The Algebra of Multi-Agent Dynamic Belief Revision

Alexandru Baltag
Oxford University Computing laboratory
baltag @comlab.ox.ac.uk

Mehrnoosh Sadrzadeh
Université du Québec À Montréal
sadrzadeh.mehrnoosh@courrier.uqam.ca

1 Introduction

We refine our algebraic axiomatization in [8, 9] of epistemic actions and epistemic update (notions defined in [5, 6] using a relational, Kripke-style semantics), to incorporate a mechanism for dynamic belief revision in a multi-agent setting. Our approach has a number of novel features, when compared with traditional belief revision systems such as AGM [2]. Firstly: while traditional belief revision was dealing only (the revision of theories comprised of) ”facts”, we can also revise (theories comprised of) epistemic/doxastic propositions; thus, some of traditional AGM postulates have to be modified in order to deal with non-stable epistemic propositions, such as the ones generated by the so-called *Moore sentences*[1]. Secondly, ours is a *dynamic* belief revision, i.e. a particular form of update: an action q happens, changing the original theory m to a revised theory $m \star q$. Thirdly: this dynamic character is enhanced by the fact that we allow *revision of actions, and not just of ”static” theories*; the agent has theories about the current action, and these may be revised by further actions. Fourthly: we do *multi-agent* belief revision. Our approach, though related in aim, is different in flavour from the work of [14] on Kripke-style belief revision and from the more recent work of [3, 15] on dynamic belief revision: while these approaches are *semantic* and *”quantitative”* (being based on having ”degrees of belief” as the quantitative basis of belief revision), our approach is purely *qualitative* and *axiomatic*, being thus closer in spirit to the traditional AGM approach. Indeed, our aim is simply to find the ”correct” axiomatization of the dynamic (and multi-agent) version of AGM.

2 Dynamic Revision Systems

We introduce here a *dynamic analogue* of the classical AGM axioms of belief revision. For the moment, we neglect the agents doing the revision, concentrating (as in the AGM approach) on *a purely impersonal notion of rational belief revision*. In the next section, we will re-introduce the agents (and their views of the world) into the picture. Our setting is based on the notion of *system*, structure composed of a pair quantale-module (M, Q) linked by an action $\otimes : M \times Q \to M$, subject to some conditions. Systems are very general algebraic settings for modelling dynamical phenomena. Quantales have been used to study different phenomena such as models of concurrent processes in [1] and physical properties of Quantum Mechanics in [10]. They have also been studied as models of Dynamic Logic [12]. For a detailed mathematical study of quantales and their properties refer to [13].

Definition 2.1 *A system* $(\mathbf{M}, \mathbf{Q}, \otimes)$ *consists of: (1) a complete sup-lattice* $\mathbf{M} = (M, \bigvee)$ *(with* $\bigvee$ *as the supremum operator); (2) a quantale* $\mathbf{Q} = (Q, \bigvee, \bullet, 1)$*, i.e. a complete sup-lattice* $(Q, \bigvee)$ *endowed*

[1] E.g. the agent (”you”) is informed that the ”The cat is on the mat, but you don’t know it”.

with an additional monoidal structure $(Q, \bullet, 1)$*; (3) a right-module structure on* $\mathbf{M}$*, i.e. an action* $- \otimes - : M \times Q \to M$ *on the quantale, which is sup-preserving in both arguments and satisfies:*

$$m \otimes (q \bullet q') = (m \otimes q) \otimes q' \text{ for } m \in M \text{ and } q, q' \in Q$$

We call the elements of the module M *theories*, or *propositions*, or also *situations*. These situations are not neccessarily deterministic, as our theories might not be complete. There exists exactly *one inconsistent theory*, namely $\perp := \bigvee \emptyset$. Similarly, there exists exactly one trivial (tautological) theory, namely $\top := \bigvee M$. In a possible-world model, elements of M can be modelled as *sets of possible worlds*, representing some theory (or belief)[2] about the real world: it is believed[3] that the real world belongs to this set. However, one can also represent theories in an *intensional way* (as it is standard in belief revision), as *sets of sentences (in some language), closed under logical consequence*[4]. The lattice order $m \leq m'$ on the module M represents the *consequence relation* (logical entailment) $T \vdash T'$ between theories. In terms of theories as sets of possible worlds, this is simply *set inclusion* $\subseteq$. But, in terms of deductively closed theories as sets of sentences, it is *reversed inclusion* $T \supseteq T'$. So, in particular, our join $m \vee m'$ corresponds to union in terms of sets of possible worlds, i.e. is logical disjunction of propositions; but notice that it also corresponds to *intersection* $T \cap T'$ of sets of sentences; dually, our meet $m \wedge m'$ is the logical *conjunction* of propositions, i.e. the intersection of the two sets of worlds; but it also corresponds to the *deductive closure of the union* $\mathcal{C}l(T \cup T')$ of the (sets of sentences of the) two theories.

