Synthese (2011) 179:211–221
DOI 10.1007/s11229-010-9778-3

Beth definability, interpolation and language splitting

Rohit Parikh

Received: 6 May 2009 / Accepted: 23 July 2009 / Published online: 31 August 2010

Abstract Both the Beth definability theorem and Craig's lemma (interpolation theorem from now on) deal with the issue of the *entanglement* of one language L_1 with another language L_2, that is to say, information transfer—or the lack of such transfer—between the two languages. The notion of splitting we study below looks into this issue. We briefly relate our own results in this area as well as the results of other researchers like Kourousias and Makinson, and Peppas, Chopra and Foo. Section 3 does contain one apparently new theorem.

Keywords Beth definability · Interpolation · Belief revision · Language splitting · Information

1 Introduction

One way to prove that a theory is incomplete is to provide two models of the theory which are not equivalent, i.e., have different properties. To prove that a theory is not *categorical*, the notion of equivalence used above is that of isomorphism.

Padoa's method in the theory of definitions suggests a similar procedure to show that some *notion* P is independent of some other notions $Q_1, \ldots, Q_n$ in the context that some theory T about $P, Q_1, \ldots, Q_n$ is already presupposed. One finds two models $\mathcal{M}_1, \mathcal{M}_2$ of T which coincide on $Q_1, \ldots, Q_n$ but disagree on P. Clearly if P were *definable* in terms of $Q_1, \ldots, Q_n$ (given T), this could not happen. For P

R. Parikh
Brooklyn College of CUNY, Brooklyn, NY, USA

R. Parikh (✉)
CUNY Graduate Center, New York, NY, USA
e-mail: RParikh@gc.cuny.edu

would be equivalent to some formula $\phi(Q_1, \ldots, Q_n)$ and if the two models agreed on $Q_1, \ldots, Q_n$, then they must also agree on ϕ and hence on P as well. Ergo, if $\mathcal{M}_1, \mathcal{M}_2$ exist which agree on the Q_i but not on P, then P could not be definable in terms of the Q_i.

Beth (1953) published a result[1] which showed that at least for first order logic Padoa's method is complete. *If no two models $\mathcal{M}_1, \mathcal{M}_2$ exist as supposed above then there is an actual definition of P in terms of the Q_i.*

The result which Beth proved has become known as the Beth definability theorem, a result which essentially every graduate student of logic knows.

Beth's paper was reviewed by Craig (1956), who then used his interpolation theorem (the last part of the theorem below, Craig 1957), to give his own proof of Beth's result.

Theorem 1.1 *Let L_1, L_2 be first order languages, $L = L_1 \cap L_2$ and T_1, T_2 be theories in L_1, L_2, respectively such that $T_1 \cup T_2$ has no model (is inconsistent). Then there is some formula ψ of L such that $T_1 \vdash \psi$ and $T_2 \vdash \neg\psi$. In particular, if ϕ is an L_1 formula and ξ is an L_2 formula and $T_1 \cup T_2, \phi \vdash \xi$ then there exists an L-formula ψ such that $T_1, \phi \vdash \psi$ and $T_2, \psi \vdash \xi$.*

Intuitively the first part means that if the two theories quarrel, then there must be something specific in their *common language* which they quarrel about. The second part says that any information being sent by T_1 to T_2 must go through some fact in the common language. The second part follows from the first by taking $T_1' = Con(T_1 \cup \{\phi\})$, $T_2' = Con(T_2 \cup \{\neg\xi\})$ and applying the first part to T_1', T_2'.

The interpolation theorem easily implies the Beth definability theorem via the following argument.

