Obligationes as formal dialogue systems

Uckelman, S.L.

Published in:
STAIRS 2010

Publication date:
2011

Obligationes as Formal Dialogue Systems

Sara L. UCKELMAN [a,1,2]

[a] *Institute for Logic, Language, and Computation*

Abstract. Formal Dialogue Systems (FDSs) model rule-based interaction between agents. Their conceptual roots go back to Hamblin's [11,12], which cites the medieval theory of *obligationes* as inspiration for his development of a formal system of argumentation. In an *obligatio*, two agents, Opponent and Respondent, engage in an alternating-move dialogue, where 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 [20] using Dynamic Epistemic Logic [26]. The result is a new type of inter-agent dialogue, *consistency-checking*, and analyzing *obligationes* in this way also sheds light on interpretational and historical questions concerning their use and purpose in medieval academia.

Keywords. dialogue protocol, disputation, Formal Dialogue Systems, *obligationes*

1. Introduction

Rule-based interactions such as dialogues are ubiquitous and diverse; they are the basic method of communication between agents. In the context of AI and computer science, formal dialogue systems (FDSs) such as those developed in, e.g., [14,20] give formal, and hence potentially implementable, methods for modeling real-life dialogue situations, for example the complex reasoning in legal domains. The conceptual roots of FDSs are found in philosophical logic, argumentation theory, and, more broadly, the role of dialogue or argumentation in law and philosophy. One of the earliest attempts to provide a theory of formal dialogues is Hamblin's [12,11]. In [12], 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 that he considers in particular is the medieval theory of *obligationes*, developed in the 13th and 14th centuries. In an *obligatio*, two agents, Opponent and Respondent, engage in an alternate-move dialogue, where 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, and

[1] Institute for Logic, Language, and Computation, University of Amsterdam, The Netherlands, email: `S.L.Uckelman@uva.nl`

[2] This research was funded by the 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 helpful comments on an earlier version of this paper.

that they determine a new type of dialogue system different from those discussed in the AI literature.

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. 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 and show generally how *obligationes* fit within this set-up. In order to give a precise specification, we must first outline the logic used in the argumentation, which we do in §5. In §6 we then give a precise characterization of *obligationes* as FDSs. In §7, we compare the result with standard types of FDSs, and define a new dialogue protocol for *consistency-checking*.

2. The medieval theory of *obligationes*

An *obligatio* is a dialogue between two agents, Opponent and Respondent, where Opponent puts forward a sequence of propositions, and Respondent is obligated (hence the name) to follow certain rules in his responses to Opponent's propositions. More precisely, Opponent puts forward an initial statement, called the *positum*, which 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, Opponent puts forward propositions and 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 the anonymous author of the *Obligationes Parisienses* [7], mention a fourth option, which is to 'draw distinctions', that is, to clarify an ambiguity on the part of Opponent.) The *obligatio* continues until Opponent calls "*Cedat tempus*" ("Time's up"), whereupon the responses of Respondent are analysed with respect to 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 [6,7,8], 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. [29, §II.A]), the systematic development of the theory of *obligationes* over the course of the 13th 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 [29].

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 [10] and a partial translation of the text, including the section on *positio* in its entirety, is found in [4]. 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* [4, p. 370].[3]

Thus, it is clear that in an *obligatio* the goal is consistency, not logical truth or validity. In *positio*, the primary obligation of Respondent is to grant, that is, to hold as true, the *positum*. If 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 [4, p. 379].[4]

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 [4, p. 381].[5]

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 [4, p. 381].[6]

In Rule 1, 'in the same form as' should be understood syntactically; if the *positum* is 'Marcus is Roman', then 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 [4, p. 375].[7]

I.e., Respondent should concede the proposition if it is true, deny if it is false, and doubt if he is not sure.

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

[3] *Opus opponentis est sic inducere orationem ut faciat respondentem concedere impossibilia quae propter positum non sunt necessaria concedere. Opus autem respondentis est sic sustinere positum ut propter ipsum non videatur aliquod impossibile sequi, sed magis propter positum* [10, p. 34].

[4] *Omne positum, sub forma positi propositum, in tempore positionis, est concedendum* [10, p. 46].

[5] *Omne sequens ex posito est concedendum. Omne sequens ex posito cum concesso vel concessis, vel cum opposito bene negati vel oppositis bene negatorum, scitum esse tale, est concedendum* [10, p. 48].

[6] *Omne repugnans posito est negandum. Similiter omne repugnans posito cum concesso vel concessis, vel opposito bene negati vel oppositis bene negatorum, scitum esse tale, est negandum* [10, p. 48].

