Obligationes as Formal Dialogue Systems[1]

Sara L. Uckelman[2]

Abstract. Formal Dialogue Systems (FDSs) model rule-based interaction between agents and as such have multiple applications in multi-agent systems and AI more generally. Their conceptual roots are in formal theories of natural argumentation, of which Hamblin's formal systems of argumentation in [15, 16] are some of the earliest examples. Hamblin cites the medieval theory of *obligationes* as inspiration for his development of formal argumentation. In an *obligatio*, two agents, the Opponent and the Respondent, engage in an alternating-move dialogue, where the Respondent's actions are governed by certain rules, and the goal of the dialogue is establishing the consistency of a proposition. We implement *obligationes* in the formal dialogue system framework of [31] using Dynamic Epistemic Logic [38]. The result is a new type of inter-agent dialogue, for *consistency-checking*, and analyzing *obligationes* in this way also sheds light on interpretational and historical questions concerning their use and purpose in medieval academia.

1 Introduction

Rule-based interactions such as dialogues or arguments are ubiquitous and diverse; they are the basic method of communication between agents. Such interactions are modeled in artificial intelligence and computer science by formal dialogue systems (FDSs) and dialogue games [18, 22, 23, 24, 31]. These systems give formal, and hence potentially implementable, methods for modeling real-life dialogue situations, such as complex reasoning in legal domains. More generally, dialogue systems and games are used in multi-agent systems to model distributed cognition and interaction between intelligent agents, and they can also be used in the specification of complex software systems and programs [24].

The conceptual roots of formal dialogues, however, come not from within AI but from without. They are located in the sphere of natural argumentation, that is, philosophical logic, argumentation theory, and rhetoric. One of the earliest attempts to provide a theory of formal dialogues is Hamblin's [16, 15]. In [16], Hamblin locates part of the motivation for his development of formal argumentation in historical formal dialogue systems, that is, dialogical or disputational settings where explicit rules are given governing the actions of the participants. One such system of natural argumentation that he considers in particular is the medieval theory of *obligationes*, developed in the 13th and 14th centuries. In an *obligatio*, two agents, the Opponent and the Respondent, engage in an alternate-move dialogue, where the Respondent's actions are governed by certain rules, and the goal of the dialogue is, in the most basic case, to establish the consistency of a proposition. We argue that *obligationes* are best modeled by FDSs because of their intrinsic dialogical nature. Furthermore, taking *obligationes* from the realm of natural argumentation to the realm of formal dialogue systems results in the determination of a new type of dialogue system different from those generally discussed in the AI literature, and thus they make a novel contribution to formal modeling of dialogue and interaction in multi-agent systems.

The plan of the paper is as follows. In §2 we present the medieval theory of *obligationes*, focusing specifically on the works of one author, Walter Burley, and give examples. In §3 we briefly survey previous work on *obligationes*, both formal and philosophical, and motivate modeling *obligationes* as FDSs by showing how they can make sense of King's interpretation of *obligationes* as a meta-disputational framework. In §4 we introduce formal dialogue systems, following the presentation in [31], and show generically how *obligationes* can be viewed as a formal dialogue system. Before we give a precise specification of *obligationes* as FDSs in §6, we first outline the logic used in the argumentation, a type of Dynamic Epistemic Logic, in §5. We discuss the properties of the protocol we introduce, compare our results with standard types of FDSs, and define a new type of inter-agent dialogue, for *consistency-* or *feasibility-checking*, in §7. In §8 we conclude and point towards future work.

2 Medieval theories of *obligationes*

An *obligatio* is a dialogue between two agents, the Opponent and the Respondent, where the Opponent puts forward a sequence of propositions, and the Respondent is obligated (hence the name) to follow certain rules in his responses to the Opponent's propositions. More precisely, the Opponent puts forward an initial statement, called the *positum*, which the Respondent can either accept or refuse to accept. If he accepts, the *obligatio* begins. If he does not, no *obligatio* begins. If the *obligatio* begins, the Opponent puts forward propositions and the Respondent has three ways that he can respond: He can grant or concede the proposition, he can deny the proposition, or he can doubt it, where 'doubt' should be understood as 'remain agnostic about'; doubting φ does not entail any commitment to $\neg\varphi$. (Some authors, such as William of Ockham [26] and the anonymous author of the *Obligationes Parisienses* [11], mention a fourth option, which is to 'draw distinctions', that is, to clarify an ambiguity on the part of the Opponent.) The *obligatio* continues until the Opponent calls "*Cedat tempus*" ("Time's up"), whereupon the responses of the Respondent are analysed with respect to the Respondent's obligations, to determine whether he has responded well or badly.

The earliest texts on *obligationes* date from the beginning of the 13th century [10, 11, 12], and many of the leading logicians from that century and the next wrote treatises on the subject. While the roots of obligational disputations are clearly grounded in Aristotle's discussion of dialectical exchanges in the *Topics* VIII, 4 (159a15–24) and in the *Prior Analytics* I, 13 (32a18–20) (cf. [41, §II.A]), the systematic development of the theory of *obligationes* over the course of the 13th

[1] A short version of this paper will appear in the proceedings of STAIRS'10.
[2] Institute for Logic, Language, and Computation, University of Amsterdam, The Netherlands, email: S.L.Uckelman@uva.nl

and 14th centuries tends to show little adherence to the Aristotelian definitions. While the specific details vary from author to author, a number of distinct types of *obligationes* discussed by multiple authors can be identified. The six most common are *positio*, *depositio*, *dubitatio*, *sit verum* or *rei veritatis*, *institutio*, and *petitio*. Of these six, *positio* is universally the most widely studied, both by medieval and modern authors; as a result, it is the focus of the current paper. For further information on *obligationes*, including a discussion of their purpose and their role in medieval philosophy, see [41].

To make the above more precise, we look at the theory of *obligationes* of a specific writer, Walter Burley. Burley's treatise *De obligationibus*, written around 1302, gives a standard treatment of *positio*. The text of this treatise is edited in [7] and a partial translation of the text, including the section on *positio* in its entirety, is found in [6]. (In the interests of conserving space, we quote all medieval texts in translation only, rather than in the original Latin.) Burley defines the general goal of an *obligatio* as follows:

> The opponent's job is to use language in a way that makes the respondent grant impossible things that he need not grant because of the *positum*. The respondent's job, on the other hand, is to maintain the *positum* in such a way that any impossibility seems to follow not because of him but rather because of the *positum* [6, p. 370].

Thus, it is clear that in an *obligatio*, the goal is consistency, not logical truth or validity. In *positio*, the primary obligation of the Respondent is to grant, that is, to hold as true, the *positum*. If the Respondent accepts the *positum* and the *obligatio* begins, he is obliged to follow the following rules:

Rule 1 Everything that is posited and put forward in the form of the positum during the time of the positio must be granted [6, p. 379].

Rule 2 Everything that follows from the positum must be granted. Everything that follows from the positum either together with an already granted proposition (or propositions), or together with the opposite of a proposition (or the opposites of propositions) already correctly denied and known to be such, must be granted [6, p. 381].

Rule 3 Everything incompatible with the positum must be denied. Likewise, everything incompatible with the positum together with an already granted proposition (or propositions), or together with the opposite of a proposition (or the opposites of propositions) already correctly denied and known to be such, must be denied [6, p. 381].

In Rule 1, 'in the same form as' should be understood syntactically; if the *positum* is 'Marcus is Roman', then the Respondent doesn't have an obligation to accept 'Tullius is Roman' unless it is explicit (either through common knowledge or through previous concessions) that Marcus is Tullius.

Burley also defines a notion of relevance of propositions which applies to all types of *obligatio*. A proposition is *irrelevant* or *impertinent* if neither it nor its negation follows from the set of propositions which have already been conceded (which includes the negations of propositions which have been denied).

