FORMAL EPISTEMOLOGY: Representing the fixation of belief and its undoing

Horacio Arló Costa

June 20, 2006

1 Introduction

Formal epistemology is a discipline with various branches and sub-branches dealing with precise representations of attitudes and their flux. The formal tools utilized for this purpose go from standard probability theory to non-standard (or infinitesimal) probability, to decision theoretical tools to logical tools.

These lectures are divided in two parts. The first part deals with static representations of knowledge and belief via epistemic and doxastic operators. This is perhaps the oldest sub-branch of formal epistemology, as well as one of the parts of the discipline that has witnessed a strong renaissance in recent years, mainly due to related work in computer science and mathematical economics.

Even in this first part we will see that apparently subtle philosophical distinctions between knowledge as a standard for serious possibility and a more external conception of knowledge where agents are seen as potential sources of knowledge, leads to very different formal models. We will explore both while considering mainly propositional languages. Pointers to richer models are nevertheless provided and, time-permitting, we will devote some sessions to consider models of knowledge and belief of more expressive power where quantification plays an essential role.

Under the logical point of view the first part provides a self-contained introduction to the study of salient examples of epistemic logics and various semantics for them. We will start with Jaakko Hintikka's models of epistemic and doxastic operators, and then we will show that some of the primitives of these semantics can be clarified in the context of a semantic approach of greater scope and power, deriving from early work by Dana Scott and Richard Montague. The epistemic operators interpreted via these semantic frameworks deal with formal models of epistemic or doxastic attributions to a third person, rather than with knowledge or belief claims (of the first person). The final part of this first section deals with knowledge and belief claims by appealing to a less committal theory of acceptance of epistemic claims, whose logical underpinnings have been independently studied in recent years by researchers in the area of Artificial Intelligence and by various philosophers (Isaac Levi, Robert Stalnaker, and myself, among others). We will close the first section, by considering also probabilistic models and models in terms of ranking functions deriving from the work of Wolfgang Spohn. More information about the last two topics will be provided in class.

The second part is divided in two parts. The first part offers a first decision theoretical model for the process of changing view either by incorporating information (via expansions or revisions) or by effectively withdrawing information from the current view (contraction). This first model springs from recent joint work with Isaac Levi and therefore it reflects his most recent views about belief change articulated in a recent book [48].

The second part offers a different account of belief change, which is also inspired by work in rational choice. It is a generalization of a decision theoretical view of belief change first explored by Hans Rott in a recent book [66] and various articles. The model deals with the problem of indeterminate or imprecise epistemic values and is therefore related to work recently done by Isaac Levi [48] and by Amartya Sen [71], independently.

The two main parts of these lectures have many interesting connections, some of which we will consider in certain detail at the end of the course. The complete picture and some of its interconnections should provide an introduction to some of the most interesting problems discussed in this area. A list of open problems will be discussed in class the last day of classes as well as a list of related topics which we will not have time to consider, like the conceptual and logical problems deriving from the study of multi-agent interaction.

2 Logical preliminaries

Let L_0 be a language containing a complete set of Boolean connectives, including *falsum* and *verum* constants $\perp$ and T. The set of wff of L_0 are defined in the usual manner and $K, J, ...$denote sets of sentences. Let Cn be a function from sets of sentences to set of sentences, obeying the following Tarskian postulates:

Inclusion $A \in Cn(\{A\})$

Iteration $Cn(K) = Cn(Cn(K))$

Monotony If $K \subseteq H$, then $Cn(K) \subseteq Cn(H)$

as well as (at least);

Compactness For all X $\subseteq L_0$, Cn(X) = $\cup${C(Y): Y is a finite subset of X}

Consistency $\perp \notin$ Cn($\emptyset$)

Deduction $A \rightarrow B \in Cn(K)$ iff B $\in Cn(K \cup \{A\})$.

We suppose also that $Cn(\emptyset)$ contains all substitution instances of classical tautologies expressible in L_0. A *theory* is any set of sentences K such that $Cn(K) = K$.

2.1 Consistency and model sets

Consider sets M of sentences satisfying the following conditions:

(C.¬) If $A \in M$, then $\neg A \notin M$.

(C.∧) If $A \wedge B \in M$, then $A \in M$ and $B \in M$.

(C.∨) If $A \vee B \in M$, then $A \in M$ or $B \in M$.

(C. ¬¬) If $\neg\neg A \in M$, then $A \in M$.

(C.¬∧) If $\neg(A \wedge B) \in M$, then $\neg A \in M$ or $\neg B \in M$.

(C.¬∨) If $\neg(A \vee B) \in M$, then $\neg A \in M$ and $\neg B \in M$.

Jaakko Hintikka argued almost half a century ago that these sets have interesting properties. The main idea is that a set of sentences is inconsistent if it cannot be embedded into a model set. He also argued convincingly that therefore model sets are ‘a very good formal counterpart of the informal idea of (partial) description of a possible state of affaires or (following Liebniz) a possible world.

3 The logical structure of an epistemic state

At any point in time we believe some propositions explicitly and we might believe others implicitly. But one can argue (see [43]) that we are *committed* to believe the logical consequences of what we believe explicitly. So, utilizing the previous notation, if we are aware of our belief in each of a set S of propositions, then we are also committed to believe $Cn(S)$. This requirement of *deductive cogency* – encompassing consistency and logical closure – requires that the only feasible candidates for representing doxastic commitments are theories in the logical sense just specified.

3.1 Hintikka and epistemic operators

A very influential way of specifying formally what are our doxastic and epistemic commitments appeals to *epistemic* (or *doxastic*) operators. The idea was first articulated by Jaakko Hintikka in his seminal essay [31].

For this purpose we can enrich our language by adopting the following grammar.

$$p \mid \neg\phi \mid \phi \wedge \phi \mid \Box\phi \mid \Diamond\phi$$

where $p \in \mathsf{At}$, the set of atomic sentences. The operator $\Box$ can be interpreted here as 'agent a knows ϕ'. A doxastic interpretation is also possible and we will make that explicit when needed. In addition we have the operator '$\Diamond$' which can be interpreted as 'it is epistemically possible that ϕ'.

In [31] Hintikka asks the following important question:

> [...] how the properties of a model sets are affected by the presence of the notions of knowledge and belief; how, in other words, the notion of model set can be generalized in such a way that the consistency of a set of statements remains tantamount to its capacity of being imbedded in a model set. What additional conditions are needed when the notions of knowledge and belief are present.

In particular consider the statement that says that a sentence A is epistemically possible at a sate of affaires w. In other words consider a state of affaires w where it is true to say, for a fixed person a, that A is possible. A central idea about how to evaluate these statements is presented as follows by Hintikka:

> Clearly the content of this statement cannot be adequately expressed by speaking only of the state of one state of affaires. The statement in question can be true only if there is a possible state of affaires where in which A would be true: but this state of affaires need not be identical with the one where the statement was made. A description of such state of affaires will be called an *alternative* to w, with respect to the agent a. Hence we have to impose the following conditions on the description of a model set w:

If $\Diamond A \in w$ then there is at least an alternative $w*$ to w such that $A \in w*$

(C.KK*) If $\Box A \in w$ then for all alternatives $w*$ to w we have that $\Box A \in w*$

Let Ω be a *model system* understood as a set of model sets (closed under the previous C-rules, CKK* and C.K – presented below). Then we have:

If $\Diamond A \in w$ and w in a model system Ω, then there is at least an alternative $w*$ to w such that $A \in w*$

From now on we will tacitly assume we are working with model systems. We also need:

(C.K) If $\Box A \in w$ then $A \in w$.

The previous account entails:

If $\Box A \in w$ then for all alternatives $w*$ to w we have that $A \in w*$

To which, in the case of knowledge we have to add that the alternativeness realtion is transitive, reflexive and closed under its ancestral.

Hintikka admits that the statements above are not quite explicit, given that the notion of epistemic alternative and the notion of existence of such alternative are left unspecified. Nevertheless, this kind of analysis of epistemic operators has been prevalent for almost more than 60 years now. The analysis was also proposed by other logicians and metaphysicians like Saul Kripke in his influential essay [39] for the notion of metaphysical necessity.

3.1.1 Truth sets

So far we have talked informally about states of affaires. We can now introduce formally a set of points W endowed with the informal interpretation of encoding the space of available such states. Then we can define a truth set for a sentence A as the set of states where A is true. We will denote this truth set as $|A|$. B

3.2 Epistemic functions

Every possible state can be associated with agent-dependent function D that for each state w yields as values a set of truth sets $D(w)$ that are fully believed by the agent. In particular, if the epistemic state of the given agent is represented by a theory K, the we have that:

$$|A| \in D_K(w) \text{ if and only if } A \in K$$

Notice that as long as K is consistent we have that $\bigcap D(w)$ is non empty and:

$$\bigcap D_K(w) \subseteq |A| \text{ if and only if } A \in K$$

Notice also that the set $\bigcap D_K(w)$ has also a natural interpretation here: it is the set of all states where all the sentences in K are true. If we extend the notion of truth set to sets of sentences in the obvious way, we will also have that $|K| = \bigcap D_K(w)$. We can put this in yet a different way: $\bigcap D_K(w)$ represents the strongest proposition fully believed by our agent.

If we want to deal with the notion of knowledge, then we can have an epistemic function E instead, only that here a constant function will not do (why?). If the body of known propositions is also represented by a theory K here we can then have that for each $w \in \bigcap E_K(w)$, $E(w) = \{X \in 2^W\colon \bigcap E_K(w) \subseteq X\}$, and $E(v) = \{X \in 2^W\colon \{v\} \subseteq X\}$ otherwise.

But now we have an natural candidate for the notion of epistemic alternative (among possible worlds not model sets), namely v is an alternative to w if and only if the agent in question knowns (in the sense of fully believes truly) all that he knowns at w. In this case it is clear that the only epistemic alternatives to w are exaclty the states in $\bigcap E_K(w)$. So, we have that:

$$\bigcap E_K(w) \subseteq |A| \text{ if and only if } |A| \text{ is known (by Bthe agent represented by K) if and only if } \Box A \text{ holds at } w$$

But now, armed with the more precise notion of epistemic alternative we can *derive*:

If $\Box A \in w$ then for all alternatives $w*$ to W we have that $\Box A \in w*$

For only epistemic alternatives $v \in \bigcap E_K(w)$ (to w), and only those, are states where any known proposition at w is also known (in the sense of fully believed truly) at v. Notice also that if the notion of epistemic alternative to a state w is clarified in some way then we can always extract from it a set $\bigcap E_K(w)$. In fact, one just can take all states $w*$ related to w and identify $\bigcap E_K(w)$ with this set, which also should coincide with $|K|$.

3.2.1 Relational semantics for knowledge

A **relational frame** is a pair $\langle W, R\rangle$ where R is an alternativeness relation on W (i.e., $R \subseteq W \times W$). A **relational model** based on a frame $\mathbb{F}$ is a pair $\langle \mathbb{F}, V\rangle$ where $V : \mathsf{At} \rightarrow 2^W$ is a valuation function. Formulas from $\mathcal{L}$ are interpreted at states from W. The intuition is that each state should be thought of as a boolean valuation[1] and boolean connectives are interpreted locally. The additional structure in a relational model (i.e., the relation R) is used to give a truth-value to formulas where the knowledge (belief) operator is the main connective. Formally, truth in a relational model $\mathbb{M} = \langle KW, R, V\rangle$ is defined inductively as follows. Let $w \in W$ and $\phi \in \mathcal{L}$,

1. $\mathbb{M}, w \models p$ iff $w \in V(p)$
2. $\mathbb{M}, w \models \neg\phi$ iff $\mathbb{M}, w \not\models \phi$
3. $\mathbb{M}, w \models \phi \wedge \psi$ iff $\mathbb{M}, w \models \phi$ and $\mathbb{M}, w \models \psi$
4. $\mathbb{M}, w \models \Box\phi$ iff for each $v \in W$, if wRv then $\mathbb{M}, v \models \phi$
5. $\mathbb{M}, w \models \Diamond\phi$ iff there is a $v \in W$ such that wRv and $\mathbb{M}, v \models \phi$

Of course, the previous account can also be interpreted in a different way. For example, the main operator $\Box$ can be seen as representing the notion of metaphysical necessity.

We say that a formula $\phi \in \mathcal{L}$ is valid in a model $\mathbb{M}$, written $\mathbb{M} \models \phi$, if for each state w in $\mathbb{M}$, $\mathbb{M}, w \models \phi$. The formula ϕ is valid in a frame $\mathbb{F}$, written $\mathbb{F} \models \phi$, if for all models $\mathbb{M}$ based on that frame, $\mathbb{M} \models \phi$. Using relations to represent this additional information forces us to accept a number of principles. For example,

Exercise 1 *Prove that the following formulas are valid on any relational frame*

[1] To make this precise, think of the valuation function V as a function from $W \times \mathsf{At}$ to $\{0, 1\}$.

1. $\Box(\phi \wedge \psi) \rightarrow \Box\phi \wedge \Box\psi$

2. $\Box\phi \wedge \Box\psi \rightarrow \Box(\phi \wedge \psi)$

3. $\Box\top$

4. $\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$

5. $\Box\phi \leftrightarrow \neg\Diamond\neg\phi$

Of course, if the modality is knowledge we will need to impose adequate constrains on R. Traditionally in this setting it is required that the accessibility relation is an equivalence relation.

Exercise 2 *Which are the most salient similarities and differences between this account and Hintikka's presentation in terms of model sets? Can Hintikka impose symmetry in his characterization of knowledge aside from transitivity and reflexivity?*

What about the ideas that we presented informally via the use of the epistemic function E? We can start with a set of primitive states W and then use an **epistemic function** function $E : W \rightarrow 2^{2^W}$.

3.2.2 Scott-Montague-style semantics

Definition 1 *A pair $\langle W, E\rangle$ is a called a* **neighborhood system**, *or a* **epistemic frame**, *if W a non-empty set and E is an epistemic function.*

Recall the definition of the basic modal language, denoted by $\mathcal{L}(\mathsf{At})$, where At is a set of atomic sentences:

$$p \mid \neg\phi \mid \phi \wedge \phi \mid \Box\phi \mid \Diamond\phi$$

where $p \in \mathsf{At}$. We can now interpret formulas of $\mathcal{L}(\mathsf{At})$ in epistemic models.

Definition 2 *Let $\mathbb{F} = \langle W, N\rangle$ be an epistemic frame. A* **model** *based on $\mathbb{F}$ is a tuple $\langle W, N, V\rangle$ where $KV : \mathsf{At} \rightarrow 2^W$ is a valuation function.*

Let $\mathbb{M} = \langle W, E, V\rangle$ be a model and $w \in W$. Truth of a formula $\phi \in \mathcal{L}(\mathsf{At})$ is defined inductively as follows:

- $\mathbb{M}, w \models p$ iff $w \in V(p)$

- $\mathbb{M}, w \models \neg\phi$ iff $\mathbb{M}, w \not\models \phi$
- $\mathbb{M}, w \models \phi \wedge \psi$ iff $\mathbb{M}, w \models \phi$ and $\mathbb{M}, w \models \psi$
- $\mathbb{M}, w \models \Box\phi$ iff $(\phi)^{\mathbb{M}} \in N(w)$
- $\mathbb{M}, w \models \Diamond\phi$ iff $W - (\phi)^{\mathbb{M}} \notin E(w)$

where $(\phi)^{\mathbb{M}}$ denotes the **truth set** of ϕ. That is $(\phi)^{\mathbb{M}} = \{w \mid \mathbb{M}, w \models \phi\}$. The following properties of the truth set will be used throughout these notes. We first need some notation. Let $E : W \to 2^{2^W}$ be an epistemic function. Define $m_E : 2^W \to 2^W$ as follows: for $X \subseteq W$,

$$m_E(X) = \{w \mid X \in E(w)\}$$

Intuitively, $m_E(X)$ is the set of states in which X is known. Let $\mathbb{M} = \langle W, E, V\rangle$ be a neighborhood model.

1. $(p)^{\mathbb{M}} = V(p)$ for $p \in \mathsf{At}$
2. $(\neg\phi)^{\mathbb{M}} = W - (\phi)^{\mathbb{M}}$
3. $(\phi \wedge \psi)^{\mathbb{M}} = (\phi)^{\mathbb{M}} \cap (\psi)^{\mathbb{M}}$
4. $(\Box\phi)^{\mathbb{M}} = m_E((\phi)^{\mathbb{M}})$
5. $(\Diamond\phi)^{\mathbb{M}} = W - m_E(W - (\phi)^{\mathbb{M}})$

The proof of the above statements is an easy application of the definition of truth and is left to the reader. We say ϕ is **valid** in $\mathbb{M}$, denoted $\mathbb{M} \models \phi$, if for each $w \in W$, $\mathbb{M}, w \models \phi$. Also, ϕ is **valid** in a frame $\mathbb{F}$ if for each model $\mathbb{M}$ based on $\mathbb{F}$, $\mathbb{M} \models \phi$.

Exercise 3 *Prove that the following formula and rules are valid on all epistemic models.*

1. $\Box\phi \leftrightarrow \neg\Diamond\neg\phi$
2. *From* $\phi \leftrightarrow \psi$ *infer* $\Box\phi \leftrightarrow \Box\psi$

Notice that many of the axioms that are true in all relational frames can perfectly be invalid here:

Exercise 4 *Prove that the following axiom schemes and rules are not valid in the class of all epistemic models*

1. $\Box\phi \wedge \Box\psi \to \Box(\phi \wedge \psi)$
2. $\Box\top$
3. $\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$
4. *From* $\phi \to \psi$ *infer* $\Box\phi \to \Box\psi$
5. $\Box\phi \to \phi$

Notice also that in our previous analysis we focused on a particular subset of epistemic frames, for which there are also elementary equivalent relational models. To make this clear we will present here some basic possible constraints on epistemic models, and then we will return to this issue.

3.2.3 Constraints on epistemic models

Let W be a non-empty set. We will be working with collections of subsets of W. I.e., elements of 2^{2^W}. The following properties of $\mathcal{F} \subseteq 2^W$ will be relevant for our study:

1. We say $\mathcal{F}$ is **closed under intersections** if for any collections of sets $\{X_i\}_{i\in I}$ such that for each $i \in I$, $X_i \in \mathcal{F}$, then $\bigcap_{i\in I} X_i \in \mathcal{F}$. For any cardinal κ, we say that $\mathcal{F}$ is **closed under $\leq \kappa$-intersections** if for each collections of sets $\{X_i\}_{i\in I}$ from $\mathcal{F}$ with $|I| \leq \kappa$, $\cap_{i\in I} X_i \in \mathcal{F}$.
2. We say $\mathcal{F}$ is **closed under unions** if for any collections of sets $\{X_i\}_{i\in I}$ such that for each $i \in I$, $X_i \in \mathcal{F}$, then $\cup_{i\in I} X_i \in \mathcal{F}$. For any cardinal κ, we say that $\mathcal{F}$ is **closed under $\leq \kappa$-unions** if for each collections of sets $\{X_i\}_{i\in I}$ from $\mathcal{F}$ with $|I| \leq \kappa$, $\cup_{i\in I} X_i \in \mathcal{F}$.
3. We say that $\mathcal{F}$ is **closed under complements** if for each $X \subseteq W$, if $X \in \mathcal{F}$, then $X^C \in \mathcal{F}$.
4. We say $\mathcal{F}$ is **supplemented**, or **closed under supersets** provided for each $X \subseteq W$, if $X \in \mathcal{F}$ and $X \subseteq Y \subseteq W$, then $Y \in \mathcal{F}$.
5. We say $\mathcal{F}$ **contains the unit** provided $W \in \mathcal{F}$; and $\mathcal{F}$ **contains the empty set** if $\emptyset \in \mathcal{F}$.

6. Call the set $\bigcap_{X \in \mathcal{F}} X$ the **core of** $\mathcal{F}$. We say that $\mathcal{F}$ **contains its core** provided $\cap_{X \in \mathcal{F}} X \in \mathcal{F}$.

7. We say $\mathcal{F}$ is **consistent** if $\emptyset \notin \mathcal{F}$ and $\mathcal{F} \neq \emptyset$.

Exercise 5 *Prove that $\mathcal{F}$ is supplemented iff if $X \cap Y \in \mathcal{F}$ then $X \in \mathcal{F}$ and $Y \in \mathcal{F}$.*

Property 6 above deserves some comments. If $\mathcal{F}$ contains the unit, then intuitively $\mathcal{F}$ contains a maximal element (under the subset relation). Similarly, if $\mathcal{F}$ contains its core, then $\mathcal{F}$ contains a minimal element (under the subset relation). The following definition lists some well-known collections of sets.

Definition 3 *Let W be a set and $\mathcal{F} \subseteq 2^W$, then*

1. *$\mathcal{F}$ is a* **filter** *if $\mathcal{F}$ contains the unit, closed under finite intersections and supplemented. $\mathcal{F}$ is a proper filter if in addition $\mathcal{F}$ does not contain the emptyset.*

2. *$\mathcal{F}$ is a* **topology** *if $\mathcal{F}$ contains the unit, the emptyset, is closed under finite intersections and arbitrary unions.*

3. *$\mathcal{F}$ is* **augmented** *if $\mathcal{F}$ contains its core and is supplemented.*

Let's look at augmented collections in a bit more detail. The following fact is straightforward and is left as an exercise.

Lemma 4 *If $\mathcal{F}$ is augmented, then $\mathcal{F}$ is closed under arbitrary intersections.*

Proof Left as an exercise for the reader.

Of course, the converse is false. But that is not very interesting since it is easy to construct collections of sets closed under intersections that are not supplemented. What is more interesting is that there are consistent filters that are not augmented.

Lemma 5 *There are consistent filters that are not augmented.*

Proof Left as an exercise for the reader.

It should be now clear that in our previous analysis we focused only on *augmented* epistemic frames. In fact, it is easy to see that:

Lemma 6 *If $\mathcal{F}$ is augmented, then $X \in \mathcal{F}$ if and only if $\bigcap \mathcal{F} \subseteq X$.*

Proof Left as an exercise for the reader.

Definition 7 *Suppose that $R \subseteq W \times W$ is a relation. The pair $\langle W, R\rangle$ is called a* **relational frame**, *or a* **relational structure**.

Given a relation R on a set W (i.e., $R \subseteq W \times W$) we can define the following functions:

1. $R^{\rightarrow} : W \rightarrow 2^W$ defined as follows. For each $w \in W$, let $R^{\rightarrow}(w) = \{v \mid wRv\}$.
2. $R^{\leftarrow} : 2^W \rightarrow 2^W$ defined as follows. For each $X \subseteq W$, $R^{\leftarrow}(X) = \{w \mid \exists v \in X$ such that $wRv\}$.

Definition 8 *Given a relation R on a set W and a state $w \in W$ a set $X \subseteq W$ is* **known at** *w if $R^{\rightarrow}(w) \subseteq X$. Let $\mathcal{N}_w$ be the set of sets that are known at w. That is,*

$$\mathcal{N}_w = \{X \mid R^{\rightarrow}(w) \subseteq X\}$$

The following lemma shows that the collection of known sets have very nice properties.

Lemma 9 *Let R be a relation on W. Then for each $w \in W$, $\mathcal{N}_w$ is augmented.*

Proof Left as an exercise.

□

Furthermore, properties of R are reflected in these collection of sets.

Observation 10 *Let W be a set and $R \subseteq W \times W$.*

1. *If R is reflexive, then for each $w \in W$, $w \in \bigcap \mathcal{N}_w$*
2. *If R is transitive then for each $w \in W$, if $X \in \mathcal{N}_w$, then $\{v \mid X \in \mathcal{N}_v\} \in \mathcal{N}_w$.*

Proof Suppose that R is reflexive. Let $w \in W$ be an arbitrary state. Suppose that $X \in \mathcal{N}_w$. Then since R is reflexive, wRw and hence $w \in R^{\rightarrow}(w)$. Therefore by the definition of $\mathcal{N}_w$, $w \in X$. Since X was an arbitrary element of $\mathcal{N}_w$, $w \in X$ for each $X \in \mathcal{N}_w$. Hence $w \in \bigcap \mathcal{N}_w$.

Suppose that R is transitive. Let $w \in W$ be an arbitrary state. Suppose that $X \in \mathcal{N}_w$. We must show $\{v \mid X \in \mathcal{N}_v\} \in \mathcal{N}_w$. That is, we must show $R^{\rightarrow}(w) \subseteq \{v \mid X \in \mathcal{N}_v\}$. Let $x \in R^{\rightarrow}(w)$. Then wRx. To complete the proof we need only show $X \in \mathcal{N}_x$. That is, we must show $R^{\rightarrow}(x) \subseteq X$. Since R is transitive, $R^{\rightarrow}(x) \subseteq R^{\rightarrow}(w)$ (why?). Hence since $R^{\rightarrow}(w) \subseteq X$, $R^{\rightarrow}(x) \subseteq X$.

