Iterated Belief Revision, Reliability, and Inductive Amnesia

Kevin T. Kelly

Department of Philosophy

Carnegie Mellon University

September 24, 2001

Abstract

Belief revision theory concerns methods for reformulating an agent's epistemic state when the agent's beliefs are refuted by new information. The usual guiding principle in the design of such methods is to preserve as much of the agent's epistemic state as possible when the state is revised. Learning theoretic research focuses, instead, on a learning method's reliability or ability to converge to true, informative beliefs over a wide range of possible environments. This paper bridges the two perspectives by assessing the reliability of several proposed belief revision operators. Stringent conceptions of "minimal change" are shown to occasion a limitation called *inductive amnesia*: they can predict the future if and only if they cannot remember the past. Avoidance of inductive amnesia can therefore function as a plausible and hitherto unrecognized constraint on the design of belief revision operators.

0.1 Introduction

According to the familiar, Bayesian account of probabilistic updating, full beliefs change by *accretion*: in light of new information consistent with one's current beliefs, one's new belief state is the the result of simply adding the new information to one's current beliefs and closing under deductive consequence. Inductive generalizations that extend both one's current beliefs and the new information provided are not licensed, although the new information may increase the agent's *degree* of belief in such a proposition.[1] This account breaks down when new information contradicts the agent's current beliefs, for accretive updating leads, in this case, to a contadictory belief state from which further accretion can never escape. *Belief revision theory* aims to provide an account of how to update full belief so as to preserve consistency when one's current beliefs are refuted by the new information provided. Belief revision theory has attracted attention in a number of areas, including data base theory (Katsuno and Mendelson 1991), the theory of conditionals (Boutilier 93; Levi 96; Arlo-Costa 1997), the theory of causation (Spohn 1988, 1990; Goldszmidt and Pearl 94), and game theory (Samet 1996).

A *belief revision operator* is a rule for modifying an agent's overall *epistemic state* in light of new information. An agent's epistemic state determines an assignment of *degrees of implausibility* to possible worlds. The agent's *belief state* is taken to be the proposition satisfied by all and only the possible worlds of implausibility degree zero. Proposed belief revision operators differ markedly as to how they update the overall epistemic state, but they all agree about how to revise the current belief state: the new belief state is the proposition satisfied by all and only the most plausible possibilities satisfying the newly

[1]A broadly Bayesian perspective is not bound to identify inductive methodology with updating by conditionalization. Genuinely inductive expansions may be justified by decision-theoretic considerations based on epistemic utility (e.g., Levi 81).

received information. According to this rule, the character of the revised belief state depends on the character of the agent's initial epistemic state. If all possible worlds are assigned implausibility degree zero, the agent starts out as a *tabula rasa* with vacuous beliefs and updates by mere accretion, without taking any inductive risks. At the opposite extreme, consider an agent whose initial epistemic state is maximally *refined*, in the sense that all possible worlds are assigned distinct degrees of implausibility. Such an agent starts out fully convinced of a complete theory and retains this conviction until the theory is refuted, at which point she replaces it with the complete theory of the most plausible world consistent with the new information. The new theory may differ radically from its predecessor. Described this way, belief revision sounds like a process of "eliminative" induction, in which a "bold conjecture" is retained until it is refuted, after which it is replaced with the first alternative theory (in a subjective "plausibility ranking") that is consistent with the new information provided (Popper 68; Kemeny 53; Putnam 63; Gold 67; Earman 92). Between these two extremes are agents with moderately refined initial states whose inductive leaps from one theory to another are correspondingly weaker.

The belief revision literature has focused on the aim of minimizing change in the agent's epistemic (or belief) state when new information contradicting the agent's beliefs is received. The similarity between belief revision and eliminative induction suggests a natural, alternative aim for belief revision: namely, to arrive at informative, true, empirical beliefs on the basis of increasing information. This aim is largely unexplored in the belief revision literature,[2] but it has long been the principal focus of *formal learning theory*, the study of processes of sequential belief change that are *reliable*, or guaranteed to stabilize to true, informative beliefs on the basis of increasing information. The purpose of this paper is to bring learning theoretic analysis to bear on a variety of iterated belief

[2]It is raised, informally, in (Gärdenfors 1988).

revision operators proposed by Spohn (1988), Boutilier (1993), Nayak (1994), Goldszmidt and Pearl (1994), and Darwiche and Pearl (1997).[3] A very simple model of learning is employed, in which the successive propositions received by the agent are true reports of successive outcomes of some discrete, sequential experiment. An *inductive problem* specifies (1) what counts as a sufficiently informative belief state and (2) how the outcome sequence might possibly evolve in the unbounded future. The agent's task is to stabilize to sufficiently informative, true beliefs about the outcome sequence, for *each* outcome sequence admitted by the inductive problem.

The investigation yields an interesting mixture of positive and negative results. Some of the operators are *empirically complete*, in the sense that for each solvable learning problem, there exists an initial epistemic state for which the operator solves it. Others *restrict* reliability, in the strong sense that there are solvable learning problems that they cannot solve no matter how cleverly we adjust their initial epistemic states. All of the restrictive belief revision operators considered can have their initial epistemic states adjusted so that they remember the past, and nearly all of them can be adjusted to eventually predict the future. So such an operator has the odd property that it can remember the past perfectly but then it cannot eventually predict the future and it can eventually predict the future, but then it forgets some of the past. I refer to this limitation as *inductive amnesia.* Inductive amnesia is the sort of thing we would like rules of rationality to protect us from rather than impose on us.[4] Avoiding it can therefore function as a well-motivated constraint on proposed methods and principles of belief revision.

[3]For earlier applications of learning theoretic analysis to belief revision theory, cf. (Martin and Osherson 1995, 1996) and (Kelly, Schulte and Hendricks 1996).

[4]The fact that the operators can all be programmed to possess perfect memory blocks the response that inductive amnesia is a matter of resource bounds rather than a defect in the updating rules subject to it.

Among the inductively amnestic belief revision operators, it is of interest to determine which are more restrictive than others. To answer these questions, I introduce a hierarchy of increasingly difficult inductive problems based on the number of applications of Nelson Goodman's (1983) "grue" operation, which reverses the binary outcomes in a data stream from a given point onward. For each of the belief revision operators considered, I determine the hardest problem in this *grue hierarchy* that it can solve, obtaining, thereby, an objective measure of its reliability.

It might be expected that a global consideration such as eventually finding the truth would impose only the loosest short-run constraints on concrete belief revision procedures. However, sharp and unexpected recommendations are obtainable. For example, some proposed belief revision operators are equipped with a parameter α, which is the amount by which the implausibility of a possibility is increased when the possibility is refuted. Lower values of α may be interpreted as more stringent notions of "minimal" change since they correspond, in a sense, to less distortion of the original epistemic state.[5] Two of these operators (Spohn 1988, 1990; Darwiche and Pearl 1997) turn out to fail by the second level of the grue hierarchy if $\alpha = 1$ but succeed over the entire, infinite grue hierarchy if α is incremented to 2. So although the difference between 1 and 2 is innocuous in light intuitive coherence and symmetry considerations, it marks an infinite improvement in learning power. It will be argued, moreover, that this result reflects a deep, epistemological tension between memory and prediction faced by iterated belief revision operators of the sort under consideration.

The purpose of this paper is not to argue that reliability considerations always win

[5]The parameter α may also be viewed as an assessment of the quality or reliability of the input information. Under that interpretation, the following results concern the minimal quality of data necessary for finding the truth.

when they conflict with coherence, symmetry, or minimality of belief change. As in every case of conflicting aims, a personal balance must be sought. But if the ultimate balance is subjective, structural conflicts between intuitive rationality considerations and reliability are objective. The isolation and investigation of such conflicts is therefore a suitable aim for objective, epistemological analysis. The following results are preliminary and subject to generalization and refinement along a number of dimensions. Nonetheless, they illustrate how reliability analyses can usefully and routinely be carried out for proposed theories of iterated belief revision.

0.2 Ordinal Implausibility

Let W be a set of possible worlds.[6] The agent's *epistemic state* at a given time is modelled as an *implausibility assignment (IA)*, which is a (possibly partial) ordinal-valued function r defined on W.[7] Possibilities that are not even in the domain of r are "beyond possible consideration" in the strong sense that they will never be consistent with the agent's belief state, no matter what information the agent might encounter in the future. For a given world w, let $[w]_r$, $[w]_r^{\leq}$, and $[w]_r^{<}$ denote, respectively, the set of all worlds equally, no more, or less implausible than w.

A proposition is identified with the set of all possible worlds satisfying it. The full belief state of r is defined to be the proposition satisfied exactly by the possible worlds of implausibility zero.

$$b(r) = r^{-1}(0).$$

[6]The approach adopted in this section follows Spohn (1988).

[7]It is not generally accepted that degrees of implausibility are well-ordered. This assumption will be dropped in section 0.8.

Define the minimum degree of implausibility of worlds in E as follows:

$$r_{\text{min}}(E) = \min\{r(w) : w \in E \ \cap \ \text{dom}(r)\}.$$

It will also be convenient to refer to the lowest degree of implausibility that is strictly greater than the implausibility of each world in E:

$$r_{\text{above}}(E) = \min\{\alpha : \forall w \in E \cap \ \text{dom}(r), r(w) < \alpha\}.$$

If $\alpha \leq \beta$ then $-\alpha + \beta$ denotes the unique γ such that $\alpha + \gamma = \beta$ (i.e., $-\alpha + \beta$ is the order type of the "tail" that remains when the initial segment α is "deleted" from β). Given implausibility assignment r, we may define $r(.|E)$ to be an ordinal valued function with domain $\text{dom}(r) \cap E$ such that for each w in this domain:

$$r(w|E) = -r_{\text{min}}(E) + r(w).$$

Then $r_{\text{min}}(A|E)$ and $r_{\text{above}}(A|E)$ may be defined as follows:

$$r_{\text{min}}(A|E) = (r(.|E))_{\text{min}}(A).$$

$$r_{\text{above}}(A|E) = (r(.|E))_{\text{above}}(A).$$

0.3 Some Iterated Belief Revision Operators

An *iterated belief revision operator* takes an IA r together with an input proposition E and returns an updated IA r'.

I will analyze the following examples. Perhaps the most obvious idea is simply to eliminate refuted worlds from one's ranking and to lower all the other worlds, keeping intervals of relative implausibility fixed, until the most plausible world touches bottom. This is what Spohn (1988) refers to as the conditional implausibility ranking given the data.

Definition 1 (conditioning) $r *_C E = r(.|E)$.

Conditioning throws away refuted worlds, so it cannot recover when later data contradict earlier data. The remaining proposals boost the implausibility of refuted worlds rather than disposing with them altogether. An idea very similar to conditioning retains the refuted worlds but sends them all to a safe "point at infinity" never to be seen again unless past data are contradicted by future data. It will prove interesting to analyze a generalization of this proposal in which all refuted worlds are assigned a fixed ordinal α.

Definition 2 (The "all to α" operator)

$$(r *_{A,\alpha} E)(w) = \begin{cases} r(w|E) & \textit{if } w \in \ \textit{dom}(r) \cap E \\ \alpha & \textit{if } w \in \ \textit{dom}(r) - E \\ \uparrow & \textit{otherwise.} \end{cases}$$

Another proposal boosts all refuted worlds just above all the non-refuted worlds, maintaining intervals of implausibility among refuted worlds and among non-refuted worlds but not between the two classes.

Definition 3 (The lexicographic operator)

$$(r *_L E)(w) = \begin{cases} r(w|E) & \textit{if } w \in \ \textit{dom}(r) \cap E \\ r_{above}(E|E) + r(w|W - E) & \textit{if } w \in \ \textit{dom}(r) - E \\ \uparrow & \textit{otherwise.} \end{cases}$$

A variant of this operator was defined by Spohn (1988), who rejected it because it is irreversible, fails to commute (the resulting IA depends on the order in which the data arrive) and places extreme importance on the data (the refuted worlds are put above all the non-refuted worlds rather than being shuffled in). Against these considerations, S. M. Glaister (1997) has argued that a generalization of this rule due to Nayak (1994) is uniquely characterized by plausible symmetry conditions.

At the opposite extreme, consider the operator that drops the lowest worlds consistent with the new information to the bottom level, and that rigidly elevates all other worlds by one step, keeping their relative positions to one another fixed.

Definition 4 (The "minimal" operator)

$$(r *_M E)(w) = \begin{cases} 0 & \textit{if } w \in E \cap b(r(.|E)) \\ r(w)+1 & \textit{if } w \in \ \textit{dom}(r) - (E \cap b(r(.|E))) \\ \uparrow & \textit{otherwise.} \end{cases}$$

In a sense, this is the minimum alteration of the epistemic state consistent with the principle that one's new belief state be the set of all most plausible possibilities consistent with the new information. Boutilier's "natural operator" (1993) generalizes this operator to apply to total pre-orders on worlds rather than IAs.[8] Spohn (1998) describes an operator of this kind and rejects it. It doesn't fare better in terms of reversibility and commutativity and, in Spohn's opinion, places too little importance on the data, since the operator can easily end up admitting possibilities excluded by the information received at the previous stage.