Rule for Irrelevant Propositions One must reply to what is irrelevant in accordance with its own quality [6, p. 375].

I.e., the Respondent should reply by conceding the proposition if it is true, by denying if it is false, and by doubting if he does not know which is the case. The notion of 'relevance' in *obligationes* parallels the concept of 'support' in formal argumentation theory where an argument is defined as a pair $A = (S, p)$ where S is a consistent subset of the knowledge base such that $S \vDash p$ (cf., e.g., [9, p. 155]).

	Opponent	Respondent	
1	φ.	I admit it.	$\Phi_0 := \{\varphi\}$.
2	$\neg\varphi \vee \psi$.	I grant it.	$\Phi_1 := \{\varphi, \neg\varphi \vee \psi\}$.
3	ψ	I grant it.	

Figure 1. An example *obligatio*.

	Opponent	Respondent
1	φ or φ must be granted.	I admit it.
2	φ must be granted.	I deny it.
3	φ follows from the positum and the opposite of something correctly denied	I grant it.
4	φ must be granted.	??

Figure 2. A more interesting example.

A simple example illustrating Burley's rules for *positio* is given in Figure 1. Suppose φ does not imply $\neg\psi$ and φ is known to be contingently false. In the first round, the Opponent puts forward a contingent (but false) proposition; the Respondent grants it in accord with Rule 1. In the second round, either φ implies ψ, then the sentence is relevant and follows from Φ_0 (the set of propositions conceded so far along with the negations of propositions denied to this point); or it doesn't, in which case it is irrelevant and true (since φ is false). In both cases, the Respondent is required to concede; the first case falls under Rule 2, and the second under the Rule for Irrelevant Propositions. In the third round, the Respondent likewise must concede because ψ follows from Φ_1. This example *obligatio* shows how, given a *positum* which is false but not contradictory, the Opponent can force the Respondent to concede any other consistent proposition.

More interesting examples, such as the example in Figure 2, involve statements about the obligational rules themselves. Let φ be the proposition 'you are in Rome' (spoken by the Opponent to the Respondent). The *positum* is a disjunction between a simple proposition and the assertion that that proposition must be granted. Because the disjunction is not a logical contradiction (in particular the first disjunct is possible, though it is in fact false), the Respondent is correct in accepting the *positum*. The second disjunct is irrelevant, as it is not a logical consequence of the *positum*, and furthermore it is false: Since φ is false, and φ is also irrelevant, the Respondent is not under any obligation to accept φ. Thus it is false that φ must be granted, so he correctly denied the second proposition. The third proposition expresses a logical necessity, about the validity of disjunctive syllogism, and so is accepted. But now it is unclear how the Respondent should respond to the re-assertion that φ must be granted. On the one hand, this proposition has been put forward before, and was denied, and so it should continue to be denied. On the other hand, once the third proposition has been granted, by Rule 2, φ must be granted. So superficially it appears that the Respondent is obliged to both accept and deny this final statement. Burley's resolution to the problem is to argue that (3) is not only not necessary, but it is repugnant, since it is inconsistent with the opposite of (2). Since it is repugnant, the Respondent should have in fact denied, and thus (4) can also be denied without contradiction [41, pp. 152–155].

3 Previous work on *obligationes*

Green's Ph.D. dissertation [7], containing an edition of and commentary on two treatises on *obligationes*, now generally ascribed to William of Sherwood and Walter Burley, marks the beginning of modern research on *obligationes*. Since then, many philosophers and historians have devoted themselves to the question of the goal or purpose of obligational disputations and the role they played in medieval academic life, while somewhat fewer have focused on the logical properties of *obligationes*. Despite this, the purpose of *obligationes* and their role in medieval academic life remains stubbornly unclear [34]. Two recent sources which discuss the various views, which range from the view that *obligationes* were mere academic exercises, that they were used for counter-factual reasoning, that they were a tool for evaluating *sophismata* and *insolubilia* (paradoxical and problematic sentences), that they are thought experiments, and that they provide a theory of belief revision, are [41] and [42].

Hamblin is the first modern author to attempt to formalize *obligationes* [16, pp. 260–263]. Given his interest in formalizing argumentation generally, he focuses on the dialogical aspects of *obligationes*. His formalization is rudimentary and models only one variant, that given by William of Sherwood[3], but it marks the beginning of modern scholarship on the formal properties of *obligationes*. Recent scholarship has focused on the game-like nature of *obligationes*, e.g., [11, 13, 41]. In particular, there is an immediate apparent similarity between *obligationes* and Lorenzen's dialogical logic [20]. It may therefore seem natural to look to game-based structures in logic to provide a general framework for modeling different types of *obligationes*. However, there are a number of aspects which do not immediately lend themselves nicely to a game-like interpretation (e.g., the notion of a winning strategy for an *obligatio*, for example, is difficult to define[4]), and despite the strongly logical component of *obligationes*, to date relatively little work has been done on the formal properties of the logic involved and few attempts have been made to provide an explicit specification of the game(s) involved.

The most extensive attempt is in [13]. In this book, Dutilh Novaes analyses the obligational theories of three authors, Walter Burley, Richard Swyneshed (c.1330), and Ralph Strode (second half of the 14th C), giving separate formalizations for each one. Her formalizations are based on models $\mathfrak{M} = \langle K_c, \Phi, \Gamma, R(\varphi)\rangle$ where K_c is the set of common knowledge among the participants of the disputation (expressed as a set of propositions); Φ is a sequence of propositions, which keeps track of the assertions of the Opponent; Γ is a sequence of propositions, which keeps track of the responses of the Respondent; and $R(\varphi)$ is a function from φ to 1 (standing for 'concede'), 0 ('deny'), and ? ('doubt'). These formalizations are not very satisfying for a number of reasons. Each obligational theory studied is provided with a different framework, which means that it is difficult to make cross-theory comparisons. Further, only *positio* is studied; the other types of *obligationes* are not discussed. Additionally, the frameworks all presuppose a significant amount of background information which is taken for granted and never specified: the semantic model(s) in which truth of propositions (particularly the *positum* and irrelevant propositions) and the Respondent's knowledge of both individual propositions as well as how the consequence relations are to be evaluated, and the syntactic rules governing $\vdash$ (which is used to generate Γ). For example, the set of common knowledge K_C is not defined in any explicit fashion, and there is nothing which grounds the knowledge of the participants. Finally, since the nature of the proof system being used in the definition of $R(\varphi)$ is never specified, her frameworks are essentially incomplete; it is impossible to implement the logical model without making the proof-system explicit [13, p. 169].

When looking for an alternative to a game-theoretic approach to modeling *obligationes*, one fruitful suggestion that presents itself can be rooted in a particularly interesting interpretation of *obligationes* given by King [19]. King takes his starting point from Spade, who, in [35], looked to the textual evidence for actual uses of *obligationes* to understand how they were used by the medievals. While to date there is no historical record for actual obligational disputations, we have many examples of philosophers using obligational techniques as part of their argumentation [19, p. 1]. King explains the apparent "content-freeness" of obligational disputations by pointing out that "they operate at a higher level of logical generality than that at which substantive debate occurs. If this is correct, then actual obligational moves—perhaps even recognized as such—are the vehicle whereby real argument takes place" [19, p. 6], and thus *obligationes* provide a "meta-methodology" for reasoning [19, p. 7].