[7] *[A]d impertinens respondendum est secundum sui qualitatem* [10, p. 42].

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

Figure 1. An example *obligatio*.

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 rules require the Respondent to concede; in the first case, Rule 2 is operational, in the second, 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 necessarily inconsistent, Opponent can force Respondent to concede any other consistent proposition.

Other examples commonly found in *obligationes* treatises are more interesting, because they involve *posita* that are not propositional but instead include statements about the players obligations in the games (e.g., the example discussed in [29, pp. 152–155]).

3. Previous work on *obligationes*

Green's Ph.D. dissertation [10], 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*. Hamblin is the first modern author to attempt to formalize *obligationes* [12, 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[8], but it marks the beginning of modern scholarship on the formal properties of *obligationes*. More recent scholarship has focused on the game-like nature of *obligationes*, e.g., [7,9,29]. 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, despite the strongly logical component of *obligationes*, to date relatively little work has been done on the formal properties of the logic and few attempts have been made to provide an explicit specification of the game(s) involved[9], and 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* is difficult to define[10]).

[8] 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 [2].

[9] The most extensive attempt is [9], which contains analyses of the obligational theories of Walter Burley, Richard Swyneshed (c.1330), and Ralph Strode (second half of the 14th C). However, Dutilh Novaes's framework is not very *game*-like; there are players and winning conditions, but no concept of, e.g., strategy. Other drawbacks of her framework for modeling *obligationes* are discussed in [24].

[10] 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" [29, pp. 9–10].

In the last four decades, 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 [23,30,29]. One particularly interesting interpretation is given by King [15], who takes his starting point from Spade, who, in [22], 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 [15, 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" [15, p. 6], and thus *obligationes* provide a "meta-methodology" for reasoning [15, 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 [17]. 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 formal 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 can take place.

4. Formal dialogue systems

In this section, we follow the presentation of formal dialogue systems given by Prakken in [20], 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 [20, pp. 170, 173]. Thus, it is appropriate to use it to consider *obligationes*.

A *formal dialogue system* contains the following elements [20, 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* $\mathrm{C}_a(d_n) : M^{<\infty} \rightarrow \mathcal{P}(\mathcal{L}_t)$ 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 $T : D \rightarrow \mathcal{P}(\mathcal{A})$, 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.

We can identify a number of properties of protocols [20, 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 FDS where the protocol for Respondent is fully deterministic.

We now show how generically *obligationes* can be viewed as FDSs; we give precise examples in §6. In *obligationes*, there are two designated roles O (Opponent) and R (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_{O}, Σ_{R}, C_{O}, C_{R}, and the context K are generated. In Burley-style *positio*, the dialogue purpose is consistency: If we take R'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 R is to maintain the consistency of his commitment set, and the goal for O is to force R 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, O may ask R to respond to the claim

“You deny φ”.) The turn-taking protocol in an *obligatio* is unique-move: $T(\emptyset) = \mathsf{O}$, $T(d_n) = \mathsf{O}$ if n is odd, and $T(d_n) = \mathsf{R}$ if n is even. (Throughout we assume that we label the steps in the sequence from 0, so in an *obligatio* it is always O that goes first.) The protocol P will be such that the moves of O are not constrained in any way, but R's moves must be made in reaction to the move of O at the previous stage. The same will be true for the effect rules E; in a disputation, O makes a series of claims or assertions, but these actions have no effect on his commitment store. On the other hand, R is constrained to be reactive only: He can only concede statements claimed by O, concede their negations, or remain ambivalent. R never asserts any statement of his own devising, he only ever responds to propositions put forward by O. Thus, *obligationes* are essentially asymmetric, in that the rules governing the behavior of the O and R are disjoint[11], and so are their actions.

The outcome rules for *obligationes* are simple: If R realizes the goal, then he wins. If O 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 FDS, 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 later 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 his commitments be revised in order to maintain consistency (cf. [1,3]). In AI contexts, the ability to simulate non-monotonic reasoning is of great importance; in philosophical contexts, dialogues and disputations are more likely to be monotonic. One of the benefits of Prakken's approach to FDS is that it can handle both approaches, merely by the specification of the underlying logic [20, p. 173].

