A Dynamic Analysis of Interactive Rationality

Eric Pacuit Olivier Roy

March 17, 2011

Abstract

Epistemic game theory has shown the importance of informational contexts to understand strategic interaction. We propose a general framework to analyze how such contexts may arise. The idea is to view informational contexts as the fixed-points of iterated, "rational responses" to incoming information about the agents' possible choices. We show general conditions for stabilization of such sequences of rational responses, in terms of structural properties of both the decision rule and the information update policy. In the process, we generalize existing rules for information updates used in the dynamic-epistemic logic literature. We then apply this framework to admissibility. We give a dynamic analysis of a well-known "paradox" arising from this choice rule, characterize stabilization of iterated rational response to admissibility under two different information update rules, and argue that these embody two different ways to respond to reasons in games.

1 Background and Motivation

An increasingly popular[1] view is that "*the* fundamental insight of game theory [is] that a rational player must take into account that the players reason about each other in deciding how to play" [4, pg. 81]. Exactly how the players (should) incorporate the fact that they are interacting with other (actively reasoning) agents into their own decision making process is the subject of much debate. A variety of frameworks have been put forward to explicitly model the *reasoning* of rational agents

[1] But, of course, not uncontroversial: Consider, for example, the following quote of Kadane and Larkey [20, pg. 239]: "It is true that a subjectivist Bayesian will have an opinion not only on his opponent's behavior, but also on his opponent's belief about his own behavior, his opponent's belief about his belief about his opponent's behavior, etc. (He also has opinions about the phase of the moon, tomorrow's weather and the winner of the next Superbowl.) However, in a single-play game, all aspects of his opinion except his opinion about his opponent's behavior are irrelevant, and can be ignored in the analysis by integrating them out of the joint opinion."

in a strategic situation. Key examples include Brian Skyrms' models of "dynamic deliberation" [30], Ken Binmore's analysis of "eductive reasoning" [9], and Robin Cubitt and Robert Sugden's "common modes of reasoning" [15]. Although the details of these frameworks are quite different[2], they share a common line of thought: Contrary to classical game theory, *solution concepts* are no longer the basic object of study. Instead, the "rational solutions" of a game are the result of individual (rational) decisions in specific informational "contexts".

This perspective on the foundations of game theory is best exemplified by the so-called epistemic program in game theory (cf. [12]). The central thesis here is that the basic mathematical model of a game should include an explicit parameter describing the players' *informational attitudes*. However, this broadly decision-theoretic stance does not simply *reduce* the question of decision-making in interaction to that of rational decision making in the face of uncertainty or ignorance. Crucially, *higher-order* information (belief about beliefs, etc.) are key components of the informational context of a game[3]. Of course, different contexts of a game can lead to drastically different outcomes. But this means that the informational contexts themselves are open to rational criticism:

> "It is important to understand that we have two forms of irrationality [...]. For us, a player is rational if he optimizes *and* also rules nothing out. So irrationality might mean not optimizing. But it can also mean optimizing while not considering everything possible." [13, pg. 314, our emphasis]

Thus, a player can be rationally criticized for not choosing what is *best given one's information*, but also for not reasoning *in* or *to* a "proper" context. Of course, what counts as a "proper" context is debatable. There might be rational pressure for or against making certain *substantive assumptions*[4]

[2]Skyrms assumes the players deliberate by calculating their *expected utility* and then use this new information to recalculate their probabilities about the states of the world and recalculate their expected utilities. Binmore models the players as *Turing machines* that can *compute* their rational choices. And, Cubitt and Sugden build on David Lewis' analysis of common knowledge in terms (inductive/deductive) rules that are commonly accepted among all the players.

[3]And so, strategic behavior *depends*, in part, on the players' higher-order beliefs. However, some care is needed about what precisely is being claimed. The well-known *email game* of Ariel Rubinstein [28] demonstrates that misspecification of arbitrarily high-orders of beliefs can have a large impact on (predicted) strategic behavior. So, there are simple examples where (predicted) strategic behavior is *too sensitive* to the players higher-order beliefs. We are not claiming that a rational agent is *required* to consider *all* higher-order beliefs, but only that a rational player recognizes that her opponents are actively reasoning rational agents, and this means that a rational player does take into account *some* of her higher-order beliefs (eg., what she believes her opponents believes that she will do) as she deliberates. Precisely, "how much" higher-order information should be taken into account is a very interesting open question which we set aside for this paper.

[4]The notion of substantive assumption is explored in more detail in [27], and the references therein.

about the beliefs of one's opponents, for instance to always entertain the possibility that one of the players might not choose optimally.

Recently, researchers using methods from dynamic-epistemic logic have taken steps into understanding this idea of reasoning *to* a "proper" or "rational" context [8, 7, 6, 34]. Building on this literature, and, more generally, on the rich repertoire of notions of belief (eg., safe belief, strong belief) and informative actions (eg., radical upgrade, conservative upgrade) from modern dynamic-epistemic logic[5], we provide here a general characterization of when players can or cannot rationally reason to specific informational contexts (Section 2). We then apply this framework to issues surrounding the epistemic characterization of *iterated elimination of weakly dominated strategies* (IEWDS), aka *iterated admissibility* (Section 3).

2 Belief Dynamics for Strategic Games

The main idea of this paper is to understand well-known solutions concepts not in terms of fixed informational contexts—for instance, models (eg., type spaces or epistemic models) satisfying rationality and common belief of rationality—but rather as a result of a dynamic, interactive process of information exchanges. It is important to note that the goal is *not* to represent some type of "pre-play communication" or form of "cheap talk". Instead, the goal is to represent the process of *rational deliberation* that takes the players from the *ex ante* stage to the *ex interim* stage of decision making. Thus, the "informational exchanges" are the result of the players *practical reasoning* about what they should do, given their current beliefs. In this section, we introduce our framework incorporating ideas from the extensive literature on dynamic logics of belief revision (cf. [33, 6]) and recent work on a "reasoning-based approach to game theory" [16, 15].

2.1 Describing the Informational Context

Let $G = \langle N, \{S_i\}_{i \in N}, u_i \rangle$ be a strategic game.[6] The informational context of a game describes the players' *hard* and *soft* information about the possible outcomes of the game. The players have opinions about which of states (each state is associated with an outcome of the game) are more

[5]The reader not familiar with this area can consult the recent textbook [33] for details.

[6]We assume the reader is familiar with the basic concepts of game theory.(for example, strategic games and various solution concepts, such as iterated removal of strictly (weakly) dominated strategies).

or less *plausible* among the ones they have not ruled out. Since we are representing the rational deliberation process, we do not assume that the players have made up their minds about which actions they will choose. We start with the simplest model of beliefs: a set of states where each state is associated with a possible outcome of the game and a single plausibility ordering (which is reflexive, transitive and connected) $w \preceq v$ that says "v is at least as plausible as w."

Originally used as a semantics for conditionals (cf. [22]), these *plausibility models* have been extensively used by logicians [32, 33, 6], game theorists [10] and computer scientists [11, 21] to represent rational agents' (all-out) beliefs. Thus, we take for granted that they provide natural models of (multiagent) beliefs and focus on how they can be used to represent "rational deliberation" in a game situation. We first settle on some notation.

