Coalgebraic Epistemic Update Without Change of Model*

Corina Cîrstea and Mehrnoosh Sadrzadeh

School of Electronics and Computer Science, University of Southampton
cc2,ms6@ecs.soton.ac.uk

Abstract. We present a coalgebraic semantics for reasoning about information update in multi-agent systems. The novelty is that we have one structure for both states and actions and thus our models do not involve the "change-of-model" phenomena that arise when using Kripke models. However, we prove that the usual models can be constructed from ours by categorical adjunction. The generality and abstraction of our coalgebraic model turns out to be extremely useful in proving preservation properties of update. In particular, we prove that positive knowledge is preserved and acquired as a result of epistemic update. We also prove common and nested knowledge properties of epistemic updates induced by specific epistemic actions such as public and private announcements, lying, and in particular unsafe actions of security protocols. Our model directly gives rise to a coalgebraic logic with both dynamic and epistemic modalities. We prove a soundness and completeness result for this logic, and illustrate the applicability of the logic by deriving knowledge properties of a simple security protocol.

1 Introduction

Modelling interactive multi-agent systems has a wide range of applications, e.g. in Artificial Intelligence, computer security and e-commerce. In such systems agents communicate and as a result their knowledge gets updated, and therefore one has to model the epistemics and dynamics of the system. The Kripke and algebraic models of these settings have been presented in [9,10,3,2,14]. The Kripke models have the advantage of being intuitive and concrete, while the algebraic setting benefits from high level features that result from mathematical abstraction.

In this paper we develop a coalgebraic semantics for dynamic epistemic systems, which combines the advantages of both the Kripke and the algebraic setting. Our model reasons about such systems in a uniform way, by treating both actions and agents as state transformers. Thus, we have only one structure that captures both dynamics and epistemics. This is contrary to the models of e.g. [10,3,1] that require subsequent "changes" to the epistemic structure to

* Research supported by EPSRC grant EP/D000033/1.

T. Mossakowski et al. (Eds.): CALCO 2007, LNCS 4624, pp. 158–172, 2007.

model the dynamics. By "change" we mean either the *update product* between an epistemic Kripke structure and an action Kripke structure [3], or the *update functor* on the category of epistemic coalgebras [10,1]. In either case the epistemic structure is taken to be primitive and the dynamics is captured by operations on it. This brings us to the other novelty of our approach: we start our modelling task by fixing the epistemic actions, and then define the epistemic states based on these actions and on the agents participating in them. Again, this is contrary to the models of [10,3,1], which involve first fixing the epistemic states and then defining all possible epistemic actions on these states. Although at first sight our approach seems very different from the approach of [10,3,1], the two are strongly connected. In the main theorem of our paper we show how to construct from our models the models of [1] and vice versa, and prove that these constructions form a categorical adjunction[1].

Our approach has all the advantages of the approach in [3,1], for instance it benefits from a general updating schema and it reflects the epistemic structure of actions. Moreover, our approach does not have the usual weaknesses, for example operations on actions are a natural part of our models, e.g. sequential composition is simply unfolding the coalgebra maps twice and does not need to be defined separately. The generality and abstraction of our coalgebraic models turns out to be extremely useful in proving preservation properties of update. In particular, we prove that positive knowledge is preserved and acquired as a result of epistemic update. We also prove common and nested knowledge properties of epistemic updates induced by specific epistemic actions such as public and private announcements, lying, and in particular unsafe actions of security protocols. Finally, our model directly gives rise to a coalgebraic logic with both epistemic and dynamic modalities. This, for instance, cannot be done for the models of [10,1]. We prove a soundness and completeness result for the resulting logic. As an example of application, we derive the authentication properties of a security protocol.

An extended version of this paper is available electronically [8]. There we illustrate the applicability of our models to general scenarios involving both positive and negative knowledge, by presenting a new coalgebraic proof of the muddy children puzzle and a version of it with cheating children. Our proofs are based on restrictive recursion rather than the usual induction.

2 Coalgebraic Semantics for Actions and Agents

We consider coalgebras of the following signature functor $T : Set \to Set$

$$TX = \mathcal{P}_\kappa(X)^{Ag} \times (1 + X)^{Ac} \times \mathcal{P}(At)$$

where κ is a regular cardinal. A coalgebra map for the above functor is thus a triple $\gamma = \langle ap, up, val \rangle : S \to \mathcal{P}_\kappa(S)^{Ag} \times (1+S)^{Ac} \times \mathcal{P}(At)$. The valuation map val

[1] As noted by one of our referees, our model might share ideas with the recent *epistemic temporal models* of [4].

assigns to each state the facts that are true in that state. The (nondeterministic) appearance of states in S to agents in Ag is modelled by the function $ap : S \to \mathcal{P}(S)^{Ag}$, whereas the (deterministic) effect of actions in Ac on states in S is modelled by a function $up : S \to (1 + S)^{Ac}$. Thus, $up(s)(a)$ stands for the effect of the action a on the state s, or the update of s by a. If this effect is the unique element $*$ of 1, that is, if $up(s)(a) = *$ [2], we say that action a does not apply in state s; this should be the case, for instance, when a is the announcement of a fact that does not belong to $val(s)$. The choice of functor T automatically yields notions of T-*bisimulation* and T-*bisimilarity* for T-coalgebras (see e.g. [16]).

2.1 Restrictions to the Coalgebras

We are interested in using T-coalgebras to model the effect of communication actions on the information state or knowledge of agents. Examples of such actions are public or secret announcements, and message passing actions in a multi-agent system. We want to model the effect of updates with such actions on the appearances of states to the agents and on the valuations of states. In order to limit the behaviour of our systems to the effect of these actions, we require that the coalgebra maps satisfy some additional conditions, detailed in the following.

The communication actions that we model are *epistemic*, that is, they only affect the information states of agents, while leaving the facts of the world unchanged. Our first restriction, called *preservation of facts*, reflects this point:

$$val(up(s)(a)) = val(s) \quad \text{whenever} \quad up(s)(a) \neq *$$

It says that, if applicable to a state, an action does not change the valuation of that state. So the valuation of the effect of the action is the same as the valuation of the state before the action. We need this restriction to prove the preservation results in Section 2.2. In a more general approach, one can divide the set of actions into two subsets, namely *information-changing* actions and *fact-changing* actions, and only require this restriction for actions of the first type.

