Bayesian updating rules and AGM belief revision

Pathikrit Basu

February 26, 2018

Abstract

We interpret the problem of updating beliefs as a choice problem (selecting a posterior from a set of admissible posteriors) with a reference point (prior). We use AGM belief revision to define the support of admissible posteriors after arrival of information, which applies also to zero probability events. We study two classes of updating rules for probabilities: 1) "lexicographic" updating rules where posteriors are given by a lexicographic probability system 2) "minimum distance" updating rules which select the posterior closest to the prior by some metric. For the former, we show that an updating rule is lexicographic if and only if it is Bayesian, AGM-consistent and satisfies a weak form of path independence. For the latter, we show Bayesian updating rules may or may not have a minimum distance representation. For specific AGM belief revisions, we provide necessary and sufficient conditions for a minimum distance representation. Lastly, we study a sub-class of lexicographic updating rules, which we call "support-dependent" rules. We show that such updating rules have a minimum distance representation.

1 Introduction

Updating beliefs in light of newly acquired information is a problem that is relevant and occurs in many situations in economic theory. In most environments, an agent’s belief is represented by a probability measure over a state space, elements of which are payoff relevant. As a result, actions of the agent depend crucially on his opinion or belief over the


†California Institute of Technology. E-mail : pathkrtb@caltech.edu.
state variable and consequently also on how he chooses to update his belief upon learning an event. A dominant principle used to update probabilities in most models is Bayesian updating. Starting with a prior $\pi$ and observing a positive probability event $A$, Bayesian updating suggests the posterior $\pi'(E) = \pi(E \cap A)/\pi(A)$. However, in the event of a surprise i.e. observing a zero probability event, Bayesian updating remains silent and is not well-defined. Such situations arise in games of imperfect information where a player’s strategy must specify which action to choose in an information set that is reached with zero probability and the updating problem is one of assigning probabilities to nodes in the information set. The solution concepts of sequential equilibrium and trembling-hand perfect equilibrium (see ? and ?) both place restrictions on admissible beliefs on information sets which lie off path (using Bayesian updating otherwise). These put restrictions on off-path beliefs by requiring them to be the limit of on-path beliefs corresponding to a sequence of perturbations of the equilibrium strategy profile. Abstracting away from the game-theoretic scenario, we ask whether in the probabilistic model itself, there exists a systematic way to extend Bayesian updating to zero probability events.

In this paper, we interpret the updating problem as a choice problem with a reference point (see ?)). The reference point here is a prior $\pi$ on the state space and given an event $E$ the choice problem is one of choosing a posterior from an admissible set of posteriors which have a common support in $E$. But how should the “admissible set of posteriors” be selected? Since we wish to extend Bayesian updating, if $\pi(E) > 0$ we would want this set to be all probability measures whose support is the intersection of the support of the prior and the observed positive probability event and choice would be the Bayesian posterior. But what happens if $E$ has zero probability? Is there a consistent way to select the set of admissible posteriors for all events $E$?

To address these questions, we use a non-probabilistic theory of belief revision from propositional logic, namely, AGM belief revision (see ?). In this theory, an agent’s primitive is a belief set, which is a set of propositions or events that the agent believes to be true (in the present context this would be all probability one events according to the prior). Upon learning an event or proposition, AGM offers a sound and well-defined procedure to update belief sets, even when the observed information is inconsistent with one’s belief set. This property of the theory is, as we shall see, intimately linked with the problem of updating over zero probability events. We use this procedure to define the admissible set of posteriors in the following way: 1) We associate with each prior its belief set i.e all probability one events. 2) Upon observing an event, use AGM revision to update the be-
lief set. 3) Define the set of admissible posteriors to be the set of all probability measures whose belief set equals the updated belief set. The advantage of this procedure is that it is well defined for all events including zero-probability events. It also has the desirable property that all admissible posteriors have common support in the event observed and hence, updating satisfies consequentialism. When an updating rule abides by the AGM procedure, we say that it AGM-consistent.

We focus on two classes of updating rules. The first class of updating rules we study are lexicographic updating rules where the posteriors are defined by a lexicographic probability system (see ?, ? and ?). We show that the an updating rule is lexicographic if and only if it is Bayesian, AGM-consistent and satisfies a weak form of path-independence (order in which information arrives does not matter). The weakening is in the sense that order independence is satisfied only for certain pairs of events. It also turns out that this weakening is crucial. There exist no updating rules which are Bayesian and satisfy strong path-independence (order-independence on all consistent pairs of events).

The second class of updating rules, which we call minimum-distance, picks the posterior closest to the prior according to a metric defined on the space of probability measures. Not all Bayesian updating rules have a minimum distance representation. We show how in some cases a metric may be constructed to give us the Bayesian posterior as the unique minimiser of distance. In a special case of the AGM procedure, provide necessary and sufficient condition for a minimum distance representation. Lastly, we study the relationship between lexicographic and minimum distance updating rules. A sub-class of lexicographic rules, which we call support-dependent updating, admit minimum distance representations. In these rules, the alternative hypotheses that are used to update over zero probability, depend only on the support of the prior. Hence, they depend only on the set of states which have ex-ante zero probability.

For an exposition of AGM theory, ?) and ?) provide excellent surveys. The problem of choosing supports of posteriors in a manner consistent with AGM belief revision has been studied by ?) where a choice correspondence given an event, chooses as a subset of the event, the support of the posterior. The framework is non-probabilistic and it is shown that rationalisability of the choice correspondence is equivalent to AGM-consistency. In the present framework, we derive a counterpart of this result and it provides a very useful characterisation of AGM-consistency of an updating rule. The relationship between

---

1Similar choice functions have also been studied by ?.
lexicographic probability systems and AGM consistency has also been discussed in ?) in terms of the belief operator for revising belief sets, again, in a non-probabilistic framework. We derive and build on that observation in our framework and also establish a complete characterisation of lexicographic updating rules.

In the theory of decision under uncertainty, in addition to ?), who consider lexicographic probabilities, there has been some attention on alternative approaches to dealing with zero probability events. ?) provides axiomatic foundations for conditional probability systems. ? studies an alternative approach where once a zero probability occurs, the agent uses a belief over a beliefs and the maximum likelihood rule to obtain posteriors. In the present work, such updating rules may violate AGM consistency, which is central to the analysis considered here. ? study a framework where the set of states itself can expand due to growing awareness and they consider the phenomenon of reverse Bayesianism. This requires that whenever the state space grows and the support of the prior belief is contained in that of the posterior, the prior can be obtained by applying Bayes rules to the posterior. The issue of updating ambiguous beliefs as defined in ? and ? has been discussed in ?. They consider the problem of updating convex non-additive probabilities and establish the equivalence of the dempster-shafer rule for conditioning and the maximum likelihood update. Though in the present work we do not discuss ambiguous beliefs, our approach may be used to define updating rules for it. This extension could potentially provide us with a connection between ambiguity aversion and agents’ attitudes to zero probability events.

The outline of this paper is as follows: In section 2, we introduce the framework and provide a brief summary of AGM belief revision. In section 3, we study lexicographic updating rules and minimum distance updating rules and also, how they are related. Section 4 is the conclusion. Some proofs are in the appendix.

2 Framework

2.1 Updating Rules

In this section, we describe the environment and introduce the choice problem faced by the agent. We denote as $\Omega$, a non-empty finite set, to be the underlying state space on which the agent has beliefs. An agent’s probabilistic belief about the state space is defined by a probability measure $\pi \in \Delta(\Omega)$. We shall sometimes call $\pi$, the agent’s prior
to emphasise that \( \pi \) is his belief prior to arrival of information. Similarly, we shall refer to a belief \( \pi' \) as the agent’s posterior when it is formed after information arrives. Throughout the paper, we shall denote as \( \text{supp}(\pi) := \{\omega \in \Omega : \pi(\omega) > 0\} \), the support of a probability measure \( \pi \). Upon learning that an event \( A \subseteq \Omega \) has taken place, the agent updates his prior belief \( \pi \) over the state space \( \Omega \) to a posterior belief \( \pi' \in \Delta(\Omega) \). The agent performs this task in two stages. Upon observing \( A \), he lists a menu of possible posteriors \( \Pi(\pi, A) \subseteq \Delta(\Omega) \) to choose from. From the menu, he selects a single posterior \( \pi' \in \Pi(\pi, A) \). This two-stage choice completely describes his belief updating procedure and we define it formally below:

**Definition 2.1.** An updating rule is a tuple \( < c, \Pi > \) where \( c : \Delta \Omega \times 2^{\Omega \setminus \{\emptyset\}} \to \Delta \Omega \) and \( \Pi : \Delta \Omega \times 2^{\Omega} \to 2^{\Delta \Omega} \) such that:

- For all \( \pi', \pi'' \in \Pi(\pi, A) : \text{supp}(\pi') = \text{supp}(\pi'') \subseteq A \).
- \( c(\pi, A) \in \Pi(\pi, A) \).

