficient that only one participant thinks that the act can be performed. Therefore, we have to build intersections of the derived classes. We get the four groups of derived cases as shown in Figure 2.

	direct H	indirect H
direct S	$\Box_S M \cap \Diamond_H M$	$\Box_S M \cap \Diamond_H \Box_S M$
indirect S	$\Box_S \Diamond_H M \cap \Diamond_H M$	$\Box_S \Diamond_H M \cap \Diamond_H \Box_S M$

Figure 2: The four derived classes.

3.2 Considerations concerning Speaker's Goals

So far we have not taken speaker's intentions into account. In Example **(4)** Helga will interpret Stephan's utterance as a lie because she knows that it fits his intentions. This contrasts with Example **(3)** where she will interpret the same utterance as a mistake because she has no reason to assume that he wants to mislead her. Also in a basic situation like in Example **(1)** the hearer not only learns that some fact ψ holds and that the speaker believes in ψ but also that he wants to inform her about this fact. The following example is identical with Example **(1)** except for the fact that it is common knowledge that the speaker wants to mislead the hearer in one possible context:

(6) Helga calls up her son Stephan. The last time she invited him, he pretended that he could not come because of the bad weather, and she knows that he does not like to come this time either. They both know that Stephan would like to use the same excuse this time.

If we don't consider speaker's goals then this situation receives the same representation as Example **(1)**. But, of course, if Stephan says that *it is snowing in the mountains*, then this should lead to a different update this time. He should not be able to convince Helga. This means that our theory should predict that it is not rational to make the assertion in this context. To this end we introduce speaker's intentions and rationality constraints connected to goals. We represent speaker's goals by a function which gives us the desired states of the world for all his epistemic alternatives, i.e. we redefine possibilities as triples of the form $\langle s_w, \langle w(S), G_S^w\rangle, \langle w(H)\rangle\rangle$ where G_S is a function with domain $w(S)$ and values $G_S(v) \subseteq \dot{\mathcal{I}}$.[7]

Our model does not allow to rank possible outcomes, so we can't expect it to provide a real criterion for rational *choice*. Our goal–functions divide the class of all possibilities into the class of possibilities where the goal is achieved, and the class where it is not achieved. We formulate criteria which tell us whether it is reasonable to choose an action as a *means* to reach the goal.

We formulate two elementary constraints which tell us whether it is rational to assert a sentence relative to some goal G. Assume that there are given a situation w, a mutual update a and a goal $G(w)$. Let w be in the domain of a. Then we postulate the following rationality constraints:

($\mathbf{R_1}$) $w \notin G(w)$

($\mathbf{R_2}$) $a(w) \in G(w)$

The first axiom claims that we should only perform an action if our goal is not yet achieved. If it is, nothing should be done. The second axiom states that we should only choose actions which allow us to reach our goal, i.e. the result $a(w)$ of the performance of a in w should be in $G(w)$. We can also see that ($\mathbf{R_1}$) and ($\mathbf{R_2}$) imply that $a(w) \neq w$. We can incorporate these constraints in our operators but this will have to wait until Section 5. If (1) we know that there is a class M where we have already established that a felicitous use of an assertion is possible, and if (2) w is an element of one of the classes which can be derived by a combination of perspectival operators for speaker and hearer, and if (3) the update with the information connected to the fact that ψ has been asserted, i.e. with $\Delta(\mathcal{B})$, satisfies the rationality conditions, then this *proves* that a felicitous use of the assertion is also possible in w. (3) makes the criterion circular.

[7]The precise definition will follow in Section 5.

3.3 Characterising the Intended Applications

We start with the felicity conditions and update effects of assertions as they are defined by pure semantics, i.e. we assume that a sentence with semantic content ψ can be asserted exactly iff the actual situation belongs to $[\psi] := \{w \in \mathcal{T} \mid w \models \psi\}$. Its effect is described by the normal mutual update with this semantic content, i.e. by $a: [\psi] \longrightarrow \dot{\mathcal{I}}$. The pragmatically *basic* cases build the subclass $\mathcal{B}$ of $[\psi]$ where the additional pragmatic constraints linked to perspectives and intentions hold. We assume here (1) that the speaker is sincere and must know that he is successful, and (2) that the hearer must be able to make sense of the utterance, i.e. updating with the information which is connected to the fact that ψ was uttered does not lead to a contradiction. The implicit circularity of these constraints makes the task of finding $\mathcal{B}$ a non–trivial one.

We extend $\mathcal{B}$ to non–ideal dialogue situations by applying a combination of the perspectival operators. Let us assume now that we have applied the extending operation α times and got $\Delta^\alpha(\mathcal{B})$. $\Delta^\alpha(\mathcal{B})$ should also represent the information which is connected to the fact that ψ was uttered, and we can identify the update potential with the normal mutual update $a: \Delta^\alpha(\mathcal{B}) \longrightarrow \dot{\mathcal{I}}$ defined by $\Delta^\alpha(\mathcal{B})$. What we want is to define $\Delta^{\alpha+1}(\mathcal{B})$ as $\Delta^\alpha(\mathcal{B}) \cup D(\Delta^\alpha(\mathcal{B}))$ where D denotes the operation which derives extended uses. Then we can identify the new update effect of asserting ψ with the normal mutual update defined by $\Delta^{\alpha+1}(\mathcal{B})$. In the end we want to collect all $\Delta^\alpha(\mathcal{B})$ in a class $\Delta(\mathcal{B})$. This then proves that the pragmatic principles connected to perspectives and the rationality constraints allow extended uses of the assertion of ψ exactly iff the utterance situation belongs to $\Delta(\mathcal{B})$. But this does not work out so smoothly. E.g. we have said that $\mathcal{B}$ should characterise the pragmatically basic uses, i.e. a sentence ψ can be asserted in an ideal dialogue situation $w \in \mathcal{T}$ *if, and only if* $w \in \mathcal{B}$. But if we now apply our extending operators, this may add new elements from $\mathcal{T}$, and this first contradicts the claim that we can assert ψ in $\mathcal{T}$ *only if* it is an element in $\mathcal{B}$, and second the update of a situation in $\mathcal{B}$ with the larger set $\Delta^{\alpha+1}(\mathcal{B})$ may now lead to a different result. This leads to problems in connection with the rationality constraints, Example **(7)** below. We solve these problems by restricting our applications to a class $\mathcal{H}(\mathcal{T})$ of candidates with an internal hierarchical structure. There is a *hard* reason for introducing $\mathcal{H}(\mathcal{T})$ which is related to speaker's intentions. If we allow for possibilities with arbitrary circular structure, then we find situations where our criterion for extended uses leads into an irresolvable circle:

(7) Let w be the situation from Example **(1)**, i.e. $w = \langle \psi, \{w\}, \{w, v\} \rangle$ and $v = \langle \neg\psi, \{v\}, \{w, v\} \rangle$. Then we consider

$$\begin{aligned} t &= \langle \neg\psi, \{t\}, \{t, u\} \rangle, \\ u &= \langle \psi, \{w\}, \{t, u\} \rangle, \end{aligned}$$

We further assume that in t and w the speaker wants to convince the hearer that ψ holds. Furthermore, we assume that in w it is common knowledge that the speaker is sincere. Hence we assume $w \in \mathcal{B}$. Now we get the following problem: Neither t nor u are elements of $\mathcal{T}$, hence, in the initial situation no use of an assertion is defined. Then we try to extend $\mathcal{B}$ step by step. First we see that $u \in \Box_S \mathcal{B} \cap \Diamond_H \Box_S \mathcal{B}$. It follows that a derived use is possible in u. As t and u do not belong to $\mathcal{B}$ it follows that t is not contained in the first extension $\Delta^1(\mathcal{B})$. Hence the related update with $\Delta^1(\mathcal{B})$ should eliminate t from H's information state in u. But this means that the update results in a state where H is convinced that ψ holds. Now we apply our operators once more, and we see that $t \in \Box_S \Diamond_H \Delta^1(\mathcal{B}) \cap \Diamond_H \Delta^1(\mathcal{B})$. As S can also see that an update with $\Delta^1(\mathcal{B})$ leads to a desired situation, it follows that $t \in \Delta^2(\mathcal{B})$. But now the mutual update with $\Delta^2(\mathcal{B})$ does not remove t from H's information state, hence, she will no more be convinced that ψ holds. This means that we should remove t again if we define $\Delta^3(\mathcal{B})$ because a rationality condition is violated. But then the update with $\Delta^3(\mathcal{B})$ eliminates t in $u(H)$ and we are in the same situation we have been in after the first derivation.

We are going to characterise the intended applications by a class $\mathcal{H}(\mathcal{T})$. The elements of $\mathcal{H}(\mathcal{T})$ will have an *internal hierarchical structure* which allows us to measure their complexity by ordinal numbers. We then construct $\Delta^{\alpha+1}(\mathcal{B})$ in such a way that our operators add only elements with complexity $\alpha+1$. Then our claim will be that a sentence ψ can be asserted in a situation $w \in \mathcal{H}^{\alpha+1}(\mathcal{T})$, i.e. in situations in $\mathcal{H}(\mathcal{T})$ with complexity at most $\alpha+1$, if and only if $w \in \Delta^{\alpha+1}(\mathcal{B})$. $\mathcal{H}(\mathcal{T}) \setminus \mathcal{T}$ is the class of candidates for a derived extended use in the same way as $\mathcal{T}$ was for the basic uses.

For the direct derivation, our intuition has been that the participant believes to be in a situation where some sentence can be asserted successfully. For the indirect case, he has to be convinced that the other one is or might be in such a situation. In both cases, we think that the possibilities in the exploited information states have to be more simple than the newly derived possibilities. In the ideal case of Example **(1)** we see that the hearer not only believes that they may be in a basic situation $w \in \mathcal{B}$ for ψ but also that she is *convinced* to be in an ideal situation, i.e. the actual situation belongs to $\Diamond_H \mathcal{B} \cap \Box_H \mathcal{T}$. In Example **(4)** the speaker believes that he can mislead the hearer. But this is only the case because he is convinced that the hearer is convinced to be in an ideal situation, i.e. the actual situation belongs to $\Box_S \Diamond_H \mathcal{B} \cap \Box_S \Box_H \mathcal{T}$. In Example **(3)** the hearer interprets the assertion as a mistake because she is convinced that he acts according to the conditions of ideal dialogue, i.e. the actual situation is in $\Diamond_H \Box_S \mathcal{B} \cap \Box_H \Box_S \mathcal{T}$. In case of Example **(7)** we see that $u \in \Diamond_H \mathcal{B}$ but $u \notin \Box_H \mathcal{T}$. These observations lead us to the following generalisation: If we have established that a sentence ψ can be asserted in situation $w \in G$ if, and only if $w \in M$, then we find the *candidates* for an extended use by looking at the table in Figure 3: This means, if an extended use is possible in a situation v,

	direct	candidate	indirect	candidate
speaker	$\Box_S M$	$\Box_S G$	$\Box_S(M \cup \Diamond_H M)$	$\Box_S(G \cup \Box_H G)$
hearer	$\Diamond_H M$	$\Box_H G$	$\Diamond_H \Box_S M$	$\Box_H(G \cup \Box_S G)$

Figure 3: The four derived uses and their candidate classes.

then v has to be e.g. an element of $(\Box_S M \cap \Box_S G) \cap (\Diamond_H \Box_S M \cap \Box_H \Box_S G)$. The operators in the second and fourth column provide us with the candidates where speaker and hearer can search for extended uses. We can construct the class $\mathcal{H}(\mathcal{T})$ by an iterated application of these operators. We will prove in Lemma 4.3 and 5.2 that, in fact, the problems of Example **(7)** do not occur if we restrict our applications to this class.