Proof Suppose that theory T is such that every two models $\mathcal{M}_1, \mathcal{M}_2$ of T which agree on the Q_i agree on P. Let T' be a theory just like T except that it has P' everywhere that T had P. If L_1 is the language of T and L_2 is the language of T', then $L_1 \cap L_2 = L = \{Q_1, \ldots, Q_n\}$. Clearly $T \cup T' \vdash P \rightarrow P'$ since in every model of $T \cup T'$, P and P' have the same extension. Hence by the interpolation theorem, there is a formula ϕ in the language $Q_1, \ldots, Q_n$ alone such that $T \vdash P \rightarrow \phi$ and $T' \vdash \phi \rightarrow P'$. Replacing P' by P in these proofs makes T and T' the same, and makes P and P' the same. Thus we get $T \vdash P \leftrightarrow \phi$. □

Both Beth and Craig used rather syntactic arguments, but the result can also be obtained model theoretically, e.g., from the Robinson joint consistency theorem (Hodges 1993, p. 301; Robinson 1956).

Theorem 1.2 *Let L_1, L_2 be two languages and $L = L_1 \cap L_2$. Suppose T is a complete theory in L and T_1, T_2 are consistent extensions of T in L_1, L_2, respectively. Then $T_1 \cup T_2$ is consistent in the language $L_1 \cup L_2$.*

The interpolation theorem can be generalized via an easy induction on n to the parallel interpolation theorem which goes as follows.

[1] Beth imposed the condition that the n above is not 0. But this condition is quite weak—even the truth functional constant T as one of the Q_i will serve.

Theorem 1.3 Parallel interpolation theorem (Kourousias and Makinson 2007): *Suppose that $\phi_1, \ldots, \phi_n$ are formulas in languages $L_1, \ldots, L_n$, respectively and the L_i are pairwise disjoint. Suppose $\{\phi_1, \ldots, \phi_n\} \vdash \psi$ where ψ is an L-formula. Then there exist formulas $\xi_1, \ldots, \xi_n$ in languages $L_i \cap L$ such that $\phi_i \vdash \xi_i$ for each i, and $\{\xi_1, \ldots, \xi_n\} \vdash \psi$.*

To use a somewhat colorful metaphor, if a king is receiving advice from a general and an economist, it would be sufficient for the general to submit a battle plan and for the economist to submit a budget. Any advice which the general gave about budgets or the economist gave about battles could simply be ignored by the king.

Question: Is there also a parallel counterpart to the Beth definability theorem?

For more recent reviews of the material, please see Craig (2008) and van Benthem (2008).

2 Belief revision

Much current work in the study of belief revision goes back to a now classic paper due to Alchourron et al. (1985). The central issue is how to revise an existing set of beliefs T to a new set of beliefs $T * A$ when a new piece of information A is received. If A is consistent with T, then it is easy; we just add A to T and close under logical inference to get the new set of beliefs. The harder problem is how to revise the theory T when a piece of information A *inconsistent* with T is received. Clearly, as Levi has suggested, T must first be contracted to a smaller theory $T' = T - \neg A$ which *is* consistent with A and then A added to T'. However, it is not clear how $T - \neg A$ should be obtained. The mere deletion of $\neg A$ from T will clearly not leave us with a theory and there is in general no unique way to get a theory T' which is contained in T and does not contain $\neg A$.

Suppose, for example, that I believe that country Saturnia is hot and country Urania is cold. Now I discover that the two countries have very similar climates. Do I drop my belief that Saturnia is hot or that Urania is cold? Clearly I cannot retain them both, and yet there seems no obvious basis for a preference either way.

The AGM approach does not actually tell us what to think about the two lands in question. What it does tell us is *if* we do have some procedure for updating, what logical properties such a procedure should satisfy. These properties (the AGM axioms) have been widely studied and model theoretic results proved for them. Yet some issues remain.

Notation In the following, L is a finite propositional language.[2] We assume that the constants *true, false* are in L.