Spohn recommends, instead, the following kind of operator. As usual, sort the worlds at each level into those that are refuted by the current evidence and those that are not. Lower both groups of worlds, preserving distances within the two groups, until the lowest worlds in each group are at the bottom level. Now raise all of the refuted worlds together so that the lowest refuted words end up at level α.[9] Spohn shows that this rule can be

[8] Boutilier's operator is considered in section 0.8 below. Boutilier also considers the problem of updating on conditionals, which is not addressed in this paper.

[9] This is actually a special case of Spohn's proposal. In general, Spohn's rule updates on a partition of possible worlds, with a separate α for each cell of the partition. It is assumed that one such α is zero. Here I present only the special case of binary partitions.

represented as updating a nonstandard probability measure by Jeffrey's rule, so long as there are but countably many possible worlds mapped to each degree of implausibility. The rule is also shown by Spohn to be both reversible and commutative (if α is understood to be an adjustable parameter). Nor is it as "extreme" as the preceding rules. But according to this rule, the implausibility of a refuted world may actually go down *when it is refuted*, if α is less than the world's current implausibility.

Definition 5 (The Jeffrey operator)

$$(r *_{J,\alpha} E)(w) = \begin{cases} r(w|E) & \text{if } w \in \ dom(r) \cap E \\ r(w|W - E) + \alpha & \text{if } w \in \ dom(r) - E \\ \uparrow & \text{otherwise.} \end{cases}$$

Goldszmidt and Pearl (1995) and Darwiche and Pear (1996) propose an interesting modification of Spohn's Jeffrey conditioning operator. Instead of dropping the refuted worlds to the bottom level before elevating them by α, the new proposal lifts the refuted worlds by α from their *current* position, whatever that might be. Since refuted worlds cannot backslide from their current position, I refer to this as the "ratchet" method.

Definition 6 (The ratchet operator) *Let α be an ordinal.*

$$(r *_{R,\alpha} E)(w) = \begin{cases} r(w|E) & \text{if } w \in \ dom(r) \cap E \\ r(w) + \alpha & \text{if } w \in \ dom(r) - E \\ \uparrow & \text{otherwise.} \end{cases}$$

Proponents of different belief revision operators have in mind different conceptions of minimal change and different assessments of the relative importance of minimality as opposed to other symmetry conditions. Such debates may be irresolvable. My purpose is to shift the focus of such debates to the relative abilities of the various operators to generate true, informative beliefs; a natural goal that distinguishes sharply and objectively between the above proposals.

0.4 Iterated Implausibility Revision as Inductive Inquiry

Iterated belief revision involves successive modifications of one's epistemic state as successive input propositions are received. Iteration of a belief revision operator over a sequence of propositions is defined recursively as follows:

$$\begin{aligned} r * () &= r. \\ r * (E_0, \ldots, E_n, E_{n+1}) &= (r * (E_0, \ldots, E_n)) * E_{n+1}. \end{aligned}$$

A belief revision agent starts out with an initial epistemic state r and sequentially updates her beliefs using a belief revision operator $*$, so we may identify the agent with a a pair $(r, *)$, which I refer to as an *implementation* of $*$. Such an agent determines a unique map from finite sequences of input propositions to new belief states as follows:

$$(r, *)((E_0, \ldots, E_n)) = b(r * (E_0, \ldots, E_n)).$$

In general, an *inductive method* is a rule that produces an empirical hypothesis in response to a finite sequence of input propositions:

$$f((E_0, \ldots, E_n)) = B_{n+1}.$$

Inductive methods are the usual objects of learning theoretic analysis. Since an implementation $(r, *)$ of a belief revision operator $*$ is a special kind of inductive method, it is directly subject to learning theoretic analysis.

0.4.1 Data Streams

Suppose a scientist who uses an inductive method f is faced with the task of studying the successive outcomes of experiments on some unknown system. We will suppose that the outcomes are discretely recognizable, and hence may be encoded by natural numbers.

The *data stream* generated by the system under study is just an infinite tape on which the code numbers of the successive outcomes of the experiment are written. The first datum arrives at stage 0, so a data stream is a total function e defined on the natural numbers. Let U denote the set of all data streams. An *empirical proposition* is a subset of U. In other words, the truth of an empirical proposition supervenes on the actual data stream.

Consider the scientist's idealized situation at stage n of inquiry. At that stage, she observes that the outcome for stage n is $e(n)$ and updates on the empirical proposition $[n, e(n)]$, which is defined to be the set of all data streams e' such that $e'(n) = e(n)$. The *initial segment* of the data stream scanned by stage n is

$$e|n = (e(0), \ldots, e(n-1)).$$

The *length* of of this sequence is defined to be n:

$$\mathrm{lh}(e(0), \ldots, e(n-1)) = n.$$

The *tail* of the data stream remaining to be scanned from stage n is:

$$n|e = (e(n), e(n+1), \ldots).$$

By stage n, the scientist updates on the sequence of empirical propositions

$$[[e|n]] = ([0, e(0)], \ldots, [n-1, e(n-1)]).$$

Then her inductive method's output at stage n is just

$$f([[e|n]]) = f(([0, e(0)], \ldots, [n-1, e(n-1)])).$$

Note that $[[e|n]]$ is not the same thing as the empirical proposition

$$[e|n] = \{e' \in U : e|n \text{ is extended by } e'\},$$

which states that the finite outcome sequence $e|n$ has occured. Rather, $[e|n]$ is the intersection of all the propositions $[i, e(i)]$ occurring in the sequence of propositions $[[e|n]]$. Now that these distinctions are clear, I will simplify notation by writing

$$f(e|n) = f([[e|n]]).$$

0.4.2 Empirical Questions

Inquiry has two cognitive aims, seeking truth and avoiding error.[10] Seeking truth involves relief from ignorance. One simple way to specify nontrivial content is to partition possibilities and to require that the outputs of the method eventually entail the true cell of this "target" partition. We may think of the partition as an *empirical question* and the cells of the partition as the potential *answers* to the question. Let Θ_0 denote the singleton partition $\{\{e\} : e \in U\}$, which corresponds to the hardest empirical question "what is the complete empirical truth?" and let Θ_1 denote the trivial question $\{U\}$, answered by vacuously true beliefs.

0.4.3 Reliability in the Limit

Given an empirical question Θ, one may hope that one's method is guaranteed to halt with a correct answer to Θ. But no bell rings when science has found the truth,[11] suggesting the weaker requirement that inquiry eventually stabilize to a correct answer to Θ, perhaps without ever knowing when it has done so. Then we say that the method *identifies* an answer to Θ on e, or that the method *identifies* Θ on e for short.

It is not enough that a method *happen* to stabilize to the right answer in the actual world: scientific success should be more than opinionated luck. *Reliability* demands that

[10] William James (1948), (Levi 1981).

[11] This charming phrase is from William James (1948).

a method succeed over some broad range K of possible data streams. One may think of K as the domain of the agent's initial epistemic state (i.e., the set of worlds that the agent might possibly admit as serious possibilities in the future). But one may also conceive of K simply as a range of possibilities over which the method can be shown to succeed, so that the method is more reliable insofar as K is larger (weaker). When the method identifies Θ on every data stream in K, we say that it identifies Θ *given* K. In the special case when the target partition is Θ_0, one may speak simply of identifying K.

Definition 7

1. *Method f identifies partition Θ given K just in case for each $e \in K$, for all but finitely many n, $e \in f(e|n) \subseteq \Theta(e)$.*

2. *Method f identifies K just in case f* identifies Θ_0 *given K.*

Identification of K requires that inquiry eventually arrive at complete, true beliefs both about the future and the past. One may weaken this requirement by countenancing incorrect or incomplete memories of the past, so long as these do not compromise predictive power. Then it will be said that the method *projects* K.

Definition 8 *Method f projects K just in case for each e in K and for all but finitely many $n, \emptyset \neq f(e|n) \subseteq [n|e]$.*

If projection looks forward, we may also look backward and ask if the method's conjecture at each stage consistently entails the data received thus far.

Definition 9 *Method f remembers K just in case for each e in K, for each $n, \emptyset \neq f(e|n) \subseteq [e|n]$.*

Clearly, f identifies K just in case f remembers K and f projects K. Intuitively, it seems as though perfect memory would only make reliable projection of the future easier.

But for some of the belief revision operators introduced above, perfect memory *prevents* projection, as will be apparent shortly.

0.4.4 Identifiability, Restrictiveness and Completeness

Let M be the set of all inductive methods and let $M' \subseteq M$. Think of M' as a a proposed *architecture* or restriction on admissible inductive methods. For example, M' may reflect someone's "intuitive" ideas about rationality (e.g., that $f = (r, *)$, for some choice of $r, *$). Then we may say that partition Θ is *identifiable* by M' given K just in case there is an $f \in M'$ such that f identifies Θ given K, and similarly for the identifiability or projectability of K by M'. When $M' = M$, the explicit reference to M will be dropped.

Architecture M' is inductively *complete* just in case each identifiable partition Θ is identified by some method in M'. Otherwise, M' is inductively *restrictive*, in the sense that it prevents us from solving inductive problems we could have solved by other means.[12] In a similar manner, we may speak of completeness and restrictiveness with respect to function identification, projection, or memory. Restrictiveness raises serious questions about the normative standing of a proposed account of rational inquiry, since it seems that rationality ought to augment rather than inhibit the search for truth.[13]

The main question before us is whether insistence on a particular belief revision operator * is restrictive (i.e., prevents us from answering inductive questions we could have answered otherwise). Let M^* denote the set of all inductive methods that implement the plausibility revision operator * (i.e., $M^* = \{(r, *) : r \in \text{IA}\}$). I say that * is complete or restrictive (in any of the above senses) just in case M^* is.

[12]The term "restrictiveness" is due to Osherson et al. (1986).

[13]The principle that restrictiveness calls into qusestion the normative standing of rules of rationality is enunciated in (James 1948) and (Putnam 1963). This principle motivates much learning theoretic work (e.g., Osherson et al.1986, Kelly 1996).

Some of the belief revision operators introduced above are restrictive. But their restrictiveness is manifested in a curious way: they are complete with respect to projection and they are complete with respect to memory, but they are restrictive with respect to identification. In other words, such operators can be implemented to remember or to project the future, but cannot be implemented to do both. Such a method will be said to suffer from *inductive amnesia*. Inductive amnesia implies that those who don't want to repeat history *should* forget it!

Since restrictiveness is a matter of preventing the solution of solvable learning problems, it is useful to characterize the set of solvable problems. Identifiability has an elegant topological characterization. Let K be a collection of data streams. Recall that for finite sequence ϵ, $[\epsilon] = \{e \in U : \epsilon \text{ is extended by } e\}$. A K-fan is a proposition of form $[\epsilon] \cap K$. Then we say S is K*-open* (or open in K) just in case S a union of K- fans. S is K*-closed* just in case $K - S$ is K-open.

Proposition 1 (characterization theorem for partition identification) *Let $\Theta[K]$ denote the restriction of partition Θ to K (i.e, $\{C \cap K : C \in \Theta\}$). Then Θ is identifiable given K just in case $\Theta[K]$ is countable and each cell in $\Theta[K]$ is a countable union of K- closed sets.*[14]

Proof: (Kelly 96). $\dashv$

The characterization of function identifiability is even simpler:

Proposition 2 (characterization theorem for identification) *The following propositions are equivalent:*

1. K is identifiable; 2. K is projectable; 3. K is countable.

Proof: In Appendix I. $\dashv$

[14] I.e., each cell is Σ^0_2 in the Borel hierarchy over K (Kelly 96).

Projectability and identifiability are equivalent with respect to the collection of all possible inductive methods, but not when we restrict attention to methods implementing an inductively amnestic revision operator *.

0.4.5 Counting Retractions

No scientist likes to retract. The social stigma associated with retraction reflects the painful choices and costly conceptual retooling that scientific revolutions entail (Kuhn 1970). If partition Θ is identifiable, we can ask whether some method identifies Θ with an a priori bound n on the number of retractions performed prior to convergence.

Definition 10

1. *retractions*$(f, e) = |\{k : f(e|k+1) \not\subseteq f(e|k)\}|$.
2. *Method f identifies K with n retractions just in case f identifies K and for each e in K, retractions$(f, e) \leq n$.*
3. *The set K is identifiable with n retractions just in case there is a method f such that f identifies K and for each e in K, retractions$(f, e) \leq n$.*

Identification with n retractions has a natural characterization in terms of Spohn's implausibility assignments independently of any choice of operator, a pleasant and revealing connection between learning theory and belief revision.

