Diversity of Agents and their Interaction

Fenrong Liu*

January 20, 2007

Abstract

Diversity of agents occurs naturally in epistemic logic, and dynamic logics of information update and belief revision. We provide a systematic discussion of different sources of diversity, such as introspection ability, powers of observation, memory capacity, and revision policies, and we show how these can be encoded in dynamic epistemic logics allowing for individual variation among agents. Next, we explore the interaction of diverse agents by looking at some concrete scenarios of communication and learning, and we propose a logical methodology to deal with these as well. We conclude with some further questions on the logic of diversity and interaction.

1 Diversity Inside Logical Systems

Logical systems seem to prescribe one norm for an "idealized agent". Any discrepancies with actual human behavior are then irrelevant, since the logic is meant to be normative, not descriptive. But logical systems would not be of much appeal if they did not have a plausible link with reality. And this is not just a matter of confronting one ideal norm with one kind of practical behavior. The striking fact is that human and virtual agents are not all the same: actual reasoning takes place in societies of diverse agents.

This diversity shows itself particularly clearly in *epistemic logic*. There have been long debates about the appropriateness of various basic axioms, and they have to do with agents' different powers. In particular, the ubiquitous modal Distribution Axiom has the following epistemic flavor:

Example 1.1 Logical omniscience: $K(\varphi \to \psi) \to (K\varphi \to K\psi)$.

Do rational agents always *know the consequences* of what they know? Most philosophers deny this. There have been many attempts at bringing the resulting diversity into the logic as a legitimate feature of agents. Some authors have used "awareness" as a sort of restriction on short-term memory ([FH85]), others have concentrated on the stepwise dynamics of making inferences ([Kon88], [Dun95]). A well-informed up-to-date philosophical summary is found in [Egr04].

The next case for diversity lies in a different power of agents:

Example 1.2 Introspection axioms: $K\varphi \to KK\varphi$, $\neg K\varphi \to K\neg K\varphi$.

Do agents *know when they know* (or *do not know*)? Many philosophers doubt this, too. This time, there is a well-established way of incorporating different powers into

*Institute for Logic, Language and Computation (ILLC), University of Amsterdam, Netherlands. Institute of Philosophy, Chinese Academy of Social Sciences, China.

the logic, using different accessibility relations between possible worlds in Kripke models. Accordingly, we get different modal logics: K, T, $S4$, or $S5$. Each of these modal logics can be thought of as describing one sort of agents. The interesting setting is then one of combinations. E.g., a combined language with two modalities K_1, K_2 describes a two-person society of introspectively different agents! This gives an interestingly different take on current logic combinations ([GS98], [KZ03]): the various ways of forming combined logics, by "fusions" $S5 + S4$ or "products" $S5 \times S4$, correspond to different assumptions about how the agents *interact* in an abstract sense. Effects may be surprising here. E.g., later on, in our discussion of memory-free agents, we see that knowledge of memory-free agents behaves much like "universal modalities". But in certain modal logic combinations, adding a universal modality drives up complexity, showing how the interplay of more clever and more stupid agents may itself be very complex...

Thus, we have seen how *diversity exists inside standard epistemic logic*, and hence likewise in doxastic logic. The purpose of this paper is to bring to light some further sources of diversity in existing logics of information. Eventually, we would want to move from complaints about 'limitations" and "bounds" to a positive understanding of how societies of diverse agents can perform difficult tasks ([GTtARG99]). In addition to identifying diversity of behavior, this also requires a study of *interactions* between different agents: e.g., how one agent learns the types of the agents she is encountering and makes use of such knowledge in communication. This paper is structured as follows. Section 2 briefly identifies some further parameters of variation for agents beyond the well-known, and somewhat over-worked, concerns of standard epistemic logic. These are: powers of observation, powers of memory, and policies for belief revision. Section 3 then looks at dynamic epistemic logics of information update, showing how limited powers of observation for different agents are already accounted for, while we then add some new update systems which also describe varieties of bounded memory. Moving on to correcting beliefs on the basis of new information, Section 4 takes a parallel look at dynamic doxastic logics for belief revision, and shows how different revision policies can be dealt with inside one logical system. Section 5 is a brief summary of sources of diversity, and a transition to our next topic: that of interaction between different agents. In particular, Section 6 discusses several scenarios where different sorts of agent meet, involving identification of types of speaker (liars versus truth-tellers), communication with agents having different introspective powers, and encounters between belief revisers following different policies. We show how these can be dealt with in plausible extensions of dynamic-epistemic and dynamic-doxastic logics. Finally, in Section 7, we summarize, and pose some further more ambitious questions.

This paper is based on existing literature, unpublished work in the author's Master's Thesis ([Liu04]) plus some new research in the meantime. We will mainly cite the relevant technical results without proof, and put them into a hopefully fresh story.

2 Sources of Diversity

The diversity of logical agents seems to stem from different sources. In what follows, we shall mainly speak about "limitations", even though this is a loaded term suggesting "failure". The more cheerful reality is of course that agents have various resources, and they use these positively to perform tasks, often highly successfully.

Our epistemic axioms point at several "parameters" of variation of agents, and indeed, we already identified two of them:

(a) *inferential/computational power*: making all possible proof steps,

(b) *introspection*: being able to view yourself in "meta-mode".

One further potential parameter relevant to epistemic logic is the "awareness" studied by some authors ([FH85]), which suggests some resource like limited attention span, or short-term memory.

Next, consider modern dynamic logics of information, whose motivation sounds closer to actual cognitive practice. These also turn out to incorporate idealizations that suggest further parametrization for diversity. We start with the case of information update.

Consider the basics of *public announcement logic* (PAL): the event $!\varphi$ in this language means "the fact φ is truthfully announced". PAL considers the epistemic effects these announcement actions bring about. In addition to static epistemic axioms that invite diversity, here is a new relevant issue which merges only in such a dynamic setting. The following principle is crucial to the way PAL analyzes epistemic effects of public assertions, say, in the course of a conversation, or a sequence of experiments with public outcomes:

$$[!\varphi]K_a\psi \leftrightarrow \varphi \rightarrow K_a[!\varphi]\psi \quad \textit{Knowledge Prediction Axiom}$$

But the validity of this axiom presupposes several things, notably *Perfect Observation* and *Perfect Recall* by agents. The event of announcement must be clearly identifiable by all, and moreover, the update induced by the announcement only works well on a unique current information state recording all information received so far. More technically, these points show in a detailed soundness proof for the Knowledge Prediction Axiom in its intended semantics. We will discuss this in Section 3, in the more general framework of "product update" for dynamic epistemic languages ([BMS98]). Thus, we have found two more parameters of diversity in logic. Agents can also differ in their powers of:

(c) *observation*: variety of agents' powers for observing current events,

(d) *memory*: agents may have different memory capacities, e.g., storing only the last k events observed, for some fixed k.

Can one deal with these additional forms of diversity inside the logic? As we will see, dynamic epistemic logic with product update can itself be viewed as a calculus of observational powers. And as to memory, [BL04] have shown how to incorporate this into dynamic epistemic logic (DEL) for memory-free agents, and we will extend their style of analysis below to arbitrary finite memory bounds.

Yet another source for diversity of agents lies in *belief revision theory* ([AGM85]). This time, agents must revise their beliefs on the basis of incoming information which may contradict what they believed so far. This scenario is different from the preceding one, as has been pointed out from the start in this area ([GR95]). Even for agents without limitations of the earlier sorts, there is now another legitimate source of diversity, viz. their habits that create diversity:

(e) *revision policies*: varying from conservative to radical revision.

Different agents may react differently towards new information: some behave conservatively and try to keep their original beliefs as much as possible, others may be radical, easily accepting new information without much deliberation. However, these policies are not explicitly part of belief revision theory, except for some later manifestations ([Was00]).

We will show in this paper, following [Liu04], [BL07], how they can be brought explicitly into dynamic logic as well.

This concludes the list of parameters of diversity that we see in current dynamic-epistemic and dynamic-doxastic logics. It is important to mention that acknowledging this diversity inside logical systems is not a concession to the ugliness of reality. It is rather an attempt to get to grips with the most striking aspect of human cognition: despite our differences and limitations, societies of agents like us manage to cooperate in highly successful ways! Logic should not ignore this, but rather model it and help explain it. Our paper is a modest attempt at systematization toward this goal.