We now look for a description which can be better used in the subsequent proofs. For this reason, we provide for an *axiomatic* characterisation of $\mathcal{H}(\mathcal{T})$. In Section A, Lemma A.7, we will prove that the recursive construction and the axiomatic characterisation define the same structure $\mathcal{H}(\mathcal{T})$. The construction by an iterated application of the above operators suggests that we can characterise $\mathcal{H}(\mathcal{T})$ in terms of *paths*. And, in fact, we can do this: $\mathcal{H}(\mathcal{T})$ is the largest transitive subclass of $\dot{\mathcal{I}}$ such that for all situations $w \in \mathcal{H}(\mathcal{T}) \setminus \mathcal{T}$:

1. There is at most one participant who believes the real situation to be possible.

2. There are no long circular paths going from one participant to the other and coming back to the original situation.

3. If there is a path starting at the real situation which goes from one participant to the other, then this path should ultimately reach the ideal situations in $\mathcal{T}$.

I.e. we don't allow for structures where we have $(\mathbf{H_{1'}})$ $w \in w(S) \cap w(H)$, or $(\mathbf{H_{2'}})$ sequences $v_1 \in v_0(X_0) \,\&\, v_2 \in v_1(X_1), \ldots, v_0 \in v_n(X_n)$ where $v_0 \in \mathcal{H}(\mathcal{T}) \setminus \mathcal{T}$, for all i it holds that $X_i \neq X_{i+1}$ and $n > 0$, or $(\mathbf{H_{3'}})$ sequences $v_1 \in v_0(X_0) \,\&\, v_2 \in v_1(X_1), \ldots$ where for all i holds that $X_i \neq X_{i+1}$ and where no $v_i \in \mathcal{T}$. These quite intuitive conditions may help to understand our final characterisation of $\mathcal{H}(\mathcal{T})$. We will see in Section 2 that the following axioms $(\mathbf{H_1})$–$(\mathbf{H_3})$ capture the content of $(\mathbf{H_{1'}})$–$(\mathbf{H_{3'}})$. We provide for a more fine–grained structure because it allows us do define extended uses by recursion over the *complexity* of situations.

Let $T(w)$ be the smallest transitive superset of $\{w\}$. We call $T(w)$ the *transitive hull* of the possibility w. In case of Example **(1)** $T(w)$ is $\{w, v\}$, for Example **(2)** $T(u) = \{u, w, v\}$, and for Example **(3)** $T(s) = \{s, w, v\}$. We can see that $T(w) \subseteq T(u), T(s)$ but $T(u), T(s) \not\subseteq T(w)$. In general, we find e.g. for $w \in \Box_S\mathcal{T} \cap \Box_H\mathcal{T}$ with $w \notin \mathcal{T}$ that for all $v \in w(S) \cup w(H)$: $T(w) \not\subseteq T(v)$. In case of Example **(7)** we see that $T(t) = T(u)$.

In a first step we restrict the intended applications to cases where the subset relation between transitive hulls defines a well–founded partial order on dialogue situations:

- $[w] := \{v \in \dot{\mathcal{I}} \mid T(w) = T(v)\}$,
- $[v] \leq [w]$ *iff* $T(v) \subseteq T(w)$.

Let $\mathcal{F}$ be the class of all possibilities $w \in \dot{\mathcal{I}}$ where $\leq$ is a well founded partial order on $\{[v] \mid v \in T(w)\}$. We can see that Example **(7)** is still an element of $\mathcal{F}$. For $\mathcal{F}$ we can define an *order type* for the possibilities. This order type provides us with a measure for the complexity of situations.

- $\text{otp}(w) = 0$ iff $\{[v] \mid [v] < [w]\} = \emptyset$.
- $\text{otp}(w) := \sup\{\text{otp}(v) + 1 \mid [v] < [w]\}$, else.

For $M \subseteq \mathcal{F}$ let $M^\alpha := \{w \in M \mid \text{otp}(w) \leq \alpha\}$. This measure of complexity is still quite rough. E.g. all possibilities in $\mathcal{T}$ have order type $\text{otp}(w) = 0$. We note the following fact, which follows by definition of otp.

Fact 3.1 *For $w \in \mathcal{F}$ we have:* $\forall v \in T(w)\ (\text{otp}(v) = \text{otp}(w) \Leftrightarrow w \in T(v))$.

Now, we can provide our final characterisation of $\mathcal{H}(\mathcal{T})$: It is the largest subclass of $\mathcal{F}$ such that for all $w \in \mathcal{H}(\mathcal{T}) \setminus \mathcal{T}$, and for all $X \in \mathrm{DP}$:

($\mathbf{H_1}$) $w \notin w(S) \cap w(H)$,

($\mathbf{H_2}$) $\forall v \in T(w)\ (w \notin v(X) \Rightarrow \forall u \in v(X)\, \mathrm{otp}(v) < \mathrm{otp}(w))$,

($\mathbf{H_3}$) $w(X) \subseteq \mathcal{H}(\mathcal{T})$.

The second axiom says that for all v in the transitive hull of w, if w is not an element of $v(X)$, then all possibilities in $v(X)$ have a complexity lower than the complexity of w. Following from Fact 3.1 ($\mathbf{H_2}$) is equivalent to:

($\mathbf{H_2}$) $\forall v \in T(w)\ (w \notin v(X) \Rightarrow \forall u \in v(X)\, w \notin T(u))$.

4 Perspectives and Assertions

In this section we show how to derive the basic cases and the extended cases for asserting a sentence ψ using our perspectival operators. We concentrate on the *epistemic perspectives* of the dialogue participants and the role they play for assertions. The additional problems posed by constraints linked to speaker's goals are the topic of Section 5. We hope that the essentials of our theory become clearer if we separate the discussion of perspectives and speaker's intentions.

4.1 The Base Case Problem

We look again at the assertion (ψ) *It is snowing in the mountains* in **(1)**.

In order to make his utterance, S should be convinced that it is really snowing in the mountains. If it in fact does but S has no evidence for ψ, then his assertion in **(1)** is not justified.

Suppose now it is snowing in the mountains, S is convinced of it, and H happens to know the truth of it too. But assume also that H is convinced that S can't know whether ψ. In this case the fact that the speaker asserts ψ contradicts his previous beliefs, and either he revises them or he has to assume that S is insincere or mistaken.

We make the assumption that *agents* are *justified* to perform an action iff they are convinced that it must be successful, and we make the assumption that hearers can make sense of an assertion if the fact that the act is performed by the speaker does not contradict the hearer's beliefs about the dialogue situation.

Assume that we have a condition γ which specifies the conditions of success for a certain dialogue act a. This means that the act can be performed successfully in all situations s where $s \models \gamma$. If a speaker S wants to perform a he has to be sure that γ really holds. Therefore, in a situation s where S performs a we must have $s \models \Box_S \gamma$. But this then becomes part of the information carried by the fact that a was performed. Hence, if the hearer recognises that a was performed by S, it should be the case that $s \models \Diamond_H(\gamma \wedge \Box_S \gamma)$. Otherwise the fact of S performing a will contradict H's beliefs.

But then, S has to be sure that this is the case. So we need in addition $s \models \Box_S \Diamond_H(\gamma \wedge \Box_S \gamma)$. We can go on with this way of reasoning and get an infinite number of new conditions for s. We can describe this observation by: Let Σ be the smallest set of formulas which contains γ and which is closed under:

$p, q \in \Sigma$ then $p \wedge q, \Box_S p, \Diamond_H p \in \Sigma$. Then, we call the performance of dialogue act a in a situation s *mutually justified* iff $s \models \Sigma$. We claim that the situations in $\mathcal{T}$ where a dialogue act is mutually justified are the *pragmatically basic* cases for this act. The following proposition gives us a simple criterion for deciding which situations in $\mathcal{T}$ support Σ.

Proposition 4.1 *With Σ defined as above and $s \in \mathcal{T}$ we get:*

$$s \models \Sigma \iff s \models \gamma \wedge \Box_S \gamma$$

Proof: Let Σ be the smallest set containing γ and closed under $p, q \in \Sigma \Rightarrow p \wedge q, \Box_S p, \Diamond_H p \in \Sigma$. Let $M := \{s \in \mathcal{T} \mid s \models \gamma \wedge \Box_S \gamma\}$. It is clear that $s \models \Sigma$ implies $s \in M$. Hence, assume $s \in M$. Then $s \models \gamma$ and if $s \models p$ and $s \models q$ then, of course, $s \models p \wedge q$. Furthermore, $s \in \mathcal{T}$ implies $s \models p \Rightarrow s \models \Diamond_H p$. Hence, it remains to show that $s \models p$ implies $s \models \Box_S p$ for $p \in \Sigma$. If $p \equiv \gamma$, it follows by $s \in M$. Of course, we have $s \models \Box_S p$ & $s \models \Box_S q$ then $s \models \Box_S (p \wedge q)$. Hence, assume $p \equiv \Box_S q$ or $p \equiv \Diamond_H q$. The first case is clear due to introspection, and for the second part we have $s \models \Box_S q \Rightarrow s \models \Box_S \Diamond_H q$ because $s \in \mathcal{T}$. Therefore, we have $s \models \Box_S p$ for all $p \in \Sigma$. This finally proves that $s \models \Sigma$. ■

Let $M \subseteq \mathcal{T}$ be a class which represents some property of possibilities. Then, we can formulate our result as follows: Let $M^0 := M$, $M^{2n+1} := M^{2n} \cap \Box_S M^{2n}$, $M^{2n+2} := M^{2n+1} \cap \Diamond_H M^{2n+1}$, and

$$\nabla M := \bigcap_{n \in \mathbf{N}} M^n,$$

then

Lemma 4.2 *If $M \subseteq \mathcal{T}$, then $\nabla M = M \cap \Box_S M$.*

If a is the act where the speaker asserts a sentence ψ, then $\nabla[\psi]$ is the class of possibilities where the assertion is mutually justified. It is the class of basic cases.