Definition 2.1 (Strategy Functions and Propositions) Let W be a set of states and $G = \langle N, \{S_i\}_{i\in N}, u_i\rangle$ a strategic game. A **strategy function** on W for G is a function $\sigma : W \to \Pi_i S_i$ assigning strategy profiles to each state. To simplify notation, we write $\sigma_i(w)$ for $(\sigma(w))_i$ (similarly, write $\sigma_{-i}(w)$ for the sequence of strategies of all players except i). For each $a \in S_i$, we write $P^i_a = \{w \in W \mid \sigma_i(w) = a\}$ for the set of states where player i chooses a (this is the proposition "i selects strategy a"). Furthermore, we write $P^i_a \wedge P^j_b$ for the set-theoretic intersection of P^i_a and P^j_b, $P^i_a \vee P^j_b$ for the set-theoretic union of P^i_a and P^j_b and $\neg P^i_a$ for the set-theoretic complement of P^i_a. $\triangleleft$

We can now define the informational context of a game.

Definition 2.2 (Informational Context of a Game) Let $G = \langle N, \{S_i\}_{i\in N}, u_i\rangle$ be a strategic form game. An **informational context** of G is a plausibility model $\mathcal{M}_G = \langle W, \preceq, \sigma\rangle$ where $\preceq$ is a connected, reflexive, transitive and well-founded[7] relation on W and σ is a strategy function on W for G. We also say $\mathcal{M}_G$ is a model of G. $\triangleleft$

Note that there is only one plausibility ordering in the above model, yet we are interested in games with more than one player. There are different ways to interpret the fact that there is only one plausibility ordering. The models can represent the beliefs of one of the players before she has made up her mind about which option to choose. Alternatively, we can think the model as

[7]Well-foundedness is only needed to ensure that for any set X, the set of minimal elements in X is nonempty. This is important only when W is infinite – and there are ways around this in current logics. Moreover, the condition of connectedness can also be lifted, but we use it here for convenience.

describing a stage of the rational deliberation of *all* the players starting from an initial model where all the players have the same beliefs (i.e., the *common prior*). The private information about which possibilities the players consider possible given their actual choice can then be defined from the *conditional beliefs*. For $X \subseteq W$, let $Min_{\preceq}(X) = \{v \in X \mid v \preceq w \text{ for all } w \in X\}$ be the set of minimal elements of X according to $\preceq$.

Definition 2.3 (Belief and Conditional Belief) Let $\mathcal{M}_G = \langle W, \preceq, \sigma \rangle$ be a model of a game G. We define the following. Let E and F be subsets of W, we say

- E is **believed conditional on** F in $\mathcal{M}_G$, denoted $\mathcal{M}_G \models B^F(E)$, provided $Min_{\preceq}(F) \subseteq E$.

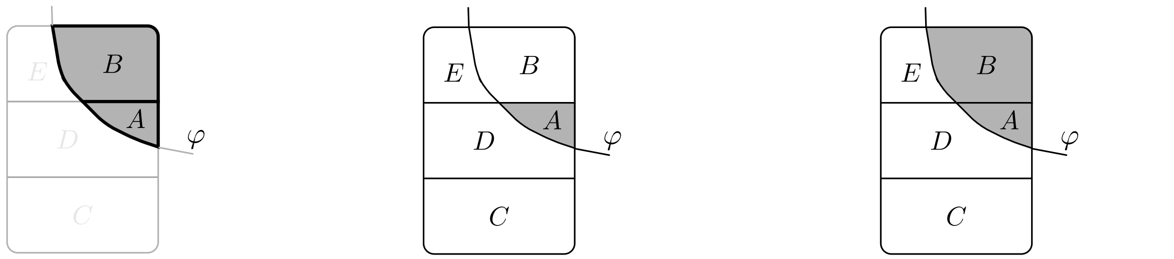
Also, we say E is **believed** in $\mathcal{M}_G$ if E is believed conditional on W. Thus, E is believed provided $Min_{\preceq}(W) \subseteq E$ ◁

Other notions of beliefs have been studied (eg. *strong belief* and *safe belief*), but, for this paper, we keep things simple and focus on the above standard notions.

Of course, this model can (and has: see [6, 33]) be extended to include beliefs for each of the players, an explicit relation representing the player(s) hard information or by making the plausibility orders state-dependent. To keep things simple, we focus on models with a single plausibility ordering.

2.2 A Primer on Belief Dynamics

We are not interested in informational contexts *per se*, but rather how the informational context changes during the process of rational deliberation. The type of change we are interested in is how a model $\mathcal{M}_G$ of a game G incorporates new information about what the players *should* do (according to a particular choice rule). As is well-known from the belief revision literature, there are many ways to transform an epistemic-doxastic model given some new information [26]. We do not have the space to survey this entire literature here (see [33, 5] for modern introductions). Instead we sketch some key ideas. The picture below illustrates different ways a plausibility model can incorporate the proposition φ (thought of a set of states).



$!(\varphi) : A \prec B$ $\quad\quad$ $\uparrow(\varphi) : A \prec C \prec D \prec B \cup E$ $\quad\quad$ $\Uparrow(\varphi) : A \prec B \prec C \prec D \prec E$

The general approach is to define a way of *transforming* a plausibility model $\mathcal{M}_G$ given a proposition φ. That is, the idea is to define a transformation τ that maps plausibility models and propositions to plausibility models (we write $\mathcal{M}_G^{\tau(\varphi)}$ for $\tau(\mathcal{M}_G, \varphi)$). So, given a model $\mathcal{M}_G$ of a game G and a proposition φ describing what the players (might/should/will) do, $\mathcal{M}_G^{\tau(\varphi)}$ is a new informational context taking this information into account. Different definitions of τ represent the different attitudes an agent can take towards the incoming information. For example, the operation on the left is the well-known *public announcement* operation [23, 17] which assumes that the players considers the source of φ *infallible* ruling out any possibilities that are inconsistent with φ. For the other transformations, while the players do *trust* the source of φ, they do not treat the source as infallible. These dynamic operations satisfy a number of interesting logical principles [33, 5] which we do not discuss in this paper.

We are interested in the operations that transform the informational context as the players deliberate about what they should do in a game situation. In each informational context (viewed as describing one stage of the deliberation process), the players determine which options are "*rationally permissible*" and which options the players ought to avoid (which is guided by some fixed rules of practical reasoning or a choice rule, cf. the discussion in the next section). This leads to a transformation of the informational context as the players adopt the relevant beliefs about the outcome of their *practical reasoning*. In this new informational context, the players again think about what they should do leading to another transformation. The main question is does this process *stabilize*?

