Hidden Protocols

Hans van Ditmarsch
Dept. of Logic
University of Sevilla
Calle Camilo José Cela s/n
41018 Sevilla, Spain
hvd@us.es

Sujata Ghosh
Dept. of Artificial Intelligence
University of Groningen
P.O. Box 407
9700 AK Groningen,The Netherlands
sujata@ai.rug.nl

Rineke Verbrugge
Dept. of Artificial Intelligence
University of Groningen
P.O. Box 407
9700 AK Groningen,The Netherlands
rineke@ai.rug.nl

Yanjing Wang
Dept. of Philosophy
Peking University
100871, Beijing, China
y.wang@pku.edu.cn

ABSTRACT

When agents know a protocol, this leads them to have expectations about future observations. Agents can update their knowledge by matching their actual observations with the expected ones. They eliminate states where they do not match. In this paper, we study how agents perceive protocols that are not commonly known, and propose a logic to reason about knowledge in such scenarios.

Categories and Subject Descriptors

F.4.1 [**Mathematical Logic**]: Modal Logic; D.4.7 [**Organization and Design**]: Distributed Systems; I.2.3 [**Artificial Intelligence**]: Deduction and Theorem Proving

General Terms

Theory

Keywords

communication protocols, dynamic epistemic logic, information hiding

1. INTRODUCTION

Talking about knowledge and protocols, some questions that come foremost to our mind concern the following issues. *What do we mean by knowing a protocol? How does this protocol knowledge affect our knowledge of facts about the world?* The existing literature abounds with various formal models answering these questions from different angles [8, 15, 19, 27, 11]. In some situations, agents have partial knowledge of the underlying protocols that guide the behaviors of other agents. Based on their incomplete knowledge of protocols and their observations, the agents try to reason about their epistemic attitudes as well as hard facts. These protocols may occur when agents communicate using full-blown secret codes (see [18] for many intriguing historical examples). Our daily communications provide more mundane protocols.

Consider, as **Example 1**, a café in the 1950's, with three persons, Kate, Jane and Ann sitting across a table. Suppose Kate is gay and wants to know whether any of the other two is gay. She wants to convey the right information to the right person, without the other getting any idea of the information that is being communicated. She states, 'I am musical, I like Kathleen Ferrier's voice'. Jane, who is gay herself, immediately realizes that Kate is gay, whereas, for Ann, the statement just conveys a particular taste in music.[1]

Coming back to the present day, consider as **Example 2** a similar café scenario with Carl, Ben and Alice. Carl and Ben are childhood friends and know each other like the back of their hands. Carl to Ben: 'On Valentine's day I went to the pub with Mike and Sara. It was a crazy night!'. This immediately catches the attention of Alice who is in love with Mike. She asks: 'What happened?' Carl winks to Ben and says: 'Nothing'. Knowing Carl very well Ben immediately realizes that nothing has happened, whereas Alice becomes unsure of that, as she saw the wink that Carl has given to Ben.

This paper presents a dynamic epistemic logic (*DEL*, [1, 21]) that can suitably describe such scenarios. Knowing a protocol can mean 'knowing what to do according to the protocol' [8]. It can also correspond to 'understanding the underlying meaning of the actions induced by the protocol' [15]. Here, we follow the latter interpretation, because it aptly captures the notion of a protocol in situations we are modeling. Kate's making a statement like 'I am musical, I like Kathleen Ferrier's voice' corresponds to the fact that

[1] This example has been inspired by the interviews in [25], from which it appears that in 1950's Amsterdam, 'musical' was indeed a code term for 'gay', known almost exclusively by gay people. The additional mention of singer Kathleen Ferrier strengthened this 'gay' hint. Among gay women, Ferrier's low contralto voice, for example in her performance as Orfeo in Gluck's Orfeo ed Euridice, was widely popular.

'Kate is gay'. In the second situation, 'Nothing' (even if accompanied by a wink) corresponds to the fact that 'Nothing has happened'.

Our work is largely inspired by two lines of research: the work relating DEL and ETL [19, 11, 14] and the work on protocol changes [27, 28]. In [14] protocols are modeled as tree compositions, basically equating protocols with plans. In [19, 11], the notion of 'state-dependent' DEL-protocols (sets of sequences of *event models* [21]) is proposed in order to handle protocols that are not common knowledge. For example, the model

$$s : p(a) \text{ —}_{1,2}\text{— } t : \neg p(b)$$

represents an epistemic scenario where the agents are not only uncertain about the factual state of the world but also about the protocol that can be executed given some factual state; this is denoted by a state-dependent protocol assigning singleton action sets $\{a\}$ to s and $\{b\}$ to t, where a and b are actions. A system wherein the protocol can be different in any state is clearly more complex than a system wherein the protocol is a background parameter, and thus can be assumed common knowledge to all agents. But in our example we can still reclaim some form of common knowledge of the protocol, namely by describing it as follows: **if** p **then** a and **if** $\neg p$ **then** b. It seems that in order to discuss the knowledge of protocols formally we need to first fix a protocol specification language.

Given a protocol language, how do we obtain such epistemic models with protocol information from specifications of conditional protocols, and vice versa? Similar questions are addressed in [27, 28], presenting a logical framework that incorporates protocol specifications on epistemic models. However, there, protocols are assumed to be common knowledge. We do not assume that here. Our work is based on the logic developed in [27] but it uses epistemic models with procedural information as in [19, 11] to deal with uncertainties of protocols, an agent's knowledge of underlying protocols, and her current observations affecting factual uncertainty. In our framework, the protocols can be viewed as 'given by nature', so the framework does not cover interesting aspects such as how and by whom the protocols have been designed and how agents have come to agree to use them.

The ingredients of our work are: 1. epistemic models encoding state-dependent expected observations; 2. an update mechanism for eliminating impossible worlds according to the observation of agents and their expectations; 3. a formal language for specifying observations and protocols; 4. protocol models that represent agents' incomplete information about the 'real' protocols. 5. an update mechanism for incorporating protocol information (as protocol models) on epistemic (observation) models; 6. a notion of equivalence between protocol models; 7. a logic for reasoning about knowledge based on protocols.

The paper is organized as follows. Section 2 introduces the epistemic observation models and a simple PDL-style epistemic logic for reasoning about knowledge via matching the expectations and observations. Section 3 discusses how we obtain observation models from protocol models (i.e., epistemic protocols). We characterize three classes of observation models that can be generated from various epistemic models. Furthermore we give a characterization of the effective-equivalence of epistemic protocols. A logic is then given to incorporate the updates of protocols and reasoning about knowledge and observations. In the end we address incorporation of factual change actions in Section 4 and point out future work in Section 5.

2. REASONING VIA EXPECTATION AND OBSERVATION

In this section, we introduce *observation models*, which are Kripke models with expected observations, and propose a dynamic logic style epistemic logic interpreted on such models for reasoning about knowledge via matching observations with expectations.

2.1 Epistemic Observation Models

Let $\mathbf{I}$ be a finite set of agents, and let $\mathbf{P}$ be a finite set of propositions describing the facts about the world. Let $Bool(\mathbf{P})$ denote the set of all Boolean formulas over $\mathbf{P}$. To set up the semantics we first define a Kripke model in the usual sense, which models agents' epistemic uncertainties regarding the actual state of the world.

Definition 1 (Epistemic model). *An* epistemic model *$\mathcal{M}_e$ is a triple $\langle S, \sim, V\rangle$: S is a non-empty domain of states, $\sim$ stands for the set of accessibility (equivalence) relations $\{\sim_i \mid i \in \mathbf{I}\}$, $V : S \to \mathcal{P}(\mathbf{P})$ is a* valuation *assigning to each state a set of propositional variables (those that are 'true in that state').*

We will introduce the concept of epistemic observation models based on Kripke models, which captures the expected observations of agents. Agents observe what is happening around them and reason based on these observations. One way of expressing such observations is by means of 'actions', viz. the action of making statements, going to the right, nodding your head and many others. To this end, we introduce a finite set of actions, viz. $\mathbf{\Sigma}$. An observation is a finite string of actions, e.g., *abcd*. Note that an agent may expect different (even infinitely many) potential observations to happen at a given a state, e.g. she expects $a \dots ab$ to happen for any finite sequence of a preceding the terminating action b. As human beings and computers are essentially finite, we need to denote and talk about such expectations in a finitary way. To this end, we introduce the observation expressions (as regular expressions over $\mathbf{\Sigma}$):

Definition 2 (Observation expressions). *The language $\mathcal{L}_{obs}$ of* observations *is given by*

$$\pi \quad ::= \quad \varepsilon \mid \delta \mid a \mid \pi \cdot \pi \mid \pi + \pi \mid \pi^*$$

where $a \in \mathbf{\Sigma}$, and ε and δ are constants for the empty string and the empty language respectively.

The semantics for the observation expressions are given by sets of observations (strings over $\mathbf{\Sigma}$), similar to that for the regular expressions.