We shall use the letter L both for a set of propositional symbols and for the formulae generated by that set. It will be clear from the context which is intended. $A \Leftrightarrow B$ means that A and B are logically equivalent, i.e. that $A \leftrightarrow B$ is a tautology, i.e. true

[2] This restriction is only made for convenience. The results continue to hold for a countably infinite first order language without equality.

under all truth assignments. Similarly, $A \Rightarrow B$ means that $A \rightarrow B$ is a tautology. If X is a set of formulae then $Con(X)$ is the logical closure of X. In particular, X is a theory iff $X = Con(X)$. Suppose we take $Mod(X)$ to be the set of all models of a set X of formulas, and, given a set S of truth assignments, we let $Th(S)$ to be the set of formulas which hold in all of S. Then $Con(X) = Th(Mod(X))$.

We shall use letters T, T' etc. for theories. $T * A$ is the revision of T by A, and finally, $T + A$ is $Con(T \cup \{A\})$, i.e. the result of a brute addition of A to T (followed by logical closure) without considering the need for consistency.

AGM have proposed the following widely accepted axioms for the revision operator $*$:

1. $T * A$ is a theory.
2. $A \in T * A$
3. If $A \Leftrightarrow B$, then $T * A = T * B$.
4. $T * A \subseteq T + A$
5. If A is consistent with T, i.e. it is not the case that $\neg A \in T$, then $T * A = T + A$.
6. $T * A$ is consistent if A is.
7. $T * (A \wedge B) \subseteq (T * A) + B$
8. If $\neg B \notin T * A$ then $(T * A) + B \subseteq T * (A \wedge B)$

Unfortunately, the AGM axioms are consistent with the trivial update, which is defined by:
If A is consistent with T, then $T * A = T + A$, otherwise $T * A = Con(A)$.
Thus in case A is inconsistent with T, under this update, all information in T is simply discarded.[3] Clearly this is unsatisfactory because we would like to keep as much of the old information as is feasible. Hence the AGM axioms need to be supplemented to rule out the trivial update. However, various actual proposals have run into trouble, either by being too flexible and allowing implausible update operators, or worse, by allowing *only* the trivial update.

We propose axioms for update operators which are consistent with the AGM axioms and which block the trivial update. The axioms are based on the notion of *splitting languages*. We shall explain first the intuitive idea behind this. The existing set of beliefs T may contain information about various matters. E.g. my current state of beliefs contains beliefs about the location of my children, the state of health of my teeth, and beliefs about the forthcoming election in India. In case one of my beliefs about the location of my children turns out to be false, it surely ought not to affect my beliefs about the election, since the subject matters of the two beliefs do not interact in any way. In order to model this intuition mathematically, we need to define in a rigorous way what it means to say that some given set of beliefs can be split among various unrelated matters. The notion of splitting languages does this for us.[4] Intuitively a theory T in language L *splits* if L is a union of two or more disjoint sub-languages, and the beliefs in T are generated by separate beliefs in the various sub-languages.

[3] This fact was noticed independently by Ryan (1996). A much stronger triviality result appears in Tennant (2006).

[4] This is a very natural notion and indeed we suspect that the notion of splitting has a wider application than just in belief revision.

Definition 1 (1) Suppose T is a theory in the language L and let $\{L_1, L_2\}$ be a partition of L. We shall say that L_1, L_2 *split* the theory T if there are formulae A, B such that A is in L_1, B is in L_2 and $T = Con(A, B)$. Similarly we say that (mutually disjoint) languages $L_1, L_2, \ldots, L_n$ split T if there exist formulae $A_i \in L_i$ such that $T = Con(A_1, \ldots, A_n)$. We may also say that $\{L_1, \ldots, L_n\}$ is a T-splitting.

(2) If $L_1 \subset L$ then we say that T is *confined* to L_1 if $T = Con(T \cap L_1)$. Note that in that case T also splits between L_1 and $L - L_1$, with the $L - L_1$ part being trivial, i.e. any formula of $L - L_1$ which is a theorem of T will be a tautology.

In part 1 of the definition, we can think of T as being generated by the various T_i in languages L_i. Then the condition implies that T contains no "cross-talk" between L_i and L_j for distinct i, j. Part 2 of the definition says that T knows nothing about the part $L - L_1$ of L.