Our second restriction concerns the appearance of an update to each agent involved in the corresponding action. For applicable updates $up(s)(a) \neq *$, this will be related to the update of each of the agent's appearances $t \in ap(s)(A)$ with a finite subset of actions $Ac_{a,A} \subseteq Ac$, as follows

$$ap(up(s)(a))(A) = \{up(t)(a') \mid t \in ap(s)(A),\ a' \in Ac_{a,A},\ up(t)(a') \neq *\}$$

where the actions $Ac_{a,A}$ depend both on the action a and on agent A's involvement in it, and are intended to capture agent A's appearance of the action a. This relation says that if an action a applies to a state s, then the appearance of its effect to an agent A is the same as the effect of one of the actions in $Ac_{a,A}$

[2] Here and in what follows, we assume that $1 \cap S = \emptyset$. Under this assumption, and to simplify the notation, we regard elements of the set $1 + S$ as being either $*$ or elements of S; that is, we make implicit the isomorphism between $1 + S$ and $1 \cup S$.

on one of the appearances to A of the original state s. The case when $Ac_{a,A}$ is a singleton $\{a'\}$ corresponds to a deterministic view of A about the real action a (with A thinking that a' is happening when in fact a is happening), whereas any non-singleton set $Ac_{a,A}$ captures A's uncertainty about the action taking place. We refer to the collection of all instances of this restriction (one for each action in Ac) as *rationality*.

The *content* of an epistemic action, as its name suggests, describes the information that is being transmitted as a result of the action taking place. We use the following syntax to denote specific contents[3]:

$$\mu := p \mid \Box_A \mu \mid \mathsf{tt} \mid \neg\mu \mid \bigwedge_{i\in\mathbb{I}} \mu_i$$

with $p \in At$ and $\mathbb{I}$ an arbitrary set. That is, the content of an action can be a fact, the knowledge or belief of some other content by an agent[4], the true proposition, the negation of a content, or a potentially infinite conjunction of contents[5]. In particular, the content can involve nested knowledge, as in $\Box_A\Box_B p$. We do not allow contents to refer to (the effect of) actions, as in $[q]-$; this avoids a circularity between requiring each action to have a content and allowing contents to depend on actions. Contents whose only occurrences of the negation operator immediately precede a fact are called *positive contents*, otherwise they are referred to as *negative contents*.

From now on, we assume that each action $a \in A$ has a content μ_a associated to it. Then, a should be applicable precisely to those states where its content μ_a is satisfied. This is encoded as a further restriction on T-coalgebras, referred to as the *content restriction*:

$$up(s)(a) \neq * \quad \text{iff} \quad s \models \mu_a$$

where the relation $\models$ between states and contents of actions is defined by structural induction on contents:

- $s \models p$ iff $p \in val(s)$
- $s \models \Box_A\mu$ iff $t \models \mu$ for all $t \in ap(s)(A)$

and the usual clauses for the true proposition, negation and conjunction.

Definition 1. *An* appearance-update coalgebra *is a T-coalgebra additionally satisfying the preservation of facts, content, and rationality restrictions. We denote the set of all of these restrictions by $\mathcal{R}$.*

[3] We emphasise that this is just a syntax for expressing our second restriction on the content s of actions. The logic will be presented in section 5.

[4] Similarly to [3] and as a result of accommodating misinformation actions, our knowledge $\Box_A$ is not necessarily truthful. Indeed, one can also think of $\Box_A$ as belief in contexts where no wrong knowledge is allowed.

[5] The infinite contents are just a technicality that is needed later in order to establish a connection to the model of [1].

2.2 Preservation and Acquisition of Knowledge

An important consequence of the restrictions in $\mathcal{R}$ is the so-called *preservation of positive contents by updates*, made formal in the next result.

Proposition 1. *Let $(S, \langle ap, up, val\rangle)$ be an appearance-update coalgebra. Then for all positive contents μ, all states $s \in S$, and all actions $a \in Ac$ such that $up(s)(a) \neq *$, we have*

$$s \models \mu \quad \Longrightarrow \quad up(s)(a) \models \mu$$

Proof. The statement is proved by induction on μ. If μ is a fact or the negation of a fact, the conclusion follows directly from the preservation of facts. Now suppose that $s \models \mu'$ implies $up(s)(a) \models \mu'$ for all states $s \in S$ and applicable actions $a \in Ac$. Also, let $s \in S$ and $A \in Ag$ be such that $s \models \Box_A \mu'$. To show that $up(s)(a) \models \Box_A \mu'$ for any applicable action a, we use the rationality restriction to reduce $ap(up(s)(a))(A)$ to $\{up(t)(a') \mid t \in ap(s)(A),\ a' \in Ac_{a,A},\ up(t)(a') \neq *\}$. Thus, we must show that $up(t)(a') \models \mu'$ whenever $t \in ap(s)(A)$ and $a' \in Ac_{a,A}$ are such that $up(t)(a') \neq *$. But this follows from the induction hypothesis and the assumption that $s \models \Box_A \mu'$. The cases when μ is the true proposition or a conjunction of contents are trivial.

The above result does not hold for negative contents. That is, there exists an appearance-update coalgebra $(S, \langle ap, up, val\rangle)$, a state $s \in S$ with an applicable action $a \in Ac$ and a negative content μ such that $s \models \mu$ but $up(s)(a) \models \neg\mu$. For an example of such a situation, which gives rise to the epistemic puzzle of *muddy children*, see [8]. It is also not possible to generalise the above result to an exclusive one for positive contents. In particular, any appearance-update coalgebra that contains in its set of actions a neutral action τ with $\mu_\tau = tt$ and $Ac_{\tau,A} = \{\tau\}$ for all $A \in Ag$ is such an example. To see why, we refer the reader to the next section where we prove that such an action preserves all contents.