The answer to this question will depend on a number of factors. The general picture is

$$\mathcal{M}_0 \overset{\tau(D_0)}{\Longrightarrow} \mathcal{M}_1 \overset{\tau(D_1)}{\Longrightarrow} \mathcal{M}_2 \overset{\tau(D_2)}{\Longrightarrow} \cdots \overset{\tau(D_n)}{\Longrightarrow} \mathcal{M}_{n+1} \Longrightarrow \cdots$$

where each D_i is some proposition and τ is model transformer (eg., public announcement, radical upgrade or conservative upgrade). Two questions are important for the analysis of this process. First, what type of transformations are being used by the players? For example, if τ is a public announcement, then it is not hard to see that, for purely logical reasons, this process must eventually stop at a limit model (see [6] for a discussion and proof). The second question is where do the propositions D_i come from? To see why this matters, consider the situation where you iteratively perform a radical upgrade with p and $\neg p$ (i.e., $\Uparrow(p), \Uparrow(\neg p), \ldots$). Of course, this sequence of upgrades never stabilizes. However, in the context of reasoning about what to do in a game situation, this situation may not arise because of special properties of the choice rule that is being used to describe (or guide) the players' decisions.

Suspending Judgement It is well-known that, in general, there are no rational principles of decision making (under ignorance or uncertainty) which *always* recommend a *unique* choice. In particular, it is not hard to find a game and an informational context where there is at least one player without a *unique* "rational choice" (see the discussion in the next section for concrete examples). How should a rational player incorporate the information that more than one action is classified as "choice-worthy" or "rationally permissible" (according to some choice rule) for her opponent(s)? One natural response is to *suspend judgement* about which options the relevant players will *pick*. Making use of a well-known distinction of Edna Ullmann-Margalit and Sidney Morgenbesser [31], the assumption that all players are rational can help determine which options the player will *choose*, but rationality alone does not help determined which of the rationally permissible options will be "picked". This line of thought led Cubitt and Sugden to impose a "privacy of tie breaking" property which says that players cannot *know* that her opponent will not pick an option that is classified as "choice-worthy" [15, pg. 8]. Wlodeck Rabinovich takes this even further and argues that from the principle of indifference, players must assign equal probability to all choice-worthy options [25]. What interests us is how to transform a plausibility model to incorporate the fact that there is a *set* of choice-worthy options for (some of) the players.

We do not offer an extended discussion of belief suspension here, but suggest that a generalization of *conservative upgrade* is the notion we are looking for (Wes Holliday makes the same point in [19]). The idea is to do an upgrade with a *set* of propositions $\{\varphi_1, \ldots, \varphi_n\}$ by letting the most plausible worlds be the union of each of the most plausible φ_i worlds. Formally,

Definition 2.4 (Generalized Conservative Upgrade.) Let $\mathcal{M} = \langle W, \preceq, \sigma \rangle$ be a plausibility model and $\{\varphi_1, \ldots, \varphi_n\}$ a set of propositions. Define $\mathcal{M}^{\uparrow\{\varphi_1 \ldots, \varphi_n\}} = \langle W^{\uparrow\{\varphi_1 \ldots, \varphi_n\}}, \preceq^{\uparrow\{\varphi_1 \ldots, \varphi_n\}}, \sigma^{\uparrow\{\varphi_1 \ldots, \varphi_n\}} \rangle$ as follows: $W^{\uparrow\{\varphi_1, \ldots, \varphi_n\}} = W$, $\sigma^{\uparrow\{\varphi_1 \ldots, \varphi_n\}} = \sigma$ and for all $i \in N$ and $w \in W^{\uparrow\{\varphi_1, \ldots, \varphi_n\}}$ we have: let $B = Min_{\preceq}(\varphi_1) \cup Min_{\preceq}(\varphi_2) \cup \cdots \cup Min_{\preceq}(\varphi_n)$

1. If $v \in B$ then $v \preceq^{\uparrow\{\varphi_1, \ldots, \varphi_n\}} x$ for all $x \in W$, and
2. for all $x, y \in W - B$, $x \preceq^{\uparrow\{\varphi_1, \ldots, \varphi_n\}} y$ iff $x \preceq y$. ◁

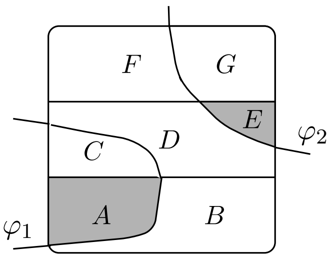
Note that this is indeed a generalizaiton of conservative upgrade, since in general, $best(\varphi_1) \cup \cdots best(\varphi_n) \neq best(\varphi_1 \vee \cdots \vee \varphi_n)$. Thus, this is not the same as performing a conservative upgrade with $\varphi_1 \vee \cdots \vee \varphi_n$ (where the most plausible worlds are the most plausible worlds satisfying at least one of the φ_i). A simple suspension of a belief in φ can be represented as $\uparrow \{\varphi, \neg\varphi\}$. This transformation is illustrated below.

Remembering Reasons A generalized conservative upgrade of $\{\varphi_1, \ldots, \varphi_n\}$ "flattens" out the players belief relative to this set of propositions. After the upgrade, the player will consider each of the φ_i equally plausible. But this means that, if w is a most plausible φ_i-world and v is a most plausible φ_j-world, the player forgets whatever reason she had for considering state w more plausible than v (or vice versa). This suggests a generalization of *radical upgrade* (cf. the picture above on the right) where the player(s) remember their earlier reasons for considering some states more plausible than others. Again the idea is to upgrade with a set of propositions $\{\varphi_1, \ldots, \varphi_n\}$ as above, but maintain the original ordering within the union of the most plausible φ_i-worlds.

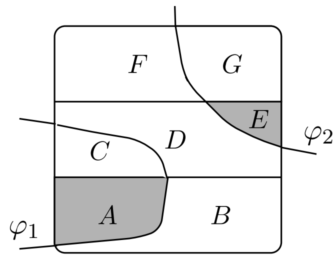
Definition 2.5 (Generalized Radical Upgrade) Let $\mathcal{M} = \langle W, \preceq, \sigma \rangle$ be a plausibility model and $\{\varphi_1, \ldots, \varphi_n\}$ a set of propositions. Define $\mathcal{M}^{\Uparrow\{\varphi_1 \ldots, \varphi_n\}} = \langle W^{\Uparrow\{\varphi_1 \ldots, \varphi_n\}}, \preceq^{\Uparrow\{\varphi_1 \ldots, \varphi_n\}}, \sigma^{\Uparrow\{\varphi_1 \ldots, \varphi_n\}} \rangle$ as follows: $W^{\Uparrow\{\varphi_1, \ldots, \varphi_n\}} = W$, $\sigma^{\Uparrow\{\varphi_1 \ldots, \varphi_n\}} = \sigma$ and for all $i \in N$ and $w \in W^{\Uparrow\{\varphi_1, \ldots, \varphi_n\}}$ we have: let $B = Min_{\preceq}(\varphi_1) \cup Min_{\preceq}(\varphi_2) \cup \cdots \cup Min_{\preceq}(\varphi_n)$

1. for all $v \in B$, $v \preceq^{\Uparrow\{\varphi_1, \ldots, \varphi_n\}} x$ for all $x \in W - B$,
2. for all $x, y \in B$, $x \preceq^{\Uparrow\{\varphi_1, \ldots, \varphi_n\}} y$ iff $x \preceq y$, and
3. for all $x, y \in W - B$, $x \preceq^{\Uparrow\{\varphi_1, \ldots, \varphi_n\}} y$ iff $x \preceq y$. ◁

We illustrate both transformations below:



$\uparrow\{\varphi_1, \varphi_2\} : A \cup E \prec B \prec C \cup D \prec F \cup G$



$\Uparrow\{\varphi_1, \varphi_2\} : A \prec E \prec B \prec C \cup D \prec F \cup G$

We will see other examples of these transformation in the next section. These transformations can be logically analyzed using standard techniques from dynamic epistemic/doxastic logic literature (eg., the "reduction axiom method").