Remark If P and P' are partitions of L, P is a T-splitting and P refines P' then P' will also be a T splitting.[5] For example suppose that $P = \{L_1, L_2, L_3\}$ is a T-splitting and let $P' = \{L_1 \cup L_2, L_3\}$. Then P' is a 2-element partition, P is a 3-element partition which refines P' and P' is also a T-splitting. For let $T = Con(A_1, A_2, A_3)$ where $A_i \in L_i$ for all i. Then $T = Con(A_1 \wedge A_2, A_3)$ and $A_1 \wedge A_2 \in L_1 \cup L_2$ so that P' is also a T-splitting.

Example Let $L = \{P, Q, R, S\}$, and $T = Con(P \wedge (Q \vee R))$. Then $T = Con(P, Q \vee R)$, and the partition $\{\{P\}, \{Q, R\}, \{S\}\}$ will be (the finest) T-splitting. $\{\{P, Q, R\}, \{S\}\}$ is also a T-splitting, but not the finest. Also, T is confined to the language $\{P, Q, R\}$ and knows nothing about S. Note that Q and R are entangled and cannot be separated.

Lemma 2.1 (Kourousias and Makinson 2007; Parikh 1999): *Given a theory T in the language L, there is a unique finest T-splitting of L, i.e. one which refines every other T-splitting.*

This lemma says that there is a unique way to think of T as being composed of disjoint information about certain subject matters. The proof is heavily dependent on the interpolation theorem. Our original result in Parikh (1999) considered only the case where L is finite. Kourousias and Makinson (2007) extended this result to the case where L is infinite.

Lemma 2.2 (Parikh 1999): *Given a formula A, there is a smallest language L' in which A can be expressed, i.e., there is $L' \subseteq L$ and a formula $B \in L'$ with $A \Leftrightarrow B$, and for all L'' and B'' such that $B'' \in L''$ and $A \Leftrightarrow B''$, $L' \subseteq L''$.*

Although A is equivalent to many different formulas in different languages, lemma 2 tells us that nonetheless, the question, "What is A actually *about*?" can be uniquely

[5] P *refines* P' if every element of P is a subset of some element of P'. Equivalently, the equivalence relation corresponding to P extends the equivalence relation corresponding to P'. P will have smaller members than P' does and more of them.

answered by providing a smallest language in which (a formula equivalent to) A can be stated.

We are skipping most of the proofs since they already occur in the originals, but the proof of the preceding lemma is simple and shows how the interpolation theorem enters, so we give it here.

Proof The languages in which A is expressible are partially ordered by inclusion, and some are finite. Thus there is a minimal language L in which A can be expressed. To show that L is in fact a *minimum* we argue as follows. Let L' be another minimal language in which A can also be expressed. Specifically, let B be an L-formula equivalent to A, and C be an L'-formula equivalent to A. Then $B \vdash C$ and by the interpolation theorem, there is an $L \cap L'$-formula D such that $B \vdash D \vdash C$. Now B, C, D are all equivalent, so D is also equivalent to A and A is expressible in $L \cap L'$. Now, since L was supposed to be minimal, $L \cap L'$ cannot be strictly smaller than L. Hence $L \subseteq L'$ and L was minimum, not just minimal. □

The axioms:

The general rationale for the axioms is as follows. If we have information about two subject matters which, as far as we know, are unrelated (are split) then when we receive information about *one* of the two, we should only update our information in that subject and leave the rest of our beliefs unchanged. E.g. suppose I believe that Barbara is rich and Susan is beautiful and only that. Later on I meet Susan and realize that she is not beautiful. My beliefs about Barbara should remain unchanged since I do not connect Susan and Barbara in any way. If on the other hand I had initially believed that Barbara had made her money *as the agent* for Susan who was a beautiful model, then the two beliefs would be *connected* and finding that Susan was not beautiful could require an adjustment also of my beliefs about Barbara's wealth.