□

Exercise 6 *State and prove analogous results for the situations when R is serial, Euclidean and symmetric respectively.*

From this point of view, we can think of relational frames and (augmented) epistemic frames as two different ways of presenting the same information. That is, we are after a mathematical structures that can represent for each state, the set of known propositions at each state. It should be clear that with epistemic frames, there is more freedom in which collection of sets can be known at a particular state. A natural question to ask is under what circumstances do an epistemic frame and a relational frame represent the same information.

Definition 11 *Let $\langle W, E\rangle$ be an epistemic frame and $\langle V, R\rangle$ be a relational frame. We say that $\langle W, E\rangle$ and $\langle V, R\rangle$ are* **equivalent** *if there is a function $f : 2^W \rightarrow 2^V$ such that for each $X \subseteq W$, $X \in E(w)$ iff $f[X] \in \mathcal{N}_w$.*

Of particular interest are situations where $W = V$ and f is the identity function. The following theorems give the situations in which relational frames and neighborhood frame are equivalent.

Theorem 12 *Let $\langle W, R\rangle$ be a relational frame. Then there is an equivalent augmented neighborhood frame.*

Proof The proof is trivial given the previous Lemma. For each $w \in W$, let $E(w) = \mathcal{N}_w$.

□

Theorem 13 *Let $\langle W, E\rangle$ be an augmented epistemic frame. Then there is an equivalent relational frame.*

Proof Let $\langle W, E\rangle$ be an epistemic frame. We must define a relation R_E on W. Since $\langle W, E\rangle$ is augmented, for each $w \in W$, $\bigcap E(w) \in E(w)$. For each $w, v \in W$, we say wR_Ev iff $v \in \bigcap E(w)$. To show $\langle W, R_E\rangle$ and $\langle W, E\rangle$ are equivalent, we must show for each $w \in W$, $\mathcal{N}_w = E(w)$. Let $w \in W$ and $X \subseteq W$. If $X \in \mathcal{N}_w$. Then $R_E^{\rightarrow}(w) \subseteq X$. Since $R_E^{\rightarrow}(w) = \bigcap E(w)$ and E contains its core, $R_E^{\rightarrow}(w) \in E(w)$. Furthermore, since E is supplemented and $R_E^{\rightarrow}(w) = \bigcap E(w) \subseteq X$, $X \in E(w)$. Suppose that $X \in E(w)$. Then clearly $\cap E(w) \subseteq X$. Hence $X \in \mathcal{N}_w$.

□

3.2.4 Definability Results

When working over relational frames, modal formulas **define** properties of a relation in the following sense. A relational frame $\mathbb{F} = \langle W, R\rangle$ validates $\Box\phi \to \phi$ iff R is reflexive. Similar observations are true when working with neighborhood models (here we will use the generic letter N to refer to neighborhoods, rather than D or E (reserved for doxastic or epistemic models).

Lemma 14 *Let $\mathbb{F} = \langle W, N\rangle$ be a neighborhood frame. Then $\mathbb{F} \models \Box\phi \wedge \Box\psi \to \Box(\phi \wedge \psi)$ iff $\mathbb{F}$ is closed under finite intersections.*

Proof Suppose that $\mathbb{F} = \langle W, N\rangle$ is a neighborhood frame that is closed under finite intersections. We must show $\mathbb{F} \models \Box\phi\wedge\Box\psi \to \Box(\phi\wedge\psi)$. Let $\mathbb{M} = \langle W, N, V\rangle$ be any model based on $\mathbb{F}$ and $w \in W$. Suppose that $\mathbb{M}, w \models \Box\phi \wedge \Box\psi$. Then $(\phi)^{\mathbb{M}} \in N(w)$ and $(\psi)^{\mathbb{M}} \in N(w)$. Since $N(w)$ is closed under finite intersections, $(\phi)^{\mathbb{M}} \cap (\psi)^{\mathbb{M}} \in N(w)$. Hence $(\phi \wedge \psi)^{\mathbb{M}} \in N(w)$ and therefore $\mathbb{M}, w \models \Box(\phi \wedge \psi)$.

Suppose that $\mathbb{F} = \langle W, N\rangle$ is not closed under finite intersections. Then there are sets $X_1, \ldots X_k$ such that $X_i \in N(w)$ for $i = 1, \ldots, k$ but $\cap_{1\leq i\leq k} X_i \notin N(w)$. We claim that there are two sets Y and Y' such that $Y, Y' \in N(w)$ but $Y \cap Y' \notin N(w)$. If $\cap_{i=1}^{k-1} X_i \in N(w)$, then let $Y = \cap_{i=1}^{k-1} X_i \in N(w)$ and $Y' = X_k$. If not, then $\cap_{i=1}^{k-1} \notin N(w)$. If $\cap_{i=1}^{k-2} X_i \in N(w)$, then let $Y = \cap_{i=1}^{k-2} X_i \in N(w)$ and $Y' = X_{k-1}$. If not, then $\cap_{i=1}^{k-2} X_i \notin N(w)$. Continue in this manner until we find and $l \leq k$ such that $\cap_{i=1}^{l} X_i, X_{l+1} \in N(w)$ but $\cap_{i=1}^{l+1} X_i \notin N(w)$. Notice that such an l must exists since $X_i \in N(w)$ for each $i = 1, \ldots, k$. Then define a valuation function so that $V(p) = Y$ and $V(q) = Y'$. Hence, $\mathbb{M}, w \models \Box p \wedge \Box q$. However, since $Y \cap Y' \notin N(w)$, $\mathbb{M}, w \not\models \Box(p \wedge q)$.

□

Lemma 15 *Let $\mathbb{F} = \langle W, N \rangle$ be a neighborhood frame. Then $\mathbb{F} \models \Box(\phi \wedge \psi) \rightarrow \Box\phi \wedge \Box\psi$ iff $\mathbb{F}$ is closed under supersets.*

Proof The right to left direction is left as an exercise for the reader. Suppose that $\mathbb{F} = \langle W, N \rangle$ is not closed under supersets. Then there are sets X and Y such that $X \subseteq Y$, $X \in N(w)$ but $Y \notin N(w)$. Define a valuation V such that $V(p) = X$ and $V(q) = Y$. Then since $X \subseteq Y$, $(p \wedge q)^{\mathbb{M}} = X \in N(w)$. Hence $\mathbb{M}, w \models \Box(p \wedge q)$. However since, $(q)^{\mathbb{M}} = Y \notin N(w)$, $\mathbb{M}, w \not\models \Box q$. Hence $\mathbb{M}, w \not\models \Box p \wedge \Box q$.

□

Lemma 16 *Let $\mathbb{F} = \langle W, N \rangle$ be a neighborhood frame. Then $\mathbb{F} \models \Box\top$ iff $\mathbb{F}$ contains the unit.*

Proof Left as an exercise for the reader.

Lemma 17 *Let $\mathbb{F} = \langle W, N \rangle$ be a neighborhood frame such that for each $w \in W$, $N(w) \neq \emptyset$.*

1. *$\mathbb{F} \models \Box\phi \rightarrow \phi$ iff for each $w \in W$, $w \in \bigcap N(w)$*
2. *$\mathbb{F} \models \Box\phi \rightarrow \Box\Box\phi$ iff for each $w \in W$, if $X \in N(w)$, then $\{v \mid X \in N(v)\} \in N(w)$*

Proof Suppose that $\mathbb{F} = \langle W, N \rangle$ is a neighborhood frame. Suppose that for each $w \in W$, $w \in \cap N(w)$. Let $\mathbb{M}$ be any model based on $\mathbb{F}$ and $w \in W$. Suppose that $\mathbb{M}, w \models \Box\phi$. Then $(\phi)^{\mathbb{M}} \in N(w)$. Since $w \in \cap N(w) \subseteq (\phi)^{\mathbb{M}}$, $w \in (\phi)^{\mathbb{M}}$. Hence $\mathbb{M}, w \models \phi$. As for the converse, suppose that $w \notin \cap N(w)$. Since $N(w) \neq \emptyset$, there is an $X \in N(w)$ (note that X may be empty) such that $w \notin X$, otherwise $w \in \cap N(w)$. Define a valuation V such that $V(p) = X$. Then it is easy to see that $\mathbb{M}, w \models \Box p$ but $\mathbb{M}, w \not\models X$.

Suppose that for each $w \in W$, if $X \in N(w)$, then $\{v \mid X \in N(v)\} \in N(w)$. Suppose that $\mathbb{M}$ is any model based on $\mathbb{F}$ and $\mathbb{M}, w \models \Box\phi$. Then $(\phi)^{\mathbb{M}} \in N(w)$. Therefore, by assumption $\{v \mid (\phi)^{\mathbb{M}} \in N(v)\} \in N(w)$. Since $(\Box\phi)^{\mathbb{M}} = \{v \mid (\phi)^{\mathbb{M}} \in N(w)\}$, $\mathbb{M}, w \models \Box\Box\phi$. For the other direction, suppose that there is some state $w \in W$ and set X such that $X \in N(w)$ but $\{v \mid X \in N(v)\} \notin N(w)$. The define a valuation V such that $V(p) = X$. It is easy to verify that $\mathbb{M}, w \models \Box p$ but $\mathbb{M}, w \not\models \Box\Box p$.

□

Exercise 7 *Find properties on frames that are defined by the following formulas:*

1. $\Diamond\top$

2. $\neg\Box\phi \rightarrow \Box\neg\Box\phi$

3. $\Box\phi \rightarrow \Diamond\phi$

4 Topology and Modal Logic

Much of the original motivation for neighborhood frames as a semantics form modal logics comes from elementary point-set topology. The idea is to think of the propositions in $N(w)$ to be *close* to the point w. In this section we consider an alternative semantics for modal logics inspired by point-set topology due to McKinsey and Tarski, We begin by reviewing some very basic point-set topology. More information can be found in any point-set topology text book (Dugundji [20] is an excellent choice).

4.1 Basic Topological Notions and Topological Semantics

In this section, we discuss some basic point-set topology. Readers already familiar with such notions can easily skip to the next subsection. Recall the definition of a topology.

Definition 18 *A* **topological space** *is a neighborhood frame $\langle W, \mathcal{T}\rangle$ where W is a nonempty set and*

1. $W \in \mathcal{T}$, $\emptyset \in W$

2. *$\mathcal{T}$ is closed under finite intersections*

3. *$\mathcal{T}$ is closed under arbitrary unions.*

The collection $\mathcal{T}$ is called a **topology**. Elements $O \in \mathcal{T}$ are called **opens**. A set C such that $W - C \in \mathcal{T}$ is called **closed**. Given a topology $\langle W, \mathcal{T}\rangle$, let $\mathcal{T}_C$ be the collection of closed subsets of W, i.e., $\mathcal{T}_C = \{C \mid W - C \in \mathcal{T}\}$. The following observation is an easy consequence of the above definition.

Observation 19 *Let $\langle W, \mathcal{T}\rangle$ be a topological space. Then $\mathcal{T}_C$ has the following properties:*

1. $\emptyset, W \in \mathcal{T}_C$

2. *$\mathcal{T}_C$ is closed under finite unions*

3. *$\mathcal{T}_C$ is closed under arbitrary intersections*

Proof Left as an exercise for the reader.

Given a topological space $\langle W, \mathcal{T}\rangle$ and a point $w \in W$, a **neighborhood of** w is any open set that contains w. Let $\mathcal{T}_w = \{O \mid O \in \mathcal{T} \text{ and } w \in O\}$ be the collection of all neighborhoods of w.

Lemma 20 *Let $\langle W, \mathcal{T}\rangle$ be a topological space. Then for each $w \in W$, the collection $\mathcal{T}_w$ contains W, is closed under finite intersections and closed under arbitrary unions.*

Proof The proof is left as an exercise for the reader.

□

Definition 21 *Let $\langle W, \mathcal{T}\rangle$ be a topological space. A pair $\langle W, N\rangle$ is called a* **neighborhood system** *provided $N : W \to \mathcal{T}$ is defined as follows: $N(w) = \mathcal{T}_w$.*

Let $\langle W, \mathcal{T}\rangle$ be a topological space and $X \subseteq W$ any set. The largest open subset of X is called the **interior** of X, denoted $Int(X)$. Formally,

$$Int(X) = \cup\{O \mid O \in \mathcal{T} \text{ and } O \subseteq X\}$$

The smallest closed set containing X is called the **closure** of X, denoted $Cl(X)$. Formally,

$$Cl(X) = \cap\{C \mid W - C \in \mathcal{T} \text{ and } X \subseteq C\}$$

It is easy to see that a set X is open if $Int(X) = X$ and closed if $Cl(X) = X$. The following Lemma will be helpful when studying the topological semantics of the next section.

Lemma 22 *Let $\langle W, \mathcal{T}\rangle$ be a topological space and $X \subseteq W$. Then*

1. $Int(X \cap Y) = Int(X) \cap Int(Y)$

2. $Int(\emptyset) = \emptyset$, $Int(W) = W$

3. $Int(X) \subseteq X$

4. $Int(Int(X)) = Int(X)$

5. $Int(X) = W - Cl(W - X)$

Exercise 8 *Use the last fact in the above lemma to derive corresponding properties for the closure operator.*

The famous result of Tarski and McKinsey shows that **S4** is sound and complete with respect to the class of all topologies `Discuss results`.

Let $\langle W, N, V\rangle$ be a neighborhood models. Suppose that N satisfies the following properties

- for each $w \in W$, $N(w)$ is a filter
- for each $w \in W$, $w \in \bigcap N(w)$
- for each $w \in W$ and $X \subseteq W$, if $X \in N(w)$, then $m_N(X) \in N(w)$

We can now show that there is a topological model that is point-wise equivalent to $\mathcal{M}$. Consider the set $\mathcal{B} = \{m_N(X) \mid X \subseteq W\}$. We will show that $\mathcal{B}$ is a base. That is we must show that

1. $\cup\mathcal{B} = W$ and 2. for each $X, Y \in \mathcal{B}$ and each $x \in X \cap Y$ there is a $Z \in \mathcal{B}$ such that $x \in Z \subseteq X \cap Y$.

Suppose that $X = m_N(x_1)$ and $Y = m_N(X_2)$ and $x \in X \cap Y$. Since $N(w)$ is a filter, $m_N(X_1) \cap m_N(X_2) = m_N(X_1 \cap X_2)$.

Thus $x \in m_N(X_1 \cap X_2) \subseteq m_N(X_1) \cap m_N(X_2)$. Thus we are done if we can show that $m_N(X_1 \cap X_2) \in \mathcal{B}$. But this follows form thethird property above since $X_1 \cap X_2 \in N(w)$.

4.2 Topological Semantics

In this section, we briefly introduce topological semantics for modal logic. This semantics has been around for nearly 60 years and is usually attributed to McKinsey and Tarski [51]. A **topological model** is a tuple $\mathbb{M}^T = \langle W, \mathcal{T}, V\rangle$, where $\langle W, \mathcal{T}\rangle$ is a topological space and V is a valuation function. Formulas in $\mathcal{L}$ are interpreted at states $w \in W$. The boolean connectives and atomoic propositions are interpreted as usual. We only give the definition of truth of the modal operator:

$$\mathbb{M}^T, w \models \Box\phi \text{ iff } \exists O \in \mathcal{T}, w \in O \text{ such that } \forall v \in O, \mathbb{M}^T, v \models \phi$$

Notice the similarity between this definition an the definition of truth of the modal operator $\langle\ ]$. The only difference is the extra clause $w \in O$. However, recall the function $w \mapsto \mathcal{T}_w$, where $\mathcal{T}_w$ is the set of neighborhoods of w. Then, the above clause can be written as

$$\mathbb{M}^T, w \models \Box\phi \text{ iff } \exists O \in \mathcal{T}_w \text{ such that } \forall v \in O, \mathbb{M}^T, v \models \phi$$

Although this differnece is a trivial change in terminology, it demonstrates a close connection between neighborhood frames and topological frames. Finally it is easy to see that

$$(\Box\phi)^{\mathbb{M}^T} = Int((\phi)^{\mathbb{M}^T})$$

4.3 Logics for Neighborhood Frames

In this section we will be interested in the following axiom schemas and rules.

PC Any axiomatization of propositional calculus

E $\Box\phi \leftrightarrow \neg\Diamond\neg\phi$

M $\Box(\phi \wedge \psi) \rightarrow (\Box\phi \wedge \Box\psi)$

C $(\Box\phi \wedge \Box\psi) \rightarrow \Box(\phi \wedge \psi)$

N $\Box\top$

K $\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$

RE $\dfrac{\phi \leftrightarrow \psi}{\Box\phi \leftrightarrow \Box\psi}$

Nec $\dfrac{\phi}{\Box\phi}$

MP $\dfrac{\phi \quad \phi \rightarrow \psi}{\psi}$

Let **E** be the smallest set of formulas closed under instances of PC, E and the rules RE and MP. **E** is the smallest classical modal logic. The logic **EC** extends **E** by adding the axiom scheme C. Similarly for **EM**, **EN**, **ECM**, and **EMCN**. The logic **K** is the smallest set of formulas closed under instances of PC, K, E and the rules Nec and MP. Let **S** be any of the above logics, we write $\vdash_{\mathbf{S}} \phi$ if $\phi \in \mathbf{S}$.

Our first observation about these axiom systems, is that we can prove a uniform substitution theorem in **E**. Given a formulas $\phi, \psi, \psi' \in \mathcal{L}$, let $\phi[\psi/\psi']$ be the formula ϕ but replace *some* occurrences of ψ with ψ'. For example, suppose that ϕ is the formula $\Box(\Diamond p \wedge \Box\Box q) \wedge \Box p$, ψ is the formula p and ψ' is the formula $\Box p$. Then $\phi[\psi/\psi']$ can be any of the following

- $\Box(\Diamond p \wedge \Box\Box q) \wedge \Box p$
- $\Box(\Diamond\Box p \wedge \Box\Box q) \wedge \Box p$
- $\Box(\Diamond p \wedge \Box\Box q) \wedge \Box\Box p$
- $\Box(\Diamond\Box p \wedge \Box\Box q) \wedge \Box\Box p$

The uniform substitution theorem states that we can always replace logically equivalent formulas.

Theorem 23 (Uniform Substitution) *The following rule can be derived in* **E**

$$\frac{\psi \leftrightarrow \psi'}{\phi \leftrightarrow \phi[\psi/\psi']}$$

Proof Suppose that $\vdash_{\mathbf{E}} \psi \leftrightarrow \psi'$. We must show $\vdash_{\mathbf{E}} \phi \leftrightarrow \phi[\psi/\psi']$. First of all, note that if ϕ and ψ are the same formula. Then either $\phi[\psi/\psi']$ is ϕ (when ψ is not replaced) or $\phi[\psi/\psi']$ is ψ' (when ψ is replaced). In the first case, $\phi \leftrightarrow \phi[\psi/\psi']$ is the formula $\phi \leftrightarrow \phi$ and so trivially, $\vdash_{\mathbf{E}} \phi \leftrightarrow \phi[\psi/\psi']$. In the second case, $\phi \leftrightarrow \phi[\psi/\psi']$ is the formula $\psi \leftrightarrow \psi'$, which can be deduced in **E** by assumption. Thus we may assume that ϕ and ψ are distinct formulas.

The proof is by induction on ϕ. The base case and boolean connectives are left as an exercise for the reader. We demonstrate the modal operator. Suppose that ϕ is $\Box\gamma$ and $\vdash_{\mathbf{E}} \psi \leftrightarrow \phi'$. The induction hypothesis is $\vdash_{\mathbf{E}} \gamma \leftrightarrow \gamma[\psi/\psi']$. Using the *RE* rule, $\vdash_{\mathbf{E}} \Box\gamma \leftrightarrow \Box(\gamma[\psi/\psi'])$. Note that $\Box(\gamma[\psi/\psi'])$ is the same formula as $\Box\gamma[\psi/\psi']$. Hence $\vdash_{\mathbf{E}} \Box\gamma \leftrightarrow \Box\gamma[\psi/\psi']$.

$\Box$

The above theorem will often be used (without reference). We will now proceed to show a number of basic facts about the above axiom systems. Much more information can be found in [16], Chapter 8. The first is an alternative characterization of **EM**. The alternative axiomatization is in terms of a rule. The rule, called **right monotonicity** (*RM*) is

$$\frac{\phi \to \psi}{\Box\phi \to \Box\psi}$$

Lemma 24 *The logic* **EM** *equals the logic* **E** *plus the rule RM.*

Proof The proof follows easily if we can show that RM is a derived rule in **EM** and M can be derived in the logic **E** plus the rule RM. We first show that RM can be derived in **EM**.

1.	$\phi \to \psi$	Assumption
2.	$\phi \leftrightarrow (\phi \wedge \psi)$	Follows from 1 and propositional reasoning
3.	$\Box\phi \leftrightarrow \Box(\phi \wedge \psi)$	2 and RE
4.	$\Box(\phi \wedge \psi) \to \Box\phi \wedge \Box\psi$	Instance of M
5.	$\Box\phi \to \Box\phi \wedge \Box\psi$	Follows from 3,4 using propositional reasoning
6.	$\Box\phi \to \Box\psi$	Follows form 5 using propositional reasoning

Thus RM is a derived rule of **EM**. We now show that we can derive M in the logic **E** plus the rule RM.

1.	$(\phi \wedge \psi) \to \phi$	Propositional Tautology
2.	$\Box(\phi \wedge \psi) \to \Box\phi$	Follows from 1 using RM
3.	$(\phi \wedge \psi) \to \psi$	Propositional Tautology
4.	$\Box(\phi \wedge \psi) \to \Box\psi$	Follows from 3 using RM
5.	$\Box(\phi \wedge \psi) \to \Box\phi \wedge \Box\psi$	Follows from 2,4 using propositional reasoning

Thus M can be derived using the rule RM.

□

The logic **K** is the smallest **normal**, or Kripkean, modal logic. The logics **E**, **EM**, etc. are strictly *weaker* than **K**.

Lemma 25 *The logic* **EMC** *equals the logic* **E** *plus the axiom scheme K.*

Proof The fact that $\vdash_{\mathbf{K}} M$ and $\vdash_{\mathbf{K}} C$ is left for the reader. We only show that K can be derived in **EMC**. First of all note that since **EMC** contains M, the rule RM is derivable (see Lemma 24).

1.	$((\phi \to \psi) \wedge \phi) \to \psi$	Propositional Tautology
2.	$\Box((\phi \to \psi) \wedge \phi) \to \Box\psi$	From 1 and RM
3.	$(\Box(\phi \to \psi) \wedge \Box\phi) \to \Box((\phi \to \psi) \wedge \phi)$	Instance of C
4.	$(\Box(\phi \to \psi) \wedge \Box\phi) \to \Box\psi$	Follows from 2,3 using[2] PR
5.	$\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$	Follows from 4 using PR

$\Box$

Lemma 26 *The logic* **EN** *equals the logic* **E** *plus the rule Nec.*

Proof It is easy to see that using Nec, we can prove $\Box\top$. Suppose that $\vdash_{\mathbf{EN}} \phi$. We must show that $\vdash_{\mathbf{EN}} \Box\phi$. Using propositional reasoning since ϕ is derivable in **EN**, $\vdash_{\mathbf{EN}} (\top \leftrightarrow \phi)$. Using RE, $\vdash_{\mathbf{EN}} \Box\top \leftrightarrow \Box\phi$. Hence $\vdash_{\mathbf{EN}} \Box\top \to \Box\phi$. Since $\Box\top$ is obviously provable in **EN**, $\vdash_{\mathbf{EN}} \Box\phi$.

$\Box$

Putting these two lemmas together, we see that **EMCN** is the smallest normal modal logic.

Corollary 27 *The logic* **EMCN** *equals the logic* **K***.*

In this section we have define eight modal logics between **E** and **K**. Of course, there still remains a question as to whether all eight logics are in fact distinct. This will be answered in the following section.

4.3.1 Soundness and Completeness

In this section we will prove soundness and completeness results. It is assumed that the reader is familiar with basic soundness and completeness results in modal logic (with respect to relational frames), see [14, 16] for more information. We quickly review some some basic terminology.