We use this suggestion as the motivation for our approach to modeling *obligationes*. An *obligatio* is essentially a dialogue; and any dialogue can be seen as a game played according to the rules specified by an FDS [22]. We believe that viewing *obligationes* as FDSs, which require that we explicitly specify the logic of argumentation/inference and the models against which the dialogue is to be evaluated, provide a more fruitful approach to modeling *obligationes*. On this view, Hamblin's modeling approach has the advantage over others proposed in recent literature because it takes the dialogical nature of the disputation seriously. By varying the rules governing the disputation, radically different types of *obligationes* arise, which result in radically different types of dialogues/disputations. Despite the wide range of difference that can be found, the basic structure of an *obligatio* remains the same, making the general framework of FDSs an appropriate modeling choice. Specifying *obligationes* from within the context of FDSs allows us to situate them formally in current research on formal dialogues, which in turn can help to clarify the interpretational question, by helping us understand the possible purposes to which *obligationes* could be disposed. In particular, we argue that the naturalness of modeling *obligationes* as dialogue systems supports King's suggestion that *obligationes* provide agents with a meta-methodology for argumentation. That is, *obligationes* give frameworks within which dialectical argumentation—dialogue—can take place.

4 Formal dialogue systems

The standard taxonomy of formal dialogue systems is based on the argumentation-based typology given by Walton and Krabbe [39], who identify six different basic types of dialogues: information seeking, inquiry, persuasion, negotiation, deliberation, and eristic. The division is made on the basis of the preconditions and postconditions that must hold for successful dialogue to take place [9, 24]. Three of these classes, information seeking, inquiry, and persuasion, deal with beliefs and knowledge, and thus are of particular interest. Walton and Krabbe make no claim as to the comprehensiveness of their classi-

[3] Hamblin routinely questions the attribution to Sherwood of the text he is considering; however, more recent scholarship is agreed that the text was almost certainly written by Sherwood, sometime in the middle of the 13th century [4].

[4] Yrjönsuuri mentions the possibility of modeling *obligationes* as games, but he says that "defining the results of the game in any manner appropriate to modern game-theory seem utterly problematic" though despite this "[i]n the following I will keep to the English word *game*, assuming that the problems pointed out above can just be left unsolved" [41, pp. 9–10].

fication, and others [2, 8, 9, 14, 32, 33] have researched types not covered in the Walton and Krabbe typology. In particular, Cogan et al. argue that "there remain several situations in which it seems natural to engage in dialogues, but to which the basic Walton and Krabbe dialogue types do not apply" [9, p. 161]. In [8, 9], they take a systematic approach to defining dialogue types on the basis of pre- and post-conditions, and extend Walton and Krabbe's list of belief-based dialogues with four new types, verification and three types of queries. In §7, we discuss how *obligationes* fit into these different types. First, we make precise what we mean by a formal dialogue system.

In this section, we follow the presentation of formal dialogue systems given in [31], an overview paper which discusses different formal argumentation systems that have been proposed for the analysis of persuasion dialogues and provides a unified approach within which each of these different systems can be modeled. While Prakken focuses on persuasion dialogues, his framework is in fact general enough to handle other types as well [31, pp. 170, 173]. Thus, it is appropriate to use it to consider *obligationes*.

The specification of a *formal dialogue system* contains the following elements [31, p. 166]:

- A *topic language* $\mathcal{L}_t$, closed under classical negation.
- A *communication language* $\mathcal{L}_c$. We denote the set of *dialogues*, that is, the set sequences of $\mathcal{L}_c$, by $M^{\leq\infty}$, and the set of finite sequences of $\mathcal{L}_c$ by $M^{<\infty}$. For a dialogue $d = m_0, \dots, m_n, \dots$, the subsequence $m_0, \dots, m_i$ is denoted d_i.
- A *dialogue purpose* or *goal*.
- A set $\mathcal{A}$ of *agents* (participants) and a set $\mathcal{R}$ of *roles* that the participants can occupy. Each participant a has a (possibly empty) *belief base* $\Sigma_a \subseteq \mathcal{L}_t$ and a (possibly empty) *commitment set* $\mathrm{C}_a(d_n) \subseteq \mathcal{L}_t$. The belief base may or may not change during the dialogue; the commitment set usually does.
- A *context* $K \subseteq \mathcal{L}_t$, representing the (shared, consistent, and unchanging) knowledge of the agents specified at the outset.
- A *logic* L for $\mathcal{L}_t$.
- A set E of *effect rules* $\mathrm{C}_a(d_n) : M^{<\infty} \to \mathcal{P}(\mathcal{L}_t)$ for $\mathcal{L}_c$, specifying how utterances $\varphi \in \mathcal{L}_c$ in the dialogue affect the commitment stores of the agents. The effect rules are such that if $d = d'$ then $\mathrm{C}_a(d, m) = \mathrm{C}_a(d', m)$, that is, the changes in commitments are determined solely by the most recent move in the dialogue along with the commitments at that step.
- A *protocol* P for $\mathcal{L}_c$, specifying the legal moves of the dialogue, which is a function from the context and a non-empty $D \subseteq M^{<\infty}$ to $\mathcal{P}(\mathcal{L}_c)$, satisfying the requirement that if $d \in D$ and $m \in \mathrm{P}(d)$, then $d, m \in D$. The elements of D are called *legal finite dialogues*, and $\mathrm{P}(d)$ is the set of moves allowed after move d. At any stage, if $\mathrm{P}(d) = \emptyset$, then the dialogue has *terminated*. A protocol will often be accompanied by a *turn-taking* function $T : D \to \mathcal{P}(\mathcal{A})$, which takes a finite dialogue d_n and specifies who governs move m_{n+1}, and *termination* conditions, which specify when $\mathrm{P}(d) = \emptyset$.
- A set of *outcome rules* O.

Dialogue systems can be explicitly connected with the games they specify by defining the agents' strategies in the standard game-theoretical way. Formally, a *strategy* s_a for agent a is a function $D_a \to \mathcal{L}_c$, where $D_a \subseteq D$ is the set of all finite legal dialogues d_n in which $T(d_n) = a$. A strategy s_a is called *winning* if in every dialogue where a follows this strategy, he realizes his dialogue goal.

We can identify a number of properties of protocols [31, p. 170]:

- A protocol has *public semantics* iff the set of legal moves is always independent from the agents' belief bases.
- A protocol is *context-independent* iff the set of legal moves and the outcome is always independent of the context, that is, $\mathrm{P}(K, d) = P(\emptyset, d)$.
- A protocol is *fully deterministic* iff P always returns a singleton or the empty set.
- A protocol is *unique-move* iff the turn shifts after each move; it is *multiple-move* otherwise.

Protocols which are not fully deterministic are *permissive*, that is, they specify what moves are *legal* or *allowed* for the agent, rather than specify what moves are *required*. Thus, *obligationes* are a type of DS where the protocol for the Respondent is fully deterministic; for each proposition the Opponent puts forward, there will be exactly one correct move that the Respondent can make.

We now show how generically *obligationes* can be viewed as DSs; we give precise examples in §6. In *obligationes*, there are two designated roles Opp (Opponent) and Res (Respondent) that members of $\mathcal{A}$ can have; those members of $\mathcal{A}$ which do not fill either role are irrelevant for modeling the disputation. We explain below how Σ_{Opp}, Σ_{Res}, $\mathrm{C}_{\mathsf{Opp}}$, $\mathrm{C}_{\mathsf{Res}}$, and the context K are generated. In Burley-style *positio*, the dialogue purpose is consistency: If we take Res's commitment set to be the set of formulas he has conceded along with the negation of those that he's denied over the course of a *positio*, then the goal for Res is to maintain the consistency of his commitment set, and the goal for Opp is to force Res into contradiction.