3 Dynamic Logics of Information Update

3.1 Preliminaries in dynamic epistemic logic

To model knowledge change due to incoming information, a powerful current mechanism is dynamic epistemic logic, which has been developed intensively by [Pla89], [Ben96], [BMS98], [Ger99], [DHK07], etc. Since our discussions in this paper will be heavily based on DEL, we briefly recall its basic ideas and techniques.

The language is an extension of the one for standard epistemic logic.

Definition 3.1 *Let a finite set of propositional variables Φ, a finite set of agents G, and a finite set of events E be given. The* ***dynamic epistemic language*** *is defined by*

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_a\varphi \mid [\mathcal{E}, e]\varphi$$

where $p \in \Phi$, $a \in G$, and $e \in E$.

As usual, $K_a\varphi$ stands for 'agent a knows that φ'. There are also new well-formed formulas of the type $[\mathcal{E}, e]\varphi$, which intuitively mean 'after event e takes place, φ will hold'. Here the $[\mathcal{E}, e]$ act as dynamic modalities. Thus, the expressiveness of the language is expanded in comparison with that of epistemic logic. One could also add the usual program operations of composition, choice, and iteration from propositional dynamic logic to the event vocabulary to deal with more complex situations like two events happening in sequence, choice of two possible events, and events taking place repeatedly. However in the current context, we will only consider a language without these operations.

Definition 3.2 *An* ***epistemic model*** *is a tuple $\mathcal{M} = (S, \{\sim_a \mid a \in G\}, V)$ [1] such that S is a non-empty set of states, G is a group of agents, each $\sim_a$ is a binary epistemic relation on S, and V is a map assigning to each propositional variable p in Φ a subset $V(p)$ of S.*

Definition 3.3 *Given an epistemic model $\mathcal{M} = (S, \{\sim_a \mid a \in G\}, V)$, we define $\mathcal{M}, s \models \varphi$ (formula φ* ***is true in*** *$\mathcal{M}$* ***at*** *s) by induction on φ:*

1. *$\mathcal{M}, s \models \top$ always*
2. *$\mathcal{M}, s \models p$ iff $s \in V(p)$*
3. *$\mathcal{M}, s \models \neg\varphi$ iff not $\mathcal{M}, s \models \varphi$*
4. *$\mathcal{M}, s \models \varphi \wedge \psi$ iff $\mathcal{M}, s \models \varphi$ and $\mathcal{M}, s \models \psi$*

[1] We will sloppily write $\mathcal{M} = (S, \sim_a, V)$ when G is clear from the context.

5. $\mathcal{M}, s \models K_a\varphi$ *iff for all* $t : s \sim_a t$ *implies* $\mathcal{M}, t \models \varphi$.

We also have explicit models for our special citizens, the 'events'. Abstractly speaking, it has a similar structure as the epistemic model.

Definition 3.4 *An **event model** is a tuple* $\mathcal{E} = (E, \sim_a, PRE)$ *such that* E *is a non-empty set of events,* $\sim_a$ *is a binary epistemic relation on* E*, and* PRE *is a function from* E *to the collection of epistemic propositions.*

The new function PRE in an event model gives the so-called *precondition* for each event: an event e occurs at world s if and only if s fulfills $PRE(e)$.

Definition 3.5 *Consider an epistemic model* $\mathcal{M} = (S, \sim_a, V)$ *and an event model* $\mathcal{E} = (E, \sim_a, PRE)$*. The **product update model** is* $\mathcal{M} \otimes \mathcal{E} = (S \otimes E, \sim'_a, V')$ *with:*

- $S \otimes E = \{(s, e) \in S \times E : (\mathcal{M}, s) \models PRE(e)\}$.
- $(s, e) \sim'_a (t, f)$ *iff both* $s \sim_a t$ *and* $e \sim_a f$.
- $V'(p) = \{(s, e) \in S \otimes E\colon s \in V(p)\}$.

The *actual world* of the new model is the pair consisting of the actual world in $\mathcal{M}$ and the actual event or action in $\mathcal{E}$. The product rule says that uncertainty among new states can only come from existing uncertainty via indistinguishable actions.

We can now add an item for the truth definition of the formulas $[\mathcal{E}, e]\varphi$ to Def. 3.3:

6. $\mathcal{M}, s \models [\mathcal{E}, e]\varphi$ iff, if $\mathcal{M}, s \models PRE(e)$, then $\mathcal{M}\otimes\mathcal{E}, (s, e) \models \varphi$.

Next, so called *reduction axioms* in *DEL* play an important role in encoding the epistemic changes. In particular, the following principle describes knowledge change of agents following some observed event in terms of what they knew beforehand:

$$[\mathcal{E}, e]K_a\varphi \leftrightarrow PRE(e) \to \bigwedge_{f\in\mathcal{E}}\{K_a[\mathcal{E}, f]\varphi : e \sim_a f\}.$$

Intuitively, after an event e takes place the agent a knows φ, is equivalent to saying that if the event e can take place, a knows beforehand that after e (or any other event f which a can not distinguish from e) happens φ will hold. Such a principle is of importance in that it allows us to relate our knowledge after an action takes place to our knowledge beforehand, which plays a crucial role in communication and general interaction.

This concludes our brief review of dynamic epistemic logic. We are ready to move to more complex situations where different agents live and interact. Public announcement logic is the simplest logic which is relevant here, as it describes agents who communicate via public assertions. This is the special case of *DEL* in the sense that the event model contains just one single event. The precondition of $!\varphi$ boils down to the fact that φ is true, as we will see in the formulas in the next section. In this paper, for easy understanding, we use simple variants of *PAL* to motivate our claims, though we also consider a few scenarios using full-fledged *DEL* with a general mechanism of product update.

3.2 Public announcement, observation, and memory

First, we recall the complete axiom system for public announcement.

Theorem 3.6 ([*Pla89*][*Ger99*]) *PAL is axiomatized completely by the usual laws of epistemic logic plus the following reduction axioms:*

(!p). $[!\varphi]p \leftrightarrow \varphi \rightarrow p$ *for atomic facts* p

(!¬). $[!\varphi]\neg\psi \leftrightarrow \varphi \rightarrow \neg[!\varphi]\psi$

(!∧). $[!\varphi](\psi \wedge \chi) \leftrightarrow [!\varphi]\psi \wedge [!\varphi]\chi$

(!K). $[!\varphi]K_a\psi \leftrightarrow \varphi \rightarrow K_a[!\varphi]\psi$.

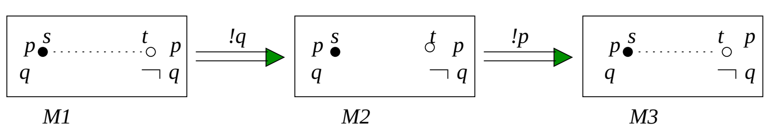
Next, to introduce variety in *observation*, we need to assume a set of possible announcements $!\varphi$, $!\psi$, ... where an agent a need not be able to distinguish all of them. This uncertainty can be modelled by a simple event models with equivalence relation $\sim_a$ between statements which a cannot distinguish. The following principle – a special case of the above general DEL reduction axiom – then describes what agents know on the basis of partial observation:

Fact 3.7 *The following reduction axiom is valid for agents with limited observation power:*

$$[!\varphi]K_a\chi \leftrightarrow (\varphi \rightarrow K_a \bigwedge_{!\psi\sim_a!\varphi}[!\psi]\chi)$$

But there is another sources of diversity not dealt with by PAL or DEL. As we have seen in the previous section, *Perfect Recall* assumes that agents can remember all the events happened so far. But in reality there are agents with bounded memory, who can only remember a fixed number of previous events. It is much harder in PAL to model memory difference because the world elimination update procedure shifts agents to ever more informed states. How can they forget? Here is one option (suggested by [BL04]). First, we can reformulate PAL semantics as in [BL07] to never eliminate worlds. The idea is to let announcement $!\varphi$ cut all links between φ-worlds and $\neg\varphi$-worlds, but keep all worlds in. Now, the resulting "unreachabilities" represent the information agents have so far. One way of describing a memory-restricted agent is then as having forgotten part or all of the "missing links". In the most extreme case, a *memory-free* agent will only acknowledge distinctions made by the last announcement – which may reinstate indistinguishability links that had been cut before. Thus, worlds may also become indistinguishable again: modelling the 'forgetting'. Forgetful agents like this do not satisfy the earlier reduction axiom (!K), as is shown in the following example.