Let F be a collection of neighborhood frames. A formula $\phi \in \mathcal{L}$ is **valid in** F, or F**-valid** if for each $\mathbb{F} \in \mathsf{F}$, $\mathbb{F} \models \phi$. We say that a logic **L** is **sound** with respect to F, provided $\mathbf{L} \subseteq \mathsf{F}$. That is for each formula $\phi \in \mathcal{L}$, $\vdash_{\mathbf{L}} \phi$ implies ϕ is valid in F. Given a set of formulas Γ, a formula ϕ and a collection of frames F, we say Γ **semantically entails** ϕ with respect to F, denoted $\Gamma \models_{\mathsf{F}} \phi$, if for each $\mathbb{F} \in \mathsf{F}$, if $\mathbb{F} \models \Gamma$ then $\mathbb{F} \models \phi$. Here $\mathbb{F} \models \Gamma$ means for each $\phi \in \Gamma$, $\mathbb{F} \models \phi$. Finally we write $\models_{\mathsf{F}} \phi$ if for each $\mathbb{F} \in \mathsf{F}$, $\mathbb{F} \models \phi$. A logic **L** is **weakly complete** with respect to a class of frames F, if $\models_{\mathsf{F}} \phi$ implies $\vdash_{\mathbf{L}} \phi$. A logic **L** is **strongly complete** with respect to a class of frames F, if for each set of formulas Γ, $\Gamma \models_{\mathsf{F}} \phi$ implies $\Gamma \vdash_{\mathbf{L}} \phi$.

Let **L** be any modal logic. A formula $\phi \in \mathcal{L}$ is said to be **inconsistent in L**, or**L-inconsistent** if $\vdash_{\mathbf{L}} \neg\phi$. A set of formulas Γ is said to be **L-inconsistent** if $\Gamma \vdash_{\mathbf{L}} \bot$. The set Γ is **L-consistent** if it is not inconsistent. A consistent set of formulas Γ is called a **maximally consistent set** if for each formula ϕ, either $\phi \in \Gamma$ or $\neg\phi \in \Gamma$.

Let $M_{\mathbf{L}}$ be the set of **L**-maximally consistent sets of formulas. Recall that *Lindenbaum's Lemma* (see [16] and [14] for an extended discussion) states that given any consistent collection of formulas Γ' there is a maximally consistent set of formulas that contains Γ'. Given a formula $\phi \in \mathcal{L}$, let $|\phi|_{\mathbf{L}}$ be the **proof set** of ϕ in **L**. Formally, $|\phi|_{\mathbf{L}} = \{\Delta \mid \Delta \in W_{\mathbf{L}} \text{ and } \phi \in \Delta\}$. We first note that proof sets share a number of properties in common with truth sets.

Lemma 28 *Let* **L** *be a logic and* $\phi, \psi \in \mathcal{L}$. *Then*

1. $|\phi \wedge \psi|_{\mathbf{L}} = |\phi|_{\mathbf{L}} \cap |\psi|_{\mathbf{L}}$
2. $|\neg\phi|_{\mathbf{L}} = M_{\mathbf{L}} - |\phi|_{\mathbf{L}}$
3. $|\phi \vee \psi|_{\mathbf{L}} = |\phi|_{\mathbf{L}} \cup |\psi|_{\mathbf{L}}$
4. *If* $|\phi|_{\mathbf{L}} \subseteq |\psi|_{\mathbf{L}}$ *then* $\vdash_{\mathbf{L}} \phi \rightarrow \psi$
5. $\vdash_{\mathbf{L}} \phi \leftrightarrow \psi$ *iff* $|\phi|_{\mathbf{L}} = |\psi|_{\mathbf{L}}$

Proof The proofs are standard facts about maximally consistent sets and left for the reader.

□

Finally, one more fact about proof sets that will be of interest.

Lemma 29 *For each* $\phi \in \mathcal{L}$, $\psi \in \bigcap |\phi|_{\mathbf{L}}$ *iff* $\vdash_{\mathbf{L}} \phi \rightarrow \psi$.

Proof Suppose that $\vdash_{\mathbf{L}} \phi \rightarrow \psi$. Then for each maximally consistent set Γ, $\phi \rightarrow \psi \in \Gamma$. Hence since for each $\Gamma \in |\phi|_{\mathbf{L}}$, $\phi \in \Gamma$, we have $\psi \in \Gamma$. Thus $\phi \in \cap|\phi|_{\mathbf{L}}$.

Suppose that $\psi \in \cap|\phi|_{\mathbf{L}}$ but it is not the case that $\vdash_{\mathbf{L}} \phi \rightarrow \psi$. Then $\neg(\phi \rightarrow \psi)$ is consistent in **L**. Using Lindenbaum's Lemma, we can construct a maximally consistent set Γ such that $\neg(\phi \rightarrow \psi) \in \Gamma$. That is, $\phi, \neg\psi \in \Gamma$. Since $\phi \in \Gamma$, $\Gamma \in |\phi|_{\mathbf{L}}$. But then $\neg\psi \in \Gamma$ contradicts the fact that $\psi \in \cap|\phi|_{\mathbf{L}}$.

□

Given any model $\mathbb{M} = \langle W, N, V\rangle$ and a set $X \subseteq W$, we say that X is **definable** in $\mathbb{M}$ if there is a formula $\phi \in \mathcal{L}$ such that $(\phi)^{\mathbb{M}} = X$. Let $\mathcal{D}_{\mathbb{M}}$ be the set of all sets that are definable in $\mathbb{M}$. First of all note, that $\mathcal{D}_{\mathbb{M}} \neq 2^W$. A simple counting argument will show this fact. For there are only countably many formulas of $\mathcal{L}$

and hence only countably many definable subsets. However (if W is countable) 2^W is uncountable. A subset $X \subseteq M_{\mathbf{L}}$ is called a **proof set** provided there is some formula $\phi \in \mathcal{L}$ such that $X = |\phi|_{\mathbf{L}}$. Again notice that there are only countably many proof sets; however, if At is countable, then $M_{\mathbf{L}}$ is uncountable, and hence there are uncountably many subsets of $M_{\mathbf{L}}$.

As usual, when constructing a canonical model the states of the world will be maximally consistent sets, i.e., elements of $M_{\mathbf{L}}$. Suppose that the logic $\mathbf{L}$ contains $\mathbf{E}$. Consider any function $N_{\mathbf{L}} : M_{\mathbf{L}} \to 2^{2^{M_{\mathbf{L}}}}$ such that for each $\Gamma \in M_{\mathbf{L}}$

$$|\phi|_{\mathbf{L}} \in N_{\mathbf{L}}(\Gamma) \text{ iff } \Box\phi \in \Gamma$$

Intuitively, given a set Γ, $N_{\mathbf{L}}(\Gamma)$ contains at least all the proof sets of necessary formulas from Γ. When defining any function, the first question one should ask is *is the function well-defined*? This may be a strange question to ask about $N_{\mathbf{L}}$, since we were *assuming* that it is in fact a function. So, what can go wrong? Well, $N_{\mathbf{L}}$ is any function that maps maximally consistent sets Γ to sets of subsets of $M_{\mathbf{L}}$. The only requirement on $N_{\mathbf{L}}$ is that for each $\Gamma \in M_{\mathbf{L}}$, $|\phi|_{\mathbf{L}} \in N(\Gamma)$ iff $\Box\phi \in \Gamma$. In particular, this requirement ensures that for each set $X \in N_{\mathbf{L}}(\Gamma)$ such that $X = |\phi|_{\mathbf{L}}$, $\Box\phi \in \Gamma$. Say that a proof set $|\phi|_{\mathbf{L}}$ is **necessary at** Γ provided $\Box\phi \in \Gamma$. Thus the functions we are interested in map maximally consistent sets Γ to collections of subsets of $M_{\mathbf{L}}$ that can be broken into to parts: the proof sets that are necessary at Γ and non-proof sets (i.e., sets not of the form $|\phi|_{\mathbf{L}}$ for some ϕ). However, as we see from Lemma 28 it is possible that a proof set X is defined by two different formulas. Thus to ensure that any functions can be defined that satisfy the above property, we must check that if $|\phi|_{\mathbf{L}} \in N_{\mathbf{L}}(\Gamma)$ and $|\phi|_{\mathbf{L}} = |\psi|_{\mathbf{L}}$, then $\Box\psi \in \Gamma$.

Lemma 30 *For any logic* $\mathbf{L}$ *containing the rule RE, if* $N_{\mathbf{L}} : M_{\mathbf{L}} \to 2^{2^{M_{\mathbf{L}}}}$ *is a function such that for each* $\Gamma \in M_{\mathbf{L}}$*,* $|\phi|_{\mathbf{L}} \in N_{\mathbf{L}}(\Gamma)$ *iff* $\Box\phi \in \Gamma$*. Then if* $|\phi|_{\mathbf{L}} \in N_{\mathbf{L}}(\Gamma)$ *and* $|\phi|_{\mathbf{L}} = |\psi|_{\mathbf{L}}$*, then* $\Box\psi \in \Gamma$*.*

Proof Let ϕ and ψ be two formulas such that $|\phi|_{\mathbf{L}} = |\psi|_{\mathbf{L}}$. Further suppose that $\Box\phi \in \Gamma$ and $|\phi|_{\mathbf{L}} \in N_{\mathbf{L}}(\Gamma)$. Since $|\phi|_{\mathbf{L}} \in N_{\mathbf{L}}(\Gamma)$, $\Box\phi \in \Gamma$. Also, by Lemma 28, since $|\phi|_{\mathbf{L}} = |\psi|_{\mathbf{L}}$, $\vdash_{\mathbf{L}} \phi \leftrightarrow \psi$. Using the *RE* rule, $\vdash_{\mathbf{L}} \Box\phi \leftrightarrow \Box\psi$. Hence $\Box\phi \leftrightarrow \Box\psi \in \Gamma$. Hence $\Box\psi \in \Gamma$.

$\Box$

Define the **canonical valuation**, $V_{\mathbf{L}} : \mathsf{At} \to 2^{M_{\mathbf{L}}}$ as follows. Let $p \in \mathsf{At}$, then $V_{\mathbf{L}}(p) = |p|_{\mathbf{L}} = \{\Gamma \mid \Gamma \in M_{\mathbf{L}} \text{ and } p \in \Gamma\}$.

Definition 31 *A neighborhood model* $\mathbb{M} = \langle W, N, V \rangle$ *is called a* **canonical for L** *if*

1. $W = M_{\mathbf{L}}$
2. *for each* $\Gamma \in W$ *and each formula* ϕ, $|\phi|_{\mathbf{L}} \in N(\Gamma)$ *iff* $\Box\phi \in \Gamma$
3. $V = V_{\mathbf{L}}$

For example, let $\mathbb{M}^{min}_{\mathbf{L}} = \langle M_{\mathbf{L}}, N^{min}_{\mathbf{L}}, V_{\mathbf{L}} \rangle$, where for each $\Gamma \in M_{\mathbf{L}}$, $N^{min}_{\mathbf{L}}(\Gamma) = \{|\phi|_{\mathbf{L}} \mid \Box\phi \in \Gamma\}$. The model $\mathbb{M}^{min}_{\mathbf{L}}$ is easily seen to be a canonical for **L**. Furthermore, it is the minimal canonical for **L** in the sense that for each Γ, $N_{\mathbf{L}}(\Gamma)$ is defined to be the smallest set still satisfying the requisite property. Let $P_{\mathbf{L}}$ be the set of all proof sets of $M_{\mathbf{L}}$. Alternatively the largest canonical, $\mathbb{M}^{max}_{\mathbf{L}}$ can be defined as $\langle M_{\mathbf{L}}, N^{max}_{\mathbf{L}}, V_{\mathbf{L}} \rangle$, where for each $\Gamma \in M_{\mathbf{L}}$, $N^{max}_{\mathbf{L}}(\Gamma) = N^{min}_{\mathbf{L}}(\Gamma) \cup \{X \mid X \subseteq M_{\mathbf{L}}, X \notin P_{\mathbf{L}}\}$

Lemma 32 (Truth Lemma) *For any consistent logic* **L** *and any consistent formula* ϕ, *if* $\mathbb{M}$ *is canonical for* **L**,

$$(\phi)^{\mathbb{M}} = |\phi|_{\mathbf{L}}$$

Proof The boolean connectives are as usual and left for the reader. We focus on the modal case. Let $\mathbb{M} = \langle W, N, V \rangle$ be canonical for **L**. Suppose that $\Gamma \in (\Box\phi)^{\mathbb{M}}$, then by definition $(\phi)^{\mathbb{M}} \in N(\Gamma)$. By the induction hypothesis, $(\phi)^{\mathbb{M}} = |\phi|_{\mathbf{L}}$, hence $|\phi|_{\mathbf{L}} \in N(\Gamma)$. By part 2 of Definition 31, $\Box\phi \in \Gamma$. Hence $\Gamma \in |\Box\phi|_{\mathbf{L}}$. Conversely, suppose that $\Gamma \in |\Box\phi|_{\mathbf{L}}$. Then by definition of a truth set, $\Box\phi \in \Gamma$. Hence by part 2 of Definition 31, $|\phi|_{\mathbf{L}} \in N(\Gamma)$. By the induction hypothesis, $|\phi|_{\mathbf{L}} = (\phi)_{\mathbf{L}}$, hence $(\phi)^{\mathbb{M}} \in N(\Gamma)$. Hence $\Gamma \in (\Box\phi)^{\mathbb{M}}$.

□

Theorem 33 *The logic* **E** *is sound and strongly complete with respect to the class of all neighborhood frames.*

Proof The proof is standard and so will only be sketched. Soundness is straightforward (and in fact already shown in earlier exercises). As for strong completeness, we will show that every consistent set of formulas can be satisfied in some model. Before proving this, we briefly explain why this implies strong completeness. The proof is by contrapositions. Suppose that it is not the case that $\Gamma \vdash_{\mathbf{L}} \phi$. Then $\Gamma \cup \{\neg\phi\}$ is consistent. Since any such set has a model, there $\Gamma \cup \{\neg\phi\}$ is true in some model. But then Γ cannot semantically entail ϕ. Thus if $\Gamma \not\vdash_{\mathbf{L}} \phi$ then $\Gamma \not\models_{\mathsf{F}} \phi$ (where F is the class of all neighborhood frames).

Let Γ be a consistent set of formulas. By Lindenbaum's Lemma, there is a maximally consistent set Γ' such that $\Gamma \subseteq \Gamma'$. Then consider the model $\mathbb{M}^{min}_{\mathbf{E}}$. By the Truth Lemma (Lemma 32), $\mathbb{M}^{min}_{\mathbf{L}}, \Gamma' \models \Gamma'$. Thus Γ is true is some model, namely the minimal canonical model.

□

Notice that in the above proof, the choice to use the *minimal* canonical model for **E** was somewhat arbitrary. It is easy to see that the proof would go through if we had used $\mathbb{M}^{max}_{\mathbf{E}}$ instead of $\mathbb{M}^{min}_{\mathbf{E}}$. Indeed, *any* canonical model for **E** could have been used in the above proof. That there is such a choice of canonical models, will be of great use when proving completeness of systems above **E**. The strategy for proving strong completeness for systems above **E** is similar to the strategy for proving strong completeness of some well-known normal modal logics, such as **S4** or S5. Given the above construction and proof lemma, all that remains is to show that a particular canonical model belongs to the class of frames under consideration. This argument, called by [14] *completeness-via-canonicity*, can be adapted to the neighborhood setting. For example, consider the system **EC**. We argued in the previous section that **C** corresponds to the neighborhoods being closed under (finite) intersections. We now show that **EC** is sound and strongly complete with respect to neighborhood frames that are closed under intersections. We first show that C is canonical for this property (see [14] Chapter 4 for an extended discussion of this notion).

Lemma 34 *If $C \in \mathbf{L}$, then $\langle M_{\mathbf{L}}, N^{min}_{\mathbf{L}} \rangle$ is closed under finite intersections.*

Proof Suppose that $C \in \mathbf{L}$. Further, suppose that $X, Y \in N^{min}_{\mathbf{L}}(\Gamma)$. By definition of $N^{min}_{\mathbf{L}}$, $X = |\phi|_{\mathbf{L}}$ and $Y = |\psi|_{\mathbf{L}}$ where $\Box\phi \in \Gamma$ and $\Box\psi \in \Gamma$. Hence $\Box\phi \wedge \Box\psi \in \Gamma$ and so using C, $\Box(\phi \wedge \psi) \in \Gamma$. Hence, $|\phi \wedge \psi|_{\mathbf{L}} \in N^{min}_{\mathbf{L}}(\Gamma)$. Therefore, since $|\phi|_{\mathbf{L}} \cap |\psi|_{\mathbf{L}} = |\phi \wedge \psi|_{\mathbf{L}}$, $N^{min}_{\mathbf{L}}$ is closed under intersections.

□

Given the above proof, strong completeness is straightforward.

Theorem 35 *The logic* **EC** *is sound and strongly complete with respect to the class of neighborhood frames that are closed under intersections.*

Proof The proof is left as an exercise for the reader.

□

Exercise 9 *Prove that* **EN** *is sound and strongly complete with respect to neighborhood frames that contain the unit.*

Moving on to the logic **EM**, we see that the argument is not quite so straightforward. It is here where we make use of the fact that we have a choice of canonical models. The main roadblock is that $\langle M_{\mathbf{EM}}, N^{min}_{\mathbf{EM}}\rangle$ is not closed under supersets.

Observation 36 *Show that* $\langle M_{\mathbf{EM}}, N^{min}_{\mathbf{EM}}\rangle$ *is not closed under supersets.*

Proof Let p be a propositional variable and let Γ be a maximally consistent set such that $\Box p \in \Gamma$ (such a set exists by Lindenbaum's Lemma since $\Box p$ is consistent). Then $|p|_{\mathbf{EM}} \in N^{min}_{\mathbf{EM}}(\Gamma)$. Let Y be any non-proof set that extends $|p|_{\mathbf{EM}}$. To see that such a set exists, let Y' be any non-proof set (there are uncountably many subsets of $M_{\mathbf{EM}}$ but only countably many proof sets.) Then $Y = Y' \cup |p|_{\mathbf{EM}}$ is not a proof set. For if $Y = |\psi|_{\mathbf{EM}}$ for some formula ψ, then $Y' = |\psi \wedge \neg p|_{\mathbf{EM}}$ (why?), which contradicts the fact that Y' is a non-proof set. Clearly $Y \notin N^{min}_{\mathbf{EM}}(\Gamma)$ (why?). But then we have found a set X in $N^{min}_{\mathbf{EM}}(\Gamma)$ such there is a superset of X not contained in $N^{min}_{\mathbf{EM}}(\Gamma)$.

□

However, we can easily skirt this difficulty by choosing a different, better behaved, canonical models. Recall from Section 2, that if $\mathcal{F}$ is any collection of subsets of W, then $sup(\mathcal{F}) = \{X \mid \exists Y \in \mathcal{F} \text{ where } Y \subseteq X\}$. Given any frame $\mathbb{F} = \langle W, N\rangle$, let the **supplementation** of $\mathbb{F}$, denoted $sup(\mathbb{F})$, be the frame $\langle W, N^{sup}\rangle$, where for each $w \in W$, $N^{sup}(w) = sup(N(w))$. Use a similar definition for models, i.e., given $\mathbb{M} = \langle W, N, V\rangle$, the $sup(\mathbb{M}) = \langle W, N^{sup}, V\rangle$. The main argument is to show that the supplementation of the minimal canonical model is a canonical for **EM**.

Lemma 37 *Suppose that* $\mathbb{M} = sup(\mathbb{M}^{min}_{\mathbf{EM}})$. *Then* $\mathbb{M}$ *is canonical for* **EM**.

Proof Suppose that $\mathbb{M} = \langle W, N, V\rangle$, where $W = M_{\mathbf{EM}}$ and for each $\Gamma \in W$, $N(\Gamma) = sup(N^{min}_{\mathbf{EM}}(\Gamma)$, and $V = V_{\mathbf{EM}}$. Let $\Gamma \in W$ be arbitrary. We must show for each formula ϕ,

$$|\phi|_{\mathbf{EM}} \in N(\Gamma) \text{ iff } \Box\phi \in \Gamma$$

The right to left direction is trivial since $\mathcal{N}^{min}_{\mathbf{EM}}(\Gamma) \subseteq N(\Gamma)$. Suppose that $|\phi|_{\mathbf{EM}} \in N(\Gamma)$. Then there is some proof set X such that $X \subseteq |\phi|_{\mathbf{EM}}$. Since X is a proof set, there is some formula ψ such that $X = |\psi|_{\mathbf{EM}}$. Since $|\psi|_{\mathbf{EM}} \subseteq |\phi|_{\mathbf{EM}}$, by Lemma **??** $\vdash_{\mathbf{EM}} \phi \rightarrow \psi$. Since **EM** satisfies the RM rule, $\vdash_{\mathbf{EM}} \Box\phi \rightarrow \Box\psi$. Thus $\Box\phi \rightarrow \Box\psi \in \Gamma$. Hence $\Box\psi \in \Gamma$.

□

Theorem 38 *The logic* **EM** *is sound and strongly complete with respect to the class of supplemented frames.*

Proof Left as an exercise for the reader.

□

Putting everything together, we have a characterization of the smallest normal modal logic **K**.

Theorem 39 *The logic* **K** *is sound and strongly complete with respect to the class of filters.*

Exercise 10 *Prove that* **K** *is sound and strongly complete with respect to the class of augmented frames.*

4.3.2 General Neighborhood Frames

General frames are an important tool for modal logicians. There is no inherent difficulty in adapting the definition to the neighborhood setting.

Definition 40 *A* **general neighborhood frame** *is a tuple* $\mathbb{F}^g = \langle W, N, \mathcal{A}\rangle$*, where* W *is a non-empty set of states,* N *is a neighborhood function, and* $\mathcal{A}$ *is a collection of subsets of* W *closed under intersections, complements, and the* m_N *operator.*

We say a valuation $V : \mathsf{At} \to 2^W$ is admissible for a general frame $\langle W, N, \mathcal{A}\rangle$ if for each $p \in \mathsf{At}$, $V(p) \in \mathcal{A}$.

Definition 41 *Suppose that* $\mathbb{F}^g = \langle W, N, \mathcal{A}\rangle$ *is a general neighborhood frame. A general modal based on* $\mathbb{F}^g$ *is a tuple* $\mathbb{M}^g = \langle W, N, \mathcal{A}, V\rangle$ *where* V *is an admissible valuation.*

Truth in a general model is defined as in the beginning of this section.

Lemma 42 *Let* $\mathbb{M}^g$ *be an general neighborhood model. Then for each* $\phi \in \mathcal{L}$*,* $(\phi)^{\mathbb{M}^g} \in \mathcal{A}$*.*

Proof The proof follows from an easy induction over the structure of ϕ.

□

Given a logic $\mathbf{L}$, it is easy to show that the set $\mathcal{A}_{\mathbf{L}} = \{|\phi|_{\mathbf{L}} \mid \phi \in \mathcal{L}\}$ is a boolean algebra and closed under the m_N operator. A general frame is called a $\mathbf{L}$-frame, if $\mathbf{L}$ is valid on that frame. We show that for each modal logic $\mathbf{L}$ the canonical frame is a $\mathbf{L}$-frame.

Lemma 43 *Let* $\mathbf{L}$ *be any logic extending* $\mathbf{E}$*. Then the general canonical fram* $\mathbb{F}^g_{\mathbf{L}} \models \mathbf{L}$.

Proof Let $\phi \in \mathbf{L}$ and V an arbitrary admissible valuation. We must show that $\mathbb{M}^g = \langle M_{\mathbf{L}}, N, V\rangle$ validates ϕ. Since V is admissible, for each propositional letter p_i occurring in ϕ, $V(p_i) \in \mathcal{A}_{\mathbf{L}}$. Hence for each (there are only finitely many), p_i, $V(p_i) = |\psi_i|_{\mathbf{L}}$ for some formula ψ_i. Let ϕ' be ϕ where each p_i is replaced with ψ_i. We prove by induction of ϕ that $(\phi)^{\mathbb{M}^g} = (\phi')^{\mathbb{M}^g_{\mathbf{L}}}$.

The base case is when $\phi = p$. Then $\phi' = \psi$ for some $\psi \in \mathcal{L}$ where $V(p) = |\psi|_{\mathbf{L}} \in \mathcal{A}_{\mathbf{L}}$. Then $\Gamma \in (p)^{\mathbb{M}^g}$ iff $\Gamma \in V(p) = |\psi|_{\mathbf{L}}$ iff $\Gamma \in (p)^{\mathbb{M}^g_{\mathbf{L}}}$. The boolean connectives are straightforward. Suppose that ϕ is of the form $\Box\gamma$ and $(\gamma)^{\mathbb{M}^g} = (\gamma')^{\mathbb{M}^g_{\mathbf{L}}}$. Note that $\phi' = \Box\gamma'$. Hence $\Gamma \in (\phi)^{\mathbb{M}^g}$ iff $(\gamma)^{\mathbb{M}^g} \in N(\Gamma)$ iff $(\gamma')^{\mathbb{M}^g_{\mathbf{L}}} \in N(\Gamma)$ iff $\Gamma \in \phi'$.