Definition 3 (Observations). *Given an observation expression π, the corresponding* set of observations, *denoted by $\mathcal{L}(\pi)$, is the set of finite strings over $\mathbf{\Sigma}$ defined as follows.*

$\mathcal{L}(\varepsilon) = \{\epsilon\} \quad \mathcal{L}(\delta) = \emptyset \quad \mathcal{L}(a) = \{a\}$

$\mathcal{L}(\pi \cdot \pi') = \{wv \mid w \in \mathcal{L}(\pi) \text{ and } v \in \mathcal{L}(\pi')\}$

$\mathcal{L}(\pi + \pi') = \mathcal{L}(\pi) \cup \mathcal{L}(\pi')$

$\mathcal{L}(\pi^*) = \{\epsilon\} \cup \bigcup_{n>0}(\mathcal{L}(\underbrace{\pi \cdots \pi}_{n}))$

Definition 4 (Epistemic observation model). *An* epistemic observation model $\mathcal{M}_{obs}$ *is a quadruple*

$$\langle S, \sim, V, Obs \rangle,$$

where $\langle S, \sim, V \rangle$ *is an epistemic model (the* epistemic skeleton *of* $\mathcal{M}_{obs}$*) and* $Obs : S \to \mathcal{L}_{obs}$ *is an observation function assigning to each state an observation expression* π *such that* $\mathcal{L}(\pi) \neq \emptyset$ *(non-empty set of finite sequences of observations). An* epistemic observation state *is a pointed epistemic observation model. Intuitively, Obs assigns to each state a set of potential or expected observations.*

Given an epistemic observation model $\mathcal{M}_{obs} = \langle S, \sim, V, Obs \rangle$, note that $\langle S, \sim, V \rangle$ is an epistemic model in the usual sense. Hence, sometimes, we also denote an observation model as $(\mathcal{M}_e, Obs)$, where $\mathcal{M}_e$ is the corresponding epistemic model. An epistemic model $\mathcal{M}_e$ can be considered as an epistemic observation model $\mathcal{M}_{obs}$ where for all $s \in S$, $Obs(s) = \mathbf{\Sigma}^*$(shorthand for $(a_0 + a_1 + \cdots + a_k)^*$ where $\{a_0, \ldots, a_k\} = \mathbf{\Sigma}$), that is, in an epistemic model the observations possible at each state are not specified; one can observe anything. In this sense, $\mathcal{M}_e$ lacks in providing certain information about the world, and $\mathcal{M}_{obs}$ fills up that gap. In what follows we often leave out the subscripts, whenever the respective models are clear from the context.

Example 5 (Dutch or not Dutch). *In the Netherlands, people often greet each other by kissing three times on the cheek (left-right-left) while in the rest of Europe people usually kiss each other only once or twice. We can reason whether a person is 'Dutch-related' by observing his behavior. Let* p_D *be the proposition meaning someone is Dutch-related;* a *and* b *are two actions denoting kissing the left cheek and kissing the right cheek, respectively. The following model is what we expect (reflexive arrows are omitted):*

$$s : p_D(a \cdot b \cdot a) \overset{1}{\text{——}} t : \neg p_D(a \cdot b)$$

The indistinguishability relation above depicts that agent 1 does not know whether p_D. The associated observations are those that the agents might expect in each state. Intuitively if we observe someone kissing three times (observation $a \cdot b \cdot a$), then we can infer that he or she is Dutch-related. In the next section a simple logic is defined to handle such reasoning based on actual observations.

2.2 Public Observation Logic

In this subsection we define a simple dynamic logic with knowledge operators to reason about knowledge via the matching of observations and expectations. The idea is similar to the one behind public announcement logic where people update their information by deleting the impossible scenarios according to what is publicly announced. Here we relax the link between the meaning and public actions (like an announcement), and assume that when observing an action, people delete some impossible scenarios where they wouldn't expect such an observation to happen.To make such reasoning formal, we first define the update of observation models according to some observation w. The idea behind $\mathcal{M}|_w$ is that we delete the states where the observed execution could not have been performed.

Definition 6 (Update by observation). *Let* w *be an observation over* $\mathbf{\Sigma}$*, let* $\mathcal{M} = (S, \sim, V, Obs)$ *be an observation model. The updated model* $\mathcal{M}|_w = (S', \sim', V', Obs')$*. Here,* $S' = \{s \mid \mathcal{L}(Obs(s)\backslash w) \neq \emptyset\}$, $\sim'_i = \sim_i|_{S' \times \mathbf{I} \times S'}$, $V' = V|_{S'}$*, and* $Obs'(s) = Obs(s)\backslash w$*, where* $\pi\backslash w$ *is defined as the regular expression denoting the language* $\{v \mid wv \in \mathcal{L}(\pi)\}$*.*

$\pi\backslash w$ is a regular language [5] and can be axiomatized with the output function o from the set of regular expressions over $\mathbf{\Sigma}$ to $\{\delta, \varepsilon\}$ as follows (cf. [3, 5]):

$$\begin{array}{l}
\pi\backslash a_0 \ldots a_n = \pi\backslash a_0\backslash a_1 \ldots \backslash a_n \\
\pi = o(\pi) + \sum_{a \in \mathbf{\Sigma}}(a \cdot \pi\backslash a) \\
\varepsilon\backslash a = \delta\backslash a = b\backslash a = \delta \quad (a \neq b) \\
a\backslash a = \varepsilon \\
(\pi \cdot \pi')\backslash a = (\pi\backslash a) \cdot \pi' + o(\pi) \cdot (\pi'\backslash a) \\
(\pi + \pi')\backslash a = \pi\backslash a + \pi'\backslash a \\
\pi^*\backslash a = \pi\backslash a \cdot \pi^* \\
o(\pi \cdot \pi) = o(\pi) \cdot o(\pi') \\
o(\pi^*) = \varepsilon \\
o(\varepsilon) = \varepsilon \\
o(\delta) = o(a) = \delta \\
o(\pi + \pi') = o(\pi) + o(\pi')
\end{array}$$

These are used for the computation of observations in a syntactic way.

We design a logic to reason about the observations, *Public observation logic* (*POL*):

Definition 7 (public observation logic). *The* formulas φ *of POL are given by:*

$$\varphi \quad ::= \quad \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid [\pi]\varphi$$

where $p \in \mathbf{P}$*,* $i \in \mathbf{I}$*, and* $\pi \in \mathcal{L}_{obs}$*.*

Definition 8 (Truth definition for POL). *Given an epistemic observation model* $\mathcal{M} = (S, \sim, V, Obs)$*, a state* $s \in S$*, and a POL-formula* φ*, the truth of* φ *at* s*, denoted by* $\mathcal{M}, s \vDash \varphi$*, is defined as follows:*

$$\begin{array}{rcl}
\mathcal{M}, s \vDash p & \Leftrightarrow & p \in V(s) \\
\mathcal{M}, s \vDash \neg\varphi & \Leftrightarrow & \mathcal{M}, s \nvDash \varphi \\
\mathcal{M}, s \vDash \varphi \wedge \psi & \Leftrightarrow & \mathcal{M}, s \vDash \varphi \text{ and } \mathcal{M}, s \vDash \psi \\
\mathcal{M}, s \vDash K_i\varphi & \Leftrightarrow & \text{for all } t : s \sim_i t \implies \mathcal{M}, s \vDash \varphi \\
\mathcal{M}, s \vDash [\pi]\varphi & \Leftrightarrow & \text{for each } w \in \mathcal{L}(\pi) : (w \in init(Obs(s)) \\
 & & \quad \text{implies } \mathcal{M}|_w, s \vDash \varphi)
\end{array}$$

where $w \in init(\pi)$ *iff* $\exists v \in \Sigma^*$ *such that* $wv \in \mathcal{L}(\pi)$ *iff* $\mathcal{L}(\pi\backslash w) \neq \emptyset$.

Consider the model $\mathcal{M}$ in Example 5. If we observe one or two kisses, we still cannot tell whether the person is Dutch-related, but if there is one more kiss to follow then we know. Formally, it can be verified that $\mathcal{M}, s \vDash [a \cdot b](\neg Kp_D \wedge [a]Kp_D)$. More complicated π can be used to express (infinite) sets of observations, e.g., $[\mathbf{\Sigma}^* \cdot a \cdot \mathbf{\Sigma}^*]K_i\varphi$ says 'as long as a is observed at some point, i knows φ'.

Clearly, standard bisimulation between epistemic models is not an invariance of the above logic: POL can reason about what may happen at each state. We now define bisimulation between observation models, which facilitates characterization results in later sections.

Definition 9 (Observation Bisimulation). *A binary relation R between the domains of two observation models $\mathcal{M} = (S, \sim, V, Obs)$ and $\mathcal{N} = (S', \sim', V', Obs')$ is called a bisimulation if for any $s \in S, s' \in S'$: $(s, s') \in R$ implies that the following conditions hold:*

Propositional Invariance $V(s) = V'(s')$

Observational Invariance $\mathcal{L}(Obs(s)) = \mathcal{L}(Obs'(s'))$.

Zig *if $s \sim_i t$ in $\mathcal{M}$ then there exists a t' in $\mathcal{N}$ such that $s' \sim_i t'$ and tRt'.*

Zag *if $s' \sim_i t'$ in $\mathcal{N}$ then there exists a t in $\mathcal{M}$ such that $s \sim_i t$ and tRt'.*

$\mathcal{M}$ and $\mathcal{N}$ are said to be bisimilar ($\mathcal{M} \leftrightarroweq_o \mathcal{N}$) if there is a bisimilation R between them. $(\mathcal{M}, s)$ and $(\mathcal{N}, s')$ are said to be bisimilar ($\mathcal{M}, s \leftrightarroweq_o \mathcal{N}, s'$) if there is a bisimulation R between them such that $(s, s') \in R$.

Note that the standard bisimulation (notation $\leftrightarroweq$) is defined as $\leftrightarroweq_o$ without the condition for the invariance for observations. It is not hard to show the following.