4.2 The Derived Extended Uses

We now show how to derive extended uses. We proceed by recursion over the order type of possibilities. For this end, we define restricted versions of the operators. They produce only possibilities of a certain maximal complexity. Our aim is to construct in every step α exactly all possible extensions of complexity α. That we reached our aim will be proved in Lemma 4.3. For Example **(7)** it is essential that we can derive in later steps new extensions with the same complexity. In Lemma 4.3 we show that this can not happen. Let Q denote one of the operators $\Box$ or $\Diamond$ and $\mathcal{H}_*(\mathcal{T}) := \mathcal{H}(\mathcal{T}) \setminus \mathcal{T}$.

$$\begin{aligned} Q_X^{\leq \alpha} M &:= Q_X M \cap \mathcal{H}_*^{\alpha}(\mathcal{T}) \\ Q_X^{< \alpha} M &:= Q_X M \cap \{w \in \mathcal{H}_*^{\alpha}(\mathcal{T}) \mid \forall v \in w(X)\, \mathrm{otp}(v) < \alpha\} \end{aligned}$$

The elements of $Q_X^{\leq \alpha} M$ all have maximal complexity α and an internal hierarchical structure. No element of $\mathcal{T}$ is an element of $Q_X^{\leq \alpha} M$. $Q_X^{< \alpha} M$ adds the restriction that all epistemic alternatives of participant X have a complexity *lower* than α. Hence, if $w \in Q_X^{<\alpha}M$ and $\text{otp}(w) = \alpha$, then Fact 3.1 implies that $w \notin w(X)$.

With these operators we can define the operations which give us all extensions of a certain complexity. The following four operations correspond to the four possible intersections of classes, which we can derive using the four operators of Figure 2. For $\alpha > 0$ we define

$$\begin{aligned}
\delta_1^\alpha M &:= \Box_S^{\leq\alpha} M \cap \Diamond_H^{<\alpha} M \\
\delta_2^\alpha M &:= \Box_S^{<\alpha} M \cap \Diamond_H^{\leq\alpha} \Box_S^{<\alpha} M \\
\delta_3^\alpha M &:= \Box_S^{\leq\alpha} \Diamond_H^{<\alpha} M \cap \Diamond_H^{<\alpha} M \\
\delta_4^\alpha M &:= \Box_S^{\leq\alpha} \Diamond_H^{<\alpha} M \cap \Diamond_H^{\leq\alpha} \Box_S^{<\alpha} M
\end{aligned}$$

For example, $w \in \delta_1^\alpha M$ means that the speaker is convinced that the order type of w is smaller than α and that it belongs to M, and that the hearer is also convinced that the order type of w is lower than α and that it might belong to M.

We define $\Delta_0(M) := M$, and $\Delta_{<\alpha}(M) := \bigcup_{\beta<\alpha} \Delta_\beta(M)$. For $\alpha > 0$ we set

$$\Delta_\alpha(M) := \Delta_{<\alpha}(M) \cup \bigcup_{i=1}^{4} \delta_i^\alpha \Delta_{<\alpha}(M).$$

We set $\Delta(M) := \bigcup_\alpha \Delta_\alpha(M)$. By definition of δ_i^α and $\Delta_{<\alpha}(M)$ it follows that $\Delta_\alpha(M) \subseteq \mathcal{H}^\alpha(\mathcal{T})$. The following lemma shows that we get for all α exactly all derived possible extensions of M of complexity α.

Lemma 4.3 *Let $M \subseteq \mathcal{T}$. For all ordinals α:*

$$\forall w \in \Delta_\alpha(M) \forall \beta < \alpha \ (w \in \Delta_\beta(M) \Leftrightarrow \text{otp}(w) \leq \beta).$$

Proof: Let $w \in \Delta_\alpha(M)$ and $\beta < \alpha$. The direction from left to right is trivial. Assume that $\text{otp}(w) = 0$. By definition of the operators $\Box^{\leq}$, $\Diamond^{\leq}$ etc. it follows that $\Delta(M) \subseteq M \cup \mathcal{H}_*(\mathcal{T}) = M \cup (\mathcal{H}(\mathcal{T}) \setminus \mathcal{T})$. $w \in \mathcal{H}(\mathcal{T})$ and $\text{otp}(w) = 0$ implies $w \in \mathcal{T}$ by Lemma A.1. Hence, $w \in \Delta(M)$ & $\text{otp}(w) = 0$ implies $w \in M = \Delta_0(M)$. The next fact follows from $(\mathbf{H_2})$ and from $w \in \mathcal{H}_*(\mathcal{T}) \Rightarrow \text{otp}(w) > 0$:

Fact 4.4 *$w \in \mathcal{H}_*(\mathcal{T})$ implies (1) $\forall v \in T(w)\, w \notin v(S) \cap v(H)$ and (2) $\forall v \in T(w) \neg\exists u, u' \, (u \in v(S)$ & $u' \in v(H)$ & $w \in T(u) \cap T(u'))$.*

So assume $\text{otp}(w) \leq \beta$, $\alpha = \beta + 1$. Assume further that we have proved that $w \in \Delta_{\beta'}(M)$ for $\text{otp}(w) = \beta' < \beta$. Hence, let $\text{otp}(w) = \beta$. We consider only the case $w \in \delta_4^\alpha \Delta_{<\alpha}(M)$. Therefore, $w \in \Box_S^{\leq\beta} \Diamond_H^{\leq\beta} \Delta_\beta(M) \cap \Diamond_H^{\leq\beta} \Box_S^{\leq\beta} \Delta_\beta(M)$.

Let $v \in w(S)$, $\text{otp}(v) = \beta$. Then $v \in \Diamond_H^{\leq\beta} \Delta_\beta(M)$. Let $u \in v(H)$. Suppose $\text{otp}(u) = \beta$. Then $w \in T(u)$ by Fact 3.1, and therefore $w \in v(H)$ by $(\mathbf{H_2})$. By $(\mathbf{H_2})$ it follows also that $w \in w(S)$, and by introspection $w \in v(S)$. But then $w \in v(S) \cap v(H)$, in contradiction with Fact 4.4. Hence, for all $u \in v(H)\, \text{otp}(u) < \beta$. Therefore, $v \in \Diamond_H^{<\beta} \Delta_{<\beta}(M)$. As v was arbitrary, it follows that $w \in \Box_S^{\leq\beta} \Diamond_H^{<\beta} \Delta_{<\beta}(M)$. If for all $v \in w(S)\, \text{otp}(v) < \beta$, then $w \in \Box_S^{<\beta} \Diamond_H^{<\beta} \Delta_{<\beta}(M) \subseteq \Box_S^{\leq\beta} \Diamond_H^{<\beta} \Delta_{<\beta}(M)$.

Let $v \in w(H) \cap \Box_S^{\leq\beta} \Delta_\beta(M)$. Assume that $\text{otp}(v) = \beta$. Let $u \in v(S)$. Suppose $\text{otp}(u) = \beta$. Then $w \in T(u)$, and therefore $w \in v(S)$ by $(\mathbf{H_2})$. But

$v \in w(H) = v(H)$, therefore we also have $w \in v(H)$, in contradiction with Fact 4.4. Hence, for all $u \in v(S)\,\mathrm{otp}(u) < \beta$. Hence, $v \in w(H) \cap \Box_S^{<\beta}\Delta_{<\beta}(M)$, and therefore $w \in \Diamond_H^{\leq\beta}\Box_S^{<\beta}\Delta_{<\beta}(M)$. If for all $v \in w(H)\,\mathrm{otp}(v) < \beta$, then $w \in \Diamond_H^{<\beta}\Box_S^{<\beta}\Delta_{<\beta}(M) \subseteq \Diamond_H^{\leq\beta}\Box_S^{<\beta}\Delta_{<\beta}(M)$.

Hence, $w \in \Box_S^{\leq\beta}\Diamond_H^{<\beta}\Delta_{<\beta}(M) \cap \Diamond_H^{\leq\beta}\Box_S^{<\beta}\Delta_{<\beta}(M)$. Therefore $w \in \Delta_\beta(M)$.

Next, assume that α is a limit ordinal. Again, we consider only the case $w \in \delta_4^\alpha\Delta_{<\alpha}(M)$. Let $\mathrm{otp}(w) = \beta < \alpha$. We suppose that we have proven that $v \in \Delta_{\beta'}(M)$ for $\mathrm{otp}(v) = \beta' < \alpha$ and $v \in \Delta_{<\alpha}(M)$. This allows us to conclude that $w \in \Box_S^{\leq\beta+1}\Diamond_H^{<\beta+1}\Delta_{<\beta+1}(M) \cap \Diamond_H^{\leq\beta+1}\Box_S^{<\beta+1}\Delta_{<\beta+1}(M) = \delta_4^{\beta+1}\Delta_{<\beta+1}(M) \subseteq \Delta_{\beta+1}(M)$. By I.H. it follows that $w \in \Delta_\beta(M)$. ■

Remark 4.5 *Lemma 4.3 remains valid if we replace $\Box_S$ in the definition of $\Delta(M)$ by an operator Q which has the form $Q\,M = \Box_S(M \cap C)$ for some class $C \subseteq \mathcal{H}(\mathcal{T})$.*

We can go through the proof of Lemma 4.3 and see that all inferences remain valid.

5 Speaker's Goals, Perspectives and Assertions

We introduce now explicit representations for the goals of the speaker. We model goals of a participant as a function mapping his epistemic possibilities into subsets of all possibilities, i.e. this function tells us for each of his epistemic possibilities which situations are desirable for him. We have seen in Section 3.2 why we have to introduce speaker's goals. Example **(7)** shows that these goals together with some rationality constraints may lead to difficult problems. In case of extended uses we can avoid these problems by constructing the extensions relative $\mathcal{H}(\mathcal{T})$. In this class all possibilities have an internal hierarchical structure except for those possibilities which belong to $\mathcal{T}$. We will see that we can use essentially the same construction of extended uses for possibilities with speaker's goals. But the circular structures of $\mathcal{T}$ will lead to serious problems for the characterisation of basic cases.

Let $\mathcal{M}$ denote a class of models for the possible outer situations.

- A *possibility* is a triple $w = \langle s_w, \langle w(S), G_S^w\rangle, \langle w(H)\rangle\rangle$ such that
  - $s_w \in \mathcal{M}$,
  - $w(S)$ and $w(H)$ are information states,
  - G_S^w is a function with: (1) $\mathrm{dom}\,G_S^w = w(S)$, and (2) $\forall v \in w(S)\,G_S^w(v)$ is an information state.
- An *information state* is a set of possibilities.

