Bayesian updating rules and AGM belief revision*

Pathikrit Basu

Department of Economics
Pennsylvania State University

March 9, 2017

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 observing zero probability events and investigate two classes of updating rules for probabilities: 1) "minimum distance" updating rules which select the posterior closest to the prior by some metric. 2) "lexicographic" updating rules where posteriors are given by a lexicographic probability system. For the former, we show Bayesian updating as a special case and for specific AGM belief revisions, provide necessary and sufficient conditions for a minimum distance representation. For the latter, 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. 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

---

*I am deeply grateful to Kalyan Chatterjee, Edward Green, Neil Wallace, Stephen Simpson, Ariel Rubinstein, Arunava Sen, Debasis Mishra, Yuval Salant, Willemien Kets, Ryan Fang and Bruno Salcedo for providing useful comments and suggestions.
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 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 Kreps and Wilson(1982) [10] and Selten(1975)[16]) both place restrictions on admissible beliefs on information sets which lie off path (using Bayesian updating otherwise). The former achieves this by placing a consistency requirement on the belief system and the latter considers perturbations of the strategy profile. Abstracting away from the game-theoretic scenario, we ask whether in the probabilistic model itself, there exists a coherent way to extend Bayesian updating to zero probability events.

The preceding question constitutes the heart of this paper and any suggestions prescribed in answering it would entail implications for belief updating in models where an agent observes an unexpected event (extensive form games being an example of a relevant scenario). In this paper, we interpret the updating problem as a choice problem with a reference point (See Rubinstein and Zhou(1999)[13]). 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 logically 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 called AGM belief revision (See Alchourron et.al.(1985)[2]). In this theory, an agent’s primitive is a belief set i.e. a set of propositions or events that he believes to true (in the present context this would be all the full probability events of the prior). Upon learning an event or proposition, AGM offers a sound and well-defined method to update the belief set (the ex-post believed formulas or events). We use this
method to define the admissible set of posteriors and do so in the following way: 1) We associate to each prior its belief set i.e all the full probability events. 2) Upon observing an event, use AGM revision to update the belief 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. Hence, AGM allows us to clearly define the choice problem and more importantly generates in a logical manner, the support of the posterior when completely surprised.

Having set up the choice problem, we define an updating rule as choice function which selects a posterior and focus on two classes of updating rules. The first class (called "minimum distance") picks the posterior closest to the prior according to a metric defined on the space of probability measures. Such choice functions have been studied by Rubinstein and Zhou(1999)[13]. We show how a metric may be constructed to give us the Bayesian posterior as the unique minimiser of distance and in a special case of the AGM procedure, provide necessary and sufficient condition for a minimum distance representation. The second class of updating rules we study are lexicographic updating rules where the posteriors are selected using a lexicographic probability system (See Blume et.al(1991) [4], Halpern [8] and Hammond (1994) [9]) and provide a complete characterisation of such updating rules.

The problem of choosing supports of the posterior in a manner consistent with AGM belief revision has been studied by Bonanno(2009)[5] 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. The counterpart of this result is derived in the present context and serves as a useful characterisation of AGM-consistency of an updating rule. The relationship between lexicographic probability systems and AGM consistency has been discussed in Shoham et.al(2009)[17] in terms of the belief operator for revising belief sets. We derive and build on that observation in our framework and also establish a complete characterisation of lexicographic updating rules. The issue of updating ambiguous beliefs as defined in Schmeidler(1989)[15] and Gilboa and Schmeidler(1989)[6] has been discussed in Gilboa and Schmeidler(1993)[7]. 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 and AGM belief revision.

The outline of this paper is as follows: In section 2, we provide preliminaries from propositional logic and provide a brief summary of AGM belief revision. In section 3, using the framework of propositional logic, we bridge AGM belief revision with the problem of updating probabilities and choosing supports of the posterior. In section 4, we abstract away from the propositional logic framework and define AGM belief revision for more general state spaces (finite) and derive all counterparts of the results in section 3. In section 5, using the general framework, we define updating rules and define and investigate properties of two classes of updating rules: minimum distance and lexicographic. In the remaining sections, we discuss our results and directions for future work.

2 Framework

Let \( L \) be a propositional language. \( \mathcal{P} \) is a finite set of propositional atoms. Let \( \mathcal{F}(\mathcal{P}) \) be the set of all propositional formulae derived from the binary and unary propositional connectives. We shall use symbols \( p, q, r \ldots \) to denote atoms and greek alphabets \( \alpha, \beta, \gamma \ldots \) to denote propositional formulae. \( Cn : 2^\mathcal{F}(\mathcal{P}) \rightarrow 2^\mathcal{F}(\mathcal{P}) \) denote the classical consequence operator. For any \( K \subseteq \mathcal{F}(\mathcal{P}) \), this operator is defined as \( Cn(K) = \{ \alpha : K \vdash \alpha \} \).

2.1 AGM revision

In this section we discuss first, the idea behind the general method of AGM belief revision, present the AGM postulates and then examine special cases of the revision method.

Let \( K \) be a subset of formulae and \( \alpha \) be an arbitrary formula. AGM belief revision involves two stages. First, we contract our belief set (not necessarily to a logically closed set) by selecting subset \( \Gamma \subseteq K \) which is consistent with \( \alpha \) and denote it by \( K \vdash \neg \alpha \). Formally, this is done as follows: We select a non-empty subset \( \Gamma \subseteq K \) such that \( \Gamma \vdash \neg \alpha \) (if no such set exists then define \( K \vdash \neg \alpha = K \)) and define \( K \vdash \neg \alpha = \Gamma \). The final selection \( K \vdash \alpha \) is the result of the contraction stage. Second, we add the new information \( \alpha \) set theoretically to the \( K \vdash \neg \alpha \) and then take the logical closure, thus deriving our revised belief set. This stage is called the revision stage. So AGM revision of \( K \) given information \( \alpha \) is defined:
\[ K \ast \alpha = \text{Cn}((K \dashv \neg \alpha) \cup \{ \alpha \}) \]  

(1)

Note here that the contraction stage has not been clearly specified. In fact, there exist many different ways to make the selection of the information consistent subset \( \Gamma \). In order to decide which contractions are plausible and which are not, we rely on some postulates that the revised belief set \( K \ast \alpha \) should satisfy. These postulates are supplied by [2]. The revision operation (however so it may have been defined) \( K \ast \alpha \subseteq \mathcal{F} \mathcal{P} \) should satisfy the following properties:

1. \( K \ast \alpha = \text{Cn}(K \ast \alpha) \)
2. \( \alpha \in K \ast \alpha \)
3. \( K \ast \alpha \subseteq \text{Cn}(K \cup \{ \alpha \}) \)
4. If \( \neg \alpha \notin K \), then \( \text{Cn}(K \cup \{ \alpha \}) \subseteq K \ast \alpha \)
5. If \( \alpha \) is consistent then so is \( K \ast \alpha \)
6. If \( \text{Cn}(\alpha) = \text{Cn}(\beta) \), then \( K \ast \alpha = K \ast \beta \)
7. \( K \ast (\alpha \land \beta) \subseteq \text{Cn}((K \ast \alpha) \cup \{ \beta \}) \)
8. If \( \neg \beta \notin K \ast \alpha \), then \( \text{Cn}(K \ast \alpha \cup \{ \beta \}) \subseteq K \ast (\alpha \land \beta) \)

The above postulates on the revision procedure have come to be known in the literature as the AGM postulates. The first six are referred to as the “basic AGM postulates”. We now look at some contraction procedures \( \dashv \) such that the revision operation defined by (1) satisfies the AGM postulates and the basic AGM postulates.

### 2.2 Partial meet contraction

For a given belief set \( K \) and a formula \( \phi \), define the set \( K \perp \phi = \max \{ \Gamma \subseteq K : \Gamma \vdash \phi \} \) as the set of all maximal subsets of \( K \) which do not imply \( \phi \) (if there exist no such sets then simply define \( K \perp \phi = K \)).

**Definition** Given a belief set \( K \), a selection function on \( K \) is defined as a correspondence \( \gamma : \bigcup_{\phi \in \mathcal{F}(\mathcal{P})} \{ K \perp \phi \} \Rightarrow 2^\mathcal{F}(\mathcal{P}) \) such that \( \gamma(K \perp \phi) \subseteq K \perp \phi \)
Definition Given a belief set $K$ and a selection function $\gamma$ on $K$, a partial meet contraction operation $\dot{\cdot}$ is defined as

$$K\dot{\cdot}\phi = \bigcap_{\Gamma \in \gamma(K\bot\phi)} \Gamma$$

The idea behind partial meet contraction is as follows. Given a belief set $K$ and a formula $\alpha$, select by an external procedure, a sub-collection of $K\bot\phi$ and conclude the contracted belief set as the intersection of this subcollection. It has been established by [2] that a belief procedure $K\ast\alpha$ satisfies the basic AGM postulates 1-6 if and only if $K\ast\alpha$ can be represented as (1) for some partial meet contraction operation $\dot{\cdot}$.

We now look at some specific instances of partial meet contraction

Definition A partial meet contraction operation $\dot{\cdot}$ is said to be maxichoice if the underlying selection function defining $\gamma$ is a singleton.

Definition A partial meet contraction operation $\dot{\cdot}$ is said to be full meet if the underlying selection function defining $\gamma$ satisfies:

$$\gamma(K\bot\phi) = K\bot\phi$$

From the result cited above, it follows that both maxichoice and full meet contraction operations satisfy the basic AGM postulates 1-6. Also, in general it may be the case partial meet contraction might violate the last two postulates. In order to satisfy them, a restriction on the selection function $\gamma$ needs to be imposed. We discuss this restriction in the next section

2.3 Transitive Relationality

Given a belief set $K$ and a transitive binary relation $\geq_K$ on $2^K$, define the selection function

$$\gamma^{\geq_K}(K\bot\phi) = \{\Gamma \in K\bot\phi : \Gamma \geq_K \Gamma' \text{ for all } \Gamma' \in K\bot\phi\}$$

Definition: Given a belief set $K$, a selection function $\gamma$ on $K$ is said to be transitively relational if there exists a transitive binary relation $\geq_K$ on $2^K$ such that $\gamma = \gamma^{\geq_K}$.

Definition: A contraction operation $\dot{\cdot}$ is said to be transitively relational if it is a partial meet contraction under a transitively relational selection function. A revision operation
is said to transitively relational if it induced by a transitively relation contraction operation \(\dot{}\) via (1).

It turns out that the above additional restriction is exactly the one we need in order to satisfy the last two AGM postulates. In fact transitively relationality provides a complete characterisation of the eight postulates from the following result: The revision operation \(\ast\) satisfies all the eight AGM postulates if and only if it is defined by a partial meet contraction operation induced by a transitively relational selection function. Note that this implies that maxichoice satisfies the additional AGM postulates if and only if its selection function picks a maximal subset using an "underlying" asymmetric relation. Also, it can be shown that there exist maxichoice contractions which violate both the additional postulates. On the other hand full meet contraction always satisfies the additional postulates where the relation underlying the selection function simply treats all maximal subsets as equivalent.