Example 3.8 *Consider the two model changes depicted here:*



There are two possible worlds, s and t in $\mathcal{M}_1$, p and q hold at s, p and $\neg q$ hold at t. After q is announced, we get a new model $\mathcal{M}_2$, in which there is no uncertainty link between s and t. Then we have $(\mathcal{M}_2, s) \models p \rightarrow K_a(p \rightarrow q)$, i.e. $(\mathcal{M}_2, s) \models p \rightarrow K_a[!p]q$. After that, p is announced, and we have $\mathcal{M}_3 \not\models K_aq$, since the agent forgot $!q$ already. We look back at $\mathcal{M}_2$: $(\mathcal{M}_2, s) \not\models [!p]K_aq$. The reduction axiom does not hold!

Fact 3.9 ([*BL04*]) *The correct update reduction axiom valid for memory-free agents is*

$$[!\varphi]K_a\psi \leftrightarrow (\varphi \to U[!\varphi]\psi)$$

In this new principle, the *universal modality* $U\varphi$ says that φ holds in all worlds. With the above picture, it is easy to check that the axiom is correct. To restore the harmony of an update logic, one should also extend the reduction axioms with a clause for the new operator U. The following one is valid:

$$[!\varphi]U\psi \leftrightarrow \varphi \to U[!\varphi]\psi$$

Note that it looks like the above $(!K)$ clause.

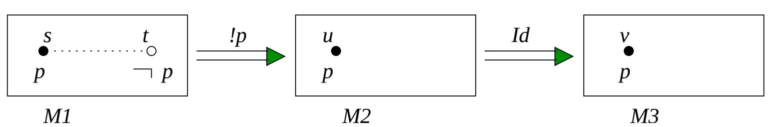
Thus, "logic of public announcement" is actually a family of formal systems, with different update rules depending on the memory type of the agents, and correspondingly, different reduction axioms and reasoning styles.

3.3 Adding memory to product update

The previous section shows that the reduction axiom for knowledge under product update fails for memory-free agents. We will now propose a correct update rule for agents having a bounded memory for the last observed events. By a *k-memory* agent, we mean an agent that remembers only the last k events before the most recent one. A 0-memory or memory-free agent, then, knows only what she learned from the most recent action; a 1-memory agent knows only what she learned from the last two actions, and so on. Modelling this diversity requires some care, witness the following scenario:

Example 3.10 *Memory-free agent a is uncertain about p at first. Then p is announced, and afterwards, an "idle" action Id takes place. Then a should not know p any more since she only remembers the last action.*

Now, according to Def. 3.5, the model changes in the following way:



There are two possible worlds in the original model $\mathcal{M}_1$, the agent a is uncertain about p. After p is announced, we get $\mathcal{M}_2$. Since p does not hold at the world t, the action $!p$ only possibly happens at the world s, so we have only one world u in the model. Intuitively, after the announcement of p, agent a should now know that p, and indeed the agent a *does* know p in $\mathcal{M}_2$. Now the Id action happens, since Id takes place everywhere. We get $\mathcal{M}_3$, abbreviating (u, Id) as v. Intuitively, the action Id being performed, the memory-free agent a should no longer know whether p, because she has 0 memory, she already forgot what had happened one step ago, and she should be uncertain again whether p. But by the above illustration, the agent a *knows* p. This is counter-intuitive!

The difficulty here is that eliminating worlds is a form of hard-wired memory: worlds that have been removed do not come back, so one is 'forced to know'. To get this right in a more sensitive manner, we now present a proposal for product update with general memory free agents. The first source for this is as follows:

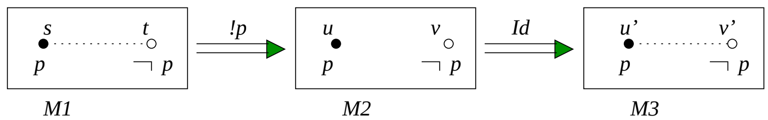
Definition 3.11 ([*Sny04*]) *Let an epistemic model $\mathcal{M} = (S, \sim_a, V)$ and an event model $\mathcal{E} = (E, \sim_a, PRE)$ be given. The* ***product update for memory-free agents*** *is $\mathcal{M} \otimes \mathcal{E}$ $= (S \otimes E, \sim'_a, V')$ with:*

(i) $S \otimes E = \{(s,e) : (s,e) \in S \times E\}$.

(ii) $(s,\ e) {\sim'_a} (t,f)$ *iff* $(\mathcal{M},\ s \models PRE(e)$ *iff* $\mathcal{M}, t \models PRE(f))$ *and* $e \sim_a f$.

(iii) $V'(p) = \{(s,e) \in S \otimes E\colon\ s \in V(p)\}$.

Compared with standard product update, item (i) in the above definition leaves out the precondition restriction. This keeps all worlds around. Item (ii) then defines the uncertainty relation on all worlds ('active', or not) in the new models. (iii) remains the same, and we will ignore this valuation clause henceforth. To understand this new definition, we look at the example again, now updating models according to the new definition:



This is like Example 3.8 – but now, the original state model remains the same. According to Def. 3.11, we obtain a different model $\mathcal{M}_2$, abbreviating $(s, !p)$ as u and $(t, !p)$ as v. There is no uncertainty link between them. So the agent a knows that p in $\mathcal{M}_2$. Now the 'idle' identity event Id happens, and we get a new state model $\mathcal{M}_3$, abbreviating (u, Id) as u' and (v, Id) as v'. The agent a is uncertain whether p. This is what we expect for a 0-memory agent.

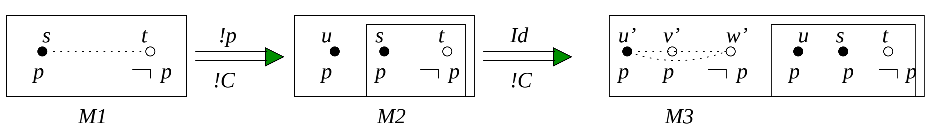
[Sny04] also extended this proposal to the k-memory case. Here, however we propose an alternative for modelling forgetting. We introduce an auxiliary *copy action* $!C$ which always takes an old possible world into the new model. Essentially it puts those worlds which were previously deleted into a stack, and makes sure agents can always retrieve them when needed.

Definition 3.12 ([*Liu04*]) *Let an epistemic model* $\mathcal{M} = (S, \sim_a, V)$ *and an event model* $\mathcal{E} = (E, \sim_a, PRE)$ *be given. The* ***product update for memory-free agents*** *is* $\mathcal{M} \otimes \mathcal{E} = (S \otimes E, \sim'_a, V')$ *with:*

(i) $S \otimes E = \{(s,e) \in S \times E\colon\ \mathcal{M}, s \models PRE(e)\}$.

(ii) For $e, f \neq !C$, $(s,e) \sim'_a (t,f)$ *iff* $e \sim_a f$.

To see how this new proposal works, we go back to the above example, but now update with an additional copy action:



From the original model, by Def. 3.12, we get model $\mathcal{M}_2$, with the abbreviations: $u \hookleftarrow (s, !p)$; $s \hookleftarrow (s, !C)$; $t \hookleftarrow (t, !C)$. The agent a then knows that p. The rectangular box inside contains the states coming from the copy action. After the Id action, similarly, we obtain the new model $\mathcal{M}_3$ with the abbreviations: $u' \hookleftarrow (u, Id)$; $v' \hookleftarrow (v, Id)$; $w' \hookleftarrow (w, Id)$; $u \hookleftarrow (u, !C)$; $s \hookleftarrow (s, !C)$; $t \hookleftarrow (t, !C)$. Again, the agent a is uncertain whether p. This idea is similar to the usual design of operation systems ([SGG03]), where the working memory does the jobs while carrying a stack of old information to be be visited when necessary. [Liu04] has a more restrictive variant of the above definition copying worlds only when necessary. This makes the above models less over-loaded.