We denote the class of all possibilities with representations for the goals by $\mathcal{W}_G$. Transitivity of a class is defined in the same way as in Section 2: $M \subseteq \mathcal{W}_G$ is *transitive* iff $\forall w \in M\,\forall X \in \mathrm{DP}\,w(X) \subseteq M$. I.e. if $w \in M$, and if M is transitive, then the information states $G_S^w(v)$ don't need to be subsets of M.

For *introspectivity* we have to add a condition which guarantees that the real goals of the speaker and those of all his epistemic alternatives are the same. Let $\mathcal{I}_G$ denote the largest transitive subclass of $\mathcal{W}_G$ such that for all $X \in \mathrm{DP}$ for all $w \in \mathcal{I}_G$

- $\forall v \in w(X)\, w(X) = v(X)$,
- $\forall v \in w(S)\, G_S^w = G_S^v$.

We denote by $\dot{\mathcal{I}}_G$ the largest transitive subclass of $\mathcal{I}_G$ such that $\forall w \in \dot{\mathcal{I}}_G\, \forall X \in \mathrm{DP}\, w(X) \neq \emptyset$. We take $\mathcal{T}_G$ to be the largest transitive subclass of $\dot{\mathcal{I}}_G$ such that for all $w \in \mathcal{T}_G$

- $w \in w(S) \cap w(H)$
- $\forall v \in w(S)\, G_S^w(v) \subseteq \mathcal{T}_G$.

The second condition is a kind of *sincerity* condition, i.e the speaker does not want to mislead the hearer.[8]

We modify our definition of *normal mutual update* for possibilities with speaker's goals: Let a be a function with domain $D(a)$ then a is a *normal mutual update* if (**NMU**) holds, and if

$$\forall w \in D(a)\, \forall v \in w(S) \cap D(a)\; G_S^{a(w)}(a(v)) = G_S^w(v).$$

I.e. we assume that an update with the information represented by $D(a)$ does not change speaker's goals.

The definitions of $T(w)$, otp and $\mathcal{H}(M)$ remain the same as in the previous sections except that $\mathcal{W}$ is replaced by $\mathcal{W}_G$. $\mathcal{F}_G$ denotes the subclass of $\dot{\mathcal{I}}_G$ where otp is defined.

In Section 3.2 we have introduced rationality constraints linked to speaker's goals. For an action a and a goal $G(w)$ the constraints looked as follows:

$$(\mathbf{R_1})\;\; w \notin G(w) \qquad (\mathbf{R_2})\;\; a(w) \in G(w)$$

I.e. we should only perform an action if our goal is not yet achieved, and we should only choose actions which allow us to reach our goal.

We modify the $\Box$ operator in such a way that we check the rationality constraints at the time when we apply the operator. Let $M \subseteq \mathcal{F}_G$ be an arbitrary class:

$$\begin{aligned}[\mathrm{R}]_S M &:= \{w \in \dot{\mathcal{I}}_G \mid \forall v \in w(S)\, (v \in M \,\&\, v \notin G_S^v(v) \,\&\, a(v) \in G_S^v(v))\} \\ &= \Box_S \Big(M \cap \{v \in \dot{\mathcal{I}}_G \mid v \notin G_S^v(v) \,\&\, a(v) \in G_S^v(v)\}\Big)\end{aligned}$$

Remember that $G_S^w = G_S^v$ for all $v \in w(S)$. Therefore, we could also have written G_S^w instead of G_S^v in the definition of the operator. $[\mathrm{R}]_S M$ is the class of all possibilities where S is convinced that the real situation belongs to M, a can be performed successfully in order to reach his goal, and where this goal is not already reached. The axioms ($\mathbf{R_1}$) and ($\mathbf{R_2}$) hold by definition of the operator for all epistemic alternatives of S. It follows that $\Diamond_H [\mathrm{R}]_S M$ denotes the class of all possibilities where H thinks it is possible that S can choose a in order to reach his goal.

[8] We will see in Example **(11)** why we need such a condition.

5.1 The Base Case Problem

In Section 4.1 we argued that the speaker should be convinced that his assertion is successful, and the hearer should know that this can be the case. Then the speaker should be convinced that both conditions hold, and the hearer should again know that this is possible. We have argued that, in principle, we must impose these conditions again and again. If M denotes the class where some sentence ψ is true, then we denote the class where its assertion is *mutually justified* as ∇M. For possibilities without speaker's goals we could provide a simple characterisation of ∇M. In this section we will see that this is not possible if we add goals.

We see that the condition for the hearer does not impose a real restriction because $M \subseteq \mathcal{T}_G$ implies $M \cap \Diamond_H M = M$. Hence, $M \cap [\mathrm{R}]_S M \cap \Diamond_H M = M \cap [\mathrm{R}]_S M$. Let a_M denote the normal mutual update determined by M. With $w \in \mathcal{T}_G \Rightarrow w \in w(S)$ and introspection we find:

$$\begin{aligned}
&M \cap [\mathrm{R}]_S(M \cap [\mathrm{R}]_S M) \; = \\
&= \; M \cap [\mathrm{R}]_S\{w \in M \mid w(S) \subseteq M \;\&\; \forall v \in w(S)(v \notin G^w_S(v) \;\&\; a_M(v) \in G^w_S(v))\} \\
&= \; \{w \in M \mid w(S) \subseteq M \;\&\; \forall v \in w(S)(v \notin G^w_S(v) \;\&\; a_M(v) \in G^w_S(v))\} \\
&= \; M \cap [\mathrm{R}]_S M.
\end{aligned}$$

If we adopt the definition from Section 4.1 and define ∇M by $M^1 := M \cap [\mathrm{R}]_S M$, $M^{n+1} := M^n \cap [\mathrm{R}]_S M^n \cap \Diamond_H M^n$, $\nabla M := \bigcap_{n \in \mathbf{N}} M^n$, then the considerations above seem to show that we can again characterise ∇M as $M \cap [\mathrm{R}]_S M$. But the following example shows that this is not correct:

(8) Helga calls up her son Stephan early on Sunday morning and asks him whether he wants to visit her. They both know that Stephan would have checked the weather at this time only if he needed an excuse for not accepting the expected invitation. It happens that he knows that it is snowing, and he does not want to let her know that he would prefer to stay at home this day. Should he tell his mother that he can't come because *it is snowing in the mountains*?

In Example **(8)** it is true that (ψ) *it is snowing in the mountains*, and we assume that, if Helga learns this fact, she will excuse her son for not visiting her, and that she can't know whether Stephan likes to visit her or not. Hence, if M denotes the set of all $w \in \mathcal{T}_G$ where *it is snowing in the mountains*, then the situation is an element of $M \cap [\mathrm{R}]_S M$, and we would predict that Stephan will be successful if he tells her that ψ. But of course this implies that Helga also learns that Stephan knows that ψ, hence she can conclude that he does not like to visit her, which is an undesired result for Stephan. This means that Stefan should check in the second step whether the mutual update with $M \cap [\mathrm{R}]_S M$ also leads to a desired situation. The problem is that the above definition of ∇M does only capture the fact that the hearer learns ψ but not that she also learns that the speaker knows that ψ and that it is reasonable for him to assert ψ, and that he knows that she knows this, etc.

In general, we have to find a fixed–point of the update operations. This is a non–trivial task. First of all, we have to find the correct fixed–point condition. In order to capture Example **(8)**, we have to search for sets M which are fixed–

points of the following operator. Let $M \subseteq \mathcal{T}_G$:

$$J\,M := \{w \in M \mid w(S) \subseteq M \;\&\; \forall v \in w(S)(v \notin G_S^w(v) \;\&\; a_M(v) \in G_S^w(v))\}$$

Let $M_0 := M$ and $M_{n+1} := J\,M_n$. Then, M_1 is the set of worlds where the rationality conditions for asserting M are satisfied. It follows especially that $v \in M_1$ implies that mutually updating with M leads to a desired outcome. M_1 may be a proper subset of M, hence, mutually learning that M_1 may lead to a different outcome than simply mutually learning that M. Hence, we have to check in the next step whether mutually learning M_1 is desired by the speaker. In general, in step $n+1$ we check whether mutually learning M_n is desirable for the speaker.

In this paper, we use a strong assumption about speaker's goals which guarantees that a fixed–point can easily be constructed. The assumption states: if the speaker can reach his goal by a mutual update with some information M, then he can also reach it by updating with any stronger information $N \subseteq M$. Hence, in the following, we consider only sets $M \subseteq \mathcal{T}_G$ such that for all $w \in M$ and $v \in w(S)$:

$$(\mathbf{R_3})\;\; \forall\, N \subseteq M:\; a_M(v) \in G_S^w(v) \;\&\; v(S) \subseteq N \Rightarrow a_N(v) \in G_S^w(v)$$

With the foregoing, we can easily show the following lemma.

Lemma 5.1 *Let $M \subseteq \mathcal{T}_G$ be such that $(\mathbf{R_3})$ holds for all $v \in M$. Then, $\nabla M := M \cap [\mathrm{R}]_S M$ is a fixed–point the J operator.*

5.2 The Derived Extended Uses

In this section we show how to derive extended uses of assertion if we add rationality constraints linked to speaker's goals to our considerations. In principle we proceed in the same way as in Section 4.2 by recursively defining extensions $\Delta_\alpha(\mathcal{B})$ of some base set $\mathcal{B}$. If we know $\Delta_\alpha(\mathcal{B})$ and want to decide whether some w belongs to $\Delta_{\alpha+1}(\mathcal{B})$ then we have to check wether the update with the information carried by the assertion leads to a desirable state. But this means that we have to know $\Delta_{\alpha+1}(\mathcal{B})$ in order to check the rationality principles. As a consequence, we cannot directly define $\Delta_{\alpha+1}(\mathcal{B})$ by using the operators linked to perspectives. Therefore we proceed as follows: Given some set $M \subseteq \mathcal{T}_G$ we provide an axiomatic characterisation of possible extensions $\Delta_\alpha(M)$ of M and then show that there exists exactly one sequence $(\Delta_\alpha(M))_{\alpha \in \mathbf{On}}$ which satisfies the following axioms $(\mathbf{D_1})$ and $(\mathbf{D_2})$. Let $(\Delta_\alpha(M))_{\alpha \in \mathbf{On}}$ be a sequence of subclasses of $\mathcal{H}(\mathcal{T})$. The first axiom $(\mathbf{D_1})$ states that for each α the elements of $\Delta_\alpha(M)$ with order type $\beta \leq \alpha$ are already elements of $\Delta_\beta(M)$.

$$(\mathbf{D_1})\;\; \forall \beta \leq \alpha\; \Delta_\alpha(M) \cap \mathcal{H}^\beta(\mathcal{T}) = \Delta_\beta(M).$$

Note that $(\mathbf{D_1})$ implies $\Delta_0(M) = M$ and $\Delta_\alpha(M) \subseteq \mathcal{H}^\alpha(\mathcal{T})$.