It has been argued in [3] and will be demonstrated in the present context explicitly that maxichoice and full meet contraction may have some disconcerting properties. The former demands that we contract (or equivalently throw way propositions) as little as possible whereas as the latter demands that we contract as much as possible in order to accomodate the new information. In the case where the negation of the information \(\alpha\) is believed a priori, maxichoice results in the revised belief set being complete (even though the prior belief set was not) and full meet contraction results in the revised being simply the closure of \(\alpha\). As will be shown later, when applying AGM to the problem of updating probabilities, similar consequences will ensue. When updating over zero probability events, maxichoice will allow only delta measures as admissible posteriors and full meet contraction will allow all posteriors whose support is \(\alpha\).

Despite this, maxichoice provides a useful interpretation (this is discussed in [3]). This interpretation is in the context of the theory of conditionals and counterfactuals. As discussed in [18] and [11], one plausible way to believe in a conditional statement \(\phi \rightarrow \psi\) under currently held beliefs \(K\) would be to say that in the theory "closest" to \(K\) where \(\phi\) is believed, \(\psi\) is also believed. Maxichoie contraction and in general partial meet contraction (under transitive relationality) both provide a concrete procedure to achieve this. The underlying relation \(\geq_K\) may be interpreted as a "closer to \(K\) than" relation over theories and with the closest being the updated set \(K \ast \phi\) where \(\phi\) is believed. In the
present context, this interpretation will allow us to bridge the theory of conditionals and counterfactuals to the problem of updating probabilities via AGM revision.

3 Updating probabilities

In this section we define an updating rule for probability measures that is consistent with AGM belief revision (in a sense that we describe later). But first, we deal with some preliminaries and define notions that will assist us in setting up the updating rule. Since we wish to talk about probabilities and AGM belief revision in the same framework we must define an appropriate state space. A natural state space (also sometimes called the canonical state space) is \( \Omega = \{0, 1\}^P \) i.e the set of all truth assignments to propositional atoms (it will be shown later that AGM is well defined even for arbitrary state spaces, but this is perhaps a good starting point to convey the main idea). The interpretation of a state as a truth assignment is valid in the following sense: once a truth assignment is known, the truth or falsehood of a formula (event) is known i.e all uncertainty about the environment is resolved. For a given state (assignment) \( \omega \) denote as \( \tilde{\omega} \) the valuation extending \( \omega \) to all formulae. Note that there is a equivalence between formulae and events in this set-up. This is made clear by the following definition.

**Definition**: For a formula \( \phi \in \mathcal{F}(\mathcal{P}) \), define \( E(\phi) = \{ \omega : \tilde{\omega}(\phi) = 1 \} \). \( E(\phi) \) is interpreted as the event that \( \phi \) is true.

Conversely, since \( \mathcal{P} \) is finite, every event \( E \subseteq \Omega \) can be expressed by some formula \( \phi \) i.e. \( E(\phi) = E \). We now define the notion of a belief set elicited from a probability measure on \( \Omega \).

**Definition**: Let \( \pi \) be a probability measure on \( \Omega \). Define the belief set elicited from \( \pi \) as the set \( K(\pi) = \{ \phi \in \mathcal{F}(\mathcal{P}) : \pi(E(\phi)) = 1 \} \)

For \( E \subseteq \Omega \) define \( \Phi(E) = \{ \phi : E(\phi) = E \} \) as the set of all formulae which expresses the event \( E \). Now given a \( \pi \in \Delta \Omega \), define \( E^*(\pi) = \text{supp}(\pi) = \{ \omega : \pi(\omega) > 0 \} \) to be the support of \( \pi \) and define \( \mathcal{E}(\pi) = \{ E : \pi(E) = 1 \} \). It shall be convenient to know the following claim

**Claim 1** Let \( \pi \in \Delta \Omega \). The following statements are true

1. \( K(\pi) = \bigcup \Phi(E) \)
2. \((\forall E)(E^*(\pi) \subseteq E \iff E \in \mathcal{E}(\pi))\)

3. \((\exists \phi^* \in K(\pi))(\phi^* \vdash \phi \text{ for all } \phi \in K(\pi))\)

4. \(E^*(\pi) = \bigcap_{\phi \in K(\pi)} E(\phi)\)

**Proof**

1. Let \(\phi \in K(\pi)\). Clearly \(\phi \in \Phi(E(\phi))\) and \(E(\phi) \in \mathcal{E}(\pi)\). Hence \(\phi \in \bigcup_{E \in \mathcal{E}(\pi)} \Phi(E)\). On the other hand, if \(\phi \in \bigcup_{E \in \mathcal{E}(\pi)} \Phi(E)\), then \(\pi(E(\phi)) = 1\) and hence \(\phi \in K(\pi)\).

2. This is clearly true since \(E^*(\pi) \subseteq E \iff \pi(E) = 1 \iff E \in \mathcal{E}(\pi)\)

3. Since the language \(\mathcal{P}\) is assumed to be finite, \(E^*(\pi)\) can be expressed by a finite propositional formula. So let \(\phi^*\) be such that \(E(\phi^*) = E^*(\pi)\). Clearly \(\phi^* \in K(\pi)\) because \(\pi(E^*(\pi)) = 1\). Now, let \(\phi \in K(\pi)\) which from 2 implies that \(E(\phi^*) \subseteq E(\phi)\). Hence \(\phi^* \vdash \phi\)

4. We from the above parts that \(E^*(\pi) \subseteq E(\phi)\) for all \(\phi \in K(\pi)\). From the proof of 3, it is clear that \(\Phi(E^*(\pi)) \subseteq K(\pi)\). Hence we get \(E^*(\pi) = \bigcap_{\phi \in K(\pi)} E(\phi)\). □

### 3.1 Updating procedure and positive probability events

We are now ready to define the updating procedure. Given a prior \(\pi\), we elicit the underlying belief set \(K(\pi)\). Upon receiving information that \(\alpha\) is true (that is the event \(E^*(\alpha)\) has occurred) we update the belief set using AGM belief revision (with partial meet contraction using a selection function \(\gamma\)) and consider the set of admissible posteriors \(\Pi(\pi, \alpha) = \{\pi' : K(\pi') = K(\pi) \ast \alpha\}\). From \(\Pi(\pi, \alpha)\) we pick a posterior with the reference point being \(\pi\). This procedure is summarized in the following diagram:

\[
\begin{array}{c}
\pi \\
\xrightarrow{\alpha} \pi \ast \alpha \\
\downarrow \\
\{\pi' : K(\pi') = K(\pi) \ast \alpha\}
\end{array}
\]
We shall now investigate some details of this procedure and identify the set \( \Pi(\pi, \alpha) \) for the case when the information \( \alpha \) has prior positive probability i.e. \( \pi(E(\alpha)) > 0 \)

**Claim 2** Let \( \pi \in \Delta \Omega \) and let \( \alpha \in \mathcal{F}(\mathcal{P}) \) such that \( \pi(E(\alpha)) > 0 \). The following statements are true:

1. \( K(\pi)^\circ (-\alpha) = K(\pi) \)
2. \( \phi \in K(\pi)^\ast \alpha \iff E^*(\pi) \cap E(\alpha) \subseteq E(\phi) \)
3. \( \Pi(\pi, \alpha) = \{ \pi' : K(\pi') = K(\pi)^\ast \alpha \} = \{ \pi' : E^*(\pi') = E^*(\pi) \cap E(\alpha) \} \neq \emptyset \)
4. \( \bigcap_{\phi \in K(\pi)^\ast \alpha} E(\phi) = E^*(\pi) \cap E(\alpha) \)

**Proof**

1. Since \( \pi(E(\alpha)) > 0 \), we have \( E^*(\pi) \setminus E(-\alpha) \neq \emptyset \). Hence at state \( \omega \in E^*(\pi) \setminus E(-\alpha) \), we have that \( \omega(\phi) = 1 \) for all \( \phi \in K(\pi) \) (follows from part 4 of the previous claim) but \( \omega(-\alpha) = 0 \). So \( K(\pi) \nvdash (-\alpha) \). Hence \( K(\pi) = K(\pi)^\circ (-\alpha) \)

2. Note that a consequence of part 1 is that \( K(\pi)^\ast \alpha = Cn(K(\pi) \cup \{ \alpha \}) \). Now note that \( \phi \in K(\pi)^\ast \alpha \iff \phi \in Cn(K(\pi) \cup \{ \alpha \}) \iff \left( \bigcap_{\phi' \in K(\pi)} E(\phi')) \cap E(\alpha) \subseteq E(\phi) \iff E^*(\pi) \cap E(\alpha) \subseteq E(\phi) \right) \) where the last double implication follows from part 4 of the previous claim.

3. Let \( \hat{\pi} \in \{ \pi' : K(\pi') = K(\pi)^\ast \alpha \} \). It is clear that \( \hat{\pi}(E(\phi)) = 1 \iff \phi \in K(\pi)^\ast \alpha \). Now let \( \phi^* \) such that \( E(\phi^*) = E^*(\pi) \). Now, \( \{ \phi^*, \alpha \} \subseteq K(\pi)^\ast \alpha \). Hence \( \hat{\pi}(E^*(\pi) \cap E(\alpha)) = 1 \). Hence from part 2 of the previous claim, \( E^*(\pi) \cap E(\alpha) \supseteq E^*(\hat{\pi}) \). Now let \( \hat{\phi} \) such that \( E(\hat{\phi}) = E^*(\hat{\pi}) \). Since \( \hat{\pi}(E(\hat{\phi})) = 1 \), we get that \( \hat{\phi} \in K(\hat{\pi}) = K(\pi)^\ast \alpha = Cn(K(\pi) \cup \{ \alpha \}) \). Hence, from part 2 we get \( E^*(\pi) \cap E(\alpha) \subseteq E(\hat{\phi}) = E^*(\hat{\pi}) \).

Now, let \( \hat{\pi} \in \{ \pi' : E^*(\pi') = E^*(\pi) \cap E(\alpha) \} \). We want to show \( K(\hat{\pi}) = K(\pi)^\ast \alpha = Cn(K(\pi) \cup \{ \alpha \}) \). Let \( \phi \in Cn(K(\pi) \cup \{ \alpha \}) \). Hence, \( E^*(\pi) \cap E(\alpha) \subseteq E(\phi) \). But since \( \hat{\pi}(E^*(\pi) \cap E(\alpha)) = 1 \), we get that \( \hat{\pi}(E(\phi)) = 1 \). Hence, \( \phi \in K(\hat{\pi}) \). Now consider a formula \( \phi \in K(\hat{\pi}) \). We know that \( \hat{\pi}(E(\phi)) = 1 \) so as result we get \( E^*(\pi) \cap E(\alpha) = E^*(\pi') \subseteq E(\phi) \). This implies from part 2 that \( \phi \in K(\pi)^\ast \alpha \).

Now note that since \( \pi(E(\alpha)) > 0 \), the set \( E^*(\pi) \cap E(\alpha) \) is non-empty. As a result, \( \{ \pi' : E^*(\pi') = E^*(\pi) \cap E(\alpha) \} \neq \emptyset \).
4. This follows directly from the previous part. □

For information $\alpha$ with positive prior probability, the claim above gives us a useful characterisation of set $\Pi(\pi, \alpha)$ in terms of $E^*(\pi)$ and $E(\alpha)$ and also establishes that it is a non-empty set. What we learn from part 3 is that the set of admissible posteriors are exactly those probability measures which have support $E^*(\pi) \cap E(\alpha)$. Also note that the selection function $\gamma$ played no role in defining $\Pi(\pi, \alpha)$. However for zero probability events, it will indeed play an important role as we shall see next.