Proposition 3 (characterization of n retraction identifiability) *Partition Θ is identifiable given K with at most n retractions just in case there is an r such that rng$(r) = \{0, \ldots, n\}$, $K \subseteq$ dom(r), and for each cell $C \in \Theta$, for each $k \leq n$,*

1. *C is open (and hence clopen) in $r^{-1}(k)$ and*

2. $\bigcup_{i=1}^{k} r^{-1}(i)$ *is closed in* $dom(r)$*.*

Proof: In Appendix I. ⊣

Data stream e is *isolated* in $S \subseteq U$ just in case for some n, $[e|n] \cap S \subset \{e\}$ (i.e., $\{e\}$ is open in S).

Proposition 4 (characterization of n retraction function identifiability) *The set* K *of data streams is identifiable with* n *retractions just in case there is an* r *such that* $rng(r) = \{0, \ldots, n\}$*,* $K \subseteq dom(r)$ *and for each* $e \in K$*,* e *is isolated in* $[e]_r^{\leq}$*.*

Proof: In Appendix I. ⊣

0.5 Some Diachronic Properties of Implausibility Revision

Three diachronic properties of implausibility revision operators have particular relevance for reliability considerations. The first requires that the operator always produce new beliefs consistent with the current datum and the domain of the current IA. All belief revision theorists insist on this requirement and all the operators under consideration satisfy it.

Definition 11 (local consistency) *The pair* $(r, *)$ *is locally consistent just in case for all* $(A_1, \ldots, A_{n+1})$ *such that* $dom(r * (A_1, \ldots, A_n)) \cap A_{n+1} \neq \emptyset$,

$$A_{n+1} \cap b(r * (A_1, \ldots, A_{n+1})) \neq \emptyset.$$

The next property requires preservation of the implausibility ordering among worlds satisfying all the input propositions received so far. This does not entail that the ordinal distances between such possibilities are preserved (gaps may appear or disappear).

Definition 12 (positive order-invariance) *The pair* $(r, *)$ *is positively order-invariant just in case for all* $(A_1, \ldots, A_n)$ *such that* $n > 0$*, for all* $w, w' \in dom(r) \cap A_1 \cap \ldots \cap A_n$*,*

1. $w, w' \in dom(r * (A_1, \ldots, A_n))$ *and*

2. $r(w) \leq r(w') \Leftrightarrow (r * (A_1, \ldots, A_n))(w) \leq (r * (A_1, \ldots, A_n))(w')$.

A stricter property requires preservation of the ordinal distances among worlds consistent with all the data received so far.

Definition 13 (positive invariance) *The pair* $(r, *)$ *is positively invariant just in case for all* $(A_1, \ldots, A_n)$ *such that* $n > 0$*, for all* $w, w' \in dom(r) \cap A_1 \cap \ldots \cap A_n$*,*

1. $w, w' \in dom(r * (A_1, \ldots, A_n))$ *and*

2. $r(w) - r(w') = (r * (A_1, \ldots, A_n))(w) - (r * (A_1, \ldots, A_n))(w')$.

Local consistency and positive order-invariance say nothing about what to do with worlds that do not satisfy E. One requirement, reflecting high respect for the data, demands that each world satisfying E be strictly more plausible than every world failing to satisfy E. This property goes much farther than the requirement that the updated belief set $b(r * E)$ entail E. It governs the overall implausibility structure concerning even remotely plausible worlds.

Definition 14 (positive precedence) *The pair* $(r, *)$ *is positively precedent just in case for all* $(A_1, \ldots, A_n)$*, for all* $w \in dom(r) \cap A_1 \cap \ldots \cap A_n$*, for all* $w' \notin dom(r) \cap A_1 \cap \ldots \cap A_n$*,*

1. $w' \in dom(r * (A_1, \ldots, A_n))$ *and* $w' \notin dom(r * (A_1, \ldots, A_n))$ *or*

2. $w, w' \in dom(r * (A_1, \ldots, A_n))$ *and* $(r * (A_1, \ldots, A_n))(w') > (r * (A_1, \ldots, A_n))(w)$.

For each of the properties just defined, we say that $*$ has the property just in case $(r, *)$ has the property, for each IA r.

Local consistency, positive order-invariance and positive precedence are logically independent. Together, they force a belief revision operator to behave in a manner that makes a great deal of sense if sufficiently informative truth is the goal of inquiry. Consider an operator with all three properties. It starts out with a fixed IA r on worlds. Upon updating on E, positive precedence requires that all the non-E worlds are either weeded out altogether (they are not even in the domain of $(r * E)$ or are sent to a "safe" place beyond all the E worlds). By positive order-invariance, the E worlds remain ranked as they were before (the ordinal intervals between two E-worlds may stretch or contract, however). By local consistency, the lowest of these E-worlds must drop to the bottom of the revised IA. As inquiry proceeds, such an operator continues to weed out non-E worlds and to conjecture the most plausible remaining worlds, according to a fixed implausibility ranking, so eventually the actual world migrates to the bottom of the ranking and the operator's belief state is true forever after. The informativeness of this true belief state will depend on how informative the individual "levels" $r^{-1}(k)$ of r are at the outset.

In light of the preceding discussion, it is natural to say that $(r, *)$ *enumerates and tests* just in case $(r, *)$ is locally consistent, positively order-invariant, and positively precedent.[15] Then we have:

Proposition 5 *If partition Θ is identifiable given K then there exists an r such that*

1. *$rng(r) \subseteq \omega$ and*

[15]This kind of procedure has long been entertained under a variety of headings. In the philosophy of science it has been referred to as the method of *bold conjectures and refutations* (Popper 1968) or as the *hypothetico-deductive* method (Kemeny 1953, Putnam 1963). In the learning theoretic literature it is referred to as the *enumeration method* (Gold 1967).

2. for each $*$ such that $(r, *)$ enumerates and tests, $(r, *)$ identifies Θ given K.

Proof: In Appendix II. $\dashv$

Now suppose that $(r, *)$ is locally consistent and positively order-invariant but does not satisfy positive precedence. Then the method still maintains a fixed ranking of implausibility over the E worlds, but some non-E world w may fail to rise above all the E worlds. Hence, it is possible for them to return, eventually, to the bottom of the ranking as inquiry continues. When this happens, the belief state of the agent no longer entails E, so E is "forgotten". It is not difficult to choose *particular* initial epistemic states that lead such a method to forget. Inductive amnesia is the much less trivial situation in which *every* initial epistemic state that ensures that the method reliably predicts the future also causes it to forget some past datum.

Although $*_{R,n}$ does not satisfy positive precedence, it satisfies a weakened version of positive precedence. To define the property, first define the *difference set* of all positions on which two data streams differ:

$$\Delta(e, e') = \{i \in \omega : e(i) \neq e'(i)\}.$$

Then define *Hamming distance* to be the size of the difference set.

$$\rho(e, e') = |\Delta(e, e')|.$$

Finally, define restricted Hamming distance as the number of positions up to k at which two data streams differ:

$$\rho_k(e, e') = |\Delta(e, e') \cap \{0, \ldots, k-1\}|.$$

Now we have:

Proposition 6 (climbing lemma) *Suppose $r(e), r(e'), n$ are finite. Then*

$$(r *_{R,n} [[e|k]])(e) - (r *_{R,n} [[e|k]])(e') \leq (r(e) - r(e')) - n\rho_k(e, e').$$

Proof: By induction on $\rho_k(e, e')$. $\dashv$

The operator $*_{J,n}$ lacks this property because possibilities may backslide upon refutation, making its convergent behavior much more difficult to analyze.

The following properties, like local consistency, are axioms of the AGM theory of belief revision (Gärdenfors 88) and are satisfied by all the belief revision operators under consideration.

Definition 15 (timidity and stubbornness) *The pair* $(r, *)$ *is timid [stubborn] just in case for each* $(A_1, \ldots, A_{n+m})$ *such that* $(A_{n+1} \cap \ldots \cap A_{n+m}) \cap b(r * (A_1, \ldots, A_n)) \neq \emptyset$, $b(r * (A_1, \ldots, A_n)) \cap (A_{n+1} \cap \ldots \cap A_{n+m}) \subseteq [\supseteq] b(r * (A_1, \ldots, A_{n+m}))$.

A *timid* method refuses to draw conclusions that go beyond the data unless its current belief state is refuted. A *stubborn* method retains its current beliefs until they are refuted. Together, these properties force full belief to evolve by mere accretion (according to the standard Bayesian approach) until one's full beliefs are refuted by new information. All enumerate-and-test operators are timid and stubborn[16] and are also complete inductive architectures (proposition 5), which provides something of a reliabilist motivation for timidity and stubbornness. But when positive precedence is dropped in favor of a more "minimal" conception of epistemic change, timidity and stubbornness assume a more sinister aspect, serving a pivotal role in each of the negative arguments presented below.[17]

Proposition 7 *Table I in figure 1 specifies which of the above properties hold of the*

[16]Positive invariance keeps unrefuted worlds at the bottom of the ranking below all other non-refuted worlds. Positive precedence sends all refuted worlds permanently above the non-refuted worlds. And local consistency ensures that the lowest of the non-refuted worlds stay down, so we have timidity and stubbornness.

[17]This observation raises the very interesting question whether belief revision theorists should be so keen to preserve the accretive, Bayesian image of inquiry when the belief state is not refuted.

Table I	C	L	R,α	J,α	A,α	M
pos. order-invariance	yes	yes	yes	yes	yes	yes
pos. invariance	yes	yes	yes	yes	yes	no
local consistency	yes	yes	yes	yes	yes	yes
positive precedence	yes	yes	no	no	no	no
timidity	yes	yes	yes	yes	yes	yes
stubbornness	yes	yes	yes	yes	yes	yes

Table II	C	L	R,α	J,α	A,α	M
positive precedence	yes	yes	yes	yes	yes	no

Figure 1: Proposition 7

operators under consideration regardless of the choice of r and of α. Table II in figure 1 summarizes the changes in the first table when it is assumed that $\alpha \geq r_{above}(dom(r))$.

Proof: Induction on the stage of inquiry and some simple examples. $\dashv$

0.5.1 Inductive Completeness Theorems

The following completeness result follows immediately from propositions 5 and 7 above.

Proposition 8 (complete partition identification operators) *If partition Θ is identifiable given K, then $*_C, *_L, *_B, *_{R,\omega}, *_{J,\omega}, *_{A,\omega}$ can identify Θ given K.*

The next result concerns operators that are complete architectures for identification with n retractions. Recall that problems solvable with n retractions can be packed into an initial epistemic state whose highest level is n (proposition 3). Operators $*_{A,n+1}, *_C, *_L, *_{R,n+1}$, and $*_{J,n+1}$ safely launch refuted worlds above all non-refuted worlds in such an ordering.

Since the truth drops at least one level at each retraction, convergence occurs by the nth retraction.

Proposition 9 (n retraction completeness for partitions) *If partition Θ is identifiable given K with at most n retractions, then $*_C, *_L, *_{R,n+1}, *_{J,n+1}, *_{A,n+1}$ can identify Θ given K with at most n retractions.*

Proof: In Appendix II. $\dashv$

The following results concern the narrower problem of function identification.

Proposition 10 (complete function identification operators) *If K is identifiable (i.e., projectable) then*

1. *K is identifiable by $*_C$, $*_L$, $*_{J,\omega}$, $*_{A,\omega}$, $*_{R,2}$, and*
2. *K is projectable by $*_{R,1}$, $*_{J,1}$.*

Proof: In Appendix II. $\dashv$

Most of these equivalences follow from the preceding proposition and concern operators that boost refuted possibilities above all "live" possibilities. A surprising exception is the fact that $*_{R,2}$ is a complete function identification architecture.[18] To prove completeness, one must construct, for each K, an epistemic state r such that $(r, *_{R,2})$ identifies K. By way of illustration, here is how it can be done in the special case in which all elements of K are finite variants of one another. Given a fixed data stream e_0 we can construct an epistemic state

$$r^H_{e_0}(e) = \rho(e_0, e)$$

on K, where it will be recalled that $\rho(e_0, e)$ is the Hamming distance between e_0 and e, which is just the number of positions i such that $e_0(i) \neq e(i)$. This "Hamming" state has

[18] It is left open whether this can be extended to the case of partition identification.

the nice property that a data stream e' that is k steps below the true data stream e differs from e in at least k positions. When $\alpha = 2$, e' moves up with respect to e at least two steps each time one of the k differences between e and e' is seen, so e' ends up at least one step above e after all of these positions have been observed. The full completeness theorem is proved by means of a generalization of this construction.

0.6 The Grue Hierarchy

To show that a methodological recommendation restricts reliability, one must find an otherwise solvable problem that the recommended method fails to solve, no matter how its initial epistemic state is arranged. This end is served admirably by an unfamiliar application of a familiar idea due to the philosopher Nelson Goodman (1983).[19] Let 0 represent a "green" outcome and let 1 represent a "blue" outcome. Then a "grue$_n$" outcome is either a green outcome by stage n or a blue outcome after stage n. The everywhere green data stream is the everywhere 0 sequence and the everywhere grue$_n$ sequence is a sequence of n 0s followed by all 1s. More generally, let $\neg b$ denote the Boolean complement of b. Let B denote the set of all Boolean-valued data streams. Then if $e \in B$, let $\neg e$ denote the outcome stream in which each outcome occurring in e is reversed (i.e., $(\neg e)(n) = \neg e(n)$). Now define the *grue* operation as follows:

$$e \ddagger n = (e|n)\neg(n|e).$$

In other words, $(e \ddagger n)(i) = e(i)$ if $i < n$ and $= \neg e(i)$ otherwise.