The first condition says that the menu of posteriors selected by the agent has a common support that is contained completely in the event learned \( A \). The common support assumption stems from the rationale that the set of states the agent cannot disregard must be independent of the choice of the posterior and depends solely on \( (\pi, A) \), otherwise additional information must have influenced him to do so. Inclusion of the support in \( A \) captures the idea that the event \( A \) has been observed to have taken place and the agent disregards as improbable, states outside it. This latter condition may be viewed as updating satisfying consequentialism. In what follows, we shall sometimes write \( \text{supp}\Pi(\pi, A) \) to denote the common support of the probability measures in \( \Pi(\pi, A) \).

### 2.2 Bayesian Updating

We now discuss updating rules defined by Bayesian updating. Such updating rules apply Bayes’ rule when positive probability events are observed. The support of the posterior then equals the intersection of the support of the prior and the event observed.

**Definition 2.2.** An updating rule \( < c, \Pi > \) is said to be Bayesian if:

- Whenever \( \pi(A) > 0 \), \( \Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = \text{supp}(\pi) \cap A\} \).
- Moreover, when \( \pi(A) > 0 \), \( c(\pi, A)(\omega) = \frac{\pi(\{\omega\} \cap A)}{\pi(A)} \).
Notice in the above definition, no restriction has been placed on how the updating rule behaves on zero probability events. Additionally, in this case, we make no restrictions on the common support of the menu of posteriors chosen by the individual. This makes the class of Bayesian updating rules very large and potentially allows for varied attitudes in updating when agents are completely surprised. Restrictions on \( c(\pi, A) \) and \( \Pi(\pi, A) \) when \( \pi(A) = 0 \) can potentially have very strong implications and is indeed the main subject of the remainder of the paper. We impose consistency requirements on updating rules to restrict behaviour on zero probability events. One would hope for such additional requirements to not be dictated by arbitrary exogenous rules but arise from the primitives \((\pi, A)\) in a manner that is plausible. In doing so, we impose that the updating rule be consistent with a belief revision procedure proposed by ? (henceforth AGM), that arises in propositional logic in the context of updating sets of propositions ("belief sets" in the parlance of the belief revision literature). We introduce a notion called AGM-consistency for updating rules based on this requirement. The next section is devoted to its formulation and interpretation in the present environment.

2.3 AGM Consistency

In this section, we introduce a notion of consistency of updating rules based on AGM belief revision. It shall be useful to interpret subsets of the state space \( \Omega \) as an algebra of propositions with an event \( E \) being treated as a proposition. The set of all propositions is thus \( 2^\Omega \). In this context, the notion of logical consequence may be defined as follows:

**Definition 2.3.** Let \( G \subseteq 2^\Omega \) and let \( A \subseteq \Omega \). We say \( G \vdash A \) (meaning \( A \) is a logical consequence of \( G \)) if \( \bigcap_{G \in G} G \subseteq A \) and define the consequence operator as \( Cn(G) = \{ A : G \vdash A \} \). We say a set of events \( G \) is logically closed if \( Cn(G) = G \) and consistent if \( \bigcap_{G \in G} G \neq \emptyset \).

Let us discuss the notion of logical consequence defined above. A set \( A \) is said to be a logical consequence of \( G \) if it impossible that all events in \( G \) occur but \( A \) does not occur i.e there does not exist any state \( \omega \) such that \( \omega \in \bigcap_{G \in G} G \) but \( \omega \notin A \). This is a definition which is natural and analogous to the definition of logical consequence in propositional logic if

\[ \text{Consider, for example, tossing a coin one hundred times. The state space would be the set } \{H, T\}^{100}. \text{ Now, the propositions "the 24th coin toss led to Heads" and "the 26th coin toss led to tails" can both be expressed as events } \{ \omega | \omega_{24} = H \} \text{ and } \{ \omega | \omega_{26} = T \}. \text{ The conjunction of the two propositions would correspond to the intersection of the two events.} \]
we were to interpret the set of all truth valuations as the state space $\Omega$. In the remainder of the paper, a logically closed and consistent set of events will often be referred to as a theory or belief set and will be denoted as $\mathcal{K}$. Corresponding to each $\pi \in \Delta(\Omega)$, we can define a belief set consisting of all the probability one events.

**Definition 2.4.** Let $\Omega$ be finite and let $\pi \in \Delta\Omega$. The belief set corresponding to $\pi$ is defined as the set $\mathcal{K}(\pi) = \{ E \subseteq \Omega : \pi(E) = 1 \}$

It follows that $\mathcal{K}(\pi)$ is logically closed and consistent. Clearly, $\bigcap_{E \in \mathcal{K}(\pi)} = \text{supp}(\pi)$. The set $\mathcal{K}(\pi)$ is logically closed since all probability one events contain the support $\text{supp}(\pi)$. Consistency follows from the fact that the support is non-empty.

### 2.3.1 AGM Postulates

The theory of belief revision (see ? and ? for introductory surveys) is devoted to studying the revision of belief sets $\mathcal{K}$ upon learning a new event $A$ by means of revision operator $*$. Formally, letting $T$ denote the set of all logically closed and consistent belief sets, a revision operator is a function $*: T \times 2^\Omega \setminus \{\emptyset\} \rightarrow T$. Hence, revision of the belief set $\mathcal{K}$ leads to the new belief set $\mathcal{K}*A$. AGM study revision operators $*$ that satisfy the following postulates:

1. $\mathcal{K}*A = Cn(\mathcal{K}*A)$.
2. $A \in \mathcal{K}*A$.
3. $\mathcal{K}*A \subseteq Cn(\mathcal{K} \cup \{A\})$.
4. If $A^c \notin \mathcal{K}$, then $Cn(\mathcal{K} \cup \{A\}) \subseteq \mathcal{K}*A$.
5. If $A$ is consistent then so is $\mathcal{K}*A$.
6. If $Cn(A) = Cn(B)$, then $\mathcal{K}*A = \mathcal{K}*B$.
7. $\mathcal{K}*(A \cap B) \subseteq Cn(\mathcal{K}*A \cup \{B\})$.
8. If $B^c \notin \mathcal{K}*A$, then $Cn(\mathcal{K}*A \cup \{B\}) \subseteq \mathcal{K}*(A \cap B)$.

---

3See, for example, Chapter 2 of ?. For a set of atomic propositions, $\mathbb{P}$, the state space would be $\{T,F\}$. Each proposition $\alpha$ derived from $\mathbb{P}$ through elementary operations of conjunction, disjunction and negation, has an event corresponding to it. Namely, it is the event $E(\alpha) = \{\omega|\omega(p) = T\}$. The converse is true as well. For any event $E$, there exists a proposition $\alpha$, such that $E = E(\alpha)$. 

The above axioms on the revision procedure have come to be known in the literature as the *AGM postulates*. In their seminal paper, AGM (?), establish that adherence to the postulates is equivalent to following a two stage procedure for belief revision involving a *contraction* stage and *expansion* stage. The contraction step deletes events from $\mathcal{K}$ by selecting from maximal subsets (according to set inclusion) of $\mathcal{K}$ consistent with $A$ via a selection function, and treats the intersection of the selected maximal subsets as the set of retained events. The expansion stage set theoretically adds $A$ to the retained set of events from the contraction stage and treats the logical closure of the resultant set as the revised belief set $\mathcal{K} \ast A$.

### 2.3.2 Incorporating AGM as a consistency requirement

We now define what it means for an updating rule to be AGM-consistent.