Proposition 10 (Bisimulation Invariance). *For any two finite epistemic observation states $\mathcal{M}, s$ and $\mathcal{N}, s'$, the following statements are equivalent:*

- $\mathcal{M}, s \leftrightarroweq_o \mathcal{N}, s'$
- *For any formula $\varphi \in POL : \mathcal{M}, s \vDash \varphi \iff \mathcal{N}, s' \vDash \varphi$*

Proof. [$\leftrightarroweq_o \implies \equiv_{POL}$]: by induction on φ: Boolean and $K_i\psi$ cases are trivial. Now consider $\varphi = [\pi]\psi$: suppose $\mathcal{M}, s \leftrightarroweq_o \mathcal{N}, s'$ but $\mathcal{M}, s \vDash [\pi]\psi$ and $\mathcal{N}, s' \nvDash [\pi]\psi$. Then there exists a $w \in \mathcal{L}(\pi)$ such that $w \in init(Obs(s'))$ and $\mathcal{N}|_w, s' \vDash \neg\psi$.

By the definition of $\leftrightarroweq_o$, $\mathcal{L}(Obs(s)) = \mathcal{L}(Obs(s'))$ therefore $w \in init(Obs(s))$. Thus $\mathcal{M}|_w, s$ exists. We now show that $\mathcal{M}|_w, s \leftrightarroweq_o \mathcal{N}|_w, s'$. Let R be $\{(t, t') \in S_{\mathcal{M}|_w} \times S_{\mathcal{N}|_w} \mid \mathcal{M}, t \leftrightarroweq_o \mathcal{N}, t'\}$. Clearly $(s, s') \in R$. Note that if $\mathcal{L}(Obs(t)) = \mathcal{L}(Obs(t'))$ then $\mathcal{L}(Obs(t)\backslash w) = \mathcal{L}(Obs(t')\backslash w)$; this proves the invariance for observations. Based on this invariance, it is not hard to verify that R is indeed an observation bisimulation between $\mathcal{M}|_w$ and $\mathcal{N}|_w$.

Since $\mathcal{M}|_w, s \leftrightarroweq_o \mathcal{N}|_w, s'$, by induction hypothesis

$$\mathcal{M}|_w, s \vDash \neg\psi.$$

Clearly, this contradicts the assumption that $\mathcal{M}, s \vDash [\pi]\psi$.

[$\equiv_{POL} \implies \leftrightarroweq_o$]: Let $R = \{(t, t') \in S_{\mathcal{M}} \times S_{\mathcal{N}} \mid \mathcal{M}, t \equiv_{POL} \mathcal{N}, t'\}$. We can show that R is an observation bisimulation. All the conditions are standard and thus can be handled by standard techniques except the new clause about the invariance for observations: we need to show $tRt' \implies \mathcal{L}(Obs(t)) = \mathcal{L}(Obs(t'))$; however, this is trivial since in the language of POL we can express $\langle w\rangle\top$ such that $M, t \vDash \langle w\rangle\top \iff w \in \mathcal{L}(Obs(t))$. $\square$

Intuitively, these observation models can be seen as compact representations of certain epistemic temporal models [15, 19]. To make the link more precise, we can relate *POL* on observation models to the same language on epistemic temporal models with the usual *PDL*-style interpretation of $[\pi]\varphi$ formulas.

Definition 11. *Let $\mathcal{M}$ be an epistemic observation model $\langle S, \sim_i, V, Obs\rangle$. The $\mathcal{M}$-generated epistemic temporal model is defined as $ETL(\mathcal{M}) = \langle H, \xrightarrow{a}, \sim'_i, V'\rangle$ where: $H = \{(s, w) \mid s \in S, w = \epsilon$ or $w \in \mathcal{L}(Obs(s))\}$; $(s, w) \xrightarrow{a} (t, v) \iff s = t$ and $v = wa, a \in \mathbf{\Sigma}$; $(s, w) \sim_i (t, v) \iff s \sim_i t$ and $w = v$; $p \in V'(s, w) \iff p \in V(s)$.*

The formula $[\pi]\varphi$ is true at a pointed epistemic temporal model $\mathcal{N}, h$ iff for any $w \in \mathcal{L}(\pi)$, $h \xrightarrow{w} h'$ implies $\mathcal{N}, h \vDash \varphi$. The truth definitions for observation-free formulas are as usual. We call this logic *EPDL* (Epistemic-*PDL*). To establish the precise link between observation models and epistemic temporal models, we can prove the following.

Proposition 12. *Given a pointed POL model $\mathcal{M}, s$, and a POL formula φ, it can be shown that:*
$\mathcal{M}, s \vDash \varphi \iff ETL(\mathcal{M}), (s, \epsilon) \vDash_{EPDL} \varphi$.

Proof. We need to show for any observation model $\mathcal{M}, s$ and any *POL* formula φ:

$$\mathcal{M}, s \vDash \varphi \iff ETL(\mathcal{M}), (s, \epsilon) \vDash_{EPDL} \varphi$$

We prove this by induction on φ.

The Boolean case and the $K_i\psi$ case are trivial. Now consider the case $[\pi]\psi$. Suppose without loss of generality that there is an observation model $\mathcal{M}, s \vDash [\pi]\psi$ and $ETL(\mathcal{M}), (s, \epsilon) \nvDash [\pi]\psi$. Then there exists a $w \in \mathcal{L}(\pi)$ such that $ETL(\mathcal{M}), (s, w) \nvDash \psi$. By the definition of $ETL(\mathcal{M})$, $w \in Obs(s)$ thus $\mathcal{M}|_w$ exists. Based on the definition of $ETL(\mathcal{M})$, it is not hard to show that $ETL(\mathcal{M}|_w), (s, \epsilon)$ is bisimilar (w.r.t. both $\sim$ and $\to$) to $ETL(\mathcal{M}), (s, w)$. Since EPDL is clearly invariant under bisimulation,

$$ETL(\mathcal{M}|_w), (s, \epsilon) \vDash \neg\psi.$$

By induction hypothesis, $\mathcal{M}_w, s \vDash \neg\psi$ which contradicts the assumption that $\mathcal{M}, s \vDash [\pi]\psi$. $\square$

Note that the generated epistemic temporal models can be infinite, and thus the above result does not give a straightforward model checking procedure for *POL*. According to the semantics of $[\pi]\varphi$ we need to check infinitely many $w \in \mathcal{L}(\pi)$. Fortunately, this can be handled by partitioning $\mathcal{L}(\pi)$ into a finite number of regular expressions $\pi_0 \ldots \pi_k$ such that for any $0 \le i \le k$, any $w, v \in \mathcal{L}(\pi_i)$, $\mathcal{M}|w = \mathcal{M}|v$, providing decidability of model checking after all (see [28] for details in a similar setting).

3. EXPECTATION COMES FROM PROTOCOLS

Observation models describe the agents' expected observations, which in turn influence their reasoning. We investigate how agents acquire and change their expectations, by looking at protocols and protocol models as sources for the expected observations.

3.1 Protocol models

A protocol is a rule telling us what we should do under what conditions. We can specify protocols in the following language of protocol expressions $\mathcal{L}_{prot}$:

Definition 13 (Protocol expression). *The language $\mathcal{L}_{prot}$ of* protocols *is given by*

$$\eta \quad ::= \quad \varepsilon \mid \delta \mid a \mid ?\varphi \mid \eta \cdot \eta \mid \eta + \eta \mid \eta^*$$

where $\varphi \in Bool(\mathbf{P})$.

The above language of protocol expressions is obtained by adding Boolean tests to observation expressions. For example, $(?love \cdot stay)^* \cdot (?\neg love \cdot separate)$ expresses 'we should stay together as long as we are in love'. For a discussion on more complicated test scenarios (e.g. considering agents' knowledge) see Section 5.

In the story of Example 5, there seems to be an underlying protocol: *if* you are Dutch *then* you kiss three times and *if* you are non-Dutch *then* you kiss two times. It is the reason for the agent to have the corresponding expectations of the observations. This protocol (call it π_K) can be expressed as $?p_D \cdot a \cdot b \cdot a + ?\neg p_D \cdot a \cdot b$. We would like to generate the observation model in Example 5 from the protocol π_K and the epistemic model

$$p_D \text{---}1\text{---} \neg p_D$$

Intuitively, the information of the protocol π_K can be incorporated by adding to each state the possible observations allowed by the protocol. We now move on to the technical details.

To compute the expected observations corresponding to a given protocol we first define the semantics of protocol expressions. Intuitively, we associate to each protocol η a set $\mathcal{L}_g(\eta)$ of *conditional observations* in the form of

$$\rho_0 a_0 \rho_1 a_1 \ldots \rho_k a_k$$

where each $\rho_i \subseteq \mathbf{P}$ denotes a state of affairs (the basic propositions $p \in \rho$ are true while the others are false), encoding the conditions for the later observations to happen. For Boolean formulas φ, we write $\rho \vDash \varphi$ if φ is true under ρ (viewed as a valuation).