[19] Goodman was not interested in constructing unsolvable inductive problems. His purpose was to show that the constancy of the data stream is not preserved under translation (from the "green" to the "grue" language), and hence that no purely *logical* theory of scientific confirmation can underwrite a bias in credibility for constant data streams.

Grue operations are commutative:[20]

$$(e \ddagger n) \ddagger m = (e \ddagger m) \ddagger n.$$

Also, gruing twice in the same place yields the original data stream. Hence, each composed grue operation can be represented by the set S of positions that have grue operations applied an odd number of times. Let $e \ddagger S$ denote the (unique) data stream that results from applying, in any order, any odd number of grue operations at positions in S and any even number of grue operations (possibly zero) at all other positions.

Now given $K \subseteq B$, we can define a hierarchy of ever more complex inductive problems as follows:

Definition 16 (The Grue Hierarchy) *Let $K \subseteq B$.*

1. $g^n(K) = \{e \ddagger S : |S| = n \text{ and } e \in K\}$.

2. $G^n(K) = \bigcup_{i \leq n} g^n(K)$.

3. $G^\omega(K) = \bigcup_{i < \omega} g^n(K)$.

The even grue hierarchy $G^n_{\text{even}}(K)$ is defined similarly, except that (1) is replaced with:

$$g^n_{\text{even}}(K) = \{e \ddagger S : |S| = 2n \text{ and } e \in K\}.$$

The distinction between the even and the full grue hierarchies is important for some of the belief revision operators under consideration, but it makes no difference to identifiability when no extra constraints are imposed on the scientist's inductive method:

Proposition 11 *For all $n \in \omega, e_0 \in B$,*

[20]Let $\mathcal{G}$ denote the closure under composition of the set of grue all functions: $\{(. \ddagger k) : k \in \omega\}$. Then $(\mathcal{G}, \circ)$ is an Abelian group in which each element is its own inverse.

1. *$G^n(e_0), G^n_{even}(e_0)$ are identifiable with n retractions but not with $n-1$ retractions.*

2. *$G^\omega(e_0), G^\omega_{even}(e_0)$ are identifiable but not under any fixed bound on the number of retractions performed.*

Proof: In Appendix I. $\dashv$

0.7 The Main Result

The following proposition determines exactly which problems in the grue hierarchy and in the even grue hierarchy each of the operators under consideration can solve.

Proposition 12 (The grue scale) *The table in figure 2 specifies which problems in the grue and even grue hierarchies generated from an arbitrary e_0 each of the operators under consideration can identify. The classifications are optimal, in the sense that no lower value of α than the one reported suffices for identification of the corresponding problem.*

Proof: Propositions 10 and 17 in Appendix II, proposition 20 in Appendix III, and propositions 27, 28, 29, and 30 in Appendix IV. $\dashv$

Most of the positive results in the table follow from more general completeness results already discussed. Noteworthy exceptions are the abilities of $*_{J,1}$ and $*_{R,1}$ to identify $G^\omega_{\text{even}}(e_0)$. This contrasts markedly with the situation in the full grue hierarchy, in which these operators all fail by level three. The situation is quite different for $*_M$ and $*_{A,n}$, which see no improvement in the even grue hierarchy (proposition 28.2). The problem $G^\omega_{\text{even}}(e_0)$ is just the set of all finite variants of a given data stream e_0. The evolution of $*_{J,1}$ and of $*_{R,1}$ in this problem can be pictured as follows. Suppose the method starts out with an initial epistemic state ranking each data stream according to its Hamming distance from a given data stream e_0. Suppose e is the truth. Then the set of data streams

problem	M	A, α	J,α	R,α	L	C
$G^{\omega}(e_0)$	no	$\alpha = \omega$	$\alpha = 2$	$\alpha = 2$	yes	yes
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$G^{n}(e_0)$	no	$\alpha = n+1$	$\alpha = 2$	$\alpha = 2$	yes	yes
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$G^{2}(e_0)$	no	$\alpha = 3$	$\alpha = 2$	$\alpha = 2$	*yes*	*yes*
$G^{1}(e_0)$	no	$\alpha = 2$	$\alpha = 2$	$\alpha = 1$	yes	yes
$G^{0}(e_0)$	*yes*	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	yes	yes
$G^{\omega}_{\text{even}}(e_0)$	no	$\alpha = \omega$	$\alpha = 1$	$\alpha = 1$	yes	yes
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$G^{n}_{\text{even}}(e_0)$	no	$\alpha = n+1$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$G^{2}_{\text{even}}(e_0)$	no	$\alpha = 3$	$\alpha = 1$	$\alpha = 1$	*yes*	*yes*
$G^{1}_{\text{even}}(e_0)$	no	$\alpha = 2$	$\alpha = 1$	$\alpha = 1$	yes	yes
$G^{0}_{\text{even}}(e_0)$	yes	$\alpha = 0$	$\alpha = 0$	$\alpha = 0$	yes	yes

Figure 2: Proposition 12

rotation.eps

Figure 3: Learning as hypercube rotation

differing from e_0 only where e_0 differs from e may be viewed as a finite dimensional hypercube whose dimensionality matches the total number of differences between e_0 and e (cf. the upper part of fig. 3). Think of this hypercube as resting balanced on the vertex labelled with e_0. To find the initial implausibility of a vertex, find the shortest path from the bottom of the cube to that vertex. It is shown (proposition 30) that the sequential operation of both $*_{J,1}$ and $*_{R,1}$ can be viewed as the rigid rotation of the hypercube from one vertex to another on a direct path to the true (originally uppermost) vertex e. After e is rotated to the bottom, the rotation stops and the method has converged to the truth!

The pleasing image of learning as rigid hypercube rotation cannot be extended to the full grue hierarchy. Indeed, no matter where we insert the single data stream $\neg e_0$ into the restriction of the Hamming ranking to $G^2_{\text{even}}(e_0)$, $*_{J,\alpha}$ fails even when $\alpha = 2$.

Proposition 13 *For each* $r \supseteq r^H_{e_0}|G^2_{even}$, $(r, *_{J,2})$ *does not identify* $G^2_{even}(e_0) \cup \{\neg e_0\}$.

Proof: Appendix III. $\dashv$

By way of illustration, consider how the hypercube's rotation is spoiled in the special case in which $\alpha = 1$. If we insert $\neg e_0$ into an infinite level of the Hamming ranking, then by positive invariance, it never falls the infinite distance to the bottom of the ranking, since the infinitely many elements of $G^1(e_0)$ occurring lower than $\neg e_0$ are never all refuted. So suppose we insert $\neg e_0$ at a finite level. Then $\neg e_0$ is below some other data stream e agreeing with $\neg e_0$ as far as we please. If e happens to be the truth, then the cube rotates as usual until $\neg e_0$ arrives at the bottom, along with some vertex e'. By timidity and stubbornness, $\neg e_0$ remains at the bottom until it is seen to disagree with e. During this time, e' will be refuted and will rise one step along with all the nodes that disagree with e at the currently observed position. Since $\neg e_0$ still remains at the bottom, positive invariance prevents e and the data streams agreeing with e at the currently scanned position from dropping. Thus, instead of fully rotating onto a new vertex, the cube partially tips up so that an edge is parallel to the floor (cf. the lower part of fig. 3). This carries some currently refuted unit variant v of e up, just even with e. Since v never again differs from e (its single difference from e has already been used up), no future data will ever drop e below v. So if the method succeeds in lowering e to the bottom, it will also lower v to the bottom, and hence will have forever forgotten the datum refuting v.

By proposition 13, the initial state $*_{J,2}$ employs to identify $G^{\omega}(e_0)$ cannot be an extension of the Hamming ranking over $G^{\omega}_{\text{even}}(e_0)$. Instead I employ a ranking based on *grue distance*, or the number of grue operations required to transform one data stream into another (cf. Appendix III).[21] Define the *grue set* for two data streams as follows:

$$\Gamma(e, e') = \{n \in \omega : [n = 0 \wedge e(n) \neq e'(n)] \vee [n > 0 \wedge \neg(e(n-1) = e'(n-1) \Leftrightarrow e(n) = e'(n))]\}.$$

The terminology is justified by the following fact.

Proposition 14 $\Gamma(e, e')$ *is the least* $S \subseteq \omega$ *such that* e' *can be obtained from* e *by applying grue operations only at positions in* S*.*

Proof: Omitted. $\dashv$

Now define the *grue distance* on B as follows:

$$\gamma(e, e') = |\Gamma(e, e')|.$$

[21] This idea is familiar in computer science as a way to compress image files. Instead of recording the intensity of each pixel separately, one records the places at which intensity changes, which saves space if many adjacent pixels have the same intensity.

In light of the preceding proposition, grue distance is the least number of grue operations required to transform e into e'. It is readily verified that grue distance is an extended metric over B. The initial epistemic state on B induced by grue distance from e_0 is just:

$$r^G_{e_0} = \gamma(e_0, e').$$

For an algebraic perspective on the relationship between $r^G_{e_0}$ and $r^H_{e_0}$, define the Hamming and grue orders as follows (figure 0.7):

1. $e' \leq^{e_0}_G e'' \Leftrightarrow \Gamma(e_0, e') \subseteq \Gamma(e_0, e'')$.

2. $e' \leq^{e_0}_H e'' \Leftrightarrow \Delta(e_0, e') \subseteq \Delta(e_0, e'')$.

These orderings are isomorphic copies of the inclusion ordering on the power set of ω and hence are isomorphic Boolean algebras, but they label this structure very differently (e.g., adjacent elements of the grue algebra are complements in the Hamming algebra). Moreover, by proposition 14, $G^\omega(e_0)$ is the union of the finite levels of $r^G_{e_0}$, whereas the union of the finite levels of $r^H_{e_0}$, is just $G^\omega_{\text{even}}(e_0)$.

The method $(r^G_{e_0}, *_{J,2})$ identifies $G^\omega(e_0)$ in an intuitively attractive manner. It starts out assuming that the true data stream is e_0. When it encounters a surprise at stage n, it then assumes that the true data stream is $e_0 \ddagger n$, and so forth, adding successive grue operations to e_0 only when the data require them (propositions 20.2 and 19). Recall that $*_{J,2}$ has the objectionable property that a possibility can become more plausible when it is refuted if only very implausible worlds are refuted by the current datum. The implausibility assignment based on grue distance prevents this possibility from ever occurring over possibilities in G^ω_{even}. This assignment has the property that, at each stage prior to convergence, a highly plausible (degree 0 or 1) possibility is refuted. Since $\alpha = 2$, all refuted possibilities are pushed up at least one step by $*_{J,2}$. When $\alpha < 2$, refuted possibilities do not rise when the agent's current beliefs are not refuted, so the same argument does not work and in fact cannot be made to work since even the very easy problem $G^1(e_0)$ is not identifiable by $*_{J,1}$.

Turning to the negative results, it is remarkable that $*_M$, $*_{J,1}$ and $*_{A,1}$ cannot even identify $G^1(e_0)$ (proposition 28), and hence cannot cope with the possibility of even a single reversal in the data stream! Operator $*_{R,1}$ survives just one level higher, failing on $G^2(e_0)$ (proposition 27). Operator $*_{A,\alpha}$ compares unfavorably with $*_{J,\alpha}$ and $*_{R,\alpha}$, because $*_{A,n+1}$ fails on $G^{n+1}(e_0)$, for each n, whereas $*_{J,2}$ and $*_{R,2}$ succeed on $G^\omega(e_0)$. By proposition 11, $G^n(e_0)$ can be solved with just n retractions by the obvious method that starts out conjecturing e_0 and that refuses to believe in grue operations until they are observed. The negative results imply that this sensible behavior cannot be obtained from $*_M$, $*_{J,1}$ or $*_{A,1}$, no matter how cleverly the initial epistemic state is arranged.

By the following proposition, nearly all of the negative results in the table are examples of inductive amnesia.

Proposition 15 *Let $e_0 \in B$.*

1. *All of the operators under consideration can remember the past.*
2. *All of the operators under consideration can project $G^\omega(e_0)$ so long as $\alpha > 0$. Among these, only M fails to be a complete projector.*

Proof: propositions 18, 20, and 10. $\dashv$

Inductive amnesia illustrates a fundamental, epistemic dilemma for iterated belief revision operators. Recall that belief revision theory can be stretched in two directions. Lumping all possible worlds together at one level of implausibility makes a belief revision agent behave like an accretive tabula rasa that takes no inductive risks and never has its beliefs contradicted so long as the successive data are mutually consistent. Spreading worlds out at distinct levels of implausibility makes belief revision look more like Popper's methodology of bold conjectures and refutations. The former extreme secures perfect memory at the price of refusing to predict the future, whereas the latter guarantees convergence to correct predictions at the price of possibly forgetting the past when α is low. A crucial epistemological question for belief revision theory is therefore to find the least α for which these competing demands are jointly satisfiable for a given empirical problem. Perhaps the most striking result of this investigation is that the operators $*_{J,\alpha}$ and $*_{R,\alpha}$ enjoy an *infinite* jump in reliability when α is incremented from *one* to *two*. For $\alpha \geq 2$, the methods succeed over the entire, infinite, grue hierarchy. For $\alpha < 2$, neither can cope with more than two grue operations.