**Definition 2.5.** An updating rule $< c, \Pi >$ is said to be *AGM-consistent* if there exist revision operators $\{ \ast_{\pi} \}_{\pi \in \Delta \Omega}$ (one for each prior $\pi$) satisfying the AGM postulates, such that for all $(\pi, A)$

$$\Pi(\pi, A) = \{ \pi' : \mathcal{K}(\pi') = \mathcal{K}(\pi) \ast_{\pi} A \}. \quad (1)$$

The requirement that the choice of posterior must be such that its support is consistent with AGM belief revision i.e the posterior $\pi'$ should be such that the belief set corresponding to it equals the belief set obtained by revising the belief set corresponding to the prior $\pi$ based on an event $A$. This can be depicted by means of the following diagram

![Diagram](image.png)

Figure: AGM-Consistency
AGM consistency requires the use of revision operators, which as mathematical objects, could be potentially be very difficult to handle. The following results provide a simple characterisation of AGM-consistent updating rules:

**Proposition 1.** An updating rule \(<c, \Pi>\) is AGM-consistent if and only if there exists a family \(\{\pi \geq \pi\}_{\pi \in \Delta\Omega}\) of complete and transitive relations on \(\Omega\) such that

\[
\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = M_{\pi}(A)\},
\]

where \(M_{\pi}(A) = \{\omega \in A : \omega \geq_{\pi} \omega' \text{ for all } \omega' \in A\}\) is the set of \(\geq_{\pi}\)-greatest elements in \(A\); for all \(\pi\), we have \(M_{\pi}(\Omega) = \text{supp}(\pi)\).

## 3 Lexicographic Updating Rules

In this section, we study updating rules based on updating using lexicographic probability systems. The main result is that an updating rule is lexicographic if and only if it is Bayesian, AGM consistency and satisfies a weak form of path independence i.e the order in which information arrives does not influence the final posterior. Path independence is a key property in the characterisation. We first show that when a stronger version is imposed, there exist no Bayesian updating rules which satisfy path independence (Claim 1). Propositions 2 and 3 correspond to the main result of this section.

### 3.1 Path Independence

Note that for non-Bayesian updating rules it is permitted that \(\Pi(\pi, A)\) have any desirable property of an updating rule to have is path independence. This means that the order in which information is received does not affect the final posterior. We define this formally.

**Definition 3.1.** An updating rule \(<c, \Pi>\) satisfies strong path independence if for all \(A, B \subseteq \Omega\) such that \(A \cap B \neq \emptyset\) and for all \(\pi \in \Delta\Omega\) it is true that

\[
c(c(\pi, A), B) = c(c(\pi, B), A).
\]

We now prove the following result.

**Claim 1.** If \(|\Omega| \geq 3\) then there exists no Bayesian updating rule which satisfies strong path independence.
Proof. Let \( < c, \Pi > \) be a Bayesian updating rule. Consider three distinct states \( \{\omega_1, \omega_2, \omega_3\} \subseteq \Omega \) and consider the prior \( \delta_{\omega_1} \). Now define the sets \( A = \{\omega_1, \omega_2\}, B = \{\omega_1, \omega_3\}, C = \{\omega_2, \omega_3\} \).

Since \( < c, \Pi > \) is Bayesian it is true that \( c(\delta_{\omega_1},A) = c(\delta_{\omega_1},B) = \delta_{\omega_1} \). Now define \( \hat{\pi} = c(\delta_{\omega_2},C) = c(\delta_{\omega_2},A), C) = c(\delta_{\omega_2},B), C) \). Now by definition it is the case that \( \text{supp}(\hat{\pi}) \subseteq C = \{\omega_2, \omega_3\} \). So there are now three cases

\begin{itemize}
  \item Case 1 : \( \text{supp}(\hat{\pi}) = C = \{\omega_2, \omega_3\} \). Now since \( < c, \Pi > \) is Bayesian, \( c(\delta_{\omega_1},C), B) = \delta_{\omega_3} \).

  But, we have \( c(\delta_{\omega_1},B), C) = c(\delta_{\omega_1},C) = \hat{\pi} \neq \delta_{\omega_3} \).
  \item Case 2 : \( \text{supp}(\hat{\pi}) = \{\omega_2\} \). This implies that \( \hat{\pi} = \delta_{\omega_2} \). Now \( c(\delta_{\omega_1},C), B) = c(\delta_{\omega_2},B) \).

  Note that \( \text{supp}(c(\delta_{\omega_2},B)) \subseteq B \). So clearly \( c(\delta_{\omega_2},B) \neq \delta_{\omega_2} \). But we have \( c(\delta_{\omega_1},B), C) = c(\delta_{\omega_1},C) = \delta_{\omega_2} \).
  \item Case 3 : \( \text{supp}(\hat{\pi}) = \{\omega_3\} \). This implies that \( \hat{\pi} = \delta_{\omega_3} \). Now \( c(\delta_{\omega_1},C), A) = c(\delta_{\omega_3},A) \).

  Note that \( \text{supp}(c(\delta_{\omega_3},A)) \subseteq A \). So clearly \( c(\delta_{\omega_3},A) \neq \delta_{\omega_3} \). But we have \( c(\delta_{\omega_1},A), C) = c(\delta_{\omega_1},C) = \delta_{\omega_3} \).
\end{itemize}

\( \square \)

The above result may be viewed as unsettling. Note that when restricted to positive probability events, Bayesian updating satisfies path independence. However, when we expand to the domain of updating to allow for zero probability events, path independence cannot be satisfied. Note also that with two states of the world, the result does not apply. Moreover, with two states, one can verify that every Bayesian updating rule satisfies strong path independence.

### 3.2 Lexicographic Updating Rules : Characterisation

In ?, updating via lexicographic probability systems (LPS) is defined and axiomatic foundations are studied for a model of decision making under uncertainty where agents rank acts according to lexicographic expected utility. LPS’s provide a new class of updating rules where the posterior \( c(\pi, A) \) is provided by the Bayesian update of the first probability measure in a hierarchy \( < p_1, ..., p_K > \) with \( p_1 = \pi \), where \( A \) assumes a positive probability. We now investigate the relationship between LPS’s and AGM-consistent updating rules. First, we define formally what it means for an updating rule to be lexicographic.

**Definition 3.2.** A lexicographic probability system (LPS) is a finite sequence \( p = < p_1, ..., p_K > \) of probability measures on \( \Omega \) such that for any non-empty event \( A \subseteq \Omega \), there exists an \( 1 \leq i \leq K \) such that \( p_i(A) > 0 \). A lexicographic conditional probability system (LCPS) is an LPS where the supports of the probability measures in \( < p_1, ..., p_K > \) are pairwise disjoint.
LPS’s can be used to define posteriors for all non-empty events in the following way: For $A \subseteq \Omega$ non-empty define the posterior to be $p(\cdot|A) := p_k(\cdot|A)$ where $k^* = \min\{k : p_k(A) > 0\}$.

Claim 2. For every LPS $p$ there exists an LCPS $q$ such that for all non-empty $A \subseteq \Omega$, $p(\cdot|A) = q(\cdot|A)$.

Proof. Let $p =< p_1, ..., p_K >$ be an LPS. Define $q$ as follows:

- Set $q_1 = p_1$.
- If $\Omega \setminus \bigcup_{i=1}^k \text{supp}(q_i) \neq \emptyset$, then define $q_{k+1} = p_l(|\text{supp}(p_l) \setminus \bigcup_{i=1}^k \text{supp}(q_i))$ where $l = \min\{j : \text{supp}(p_j) \setminus \bigcup_{i=1}^k \text{supp}(q_i) \neq \emptyset\}$. If $\Omega \setminus \text{supp}(q_k) = \emptyset$, then stop.

Since $\Omega$ this process terminates at some step $K$. Define $q =< q_1, ..., q_K >$. It can now be shown that indeed $p(\cdot|A) = q(\cdot|A)$ for all non-empty $A \subseteq \Omega$. □

It can also be shown that if two LCPS’s $q$ and $q'$ induce the same conditional probabilities, then $q = q'$. The implication of this is that for any LPS $p$ there exists a unique LCPS $q$ which induces the conditional probabilities as $p$.

Definition 3.3. An updating rule $< c, \Pi >$ is said to be lexicographic if there exists a family $\{p^\pi\}_{\pi \in \Delta \Omega}$ of LPS’s such that:

- $p_1^\pi = \pi$ for all $\pi \in \Delta \Omega$.
- $\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = \text{supp}(p^\pi(\cdot|A))\}$.
- $c(\pi, A) = p^\pi(\cdot|A)$.

Let us now make some useful observations. Firstly, note that any family of a LPS’s $\{p^\pi\}_{\pi \in \Delta \Omega}$ induces an updating rule defined by the conditions above. Secondly, if for two families $\{p^\pi\}_{\pi \in \Delta \Omega}, \{q^\pi\}_{\pi \in \Delta \Omega}$ we have that $p^\pi(\cdot|A) = q^\pi(\cdot|A)$ for all non-empty $A \subseteq \Omega$ and for all $\pi \in \Delta$, then both families induce the same updating rule. Having made these observations, we are now ready to prove our first result.

Proposition 2. Every lexicographic updating rule $< c, \Pi >$ is Bayesian and AGM-consistent.