In general, the topic language $\mathcal{L}_t$ and the communication language $\mathcal{L}_c$ are the same. This allows, among other things, the participants in an *obligatio* to dispute about the allowed moves of the other players. (For example, Opp may ask Res to respond to the claim "You deny φ".) The turn-taking protocol in an *obligatio* is unique-move: $T(\emptyset) = \mathsf{Opp}$, $T(d_n) = \mathsf{Opp}$ if n is odd, and $T(d_n) = \mathsf{Res}$ if n is even. (Throughout we assume that we label the steps in the sequence from 0, so in an *obligatio* it is always Opp that goes first.) The protocol P will be such that the moves of Opp are not constrained in any way, but Res's moves must be made in reaction to the move of Opp at the previous stage. The same will be true for the effect rules E; in a disputation, Opp makes a series of claims or assertions, but these actions have no effect on his commitment store. On the other hand, Res is constrained to be reactive only: He can only concede statements claimed by Opp, concede their negations, or remain ambivalent. Res never asserts any statement of his own devising, he only ever responds to propositions put forward by Opp. Thus, *obligationes* are essentially asymmetric, in that the rules governing the behavior of the Opp and Res are disjoint[5], and so are their actions.

The outcome rules for *obligationes* are simple: If Res realizes the goal, then he wins. If Opp realizes the goal, then he wins. There is nothing further that hinges upon winning or losing an obligational disputation (except, of course, the individual prestige or embarrassment of the participants!).

Above we noted that in an arbitrary DS, the commitment set of an agent will generally change during the course of the dialogue. It can either strictly grow, so that the agents are only adding new propositions to their commitment-base at each turn, or they can also revise their commitments by rejecting previous commitments in favor of new ones. This latter case arises in ordinary circumstances when agents utilize a form of default reasoning, which is defeasible and non-monotonic, in that an agent can be forced to accept information which contradicts his previous commitments, requiring that

[5] In fact, in most texts, no rules for Opp are given. One exception is the early text *Tractatus Emmeranus* [10], which gives some rules (better thought of as guidelines, or strategic advice) to the Opponent.

his commitments be revised in order to maintain consistency (cf. [3, 5]). In AI contexts, the ability to simulate non-monotonic reasoning is of great importance; monotonic dialogues and discussions are more commonly found in philosophical contexts. One of the benefits of Prakken's approach to DS is that it can handle both approaches, merely by the specification of the underlying logic [31, p. 173].

5 The underlying logic

By specifying the logic L and its underlying models, we are able to explicitly generate Σ_{Opp}, Σ_{Res}, C_{Opp}, C_{Res}, and K satisfying desired properties. In our approach to modeling *obligationes* as DSs, the underlying logic is multi-agent Dynamic Epistemic Logic (DEL, [38]). This logic is monotonic and not argument based. An epistemic logic is an extension of propositional logic with a family of modal operators K_a for $a \in \mathcal{A}$. We are interested in a particular extension of standard epistemic logic, namely, *epistemic logic with common knowledge*, which has a further family of operators C_G, for $G \subseteq \mathcal{A}$. For a set Φ_0 of propositional letters and set $\mathcal{A}$ of agents, the set $\Phi^{\mathcal{A}}_{\mathrm{EL}}$ of wffs of EL is defined by:

$$\varphi \quad := \quad p \in \Phi_0 \mid \neg\varphi \mid \varphi \vee \varphi \mid K_a\varphi : a \in \mathcal{A} \mid C_G\varphi : G \subseteq \mathcal{A}$$

$K_a\varphi$ is read 'agent a knows that φ'. $C_G\varphi$ is read 'it is common knowledge amongst the group of agents G that φ'. We will use C_G to represent explicitly the knowledge of the two agents at the beginning of the disputation.

Epistemic logic is interpreted on Kripke frames. A structure $\mathfrak{M} = \langle W, w^*, \{\sim_a : a \in \mathcal{A}\}, V\rangle$ is an *epistemic model* if

- W is a set (of possible worlds), with $w^* \in W$ a designated point (representing the actual world).
- $\{\sim_a : a \in \mathcal{A}\}$ is a family of equivalence relations on W, one for each member of $\mathcal{A}$. The relation $w \sim_a w'$ is interpreted as 'w and w' are epistemically equivalent for agent a'. $\sim_G : G \subseteq \mathcal{A}$ is defined as the reflexive and transitive closure of $\bigcup_{a\in G}\{\sim_a\}$.
- $V : \Phi_0 \to 2^W$ is a valuation function associating atomic propositions with subsets of W. For $p \in \Phi_0$, if $w \in V(p)$, we say that 'p is true at w'.

The semantics for the propositional connectives are as expected. We give just the semantics for the epistemic operators.

$$\begin{array}{lll} \mathfrak{M}, w \vDash K_a\varphi & \text{iff} & \forall w'(\text{if } \langle w, w'\rangle \in \sim_a \text{ then } \mathfrak{M}^E, w' \vDash \varphi) \\ \mathfrak{M}, w \vDash C_G\varphi & \text{iff} & \forall w'(\text{if } \langle w, w'\rangle \in \sim_G \text{ then } \mathfrak{M}^E, w' \vDash \varphi) \end{array}$$

EL models cover the knowledge of the agents; to model their actions, we add dynamics, via Propositional Dynamic Logic (PDL, [17]). PDL is an extension of propositional logic by a family of modal operators $[\alpha]$ for $\alpha \in \Pi$, a set of programmes (or more generally, a set of actions or events). The language of PDL is two-sorted, with a set Φ_0 of atoms and a set Π_0 of atomic actions. We do not need the full expressivity of PDL to model *obligationes*, so we introduce only the fragment we require. We let $\Pi_0 = \emptyset$, and the sets Φ_{Ob} and Π_{Ob} of complex well-formed formulas and programmes are defined by mutual induction:

$$\begin{array}{lll} \varphi & := & \varphi \in \Phi^{\mathcal{A}}_{\mathrm{EL}} \mid [\alpha]\varphi : \alpha \in \Pi_{\mathrm{Ob}} \\ \alpha & := & \varphi? : \varphi \in \Phi^{\mathcal{A}}_{\mathrm{EL}} \end{array}$$

The programme $\varphi?$ is to be interpreted as a test operator, which tests for the truth of φ. Note that the only programmes that we allow are testing of formulas which do not themselves contain any programmes. The semantics for the new $[\varphi?]$ operator are given in terms of model reduction. Let $\mathfrak{M} \restriction \varphi := \langle W^{\mathfrak{M},\varphi}, \{\sim^{\mathfrak{M},\varphi}_a : a \in \mathcal{A}\}, V^{\mathfrak{M},\varphi}\rangle$, where $W^{\mathfrak{M},\varphi} := \{w \in W : \mathfrak{M}, w \vDash \varphi\}$, and the relations and valuation functions are just restrictions of the originals. For a set of ordered propositions Γ_n, let $\mathfrak{M} \restriction \Gamma_n = \mathfrak{M} \restriction \gamma_0 \restriction \cdots \restriction \gamma_n$, that is, $\mathfrak{M} \restriction \Gamma_n$ is the result of the sequential restriction of $\mathfrak{M}$ by the elements of Γ_n. Then:

$$\mathfrak{M}, w \vDash [\varphi?]\psi \quad \text{iff} \quad \forall v \in \mathfrak{M} \restriction \varphi, v \vDash \psi$$

One advantage of using an epistemic logic for our disputation logic is that it allows us to model the epistemic bases of the agents, and the context of the disputation, explicitly (for a fuller discussion of the advantages, which includes the ability to use this framework to model different types of *obligationes* beyond just the one considered here, see [36]). While above we defined the concept of a "belief base" in a dialogue, in the context of *obligationes* it is the agent's knowledge, not their beliefs that is important. Given an epistemic model $\mathfrak{M}$, the knowledge bases of Opp and Res are defined as follows:

$$\begin{array}{lll} \Sigma^{\mathfrak{M}}_{\mathsf{Opp}} & := & \{\varphi : \mathfrak{M}, w^* \vDash K_{\mathsf{Opp}}\varphi\} \\ \Sigma^{\mathfrak{M}}_{\mathsf{Res}} & := & \{\varphi : \mathfrak{M}, w^* \vDash K_{\mathsf{Res}}\varphi\} \end{array}$$

In an arbitrary model $\mathfrak{M}$, the set of propositions which are common knowledge amongst a group of agents is not explicitly specified. In an *obligatio*, the set of common knowledge, against which the truth of irrelevant propositions is evaluated, is likewise often left implicit. In some cases, before the *obligatio* begins, a *casus* is introduced. A *casus* is a hypothesis about how the world is, or extra information about how the *positum* should be analyzed [40]. A common example of a *positum* introduced with a *casus* (understood in the first sense) is "In truth Socrates is black. It is posited that Socrates is white." The first sentence is the *casus*; it tells the participants not only that Socrates is black, but also that he exists, and is colored, all of which facts the Respondent must take into account when responding to the Opponent's *posita*. Thus, in the first sense, the *casus* can be understood as a set of literals expressing the *explicit common knowledge* at the start of the dialogue, so the *casus* can be implemented by a restriction on V.