We call the elements of the quantale Q *epistemic actions*, or *experiments*, or also *announcements*. They are information-changing actions, which *do not change the objective facts of the world, but only uncover, discover, or communicate them.* Think of elements of quantale as scientific experiments, or even better as actions of *communicating the results of an experiment*. More generally, they may cover *any communication*, of *facts* or also of *hypothesis about the world, or even about what is (not) known about the world* etc. Since these actions are not necessarily deterministic, they might also be thought of as "theories about (the current) action(s)": there might be uncertainty about what exactly has been discovered, or announced. The lattice order $q \leq q'$ on the quantale Q is the order of non-determinism: q' contains less information about the current action (and thus is more non-determined) than q. The quantale multiplication $q \bullet q'$ represents *sequential composition* of actions: *first do* q*, then do* q'. The unit 1 of multiplication stands for the *skip* ("do-nothing") action.

Observe the lack of a Boolean structure: there is no negation in M, nor in Q. One can of course define some pseudo-complement, as is usually done in complete lattices, but this will not necessarily behave as a Boolean negation. This lack is the result of a conscious choice: we think this algebraic setting is simpler, more transparent and more general (than the one that we would obtain by adding negation), being applicable to cases of partial information.[5] In particular, when thinking of M as consisting of "theories", there is no meaningful notion of "negation of a theory"; similarly, there is no action that can be called "the negation" of action q. (However, as we show bellow, there is a natural notion of a static situation m representing in some sense a "negation" of an action!)

The operation $\otimes$ is called *update product*, and it encodes the way in which an action *changes* a situation or theory: if a theory m correctly describes the current situation and q is a correct description of the current action that is happening, then the theory $m \otimes q$ will correcty describe the situation after the action.

[2]This theory may be of course held by some (arbitrary) agent. This agent is left implicit here, but in the next section we will introduce it explicitly.

[3]See the previous footnote.

[4]Algebraically, theories in this sense correspond to *filters* in a complete lattice, or Boolean algebra.

[5]However, it would be easy to extend our setting to intuitionistic or Boolean settings, by requiring the module to be a Heyting or a Boolean Algebra.

Kernel. In case that the resulting situation is inconsistent, i.e. $m \otimes q = \bot$, then this simply means that the *action* q *is impossible in (the situation described by)* m. The condition of (im)possibility of a given action can be encoded as a proposition in M, by defining the *kernel* of action q as: $ker(q) = \bigvee\{m \in M \mid m \otimes q = \bot\}$ .This is the weakest theory/situation that makes this action impossible. The kernel can be thought of as a "strong negation of an action": $ker(q)$ is the proposition asserting the *impossibility of action* q.[6]

Image Maps. Any action $q \in Q$ "acts" on M via the update product $\otimes$, and "acts" on Q itself via multiplication $\bullet$. As a consequence, we can define two notions of *image of an action*: $Im^M(q) = \{m \otimes q : m \in M\}$ and $Im^Q(q) = \{q' \bullet q : q' \in Q\}$.

Atoms. If existing, the *atoms*[7] of the module M are called "states", and can be thought of as *complete (fully determined) situations*, i.e. complete descriptions of the world. In terms of sets of states (or possible worlds), they simply are the states (or rather the singleton-sets consisting of only one state). In terms of theories, they are the *complete consistent* (i.e. maximally consistent) theories.[8] Similarly, the *atoms of the quantale* Q represent *deterministic* actions (or, alternatively, complete theories about actions).

Atomicity. A system is called *atomistic* if both the module M and the quantale Q are atomistic lattices[9] and if any update product or multiplication of atoms is either an atom or inconsistent, i.e.:

$$s \in Atm(M), \sigma, \sigma' \in Atm(Q) \text{ imply } s \otimes \sigma \in Atm(M) \cup \{\bot\} \text{ and } \sigma \bullet \sigma' \in Atm(Q) \cup \{\bot\}$$

This condition expresses a natural property of determinism: a deterministic action acts as a partial function on states, i.e. it transforms any (fully determined) state into a (at most one, fully determined) state (or else, it fails); and the sequential composition of two deterministic actions (if not failing) is itself a deterministic action.

Finally, a system is called *extensional* if actions are uniquely determined by their behavior on situations, i.e.: if $m \otimes q = m \otimes q'$ for all $m \in M$ then $q = q'$. As a consequence of this, we have that: $ker(q) = \top \rightarrow q = \bot$. This is because if the kernel of an action q is $\top$ then the action cannot be applied to any proposition. That is if $ker(q) = \top$ then for all $m \in M$ we have $m \otimes q = \bot$. But epistemic update is bottom-preserving in the sense that $\bot = m \otimes \bot$. So we have $m \otimes q = m \otimes \bot$ and if the system is extensional we get $q = \bot$.

"Facts". Since we think of our actions as "purely epistemic" (i.e. actions of discovery, belief-change or communication), they do not affect the "objective" features of the world, they do not change the "facts". We can turn things around, by defining "facts" to be the propositions that are invariant under any actions, i.e. the ones that are "stable" under any update. In other words, the set of "facts" is defined as the "stabilizer" of all actions: $Stab(Q) = \{m \in M \mid \forall q \in Q, m \otimes q \leq m\}$. Note that in our system any proposition that is invariant under actions is called a fact. For example since update preserves all joins including the empty join $\vee\emptyset = \bot$, we have $\bot \otimes q = \bot$, which says bottom is invariant under any action and thus a fact. But it is a fact in which we are not interested, since it is wrong!