3.2 Zero probability events

The case of zero probability events (information $\alpha$) is more interesting. Since $\pi(E(\alpha)) = 0$, we observe that $E^*(\pi) \subseteq E(\neg \alpha)$. So clearly, $K(\pi) \vdash (\neg \alpha)$. However since $K(\pi)$ contains all tautologies, if $E(\alpha) \neq \emptyset$, then there exist states where $\neg \alpha$ will be false but of course any tautology will be true. Hence the set $\{\Gamma \subseteq K(\pi) : \Gamma \vdash (\neg \alpha)\}$ is non-empty so by Zorn’s Lemma a maximal exists. So, $K(\pi) \perp (\neg \alpha) = \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash (\neg \alpha)\} \neq \emptyset$ (the set of all maximals subsets of $K(\pi)$ which do not imply $\neg \alpha$) is non-empty. The maximal subsets $\Gamma$ may not be unique and the selection function $\gamma(K(\pi) \perp (\neg \alpha))$ picks out a subcollection of $K(\pi) \perp (\neg \alpha)$. The contraction step is concluded as $K(\pi) \neg (\neg \alpha) = \bigcap \gamma((\pi) \perp (\neg \alpha))$

Claim 3: Consider $\pi \in \Delta \Omega$ and $\alpha$ be such that $\pi(E(\alpha)) = 0$ and $E(\alpha) \neq \emptyset$. The following statements are true.

1. $\max\{\Gamma \subseteq K(\pi) : \Gamma \vdash (\neg \alpha)\} \neq \emptyset$ and contains more than one element if $E(\alpha)$ contains more than element.

2. $\max\{\Gamma \subseteq K(\pi) : \Gamma \vdash (\neg \alpha)\} = \bigcup_{\omega \in E(\alpha)} \{\phi \in K(\pi) : E^*(\pi) \cup \{\omega\} \subseteq E(\phi)\}.$

3. Define $\Gamma_\omega^\pi = \{\phi \in K(\pi) : E^*(\pi) \cup \{\omega\} \subseteq E(\phi)\}$ and $E_{\alpha}^\pi = \{\omega : \Gamma_\omega^\pi \in \gamma((\pi) \perp (\neg \alpha))\}$ Then, $\Pi(\pi, \alpha) = \{\pi' : E^*(\pi') = E_{\alpha}^\pi\}$

Proof:

1. Clearly non-emptiness follows from the argument presented in the above paragraph. Now suppose $E(\alpha)$ contains more than element. So, let $\omega_1 \neq \omega_2$ such that $\{\omega_1, \omega_2\} \subseteq E(\alpha)$. Consider $\Gamma^*_1 = \{\phi \in K(\pi) : E^*(\pi) \cup \{\omega_1\} \subseteq E(\phi)\}$. We shall show that $\Gamma^*_1 \in \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash (\neg \alpha)\}$. Now note that $\Gamma^*_1 \neq \emptyset$ because clearly any tautology is in $\Gamma^*_1$. Let $\phi^*$ such that $E(\phi^*) = E^*(\pi) \cup \{\omega_1\}$. Now notice $\phi^* \in K(\pi)$ and hence clearly in $\Gamma^*_1$, so we have $\cap_{\phi \in \Gamma^*_1} E(\phi) = E^*(\pi) \cup \{\omega_1\}$. So $\Gamma^*_1 \vdash (\neg \alpha)$. Now suppose
In light of the above claim, some important observations come to mind. Firstly, in the
situation the \( \Pi_\alpha \subseteq \Omega \). That the latter is a subset of the former has already been shown in the proof of the
K-case. We now show inclusion in the other direction. Let \( \Gamma \subseteq \bigcup_{\phi \in \Gamma} E(\phi) \). Now consider a formula \( \phi^* \) such that \( E(\phi^*) = E^*(\pi) \subseteq \Omega \). Clearly, \( \phi^* \notin \Gamma' \). Note that \( \phi^* \in K(\pi) \). Consider the set \( \Gamma_0 = \Gamma' \cup \{ \phi^* \} \). Clearly \( \bigcap_{\phi \in \Gamma_0} E(\phi) = E^*(\pi) \subseteq \Omega \). But \( \Gamma' \subseteq \Gamma' \) which contradicts the fact that \( \Gamma \subseteq \bigcup_{\phi \in \Gamma} E(\phi) \). Hence, \( (\bigcap_{\phi \in \Gamma} E(\phi)) \cap E(\pi) \) has to be a singleton and we get our desired result.

3. It is clear that \( K(\pi)^\perp(-\alpha) = \bigcap_{\gamma} \gamma((\pi) \perp (-\alpha)) = \{ \phi \in K(\pi) : E^*(\pi) \subseteq \Omega \} \). Now, let \( \pi' \in \Pi(\pi, \alpha) \). We know that \( K(\pi') = K(\pi) \times \alpha \). Hence, \( E^*(\pi') = \bigcap_{\phi \in K(\pi)^\perp(-\alpha)} E^*(\phi) \cap E^*(\alpha) = E^*(\pi) \cup E^*_{\alpha} \cap E^*(\alpha) = E^*_{\alpha} \). Now let \( \pi' \in \Pi(\pi, \alpha) \). Then, \( K(\pi') = \{ \phi : E^*_{\alpha} \subseteq E^*(\phi) \} \). Hence \( K(\pi') = K(\pi) \times \alpha \). Hence \( K(\pi') = K(\pi) \times \alpha \). □

In light of the above claim, some important observations come to mind. Firstly, in the special case of maxichoise contraction, the admissible set of posteriors is just a singleton comprising a delta measure \( \delta_\omega \) where \( \omega \in E(\pi) \). And, in the case of full meet contraction the \( \Pi(\pi, \alpha) \) is simply the set of all possible probability measure whose support is \( E(\pi) \). Secondly, the maximal subsets of \( K(\pi) \) that are consistent with \( \alpha \) are exactly sets of the form \( \Gamma^\pi_\omega \) where are \( \omega \in E^*(\pi) \). The selection function \( \gamma \) picks the collection \( \{ \Gamma^\pi_\omega \}_{\omega \in E^*_{\alpha}} \).
So essentially what the selection function selects is a subset \( E^\pi_a \subseteq E(\alpha) \) as the support of the posterior. In the special case of transitively relational selection functions, \( E^\pi_a \) are "highest" states amongst the states in \( E(\alpha) \) according to \( \geq_{K(\pi)} \). This implies that the choice correspondence \( \alpha \mapsto E^\pi_a \subseteq E(\alpha) \) is rationalisable. Moreover, this even extends to positive probability events i.e the correspondence \( \alpha \mapsto \text{supp}\Pi(\pi, \alpha) \) is rationalisable (here \( \text{supp}\Pi(\pi, \alpha) \) denotes the common support of all the posteriors in \( \Pi(\pi, \alpha) \)). Our next result establishes exactly the results made in this observation.

**Claim 4**: Let \( \pi \in \Delta\Omega \) and let \( \gamma \) be a transitively relational selection function according to the transitive relation \( \geq_{K(\pi)} \). Then the choice correspondence \( \alpha \mapsto \text{supp}\Pi(\pi, \alpha) \subseteq E(\alpha) \) is rationalisable i.e there exists a complete, transitive relation on \( \geq \) on \( \Omega \) such that \( \text{supp}\Pi(\pi, \alpha) = \{ \omega \in E(\alpha) : \omega \geq \omega' \text{ for all } \omega' \in E(\alpha) \} \)

**Proof**: For every \( \omega \in \Omega \), define \( \Gamma^\pi_\omega = \{ \phi \in K(\pi) : E^*(\pi) \cup \{ \omega \} \subseteq E(\phi) \} \). Clearly, for all \( \omega \in E^*(\pi) \), we have \( \Gamma^\pi_\omega = K(\pi) \). Now define the relation \( \geq \) as follows:

- \( \omega \geq \omega' \) if \( \Gamma^\pi_\omega \geq_{K(\pi)} \Gamma^\pi_\omega' \) and \( \{ \omega, \omega' \} \subseteq E^*(\pi) \)
- \( \omega \geq \omega' \) if \( \Gamma^\pi_\omega \geq_{K(\pi)} \Gamma^\pi_\omega' \) and \( \{ \omega, \omega' \} \subseteq E^*(\pi)^c \)
- \( \omega \geq \omega' \) if \( \omega \in E^*(\pi) \) and \( \omega' \notin E^*(\pi) \)

Note that \( \geq \) is transitive since \( \geq_{K(\pi)} \) is transitive. Note also that \( \omega > \omega' \) if \( \omega \in E^*(\pi) \) and \( \omega \notin E^*(\pi) \). We now show that \( \geq \) is complete and indeed rationalises the choice correspondence \( \alpha \mapsto \text{supp}\Pi(\pi, \alpha) \). Let \( \omega, \omega' \in \Omega \)

- **Case 1**: \( \{ \omega, \omega' \} \subseteq E^*(\pi) \).
  Then clearly, \( \Gamma^\pi_\omega = \Gamma^\pi_\omega' = K(\pi) \). Now let \( \psi \) be a formula which expresses the event \( \{ \omega, \omega' \} \). Then we have \( \gamma(K(\pi) \perp \psi) = K(\pi) \perp \psi = \{ K(\pi) \} \). Since \( \gamma \) is transitively relational under \( \geq_{(K(\pi))} \), we have that \( \Gamma^\pi_\omega \geq_{K(\pi)} \Gamma^\pi_\omega' \). Hence \( \omega \geq \omega' \). Note that this argument implies that \( \omega \geq \omega' \) for all \( \omega, \omega' \in E^*(\pi) \).

- **Case 2**: \( \omega \in E^*(\pi) \) and \( \omega \notin E^*(\pi) \)
  By construction we have that \( \omega > \omega' \)

- **Case 3**: \( \{ \omega, \omega' \} \subseteq E^*(\pi)^c \).
  Let \( \psi \) be a formula expressing the event \( \{ \omega, \omega' \} \). In this case \( K(\pi) \perp \psi = \{ \Gamma^\pi_\omega, \Gamma^\pi_\omega' \} \). Since \( \gamma \) selects at least one element from this set, we have either \( \Gamma^\pi_\omega \geq_{K(\pi)} \Gamma^\pi_\omega' \) or \( \Gamma^\pi_\omega' \geq_{K(\pi)} \Gamma^\pi_\omega \). Hence either \( \omega \geq \omega' \) or \( \omega' \geq \omega \).
Hence, we have shown that $\geq$ is complete and transitive. Now we show that it indeed rationalises the correspondence $\alpha \mapsto \text{supp}\Pi(\pi, \alpha)$. Let $\alpha \mapsto \text{supp}\Pi(\pi, \alpha)$.

- **Case 1**: $\pi(E(\alpha)) > 0$.