In fact, the notion of language splitting seems intrinsic to any attempt to form a theory of anything at all. Any observation or experiment gives us an enormous amount of information. E.g. when we are dealt a hand of cards, we are dealt them in a certain order, either by the right hand or the left hand of the dealer, who may have grey or brown or blue eyes. We usually ignore all this extra information and concentrate on the *set* of cards received. There is a tacit assumption, for instance, that the color of the dealer's eyes will not affect the probability that the hand contains two aces. This assumption that we can ignore some aspects while we are considering others is inherent in almost all intellectual activity.[6]

Now we give our new axioms P1–P3, giving intuitive justification for each. We also give a single axiom P which implies all of P1–P3.

Axiom P1: If T is split between L_1 and L_2, and A is an L_1 formula, then $T * A$ is also split between L_1 and L_2.

[6] The third chapter of Cherniak's (1986) book gives arguments why beliefs must thus be divided into subsets, and cites supporting statements from van Orman Quine and Ullian (1978) as well as from Simon (1947).

Justification: The two subject areas L_1 and L_2 were unconnected. We have not received any information which *connects* these two areas, so they remain separate.

Axiom P2: If T is split between L_1 and L_2, A, B are in L_1 and L_2, respectively, then $(T * A) * B = (T * B) * A$.

Justification: Since A and B are unrelated, they do not affect each other and so it should not matter in which order they are received.

The condition that L_1, L_2 are disjoint is essential. For suppose they were not disjoint and A and B were inconsistent with each other. Then $(T * A) * B$ contains B and hence cannot contain A. On the other hand $(T * B) * A$ contains A and hence cannot contain B. Thus they are different.

Axiom P3: If T is confined to L_1 and A is in L_1 then $T * A$ is just the consequences in L of $T *' A$ where $*'$ is the update of T by A in the sub-language L_1.

Justification: Since we had no information about $L - L_1$ and have received none in this round, we should update as if we were in L_1 only. $L - L_1$, about which we have no prior opinions and no new information, should simply not have any impact.

All these axioms follow from axiom P, below.

Axiom P: If $T = Con(A, B)$ where A, B are in L_1, L_2, respectively and C is in L_1, then $T * C = Con(A) *' C + B$, where $*'$ is the update operator for the sub-language L_1.

Justification: We have received information only about L_1 which does not pertain to L_2 so we should revise only the L_1 part of T and leave the rest alone.

It is easy to see that P implies P1. To see that P3 is implied, we use the special case of P where the formula B is the trivial formula *true*. To see that P implies axiom P2, suppose T is split between L_1 and L_2, and A, B are in L_1 and L_2, respectively. Let $T = Con(C, D)$ where $C \in L_1$ and $D \in L_2$. Then we get $T * A * B = (T * A) * B =_P [(Con(C) *' A) + D] * B =_P (Con(C) *' A) + (Con(D) *'' B)$. The two occurrences of $=_P$ indicate where we used the axiom P. Now the last expression $(Con(D) * B) + (Con(C) * A)$ is symmetric between the pairs (C, A) and (D, B) and calculating $T * B * A$ yields the same result.

Remark The trivial update procedure cannot satisfy P2 (or P), though it does satisfy P1 and P3. It follows that any procedure that does satisfy P cannot be the trivial procedure.

Justification: Let $T = Con(P, Q)$ and let $A = P$ and $B = \neg Q$. Then the trivial update yields $T * A * B = T * B = Con(\neg Q)$ and $T * B * A = Con(B) * A = Con(\neg Q, P)$. This violates P2. Also, $T * B = Con(\neg Q)$ which violates P. Thus P, or P2 alone, rules out the trivial update.