Definition 14. *The set of conditional observations corresponding to η is defined as follows:*

$$\begin{array}{l}
\mathcal{L}_g(\delta) = \emptyset, \quad \mathcal{L}_g(\varepsilon) = \{\rho \mid \rho \subseteq \mathbf{P}\}, \\
\mathcal{L}_g(a) = \{\rho a \rho \mid \rho \subseteq \mathbf{P}\}, \quad \mathcal{L}_g(?\psi) = \{\rho \mid \rho \vDash \psi\}, \\
\mathcal{L}_g(\eta_1 \cdot \eta_2) = \{w \diamond v \mid w \in \mathcal{L}_g(\eta_1), v \in \mathcal{L}_g(\eta_2)\}, \\
\mathcal{L}_g(\eta_1 + \eta_2) = \mathcal{L}_g(\eta_1) \cup \mathcal{L}_g(\eta_2), \\
\mathcal{L}_g(\eta^*) = \{\rho \mid \rho \subseteq \mathbf{P}\} \cup \bigcup_{n>0}(\mathcal{L}_g(\eta^n)),
\end{array}$$

where $\diamond$ is the fusion product*: $w \diamond v = w'\rho v'$ when $w = w'\rho$ and $v = \rho v'$, and not defined otherwise.*

Note that the ρ_i's in a conditional observation remain unchanged since no *factual change* is introduced by the execution of the actions (see Section 4 for a detailed discussion of fact changing actions). In the following we show how to derive the set of observations to be expected under the same condition ρ according to η, by defining the conversion function $f_\rho : \mathcal{L}_{prot} \to \mathcal{L}_{obs}$,

$$\begin{array}{lcllcl}
f_\rho(\varepsilon) & = & \varepsilon & f_\rho(\delta) & = & \delta \\
f_\rho(a) & = & a & f_\rho(?\varphi) & = & \varepsilon \text{ if } \rho \models \varphi \\
f_\rho(\eta \cdot \eta') & = & f_\rho(\eta) \cdot f_\rho(\eta') & & & \delta \text{ if } \rho \not\models \varphi \\
f_\rho(\eta + \eta') & = & f_\rho(\eta) + f_\rho(\eta') & f_\rho(\eta^*) & = & f_\rho(\eta)^*
\end{array}$$

Proposition 15. *For any $\eta \in \mathcal{L}_{prot}$, $\mathcal{L}(f_\rho(\eta)) = \{w \mid w = a_0 \ldots a_k,$ where $a_i \in \mathbf{\Sigma} \cup \{\varepsilon\}$ and $\rho a_0 \rho a_1 \ldots a_k \rho \in \mathcal{L}_g(\eta)\}$.*

Therefore, every η has a normal form

$$\eta^\circ = \sum_{\rho \subseteq \mathbf{P}} (?\varphi_\rho \cdot f_\rho(\eta))$$

such that $\mathcal{L}_g(\eta) = \mathcal{L}_g(\eta^\circ)$, where φ_ρ is a characteristic formulas for $\rho \subseteq \mathbf{P}$ (e.g. $\varphi_{\{p\}} = p \wedge \neg q$ if $\mathbf{P} = \{p, q\}$).

Proof. We first show that $\mathcal{L}(f_\rho(\eta))$ is equal to

$$\{w \mid w = a_0 \ldots a_k, a_i \in \mathbf{\Sigma} \cup \{\epsilon\} \text{ and } \rho a_0 \rho a_1 \ldots a_k \rho \in \mathcal{L}_g(\eta)\}$$

. We prove this by induction on $\eta \in \mathcal{L}_{Prot}$.

The atomic cases are straightforward. Now we check the complex cases:

$\eta = \eta_1 + \eta_2$:

$$\begin{array}{cl}
 & \mathcal{L}(f_\rho(\eta)) = \mathcal{L}(f_\rho(\eta_1 + \eta_2)) = \mathcal{L}(f_\rho(\eta_1) + f_\rho(\eta_2)) \\
= & \mathcal{L}(f_\rho(\eta_1)) \cup \mathcal{L}(f_\rho(\eta_2)) \\
= & \{w \mid w = a_0 \ldots a_k, \text{ and } \rho a_0 \rho \ldots \rho a_k \rho \in \mathcal{L}_g(\eta_1)\} \cup \\
 & \{w \mid w = a_0 \ldots a_k, \text{ and } \rho a_0 \ldots a_k \rho \in \mathcal{L}_g(\eta_2)\} \text{(by IH)} \\
= & \{w \mid w = a_0 \ldots a_k, \text{ and } \rho a_0 \ldots a_k \rho \in \mathcal{L}_g(\eta_1 + \eta_2)\}
\end{array}$$

$\eta = \eta_1 \cdot \eta_2$:

$$\begin{array}{cl}
 & \mathcal{L}(f_\rho(\eta)) = \mathcal{L}(f_\rho(\eta_1 \cdot \eta_2)) = \mathcal{L}(f_\rho(\eta_1) \cdot f_\rho(\eta_2)) \\
= & \{wv \mid w \in \mathcal{L}(f_\rho(\eta_1) \text{ and } v \in \mathcal{L}(f_\rho(\eta_2)\} \\
= & \{wv \mid w = c_0 \ldots c_m \text{ st. } \rho c_0 \ldots c_m \rho \in \mathcal{L}_g(\eta_1) \\
 & \text{ and } v = b_0 \ldots b_n \text{ st. } \rho b_0 \ldots b_n \rho \in \mathcal{L}_g(\eta_2)\} \text{(by IH)} \\
= & \{u \mid u = a_0 \ldots a_k, \text{ and } \rho a_0 \ldots a_k \rho \in \mathcal{L}_g(\eta_1 \cdot \eta_2)\} \\
 & \text{(by fusion product)}
\end{array}$$

$\eta = \eta_1^*$:

$$\begin{array}{cl}
 & \mathcal{L}(f_\rho(\eta)) = \mathcal{L}(f_\rho(\eta_1^*)) = \mathcal{L}((f_\rho(\eta_1))^*) \\
= & \{\epsilon\} \cup \bigcup_{n>0} \mathcal{L}((f_\rho(\eta_1))^n) \\
= & \{u \mid u = a_0 \ldots a_k, \text{ and } \rho a_0 \ldots a_k \rho \in \{\rho \mid \rho \subseteq \mathbf{P}\} \\
 & \cup \bigcup_{n>0} \mathcal{L}_g(\eta_1^n)\} \text{(by IH)} \\
= & \{u \mid u = a_0 \ldots a_k, \text{ and } \rho a_0 \ldots a_k \rho \in \mathcal{L}_g(\eta_1^*)\}
\end{array}$$

This completes the proof for the following statement : For all η in $\mathcal{L}_{prot}$, for all $\rho \subseteq \mathbf{P}$, $\mathcal{L}(f_\rho(\eta)) = \{w \mid w = a_0 \ldots a_k$, where $a_i \in \mathbf{\Sigma} \cup \{\epsilon\}$ and $\rho a_0 \rho a_1 \ldots a_k \rho \in \mathcal{L}_g(\eta)\}$.

From the result above and the definition of $\mathcal{L}_g$, it follows,

$\mathcal{L}_g(f_\rho(\eta)) = \{\rho' a_0 \ldots a_k \rho' \mid \rho' \subseteq \mathbf{P}$ and $\rho a_0 \ldots a_k \rho \in \mathcal{L}_g(\eta)\}$.

Let $G^\eta_\rho = \{\rho a_0 \rho a_1 \ldots a_k \rho \mid \rho a_0 \rho a_1 \ldots a_k \rho \in \mathcal{L}_g(\eta)\}$, the set of all ρ-guarded expressions in $\mathcal{L}_g(\eta)$. Then, by fusion product, it follows that $\mathcal{L}_g(?\varphi_\rho \cdot f_\rho(\eta)) = G^\eta_\rho$. Thus,

$\mathcal{L}_g(\eta^\circ) = \mathcal{L}_g(\sum_{\rho \subseteq \mathbf{P}}(?\varphi_\rho \cdot f_\rho(\eta))) = \bigcup_{\rho \subseteq \mathbf{P}} \mathcal{L}_g(?\varphi_\rho \cdot f_\rho(\eta)) = \bigcup_{\rho \subseteq \mathbf{P}} G^\eta_\rho = \mathcal{L}_g(\eta)$. $\Box$

From Proposition 15, according to the protocol η, the expected observations on a state s in an epistemic model $\mathcal{M}$ can be computed by $f_{V_M(s)}(\eta)$. For example, $f_{\{p\}}(?p \cdot a + ?\neg p \cdot b) = a$. However, not every observation model can be generated by a single protocol.

Example 16. *Consider the observation model:*

$$s : p(a) \text{ —}1,2\text{— } t : p(b)$$

We cannot associate a protocol η to its epistemic skeleton such that $f_{V_M(s)}(\eta) = a$ and $f_{V_M(t)}(\eta) = b$, since $V_\mathcal{M}(s) = V_\mathcal{M}(t)$. Note that taking $?p(a + b)$ for η does not work.

In this example agents are uncertain about the protocol. Their uncertainty may be seen as follows:

$$?p \cdot a \text{ —}1,2\text{— } ?p \cdot b$$

The following definition formalizes this idea.

Definition 17 (Epistemic Protocol model). *An epistemic protocol model $\mathcal{A}$ is a triple $\langle T, \sim, Prot\rangle$ where T is a domain of abstract objects, $\sim$ stands for a set of accessibility (equivalence) relations $\{\sim_i \mid i \in \mathbf{I}\}$, and $Prot : T \to \mathcal{L}_{prot}$ assigns to each domain object a protocol. We call a pointed epistemic protocol model an* epistemic protocol *and a singleton epistemic protocol model a* public protocol.

Note that public protocols are (implicitly) commonly known by all the agents.

We will now proceed towards our main result in this section, namely that an epistemic observation state uniquely determines an epistemic protocol, and that an epistemic protocol and an epistemic state uniquely determine an epistemic observation state. To show the correspondence, we need one more semantic operation, that is a *modal product operation* of an epistemic observation model and a protocol model. It formalizes the change in possible observations induced by a protocol. We should see this definition as installing a new protocol, by means of *novel* observations, into the epistemic observation model, and thus completely obliterating the *current* expected observations.