Let a_α be the normal mutual update determined by $\Delta_\alpha(M)$, and a_Δ the normal mutual update determined by $\Delta(M) := \bigcup_\alpha \Delta_\alpha(M)$. $(\mathbf{D_1})$ implies $a_\alpha = a_\Delta|_{\mathcal{H}^\alpha(\mathcal{T})}$, and especially it follows $w \in \Delta_\alpha(M) \Rightarrow a_\alpha(w) = a_\Delta(w)$. We define the operator [R] which combines the $\Box$ operator and the checking of rationality axioms $(\mathbf{R_1})$, $(\mathbf{R_2})$ relative to a_Δ:

$$[\mathrm{R}]_S N := \{w \in \dot{\mathcal{I}}_G \mid \forall v \in w(S)\,(v \in N \;\&\; v \notin G_S^v(v) \;\&\; a_\Delta(v) \in G_S^v(v))\}.$$

The next axiom demands that the new elements of $\Delta_\alpha(M)$ are those which can be derived by application of the restricted versions of perspectival operators motivated in Section 3.

($\mathbf{D_2}$) Let $w \in \mathcal{H}^\alpha(\mathcal{T})$ and $\Delta_{<\alpha}(M) := \bigcup_{\beta<\alpha} \Delta_\alpha(M)$. Then $w \in \Delta_\alpha(M)$, iff $w \in \Delta_{<\alpha}(M)$ or there exists $N \subseteq \Delta_{<\alpha}(M)$ such that w is an element of one of the following sets:

$$\delta_1^\alpha N := [\mathrm{R}]_S^{<\alpha} N \cap \diamond_H^{<\alpha} N; \qquad \delta_2^\alpha N := [\mathrm{R}]_S^{<\alpha} N \cap \diamond_H^{\leq\alpha} [\mathrm{R}]_S^{<\alpha} N$$
$$\delta_3^\alpha N := [\mathrm{R}]_S^{\leq\alpha} \diamond_H^{<\alpha} N \cap \diamond_H^{<\alpha} N; \quad \delta_4^\alpha N := [\mathrm{R}]_S^{\leq\alpha} \diamond_H^{<\alpha} N \cap \diamond_H^{\leq\alpha} [\mathrm{R}]_S^{<\alpha} N$$

We call $(\Delta_\alpha)_{\alpha\in\mathbf{On}}$ a *possible derived extension* of M if the axioms ($\mathbf{D_1}$) and ($\mathbf{D_2}$) hold.

The condition of ($\mathbf{D_2}$) does not allow for a direct definition of $\Delta_\alpha(M)$ because the definition of the [R] operator presupposes that $\Delta_\alpha(M)$ is already defined. We show that for every $M \subseteq \mathcal{T}_G$ there exists a unique possible derived extension $\Delta(M)$, see Theorem 5.3. We define an extension $\Delta(M)$ and a_Δ simultaneously by recursion over α.

Assume that $\Delta_\beta(M)$ and a_β are defined for $\beta < \alpha$ and that the restricted versions of ($\mathbf{D_1}$) and ($\mathbf{D_2}$) hold for $(\Delta_\beta(M))_{\beta<\alpha}$. We show that this already implies that $[\mathrm{R}]_S^{<\alpha}N$ and $[\mathrm{R}]_S^{\leq\alpha}\diamond_H^{<\alpha}N$ are uniquely defined for $N \subseteq \Delta_{<\alpha}(M)$. First we see that by definition $[\mathrm{R}]_S^{<\alpha}N$ must be equal to:

$$\{w \in \mathcal{H}^\alpha(\mathcal{T}) \mid w(S) \subseteq N \,\&\, \forall v \in w(S)(a_\Delta(v) \in G^v(v) \,\&\, v \notin G^v(v))\}$$

If we write $a_{<\alpha}$ for the normal mutual update determined by $\Delta_{<\alpha}(M)$, then ($\mathbf{D_1}$) implies that this is equal to:

$$\{w \in \mathcal{H}^\alpha(\mathcal{T}) \mid w(S) \subseteq N \,\&\, \forall v \in w(S)(a_{<\alpha}(v) \in G^v(v) \,\&\, v \notin G^v(v))\}.$$

But $a_{<\alpha}$ is given by I.H., hence, $[\mathrm{R}]_S^{<\alpha}N$ is uniquely defined.

Next we consider $[\mathrm{R}]_S^{\leq\alpha}\diamond_H^{<\alpha}N$. By definition it must be equal to:

$$\{w \in \mathcal{H}^\alpha(\mathcal{T}) \mid w(S) \subseteq \diamond_H^{<\alpha}N \,\&\, \forall v \in w(S)(a_\Delta(v) \in G^v(v) \,\&\, v \notin G^v(v))\}$$

Hence, due to introspection, $w \in [\mathrm{R}]_S^{\leq\alpha}\diamond_H^{<\alpha}N$ implies $w(S) \subseteq [\mathrm{R}]_S^{\leq\alpha}\diamond_H^{<\alpha}N$. This implies that $a_\Delta(w(S)) = \{a_\Delta(v) \mid v \in w(S)\}$. Hence, it must hold that

$$(*) \quad a_\Delta(w) = \langle s_w, \langle \{a_\Delta(v) \mid v \in w(S)\}, G^{a_\Delta(w)}\rangle, \langle a_\Delta(w(H))\rangle\rangle.$$

We can assume that $w \in w(S)$, hence $w(H) \subseteq \diamond_H^{<\alpha}N$. As $a_\Delta|_{w(H)}$ is defined by I.H. it follows that $a_\Delta(w)$ is uniquely defined by this equation[9]. Hence, we can define $[\mathrm{R}]_S^{\leq\alpha}\diamond_H^{<\alpha}N$. This shows that $\Delta_\alpha(M)$ exists and is uniquely determined by ($\mathbf{D_2}$) and I.H. We have to check that ($\mathbf{D_1}$) and ($\mathbf{D_2}$) hold also for the sequence $(\Delta_\alpha(M))_{\beta\leq\alpha}$. First, we show ($\mathbf{D_1}$), i.e. we show that we get in step α exactly all new situations of complexity α:

[9]The last equation defines a *system of equations*, where the $a_\Delta(v)$ for $v \in T(w)$ with $\mathrm{otp}(v) \geq \alpha$ denote new *urelements* which function as unknown parameters. (AFA) set theory guarantees that it has a unique solution for *all* w. Then, the set of new possibilities in $[\mathrm{R}]_S^{\leq\alpha}\diamond_H^{<\alpha}N$ is equal to the set of all $w \in \mathcal{H}^\alpha(\mathcal{T})$ such that (1) $w(H) \subseteq \diamond_H^{<\alpha}N$ and such that (2) the solution a_Δ of $(*)$ satisfies $\forall v \in w(S)(a_\Delta(v) \in G^v(v) \,\&\, v \notin G_S^v(v))$.

Lemma 5.2 $\forall\gamma \leq \beta \leq \alpha\ \Delta_\alpha(M)^\gamma = \Delta_\beta(M)^\gamma$

But this follows from Remark 4.5 because of:

$$[R]_S N = \Box_S \left(N \cap \{v \in \dot{\mathcal{I}}_G \mid v \notin G_S^v(v)\ \&\ a_\Delta(v) \in G_S^v(v)\} \right)$$

This shows that $(\mathbf{D_1})$ holds. But this means that $a_\alpha|_{\mathcal{H}^{<\alpha}(\mathcal{T})} = a_{<\alpha}$. Therefore, it follows that a_α is also a solution for $(*)$. This implies that $(\mathbf{D_2})$ holds. Hence, this construction defines recursively a unique extension a_Δ simultaneously with $\Delta(M)$. We summarise the result as:

Theorem 5.3 *Let $M \subseteq \mathcal{T}_G$. Then there exists a unique possible extension $\Delta(M)$.*

6 Applications

We apply our theory to examples introduced in the previous sections. A possibility w is a triple $\langle s_w, \langle w(S), G_S^w \rangle, \langle w(H) \rangle \rangle$ where s_w is a model for the *outer* situation. As we are only interested in the truth or falsity of the sentence (ψ) '*It is snowing in the mountains*', we use the respective formula to denote this model. We denote by $[\psi]$ the set of all possibilities in $\mathcal{T}_G$ where ψ is true. We write a_M for the normal mutual update determined by some M. We can summarise the results of the last sections as follows:

- In a basic case an assertion *that* ψ must be mutually justified. We denote the class of all these basic situations by $\mathcal{B} = \nabla[\psi]$.
- The maximal class where the speaker can reasonably assert that ψ, and where the hearer can interpret this assertion, is given by $\Delta(\mathcal{B})$.
- The update effect is given by $a_{\Delta(\mathcal{B})}$, the normal mutual update determined by $\Delta(\mathcal{B})$.

Theorem 5.3 guarantees that $\Delta(\mathcal{B})$ and $a_{\Delta(\mathcal{B})}$ are always defined.

The basic case is exemplified by Example **(1)**. We can represent the utterance situation w_1 by the following equations[10]:

$$\begin{aligned} w_1 &= \langle \psi, \langle \{w_1\}, G_S^{w_1} \rangle, \langle \{w_1, v_1\} \rangle \rangle \\ v_1 &= \langle \neg\psi, \langle \{v_1\}, G_S^{v_1} \rangle, \langle \{w_1, v_1\} \rangle \rangle, \end{aligned}$$

i.e. it is the case that ψ, H does not know it but knows that S knows whether ψ. The intended resulting state is described by the equation

$$s_1 = \langle \psi, \langle \{s_1\}, G_S^{s_1} \rangle, \langle \{s_1\} \rangle \rangle,$$

where $G_S^{s_1}(s_1) = G_S^{w_1}(w_1)$. Hence, we assume that $\{s_1\} = G_S^{w_1}(w_1)$, and $\{t_1\} = G_S^{v_1}(v_1)$, where $t_1 = \langle \neg\psi, \langle \{t_1\}, G_S^{t_1} \rangle, \langle \{t_1\} \rangle \rangle$, i.e. if ψ holds, then S wants that they mutually know that ψ, and if $\neg\psi$ holds, then he wants that they mutually

[10] Of course, this interpretation is not fully justified by the way the example was stated. There are a lot of dialogue situations where this can be part of the description. We don't want to explain how we arrive at such a strong reading, but only, given the reading, why the utterance of ψ is reasonable for the speaker and can be interpreted by the hearer.

know that $\neg\psi$ holds. Clearly, $w_1, v_1, s_1 \in \mathcal{T}_G$. We find that $w_1 \notin G_S^{w_1}(w_1)$, and $s_1 \in G_S^{w_1}(w_1)$. Hence, $w_1 \in [\mathrm{R}]_S\{w_1\}$, and therefore by Lemma 5.1: $w_1 \in \nabla\mathcal{B}$.