Corollary *If T in language L splits among theories $T_1, \ldots, T_k$ in languages $L_1, \ldots, L_k$, and formulas $A_1, \ldots, A_k$ in $L_1, \ldots, L_k$, respectively are received, then the order in which the A_i are received will not matter, and the result of revising T by all the A_i will be the same as the result of separately revising each T_i with the corresponding A_i and then adding up the results.*

In the next theorem we shall restrict the AGM axioms to the case of those updates where both T and A are individually consistent and only their union might not be. This is because we can suppose that our current state of belief T about the subject matter of L originates in a state T_0 where we only believe tautologies (or if not, T_0 is at least consistent) and T is obtained from T_0 through zero or more revisions. Suppose that at some stage we are told an inconsistent formula A. Then axiom 3 (the irrelevance of syntax) tells us that this is equivalent to being told a blatant contradiction like $P \wedge \neg P$ and we would simply not believe A in that case. Hence if A is inconsistent, then $T * A$ should be just T. The AGM axioms in their original form *force* that $T * true = T$ for consistent T and *disallow* it for an inconsistent T. Our restriction has the fortunate consequence that $T * true$ always equals T.

Definition 2 Given a theory T, language L and formula A, let L'_A be the smallest language in which A can be expressed and L^T_A be the smallest language containing L'_A such that $\{L^T_A, L - L^T_A\}$ is a T-splitting. Thus L^T_A is a union of certain members of the finest T-splitting of L, and in fact the smallest in which A can be expressed.

Example Let $L = \{P, Q, R, S\}$, and $T = Con(P \wedge (Q \vee R) \wedge S)$. Then $T = Con(P, Q \vee R, S)$ and $\{\{P\}, \{Q, R\}, \{S\}\}$ will be the finest T-splitting. If A is the formula $P \vee \neg Q$, then L'_A is the language $\{P, Q\}$. But L^T_A is the smallest language *compatible* with the T-splitting, in which A can be expressed. Thus it will be the larger language $\{P, Q, R\}$ which is the union of the sets $\{P\}$ and $\{Q, R\}$ of the finest T-splitting.

Theorem 2.3 (Parikh 1999; Peppas et al. 2004)*: There is an update procedure which satisfies the eight AGM axioms and axiom P.*

Example Let, as before, $T = Con(P, Q \vee R, S)$. Then the partition $\{\{P\}, \{Q, R\}, \{S\}\}$ is the finest T-splitting. Let A be the formula $\neg P \wedge \neg Q$, then L^T_A is the language $\{P, Q, R\}$. Thus B will be the formula $P \wedge (Q \vee R)$ and $C = S$. B represents the part of T incompatible with the new information A. Thus $T * A$ will be $Con((\neg P \wedge \neg Q), S)$. The update procedure notices that A has no quarrel with S and keeps it. As we will see, axiom P requires us to keep S.

Remark In this update procedure we used the trivial update on the sub-language L^T_A, but we did not need to. Thus suppose we are given certain updates $*_{L'}$ for sub-languages L' of L. We can then build a new update procedure $*$ for all of L by letting $T * A = (B *_{L'} A) + C$ in the proof above, where $L' = L^T_A$. What this does is to update B by A on L^T_A according to the old update procedure, but preserves all the information C in $L - L^T_A$.

3 Information content

Given a (consistent) theory T on a finite language L we can define the information content Inf(T) of T to be $log(2^n/|Mod(T)|)$, i.e., the logarithm (to base 2) of the total number of truth assignments on L (namely 2^n) divided by the number of models of T. If T is complete, then Inf(T) is n. If T consists only of tautologies, then Inf(T) is 0. Note that if T is expressed in language L, then Inf(T) will be the same for *any* language L' which includes L. If T is a subtheory of T' then as expected, T' has more information.

Suppose now that T is an L theory which splits into theories $T_1, \ldots, T_k$ in languages $L_1, \ldots, L_k$, respectively. It is easily seen that $\text{Inf}(T) = \Sigma_{i\le k} \text{Inf}(T_i)$. Moreover, if T is revised by (say) an L_k formula A, and axiom P is obeyed, then $\text{Inf}(T * A)$ will be *at least* $[\Sigma_{i<k} \text{Inf}(T_i)]+\text{Inf}(A)$. Thus simply in terms of the *amount* of information, axiom P forces most of the information to be saved and adds the information which A brought.