Theorem 2.6 Both generalized conservative upgrade and generalized radical upgrade can be completely axiomatized over class of plausibility models over a static language including operators for belief, conditional belief and knowledge.

For the proof, we need to generalize the reduction axioms for radical and conservative upgrade (see [33] for a discussion of this method of proving completeness). Since the logic of these transformations is not the focus of this paper, we do not include the proof here.

2.3 Practical Reasoning in Games

We do not intend the dynamic operations of belief change discussed in the previous section to directly represent the (practical) *reasoning* of the players as they deliberate about what to do in a game situation. In fact, we do not represent directly any formal model of practical reasoning. Instead, following Cubitt and Sugden [15], we assume that during each stage of rational deliberation, the players can *categorize* their available options. Thus, we treat practical reasoning as a "black box" and focus on general *choice rules* that are intended to describe rational decision making (under ignorance). To make this precise, we need some notation:

Definition 2.7 (Strategies in Play) Let $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ be a strategic game and $\mathcal{M}_G = \langle W, \preceq, \sigma\rangle$ an informational context of G. For each $i \in N$, the strategies **in play for** i **at** w is the set

$$S_{-i}(\mathcal{M}_G) = \{s_{-i} \in \Pi_{j \neq i} S_j \mid \text{ there is a } w \in Min_{\preceq}(W) \text{ such that } \sigma_{-i}(w) = s_{-i}\} \qquad \triangleleft$$

This set $S_{-i}(\mathcal{M}_G)$ is the strategies that player i still *believes* are possible at some stage of the deliberation process represented by the model $\mathcal{M}_G$. Given these beliefs, we assume that each player can *categorize* her available options:

Definition 2.8 (Categorization) Let $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ be a strategic game and $\mathcal{M}_G = \langle W, \preceq, \sigma \rangle$ an informational context of G. A **categorization** for player i in $\mathcal{M}_G$ is a pair $\mathbf{S}_i(\mathcal{M}_G, w) = (S_i^+, S_i^-)$ where $S_i^+ \cup S_i^- \subseteq S_i$ and

$$\text{for each } a \in S_i, \text{ if there is no } v \in W \text{ with } \sigma_i(v) = a \text{ then } a \in S_i^-$$

If $\mathbf{S}_i(\mathcal{M}_G) = (S_i^+, S_i^-)$, we write $\mathbf{S}_i^+(\mathcal{M}_G)$ for S_i^+ and $\mathbf{S}_i^-(\mathcal{M}_G)$ for S_i^-. Also, we write $\mathbf{S}(\mathcal{M}_G)$ for the sequence of categorizations $(\mathbf{S}_1(\mathcal{M}_G), \ldots, \mathbf{S}_n(\mathcal{M}_G))$. $\triangleleft$

Note that, in general, a categorization need not be a partition (i.e., $S_i^+ \cup S_i^- = S_i$). See [16] for an example of such a categorization. However, in the remainder of this paper, we focus on familiar choice rules where the categorization does form a partition. Typically, we are interested in a "categorization method", i.e., rules for defining a categorization given any game. Two standard examples are weak and strong dominance: Let $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ be a strategic game and $\mathcal{M}_G$ an model of G. Then,

Strong Dominance (pure strategies) For each i and $a \in S_i$,

$$a \in S_i^- \text{ iff there is } b \in S_i \text{ such that for all } s_{-i} \in S_{-i}(\mathcal{M}_G),\ u_i(s_{-i}, b) > u_i(s_{-i}, a)$$

and $S_i^+ = S_i - S_i^-$.

Weak Dominance (pure strategies) For each i and $a \in S_i$,

$$a \in S_i^- \text{ iff there is } b \in S_i \text{ such that for all } s_{-i} \in S_{-i}(\mathcal{M}_G),\ u_i(s_{-i}, b) \geq u_i(s_{-i}, a) \text{ and there is some } s_{-i} \in S_{-i}(\mathcal{M}_G) \text{ such that } u_i(s_{-i}, b) > u_i(s_{-i}, a)$$

and $S_i^+ = S_i - S_i^-$.

Both of the above definitions can be modified to cover strict/weak dominance by mixed strategies, but we leave issues about how to incorporate probabilities into the framework sketched in this paper for another time.

2.4 Responding to Practical Reasoning

In this section, we merge the two perspectives (rational dynamics of belief from Section 2.2 and context-dependent practical reasoning of Section 2.3. The idea is that the "rational response" for a player to a given categorization is to transform the current informational context (using one of the transformations from the Section 2.2) by incorporating this information. To make this precise, we need to *describe* a categorization.

Definition 2.9 (Language for a Game) Let $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ be a strategic game. Without loss of generality, assume that each of the S_i are disjoint and let $\mathsf{At}_G = \{P^i_a \mid a \in S_i\}$ be a set of atomic formulas (one for each $a \in S_i$). The propositional language for G, denoted $\mathcal{L}_G$, is the smallest set of formulas containing At_G and closed under the boolean connectives $\neg$ and $\wedge$. The other boolean connectives ($\vee$, $\rightarrow$, $\leftrightarrow$) are defined as usual. $\triangleleft$

Formulas of $\mathcal{L}_G$ are intended to describe possible outcomes of the game. Given an informational context of a game $\mathcal{M}_G$, the formulas $\varphi \in \mathcal{L}_G$ is can be associated with subsets of the set of states in the usual way:

Definition 2.10 (Interpretation of $\mathcal{L}_G$) Let G be a strategic game, $\mathcal{M}_G = \langle W, \preceq, \sigma\rangle$ an informational context of G and $\mathcal{L}_G$ a propositional language for G. We define a map $[\![\cdot]\!]_{\mathcal{M}_G} : \mathcal{L}_G \to \wp(W)$ by induction on the structure of $\mathcal{L}_G$ as follows: $[\![P^i_a]\!]_{\mathcal{M}_G} = \{w \mid \sigma(w)_i = a\}$, $[\![\neg\varphi]\!]_{\mathcal{M}_G} = W - [\![\varphi]\!]_{\mathcal{M}_G}$ and $[\![\varphi \wedge \psi]\!]_{\mathcal{M}_G} = [\![\varphi]\!]_{\mathcal{M}_G} \cap [\![\psi]\!]_{\mathcal{M}_G}$. $\triangleleft$