Proof. Let $< c, \Pi >$ be lexicographic with the corresponding family $\{p^\pi\}_{\pi \in \Delta \Omega}$. From the above claim we can construct a family of LCPS’s $\{q^\pi\}_{\pi \in \Delta \Omega}$ such that $p^\pi(\cdot|A) = q^\pi(\cdot|A)$ for all non-empty $A \subseteq \Omega$ and for all $\pi \in \Delta$. Hence $\{q^\pi\}_{\pi \in \Delta \Omega}$ the same updating rule.

Let $\pi$ and $A \subseteq \Omega$ be such that $\pi(A) > 0$. Now $\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = \text{supp}(\pi) \cap A\}$ and $c(\pi, A) = p_1^\pi(\cdot|A) = q_1^\pi(\cdot|A) = \pi(\cdot|A)$. So $< c, \Pi >$ is Bayesian. We now show that it is AGM-consistent. For $\pi \in \Delta \Omega$, define the relation $\Omega$.
\[\omega \geq_{\pi} \omega' \iff \omega \in \text{supp}(q^n) \text{ and } \omega' \in \text{supp}(q^n) \text{ for } k \leq l.\]

Since the collection \(\{\text{supp}(q^n)\}_k\) constitutes a partition of \(\Omega\), the order \(\geq_{\pi}\) is complete and transitive. Denote as \(M^k_{\geq_{\pi}}(\Omega)\), the set of all \(k\)-th highest states according to \(\geq_{\pi}\). Clearly, \(M^k_{\geq_{\pi}}(\Omega) = \text{supp}(q^n_{k})\) for all \(k\). Let \(k(\pi, A) = \min\{k : q^n(A) > 0\}\). From this observation we get that \(\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = \text{supp}(q^n\mid A)\} = \{\pi' : \text{supp}(\pi') = M_{\geq_{\pi}}^{k(\pi, A)}(\Omega) \cap A = M_{\geq_{\pi}}(A)\}\). Hence, from Proposition 1, it follows that \(< c, \Pi >\) is AGM-consistent. \(\square\)

We have established that every lexicographic updating rule is Bayesian and AGM-consistent. However, there exist Bayesian and AGM-consistent updating rules which are not lexicographic. One natural question to address at this point would be : under what additional conditions are Bayesian AGM-consistent updating rules lexicographic? It turns out that the only additional condition we need is a weak form of path-independence. We first define path-independence and then derive our result.

**Definition 3.4.** An updating rule \(< c, \Pi >\) satisfies weak path independence if for all \(\pi \in \Delta\Omega\) and for all \(A, B \subseteq \Omega\) such that \(\text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) \neq \emptyset\) it is true that

\[c(c(\pi, A), B) = c(c(\pi, B), A).\]

Some remarks about weak path independence are in order. If an updating rule is path independent and AGM consistent, then it is true that for all \(\pi \in \Delta\Omega\) and for all \(A, B \subseteq \Omega\) such that \(\text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) \neq \emptyset, \text{supp}(c(c(\pi, A), B)) = \text{supp}(c(c(\pi, B), A))\). This is because if the support of the posteriors under \(A\) and \(B\) intersect, then \(A\) and \(B\) both must be intersecting the same equivalence class of \(\geq_{\pi}\). This observation also implies that under AGM-consistency, for any pair of events \((A, B)\), we have \(\text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) \neq \emptyset\) if and only if \(\text{supp}\Pi(\pi, A) \cap B \neq \emptyset\) and \(A \cap \text{supp}\Pi(\pi, B) \neq \emptyset\). It can hence be shown that lexicographic updating rules satisfy weak path-independence. We now prove the main result.

**Proposition 3.** Let \(< c, \Pi >\) be a updating rule which is Bayesian, AGM-consistent and satisfies weak path independence. Then \(< c, \Pi >\) is lexicographic.

**Proof.** We wish to define a family \(\{q^n\}_{\pi \in \Delta\Omega}\) of LPS’s. We do this as follows. Let \(\pi \in \Omega\). Since \(< c, \Pi >\) there exists a complete transitive relation \(\geq_{\pi}\) on \(\Omega\) which determines \(\Pi(\pi, .)\). Now define \(q^n\) in the following way :

Set \(q^n_k = c(\pi, M^k_{\geq_{\pi}}(\Omega))\) for all \(k\).

Define \(q^n = \langle q^n_k \rangle_k\). Note that the support of \(q^n_{k}\) is \(M^k_{\geq_{\pi}}(\Omega)\). Now let \(A \subseteq \Omega\) be non-empty. Let \(k(\pi, A) = \min\{k : q^n(A) > 0\}\). So now we get that \(M_{\geq_{\pi}}(A) = M^{k(\pi, A)}_{\geq_{\pi}}(\Omega) \cap A = \)
\[ \text{supp}(q^\pi_{k(\pi,A)}) \cap A = \text{supp}(q^\pi(\cdot|A)). \] Hence, we get \( \Pi(\pi,A) = \{ \pi' : \text{supp}(\pi') = M_{\geq \pi}(A) \} = \{ \pi' : \text{supp}(\pi') = \text{supp}(q^\pi(\cdot|A)) \}. \] Also:

\[
q^\pi(\cdot|A) = q^\pi_{k(\pi,A)}(\cdot|A) \\
= c(q^\pi_{k(\pi,A)},A) \\
= c(c(\pi,M^k_{\geq \pi}(\Omega)),A) \\
= c(c(\pi,A),M^k_{\geq \pi}(\Omega)) \\
= c(\pi,A)
\]

Where 3 and 5 follow from the fact that \( < c, \Pi > \) is Bayesian and 4 follows from weak path independence. Hence, we have shown that \( < c, \Pi > \) is lexicographic.

Combining the results above we get that an updating rule is Bayesian, AGM-consistent and satisfies path-independence if and only if it is lexicographic. We now show that the properties “AGM-consistency”, “bayesian” and “weak path independence” are independent from each other. For convenience, let us call them AGM,BA and WPI respectively.

- **AGM, BA but not WPI**: Let \( \Omega = \{x,y,z,w\} \) and let \( < c, \Pi > \) be any Bayesian updating rule such that \( \Pi(\pi,A) = A \) whenever \( \pi(A) = 0 \), \( c(\delta_x,\{y,z\}) = 1/2(y) + 1/2(z) = (0,1/2,1/2,0) \) and \( c(\delta_x,\{y,z,w\}) = 1/2(y) + (1/4)(z) + 1/4(w) \). One can verify that this is AGM-consistent. For each \( \pi \), define \( \geq \pi \) as the complete and transitive with two equivalence classes, namely \( \text{supp}(\pi) \) and its complement, where the former is ranked higher than the latter. Now let \( \pi = \delta_x, A = \{y,z\}, B = \{y,z,w\} \). Then, \( c(c(\pi,A),B) = (0,1/2,1/2,0) \) but \( c(c(\pi,B),A) = (0,2/3,1/3,0) \).

- **AGM, WPI but not BA**: Let \( \Omega = \{x,y,z\} \) and let \( < c, \Pi > \) be a full meet updating rule which always picks the posterior which has uniform distribution on event observed \( A \). Since we are always updating to a uniform distribution, weak path independence is trivially satisfied. AGM consistency is also satisfied. Note however, that such an updating rule cannot be Bayesian since for \( \pi = (1/2,1/3,1/6) \), and \( A = \{x,y\} \), we get \( c(\pi,A) = (1/2,1/2,0) \) which is not the Bayesian posterior.

- **WPI,BA but not AGM** Let \( < c, \Pi > \) be any Bayesian updating rule such that \( \Pi(\pi,A) \) is a singleton whenever \( \pi(A) = 0 \). Note in this case, we must necessarily have \( c(\pi,A) = \delta_\omega \) where \( \{\omega\} = \text{supp}\Pi(\pi,A) \). Such updating rules may or not satisfy AGM-consistency. It is AGM-consistent if and only if there exist relations \( \{\geq \pi\} \) such that \( \geq \pi \) ranks all states outside \( \text{supp}(\pi) \) strictly and \( M_{\geq \pi}(A) = \text{supp}\Pi(\pi,A) \) whenever \( \pi(A) = 0 \). If the family of relations doesn’t exist, then AGM-consistency is violated.
We can also show that it always satisfies weak path independence. Let \( \pi \in \Delta \Omega \) and \( A, B \subseteq \Omega \) such that \( \text{supp} \Pi(\pi, A) \cap \text{supp} \Pi(\pi, B) \neq \emptyset \).