  Now, $\text{supp}\Pi(\pi, \alpha) = E^*(\pi) \cap E(\alpha)$. Since by construction $E^*(\pi)$ is the highest equivalence class under $\geq$, $\text{supp}\Pi(\pi, \alpha) = \{\omega \in E(\alpha) : \omega \geq \omega' \text{ for all } \omega' \in E(\alpha)\}$.

- **Case 2**: $\pi(E(\alpha)) = 0$ and $E(\alpha) \neq \emptyset$.

  Now $\text{supp}\Pi(\pi, \alpha) = E^*_\alpha$. We know that $\gamma(K(\pi) \perp \neg \alpha) = \{\Gamma^\pi_\omega \in K(\pi) \perp \neg \alpha : \Gamma^\pi_\omega \geq_{K(\pi)} \Gamma^\pi_\omega' \text{ for all } \omega' \in E(\alpha)\}$. Hence, by construction $\text{supp}\Pi(\pi, \alpha) = E^*_\alpha = \{\omega \in E(\alpha) : \omega \geq \omega' \text{ for all } \omega' \in E(\alpha)\}$.

\[\square\]

We now argue that the converse is true as well. We state this formally:

**Claim 5**: Let $\mathcal{F}^*$ be the set of all satisfiable formulae and let $f : \mathcal{F}^* \to 2^\Omega$ be a rationalisable choice function i.e $f(\alpha) \subseteq E(\alpha)$ is a non-empty selection. Then there exists a $\pi \in \Delta\Omega$ and a transitive $\geq_{K(\pi)}$ on $2^{K(\pi)}$ such that the revision operation induced by it satisfies $f(\alpha) = \text{supp}\Pi(\pi, \alpha)$ for all $\alpha \in \mathcal{F}^*$.

**Proof** : Since $f$ is rationalisable and since $\Omega$ is finite, there exists a function $u : \Omega \to \mathbb{R}$ which rationalises $f$. Now define $\pi$ to be the uniform measure with support $f(p \lor \neg p)$ and define a function $v : 2^{K(\pi)} \to \mathbb{R}$ as follows:

- If $L = \Gamma^\pi_\omega$ for some $\omega \in \Omega$, then define $v(L) = u(\omega)$
- If $L \neq \Gamma^\pi_\omega$ for all $\omega \in \Omega$, then define $v(L) = \min_{\omega \in \Omega} u(\omega) - 1$.

Note that the above construction is indeed well defined because $u(\omega) = u(\omega')$ for all $\omega, \omega' \in E^*(\pi)$. Now define $\geq_{K(\pi)}$ as follows

$$J \geq_{K(\pi)} L \iff v(J) \geq v(L)$$

It is now straightforward that the above define $\geq_{K(\pi)}$ indeed delivers the desired result. The proof is along the lines of the previous claim.

\[\square\]

The above results essentially demonstrate that any choice function which selects the support of the posterior is rationalisable if and only if it is "AGM-consistent" in some sense.
In a different framework, a version of this result appears in [5]. In [5] the notion of "AGM consistency" is defined in a different manner and also there are no probabilities in the model.

4 General Finite State Spaces

The AGM procedure described in section 2 can be extended to the case of a general finite state space (not necessarily \(\{0,1\}^\mathbb{P}\)). Here, instead of dealing with the boolean algebra of formulae of a propositional language, we shall deal with a set algebra comprising events in \(\Omega\) (note that from Stone's representation theorem, any boolean algebra is isomorphic to a set algebra and hence AGM belief revision shall be well-defined even for Boolean algebras). In order to talk about AGM belief revision for a general state space \(\Omega\) we need only to define notions of a belief set and logical consequence. We do so in the following two definitions:

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

**Definition** 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 is logically closed if \(Cn(G) = G\).

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. Note that in fact propositional logic will serve as a special case when \(\Omega = \{0,1\}^\mathbb{P}\). In the following subsections, we shall define AGM belief revision and show that all the results about positive and zero probability events from section 3 go through here as well.

4.1 AGM belief revision and the updating rule

In this subsection, we define AGM belief revision and the updating procedure for general state spaces. Given a logically closed belief set \(K \subseteq 2^\Omega\) and event \(A\), the contraction stage
is defined as follows: consider the set of maximal subsets of $K$ which do not imply $A^c$, $K \perp A^c = \max\{\Gamma \subseteq K : \Gamma \vdash A\}$. A selection function $\gamma$ selects a subcollection $\gamma(K \perp A^c) \subseteq K \perp A^c$ and we conclude the contraction step as $K^c = \bigcap \gamma(K \perp A^c)$. There may or may not exist a maximal subset. If there doesn’t, define $K^c = K$. So the revised belief set is defined as:

$$K \ast A = Cn((K^c) \cup \{A\})$$

The notion of a transitively relational selection function is defined in the same way as earlier and it can be shown that any revision operator $\ast$ satisfies the AGM postulates if and only if there exists a transitively relation selection function which induces it via partial meet contraction. We shall now show that the results of section 3.1 and 3.2 go through. Finally the updating procedure is summarised in the following diagram:

\[ \begin{array}{ccc}
\pi & \rightarrow & K(\pi) \\
\downarrow & & \downarrow \\
A & \rightarrow & K(\pi) \ast A \\
\downarrow & & \downarrow \\
\{\pi' : K(\pi') = K(\pi) \ast A\} & & \{\pi' : E(\pi') = E(\pi) \cap A\} \\
\end{array} \]

### 4.2 Positive and zero probability events

The following results, which are natural analogs of claims 4 and 5 in section 3 go through in the case of general state spaces. They are stated as follows:

**Claim 2′**: Let $\pi \in \Omega$. Let $A \subseteq \Omega$ such that $\pi(A) > 0$. Then the following statements are true:

1. $K(\pi) \vdash (A^c) = K(\pi)$
2. $E \in K(\pi) \ast A \iff E(\pi) \cap A \subseteq E$
3. $\Pi(\pi, A) = \{\pi' : K(\pi') = K(\pi) \ast A\} = \{\pi' : E(\pi') = E(\pi) \cap A\} \neq \emptyset$
4. $\bigcap_{E \in K(\pi) \ast A} E = E(\pi) \cap A$
Proof The proof can be found in the appendix.

Claim 3’ : Let \( \pi \in \Delta \Omega \) and let \( A \subseteq \Omega \) such that \( \pi(A) = 0 \) and \( A \neq \emptyset \). The following statements are true :

1. \( \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\} \neq \emptyset \) and contains more than one element if \( A \) contains more than one element

2. \( \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\} = \bigcup_{\omega \in A} \{E \in K(\pi) : E^*(\pi) \cup \{\omega\} \subseteq E\} \)

3. Define \( \Gamma^\pi_\omega = \{E \in K(\pi) : E^*(\pi) \cup \{\omega\} \subseteq E\} \) and \( E^\pi_A = \{\omega : \Gamma^\pi_\omega \in \gamma((\pi) \perp A^c)\} \) Then, \( \Pi(\pi, A) = \{\pi' : E^*(\pi') = E^\pi_A\} \)

Proof : The proof can be found in the appendix.

The counterparts of claim 4 and 5 also go through in the general case. They are stated as follows :

Claim 4’ : Let \( \pi \in \Delta \Omega \) and let \( \gamma \) be a transitivity relational selection function according to the transitive relation \( \geq_{K(\pi)} \). Then the choice correspondence \( A \mapsto \text{supp} \Pi(\pi, A) \subseteq A \) is rationalisable i.e there exists a complete, transitive relation on \( \geq \) on \( \Omega \) such that \( \text{supp} \Pi(\pi, A) = \{\omega \in A : \omega \geq \omega' \text{ for all } \omega' \in A\} \)

Proof : Proof is along the lines of claim 4

Claim 5’ : Let \( f : 2^{\Omega \setminus \{\emptyset\}} \rightarrow 2^{\Omega} \) be a rationalisable choice function i.e \( f(A) \subseteq A \) is a non-empty selection induced by a complete transitive relation on \( \Omega \). Then there exists a \( \pi \in \Delta \Omega \) and a transitive \( \geq_{K(\pi)} \) on \( 2^{K(\pi)} \) such that the revision operation induced by it satisfies \( f(A) = \text{supp} \Pi(\pi, A) \) for all \( A \neq \emptyset \).

Proof : Proof is along the lines of claim 5

5 Updating rules and Bayes rule

5.1 Updating Rule

We now turn to the question of updating via the updating rule defined by the procedure summarized in the diagram in section 3.1. We first define what an updating rule is in
Definition: An updating rule is a tuple \( < c, \Pi > \) where \( c : \Delta \Omega \times 2^{\Omega} \setminus \{ \emptyset \} \rightarrow \Delta \Omega \) and \( \Pi : \Delta \Omega \times 2^{\Omega} \rightarrow 2^{\Delta \Omega} \) such that:

- For all \( \pi', \pi'' \in \Pi(\pi, A) : E^*(\pi') = E^*(\pi'') \subseteq A \)
- \( c(\pi, A) \in \Pi(\pi, A) \)

Definition: An updating rule \( < c, \Pi > \) is said to be AGM-consistent if there exist transitively relational selection functions \( \gamma_\pi \) on \( 2^{K(\pi)} \) for each \( \pi \in \Delta \Omega \) such that the corresponding revision operations \( \{ \ast_{\gamma_\pi} \}_{\pi \in \Delta \Omega} \) induced by \( \{ \gamma_\pi \}_{\pi \in \Delta \Omega} \) satisfy:

\[
\Pi(\pi, A) = \{ \pi' : K(\pi') = K(\pi) \ast_{\gamma_\pi} A \}
\]

Note we can naturally define now maxichoice and full meet updating rules as AGM-consistent updating rules where the respective selection functions induce maxichoice and full meet contractions. For zero probability event \( \pi(A) = 0 \) : In the former case, \( \Pi(\pi, A) \) is a singleton consisting of a delta measure on some state in \( A \) and in the latter case, \( \Pi(\pi, A) \) is simply the set of all probability measures whose support is \( A \). We now define our desired class of updating rules.

Remark: In light of claims 4’ and 5’, an updating rule is AGM-consistent if and only if there exists a family \( \{ \geq_{\pi} \}_{\pi \in \Delta \Omega} \) of complete and transitive relations on \( \Omega \) such that:

- \( M_{\geq_{\pi}}(\Omega) = E^*(\pi) \)
- \( \Pi(\pi, A) = \{ \pi' : E^*(\pi') = M_{\geq_{\pi}}(A) \} \)

where \( M_{\geq_{\pi}}(A) = \{ \omega \in A : \omega \geq_{\pi} \omega' \text{ for all } \omega' \in A \} \). In the case of maxichoice rules, \( \geq_{\pi} \) is linear outside the support of \( \pi \) as opposed to full meet updating where \( \geq_{\pi} \) is trivial outside the support. This alternative characterisation of AGM-consistent updating rules is easier to deal with and hence will be used in the remainder of this paper.

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

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

- Whenever \( \pi(A) > 0 \), \( \Pi(\pi, A) = \{ \pi' : E^*(\pi') = E^*(\pi) \cap A \} \)
• Moreover, when \( \pi(A) > 0 \), \( c(\pi, A)(\omega) = \frac{\pi(\omega \cap A)}{\pi(A)} \)

Note that from claim any basic AGM-consistent updating rule satisfies \( \Pi(\pi, A) = \{ \pi' : E^*(\pi') = E^*(\pi) \cap A \} \) for all \( A \) such that \( \pi(A) > 0 \). We now investigate the relationship between maxichoice and full meet updating rules and minimum distance updating rules in the following section.