Definition 5.1. (*Casus*). Let Lit_{Φ_0} be the set of literals formed from Φ_0, and $K \subseteq \mathrm{Lit}_{\Phi_0}$ be the *casus*. Then $\mathfrak{M}$ *models the casus* if there is a $P_c \subseteq P$ of W with $w^* \in P_c$, such that if $w \sim_{\mathsf{Res}} w^*$, then $w \in P_c$, if $v \sim_{\mathsf{Opp}} w^*$, then $v \in P_c$, and for all $w, v \in P_c$, $w \sim_{\mathsf{Res}} v$ and $w \sim_{\mathsf{Opp}} v$; and for every positive literal $p \in K$ and every $w \in P_c$, $w \in V(p)$, and for every negative literal $\neg q \in K$ and every $w \in P_c$, $w \notin V(q)$.

Unlike contexts in dialogue systems, it is not assumed that the *casus* of an *obligatio* is consistent, but if it is not, then Res should not accept the *positum*, since Opp could easily force him into conceding a contradiction. However, if the *casus* is consistent, we can easily show that if $\mathfrak{M}$ models a *casus* K, then for every $\varphi \in K$, $\mathfrak{M} \vDash C_{\{\mathsf{Opp},\mathsf{Res}\}}\varphi$, and so $K \subseteq \Sigma^{\mathfrak{M}}_{\mathsf{Opp}}$ and $K \subseteq \Sigma^{\mathfrak{M}}_{\mathsf{Res}}$.

6 Protocols, effect rules, and outcomes

Different types of *obligationes* can be modeled by changing the protocols, effect rules, and outcome conditions. First, we specify the general properties shared by all *obligationes*. We identify our set of agents with their roles, i.e., our set of agents is $\mathcal{A} = \{\mathsf{Opp}, \mathsf{Res}\}$, and our topic language and commitment language is the language of dynamic epistemic logic $\mathcal{L}_{\mathrm{DEL}}$ introduced in the previous section. Let

α be a designated formula representing "*cedat tempus*". We can identify two types of protocols used in *obligationes*. The first type of protocol is uniform throughout all different systems; the second varies from author to author and type to type. The uniform protocol P_u is invariant over all contexts and is defined for a finite dialogue d_n:

$$\begin{array}{ll} & \mathrm{P}_u(\emptyset) = \mathcal{L}_c \\ \text{if } m_n = \alpha & \mathrm{P}_u(d_n) = \emptyset \\ \text{otherwise, if } n \text{ is odd,} & \mathrm{P}_u(d_n) = \mathcal{L}_c \\ \text{and if } n \text{ is even,} & \mathrm{P}_u(d_n) = \{[m_n?]\top, [\neg m_n?]\top, [\top?]\top\} \end{array}$$

That is, if it is Opp's turn, he is allowed to assert any statement in the communication language (we allow repetitions). If it is Res's turn, he must either concede, deny, or doubt Opp's statement from the previous round. And if "*cedat tempus*" has been called, the dialogue ends and there are no more legal moves available. Since m_n, the move of Opp, will always be a statement in the communication language $\mathcal{L}_c$, and the communication language allows for the embeddings of the test programme, this protocol is well-defined. This protocol has public semantics and is context-independent, but it is not fully deterministic, since whenever it is Res's turn, he has a choice of actions. For ease of future reference, we introduce meta-names for the actions of Res: **concede**:$\varphi := [\varphi_n?]\top$, **deny**:$\varphi := [\neg\varphi_n?]\top$, and **doubt**:$\varphi := [\top?]\top$. The last clause is equivalent to saying "I don't know"; $[\top?]\top$ will always be valid, in any model.

The rules governing the commitment sets $\mathrm{C}_{\mathsf{Opp}}$ and $\mathrm{C}_{\mathsf{Res}}$ are defined as follows:

$$\begin{array}{ll} \text{for all } n & \mathrm{C}_{\mathsf{Opp}}(d_n) = \emptyset \\ \text{if } n \text{ is even} & \mathrm{C}_{\mathsf{Res}}(d_n) = \mathrm{C}_{\mathsf{Res}}(d_{n-1}) \\ \text{if } n \text{ is odd} & \mathrm{C}_{\mathsf{Res}}(d_n) = \mathrm{C}_{\mathsf{Res}}(d_{n-1}) \cup \{m_n\} \end{array}$$

That is, Opp has no commitments, Opp's moves do not change Res's commitments, and Res's commitment store strictly grows on the basis of his actions, and thus obligational dialogues are monotonic. As above, since $\mathcal{L}_c$ and $\mathcal{L}_t$ coincide, the final clause of the definition is well-defined. Note that in general, $\mathrm{C}_{\mathsf{Res}}$ and Σ_{Res} will be disjoint, and similarly for $\mathrm{C}_{\mathsf{Res}}$ and K (contra, e.g., [27, §3], where "an agent's commitment store is just a subset of its knowledge base").

The general protocol defined above specifies what the possible moves of Res are. In an *obligatio*, however, we want to say more than what moves are *allowed*, we also want to specify a set of possible moves which are in fact *required*, since in an obligational disputation Res is under obligation to respond to Opp in certain ways. This is done by specifying a more refined protocol. Such a protocol, because it makes reference to the agents' knowledge bases, will always be defined with respect to a particular DEL model $\mathfrak{M}$. We give as an example Burley's protocol $\mathrm{P}^{\mathsf{Bur}}$ for *positio*, introduced in §2. Let Γ_n be the sequence of Res's move in a dialogue d_n. For a DEL model $\mathfrak{M}$ and context K, $\mathrm{P}^{\mathsf{Bur}}(K,\emptyset) = P_u(\emptyset)$ and if n is odd, $\mathrm{P}^{\mathsf{Bur}}(K,d_n) = P_u(d_n)$. For n even,

- For $d_0 = m_0 =$ the *positum*,

$$\mathrm{P}^{\mathsf{Bur}}(K,d_0) = \begin{cases} \textbf{concede}{:}m_0 & \text{iff } \exists w \in W, \mathfrak{M}, w \vDash m_0 \\ \textbf{deny}{:}m_0 & \text{iff } \forall w \in W, \mathfrak{M}, w \nvDash m_0 \end{cases}$$

- For $d_n, n > 0$:

$$\begin{array}{ll} \text{If } \mathfrak{M} \restriction \Gamma_n \vDash m_n\text{:} & \mathrm{P}^{\mathsf{Bur}}(K,d_n) = \textbf{concede}{:}m_n \\ \text{If } \mathfrak{M} \restriction \Gamma_n \vDash \neg m_n\text{:} & \mathrm{P}^{\mathsf{Bur}}(K,d_n) = \textbf{deny}{:}m_n \\ \text{Otherwise:} & \\ \text{If } \mathfrak{M}, w^* \vDash K_{\mathsf{Res}} m_n\text{:} & \mathrm{P}^{\mathsf{Bur}}(K,d_n) = \textbf{concede}{:}m_n \\ \text{If } \mathfrak{M}, w^* \vDash K_{\mathsf{Res}} \neg m_n\text{:} & \mathrm{P}^{\mathsf{Bur}}(K,d_n) = \textbf{deny}{:}m_n \\ \text{If } \mathfrak{M}, w^* \vDash \neg(K_{\mathsf{Res}} m_n \vee K_{\mathsf{Res}} \neg m_n)\text{:} & \mathrm{P}^{\mathsf{Bur}}(K,d_n) = \textbf{doubt}{:}m_n \end{array}$$