- Case 1: W.l.o.g suppose \( \pi(A) > 0 \). Then \( \emptyset \neq \text{supp} \Pi(\pi, A) \cap \text{supp} \Pi(\pi, B) = \text{supp}(\pi) \cap A \cap \text{supp} \Pi(\pi, B) \subseteq \text{supp}(\pi) \cap A \cap B \subseteq A \cap B \). This implies that \( \pi(A \cap B) > 0 \). Since \( < c, \Pi > \) is Bayesian, \( c(c(\pi, A), B) = c(c(\pi, B), A) \).

- Case 2: \( \pi(A) = 0 \) and \( \pi(B) = 0 \). By assumption, \( \emptyset \neq \text{supp} \Pi(\pi, A) \cap \text{supp} \Pi(\pi, B) = \text{supp} \Pi(\pi, A) = \text{supp} \Pi(\pi, B) = \{ \omega \} \), for some \( \omega \in A \cap B \). Hence, \( c(\pi, A) = c(\pi, B) = \delta_\omega \). Since \( < c, \Pi > \) is Bayesian, \( c(c(\pi, A), B) = c(c(\pi, B), A) = \delta_\omega \).

### 3.3 Minimum Distance Updating Rules

We now study minimum distance updating rules. Such updating rules select posteriors in \( \Pi(\pi, A) \) which are “closest” to the prior.

**Definition 3.5.** An updating rule \( c \) is said to be minimum distance if there exists a metric \( d \) on \( \Delta \Omega \) such that \( \{ c(\pi, A) \} = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi') \).

Not all updating rules have a minimum distance representation. However, certain restrictions on the AGM procedure allow Bayesian updating rules to be minimum distance. We focus on two extreme kinds of AGM revision in this section: maxichoice and full-meet updating. A maxichoice (full-meet) updating rule is an AGM-consistent updating rule such that whenever \( \pi(A) = 0 \), the posterior \( c(\pi, A) \) has support that is a singleton (equal to \( A \)). These restrictions corresponding to the nature of the underlying AGM revision procedure. Recall, the two stage representation discussed in Section 2.3. Maxichoice (full-meet) eliminates as few (as many) propositions from \( K(\pi) \) to accommodate the new information \( A \). The conditions are discussed in detail in \( ? \) and \( ? \).

**Proposition 4.** Every Bayesian maxichoice updating rule is a minimum distance updating rule for some metric \( d \).

**Proof.** Let \( < c, \Pi > \) be a Bayesian maxichoice updating rule. We define the metric \( d \) on \( \Delta \Omega \). For any pair of states \( \omega, \omega' \in \Omega \), define the function \( S_{\omega\omega'} : \Omega \times \Omega \to \mathbb{R}_+ \) as follows:

- If \( \pi \neq \pi' \) and \( \min \{ \pi(\{ \omega, \omega' \}), \pi'(\{ \omega, \omega' \}) \} > 0 \):
  \[
  S_{\omega\omega'}(\pi, \pi') = \left| \frac{\pi(\omega)}{\pi'(\omega)} \right| - \left| \frac{\pi'(\omega)}{\pi(\omega)} \right|.
  \]

- If \( \pi \neq \pi' \) and \( \min \{ \pi(\{ \omega, \omega' \}), \pi'(\{ \omega, \omega' \}) \} = 0 \):
\[ S_{\omega \omega'}(\pi, \pi') = c \text{ where } c \geq 1. \]

- If \( \pi = \pi' \) then : 
\[ S_{\omega \omega'}(\pi, \pi') = 0. \]

Notice that for any pair \( \omega, \omega' \in \Omega, S_{\omega \omega'}(\pi, \pi') = S_{\omega' \omega}(\pi, \pi') \). Also, it can be shown that \( S_{\omega \omega'} \) defines a pseudometric on \( \Delta \Omega \) and that for any \( \pi \neq \pi' \), there exist states \( \omega, \omega' \in \Omega \) such that \( S_{\omega \omega'}(\pi, \pi') > 0 \) (proofs of these facts can be found in the appendix in Claims 3 and 4). Combining these two facts we get that the following is a metric on \( \Delta \Omega \)
\[ d(\pi, \pi') = \sum_{\omega \neq \omega'} S_{\omega \omega'}(\pi, \pi'). \]

The above metric is indeed exactly the one we need. So we shall now show that for any \( A \subseteq \Omega \) such that \( \pi(A) > 0 \), we have \( c(\pi, A) = |\pi_B(\pi, A)| = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi') \). Now notice if \( \operatorname{supp}(\pi) \cap A = \{\omega\} \), then \( \Pi(\pi, A) = \{\delta_{\omega}\} \), so clearly \( |\pi_B(\pi, A)| = |\delta_{\omega}| = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi') \).

So we only consider the case when \( |\operatorname{supp}(\pi) \cap A| \geq 2 \).

- **Case 1 :** \( \pi \in \Pi(\pi, A) \) :

Notice in this case, \( \operatorname{supp}(\pi) = \operatorname{supp}(\pi) \cap A \), hence \( \pi(A) = 1 \). So we have \( \pi = \pi_B(\pi, A) \). Now clearly, \( d(\pi, \pi_B(\pi, A)) = 0 \). Now consider \( \pi' \in \Pi(\pi, A) \backslash \{\pi_B(\pi, A)\} \). Since \( \pi \neq \pi_B(\pi, A) = \pi \), there exists a pair of states \( \omega_1, \omega_2 \in \Omega \) such that \( S_{\omega_1 \omega_2}(\pi, \pi') > 0 \). Hence \( d(\pi, \pi') > 0 \) and we get \( |\pi_B(\pi, A)| = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi') \).

- **Case 2 :** \( \pi \notin \Pi(\pi, A) \) :

Now partition pairs of states in the following way :

- \( S^o = \{ (\omega, \omega') : \omega \in \operatorname{supp}(\pi) \cap A \text{ but } \omega' \notin \operatorname{supp}(\pi) \cap A \} \).
- \( S^- = \{ (\omega, \omega') : \omega \in \operatorname{supp}(\pi) \cap A \text{ and } \omega' \notin \operatorname{supp}(\pi) \cap A \} \).
- \( S^+ = \{ (\omega, \omega') : \omega \notin \operatorname{supp}(\pi) \cap A \text{ and } \omega' \notin \operatorname{supp}(\pi) \cap A \} \).

1. Consider \( (\omega, \omega') \in S^o \). Since \( \omega \in \operatorname{supp}(\pi) \cap A \), \( \pi([\omega, \omega']) > 0 \) and \( \pi'([\omega, \omega']) > 0 \) for all \( \pi' \in \Pi(\pi, A) \). Hence, \( \min \{\pi[\omega, \omega'], \pi'([\omega, \omega'])\} > 0 \) and \( \pi \neq \pi' \) for all \( \pi' \in \Pi(\pi, A) \) (since \( \pi \notin \Pi(\pi, A) \)). Hence :
\[ S_{\omega \omega'}(\pi, \pi') = \left| \frac{\pi([\omega, \omega'])}{\pi_S[\omega, \omega']} - \frac{\pi'([\omega, \omega'])}{\pi_S[\omega, \omega']} \right| = \left| \frac{\pi([\omega, \omega'])}{\pi_S[\omega, \omega']} - 1 \right| \text{ for all } \pi' \in \Pi(\pi, A). \]

So \( S_{\omega \omega'}(\pi, \pi') = S_{\omega' \omega}(\pi, \pi') \) for all \( \pi \pi'' \in \Pi(\pi, A) \). As a result we get
\[ \sum_{(\omega, \omega') \in S^o} S_{\omega \omega'}(\pi, \pi_B(\pi, A)) = \sum_{(\omega, \omega') \in S^o} S_{\omega' \omega}(\pi, \pi'). \]
for all $\pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}$.

2. Consider $(\omega, \omega') \in S^-$. Then clearly $\pi'(\{\omega, \omega'\}) = 0$ and $\pi \neq \pi'$ for all $\pi' \in \Pi(\pi, A)$. Hence, $S_{\omega, \omega'}(\pi, \pi') = c$ for all $\pi' \in \Pi(\pi, A)$. As a result we get

$$\sum_{(\omega, \omega') \in S^-} S_{\omega, \omega'}(\pi, \pi_B(\pi, A)) = \sum_{(\omega, \omega') \in S^-} S_{\omega, \omega'}(\pi, \pi') \text{ for all } \pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}.$$ 

3. Consider $(\omega, \omega') \in S^+$. Then clearly $\pi(\{\omega, \omega'\}) > 0$ and $\pi'(\{\omega, \omega'\}) > 0$ for all $\pi' \in \Pi(\pi, A)$. Hence, $\min\{\pi(\omega, \omega'), \pi'(\omega, \omega')\} > 0$ and $\pi \neq \pi'$ for all $\pi' \in \Pi(\pi, A)$.