### 5.2 Minimum Distance Updating Rules

**Proposition 1**: Every Bayesian maxichoice updating rule is a minimum distance updating rule for some metric \( d \).

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

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

- 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). Combining these two facts we get that the following is a metric on \( \Delta \Omega \)

\[
\sum_{\omega \neq \omega'} S_{\omega \omega'}(\pi, \pi')
\]

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 \( E^*(\pi) \cap A = \{\omega\} \), 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 \( |E^*(\pi) \cap A| \geq 2 \).
• **Case 1**: $\pi \in \Pi(\pi, A)$:

Notice in this case, $E^*(\pi) = E^*(\pi) \cap A$, hence $\pi(A) = 1$. So we have $\pi = \pi_B(\pi, A)$. Now clearly, $d(\pi, \pi_B(\pi, A)) = 0$. Now consider $\pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}$. 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 $\{\pi_B(\pi, A)\} = \text{arg } \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi')$

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

Now partition pairs of states in the following way:

- $S^o = \{\{\omega, \omega'\} : \omega \in E^*(\pi) \cap A \text{ but } \omega' \notin E^*(\pi) \cap A\}$
- $S^- = \{\{\omega, \omega'\} : [\omega, \omega'] \subseteq [E^*(\pi) \cap A]^c \text{ and } \omega \neq \omega'\}$
- $S^+ = \{\{\omega, \omega'\} : [\omega, \omega'] \subseteq E^*(\pi) \cap A \text{ and } \omega \neq \omega'\}$

1. Consider $\{\omega, \omega'\} \in S^o$. Since $\omega \in E^*(\pi) \cap A$, $\pi(\{\omega, \omega'\}) > 0$ 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:

$$S_{\omega,\omega'}(\pi, \pi') = \frac{\pi(\omega)}{\pi(\omega', \omega')} - \frac{\pi'(\omega)}{\pi'(\omega', \omega')} = \frac{\pi(\omega)}{\pi(\omega', \omega')} - 1 \text{ for all } \pi' \in \Pi(\pi, A)$$

So $S_{\omega,\omega'}(\pi, \pi') = S_{\omega,\omega'}(\pi, \pi''')$ for all $\pi''\pi''\in \Pi(\pi, A)$. As a result we get that

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

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 that

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

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 $\omega, \omega' \in E^*(\pi) \setminus A$ such that $S_{\omega,\omega'}(\pi, \pi') > 0$.

As a result we get $\sum_{\{\omega, \omega'\} \in S^+} S_{\omega,\omega'}(\pi, \pi_B(\pi, A)) < \sum_{\{\omega, \omega'\} \in S^+} S_{\omega,\omega'}(\pi, \pi') \text{ for all } \pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}.$
Combining the results of 1 2 and 3 above, we get 
\[ d(\pi, \pi_B(\pi, A)) = \sum_{\omega \neq \omega'} S_{\omega \omega'}(\pi, \pi_B(\pi, A)) < \sum_{\omega \neq \omega'} S_{\omega \omega'}(\pi, \pi') = d(\pi, \pi') \] for all \( \pi' \in \Pi(\pi, A) \backslash \{\pi_B(\pi, A)\} \). Hence, we obtain our desired result that 
\[ \{\pi_B(\pi, A)\} = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi') \] □

**Proposition 2** : Let \(|\Omega| \geq 4\). Then, there exist Bayesian AGM-consistent 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 : 
\[ a : \Delta A \rightarrow [0, 1] \text{ and } b : \Delta B \rightarrow [0, 1] \]. 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 \( \pi \in \Delta A \)
- \( c(\pi', A) = a^{-1}(1 - b(\pi')) \) 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. □

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**: 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**: 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 3** 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.

### 5.3 Path Independence

In the above section we restricted the set of admissible posterior $\Pi(\pi,A)$ so that it is consistent with AGM given $\pi$ and $A$. We now relax these assumption and deal with updating rules more generally where the only restriction on $\Pi(\pi,A)$ will be that all posteriors in it have a common support in $A$. Notice that $\Pi(\pi,A)$ derived from AGM satisfies this condition and hence shall serve as a special case.

Note that for non-Bayesian updating rules it is permitted that $\Pi(\pi,A)$ have any arbitrary common support in $A$. Moreover, even within the class of Bayesian updating rules, for a zero probability $A$, the admissible set of posteriors $\Pi(\pi,A)$ is allowed to have an
arbitrary support in $A$. Hence, we see that for Bayesian updating rules, $\Pi(\pi, A)$ is the AGM induced set of admissible posteriors for positive probability events but may differ on zero probability events. In this sense, Bayesian updating rules have now been defined more generally than was done in the previous subsection.

One 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** An updating rule $< c, \Pi >$ satisfies *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 (rather disappointing) result:

**Claim 6**: If $|\Omega| \geq 3$ then there exists no Bayesian updating rule which satisfies 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 $\hat{\pi} = c(\delta_{\omega_1}, C) = c(c(\delta_{\omega_1}, A), C) = c(c(\delta_{\omega_1}, B), C)$. Now by definition it is the case that $E^*(\hat{\pi}) \subseteq C = \{\omega_2, \omega_3\}$. So there are now three cases

- **Case 1**: $E^*(\hat{\pi}) = C = \{\omega_2, \omega_3\}$. Now since $< c, \Pi >$ is Bayesian, $c(c(\delta_{\omega_1}, C), B) = \delta_{\omega_1}$. But, we have $c(c(\delta_{\omega_1}, B), C) = c(\delta_{\omega_1}, C) = \hat{\pi} \neq \delta_{\omega_3}$

- **Case 2**: $E^*(\hat{\pi}) = \{\omega_2\}$. This implies that $\hat{\pi} = \delta_{\omega_2}$. Now $c(\delta_{\omega_1}, C), B) = c(\delta_{\omega_2}, B)$. Note that $E^*(c(\delta_{\omega_2}, B)) \subseteq B$. So clearly $c(\delta_{\omega_2}, B) \neq \delta_{\omega_2}$. But we have $c(c(\delta_{\omega_1}, B), C) = c(\delta_{\omega_1}, C) = \delta_{\omega_2}$

- **Case 3**: $E^*(\hat{\pi}) = \{\omega_3\}$. This implies that $\hat{\pi} = \delta_{\omega_3}$. Now $c(\delta_{\omega_1}, C), A) = c(\delta_{\omega_3}, A)$. Note that $E^*(c(\delta_{\omega_3}, A)) \subseteq A$. So clearly $c(\delta_{\omega_3}, A) \neq \delta_{\omega_3}$. But we have $c(c(\delta_{\omega_1}, A), C) = c(\delta_{\omega_1}, C) = \delta_{\omega_3}$

□

The above result is very dissettling. One would have hoped to establish a rule for selecting maximal subsets in the AGM contraction step for zero probability events in such
a way that would yield path independence of the updating rule. This would have allowed us to have a natural restriction on the choice of maximal subsets. However, this is completely ruled out by the above result which in fact conveys something even stronger: there is no way to extend bayes rule to zero probability events in order to achieve path independence. Hence, a different criterion must be chosen to deal with non-determinacy in the contraction stage.

5.4 Lexicographic Updating Rules

In [4], updating via lexicographic probability systems (LPS) is defined and decision theoretic foundations for it have been supplied. In the present context, LPS provides 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**: 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 are pairwise disjoint.

LPS’s can be used to provide posteriors for all non-empty events in the following way: For $A \subseteq \Omega$ non-empty define the posterior to be $p(.)|A) := p_{k^*}(.|A) \text{ where } k^* = \min\{k : p_k(A) > 0\}$

**Claim 7**: For every LPS $p$ there exists an LCPS $q$ such that for all non-empty $A \subseteq \Omega$, $p(.|A) = q(.|A)$.

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

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

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

24
It can also be shown that if two LCPS's \( q \) and \( q' \) induce the same conditional probability, then \( q = q' \). The implication of this is that for any LPS \( p \) there exists a unique LCPS \( q \) which induces the conditional probability as \( p \).

**Definition**: 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 \( \pi \in \Delta \Omega \)
- \( \Pi(\pi, A) = \{ \pi': E^*(\pi') = E^*(p_\pi^\pi (|A)) \} \)
- \( c(\pi, A) = p_\pi^\pi (|A) \)

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 \( p_\pi^\pi (|A) = q_\pi^\pi (|A) \) 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 and in light of Claim 7, we shall without loss of generality, focus on lexicographic updating rules for which the corresponding family \( \{ p_\pi \}_{\pi \in \Delta \Omega} \) consists of LCPS's. We are now ready to prove our first result.

**Proposition 4**: 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 \( p_\pi^\pi (|A) = q_\pi^\pi (|A) \) 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 \( \Pi(\pi, A) = \{ \pi' : E^*(\pi') = E^*(\pi) \cap A \} \) and \( c(\pi, A) = p_\pi^\pi (|A) = q_1^\pi (|A) = \pi(|A) \). So \(< c, \Pi >\) is Bayesian. We now show that it is AGM-consistent. For \( \pi \in \Delta \Omega \), define the relation:

\[
\omega \geq_\pi \omega' \iff \omega \in E^*(q_k^\pi) \text{ and } \omega' \in E^*(q_l^\pi) \text{ for } k \leq l
\]

Since the collection \( \{ E^*(q_k^\pi) \}_k \) partitions \( \Omega \), the order \( \geq_\pi \) is complete and transitive. Denote as \( M_{\geq_\pi}^k (\Omega) \), the set of all k-th highest states according to \( \geq_\pi \). Clearly, \( M_{\geq_\pi}^k (\Omega) = E^*(q_k^\pi) \) for all \( k \). Let \( k(\pi, A) = \min\{ k : q_k^\pi (A) > 0 \} \). From this observation we get that \( \Pi(\pi, A) = \{ \pi' : E^*(\pi') = E^*(q^\pi (|A)) \} = \{ \pi' : E^*(\pi') = M_{\geq_\pi}^{k(\pi, A)} (\Omega) \cap A = M_{\geq_\pi} (A) \} \). Hence \(< c, \Pi >\) is AGM-consistent.
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** An updating rule \(<c, \Pi>\) satisfies *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 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 \(\text{supp} \Pi(\pi, A) \cap \text{supp} \Pi(\pi, B) \neq \emptyset\), \(E^*(c(\pi, A), B)) = E^*(c(\pi, B), A))\). 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 be shown that lexicographic updating rules satisfy path-independence. We now prove our desired result.

**Proposition 5**: Let \(<c, \Pi>\) be a updating rule which is Bayesian, AGM-consistent and satisfies path-independence. Then \(<c, \Pi>\) is lexicographic.