□

4.4 Risky knowledge

Many consider that belief comes in degrees rather than being the 'all or nothing' binary notion we have studied so far. The probability calculus is the formal tool used to capture this idea. So, the doxastic state of the agent can be represented via a probability function and the binary notion of belief can be then derived from it by stipulating that a proposition X is believed if its probability is greater or equal that a certain established threshold (greater than .5). Moreover, some authors speak of probable knowledge or *risky knowledge* and characterize the notion in a similar way, like Henry Kyburg in varies writings (most recently in [41]).

Exercise 11 *If our operator* $\Box$ *is interpreted as 'it is highly probable that ...' which properties continue to be admissible? Construct a counterexample for the axiom (C):* $\Box(A) \wedge \Box(B) \rightarrow \Box(A \wedge B)$

It is clear that any augmented frame satisfies C above. Therefore it should be clear that epistemic models have a representational range that exceeds the range of relational models. The notion of risky knowledge can be represented via the former models but it cannot be represented via the latter.

4.5 The varieties of probabilism

There is an important historical reason motivating a widening gulf between traditional and formal epistemology. The reason in question can be traced back to the widespread influence of the use of probabilistic methods both in the social and physical sciences since at least the first quarter of the XX century.

Influenced in part by the writings of Rudolf Carnap, many epistemologists lost interest since the 1950s in the usual constructs of traditional epistemology like knowledge or even qualitative belief, and converted to some form of *probabilistic epistemology*. A particularly influential kind of probabilism of this sort was the *radical probabilism* proposed by Richard Jeffrey [32].

> 'Radical probabilism doesn't insists that probabilities be based on certainties; it can be probabilities all the way down, to the roots.' [...] Radical probabilism adds the 'non-foundational' thought that there is no bedrock of certainty underlying our probability judgments. ([33])

Jeffrey's garden variety of probabilism rejects the admission of a primitive notion of qualitative full belief or certainty. Other epistemological stances (which do appeal to personal probability) admit this primitive notion and then proceed to distribute probabilities over the field of possibilities determined by this basic attitude (where an event is possible if compatible with primitive certainties).

One of the central ideas of radical probabilism is that any notion of traditional epistemology, if legitimate, should be derivable probabilistically. And if most of these epistemological notions turn not to be derivable in this way, the worst for them. They should be abandoned as feeble illusions created by folk psychology. For example, radical probabilists tend to be rather hostile towards the notion of *acceptance*, which plays a crucial role in other exact epistemologies.

Bas van Fraassen's himself has defended a variant of probabilism in recent years, which is also radical (in the sense that only admits a unique probabilistic primitive) but that attempts to close the gulf separating traditional and formal epistemology.[3] The central idea is to take conditional (rather than monadic) probability as the main (and unique) epistemological primitive. The change is not a mild one. The price is to abandon the usual measure-theoretic presentation of probability inherited from the foundational work of Kolmogorov (and replacing it with a view that in many ways resembles the account defended by De Finetti, Savage and others).

[3] This work was carried out after the publication of [79]. One of the relevant articles is [80].

4.5.1 Two traditions in the study of conditional probability

There are at least two dominant traditions in the theory of conditional probability which are able to deal with conditioning events of measure zero. One is represented by Dubins' principle of Conditional Coherence [19]: For all pairs of events A and B such that $A \cap \mathrm{B} \neq \emptyset$:

(1) $Q(.) = P(.|A)$ is a finitely additive probability.
(2) $Q(A) = 1$, and
(3) $Q(.|B) = Q[B](.) = P(.|A \cap B)$

When $P(A \cap B) > 0$, Conditional Coherence captures some aspects of De Finetti's idea of conditional probability *given an event*, rather than *given a σ-field*.[4]

The well-known Kolmogovorian alternative to the former view operates as follows. Let $\langle \Omega, \mathcal{B}, P \rangle$ be a measure space where Ω is a set of points, $\mathcal{B}$ a σ-field of sets of subsets of Ω, with points w. Then when $P(A) > 0$, $A \in \mathcal{B}$, the conditional probability over $\mathcal{B}$ given A is defined by: $P(.|A) = P(. \cap A) \; / \; P(A)$. Of course, this does not provide guidance when $P(A) = 0$. For that the received view implements the following strategy. Let $\mathcal{A}$ be a sub-σ-field of $\mathcal{B}$. Then $P(.|\mathcal{A})$ is a *regular conditional distribution* [rcd] of $\mathcal{B}$, given $\mathcal{A}$ provided that:

(4) For each $w \in \Omega$, $P(.|\mathcal{A})(w)$ is a probability on $\mathcal{B}$.
(5) For each $B \in \mathrm{B}$, $P(B|\mathcal{A})(.)$ is an $\mathcal{A}$-measurable function.
(6) For each $A \in \mathrm{A}$, $P(A \cap B) = \int_A P(B|\mathcal{A})(w) dP(w)$

Kolmogorov illustrates with a version of the so-called 'Borel paradox', that $P(.|\mathcal{A})$ is probability not given an event, but given a σ-field. Blackwell and Dubins discuss in [19] conditions of *propriety* for rdcs. An rcd $P(.|\mathcal{A})(w)$ on $\mathcal{B}$ given $\mathcal{A}$, is proper at w, if $P(.|\mathcal{A})(w) = 1$, whenever $\mathrm{w} \in \mathrm{A} \in \mathcal{A}$. $P(.|\mathcal{A})(w)$ is improper otherwise. Recent research has shown that when $\mathcal{B}$ is countably generated, almost surely with respect to P, the rcd's on $\mathcal{B}$ given $\mathcal{A}$ are maximally improper [69]. This is so in two senses. On the one hand the set of points where propriety fails has measure 1 under P. On the other hand we have that $P(a(w)|\mathcal{A})(w) = 0$, when propriety requires that $P(a(w)|\mathcal{A})(w) = 1$.

[4]Conditional Coherence does not capture, nevertheless, important aspects of De Finett's ideas about primitive conditional probabilities. Unlike probabilists, De Finetti uses a primitively given notion of information in order to define conditional probability. Such notion of information set does not admit, in his construction, a probabilistic account. Some repercussions of this aspect of De Finetti's notion will be discussed below.

It seems that failures of propriety conspire against any reasonable epistemological understanding of probability of the type commonly used in various branches of mathematical economics, philosophy and computer science. To be sure finitely additive probability obeying Conditional Coherence is not free from foundational problems,[5] but, by clause 2 of Conditional Coherence, each coherent finitely additive probability is proper. In addition Dubins shows in [19] that each unconditional finitely additive probability carries a full set of coherent conditional probabilities.

In this section I shall only consider probabilities respecting propriety. So, I shall start with Conditional Coherence and I shall add the axiom of Countable Additivity only to restricted applications where the domain Ω, when infinite, is at most countable. Then I shall define qualitative belief from conditional probability by appealing to a procedure studied in [80], [9], [8]. The framework that thus arises is immune to standard Bayesian arguments against the use of inductive and abductive rules (of the type offered in [79]).

First we will add a resource in order to keep track on inconsistency as well as an intuitive constraint on conditional probability (compatible with Conditional Coherence):

(I) for any fixed A, the function $P(X|A)$ as a function of X is either a (finitely additive) probability measure, or has constant value 1.

(II) $P(B \cap C|A) = P(B|A)P(C|B \cap A)$ for all A, B, C in F.

The probability (*simpliciter*) of A, $pr(A)$, is $P(A|U)$. We will follow established terminology by referring to (II) as the *Multiplication Axiom*.[6] If $P(X|A)$ is a probability measure as a function of X, then A is *normal* and otherwise *abnormal*. Conditioning with abnormal events puts the agent in a state of incoherence represented by the function with constant value 1. Thus A is normal iff $P(\emptyset|A) = 0$. van Fraassen shows in [80] that supersets of normal sets are normal and that subsets of abnormal sets are abnormal. Assuming that the whole space is normal, abnormal sets have measure 0, though the converse need not hold (why?). In the following we shall confine ourselves to the case where the whole space U is normal.

We can now introduce the notion of *probability core*. I follow here ideas presented in [8], which, in turn, slightly modify the schema first proposed in [80].

[5] Perhaps the most poignant result is presented in [68], where it is shown that each finitely additive probability fails the property that De Finetti called *conglomerability* in some denumerable partition. See also [19].

[6] This axiom appears, under the name 'W. E. Johnson's product rule' in [34].

A *core* as a set K which is normal and satisfies the *strong superiority condition* (SSC) i.e. if A is a nonempty subset of K and B is disjoint from K, then $P(B|A \cup B) = 0$ (and so $P(A|A \cup B) = 1$). Thus any non-empty subset of K is more "believable" than any set disjoint from K. It can then be established that all non-empty subsets of a core are normal (Finesse).

Exercise 12 *Show that probability cores are nested.*

Second, it is possible to show that every countable additive function P is such that the chain of cores induced by it cannot contain an infinitely descending chain of cores. First we need a lemma which is left as an exercise:

Exercise 13 *Show that all non-empty subsets of a probabilistic cores are normal (Finesse).*

Lemma 44 *(Descending Chains). The chain of belief cores induced by a non-coreless 2-place function P cannot contain an infinitely descending chain of cores.*

Proof Assume by contradiction that there is a 2-place P, such that the chain of belief cores induced by it contains a core B_0 and a countable chain of cores B_0, B_1, B_2, ..., B_n, ..., such that B_0 is the outermost belief core of this subsystem of cores for P. Consider in addition a set of points x_j, with j in N, such that for every index i and j in N, $x_i \in B_i$, and $x_i \notin B_j$, if $j > i$.

For every index m, we have (by axiom A2 of belief cores) that

$$\mathrm{P}(\{x_{m+1}\} \mid \{x_{m+1}\} + \{x_m\}) = 1$$

By axiom A4 the 1-place function P(... $\mid \{x_{m+1}\} + \{x_m\}$) is a probability function obeying finite additivity. Therefore:

$$(1)\ \mathrm{P}(\{x_m\} \mid \{x_{m+1}\} + \{x_m\}) = 0$$

Now notice that the Multiplication Axiom guarantees that for all propositions A, B, C in the sigma-field F

$$\text{(E) } \mathrm{P}(A \mid A \cup B \cup C) = 0 \text{ if } \mathrm{P}(A \mid A \cup B) = 0.$$

To see that E is true it is enough to consider the following instance of the Multiplication Axiom:

$$\mathrm{P}((A \cup B) \cap A \mid A \cup B \cup C) = \mathrm{P}(A \cup B \mid A \cup B \cup C) \,.\, \mathrm{P}(A \mid (A \cup B) \cap (A \cup B \cup C))$$

Call C the countable set of points $\{x_0, x_1, ..., x_m,...\}$. Therefore (1) and (E) guarantee that for every x_m in U:

$$(1') \ P(\{x_m\} \mid C) = 0$$

Since $C \subseteq B_0$, A4 implies that P(... | C) is a 1-place probability function. By the Multiplication Axiom and the fact that P(... | C) is *countably additive* we have the desired contradiction: $1 = P(C \mid C) = \sum P(x_i \mid C) = 0$.

□

4.5.2 Belief, probability and paradox

What is full belief? One can say that it is our strongest doxastic attitude. Probabilists have considered the following (more specific) alternatives:

(1) Full belief should be identified with maximal personal probability[7]

(2) Full belief should be identified with probability one.

The second reductive strategy is popular en Economics (where it is usually taken for granted). But unfortunately both strategies suffer from problems first pointed out by Henry Kyburg in [40].

EXAMPLE 1 (Kyburg): Think about a lottery with so many tickets that, for any i, an agent considers maximally probable that ticket number i will not win. Then it seems that it would be reasonable for this agent to have the following expectations:[8]

(3) Each ticket is expected to lose.

(4) Some ticket wins.

[7]A terminological clarification: the rules of the form 'X fully believes the propositions that he deems maximally probable' are usually called *rules of acceptance*. This notion of acceptance bridging probability and qualitative attitudes should be distinguished from the notion used in the first part of this article. One can provide 'acceptability conditions' in the qualitative manner in which we provide them in section 1, without appealing at all to the notion of probability. In order to avoid confusions, we will not use the term 'acceptance rules' to refer to high probability rules – or any other rule of that sort, like (2). Our only use of the term 'acceptance' will be the qualitative one introduced in section 1.

[8]I remind the reader that I am using the term 'expectation' as a shorthand for 'practical certainty' or 'full belief'.

Now, in the presence of the following rule we have a contradiction

(Adjunction) If agent X fully believes A at t and X fully believes B at t, then X fully believes A and B at t.

For in the presence of Adjunction, (3) entitles us to expect that:

(5) Every ticket will lose.

So, finally a self-contradictory expectation is derived: 'some ticket wins and no ticket wins'.•

The only controversial step in the derivation of the paradox is (Adjunction), which is impossed in all relational models of knowledge and belief.

The second proposal suffers from what has been called a transfinite version of the lottery paradox (see [50]). The following example is a measure-theoretical version of the lottery paradox (that defeats (2)).

EXAMPLE 2. Say that agent a assumes that the weight of a stone is representable by a real number in some interval, say between .5 and 1 pounds. a is practically certain of the hypothesis that says that the weight lies somewhere in the interval [.5, 1]. In other words, a does not reject any of the hypothesis H_n: 'The weight is exactly x pounds, where x is a real value between .5 and 1'. I.e. a does not fully believe $\overline{H_n}$, for any x between .5 and 1'.[9] Yet a might coherently assign 0 probability to the uncountably many rival hypothesis of the form H_n.•

EXAMPLE 2 has been used by Patrick Maher (in [50], pages 388-89) in order to argue that probability one is not a sufficient condition for full belief.[10] This is so at least if one insists that doxastic representations should be consistently

[9]Here we are using informally some features of the logic of full belief developed in the former sections. Say that σ is the current 'belief set ' of a. Then the *rejection* of H_n is represented as: $\overline{H_n} \in \sigma$. Failure to reject H_n as $\overline{H_n} \notin \sigma$. But then (A2) gives us that $\neg B\overline{H_n} \in \sigma$. This argument is informally carried out in [50], page 387.

[10]Maher argued in his piece that probability 1 is not a sufficient condition for *acceptance*, not for full belief. Maher characterizes acceptance of a hypothesis H as the state of mind expressed by the sincere intentional assertion of H. Full belief, in turn, is characterized in terms of maximal personal probability. Then, he argues that probability one is neither a sufficient not a necessary condition for acceptance. Here we have not appealed to the cuasi-pragmatical characterization of acceptance used by Maher. But we are assuming as intuitive that an agent might fail to fully believe hypothesis to which he assigns maximal probability. One therefore needs a precise characterization of this intuitive notion. The following sections of the article will be devoted to study a proposal recently developed by Bas van Fraassen. A final historical clarification.

completable. Otherwise, it would not be unreasonable for a to give probability one to the hypotheis that one of the H_n is true and probability one to all the $\overline{H_n}$.

4.5.3 Probability Cores, Full Belief and Expectation

There are at least two main lines of response to loterry examples . One of them is based on a line of inquiry that I started various years ago and that is now the official response to lottery problems adopted by Henry Kyburg and collaborators. It is well known today (due to a famous argument by Kurt Godel) that, for example, intuitionistic logic, a paradigmatic example of a 'deviant' logic, is mappable to an extension of classical logic, namely the system S4. Similarly non-Adjunctive logics in general, and logics of high probability in particular, are mappable to classical modal logics weaker than the weakest normal system K. The connection with extensions of classical logic is here particularly appealing, for the added modality H is directly interpreted as 'It is highly likely that...'. So the principle that has to be given up is the axiom that has as antecedent H(A) and H(B) and as consequent H(A $\wedge$ B), rather than the classical principle of Adjunction which has A and B as antecedent and concludes (A $\wedge$ B). At the first order level the principle that corresponds to the lottery paradox is the so-called Barcan Schema (which is this case says that if it is highly likely that each ticket is a loser then it is highly likely that all are). Interestingly enough there are no Kripke-style models for a modality of this kind, but there are 'neighborhood' models of the sort first proposed by Dana Scott and Richard Montague (why?). So, one line of response here is that lottery and preface problems can be dealt with by appealing to more deductive logic (i.e. an extension of classical logic) rather than less (i.e. a deviant logic with respect to classical logic). A second line of response consists in defining full belief explicitly from conditional probability in such a way that it obeys closure requirements:

When the universe of points is at most countable, very nice properties of cores and conditional measures hold, which can be used to define full belief and expectation in a paradox free manner. In general it can be shown that for each function P there is a smallest as well as Ca largest core and that the smallest core has measure 1. In addition, when the universe is countable we can add Countable Additivity without risking failures of propriety. In this case we have that the

van Fraassen has argued that full acceptance does not entail full belief. But, again, here full belief is the notion characterized in terms of maximal personal probability. In the new model developed by van Fraassen full belief no longer will be characterized in this manner, and in this new context the eventual remaining differences between full acceptance and full belief seem to be minimal.

smallest core is constituted exactly by the points carrying positive probability [8].

All cores carry probability one, but, of course, only the innermost core lacks subsets of zero measure. There is, in addition, a striking bdifference between the largest and the smallest core (and between the blargest and any other core). In fact, any set S containing the largest core is robust with respect to suppositions in the sense that, $P(S \mid X) = 1$ for all X, and the complement of S is abnormal. So, the largest core encodes a strong doxastic notion of certainty or *full belief*, while the smallest encodes a weaker notion of 'almost certainty'. Both formal and conceptual reasons are provided in [8] for understanding the weaker notion as a notion of qualitative *expectation*.[11] So, when the universe is countable and countable additivity is imposed we can define two main attitudes as follows: *an event A is expected if it contains the smallest core, whereas it is fully believed if it contains the largest.*

In the general case there is still enough structure to define both attitudes. In fact, in this case the existence of the innermost core cannot be guaranteed. But the definition of full belief needs no modification and the notion of expectation can be characterized as follows: *an event A is expected if it is entailed by some core.*

In the following sections I'll show how the notions just presented can be used in order to provide a model of abductive operators. The following section is devoted to an axiomatic characterization of such operators.

Exercise 14 *Show via examples how the former tools can be used to dissolve both finite and trans-finite cases of the lottery paradox.*

In spite of its elegance the previous solution is not a compelling one in view of problems that arise when iterated update is considered. This point will be discussed in class.

4.5.4 Beyond probabilism and the prospects for unification

Extending the set of admissible epistemological notions can lift the strong constraints imposed by radical probabilism on admissible forms of belief change. The sole introduction of a qualitative form of full belief (as an additional epistemological primitive alongside probability) gives us enough degrees of freedom for

[11] The term 'expectation' is not decision theoretically motivated. Its motivation comes from the field of non-monotonic logic, where 'expectation' models of defeasible reasoning are usual – see [9] for details.

enriching considerably the previous account. Of course this is tantamount to abandon one of the central ideas of radical probabilism. Now a purely qualitative notion (familiar in traditional epistemology) is assumed primitively without probabilistic foundations. Still the resulting account can use probabilities as a way of encoding degrees of belief. So, this is a way of *supplementing* rather than *replacing* traditional epistemology.

The basic idea is that full belief determines what Levi calls the standard for serious possibility [43]. In other words a set of *possibilities* is determined primitively as the set of events compatible with full belief. Probability is then distributed over this set of possible events. But this body of full belief can change. It can be *contracted*, by giving up held beliefs; or it can be *expanded*, by adding beliefs compatible with held ones (and taking the corresponding logical closure), or it can be *revised*, when the input contradicts background knowledge and one wants to shift from the initial view to a new consistent view including the incoming piece of information.

The idea of adopting a notion of full belief or certainty as a primitive on which the very characterization of probability (especially conditional probability) depends appears clearly articulated already by B. De Finetti in [18].[12]More recently Isaac Levi adopted it as one of the central pillars of his epistemology [43].

Notice that now probability cannot help us determining the form that such changes have to adopt. In unified probabilism full belief supervenes to an underlying conditional probability function and changes of this derived notion of full belief supervene as well on the standard method for updating the underlying probability functions. But once the connection between full belief and probability is severed and both (conditional) probability and (full) belief have equal status as epistemological primitives, changes of qualitative belief cannot be derived from probability kinematics.

What can then guide us in determining what counts as a rational change of view? There is no unique response to this question. Many criteria have been

[12]The following passage is revealing and perhaps surprising for readers not familiar with the foundational ideas of De Finetti:

> In almost all circumstances, and at all times, we find ourselves in a state of uncertainty. Uncertainty in every sense. [...] It would therefore seem natural that the customary modes of thinking, reasoning and deciding should hinge explicitly and systematically on the factor uncertainty as the conceptually pre-eminent and determinative element. The opposite happens however: there is no lack of expressions referring to uncertainty, but it seems that these expressions, by and large, are no more than verbal padding. The solid, serious, effective and essential part of arguments, on the other hand, would be the nucleus that can be brought wibthin the language of certainty - of what is certainly true or certainly false. It is in this ambit that our faculty of reasoning is exercised, habitually, intuitively and often unconsciously ([18]. p. 24).

Immediately, De Finetti makes clear that his set of certainties contains more than mere tautologies, and that its main role is to determine a space of possibilities:

> Thinking of a subset of truths as given (knowing, for instance, that certain facts are true, certain quantities have given values, or values between certain limits, certain shapes, bodies or graphs of given phenomena enjoy certain properties, and so on), we will be able to ascertain which conclusions, among those of interests, will turn to be - on the basis of the data - either certain (certainly true), or impossible (certainly false), or else possible ([18], p. 25).

What about probability? According to De Finetti ”probability is something that can be distributed over the field of possibility:”

> Using a visual image, which at a later stage could be taken as an actual representation, we could say that the logic of certainty reveals to us a space in which the range of possibilities is seen in outline, whereas the logic of the probable will fill in this blank outline by considering a mass distributed upon it ([18]. p. 25).

applied, from symmetry and the use of computational constraints to reflective equilibrium, as well as various attempts to ground a theory of belief change into broader frameworks from rational choice to learning theory. These methodologies need not produce convergent accounts and, in fact, what we have today is a variety of standards for belief change, each one fully axiomatized and grounded in the use of the aforementioned methodological frameworks.

Far from constituting an embarrassment, this diversity of standards for belief change is perhaps something that one should expect. Laws constraining the rationality of belief change are likely to be heavily dependent on the general epistemic values guiding inquiry. For example, for a learning theorist a central goal of inquiry is to learn the (whole) truth about the world in the long run. A perspective more influenced by Bayesian methodology will insist in regimenting what counts as rational in the next step of inquiry or in the proximity of that change. A radical form of this Bayesian view is Levi's idea that inquiry has to be 'myopic'. It has to care exactly about the next step of inquiry. So, the complete determination of an axiom system for belief change is likely to be relative to the exact specification of the epistemic values that guide inquiry. Many leading philosophers have been recently more prone to recognize this dependency and to make it explicit the important role played by epistemic values in contemporary models of inquiry.[13]

Other aspects of the model are also quite crucial. For example, should one consider changes of view as cognitive decisions? Obviously some changes of view, like the ones triggered by perception, should not be modeled this way. But other changes of view could be modeled decision-theoretically, like the changes implemented in a legal code (promulgation or derogations of laws in logically closed codes) or in a scientific theory which is revised in the face of new evidence. And this can also be the case when one considers everyday decisions, from small economic choices to managerial decisions to various forms of planning, etc. Unfortunately the dominant tradition in belief change has not modeled explicitly changes

[13] Two recent books can be mentioned here. One is a recent book by van Fraassen on empiricism [81] where he deploys a model of scientific change, and in particular of scientific revolutions, inspired by the early pragmatist philosophy of W. James. The book sketches a form of cognitive decision theory compatible with the instrumentalism that van Fraassen has defended for decades. Hillary Putnam has also written a recent book [58] elaborating on the role of epistemic values in inquiry. The book is also influenced by early American pragmatism (through the influence of Dewey) and its main focus is to deflate the fact-value dichotomy inherited from the writings of David Hume. Putnam explicitly proposes to treat notions like epistemic coherence or simplicity as epistemic values, and devotes various chapters to consider the problem of value indeterminacy. Much of these chapters are also heavily influenced by the recent work of Sen on rationality.

of view as decisions. This has not helped to make clear the important role played by epistemic values in inquiry. One of the few exceptions to this pattern is the work of Isaac Levi in a series of writings going back at lest to [43]. His account of contraction has been revised and perfected various times in subsequent books. The most recent model can be found in [48]. One of the central ideas of the model is that what is minimized in contracting a held belief is the loss of information value, rather than just the loss of information. Both in a recent joint paper [12] as well as in [48] three axioms are proposed in order to characterize a value function V encoding the relevant notion of information value.