Suppose now that L is the disjoint union of $L_1, \ldots, L_k$ but T does not split into subtheories. In that case, letting $T_i = T \cap L_i$ we get $\text{Inf}(T) > \Sigma_{i\le k} \text{Inf}(T_i)$. For instance let $T = Con(P \vee Q)$ and $L = \{P, Q\}$ with $L_1 = \{P\}$ and $L_2 = \{Q\}$. Then T_1, T_2 are both trivial and $\text{Inf}(T_1) = \text{Inf}(T_2) = 0$. However, T only has three models from the four truth assignments and hence $\text{Inf}(T) > 0.4$.

Theorem 3.1 *Suppose T is a theory in language L which is the disjoint union of $L_1, \ldots, L_k$. Let $T_i = T \cap L_i$. Then $T_1, \ldots, T_k$ split T (i.e. T splits in $L_1, \ldots, L_k$) iff $Inf(T) = \Sigma_{i\le k} Inf(T_i)$.*

Proof Let $\mathcal{M}$ be a model of T and $\mathcal{M}_i$ be $\mathcal{M}$ restricted to L_i. Then each of the $\mathcal{M}_i$ must satisfy T_i. Thus every model of T is put together from models of the T_i, but not every such 'putting together' will necessarily yield a model of T, *unless* the T_i actually split T. Hence $|Mod(T)| \le |Mod(T_1)| \times \cdots \times |Mod(T_k)|$. The two sides are equal iff $T = Con(T_1 \cup T_2 \cdots \cup T_k)$, i.e. if T splits into the languages L_i. Otherwise T has strictly more information that the T_i put together. □

4 Belief revision, language splitting and grove spheres

A technique invented by Adam Grove related belief revision with what are now called Grove spheres. Given a theory T, we can classify all models (equivalently all complete theories) into various spheres such that

(1) The spheres are linearly ordered by inclusion.
(2) $Mod(T)$ is the smallest Grove sphere. The largest Grove sphere consists of all models.
(3) Given a consistent formula A, there is a *smallest* Grove sphere which contains a model of A.

Such a system of spheres for T will be denoted S_T. *Caution:* S_T is not unique but corresponds to a particular update function *.

Now, given a theory T and a consistent formula A, we consider the smallest Grove sphere S which contains a model of A. Such a sphere must exist by (3) above. Then we define $T * A = Th(S \cap Mod(A))$, i.e., $T * A$ is the set of all formulas which are true in all members of S which satisfy A.

Now note that $T * A$ automatically contains A. Moreover, if A is consistent with T, then S will be simply $Mod(T)$ and $T * A = T + A$.

Grove (1988) showed that the update operation * defined this way satisfies all eight AGM axioms, and that every such update operation can be obtained from a suitable system of Grove spheres.

Peppas et al. (2004) correlate our axiom P with Grove spheres in the following manner. In the following, L_D is the smallest language in which some formula D (or equivalent) is expressible and L_D^c is the complement of L_D. The axiom P implies condition (R1) below.

(R1) If $T = Con(B, D)$, $L_B \cap L_D = \emptyset$ and $A \in L_D$, then $(T * A) \cap L_D^c = T \cap L_D^c$.

In other words, the part of T outside L_D is unaffected by the update. Then for complete theories T, Peppas et al. (2004) show a correlation (the next theorem) between (R1) and the principle (PS) below. Let us use letters r, s to denote truth assignments, (i.e., complete theories), and let $\text{Diff}(r, s) = \{P : r(P) \neq s(P)\}$. In other words, the difference set is the set of those propositional variables to which r, s assign different values. Let S_T be some family of Grove spheres for T. We consider now the condition (PS):

(PS) If $\text{Diff}(T, r) \subset \text{Diff}(T, s)$ then there is a Grove sphere in S_T which contains r but not s.

Note that T, being complete, is also itself a truth assignment.