Definition 18 (Protocol Update). *Given an epistemic observation model $\mathcal{M}_{obs} = \langle S, \sim, V, Obs\rangle$ and an epistemic protocol model $\mathcal{A} = \langle T, \sim, Prot\rangle$. We define the product $(\mathcal{M}_{obs} \otimes \mathcal{A}) = (S', \sim', V', Obs')$ as follows:*

- $S' = \{(s, t) \in S \times T : \mathcal{L}(f_{V_{\mathcal{M}}(s)}(Prot(t))) \neq \emptyset\}$;
- $(s, t) \sim'_i (s', t')$ *iff* $s \sim_i s'$ *in* $\mathcal{M}_{obs}$ *and* $t \sim_i t'$ *in* $\mathcal{A}$;
- $V'(s, t) = V(s)$;
- $Obs'((s, t)) = f_{V_{\mathcal{M}}(s)}(Prot(t))$.

We mentioned that epistemic models can be seen as special cases of epistemic observation models, namely with the 'anything goes' protocol. Therefore, also in that case the product operation between an epistemic model and a protocol model corresponds to the installation of a protocol.

We now illustrate the definition by the two example scenarios of the introduction. In the pictures, assume reflexivity and transitivity of access. In the first scenario, at the beginning neither of Jane or Ann knows the fact g (Kate is gay). However, one of them, Jane, is aware of the protocol that: **if** Kate is gay **then** she will make the statement 'I am musical, I like Kathleen Ferrier's voice' (action a); and **if** she is not gay, **then** she will talk about something else (action b). However, Ann has no idea whether a and b can carry such information. The scenario is modeled as follows, where the last model is the observation model resulting from the update of the protocol on the first epistemic model ('real states' are underlined):

$$\begin{array}{c} \underline{g(\mathbf{\Sigma}^*)} \\ \Big|\,{\scriptstyle \text{Jane,Ann}} \\ \neg g(\mathbf{\Sigma}^*) \end{array} \otimes \begin{array}{c} \underline{?g \cdot a + ?\neg g \cdot b} \\ \Big|\,{\scriptstyle \text{Ann}} \\ a + b \end{array} = \begin{array}{ccc} \underline{g(a)} & \text{—}{\scriptstyle \text{Jane,Ann}}\text{—} & \neg g(b) \\ \Big|\,{\scriptstyle \text{Ann}} & & \Big|\,{\scriptstyle \text{Ann}} \\ g(a+b) & {\scriptstyle \text{Jane,Ann}} & \neg g(a+b) \end{array}$$

where g denotes the fact that 'Kate is gay', a denotes the observation of Kate making the 'musical statement' and b stands for Kate saying something else.

We now consider the second example. After Carl's first description of the night of the Valentine's day, Ben and Alice still do not know what has happened. This prompts Alice's question. Now, the wink from Carl creates an uncertainty in Alice regarding how to interpret Ben's statements, while knowing Carl so well, Ben immediately gets the idea of the protocol Carl is using. The modeling is as follows:

$$\begin{array}{c} p(\mathbf{\Sigma}^*) \\ \Big|\,{\scriptstyle \text{Ben,Alice}} \\ \underline{\neg p(\mathbf{\Sigma}^*)} \end{array} \otimes \begin{array}{c} \underline{?p \cdot Y + ?\neg p \cdot N} \\ \Big|\,{\scriptstyle \text{Alice}} \\ ?\neg p \cdot Y + ?p \cdot N \end{array} = \begin{array}{ccc} p(Y) & {\scriptstyle \text{Ben,Alice}} & \underline{\neg p(N)} \\ \Big|\,{\scriptstyle \text{Alice}} & & \Big|\,{\scriptstyle \text{Alice}} \\ \neg p(Y) & {\scriptstyle \text{Ben,Alice}} & p(N) \end{array}$$

where p denotes the fact that 'Something has happened involving Mike and Sara on V-day night', 'Y' corresponds to Carl's saying something in the affirmative to Alice's question, and 'N' the opposite.

According to our definition, an epistemic protocol model acts on an epistemic model determining a unique observation model. In the rest of this section we will investigate the converse: whether an arbitrary observation model can be generated by updating an epistemic model by an epistemic protocol model.

Proposition 19. *Given an epistemic observation model $\mathcal{M} = (\mathcal{N}, Obs)$, there is an epistemic model $\mathcal{N}'$ and a protocol model $\mathcal{A}$ such that $\mathcal{M} \leftrightarrows_o \mathcal{N}' \otimes \mathcal{A}$.*

Proof. Let $\mathcal{N}' = (S', \sim', V')$ be the universal ignorance model, i.e., $S' = \mathcal{P}(\mathbf{P})$, for each $i, \sim'_i = S' \times S'$, and $V(\rho) = \rho \subseteq \mathbf{P}$. Given $\mathcal{M} = (S, \sim, V, Obs)$, let $A = (S, \sim, Prot)$ such that $Prot(s) = ?\varphi_{V(s)} \cdot Obs(s)$. where $\varphi_{V(s)}$ is the characteristic formula of $V(s) \subseteq \mathbf{P}$ (e.g. $p \wedge \neg q$ is a characteristic formula for $\{p\}$ if $\mathbf{P} = \{p, q\}$). Now we show $\mathcal{M} \leftrightarrows_o \mathcal{N}' \otimes \mathcal{A}$ by proving that $R = \{(s, (\rho, s)) \mid V(s) = \rho\}$ is a bisimulation relation.

The invariance conditions are immediate. Now suppose $s \sim_i t$ in $\mathcal{M}$ then $(\rho, s) \sim_i (V(t), t)$ in $N' \times A$ by the definition of the product. Obviously, $tR(\rho', t)$, where $\rho' = V(t)$.

Suppose $(\rho, s) \sim_i (\rho', t)$. Then $V(t) = \rho'$. Therefore $s \sim_i t$ and $tR(\rho', t)$. $\square$

This result shows that every observation model is reasonable in the sense that it can be generated from an epistemic model by some epistemic protocol model. Note that in the above proposition, we consider an arbitrary epistemic model. However, it is more intuitive to consider the particular epistemic model $\mathcal{N}$ in $\mathcal{M} = (\mathcal{N}, Obs)$, and ask if there is a protocol model $\mathcal{A}$ such that $\mathcal{N} \otimes \mathcal{A} \leftrightarrows_o \mathcal{M}$. For singleton protocol models, we have a characterization result.

Definition 20. *An observation model $\mathcal{M}$ is said to be Boolean normal if for any two worlds s, t in it, $V_{\mathcal{M}}(s) = V_{\mathcal{M}}(t) \implies \mathcal{L}(Obs(s)) = \mathcal{L}(Obs(t))$.*

Theorem 21. *Given an epistemic observation model $\mathcal{M} = (\mathcal{N}, Obs)$, $\mathcal{M}$ is Boolean normal iff there exists a singleton protocol model $\mathcal{A}$ such that $\mathcal{N} \otimes \mathcal{A} \leftrightarrows_o \mathcal{M}$, $\leftrightarrows_o$ being a total bisimulation.*

Proof. $\Rightarrow$: Let φ_s be the Boolean characterization formula corresponding to $V_{\mathcal{N}}(s)$. Let $\eta_{\mathcal{M}} = \sum_{s \text{ in } \mathcal{N}} ?\varphi_s \cdot Obs(s)$. Because of the finiteness of $\mathbf{P}$ and Boolean normality, $\eta_{\mathcal{M}}$ has a finite representation. Let $\mathcal{A}_{\eta_{\mathcal{M}}}$ be the singleton pointed protocol model with *Prot* assigning $\pi_{\mathcal{M}}$ to the single point. We can verify that $\mathcal{N} \otimes \mathcal{A}_{\pi_{\mathcal{M}}} \leftrightarroweq_o \mathcal{M}$.

$\Leftarrow$: suppose $\mathcal{M}$ is not Boolean normal then there are s, t in $\mathcal{M}$ such that $V(s) = V(t)$ and $Obs(s) \neq Obs(t)$. Due to the normal form of protocols, updating a single pointed protocol on s, t will result in the same observations. So there cannot be any single pointed protocol model to do the job. $\square$

Clearly, not every epistemic observation model is Boolean normal, thus not every observation model can be generated from a public protocol.

EXAMPLE 22. *Consider the following epistemic observation model $\mathcal{M}$, we will show that $\mathcal{M}$ cannot be generated by any epistemic protocol on its epistemic skeleton:*

$$p(b) \overset{1}{\text{---}} p(a) \overset{2}{\text{---}} \neg p(b)$$

Suppose towards contradiction that there is a protocol model $\mathcal{A}$ such that the execution of $\mathcal{A}$ on the epistemic skeleton of $\mathcal{M}$ gives an observation model which is bisimilar to $\mathcal{M}$. To compose the middle world in the observation model we need a state t in the protocol model such that $Prot(t)$ allows a to happen if p is true. Then t can be composed with the leftmost p world above as well, since the left world and middle world are Boolean indistinguishable. Therefore there will be a $p(a)$-world in the resulting model which cannot reach any $\neg p$ world in one step, due to the definition of $\otimes$ (the leftmost state above cannot reach any $\neg p$ world in one step).

This leads us to consider a subclass of the observation models given as follows.