For a proof of $a_\mathcal{B}(w_1) = s_1$ we would need some additional techniques from (AFA) set theory. The general properties of normal mutual updates imply

$$a_\mathcal{B}(w_1) = \langle \psi, \langle \{a_\mathcal{B}(w_1)\}, G_S^{a_\mathcal{B}(w_1)} \rangle, \langle \{a_\mathcal{B}(w_1)\} \rangle \rangle.$$

But this equation is structurally identical with the equation for s_1. Using the Solution Lemma, Theorem 2.1, it is provable that the two equations have the same solution.

Hence, the theory predicts that it is reasonable for S to say that ψ, and that it will be successful.

In Example **(2)** the beliefs and goals are the same as in **(1)**:

$$w_2 = \langle \neg\psi, \langle \{w_1\}, G_S^{w_1} \rangle, \langle \{w_1, v_1\} \rangle \rangle.$$

Hence, it is an element of $\Box_S^{\leq 1}\mathcal{B} \cap \Diamond_H^{\leq 1}\mathcal{B} = \delta_1^1\mathcal{B} \subseteq \Delta_1(\mathcal{B})$. Hence it is reasonable for S to say that ψ, and the hearer will interpret his utterance in the same way as in situation **(1)**. Next, we consider Example **(3)**.

$$w_3 = \langle \neg\psi, \langle \{w_1\}, G_S^{w_1} \rangle, \langle \{w_3\} \rangle \rangle.$$

Here the information of the speaker is the same as in the basic situation w_1, and as in w_2. But this time, the hearer knows that it is not snowing in the mountains ($\neg\psi$), and she is aware of the entire dialogue situation. w_3 is an element of $\Box_S^{\leq 1}\{w_1\} \cap \Box_H^{\leq 1}\Box_S^{\leq 1}\{w_1\} \subseteq \delta_2^1\mathcal{B} \subseteq \Delta_1(\mathcal{B})$. The update of w_2 and w_3 with $a_{\Delta_1(\mathcal{B})}$ leads to s_2 and s_3, where

$$\begin{aligned} s_2 &= \langle \neg\psi, \langle \{s_1\}, G_S^{s_1} \rangle, \langle \{s_1\} \rangle \rangle, \\ s_3 &= \langle \neg\psi, \langle \{s_1\}, G_S^{s_1} \rangle, \langle \{s_3\} \rangle \rangle. \end{aligned}$$

In the next example, Example **(9)**, S is no longer sincere. He is lying successfully. It is a case where the belief state of the hearer is the same as in the basic situation, and where the speaker knows this but knows also that ψ is not true. The speaker can exploit this situation and deceive the hearer.

(9) Helga calls up her son Stephan and asks him whether he wants to visit her in Munich. Stephan, who has absolutely no inclination to drive to Munich this day, answers: "It is snowing in the mountains."

Now, the hearer updates as in the basic case, and the speaker updates only his representation of the hearer's beliefs.

$$w_4 = \langle \neg\psi, \langle \{w_4\}, G_S^{w_4} \rangle, \langle \{w_1, v_1\} \rangle \rangle,$$

with $G_S^{w_4}(w_4) = \{s_4\}$, where

$$s_4 = \langle \neg\psi, \langle \{s_4\}, G_S^{s_4} \rangle, \langle \{s_1\} \rangle \rangle,$$

and $G_S^{s_4}(s_4) = \{s_4\}$. w_4 is an element of $\Box_S^{\leq 1}\Diamond_H^{\leq 1}\{w_1\} \cap \Diamond_H^{\leq 1}\{w_1\} \subseteq \delta_3^1\mathcal{B} \subseteq \Delta_1(\mathcal{B})$. We can see that $a_{\Delta_1(\mathcal{B})}(w_4) = s_4$. Hence, the speaker should be able to successfully mislead the hearer.

Example **(4)** receives the following representation:

$$w_5 \quad = \quad \langle \neg\psi, \langle \{w_4\}, G_S^{w_4} \rangle, \langle \{w_5\} \rangle \rangle.$$

I.e. the speaker remains in the same situation as in Example **(9)** but now the hearer knows that the speaker wants to mislead him. w_5 is an element of $\Box_S^{\leq 1} \Diamond_H^{\leq 1} \{w_1\} \cap \Diamond_H^{\leq 2} \Box_S^{\leq 2} \{w_4\} \subseteq \delta_4^2 \{w_4\} \subseteq \Delta_2(\mathcal{B})$. We can see that $a_{\Delta_2(\mathcal{B})}(w_5) = s_5$ with $s_5 = \langle \neg\psi, \langle \{s_4\}, G_S^{s_4} \rangle, \langle \{s_5\} \rangle \rangle$. Hence, the speaker thinks that he has successfully misled the hearer, and the hearer knows this.

The next example is a case where the hearer *suspects* that the speaker might lie.

(10) Helga calls up her son Stephan. She knows that he does not like to visit her, hence she suspects that he will not be honest. But she knows also that Stephan can't know this. She asks him whether he wants to visit her, and Stephan replies that *it is snowing in the mountains.* And, indeed, it was not a lie.

$$\begin{aligned} w_7 &= \langle \psi, \langle \{w_1\}, G_S^{w_1} \rangle, \langle \{w_7, u_7, v_7\} \rangle \rangle \\ u_7 &= \langle \neg\psi, \langle \{v_1\}, G_S^{v_1} \rangle, \langle \{w_7, u_7, v_7\} \rangle \rangle \\ v_7 &= \langle \neg\psi, \langle \{w_4\}, G_S^{w_4} \rangle, \langle \{w_7, u_7, v_7\} \rangle \rangle. \end{aligned}$$

This is an element of $\Delta_2(\mathcal{B})$, and the update with $a_{\Delta_2(\mathcal{B})}$ results in s_7:

$$\begin{aligned} s_7 &= \langle \psi, \langle \{s_1\}, G_S^{s_1} \rangle, \langle \{s_7, t_7\} \rangle \rangle \\ t_7 &= \langle \neg\psi, \langle \{s_4\}, G_S^{s_4} \rangle, \langle \{s_7, t_7\} \rangle \rangle. \end{aligned}$$

I.e. Stephan believes that they now mutually believe that it is snowing in the mountains. His mother knows this to be the case if it were really snowing, and she believes that, if it were not snowing, he would think that he could deceive her successfully.

The following example shows that we need a *sincerity* condition for our basic situations. In **(11)** it is common knowledge that the hearer thinks that the speaker might lie.

(11) Helga calls up her son Stephan. The last time she invited him, he pretended that he could not come because of the bad weather. They both know that she will be suspicious this time, if he replies to her question that *it is snowing in the mountains.*

$$\begin{aligned} w_8 &= \langle \psi, \langle \{w_8\}, G_S^{w_8} \rangle, \langle \{w_8, v_8\} \rangle \rangle \\ v_8 &= \langle \neg\psi, \langle \{v_8\}, G_S^{v_8} \rangle, \langle \{w_8, v_8\} \rangle \rangle, \end{aligned}$$

where $G_S^{w_8}(w_8) = \{s_1\}$ and $G_S^{v_8}(v_8) = \{s_4\}$. Clearly, $\text{otp}(w_8) = 0$. But w_8 is not an element of $\mathcal{T}_G$ because $G_S^{v_8}(v_8) \not\subseteq \mathcal{T}_G$. Hence, $w_8 \notin \mathcal{T}_G \supseteq \nabla[\psi] = \mathcal{B}$. Therefore, it can't be an element of $\Delta(\mathcal{B})$. Without sincerity condition, i.e. without the condition $G_S^w(v) \subseteq \mathcal{T}_G$ for $w \in TG$ and $v \in w(S)$, w_8 would be an element of $\mathcal{T}_G$, and the general theory would predict that the hearer would update his belief state with the information that ψ. Intuitively, Stephan has to make clear that he is sincere to make his mother believe him. I.e. they mutually have to redefine $G_S^{v_8}(v_8)$ as t_1, but this results in w_1 and then, indeed, the assertion of ψ will result in s_1.

7 Conclusion

We searched for a characterisation of dialogue situations where a felicitous use of an assertion is possible. Felicity is taken here in the sense that (1) the speaker is convinced that the update triggered by the assertion leads to a desirable situation, and (2) the fact that the sentence was asserted does not contradict the beliefs of the interpreter. We distinguished between the use of assertions in *ideal* and *non–ideal* situations, where in ideal situations it holds that it is common knowledge that (1) everybody believes the real situation to be possible (*has knowledge*), and (2) the speaker does not want to mislead the hearer. We started with the felicity conditions and update effects of assertions as they are defined by pure semantics, i.e. we assume that a sentence with semantic content ψ can be asserted exactly iff ψ is true in the actual situation. We described the effect of an assertion by a normal mutual update with its semantic content. AFA set theory was an indispensable tool for working out our theory.

The pragmatically *basic* cases build the subclass $\mathcal{B}$ where the pragmatic constraints linked to perspectives and intentions hold. We saw that the implicit circularity of these constraints makes the task of finding $\mathcal{B}$ a non–trivial one.

We have argued that perspectives play an essential role in explaining extended uses in non–ideal situations. The idea was to start with a characterisation of basic situations, and then to derive new uses from ideal situations by systematic application of operators which reflect the way how partiality of information can give rise to uses in new situations.

The theory presented in this paper evolved out of our research concerning the felicity conditions for the referential use of definite descriptions (Benz, 1999). There we used the same ideas to explain the referential use in situations with arbitrarily complex belief structures. The present paper contains a much improved description of the underlying mathematical structures. In (Benz, 2000) we outlined on an informal level how the two papers combine to form a single theory of perspectives. Assertions are examples of speech acts, and speech acts are examples of goal directed linguistic actions. If a linguistic act can be characterised by its preconditions, the speaker's goals and if its effects can be described by mutual learning, then our approach should be applicable and explain how this linguistic act can have extended uses in defective dialogue situations. If successful, this theory greatly simplifies the analysis of linguistic acts. To find a characterisation of linguistic acts in ideal situations is normally a less complicated task. The theory of perspectival derivations then automatically predicts their use and effect in defective situations.