5. The underlying logic

By specifying the logic L and its underlying models, we are able to explicitly generate Σ_{O}, Σ_{R}, C_{O}, C_{R}, and K satisfying desired properties. In our approach to modeling *obligationes* as FDSs, the underlying logic is multi-agent Dynamic Epistemic Logic (DEL, [26]). This logic is monotonic and not argument based (we discuss below our motivations for selecting this type of logic for our underlying logic). 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 as follows:

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

[11] In fact, in most texts, no rules for O are given. One exception is the early text *Tractatus Emmeranus* [6], which gives some rules (better thought of as guidelines, or strategic advice) to Opponent.

$K_a\varphi$ 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 can use C_G to represent explicitly the knowledge of the two agents at the beginning of the disputation, if so required.

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 a designated point $w^* \in W$ (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.

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

EL models cover the knowledge of the agents; to model their actions, we add dynamics, via Propositional Dynamic Logic (PDL, [13]). 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:

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

The programme $\varphi?$ is similar to the ordinary test operator in PDL, in that it tests for the truth of φ, but it differs in that it does not, as the truth conditions below make clear, require the truth of φ at the actual world. 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 $\mathfrak{M} \upharpoonright \varphi := \langle W^{\mathfrak{M},\varphi}, \{\sim^{\mathfrak{M},\varphi}_a : a \in \mathcal{A}\}, V^{\mathfrak{M},\varphi}\rangle$, 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 $\mathfrak{M} \upharpoonright \Gamma_n = \mathfrak{M} \upharpoonright \gamma_0 \upharpoonright \cdots \upharpoonright \gamma_n$, that is, $\mathfrak{M} \upharpoonright \Gamma_n$ is the result of the sequential restriction of $\mathfrak{M}$ by the elements of Γ_n. Then:

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

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 other advantages see [24]). We are in general not interested in the *belief* bases of the agents, but rather their *knowledge* bases; while beliefs may be false, we follow the standard definition of knowledge, on which it is veridical. Given an epistemic model $\mathfrak{M}$, the knowledge bases of O and R are defined as follows:

$$\Sigma_{\mathsf{O}}^{\mathfrak{M}} := \{\varphi\colon \mathfrak{M}, w^* \vDash K_{\mathsf{O}}\varphi\}$$
$$\Sigma_{\mathsf{R}}^{\mathfrak{M}} := \{\varphi\colon \mathfrak{M}, w^* \vDash K_{\mathsf{R}}\varphi\}$$

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 [28]. 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 $w \sim_{\mathsf{R}} w^*$, then $w \in P_c$, if $v \sim_{\mathsf{O}} w^*$, then $v \in P_c$, and for all $w, v \in P_c$, $w \sim_{\mathsf{R}} v$ and $w \sim_{\mathsf{O}} v$; 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 FDS, it is not assumed that the *casus* of an *obligatio* is consistent, but if it is not, then R should not accept the *positum*, since O 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 $\varphi \in K$, $\mathfrak{M} \vDash C_{\{\mathsf{O},\mathsf{R}\}}\varphi$, and so $K \subseteq \Sigma_{\mathsf{O}}^{\mathfrak{M}}$ and $K \subseteq \Sigma_{\mathsf{R}}^{\mathfrak{M}}$.

We close this section by briefly commenting on our choice of DEL for the underlying logic, as opposed to an argument-based logic or one that allows for nonmonotonicity. First, as we noted above, one advantage of using an epistemic logic rather than an argument-based logic is that we can explicitly discuss the knowledge bases of the agents, and the context. While the applications that we discuss in this paper do not exploit this expressivity, dealing with individual knowledge and common knowledge, other types of *obligationes* which we do not consider here make crucial use of this information during the disputation. Building epistemic characteristics into our underlying logic from the start means that this framework can be used to model a much wider range of types of *obligationes* than we consider in this paper (cf. [24,25] for two further applications).

Second, a general characteristic shared by almost all types of *obligationes*, both *positio* and otherwise, is that if R follows the rules correctly, the only time he is required to give different answers to the same proposition at different rounds of the disputation is if he first doubts the proposition, and then at a later stage concedes or denies it. Thus, the obligational systems are monotonic in so far as the only change in R's response is moving from doubt concerning a specific proposition to certainty (that is, concession or denial). Thus, since the rules are monotonic, we do not need to consider defeasibility.