Another consequence of the restrictions in $\mathcal{R}$ is the following *acquisition of knowledge* after updates:

Proposition 2. *Let $(S, \langle ap, up, val\rangle)$ be an appearance-update coalgebra. Then for all agents $A \in Ag$, all states $s \in S$, and all applicable actions $a \in Ac$ with positive contents $\mu_{a'}$ for all $a' \in Ac_{a,A}$, we have*

$$up(s)(a) \models \Box_A \bigvee_{a' \in Ac_{a,A}} \mu_{a'}$$

Proof. Let $s \in S$ and $a \in Ac$ be such that $up(s)(a) \neq *$. We must show that for $A \in Ag$ we have $s' \models \bigvee_{a' \in Ac_{a,A}} \mu_{a'}$ for all $s' \in ap(up(s)(a))(A)$. By the rationality restriction on $ap(up(s)(a))(A)$, we must show that $up(t)(a') \models \bigvee_{a' \in Ac_{a,A}} \mu_{a'}$ whenever $t \in ap(s)(A)$ and $a' \in Ac_{a,A}$ are such that $up(t)(a') \neq *$. By the content restriction, the positivity of $\mu_{a'}$ and the preservation result we obtain $up(t)(a') \models \mu_{a'}$, which implies $up(t)(a') \models \bigvee_{a' \in Ac_{a,A}} \mu_{a'}$.

The known preservation results in the literature are special cases of our general results. For instance, it has been shown in [3] that contents that do not contain the epistemic modality are preserved under any update.

3 Epistemic Actions

In this section, we present epistemic actions, describe their contents and appearances, and prove their knowledge acquisition effects on agents[6].

Skip. This is the action τ in which nothing happens. We have $\mu_\tau = \mathtt{tt}$ and $Ac_{\tau,A} = \{\tau\}$ for all $A \in Ag$. This particular choice of μ_τ and $Ac_{\tau,A}$ is sufficient to guarantee that, in any appearance-update coalgebra, the skip action does not affect the epistemic content of states; that is, no knowledge is lost or acquired as a result of this action. This is formalised in the next two results, where we write $F : Set \to Set$ for the functor defined by $F(S) = \mathcal{P}_\kappa(S)^{Ag} \times \mathcal{P}(At)$.

Proposition 3. *In any appearance-update coalgebra $(S, \langle ap, up, val\rangle)$ where the set Ac of actions includes the τ action, $up(s)(\tau) \sim_F s$ for any state $s \in S$, where $\sim_F \subseteq S \times S$ denotes the F-bisimilarity relation on the F-coalgebra $(S, \langle ap, val\rangle)$.*

Proof. The statement follows by coinduction, namely by showing that the relation $R \subseteq S \times S$ given by $\{\, (s, up(s)(\tau)) \mid s \in S \,\}$ is an F-bisimulation. The preservation of facts ensures that R only relates states with the same valuations, whereas the rationality restriction guarantees closure of R under appearances.

Since F-bisimilar states satisfy the same content formulas, a stronger preservation of knowledge result can now be formulated for the τ action.

Corollary 1. *Let $(S, \langle ap, up, val\rangle)$ be an appearance-update coalgebra. Then for all contents μ and all states $s \in S$, we have*

$$s \models \mu \quad \Longleftrightarrow \quad up(s)(\tau) \models \mu$$

Public Announcements. The public announcement of a content μ is denoted $\mu!$, and has $Ac_{\mu!,A} = \{\mu!\}$ for all $A \in Ag$. We define *truthful common knowledge* of a content μ among a group β of agents as follows

$$CK_\beta\, \mu := \bigwedge_{\langle A_0, A_1, \ldots, A_n\rangle \in \beta^*} \Box_{A_0}\Box_{A_1} \ldots \Box_{A_n}\mu$$

where $\beta^* = \cup_{i\in\mathbb{N}}\, \beta^i$ is the set of all finite sequences of agents in β, including the empty sequence. Excluding the empty sequence provides us with the notion of *not necessarily truthful common knowledge*, denoted $\Box^*_\beta \mu$.

We now show that the public announcement of a positive content results in truthful common knowledge of that content.

[6] To be in line with the existing literature, we consider contents rather than preconditions of actions. The difference between the two is best seen in an example: the content of a public announcement is simply the announced proposition μ, whereas its precondition is the conjunction of μ with the knowledge of the announcer about μ.

Proposition 4. *For a state s of an appearance-update coalgebra in which the public announcement $\mu!$ with positive μ is possible, we have $up(s)(\mu!) \models CK_{Ag}\,\mu$.*

Proof. We must show that for any state s and any state s' connected to the applicable update $up(s)(\mu!) \neq *$ via any sequence of appearance maps we have $s' \models \mu$. Thus, we have a sequence of states $up(s)(\mu!) = s_0, s_1, \ldots, s_m = s'$ such that for $0 \leq j < m$ and some agent $A_j \in Ag$ we have $s_{j+1} \in ap(s_j)(A_j)$. For $m = 0$, $s_0 \models \mu$ follows from the applicability of update $up(s)(\mu!) \neq *$, the content restriction, and the preservation result. For $m > 0$, we have that s_m is in the following set of nested appearances

$$ap(\ldots(ap(ap(up(s)(\mu!))(A_0))(A_1))\ldots)(A_{m-1})$$

which, by applying the rationality restriction m times, is equal to

$$\{\, up(t_m)(\mu!) \mid t_m \in ap(t_{m-1})(A_{m-1}),\, \ldots,\, t_2 \in ap(t_1)(A_1),\, t_1 \in ap(s)(A_0),\\ \text{and}\;\; up(t_1)(\mu!) \neq *,\, up(t_2)(\mu!) \neq *,\, \ldots,\, up(t_m)(\mu!) \neq * \,\}$$

By the content restriction $up(t_m)(\mu!) \neq *$ is equivalent to $t_m \models \mu$, and from this by the preservation result it follows that $up(t_m)(\mu!) \models \mu$.

The closest special case to this proposition is that of [3], where the authors show that common knowledge of a fact implies preservation of any content under the public announcement of that fact.

Private Announcements. A private announcement $\mu!_\beta$ is the action of announcing the content μ to a subgroup of agents $\beta \subseteq Ag$ with $Ac_{\mu!_\beta,B} = \{\mu!_\beta\}$ for $B \in \beta$ and $Ac_{\mu!_\beta,A} = \{\tau\}$ for $A \notin \beta$.