References

P. Aczel: *Non–Well–Founded Sets*; CSLI–Lecture Note 14, Stanford 1988

A. Baltag: *A Logic of Epistemic Actions*; Proceedings of the ESSLLI99 Workshop on Foundations and Applications of Collective Agent Based Systems, Utrecht 1999

A. Baltag, H. van Ditmarsch, L. Moss: *Epistemic Logic and information update*; to appear in: J. van Bentham, P. Adriaans (eds.): *Handbook on the Philosophy of Information*, Elsevier Science Publishers 2006

A. Baltag, L. Moss, S. Solecki: *The Logic of Public Announcements, Common Knowledge and Private Suspicion*; CWI Reports, SEN-R9923 1999

J. Barwise, L. Moss: *Vicious Circles*; Stanford, CSLI, 1996

A. Benz: *Perspectives and the Referential Use of Definite Descriptions in Dialogue*; ms., Berlin 1999

A. Benz: *Perspectives and Derived Extensions of Dialogue Acts*; in: Proceedings of the *Communicating Agents* Workshop, Bonn, in: *LDV-Forum 17*, Vol. 1/2 2000

R. Fagin, J.Y. Halpern, Y. Moses, M.Y. Vardi: *Reasoning About Knowledge*; Cambridge, Massachusetts 1995

A.A. Fraenkel, Y. Bar-Hillel: *Foundations of Set Theory*; North-Holland Publishing Company, Amsterdam 1958

J. Gerbrandy: *Bisimulations on Planet Kripke*; ILLC Dissertation Series, Institute for Logic, Language and Computation, Universiteit van Amsterdam 1998

J. Gerbrandy, W. Groeneveld: *Reasoning about Information Change*; Journal of Logic, Language and Information 6, pp. 147–169, 1997

W. Groeneveld: *Logical Investigations into Dynamic Semantics*; PhD thesis, ILLC dissertation series 1995–18, University of Amsterdam 1995

J. Hintikka: *Knowledge and Belief*; Ithaka 1962

J. Jaspars: *Calculi for Constructive Communication. A Study of the Dynamics of Partial Sets*; PhD thesis, ITK, Katholieke Universiteit Brabant 1994

H. Zeevat: *The Common Ground as a Dialogue Parameter*; in: A. Benz & G. Jäger (ed.), *Proceedings of the Munich Workshop on Formal Semantics and Pragmatics of Dialogue MunDial'97*, pp.195–214 München 1997

A The Structure of $\mathcal{H}(\mathcal{T})$

In this section we prove the claims we have made about $\mathcal{H}(\mathcal{T})$. We first show that "$(\mathbf{H_1})$–$(\mathbf{H_3})+\mathcal{H}(\mathcal{T}) \subseteq \mathcal{F}$" defines the same class as "$(\mathbf{H_{1'}})$–$(\mathbf{H_{3'}})+\mathcal{H}(\mathcal{T}) \subseteq \dot{\mathcal{I}}$". In the second part we show that we can construct $\mathcal{H}(\mathcal{T})$ by an iterated application of a combination of perspectival operators. The proofs in the subsequent sections do only make use of "$(\mathbf{H_1})$–$(\mathbf{H_3})+\mathcal{H}(\mathcal{T}) \subseteq \mathcal{F}$"[11]. The following lemma shows that this implies that (1) the complexity of the possibilities $w \in \mathcal{H}(\mathcal{T}) \setminus \mathcal{T}$ is greater than at least one participant believes it to be, (2) all situations with complexity 0 must be elements of $\mathcal{T}$, and (3) $(\mathbf{H_{3'}})$ holds.

Lemma A.1 *Let $w \in \mathcal{H}(\mathcal{T})$. Then*

1. *If $w \notin \mathcal{T}$, then $\exists X \in \mathrm{DP}$ such that $\forall v \in w(X)\, \mathrm{otp}(v) < \mathrm{otp}(w)$.*

[11] Hence, the reader may skip this rather technical section.

2. $\mathrm{otp}(w) = 0 \vee w \in w(S) \cap w(H) \Rightarrow w \in \mathcal{T}$.

3. *Let $X \neq Y$. If p is a function with $p(0) = w$, $p(2n+1) \in p(2n)(X)$, and $p(2n+2) \in p(2n+1)(Y)$, then there is an $n \in \mathbf{N}$ with $p(n) \in \mathcal{T}$.*

Proof: (1) By $(\mathbf{H_1})$ there exists $X \in \mathrm{DP}$ such that $w \notin w(X)$. With $(\mathbf{H_2})$ it follows for all $v \in w(X)$ that $w \notin T(v)$, which implies $\mathrm{otp}(w) \geq \mathrm{otp}(v) + 1$.

(2) If $w \in w(S) \cap w(H)$, it immediately follows from $(\mathbf{H_1})$ that $w \in \mathcal{T}$. If $w \notin \mathcal{T}$, it follows from *1.* that $\mathrm{otp}(w) \geq 1$.

(3) It follows from *1.* and $(\mathbf{H_2})$ that $\mathrm{otp}(p(n+2)) < \mathrm{otp}(p(n))$. Suppose that for all $n \in \mathbf{N}$ $p(n) \notin \mathcal{T}$. Then there exists an n with $\mathrm{otp}(p(n)) = 0$. But then, $p(n) \in \mathcal{T}$, which contradicts the assumption. ∎

As $(\mathbf{H_{1'}})$ is the same as $(\mathbf{H_1})$ and $(\mathbf{H_{3'}})$ implies $(\mathbf{H_{2'}})$, it follows that $(\mathbf{H_1})$–$(\mathbf{H_3}) + \mathcal{H}(\mathcal{T}) \subseteq \mathcal{F}$ imply $(\mathbf{H_{1'}})$–$(\mathbf{H_{3'}})$. Clearly, $(\mathbf{H_{3'}})$ implies for $w \in \dot{\mathcal{I}}$ that $w \in \mathcal{F}$. We need the following fact:

Fact A.2 *Let $X, Y \in \mathrm{DP}$, $X \neq Y$ and $w, v \in \mathcal{F}$. If $v \in w(X)$, then $T(v) = w(X) \cup \bigcup_{u \in v(X)} \bigcup_{u' \in u(Y)} T(u') =: T'(v)$.*

Clearly $T'(v) \subseteq T(v)$. As $T'(v)$ is transitive and $\{v\} \subseteq T'(v)$, it follows from definition of $T(v)$ that $T(v) \subseteq T'(v)$.

If $(\mathbf{H_2})$ does not hold for $w \in \mathcal{F}$, then there is a $v \in T(w)$ such that for some X $w \notin v(X)$ & $\exists u \in v(X)\, w \in T(u)$. Fact A.2 implies that $w \in u'(Y)$ for some $u' \in v(X)$, $Y \neq X$. But this means that we can construct a sequence $(v_n)_{n \in \mathbf{N}}$ which violates $(\mathbf{H_{2'}})$.

We can now provide a construction of $\mathcal{H}(\mathcal{T})$ which shows that we can derive this class by systematic application of perspectival operators. We will even show that we can apply them in such a way that we get in each step α of the recursion exactly all elements of $\mathcal{H}(\mathcal{T})$ with complexity α. We therefore introduce restricted versions of our operators:

- $\Box_X^{\leq\alpha} M := \Box_X M \cap \mathcal{F}^\alpha$,
- $\Box_X^{<\alpha} M := \{w \in \mathcal{F}^\alpha \mid w(X) \subseteq M \;\&\; \forall v \in w(X)\, \mathrm{otp}(v) < \alpha\}$.

This gives us more control over the complexity of derived candidates. We can then construct our class of candidates as suggested in the last section. Let $\alpha > 0$:

$$\begin{array}{rcl}
\mathrm{d}_1^\alpha M & := & \Box_S^{<\alpha} M \cap \Box_H^{<\alpha} M \\
\mathrm{d}_2^\alpha M & := & \Box_S^{<\alpha} M \cap \Box_H^{\leq\alpha} \Box_S^{<\alpha} M \\
\mathrm{d}_3^\alpha M & := & \Box_S^{\leq\alpha} \Box_H^{<\alpha} M \cap \Box_H^{<\alpha} M \\
\mathrm{d}_4^\alpha M & := & \Box_S^{\leq\alpha} \Box_H^{<\alpha} M \cap \Box_H^{\leq\alpha} \Box_S^{<\alpha} M
\end{array}$$

We define $\mathcal{H}_0(\mathcal{T}) := \mathcal{T}$, $\mathcal{H}_{<\alpha}(\mathcal{T}) := \bigcup_{\beta<\alpha} \mathcal{H}_\beta(\mathcal{T})$. For $\alpha > 0$ we set

$$\mathcal{H}_\alpha(\mathcal{T}) := \mathcal{H}_{<\alpha}(\mathcal{T}) \cup \bigcup_{i=1}^{4} \mathrm{d}_i^\alpha \mathcal{H}_{<\alpha}(\mathcal{T}).$$

Let $\mathcal{H}_\infty(\mathcal{T}) := \bigcup_\alpha \mathcal{H}_\alpha(\mathcal{T})$. For this construction we used the combination $\Box_X^{\leq\alpha} \Box_Y^{<\alpha} M$ instead of $\Box_X^{\leq\alpha}(M \cup \Box_Y^{<\alpha} M)$. This is justified by the following lemma which implies that $\mathcal{H}_{<\alpha}(\mathcal{T}) \subseteq \Box_Y^{<\alpha} \mathcal{H}_{<\alpha}(\mathcal{T})$.