DEFINITION 23 (BOOLEAN DISTINGUISHING). *An epistemic (observation) model $\mathcal{M}$ is said to be Boolean distinguishing if for each state $s \in \mathcal{M}$, there exists a Boolean distinguishing formula for s, that is, there is a Boolean formula which is only true at s and the states in $\mathcal{M}$, related by $\leftrightarroweq$ ($\leftrightarroweq_o$) to s.*

THEOREM 24. *Given an epistemic observation model $\mathcal{M} = (\mathcal{N}, Obs)$, if $\mathcal{N}$ is Boolean-distinguishing then there is a protocol model $\mathcal{A}$ such that $\mathcal{N} \otimes \mathcal{A} \leftrightarroweq_o \mathcal{M}$.*

PROOF. Suppose $\mathcal{N} = (W, \sim, V)$ and let $\varphi_s^{\mathcal{N}}$ be the Boolean distinguishing formula corresponding to $s \in W$. Let $\mathcal{A} = (W, \sim, Prot)$ where $Prot(s) = ?\varphi_s^{\mathcal{N}} \cdot Obs(s)$. We will show that $\mathcal{N} \otimes \mathcal{A} \leftrightarroweq_o \mathcal{M}$.

Let $R \subseteq W \times W_{\mathcal{N} \otimes \mathcal{A}}$ be the binary relation defined by setting $wR(v, t)$ iff $\mathcal{M}, w \leftrightarroweq_o \mathcal{M}, t$. We need to show that R is indeed an observation bisimulation.

Now suppose $wR(v, t)$. Since $Prot(t) = ?\varphi_t^{\mathcal{N}} \cdot Obs(t)$ and (v, t) is in $\mathcal{N} \otimes \mathcal{A}$ then $\mathcal{N}, v \vDash \varphi_t^{\mathcal{N}}$. Since $\varphi_t^{\mathcal{N}}$ is a Boolean-distinguishing formula for the world t, we have that $\mathcal{N}, v \leftrightarroweq \mathcal{N}, t$, since a state non-bisimilar to t will not satisfy $\varphi_t^{\mathcal{N}}$. Due to the fact that $\mathcal{M}, w \leftrightarroweq_o \mathcal{M}, t$ we have $\mathcal{N}, w \leftrightarroweq \mathcal{N}, t$, thus $\mathcal{N}, w \leftrightarroweq \mathcal{N}, v$. Therefore the propositional invariance condition of observation bisimulation holds. Since $\mathcal{M}, w \leftrightarroweq_o \mathcal{M}, t$, $Obs(w) = Obs(t) = Obs((v, t))$.

From the fact that $\mathcal{M}, w \leftrightarroweq_o \mathcal{M}, t$ and $\mathcal{N}, w \leftrightarroweq_e \mathcal{N}, v$ the conditions Zig and Zag of Definition 9 can be verified easily. $\square$

3.2 Equivalence of protocols

We motivated in the introduction that one observation model might be generated in different ways (even based on the same epistemic model). For example, consider the following model:

$$p(b) \overset{1,2}{\text{---}} \neg p(a)$$

It can be generated from its epistemic skeleton by updating a public protocol $?p \cdot b + ?\neg p \cdot a$ or the epistemic protocol model:

$$?p \cdot b \overset{1,2}{\text{---}} ?\neg p \cdot a$$

Actually, the announcement of $?p \cdot b + ?\neg p \cdot a$ will always yield the same result as the above epistemic protocol model on arbitrary epistemic models. On the other hand, the announcement $?p \cdot (a + b)$ has a different update result on the same epistemic model compared to the update of the following epistemic protocol:

$$?p \cdot a \overset{1,2}{\text{---}} ?p \cdot b$$

Such examples lead us to the following notion of equivalence between protocol models.

DEFINITION 25 (EFFECTIVE EQUIVALENCE). *Two protocol models $\mathcal{A}$ and $\mathcal{B}$ are said to be effective-equivalent (notation: $\mathcal{A} \equiv_{ef} \mathcal{B}$) if for any observation model $\mathcal{M}$: $\mathcal{M} \otimes \mathcal{A} \leftrightarroweq_o \mathcal{M} \otimes \mathcal{B}$.*

Inspired by the idea of action emulation in [24], we characterize the notion of effective-equivalence by the following structural equivalence. Let $\mathcal{L}^\rho(\eta)$ be $\mathcal{L}(f_\rho(\eta))$ (cf. Proposition 15).

DEFINITION 26 (PROTOCOL EMULATION). *Two protocol models $\mathcal{A} = (S, Prot)$ and $\mathcal{B} = (T, Prot)$ are said to be emulated (notation: $\mathcal{A} \approx \mathcal{B}$) if there is a binary relation $E \subseteq S \times T$ such that whenever sEt we have:*

- *there exists $\rho \subseteq \mathbf{P}$ such that $\mathcal{L}^\rho(Prot(t)) = \mathcal{L}^\rho(Prot(s))$.*
- *if $s \sim_i s'$ in $\mathcal{A}$ then there is a set $T' \subseteq T$ such that:*
  1. *for any $t' \in T'$: $t \sim_i t'$;*
  2. *for any $t' \in T'$: $s'Et'$;*
  3. *for any $\rho \subseteq \mathbf{P}$ such that $\mathcal{L}^\rho(Prot(s')) \neq \emptyset$ there exists $t' \in T'$ such that $\mathcal{L}^\rho(Prot(s')) = \mathcal{L}^\rho(Prot(t'))$*
- *if $t \sim_i t'$ in $\mathcal{B}$ then there is a set $S' \subseteq S$ such that:*
  1. *for any $s' \in S'$: $s \sim_i s'$;*
  2. *for any $s' \in S'$: $s'Et'$;*
  3. *for any $\rho \subseteq \mathbf{P}$ such that $\mathcal{L}^\rho(Prot(t')) \neq \emptyset$ there exists $s' \in S'$ such that $\mathcal{L}^\rho(Prot(s')) = \mathcal{L}^\rho(Prot(t'))$*

When restricted to public protocols, it is not hard to see that $\eta \approx \eta' \iff \mathcal{L}_g(\eta) = \mathcal{L}_g(\eta')$. In general, we have the following result.

THEOREM 27. *For any finite protocol models $\mathcal{A}$ and $\mathcal{B}$: $\mathcal{A} \equiv_{ef} \mathcal{B} \iff \mathcal{A} \approx \mathcal{B}$.*

PROOF. $\Leftarrow$: Suppose $\mathcal{A} \approx \mathcal{B}$, we need to show for any observation model $\mathcal{M}$: $\mathcal{M} \otimes \mathcal{A} \leftrightarroweq_o \mathcal{M} \otimes \mathcal{B}$. We define a binary relation between $\mathcal{M} \otimes \mathcal{A}$ and $\mathcal{M} \otimes \mathcal{B}$ as $(w, s)R(v, t) \iff w = v, sEt$ and $Obs((w, s)) = Obs((v, t))$. Now we verify the condition Zig of Definition 9 (the invariance condition is trivial by definition of R). Suppose $(w, s) \sim_i (w', s')$

then $w \sim_i w'$ in $\mathcal{M}$ and $s \sim_i s'$ in $\mathcal{A}$. Since sEt, there is a t' in $\mathcal{B}$ such that $t \sim_i t'$, $s'Et'$, and $\mathcal{L}^{\rho_0}(\mathit{Prot}(s')) = \mathcal{L}^{\rho_0}(\mathit{Prot}(t'))$ where $\rho_0 = V(s')$. Clearly (w', t') is in $\mathcal{M} \otimes \mathcal{B}$ and $\mathit{Obs}((w', t')) = \mathit{Obs}((w, s'))$. Thus we have that $(w, t) \sim_i (w', t')$ and $(w', s')R(w', t')$. The condition Zag can be proved in a similar way.

$\Rightarrow$: Suppose $\mathcal{A} \equiv_{ef} \mathcal{B}$. It is clear that for a universal ignorance model $\mathcal{M}$ (cf. the proof of Proposition 19): $\mathcal{M} \otimes \mathcal{A} \leftrightarrows_o \mathcal{M} \otimes \mathcal{B}$. We set a relation E between the state spaces of $\mathcal{A}$ and $\mathcal{B}$ as sEt iff $(w, s) \leftrightarrows_o (w, t)$. We can verify that E is a protocol emulation relation. The first (consistency) condition of protocol emulation is immediate according to the invariance condition of observation bisimulation. Now we show the second one. Suppose $s \sim_i s'$ and sEt. Now consider an arbitrary $\rho \subseteq \mathbf{P}$ such that $\mathcal{L}^{\rho}(\mathit{Prot}(s')) \neq \emptyset$. Since $\mathcal{M}$ is a universal ignorance model, there is a state w' in $\mathcal{M}$ such that $V(w') = \rho$ and $(w, s) \sim_i (w', s')$. Since sEt then by definition of E, $(w, s) \leftrightarrows_o (w, t)$. Thus there is a (v', t') in $\mathcal{M} \otimes \mathcal{B}$ such that $(w, t) \sim_i (v', t')$ and $(w', s') \leftrightarrows_o (v', t')$ (clearly $w' = v'$ since $\mathcal{M}$ is a universal ignorance model). It follows that $t \sim_i t'$ and $\mathcal{L}^{\rho}(\mathit{Prot}(s')) = \mathcal{L}^{\rho}(\mathit{Prot}(t'))$. Thus for all $\rho \subseteq \mathbf{P}$ such that $\mathcal{L}^{\rho}(\mathit{Prot}(s')) \neq \emptyset$ there is a state t' such that $t \sim_i t'$ in $\mathcal{B}$, $s'Et'$ and $\mathcal{L}^{\rho}(\mathit{Prot}(s')) = \mathcal{L}^{\rho}(\mathit{Prot}(t'))$. The third condition can be shown similarly. $\square$

We now extend the framework for *POL* to provide a *DEL*-style logical language, describing the 'installation' or 'change' of protocols, together with the effect of the observations of agents, based on the current protocol.