0.8 Dropping Well-ordering

So far, it has been assumed that epistemic states well-order the possible worlds in their domains, since epistemic states assume ordinal values. This assumption is not generally accepted in the belief revision community, and it is centrally involved in the proof that $*_M$ fails on the easy problems $G^1_{\text{even}}(e_0)$ and $G^1(e_0)$. Since $*_M$ has a straightforward extension to a wide class of non-well-ordered epistemic states (Boutilier 94), we should examine whether its modest learning abilities improve in this more general formulation.

Let $R = (D, \leq)$ be a totally ordered set. Let $\min(R, E)$ denote the set of all minimial elements of $E \cap D$. For present purposes, an *epistemic state* is a total order $R = (D, \leq)$ such that $D \subseteq U$ and for each proposition $E \in \{U\} \cup \{[i, k] : i, k \in \omega\}$, $\min(R, E) \neq \emptyset$. In other words, an epistemic state is a total order on data streams that has a least element and in which each observation of an outcome (consistent with the domain of the order) has a least element. The associated belief state of R is given by $b(R) = \min(R, U)$. Upon receiving new information $[i, k]$, $*_M$ updates the epistemic state $R_1 = (D_1, <_1)$ producing a new state $R_2 = (D_2, <_2)$, such that

1. $\min(R_2, U) = \min(R_1, [i, k])$ (which is nonempty) and

2. for all $e, e' \in D_1 - \min(R_1, [i, k])$, $e \leq_2 e' \Leftrightarrow e \leq_1 e'$.

In other words, $*_M$ brings $\min(R_1, [i, k])$ to the bottom of the new state and rigidly raises all the rest of the worlds so that the lowest are immediately above $\min(R_1, [i, k])$ in the revised ordering. An inductive method *implementing* $*_M$ is a pair of form $(R, *_M)$. Now we have:

Proposition 16 *Generalize $*_M$ and epistemic states as just described. Then:*

1. $G^1(e_0), G^1_{even}(e_0)$ *are identifiable by* $*_M$.

2. $G^2(e_0), G^2_{even}(e_0)$ *are not identifiable by* $*_M$.

Proof: in Appendix IV. $\dashv$

So the learning power of $*_M$ improves slightly in the more general setting in which the well-ordering assumption is dropped. This result illustrates how learning theoretic analysis can be employed to criticize controversial assumptions about the nature of epistemic states.

The well-ordering assumption is also involved in the negative results concerning $*_{A,n}$,[22] $*_{J,1}$, and $*_{R,1}$. But these operators were originally defined only on ordinal-valued epistemic states (Goldszmidt and Pearl 94; Darwiche and Pearl 97; Spohn 88), and it is unclear how they should be extended to arbitrary, totally ordered states.[23]

[22] Ordinal-valued implausibilities are required for the proofs of propositions 24 and 26.

[23] If, more modestly, degrees of implausibility are taken to be numbers in a non-well-ordered system (e.g., rationals, reals, or nonstandard reals) then all of these operators are enumerate-and-test operators

0.9 Conclusion

The normative principles of belief revision theory have been motivated by intuition, coherence, and symmetry considerations. The natural question whether following such a rule would help or hinder the formulation of informative, true beliefs has largely been ignored. Once this question is entertained, interesting and unanticipated issues emerge, such as (i) inductive amnesia, (ii) the essential tension between compression and rarefaction in the epistemic state, (iii) the pivotal significance of the value $\alpha = 2$ for the resolution of this tension, (iv) the idea of generating epistemic states from operations on data streams or as ranks in Boolean algebras, (v) the utility of grue distance for improving the reliability of belief revision operators, (vi) the appealing portrayal of induction as rigid rotation of a hypercube, (vii) the image of tail reversals in data streams "derailing" this rotation and (viii) the relevance for reliability of well-ordered degrees of implausibility. These issues are not drawn from a priori intuitions. They are rigorously derivable from the straightforward aim of reliably arriving at sufficiently informative truths from increasing empirical data. As such, they can serve as well- motivated constraints on theories of rational belief revision.

The results of this study should be expanded and generalized. It is left open, for example, whether $*_{J,2}$ and $*_{R,2}$ are complete architectures for partition identification. The rest of the results could be cast in more general settings, in which the order of the data may be scrambled, experimental acts may be performed, meaning shifts are possible, and so forth.[24] But even the simplified, narrow setting of the present study illustrates how a

if the domain of the initial epistemic state is confined to the $[0, 1)$ interval, and hence are empirically complete (proposition 5).

[24](Kelly 1996) provides illustrations of how such generalizations might proceed. Osherson and Martin (1995, 1996) develop a reliability analysis of belief revision operators using logical languages.

systematic logical analysis founded on the aim of finding true, informative beliefs can serve as a powerful and interesting constraint on belief revision theorizing; a constraint that, it is hoped, will become as familiar to belief revision theorists as the usual representation and equivalence results are today.

0.10 Acknowledgements

Oliver Schulte discovered the basic strategy for proving proposition 27 and provided comments on a draft of this paper. The paper benefitted from numerous discussions with Horacio Arlo-Costa and from comments and corrections by Wolfgang Spohn, Clark Glymour, and Steve Awodey.

0.11 Bibliography

C. Boutilier (1993), “Revision Sequences and Nested Conditionals”, in *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning*, pp. 519-525.

Horacio Arlo Costa (1997). “Belief Revision Conditionals: Basic Iterated Systems”, *Annals of Pure and Applied Logic*, forthcoming.

Adnan Darwiche and Judea Pearl (1997) “On the Logic of Iterated Belief Revision”, 89:1-29.

John Earman (1992). *Bayes or Bust*, Cambridge: M.I.T. Press.

Peter Gärdenfors (1988) *Knowledge in Flux*, Cambridge: MIT Press.

Stephen Murray Glaister (1997), “Symmetry and Belief Revision”, Unpublished Manuscript.

E. M. Gold (1967) "Language Identification in the Limit", *Information and Control* 10: 302-320.

Moises Goldszmidt and Judea Pearl (1994) "Qualitative Probabilities for Default Reasoning, Belief Revision, and Causal Modeling", Technical Report r-161-L, Revision II.

Nelson Goodman (1983) *Fact, Fiction and Forecast*, 4th ed. Cambridge: Harvard University Press.

Adam Grove (1988), "Two Modellings for Theory Change", em Journal of Philosohical Logic 17: 157-170.

William James (1948) "The Will to Believe", in *Essays in Pragmatism*, ed. A. Castell. New York: Macmillan.

Hirofumi Katsuno and Alberto O. Mendelzon (1991) "Propositional knowledge base revision and minimal change", *Journal of Artificial Intelligence* 52: 263-294.

Kevin T. Kelly (1996) *The Logic of Reliable Inquiry.* New York: Oxford University Press.

Kevin T. Kelly and Oliver Schulte (1995) "The Computable Testability of Theories Making Uncomputable Predictions", *Erkenntnis* 43: 29-66.

Kevin T. Kelly, Oliver Schulte, and Vincent Hendricks (1997) "Reliable Belief Revision", in *Proceedings of the 10th International Congress of Logic, Methodology, and Philosophy of Science*, ed. Maria Luisa Dalla Chiara, Dordrecht: Kluwer.

J. Kemeny (1953) "The Use of Simplicity in Induction", *Philosophical Review* 62: 391-408.

Thomas Kuhn (1970), *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

Isaac Levi (1996), *The Enterprise of Knowledge*, Cambridge: M.I.T. Press.

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

Eric Martin and Daniel Osherson (1997). "Scientific Discovery via Rational Hypothesis Revision", in *Proceedings of the 10th International Congress of Logic, Methodology, and Philosophy of Science*, ed. Maria Luisa Dalla Chiara, Dordrecht: Kluwer.

Eric Martin and Daniel Osherson (1996). "Scientific Discovery Based on Belief Revision", *Journal of Symbolic Logic*, forthcoming.

Eric Martin and Daniel Osherson (1998). *Elements of Scientific Inquiry*, Cambridge: M.I.T. Press.

Abhaya C. Nayak (1994) "Iterated Belief Change Based on Epistemic Entrenchment", *Erkenntnis* 41: 353-390.

Daniel Oshersion, Michael Stob, and Scott Weinstein (1986) *Systems that Learn.* Cambridge: MIT Press.

Karl R. Popper (1968), *The Logic of Scientific Discovery*, New York: Harper.

Hilary Putnam (1963) "'Degree of Confirmation' and Inductive Logic", in *The Philosophy of Rudolph Carnap*, ed. A. Schilpp. LaSalle, Ill.: Open Court.

Frank P. Ramsey (1990) *Philosophical Papers*, H. A. Mellor, ed., Cambridge: Cambridge University Press.

Dov Samet (1996) “Hypothetical Knowledge and Games with Perfect Information”, *Games and Economic Behavior* 17: 230-251.

Wolfgang Spohn (1988) “Ordinal Conditional Functions: A Dynamic Theory of Epistemic States”, *Causation in Decision, Belief Change, and Statistics, II*, ed. Brian Skyrms and William L. Harper, Dordrecht: Kluwer.

Wolfgang Spohn (1990) “A General Non-probabilisistic Theory of Inductive Reasoning”, *Uncertainty in Artificial Intelligence* 4: 149-159.

0.12 Appendix I: Proofs of Characterization Theorems

Proof of proposition 2 $1 \Rightarrow 3$ follows from proposition 1. $3 \Rightarrow 1$: Suppose K is countable. Then $\Theta_0[K]$ has countably many cells. Moreover, each cell $\{e\} \in \Theta_0[K]$ is K-closed, since $U - \{e\} = \bigcup\{[\epsilon] : \epsilon \not\subset e\}$. Hence, by proposition 8, Θ_0 is identifiable given K, so K is identifiable.

Proof of $1 \Rightarrow 2$. Immedate. $2 \Rightarrow 1$. Suppose that method f projects K. Now define method g, which identifies K, as follows.

$$g(\epsilon) = \begin{cases} [\epsilon] \cap [\mathrm{lh}(\epsilon)|e] & \text{if } \exists e, \emptyset \neq f(\epsilon) \subseteq [\mathrm{lh}(\epsilon)|e] \\ \{e'\} & \text{otherwise,} \end{cases}$$

where e' is an arbitrary element of $[\epsilon]$.$\dashv$

Proof of proposition 3 ($\Rightarrow$) Suppose (i) f identifies Θ given K with n retractions. Define for each $e \in K$: (ii) $r(e) = k \Leftrightarrow \text{retractions}(f, e) = k$. Thus, (iii) $\mathrm{rng}(r) = \{0, \ldots, n\}$ and $K = \mathrm{dom}(r)$. Let $C \in \Theta$. (1) Suppose for reductio that C is not open in $r^{-1}(k)$.

Then some $e \in K \cap r^{-1}(k) \cap C$ is a limit point of $K \cap r^{-1}(k) - C$. So for each i there is an $e_i \in r^{-1}(k) - C$ such that $e_i|i = e|i$. Let $w =$ the least m, such that for all $m' \geq m, f(e|m') = f(e|m)$, so by statement i above, $f(e|w) \subseteq C$ and by ii, f uses a full k retractions along e by stage w. But since $e_w \notin C$ and $e_w|w = e|w$, there is a $w' > w$ such that $f(e_w|w) \subseteq C$ but $f(e_w|w') \subseteq K - C$. Hence, f performs more than k retractions along e_w. Contradiction. So (iv) C is open in $r^{-1}(k)$.

(2) Define: $G = \{\epsilon \in \omega^{<\omega} : f$ performs at least $k+1$ retractions by the end of $\epsilon\}$. Then $\bigcup_{i=1}^{k} r^{-1}(i) = \{e \in \text{dom}(r) : \text{retractions}(f, e) \leq k\} = K - \bigcup\{[\epsilon] : \epsilon \in G\}$. Hence, (v) $\bigcup_{i=1}^{k} r^{-1}(i)$ is closed in dom(r). The proposition is established by iii, iv and v.