Dynamic Modalities. Since the update product preserves joins on both of its arguments, the join-preserving operation has a (Galois) right adjoint[10] $[q]m'$, which is the standard *dynamic (action) modality* of Propositional Dynamic Logic or PDL [12], definining the *weakest precondition* that ensures the

[6] In the "concrete" Kripke-model-based setting of [5, 6], the role of the kernel was played by its complement, the *precondition* $pre(q)$ of an action, defining its conditions of possibility.

[7] In a lattice M, the set of *atoms* is defined as $Atm(M) = \{a \in M : \forall m \in M(\bot \neq m \leq a \rightarrow m = a)\}$.

[8] Algebraically, these can be the *ultrafilters* in a Boolean algebra.

[9] A lattice M is *atomistic* if every element is the sup of all the atoms bellow it, i.e. $m = \bigvee\{s \in Atm(M) : s \leq m\}$ for all $m \in M$.

[10] defined as $[q]m' := \bigvee\{m \in M : m \otimes q \leq m'\}$.

fulfillment of (the post-condition) m' after action q. Note that the adjunction $-\otimes q \dashv [q]-$ implies the equivalence:

$$m \otimes q \leq m' \text{ iff } m \leq [q]m'$$

Residuals. The *right residual* of multiplication $\bullet$ is the right adjoint of the join-preserving operation $-\bullet q' : Q \to Q$, denoted by $q/q' := \bigvee\{q'' \in Q : q'' \bullet q' \leq q\}$

We enrich now the notion of system with a way to *revise old theories (about the world or about actions) in the view of new experiments*.

Definition 2.2 *A dynamic revision operator* on a system $(\mathbf{M}, \mathbf{Q}, \otimes)$ is a pair $* = (*^M, *^Q)$ of maps $*^M : M \times Q \to M$, $*^Q : Q \times Q \to Q$, satisfying a list of conditions (to be given). In practice, we skip the superscripts whenever possible. The required conditions are the following:

1. $m * q \in Im^M(q)$, and $q * q' \in Im^Q(q')$
2. $m \otimes q \leq m * q$ and $q \bullet q' \leq q * q'$
3. if $m \otimes q \neq \bot$ then $m * q \leq m \otimes q$; similarly: if $q \bullet q' \neq \bot$ then $q * q' \leq q \bullet q'$
4. if $m * q = \bot$ then $q = \bot$; similarly: if $q * q' = \bot$ then $q' = \bot$
5. $m * (q * q') = (m * q) * q'$ and $q * (q' * q'') = (q * q') * q''$

A *dynamic revision system* is an extensional, atomistic system endowed with a dynamic revision operator.

To explain the above conditions, note that they are a dynamic equivalent of the classical AGM conditions. The intuition about $m * q$ is that, *if m is the (possibly incorrect) theory about the world, then $m * q$ represents the (”rationally”) revised theory after the experiment q is performed*. We call this ”dynamic” revision, since $m * q$ is the actual theory *about the state of the world after the action* q: it is *not* a revised theory about the original state of the world before the action.[11] In the case the original theory is consistent with the experiment q (i.e. when $m \otimes q \neq \bot$), the theory is simply ”updated” (using update product) with the action q. Otherwise, the theory m *has been refuted* by the experiment q, and it has to be revised. The revised theory $m * q$ *has to be consistent with the experiment*: $m * q \in Im^M(q)$, i.e. there exists some $m' \in M'$ such that $m * q = m' \otimes q$. This says that the new theory has to be conceivable as being the result of updating some previously existing situation m' with the actual experiment q'. Similarly, the intuition about $q * q'$ is that, *if q is a (possibly incorrect) theory about what action is going on, then $q * q'$ represents the revised theory (about what is going on) after the experiment q' is performed*. In the case the original action theory is consistent with the experiment, the two action theories are simply composed sequentially in $q \bullet q'$: the new theory says that action q followed by action q' has been going on. In any case, the revised theory has to be consistent with the experiment: $q * q' \in Im^Q(q')$, i.e.: the revised action theory should be that some hypothetical previous action q'' happened, followed by the experiment q'. Finally, the last condition[12] (imposing two types of transitivity of revision) is a dynamic version of the classic AGM condition about revision with a conjunction $\varphi \wedge \psi$. Since we have two types of revisions: one for propositions and one for actions, we need to relate the two. The fifth axiom is thus about the consistency of revision of actions with revision of propositions. It says if we revise a proposition m with an action q that has itself been revised by another action q' we

[11] In contrast, classical ”static” belief revision deals with revised theories about the original state of the world.

[12] Technically speaking, we will not need this last condition, since all its instances that are relevant to us will follow from the axioms on belief revision endomorphisms in the next section. .

get the same revision as when we first revise m with q and then with q'. It can also be seen as a way of defining revision of actions $q * q'$ in terms of revision of propositions.