3.3 Epistemic Protocol Logic

In the language of the Epistemic protocol logic (*EPL*), we consider protocol models as first-class citizens, giving a *DEL*-like language.

Definition 28 (Language of EPL). *The* formulas φ *of EPL are given by:*

$$\varphi \quad ::= \quad \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid [\pi]\varphi \mid [!\mathcal{A}_e]\varphi$$

where $p \in \mathbf{P}$*,* $i \in \mathbf{I}$*,* $\pi \in \mathcal{L}_{obs}$*, and* $\mathcal{A}_e$ *is an epistemic protocol with the designated state* e*.*

In defining the language we restrict ourselves to finite protocol models. The models for the logic *EPL* are taken to be the epistemic observation models $\mathcal{M} = \langle S, \sim, V, \mathit{Obs}\rangle$. The truth definition is given as follows:

Definition 29 (Truth definition for EPL). *Given an epistemic observation model* $\mathcal{M} = \langle S, \sim, V, \mathit{Obs}\rangle$*, a state* $s \in S$*, and an EPL-formula* φ*, the truth conditions of* φ *at* s *coincide with POL for the formulas that they have in common. The truth condition for the new formula in EPL is defined as follows:*

$$\boxed{\begin{array}{lll} \mathcal{M}, s \vDash [!\mathcal{A}_e]\varphi & \Leftrightarrow & \mathcal{L}(f_{V(s)}(\mathit{Prot}(e))) \neq \emptyset \\ & & \quad \Longrightarrow \mathcal{M} \otimes \mathcal{A}, (s, e) \vDash \varphi \end{array}}$$

Recalling the meaning of the modal product operation, the expression '$[!\mathcal{A}_e]\varphi$' therefore stands for 'after installing the new epistemic protocol $\mathcal{A}_e$, φ is true'. As an example, let us give the model of Example 1, the observation model induced by the epistemic protocol (modelled earlier, call it $\mathcal{A}_e$), and the updated model according to observation a (in the picture, visualized by $|_a$):

$$\begin{tikzcd}
\underline{g(a)} \arrow[r, dash, "\text{Jane,Ann}"] \arrow[d, dash, "\text{Ann}"] & \neg g(b) \arrow[d, dash, "\text{Ann}"] & |_a & \underline{g(\epsilon)} \arrow[d, dash, "\text{Ann}"] & \\
g(a+b) \arrow[r, dash, "\text{Jane,Ann}"] & \neg g(a+b) & & g(\epsilon) \arrow[r, dash, "\text{Jane,Ann}"] & \neg g(\epsilon)
\end{tikzcd}$$

Recall the initial model $\mathcal{M}$:

$$\underline{g(\mathbf{\Sigma}^*)} \xrightarrow{\hspace{1em}\mathit{Jane,Ann}\hspace{1em}} \neg g(\mathbf{\Sigma}^*)$$

Suppose the actual state (underlined) was the leftmost state in $\mathcal{M}$, (say s), then we can verify:

$\mathcal{M}, s \vDash [!\mathcal{A}_e][a](K_{Jane}g \wedge \neg K_{Ann}g)$, and
$\mathcal{M}, s \vDash [!\mathcal{A}_e][a]\neg K_{Ann}(K_{Jane}g \vee K_{Jane}\neg g)$.

The picture corresponding to Example 2 is as follows ($\mathcal{A}'_{e'}$ is the corresponding epistemic protocol modelled earlier):

$$\begin{tikzcd}
p(Y) \arrow[r, dash, "\text{Ben,Alice}"] \arrow[d, dash, "\text{Alice}"] & \underline{\neg p(N)} \arrow[d, dash, "\text{Alice}"] & |_N & \underline{\neg p(\epsilon)} \arrow[d, dash, "\text{Alice}"] \\
\neg p(Y) \arrow[r, dash, "\text{Ben,Alice}"] & p(N) & & p(\epsilon)
\end{tikzcd}$$

Recall the initial model $\mathcal{N}$:

$$p(\mathbf{\Sigma}^*) \xrightarrow{\hspace{1em}\mathit{Ben,Alice}\hspace{1em}} \underline{\neg p(\mathbf{\Sigma}^*)}$$

Let the actual state (underlined) be the rightmost state in $\mathcal{N}$ (say t), we can verify:

$\mathcal{N}, t \vDash [!\mathcal{A}'_{e'}][N](K_{Ben}\neg p \wedge \neg K_{Alice}\neg p)$, but
$\mathcal{N}, t \vDash [!\mathcal{A}'_{e'}][N]K_{Alice}(K_{Ben}p \vee K_{Ben}\neg p)$.

The further investigation of this logic *EPL* and its relation to *DEL* is future work.

4. INCORPORATING FACTUAL CHANGES

So far, we presented information-changing actions, not fact-changing actions: recall that $\mathcal{L}_g(\eta)$ consists of guarded strings with uniform guards only. This is not so realistic in practice since many actions used in protocols also change the facts, e.g., 'turn on the light if you see the enemy'. Factual change can be modelled by assigning to each action a function which changes the valuation of basic propositions (as in [20, 23]). Let us now show how protocols based on factual changing actions can be incorporated in our setting.

Definition 30 (factual changing actions). *A set of factual changing actions (fc-actions) is a tuple* $(\mathbf{\Sigma}, \iota)$ *such that* $\iota : \mathbf{\Sigma} \times \mathbf{P} \to \mathit{Bool}(\mathbf{P})$.

Intuitively, after executing action $a \in \mathbf{\Sigma}$, p is assigned the truth value of $\iota(a, p)$ (evaluated before executing a). For example, let p be the proposition denoting 'the door is closed' then slamming the door (a) has the post-effect: $\iota(a)(p) = \top$. On the other hand, toggling the switch (b) has the post-effects modelled by $\iota(b)(p) = \neg p$ if p expresses the switch is on. Clearly non-factual change actions can be seen as $(\mathbf{\Sigma}, \iota_0)$ where for any $a \in \mathbf{\Sigma}$, $\iota_0(a)$ is the identity function.

For the ease in proofs, we introduce an alternative way to represent factual changing actions.

DEFINITION 31 (FACTUAL CHANGE SYSTEM). *A $\mathbf{\Sigma}$-factual change system (fc-system) $\mathcal{F}$ is a tuple $(Q, \longrightarrow)$ where $Q = \mathcal{P}(\mathbf{P})$ and $\longrightarrow: Q \times \mathbf{\Sigma} \to Q$.*

Clearly, $\longrightarrow$ is a deterministic transition function and thus it can be extended to the domain of $Q \times \mathbf{\Sigma}^*$ such that $(\rho, a_0 \cdots a_k)$ is the unique state of the fc-system that is reachable via transitions subsequently labelled by $a_0, \dots, a_k$. We show that a set of factual change actions can be seen as a factual change system:

PROPOSITION 32. *For each set of fc-actions $(\mathbf{\Sigma}, \iota)$ there is a $\mathbf{\Sigma}$-fc-system such that for each $a \in \mathbf{\Sigma}, \rho \subseteq \mathbf{P}$: $\rho \vDash \bigwedge_{p \in \rho'} \iota(a,p) \wedge \bigwedge_{p \notin \rho'} \neg\iota(a,p) \iff \longrightarrow (\rho, a) = \rho'$. For each $\mathbf{\Sigma}$-fc-system there is a set of fc-actions $(\mathbf{\Sigma}, \iota)$ such that for each $a \in \mathbf{\Sigma}, \rho \subseteq \mathbf{P}$: $\rho \vDash \bigwedge_{p \in \rho'} \iota(a,p) \wedge \bigwedge_{p \notin \rho'} \neg\iota(a,p) \iff \longrightarrow (\rho, a) = \rho'$.*

PROOF. The first part of the proposition is straightforward, just take $\rho \vDash \bigwedge_{p \in \rho'} \iota(a,p) \wedge \bigwedge_{p \notin \rho'} \neg\iota(a,p) \iff \longrightarrow (\rho, a) = \rho'$ as the definition of the transition function in the factual change system. For the second part, let $\iota(a,p) = \bigvee\{\varphi_\rho \mid p \in \longrightarrow (\rho, a)\}$. $\Box$

In the sequel, we only work with fc-systems. To interpret observation expressions w.r.t. a fc-system $\mathcal{F}$, we only need to revise the definition of $\mathcal{L}_g$ as follows:

$$\mathcal{L}_g^{\mathcal{F}}(a) = \{\rho a \rho' \mid \rho \xrightarrow{a} \rho' \text{ in } \mathcal{F}\}$$

Based on the automaton developed in [12], we can prove an analogy of Proposition 15, viz. Proposition 33.