Extending this approach, we get the following extended update rule:

Definition 3.13 ([*Liu04*]) *Let $\mathcal{M}$ be an epistemic model, $\mathcal{E}_{-k}$ be the k-th event model before the most recent one $\mathcal{E}$. The* ***product update for** k**-memory agents*** *is* $\mathcal{M} \otimes \mathcal{E}_{-k} \otimes \cdots \otimes \mathcal{E}_{-1} \otimes \mathcal{E} = (S \otimes E_{-k} \otimes \cdots \times E_{-1} \otimes E, \sim'_a, V')$ *with:*

(i) $S \otimes E_{-k} \otimes \cdots \otimes E_{-1} \otimes E = \{(s, e_{-k}, \ldots, e_{-1}, e) \in S \times E_{-k} \times \cdots \times E_{-1}\colon \mathcal{M} \otimes \mathcal{E}_{-k} \otimes \cdots \otimes \mathcal{E}_{-1}, (s, e_{-k}, \ldots, e_{-1}) \models PRE(e)\}$.

(ii) For $e_{-k}, \ldots, e_{-1}, e, f_{-k}, \ldots, f_{-1}, f \neq C!$, $(s, e_{-k}, \ldots, e_{-1}, e) \sim'_a (t, f_{-k}, \ldots, f_{-1}, f)$ *iff* $e_{-k} \sim_a f_{-k}, \ldots, e_{-1} \sim_a f_{-1}$ *and* $e \sim_a f$.

Given this update rule, it is straightforward to find a complete dynamic logic in the earlier DEL format, but now for k-memory agents. As we do not need this calculus in what follows, we omit a detailed formulation.

Of course, this is only the beginning of an array of further questions. In particular, we would like to have a more structured account of memory, as in computer science where we update data or knowledge bases. Update mechanisms are more refined there, referring to memory structure with actions such as information replacement ([Liu04]), where the agent would have a priority order in her database, so that she would know which old information should go to make room for the new. This is one instance of a more "constructive" syntactic approach to update, complementary to our abstract one in terms of model manipulation. Whether our current semantic method or a syntactic one works better for finding agents' parameters of diversity is a question worth investigating.

This section has identified two new parameters for dynamic updating agents: powers of *observation* and powers of *memory*. DEL as it stands already provides a way of modelling the former, while we have shown how it can also be modified to accommodate agents with bounded memory. We consider these two additional phenomena at least as important from an epistemological viewpoint as the usual themes of inferential power and introspection, generated by the earlier static phase of logical theorizing.

In the following section, we move to an extension of DEL treating one further crucial aspect of agents' cognitive behavior, when 'things get rough'.

4 Diversity in Dynamic Logics of Belief Change

Information flow and action based upon it is not always a matter of just smooth update. Another striking phenomenon is the way agents correct themselves when encountering evidence which contradicts their beliefs so far. *Belief revision theory* describes what happens when an agent is confronted with new information which conflicts her earlier beliefs. It has long been acknowledged that there is not one single logical rule for doing this. Indeed, different policies toward revising beliefs, from more 'radical' to more 'conservative' all fall within the compass of the famous AGM postulates.

In this paper, however, we take another approach inspired by dynamic epistemic logic. First, on the static side, we follow the common idea that beliefs are modelled by so-called *plausibility relations* between worlds, making some epistemically accessible worlds more plausible than others. Agents believe what is true in the most plausible worlds – and the same thinking may also be used to define their conditional beliefs. In this setting, one can then view belief revision on the analogy of the preceding update paradigm, viz. as a mechanism of *change in plausibility relations*. To see this, here is a concrete example of how this can be implemented technically.

4.1 Belief revision as changing plausibility relations

Example 4.1 ([*Ben07*]) ($\Uparrow$) *Radical revision*
$\Uparrow P$ is an instruction for replacing the current ordering relation $\leq$ between worlds by the following: all P-worlds become better than all $\neg P$-worlds, and within those two zones, the old ordering remains.

Note that the $\neg P$-worlds are not eliminated here: they move downward in plausibility. This reflects the fact that we may change our mind once more on the basis of further information. $\Uparrow P$ is one famous policy for belief revision, corresponding tom an 'eager response', or a 'radical revolution', or 'high trust' in the source of the information. But there are many other policies in the literature. Another famous one would just place the best P-worlds on top, leaving the further order unchanged. A more general description of such different policies can be given as definable ways of changing a current plausibility relation ([BL07], [Rot06]). Once we have such a definition for a policy of plausibility change, the corresponding dynamic logic for belief revision can be axiomatized completely in DEL style. Here is the result for the policy of radical revision:

Theorem 4.2 ([*Ben07*]) *The dynamic logic for radical revision ($\Uparrow$) is axiomatized completely by an axiom system KD45 on the static models, plus the following reduction axioms*

($\Uparrow p$). $[\Uparrow \varphi]p \leftrightarrow p$

($\Uparrow \neg$). $[\Uparrow \varphi]\neg\psi \leftrightarrow \neg[\Uparrow \varphi]\psi$

($\Uparrow \wedge$). $[\Uparrow \varphi](\psi \wedge \chi) \leftrightarrow [\Uparrow \varphi]\psi \wedge [\Uparrow \varphi]\chi$

($\Uparrow B$). $[\Uparrow\varphi]B\psi \leftrightarrow (E\varphi \wedge B([\Uparrow\varphi]\psi|\varphi)) \vee (\neg E\varphi \wedge B[\Uparrow\varphi]\psi)$ [2]

The last axiom here shows the doxastic effects of the chosen policy. In the same style, one can also axiomatize other belief revision policies. When put together, the result is a dynamic logic of belief revision which can also describe interactions between different agents, using operator combinations such as, say, ['*Radical P*']['*Conservative Q*']φ.

All this is still qualitative. But the earlier product update mechanisms also admits of a more refined quantitative version, describing agents' attitudes in a more detailed numerical manner, and allowing for further polices of changing these fine-grained beliefs. In the next subsection, we will briefly show how.

4.2 Belief revision as changing plausibility values

Following [Spo88], a κ-ranking function was introduced in [Auc03] to extend DEL with numerical beliefs. A κ-ranking function maps a given set S of possible worlds into the class of numbers up to some maximum Max. The numbers can be thought of as denoting degree of surprise. 0 denotes 'unsurprising', 1 denotes 'somewhat surprising', etc. κ represents a plausibility grading of the possible worlds, in other words, degree of beliefs.

Definition 4.3 *A **doxastic epistemic model** is a tuple $\mathcal{M} = (S, \sim_a, \kappa_a, V)$, where S, $\sim_a$ and V are defined as usual, and the plausibility function κ_a ranging from 0 to some upper limit Max is defined on all worlds.*

[2] Here E is the existial modality, dual to the earlier universal modality U. The symbol $|$ denotes a conditional belief, and it means: 'given that'. Van Benthem's full system also has complete reduction axioms for conditional beliefs, thereby solving the notorious 'Iteration Problem' of AGM theory.