In fact, the first clause in the above definition can be strengthened to an apparently stronger (but in fact equivalent) one:

Proposition 2.3 *Any dynamic revision operator satisfies the condition:*

$$1'. \quad m * q = ([q](m * q)) \otimes q \text{ and } q * q' = ((q * q')/q') \bullet q'$$

Proof. The proof is straightforward using two of the previously mentioned Galois adjunctions: the one between the update and dynamic modality $-\otimes \dashv [q]-$ and another one between the sequential composition and residuals $- \bullet q' \dashv - \setminus q'$. For instance to prove $m * q = ([q](m * q)) \otimes q$ start from the trivial inequality $(m * q) \otimes q \leq (m * q) \otimes q$ and then apply the update adjunction definition $m \otimes q \leq m'$ iff $m \leq [q]m'$ to get $(m * q) \leq [q]((m * q) \otimes q)$. For the other side start from $[q](m * q) \leq [q](m * q)$ and apply the same adjunction to get $([q](m * q)) \otimes q \leq m * q$. The proof of $q * q' = ((q * q')/q') \bullet q'$ is similar. □

This proposition says that $([q](m * q)) \otimes q$ is a possible or tentative belief that the agent might have after the action q, about the prior situation (i.e. the situation before q had happened). In fact $[q](m * q)$ is the weakest such belief[13], which is consistent with of the agent with which the revision $m * q$. As such, $[q](m * q)$ can be seen as a tentative justification of this particular belief revision $m * q$.

3 Multi-Agent Dynamic Belief Revision

We want to introduce in the picture *agents* and their views about the world (and about actions).

Definition 3.1 *A dynamic-revision endomorphism*[14] *on a dynamic revision system* $(M, Q, \otimes, *)$ *is a pair* $f = (f^M, f^Q)$ *of maps* $f^M : M \to M$, $f^Q : Q \to Q$, *satisfying some extra-conditions (to be given below). As before, we skip the superscripts whenever possible. The conditions are:*

1. *Both* f^M *and* f^Q *are sup-morphisms (i.e. sup-preserving maps), and* $f^Q(1) = 1$
2. *For atoms* $s \in Atm(M), \sigma \in Atm(Q)$ *such that* $s \otimes \sigma \neq \bot$, *we have*
$$f(s \otimes \sigma) = f(s) * f(\sigma)$$
3. *For atoms* $\sigma, \sigma' \in Atm(Q)$ *such that* $\sigma \bullet \sigma' \neq \bot$, *we have*
$$f(\sigma \bullet \sigma') = f(\sigma) * f(\sigma')$$

Intuitively, we think of $f(m)$ as the theory that some (unspecified) agent has about the situation described by m: when the real situation is given m, the agent believes that this situation is given by (his theory) $f(m)$. Similarly, we think of $f(q)$ as the theory that the (unspecified) agent has about the action q.

Definition 3.2 *A dynamic-revision endomorphism* f *is said to be doxastic (or* $D45$*) if it satisfies the following additional conditions:*

[13]Note that there may be many such beliefs (about the prior state) consistent with a given revision.

[14]Note that this notion differs essentially from the notion of *epistemic endomorphism* in [8, 9]: in the conditions corresponding to the last two conditions above, the epistemic endomorphisms had update product $\otimes$, and respectively composition $\bullet$, instead of dynamic revision $*$.

D. ("Consistency of beliefs"): $ker(f) = \bot$

45. ("Introspection"): $f \circ f = f$

A doxastic dynamic-revision endomorphism is also called an appearance map.

The intuition is that the theory $f(m)$ that the (unspecified) agent has about the situation m gives the "appearance" of this situation to the agent (or the "view" that the agent on this situation). Axiom D says that, in any consistent situation $m \neq \bot$, the agent has *consistent* theory $f(m) \neq \bot$. Axiom 45 says that the apearance mp is idempotent, i.e. the agent is introspective: the theory $f(f(m))$ that he has about (the situation represented by) his own theory $f(m)$ coincides with his theory (i.e. $f(f(m)) = f(m)$). This means that the agent knows what he believes and what he does not believe.
Note: The notion of epistemic endomorphism in [8, 9] was not required to satisfy the $D45$ conditions. Moreover, in that context, condition D was *not satisfied* by the examples, since (in specific examples) it was contradicted by the nature of update product: agents may indeed come to inconsistent believes when using the update product of [5, 6] to update their beliefs! The reason, again, was the absence of a mechanism for belief revision.

Definition 3.3 *A (multi-agent dynamic) belief revision system* $(M, Q, \otimes, *, \{f_A\}_{A\in\mathcal{A}})$ is a dynamic revision system $(M, Q, \otimes, *)$ endowed with a family of appearance maps (=doxastic dynamic-revision system endomorphisms), indexed by a set of "agents" $A \in \mathcal{A}$.

The element $f_A(m) \in M$ is called *the appearance of (situation)* m *to agent* A, or the *theory of* A *about* m. Different agents may have different *views* of the same situation, or different *interpretations* of the same theory. Similarly, $f_A(q)$ is *the appearance of action* q *to agent* A. In other words, $f_A(q)$ represents what agent A thinks is going on when in reality action q is going on. So for instance different appearance maps $f_A(q)$ might represent different interpretations (or different) views of the experiment q. An experiment might be public or private, its results might be communicated only to some of the agents, or some outsiders might be deluded by their dogmatic beliefs into rejecting, misunderstanding or misinterpreting the experiment.