**Proof**: We wish to define a family \(\{q^\pi\}_{\pi \in \Delta \Omega}\) 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 \(q^\pi\) in the following way:

\[
q_k^{\pi} = c(\pi, M^{k}_{\geq \pi}(\Omega)) \text{ for all } k
\]

Define \(q^\pi =< q_k^{\pi} >_k\). Note that the support of \(q_k^{\pi}\) is \(M^{k}_{\geq \pi}(\Omega)\). Now let \(A \subseteq \Omega\) be non-empty. Let \(k(\pi, A) = \min[k : q_k^{\pi}(A) > 0]\). So now we get that \(M_{\geq \pi}(A) = M_{k^{(\pi, A)}}(\Omega) \cap A = E^*(q_k^{\pi}(\pi, A)) \cap A = E^*(q^{\pi}(\pi, A))\). Hence, we get \(\Pi(\pi, A) = \{\pi' : E^*(\pi') = M_{\geq \pi}(A)\} = \{\pi' : E^*(\pi') = E^*(q^{\pi}(\pi, A))\}\).
Also:

\[ q^\pi(\cdot | A) = q_{k(\pi, A)}^\pi(\cdot | A) = c(q_{k(\pi, A)}^\pi, A) = c(c(\pi, M_{\geq \pi}^k (\Omega)), A) = c(c(\pi, A), M_{\geq \pi}^k (\Omega))) = c(\pi, A) \]

Where 2 and 4 follow from the fact that \(<c, \Pi>\) is Bayesian and 3 follows from path-independence. Hence, we have shown that \(<c, \Pi>\) is lexicographic.

\(\square\)

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 "path independence" are independent of each other. For convenience, let us call them AGM, BA and PI respectively.

- **AGM, BA but not PI**: Let \(\Omega = \{x, y, z, w\}\) and let \(<c, \Pi>\) be any Bayesian full meet updating rule such that \(c(\delta_x, \{y, z\}) = 1/2(y) + 1/2(z) = (0, 1/2, 1/2, 0)\) and \(c(\delta_x, \{y, z, w\}) = 1/2(y) + (1/4)(z) + 1/4(w)\). 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, PI but not BA**: Let \(\Omega = \{x, y, z,\}\) and let \(<c, \Pi>\) be a full meet updating rule which always picks the posterior has a uniform distribution on the support. By the remark made earlier about AGM-consistency and path-independence, we know that for all \(\pi \in \Delta \Omega\) and for all \(A, B \subseteq \Omega\) such that \(\text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) \neq \emptyset\), \(E^*(c(c(\pi, A), B)) = E^*(c(c(\pi, B), A))\). And since we are also updating to a uniform distribution, path-independence is 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.

- **PI, BA but not AGM**: Let \(<c, \Pi>\) be any Bayesian maxichoice updating rule. We can show that it always satisfies 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 \(\pi(A) > 0\) : Then \(\emptyset \neq \text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) = E^*(\pi) \cap A \cap B\)
suppΠ(π,B) ⊆ E∗(π) ∩ A ∩ B ⊆ A ∩ B. This implies that π(A ∩ B) > 0. Since < c,Π > is Bayesian, c(c(π,A),B) = c(c(π,B),A).

- Case 2 : π(A) = 0 and π(B) = 0. Then, since < c,Π > is maxichoice, ∅ ̸= suppΠ(π,A) ∩ suppΠ(π,B) = suppΠ(π,A) = suppΠ(π,B) = {ω}, for some ω ∈ A ∩ B. Hence, c(π,A) = c(π,B) = δω. Since < c,Π > is Bayesian, c(c(π,A),B) = c(c(π,B),A) = δω.

But we know that there exist maxichoice updating rules which are not AGM-consistent.

5.5 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** : A lexicographic updating rule < c,Π > is **support dependent** if the family of lexicographic updating rules {pπ}π generating it satisfies :

For all π,π′ ∈ ∆(Ω), E∗(π) = E∗(π′) implies pπ\{π} = pπ′\{π′}

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

**Proposition 6** : Every support dependent lexicographic updating rules has a minimum distance representation.

**Proof** : Let < c,Π > be a support dependent lexicographic updating rule and let {pπ}π be the family generating it. Consider the metric d constructed in Proposition 3. Now we define a function ḅ : ∆(Ω) × ∆(Ω) → R+ as follows :

1. For π = π’, set ḅ(π,π’) = 0.

2. For π ̸= π’, let pπ = {pπk}k and pπ’ = {pπ’l}l define the natural number K(π,π’) as :

   K(π,π’) := min{h : E∗(πh) ∩ E∗(π’) ̸= ∅}
And now define:

\[ \hat{d}(\pi, \pi') = d(p_{I(\pi, \pi')}, \pi') \]

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 \( k(\pi, A) = \min\{k : p_{I(\pi, A)} > 0\} \). If \( E^*(p_{I(\pi, A)}^\pi) \cap A \) is a singleton, we are done. Assume, \( |E^*(p_{I(\pi, A)}^\pi) \cap A| \geq 2 \). Also, since \( < c, \Pi > \), is lexicographic, we have \( \Pi(\pi, A) = \{\pi' : E^*(\pi) = E^*(p_{I(\pi, A)}^\pi) \cap A\} \).

From support dependence, \( \hat{d}(\pi', \pi) = \hat{d}(\pi'', \pi) \) for all \( \pi', \pi'' \in \Pi(\pi, A) \). Also from the definition of \( d \), \( \hat{d}(\pi, \pi') = d(p_{I(\pi, A)}^\pi, \pi') > d(p_{I(\pi, A)}^\pi, p_{I(\pi, A)}^\pi(\mid A)) = \hat{d}(p_{I(\pi, A)}^\pi, c(\pi, A)) \) for all \( \pi' \in \Pi(\pi, A) \setminus \{c(\pi, A)\} \). Together, these imply the desired result: \( d_L(\pi, \pi') > d_L(\pi, c(\pi, A)) \) for all \( \pi' \in \Pi(\pi, A) \setminus \{c(\pi, A)\} \). Hence, \( < c, \Pi > \) is minimum-distance under the defined metric \( d_L \)

\[ \square \]

6 Discussion and Conclusion

6.1 Axiomatisation

Having defined updating rules, it would be desirable to have an axiomatisation of minimum distance updating rules and the Bayesian updating rule. For Bayesian injective updating rules, we have provided necessary and sufficient conditions (no-cycle condition) for a minimum distance representation. An axiomatisation of Bayes Rule has been provided by Majumdar (2005)([12]).

6.2 Decision Theoretic Framework

We shall focus on the savage framework. For an arbitrary \( \mathcal{P} \), the state space would need to be endowed with an algebra or \( \sigma \)-algebra. A natural algebra to endow \( \Omega \) with is
\[ G = \{ E(\phi) : \phi \in \mathcal{F}(P) \} \]

**Definition**: Let \( C \) be a set of consequences. An *act* \( f \) is defined as function \( f : \Omega \to C \).

Ideally an act should be required to be measurable. Let \(( C, \mathcal{C} )\) be the space of consequences where \( \mathcal{C} \) is an algebra or a \( \sigma \)-algebra on \( C \). An act could be defined as a measurable function \( f : (\Omega, G) \to (C, \mathcal{C}) \). However, in this paper, we shall deal with a finite \( P \). As a result, measurability issues will not be encountered. Denote as \( \mathcal{F} \), the set of all possible acts.

**Definition**: A *preference relation* \( \succsim \) on \( \mathcal{F} \) is defined as a complete and transitive binary relation on \( \mathcal{F} \).

Let \( f \) and \( g \) be two acts and \( E \) be an event. The act \( h = f_{E} g_{E^c} \) denotes the act which equals \( f \) on \( E \) and \( g \) on \( E^c \). More generally, let \( f_1, f_2, \ldots, f_n \) be acts and \( E_1, E_2, \ldots, E_n \) be disjoint events whose union is \( \Omega \). Then \( f_{E_1} f_{E_2} \ldots f_{E_n} \) denotes the act that takes value \( f_i \) on \( E_i \). For a preference relation \( \succsim \), we may define conditional preferences as follows. Given an event \( E \) define the conditional preference relation \( \succsim_E \) as:

\[
[f \succsim_E g] \iff [f_{E} h_{E^c} \succsim g_{E} h_{E^c} \text{ for all } h \in \mathcal{F}] 
\]

**Definition**: An event \( E \) is said to be *null* if \( f \sim_E g \) for all \( f, g \in \mathcal{F} \).

The interpretation of a null event is an event that is believed to never occur. If a decision maker knows \( E \) is never going occur, then he should be indifferent between any two acts which are equal outside \( E \) i.e. \( E^c \).

### 6.2.1 Belief sets

In the theory of belief revision under the framework of propositional logic, an agent’s primitive is what is known as a belief set. A belief set is any subset of formulae that the agent believes to be true. One natural requirement on a belief set is that it be logically closed. This means that if \( K \) is a belief set, we want \( K = \text{Cn}(K) \). We now define how from an agent’s preferences, his belief set could be elicited. In terms of a preference, the decision makers belief set should be the set of all formulae \( \phi \) whose negation \( \neg \phi \) is null (that is the event that \( \neg \phi \) is true).

**Definition** Let \( \succsim \) be a preference relation. Define the *belief set elicited from* \( \succsim \) as the set...
\[ K(\succeq) = \{ \phi \in \mathcal{F}(\mathcal{P}) : f \sim_{E(\neg\phi)} g \text{ for all } f, g \in \mathcal{F} \} \]

We now show that belief sets elicited from a decision maker’s preferences are logically closed. This is expressed in the following claim.

**Claim 8**: For any preference relation \( \succeq \), \( K(\succeq) \) is logically closed.

**Proof**: It is, of course, clear that \( K(\succeq) \subseteq \text{Cn}(K(\succeq)) \). We now show that \( \text{Cn}(K(\succeq)) \subseteq K(\succeq) \).

Let \( \phi \in \text{Cn}(K(\succeq)) \). We need to show that \( E(\neg\phi) \) is null. From the compactness of the operator \( \text{Cn} \), there exists a finite set of formulae \( J = \{ \alpha_1, \alpha_2, \ldots, \alpha_n \} \) such that \( J \vdash \phi \). So we have

\[ \bigcap_{i=1}^{n} E(\alpha_i) \subseteq E(\phi) \].

Hence, \( E(\neg\phi) = E(\phi)^c \subseteq \bigcup_{i=1}^{n} E(\neg\alpha_i) \). Since each \( E(\neg\alpha_i) \) is null, the result follows from the fact that finite union of null events is null and subsets of null events are null. \( \square \)

### 6.2.2 AGM revision and updating preferences

In this section we discuss first, the AGM method of belief revision; second, define what we mean by an updating rule for preference relations and lastly address the question whether an updating rule consistent with AGM exists.

We now discuss updating of preferences. We define a general updating rule for preferences as follows

**Definition**: Let \( \Theta \) be a domain of preferences relations. An *updating rule* is defined as mapping \( U : \Theta \times \mathcal{F}(\mathcal{P}) \to 2^{\mathcal{F} \times \mathcal{F}} \) such that \( U(\succeq, \phi) \) is a preference relation for every \( (\succeq, \phi) \).

We now ask whether there exists an updating rule \( \hat{U} \) such that \( K(\hat{U}(\succeq, \phi)) = \text{Cn}(K(\succeq) \cup \{ \alpha \}) \). The answer is in the affirmative if it can be established that for any logically closed set \( \hat{K} \subseteq \mathcal{F}(\mathcal{P}) \), there exists a preference relation \( \succeq \) such that \( K(\succeq) = \hat{K} \).