Definition 45 *Let $S(K, A)$ be the family of A-saturatable sets of K. I.e. if K is a theory, $X \in S(K, A)$ if and only if $X \subseteq K$, X is closed, and $Cn(X \cup \{\neg A\})$ is a maximal and consistent set.*

Then one can define a contraction operator as follows:

Definition 46 *$\div$ is a Levi-contraction of a theory K if and only if there exists a choice function G for K such that for all sentences A: if $A \in K$, then $K \div A = \bigcap G(S(K, A)) = \{X \in S(K, A)$: $V(X) \leq V(Y)$ for all $Y \in S(K, A)\}$, and if $A \notin K$, $K \div A = K$.*

The formal properties of a contraction of this form, when the value function obeys adequate postulates for information value are studied in [12]. This account of contraction diverges in many ways from other standard accounts. First it focuses on the feasible set $S(K, A)$ when considering an A-contraction of a theory K. This is a larger feasible set of options than the one usually considered to compute an A-contraction of K. Usually researchers focus instead on the set of maximal subsets of K that fail to entail A. Obviously this set should be included in $S(K, A)$, but the inclusion can perfectly well be proper. In addition the use of a value function makes possible to discuss explicitly the role played by epistemic values in inquiry.

This model abstracts away from the possibility of value indeterminacy. It also presupposes that different dimensions of epistemic value (from coherence to simplicity and more) are integrated in a single value function inducing a weak order on the feasible set of epistemic options. Nevertheless if each dimension of value induces by itself a weak ordering of epistemic options a preference relation integrating these different dimensions will not be complete. By the same token value indeterminacy will also induce the same kind of incompleteness. This leads us to consider a very important dimension of rationality. Rationality is usually equated to some form of maximization process. But there is a crucial difference between strict maximization (sometimes called optimization, where the idea is to choose the best feasible option) and liberal maximization (where the idea is to select un-dominated feasible options). While liberal maximization is able to tolerate incompleteness, strict maximization is not. We will return to this problem in the second part of these notes devoted to build decision-theoretic models of belief change.

4.6 Fixed point characterization of full belief

Knowledge understood as a source of knowledge and knowledge as evidence or *standard of serious possibility* [43] are not one and the same. If the emphasis is on the second aspect of knowledge then we are interested in the reasoning capabilities of an agent. Interest on an agent as an authoritative source of knowledge leads to a different stance of the type we have explored so far where the game are knowledge attributions rather than the study of knowledge claims.

How knowledge and full belief should be encoded when knowledge is understood as a standard of serious possibility?

Acceptance as true and *full belief* are closely related epistemic attitudes. Commitment to full belief is mirrored by commitment to accept as true; and commitment not to accept is mirrored by commitment not to fully believe. These ideas can be rendered formally by appealing to a regimented language L containing a modal operator B. In order to use a uniform notation we can focus on the set of *consistent theories* σ of L obeying the following two constraints:

(A1) A $\in \sigma$ iff B(A) $\in \sigma$
(A2) A $\notin \sigma$ iff $\neg$ B(A) $\in \sigma$

σ can be seen in this context as a *commitment set* representing the doxastic commitments of an agent at certain instant t. One can say that an agent explicitly believes a finite set of sentences M, but that he is doxastically committed to the closure of M – under the classical notion of logical consequence. Then membership in σ represents commitment to accept and lack of membership in σ represents commitment no to accept.

The account of full belief summarized above was first articulated by Isaac Levi in an essay written in reponse to Hintikka's account of knowledge and belief operators [42]. The account in question was later refined in [46] and [47] as well as in [6].[14]

[14]The presentation here mainly follows the one I used in the first part of [4].

The main intuition in autoepistemic logic is as follows: one starts with a *premise set*, which typically is not closed. Then the idea is to capture the final states that an agent might reach by reflecting on his beliefs and by making inferences from them and about them. As Stalnaker explains in [76] these final states must meet two intuitive conditions: first they must be stable, in the sense that no further conclusions can be drawn from them; and secondly they should be *grounded* in the initial (usually not closed) premise set. A theory T over the language L is *stable* as long as it is deductively closed and it is also closed under A1-2. *Groundedness* is usually captured via a syntactic definition of a fixed point.

Most of the work in autoepistemic logic has consisted in specifying useful formal definitions of groundedness. More recently a possible world semantics has been developed (see [53] for details and [21] for the first completeness result using possible worlds semantics). For the view presented here the notion of groundedness is unimportant and the space of all stable theories plays a different role than the one that they play in eutoepistemic theories. We are interested in formulating *acceptance* rules for full belief, not in specifying truth conditions for (an eventually indexical) epistemic operator. Therefore we prefer to develop a purely syntactical modelling.

According to the view defended here the stable theories represent the *informational commitments* of the agent rather than the information explicitly available for the purposes of reasoning and acting. A thorough presentation of the importance of adopting the previous interpretation of stable theories rather than the one prevalent in autoepistemic logic is presented in [7]. I will only summarize some of the features of the view of saturated theories defended here. *First*, we characterize an operator of certainty or full belief (via acceptability conditions), not a doxastic notion used to represent the uncertainty of the agent. *Second*, we are tacitly assuming that both acceptance and belief carry commitments. *Thirdly* we focus on the linguistic description of *mental acts*. We are not giving an account of *speech acts* or an account of *doxastic reports*.[15]

An *information model* is a set of theories of L closed under A1-2. In order to distinguish the view proposed here from the more dynamical one used in autoepistemic logic we will call such theories *saturated*. They can be seen as the linguistic representation of the commitments carried by the set of feasible states of information of a rational agent at a given instant. A1 and A2 are now seen as

[15]The last distinction has several formal and conceptual consequences. For example, one can define the following notion of consequence: Γ entails A iff A is believed in any stable theory grounded on Γ. Stalnaker proposed in [76] to use such notion of consequence to formalize the (Gricean) notion of (generalized) conversational implicature. A criticism of this idea is presented in [**?**] where a different view of the former notion of entailment is offered.

acceptance rules for full belief. One can then ask which are the modal formulae that are accepted in all saturated theories of a given model. Following the notation used in [46], [5] and [6], as well as [47], we can call such formulae *positively valid in a model*, and *positively valid* when positively valid in all models. Let Th(S5) denote the set of theorems of the system S5 of modal logic defined as follows:

We will use the meta-letters A, B, C,... to refer to formulas of L. S5 consists of the following axioms (1) All tautologies, (2) $(B(\mathrm{A}) \wedge B(\mathrm{A} \rightarrow \mathrm{B})) \rightarrow B(\mathrm{A})$, (3) the alethic axiom: $B(\mathrm{A}) \rightarrow \mathrm{A}$, (4) the positive introspection axiom: $B(\mathrm{A}) \rightarrow B(B(\mathrm{A}))$, (5) the negative introspection axiom: $\neg B(\mathrm{A}) \rightarrow B(\neg B(\mathrm{A}))$. The rules of inference are modus ponens and the rule that guarantees that $B(\mathrm{A})$ is a theorem of S5 whenever A is a theorem of S5.

Exercise 15 *Prove that all axioms of S5 are positively valid.*

It is easy to see that this account leads to a very specific logical theory of what is the logic of full belief. Weakening some of the fixed point equations can provide weaker theories but then one needs philosophical justifications for such weaker theories.

Exercise 16 *Call* negatively valid *the modal formulas whose negations are not accepted in any saturated theory. Are all positively valid theses also negatively valid?*

Theorem 47 *All positively valid theses are exaclty axioms of S5.*

Proof We need to show that a formula A of L is positively valid if and only if A is a theorem of S5. From left to right the proof is not demanding. From right to left the proof is slightly harder.

We will assume that α is not a S5 theorem and we will show that α does not belong to some saturated theory T. Take K = S5 $\cup \neg B(\alpha)$. K is consistent. For suppose that K were inconsistent. Then, by axiom (3), α is an S5 theorem.

We form now a list **L** of all formulae of L: g_1, g_2,..., g_n,... With respect to this list, we construct an infinite sequence of sets

I_0, I_1,..., I_n,...

as follows. As I_0 we take K, i.e.,

$I_0 = \mathrm{K}$

Then, for each positive integer n we set:

$$I_{i+1} = \begin{cases} Cn(I_i \cup \{g_{i+1}\}) & \text{if } I_i, g_{i+1} \text{ is consistent} \\ I_i & \text{otherwise} \end{cases}$$

Form, then, a Lindenbaum-type set: $\overline{K}$ = C($\cup I_i$), where $\cup I_i$ denotes the union of all the infinitely many sets I_i. $\overline{K}$ is a complete S5-theory. Nevertheless $\overline{K}$ is not saturated. It is interesting to see why not. Consider any formula A added at stage i. At stage i + j one could consistently add $\neg B$(A) - if **L** is such that B(A) has not been added yet. This is so because the formula A $\wedge$ $\neg B$(A), although epistemically problematic, is logically consistent. To put it in a different way, the modal system S5 is unable to capture syntactically the paradoxical nature of Moore's paradox ('this is my hand, but I do not fully believe it').[16]

But the formula A $\wedge$ $\neg B$(A) does not belong to any consistent saturated theory. For say (by contradiction) that A $\wedge$ $\neg B$(A) is in a consistent saturated theory T. Then $\neg B$(A) $\in$ T and hence (by A2) A $\notin$ T which leads to an immediate contradiction. So, the standard Henkin method used in modal logic to prove completeness does not suffice. We propose the following *Moore saturation*[17] of K:

$$\mathcal{M} = \{\text{A: } B(\text{A}) \in \overline{K}\ \}$$

$\mathcal{M}$ is consistent. For suppose by contradiction that $\perp \in \mathcal{M}$. Then, $B(\perp) \in \overline{K}$. But since S5 $\subseteq \overline{K}$, $\neg B(\perp) \in \overline{K}$. Therefore we have a contradiction because it is easy to check that $\overline{K}$ is consistent. $\mathcal{M}$ is a S5 theory. It is easy to see that all S5 theorems are in $\mathcal{M}$. In fact, B(A) is a S5-theorem whenever A is a S5-theorem and, by construction, all S5-theorems are in K . In general, assume that A entails B. Then B(A $\rightarrow$ B) is in $\overline{K}$. Assume that A is in $\mathcal{M}$. Then B(A) $\in \overline{K}$. Therefore B(B) $\overline{K}$ – $\overline{K}$ is a complete S5-theory – and this guarantees that B is in $\mathcal{M}$. Assume that A, B are in $\mathcal{M}$. Then B(A), B(B) are in $\overline{K}$. Now the formula B(A) $\wedge$ B(B) $\rightarrow$ B(A $\wedge$ B) is a S5-theorem, and therefore A $\wedge$ B is in $\mathcal{M}$.

[16]The philosopher G. E. Moore offered different variants of his paradox. A cannonical version appears in [52].

[17]Notice that the Moore saturation of a set K does not need to be an *extension* of $\overline{K}$ – although it is constructed as a function of $\overline{K}$. For consider the case where the initial K is paradoxical (in the sense of G.E. Moore), i.e. it contains both A and $\neg B$(A). The Lindenbaum extension of K will contain both A and $\neg B$(A), but the saturation of K will only contain $\neg B$(A) - and it will *not* contain A.

We will check now that $\mathcal{M}$ is closed under A1 and A2. First we will check A1. Assume that $B(\mathrm{A}) \in \mathcal{M}$. Assume now by contradiction that $\mathrm{A} \notin \mathcal{M}$. Then $B(\mathrm{A}) \notin \overline{K}$, and since $\overline{K}$ is a complete S5 theory, $\neg B(\mathrm{A}) \in \overline{K}$. Therefore, by negative introspection, $B(\neg B(\mathrm{A})) \in \overline{K}$. This, in turn, entails that $\neg B(\mathrm{A}) \in \mathcal{M}$, against the consistency of $\mathcal{M}$. Assume now that $\mathrm{A} \in \mathcal{M}$. Then $B(\mathrm{A}) \in \overline{K}$. By positive introspection, $B(B(\mathrm{A})) \in \overline{K}$. This yields the desired result, namely that $B(\mathrm{A}) \in \mathcal{M}$.

Secondly we should check A2. Assume that $\mathrm{A} \notin \mathcal{M}$. Then $B(\mathrm{A}) \notin \overline{K}$, and since $\overline{K}$ is a complete S5 theory, $\neg B(\mathrm{A}) \in \overline{K}$. Therefore, by negative introspection, $B(\neg B(\mathrm{A})) \in \overline{K}$. This, in turn, entails that $\neg B(\mathrm{A}) \in \mathcal{M}$, as desired. Finally assume that $\neg B(\mathrm{A}) \in \mathcal{M}$. Assume by contradiction that $\mathrm{A} \in \mathcal{M}$. Then $B(\mathrm{A}) \in \overline{K}$. Positive introspection guarantees that $B(B(\mathrm{A})) \in \overline{K}$. But then $B(\mathrm{A}) \in \mathcal{M}$, against the consistency of $\mathcal{M}$.

So, $\mathcal{M}$ is a consistent and saturated S5 theory. But, by construction, $\neg B(\alpha) \in \mathcal{M}$. In fact, $\neg B(\alpha) \in \mathrm{K} \subseteq \overline{K}$. Therefore, by negative introspection, $B(\neg B(\mathrm{A})) \in \overline{K}$, which is enough to guarantee that $\neg B(\alpha) \in \mathcal{M}$. But now, by A2, $\alpha \notin \mathcal{M}$. This completes the proof.

□

Exercise 17 *Consider Moore's paradox 'This is my hand but I do not believe it'. How will you represent it in this framework? Is it consistent statable in any consistent saturated theory?*

4.7 Other theories about the structure of epistemic states

The previous accounts reviewed different ways of representing knowledge via epistemic operators. In one case (the last section) knowledge is viewed as a standard for serious possibility. In another (the previous sections) the agent is seen externally as a potential informant or as a source of knowledge. The accounts are different both conceptually and logically and we also saw that as long as one sees knowledge as a source of knowledge (and as knowledge of propositions rather tahn sentences) there are two ways of dealing with epistemic and doxastic operators. One is related to the seminal work by Hintikka, Kripke and others and another (more comprehensive) stance derives from initial work from Scott and Montague. And there is still a third view where knowledge is seen as a standard for serious possibility, deriving in this case initially from the work of Levi and later on from work by Moore, Stalnaker and others. We also mentioned the view that intends to unify belief and probability by seeing belief as a particular case of probabilistic belief.

But these accounts are hardly the only ones that one can use to represent epistemic states. There are indeed many more. One of them are Spohn's *ranking systems* [74], [75].

A ranking function κ is a function from $\mathcal{M}$ to the set of extended non-negative integers $\mathcal{N}^+ = \mathcal{N} \cup \{\infty\}$, such that $\kappa(w) = 0$, for some $w \in \mathcal{M}$. For each proposition $P \subseteq \mathcal{M}$ the *rank* $\kappa(P)$ of P is defined by $\kappa(P) = \min \{\kappa(w)\colon w \in P\}$ and $\kappa(\emptyset) = \{\infty\}$.

Spohn proposes to interpret ranks as *grades of disbelief.* $\kappa(P) = 0$ says that P is not disbelieved at all. It does not say that P is believed; this is rather expressed by $\kappa(P^c) > 0$, i.e., that non-P is disbelieved (to some degree). The set $C_\kappa = \{w\colon \kappa(w) = 0\}$ is called the *core* of κ and C_κ is the strongest proposition believed (to be true) in κ.

This account has already a dynamic flavor absent in the previously reviewed views. So, if A^c is believed to be true in κ, one way of representing the *contraction* of A^c from C_κ is to take the union of C_κ with the set of least disbelieved $\neg A$ points, i.e. $\{w\colon \kappa(w) = \kappa(\neg A)\}$. We will pay detailed attention to this process of contraction in the coming sections, devoted to belief change.

5 Changing view as a decision problem: Part I

Students of belief change have recognized the need to give some sort of account of how to *contract* a belief state represented by a deductively closed theory or corpus K to a deductively closed subset K' when some specific sentence is to be removed. Contraction from K removing A may be achieved in many ways so that the inquirer is called upon to make a decision. There is a broad and often deceptive unanimity that the choice made should "Keep loss at a minimum" as the *Principle of Economy* as formulated by Rott and Pagnucco ([65], 502) stipulates. The unanimity unravels when the issue of evaluating "loss" is addressed. Nonetheless, the principle of economy recommends optimizing by minimizing loss of something of value.

An alternative type of principle for recommending a contraction removing A urges retaining all and only those elements of K that are better *entrenched* than A. Such a *Principle of Entrenchment* is formulated as a satisficing principle rather than a maximizing (or minimizing) one. Instead of urging the choice of a "best" contraction, one may urge the choice of a contraction consisting of sentences in K that are "good enough" to be retained. Once more there has been controversy concerning how entrenchment is to be evaluated.[18]

[18] Levi ([43], [45]) uses a notion of *degrees of incorrigibility* rather than degrees of entrench-

To obtain a viable account of contraction based on both principles as authors like Gardenfors [22] sought to do, not only must the methods of evaluating loss and of assessing entrenchment be specified but they must be specified in a manner that assures that the maximizing Principle of Economy and the satisficing Principle of Entrenchment recommend the same contraction.

In this discussion, we shall focus on the Principle of Economy. The Principle of Entrenchment will receive brief attention towards the end of this section.

The Principle of Economy is a schema. The loss to be minimized needs to be specified. The approach to contraction introduced in the classic paper by Alchourrón, Gardenfors and Makinson (AGM, [3] considered loss of information incurred in contraction. Potential state K_2 carries more information than K_1 if and only if the set of sentences in K_1 is a subset of the set of sentences in K_2 so that K_2 can be said to be logically stronger than K_1. Hence if two contractions K_1 and K_2 of K removing A are compared with respect to the loss of information incurred where K_2 is more informative than K_1, the loss incurred by shifting from K to K_2 is clearly less than that incurred by shifting from K to K_1.

When the Principle of Economy is construed as recommending the minimization of loss of information in this subset sense, it has become known as a *Principle of Conservatism* as in Harman, [29].

Conservatism played, without doubt, a motivating role in the early stages of research in the AGM tradition. It led to some precise formulations of contraction, like the so-called *maxichoice contraction* (Gärdenfors, [22]) and it paved the way towards the more sophisticated account of contraction defended in (Alchourrón et al., [3]): *partial meet contraction.*

Maxichoice contractions propose to achieve the contraction of a theory by a sentence A by selecting some maximal subset of K that does not imply A. This account is directly motivated by Conservatism, but it produces an unintuitive account of revision. In fact, if we denote the (maxichoice) contraction of K with $\neg A$, by $K/\neg A$, it seems reasonable to represent the revision of K with A as the logical closure of the set $\{K/\neg A \cup \{A\}\}$. But then all revisions of theories will be represented by maximal and consistent theories, an undesirable result.

AGM departed from the Principle of Conservatism by rejecting the recom-

ment. According to this view, the recommended contraction of K removing A is the set of incorrigible sentences in K which in turn is the set of sentences in K better entrenched than A. Gardenfors [22] uses a different condition relating contraction with entrenchment; but he presupposes the controversial Recovery Postulate (Recovery stipulates that removing an item of information from K by contraction and then adding it back to the contracted belief state yields the initial point of view K.). In the presence of the Recovery Postulate, his approach corresponds to the account of contraction given in terms of incorregibility.

mendation of maxichoice contractions as mandatory in all cases. The central idea in (AGM, [3]) was to make a selection of the 'best' elements in the set of all maximal non-A-implying subsets of K; and then take the intersection of this selection. This is what is usually called a *partial-meet contraction*. Recommending partial meet contraction is generally paradigmatic of the AGM approach.

It is clear that partial-meet contractions do not follow Conservativism. As Rott and Pagnucco ([65], 503) have recently observed: "The Principle of Economy has been severely compromised in the AGM framework." In a more recent article, Rott called the principles of Economy and Entrenchment "dogmas" "...*not* because almost all researchers kept to these principles (quite the opposite is true) but because so many authoritative voices *proclaimed* them to be the philosophical or methodological rationale for their theories [62].

As a matter of fact, the philosophical motivation for the AGM approach remains unclear. Attempts to clarify the main guiding principles allegedly used in formulating AGM contraction have ended up in many cases in the proposal of deviant notions ([65] is a perfect example of this kind of attempt). Perhaps the influence of the AGM approach is due to the fact that the AGM trio were pioneers in providing exact axiomatic formulations of their proposals. And in formal epistemology, like in other fields, the use of the axiomatic method promotes progress by systematization of ideas. But the mere use of the axiomatic method does not guarantee conceptual clarity.

Usually the AGM approach has been criticized in a piecemeal fashion, by pointing out to the eventual lack of intuitiveness of some of the postulates it proposes. For example, the so-called axiom of *recovery* – presented below – has been copiously criticized, in part by the members of AGM. But this process of criticism itself requires a certain initial level of conceptual clarity as to what contraction is supposed to achieve. The conceptual soundness of particular axioms can only be judged in terms of a set of basic principles regulating contractions. We want to propose in this section that the two principles formulated above are indeed these guiding principles. The principle of Economy is basically correct, except for the fact that the main idea in contraction is not to minimize losses of information, but to minimize *informational value* [43]. As we shall see later, the troubles encountered by the principle of entrenchment are removed by this suggestion as well.

The principle of Cognitive Economy
Keep loss of informational value to a minimum in contraction.

We propose to use this instance of the principle of Economy together with the principle of Entrenchment as the main foundational guidance for articulating the notion of contraction. Some questions seem pertinent even at this early stage in the analysis. For example, what is the connection between information measured by set inclusion and informational value? Let V be a real value index on theories. Then we will appeal to the following principle:

Weak Monotony
If $X \subset Y$, then $V(X) \leq V(Y)$

In this discussion, we consider comparisons of theories with respect to informational value that constitute a weak ordering of the set of theories while satisfying Weak Monotony. According to this principle a theory Y can be a strict superset of another theory X, i.e. a theory Y can carry strictly more information than another theory X, but the informational value of the two theories can be the same. The extra information might not matter. Of course, one gets more specialized instances of the principle of Cognitive Economy by telling a more detailed story about informational value. Here we will propose a specific notion of informational value and tackle the problem of axiomatizing the resulting notion.

The idea of using informational value as the central notion in belief change is not new. Isaac Levi proposed it in his early writings [43] and he has refined it progressively in more recent writings [45], [48]. The axiomatic account of AGM clearly does not fit this approach. Our task here is to go from first principles to axiomatization, by finding the axioms that *completely* characterize the contractions obeying appropriate instances of the principles of Economy and Entrenchment.

Of course these axioms will not coincide with axioms for AGM contraction or with other versions of contraction that have been proposed. We will focus on a salient notion of informational value, but our account will be parametrical. We will therefore point to other permissible notions of informational value and to their respective axiomatic encodings.

We agree with Rott (and Pagnucco) that many authors in the field have proclaimed allegiances to instances of the two principles while at the same time developing theories that do not obey those principles. The AGM tradition is only one instance of this mismatch between foundations and axiomatic proposals. We, nevertheless, do not agree that the principles should be abandoned as dogmas. We think that, when appropriately formulated, they are sound. The main task of this section is to characterize these sound principles by a complete set of axioms.

We will proceed as follows. First we will define an operator $\div$ of informational value encoding decision theoretically the principle of Cognitive Economy. That is to say $K \div A$ is a theory removing A from K with minimum loss of informational value. Then we will propose a set of axioms characterizing what we call *mild contractions* [48]. We will prove both soundness and completeness for these postulates. I.e. we will show that $\div$ obeys the postulates of mild contractions; and we will show that any mild contraction operator can be represented as an operator of informational value.

Once this is done we will show that mild contractions fit the principle of Entrenchment, while AGM contractions do not. We will conclude by comparing our account to other foundational accounts, especially the one presented in [65] which shares our axiomatic base - but whose foundations depart from our proposal.

5.1 Operators of informational value

We understand the problem of how to contract by removing A from a theory K to be a decision problem where one is called upon to choose a contraction removing A from K *from among all the contraction strategies removing A from K available in the context*.

Let K be a theory (representing the current commitments for full belief) and let LK be a *minimal theory* such that $LK \subseteq K$. The *basic partition* Π is a set of *expansions*[19] of LK, not necessarily all of them and not necessarily all (or some of) the maximal and consistent ones. A necessary constraint on the admissibility of Π is that should be formed by expanding LK with sentences that are relevant answers to questions under investigation and that the expansions are restricted to expansions by adding to LK elements of a set of sentences such that LK entails that exactly one of them is true and each element of the set is consistent with LK.