The essential difference between this notion and the notion of *epistemic system* in [8, 9] comes from the clauses for the appearance of an updated situation, and of a composition of actions:

$$f_A(s \otimes \sigma) = f_A(s) * f_A(\sigma)$$

means that the theory A has about the world after the experiment σ is obtained by *dynamically revising* A's old theory about the world $f_A(s)$ with A's theory $f_A(\sigma)$ about the experiment. This means that, in case the experiment appears to contradict the old theory, the contradiction is solved by the agent in favour of the experiment (using the dynamic revision operator). Similarly, the identity

$$f_A(\sigma \bullet \sigma') = f_A(\sigma) * f_A(\sigma')$$

means that the theory A has about the composed action $\sigma \bullet \sigma'$ is obtained by dynamically revising A's theory about the first action with A's theory about the second action.

Belief. The *belief* modality $\Box_A$ can be defined as the *right-adjoint of the appearance map*: $f_A(-) \dashv \Box_A -$. Indeed, since the appearance maps are join-preserving maps, they have meet-preserving Galois right adjoints[15]. Formally that is to say $f_A(m) \leq m'$ iff $m \leq \Box_A m'$. We read $\Box_A m$ as "agent A

[15]definable as: $\Box_A m' = \bigvee\{m \in M : f_A(m) \leq m'\}$.

believes theory m ". A similar notion $\Box_A q$ can be defined for actions. One can easily see that the belief modality has the properties of a *normal modality*, and moreover (due to the $D45$ conditions above), it satisfies "Introspection" ($\Box_A m = \Box_A \Box_A m$) and "Consistency of beliefs" ($\Box_A \bot = \bot$).

Multi-Agent Learning Systems. To deal with applications, it is useful to enrich our structure a bit, allowing it to deal with *positive and negative tests*, and with *learning actions* (and as a consequence, with public/private announcements and refutations).

Definition 3.4 *A multi-agent learning system* is a multi-agent dynamic belief revision system $(M, Q, \otimes, *, \{f_A\}_{A\in\mathcal{A}})$, endowed with maps $? : M \times M \to Q$ and $L_{\mathcal{B}} : Q \to Q$, for each set $\mathcal{B} \subseteq \mathcal{A}$ of agents. These maps satisfying certain conditions (to be given below). As a notation, we put $m?\overline{n} :=?(m,n)$. We read $m?\overline{n}$ as the experiment: *"test" m and refute n*. This is a *non-epistemic*, PDL-like test: theory m is "tested", while theory n is "refuted", but *without any of the relevant agents being announced* (of the result of the test/refutation). The action $L_{\mathcal{B}}q$ is read as *learning of action q by group $\mathcal{B}$* . More precisely, it can be described as: "while an action q is happening, the agents in group $\mathcal{B}$ privately get together and learn (in common, by mutual update) that q is happening". This learning action is so private that none of the other agents $C \in \mathcal{A} \setminus \mathcal{B}$ suspect that it is happening. The required conditions are:

- If $s \in Atm(M)$ then $s \otimes (m?\overline{n}) = \begin{cases} s & s \leq m, s \not\leq n \\ \bot & \text{otherwise} \end{cases}$
- $f_A(m?\overline{n}) = 1$
- $ker(L_{\mathcal{B}}q) = ker(q)$, $f_B(L_{\mathcal{B}}q) = L_{\mathcal{B}}q$ for $B \in \mathcal{B}$, and $f_C(L_{\mathcal{B}}q) = 1$ for $C \in \mathcal{A} \setminus \mathcal{B}$.

The first clause says that a state survives a test/refutation $m?\overline{n}$ iff it is satisfies the tested property but it doesn't satisfy the refuted property; in which case the state is left unchanged by the state. The second clause says the appearance of a pure test/refutation action to all agents is 1, i.e. the action $skip$ in which "nothing happens"; finally, the third clause says that learning an action q is (im)possible iff the learned action (q) is (im)possible, that the learning action appears as "learning" to all the agents involved in it, and it appears as $skip$ to all the outsiders.

Public/Private Announcements/Refutations. In a multi-agent learning system, we can define a *mutual announcement-and-refutation to a group $\mathcal{B}$* of agents, by putting: $m!_{\mathcal{B}}\overline{n} := L_{\mathcal{B}}(m?\overline{n})$. Notice that this is a "combined" action, in which something (m) is mutually announced, while something else (n) is mutually refuted. More "pure" special cases of this are *mutual announcements* $m!_{\mathcal{B}} := m!_{\mathcal{B}}\overline{\bot}$ and public refutations $!_{\mathcal{B}}\overline{m} := \top!_{\mathcal{B}}\overline{m}$. Even more special cases are: *public announcements* $m! := m!_{\mathcal{A}}$ and public refutations $!\overline{m} :=!_{\mathcal{A}}\overline{m}$; and *private announcements to an agent* $m!_A := m!_{\{A\}}$ and *private refutations* $!_A\overline{m} :=!_{\{A\}}\overline{m}$.