Lemma A.3 *$\mathcal{H}_\alpha(\mathcal{T})$ is transitive for all α.*

Before we prove the lemma, we note the following fact:

Fact A.4 *Let $X, Y \in$ DP, $X \neq Y$. Assume that $w \in \mathcal{H}_\alpha(\mathcal{T}) \setminus \mathcal{T}$, $v \in w(X)$, and $\mathrm{otp}(w) = \alpha$. Then $\mathrm{otp}(v) = \alpha \Rightarrow w \in \Box_X^{\leq\alpha}\Box_Y^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$.*

Proof: If $w \notin \Box_X^{\leq\alpha}\Box_Y^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$, then $w \in \Box_X^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$. But then $\mathrm{otp}(v)$ would have to be smaller than α. ■

Proof of Lemma A.3: We have to show that for all α $w \in \mathcal{H}_\alpha(\mathcal{T})$ implies $w(S) \cup w(H) \subseteq \mathcal{H}_\alpha(\mathcal{T})$. By induction over α: For $\mathcal{H}_0(\mathcal{T}) = \mathcal{T}$ the claim holds by definition. Assume that $w \in \mathcal{H}_\alpha(\mathcal{T}) \setminus \mathcal{H}_{<\alpha}(\mathcal{T})$. Then, suppose $w(S) \not\subseteq \mathcal{H}_\alpha(\mathcal{T})$. Let $v \in w(S) \setminus \mathcal{H}_\alpha(\mathcal{T})$. By introspection $v(S) = w(S)$. As $w(S) \setminus \mathcal{H}_\alpha(\mathcal{T}) \neq \emptyset$, it follows that $w \in \Box_S^{\leq\alpha}\Box_H^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$. Therefore $v \in \Box_S^{\leq\alpha}\Box_H^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T}) \cap \Box_H^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T}) = \mathrm{d}_3^\alpha\mathcal{H}_{<\alpha}(\mathcal{T})$. Therefore, $v \in \mathcal{H}_\alpha(\mathcal{T})$, which contradicts the assumption. The case for $w(H)$ is symmetric. ■

For the proofs of the following lemmas we need one more technical fact:

Fact A.5 *Let $X, Y \in$ DP, $X \neq Y$. Assume that $w \in \mathcal{H}_\alpha(\mathcal{T}) \setminus \mathcal{T}$, $v \in w(X)$, and $\mathrm{otp}(w) = \alpha$. Then $\forall u \in T(v)\ (\mathrm{otp}(u) = \alpha \Rightarrow u \in w(X))$.*

Proof: Let $\mathrm{otp}(u) = \alpha$, then (1) implies $w \in \Box_X^{\leq\alpha}\Box_Y^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$. Therefore, $\forall u \in v(X) = w(X)\ \forall u' \in u(Y)\ \mathrm{otp}(u') < \alpha$. Now it follows by Fact A.2 that $u \in w(X)$. ■

The next lemma shows that our construction provides in step α really all possibilities with complexity α:

Lemma A.6 *$\forall w \in \mathcal{H}_\infty(\mathcal{T})\ \mathrm{otp}(w) = \alpha \Rightarrow w \in \mathcal{H}_\alpha(\mathcal{T})$.*

Proof: Assume that we have shown by induction that $w \in \mathcal{H}_{<\beta}(\mathcal{T})\ \&\ \mathrm{otp}(w) = 0$ implies $w \in \mathcal{T}$. Suppose $w \in \mathcal{H}_\beta(\mathcal{T}) \setminus \mathcal{H}_{<\beta}(\mathcal{T})$, $\mathrm{otp}(w) = 0$. Hence, $w \in \Box_S^{<\beta}\mathcal{H}_{<\beta}(\mathcal{T})$ or $w \in \Box_S^{\leq\beta}\Box_H^{<\beta}\mathcal{H}_{<\beta}(\mathcal{T})$. Therefore, either $\exists v \in w(S)\, v \in \mathcal{H}_{<\beta}(\mathcal{T})$, or $\exists v' \in w(S)\ \exists u \in v'(H)\, u \in \mathcal{H}_{<\beta}(\mathcal{T})$. Due to the transitivity of $\mathcal{H}_{<\beta}(\mathcal{T})$ it follows that $T(v), T(u) \subseteq \mathcal{H}_{<\beta}(\mathcal{T})$, and $\mathrm{otp}(w) = 0$ implies that $w \in T(w) = T(v) = T(u) \subseteq \mathcal{H}_{<\beta}(\mathcal{T})$. Hence, it follows by I.H. that $w \in \mathcal{T}$.

Assume that $\alpha > 0$, and that the proposition holds for all $\beta < \alpha$. Let $w \in \mathcal{H}_\infty(\mathcal{T})$, $\mathrm{otp}(w) = \alpha$. Then, for some $\gamma \geq \alpha$ $w \in \mathcal{H}_\gamma(\mathcal{T})$. Let $\gamma = \alpha + 1$. We first consider the case where $w \in \Box_S^{\leq\alpha+1}\Box_H^{<\alpha+1}\mathcal{H}_{<\alpha+1}(\mathcal{T})$, i.e. $w \in \Box_S^{\leq\alpha}\Box_H^{\leq\alpha}\mathcal{H}_\alpha(\mathcal{T})$. Suppose that $\exists v \in w(S) \exists u \in v(H)\, \mathrm{otp}(u) = \alpha$. Then $w \in T(u)$ by Fact 3.1, and by Fact A.4 it follows that $u \in \Box_H^{\leq\alpha}\Box_S^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$ because $u \in \mathcal{H}_\alpha(\mathcal{T})$, $u \in u(H)$ and $\mathrm{otp}(u) = \alpha$. It follows from Fact A.5 that $w \in u(H)$. Hence, $w \in \Box_S^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$. But then it follows that $\mathrm{otp}(v) < \alpha$, and therefore $\mathrm{otp}(u) < \alpha$, in contradiction to the assumption. Hence, $\forall v \in w(S) \forall u \in v(H)\, \mathrm{otp}(u) < \alpha$. By I.H. it follows that $\forall v \in w(S)\ \forall u \in v(H)\, u \in \mathcal{H}_{<\alpha}(\mathcal{T})$. Together, this implies $w \in \Box_S^{\leq\alpha}\Box_H^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$.

Now we consider the case $w \in \Box_S^{\leq\alpha+1}\mathcal{H}_{<\alpha+1}(\mathcal{T})$. Then $w \in \Box_S^{\leq\alpha}\mathcal{H}_\alpha(\mathcal{T})$. Let $v \in w(S)$, $\mathrm{otp}(v) = \alpha$. Then $v(S) = w(S)$ and $v \in \Box_S^{\leq\alpha}\Box_H^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$. Hence $w \in \Box_S^{\leq\alpha}\Box_H^{\leq\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$.

The case for H is symmetric. Therefore $w \in \mathcal{H}_\alpha(\mathcal{T})$. For $\gamma > \alpha + 1$ the proposition follows easily with the I.H. ■

Now we can prove the central claim, that the hierarchy of the $\mathcal{H}_\alpha(\mathcal{T})$ is identical to $\mathcal{H}(\mathcal{T})$.

Lemma A.7 *For all α: $\mathcal{H}^\alpha(\mathcal{T}) = \mathcal{H}_\alpha(\mathcal{T})$. Moreover, this claim remains valid if we use only d_1^α to d_3^α in the above construction of $\mathcal{H}_\alpha(\mathcal{T})$.*

Proof: We prove the lemma by induction over α. We write $\mathcal{H}^{<\alpha}(\mathcal{T})$ for $\bigcup_{\beta<\alpha} \mathcal{H}^\beta(\mathcal{T})$.

For $\alpha = 0$ the claim follows by definition and Lemma A.1. Assume that $\alpha > 0$, and that $w \in \mathcal{H}^\alpha(\mathcal{T})$. If $\mathrm{otp}(w) < \alpha$, then $w \in \mathcal{H}^\beta(\mathcal{T})$ for some $\beta < \alpha$, and it follows by I.H. that $w \in \mathcal{H}_\beta(\mathcal{T})$. Assume $\mathrm{otp}(w) = \alpha$, hence, $w \notin \mathcal{T}$. Lemma A.1 shows that $w \in \Box_S^{<\alpha}\mathcal{H}^{<\alpha}(\mathcal{T})$, or that $w \in \Box_H^{<\alpha}\mathcal{H}^{<\alpha}(\mathcal{T})$. If $w \notin w(S) \cup w(H)$, then w is an element of the intersection of both classes, and therefore we find by I.H. $w \in \Box_S^{<\alpha}\mathcal{H}_{<\alpha}(\mathcal{T}) \cap \Box_H^{<\alpha}\mathcal{H}_{<\alpha}(\mathcal{T}) = \mathrm{d}_1\mathcal{H}_{<\alpha}(\mathcal{T})$. Assume that $w \in w(S)$. By introspection it holds that for all $v \in w(S)$: $v \in v(S)$, hence, by Fact 3.1 and Lemma A.1 that for all $v \in w(S) \setminus \mathcal{T}$: $v \in \Box_H^{<\alpha}\mathcal{H}^{<\alpha}(\mathcal{T})$. Hence, w must be an element of $\Box_S^{\leq\alpha}\Box_H^{<\alpha}\mathcal{H}^{<\alpha}(\mathcal{T})$, therefore $w \in \Box_S^{\leq\alpha}\Box_H^{<\alpha}\mathcal{H}^{<\alpha}(\mathcal{T}) \cap \Box_H^{<\alpha}\mathcal{H}^{<\alpha}(\mathcal{T})$ and with I.H. it follows that $w \in \mathrm{d}_3^\alpha\mathcal{H}_{<\alpha}(\mathcal{T}) \subseteq \mathcal{H}_\alpha(\mathcal{T})$. This proves that $\mathcal{H}^\alpha(\mathcal{T}) \subseteq \mathcal{H}_\alpha(\mathcal{T})$, and, moreover, that d_4^α was not necessary for the definition of $\mathcal{H}_\alpha(\mathcal{T})$.

Next assume that $w \in \mathcal{H}_\alpha(\mathcal{T})$. If $\mathrm{otp}(w) < \alpha$, the claim follows by I.H. and Lemma A.6. Hence, assume that $\mathrm{otp}(w) = \alpha$, which implies $w \notin \mathcal{T}$. We have to show that $(\mathbf{H_1})$ to $(\mathbf{H_3})$ hold.

Suppose $w \in w(S) \cap w(H)$. From $w \in w(S)$ and $\mathrm{otp}(w) = \alpha$ it

follows by Fact A.4 that w must be an element of $\Box_S^{\leq\alpha}\Box_H^{<\alpha}\mathcal{H}_{<\alpha}(\mathcal{T})$, and therefore in $\mathrm{d}_3^\alpha\mathcal{H}_{<\alpha}(\mathcal{T})$. But then, $\forall v \in w(H)\ \mathrm{otp}(v) < \alpha$, which contradicts $w \in w(H)$.

Suppose next that there is a $v \in T(w)$ such that there exists a $u \in v(S)$ with $w \in T(u)$. It follows by Fact 3.1 that $\mathrm{otp}(v) = \mathrm{otp}(u) = \mathrm{otp}(w) = \alpha$. But then, it follows by Fact A.5 that $w \in v(S)$.

Finally, $w \in \mathcal{H}(\mathcal{T}) \Rightarrow w(S), w(H) \subseteq \mathcal{H}(\mathcal{T})$ follows by transitivity of $\mathcal{H}(\mathcal{T})$, Lemma A.3.

As $\mathcal{H}(\mathcal{T})$ is the largest subclass of $\mathcal{F}$ where $(\mathbf{H_1})$ to $(\mathbf{H_3})$ hold, it follows that $\mathcal{H}_\alpha(\mathcal{T}) \subseteq \mathcal{H}^\alpha(\mathcal{T})$.

■