Notice that $S_{\omega, \omega'}(\pi, \pi_B(\pi, A)) = 0$ since $\frac{\pi_B(\pi, A)(\omega)}{\pi_B(\pi, A)(\omega, \omega')} = \frac{\pi(\omega)}{\pi(\omega, \omega')} = \frac{\pi(\omega)}{\pi(\omega, \omega')}$. So we get $S_{\omega, \omega'}(\pi, \pi_B(\pi, A)) = 0$ for all $(\omega, \omega') \in S^+$.

Now consider $\pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}$. Now there exist states $\omega, \omega' \in E \ast (\pi) \cap A$ such that $S_{\omega, \omega'}(\pi, \pi') > 0$.

As a result we get $\sum_{(\omega, \omega') \in S^+} S_{\omega, \omega'}(\pi, \pi_B(\pi, A)) < \sum_{(\omega, \omega') \in S^+} S_{\omega, \omega'}(\pi, \pi') \text{ for all } \pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}$. Hence, we obtain our desired result that $\{\pi_B(\pi, A)\} = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi')$. \hfill \qed

The following result shows that some full-meet updating rules may not have a minimum distance representation.

**Proposition 5.** Let $|\Omega| \geq 4$. Then, there exist Bayesian full-meet updating rules which cannot be characterised as a minimum distance updating rule.

**Proof.** For any non-empty subset $F \subseteq \Omega$, denote as $\Delta F$ as the set of all probability measure in $\Delta \Omega$ whose support is $F$. Now, let $A, B \subseteq \Omega$ such that $A \cap B = \emptyset$ and $\min\{|A|, |B|\} \geq 2$. Note that both $\Delta A$ and $\Delta B$ have the same cardinality as $[0, 1]$. So there exist bijections : $a : \Delta A \to [0, 1]$ and $b : \Delta B \to [0, 1]$. Now consider any updating rule $< c, \Pi >$ such that

- $\Pi(\pi, B) = \Delta B$ for all $\pi \in \Delta A$.
- $\Pi(\pi', A) = \Delta A$ for all $\pi' \in \Delta B$.

and

- $c(\pi, B) = b^{-1}(a(\pi))$ for all $\pi \in \Delta A$. 

16
• \( c(\pi',A) = a^{-1}(1 - b(\pi')) \) for all \( \pi' \in \Delta B \).

Notice that there exists Bayesian AGM-consistent updating which satisfy the above restriction. Any Bayesian full meet updating rule with the above properties will do. Now, let \( \pi \) be such that \( a(\pi) = 0 \) and suppose for contradiction the above updating rules can characterised as a minimum distance updating rule with a metric say \( d \). Then, we get the following:

\[
\begin{align*}
    d(\pi, c(\pi, B)) &= d(\pi, b^{-1}(0)) \\
                       &< d(\pi, b^{-1}(1)) \\
                       &= d(c(b^{-1}(1), A), b^{-1}(1)) \\
                       &< d(a^{-1}(1), b^{-1}(1)) \\
                       &= d(a^{-1}(1), c(a^{-1}(1), B)) \\
                       &< d(a^{-1}(1), b^{-1}(0)) \\
                       &= d(c(b^{-1}(0), A), b^{-1}(0)) \\
                       &< d(a^{-1}(0), b^{-1}(0)) \\
                       &= d(\pi, c(\pi, B)).
\end{align*}
\]

This is a contradiction and we achieve our desired result. \( \square \)

The above counterexample shows that there exist Bayesian full meet updating rules which have no minimum distance representation. We now define a special class of full meet updating rules and provide necessary and sufficient conditions updating rules in this class to have a minimum distance representation.

**Definition 3.6.** A full meet updating rule \( < c, \Pi > \) is said to be bijective if for all \( A, B \subseteq \Omega \) with \( A \cap B = \emptyset \) and \( \min(|A|, |B|) \geq 2 \), the following functions are bijective.

- \( c(.,B) : \Delta A \rightarrow \Delta B \).
- \( c(.,A) : \Delta B \rightarrow \Delta A \).

**Definition 3.7.** A full meet updating rule satisfies the no cycle condition if for all \( A, B \subseteq \Omega \) with \( A \cap B = \emptyset \) and \( \min(|A|, |B|) \geq 2 \) and for all \( \pi \in \Delta A \) the following sequence does not contain any cycles of length \( \geq 2 \):

- \( \pi_0 = \pi \).
- \( \pi_{t+1} = c(c(\pi_t, B), A) \).


**Proposition 6.** Any Bayesian bijective full meet updating rule has a minimum distance representation if and only if it satisfies the no cycle condition.

*Proof.* The proof can be found in the appendix.

### 3.4 Relating Minimum distance and Lexicographic updating rules

This section is devoted to explaining the connection between lexicographic and minimum distance updating rules. We show that a sub-class of lexicographic updating rules, which we call "support dependent" updating rules necessarily have a minimum distance representation. Additionally, we demonstrate examples of lexicographic updating rules outside this class with no minimum distance representation.

**Definition 3.8.** A lexicographic updating rule \( <c, \Pi> \) is *support dependent* if the family of lexicographic updating rules \( \{p_{\pi}\}_{\pi} \) generating it satisfies:

For all \( \pi, \pi' \in \Delta(\Omega) \), \( \text{supp}(\pi) = \text{supp}(\pi') \) implies \( p_{\pi} \setminus \{\pi\} = p_{\pi'} \setminus \{\pi'\} \)

The above condition is interpreted as follows: If two decision makers have priors with the same support, then they use the same secondary hypothesis when they are completely surprised. The next result shows that support dependent lexicographic updating rules have a minimum distance representation.

**Proposition 7.** Every support dependent lexicographic updating rules has a minimum distance representation.

*Proof.* Let \( <c, \Pi> \) be a support dependent lexicographic updating rules and let \( \{p_{\pi}\}_{\pi} \) be the family generating it. Consider the metric \( d \) constructed in Proposition 4. Now we define a function \( \hat{d} : \Delta(\Omega) \times \Delta(\Omega) \rightarrow \mathbb{R}_+ \) as follows:

1. For \( \pi = \pi' \), set \( \hat{d}(\pi, \pi') = 0 \).
2. For \( \pi \neq \pi' \), let \( p_{\pi} = \{p_{\pi}^k\}_k \) and \( p_{\pi'} = \{p_{\pi'}^l\}_l \) define the natural number \( K(\pi, \pi') \) as:

\[
K(\pi, \pi') := \min\{l : \text{supp}(\pi_l) \cap \text{supp}(\pi') \neq \emptyset\}
\]

And now define:

\[
\hat{d}(\pi, \pi') = d(p_{K(\pi, \pi')}^{\pi}, \pi')
\]
Now define the function $\hat{S}(\pi, \pi') = \hat{d}(\pi, \pi') + \hat{d}(\pi', \pi)$. Note that $\hat{S}$ is symmetric and $0 < \sup_{\pi, \pi'} \hat{S}(\pi, \pi') < +\infty$. Now, let $M > 0$ be a real number such that $\sup_{\pi, \pi'} \hat{S}(\pi, \pi') < M$. We now define the desired metric $d_L$ as follows:

1. For $\pi = \pi'$, set $d_L(\pi, \pi') = 0$.
2. For $\pi \neq \pi'$, set $d_L(\pi, \pi') = \hat{S}(\pi, \pi') + M$

It can be checked that $d_L$ satisfies the axioms of a metric. We now show that this metric generates $< c, \Pi >$. Let $\pi \in \Delta(\Omega)$ and $A \in 2^\Omega \setminus \{\emptyset\}$ and as in the proof of proposition 5, define $k(\pi, A) = \min\{k : p^{\pi}_k(A) > 0\}$. If $\text{supp}(p^{\pi}_k) \cap A$ is a singleton, we are done. Assume, $|\text{supp}(p^{\pi}_k) \cap A| \geq 2$. Also, $\Pi(\pi, A) = \{\pi' : \text{supp}(\pi) = \text{supp}(p^{\pi}_k) \cap A\}$.

From support dependence, $\hat{d}(\pi', \pi) = \hat{d}(\pi'', \pi)$ for all $\pi, \pi' \in \Pi(\pi, A)$. Also from the definition of $d$, $\hat{d}(\pi, \pi') = d(p^{\pi}_k(\pi, A), \pi') > d(p^{\pi}_k(\pi, A), p^{\pi}_k(\pi', A), |A|) = \hat{d}(p^{\pi}_k(\pi, A), c(\pi, A))$ for all $\pi' \neq c(\pi, A)$.