As expected, one can prove that the private announcement of a positive content to a subgroup of agents results in truthful common knowledge of that content among the subgroup, and has no visible effect outside the subgroup.

Proposition 5. *For $\beta \subseteq Ag$ and a state s of an appearance-update coalgebra in which the private announcement $\mu!_\beta$ with positive μ is possible, we have $up(s)(\mu!_\beta) \models CK_\beta\,\mu$ and $ap(up(s)(\mu!_\beta))(A) \sim^{\mathcal{P}} ap(s)(A)$ for $A \notin \beta$*[7].

Lying. We write $\mu\dagger_A$ for the action with content $\neg\mu$ in which an agent A lies that μ to the rest of the agents. We have $Ac_{\mu\dagger_A,A} = \{\mu\dagger_A\}$ and $Ac_{\mu\dagger_A,B} = \{\mu!\}$ for any $B \neq A$.

Proposition 6. *For any agent $A \in Ag$, $\beta = Ag \setminus \{A\}$, and any state s of an appearance-update coalgebra in which the lying action $\mu\dagger_A$ with a positive μ is possible, we have $up(s)(\mu\dagger_A) \models \Box^*_\beta\,\mu$ and $up(s)(\mu\dagger_A) \models \Box_A\Box^*_\beta\,\mu$.*

Proof. In order to show that $up(s)(\mu\dagger_A) \models \Box^*_\beta\,\mu$, we must show that for any state s and any state s' connected to the applicable update $up(s)(\mu\dagger_A) \neq *$

[7] Here $\sim^{\mathcal{P}}$ denotes the lifting of the bisimilarity relation on S to $\mathcal{P}(S)$, see e.g. [12].

via any sequence of length more than 1 of appearance maps of agents in β, we have $s' \models \mu$. Consider a sequence of states $up(s)(\mu\dagger_A) = s_0, s_1, \ldots, s_m = s'$ with $1 \leq m$, such that for $0 \leq j < m$ and some agent $B_j \in Ag \setminus \{A\}$ we have $s_{j+1} \in ap(s_j)(B_j)$. It follows that s_m is in the following set of nested appearances

$$ap(\ldots(ap(ap(up(s)(\mu\dagger_A))(B_0))(B_1))\ldots)(B_{m-1})$$

which, by applying the rationality restriction m times (once for the lying action $\mu\dagger_A$ and B_0 and $m-1$ times for the public announcement $\mu!$ and B_1 to B_{m-1}), is equal to

$$\{\, up(t_m)(\mu!) \mid t_m \in ap(t_{m-1})(B_{m-1}),\ \ldots,\ t_2 \in ap(t_1)(B_1),\ t_1 \in ap(s)(B_0),\\ \text{and}\ \ up(t_1)(\mu!) \neq *,\ up(t_2)(\mu!) \neq *,\ \ldots,\ up(t_m)(\mu!) \neq * \,\}$$

By the content restriction $up(t_m)(\mu!) \neq *$ is equivalent to $t_m \models \mu$, and from this by the preservation result it follows that $up(t_m)(\mu!) \models \mu$.

Now to show that $up(s)(\mu\dagger_A) \models \Box_A \Box^*_\beta \mu$, we must show that $t \models \Box^*_\beta \mu$ for all $t \in ap(up(s)(\mu\dagger_A))(A)$. By the rationality restriction we have

$$ap(up(s)(\mu\dagger_A))(A) = \{\, up(w)(\mu\dagger_A) \mid w \in ap(s)(A),\ up(w)(\mu\dagger_A) \neq * \,\}$$

Since $up(w)(\mu\dagger_A) \neq *$ and μ is positive, it follows from $up(s)(\mu\dagger_A) \models \Box^*_\beta \mu$ that $up(w)(\mu\dagger_A) \models \Box^*_\beta \mu$.

Security Actions. A security action $\mu \star \mu'_{\{A\},\beta,\gamma}$ is a private announcement in an unsafe communication channel, where the intruders in γ change the original content μ, sent by A to the agents in β, to a fake one μ'. In this case we have $Ac_{\mu\star\mu'_{\{A\},\beta,\gamma},A} = \{\mu!_{\beta\cup\{A\}}\}$, $Ac_{\mu\star\mu'_{\{A\},\beta,\gamma},B} = \{\mu'!_{\beta\cup\{A\}}\}$ for agents $B \in \beta$, $Ac_{\mu\star\mu'_{\{A\},\beta,\gamma},C} = \{\mu \star \mu'_{\{A\},\beta,\gamma}\}$ for the intruders $C \in \gamma$, while $Ac_{\mu\star\mu'_{\{A\},\beta,\gamma},D} = \{\tau\}$ for any other agent $D \in Ag \setminus (\{A\} \cup \beta \cup \gamma)$.

Proposition 7. *For any agents $B \in \beta$, $C \in \gamma$, and any state s of an appearance-update coalgebra in which the security action $\mu\star\mu'_{\{A\},\beta,\gamma}$ with positive μ and μ' is possible, we have $up(s)(\mu\star\mu'_{\{A\},\beta,\gamma}) \models \Box_A CK_\beta\, \mu$, $up(s)(\mu\star\mu'_{\{A\},\beta,\gamma}) \models \Box^*_\beta \Box_A\, \mu'$ and $up(s)(\mu \star \mu'_{\{A\},\beta,\gamma}) \models CK_\gamma(\Box_A CK_\beta\, \mu \wedge \Box^*_\beta \Box_A\, \mu')$.*

4 Comparison with Baltag's Coalgebraic Model

We now compare our coalgebraic semantics with that of [1]. In loc. cit., both *epistemic states* and *epistemic actions* are defined via final coalgebras. Two different functors of a similar shape are used to achieve this. However, none of these functors accounts for epistemic updates, which are instead modelled using a partial product between coalgebras of states and coalgebras of actions.