($\Leftarrow$) Deferred to proposition 9. $\dashv$

Proof of proposition 4 ($\Rightarrow$) Let r be as guaranteed by proposition 3. Let $e \in K$. So there exists a $k \leq n$ such that $r(e) = k$. Since $\{e\}$ is open in $r^{-1}(k)$, $\exists n \forall m \geq n, [e|m] \cap r^{-1}(k) = \{e\}$. Also, $\bigcup_{i=1}^{k-1} r^{-1}(i)$ is closed in dom(r). So $\exists n' \forall m \geq n', [e|m] \cap \bigcup_{i=1}^{k-1} r^{-1}(i) = \emptyset$. So $\forall m \geq \max(n, n'), [e|m] \cap \bigcup_{i=1}^{k} r^{-1}(i) = \{e\}$. Hence e is isolated in $[e]_r^{\leq}$.

($\Leftarrow$) Deferred to proposition 9. $\dashv$

Proof of proposition 11 Proof of (1). For the $G^n(e_0)$ case, let $r^{-1}(i) = G^i(e_0)$, for each $i \leq n$. For the $G^n_{\text{even}}(e_0)$ case, let $r^{-1}(i) = G^i_{\text{even}}(e_0)$. Observe that for each $e \in r^{-1}(i)$, $\{e\}$ is isolated in $[e]_r^{\leq}$ and apply proposition 4.

For the negative claim, assume for reductio that g succeeds on $G^{n+1}(e_0)$ with n mind-changes. Feed to g data drawn from e_0 until a stage k_0 is reached at which g outputs $\{e_0\}$, which g must do since $e_0 \in G^n(e_0)$. Then proceed by feeding $e_1 = e_0 \ddagger k$ until g outputs $\{e_1\}$. This procedure can be continued until $n+1$ grue operations have been applied. But then $n+1$ retractions are perfomed by g on the resulting data stream e_{n+1}. Contradiction. The negative argument for $G^n_{\text{even}}(e_0)$ is similar except that we let $e_{i+1} = (e_i \ddagger k) \ddagger k+1$.

Proof of (2). The following method identifies $G^\omega(e_0)$: enumerate the whole set and output the first data stream in the set that is consistent with the finite outcome sequence ϵ seen so far. The negative claim follows from part (1). $\dashv$

0.13 Appendix II: Completeness Proofs

Proposition 17 *For each $e, \{e\}$ is identifiable with 0 retractions by $*_M, *_{A,0}, *_{J,0}, *_{R,0}$.*

Proof: Choose r so that $\text{dom}(r) = \{e\}$ and $r(e) = 0$. $\dashv$

Proposition 18 *Let $e_0 \in B$. $*_M$ can project $G^\omega(e_0)$.*

Proof: Let $\text{dom}(r) = \{e_0, \neg e_0\}$, let $r^{-1}(0) = \{e_0\}$ and let $r^{-1}(1) = \{\neg e_0\}$. If $e \in G^\omega(e_0)$, then e is a finite variant of either e_0 or $\neg e_0$. It is easy to check that $(r, *_M)$ succeeds. $\dashv$

Proof of proposition 5 Suppose Θ is identifiable given K. By proposition 1, we may suppose that for each cell $C_j \in \Theta[K]$, there exists a countable union $B_j = \bigcup_i^\infty S_i^j$ of K-closed sets such that $B_j \cap K = C_j \cap K$. Enumerate $\{S_i^j \cap K : i, j \in \omega\}$ as $R_0, \ldots, R_n, \ldots$. Define (i) $r(e) = (\mu i)(e \in R_i')$. Evidently, $\text{rng}(r) \subseteq \omega$. Let $*$ be such that $(r, *)$ enumerates and tests. Hence, $(r, *)$ is locally consistent, positively invariant, and satisfies positive precedence. It remains to show that $(r, *)$ identifies Θ given K. Let $e \in K$. Since the R_i cover K, $r(e)$ is defined and $e \in R_{r(e)}$. Since each R_i is K-closed, so is $[e]_r^< = \bigcup_{i=0}^{r(e)-1} R_i$. So $U - [e]_r^<$ is K-open. So there is a set S of finite sequences such that $K - [e]_r^< = K \cap \bigcup_{\epsilon \in S}[\epsilon]$. Since $e \in K - [e]_r^<$, there is a k such that $e|k \in S$. So since $\text{dom}(r) = K$, we have for each $k' \geq k$, (ii) $[e]_r^< \cap [e|k'] = \emptyset$. By ii and positive precedence, we have that $e \in \text{dom}(r * [[e|k']])$ and for each $e' \in [e]_r^<$, (iii) either $e' \notin \text{dom}(r * [[e|k']])$ or $(r * [[e|k']])(e) < (r * [[e|k']])(e')$. Since $e \in [e|k']$, we have by positive order-invariance that for each $e'' \in \text{dom}(r * [[e|k']]) - [e]_r^<$, (iv) $(r * [[e|k']])(e) \leq (r * [[e|k']])(e'')$. But by local consistency, (v)

$(r * [[e|k']])^{-1}(0) \neq \emptyset$. By iii-v, we have (vi) $e \in b(r * [[e|k']]) = (r, *)(e|k')$. Let C be the cell of Θ to which e belongs. It remains only to establish that (vii) $b(r * [[e|k']]) \subseteq C$. Let $e' \in \mathrm{dom}(r) - C$. Then $e' \notin R_{r(e)} \subseteq C$, so (viii) $r(e) \neq r(e')$.

Case I. Suppose $e' \notin [e|k']$. Then by positive precedence, either $e' \notin \mathrm{dom}(r * [[e|k']])$ or $(r * [[e|k']])(e) < (r * [[e|k']])(e')$. Thus, $e' \notin b(r * [[e|k']])$.

Case II. Suppose $e' \in [e|k']$. Then by ii, and viii, $r(e) > r(e')$. By positive invariance, $(r *_{CL} [[e|k']])(e') > (r *_{CL} [[e|k']])(e)$, so again $e' \notin b(r * [[e|k']])$. $\dashv$

Proof of proposition 9 Suppose Θ is identifiable given K with at most n retractions. Then by proposition 3, there exists an r such that (i) $\mathrm{rng}(r) \subseteq \{0, \ldots n\}$ and (ii) $K \subseteq \mathrm{dom}(r)$ and for each cell $C \in \Theta$, for each $k \leq n$, (iii) C is open (and hence clopen) in $r^{-1}(k)$ and (iv) $\bigcup_{i=1}^{k} r^{-1}(i)$ is K-closed. Let $*$ range over $*_C$, $*_L$, $*_{R,k+1}$, $*_{J,k+1}$, $*_{A,k+1}$. By i, $k + 1 \geq r_{\text{above}}(\mathrm{dom}(r))$, so by proposition 7, * generates and tests. So by proposition 5, $(r, *)$ identifies Θ given K.

Let $e \in K$. It remains to show that each of these operators performs at most k retractions along e when started out on r. Suppose that $(r, *)(e|k) \not\subseteq (r, *)(e|k+1)$. Then $b(r * [[e|k + 1]]) \not\subseteq b(r * [[e|k]])$. So there exists an e' such that (v) $(r * [[e|k + 1]])(e') = 0$ but $(r * [[e|k]])(e') \neq 0$. So by the definition of $*$, $(r * [[e|k]])(e'|[k, e(k)]) = 0$. So (vi) $-\min\{(r * [[e|k]])(e'') : e'' \in \mathrm{dom}(r * [[e|k]]) \cap [k, e(k)]\} + (r * [[e|k]])(e') = 0$. By v and vi, $\min\{(r * [[e|k]])(e'') : e'' \in \mathrm{dom}(r * [[e|k]]) \cap [k, e(k)]\} > 0$. Hence, $-\min\{(r * [[e|k]])(e'') : e'' \in \mathrm{dom}(r * [[e|k]]) \cap [k, e(k)]\} + (r * [[e|k]])(e) < (r * [[e|k]])(e)$, so $(r * [[e|k + 1]])(e) < (r * [[e|k]])(e)$. So we have that for each k such that $(r, *)(e|k) \not\subseteq (r, *)(e|k+1)$, $(r * [[e|k + 1]])(e) < (r * [[e|k]])(e)$. But by hypothesis, $r(e) \leq n$. Hence, $(r, *)$ performs at most n retractions along e. $\dashv$

Proof of proposition 10 The $*_C$, $*_L$, $*_{J,\omega}$, $*_{A,\omega}$ cases are instances of proposition 8.

The equivalence of identifiability and projectability is due to proposition 2.

Proof that operator $*_{J,1}$ is a complete projection architecture. Let $* = *_{J,1}$. Let K be projectable. Then K is countable. Enumerate K as $e_0, e_1, \ldots$. Let $r^{-1}(i) = \{e_i\}$. Let $e \in K$, so for some $i, e = e_i$. First it is established that: (i) $\forall n \forall i, (r *_{J,1} [[e|n]])^{-1}(i)$ is finite. This is evident by the definition of r when $n = 0$. Suppose statement i holds up to n. Let $m = \min\{(r*[[e|n]])(e') : e' \in \text{dom}(r*[[e|n]]) \wedge e' \in [n, e(n)]\}$ and let $m' = \min\{(r* [[e|n]])(e') : e' \in \text{dom}(r*[[e|n]]) \wedge e' \in U - [n, e(n)]\}$. Then by the definition of $*$, we have $(r* [[e|n+1]])^{-1}(i) = ([n, e(n)] \cap (r*[[e|n]])^{-1}(m+i)) \cup ((U - [n, e(n)]) \cap (r*[[e|n]])^{-1}(m'+i-1))$, under the convention that $(r*[[e|n]])^{-1}(z) = \emptyset$ if $z < 0$. This set is finite by the induction hypothesis. So we have statement i. Next, we establish (ii) if $b(r*[[e|n]]) \cap [n, e(n)] = \emptyset$ then $(r*[[e|n+1]])(e) \leq (r*[[e|n]])(e) - 1$. For suppose $b(r*[[e|n]]) \cap [n, e(n)] = \emptyset$. Then since $e \in [n, e(n)]$, we have $(r*[[e|n+1]])(e) = -\min\{(r*[[e|n]])(e') : e' \in \text{dom}(r*[[e|n]]) \cap [n, e(n)]\} + (r*[[e|n]])(e) \leq -1 + (r*[[e|n]])(e)$. So we have ii. Next we establish: (iii) if $\forall e' \in b(r*[[e|n]]), n|e' \neq n|e$, then $\exists m \geq n, b(r*[[e|m]]) \cap [m, e(m)] = \emptyset$. For suppose that for all $e' \in b(r*[[e|n]]), n|e' \neq n|e$. Suppose for reductio that for all $m \geq n$, $b(r*[[e|m]]) \cap [m, e(m)] \neq \emptyset$. Then by timidity and stubbornness (proposition 7), (iv) $\forall m \geq n$, $b(r*[[e|m]]) = b(r*[[e|n]]) \cap [n, e(n)] \cap \ldots \cap [m-1, e(m-1)]$. $b(r*[[e|n]])$ is finite by statement i. So by the hypothesis of iii, there exists an $m' \geq n$ such that $b(r*[[e|n]]) \cap [n, e(n)] \cap \ldots \cap [m-1, e(m-1)] = \emptyset$. By iv, $b(r*[[e|m']]) = \emptyset$, contradicting local consistency and establishing iii. Next we need (v) if $\exists e' \in b(r*[[e|n]])$ such that $n|e' = n|e$, then $\exists m \geq n$ such that $b(r*[[e|m]]) = \{e'' \in b(r*[[e|n]]) : m|e'' = m|e\}$. For by i, there is an $m \geq n$ such that $b(r*[[e|n]]) \cap [n, e(n)] \cap \ldots \cap [m-1, e(m-1)] = \{e'' \in b(r*[[e|m]]) : m|e = m|e''\}$. But by timidity and stubbornness, $b(r*[[e|m]]) = b(r*[[e|n]]) \cap [n, e(n)] \cap \ldots \cap [m-1, e(m-1)]$, establishing v. Finally, it is shown that (vi) if $\forall e' \in b(r*[[e|n]]), n|e' = n|e$, then $\forall m \geq n, b(r*[[e|m]]) = b(r*[[e|n]])$. For by local consistency, $b(r * [[e|n]]) \neq \emptyset$. So by timidity and stubbornness, (iv) holds at each stage $m \geq n$, yielding vi.

Consider the following procedure: Start out at stage 0 with r and let $n_0 = 0$. At stage k, if $b(r)$ contains no e' such that $n_k|e' = n_k|e$, apply iii to obtain an n_{k+1} such that $b(r * [[e|n_{k+1}]]) \cap [n_{k+1}, e(n_{k+1})] = \emptyset$. Otherwise stop the procedure.

The procedure halts by stage $r(e)$, for by ii, $(r*[[e|n_{k+1}]])(e) \leq (r*[[e|n_k]])(e)-1$ (i.e., e drops by at least one step at each stage) and when $e \in b(r * [[e|n_{r(e)}]])$, the condition for continuing is no longer satisfied. Let k be the last stage and let $m = n_k$. Then by the halting condition, we have $b(r)$ contains an e' such that $m|e' = m|e$. By v, there is an $m' \geq m$ such that $\emptyset \subset b(r * [[e|m']]) \subseteq [m'|e]$ By vi, this situation remains for each $m'' \geq m'$. So $(r, *)$ projects K.