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 *obliga-*

tiones. We identify our set of agents with their roles, i.e., our set of agents is $\mathcal{A} = \{\mathsf{O}, \mathsf{R}\}$, 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 O's turn, he is allowed to assert any statement in the communication language (we allow repetitions). If it is R's turn, he must either concede, deny, or doubt O's statement from the previous round. Since m_n, the move of O, 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 R's turn, he has a choice of actions. For ease of future reference, we introduce meta-names for the actions of R: **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 C_{O} and C_{R} are defined as follows:

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

That is, O has no commitments, O's moves do not change R's commitments, and R's commitment store strictly grows on the basis of his actions, and thus obligational dialogues are monotonic (cf. above). As above, since $\mathcal{L}_c$ and $\mathcal{L}_t$ coincide, the final clause of the definition is well-defined. Note that in general, C_{R} and Σ_{R} will be disjoint, and similarly for C_{R} and K (contra, e.g., [19, §3]).

The general protocol defined above specifies what the possible moves of R 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 R is under obligation to respond to O 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 R'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*,

$$\mathrm{P}^{\mathsf{Bur}}(K, d_0) = \begin{cases} \textbf{concede}{:}m_0 & \text{iff } \exists w \in W, \mathfrak{M}, w \models m_0 \\ \textbf{deny}{:}m_0 & \text{iff } \forall w \in W, \mathfrak{M}, w \not\models m_0 \end{cases}$$

- For $d_n, n > 0$:

If $\mathfrak{M} \upharpoonright \Gamma_n \vDash m_n$:	$\mathrm{P}^{\mathrm{Bur}}(K, d_n) =$ **concede**:m_n
If $\mathfrak{M} \upharpoonright \Gamma_n \vDash \neg m_n$:	$\mathrm{P}^{\mathrm{Bur}}(K, d_n) =$ **deny**:m_n
Otherwise:	
If $\mathfrak{M}, w^* \vDash K_{\mathsf{R}} m_n$:	$\mathrm{P}^{\mathrm{Bur}}(K, d_n) =$ **concede**:m_n
If $\mathfrak{M}, w^* \vDash K_{\mathsf{R}} \neg m_n$:	$\mathrm{P}^{\mathrm{Bur}}(K, d_n) =$ **deny**:m_n
If $\mathfrak{M}, w^* \vDash \neg(K_{\mathsf{R}} m_n \vee K_{\mathsf{R}} \neg m_n)$:	$\mathrm{P}^{\mathrm{Bur}}(K, d_n) =$ **doubt**:m_n

This protocol is semi-public, as it depends on R's knowledge, but does not depend on O's; context-dependent; and fully deterministic. It also meets all but four of the 13 desiderata for agent argumentation protocols given in [18].[12] It thus scores as well, or better, than the protocols that they analyse.

We can define two outcome rules for Burley-style *positio*, governing who wins. Generally speaking, O wins if he can force R into inconsistency, and R 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 O *wins* if $\mathfrak{M} \upharpoonright \Gamma_n = \langle \emptyset, \{\sim_a^{\mathfrak{M},\Gamma_n} : a \in A\}, V^{\mathfrak{M},\Gamma_n} \rangle$ and R wins otherwise. But even though individual *obligationes* are finite, they are all potentially infinite. This view gives rise to a "global" winning condition: O *wins* if there is some n such that $\mathfrak{M} \upharpoonright \Gamma_n = \langle \emptyset, \{\sim_a^{\mathfrak{M},\Gamma_n} : a \in A\}, V^{\mathfrak{M},\Gamma_n} \rangle$. R wins otherwise. In both cases, the only time W will be empty is when $\mathrm{C}_{\mathsf{R}} \vDash \varphi \wedge \neg\varphi$, that is, over the course of the disputation R has conceded an inconsistent set, and has thus "responded badly". Thus, protocol $\mathrm{P}^{\mathrm{Bur}}$ ensures the dialogical consistency of R (cf. [20, p. 171] and [9, 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 O or R has locally won, that is, whether O has been able to force R to concede a contradiction, or whether R has remained consistent in his answers. This is the view generally considered by medieval authors.

7. Discussion

7.1. *Understanding* obligationes

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 to the theories of other medieval authors (and we intend to do so in future work). 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 argu-

[12] 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 (7), and because O can continually put forward the same proposition and, on some obligational theories, R can always doubt, they do not satisfy rule-consistency (8) and discouragement of disruption (10) as they define them. 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, so we do not know yet if it satisfies computational simplicity (13).

mentation or reasoning to follow. In particular, the two-tiered nature of the protocols involved in *obligationes*, with both the general uniform protocol P_u and then a specific protocol for a particular type of *obligationes*, such as $\mathrm{P}^{\mathrm{Bur}}$ for Burley-style *positio*, helps us understand King's analysis of *obligationes* as a meta-methodology of argumentation. The specific protocol is the methodology—it tells R 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). Thus, this new approach to *obligationes* provides formal support for King's interpretation of *obligationes* as functioning at the meta-level, rather than at the content level. That is, by specifying the protocols and rules of an FDS, a particular obligational theory gives participants a framework within which to do philosophical analysis.