Examples: Secret, Secure Communication. Secret Interception. The *passing of a secret (truthful) message m from A to B* (over a secure channel) is represented by the action $m!_{A,B}$. If the channel is not really secure and in fact *the message is secretely intercepted (and read, but allowed to go further) by an agent C*, then the action is represented by $L_C(m!_{A,B})$.

Using the definition of the learning action, we can calculate the kernels of the actions defined above. For example we can prove that the refuted proposition of a public refutation constitutes its kernel. That is, a public refutation cannot be applied on the states where the refuted proposition holds. In other words:

Lemma 3.5 *For a public refutation $!\overline{m}$ and an atom $s \in Atm(M)$ we have $s \otimes !\overline{m} = \bot$ iff $s \leq m$.*

On the other hand, for a public announcement, every proposition but the announced one is in the kernel. That is a public announcement action can only be applied on the states where the announced proposition holds. In other words:

Lemma 3.6 *For a public announcement* $m!$ *and an atom* $s \in Atm(M)$ *we have* $s \otimes m! \neq \bot$ *iff* $s \in m$.

Proofs are easy and follow directly from the definition of the kernel of the learning action and how public announcements and refutation actions are defined in terms of the learning action.

4 Example: ”Cheating Muddy Children”, with belief revision.

Consider the original muddy children puzzle fully discussed in [11]. The puzzle has been solved analyzed using the update product in [5, 6]. Following the dynamic approach, it has also been algebraically dealt with, in its general version, using the algebraic setting of epistemic systems [8, 9]. For the purpose of this abstract, we shall deal with with four children, three of whom are dirty, in a belief revision system. Suppose children 1, 2, and 3 are the dirty ones and child 4 is clean. We assume given a learning system which encodes the puzzle as follows. The children $\{1, 2, 3, 4\} \subseteq \mathcal{A}$ are in the set of agents. The module M includes some atomic situations (”states”) $s_{\mathcal{B}}$, for each set $\mathcal{B} \subseteq \mathcal{A}$ of agents. Each $s_{\mathcal{B}}$ represents the situation in which the dirty children are precisely the ones in the group $\mathcal{B}$. For instance $s_{1,2,3} \in Atm(M)$ represents the real state, in which the first three children have dirty foreheads. We also have some stable propositions $D_i \in Stab(Q) \subseteq M$ (for $i = 1, 2, 3, 4$), saying ‘child i is dirty’, and some stable propositions $\overline{D}_i \in Stab(Q)$, saying ‘child i is clean’. These are “facts” that cannot be changed by epistemic actions, so that is why we assume them to belong to $Stab(Q)$. Each state satisfies its corresponding facts, i.e. we put $s_{\mathcal{B}} \leq D_i$ iff $i \in \mathcal{B}$, and $s_{\mathcal{B}} \leq \overline{D}_i$ otherwise. In any state $s_{\mathcal{B}}$, each child sees the faces of other children but not his own, so he doesn’t know if he’s dirty or not: this is encoded in the appearance maps, by putting $f_i(s_{\mathcal{B}}) = s_{\mathcal{B} \cup \{i\}} \vee s_{\mathcal{B} \setminus \{i\}}$. For example $f_1(s_{1,2,3}) = s_{1,2,3} \vee s_{2,3}$.

Father’s first announcement that “At least some one is dirty” can be represented as a *public announcement* (as defined above) $q_0 := (\vee_{i=1}^{4} D_i)!$, which is assumed to be an atomic element of our quantale Q. By the definition of public announcements, we have $f_i(q_0) = q_0$ for all $1 \leq i \leq 4$, i.e. every child hears this announcement. Every round in which all children answer “I don’t know that I am dirty” is represented by a public refutation (also an atomic element of the quantale) $q :=!\overline{(\vee_{i=1}^{4} \Box_i D_i)}$ (with the refuted proposition as its kernel).

We are interested in a cheating version of the puzzle, originally presented in [4], also discussed in [8]. In this cheating scenario, father’s first announcement and first round of no answers of children go as usual. The state of the system at this moment is

$$s' = s_{1,2,3} \otimes q_0 \otimes q \,.$$

But after the first round, children 2 and 3 cheat by secretly communicating to each other that they are dirty. Their cheating is encoded as a secret message passing $\pi := (D_2 \wedge D_3)!_{\{2,3\}}$ (as discussed before). Notice that this action appears to children 1 and 4 as $skip$: $f_1(\pi) = f_4(\pi) = 1$. The new state of the system after this cheating is

$$s'' := s' \otimes \pi \,.$$

Now the cheating children know they are dirty! Thus, in the second round of answering, while children 1 and 4 proceed as usual (that is they refute that they know they are dirty $!\overline{(\Box_1 D_1 \vee \Box_4 D_4)}$, the cheating children announce that they know they are dirty $(\Box_2 D_2 \wedge \Box_3 D_3)!$. So this second round of answers is a combination of yes and no anwers, that is public refutation of children 1 and 4, and public announcement of children 2 and 3 at the same time. This can be encoded as a *mutual announcement-and-refutation to a group* action $q' = (\Box_2 D_2 \wedge \Box_3 D_3)!\overline{\Box_1 D_1 \vee \Box_4 D_4}$.