Definition 4.4 *A **doxastic epistemic event model** is a tuple* $\mathcal{E} = (E, \sim_a, \kappa^*_a, PRE)$, *with* E, $\sim_a$ *and* PRE *defined as usual,* κ^*_a *ranges from 0 to* Max, *defined on all events.*

The κ^*_a-value describes the agent's detailed view on which event is taking place. With plausibilities assigned to states and events, 'graded beliefs' will change via a suitable rule for product update. Here is the key proposal in [Auc03], the first of its kind in the DEL-style literature:

$$\kappa'_a(s,e) = Cut_{Max}(\kappa_a(s) + \kappa^*_a(e) - \kappa^s_a(\varphi)),$$

where $\varphi = PRE(e)$, $\kappa^s_a(\varphi) = min\{\kappa_a(t) : t \in V(\varphi) \text{ and } t \sim_a s\}$, and

$$Cut_{Max}(x) = \begin{cases} x & \text{if } 0 \leq x \leq Max \\ Max & \text{if } x > Max. \end{cases}$$

A more perspicuous version of this approach uses an epistemic-doxastic language with propositional constants to describe the plausibility change ([Liu04]):

Definition 4.5 *The **epistemic-doxastic language** is defined as*

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_a\varphi \mid q^\delta_a$$

where $p \in \Phi$, *a set of propositions,* $a \in G$, *a set of agents, and* δ *is a* κ*-value in* $\mathbb{N}$, q^δ_a *are a special type of propositional constants.*

The interpretation is as usual, but now with the following simple truth condition for the additional propositional constants:

$$(\mathcal{M}, s) \models q^\delta_a \text{ iff } \kappa_a(s) \leq \delta$$

The update mechanism can now be defined by merely specifying the new κ-value in the product model $\mathcal{M} \times \mathcal{E}$. To keep our discussion simple, we use just this stipulation:

Definition 4.6 (Bare addition rule) *The new plausibilities for pair-worlds* (s, e) *in product models are defined by the following rule:*

$$\kappa'_a(s,e) = \kappa_a(s) + \kappa^*_a(e)$$

Theorem 4.7 ([*Liu04*]) *The complete dynamic logic of plausibility belief revision consists of the key reduction axioms in Theorem 3.6 plus the following new one:*

$$[!\varphi]q^\delta_a \leftrightarrow q_a^{\delta - \kappa_a(!\varphi)}$$

More generally, different update functions will account for different numerical revision policies. If such an update rule is simply expressible, we can still get a complete dynamic logic for it, though the simple substraction rule may not work anymore. To illustrate this diversity of behavior for agents, we now present an update rule which incorporates further 'degrees of freedom':

Definition 4.8 ([*Liu04*]) *Let agent* a *assign weight* λ *to world* s, *and weight* μ *to the event* e. *The **plausibility of the new world*** (s, e) *is calculated by the parametrized rule*

$$\kappa'_a(s,e) = \frac{1}{\lambda + \mu}(\lambda\kappa_a(s) + \mu\kappa^*_a(e)) \quad (\natural).$$

Variations of parameter λ and μ then describe a range of various agents. For instance, when μ=0, we get *highly conservative agents*, and the ($\natural$) rule turns into $\kappa'_a(s,e) = \kappa_a(s)$. This means that the agent does not consider the effect of the action. Similarly, when λ=0, the agents are *highly radical*, and $\kappa'_a(s,e) = \kappa^*_a(e)$. When $\lambda = \mu$, we get *Middle of the Road agents* who let plausibility of states and actions play an equally important role in determining the plausibility of the new state. We obtain *conservative agents* when $\lambda > \mu$ and *radical agents* when $\mu > \lambda$. In this manner, we have distinguished five types of agents in dynamic logic. For an even more general view of agents' behavior towards incoming information, see [Liu06]. Summing up, we may regard our numerical update rule as a refinement of the qualitative dynamic logics for belief change in section 4.1 (cf. [Ben07] and [BL07]). We defer a more detailed comparison to another occasion.

4.3 Some further observations

Our treatment of belief revision provides a simple format of plausibility change, where different policies show naturally in the update rules for either plausibility relations or value constants, and their matching reduction axioms in the dynamic doxastic logic. Moreover, our treatment also goes beyond the standard *AGM* paradigm, in that more complex event models allow agents to doubt the current information in various ways. Here are a few further issues that come up in this setting, some conceptual, some technical.

First, doubting the current information might also make sense for *PAL* and *DEL* scenarios without belief revision. It is easy to achieve this by adding further events to an event model, providing, say, a public announcement $!\varphi$ with a counterpart $!\neg\varphi$ with some plausibility value reflecting the strength of the "dissenting voice". Likewise, policies with weights for various factors in update make much sense in recently proposed dynamic logics of probabilistic update ([Auc05], [BGK06]).

Incidentally, this *DEL* approach via modified event models for different policies may also suggest that we can *relocate* policies from "modified update rules" to "modified event models" with a standard update rule. This has to do with an important more general issue: are we describing single events of update or revision 'locally' without further assumptions about the long-term behavior of the agents involved, or are we witnessing different more 'global' types of agent at work? In the former case, the diversity is in the response, rather than the type. We must leave this issue, and a comparison between the pros and cons of the two stances to another occasion.

Finally, connecting Sections 3 and 4, revision policies and memory restrictions may not be that disjoint after all. Technically speaking, the update behavior of highly radical agents is similar to that of memory-free agents, as they simply take the new information without considering what happened before (of course, for different reasons). In other words, the event that takes place completely characterizes the "next" epistemic state of the agent. This seems to be related also to notions such as "only knowing" or "minimal knowledge" in [Lev90] and [HJT90]. This final observation also provides a further challenge: viz. unifying some of our parameters of diversity discussed so far.

5 From Diversity to Interaction

We have investigated many different sources of diversity, some visible in static logics, some in dynamic ones. Besides the old parameters from epistemic logic, namely computation and introspection ability, we have added several new aspects, i.e. observation power, memory capacity and revision policy. Our discussion has been mostly in the framework of dynamic epistemic logic and we have shown how it is possible to allow for a characterization of diversity within the logic. To summarize, look at the following diagram consisting of the main components of dynamic epistemic logic:

$$\begin{cases} \textit{Static} \quad \textit{language} & \text{Epistemic model } \mathcal{M}; \\ \textit{Dynamic} \quad \textit{language} & \text{Event model } \mathcal{E}; \\ \textit{Product} \quad \textit{update} & \text{Model change } \mathcal{M} \times \mathcal{E}. \end{cases}$$

In the previous sections we have shown that the diversity of agents can be explicitly modelled in terms of these logical components. The following table is an outline: [3]

Component	Residence	Diversity
$\mathcal{M}$	relations between worlds	introspection
$\mathcal{E}$	relations between actions	observation
$\mathcal{M} \times \mathcal{E}$	update mechanism	memory, revision policy

As we can see from the table, by introducing parameters of variation in each component, we are able to describe diversity of agents inside the logic.

But recognizing and celebrating diversity is only a first step! The next important phenomenon is that diverse agents *interact*, often highly successfully. Describing this interaction raises a whole new set of issues. In particular, our logical systems can describe the behavior of various agents, but they cannot yet state in one single formula “that an agent is of a certain type” or describe what would happen when we encounter those different agents. And as they stand, they are even less equipped to describe the interplay of different agents in a compact illuminating way. Imagine, if you know the type of the agent you are encountering, can you take advantage of that knowledge? Or how could you *learn* about the type of the agent? In the following section, we will explore a few of these issues, and show in how far our current logical framework can handle these phenomena – and what features need to be added.

6 Interaction between Different Agents

Interaction between different agents is a vast area of diverse phenomena, and so, we will only discuss a few scenarios. These will show how the earlier dynamic logics can deal with some crucial aspects - though they also quickly need significant extensions. Our examples cover: reliability of sources (truth-tellers versus liars), meetings between more or less introspective agents, and interaction between belief revisers following different policies.

6.1 If I know you are a Liar ...

In this section we are challenging one of the PAL assumptions, namely, that all the announcements are truthful. What would happen if the announcer is a liar? More generally, can we figure out whether the announcer is a liar or truth-teller? In the following we will focus on such issues and explore how we update our knowledge when encountering people who should be identified first. These questions also bring us to a well-known puzzles about liars and truth-tellers. Here we consider one of its variations, high-lighting the fact that knowing what kind of type of agent you encounter makes life a lot easier:

[3] Note that we have not discussed the parameter of inferential/computational power yet in this context. A more syntactic-oriented approach on this topic can be find in [AJL06] and [Jag06]. It seems quite possible to merge the models proposed there with ours.

Example 6.1 *On a fictional island, inhabitants either always tell the truth, or always lie. A visitor to the island meets two inhabitants, Aurora and Boniface. What the visitor can do is ask questions to discover what he needs to know. His aim is to find out the inhabitants' type from their statements. The visitor asks A what type she is, but does not hear A's answer. B then says "A said that she is a liar". Can you tell who is a liar and who is a truth-teller?*

One can try to figure out the answer to the puzzle by intuitive reasoning, but we will give a precise analysis in logical terms. To describe the situation with the relevant events, the salient fact is the agent-oriented nature of the communication. To bring this out, we first need to extend the language:

Definition 6.2 *Take a finite set of proposition variables Φ, and a finite set of agents G. Predicates $L(x)$, $T(x)$ and action terms $!\varphi_a$ are now added. The* ***dynamic epistemic agent type language*** *is defined by the rule*

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_a\varphi \mid [\pi]\varphi \mid L(x) \mid T(x)$$
$$\pi := !\varphi_a$$

where $p \in \Phi$, and $a \in G$.