The functor used in [1] to model epistemic states is

$$F : Set \to Set, \qquad F(S) = \mathcal{P}_\kappa(S)^{Ag} \times \mathcal{P}(At)$$

Appearances of states to agents are encoded as elements of $\mathcal{P}_\kappa(S)^{Ag}$, while their valuations are encoded using sets of atomic propositions. Epistemic states are then defined as elements of the final F-coalgebra Ψ. Similarly, epistemic actions are defined as elements of the final coalgebra of the functor

$$G : Set \to Set, \qquad G(\Sigma) = \mathcal{P}_\kappa(\Sigma)^{Ag} \times \mathcal{P}(\Psi)$$

with $\mathcal{P}_\kappa(\Sigma)^{Ag}$ encoding the appearances of actions to agents, and $\mathcal{P}(\Psi)$ encoding the contents of actions (as sets of epistemic states where the actions are applicable). Finally, epistemic updates are modelled using a functor

$$- \otimes - : Coalg(F) \times Coalg(G) \to Coalg(F)$$

which takes a pair consisting of an F-coalgebra $(S, \langle ap_S, val_S \rangle)$ and a G-coalgebra $(\Sigma, \langle ap_\Sigma, cont_\Sigma \rangle)$ to another F-coalgebra whose elements correspond to updates of states in S with actions in Σ. Writing $!_S : S \to \Psi$ for the unique F-coalgebra morphism arising from the finality of Ψ, the coalgebra for the updated states has carrier

$$S \otimes \Sigma = \{(s, \sigma) \in S \times \Sigma \mid \ !_S(s) \in cont_\Sigma(\sigma)\}$$

That is, updated states are pairs consisting of a state $s \in S$ and an action $\sigma \in \Sigma$, with the additional property that the content of the action σ makes it applicable to the state s[8]. The coalgebra map $\langle ap_{S\otimes\Sigma}, val_{S\otimes\Sigma} \rangle : S \otimes \Sigma \to F(S \otimes \Sigma)$ is given by

$$ap_{S\otimes\Sigma}(s, \sigma)(A) = \{(s', \sigma') \in S \otimes \Sigma \mid s' \in ap_S(s)(A),\ \sigma' \in ap_\Sigma(\sigma)(A)\}$$
$$val_{S\otimes\Sigma}(s, \sigma) = val_S(s)$$

That is, the appearances of updated states to agents are computed using both the appearances of the original states and the appearances of the actions producing the updates.

In contrast to the above, our approach uses only one functor, which incorporates both the epistemic and the dynamic aspect of states. This internal modelling of updates is made possible by the fact that we apriorily fix a universe Ac of actions, together with its epistemic structure. The set Ac should be taken to contain those epistemic actions (elements of the final G-coalgebra) which are of interest to the modelling of a particular multi-agent scenario. In this setting, our choice to specify for each action $a \in Ac$ and agent $A \in Ag$, a set $Ac_{a,A}$ of actions that are perceived by A, together with for each action a a content μ_a, gives rise to a coalgebra $(Ac, \langle ap_{Ac}, \mu_{Ac} \rangle)$ of the following functor

$$H : Set \to Set, \qquad H(\Sigma) = \mathcal{P}_\kappa(\Sigma)^{Ag} \times C$$

where the set C consists of equivalence classes of content formulas. Here, two content formulas are said to be (semantically) equivalent if they are satisfied by

[8] Here it is assumed that the applicability of an action is invariant under bisimulation, and therefore an action is applicable to a state precisely when it is applicable to its image under the unique coalgebra morphism into the final F-coalgebra.

the same states of any F-coalgebra. The map ap_{Ac} of the previously mentioned H-coalgebra is given by $ap_{Ac}(a)(A) = Ac_{a,A}$ for $a \in Ac$ and $A \in Ag$, whereas the map μ_{Ac} takes actions $a \in Ac$ to the equivalence class of their content $[\mu_a]$. In this way, we do not distinguish between actions that have both the same epistemic structure and semantically equivalent contents.

In order to make precise the relationship between appearance-update T-coalgebras and the models of [1], we make the dependency of T on the set Ac of actions explicit, and write $T_{Ac} : Set \to Set$ for the functor given by

$$T_{Ac}X = (\mathcal{P}_\kappa X)^{Ag} \times (1 + X)^{Ac} \times \mathcal{P}(At)$$

Next, we let $AppUpCoalg$ denote the category whose objects are pairs (Ac, S), with $Ac = (Ac, \langle ap_{Ac}, \mu_{Ac}\rangle)$ an H-coalgebra and $S = (S, \langle ap_S, up_S, val_S\rangle)$ an appearance-update T_{Ac}-coalgebra. The H-coalgebra Ac encodes the structure on the set Ac of actions required to formulate the content and rationality restrictions of Section 2, whereas the T_{Ac}-coalgebra S specifies a set of states carrying both an epistemic structure and a dynamic structure w.r.t. the actions in Ac. To define the arrows of the category $AppUpCoalg$, we first note that any H-coalgebra morphism $f : Ac \to Ac'$ induces a functor

$$U_f : Coalg(T_{Ac'}) \to Coalg(T_{Ac})$$

which takes a $T_{Ac'}$-coalgebra $(S, \langle ap_S, up_S, val_S\rangle)$ to the T_{Ac}-coalgebra with the same carrier set and appearance and valuation maps, but with an update map w.r.t. the set Ac instead. This update is derived from the curried version $ev(up_S)\colon S \times Ac' \to (1 + S)$ of the update map up_S of the $T_{Ac'}$-coalgebra, as shown below

$$S \times Ac \xrightarrow{id_S \times f} S \times Ac' \xrightarrow{ev(up_S)} 1 + S$$

The curried version of this composition is the update map of the T_{Ac}-coalgebra

$$ev(ev(up_S) \circ (id_S \times f))\colon S \to (1 + S)^{Ac}$$

So we have $U_f(S, \langle ap_S, up_S, val_S\rangle) = (S, \langle ap_S, ev(ev(up_S) \circ (id_S \times f)), val_S\rangle)$. Now the arrows from (Ac, S) to (Ac', S') in the category $AppUpCoalg$ are pairs of maps (f, g) with $f : Ac \to Ac'$ an H-coalgebra morphism and $g : S \to U_f S'$ a T_{Ac}-coalgebra morphism. The former encodes the actions in Ac as actions in Ac', whereas the latter translates the states of the T_{Ac}-coalgebra S to states of the $T_{Ac'}$-coalgebra S'.