7.2. Comparison

Walton and Krabbe in [27] give a typology for dialogues, identifying six different basic types: information seeking, inquiry, persuasion, negotiation, deliberation, and eristic. Cogan et al. extend this division by introducing four new types of dialogues, verification and three types of queries, as they 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" [5, p. 161]. (Their new classification is based on the preconditions for dialogues.) A natural question to ask is where do *obligationes* fit in these schemes? The decempartite division of [5] 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 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 [5], 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 [16,21]. These dialogues, between a tutor system and a student, were developed in the context of collaborative e-learning. 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" [21, p. 96]. This resembles 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 complete match for *obligationes*, since persuasion is not at issue in obligational dialogues.

We conclude that the type of protocol that are generated by *obligationes* are best understood on their own merits, and not shoehorned into a type of dialogue already identified. Thus, one of the contributions of the current paper is the introduction of a new type of inter-agent dialogue, which we can term *consistency-checking*.

References

[1] 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).

[2] H.A.G. Braakhuis, 'Obligations in early 13th century Paris: the *Obligationes* of Nicholas of Paris (?)', *Vivarium*, **36**(2), 152–233, (1998).

[3] G. Brewka, 'Dynamic argument systems', *Journal of Logic and Computation*, **11**(2), 257–282, (2001).

[4] 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).

[5] 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).

[6] L. M. de Rijk, 'Some thirteenth century tracts on the game of obligation', *Vivarium*, **12**(2), 94–123, (1974).

[7] L. M. de Rijk, 'Some thirteenth century tracts on the game of obligation II', *Vivarium*, **13**(1), 22–54, (1975).

[8] L. M. de Rijk, 'Some thirteenth century tracts on the game of obligation III', *Vivarium*, **14**(1), 26–49, (1976).

[9] C. Dutilh Novaes, *Formalizing Medieval Logical Theories*, Springer, 2007.

[10] R. Green, O.F.M., *An Introduction to the Logical Treatise 'de Obligationibus'*, Université Catholique Louvain, 1963.

[11] C. Hamblin, 'Mathematical models of dialogue', *Theoria*, **37**, 130–155, (1971).

[12] C. L. Hamblin, *Fallacies*, Methuen, [1970].

[13] 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).

[14] 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).

[15] P. King, 'Opposing and responding: comments on Paul Spade'. Preprint, `http://individual.utoronto.ca/pking/presentations/Spade_Comments.pdf`, 2004.

[16] 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).

[17] N. Maudet, 'Negotiating dialogue games', *Autonomous Agents and Multi-Agent Systems*, **7**, 229–233, (2003).

[18] P. McBurney, S. Parsons, and M. Wooldridge, 'Desiderata for agent argumentation protocols', in *AAMAS'02*, pp. 402–409, (2002).

[19] S. Parsons, M. Wooldridge, and L. Amgoud, 'An analysis of formal inter-agent dialogues', in *AAMAS'02*, pp. 394–401, (2002).

[20] H. Prakken, 'Formal systems for persuasion dialogue', *Knowledge Engineering Review*, **21**(2), 163–188, (2006).

[21] 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).

[22] P. V. Spade, 'Unpublished talk on *obligationes* presented at the Midwestern Division of the American Philosophical Association'. 1993.

[23] 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.

[24] S. L. Uckelman, "A unified dynamic framework for modeling *obligationes*". Working paper, 2010.

[25] S. L. Uckelman, 'Deceit and nondefeasible knowledge: The case of *dubitatio*'. Working paper, 2010.

[26] H. van Ditmarsch, W. van der Hoek, and B. Kooi, *Dynamic Epistemic Logic*, Springer, 2007.

[27] D. N. Walton and E. C. W. Krabbe, *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*, State University of New York Press, 1995.

[28] M. Yrjönsuuri, 'The role of casus in some fourteenth century treatises on sophismata and obligations', in *Argumentationstheorie*, ed., K. Jacobi, 301–321, Brill, (1993).

[29] M. Yrjönsuuri, *Obligationes: 14th Century Logic of Disputational Duties*, Societatis Philosophia Fennica, 1994.

[30] M. Yrjönsuuri, ed., *Medieval Formal Logic*, Kluwer, 2001.