Now $L(a)$ expresses 'agent a is a Liar', and $T(a)$ expresses 'agent a is a Truth-teller'. In fact, for the above example, we only need one of these expressions, since the agent is either a liar or a truth-teller. So we can use $\neg L(a)$ to denote 'agent a is a Truth-teller'. Besides, we also want to express *who* executes some action, for instance, $!\varphi_a$ reads intuitively as 'an announcement of φ performed by agent a'. Next we enrich the structure of our models, to a first approximation, in the following structures with hard-wired known agent types.

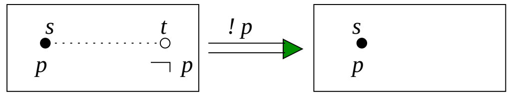
Definition 6.3 *We define the new epistemic model as $\mathcal{M} = (S, \{\sim_a \mid a \in G\}, V, V_G)$, where the new element V_G is a type function: V_G assigns x a subset of G, either G_L, the set of Liars, or G_T, the set of Truth-tellers. Moreover, given some suitable event model $\mathcal{E}$, the* ***truth conditions*** *for the new well-formed formulas are the following:*

1. $\mathcal{M}, s \models T(x)$ *iff* $x \in G_T$.
2. $\mathcal{M}, s \models L(x)$ *iff* $x \in G_L$.
3. $\mathcal{M}, s \models [!\varphi_a]\psi$ *iff ψ holds at the world $(s, !\varphi_a)$ in the product model $\mathcal{M} \otimes \mathcal{E}$.*

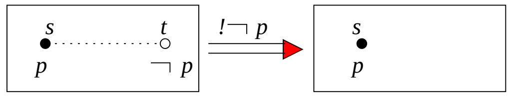
Here 1 and 2 are clear. Item 3, however, is incomplete as it stands! This is because we have not given a precise update rule for the new agent-oriented announcements, which would require *precondition* $\langle !\varphi_a\rangle\top$ for the event of agent a's saying that φ. This is a typical feature with diversity of agents: we have to give up the idealization in standard DEL that preconditions of events are common knowledge. In order to state useful preconditions, we will need more information about agent types.

Consider the example again, clearly, the reason why the visitor should first find out who belongs to what type of agents is that it immediately determines the ways he takes the incoming information. Here is a general illustration:

Case One: Agent b is does not know whether p is true, but *she knows that a is a truth-teller*. In fact, p is the case, and a says 'p is the case', after which b updates her knowledge accordingly:



Case Two: Similarly, agent b first does not know if p is true and *she knows a is a liar*. Now a says that 'p is not the case'. Agent b updates her knowledge with p instead of $\neg p$:



These examples presuppose a definition of agent types. These can be expressed more precisely as follows:

(1) **truth-teller** $T(a) \to (\langle !\varphi_a \rangle \top \leftrightarrow \varphi)$

(2) **liar** $L(a) \to (\langle !\varphi_a \rangle \top \leftrightarrow \neg\varphi)$

Clause (1) says that a truth teller a can say exactly those things φ that are true. For the liar, this reverses. Even this simple stipulation has some interesting effects. E.g., no one can say that she is a liar:

Fact 6.4 $\langle !L(a)_a \rangle \top$ *does not hold in any case.*

Proof. Suppose $\langle !L(a)_a \rangle \top$. There are two cases. Either a is a liar, $L(a)$, or a is a truth-teller, $T(a)$. In the first case, according to (1), we have $(\langle !L(a)_a \rangle \top \to L(a))$. By the assumption, we get $L(a)$. If a is a truth-teller, according (2), we get $(\langle !L(a)_a \rangle \top \to \neg L(a))$. Again, by the assumption, we have $\neg L(a)$. Contradict. So $\langle !L(a)_a \rangle \top$ does not hold. ■

This observation has some flavor of the Liar Paradox, but we do not pursue this angle here. Incidentally, another take on our scenario might take it to be about just single "lies and truths", rather than long-term liars and truth-tellers. This will not change our analysis here, but it would shift the emphasis in modelling from diversity of agents to what might be called *diversity of signals*. The latter tack is attractive, too, and sometimes simpler - but our main emphasis here is highlighting agent diversity in its own right.

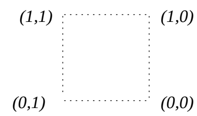
Now as for interaction, we first need to describe what agents would learn from communication if they know the type of the other agent. To compute this, we can combine the information about agent types with the general rules of dynamic epistemic logic, in particular, the earlier reduction axioms for knowledge after events have taken place. If we do this (we suppress details here), we arrive at the following principles:

(3) $K_b T(a) \to K_b(\langle !\varphi_a \rangle \top \leftrightarrow \varphi)$

(4) $K_b L(a) \to K_b(\langle !\varphi_a \rangle \top \leftrightarrow \neg\varphi)$

But this is not yet what we need for our Island Puzzle. There, and also in real life, the types of agents encountered may be *unknown*! We need to represent that as well in our static and dynamic models. There are several ways of doing this. At the very least, the above predicates L, T will no longer be fixed once and for all for agents, but they need to be made part of the specification of worlds, or events. One proposal to this effect (cf. [BGK06]) would use *pair events* of the form '(agent type, physical event)', say, "P is said by a truth-teller", or "P is said by a liar". Such events are epistemically indistinguishable if we can neither tell the agent types apart nor the actual observed events.

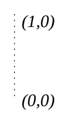
However, in the analysis of our puzzle, we do not need this rich format yet, since the conversation itself is about the types of agents, which makes things much easier. We stick with a more ad-hoc format, just to show how one can deal with the puzzle. To model the original epistemic state of the visitor, see the picture below. Since there is no information to indicate who is of what type, there are 4 possibilities in total, where for example the vertex (1, 1) is the case with A, B both truth-tellers.



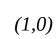
The dotted line denotes the visitor's uncertainty. Since the visitor does not hear what A says, there is no update for that. (In a more refined multi-agent scenario, there *would* be a product update for this event, as some higher-order knowledge about others changes – but we ignore this aspect here.) Then B says "A said that she is a liar". Since we already saw that is a general truth that no one can say she is a liar, what B said about A is not true. So we conclude that B is a liar. This reasoning depends on the following principle, which follows from our agent type definition:

(5) $\varphi \wedge \langle !\neg\varphi_a \rangle \top \rightarrow L(a)$.

Meanwhile, we also know that A must have said that she is a truth-teller, since she was asked what type of agent she is, and there are two possible answers. In this way, we split what B said into two statements: 'B is a liar' and 'A is a truth-teller'. To illustrate this more clearly, the update may be carried out in sequence, first with 'B is a liar':



and then with 'A is a truth-teller':



Strictly speaking, this is not quite right, since there is only one event of B's speaking, but we leave the formulation of one single update to the reader.

We have obtained the required answer: Aurora is a truth-teller, while Boniface is a liar. Our analysis is in the same spirit as when one tries to figure out what kind of color a card has according to a sequential announcements (cf. [Ben06]). What is new here is that we no longer take any incoming information automatically as truthful. Instead, we first identify the type of agent who makes the statement, then we update our knowledge. Of course, this is only the beginning, since more complex scenarios would involve our updating our ideas about the degree of reliability of the source of our information.

Here are some valid principles about agents' changing knowledge in case a proposition is announced by a source whose type they do not know yet:

(6) $K_b T(a) \rightarrow ([!\varphi_a] K_b \varphi \leftrightarrow K_b [!\varphi_a] \varphi)$

(7) $K_b L(a) \rightarrow ([!\varphi_a] K_b \neg\varphi \leftrightarrow K_b [!\varphi_a] \neg\varphi)$.