The last piece of notation we require before relating our models to those of [1] concerns *characteristic formulas* for states of F-coalgebras. These are infinitary formulas of the form used in Section 2 to specify the contents of epistemic actions, and have the additional property that they characterise individual states of F-coalgebras up to bisimulation. Their existence is guaranteed by the κ-accessibility of F. In particular, for any state ψ of the final F-coalgebra Ψ, there exists a characteristic formula ϕ_ψ with the property that, given any state s of an F-coalgebra S, we have $s \models \phi_\psi$ if and only if $!_S(s) = \psi$.

We are now ready to describe the relationship between the models of [1] and our appearance-update coalgebras. This is given by an adjunction

$$Coalg(F) \times Coalg(G) \underset{R}{\overset{L}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} AppUpCoalg$$

Definition 2 (Left adjoint). *We let* $L : Coalg(F) \times Coalg(G) \to AppUpCoalg$ *be defined by* $L(S, \Sigma) = ((\Sigma, \langle ap_\Sigma, \mu_\Sigma \rangle), (S', \langle ap_{S'}, up_{S'}, val_{S'} \rangle))$*, where*

- $\mu_\Sigma(\sigma) = \bigvee_{\psi \in cont_\Sigma(\sigma)} \phi_\psi$*, where for* $\psi \in \Psi$*,* ϕ_ψ *is the characteristic formula of* ψ*.*
- $S' = (S', \langle ap_{S'}, up_{S'}, val_{S'} \rangle)$ *is a* T_Σ*-coalgebra obtained by*
  1. *first letting* $S' = (S', \langle ap_{S'}, val_{S'} \rangle) = \cup_{i \in \omega}(S_i, \langle ap_{S_i}, val_{S_i} \rangle)$ *where*

$$S_0 = S, \qquad S_{i+1} = S_i \otimes \Sigma \quad \text{for } i \in \omega$$

  (Note that, by definition, each of the sets S_i *comes equipped with an* F*-coalgebra structure, and* S' *inherits this structure.)*
  2. *subsequently endowing the set* S' *with an update map* $up_{S'} : S' \to (1 + S')^\Sigma$*, by letting*

$$up_{S'}(s_i)(\sigma) = \begin{cases} (s_i, \sigma) & if (s_i, \sigma) \in S_{i+1} \\ * & otherwise \end{cases}, \quad \text{for } i \in \omega$$

In informal terms, the functor L constructs an H-coalgebra Σ and a T_Σ-coalgebra S' from a pair consisting of an F-coalgebra S and a G-coalgebra Σ. The H-structure of Σ is determined by the G-structure of Σ in a trivial way: appearances of actions to agents are already defined by the H-structure, whereas the content map $\mu_\Sigma : \Sigma \to C$ acts on an action $\sigma \in \Sigma$ by logically joining all the characteristic formulas of states in the content of σ. The T_Σ-coalgebra S' is obtained by performing consecutive update products with the actions in Σ, first on S, and then on the result of the preceding update product:

$$S \longmapsto S \otimes \Sigma \longmapsto (S \otimes \Sigma) \otimes \Sigma \longmapsto \ldots$$

and subsequently taking the union of the resulting F-coalgebras and endowing it with an update map.

Proposition 8. *The* T_Σ*-coalgebra* S' *is an appearance-update coalgebra.*

Proof. We have to show that S' satisfies all the restrictions in $\mathcal{R}$. The preservation of facts follows directly from the definitions of S_i and S': for $i \in \omega$, whenever $up_{S'}(s_i)(\sigma) \in S'$, that is, whenever $(s_i, \sigma) \in S_{i+1}$, we have

$$val_{S'}(up_{S'}(s_i)(\sigma)) = val_{S_{i+1}}(s_i, \sigma) = val_{S_i}(s_i) = val_{S'}(s_i)$$

For the rationality restriction, assuming $up_{S'}(s_i)(\sigma) \in S'$, that is, $(s_i, \sigma) \in S_{i+1}$, we have

$$ap_{S'}(up_{S'}(s_i)(\sigma))(A) = ap_{S_{i+1}}(s_i,\sigma)(A)$$
$$= \{(s',\sigma') \in S_i \otimes \Sigma \mid s' \in ap_{S_i}(s_i)(A),\ \sigma' \in ap_\Sigma(\sigma)(A)\}$$

and

$$\{up_{S'}(t)(\sigma') \mid t \in ap_{S'}(s_i)(A), \sigma' \in ap_\Sigma(\sigma)(A),\ up_{S'}(t)(\sigma') \neq *\} =$$
$$\{(t,\sigma') \mid t \in ap_{S_i}(s_i)(A), \sigma' \in ap_\Sigma(\sigma)(A), (t,\sigma') \in S_{i+1}\} =$$
$$\{(t,\sigma') \in S_i \otimes \Sigma \mid t \in ap_{S_i}(s_i)(A), \sigma' \in ap_\Sigma(\sigma)(A)\}$$

for each $i \in \omega$, and therefore

$$ap_{S'}(up_{S'}(s')(\sigma))(A) =$$
$$\{up_{S'}(t)(\sigma') \mid t \in ap_{S'}(s')(A), \sigma' \in ap_\Sigma(\sigma)(A),\ up_{S'}(t)(\sigma') \neq *\}$$

Finally, for the content restriction, we have, for each $i \in \omega$

$$\begin{array}{rcll} up_{S'}(s_i)(\sigma) \in S' & \text{iff} & (s_i,\sigma) \in S_{i+1} = S_i \otimes \Sigma & \text{iff} \\ !_{S_i}(s_i) \in cont_\Sigma(\sigma) & \text{iff} & !_{S_i}(s_i) \models \mu_\Sigma(\sigma) & \text{iff} \\ !_{S'}(s_i) \models \mu_\Sigma(\sigma) & \text{iff} & s_i \models \mu_\Sigma(\sigma) & \end{array}$$

and hence $up_{S'}(s_i)(\sigma) \in S'$ iff $s_i \models \mu_\Sigma(\sigma)$.