PROPOSITION 33. *Given an fc-system $\mathcal{F}$, every η has a normal form $\eta^{\mathcal{F}} = \sum_{\rho \subseteq \mathbf{P}} (?\rho \cdot \pi_\rho)$ for some $\pi_\rho \in \mathcal{L}_{obs}$ such that $\mathcal{L}_g^{\mathcal{F}}(\eta) = \mathcal{L}_g^{\mathcal{F}}(\eta^{\mathcal{F}})$.*

PROOF. (*a sketch of the proof*) In [12], Kozen gave a general semantics for guarded expressions (the η's in $\mathcal{L}_{Prot}$ as in our paper), where the only difference concerns the clause for the atomic a :

$$\mathcal{L}_g^K(a) = \{\rho a \rho' \mid \rho, \rho' \subseteq \mathbf{P}\}$$

Note that there is no constraint between ρ and ρ' in the above definition. It is not hard to see that given an fc-system $\mathcal{F}$ we can define a translation $t^{\mathcal{F}} : \mathcal{L}_{Prot} \to \mathcal{L}_{Prot}$ by replacing each a with $\sum_{\rho \subseteq \mathbf{P}} \{?\rho \cdot a \cdot ?\rho' \mid \rho \xrightarrow{a} \rho' \text{ in } \mathcal{F}\}$. It follows that $\mathcal{L}_g^K(t^{\mathcal{F}}(\eta)) = \mathcal{L}_g^{\mathcal{F}}(\eta)$. It is shown in [12] that guarded regular expressions correspond to deterministic guarded automata (finite automata with transitions labelled by atomic actions and Boolean tests) satisfying the following properties:

- Each state is either a state that only has outgoing action transitions (*action state*) or a state that only has outgoing test transitions (*test state*).
- The start state is a test state.
- The outgoing test transitions are deterministic: they are labelled by characteristic formulas of $\rho \subseteq \mathbf{P}$ and for each test state q and each ρ, q has one and only one φ_ρ-successor.

Therefore by following the different ρ transitions from the start state, we can separate the automaton that corresponds to the guarded regular expression into $|2^{\mathbf{P}}|$ zones. It is easy then to generate the corresponding regular expressions (observations) for each zone (by ignoring the test transitions). In such a way, the normal form of η can be generated. $\Box$

Based on the above proposition, we can define the 'installation' of protocols with factual changing actions on observation models, similar to Definition 18.

5. CONCLUSION AND FUTURE WORK

The information that the actions carry may depend on agents' knowledge of protocols. In this paper we studied cases where protocols are not commonly known and proposed a logic framework for updating knowledge by observations based on epistemic protocols. We consider various extensions of our work.

We only used Boolean tests in the language $\mathcal{L}_{prot}$. A more expressive protocol language includes epistemic tests. An example of such a protocol would be $(?\neg Kp \cdot (a+b))^* \cdot (?Kp \cdot c)$: as long as you do not know p, keep choosing an a or b action, until you get to know p, and then do c. As observed in [7], knowledge-based protocols are much more involved than fact-based protocols. Defining the interpretation and executability of such protocols is a challenge, because, checking epistemic formulas is then non-local. Also, the introduction of knowledge tests may make the satisfiability problem of the logic undecidable. For example, the observations may easily encode iterated public announcement, which is known as a source of undecidability in such logics [13]. On the positive side, by including more expressive tests we expect better matching between observation models and epistemic protocols (cf. Theorem 24).

Another extension is to consider less than radical update mechanisms for installing new protocols. In our current approach, when installing a new protocol, we simply ignore and overwrite the old expected observations completely. Consider a singleton observation epistemic model with observation $a+c$. Now, when updating with the protocol $a+b$ we simply replace $a+c$ by $a+b$. Instead, we could integrate $a+c$ with $a+b$, somehow. For example, such a 'non-radical' protocol update with $a+b$ could result in b (intersected refinement), or in $(b+c)\cdot(a+b)$ (concatenation), or in $(b+c)+(a+b)$ (choice), and so on. See [28] for a discussion. Finally, we can relax the assumption of public observation, e.g., some actions may not be observable to certain agents.

The subject of hidden protocols is also interesting from the point of view of language pragmatics. Speakers who intend to convey information to only some of their listeners in such a way that others will not understand what is going on, are deliberately acting against some of Grice's maxims of cooperative conversation [9]. Forms of indirect or uncooperative communication, such as veiled bribes and threats, have already been investigated from the perspective of pragmatics and cognitive science, relating them also to aspects like lack of common knowledge [4, 26, 16, 22]. Our analysis of hidden protocols in this paper, by distinguishing between observations and actions, is more fine-grained than the changes in 'standard' dynamic epistemic logic, but could benefit from taking such Gricean aspects into account. Thus, in addition to observational powers of the agents, also their assertive powers may be modeled. Finally, it would be interesting to investigate the role of the interlocutors' goals and intentions when they utter a veiled speech act that is part of a hidden protocol (cf. [2, 17, 10, 6]).

Acknowledgments

We thank the anonymous reviewers whose suggestions helped us to improve the paper considerably. Hans van Ditmarsch is also affiliated to IMSc (Institute of Mathematical Sciences, Chennai, India), as associated researcher. Part of this research was carried out by Hans van Ditmarsch and Yanjing Wang during their joint stay at the IMSc. Hans van Ditmarsch thanks the Netherlands Organization for Scientific Research (NWO) for a visiting grant 040.11.177 to Rineke Verbrugge, University of Groningen, of which this publication can be seen as the outcome. Sujata Ghosh acknowledges NWO research grant 600.065.120.08N201 and Rineke Verbrugge acknowledges NWO research grants 600.065.120.08N-201 and 227-80-001.

6. REFERENCES

[1] A. Baltag. A logic for suspicious players: Epistemic actions and belief-updates in games. *Bulletin of Economic Research*, 54(1):1–45, 2002.

[2] M. Bratman. *Intention, Plans, and Practical Reason*. Harvard University Press, Cambridge (MA), 1987.

[3] J. A. Brzozowski. Derivatives of regular expressions. *Journal of the ACM*, 11(4):481–494, 1964.

[4] H. Clark. *Using Language*. Cambridge University Press, Cambridge, 1996.

[5] J. H. Conway. *Regular Algebra and Finite Machines*. Chapman and Hall, 1971.

[6] F. Dignum, B. Dunin-Kęplicz, and R. Verbrugge. Creating collective intention through dialogue. *Logic Journal of the IGPL*, 9:145–158, 2001.

[7] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. Knowledge-based programs. *Distributed Computing*, 10(4):199–225, July 1997.

[8] R. Fagin, J. Y. Halpern, M. Y. Vardi, and Y. Moses. *Reasoning about Knowledge*. MIT Press, Cambridge, MA, USA, 1995.

[9] H. P. Grice. Logic and conversation. In P. Cole and J. L. Morgan, editors, *Syntax and semantics*, volume 3. New York: Academic Press, 1975.

[10] B. Grosz and C. Sidner. Plans for discourse. In P. Cohen, J. Morgan, and M. Pollack, editors, *Intentions in Communication*, pages 417–444. MIT Press, Cambridge (MA), 1990.

[11] T. Hoshi. *Epistemic Dynamics and Protocol Information.* PhD thesis, Stanford University, 2009.

[12] D. Kozen. Automata on guarded strings and applications. Technical report, Cornell University, Ithaca, NY, USA, 2001.

[13] J. S. Miller and L. S. Moss. The undecidability of iterated modal relativization. *Studia Logica*, 79(3):373–407, 2005.

[14] E. Pacuit and S. Simon. Reasoning with protocols under imperfect information (abstract). Short paper presentation at AiML 2010, Moscow, 2010.

[15] R. Parikh and R. Ramanujam. A knowledge based semantics of messages. *Journal of Logic, Language and Information*, 12(4):453–467, 2003.

[16] S. Pinker, M. Nowak, and J. Lee. The logic of indirect speech. *Bulletin of Economic Research*, 54(1):1–45, 2002.

[17] A. Rao and M. Georgeff. Modeling rational agents within a BDI-architecture. In R. Fikes and E. Sandewall, editors, *Proceedings of the Second Conference on Knowledge Representation and Reasoning*, pages 473–484. Morgan Kaufman, 1991.

[18] S. Singh. *The Code Book: The Evolution of Secrecy from Mary, Queen of Scots, to Quantum Cryptography*. Doubleday, New York, NY, USA, 1999.

[19] J. van Benthem, J. Gerbrandy, T. Hoshi, and E. Pacuit. Merging frameworks for interaction. *Journal of Philosophical Logic*, 38(5):491–526, 2009.

[20] J. van Benthem, J. van Eijck, and B. Kooi. Logics of communication and change. *Information and Computation*, 204(11):1620–1662, 2006.

[21] H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*, volume 337 of *Synthese Library*. Springer, 2007.

[22] H. van Ditmarsch, J. van Eijck, and R. Verbrugge. Common knowledge and common belief. In J. van Eijck and R. Verbrugge, editors, *Discourses on Social Software*, volume 5 of *Texts in Games and Logic*. Amsterdam University Press, Amsterdam, 2009.

[23] J. van Eijck. Perception and change in update logic. In J. van Eijck and R. Verbrugge, editors, *Games, Actions and Social Software*, Texts in Logic and Games (FOLLI subseries of LNCS). Springer Verlag, Berlin, 2011.

[24] J. van Eijck, J. Ruan, and T. Sadzik. Action emulation, 2008. Unpublished manuscript, available at homepages.cwi.nl/ jve/papers/08/ae/ae.pdf.

[25] A. van Kooten Niekerk and S. Wijmer. *Verkeerde Vriendschap: Lesbisch Leven in de Jaren 1920-1960*. Sara, Amsterdam, 1985.

[26] R. Verbrugge and L. Mol. Learning to apply theory of mind. *Journal of Logic, Language and Information*, 17:489–511, 2008. Special issue on formal models for real people, edited by M. Counihan.

[27] Y. Wang. *Epistemic Modelling and Protocol Dynamics*. PhD thesis, University of Amsterdam, 2010.

[28] Y. Wang. Reasoning about protocol change and knowledge. In *Proceedings of the 4th Indian Conference on Logic and its Applications (ICLA 2011), LNAI 6521*, pages 189–203. Springer, 2010.