**Claim 9**: Let \( \hat{K} \subseteq \mathcal{F}(\mathcal{P}) \) be logically closed, then there exists a preference relation \( \succeq \) such that \( K(\succeq) = \hat{K} \).

**Proof**: For any \( \psi \in \hat{K} \), let \( \succeq(\psi) = \{ (f_{E(\neg\psi)}h_{E(\psi)}), g_{E(\neg\psi)}h_{E(\psi)} : f, g, h \in \mathcal{F} \} \). Now let \( \succeq = \bigcup_{\psi \in \hat{K}} \succeq(\psi) \). We show that \( \succeq \) is an equivalence relation.
We now ask whether given a logically closed set \( K \supseteq \) preference relation \((2009)\) 

\[ \text{we shall write } \hat{\sim} \text{ to denote } \hat{\sim} \text{ in the remainder of the proof. Let } [f] = \{ g \in \mathcal{F} \mid g \hat{\sim} f \} \text{ denote the equivalence class to which the act } f \text{ belongs. Let } \mathcal{E} = \{ [f] \mid f \in \mathcal{F} \} \text{ denote the equivalence classes (partition) generated by } \hat{\sim} \text{ on } \mathcal{F}. \]

Now, clearly \( \hat{\sim} \) is an equivalence relation, its symmetric part \( \hat{\sim} \) equals \( \hat{\sim} \). For convenience, we shall write \( \hat{\sim} \) to denote \( \hat{\sim} \) in the remainder of the proof. Let \([f] = \{ g \in \mathcal{F} \mid g \hat{\sim} f \} \) denote the equivalence class to which the act \( f \) belongs. Let \( \mathcal{E} = \{ [f] \mid f \in \mathcal{F} \} \) denote the equivalence classes (partition) generated by \( \hat{\sim} \) on \( \mathcal{F} \). We assume at this stage that cardinality restrictions on \( C \) and \( \mathbb{P} \) (for eg. \(|\mathbb{P}| < \infty \) and \(|C| = |\mathbb{R}|\)) have been imposed such that \( \mathcal{E} \) has cardinality less than or equal to the set of real numbers \( \mathbb{R} \). So there exists an injective map \( \tau : \mathcal{E} \to \mathbb{R} \). Now define the binary relation \( \hat{\geq} \) as follows. Let \( f, g \in \mathcal{F} \):

- If \([f] = [g]\), then set \( f \hat{\geq} g \)
- If \([f] \neq [g]\), then either \( \tau([f]) > \tau([g]) \) or \( \tau([g]) > \tau([f]) \).
  - If \( \tau([f]) > \tau([g]) \), then set \( f \hat{\geq} g \)
  - If \( \tau([g]) > \tau([f]) \), then set \( g \hat{\geq} f \)

Note that \( \hat{\geq} \) so defined is a complete and transitive. Furthermore, since \( \tau \) is an injective map, the symmetric part of \( \hat{\geq} \) say \( \sim \) equals \( \hat{\sim} \). Now, clearly \( \hat{\sim} \subseteq K(\hat{\geq}) \). Let \( \phi \in K(\hat{\geq}) \). Now, let \( f, g \in \mathcal{F} \) such that:

- \( f|_{\mathcal{E}(\phi)} = g|_{\mathcal{E}(\phi)} \)
- \( f(\omega) \neq g(\omega) \) for all \( \omega \not\in E(\phi) \)

Clearly, \( f \sim g \). Since \( \sim = \hat{\sim} \), we get \( f \hat{\sim} g \). This means there exists \( \psi \in \hat{\mathcal{K}} \) such that \( f|_{\mathcal{E}(\psi)} = g|_{\mathcal{E}(\psi)} \). This implies that \( E(\psi) \subseteq E(\phi) \) or equivalently \( \psi \vdash \phi \). Hence \( \phi \in \hat{\mathcal{K}} \). □

We now ask whether given a logically closed set \( K \) there exists a set of outcomes \( X \) and preference relation \( \hat{\geq} \) on \( [\Delta X]^{\Omega} \) satisfying the Anscombe-Aumann axioms (See Gilboa (2009)[1]) such that \( K(\hat{\geq}) = K \). It turns out that the only additional condition we need is the satisfiability of \( K \).

**Definition**: A set \( K \subseteq \mathcal{F}(\mathbb{P}) \) is **satisfiable** if \( \bigcap_{\phi \in K} E(\phi) \neq \emptyset \)
Definition: Let $X$ be a set of outcomes. We say a preference relation $\succeq$ on $[\Delta X]^\Omega$ is an AA-preference relation if it satisfies the Anscombe-Aumann axioms.

Claim 10: Let $K \subseteq \mathcal{F}(\mathbb{P})$ be a logically closed set. Then the following is true.

$$K \text{ is satisfiable } \iff \exists X \text{ finite and an AA-preference relation on } [\Delta X]^\Omega \text{ such that } K(\succeq) = K$$

Proof: We proof the $\Leftarrow$ direction first. Suppose $X$ is a finite set and $\succeq$ is AA preference relation on $[\Delta X]^\Omega$. We show that $K(\succeq)$ is satisfiable. Now, since $\succeq$ is AA, there exists a utility function $u : \Delta X \rightarrow \mathbb{R}$ and a probability measure on $\Delta \Omega$ such that:

$$f \succeq g \iff \sum_{\omega \in \Omega} \pi(\omega)u(f(\omega)) \geq \sum_{\omega \in \Omega} \pi(\omega)u(g(\omega))$$

Now since the null sets of $\succeq$ are exactly the zero probability events of $\pi$ (we can show this using the conditional preferences), we get that $\bigcap_{\phi \in K(\succeq)} E(\phi) = \{\phi : \pi(E(\phi)) = 1\}$. This implies that $\bigcap_{\phi \in K(\succeq)} E(\phi) = E^*(\pi)$ (where $E^*(\pi)$ is the support of $\pi$). But the support of any probability measure on $\Omega$ is non-empty. Hence $K(\succeq)$ is satisfiable.

Now consider the $\Rightarrow$ direction. Suppose $K$ is logically closed and satisfiable. Then, we have $\bigcap_{\phi \in K} E(\phi) \neq \emptyset$. For convenience we shall call $E(K) = \bigcap_{\phi \in K} E(\phi) \neq \emptyset$. Now define the following probability measure on $\Omega$:

$$\pi(\omega) = I_{E(K)}(\omega). \left(\frac{1}{|E(K)|}\right)$$

This is the probability measure which give uniform weights on all states in $E(K)$ and note that $E^*(\pi) = E(K)$. Note also that since $K$ is logically closed, $K = \{\phi : E(\phi) \subseteq E(\phi)\} = \{\phi : E^*(\pi) \subseteq E(\phi)\}$.

Now consider the finite set $X = \{a, b\}$ such that $a \neq b$ and define a function $\tau : X \rightarrow \mathbb{R}$ such that $\tau(a) > \tau(b)$. Now, define the utility function $u : \Delta X \rightarrow \mathbb{R}$ as $u(p_a, p_b) = p_a \tau(a) + p_b \tau(b)$ and define the following preference relation $\succeq$ on $[\Delta X]^\Omega$:

$$f \succeq g \iff \sum_{\omega \in \Omega} \pi(\omega)u(f(\omega)) \geq \sum_{\omega \in \Omega} \pi(\omega)u(g(\omega))$$

Now clearly, $K(\succeq) = \{\phi : \pi(E(\phi)) = 1\} = \{\phi : E^*(\pi) \subseteq E(\phi)\} = K$ and hence we get our desired result.

□
6.2.3 Updating rule

We have established now that there exist updating rules which are consistent with AGM Revision i.e we have shown that for any preference relation $\succ$ and formula $\alpha$, the set $\Theta(\succ, \alpha) = \{ \succ' : K(\succ') = K(\succ) * \alpha \}$ is non-empty. The latter set is the set of all possible posterior preferences consistent with the information after AGM revision. The natural question to ask is: what ought to be the finer set (possibly a singleton) of appropriate posteriors given potentially by an updating principle? This last question essentially expresses that updating preferences based on information is equivalent to a updating problem where given a prior preference relation and information (a formula), a set of posteriors is to be prescribed out of a set of all admissible posteriors. Of interest will be the following class of updating rules:

$$c(\succ, \alpha) = \arg \min_{\succ' \in \Theta(\succ, \alpha)} d(\succ, \succ')$$

Where $d$ is a metric on the set of all possible preferences relations. The above updating rule selects the posterior preference based on how "close" they are to the prior preference. It comes with the interpretation and that agents are hesitant to change preferences and would prefer stay as close to their prior preferences as possible. A natural question to ask is whether Bayesian updating can be represented in terms of the above choice for some metric.

6.3 Non-additive Probabilities

It would be worthwhile to repeat the procedure explained in the previous subsections for the case of non-additive probabilities. Updating rules and minimum distance updating rules will be defined in the same way and one could potentially ask whether the dempster-shafer rule and the generalised baye’s rule can be characterised as a minimum distance updating rule.

Also of interest would be the question of updating preferences. Some preliminary steps towards this have been discussed in section 2. One question which suggests itself could be: What metric on preference relations would give the conditional preference defined by Savage(1954)([14]) as minimally distant from a set of admissible posterior preference relations? This will tell us the following: Both Bayesians and ambiguity averse agents who have savage-like posterior preferences are hesitant under some measure to change there preferences while arriving at posterior beliefs given by bayes and dempter shafer rule respectively.
References


7 Appendix

7.1 Proofs of claim 4’ and 5’

We present here the proofs of claims 4’ and 5’

Proof of claim 4’:

1. We show, $K(\pi) \vdash A^c$. Now since $\pi(A) > 0$, $E^*(\pi) \cap A \neq \emptyset$. Now, let $\omega \in E^*(\pi) \cap A$. Since $\bigcap_{E \in K(\pi)} E = E^*(\pi)$, we have $\omega \in \bigcap_{E \in K(\pi)} E$ but $\omega \notin A^c$. Hence, $K(\pi) \vdash A^c$.

2. Clearly we have $E \in K \ast A = Cn(K(\pi) \cup \{A\}) \iff \bigcup_{E' \in K(\pi)} E' \cap A \subseteq E \iff E^*(\pi) \cap A \subseteq E$.

3. Let $\pi' \in \{\pi': K(\pi') = K(\pi) \ast A\}$. Now from the previous part we get $E \in K(\pi') \iff E^*(\pi') \cap A \subseteq E$. This implies that $\{E : \pi'(E) = 1\} = \{E : E^*(\pi) \cap A \subseteq E\}$. From this last equality it is clear that $E^*(\pi') \subseteq E^*(\pi) \cap A$ and also since $E^*(\pi') \in K(\pi')$, $E^*(\pi') \cap A \subseteq E^*(\pi')$. Hence, $E^*(\pi') = E^*(\pi) \cap A$.