These principles again follow by combining our definitions of agent types and general dynamic-epistemic laws. These can also be, and need to be, extended to more complex event models of the earlier sort where we encode uncertainty about agent types into the events. But the earlier general dynamic-epistemic reduction axioms will still work in this setting, when combined with the preconditions supplied by different agent types.

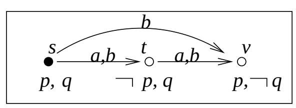
Summing up, we have see how an adequate account of different sources requires structured communicative events with agents explicitly indicated, explicit representations of agents' types, and a combination of general dynamic-epistemic reasoning principles with specific postulates about types of agent. In such a system, we can derive interesting principles about interaction between different agents. Of course, there are many more types of agent than just Liars and Truth-tellers, and Islands like the above are still logical paradise as compared to the real world. In particular, our views of the reliability of agents may change over time in subtle manners, calling for probabilistic information ([BGK06]). We will leave such further complications to future investigation.

6.2 A meeting between introspective and non-introspective agents

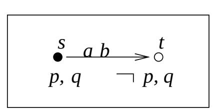
In this subsection we move to the perspective of the *addressee* instead of the *addressor* as investigated in the preceding scenario. Consider the following story.

Example 6.5 *Two agents are sitting silently on a bench in the park. One of them, a, is non-introspective, but the other: b, is. The complete epistemic situation they find themselves in is depicted below – where the actual world is called s. The agents do not communicate with each other at first. Now, the aim for both of them is to find out which world is the real one as soon as possible. They have only one chance to receive new atomic information from some passer-by, and then they are ready to communicate with each other. What information should they get? What kind of communication should they engage in?*

We picture the initial situation in the following diagram. As usual, all worlds are reflexive for each agent, but loops are omitted:

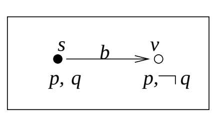


Here s is the real world where p and q are true. So, there are two possible announcements one can make, either $!p$ or $!q$. Let's compare what will happen in these two cases. First, when q is truly announced by someone, the new model becomes:



This new situation is symmetric between both agents. Both a and b are uncertain between s and t, and they do not know that s is the real world. And, given the symmetry, even if they communicate, it does not help, since they both know the same.

By contrast, once the fact p is announced, the new model becomes



Here, the effect of this announcement is different for agent a and b! Agent b learns that p, but not that q. But agent a learns that both p and q, since she has no link to the world v. And she knows this is the real situation. Now a can inform b of q, so that b would also know q and realize that s is the real situation. What is going on here? How can the less-introspective agent b learn more? Do intellectuals need help from the man in the street to get their bearings?

One analysis of what is going on here involves straight update. In terms of our epistemic models, a non-introspective agent may have fewer accessibility arrows than a corresponding introspective one,which means she is better informed, even though she does not reflect on this, and may not know everything she knows. Thus, additional information may help her more than her introspective companion.

To model the reasoning in such situations, as we have done in the previous subsection, we can introduce agent types $I(a)$ and $NI(a)$ in the language to express that 'a is an introspective agent', and 'a is an anti-introspective agent' respectively, providing type definitions like the following:

(1) $I(a) \to (K_a\varphi \to K_aK_a\varphi)$

(2) $NI(a) \to (K_a\varphi \to \neg K_aK_a\varphi)$

Clearly, because of their different introspective abilities, agents a and b may obtain different knowledge from what they learn. Intuitively, as we said already, the non-introspective agent even has an advantage in the above model, in that the following implication holds:

$$\mathcal{M}, s \models K_b\varphi \to K_a\varphi \quad (*)$$

But it is easy to think of settings where the knowledge of the agents would be incomparable. One can also analyze this type of situation more generically, using reduction axioms for informational events like before, leading to principles describing the interaction of the two agents such as the following:

(3) $NI(a) \wedge K_aI(b) \to ([!\varphi_c]K_b\psi \to K_a[!\varphi_c]\psi)$

This is the static situation, looking at the agents separately. Of course, our scenario also illustrates another phenomenon, viz. how helpful agents which differ in their capacities may still inform each other, making the group consisting of both agents together better informed than its members separately.

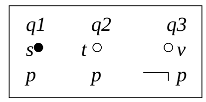
6.3 Talking with different belief revisors

In our final scenario, we consider both information update and belief revision, and we also allow for diversity of both senders and receivers of information. Can our update models and their logics handle this? The following story is a bit contrived, but it highlights some realistic issues in everyday settings.

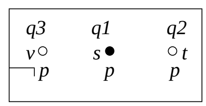
Example 6.6 *Four agents live together, and their types are common knowledge. Agent a is a radical belief revisor, and b a conservative one. Agent c is a very trustworthy person, according to a and b, but d is less so. In the initial situation, there are three possible worlds s (the actual world), t, and v, as pictured in the diagram below, which also shows the valuation for the proposition letters. As for epistemic or doxastic relations, initially, a*

and b consider all three worlds possible, and they have the same plausibility ordering over them: v is most plausible, s is least plausible, t is in between. Moreover, c happens to know that p is the case, and d happens to know that q_2 is not the case. One can only speak after the other. Does this matter? Will both orders inform a, b equally well?

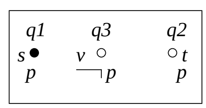
The original model is depicted as



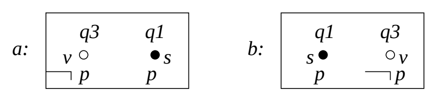
Let us first suppose that d speaks first, truly, and says that p. Because of their different attitudes towards this new information, even though she acknowledges that d might be wrong, the radical (or more trusting) agent a will change her plausibility ordering over the three worlds into:



In contrast to this, the conservative (or more suspicious) agent b would update his plausibility ordering in the following manner:



We draw these orders separately here, though they will be part of the single total epistemic-doxastic for the group, of course. Now the trusted source c tells agents a, b that q_2 is false. The above two models then change into the following ones:



Reading from the picture, we see $B_a q_1$ and $B_b q_3$. Thus, a has the right belief, but unfortunately b does not! Different revisors can get different things out of witnessing the same events, and indeed, some of them may be misled by correct information into believing false things! There is endless potential here for deceiving other agents– and even 'deception by the truth', which has already been observed by game theorists in the study of signaling games.

But dynamics of information flow is in principle order-dependent. What about the opposite order, where agent c speaks first, and only then d? Look at the original model again. After c's truthful announcement, both a and b would update their model into the model with just the two possible worlds s and v of the last picture above. Then, when agent d tells them that p, the difference between the revision policies of a and b is immaterial: they can only raise the plausibility of p in one way, putting world s on top. Thus, both acquire the right belief: as $B_a q_1$ and $B_b q_1$ hold.

If we were to analyze this scenario in detail, we could use the same methodology as in Section 5.1 to express the types of agents qua revision policies (using the dynamic logics for belief revision discussed in Section 4), and then describe the interactions using a mixture of these type definitions and the general principles of dynamic epistemic-doxastic logic.

Admittedly, as we said, the preceding scenario is a bit contrived. We also have some more appealing scenarios of this sort for Muddy Children, where children revise beliefs rather than updating their knowledge, and where both skeptical and trusting children are around. In this way, belief revision policies become concrete objects, whose workings can be determined precisely, and whose peculiarities may be exploited in more sophisticated puzzles of communication. We leave a presentation for another occasion.

With these three scenarios, our discussion of interaction between diverse agents has come to an end. The main thrust of our investigation has been this. Once we have diverse agents inside one logical system, we can talk about the way they update their information and revise their beliefs. To deal with specific scenarios, we found that we needed the following additional ingredients: (a) more structured views of relevant events, (b) language extensions with types of agent, and their properties, where one may have to distinguish between the sender and the receiver of the information, and (c) mixtures of general dynamic-epistemic reasoning with specific information about agents. All this worked for information update, but we have also indicated how it applies to belief revision, when phrased in a dynamic logic format.