Definition 3 (Right adjoint). *We define* $R : AppUpCoalg \to Coalg(F) \times Coalg(G)$ *by* $R(Ac,S) = ((S,\langle ap_S, val_S\rangle),(Ac,\langle ap_{Ac}, cont_{Ac}\rangle))$*, where the map* $cont_{Ac} : Ac \to \mathcal{P}(\Psi)$ *takes an action* $a \in Ac$ *to the set of states in the final* F*-coalgebra which satisfy the formula* $\mu_{Ac}(a)$*.*

Informally speaking, the functor R takes a pair consisting of an H-coalgebra Ac and an appearance-update T_{Ac}-coalgebra S, and produces an F-coalgebra and a G-coalgebra. The F-coalgebra is obtained from S by forgetting its update map and keeping everything else intact. The G-coalgebra has the same carrier set and epistemic structure as Ac, and a content map obtained essentially by replacing content formulas with their denotations in the final F-coalgebra.

Theorem 1. *L is left adjoint to R.*

Proof. We begin by examining the unit and counit of this adjunction. Since the categories $Coalg(H)$ and $Coalg(G)$ are naturally isomorphic, it is the move from T_{Ac}-coalgebras to F-coalgebras and back that makes the adjunction non-trivial.

For the unit of the adjunction, the inclusions $\eta_{S,\Sigma} : S \to S \cup (S \otimes \Sigma) \cup ((S \otimes \Sigma) \otimes \Sigma) \cup \ldots$ together with the natural isomorphism between $Coalg(H)$ and $Coalg(G)$ give rise to a natural transformation $\eta : Id_{Coalg(F)\times Coalg(G)} \Rightarrow R \circ L$.

For the counit, the maps $\epsilon_{Ac,S} : S \cup (S \otimes Ac) \cup ((S \otimes Ac) \otimes Ac) \cup \ldots \to S$ defined inductively by

$$\epsilon_{Ac,S}(s) = s, \qquad \epsilon_{Ac,S}(s_i,a) = up_S(\epsilon_{Ac,S}(s))(a) \quad \text{for } i \in \omega \text{ and } s_i \in S_i$$

together with the natural isomorphism between $Coalg(H)$ and $Coalg(G)$, yield a natural transformation $\epsilon : L \circ R \Rightarrow Id_{AppUpCoalg}$.

We show that η and ϵ indeed constitute the unit and counit of an adjunction $L \dashv R$. To this end, we fix $(S, \Sigma) \in Coalg(F) \times Coalg(G)$ and $(Ac, S') \in AppUpCoalg$. For $(f, g) : (S, \Sigma) \to R(Ac, S')$, the map $f^{\#} : S \cup (S \otimes \Sigma) \cup ((S \otimes \Sigma) \otimes \Sigma) \cup \ldots \to S'$ defined inductively by

$$f^{\#}(s) = f(s), \qquad f^{\#}(s_i, \sigma) = up_{S'}(f^{\#}(s_i))(g(\sigma)) \quad \text{for } i \in \omega$$

is a T_{Ac}-coalgebra morphism that satisfies $R(g, f^{\#}) \circ \eta_{(S,\Sigma)} = (f, g)$. Furthermore, any T_{Ac}-coalgebra morphism with the above property is defined in this way. For $(h, k) : L(S, \Sigma) \to (Ac, S')$, the map $k^{\flat} : S \to S'$ given by $k{\restriction}_S$ defines an F-coalgebra morphism that satisfies $\epsilon_{(Ac,S')} \circ L(k^{\flat}, h) = (h, k)$. Furthermore, this last requirement uniquely determines the definition of $k^{\flat}$.

5 Coalgebraic Dynamic Epistemic Logic

Coalgebras give rise to modal logics in different ways, for example the coalgebraic logic of Moss [13], the temporal logic of Jacobs [11], and the modular logic of Cîrstea and Pattinson [5,6]. In previous work [15], we showed how one obtains an algebraic logic from our functor by predicate lifting, and investigated the connection between this logic and the algebraic dynamic epistemic logic of [2,14]. Cîrstea and Pattinson have shown how complete and expressive coalgebraic logics can be derived in a modular fashion for an inductively-defined class of endofunctors on Set. By applying this method to our setting, we obtain a logic with a multi-sorted syntax, which is expressive – that is, two states are bisimilar if and only if they satisfy the same formulas –, and admits a sound and complete proof system. Because of the particular shape of the functor T and of the axioms in the associated proof system, the multi-sorted syntax of this logic can be simplified to the following single-sorted syntax, with no loss in expressiveness

$$\phi ::= tt \mid p \mid \neg\phi \mid \phi \wedge \phi \mid \Box_A \phi \mid [a]\phi$$

The standard knowledge and dynamic modalities, that is, $\Box_A$ (to be read as 'A knows ϕ') and $[a]$ (to be read as 'after a, ϕ'), are recovered by letting $\Box_A \phi ::= [\pi_1][A]\Box\phi$ and $[a]\phi ::= [\pi_2][a][\kappa_2]\phi$ [9]. In particular, the statement 'action a does not go through' is captured by the formula $[a]\mathsf{ff}$. Using the simplified syntax, the original proof system is equivalent to the following set of axioms and rules

$$\vdash \bigcirc tt \qquad \vdash \bigcirc\phi \wedge \bigcirc\psi \to \bigcirc(\phi \wedge \psi) \qquad \frac{\vdash \phi \to \psi}{\vdash \bigcirc\phi \to \bigcirc\psi}$$

for $\bigcirc \in \{\Box_A, [a]\}$, and

$$\vdash [a](\phi \vee \psi) \to [a]\phi \vee [a]\psi$$

on top of propositional logic[10]. As a consequence of the results in [6], this proof system is sound and complete w.r.t. T-coalgebras. However, in order to formulate

[9] See [6] for details of the multi-sorted syntax.

[10] As in [6], we include all instances of propositional tautologies and the modus ponens rule into our set of axioms and rules.

a soundness and completeness result w.r.t. appearance-update coalgebras, the restrictions defining appearance-update coalgebras must also be axiomatised. To this end, we add the following axioms to the previous proof system:

$$\vdash [a]\, p \leftrightarrow (\neg[a]\mathsf{ff} \to p) \qquad \vdash [a]\Box_A\, \phi \leftrightarrow (\neg[a]\mathsf{ff} \to \bigwedge_{a' \in Ac_{a,A}} \Box_A\, [a']\, \phi)$$

$$\phi_a \leftrightarrow \neg[a]\mathsf{ff}$$

where for an action $a \in Ac$, its content is denoted by ϕ_a. There is one such axiom for each epistemic action a and each (type of) agent A.