In other words, (PS) requires that truth assignments which differ more from T go to outer spheres compared to those which differ less.[7]

Theorem 4.1 (Peppas et al. 2004): *Let * be a revision function satisfying the eight AGM axioms, T a complete, consistent theory, and S_T be the system of spheres for T corresponding to *, then * satisfies (R1) iff S_T satisfies (PS).*

This gives us a nice connection between Grove spheres of a certain well behaved kind and axiom, Peppas et al. (2004) also have a characterization for incomplete theories. But since an incomplete theory is not a truth assignment, the notion Diff is considerably more complex and we omit their discussion here.

Note that (PS) considers atomic formulas to be more entrenched, or more central than others. We could of course consider a situation where some formula $P \leftrightarrow Q$ was better entrenched than either P or Q separately. While we do not know any results with such more general sorts of entrenchments, the issue is important. It happens quite often that compound formulas are better entrenched than atomic ones.

It is pleasant to report that some of the results on splitting described here as well as in earlier work have now appeared in an undergraduate textbook (Makinson 2008).

Acknowledgment We thank William Craig, David Makinson, Johan van Benthem and Noson Yanofsky for useful comments. This research was partly supported by a grant from the PSC-CUNY FRAP program.

References

Alchourron, C., Gärdenfors, P., & Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic, 50*, 510–530.

Beth, E. W. (1953). On Padoa's method in the theory of definition. *Indagationes Mathematicae, 15*, 330–339.

Cherniak, C. (1986). *Minimal rationality*. Cambridge: MIT Press.

[7] Here more and less is in terms of set theoretic inclusion, and not in terms of numerical size.

Chopra, S., & Parikh, R. (2000). Relevance sensitive belief structures. *Annals of Mathematics and Artificial Intelligence, 28*, 259–285.

Craig, W. (1956). Review of E. W. Beth, On Padoa's method in the theory of definition. *The Journal of Symbolic Logic, 21*(2), 194–195.

Craig, W. (1957). Three uses of the Herbrand-Gentzen theorem in relating model theory and proof theory. *The Journal of Symbolic Logic, 22*(3), 269–285.

Craig, W. (2008). The road to two theorems of logic. *Synthese, 164*, 333–339.

Grove, A. (1988). Two modellings for theory change. *Journal of Philosophical Logic, 17*, 157–170.

Heyting, A. (1966). In memoriam, Evert Willem Beth. *Notre Dame Journal of Formal Logic, 4*, 289–295.

Hodges, W. (1993). *Model theory*. Cambridge: Cambridge University Press.

Kourousias, G., & Makinson, D. (2007). Parallel interpolation, splitting, and relevance in belief change. *Journal of Symbolic Logic, 72*(3), 994–1002.

Mark, R. (1996). Belief revision and ordered theory presentations. In A. Fuhrmann, & H. Rott (Eds.), *Logic, action, information* (pp. 129–151). Berlin: De Gruyter.

Makinson, D. (2008). *Sets, logic and maths for computing (undergraduate topics in computer science)*. Belin: Springer.

Parikh, R. (1999). Beliefs, belief revision, and splitting languages. In L. Moss, J. Ginzburg, & M. de Rijke (Eds.), *Proceedings of logic, language and computation, CSLI 1999* (pp. 266–278).

Peppas, P., Chopra, S., & Foo, N. (2004). Distance semantics for relevance-sensitive belief revision. In *Proceedings of the international conference on knowledge representation and reasoning, KR2004*.

Robinson, A. (1956). A result on consistency and its application to the theory of definition. *Indagationes Mathematicae, 18*, 47–58.

Simon, H. (1947). *Administrative behaviour*. New York: Macmillan.

Tennant, N. (2006). On the degeneracy of the full AGM-theory of theory-revision. *Journal of Symbolic Logic, 71*(2), 661–676.

van Benthem, J. (2008). The many faces of interpolation. *Synthese, 164*, 451–460.

van Orman Quine, W., & Ullian, J. (1978). *The web of belief* (2nd Ed.). New York: Random house.