Let X and Y be two sets of propositions, we define $X \wedge Y := \{\varphi \wedge \psi \mid \varphi \in X, \psi \in Y\}$

Definition 2.11 (Describing a categorization) Let G be a game and $\mathcal{M}_G$ an informational context of G. Given a categorization $\mathbf{S}(\mathcal{M}_G)$, let $Do(\mathbf{S}(\mathcal{M}_G))$ denote the set of formulas that *describe* $\mathbf{S}$. This set is defined as follows: for each $i \in N$ let:

$$Do_i(\mathbf{S}_i(\mathcal{M}_G)) = \{P^i_a \mid a \in \mathbf{S}^+_i(\mathcal{M}_G)\} \cup \{\neg P^i_b \mid b \in \mathbf{S}^-_i(\mathcal{M}_G)\}$$

Then define $Do(\mathbf{S}(\mathcal{M}_G)) = Do_i(\mathbf{S}_i(\mathcal{M}_G)) \bigwedge Do_2(\mathbf{S}_2(\mathcal{M}_G)) \cdots \bigwedge Do_n(\mathbf{S}_n(\mathcal{M}_G))$. $\triangleleft$

The general project is to understand the interaction between types of categorizations (eg., choice rules) and types of model transformations (representing the rational deliberation process). One key question, is does (and under what conditions) a deliberation process *stabilize*? There are a number of ways to make precise what it means to stabilize (see [6] for a discussion). In general there are many (rational) ways to incorporate a (description of) a categorization into a plausibility model. Section 2.2 discussed two natural update rules (generalized conservative upgrade and generalized radical upgrade). In the remainder of this section, we discuss a number of abstract principles which guarantee that a rational deliberation will *stabilize*. We start by being more precise about the "rational deliberation process" and what it means to stabilize.

Definition 2.12 (Stable in Beliefs) Suppose $\mathcal{M} = \langle W, \preceq, \sigma \rangle$ and $\mathcal{M}' = \langle W, \preceq', \sigma' \rangle$ are two plausibility models based on the same set of states[8]. We say $\mathcal{M}$ and $\mathcal{M}'$ are **stable with respect to the players' beliefs** if the set of propositions that are believed in $\mathcal{M}$ is the same as those believed in $\mathcal{M}'$. Equivalent, $\mathcal{M}$ and $\mathcal{M}'$ are stable with respect to beliefs provided $Min_{\preceq}(W) = Min_{\preceq'}(W)$. We write $\mathcal{M} \equiv_B \mathcal{M}'$ is $\mathcal{M}$ and $\mathcal{M}'$ are stable with respect to beliefs. ◁

In this paper, it is enough to define stabilization in terms of the players simple beliefs. This is because, during the process of deliberation, we only incorporate *ground information* about what the players are going to do (as opposed to higher-order information[9]). We are now ready to formally define a "deliberation sequence":

Definition 2.13 (Upgrade Sequence) Given a game G and an informational context $\mathcal{M}_G$, an upgrade sequence of type τ, induced by $\mathcal{M}_G$ is an infinite sequence of plausibility models $(\mathcal{M}_m)_{m\in\mathbb{N}}$ defined as follows:

$$\mathcal{M}_0 = \mathcal{M}_G \qquad \mathcal{M}_{m+1} = \tau(\mathcal{M}_m, Do(\mathcal{M}_m))$$

An upgrade sequence **stabilizes** if there is an $n \geq 0$ such that $\mathcal{M}_n = \mathcal{M}_{n+1}$. ◁

The next section has a number of examples of upgrade streams, some that stabilize and others that do not stabilize. We now discuss some abstract principles that ensure that the categorizations are "sensitive" to the players beliefs and that the players respond to the categorizations in the appropriate way. There are two main reasons why an upgrade stream would stabilize. The first is

[8]So, we assume that the models agree about which outcomes of the game have not been ruled out.

[9]An interesting extension would be to start with a multiagent belief model and allow players to not only incorporate information about which options are "choice-worthy", but also what beliefs their opponents may have. We leave this extension for future work and focus on setting up the basic framework here.

due to the properties of the transformation (for example, it is clear that upgrade streams with public announcements always stabilize). The second is because the choice rule satisfies a monotonicity property so that, eventually, the categorizations stabilizes, and no new transformations can change the plausibility ordering.

To state these properties more precisely, we need some notation. Let U be a fixed set of states and G a fixed strategic game. We restrict attention to transformations between models of G whose states come from the universe of states U. Let $\mathbb{M}_G$ be the set of all such plausibility models. A model transformation then is a function that maps a model of G and a finite set of formulas of $\mathcal{L}_G$ to a model in $\mathbb{M}_G$:

$$\tau : \mathbb{M}_G \times \wp_{<\omega}(\mathcal{L}_G) \rightarrow \mathbb{M}_G$$

where $\wp_{<\omega}(\mathcal{L}_G)$ is the set of finite subsets of $\mathcal{L}_G$. Of course, not all transformations τ make sense in this context. We give a number of abstract principles that τ must satisfy so that the categorizations and belief transformation τ are connected in the "right way". Let $\mathcal{X} = \{\varphi_1, \ldots, \varphi_n\}$ be a finite set of $\mathcal{L}_G$ formulas $\mathcal{M} \in \mathbb{M}_G$.

A1 The operation τ depends only on the truth set of the formulas: If for each $i = 1, \ldots, n$, $[\![\varphi_i]\!]_{\mathcal{M}} = [\![\psi_i]\!]_{\mathcal{M}}$, then $\tau(\mathcal{M}, \mathcal{X}) = \tau(\mathcal{M}, \{\psi_1, \ldots, \psi_n\})$

A2 The operation τ is idempotent[10] in the language $\mathcal{L}_G$: $\tau(\mathcal{M}, \mathcal{X}) = \tau(\mathcal{M}^{\tau(\mathcal{X})}, \mathcal{X})$

Property **A1** says that the belief transformations depend only on the proposition expressed by a formula φ by treating equivalent formulas the same way. The second property **A2** says that receiving the exact same information twice does not have any effect on the players' beliefs. These are general properties of the belief transformation. Certainly, there are other natural properties that one may want to impose (for example, variants of the AGM postulates [1]), but, for the time being, we are interested in the minimal principles needed to prove a stabilization result.

The next set of properties make sure that the transformation respond "properly" to a categorization. First, we need a property to guarantee that the categorizations only depend on the players' beliefs:

C1 If $\mathcal{M} \equiv_B \mathcal{M}'$ then $\mathbf{S}(\mathcal{M}) = \mathbf{S}(\mathcal{M}')$.

[10] Here it is crucial that the language $\mathcal{L}_G$ does not contain any modalities.

Finally, we need to ensure that in all upgrade sequences respond to the categorizations in the right way.