Finally, we define two outcome rules for Burley-style *positio*, governing who wins. Generally speaking, Opp wins if he can force Res into inconsistency, and Res wins otherwise. Since any individual *obligatio* $= d_n$ for some finite n, we can define a weak notion of "local" winning: If $m_n = \alpha$, then Opp *wins* if $\mathfrak{M} \restriction \Gamma_n = \langle \emptyset, \{\sim_a^{\mathfrak{M},\Gamma_n} : a \in A\}, V^{\mathfrak{M},\Gamma_n}\rangle$ and Res wins otherwise. But even though individual *obligationes* are finite, they are all potentially infinite. This view gives rise to a "global" winning condition: Opp *wins* if there is some n such that $\mathfrak{M} \restriction \Gamma_n = \langle \emptyset, \{\sim_a^{\mathfrak{M},\Gamma_n} : a \in A\}, V^{\mathfrak{M},\Gamma_n}\rangle$. Res wins otherwise. In both cases, the only time W will be empty is when $\mathrm{C}_{\mathsf{Res}} \vDash \varphi \wedge \neg\varphi$, that is, over the course of the disputation Res has conceded an inconsistent set, and has thus "responded badly". Thus, protocol $\mathrm{P}^{\mathsf{Bur}}$ ensures the dialogical consistency of Res (cf. [31, p. 171] and [13, ch. 3]).

There are also two ways that "responded badly" can be explicated, a broad-grained way and a fine-grained way. On the broad-grained view, we are only interested in whether Opp or Res has locally won, that is, whether Opp has been able to force Res to concede a contradiction, or whether Res has remained consistent in his answers. This is the view generally considered by medieval authors.

Before we discuss some interesting properties of protocols like the ones introduced here, we close this section by formalizing the example in Figure 1. Fix a model $\mathfrak{M}$ such that $\exists w \in W, \mathfrak{M}, w \vDash \varphi \wedge \psi$, and let $K = \{\neg\varphi\}$. The first move is Opp's, and he asserts the *positum* $\varphi := d_0$. According to $\mathrm{P}^{\mathsf{Bur}}(d_0)$, Res should **concede**:$\varphi := d_1$. The protocol now allows Opp to make any assertion he likes, so he asserts $\neg\varphi \vee \psi := d_2$. Now, by the argument given above, either $\mathfrak{M} \restriction d_1 \vDash d_2$, or, if not, then $\mathfrak{M}, w^* \vDash K_{\mathsf{Res}} d_2$. By the protocol, in both cases, Res should concede and $d_3 =$ **concede**:d_2. Then Opp asserts $d_4 = \psi$, and since $\mathfrak{M} \restriction d_1, d_3 \vDash \psi$, Res is again required to respond with **concede**:$d_4 := d_5$. Then, Opp calls "*cedat tempus*", and $d_6 = \alpha$, and, by the general protocol P_u there are no more legal moves and the dialogue ends.

7 Discussion

7.1 The protocol

The protocol $\mathrm{P}^{\mathsf{Bur}}$ defined above is semi-public, as it depends on Res's knowledge, but does not depend on Opp's; context-dependent; and fully deterministic. It also meets all but four of the 13 desiderata for agent argumentation protocols given in [25]. There McBurney et al. consider dialogue protocols from the point of view of design, and identify 13 *desiderata* that a designer might want to incorporate. These are:

stated dialogue purpose A system should have one or more publicly stated purpose, and its structure should facilitate its achievement.

diversity of individual purposes Participants should be allowed to have their own (distinct) purposes, consistent with the stated dialogue purpose.

inclusiveness Any qualified participant who wants to participant can.

transparency Participants should know the rules and structure of the system prior to its commencement.

fairness Either all participants should be treated equally, or asymmetries in treatment should be explicit.

clarity of argumentation theory The dialectical system should conform to a stated theory of argumentation.

separation of syntax and semantics The syntax and the semantics of the system should be defined separately.

rule-consistency The rules and locutions should not lead to deadlocks or to infinite cycles of repeated locutions.

encouragement of resolution Resolution should not be precluded by the rules.

discouragement of disruption The rules should discourage disruptive behavior, such as repeating the same locution repeatedly.

enablement of self-transformation Participants should be able to change their preferences, knowledge, degrees of belief, etc.

system simplicity The locutions and rules of the system should be as simple as possible, consistent with the other desiderata.

computational simplicity The system should minimize computational demands on the participants.

In the FDS constructed from Burley's obligational theory, the dialogue purpose is stated explicitly (cf. §2). The purposes of Res and Opp are distinct, and anyone who is qualified may participate in an *obligatio* as either Res or Opp may. The rules are agreed upon in advance, and the asymmetry between Res and Opp are explicit. The system conforms to a stated theory of argumentation, namely Burley's rules for *positio*. The rules do not preclude resolution, and it is also quite simple, statable in a case-based structure with minimal cases to consider. Further, to the extent that the protocol is rule-consistent (which we discuss below), it discourages disruption in that Res is never able to continuously repeat the same locution, at least, not without responding badly.

The issue of self-transformation enablement requires further comment. McBurney et al. further specify that self-transformation requires that "participants should have the right to retract commitments made earlier in the same dialogue" for otherwise "in such circumstances, there would be no point for the agents to engage in dialogue" [25, p. 403]. We disagree with this assessment; the *obligationes* framework provides a counterexample, a type of dialogue which is worthwhile engaging in but in which the agents' commitments are monotonic. *Obligationes* satisfy the less restrictive view of self-transformation, given that the commitment sets and responses of Res are flexible and can change over the course of the disputation.

For the remaining three desiderata, because the topic language and the communication language coincide, it is not clear to what extent *obligationes* satisfy the requirement of the separation of syntax and semantics. It does not satisfy rule-consistency of discouragement of disruption, because Opp can continually put forward the same proposition; however, while the protocol allows Opp to act in such a fashion, there are pragmatic reasons why he will generally not do so. In Res responds well (that is, follows the protocol), the only time he will change his response to a proposition φ is to move from doubt to either concede or deny. If he has responded badly at some point, then Opp may by repeating a locution be able to force Res into losing, in which case there will be no point in Opp repeating that proposition again, as he should instead call *Cedat tempus*. Whether the final criterion is satisfied is uncertain: As we note in the final section, the computational complexity of certain decision problems that can be extracted from this protocol is not yet known. Regardless, the protocol scores quite highly—-as well, or better, than the protocols analysed in [25]. It should also be clear that this high score is not specific to Burley's protocol: Any other FDS protocol developed from a different medieval theory would score similarly high.

7.2 Comparison with other dialogue systems

We discussed Walton & Krabbe's taxonomy of dialogue types, and Cogan et al.'s extensions, in §4 before introducing Prakken's framework. Prakken himself distinguishes two types of persuasion dialogues, those that are for *pure persuasion* and those for *conflict resolution* [31, p. 169]. In his terminology, a DS S is for *pure persuasion* iff for any terminated dialogue d, agent a is a winner for topic t ($a \in w_t(d)$) iff either a is a proponent of t and $t \in \mathrm{C}_{a'}(d)$ for all a' who are either proponents or opponents of t, or a is an opponent of t and $t \notin \mathrm{C}_{a'}(d)$ for all a' who are either proponents or opponents of t. A DS S is for *conflict resolution* otherwise.