Proof that $*_{R,1}$ is a complete projection architecture. Follow the steps in the preceding argument. A shorter argument may be given using the climbing lemma.

Proof that $*_{R,2}$ is a complete identification architecture. Recall that $\rho_k(e, e') = |\Delta(e, e') \cap \{0, \ldots, k-1\}|$. We will use the fact that ρ_k satisfies the triangle inequality.

Suppose K is identifiable. So by proposition 2, K is countable. If $e \in K$ then let $[e]_K$ be the set of all finite variants of e in K. Since K is countable, we may enumerate these classes as $C_0, \ldots, C_n, \ldots$. For each i, choose a unique element $e_i \in C_i$. For each $e \in K$, let $z(e)$ denote the unique w such that $e \in C_w$. Now define the IA r as follows: $r(e) = \rho(e_{z(e)}, e) + z(e)$. Let $e \in K$ and let (i) $r(e) = m$ and $z(e) = w$. Define $P = \{i \leq m : i \neq w\}$. If $i \in P$, then there are infinitely many m such that $e_i(m) \neq e(m)$, so there is a k_i such that $\rho_{k_i}(e_i, e) > 2m$. Moreover, there is a j sufficiently large so that $\rho_j(e_w, e) = \rho(e_w, e)$. Since P is finite, let $k = \max(\{k_i : i \in P\} \cup \{j\})$. Let $k' \geq k$. So (ii) $\rho_{k'}(e_i, e) > 2m$. We now establish that (iii) $\forall k' \geq k, e' \in K, e' \neq e \Rightarrow (r *_{R,2} [[e|k']])(e') > (r *_{R,2} [[e|k']])(e)$. Let $e' \in K$, $e' \neq e$.

dpident1.eps

Figure 5: Completeness of $*_{R,2}$

Case I: $e' \in [e]^{\leq}_{r} - [e]_K$ (cf. fig. 5). So $z(e') \neq w$. Let $z(e') = i$. So by the definition of r, (iv) $\rho_{k'}(e_i, e') \leq \rho(e_i, e') \leq m$. By the triangle inequality: $\rho_{k'}(e, e') + \rho_{k'}(e_i, e') \geq \rho_{k'}(e_i, e)$, so $\rho_{k'}(e, e') \geq \rho_{k'}(e_i, e) - \rho_{k'}(e_i, e') > 2m - m = m$, by ii, iv. Hence (v) $\rho_{k'}(e, e') > m$. By the climbing lemma (proposition 6), $(r *_{R,2} [[e|k']])(e) - (r *_{R,2} [[e|k']])(e') \leq (r(e) - r(e')) - 2\rho_{k'}(e, e') < m - r(e') - 2m \leq 0$ (by i, v), so iii obtains in this case.

Case II: $e' \in [e]^{\leq}_{r} \cap [e]_K$ (cf. fig. 5). By choice of k, (vi) $\rho_{k'}(e_w, e) = \rho(e_w, e)$. By the triangle inequality, $\rho_{k'}(e, e') + \rho_{k'}(e_w, e) \geq \rho_{k'}(e_w, e)$, so (vii) $\rho_{k'}(e, e') \geq \rho_{k'}(e_w, e) - \rho_{k'}(e_w, e')$. By the definition of r:

$$\begin{aligned} r(e) - r(e') &= (\rho(e_w, e) + w) - (\rho(e_w, e') + w) \\ &= \rho(e_w, e) - \rho(e_w, e') \\ &\leq \rho_{k'}(e_w, e) - \rho_{k'}(e_w, e') \text{ (by vi)} \\ &\leq \rho_{k'}(e, e') \text{ (by vii).} \end{aligned}$$

So (viii) $r(e) - r(e') \leq \rho_{k'}(e, e')$. By proposition 6,

$$\begin{aligned} (r *_{R,2} [[e|k']])(e') - (r *_{R,2} [[e|k']])(e) &\leq (r(e') - r(e)) - 2\rho_{k'}(e, e') \\ &\leq \rho_{k'}(e, e') - 2\rho_{k'}(e, e') \text{ (by viii),} \end{aligned}$$

which quantity is negative, so long as $\rho_{k'}(e, e') > 0$.[25] So it suffices for iii to show that $\rho_{k'}(e, e') > 0$. Suppose $\rho_{k'}(e, e') = 0$. Then (ix) $e|k' = e'|k'$. By vi, we have (x) $k'|e = k'|e_w$. By the case hypothesis, $r(e) \geq r(e')$. So by the definition of r, $\rho(e_w, e) + w \geq \rho(e_w, e') + w$, so $\rho(e_w, e) \geq \rho(e_w, e')$. So by ix, x, (xi) $k'|e' = k'|e_w$. But by ix, x, xi, we

[25] For those tracing the magic of $\alpha = 2$, note that the argument would fail here if $\alpha = 1$.

have $e = e'$, contradicting the choice of e'. Hence, $\rho_{k'}(e, e') > 0$ and we have iii under this case.

Case III: $e' \notin [e]_r^{\leq}$. Then iii follows by positive order-invariance (proposition 7) and proposition 6. This concludes the argument for iii.

By proposition 7, $r{*}_{R,2}$ is locally consistent. Hence, for each $k' \geq k, b(r{*}_{R,2}[[e|k']]) \neq \emptyset$. So by iii, we have that for each $k' \geq k, b(r *_{R,1} [[e|k']]) = \{e\}$. $\dashv$

0.14 Appendix III: A Positive Result for S,2

This appendix is devoted to proving that $*_{J,2}$ can identify $G^\omega(e_0)$.

Definition 17 *Let α be an ordinal and let $e_0 \in B$. Let $r = r_{e_0}^G$. Let ϵ be a finite boolean sequence of nonzero length and let last(ϵ) denote the last item occurring in ϵ. Then define:*

1. $$\beta_\alpha(\epsilon, b) = \begin{cases} -1 & \text{if } b = \mathit{last}(\epsilon) \wedge \mathit{lh}(\epsilon) - 1 \in \Gamma(e_0, \epsilon) \\ 0 & \text{if } b = \mathit{last}(\epsilon) \wedge \mathit{lh}(\epsilon) - 1 \notin \Gamma(e_0, \epsilon) \\ \alpha & \text{if } b \neq \mathit{last}(\epsilon) \wedge \mathit{lh}(\epsilon) - 1 \in \Gamma(e_0, \epsilon) \\ \alpha - 1 & \text{if } b \neq \mathit{last}(\epsilon) \wedge \mathit{lh}(\epsilon) - 1 \notin \Gamma(e_0, \epsilon). \end{cases}$$
2. $\beta'_\alpha(\epsilon, e') = \Sigma_{i=1}^{lh(\epsilon)} \beta(\epsilon|i, e'(i-1))$.
3. $\beta^r_\alpha(\epsilon, e') = r(e') + \beta'_\alpha(\epsilon, e')$.

Proposition 19 *Let e_0, r be as in the preceding definition. Let $e, e' \in G^\omega(e_0)$. Let $e[m] = e_0 \ddagger \{i < m : i \in \Gamma(e_0, e)\}$ and let $\alpha \geq 2$. Then*

1. $\beta^r_\alpha(e|m, e') \geq 0$.
2. *If* $e|m = e'|m$ *then* $\beta^r_\alpha(e|m, e') = |\{i \geq m : i \in \Gamma(e_0, e')\}|$.
3. *If* $e' \neq e[m]$ *then* $\beta^r_\alpha(e|m, e') > 0$.

Proof: Define $M = \{i \in \omega : i < m\}; G = \Gamma(e_0, e); G' = \Gamma(e_0, e'); E = \{i \in \omega : e(i) = e'(i)\}$.

Proof of (1). Using the definition of β_α and the fact that $\alpha \geq 2$, we have:

$$\begin{aligned}
\beta^r_\alpha(e|m, e') &= r(e') + \beta'_\alpha(e|m, e') \\
&= r(e') + \Sigma_{i \in M \cap G \cap E} \beta_\alpha(e|i+1, e'(i)) \\
&\quad + \Sigma_{i \in (M \cap E) - G} \beta_\alpha(e|i+1, e'(i)) \\
&\quad + \Sigma_{i \in (M - E) \cap G} \beta_\alpha(e|i+1, e'(i)) \\
&\quad + \Sigma_{i \in (M - E) - G} \beta_\alpha(e|i+1, e'(i)) \\
&\geq |G'| - |M \cap G \cap E| + 0 + |(M - E) \cap G| + |(M - E) - G| \\
&= |G'| - |M \cap G \cap E| + |M - E| \\
&= |G'| - |M \cap G \cap E \cap G'| + |M - E| - |M \cap G \cap E - G'| \\
&\geq |M - E| - |M \cap G \cap E|,
\end{aligned}$$

so it suffices to show that $|M - E| \geq |M \cap G \cap E - G'|$. For this we construct an injection f from $|M \cap G \cap E - G'|$ to $|M - E|$. Let $i \in M \cap G \cap E - G'$. So we have (i) $i < m$, (ii) $e(i) = e'(i)$, (iii) $i \in \Gamma(e_0, e)$ and (iv) $i \notin \Gamma(e_0, e')$. Suppose for reductio that $i = 0$. Then by iv, $e_0(i) = e'(i)$ and by ii, $e(i) = e'(i)$, so $e_0(i) = e(i)$, contradicting iii. So we may assume (v) $i > 0$. Define $f(i) = i - 1$, which is evidently injective and it is also immediate that $f(i) \in M$ if $i \in M$. Suppose for reductio that $f(i) = i - 1 \in E$, so $e(i-1) = e'(i-1)$. Then by iii, iv, v, we obtain $e(i) \neq e_0(i)$, contradicting ii. Hence, $f(i) \in M - E$.

Proof of (2). Note that $r(e') = |G'|$. Suppose that $e|m = e'|m$. For each $j \leq m$, if $j \in M \cap G'$, then $\beta_\alpha(e|j+1, e'(j)) = -1$ and if $j \in M - G'$, then $\beta_\alpha(e|j+1, e'(j)) = 0$. Hence, $\beta'_\alpha(e|m, e') = -|M \cap G'|$. So $\beta^r_\alpha(e|m, e') = |G'| - |M \cap G'| = |G' - M| = |\{i \geq m : i \in \Gamma(e_0, e')\}|$.

Proof of (3). We begin by establishing (i) $\beta^r_\alpha(e|m, e') = 0 \Rightarrow M \cap G' \subseteq G$. Suppose for contraposition that $M \cap G' - G \neq \emptyset$. Let k be the least element of $M \cap G' - G$. We will construct e'' such that $\beta^r_\alpha(e|m, e'') < \beta^r_\alpha(e|m, e')$, so by 1, $\beta^r_\alpha(e|m, e') > 0$.

The construction of e'' proceeds as follows. If $e'(k) = e(k)$, let e'' be just like e' except that $e''(k-1) = \neg e'(k-1)$. Else, e'' is just like e' except that $e''(k) = \neg e'(k)$. This construction is well-defined because $e'(k) \neq e(k)$ if $k = 0$. Let $G'' = \Gamma(e_0, e'')$. We now show that (ii.a) $r(e'') \leq r(e')$ and (ii.b) $\beta'_\alpha(e|m, e'') < \beta'_\alpha(e|m, e')$. Since $k \in G' - G$, we have (iii.a) $e(k-1) = e(k) \Leftrightarrow e_0(k-1) = e_0(k)$ and (iii.b) $e'(k-1) \neq e'(k) \Leftrightarrow e_0(k-1) = e_0(k)$. So (iii.c) $e(k-1) \neq e(k) \Leftrightarrow e'(k-1) = e'(k)$.

Case: $e(k) = e'(k)$. Then $k > 0$ and e'' is just like e' except that $e''(k-1) = \neg e'(k-1)$. So (iv) $e(k) = e'(k) = e''(k)$. So by the case hypothesis and iii.c, $e(k-1) \neq e(k) \Leftrightarrow e'(k-1) = e(k)$. Hence, (v) $e''(k-1) = e(k-1) \neq e'(k-1)$. Also, by iii.b (vi) $k \notin G''$. Since e'' differs from e' only at $k-1$, we also have: (vii) for all $j \notin \{k, k-1\}, j \in G'' \Leftrightarrow j \in G'$. By vi, vii, and the fact that $k \in G'$, we have that $|G''| \leq |G'|$, which is just ii.a. Let $i < m$. If $i \notin \{k, k-1\}$, then $e'(i) = e''(i)$ and $i \in G' \Leftrightarrow i \in G''$, so (viii) $\beta_\alpha(e|i+1, e'(i)) = \beta_\alpha(e|i+1, e''(i))$.

Subcase: $k-1 \in G'$. Then $k-1 \notin G$. Using iv and v, we may calculate:

$$\begin{aligned}
\beta_\alpha(e|k, e'(k-1)) - \beta_\alpha(e|k, e''(k-1)) &= \alpha - 0 \geq 2; \\
\beta_\alpha(e|k+1, e'(k)) - \beta_\alpha(e|k+1, e''(k)) &= -1 - 0 = -1.
\end{aligned}$$

Hence by viii, $\beta'_\alpha(e|m, e') - \beta'_\alpha(e|m, e'') \geq 1$.