C2$^-$ For any upgrade sequence $(\mathcal{M}_n)_{n\in\mathbb{N}}$ in τ, if $a \in S_i^-(\mathcal{M}_n)$ then $\mathcal{M}_{n+1} \models B\neg P_i^a$

C2$^+$ For any upgrade sequence $(\mathcal{M}_n)_{n\in\mathbb{N}}$ in τ, if $a \in S_i^+(\mathcal{M}_n)$ then $\mathcal{M}_{n+1} \models \neg B\neg P_i^a$

Mon$^-$ For any upgrade sequence $(\mathcal{M}_n)_{n\in\mathbb{N}}$, for all $n \geq 0$, for all players $i \in N$, $\mathbf{S}_i^-(\mathcal{M}_n) \subseteq \mathbf{S}_i^-(\mathcal{M}_{n+1})$

Mon$^+$ Either for all models $\mathcal{M}_G$, $\mathbf{S}_i^+(\mathcal{M}_G) = S_i - \mathbf{S}_i^-(\mathcal{M}_G)$ or for any upgrade sequence $(\mathcal{M}_n)_{n\in\mathbb{N}}$, for all $n \geq 0$, for all players $i \in N$, $\mathbf{S}_i^+(\mathcal{M}_n) \subseteq \mathbf{S}_i^+(\mathcal{M}_{n+1})$

In particular, **Mon$^-$** means that once an option for a player is classified as "not rationally permissible", it cannot, at a later stage of the deliberation process, drop this classification.

Theorem 2.14 Suppose that G is a finite game and all of the above properties are satisfied. Then every upgrade sequence $(\mathcal{M}_n)_{n\in\mathbb{N}}$ stabilizes.

Proof. Let $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ be a finite strategic game. By properties **Mon$^-$** and **Mon$^+$** we have either for all upgrade streams $(\mathcal{M}_n)_{n\in\mathbb{N}}$ and players $i \in N$,

1. $\mathbf{S}_i^-(\mathcal{M}_0) \subseteq \mathbf{S}_i^-(\mathcal{M}_1) \subseteq \cdots \mathbf{S}_i^-(\mathcal{M}_n) \subseteq \cdots$ is an infinitely increasing sequence of subsets of S_i and $\mathbf{S}_i^+(\mathcal{M}_0) \supseteq \mathbf{S}_i^+(\mathcal{M}_1) \supseteq \cdots \mathbf{S}_i^+(\mathcal{M}_n) \supseteq \cdots$ is an infinite decreasing sequence of subsets of S_i; or

2. Both, $\mathbf{S}_i^-(\mathcal{M}_0) \subseteq \mathbf{S}_i^-(\mathcal{M}_1) \subseteq \cdots \mathbf{S}_i^-(\mathcal{M}_n) \subseteq \cdots$ and $\mathbf{S}_i^+(\mathcal{M}_0) \subseteq \mathbf{S}_i^+(\mathcal{M}_1) \subseteq \cdots \mathbf{S}_i^+(\mathcal{M}_n) \subseteq \cdots$ are infinite increasing sequences of subsets of S_i.

Since each S_i is assumed to be finite, for each player i, there is a n_i such that $\mathbf{S}_i^-(\mathcal{M}_{n_i}) = \mathbf{S}_i^-(\mathcal{M}_{n_i+i})$ and $\mathbf{S}_i^+(\mathcal{M}_{n_i}) = \mathbf{S}_i^+(\mathcal{M}_{n_i+i})$. Let m be the maximum of $\{n_i \mid i \in N\}$. Then, we have $\mathbf{S}(\mathcal{M}_m) = \mathbf{S}(\mathcal{M}_{m+1})$. All that remains is to show that for all $x > m$, $\mathcal{M}_x = \tau(\mathcal{M}_x)$. This follows by an easy induction on x. The key calculation is: for each $x \in \mathbb{N}$, let $\mathcal{D}_x$ be the appropriate description of $\mathbf{S}(\mathcal{M}_x)$.

$$
\begin{aligned}
\mathcal{M}_{m+2} = \tau(\mathcal{M}_{m+1}, \mathcal{D}_{m+1}) &= \tau(\mathcal{M}_n^{\tau(\mathcal{D}_m)}, \mathcal{D}_{m+1}) \\
&= \tau(\mathcal{M}_m^{\tau(\mathcal{D}_m)}, \mathcal{D}_m) \text{ (since } \mathbf{S}(\mathcal{M}_m) = \mathbf{S}(\mathcal{M}_{m+1})) \\
&= \tau(\mathcal{M}_m, \mathcal{D}_m) = \mathcal{M}_{m+1}
\end{aligned}
$$

This conclude the proof. QED

The role of monotonicity of the choice has been noticed by a number of researchers (see [3] for a discussion). We do not discuss the proof here[11], but note some interesting corollaries:

Corollary 2.15 If the categorization method is strict dominance, then any upgrade sequence of type τ stabilizes, where τ is any of the transformations discussed in this paper (eg., public announcement, (generalized) radical upgrade and (generalized) conservative upgrade).

Corollary 2.16 If τ is public announcement or (generalized) radical upgrade, then any belief sensitive categorization method stabilizes on any upgrade sequence.

This generalizes van Benthem's analysis of rational dynamics [8] to soft information, both in terms of attitudes and announcements. It also explains Apt and Zvesper's results about stabilization of beliefs, even for admissibility [2]: they use public announcements, which stabilizes beliefs.

3 Case Study: Admissibility

Larry Samuelson [29] pointed out an explicit puzzle surrounding the epistemic foundations of IEWDS - also known as the IA solution [13]. He showed (among other things) that there is no epistemic model of the following game with at least one state satisfying "common knowledge of admissibility" (i.e., a state where there is common knowledge that the players do not play a strategy that is weakly dominated).

[11] The proof is available upon request.

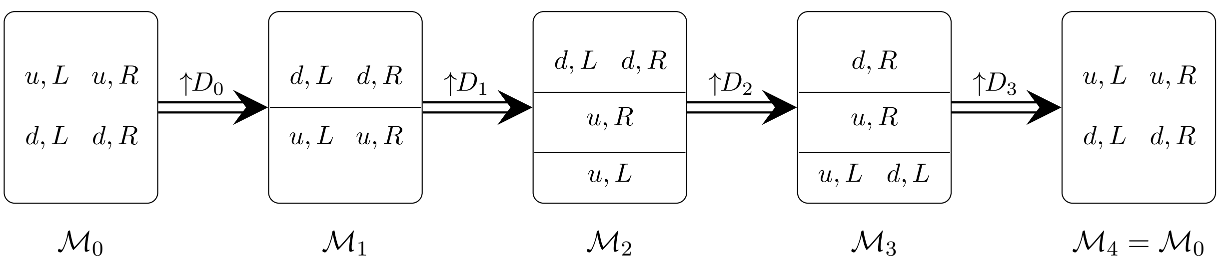
		Bob	
		L	R
Ann	u	$1,1$	$1,0$
	d	$1,0$	$0,1$

The general framework introduced above offers a new, dynamic perspective on this result, and on reasoning with admissibility more generally.[12]

3.1 Generalized Conservative Upgrade with Admissibility

Dynamically, the non-existence of a model of the above game in which admissibility is common belief[13] corresponds to non-stabilization of upgrade streams. Agents are not able to reason their way to stable, common belief in admissibility. To capture this intuition we need to work with a non-monotonic upgrade rule. We choose here generalized conservative upgrade. We explain in more detail this modeling choice in discussion at the end of Section 3.2.