Where do *obligationes* fit in these schemes? Prakken says that frameworks for persuasion dialogues can be found in the Middle Ages [31, p. 163, citing [1]]. Angelelli's discussion of medieval disputations in [1] focuses on *obligationes*, but nowhere does he make any mention of persuasion. If *obligationes* are a type of persuasion dialogue, then because Opp's commitment store is always empty in an *obligatio*, *obligationes* would be classified, on Prakken's distinction, as conflict resolution dialogues instead of pure persuasion dialogues. However, given the stated goal of *obligationes* and the fact that the players agree on the truth value of the proposition in question at the start of the dialogue, this classification does not seem appropriate.

Despite its breadth, the decempartite division of [9], incorporating Walton and Krabbe's typology, also does not accommodate *obligationes*. Because they are about the consistency of a formula, *obligationes* are not negotiation or deliberation dialogues. Because the truth value of the proposition in dispute is known to both, and the Opponent is not trying to persuade the Respondent of anything, they are not information-seeking, inquiry, or pure persuasion dialogues. Since they are not pugilistic in nature, they are not eristic dialogues. Nor are they any of the four new kinds introduced in [9], since those types require as well that at least one party not know the truth-value of the proposition.

Obligationes are somewhat similar to the 'elicit-inform' dialogue game of [21, 32]. These dialogues were developed by Ravenscroft and Matheson in the context of collaborative e-learning, where the dialogues are between a tutor system and a student. The asymmetry between the players is similar to the asymmetry between the Opponent and the Respondent. In an elicit-inform dialogue, the student is questioned by the tutorial system, and "after reasoning about the learner's contributions, the tutor system either *sanctions* their explanations by informing them they were correct, or points out that they were 'incorrect' and so *informs* them of a consistent, or 'correct' answer" [32, p. 96]. This is very similar to the behavior of the Opponent when he calls *Cedat tempus* and evaluates the actions of the Respondent to determine whether he has responded well or badly. However, as elicit-inform dialogues have as their goal the persuasion of the student to adopt a certain belief, they are not a perfect match for *obligationes*, since persuasion is not at issue in obligational dialogues.

Instead of trying to shoehorn *obligationes* into a type of dialogue system already introduced in the literature, we think it is more interesting, and highlights the unique nature of obligational reasoning better, to consider them as introducing a new type of dialogue into the typology. Thus, one of the contributions of the current paper is the introduction of a new type of inter-agent dialogue, for checking the feasibility of a set of propositions. Proving that a proposition or a set of propositions is feasible—can be consistently maintained—has various applications in constraint and allocation satisfaction, where a consistent solution meeting certain requirements has to be found. Thus, the new dialogue type that we have introduced may turn out to have useful applications beyond the context of *obligationes*, though we do not investigate this potential extension further here.

7.3 The role of obligations

One final point of interest, in comparing protocols based on *obligationes* to other dialogue protocols, is the role of obligations (in the ordinary sense of the term) in the dialogues. Generally, in an FDS, "[s]trictly speaking the only dialectical obligation that a participant has is making an allowed move when it is one's turn" [31, p. 170]. In an *obligatio*, however, Res has a two-tiered obligation: He is required to follow both the uniform protocol P_u, and the appropriate specific protocol for the type of *obligatio* he is in.[6] It is possible for Res to violate his obligation to follow the second protocol (in which case he loses), but if he does not follow the uniform protocol than no disputation even takes place. This two-tiered nature of the obligation of the Res helps us understand King's analysis of *obligationes* as a meta-methodology of argumentation. The specific protocol is the methodology—it tells Res how to respond within a particular disputation—while the general protocol constrains the types of specific protocols that are allowed, and hence can be understood as a meta-methodology (a higher order method).

Additionally, in an *obligatio* there is no connection between an agent's commitment store and his assertions; Opp has no commitments, even though all he makes are assertions, and Res makes no assertions, but his commitments are generated from his concessions and denials of Opp's assertions. This is a significant difference from standard commitment rules such as the one discussed by Prakken in his Paul and Olga example [31, p. 169]. He says "As for commitment rules, the following ones seem to be uncontroversial and can be found throughout the literature:

- If $s(m) = \mathit{claim}(\varphi)$ then $C_s(d, m) = C_s(d) \cup \{\varphi\} \ldots$"

That is, an agent's discursive commitments are generated from his assertions. A similar position is advocated when he says elsewhere that "Commitments are typically incurred by making or conceding claims and stating" [30, p. 1017].

8 Conclusion

We have now seen how at least one type of medieval obligational theory can be interpreted as giving rise to a formal dialogue system; it is straightforward to extend this analysis not just to theories of *positio* outlined by other medieval authors but also other types of *obligationes*, such as *dubitatio*, by varying the underlying logic.[7] The result of such an analysis shows that, just as a particular dialogue can be viewed as a game played according to a set of rules specified by an FDS (cf. §3), so too *obligationes* can be naturally understood as giving the participants a methodology of argumentation or reasoning to follow. By specifying the protocols and rules of an FDS, an particular obligational theory gives participants a framework within which to do philosophical analysis. This provides formal support for King's interpretation of *obligationes* as functioning at the meta-level, rather than at the content level. We close our paper by pointing towards questions that we hope to answer in future work.

In §6 we distinguished two ways "responds badly" can be explicated. Both of these give rise to decision problems whose complexity we intend to investigate in future work (cf. §7). In the first, the local winning outcome conditions corresponds to the decision problem RESPONDS-WELL$(d_n, \mathfrak{M})$ defined as follows: Given a finite *obligatio* d_n and model $\mathfrak{M}$, check whether $\mathfrak{M} \upharpoonright \Gamma_n = \langle \emptyset, \{\sim_a^{\mathfrak{M},\Gamma_n} : a \in A\}, V^{\mathfrak{M},\Gamma_n} \rangle$, that is, whether Opp has locally won.

Question. What is the computational complexity of RESPONDS-WELL$(d_n, \mathfrak{M})$?

Note: The answer is not necessarily a straightforward adaptation of complexity results for model-checking in Public Announcement Logic (PAL, [38, ch. 4]), since in PAL, attention is restricted to announcements which are *truthful*, and the introduction of announcements which may be false, but which are nonetheless effective (in that they reduce the model), is not generally considered.

The fine-grained view asks whether, at each step n, Res has followed the specific protocol he was obligated to follow. It is possible for Res to have responded correctly in the sense of winning locally on the broad-grained view, but to still have not followed the rules correctly, by choosing the wrong response for irrelevant propositions. In this case, the decision problem RULE-FOLLOWING $(d_n, \mathrm{P}, \mathfrak{M})$ is defined as follows: Given a finite *obligatio* d_n, protocol P, and model $\mathfrak{M}$ check whether the construction of Γ_n satisfies the conditions of P.

Question. What is the computational complexity of RULE-FOLLOWING$(d_n, \mathrm{P}, \mathfrak{M})$?

Note: It may be possible to extend the complexity results of [27, 28, 29], though it is not *prima facie* clear how this would be done. In particular, the results in [29, §7] are for argumentation-based logics, which DEL is not, and the results in [27, §5–6] are for protocols where repetition is not allowed.

ACKNOWLEDGEMENTS

This research was funded by the NWO project "Dialogical Foundations of Semantics" (DiFoS) in the ESF EuroCoRes programme LogICCC (LogICCC-FP004; DN 231-80-002; CN 2008/08314/GW). The author would like to thank the anonymous referees for their meticulous and useful comments on an earlier version of this paper.

