Formal Epistemology: Lecture Notes

Horacio Arló-Costa
Carnegie Mellon University
hcosta@andrew.cmu.edu

Changing view as a decision problem

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

The principle of Entrenchment

While assessing a contraction of K removing A retain all and only those elements of K that are better *entrenched* than A.

Principles of Belief Change

The principle of Preservation
Keep loss of information to a minimum in contraction.

The principle of Entrenchment
While assessing a contraction of K removing A retain all and only those elements of K that are better *entrenched* than A.

Information Value

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

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

Operators of informational value

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** of LK, not necessarily all of them and not necessarily all (or some of) the maximal and consistent ones.

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

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 **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 **L**, $[T] = \{w \in \mathcal{M}: K \subseteq w\}$. 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.

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 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 1 *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*.*

Information Value

$\Phi = \{X : X = \cap Y$, 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(.).

Weak Monotony

First V *respects 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 (why?).

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.

So, we will assume as well here the following property (which can also be deduced from a third postulate on value - Weak Intersection Equality).

(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 monotony and extended weak monotony imply the following:

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

Definition 2 *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.

We can extend here the notion of rank, by adjudicating ranks to sets $P \subseteq 2^{\Delta}$.

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

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 $|A|$.

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 3 *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, $\rho^{+}(w) =$ 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.

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 $\neg A \not\in \cup(LK)$ a sentence A rejected in K should also belong to $\mathbf{L}$, in such a way that $|A|$ is well defined for it.

Definition 4 *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 5 $\div$ *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*.

Observation 6 $|K \div \neg A| = S_A$

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

Definition 7 $P \leq_V Q$ *if and only if* $V(P) \leq V(Q)$

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 8 *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 $|Q|$.*

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^-:

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

Mild Contractions

Given K and Pi, $\div$: $K \times \mathbf{L} \rightarrow \Phi$.

($\div$ 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]

Exercise 1 *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

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

Exercise 2 *Prove that 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 9 *If '$\div$' is a mild contraction function obeying the correspondent postulates, then '$\div$' can be represented as an operator of informational value.*

Proof We need to show that starting with an operator $\div$ obeying the postulates of mild contractions we can explicitly construct a system of shells of informational value. We have to show as well that the operator $\div'$ obtained from the defined system of shells of informational value by requiring $[K \div' A] = S_{\neg A}$, where $S_{\neg A}$ is the smallest m-shell of informational value intersecting $\neg A$, is identical to $\div$. The proof proceeds in three stages. First we follow what now is a standard procedure in order to construct a Grove system in terms of the operator '$\div$'. Then we show how to build a system of shells of informational value in terms of the constructed

Grove system. In order to do so we show how to build a shell of informational value for the constructed Grove system. Finally we have to show that the operation $\div'$ obtained from the defined system of shells of informational value by requiring $[K \div' A] = S_{\neg A}$, is identical to $\div$.

Let's first focus on how to construct a Grove system in terms of the operation $\div$. As we explained above, the method for constructing a Grove system from a contraction operation is well-known. So, we will only outline here the main steps of the proof, skipping unnecessary details. The proof sketched here follows a suitable modification of Grove's original proof as presented in [**?**]. The central idea is quite simple: a Grove system of spheres **S** centered on $[X]$, is determined by identifying a sphere in **S** with the collection $[X \div A]$ for some $A \in L$. More precisely:

(d) $X_A = [X \div A]$.

In addition define the system of spheres **S** as follows;

S $= \{X_A\colon\ A \in L\} \cup \mathcal{M}$, when $K \neq L$, and **S** $= \{X_A\colon\ A \in L\} \cup \mathcal{M} \cup \emptyset$ otherwise.

The gist of this first part of the proof consists on showing that the system **S** is indeed a Grove system of spheres centered on $[X]$. Since this is important for the rest of the result we are showing, we will remind the reader immediately of the definition of a Grove system of spheres centered on $[X]$.

Let **S** be a collection of subsets of the set of all L-maximals of $\mathcal{M}$. **S** is a system of spheres, centered on $X = [\mathsf{K}] \subseteq \mathcal{M}$ and satisfying:

(1) $\mathbf{S}$ is totally ordered by $\subseteq$.

(2) X is the $\subseteq$-minimum of $\mathbf{S}$.

(3) $\mathcal{M}$ is the $\subseteq$-maximum of $\mathbf{S}$.

(4) If $A \in L$ and $\emptyset \neq [A] \in 2^{\mathcal{M}}$, then there is a smallest sphere S_A in $\mathbf{S}$ intersecting the set $[A]$.

Condition (1) is directly satisfied in virtue of (d) above and the fact that the following property can be deduced from the axioms of mild contractions:

(d1) Either $K \div A \subseteq K \div B$ or $K \div B \subseteq K \div A$.

Conditions ($\div$ 2) and ($\div$ 3) guarantee that $K \div \mathbf{true} = K$. This and (d) are enough to show that $X = [K]$ is a sphere. That this is the innermost sphere follows immediately from ($\div$ 2) and (d). This takes care of condition (2). Condition (3) is automatically satisfied by the given definition of **S**.

Condition (4) is slightly harder. Let A be such that $\emptyset \neq [A] \in 2^{\mathcal{M}}$. Now we need to show that there is a sphere $U \in$ **S**, such that $U \cap [A] \neq \emptyset$ and for every other $V \in$ **S**, such that $V \cap [A] \neq \emptyset$, we have $U \subseteq V$. The basic idea of the proof, which we skip here, is to show that $[K \div \neg A] = X_{\neg A}$ satisfies this constraint.