The *ultimate partition* is the subset Π_K of partition cells of Π whose intersection is exactly K. In addition Π - Π_K is the *dual ultimate partition* Δ.

Call $\mathcal{M}$ the set of maximal and consistent theories definable in $\mathbf{L}$. For every $A \in \mathbf{L}$, $[A] = \{w \in \mathcal{M} : A \in w\}$. By the same token for every theory T definable in $\mathbf{L}$, $[T] = \{w \in \mathcal{M}\text{: } K \subseteq w\}$. When T is a theory obtained by intersecting a set of cells of Δ, we will use the notation $|T|$ to denote the set of partition cells (of the basic partition) whose intersection determines T. Also if the theory T in question is finitely axiomatizable via a sentence $A \in L$, $|A| = |T|$. Finally let $L \subseteq \mathbf{L} = \{A \in \mathbf{L}\text{: } |A| \neq \emptyset \text{ and } |\neg A| \neq \emptyset\}$.

Every *potential contraction removing* $A \in L$ from K is the intersection with K of a nonempty subset R of $\neg A$-entailing cells of Δ and a subset R^* of A-entailing

[19] An expansion of a theory K with a sentence A is defined as $Cn(K \cap \{A\})$.

cells of Δ that may or may not be empty. A *maxichoice contraction* of K relative to Δ is the intersection of K with a single element of Δ. A *maxichoice contraction of K removing $A \in L$* relative to Δ is the intersection of K with a single element of Δ that entails $\neg A$. A *saturatable contraction* of K removing $A \in L$ relative to Δ is the intersection of a maxichoice contraction of K removing A relative to Δ with the intersection of a set of elements of Δ none of which entail $\neg A$.

Definition 48 *Let $S(K, A)$ be the family of A-saturatable sets of K. I.e. if K is a theory, $X \in S(K, A)$ if and only if $X \subseteq K$, X is closed, and $Cn(X \cup \{\neg A\})$ is an element of the partition Δ.*

$\Phi = \{X : X = \cap Y$, with $Y \in 2^{\Delta} \cup [K]\}$. With these preliminary elements we can now introduce now a *measure of informational value* $V : \Phi \rightarrow [0,1]$. V is not just any value function. As the terminology indicates V is supposed to deliver a measure of the value of *information*. As such we assume that it inherits some basic properties of classical measures of information which are probability-based. A classical manner of utilizing probability in order to measure the content of information is to utilize the measure Cont(.) = 1 - Prob(.) - see for example [43] for an account of how this measure can be used in order to construct a decision-theoretically motivated theory of *expansion*.

There are two basic properties that probability-based measures of information satisfy. First they *respect entailment* in the following sense:

(Weak Monotony) For any two sets X, that are elements of Φ, such that $X \subset Y$, $V(X) \leq V(Y)$.

The second important postulate is the following one:

(Extended Weak Monotonicity) Let $X, Y \subseteq \Phi$. If S is incompatible with both X and Y, and if $V(X) \leq V(Y)$, then $V(X \cap S) \leq V(Y \cap S)$.

Unfortunately one cannot preserve all the properties of Cont in characterizing a notion of information value useful in contraction. The trouble with Cont is that it cannot rationalize (in terms of optimality) moving to a position of suspense when there is a tie in optimality. In fact, the Cont-value of the intersection of two optimal saturatable contractions need not and, in general, will not carry maximum Cont-value. So we propose to preserve the first two postulates while adding a third that permits rationalizing suspense among optimal options as optimal. In order to present this third postulate we need an additional piece of notation. Any saturatable contraction S in $S(K, A)$ has the canonical form $K \cap T_A \cap m_{\neg A}$, where T_A is an intersection of A-cells of Δ and where $m_{\neg A}$ is a single $\neg A$-cell of Δ.

Then we can say that two saturatable contractions removing A from K are *A-equivalent* if and only if they are constituted as intersections of K with different $\neg A$-cells in Δ and the same subset T_A of the subset all of whose members entail A. A saturatable contraction S removing A is A-equivalent to an intersection of a set T of saturatable contractions removing A (including S) if S is constituted as the intersection of K, a set T_A of A-entailing cells and a $\neg A$-cell in Δ, and $[(\cap T) \cap A] = T_A$.

(Weak Intersection Equality) For every subset T of of potential contractions removing A from K each element of which is of equal informational value and such that all elements in T are A-equivalent to their intersection, for every $X \in T$, $V(\cap T) = V(X)$.

Given a set of optimal saturatable contractions removing A from K relative to Δ, the previous principle guarantees that its intersection is also an optimal saturatable contraction. Consider now the following important property entailed by these requirements.

(Weak Min) If a finite $T \subset S(K, A)$, $V(\cap T) = min(V(X) : X \in T)$.

Observation 49 *Weak monotony, extended weak monotony and weak intersection equality imply Weak Min.*

Proof Focus first on the set $S(K, A)$. List all the saturatable contractions in T $\subseteq S(K, A)$ with S_i and $1 \leq i \leq k$. Consider then:

(P) $V(\cap T) = min(V(S_i) : S_i \in T)$.

(P) holds trivially for T of cardinality 1 and we should show that if it holds for all non-empty subsets T_n of T of cardinality $n < k$, then it holds for T' of cardinality $n + 1$. Then (P) holds for all non-empty subsets of T of cardinality k, i.e. for all non-empty subsets of $S(K, A)$.

So, assume (P) holds for all subsets of T of cardinality $n < k$. Consider then a particular such T_n and $\cap T_n \cap S_j$ with $S_j \notin T_n$. Let M be the set of partition cells of Δ entailing $\neg A$ used in order to construct staturatable contractions in T_n. We will consider the most general case where $S_j = K \cap M_A \cap X$, were M_A is an intersection of partition cells of Δ entailing A and X is a partition cell entailing $\neg A$ such that $X \notin M$. We will also assume that there is a non-empty subset of elements of M_A not used in order to construct saturatable contractions in T_n. Call this subset S_A. Notice that $S = K \cap S_A \cap X$ is a contraction in $S(K, A)$ which cannot be in T. Moreover $\cap T_n \cap S_j = \cap T_n \cap S$. Finally we can construct another potential contraction removing A which will be useful below: $Y' = \cap T_n \cap S_A$. Notice that X is now a theory incompatible with all saturatable contractions in T_n, a fact that will also be useful below. Weak Min holds for $\cap T_n$. Let Y be a member of T_n such that $V(\cap T_n) = V(Y)$. We need then to show that $V(\cap T_n \cap S_j)$ $= min(X, Y)$.

Consider first the case $V(X) \leq V(\cap T_n) = V(Y)$, where Y is a member of T_n. Let Z be a partition cell entailing $\neg A$ that that does not belong to $|T_n|$ and is distinct from X and where $V(Z) = V(X)$. Z should also be incompatible with the saturatable contractions in T_n. There need not be such a Z in $S(K, A)$. If that were the case we can always embed hypothetically Δ into Δ' containing a cell for Z and such that the original structure of values remains unaltered by the partition change. In order to do so consider the (logically finite) underlying language L and its expansion $L' = L \cup t$, where t is a fresh atom not occurring in L. For every theory S in L, where $V(S) = x$, and for every S' over L', such that $S' \cap L = S$, $V(S') = x$. So, we can construct an embedding partition Π'. For each original cell of Π which is the intersection of a set of maximal and consistent theories of L, consider now the theory determined by the t-counterpart of each one of these maximals. Now let Z be the $\neg t$-counterpart of X. One does so by adding to Π' a cell determined by intersecting the $\neg t$-counterparts of each maximal and consistent theory determining the cell containing X. It is obvious that Z exists and that $V(X) = V(Z)$.

If $T'_{L'}$ is a set of saturatable contractions expressible in L' over Π', it should be clear as well that if we manage to prove that $V(\cap T'_{L'}) = min(V(R) : R \in T'_{L'})$ then the result also shows that $V(\cap T') = min(V(S_i) : S_i \in T')$ – the reason being that $V(\cap T'_{L'} \cap L) = V(\cap T')$. It is important to realize for what follows that $Z_{L'}$ is incompatible both with $X_{L'}$ and with $Y_{L'}$. In order not to inflate terminology we will drop from now on the sub-index L'.

Notice first that weak monotony guarantees that both that $V(X \cap Y') \leq V(X) \leq V(Y)$ and $V(X \cap \cap T_n) \leq V(X) \leq V(Y)$. The main supposition in this first case together with extended weak monotony give us in addition: $V(X \cap Z) \leq$

$V(\cap T_n \cap Z)$.

Now, extended weak monotony also gives us as well that $V(Y' \cap Z) = V(Y' \cap X)$. Therefore weak intersection equality and weak monotony guarantee that $V(X \cap Z) = V(X)$. This, in turn, entails that $V(X \cap Y) = V(X)$, which is enough to establish that $V(X) = V(X \cap Y) = min(V(W) : W \in \cap[T \cup \{X\}])$. This completes the proof of the first case.

For the second case consider $V(\cap T_n) = V(Y) < V(X)$. Replace Y in T_n with X. Since the result contains only n contractions, the min-rule applies. Moreover, the result must carry informational value no less than the original, it carries informational value at least as great as Y. In effect, case 2 has been converted into the first case.

□

The three postulates that we just introduced are the *core postulates* of the notion of *damped* [48] [45] informational value used in contraction (as opposed to the notion of undamped informational value characterized by the first two postulates - which is central in decision-theoretical characterizations of expansion). We will assume as well here the following stronger property.

(Strong Intersection Equality) For every subset T of Φ each element of which is of equal informational value and for every $X \in T$, $V(\cap T) = V(X)$.

Strong intersection equality combined with weak positive monotonicity and extended weak positive monotonicity imply the following (can you prove this?)

(Min) If X and Y are potential contractions from K in Φ, $V(X \cap Y) = min(V(X), V(Y))$.

It is obvious that there are other forms of contraction parametrically obtainable by relaxing some of the principles that entail Min (in particular Strong Intersection Equality). Nevertheless, the form of contraction we are studying here is salient, we would like to argue, given its compatibility with the Principle of Entrenchment. A preliminary study of the family of contraction functions obtainable by relaxing Min is offered in [10] (in particular the article proposes an axiomatization of the notion of *core contraction* obtained by only enforcing the core postulates). We will return to this issue after introducing some technical notions in the coming section.

5.1.1 Shells of informational value

The assumption of the core postulates and the stronger Min condition allows us to construct the following notion of *rank*.

Definition 50 *Let I = range(V) be a set of indices. For $x \in I$ let R^x be the non-empty set X of partition cells in Δ such that for every $Y \subseteq X$, $V((\cap Y) \cap K)$ $= x$.*

Intuitively R^x groups the partition cells of Δ such that the intersection of any subset of them with K has value x. By Min the intersection of any subset of them with K, has also value x. We can extend here the notion of rank, by adjudicating ranks to sets $P \subseteq 2^{\Delta}$.

$\rho^+(P) = \max(y\text{: } R^y \cap P \neq \emptyset)$

So, for P $\subseteq 2^{\Delta}$, such that there is $A \in L$, with $|A| = P$, we have that $\rho^+(|A|)$ $= y$, where R^y is the set of partition cells of Δ of largest rank intersecting $|A|$. Of course, we have then that $\rho^+(\{w\}) = y$ when $w \in R^y$ and for every $Y \subseteq R^y$, $\rho^+(Y) = y$.

We can now introduce the notion of *m-shell of informational value.* The idea of a m-shell is to group together all the ranks R^x where x is greater or equal than the index m.

Definition 51 *The* x-shell of informational value $S^x = \cup_{i \in I}^{i \geq x} R^i$*. The* system of shells of informational value *(SS) $\mathcal{S}$ is defined as: $\mathcal{S} = \{S^x : \cup S^x = \Delta\}$*

It should be obvious that shells of a shell system (SS) are nested. Notice in addition that for any cell $w \in \Delta$ we do not necessarily have $V(w) = \rho^+(w)$. For, by definition, $\rho^+(w) = \text{V}(K \cap w)$. The only constraint imposed by WM in this case is that $\rho^+(w) \leq V(\{w\})$. So every cell in Δ has a *value-level* which might not coincide with its rank.

A SS for a value function V determines a grading on Δ. So, none of the maximals in Π_K appear in the SS. But of course there are some constraints relating the value of K and the value of the sets in the SS. One important constraint (given by WM) is that $V(K) \geq i$, where S^i is the innermost shell of $\mathcal{S}$. Therefore $V(K)$ is greater than the value of any rank in $\mathcal{S}$.

With the help of the previous definitions we can now characterize our operator of informational value as an operation defined in systems of shells of informational value. We only need an additional definition. Let's consider $\mathbf{L} \subseteq L$ such that $\mathbf{L}$

$= \{A \in L$: there are cells $C_1, ...C_1$ in Π such that $[\cap_{i=1,n} C_i] = [A]\}$. Of course, for every $A \in \mathbf{L}$ there are cells $C_1, ...C_1$ in Π such that $|A| = \{C_1, ...C_1\}$. Let a sentence A be *rejected in* K if and only if $\neg A \in K$. Notice that as long as a sentence $\neg A \notin \cup(LK)$ a sentence A rejected in K should also belong to $\mathbf{L}$, in such a way that $|A|$ is well defined for it.

Definition 52 *Let $A \in \mathbf{L}$ be a sentence rejected in K. Then S_A is the union of $|K|$ with the set $X \in \mathcal{S}$ such that $X \cap |A| \neq \emptyset$ and for any other $Y \in \mathcal{S}$, such that $Y \cap |A| \neq \emptyset$, $X \subseteq Y$.*

S_A just picks the union of $|K|$ with the innermost shell in the SS $\mathcal{S}$ for V containing A-partition-cells of Δ. Now we can define some salient operators of informational value.

Definition 53 *$\div$ is an operator of informational value for a closed set K if and only if there is a selection function γ such that for all A in $\mathbf{L}$: (i) if $A \in K$, then $K \div A = \cap\gamma(S(K, A))$, where $\gamma(S(K, A)) = \{X \in S(K, A)$: $V(Y) \leq V(X)$ for all $Y \in S(K, A)\}$ and (ii) $K \div A = Cn(K)$ otherwise.*

When the value function V is constrained by WM, the resulting operator is called a *basic* operator of informational value. When it obeys all core postulates the resulting operator is called a *core* operator of informational value. Finally when V is constrained by all cores postulates plus Min, the resulting operator is called the *standard operator of informational value*. From now on we will mainly work with standard operators of informational value and we will use the notation '$\div$' to refer to them. Specific references and clarifications will be made otherwise.

Observation 54 $|K \div \neg A| = S_A$

Given a value function V defined on Φ it is possible to define the following useful relation:

Definition 55 *$P \leq_V Q$ if and only if $V(P) \leq V(Q)$*

In particular given a theory of reference K and a value function this relation orders all the potential contractions for the theory K. Moreover, it is immediate how to retrieve a relation $\leq_V$ from the system of shells for V and K. This can be done as follows:

Observation 56 *If P, Q are potential contractions of K then $P \leq_V Q$ if and only if there is S^x and S^y, such that $S^x \subseteq S^y$, R^x is the minimum rank intersecting $|P|$ and R^y is the minimum rank intersecting $|Q|$.*

This property flows from Min (provide a proof). Notice that if $P \leq_V Q$ this is so independently of the ranks of $|P|$ and $|Q|$ in the SS for V and K. Propositions in 2^Δ are ordered by $\leq_V$ in virtue of an index different than its rank. In fact, if $P \subseteq 2^\Delta$ we can define the following index of informational value ρ^-:

$\rho^-(P) = \min(y\colon R^y \cap P \neq \emptyset)$

Notice that for any $P \subseteq 2^\Delta$ we have that $V((\cap P) \cap K) = \rho^-(P)$. Nevertheless $V((\cap P)$ need not coincide with $\rho^-(P)$ – the theory $\cap P$ could have some value lower than $\rho^-(P)$. The index ρ^- has some obvious properties. For example: $\rho^-(P \cup Q) = \min(\rho^-(P), \rho^-(Q))$. And the index of informational value can be combined with ranks to give a simple definition of contraction. For any $A \in \mathbf{L}$ rejected in K:

Exercise 18 *Show that* $|K \div \neg A| = \cup\{P \subseteq 2^\Delta\colon \rho^-(P) = \rho^+(|A|)\} \cup |K|$.

In words, in order to construct $|K \div \neg A|$ we take the union of $|K|$ with all the propositions in 2^Δ such that their index of informational value equals the 'upper' rank of $|A|$. It is quite obvious that S_A is one of these propositions. We can now go back to some additional properties of $\leq_V$:

(d1) Either $[K \div A] \leq_V [K \div B]$, or $[K \div B] \leq_V [K \div A]$

Which, in turn, means that we can easily establish a pretty strong property of informational value contractions, namely that: (d1) Either $[K \div A] \subseteq [K \div B]$, or $[K \div B] \subseteq [K \div A]$. This property will be used later on in the proof of our main result.

Shells of informational value are structures which, at first sight at least, might be easy to conflate with Spohn's *ranking systems* [74], [75], presented above. Nevertheless ranking systems are different, both formally and conceptually from shell systems. Notice first that since we are working with finite partitions we can define shells also with range over $\mathcal{N}$, but in our case the domain is restricted to $\Phi = \{X : X = \cap Y, \text{ with } Y \in 2^\Delta\}$. Moreover in the case of rankings one proceeds by assigning first natural numbers to points (maximal and consistent theories in this case) and then ranks are assigned to propositions in an unproblematic manner. In our case an assignment of values to maximal and consistent theories does not fully determine the ranks of contractions for a theory of reference K even when the cells of the basic partition are constituted only by maximal and consistent theories in L . Foor in this limit case we can also define both $\kappa^-(P) = \min\{V(w)\colon w \in P\}$ and $\kappa^+(P) = \max\{V(w)\colon w \in P\}$. The second notion is not usually defined in Spohn's systems. But even if we were to use it notice that, given any proposition P in 2^{Δ}, nothing guarantees that $\kappa^{+}(P) = \rho^{+}(P)$ or that $\kappa^{-}(P) = \rho^{-}(P)$. As we explained before, the partition cells w in Δ receive a rank $\rho^{-}(\{w\}) = \rho^{+}(\{w\}) = x$, for R^{x} such that $w \in R^{x}$. But this rank need not coincide with w's value-level (measured by $\kappa^{+}(\{w\})$ or $\kappa^{-}(\{w\})$).[20]

In our framework the value-level of propositions is, of course, quite useful. It puts a constraint on permissible rankings ρ and it is crucial for determining iterated contractions (and revisions). But the value-level of partition cells does not fully determine ranks (ρ) of sets P in 2^{Δ}, and the 'upper' and 'lower' point-ranks κ do not play a significant role in our proposal. Moreover even if we were to restrict our attention exclusively to ranks in our sense to the detriment of value levels we would need to use *both* the 'upper' and 'lower' ranks ρ^{+} and ρ^{-}. So ranking systems and systems of shells are quite different. Both induce an indexed grading, but they induce gradings over different domains and the algorithm for assigned grades to propositions is different in each account. Spohn's ranking functions assign ranks to propositions identical to the degree of disbelief of its least disbelieved points, while in our account the rank of a proposition P (relative to K and Δ) is identical to the degree of informational value carried by the P-maxichoice contractions of K of maximal value. Notice that this notion of 'upper' rank has no operative counterpart in Spohn's system.[21]

All these formal differences flow from the central fact that the intended interpretation of grades in each account (Spohn's and ours) is fundamentally different. Spohn's account is a purely doxastic account where ranks can be (roughly) interpreted as the orders of magnitude of infinitesimal probabilities. As we explain above Spohn's main goal is to develop a non-probabilistic articulation of *degrees of disbelief*. In our account the grades are induced by a probability-based function measuring the *value* of information.

[20]The 'upper' rank κ^{+} and the 'lower' rank κ^{-} can be used in order to determine an epistemic ordering solely on the basis of point-value utility. The procedure, suggested by John Collins in [17], consists (roughly) in stipulating that proposition P is preferred to proposition Q if and only if the upper rank of P is greater than the upper rank of Q and the lower rank of P is no worse than the lower rank of Q. Or, alternatively that the lower rank of P is better than the lower rank of Q and the upper rank of P is no worse than the upper rank of Q. The procedure allows for incomparability of preference. This view of preference, nevertheless, does not satisfy postulates that are typical of probability-based notions of utility, like Weak Monotony, and therefore is quite different from the one presented here.

[21]Such notion can, of course, be defined for Spohn's ranking functions as well. According to Spohn's official interpretation the 'upper' rank of a proposition would be determined by the degree of disbelief assigned to its most disbelieved points.

5.2 Mild contractions

Here we will proceed axiomatically. The axioms used here are well known in the literature and their names are also more or less standard (see, for example, [27] – our terminology follows [48]). A contraction operator relative to K and Pi is a function from $K \times \mathbf{L} \rightarrow \Phi$.

($\div$ 0) There are cells $C_1, ..., C_n$ in Δ such that $K \div A \cap \neg A = \cap_{1,n} C_i$.

($\div$ 1) $K \div A = Cn(K \div A)$ [closure]

($\div$ 2) $K \div A \subseteq K$ [inclusion]

($\div$ 3) If $A \notin K$ or $A \in Cn(LK)$, then $K \subseteq K \div A$ [vacuity]

($\div$ 4) If $A \notin Cn(LK)$, then $A \notin K \div A$ [success]

($\div$ 6) If $Cn(A) = Cn(B)$, then $K \div A = K \div B$ [extensionality]

($\div$ 7) If $A \notin Cn(LK)$, then $K \div A \subseteq K \div (A \wedge B)$ [antitony]

($\div$ 8) If $A \notin K \div (A \wedge B)$, then $K \div (A \wedge B) \subseteq K \div A$ [conjunctive inclusion]

All the conditions, except antitony and the first structural property, are AGM properties. On the other hand there is a notorious postulate, AGM's axiom of recovery, which is not in the previous list and that is not derivable from the list:

($\div$ 5) $K \subseteq Cn((K \div A) \cup \{A\})$ [recovery]

Exercise 19 *Give a counterexample for Recovery. Prove:*
Either $[K \div A] \subseteq [K \div B]$, *or* $[K \div B] \subseteq [K \div A]$
Either $[K \div (A \wedge B)] = [K \div B]$, *or* $[K \div (A \wedge B)] = [K \div A]$
If $[K \div (A \wedge B)] \subseteq [K \div B]$, *then* $B \notin [K \div A]$, *or* $\vdash A$ *or* $\vdash B$.

Antitony is perhaps the most controversial postulate from the list. For example Hansson reports in [27] that antitony does not hold '[...] for any sensible notion of contraction'; while Rott and Pagnuco report in page 513 of [65] that '[...] intuitively antitony makes quite a bit of sense'.

The axiomatic base given here is exactly the one proposed in [65] to characterize *severe withdrawals*. Here we will show that this axiomatic base is indirectly supported by the intuitiveness of the postulates of Economy and Entrenchment. This gives, in turn, indirect support to Antitony.

5.2.1 A representation result for mild contractions

B

We will focus first on presenting some of the main lemmas that are needed in order to have a soundness result.

Exercise 20 *Prove taht any standard operator of informational value satisfies all the postulates of mild contractions.*

Aside from soundness we can also establish the following completeness result:

Theorem 57 *If '$\div$' is a mild contraction function obeying the correspondent postulates, then '$\div$' can be represented as an operator of informational value.*

5.3 The Principle of Entrenchment and informational value

So far nothing has been said about the Principle of Entrenchment invoked in the introduction. Let's first introduce a relation of entrenchment formally. Let $\leq$ be an ordering of the sentences of L. $\leq$ is a relation of entrenchment for a theory K if and only if the following postulates are satisfied

(i) If $A \leq B$ and $B \leq C$, then $A \leq C$ (transitivity)

(ii) If $A \in Cn(B)$, then $B \leq A$ (dominance)

(iii) $A \leq A \wedge B$ or $B \leq A \wedge B$ (conjunctiveness)

(iv) If $K \neq L$, then $A \leq B$ for every $B \in$ L if and only if $A \notin K$ (minimality)

(v) If $A \leq B$ for every $A \in$ L, then $B \in Cn(LK)$

Now, we remind the reader that our principle of entrenchment said that in giving up a sentence A from the current view one should preserve the sentences better entrenched than A. This translates formally into:

Definition 58 *If $A \in K$ and $A \notin Cn(LK)$, then $K \div A = K \cap \{B : A < B\}$ and K otherwise.*

This simple and elegant manner of characterizing contraction in terms of entrenchment was first proposed by Rott in [64]. AGM cannot be characterized in terms of entrenchment in this way. The non-equivalent bridge connecting entrenchment and AGM contractions stipulates that: $K \div A = K \cap \{B : A < A \vee B\}$

Exercise 21 *Prove that if $\leq$ satisfies the postulates (i) to (v) then the function $\div$ obtained from $\leq$ by definition 5.1 is a mild contraction.*

Therefore, via the completeness result offered above, the informational value contractions can be retrieved by using the Principle of Entrenchment. Moreover, we can also appeal to a result recently proved by Rott and Pagnuco (and to our completeness result) in order to retrieve the relevant notion of entrenchment from an informational value contraction.