Observation 3.1 Starting with the initial full model of the above game,[14] conservative upgrade stream of rational response with admissibility does not stabilize.



The looping stream of conservative upgrades is illustrated in the figure above. The formal details are in the appendix. Intuitively: From $\mathcal{M}_0$ to $\mathcal{M}_2$ the agents have reasons to exclude d and R, leading them to commonly believe that u, L is played. At that stage, however, d is admissible for Ann, canceling the reason that agents had to rule out this strategy. The rational response here is thus to suspend judgment on d, leading to $\mathcal{M}_3$. In this new model, the agents are similarly led to

[12] We do not to provide an alternative epistemic characterization of this solution concept. Both [13] and [18] have convincing results here. Our goal is to use this solution concept as illustration of our general approach.

[13] We use common plain beliefs here instead of common knowledge.

[14] A full model is one where it is common knowledge that each of the outcomes of the game equally likely (i.e., all outcomes of the game are in the model and are equally plausible).

suspend judgment on not playing R, bringing them back to M_0. This process loops forever; the agents' reasoning does not stabilize.

A corollary of Observation 3.1 is that common belief in admissibility is not sufficient for stabilization of upgrade streams. Stabilization also require that all *and only* those profiles that are most plausible are admissible. Admissibility needs to be a common *assumption* [14]. To show this we need the following notion.

Definition 3.2 (Tight admissible sets) A set of strategy profile S is *tightly admissible* iff, for all agents i:

1. All strategies σ_i for $\sigma \in S$ are admissible given the σ'_{-i} in S_{-i} and;

2. no other strategy s'_i is admissible given the σ'_{-i} in S_{-i}. ◁

It is a straightforward observation that if an upgrade stream based on admissibility stabilizes, then the most plausible profiles in its fixed-point constitute a tightly admissible set. The converse is also true: if the most plausible profiles in a given model constitute a tightly admissible set, then any upgrade sequence starting from that model and based on admissibility stabilizes

Tight admissible sets have a natural epistemic characterization:

Definition 3.3 (Levels of conditional beliefs) Let $\mathcal{M}$ be a model for a given game. We write B^0_i for $\{w \mid w \preceq w' \text{ for all } w' \in W\}$ and B^{n+1} for $\{w \mid w \preceq w' \text{ for all } w' \in W - B^n\}$. ◁

The set B^0 is just the set of most plausible states (the "simple" beliefs). The set B^{n+1} fixes beliefs conditional on not being in any of the state in $\bigcup_{0 \leq n} B^n$. We write $\mathcal{M} \models B^n Adm_j$ whenever j plays an admissible strategy in all most plausible states at level n.

Observation 3.4 Let $\mathcal{M}$ be a model for a given game.

(i) $\{s : \text{there is a } w \in Min_{\preceq}[W] \text{ with } \sigma(w) = s\}$ is a tight admissible set.

(ii) $\mathcal{M} \models B^0 \bigwedge_i Adm_i \wedge \bigwedge_{n>0} B^n \bigwedge_i \neg Adm_i$

Proof. Unpacking the definitions. QED

Tight admissible sets are thus those that are played exactly under, first, common belief in admissibility and, second, common conditional beliefs, for any degree $n > 0$, that no other strategies would have been admissible for any other players. Stabilization of rational, conservative upgrades with admissibility is equivalent to establishing common assumption of admissibility.

Upgrade streams based on generalized conservative upgrade and admissibility, if they stabilize, need not do so on the IA solution, nor on self-admissible sets (SAS) [13].

Observation 3.5 There are self-admissible sets that are not tightly admissible.

Proof. Consider the following game.

		Bob L	Bob R
Ann	u	$2, 2$	$2, 2$
Ann	m	$3, 1$	$0, 0$
Ann	d	$0, 0$	$1, 3$

$\{(u, R)\}$ is an SAS, but is not a closed admissible set. QED

Observation 3.6 There are games where tightly admissible sets do not coincide with the IA solution.

Proof. Consider the following game:

		Bob L	Bob M	Bob R
Ann	u	$1, 0$	$0, 1$	$0, 0$
Ann	m	$1, 0$	$1, 1$	$1, 0$
Ann	d	$0, 0$	$1, 0$	$1, 1$

(m, M) is the IA solution of that game, but $\{d, m\} \times \{M, R\}$ is a tight admissible set. QED

Finally, we note that for a given game there can be many tight admissible sets.

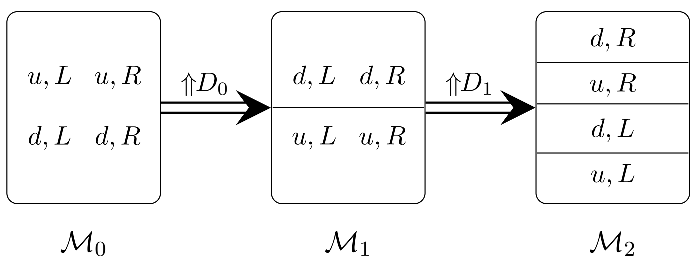
Observation 3.7 For a given game G there can be many tight admissible sets.

Proof. Take the 8 × 8 matrix constituted of one copy of the game in Observation **??** in the top-left quadrant, and another copy in the bottom-right quadrant. Both player get 0 on all the other profiles, i.e. top-right and bottom-left quadrants. The tight admissible sets in each copies of the game above are also tight admissible in this bigger game, and also the set of profiles formed by the product of strategies in the union of these two closed admissible sets[15]. QED

3.2 Generalized Radical Upgrade with Admissibility

Generalized radical upgrade stabilizes plain beliefs, and as showed above this is sufficient to guarantee stabilization of rational responses, even for non-monotonic choice rules like admissibility.

The game on page 16 is an illuminating special case. Starting with the full model of this game, the upgrade stream stabilizes on a model with common belief in admissibility. Again we only show the figure, and leave the formal details to the reader.



Intuitively, what happens is the following: Just like for conservative upgrade, $\mathcal{M}_0$ and $\mathcal{M}_1$, respectively, give the agents reasons to believe that Ann will not play d, and that Bob will not play R. This leads to M_2 where, like before, d is admissible given that Ann believes that Bob plays L. Radical update, however, doesn't allow this fact to overwrite the reason she had not to play d: her rational response is to rank u, L and d, L above all other possible outcomes, but to keep the relative ordering of these two, reflecting the fact that she previously ruled out u.