The proof of condition (4) establishes that (when A is not a tautology) $S_{\neg A} = [K \div A] = X_A$, where $S_{\neg A}$ denotes the smallest sphere intersecting $[\neg A]$ - we used basically the same notation $S_.$ for shells above, indicating the smallest m-shell intersecting [.]. This fact will be useful below.

Now we should focus on the second step of the proof, namely the construction of a system of shells of informational value for the constructed system of spheres. Here is the recipe in order to do so.

First construct a system of ranks out of the given system of spheres as follows: index first spheres with natural numbers starting with 0 assigned to the innermost core in such a way that S_i denotes the set of maximals in the sphere indexed by i. This is done via an indexing function mapping propositions to natural numbers. Then define a function δ from the range of the indexing function to propositions, such that $\delta(k) = S_{k+1}$ - S_k. As a second step assign an arbitrary V-value x to the innermost sphere of $\mathbf{S}$, $[K]$ as long as

x is greater than k, where k is the index of the outermost sphere S_k.

As a second step we need to give a value to each partition cell in Δ. In order to do so assign a uniform value $y < x$ to each partition cell in $\Delta(0)$ and, in general, for every $\Delta(i+1)$, for $i \geq 0$, assign a uniform value $z < z'$ to the elements in $\Delta(i+1)$, where z' is the uniform value of maximals in $\kappa(i)$.

As a third step we need to define ranks and m-shells of informational value. In order to do so we need to re-index the ranks we just defined from the Grove system. Assign to $\Delta(0)$ the index y and, in general, for every $\Delta(i)$, assign to it a rank equal to the value z assigned to elements of this cell of the partition. Indexes are assigned to shells accordingly.

The last definition allows us to complete the definition of the V-measure, by requiring that (i) for every $Y \subseteq R_i^x$, $V((\cap Y) \cap K)$ $= x$; (ii) that for every $Y \subseteq S_i^x$, such that R_i^x is the outermost rank such that $R_i^x \cap Y \neq \emptyset$,

$V((\cap Y)\cap K) = x$; and (iii) for any theory $T = \cap_{1,n} C_i$, for $C_i \in \Delta$, let $V(T) = min(V(C_i))$.

It is obvious that V, as defined, is a function. It is also clear it satisfies WM. In order to verify it we need to check that for any two maxichoice contractions of K, X, Y, such that $X \subset Y$, then $V(X) \leq V(Y)$. Let $R^X = R_i^m$ be the outermost rank intersecting $|X| - |K|$. Then it is clear that the outermost rank intersecting $|Y| - |K|$, $R^Y = R_{i'}^{m'}$ is such that $i' \leq i$ and $m' \geq m$. Since $|Y| - |K|$ is a subset of $S_{i'}^{m'}$ Y receives value m', and since

$|X| - |K|$ is a subset of R_i^m X receives value m. And since $m' \geq m$, weak monotony is satisfied.

Consider now $Y = \cap X$, where X is a family of maxichoice contractions of K, and its associated set $|Y|$. Consider again the outermost rank intersecting $|Y| - |K|$, $R^Y = R_i^m$. Then according to the proposed explicit definition of V, $V(Y) = m$. Take now an arbitrary $Z \in X$ such that $|Z| - |K|$ does not intersect R^Y. Then, by construction, $V(Y) < V(Z)$. And if $|Z| - |K|$ intersects R^Y,

$V(Z) = m$. So, clearly we do have that $V(Y)$ $= min\{V(Z) : Z \in X\}$.

Finally we need to check that an operation of contraction $\div'$ defined from the explicitly constructed system of shells of informational value, coincides with the operator $\div$ characterized by the postulates of mild contractions. We will define:

$|K \div' \neg A| = S_A$

where S_A is the smallest m-shell of informational value intersecting $|A|$ union $|K|$. The non-trivial case to consider is when A is not in LK. We will work under this assumption. Assume first that $B \in K \div \neg A$, $B \in \mathbf{L}$ This entails that $|K \div \neg A| \subseteq |B|$. The proof of condition (4) for Grove systems above tell us that $|K \div \neg A|$ is a sphere S_i, and since in our construction an m-shell is obtained from it by taking $S_i^m = \{C_i \in \Delta\text{:} \ \ S_i = \cup_{1,n} C_i\}$; then we have that $|K \div \neg A| - |K|$ is also a m-shell of informational value. Moreover, the proof of (4) also tells us that the smallest sphere overlapping $[A]$ is identical to $|K \div \neg A|$. Therefore the smallest m-shell of informational value overlapping $|A|$ is identical with $|K \div \neg A| - |K|$. Therefore we have that $|K \div' \neg A| \subseteq |B|$, and $B \in K \div' \neg A$, as desired. Proving that $K \div' \neg A \subseteq K \div \neg A$ only requires reversing the strategy used for the RTL inclusion$\bullet$

The Principle of Entrenchment

(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 $B \in \mathsf{L}$ 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 10 *If* $A \in K$ *and* $A \notin Cn(LK)$*, then* $K \dot{-} A = K \cap \{B : A < B\}$ *and* K *otherwise.*

Lemma 11 *If $\leq$ satisfies the postulates (i) to (v) then the function $\div$ obtained from $\leq$ by the previous definition is a mild contraction.*

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.