After q', it is easy to see that child 1 will wrongly conclude that she is clean! Indeed, as it was shown in [4], one can easily prove (in our dynamic-revision setting, as well as in the algebraic setting of epistemic systems in [8] or in the original Kripke model setting [4]) that

$$s_{1,2,3} \leq [q_0 \bullet q \bullet \pi \bullet q']\Box_1 \overline{D}_1 \,.$$

Sketch of proof. We first use the adjunction between update and dynamic modality to take $[q_0 \bullet q \bullet \pi \bullet q']$ to the left hand side:

$$s_{1,2,3} \otimes (q_0 \bullet q \bullet \pi \bullet q') \leq \Box_1 \overline{D}_1 \,.$$

Similarly, use the knoweldge-appearance adjunction to take the $\Box_1$ to the left hand side:

$$f_1 \left(s_{1,2,3} \otimes (q_0 \bullet q \bullet \pi \bullet q') \right) \leq \overline{D}_1 \,.$$

Since the sequential composition and update of atoms is an atom, we have $(q_0 \bullet q \bullet \pi \bullet q') \in Atom(Q)$ and $(s_{1,2,3} \otimes (q_0 \bullet q \bullet \pi \bullet q')) \in Atm(M)$. A consequence of the conditions two and three of the dynamic revision endomorphism (definition 3.1) tells us that for atoms $s \in Atm(M), \sigma, \sigma' \in Atm(Q)$ we have:

$$f(s \otimes \sigma) \leq f(s) * f(\sigma) \quad \text{and} \quad f(\sigma \bullet \sigma') \leq f(\sigma) * f(\sigma') \,.$$

So we can distribute our f_1 on the tensor and replace the updates and sequential compositions with our revisions operator to get:

$$f_1 \left(s_{1,2,3} \otimes (q_0 \bullet q \bullet \pi \bullet q') \right) \leq f_1(s_{1,2,3}) * f_1(q_0) * f_1(q) * f_1(\pi) * f_1(q') \,,$$

and it would be enough to show

$$f_1(s_{1,2,3}) * f_1(q_0) * f_1(q) * f_1(\pi) * f_1(q') \leq \overline{D_1} \,.$$

We replace the f_1 maps with their values (introduced above as assumptions) and we have to show

$$(s_{1,2,3} \vee s_{2,3}) * q_0 * q * 1 * q' \leq \overline{D_1} \,.$$

According to the lemma 3.6 for the public announcement of the father $q_0 = (\vee_{i=1}^4 D_i)!)$ we have $s_{2,3} \otimes q_0 \neq \bot$ since by assumption $s_{2,3} \leq D_1$. So $(s_{1,2,3} \vee s_{2,3}) \otimes q_0 \neq \bot$ and by axioms 2 and 3 of definition 2.2 we get $(s_{1,2,3} \vee s_{2,3}) * q_0 = (s_{1,2,3} \vee s_{2,3}) \otimes q_0$. By the same line of argument for actions q and q' we get

$$(s_{1,2,3} \vee s_{2,3}) \otimes q_0 \otimes q \otimes q' \leq \overline{D_1} \,.$$

Since update distributes over joins, we now have to prove two disjuncts:

$$s_{1,2,3} \otimes q_0 \otimes q \otimes q' \leq \overline{D_1} \qquad \text{and} \qquad s_{2,3} \otimes q_0 \otimes q \otimes q' \leq \overline{D_1} \,.$$

The first disjunct is proven by induction on the number of children. For the base case of induction we have $k = 0$ and have to prove that $s_\emptyset \otimes q_0 = \bot$. The proof follows from lemma 3.6 since $s_\emptyset \leq D_0$ and $D_0 \not\leq (\vee_{i=1}^4 D_i)$. The induction step is proven in full detail in [9] and as a result of it we have $s_{1,2,3} \otimes q_0 \otimes q \otimes q' = \bot$ since $s_{1,2,3} \otimes q_0 \otimes q \in ker(q')$. Now we have the first disjunct since $\bot \leq \overline{D_1}$. The second disjunct is true since by assumptions we have $s_{2,3} \leq \overline{D_1}$ and $\overline{D_1}$ is a fact. $\Box$

We take this proof to be a "success" of the epistemic system setting: we think it accurately predicts the (most likely) behavior of child 1; moreover, child 1's belief, though wrong, is justified by the appearance of the actions to her.