Stabilization of radical upgrade puts Samuelson's observation into perspective. Such upgrade forces the agents to remember the reasons they had earlier in the deliberation. Previous reasons constraint the domain of permissibility at later stages in the reasoning process. What is permissible for Ann at $\mathcal{M}_2$ depends on the reasoning process that led to this model, and in particular on the

[15]Note that this is not the case in general.

existence of an (earlier) reason not to play d. This was not the case for conservative upgrade. Reasons, at each stage, were evaluated *de novo*, without references to the preceding reasoning history. This is what led the upgrade stream for Samuelson's game into looping, to the "paradox" of admissibility. We leave it open for discussion whether this constitutes an argument to the effect that agents "should" keep track of their reasons while reasoning to a specific informational context. For now we content ourselves with the observation that there is a tight connection, on the one hand, between remembering one's reasons and stabilization of reasoning under admissibility and, on the other hand, between letting new reasons override previous ones and the possibility of never-ending reasoning chains.

4 Concluding remarks

In this paper we proposed a general framework to analyze how "proper" or "rational" information contexts my arise. We showed general conditions for stabilization of sequences of rational responses to incoming information, in terms structural properties of both the decision rule and the information update policy. In the course of doing so we generalized existing rules for information update used in the dynamic-epistemic logic literature. We then applied this framework to admissibility, giving a dynamic analysis of Samuelson's non-existence result, as well as characterizing stabilization of iterated rational response to admissibility under two different rules for information update.

Throughout the paper we worked with (logical) models of *all out* attitudes, leaving aside probabilistic, graded beliefs, even though the later are arguably the most widely used in the current literature on epistemic foundations of game theory. It is an important, but non-trivial task to transpose the dynamic perspective on informational contexts that we advocate here to such probabilistic models. We leave it for future work.

Finally, we should stress that the dynamic perspective on informational contexts is a natural complement, and not an alternative to existing epistemic characterization of solution concepts [35]. The later offer rich insights on the consequences of taking seriously the informational contexts of strategic interaction. What we proposed here is a first step towards understanding how or why such context might arise.

References

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

[2] Apt, K., and Zvesper, J. Public announcements in strategic games with arbitrary strategy sets. In *Proceedings of LOFT 2010* (2010).

[3] Apt, K., and Zvesper, J. The role of monotonicity in the epistemic analysis of strategic games. *Games 1*, 4 (2010), 381 – 394.

[4] Aumann, R., and Dreze, J. Rational expectations in games. *American Economic Review 98* (2008), 72 – 86.

[5] Baltag, A., and Smets, S. ESSLLI 2009 course: Dynamic logics for interactive belief revision. Slides available online at `http://alexandru.tiddlyspot.com/#%5B%5BESSLLI09%20COURSE%5D%5D`, 2009.

[6] Baltag, A., and Smets, S. Group belief dynamics under iterated revision: Fixed points and cycles of joint upgrades. In *Proceedings of Theoretical Aspects of Rationality and Knowledge* (2009).

[7] Baltag, A., Smets, S., and Zvesper, J. Keep 'hoping' for rationality: a solution to the backwards induction paradox. *Synthese 169* (2009), 301–333.

[8] Benthem, J. v. Rational dynamics and epistemic logic in games. *International Game Theory Review 9*, 1 (2007), 13–45.

[9] Binmore, K. Modeling rational players: Part I. *Economics and Philosophy 3* (1987), 179 – 214.

[10] Board, O. Dynamic interactive epistemology. *Games and Economic Behavior 49* (2004), 49 – 80.

[11] Boutilier, C. *Conditional Logics for Default Reasoning and Belief Revision*. PhD thesis, University of Toronto, 1992.

[12] Brandenburger, A. The power of paradox: some recent developments in interactive epistemology. *International Journal of Game Theory 35* (2007), 465–492.

[13] Brandenburger, A., Friedenberg, A., and Keisler, H. J. Admissibility in games. *Econometrica 76* (2008), 307–352.

[14] Brandenburger, A., and Keisler, H. J. An impossibility theorem on belies in games. *Studia Logica 84* (July 2006), 211–240.

[15] Cubitt, R., and Sugden, R. Common reasoning in games: A Lewisian analysis of common knowledge of rationality. CeDEx Discussion Paper, 2011.

[16] Cubitt, R., and Sugden, R. The reasoning-based expected utility procedure. *Games and Economic Behavior* (2011), In Press.

[17] Gerbrandy, J. *Bisimulations on Planet Kripke*. PhD thesis, University of Amsterdam, 1999.

[18] Halpern, J., and Pass, R. A logical characterization of iterated admissibility. In *Proceedings of the Twelfth Conference on Theoretical Aspects of Rationality and Knoweldge* (2009), A. Heifetz, Ed., pp. 146 – 155.

[19] Holliday, W. Trust and the dynamics of testimony. In *Logic and Interaction Rationality: Seminar's Yearbook 2009* (2009), ILLC Technical Reports, pp. 147 – 178.

[20] Kadane, J. B., and Larkey, P. D. Subjective probability and the theory of games. *Management Science 28*, 2 (1982), 113–120.

[21] Lamarre, P., and Shoham, Y. Knowledge, certainty, belief and conditionalisation. In *Proceedings of the International Conference on Knoweldge Representation and Reasoning* (1994), pp. 415 – 424.

[22] Lewis, D. *Counterfactuals*. Blackwell Publishers, Oxford, 1973.

[23] Plaza, J. Logics of public communications. In *Proceedings, 4th International Symposium on Methodologies for Intelligent Systems* (1989), M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, and Z. Ras, Eds., pp. 201–216 (republished as [24]).

[24] Plaza, J. Logics of public communications. *Synthese: Knowledge, Rationality, and Action 158*, 2 (2007), 165 – 179.

[25] Rabinowicz, W. Tortous labyrinth: Noncooperative normal-form games between hyperrational players. In *Knowledge, Belief and Strategic Interaction* (1992), C. Bicchieri and M. L. D. Chiara, Eds., pp. 107 – 125.

[26] Rott, H. Shifting priorities: Simple representations for 27 itereated theory change operators. In *Modality Matters: Twenty-Five Essays in Honour of Krister Segerberg* (2006), H. Lagerlund, S. Lindström, and R. Sliwinski, Eds., vol. 53 of *Uppsala Philosophical Studies*, pp. 359 – 384.

[27] Roy, O., and Pacuit, E. Substantive assumptions and the existence of universal knowledge structures: A logical perspective. Under submission, 2010.

[28] Rubinstein, A. The electronic mail game: A game with almost common knowledge. *American Economic Review 79* (1989), 385 – 391.

[29] Samuelson, L. Dominated strategies and common knowledge. *Game and Economic Behavior 4* (1992), 284 – 313.

[30] Skyrms, B. *The Dynamics of Raitonal Deliberation*. Harvard University Press, 1990.

[31] Ullmann-Margalit, E., and Morgenbesser, S. Picking and choosing. *Social Research 44* (1977), 757 – 785.

[32] van Benthem, J. Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics 14*, 2 (2004), 129 – 155.

[33] van Benthem, J. *Logical Dynamics of Information and Interaction*. Cambridge University Press, 2010.

[34] van Benthem, J., and Gheerbrant, A. Game solution, epistemic dynamics and fixed-point logics. *Fund. Inform. 100* (2010), 1–23.

[35] van Benthen, J., Pacuit, E., and Roy, O. Towards a theory of play: A logical perspective on games and interaction. *Games 2*, 1 (2011), 52–86.