Definition 59 *If $A \leq B$ if and only if $A \notin K \div B$, or $B \in Cn(LK)$.*

Exercise 22 *Prove that if $\div$ is a mild contraction then the relation $\leq$ obtained via definition 5.2 satisfies the postulates (i) to (v) for the entrenchment relation.*

So, the Principle of Entrenchment and the Principle of Economy give exactly the same account of contraction. AGM contractions are also mirrored by a corresponding notion of entrenchment, but this notion does not obey the Principle of Entrenchment. One of the consequences of this divergence is the fact that AGM contractions satisfy the controversial principle of recovery, which is not satisfied by mild contractions (alias severe withdrawals according to Rott and Pagnuco's terminology).

6 Part II: Rational Choice and Belief Change

The classical framework of *optimization* used in standard choice theory recommends choosing, among the feasible options, a *best* alternative. So, if S is the feasible set and R is a weak preference relation over S, optimization recommends focusing on the following set of maximal or best elements of S:

$$G(S, R) = \{Y \in S\colon \text{for all } Z \in S,\ ZRY\}$$

But many economists have recently pointed out that this stringent form of maximization might not be the kind of maximization that one can apply in practical problems where information is usually incomplete and sometimes scarce. For example the Nobel-winner Amartya Sen has recently pointed out that;

> The general discipline of maximization differs from the special case of optimization in taking an alternative as choosable when it is not known to be worse that any other. [...] The basic contrast between maximization and optimization arises from the possibility that the preference ranking may be incomplete ([71], p. 767).

To define a maximal set we can use the asymmetric part $P(R)$ of a binary acyclic relation R as follows:

$$M(S, R) = \{Y \in S\text{: for no } Z \in S,\ ZP(R)Y\}$$

It is easy to see that in general (for any binary relation R and any non-empty feasible set S) we have that $G(S, R) \subseteq M(S, R)$. When R is complete $G(S, R) = M(S, R)$. Moreover a maximal set $M(S, R)$ can always be replicated by optimizing a complete relation $R+$ obtained from R by transforming incomparabilities into indifferences (show this). Obviously this new relation $R+$ has to be complete but it might fail to be transitive (provide an example). In addition, although this new relation can mimic the maximizing behavior of $P(R)$ it is clear that it should not be used for representing knowledge. Sen warns against using this kind of relations in representing economic knowledge in particular, but it is clear that the problem is more general.

The discipline of belief change has been dominated by the use of optimization techniques. This is particularly transparent in the recent work of Hans Rott who has shown how some of the central axioms of the so-called AGM theory of contraction arise from constraints in G-functions [66].

In order to study the contraction of a theory K with a sentence A the AGM trio [3] has proposed to focus on the set $(K \perp A)$ of maximal subsets of K that fail to entail A as the feasible set from which one make choices. Then the idea is to utilize a *selection function* γ that when applied to $(K \perp A)$ selects a non-empty set of $(K \perp A)$. In particular *partial meet contraction* focuses on selection functions that are *relational*, i.e. selection functions for which there is a binary relation $\leq$ such that:

$$\gamma(K \perp A, \leq) = \{Y \in K \perp A\text{: for all } Z \in K \perp A,\ Z \leq Y\}$$

then $K \div A$ is defined as the intersection of the elements of $\gamma(K \perp A, \leq)$. Obviously this definition relies on a process of optimization of the sort discussed above.

Exercise 23 *Show that if contraction is defined in this way the postulate of recovery is satisfied.*

Two main criticisms can be raised against this way of articulating contraction. One concerns the feasible set, which many see as too restrictive. We will consider this problem in the final sections of this article. The second criticism concerns the use of optimization techniques. In many applications one might not have access to the binary relation needed to optimize, a relation that imposes strong demands, like completeness. In particular one might face cases of indeterminacy, which can be caused, for example, by lack of information or, alternatively, by the existence of conflicting standards of valuation.

Situations of this sort can be modeled by considering a set of binary relations that the agent considers permissible. As an alternative some researchers have proposed that contractions are not functions but relations. This seems an extreme solution that does not seem to capture the fact that in contracting one does not merely want counseling about possible options but a concrete output.

When an agent faces indeterminacy between a set of permissible orderings one standard solution is to consider the compromises between them, that is the *categorical* relation obtained by considering all ordered pairs shared by all the permissible orderings. This seems a sensible solution but, of course, the resulting categorical preference need not be complete. It has to be a quasi-ordering, but not necessarily complete. So, optimizing this relation might not be possible. But, of course, one can maximize the resulting quasi-ordering. Unfortunately the existing literature on preference relations, choice functions and revealed preference does not offer any general characterization of a notion of quasi-transitive rationality of this sort. Deceptively similar results are presented in Suzumura's monograph on rational choice [77] or in [72] - the latter result is the perhaps the closest to our interests, but the condition used in the representation is not useful for our purposes. The next section focuses on generalizing these results to the case of maximization, rather than optimization.

6.1 Choice Functions and Quasi-transitive Rationality

Choice can be analyzed in a rather abstract framework by appealing to standard techniques used in the pure theory of consumer choice. We can start with a universal set X which remains fixed throughout the analysis. Let then $\mathcal{S}$ be a distinguished non-empty collection of non-empty subsets of X. The pair $(X, \mathcal{S})$ will be called the *choice space* of the agent.

Now we can introduce a new notion that will be useful below. A *choice function* on a choice space $(X, \mathcal{S})$ is a function C defined on $\mathcal{S}$ that assigns a non-empty subset (*choice set*) $C(S)$ of S to each and every S in $\mathcal{S}$.

A weak preference relation R on X *rationalizes* (or is *a rationalization* of) a choice function C on $(X, \mathcal{S})$ if and only if, for every $S \in \mathcal{S}$, $C(S)$ consists on the greatest points of S:

(R) There is a relation R, such that for all $S \in \mathcal{S}$, $C(S) = G(S, R)$

On the other hand if a choice function C on $(X, \mathcal{S})$ has a preference relation R satisfying (R), it is customarily said that C is a *rational choice function.* Obviously this terminology reflects the dominant view according to which rational choice is choice in accordance with the *optimization* of an underlying weak preference relation.

We will follow here the terminology of Suzumura [77] according to which this notion of rationality is called G-rationality. The obvious alternative in terms of maximization will be called M-rationality. We will focus in the following subsections on characterizations of types of M-rationality. The reader is referred to [77] for classical results considering G-rationality. In the following analysis we will restrict our analysis to the case where $\mathcal{S}$ consists of all finite non-empty subsets of the universal set X.

We will start this analysis with a precise definition of quasi-transitive M-rationality.

Definition 60 *A choice function C is* quasi-transitive *M-rational if and only if it is M-rational with a quasi-transitive (reflexive and transitive) rationalization.*

The quasi-transitivity of R requires that the relation is both reflexive and that $P(R)$ is transitive. It is important to make here some comments about the domain of a choice function C on a space $(X, \mathcal{S})$. Some of the most common constraints on domains are the following:

(Finite domain) $\mathcal{S}$ consists of all finite nonempty subsets of X.

(Finitely additive domain) For any S_1, S_2 in $\mathcal{S}$, $S_1 \cup S_2$ in $\mathcal{S}$.

(General domain) $\mathcal{S}$ is a specified nonempty collection of nonempty subsets of X.

Most of the work in social choice is done with finite domains. In particular there are complete characterizations of quasi-transitive G-rationality for finite domains, but some of the most complete presentations of recent work on this area (like [77]) do not include characterizations of either quasi-transitive M-rationality for finite domains or G-rationality for general domains. We will tackle the first problem here and we will make some remarks about the second as well in the last section. More importantly we will assume from now on the Finite Domains constraint. As an important auxiliary step we will define here the *base relation* R^C for C as follows:

Definition 61 $R^C = \{(x, y) \in X \times X \colon x \in C(\{x, y\})\ \}$

Now we have that $P(R^C) = \{(x,y) \in X \times X\colon\ x \in C(\{x,y\})$ and $y \notin C(\{x,y\})$ $\}$. We will now consider some well-known axioms constraining choice functions abstractly. Most of these axioms have now more or less standard names. We will follow here again the nomenclature used by Suzumura in [77], sometimes adding historical clarifications:

(Chernoff Property) For all $S, T \in X$, if $S \subseteq T$, then $S \cap C(T) \subseteq C(S)$.

(Superset Axiom) For all S, T in $\mathcal{S}$: If $S \subseteq T$ and $C(T) \subseteq C(S)$, then $C(S) = C(T)$.

(M-Condorcet property) For all $S \in \mathcal{S}$: $M(S, R^C) \subseteq C(S)$

The M-Condorcet property is not discussed in [77]. Suzumura discuses instead a property called Generalized Condorcet, where $M(S, R^C)$ is replaced by $G(S, R^C)$ in our formulation of M-Condorcet. With the help of these definitions we can establish a preliminary result:

Observation 62 *C is M-rational if and only if R^C is a M-rationalization of C.*

Proof We only need to establish the 'only if' part. Assume that C is M-rational with a rationalization R. So, we have that for every $\{x,y\} \subseteq \mathcal{S}$, $C(\{x,y\}) = M(\{x,y\}, R)$. But then it is easy to see that $P(R^C)$ and $P(R)$ are co-extensional ($(x,y) \in P(R^C)$ if and only if $(x,y) \in P(R)$).

We can follow here the general lines of a method of proof that Suzumura used for a different purpose in [77]. This requires appealing to a nested condition called Path Independence:

(Path Independence-α) For all $S, T \in \mathcal{S}$, $C(S \cup T) = C(C(S) \cup C(T))$.

(Path Independence-β) For all $S, T \in \mathcal{S}$, $C(S \cup T) = C(C(S) \cup T)$.

Two theorems proved in [77] will be useful later on:

Theorem 63 *(Suzumura-Th. 2.4) A choice function C on a space $(X, \mathcal{S})$ satisfying the Finite Domain condition satisfies Path Independence-α if and only if it satisfies Path Independence-β.*

Theorem 64 *(Suzumura-Th. 2.5) A choice function C on $(X, \mathcal{S})$ satisfies path independence if and only if it satisfies Chernoff axiom and the Superset axiom.*

Now we can state a first theorem characterizing the notion of quasi-transitive M-rationality:

Theorem 65 *A choice function C on $(X, \mathcal{S})$ satisfying the Finite Domain condition is quasi-transitive M-rational if and only if satisfies Chernoff's axiom, the Superset axiom and the M-Condorcet property.*

Proof To prove the theorem it is enough to show that path independence (PI) and M-Condorcet are necessary and sufficient for quasi-transitive M-rationality. For sufficiency it is enough to show (in the presence of the previous observation) that R^C is quasi transitive and that for all finite subsets S of $\mathcal{S}$, $C(S) = M(S, R^C)$.

Assume by contradiction that $x \in C(S)$ but $x \notin M(S, R^C)$. If so there is $y \in S$ such that $(y, x) \in P(R^C)$. This means that $x \notin C(\{x, y\})$ and $y \in C(\{x, y\})$. Now we can use the following instance of path independence:

$C(S) = C(C(\{x, y\}) \cup C(S - \{x, y\})$

from which we can conclude that $x \notin C(S)$. Contradiction. So, this proof plus the M-Condorcet property yields that $C(S) = M(S, R^C)$ as we wanted. For transitivity of the asymmetric part of R^C assume that $(x, y) \in P(R^C)$ and $(y, z) \in P(R^C)$. This entails that $\{x\} = C(\{x, y\})$ and $\{y\} = C(\{y, z\})$. But then in virtue of Path Independence (β) we have that $C(\{x, z\}) = C(C(\{x, y\}) \cup \{z\})$ $= C(\{x\} \cup \{y, z\}) = C(\{x\} \cup C(\{y, z\})) = C(\{x, y\}) = \{x\}$.

Now we need to focus on the necessity of Path Independence and M-Condorcet for C to be quasi-transitive M-rational. It is not difficult to see that if C is quasi-transitive M-rational it should satisfy Chernoff, and this, in turn, guarantees that:

For all $S, T \in \mathcal{S}$, $C(S \cup T) \subseteq C(C(S) \cup C(T))$

For the converse assume by contradiction that $x \in C(C(S) \cup C(T))$ but that $x \notin C(S \cup T)$. Without loss of generality we can assume that $x \in C(S)$. We therefore have that:

(I) For all $z \in S$: $(z, x) \notin P(R^C)$.

Moreover we also know that there is no $z \in \{C(S) \cup C(T)\}$ such that $(z, x) \in P(R^C)$. In particular there is no $z \in C(T)$ such that $(z, x) \in P(R^C)$.

Now, since we assumed that $x \notin C(S \cup T)$ we also know that there is $y \in \{S \cup T\}$ such that $(y, x) \in P(R^C)$. Moreover y cannot belong to S in virtue of (I). And since we also established that there is no $z \in C(T)$ such that $(z, x) \in P(R^C)$, we have that y must belong to $T - C(T)$. Therefore there must be $r \in T$ such that $(r, y) \in P(R^C)$. And, again, we know that r must belong to $T - C(T)$. Repeating this procedure we can generate an infinite sequence $y, r, ...$ in a finite set T, which gives us the desired contradiction.

□

We can now focus on some familiar properties that every quasi-transitive M-rational choice function C on a space $(X, \mathcal{S})$ should obey and some which are not obeyed:

(Sen's γ) For all $S, T \in \mathcal{S}$, such that $S \cup T \in \mathcal{S}$, $C(S) \cap C(T) \subseteq C(T \cup S)$.

(Aizerman Property) For all $S, T \in \mathcal{S}$, if $S \subseteq T$, and $C(T) \subseteq S$, then $C(S) \subseteq C(T)$.

(Arrow's Axiom) For all $S, T \in \mathcal{S}$, such that $S \cup T \in \mathcal{S}$, if $C(S \cup T) \cap S \neq \emptyset$, then $C(S) \subseteq C(T \cup S)$.

The Aizerman Property is a condition that has been studied only recently in the literature on social choice. It derives from the work of M. Aizerman and A. Malishevski [1]. As stated above is attributed to Aizerman by Harve Moulin in [56], and later on by Suzumura and Xu in [78].[22] The so-called Aizerman Property has been used many times since then. For example, it appears in [66] labeled as Condition III, and more recently it has been used by J. B. Kadane et al. in [35] labeled as Axiom 2. The basic idea of the condition is that one cannot promote an inadmissible option to an admissible one by deleting inadmissible options from the option set (this is almost verbatim the paraphrases used in [35]).

It is easy to see that the Aizerman property follows from the Superset Axiom. Moreover it is also easy to see that Aizerman and Superset are co-extensional in the presence of Chernoff. This connects our results with recent results in the theory of social choice functions. In particular Aizerman and Malishevski show in [1] that if a general choice function $C(.)$ on K, which always maps $S \in K$ into a non-empty subset $C(S)$ of S, satisfies Chernoff and Aizerman then it is *pseudo-rationalizable* in the sense that, for all S, there exist n orderings $R_1, ..., R_n$ such that $C(S)$ is the union of the greatest sets of each one of these orderings.

[22] In a review published by Aizerman [2] and in [1] he actually uses a stronger condition that builds in some of condition α as well, where the consequent of the axiom is an equality rather than an inclusion. The condition is called O and Aizerman attributes it to Chernoff. It seems that what is usually known as the Aizerman Property in the recent literature is a weakening of this condition O.

And Suzumura and Xu showed recently in [77] that when the domain X is finite, the maximal sets induced by a relation Q obeying slightly stronger constraints (if a pair belongs to the strict part of the transitive relation of Q then it belongs to Q) satisfy Chernoff and Aizerman if the set of all the *strict extensions* of Q is non-empty (R is a strict extension of Q if and only if Q is contained in R and the same applies to their asymmetric parts). So, we know that the choice function that fully characterizes quasi-transitive M-rationality is (when the universe is finite) pseudo-rational.

In the case of γ it is well known that Superset is independent of γ, even for finite X and in the presence of Chernoff. But γ is entailed by Chernoff, the Superset axiom and the M-Condorcet property. As a matter of fact we can establish a stronger result:

Theorem 66 *A choice function C on a space $(X, \mathcal{S})$ satisfying the condition of Finite Domain is quasi-transitive M-rational if and only if satisfies Chernoff's axiom, the Superset axiom and γ.*

Proof The proof depends on a result first established by Sen [71]. Consider a choice function C on a finite space $(X, \mathcal{S})$, i.e. a space $(X, \mathcal{S})$ where $\mathcal{S}$ consists of and only of all finite nonempty subsets of X.

(Weak Revealed Base Preference) xR^+y if and only if $x \in C(\{x, y\})$.

(Binariness of Choice) $C(S) = G(S, R^+)$

The original choice function $C(S)$ is *binary* if and only if the revealed preference relation R^+ generated by the choice function, if used as the basis of choice, will, in turn, regenerate the choice function itself. Amartya Sen proved that a choice function is binary if and only if it satisfies Properties α (Chernoff) and Property γ.

Now, if we have Chernoff, the Superset axiom and the M-Condorcet property for a choice function defined over a space satisfying Finite Domain, we also have that for all finite subset S of $\mathcal{S}$, $C(S) = M(S, R^C)$. This flows from the first part of the proof of the previous theorem. Now, since the Finite Domain condition is satisfied, the revealed preference relation $R^+ = R^C$ should be complete. In this case we have that $M(S, R^C) = G(S, R^+)$. Therefore we have, as desired, that $C(S) = G(S, R^+)$.

So, any choice function satisfying Chernoff, the Superset axiom and the M-Condorcet property, is binary in Sen's sense, and therefore it obeys γ in virtue of his result.

Moreover, if we start with a binary choice function $C(S)$ obeying both Chernoff and γ we also know that $C(S) = G(S, R^+)$. Since R^+ is complete (by the Finite Domain condition) we have that $C(S) = G(S, R^+) = M(S, R^C)$. So, Chernoff and γ entail that a function $C(S)$ can be expressed as $C(S) = M(S, R^C)$ while Chernoff and Superset suffice to establish that R^C should be quasi transitive. These two facts give us M-Condorcet. So, quasi-transitive M-rationality is characterized by Chernoff, Superset and γ.

□

6.2 Maximizing categorical preference

We will assume a classical propositional language $\mathbf{L}$ containing the classical connectives as a representational tool. The underlying logic will be identified with its Tarskian consequence operator Cn: $2^{\mathbf{L}} \rightarrow 2^{\mathbf{L}}$ which is assumed to obey standard properties.

A tradition inaugurated by [3] studies *contractions* of a sentence A from a theory K by considering first the set of A-*remainder sets* obtained by forming all the maximal and consistent subsets of K that fail to entail A (denoted $K \perp A$). This set is tacitly considered as the menu from which a rational selection is performed and then the contraction is finally computed by taking the intersection of all the selected sets. The so-called AGM theory supposes in addition that the rational selection performed on the set of A-*remainder sets* (usually denoted as $K \perp A$) can be rationalized by selecting the *optimal* remainder sets with respect to an underlying ordering. Some facts about remainder sets are useful as well as a proporty that we will call the *upper bound property*: If $X \subseteq K$ and $Cn(X) \cap B = \emptyset$, then there is some X' such that X' extends X and $X' \in K \perp A$.

Exercise 24 *Prove the following facts about remainder sets:*

1 If A entails some element of a set of sentences S then $(K \cup \{A\}) \perp S = K \perp B$ and $K \perp (S \cup \{A\}) = K \perp S$.

2 $K \perp (A \wedge B) \subseteq K \perp A \cup K \perp B$

3 K is a theory. If $X \in K \perp A$, then $X \in K \perp B$, for all $B \in K$ minus X.

4 K is a theory and let A and B be elements of K. Then $K \perp (A \wedge B) = K \perp A \cup K \perp B$

As the recent work of Has Rott has made clear [66], the central procedure to calculate AGM contractions can be stated as follows:
B

$$\text{(AGM) } K \div A = \bigcap G((K \perp A), R)$$

Some extra constraints, like the transitivity of R can be added. Many scholars, including some of the members of the AGM trio, have criticized this approach to contraction. Decision-theoretically Isaac Levi has provided most of the criticisms in various recent papers and books. On the one hand one can question the selection of $K \perp A$ as the menu for the performance of rational choice when the problem is to contract A from K. On the other hand one can also question the idea that rational choice presupposes optimization. We will focus first on the latter criticism. So, the case that we are interested in analyzing first is the one where, given a menu of choice $K \perp A$ relative to a theory K, we have a set of binary relations $\leq_i$, with $i = 1, ..., n$ over the menu. In this case we can construct $\leq_c = \{(x, y)$: $(x, y) \in \leq_i$ for every permissible $\leq_i$ $\}$. We will follow the notation used in [48] and [44] and will call $\leq_c$ *categorical preference*.

The categorical preference $\leq_c$ should be a quasi-ordering when the binary relations $\leq_i$ are complete orderings over $K \perp A$. So, as a first step we propose the analysis of the contraction determined by the following schema for $A \in K - Cn(\emptyset)$ – the contraction is set to K otherwise.[23]

$$\text{(C) } K \div A = \bigcap M((K \perp A), <_c)$$

This is a form of *partial meet* contraction, based on maximizing rather than optimizing.[24] Which properties of contraction are satisfied by a contraction function of this form? It is easy to see that all the so-called *basic axioms* of contraction are satisfied:

[23]We are interested here in a family of orderings grounded on a set of value functions. It should be remarked, nevertheless, that the analysis we are offering is purely ordinal. Scholars who are skeptic about overusing the notion of utility either in epistemic or economic applications (or scholars who are worried about the problems related to interpersonal comparisons of utility that the traditional economic approach can engender in interactive situations) can still profit therefore from the results we are presenting. The set of orderings can perfectly arise even if their origin can be traced not to a previous appeal to epistemic utility but to the use of an algorithmic approach complemented with a broader notion of goals (epistemic goals in this case). See [57] for a view than defends the algorithmic approach rather than one based on utility in order to tackle some social and economic problems.

[24]Hans Rott considered this possibility in [66] but he did not explore it formally or conceptually.

(÷ 1) $K \div A = Cn(K \div A)$ [closure]

(÷ 2) $K \div A \subseteq K$ [inclusion]

(÷ 3) If $A \notin K$ or $A \in Cn(\emptyset)$, then $K \subseteq K \div A$ [vacuity]

(÷ 4) If $A \notin Cn(\emptyset)$, then $A \notin K \div A$ [success]

(÷ 5) $K \subseteq Cn((K \div A) \cup \{A\})$ [recovery]

(÷ 6) If $Cn(A) = Cn(B)$, then $K \div A = K \div B$ [extensionality]

Exercise 25 *Prove that the aforementioned postulates are satisfied.*

The only controversial postulate in this list is Recovery. We postpone the discussion of this postulate until the last section where we will consider issues of feasibility in cognitive choice (in other words, which menus are acceptable in characterizing contraction).

Some basic properties of 'remainder sets' of the form $(K \bot A)$ are useful at this juncture. In particular $K \bot (A \wedge B) = (K \bot A) \cup (K \bot B)$, for $A, B \in K - Cn(\emptyset)$. In addition it is also useful to notice that for $A, B \in K - Cn(\emptyset)$ we have $M((K \bot A), <_c) \cap (K \bot B) = \emptyset$ if and only if $B \in \bigcap M((K \bot A), <_c)$. These are properties of remainder sets that do not depend on maximizing or optimizing.

Now the methods of proof used in [66] can be easily adapted to establish the following results involving how to contract conjunctions and other Boolean compounds. We assume in all results that the choice menu is finite. The two conditions that interest us are Chernoff and Superset. The following postulate reflects the first constraint on rational selection:

(÷ 7) $K \div A \cap K \div B \subseteq K \div (A \wedge B)$ [conjunctive overlap]

The completion $M*$ of a choice function $M((K \bot A), <_c)$ is defined as $M*((K \bot A), <_c) = \{M \in (K \bot A)\colon \bigcap M((K \bot A), <_c) \subseteq M\}$. A choice function is *complete* if and only if $M = M*$.