4 Conclusion

In this paper, we studied the problem of updating over zero probability events. Our analysis is centered around the relationship between the problem of theory change in the belief revision literature in propositional logic and updating probabilistic beliefs about events based on evidence. In particular, we studied lexicographic and minimum-distance updating rules and their relationship with AGM consistent updating rules. As noted earlier in the introduction, there have been alternative approaches to this problem and one may also consider extensions of our analysis. One extension, as indicated in the introduced, would be to study implications of our approach to updating ambiguous beliefs. Another line of extension would be to investigate how our approach may apply to studying equilibrium refinements and issues of unawareness in games.

5 Appendix

5.1 Proof of Proposition 1

Proof. We first prove the “only if” direction. Suppose $< c, \Pi >$ is AGM-consistent with revision operators $\{*, \pi\}_{\pi \in \Delta(\Omega)}$. We shall derive a family of complete and transitive orderings
\{\pi \in \Delta(\Omega) \text{ such that :} \}

\{\pi' : \text{supp}(\pi') = M_{\geq\pi}(A)\} = \{\pi' : K(\pi') = K(\pi) \ast\pi A\}.

(6)

Now, consider some \(*\pi. Define the function \(b^\pi : 2^\Omega \setminus \{\emptyset\} \to 2^\Omega \setminus \{\emptyset\}\)

\[ b^\pi(A) = \bigcap_{E \in K(\pi) \ast\pi A} E. \]

From 2 and 5 of the AGM postulates it follows that \(b^\pi(A) \subseteq A\) and \(b^\pi(A) \neq \emptyset\). We shall now show that \(b^\pi\) satisfies the weak axiom of revealed preference (see Chapter 1 of \(\) ). This means that if \(\{\omega_1, \omega_2\} \subseteq A \cap B\), \(\omega_1 \in b^\pi(A)\) and \(\omega_2 \in b^\pi(B)\), then \(\omega_1 \in b^\pi(B)\). Now, suppose the hypothesis for the weak axiom is satisfied, then \(B^c \notin K(\pi) \ast\pi A\) and \(A^c \notin K(\pi) \ast\pi B\). Then, it follows from postulates 7 and 8 that \(\text{Cn}(K(\pi) \ast\pi A \cup \{B\}) = \text{Cn}(K(\pi) \ast\pi B \cup \{A\}) = K(A \cap B)\).

Now, since \(\omega_1 \in b^\pi(A) \cap B\), it follows that

\[ \omega_1 \in \bigcap_{E \in K(\pi) \ast\pi A \cup \{B\}} E = \bigcap_{E \in K(\pi) \ast\pi B \cup \{A\}} E. \]

Hence, \(\omega_1 \in b^\pi(B)\). This establishes that the correspondence \(b^\pi\) satisfies the weak axiom. This means that there exists a complete and transitive relation \(\geq\pi\) such that

\[ b^\pi(A) = M_{\geq\pi}(A). \]

(7)

One can now verify that the system of relations \(\{\geq\pi\}\) satisfies 1.

Now consider the "if" part. Suppose we have a system of relations \(\{\geq\pi\}\). For each \(\pi\), now define \(*\pi\) as follows

\[ K_{\ast\pi} A = \begin{cases} \{E : E \ni \bigcap_{E' \in K} E' \cap A\} & \text{if } \bigcap_{E' \in K} E' \cap A \neq \emptyset \\ \{E : E \ni M_{\geq\pi}(A)\} & \text{otherwise} \end{cases} \]

One can verify that \(*\pi\) satisfies the AGM postulates and furthermore, 6 is satisfied. \(\square\)

5.2 Define the metric

Claim 3. For any pair of states \(\omega, \omega' \in \Omega\), \(S_{\omega\omega'}\) is a psuedo-metric on \(\Delta\Omega\).

Proof. Let \(\omega, \omega' \in \Omega\) be a pair of states.

1. By definition it is true that \(S_{\omega\omega'}(\pi, \pi) = 0\) for all \(\pi \in \Delta\Omega\)
2. It also clear from the definition that $S_{\omega \omega^{'}}(\pi, \pi^{'}) = S_{\omega \omega^{'}}(\pi^{'}, \pi)$ for all $\pi, \pi^{' \in \Delta \Omega}$.

3. Let $\pi_1, \pi_2, \pi_3$ be arbitrary points in $\Delta \Omega$.

   - **Case 1 :** $\pi_1 = \pi_3$.
     Then clearly we have $S_{\omega \omega^{'}}(\pi_1, \pi_3) = 0 \leq S_{\omega \omega^{'}}(\pi_1, \pi_2) + S_{\omega \omega^{'}}(\pi_2, \pi_3)$

   - **Case 2 :** $\pi_1 \neq \pi_3$ and $\min\{\pi_1(\{\omega, \omega^\prime\}), \pi_3(\{\omega, \omega^\prime\})\} > 0$.
     - **Sub-case 1 :** $\pi_2 \neq \pi_i$ for all $i \in \{1, 3\}$ and $\min\{\pi_2(\{\omega, \omega^\prime\}), \pi_i(\{\omega, \omega^\prime\})\} = 0$ for at least one $i \in \{1, 3\}$. So we have:
       \[
       S_{\omega \omega^{'}}(\pi_1, \pi_3) \leq 1 \leq S_{\omega \omega^{'}}(\pi_1, \pi_2) + S_{\omega \omega^{'}}(\pi_2, \pi_3).
       \]
     - **Sub-case 2 :** $\pi_2 \neq \pi_i$ for all $i \in \{1, 3\}$ and $\min\{\pi_2(\{\omega, \omega^\prime\}), \pi_i(\{\omega, \omega^\prime\})\} > 0$ for all $i \in \{1, 3\}$. In this case we have:
       \[
       S_{\omega \omega^{'}}(\pi_1, \pi_3) = \left| \frac{\pi_1(w)}{\pi_1(\omega, \omega^\prime)} - \frac{\pi_3(w)}{\pi_3(\omega, \omega^\prime)} \right| \leq \left| \frac{\pi_1(w)}{\pi_1(\omega, \omega^\prime)} - \frac{\pi_2(w)}{\pi_2(\omega, \omega^\prime)} \right| + \left| \frac{\pi_2(w)}{\pi_2(\omega, \omega^\prime)} - \frac{\pi_3(w)}{\pi_3(\omega, \omega^\prime)} \right| = S_{\omega \omega^{'}}(\pi_1, \pi_2) + S_{\omega \omega^{'}}(\pi_2, \pi_3). \tag{8}
       \]

   - **Sub-case 3 :** $\pi_2 = \pi_i$ for some $i \in \{1, 3\}$.
     Then:
     \[
     S_{\omega \omega^{'}}(\pi_i, \pi_2) + S_{\omega \omega^{'}}(\pi_{-i}, \pi_2) \geq S_{\omega \omega^{'}}(\pi_{-i}, \pi_2) = S_{\omega \omega^{'}}(\pi_1, \pi_3).
     \]

   - **Case 3 :** $\pi_1 \neq \pi_3$ and $\min\{\pi_1(\{\omega, \omega^\prime\}), \pi_3(\{\omega, \omega^\prime\})\} = 0$
     - **Sub-case 1 :** $\pi_2 \neq \pi_i$ for all $i \in \{1, 3\}$. Then:
       \[
       S_{\omega \omega^{'}}(\pi_1, \pi_3) = c \leq S_{\omega \omega^{'}}(\pi_1, \pi_2) + S_{\omega \omega^{'}}(\pi_2, \pi_3).
       \]
     - **Sub-case 2 :** $\pi_2 = \pi_i$ for some $i \in \{1, 3\}$.
       This is exactly the same as sub-case 3 of case 2.

Claim 4. Let $\pi \neq \pi^\prime$, then there exist $\omega, \omega^\prime \in \Omega$ such that $S_{\omega \omega^{'}}(\pi, \pi^{'}) > 0$.