However, there is a problem with child 4 at this stage: he sees that there are at least three dirty children, so children 2 and 3 *cannot* have come to know so early that they were dirty! Indeed, we can see that there is a contradiction between the announcement q' and 4's beliefs before this announcement: *before* q', the state of the system is $s'' = s \otimes \pi$, but this state's appearance to child 4 is

$$f_4(s'') = f_4(s' \otimes \pi) = f_4(s') * f_4(\pi) = f_4(s') * 1 = f_4(s') .$$

So this state is indistinguishable to child 4 from the previous state (s'), in which the dirty children did not know they were dirty. Hence, child 4 believes that e.g. child 2 does not know that he is dirty: $f_4(s'') = f_4(s') \not\leq \Box_2 D_2$. But this is contradicted by the announcement q', by which child 2 says he knows he is dirty. In other words, the appearance of the state s'' to child 4 is incompatible with action q' happening next:

$$f_4(s'') \otimes q' = \bot .$$

Thus, after hearing the announcement q', child 4 must engage in belief revision, otherwise he will be lead to have inconsistent beliefs! Indeed, according to the "epistemic system" setting in [8] and the kripke semantics in [4], we would have

$$f_4(s'' \otimes q') = f_4(s'') \otimes f_4(q') = f_4(s'') \otimes q' = \bot ,$$

and hence $s'' \otimes q' \leq \Box_4 \bot$, i.e. at the next state, child 4 will believe the impossible. However, in our present setting, this is not the case: we have

$$f_4(s'' \otimes q') = f_4(s'') * f_4(q') = f_4(s'') * q' \neq \bot$$

since $q \neq \bot$, and thus $s'' \otimes q' \not\leq \Box_4 \bot$. So child 4 has succesfuly revised his beliefs; in fact, our revision axioms imply that he can also form a new hypothesis $m'' \not\leq ker(q')$ about the previous state s'', hypothesis that can explain his new beliefs:

$$f_4(s'') * q' = m'' \otimes q' .$$

Conclusion and Future Work. We present an algebraic setting for dynamic belief revision in multi-agent systems. Our main mathematical object is a Belief Revision System based on Epistemic Systems as algebraic models of Dynamic Epistemic Logic. We axiomatize a notion of multi-agent dynamic revision, that generalizes the update product to inconsistent pairs of a theory and and an experiment, and accommodates the revised theory with the experiment. In our setting, agents can revise with complex epistemic propositions as well as with facts. They can also revise past actions in the view of new experiments. We apply our setting to a cheating version of the muddy children puzzle and show that after the cheating, the honest children would not face any dangerous consequences, in terms of contradiction and confusion.

We are currently working on a complete sequent calculus for this logic, adapting the work in [9] to the present setting. Also, for simplicity we have chosen here to follow the AGM approach in postulating allow one "universal" revision rule: "the rational" revision operator. But there exists of course the possibility of having "personalised revision" operators: by introducing labelled revision operators $*_A$ for each agent, we can allow different agents to use different revision rules, subject only to minimal rationality constraints. We are planning to investigate this possibility in future work.

References

[1] S. Abramsky and S. Vickers, 'Quantales, observational logic and process semantics', *Mathematical Structures in Computer Science* **3**, 161-227, 1993.

[2] C. Alchourron, P. Gardenfors, and D. Makinson, 'On the logic of theory change: partial meet contraction and revision functions', *Journal of Symbolic Logic* **50**, 510-535, 1985.

[3] G. Aucher, 'Belief and Update Logics', Master thesis, ILLC, University of Amsterdam, Amsterdam, NL 2003.

[4] A. Baltag, 'Logics for Insecure Communication', Proceedings of *Theoretical Aspects of Reasoning about Knowledge*, TARK 2001.

[5] A. Baltag and L.S. Moss, 'Logics for epistemic programs', *Synthese* **139**, 2004.

[6] A. Baltag, L.S. Moss, and S. Solecki, 'The logic of public announcements, common knowledge and private suspicions', CWI Technical Report SEN-R9922, 1999.

[7] A. Baltag, 'Logics for Communication: reasoning about information flow in dialogue games', lecture notes of course presented at NASSLLI'2003.`http://www.nasslli.com/2003`.

[8] A. Baltag, B. Coecke, and M. Sadrzadeh, 'Epistemic actions as resources' in Proceedings of *Logics for Resources Programs Processes (LRPP)* workshop in LiCS 2004, `http://www.er.uqam.ca/nobel/philmath/LicsWSPROC.pdf`.

[9] A. Baltag, B. Coecke, and M. Sadrzadeh, 'Algebra and Sequent Calculus for Epistemic Action' in *Electrical Notes in Computer Science (ENTCS)* proceedings of *Logic and Communication in Multi-Agent Systems (LCMAS)* workshop in ESSLLI 2004.

[10] B. Coecke, D.J. Moore, and I. Stubbe, 'Quantaloids describing causation and propagation of physical properties', *Foundations of Physics Letters* **14**, 133-145, 2001.

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

[12] D. Harel, D. Kozen, and J. Tiuryn, *Dynamic Logic*, MIT Press, 2000.

[13] K.I. Rosenthal, *Quantales and their Applications*, Pitman Research Notes in Mathematics Series **234**, Longman, 1990.

[14] K. Segeberg, 'Two traditions in the logic of belief: bringing them together', in *Logic, Language, and Reasoning*, Vol. II, H. Ohlbach and U.Reyle (eds.), 105-134, 1999.

[15] H.P. van Ditmarsch, 'Prologomena to Dynamic Belief Revision', under submission, 2004.