Lemma 67 *If* $\div$ *is defined via (C) for some choice function* $M((K \bot A), <_c)$*, then if* M *satisfies Chernoff (Sen's* α*)* $\div$ *satisfies postulates* $(\div 1)$ *to* $(\div 7)$*. Moreover if* $\div$ *satisfies postulates* $(\div 1)$ *to* $(\div 7)$ *and the choice function that determines* $\div$ *is complete, then* M *satisfies Chernoff.*

Exercise 26 *Prove the previous lemma.*

Now we can focus on some properties first analyzed in [66]. These properties have not been usually considered in the theory of belief change. They were first considered because of their obvious formal connections with well-known postulates used in the theory of choice functions. The 'c' in the name of the first property stands for 'cumulativity' (the name should be obvious for readers familiar with the literature on non-monotonic logic).

($\div$ 8c) If $B \in K \div (A \wedge B)$, then $K \div (A \wedge B) \subseteq K \div A$.

Lemma 68 *If $\div$ is defined via (C) for some choice function $M((K \perp A), <_c)$, then if M satisfies Sen's ϵ (Superset) $\div$ satisfies postulates $(\div 1)$ to $(\div 6)$ and $(\div 8c)$. Moreover if $\div$ satisfies postulates $(\div 1)$ to $(\div 6)$ and $(\div 8c)$ and the choice function that determines $\div$ is complete, then M satisfies Sen's ϵ.*

The choice-theoretical model proposed by Rott in [66], which depends on optimization, also considers the following important axiom of AGM:

($\div$ 8) If $A \notin K \div (A \wedge B)$, then $K \div (A \wedge B) \subseteq K \div A$.

This postulate naturally corresponds to the so-called Arrow Axiom (or Sen's property β^+). But it is easy to see that Arrow's Axiom does not hold for the maximizing choice functions we are considering here (why?). So, the postulate does not hold either. Nevertheless, a weaker postulate is satisfied:

($\div$ 8r) $K \div (A \wedge B) \subseteq Cn(K \div A \cup K \div B)$

The 'r' used in the name of this property stands for 'relational'. Now one can show that:

Lemma 69 *If $\div$ is defined via (C) for some complete choice function $M((K \perp A), <_c)$, M satisfies Sen's γ if and only if $\div$ satisfies postulates $(\div 1)$ to $(\div 6)$ and $(\div 8r)$.*

Aizerman corresponds to a condition of belief revision not previously considered in the literature.[25]

($\div$ 8a) If $A \notin (K \div (A \wedge B))$ and $B \in (K \div (A \wedge B))$, then $K \div (A \wedge B) \subseteq K \div A$.

[25] Nevertheless the constraints imposed by Aizerman on non-monotonic inference have been previously studied, for example in [66].

The resulting theory of contraction is obviously weaker than the AGM theory considered by Rott in [66]. The eight AGM axiom has been criticized by many authors before via the use of intuitive examples or by pointing out that it imposes unreasonable requirements on modeling contraction (John Pollock provided similar examples in conditional logic). We can call the contraction functions obeying ($\div 1$) to ($\div 7$), plus ($\div 8c$), and ($\div 8r$), *liberal contraction*, given that it has been obtained by employing a process of liberal maximization as opposed to the stringent process called optimization.

6.3 Feasibility and Cognitive Choice

In all the previous sections we have assumed that in order to determine the content of a contraction $K \div A$ it is adequate to maximize over a feasible set determined by the contents of the remainder set ($K \perp A$). But this is a controversial issue in the foundations of belief change. Isaac Levi has proposed in various writings that one should focus instead on the larger family of *saturatable* contractions removing A.

Definition 70 *Let $S(K, A)$ be the family of A-saturatable sets of K. I.e. if K is a theory, $X \in S(K, A)$ if and only if $X \subseteq K$, X is closed, and $Cn(X \cup \{\neg A\})$ is a maximal and consistent set.*[26]

So, the proposal in many of the previous articulations on the so-called 'Levi contractions' [48] is to widen the scope of the choice function used in order to define contraction. The idea is that these choice functions take saturatable families as arguments (again, the formal presentation in [12] is more complicated given that all concepts are partition-dependent). These choice functions should be such that when applied to a family $S(K, A)$, return a non-empty subset of $S(K, A)$.

Definition 71 *$\div$ is a Levi-contraction of a theory K if and only if there exists a choice function G for K such that for all sentences A: if $A \in K$, then $K \div A = \bigcap G(S(K, A))$, and if $A \notin K$, $K \div A = K$.*

[26] As a matter of fact the presentation offered in [12] is more involved given that all contractions are relative to a basic partition conformed by a set of expansions of a theory LK included in K. Not necessarily all of them and not necessarily all (or some of) the maximal and consistent ones. A necessary constraint on the admissibility of a basic partition for K is that should be formed by expanding LK with sentences that are relevant answers to questions under investigation and that the expansions are restricted to expansions by adding to LK elements of a set of sentences such that LK entails that exactly one of them is true and each element of the set is consistent with LK. This addresses the issue of relevance and makes the contraction function problem-dependent.

A contraction operator of this kind need not obey the controversial postulate of Recovery, presented above (why? present an counterexample). Unlike other presentations of contraction this kind of contraction is decision theoretically motivated, and it is usually complemented by the explicit introduction of a value function V on the set of logically closed subsets of a theory of reference K. Since we are considering the finite case the co-domain ranges over the natural numbers. The value function is supposed to obey at least minimal structural postulates like:

(Weak Monotony) For any two sets X, Y in the range of V, such that $X \subset Y$, $V(X) \leq V(Y)$.

A more robust notion of contraction can be then introduced by further constraining the choice function G in such a way that it optimizes the underlying value function;

$$G(S(K, A)) = \{X \in S(K, A) \colon V(X) \leq V(Y) \text{ for all } Y \in S(K, A)\}$$

The corresponding notion of contraction can be further enriched by adding constraints on the function V reflecting the idea that the index that has to be minimized in contraction is the loss of information value (as opposed to just information loss). In [12] a complete representation for this notion of contraction is presented. But this is not the object of the present study. Indeterminacy can arise for this notion of contraction in the same way in which it can arise for the notion of contraction previously studied (AGM contraction). When this happens the agent might have at his disposal only the ordinalized version of *a set* of value functions. As before one can nevertheless extract a categorical preference from it and proceed in the same way that we proceeded before (this issue is considered in passing in [48] - I am here basically exploring one of the open options suggested in the book by Levi).

Of course, all the structural results about maximization of quasi-orderings can also be applied here. Nevertheless the logic of saturatable sets differs from the logic of remainder sets, and this generates some differences. The first important difference is that we only have that: $S(K, (A \wedge B)) \subseteq S(K, A) \cup S(K, B)$. But the converse does not necessarily hold. This and other logical asymmetries put constraints on immediate applications of the results on liberal contraction presented in previous sections.[27]

The characterization of the contraction function that arises in this setting is considered in detail in a companion paper [11]. It is worth mentioning here nevertheless that the basic postulates of AGM contraction minus recovery are all valid. In [11] it is verified that ($\div$ 7) also holds on the basis of very mild assumptions about the underlying value function. Since this result does not require entering into an ampler discussion about choice functions, we reproduce it here for the sake of completeness.

(WMc) For any two sets X, Y in the power set of S(K, A),, such that $X \subset Y$, $X \leq_c Y$.

Of course, all the structural results about maximization of quasi-orderings can also be applied here. Nevertheless the logic of saturatable sets differs from the logic of remainder sets, and this generates some differences. The first important difference is that we only have that: $S(K, (A \wedge B)) \subseteq S(K, A) \cup S(K, B)$. But the converse does not necessarily hold. This and other logical asymmetries put constraints on immediate applications of the results on liberal contraction presented in previous sections.[28]

In [11] it is verified that ($\div$ 7) also holds on the basis of very mild assumptions about the underlying value function. Since this result does not require entering into an ampler discussion about choice functions, we reproduce it here for the sake of completeness.

Definition 72 *$\div$ is a liberal contraction of a theory K if and only if there exists a choice function M for K such that for all sentences A: if $A \in K$, then $K \div A$ $= \bigcap M(S(K, A), \leq_c)$, and if $A \notin K$, $K \div A = K$. The notion $\leq_c$ of categorical preference is assumed here to be a quasi-ordering.*

Here is a previous lemma useful in some of the proofs that follow:

Lemma 73 *Let M be a maximizing function for $\leq_c$ obeying WMc. If $Z \in M(S(K, A), \leq_c)$, and $Z \subseteq Z' \in K \bot A$, then $Z' \in M(S(K, A), \leq_c)$.*

[27] For example, it casts doubts about the validity of some postulates like the 'cumulative' postulate ($\div 8r$) – notice that in this setting we can perfectly have $Y \in S(K, A)$ and $Y \in M(S(K, A))$, but $Y \notin S(K, (A \wedge B))$.

[28] For example, it casts doubts about the validity of some postulates like the 'cumulative' postulate ($\div 8r$) – notice that in this setting we can perfectly have $Y \in S(K, A)$ and $Y \in M(S(K, A))$, but $Y \notin S(K, (A \wedge B))$.

Proof Assume that $Z \in M(S(K, A), \leq_c)$. Then by the so-called upper-bound property (see [3] – which depends on Zorn's lemma) there is Z' extending Z such that $Z' \in K \perp A$. By WMc we have that $Z \leq_c Z'$. Now, since $Z \in M(S(K, A), \leq_c)$, and $Z' \in S(K, A)$ we have that it is not the case that $Z <_c Z'$. Therefore Z and Z' are equi-preferred.

Exercise 27 *Prove the following: Let K be a theory. Then it holds for all sentences A and B that* $S(K, A \wedge B) \subseteq S(K, A) \cup S(K, B)$.

Lemma 74 *Let* $\div$ *a notion of liberal contraction on a theory* K. *Then* $\div$ *satisfies* ($\div$ *7).*

Proof I'll consider the main case where $A, B \in K$ minus $Cn(\emptyset)$. Assume that $C \in (K \div A) \cap (K \div B)$. We need to establish that $C \in (K \div (A\wedge))$. It follows from the assumption that if $X \in M(S(K, A)$ or $X \in M(S(K, BB)$ then $C \in X$.

Now, let $Y \in M(S(K, (A \wedge B))$. It follows from the previous exercise that either $Y \in S(K, A)$ or $Y \in S(K, B)$. Assume without loss of generality that $Y \in S(K, A)$. By the upper bound property (see above) there is some Y' such that $Y \subseteq Y' \in K \perp A$. Therefore $Y' \in K \perp (A \wedge B)$ (why?). Now we can use the previous lemma to conclude that $Y' \in M(S(K, A \wedge B))$.

We need to prove that $Y \in M(S(K, A)$. To do this it is enought to establish that if $Z \in S(K, A)$, then $V(Z) \leq V(Y)$. Complete the proof arguing by contradiction.

6.4 Bounded rationality: satisficing as maximizing

The idea that rationality presupposes an optimization process has been attacked by many scholars trying to develop *bounded* accounts of rationality. The terminology was coined by Herb Simon in a series of fascinating writings on this and related topics (see for example [73]). It is interesting to notice, nevertheless, that some *bounded* methods of choice, like Simon's *satisficing* [73] can also be seen as forms of maximization (in spite superficial impressions to the contrary). Sen has argued forcefully for this view [71]. The businessman who is willing to settle for x = \$ 1.00 million without concerns about raising it to y = \$ 1.01 million, regards both x and y as acceptable, but this does not mean that he sees as 'equally good'. With respect to *his welfare function* the businessman might place y over x. On the other hand, given the bounded nature of *his choice behavior* he is ready to settle for either x or y. So, given his goals neither x is placed over y nor viceversa. Nor there is a decision to accept the options as equally good *given the agent's goals.* This has lead Sen [71] to conclude that:

> So, in terms of *the goal function* (as opposed to his *welfare function*) there is a 'tentative incompleteness' here and both x and y can be seen as 'maximal' in terms of the *operational* goals. Thus interpreted satisficing corresponds entirely to maximizing behavior.

Sen recognizes that in view of these arguments satisficing can also be seen as an *as if* optimizing exercise (by using the completed extension of the goal function). But he also correctly warns against this 'as if' account:

> But, as we discussed earlier, the use of this *as if* preference is *interpretatively* quite different. Thus the substantive gap between satisficing and optimizing remains (closable only in a *purely formal* way), whereas the gap between satisficing and maximizing is both formally and substantively absent.

The insistence on transitively relational accounts of optimizing contraction in the AGM tradition can therefore be seen as a tacit dismissal of indeterminacy and therefore also of the use of bounded methods of choice of the sort that Simon has in mind. Transitivity can be retained by focusing on maximizing accounts of contraction based on quasi orderings, but in this case the *as if* account of these functions in terms of optimization needs to drop transitivity. The use of maximization techniques seems therefore more promising than its optimizing counterpart, both in order to accommodate bounded methods and to study the indeterminacy given by either ignorance or by the multiplicity of conflicting standards of valuation that might arise in cognitive decision problems.

References

[1] M. Aizerman and A. Malishevski (1981) General theory of best variants choice: Some aspects. *IEEE Transactions of Automatic Control* 26: 1030-1040.

[2] M. Aizerman (1985) New problems in the general choice theory: Review of a research trend. *Social Choice and Welfare* 2, 235-282.

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

[4] H. Arló-Costa (1999) Qualitative and Probabilistic Models of Full Belief, *Proceedings of Logic Colloquim'98, Lecture Notes on Logic* **13**, S. Buss, P.Hajek, P. Pudlak (eds.), ASL, A. K. Peters, 1-13.

[5] H. Arló-Costa (1995) Epistemic conditionals, snakes and stars, in *Conditionals, from Philosophy to Computer Science*, vol. 5 of Studies in Logic and Computation [Series editor: D.M. Gabbay], L. Fari

[6] H. Arló-Costa and I. Levi (1996) Two notions of epistemic validity, *Synthese*, 109:217–262.

[7] H. Arló-Costa (1999) Belief revision conditionals: *basic* iterated systems, *Annals of Pure and Applied Logic* **96**, 3-28.

[8] H. Arló Costa, H. and R. Parikh (1999) Two-place probabilities, beliefs and belief revision, *Proceedings of the Twelfth Amsterdam Colloquium*, Paul Dekker (ed.), 1-6.

[9] H. Arló Costa (2001) Bayesian epistemology and epistemic conditionals: On the Status of the Export-Import Laws, *Journal of Philosophy*, Vol. XCVIII, 11, 555-598.

[10] H. Arló-Costa (2004) Decision-theoretic contraction and sequential change, in Olsson, E., ed., *Knowledge and Inquiry: Essays on the Pragmatism of Isaac Levi*. Cambridge University Press, Cambridge Studies in Probability, Induction and Decision Theory.

[11] H. Arló-Costa (2005) Decision theoretic contraction and value indeterminacy: Maximizing rather than optimizing, forthcoming, an abstract appeared in the *Working Notes 23, IJCAI-05 Workshop on Nonmonotonic reasoning, action and change*, ed. L. Morgenstern and M. Pagnucco.

[12] H. Arló-Costa and I. Levi (2005) Contraction: On the decision theoretical origins of minimal change and entrenchment, *Synthese*, forthcoming.

[13] H. Arló-Costa and R. Parikh (2005) Conditional Probability and Defeasible Inference, *Journal of Philosophical Logic* **34**, 97-119.

[14] Blackburn, P, de Rijke, M and Yde Venema (200K1) *Modal Logic*, Cambridge Tracts in Theoretical Computer Science, 58, Cambridge University Press.

[15] Carnap, R. *Meaning and Necessity; A study of semantics and modal logic*, Second edition, Midway Reprint, Chicago, 1988.

[16] Chellas, B. *Modal logic: An introduction*, Cambridge University Press, 1980.

[17] J. Collins (2002) Belief Revision Derived from Preference, Talk: Philosophical Logic and Applications. CUNY. Graduate Center.

[18] B. De Finetti (1990) *Theory of Probability*, Volume I, Wiley.

[19] Dubins, L.E. 'Finitely additive conditional probabilities, conglomerability, and disintegrations,' *Ann. Prob.* 3:89-99, 1975.

[20] J. Dugundji (1966) *Topology*, Wm. C. Brown Publishers.

[21] M. Fitting (1983) Moore's nonmonotonic logic and S5, manuscript CUNY, November 1983.

[22] P. Gärdenfors (1988) *Knowledge in Flux*, Bradford Book, MIT Press, Cambridge, Mass.

[23] A. Grove (1988) Two modelings for theory change, *Journal of Philosophical Logic*, 17, 157-170.

[24] Gärdenfors, P.:1993, 'The role of expectations in reasoning,' *Lund University Cognitive Studies*, 21.

[25] Gärdenfors, P., Makinson, D.:1994, 'Nonmonotonic inference based on expectations,' *Artificial Intelligence* 65:197-245.

[26] M. Goldszmidt and J. Pearl (1992) Rank-based systems: A simple approach to belief revision, belief update, and reasoning about evidence and actions, in B. Nebel, C. Rich, and W. Swartout (Eds.), *Proceedings of the Third International Conference on Knowledge Representation and Reasoning*, Morgan Kaufmann, San Mateo, CA, 661-672, October 1992.

[27] S.O. Hansson (1999) *A Textbook of Belief Dynamics: Theory Change and Database Updating*, Kluwer Academic Publishers.

[28] S.O. Hansson and E. J. Olsson (1995) Levi Contractions and AGM Contractions: a Comparison, *Notre Dame Journal of Formal Logic* 36, pp.103-119.

[29] G. Harman (1988) *Change in View: Principles of Reasoning*, MIT Press.

[30] H. Katsuno and A. Mendelzon (1991) On the difference between updating a knowledge base and revising it, in *Belief Revision*, Gärdenfors, P., ed.: Cambridge University Press, Cambridge, 183-203.

[31] J. Hintikka (1962) *Knowledge and Belief: An Introduction to the Logic of the Two Notions*, Cornell University Press.

[32] R. Jeffrey (1968) Probable Knowledge, in I. Lakatos (ed.), *The Problem of Induction*, North Holland, 166-180.

[33] R. Jeffrey (1992) *Probability and the Art of Judgment*, Cambridge University Press.

[34] Jeffreys, H. *Theory of Probability*, New York, Oxford, 3rd edition, 1961.

[35] J.B. Kadane, M.J. Schervish and T. Seidenfeld (2004) A Rubinesque decision theory. *Institute of Mathematical Statistics, Lecture Notes - Monograph Series* Vol. 45, 45-55.

[36] K. Kelly (1998) Iterated Belief Revision, Reliability, and Inductive Amnesia,*Erkenntnis*, 50, 11-58.

[37] K. Kelly, O. Schulte and V. Hendricks (1997) Reliable Belief Revision, in *Logic and Scientfic Methods*, M. L. Dalla Chiara, et al., eds. Dordrecht: Kluwer.

[38] Kraus, S., Lehmann, D. and Magidor, M.:1990, 'Nonmonotonic reasoning, preferential models and cumulative logics,' *44 Artificial Intelligence*, 167-207.

[39] S. Kripke (1963) Semantical considerations on Modal Logic, *Acta Philosophica Fennica*, 16, 83-94.

[40] H.E.Jr. Kyburg (1961) *Probability and the logic of rational belief*, Wesleyan University Press. Middletown.

[41] H. E. Jr. Kyburg and C. M. Teng (2002) The Logic of Risky Knowledge, proceedings of WoLLIC, Brazil.

[42] I. Levi (1979) Serious Possibility, *Essays in Honor of Jaaakko Hintikka*, Reidel, Dordrecht, 219-36.

[43] I. Levi (1983) *The enterprise of knowledge*, MIT Press.

[44] I. Levi (1986) *Hard choices*, Cambridge University Press.

[45] I. Levi (1991) *The fixation of belief and its undoing*, Cambridge University Press, Cambridge.

[46] I. Levi (1996) *For the Sake of Argument: Ramsey Test Conditionals, Inductive Inference, and Nonmonotonic Reasoning*, Cambridge University Press, Cambridge, England.

[47] I. Levi (1997) The logic of full belief, in *The Covenant of Reason: Rationality and the Commitments of Thought*, CUP, 40-69.

[48] I. Levi (2004) *Mild Contraction: Evaluating loss of information due to loss of belief*, Oxford University Press.

[49] K. Lehrer (2000) *Theory of Knowledge*, second, substantially revised edition, Westview Press.

[50] P. Maher (1990) 'Acceptance without belief,' PSA 1990, volume 1, 381-392.

[51] J.C.C. McKinsey and A. Tarksi (1944) The algebra of topology, *Annals of Mathematics* 45, 141-191.

[52] G.E. Moore (1959) Russell's 'Theory of Descriptions, in *Philosophical Papers* by G.E. Moore, ed. H.D. Lewis (London: Muirhead Library of Philosophy), [2nd ed. 1963].

[53] R. Moore (1995) *Logic and Representation*, CSLI Lecture Notes No. 39.

[54] R.D. Luce and H. Raiffa (1957) *Games and Decisions*, Willey.

[55] Montague, R. *Universal Grammar*, *Theoria 36*, 373–98, 1970.

[56] H. Moulin (1985) Choice functions over a finite set: A summary, em Social Choice and Welfare 2, 147-160.

[57] R. Parikh (2002) Social Software, *Synthese*, 132, 187-211.

[58] H. Putnam (2004) *The collapse of the fact-value dichotomy and other essays*, Harvard University Press.

[59] F.P. Ramsey (1931) Truth and probability, in *The foundations of probability and other logical essays*, ed. R.B. Braithwaite, Routledge and Kegan.

[60] R. Reiter (1980) A logic for default reasoning, *Artificial Intelligence* 13, 81-132.

[61] H. Rott (1993) Belief Contraction in the context of the general theory of rational choice, *Journal of Symbolic Logic*, 58, 1426-1450.

[62] H. Rott (2000) Two Dogmas of Belief Revision, *The Journal of Philosophy* 97, 503-22.

[63] H. Rott (2001) *Change, Choice and Inference: A study of belief revision and non-monotonic reasoning*, Oxford Science Publications.

[64] H. Rott (1993) Belief Contraction in the context of the general theory of rational choice, *Journal of Symbolic Logic*, 58, 1426-1450.

[65] H. Rott and M. Pagnucco (1999) Severe Withdrawal (and Recovery), *Journal of Philosophical Logic* 28, 501-547.

[66] H. Rott (2001) *Change, Choice and Inference: A study of belief revision and non-monotonic reasoning*, Oxford Science Publications.

[67] Scott, D. 'Advice in modal logic,' K. Lambert (Ed.) *Philosophical Problems in Logic*, Dordrecht, Netherlands: Reidel, 143–73, 1970.

[68] Seidenfeld, T. and Kadane J. 'The extent of non-conglomerability of finitely additive probability,' *Z. War.* 66:205-226, 1984.

[69] T. Seidenfeld, M. Schervish and J. Kadane (2001) Improper regular conditional distributions, *Ann. Prob.*, 1-13, 2001.

[70] A. Sen (2003) Information and Invariance in Normative Choice, in *Rationality and Freedom*, Harvard University Press.

[71] A. Sen (2003) *Rationality and Freedom*, Harvard University Press.

[72] A. Sen (1971) Choice functions and revealed preference, *The Review of Economic Studies*, Vol. 38, N0 3, 307-317.

[73] H. Simon (1982) *Models of Bounded Rationality*, MIT Press.

[74] W. Spohn (1988) A general non-probabilistic theory of inductive inference, in *Causation in Decision, Belief Change and Statistics*, eds. W. Harper and B. Skyrms, Dordrecht:Reidel, 105-134.

[75] W. Spohn (2002) Laws, Ceteris Paribus Conditions, and the Dynamics of Belief, *Forschungsberichte der DFG-Forschergruppe Logik in der Philosophie*, No. 84

[76] R. Stalnaker (1993) A note on non-monotonic modal logic, *Artificial Intelligence* 64, 183-196.

[77] K. Suzumura (1983) *Rational Choice, Collective Decisions, and Social Welfare*, Cambridge University Press.

[78] K. Suzumura and Y. Xu (2003) Recoverability of choice functions and binary relations: some duality results, *Social Choice and Welfare*, 21, 21-37.

[79] van Fraassen, B. C. *Laws and Symmetry*, Oxford, New York, Oxford University Press, 1989.

[80] van Fraassen, B. C. (1992) Fine-grained opinion, probability, and the logic of full belief, *Journal of Philosophical Logic*, XXIV: 349–77.

[81] B. C. van Fraassen (2004) *The Empirical Stance: The Terry Lectures*, Yale University Press.