7 Conclusions and Future Work

This paper has presented a more systematic discussion of different sources of diversity for rational agents than is usually found in the literature. More concretely, we showed how such diversity can be encoded in dynamic logics allowing for individual variation among agents. In particular, in the context of knowledge update, we made new proposals for modelling memory capacity, and defined a new version of product update for bounded k-memory agents. Next, in the context of belief revision, we showed how different revision policies can be put into one dynamic logic, allowing for great variation in revision and learning behavior. But allowing for and describing diversity is just one stage of ambition for logical modelling. Agents should not just 'live apart together'. Thus, we moved to the topic of interaction between agents of different types, discussing several scenarios which may arise then, having to do with different information processing, communication, and achieving of goals, when agents differ in their reliability, introspective powers, or belief revision policies. Our general conclusion was that these phenomena, too, can be modelled in our dynamic logics – but they need to be extended with explicit accounts of agents' types, and more structured informative events.

But all this is only a beginning. There are several questions we would like to explore in the future. First, back to charting the sources of diversity, there remains the issue whether one can have a *general* view of the natural "parameters" that determine differences in behavior of logical agents. Our analysis does not provide such a general account, but at least, it shows more richness and uniformity than earlier ones. Second, we have not yet found one framework for all these sources, in that current work on limitations of inferential or computational powers should be incorporated as well. Next, concerning interaction in diverse societies of agents, we have not yet looked at scenarios involving bounded memory – the way game theorists have when discussing 'bounded rationality'. Here is where our dynamic epistemic or doxastic logics should meet up with current *game logics*, if we are to describe agents' longer-term strategies for collaboration, or competition, or more realistically, their frequent mixtures of both....

Even so, we hope that our account of diversity and interaction is of use per se in placing the phenomenon on the map, while it also may provide a fresh look at current logical systems for information update and belief revision. Our cognitive and social reality is that different agents live together, and interact with each other, sometimes with remarkable success. This rich set of phenomena is not just a playground for psychologists or sociologists: it also seems a legitimate challenge to logicians!

Acknowledgement
I thank three anonymous referees, as well as the audience of the *ESSLLI Workshop on Logics for Resource Bounded Agents* for their comments and questions. I thank Natasha Alechina for her invitation to the Nottingham FOP meetings, whose audience gave me very helpful feedback. I thank two anonymous reviewers from the *International Conference on Logic Navya-Nyaya & Application* in Calcutta for their questions concerning interaction, which led to several new results. Finally, I thank Johan van Benthem for his generosity in providing questions, updates, and corrections.

References

[AGM85] C. Alchourrón, P. Gardenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.

[AJL06] N. Alechina, M. Jago, and B. Logan. Modal logics for communicating rule-based agents. In A. Perini G. Brewka, S. Coradeschi and P. Traverso, editors, *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI 2006)*. IOS Press, 2006.

[Auc03] G. Aucher. A combination system for update logic and belief revision. Master's thesis, ILLC, University of Amsterdam, 2003.

[Auc05] G. Aucher. How our beliefs contribute to interpret actions. In L. Z. Varga M. Pechoucek, P. Petta, editor, *CEEMAS 2005*, pages 276 – 286. Springer, 2005. LNAI 3690.

[Ben96] J. van Benthem. *Exploring Logical Dynamics*. CSLI Publication, Stanford, 1996.

[Ben06] J. van Benthem. 'one is a lonely number': On the logic of communication. Tech Report, PP-2002-27, ILLC, University of Amsterdam. To appear in Z. Chatzidakis, P. Koepke, W. Pohlers, Eds., *Logic Collquium, ASL Lecture Notes in Logic 27*, AMS Publications, Providence (R.I.), 2006.

[Ben07] J. van Benthem. Dynamic logic for belief revision. Tech Report, PP-2006-11, ILLC, University of Amsterdam. To appear in the *Journal of Applied Non-Classical Logic*, 2007.

[BGK06] J. van Benthem, J. Gerbrandy, and B. Kooi. Dynamic update with probabilities. ILLC Research Reports, PP-2006-21, University of Amsterdam, 2006.

[BL04] J. van Benthem and F. Liu. Diversity of logical agents in games. *Philosophia Scientiae*, 8(2):163–178, 2004.

[BL07] J. van Benthem and F. Liu. Dynamic logic of preference upgrade. Tech Report, PP-2005-29, ILLC, University of Amsterdam. To appear in the *Journal of Applied Non-Classical Logic*, 2007.

[BMS98] A. Baltag, L. S. Moss, and S. Solecki. The logic of common knowledge, public announcements, and private suspicions. In I. Gilboa, editor, *Proceedings of the 7th conference on theoretical aspects of rationality and knowledge (TARK 98)*, pages 43–56, 1998.

[DHK07] H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Springer, Berlin, 2007. To appear.

[Dun95] P. H. Dung. An argumentation-theoretic foundation for logic programming. *Journal of Logic Programming*, 22:151–177, 1995.

[Egr04] P. Egre. *Propositional Attitudes and Epistemic Paradoxes*. PhD thesis, Universite Paris 1 et IHPST, 2004.

[FH85] R. Fagin and J. Y. Halpern. Belief, awareness, and limited reasoning. In *In Proceedings of the Ninth International Joint Conference on Artificial Intelligence (IJCAI-85)*, pages 480–490, 1985.

[Ger99] J. Gerbrandy. *Bisimulation on Planet Kripke*. PhD thesis, ILLC Amsterdam, 1999.

[GR95] P. Gardenfors and H. Rott. Belief revision. In D. M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 4. Oxford University Press, 1995.

[GS98] D. Gabbay and V. Shehtman. Products of modal logics. *Logic Journal of the IGPL*, 6(1):73–146, 1998. Part 1.

[GTtARG99] G. Gigerenzer, P. Todd, and the ABC Research Group. *Simple Heuristics that Make us Smart*. New York: Oxford University Press, 1999.

[HJT90] W. van der Hoek, J. Jaspars, and E. Thijsse. A general approach to multi-agent minimal knowledge. In G. Brewka M. Ojeda-Aciego, I.P. Guzman and L.M. Pereira, editors, *Proceedings of the 7th European Workshop on Logics in Artificial Intelligence (JELIA 2000)*, pages 254–268. Springer-Verlag, Heidelberg, 1990. LNAI 1919.

[Jag06] M. Jago. *Logics for Resource-Bounded Agents*. PhD thesis, The University of Nottingham, 2006.

[Kon88] K. Konolige. On the relation between default and autoepistemic logic. *Artificial Intelligence*, 35(3):343–382, 1988.

[KZ03] A. Kurucz and M. Zakharyaschev. A note on relativised products of modal logics. In F. Wolter P. Balbiani, N.-Y. Suzuki and M. Zakharyaschev, editors, *Advances in Modal Logic*, volume 4, pages 221–242. King's College Publications, 2003.

[Lev90] H. J. Levesque. All i know: a study in autoepistemic logic. *Artificial Intelligence Journal*, 42:381–386, 1990.

[Liu04] F. Liu. Dynamic variations: Update and revision for diverse agents. Master's thesis, ILLC, University of Amsterdam, 2004.

[Liu06] F. Liu. Preference change and information processing. In *Proceedings of the 7th Conference on Logic and the Foundations of Game and Decision Theory (LOFT 06)*, Liverpool, 2006. Tech Report, PP-2006-41, ILLC, University of Amsterdam.

[Pla89] J. A. Plaza. Logics of public announcements. In *Proceedings 4th International Symposium on Methodologies for Intelligent Systems*, 1989.

[Rot06] H. Rott. Shifting priorities: Simple representations for 27 iterated theory change operators. In H. Langerlund, S. Lindström, and R. Sliwinski, editors, *Modality Matters: Twenty-Five Essays in Honour of Krister Segerberg*, pages 359–384. Uppsala Philosophical Studies 53, 2006.

[SGG03] A. Silberschatz, P. B. Galvin, and G. Gagne. *Operating System Concepts*. John Wiley & Sons, Inc., New York, NY, USA, 2003.

[Sny04] J. Snyder. Product update for agents with bounded memory. Manuscript, Department of Philosophy, Stanford University, 2004.

[Spo88] W. Spohn. Ordinal conditional functions: A dynamic theory of epistemic states. In W. L. Harper et al., editor, *Causation in Decision, Belief Change and Statistics II*, pages 105–134. Kluwer, Dordrecht, 1988.

[Was00] R. Wassermann. *Resource Bounded Belief Revision*. PhD thesis, ILLC University of Amsterdam, 2000.