Now, let $\pi' \in \{\pi': E^*(\pi') = E^*(\pi) \cap A\}$. Clearly we have $E \in K(\pi') \iff E^*(\pi') \subseteq E \iff E^*(\pi) \cap A \subseteq E \iff \bigcap_{E' \in K(\pi') \cup \{A\}} E' \subseteq E \iff E \in Cn(K(\pi) \cup \{A\}) \iff E \in K(\pi) \ast A$. Hence $K(\pi') = K(\pi) \ast A$.

Since $\pi(A) > 0$, $E^*(\pi) \cap A \neq \emptyset$. Hence, $\{\pi': E^*(\pi') = E^*(\pi) \cap A\} \neq \emptyset$.

4. This follows directly from the previous part.

Proof of claim 5’:

Clearly $\{\Omega\} \vdash A^c$ because at any state $\omega \in A$, $\Omega$ occurs but $A^c$ does not. Hence the set
\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\} is non-empty. So by Zorn’s Lemma, \(\max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\}\) (the set of maximal subsets of \(K(\pi)\) which do not imply \(A^c\)) is non-empty. We now prove the statements in claim 5.’

1. Clearly non-emptiness follows from the argument presented in the above paragraph. Now suppose \(A\) contains more than element. So, let \(\omega_1 \neq \omega_2\) such that \(\{\omega_1, \omega_2\} \subseteq A\). Consider \(\Gamma_1^* = \{E \in K(\pi) : E^*(\pi) \cup \{\omega_1\} \subseteq E\}\). We shall show that \(\Gamma_1^* \in \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\}\). Now note that \(\Gamma_1^* \neq \emptyset\) because clearly \(\Omega \in \Gamma_1^*\). Also note that since \(E^*(\pi) \cup \{\omega_1\} \subseteq K(\pi)\) and hence clearly in \(\Gamma_1^*, \cap_{\forall E \in \Gamma_1^*} E = E^*(\pi) \cup \{\omega_1\}\). So \(\Gamma_1^* \vdash A^c\). Now suppose for contradiction that there exists \(\Gamma_1^* \subseteq \Gamma' \subseteq K(\pi)\) such that \(\Gamma' \vdash A^c\). This implies that \(E^*(\pi) \subseteq \cap_{\forall E \in \Gamma'} E \subseteq E^*(\pi) \cup \{\omega_1\}\). So either \(\cap_{\forall E \in \Gamma'} E = E^*(\pi)\) or \(\cap_{\forall E \in \Gamma'} E = E^*(\pi) \cup \{\omega_1\}\). Suppose \(\cap_{\forall E \in \Gamma'} E = E^*(\pi)\). Since \(E^*(\pi) \subseteq A^c\), it cannot be the case that \(\Gamma' \vdash A^c\). So it must be that \(\cap_{\forall E \in \Gamma'} E = E^*(\pi) \cup \{\omega_1\}\). Now since \(\Gamma_1^* \subseteq \Gamma'\), there exists \(E \in \Gamma' \cap \Gamma_1^*\). For this \(E\) it must then be true that \(E^*(\pi) \cup \{\omega_1\} \subseteq E\). But this means that \(E \in \Gamma_1^*\) contradicting the fact that \(E \in \Gamma' \cap \Gamma_1^*\). Hence we get that \(\Gamma_1^* \in \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\}\).

Now notice that we could have done the same for \(\omega_2\). Instead of defining \(\Gamma_1^*\), we could have defined \(\Gamma_2^* = \{E \in K(\pi) : E^*(\pi) \cup \{\omega_2\} \subseteq E\}\) and \(\Gamma_2\) would also be in \(\max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\}\). But clearly \(\Gamma_1^* \neq \Gamma_2^*\) because there intersections are different. Hence there exists more than one elements in \(\max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\}\).

2. Let \(\Gamma' \in \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\}\). Since \(\Gamma' \vdash A^c\), \(\cap_{\forall E \in \Gamma'} E \cap A \neq \emptyset\). Suppose \(\cap_{\forall E \in \Gamma'} E \cap A = \{\omega\}\) for some \(\omega \in A\). So clearly, \(\cap_{\forall E \in \Gamma'} E = E^*(\pi) \cup \{\omega\}\). Now consider \(\Gamma^* = \{E \in K(\pi) : E^*(\pi) \cup \{\omega\} \subseteq E\}\). Clearly, \(\Gamma' \subseteq \Gamma^*\). And by the argument in the previous proof, \(\Gamma^* \in \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\}\). Since \(\Gamma' \in \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\}\), we know that it is not true that \(\Gamma' \subseteq \Gamma^*\). Hence \(\Gamma' = \Gamma^*\).

Now suppose \(\cap_{\forall E \in \Gamma'} E \cap A\) is non-singleton i.e contains more than one element. So there exist \(\omega_1 \neq \omega_2\) such that \(\{\omega_1, \omega_2\} \subseteq \cap_{\forall E \in \Gamma'} E \cap A\). Hence, \(E^*(\pi) \cup \{\omega_1, \omega_2\} \subseteq \cap_{\forall E \in \Gamma'} E\). Now consider the event \(E^*(\pi) \cup \{\omega_1\}\). Clearly, \(E^*(\pi) \cup \{\omega_1\} \notin \Gamma'\). Note that \(E^*(\pi) \cup \{\omega_1\} \subseteq K(\pi)\). Consider the set \(\Gamma_0 = \Gamma' \cup \{E^*(\pi) \cup \{\omega_1\}\}\). Clearly \(\cap_{\forall E \in \Gamma_0} E = E^*(\pi) \cup \{\omega_1\}\) so \(\Gamma_0 \vdash A^c\). But \(\Gamma' \subseteq \Gamma_0\) which contradicts the fact that \(\Gamma' \in \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\}\). Hence, \(\cap_{\forall E \in \Gamma'} E \cap A\) cannot be a non-singleton.

3. Let \(\Gamma^* \in \max\{\Gamma \subseteq K(\pi) : \Gamma \vdash A^c\}\). From the previous part we know that there exists \(\omega \in A\) such that \(\cap_{\forall E \in \Gamma^*} E = E^*(\pi) \cup \{\omega\}\). Now let \(\hat{\pi} \in \{\pi' : K(\pi') = Cn(\Gamma^* \cup \{\omega\})\}\). So \(E \in \{E' : \hat{\pi}(E') = 1\} \iff E \in Cn(\Gamma^* \cup \{\omega\}) \iff (\cap_{\forall E \in \Gamma^*} E') \cap A \subseteq E \iff \omega \in E\). But this means that \(\hat{\pi} = \delta_\omega\).
7.2 Defining the metric

Claim: 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 $\min\{|\pi_1([\omega, \omega'])\}, |\pi_3([\omega, \omega'])| > 0$.
     
     - **Sub-case 1:** $\pi_2 \neq \pi_i$ for all $i \in \{1, 3\}$ and $\min\{|\pi_2([\omega, \omega'])\}, |\pi_i([\omega, \omega'])| > 0$ for atleast 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 $\min\{|\pi_2([\omega, \omega'])\}, |\pi_i([\omega, \omega'])| = 0$ for all $i \in \{1, 3\}$.
     
     In this case we have:
     
     $S_{\omega\omega'}(\pi_1, \pi_3) = \left| \frac{\pi_1(w)}{\pi_1([\omega, \omega'])} - \frac{\pi_3(w)}{\pi_3([\omega, \omega'])} \right|$
     
     $\leq \left| \frac{\pi_1(w)}{\pi_1([\omega, \omega'])} - \frac{\pi_2(w)}{\pi_2([\omega, \omega'])} \right| + \left| \frac{\pi_2(w)}{\pi_2([\omega, \omega'])} - \frac{\pi_3(w)}{\pi_3([\omega, \omega'])} \right|$
     
     $= S_{\omega\omega'}(\pi_1, \pi_2) + S_{\omega\omega'}(\pi_2, \pi_3)$ \hspace{1cm} (5)
   
   - **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 $\min\{|\pi_1([\omega, \omega'])\}, |\pi_3([\omega, \omega'])| = 0$
     
     - **Sub-case 1:** $\pi_2 \neq \pi_i$ for all $i \in \{1, 3\}$.
       
       Then:
       
       $S_{\omega\omega'}(\pi_1, \pi_3) = \epsilon \leq S_{\omega\omega'}(\pi_1, \pi_2) + S_{\omega\omega'}(\pi_2, \pi_3)$
   
     - **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.

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

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

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

**Case 2:** $\pi(\omega) > 0$:
In this case, we have $\frac{\pi'(\omega')}{\pi(\omega)} > \frac{\pi'(\omega')}{\pi(\omega)}$ which implies that $\frac{\pi'(\omega')}{\pi(\omega)} > \frac{\pi'(\omega')}{\pi(\omega)}$. Hence $S_{\omega,\omega'}(\pi, \pi') > 0$

☐

### 7.3 Proof of Proposition 3

The following claim will be useful in establishing the result:

Claim: 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 $\{s_i(x)\}_{i \in \mathbb{Z}}$. 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 $\tau_x(s_i(x)) = \tau(t)$. Hence we have $t > r \Rightarrow \tau_x(s_i(x)) > \tau_x(s_r(x))$. Now define $S : X \times Y \to \mathbb{R}$

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

The desired result can be shown by the above function.

Proof of Proposition 3 : 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.

- $c(., B) : \Delta A \to \Delta B$
- $c(., A) : \Delta B \to \Delta A$

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 $S_{A,B} : \Delta A \times \Delta B \to [0, 1]$ 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 $\rho : \Delta \Omega \times \Delta \Omega \to \mathbb{R}_+$. Let $\pi, \pi' \in \Delta \Omega$. Define $\rho$ as follows:

- If $E^*(\pi) \cap E^*(\pi') \neq \emptyset$, then set:
  $$\rho(\pi, \pi') = d(\pi, \pi')$$
- If $E^*(\pi) \cap E^*(\pi') = \emptyset$, and $\min(|E^*(\pi)|, |E^*(\pi')|) \geq 2$ then set:
\[ \rho(\pi, \pi') = \rho(\pi', \pi) = S_{E^*(\pi), E^*(\pi')}(\pi, \pi') \]

- If \( E^*(\pi) \cap E^*(\pi') = \emptyset \), and \( \min\{|E^*(\pi)|, |E^*(\pi')|\} = 1 \), w.l.o.g let \( |E^*(\pi)| = 1 \), then set \( \rho(\pi, \pi') = \rho(\pi', \pi) = 1/2 \) if \( \pi' = c(\pi, E^*(\pi')) \) 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 \( \bar{\rho} = \sup_{\pi \neq \pi'} \rho(\pi, \pi') \) and \( \underline{\rho} = \inf_{\pi \neq \pi'} \rho(\pi, \pi') \). Clearly \( \underline{\rho} < \bar{\rho} \) and there exists a \( K > 0 \) such that \( \bar{\rho} + K < 2(\underline{\rho} + K) \). Now define the function \( \hat{d} : \Delta \Omega \times \Delta \Omega \to \mathbb{R} \) 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 \( \{\pi_t\} \) 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

\[
\begin{align*}
    d(\pi_{s+1}, c(\pi_s, B)) &> d(\pi_{s+1}, c(\pi_{s+1}, B)) \\
    &> d(\pi_{s+2}, c(\pi_{s+1}, B))
\end{align*}
\]

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. \( \Box \)