Example of Derivation. Consider a simple Man in the Middle Attack: agent A sends a message with factual content p to agent B, but on the way the intruder C changes p to another fact p' and thus B receives p' instead. If we assume that A does not suspect the interception, after sending p he believes that B believes in p. Similarly, upon receipt, B believes that A believes in p'. In security terms and since A and B do not suspect the interception, they will wrongly *authenticate* with each other. We use the encoding of the security action in Section 2 to prove that $\vdash [p \star p'_{A,B,C}]\Box_A\Box_B p$. The proof steps are sketched below:

	$\vdash \mathsf{tt}$
(propositional logic)	$\vdash \neg[p \star p'_{A,B,C}]\mathsf{ff} \to \mathsf{tt}$
(modular logic)	$\vdash \neg[p \star p'_{A,B,C}]\mathsf{ff} \to \Box_A\mathsf{tt}$
(propositional logic)	$\vdash \neg[p \star p'_{A,B,C}]\mathsf{ff} \to \Box_A(\neg[p!_B]\mathsf{ff} \to \mathsf{tt})$
(modular logic)	$\vdash \neg[p \star p'_{A,B,C}]\mathsf{ff} \to \Box_A(\neg[p!_B]\mathsf{ff} \to \Box_B\mathsf{tt})$
(propositional logic)	$\vdash \neg[p \star p'_{A,B,C}]\mathsf{ff} \to \Box_A(\neg[p!_B]\mathsf{ff} \to \Box_B(p \to p))$
(content axiom)	$\vdash \neg[p \star p'_{A,B,C}]\mathsf{ff} \to \Box_A(\neg[p!_B]\mathsf{ff} \to \Box_B(\neg[p!_B]\mathsf{ff} \to p))$
(preservation of facts axiom)	$\vdash \neg[p \star p'_{A,B,C}]\mathsf{ff} \to \Box_A(\neg[p!_B]\mathsf{ff} \to \Box_B[p!_B]p)$
(rationality for B wrt $p!_B$)	$\vdash \neg[p \star p'_{A,B,C}]\mathsf{ff} \to \Box_A[p!_B]\Box_B p$
(rationality for A wrt $p \star p'_{A,B,C}$)	$\vdash [p \star p'_{A,B,C}]\Box_A\Box_B p$

With the additional axioms, we obtain the following result:

Theorem 2 (Soundness and Completeness). *A formula holds in all appearance-update coalgebras if and only if it is derivable in the appearance-update logic.*

Proof. The proof of both soundness and completeness is detailed in [8]. Here we only sketch the completeness proof. This follows the same line as the completeness result for dynamic epistemic logic [3], and is based on a translation between our appearance-update logic (with appearance-update coalgebras as models) and ordinary epistemic logic (with F-coalgebras as models). As in [3], this translation has the property that a formula ϕ is semantically equivalent to its translation ϕ^t. Moreover, the axioms and rules of appearance-update logic ensure that $\vdash \phi \leftrightarrow \phi^t$. These properties, together with our result in [8] that the

final F-coalgebra can be extended to an appearance-update coalgebra, allow us to make use of the completeness result of [6] for F-coalgebras in order to prove completeness of appearance-update logic w.r.t. appearance-update coalgebras.

Acknowledgement. We would like to thank the anonymous referees for valuable suggestions on improving the paper.

References

1. Baltag, A.: A coalgebraic semantics for epistemic programs. In: Proceedings of Coalgebraic Methods in Computer Science. Electronic Notes in Theoretical Computer Science, vol. 82 (2003)
2. Baltag, A., Coecke, B., Sadrzadeh, M.: Epistemic actions as resources. Journal of Logic and Computation, forthcoming
3. Baltag, A., Moss, L.S.: Logics for epistemic programs. Synthese 139 (2004)
4. van Benthem, J., Pacuit, E.: The tree of knowledge in action: towards a common perspective. In: Proceedings of Advances in Modal Logic (2006)
5. Cîrstea, C.: A compositional approach to defining logics for coalgebras. Theoretical Computer Science 327(1), 45–69 (2004)
6. Cîrstea, C., Pattinson, D.: Modular construction of modal logics. In: Gardner, P., Yoshida, N. (eds.) CONCUR 2004. LNCS, vol. 3170, pp. 258–275. Springer, Heidelberg (2004)
7. Cîrstea, C.: On expressivity and compositionality in logics for coalgebras. In: Proceedings of Coalgebraic Methods in Computer Science, Electronic Notes in Theoretical Computer Science 82 (2003)
8. Cîrstea, C., Sadrzadeh, M.: Coalgebraic epistemic update without change of model `http://ecs.soton.ac.uk/~ms6/TechRep.pdf`
9. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.Y.: Reasoning about Knowledge. MIT Press, Cambridge (1995)
10. Gerbrandy, J.: Bisimulations on Planet Kripke. Ph. D. Thesis, University of Amsterdam (1999)
11. Jacobs, B.: The temporal logic of coalgebras via Galois algebras. Mathematical Structures in Computer Science 12, 875–903 (2002)
12. Jacobs, B.: Many-sorted coalgebraic modal logic: a model-theoretic study. Theoretical Informatics and Applications 35, 31–59 (2001)
13. Moss, L.S.: Coalgebraic logic. Annals of Pure and Applied Logic 96, 241–259 (1999)
14. Sadrzadeh, M.: Actions and Resources in Epistemic Logic. Ph.D. Thesis, University of Quebec at Montreal (2005), `http://www.ecs.soton.ac.uk/~ms6/all.pdf`
15. Sadrzadeh, M., Cîrstea, C.: Relating algebraic and coalgebraic logics of knowledge and update. In: Proceedings of the 7th conference on Logic and the Foundations of Game and Decision Theory, pp. 199–208, Liverpool (July 2006)
16. Rutten, J.J.M.M.: Universal coalgebra: a theory of systems. Theoretical Computer Science 249, 3–80 (2000)