Proof. Since $\pi \neq \pi^\prime$, there exists $\omega \in \Omega$ such that $\pi(\omega) \neq \pi(\omega^\prime)$. W.l.o.g, assume $0 \leq \pi(\omega) < \pi(\omega^\prime)$. Since this is true there exists $\omega^\prime \in \Omega \setminus \{\omega\}$ such that $\pi(\omega^\prime) > \pi(\omega^\prime)$. Now there are two cases

   - **Case 1 :** $\pi(\omega) = 0$:
     Here clearly, $\frac{\pi(\omega^\prime)}{\pi(\{\omega, \omega^\prime\})} = 1 > \frac{\pi^\prime(\omega^\prime)}{\pi^\prime(\{\omega, \omega^\prime\})}$. Hence $S_{\omega \omega^{'}}(\pi, \pi^{'}) > 0$
• **Case 2**: $\pi(\omega) > 0$:

In this case, we have $\frac{\pi(\omega')}{\pi(\omega)} > 0$ which implies that $\frac{\pi(\omega')}{\pi(\omega,\omega')} > 0$. Hence $S_{\omega,\omega'}(\pi,\pi') > 0$.

5.3 **Proof of Proposition 6**

The following claim will be useful in establishing the result.

**Claim 5.** Let $X, Y$ be two sets and let $f : X \to Y$ and $g : Y \to X$ be two bijections such that for all $x \in X$ the sequence $\{(gf)^n(x)\}_{n \in \mathbb{N}}$ contains no cycles of length $\geq 2$. Then there exists a function $S : X \times Y \to \mathbb{R}$ such that for all $x$ and $y$:

- $f(x) = \arg \min_{y \in Y} S(x, y)$.
- $g(y) = \arg \min_{x \in X} S(x, y)$.

**Proof.** For each $x \in X$ consider the following two sided sequence $\{s_t(x)\}_{t \in \mathbb{Z}} = \{(x_t(x), y_t(x))\}_{t \in \mathbb{Z}}$ in $X \times Y$:

- $s_0 = (x_0(x), y_0(x)) = (x, f(x))$.
- For $t > 0$: Set $s_{t+1} = (g(y_t), y_t)$ if $t$ is even and set $s_{t+1} = (x_t, f(x_t))$ if $t$ is odd.
- For $t < 0$: Set $s_{t+1} = (x_t, g^{-1}(x_t))$ if $t$ is even and set $s_{t+1} = (f^{-1}(y_t), y_t)$ if $t$ is odd.

It can be shown that the generated sequence $\{s_t(x)\}_{t \in \mathbb{Z}}$ has no cycles of length $\geq 2$. Note that this implies that either $\{s_t(x)\}_{t \in \mathbb{Z}}$ is a constant sequence or has all distinct elements.

And by construction we also have for all $t$: $s_{t+1} = (g(y_t), y_t)$ if $t$ is even and set $s_{t+1} = (x_t, f(x_t))$ if $t$ is odd.

Let $s(x)$ be the set of elements of the sequence $\{s_t(x)\}_{t \in \mathbb{Z}}$. From the argument in the last paragraph, $s(x)$ is either a singleton or countably infinite. Also, it can be shown that the collection $\{s(x)\}_{x \in X}$ is partition of $\text{gr}(f) \cup \text{gr}(g)$.

Let $\tau : \mathbb{Z} \to (0, 1)$ such that $m > n \Rightarrow \tau(m) > \tau(n)$. Now for any $x$ for which $s(x)$ is infinite, define the function $\tau_x : s(x) \to (0, 1)$ as $\tau_x(s_t(x)) = \tau(t)$. Hence we have $t > r \Rightarrow$
\(\tau_x(s_l(x)) > \tau_x(s_r(x))\). Now define \(S : X \times Y \rightarrow \mathbb{R}\)

\[
S(x, y) = \begin{cases} 
1 & (x, y) \notin \bigcup_{x \in X} s(x) \\
\tau_x(x, y) & (x, y) \in s(x) \text{ and } s(x) \text{ is infinite} \\
0 & (x, y) \in s(x) \text{ and } s(x) \text{ is a singleton}
\end{cases}
\]

The desired result can be shown by the above function. \qed

We now prove Proposition 6.

Proof of Proposition 6. Let \(c\) be a Bayesian bijective full meet updating rule satisfying the no-cycle condition and let \(d\) be the metric defined in proposition 1. Now for all \(A, B \subseteq \Omega\) with \(A \cap B = \emptyset\) and \(\min(|A|, |B|) \geq 2\), the following functions are bijective.

- \(c(. , B) : \Delta A \rightarrow \Delta B\).
- \(c(. , A) : \Delta B \rightarrow \Delta A\).

From the no-cycle condition, the two functions above satisfy the hypothesis of the previous claim and hence for any such pair \((A, B)\), we can find a function \(S_{A,B} : \Delta A \times \Delta B \rightarrow [0, 1]\) such that :

- \(\{c(\pi, B)\} = \arg \min_{\pi' \in \Delta B} S_{A,B}(\pi, \pi')\).
- \(\{c(\pi, A)\} = \arg \min_{\pi' \in \Delta A} S_{A,B}(\pi', \pi)\).

Now, define the symmetric function \(\rho : \Delta \Omega \times \Delta \Omega \rightarrow \mathbb{R}_+\). Let \(\pi, \pi' \in \Delta \Omega\). Define \(\rho\) as follows :

- If \(\text{supp}(\pi) \cap \text{supp}(\pi') \neq \emptyset\), then set :
  \[\rho(\pi, \pi') = d(\pi, \pi').\]

- If \(\text{supp}(\pi) \cap \text{supp}(\pi') = \emptyset\), and \(\min(|\text{supp}(\pi)|, |\text{supp}(\pi')|) \geq 2\) then set :
  \[\rho(\pi, \pi') = \rho(\pi', \pi) = S_{\text{supp}(\pi), \text{supp}(\pi')} (\pi, \pi').\]

- If \(\text{supp}(\pi) \cap \text{supp}(\pi') = \emptyset\), and \(\min(|\text{supp}(\pi)|, |\text{supp}(\pi')|) = 1\), w.l.o.g let \(|\text{supp}(\pi)| = 1\), then set \(\rho(\pi, \pi') = \rho(\pi', \pi) = 1/2\) if \(\pi' = c(\pi, \text{supp}(\pi'))\) and set \(\rho(\pi, \pi') = \rho(\pi', \pi) = 1\) otherwise.
It is clear that $\rho$ is symmetric and $\rho(\pi, \pi') = 0$ if $\pi = \pi'$. Also, note that $\{c(\pi, A)\} = \arg \min_{\pi' \in \Delta A} \rho(\pi, \pi')$ as desired. All we need to do now is transform $\rho$ into a metric.

Note that $\rho$ is a bounded function. Now let $\bar{s} = \sup_{\pi \neq \pi'} \rho(\pi, \pi')$ and $\underline{s} = \inf_{\pi \neq \pi'} \rho(\pi, \pi')$. Clearly $\underline{s} < \bar{s}$ and there exists a $K > 0$ such that $\bar{s} + K < 2(\underline{s} + K)$. Now define the function $\hat{d} : \Delta \Omega \times \Delta \Omega \rightarrow \mathbb{R}^+$ as $\hat{d}(\pi, \pi') = \rho(\pi, \pi') + K$ if $\pi \neq \pi'$ and $\hat{d}(\pi, \pi') = 0$ if $\pi = \pi'$. It can be shown that $\hat{d}$ is a metric and satisfies $\{c(\pi, A)\} = \arg \min_{\pi' \in \Delta A} \hat{d}(\pi, \pi')$

Now let $c$ be a Bayesian bijective full meet updating rule which has a minimum distance representation. We shall show that it satisfies the no-cycle condition. Suppose not. Let $A, B \subseteq \Omega$ with $A \cap B = \emptyset$ and $\min\{|A|, |B|\} \geq 2$ and let $\{\pi_i\}$ be the sequence as defined in the definition of the no-cycle condition. Since $c$ has a minimum distance representation, let the corresponding metric be $d$. Suppose now that the sequence contained a cycle of length $\geq 2$. Then w.l.o.g there exists $n \geq 1$ such that $\pi_0 = \pi_{n+1}$ and $\{\pi_0, ..., \pi_{n+1}\}$ is the shortest cycle in the sequence. Since $c$ is minimum distance and bijective, for $s \in \{0, ..., n-1\}$ we must have

$$d(\pi_{s+1}, c(\pi_s, B)) > d(\pi_{s+1}, c(\pi_{s+1}, B))$$

$$> d(\pi_{s+2}, c(\pi_{s+1}, B)).$$

And we also have $d(\pi_0, c(\pi_0, B)) > d(\pi_0, c(\pi_1, B))$. Together, these would imply a chain of strict inequalities that would yield $d(\pi_0, c(\pi_0, B)) > d(\pi_{n+1}, c(\pi_n, B)) = d(\pi_0, c(\pi_n, B))$. But this is contradiction because this would imply that $c(\pi_n, B)$ should be selected at $(\pi_0, A)$. Hence $c$ cannot be minimum distance. $\square$