REFERENCES

[1] I. Angelelli, 'Techniques of disputation in the history of logic', *Journal of Philosophy*, **67**(20), 800–815, (1970).
[2] M. Baker, 'The role and models in artificial intelligence and educational research: A prospective view', *International Journal of Artificial Intelligence in Education*, **11**, 122–143, (2000).
[3] A. Bondarenko, P. M. Dung, R. A. Kowalski, and F. Toni, 'An abstract, argumentation-theoretic approach to default reasoning', *Artificial Intelligence*, **93**, 63–101, (1997).
[4] H.A.G. Braakhuis, 'Obligations in early 13th century Paris: the *Obligationes* of Nicholas of Paris (?)', *Vivarium*, **36**(2), 152–233, (1998).
[5] G. Brewka, 'Dynamic argument systems: a formal model of argumentation processes based on situation calculus', *Journal of Logic and Computation*, **11**(2), 257–282, (2001).
[6] W. Burley, 'Obligations (selections)', in *Cambridge Translations of Medieval Philosophical Texts*, eds., N. Kretzmann and E. Stump, volume 1, 369–412, Cambridge University Press, (1988).
[7] W. Burley and R. Green, O.F.M., *An Introduction to the Logical Treatise 'de Obligationibus'*, volume 2, Université Catholique Louvain, 1963.
[8] E. Cogan, S. Parsons, and P. McBurney, 'What kind of argument are we going to have today?', in *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, (2005).
[9] E. Cogan, S. Parsons, and P. McBurney, 'New types of inter-agent dialogues', in *Argumentation in Multi-Agent Systems*, eds., S. Parsons, N. Maudet, P. Moraitis, and I. Rahwan, LNAI 4049, 154–168, Springer, (2006).
[10] L. M. de Rijk, 'Some thirteenth century tracts on the game of obligation', *Vivarium*, **12**(2), 94–123, (1974).

[6] The obligation to follow the uniform protocol is what Walton and Krabbe term an 'action commitment': an obligation to execute a particular course of action, given the action(s) of the other player(s) [39, ch. 1].
[7] The case of *dubitatio* is considered in [37].

[11] L. M. de Rijk, 'Some thirteenth century tracts on the game of obligation II', *Vivarium*, **13**(1), 22–54, (1975).
[12] L. M. de Rijk, 'Some thirteenth century tracts on the game of obligation III', *Vivarium*, **14**(1), 26–49, (1976).
[13] C. Dutilh Novaes, *Formalizing Medieval Logical Theories: Suppositio, Consequentiae and Obligationes*, volume 7 of *Logic, Epistemology, and the Unity of Science*, Springer, 2007.
[14] R. A. Girle, 'Commands in dialogue logic', in *Practical Reasoning: Proceedings of the First International Conference on Formal and Applied Practical Reasoning (FAPR 1996), Bonn, Germany*, eds., D. M. Gabbay and H. J. Ohlbach, volume 1085 of *LNAI*, pp. 246–260. Springer, (1996).
[15] C. Hamblin, 'Mathematical models of dialogue', *Theoria*, **37**, 130–155, (1971).
[16] C. L. Hamblin, *Fallacies*, Methuen, London, [1970].
[17] D. Harel, D. Kozen, and J. Tiuryn, 'Dynamic logic', in *Handbook of Philosophical Logic*, eds., D. M. Gabbay and F. Guenther, volume 4, Kluwer, 2 edn., (2002).
[18] N. C. Karunatillake, N. R. Jennings, I. Rahwan, and P. McBurney, 'Dialogue games that agents play within a society', *Artificial Intelligence*, **173**, 935–981, (2009).
[19] P. King, 'Opposing and responding: comments on Paul Spade'. preprint, `http://individual.utoronto.ca/pking/presentations/Spade_Comments.pdf`, 2004.
[20] P. Lorenzen and K. Lorenz, *Dialogische Logik*, Wissenschaftliche Buchgesellschaft, 1978.
[21] M. Matheson and A. Ravenscroft, 'Evaluating and investigating learning through collaborative argumentation: An empirical study', Technical Report DDRG-01–01, The Open University, Milton Keynes, (2001).
[22] N. Maudet, 'Negotiating dialogue games', *Autonomous Agents and Multi-Agent Systems*, **7**, 229–233, (2003).
[23] P. McBurney and S. Parsons, 'Games that agents play: A formal framework for dialogues between autonomous agents', *Journal of Logic, Language and Information*, **11**, 315–334, (2002).
[24] P. McBurney and S. Parsons, 'Dialogue games for agent argumentation', in *Argumentation in Artificial Intelligence*, eds., I. Rahwan and G. Simari, 261–280, Springer, (2009).
[25] P. McBurney, S. Parsons, and M. Wooldridge, 'Desiderata for agent argumentation protocols', in *Proceedings of AAMAS'02, July 15–19, 2002, Bologna, Italy*, pp. 402–409, (2002).
[26] W. of Ockham, *Opera Philosophica*, volume I, Franciscan Institute, St. Bonaventure, 1974.
[27] S. Parsons, M. Wooldridge, and L. Amgoud, 'An analysis of formal inter-agent dialogues', in *Proceedings of AAMAS'02, July 15–19, 2002, Bologna, Italy*, pp. 394–401, (2002).
[28] S. Parsons, M. Wooldridge, and L. Amgoud, 'On the outcomes of formal inter-agent dialogues', in *Proceedings of the 2nd International Conference on Autonomous Agents and Multi-Agent Systems, Melbourne*, (2003).
[29] S. Parsons, M. Wooldridge, and L. Amgoud, 'Properties and complexity of some formal inter-agent dialogues', *Journal of Logic and Computation*, **13**(3), 347–376, (2003).
[30] H. Prakken, 'Coherence and flexibility in dialogue games for argumentation', *Journal of Logic and Computation*, **15**(6), 1009–1040, (2005).
[31] H. Prakken, 'Formal systems for persuasion dialogue', *Knowledge Engineering Review*, **21**(2), 163–188, (2006).
[32] A. Ravenscroft and M. P. Matheson, 'Developing and evaluating dialogue games for collaborative e-learning', *Journal of Computer Assisted Learning*, **18**(1), 93–101, (2002).
[33] E. Sklar and S. Parsons, 'Toward the application of argumentation-based dialogues for education', in *Proceedings of the 3rd International Conference on Autonomous Agents and Multi-Agent Systems*, eds., C. Sierra and E. Sonenberg, (2004).
[34] P. V. Spade, 'Why don't medieval logicians ever tell us what they're doing? or, what is this, a conspiracy?'. preprint, `http://pvspade.com/Logic/docs/Conspiracy.pdf`, 2000.
[35] P. V. Spade. Unpublished talk on *obligationes* presented at the Midwestern Division of the American Philosophical Association, April 1993.
[36] S. L. Uckelman. 'A unified dynamic framework for modeling *obligationes*', 2010. preprint, `http://staff.science.uva.nl/~suckelma/latex/jolli-ob/jolli-ob.pdf`, in submission.
[37] S. L. Uckelman, 'Deceit and nondefeasible knowledge: The case of *dubitatio*'. preprint, `http://staff.science.uva.nl/~suckelma/latex/loft2010/loft2010.pdf`, 2010.
[38] H. van Ditmarsch, W. van der Hoek, and B. Kooi, *Dynamic Epistemic Logic*, volume 337 of *Synthese Library Series*, Springer, 2007.
[39] D. N. Walton and E. C. W. Krabbe, *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*, State University of New York Press, 1995.
[40] M. Yrjönsuuri, 'The role of casus in some fourteenth century treatises on sophismata and obligations', in *Argumentationstheorie. Scholastische Forschungen zu den logischen und semantischen Regeln korrekten Folgerns*, ed., K. Jacobi, 301–321, Brill, (1993).
[41] M. Yrjönsuuri, *Obligationes: 14th Century Logic of Disputational Duties*, volume 55 of *Acta Philosophica Fennica*, Societatis Philosophia Fennica, 1994.
[42] M. Yrjönsuuri, ed., *Medieval Formal Logic*, Kluwer, 2001.