Subcase: $k-1 \notin G'$. Then $k-1 \in G$. Using iv and v, calculate:

$$\begin{aligned}
\beta_\alpha(e|k, e'(k-1)) - \beta_\alpha(e|k, e''(k-1)) &= (\alpha - 1) - (-1) \geq 2; \\
\beta_\alpha(e|k+1, e'(k)) - \beta_\alpha(e|k+1, e''(k)) &= -1 - 0 = -1.
\end{aligned}$$

Hence, by viii, $\beta'_\alpha(e|m, e') - \beta'_\alpha(e|m, e'') \geq 1$, so ii.b follows in either case.[26]

Case: $e'(k) \neq e(k)$. Then e'' is just like e' except that $e''(k) = \neg e'(k)$. So by the case hypothesis, (ix) $e''(k) = e(k) \neq e'(k)$. Since $k \in G'$, and $e''(k) = \neg e'(k)$, it follows that (x) $k \notin G''$. Since e'' differs from e' only at k, we also have: (xi) for all $j \notin \{k, k+1\}, j \in G'' \Leftrightarrow j \in G'$. By ix, x we have $|G''| \leq |G'|$, which is just ii.a. Let $i < m$. If $i \notin \{k, k+1\}$, then $e'(i) = e''(i)$ and $i \in G' \Leftrightarrow i \in G''$. So again viii holds. Since $k \in G' - G \cup G''$, ix yields $\beta_\alpha(e|k+1, e'(k)) - \beta_\alpha(e|k+1, e''(k)) = \alpha - 0 \geq 2$. So if $k+1 = m$, we have by viii that $\beta'_\alpha(e|m, e') - \beta_\alpha(e|m, e'') \geq 1$, and hence ii.b. So we may assume that $k+1 < m$.

Subcase: $k+1 \in G'$. Then $k+1 \notin G''$. Suppose $e(k+1) = e'(k+1)$. Then $\beta_\alpha(e|k+2, e'(k+1)) - \beta_\alpha(e|k+2, e''(k+1)) = -1 - 0 = -1$. So by viii, $\beta'_\alpha(e|m, e') - \beta'_\alpha(e|m, e'') \geq 2 + (-1) = 1$.[27] Suppose, alternatively, that $e(k+1) \neq e'(k+1)$. Then $\beta_\alpha(e|k+2, e'(k+1)) - \beta_\alpha(e|k+2, e''(k+1)) = \alpha - (\alpha - 1) \geq 1$. So by viii, $\beta'_\alpha(e|m, e') - \beta'_\alpha(e|m, e'') \geq 2 + 1 = 3$.

Subcase: $k+1 \notin G'$. Then $k+1 \in G''$. Suppose $e(k+1) = e'(k+1)$. Then $\beta_\alpha(e|k+2, e'(k+1)) - \beta_\alpha(e|k+2, e''(k+1)) = 0 - (-1) = 1$. Then by viii, $\beta'_\alpha(e|m, e') - \beta'_\alpha(e|m, e'') \geq 2 + 1 = 3$. Suppose, alternatively, that $e(k+1) \neq e'(k+1)$. Then $\beta_\alpha(e|k+2, e'(k+1)) - \beta_\alpha(e|k+2, e''(k+1)) = (\alpha - 1) - \alpha = -1$. Then by viii, $\beta'_\alpha(e|m, e') - \beta'_\alpha(e|m, e'') \geq 2 + (-1) = 1$. So ii.b holds in both subcases.[28]

The next task is to establish: (xii) $\beta^r_\alpha(e|m, e') = 0 \Rightarrow G' - M = \emptyset$. Suppose that $k \geq m$ and $k \in G'$. So k contributes one unit to $r(e')$. Since $k \geq m$, k contributes nothing to the sum $\beta'_\alpha(e|m, e')$. Let $e'' = e_0 \ddagger (G' - \{k\})$. Then

$$\beta^r_\alpha(e|m, e'') \quad = \quad r(e'') + \beta'_\alpha(e|m, e'')$$

[26] Note that the value $\alpha \geq 2$ is critical in both cases.

[27] Observe that the value $\alpha \geq 2$ is critical at this step.

[28] The value $\alpha \geq 2$ is again critical at this step.

$$= \quad r(e') - 1 + \beta'_\alpha(e|m, e')$$

$$= \quad \beta^r_\alpha(e|m, e') - 1.$$

So by (1), $\beta^r_\alpha(e|m, e') > 0$.

Finally we show that (xiii) $\beta^r_\alpha(e|m, e') = 0 \Rightarrow M \cap G \subseteq G'$. Suppose that $\beta^r_\alpha(e|m, e') = 0$. Suppose for reductio that $D = (M \cap G) - G' \neq \emptyset$. By the hypothesis and i, xii, we have $G' - M = \emptyset$ and $G' \cap M \subseteq G \cap M$. So $r(e) - r(e') = |G| - |G'| = |D|$. So if we establish that (xiv) $\beta'_\alpha(e|m, e') - \beta'_\alpha(e|m, e) > |D|$, then we have $\beta^r_\alpha(e|m, e') > \beta^r_\alpha(e|m, e)$, so by (1), $\beta^r_\alpha(e|m, e') > 0$. It therefore suffices to establish xiv.

Let D be enumerated in ascending order as $\{k_1, \ldots, k_d\}$. Observe that $e|k_1 = e'|k_1$ so since $k_1 \in G - G'$, $e(k_1) \neq e'(k_1)$. Thereafter, there is constant disagreement between e and e' until k_2, where another reversal of sense yields constant agreement until k_3, etc. In general, we have for each j such that $1 \leq k \leq d$: (xv) $e(k_j) = e'(k_j) \Leftrightarrow j$ is even. Also, we have by the definition of β_α: (xvi) if $e(k_j) \neq e'(k_j)$ then $\beta_\alpha(e|k_j + 1, e'(k_j)) - \beta_\alpha(e|k_j + 1, e(k_j)) = \alpha - (-1) \geq 3$ (since[29] $\alpha \geq 2$) and (xvii) if $e(k_j) = e'(k_j)$ then $\beta_\alpha(e|k_j + 1, e'(k_j)) - \beta_\alpha(e|k_j + 1, e(k_j)) = (-1) - (-1) = 0$. By xv, xvi, xvii, we have (xviii) $\Sigma^d_{j=1}\beta_\alpha(e|k_j + 1, e'(k_j)) - \beta_\alpha(e|k_j + 1, e(k_j)) \geq 3(d + 1)/2$ if d is odd and $\geq 3d/2$ if d is even (note that $3(d + 1)/2$ is the number of odd natural numbers $\leq d$ when d is odd). Observe that (xix) for all $d > 0$, $3(d - 1)/2 > d$ if d is odd and $3d/2 > d$ if d is even.[30] We haven't yet included in the sum terms whose indices are not in D. So let $0 < k \leq m$ and suppose $k - 1 \notin D$. Then by the definition of β_α, $\beta_\alpha(e|k, e(k - 1)) \leq 0$, so we have (xx) $\beta_\alpha(e|k, e'(k - 1)) - \beta_\alpha(e|k, e(k - 1)) \geq 0$. So by xviii, xix, xx, we have, $\beta'_\alpha(e|m, e') - \beta'_\alpha(e|m, e) > d$, establishing xiv and hence xiii.

[29] $\alpha = 2$ is critical for the argument at this stage.

[30] The inequality is barely strict at $d = 1$ and would fail if $\alpha = 1$, illustrating once again the critical role of the value $\alpha \geq 2$.

Now suppose that $\beta^r_\alpha(e|m, e') = 0$. By i, xii, xiii, we infer that $e' = e_0 \ddagger \{i < m : i \in \Gamma(e_0, e)\} = e[m]$, which completes the proof of (3). $\dashv$

Proposition 20 *Let $\alpha \geq 2$ and let $e_0, r, e, e', e[n]$ be as in proposition 19 and let $m \in \omega$. Then*

1. $(r *_{J,2} [[e|n]])(e') = \beta^r_\alpha(e|n, e')$,

2. $(r, *_{J,2})(e|n) = \{e[n]\}$, *and*

3. $(r, *_{J,2})$ *identifies* $G^\omega(e_0)$.

Proof of (2). By proposition 19.2, $\beta^r_\alpha(e|n, e[n]) = 0$. By proposition 19.3, for all $e' \neq e[m]$, $\beta^r_\alpha(e|n, e') > 0$. So the result follows from (1).

Proof of (3). (3) is a consequence of (2), since for each $e \in G^\omega(e_0)$, there is a least n' such that $e = e[n']$ and $*_{J,2}$ is timid and stubborn. Note that $(r, *_{J,2})$ retracts exactly n' times prior to stabilizing to $\{e\}$.

Proof of (1). By induction on n. Let $* = *_{J,\alpha}$. $\beta^r_\alpha((), e') = r(e') + \beta'_\alpha((), e') = r(e') = (r * ())(e')$. Now suppose that for each $e' \in G^\omega(e_0), (r * [[e|n]])(e') = \beta^r_\alpha(e|n, e')$. Then since $\alpha \geq 2$, proposition 19 parts 2 and 3 yield (i) If $e|n = e'|n$ then $(r * [[e|n]])(e') = |\{i \geq n : i \in \Gamma(e_0, e')\}|$ and (ii) If $e' \neq e[n]$ then $(r * [[e|n]])(e') > 0$. Now consider $(r * [[e|n + 1]])(e')$.

Case 1: $e'(n) = e(n)$ Then

$$\begin{aligned}(r * [[e|n + 1]])(e') &= (r * [[e|n]])(e'|[n, e(n)]) \\ &= -\min\{(r * [[e|n]])(e'') : e'' \in \text{ dom}(r * [[e|n]]) \cap [n, e(n)]\} \\ &\quad +(r * [[e|n]])(e').\end{aligned}$$

Case 1.A: $n \in \Gamma(e_0, e)$. Hence, $e[n] \notin [n, e(n)]$ (recall that $e[n] = e_0 \ddagger \{i < n : i \in \Gamma(e_0, e)\}$). So by ii, we have (iii) $0 \notin \{(r * [[e|n]])(e'') : e'' \in \text{dom}(r * [[e|n]]) \cap [n, e(n)]\}$.

Since $e[n+1]|n+1 = e|n+1$ and $\{i \geq n : i \in \Gamma(e_0, e[n+1])\} = \{n\}$, we obtain by statement i that: (iv) $(r * [[e|n]])(e[n+1]) = 1$. Also, (v) $e[n+1] \in [n, e(n)]$, since $e[n] \notin [n, e(n)]$ and $n \in \Gamma(e_0, e)$. By iii, iv, v: $\min\{(r * [[e|n]])(e'') : e'' \in \mathrm{dom}(r * [[e|n]]) \cap [n, e(n)]\} = 1$. So,

$$\begin{aligned}
(r * [[e|n+1]])(e') &= (r * [[e|n]])(e') - 1 \\
&= \beta^r_\alpha(e|n, e') - 1 \text{ (by the induction hypothesis)} \\
&= \beta^r_\alpha(e|n, e') + \beta_\alpha(e|n+1, e'(n)) \text{ (by the case hypotheses)} \\
&= \beta^r_\alpha(e|n+1, e').
\end{aligned}$$

Case 1.B: $n \notin \Gamma(e_0, e)$. Hence, $e[n] \in [n, e(n)]$. So by statement i we have: $\min\{(r * [[e|n]])(e'') : e'' \in \mathrm{dom}(r * [[e|n]]) \cap [n, e(n)]\} = 0$. Hence,

$$\begin{aligned}
(r * [[e|n+1]])(e') &= (r * [[e|n]])(e') \\
&= \beta^r_\alpha(e|n, e') + 0 \text{ (by the induction hypothesis)} \\
&= \beta^r_\alpha(e|n, e') + \beta_\alpha(e|n+1, e'(n)) \text{ (by the case hypotheses)} \\
&= \beta^r_\alpha(e|n+1, e').
\end{aligned}$$

Case 2: $e'(n) \neq e(n)$. Then

$$\begin{aligned}
(r * [[e|n+1]])(e') &= (r * [[e|n]])(e'|B - [n, e(n)]) + \alpha \\
&= -\min\{(r * [[e|n]])(e'') : e'' \in \ \mathrm{dom}(r * [[e|n]]) \cap (B - [n, e(n)])\} \\
&\quad +(r * [[e|n]])(e') + \alpha.
\end{aligned}$$

Case 2.A: $n \in \Gamma(e_0, e)$. Hence, $e[n] \notin [n, e(n)]$. So by statement i we obtain, $\min\{(r * [[e|n]])(e'') : e'' \in \mathrm{dom}(r * [[e|n]]) \cap (B - [n, e(n)])\} = 0$. Hence,

$$(r * [[e|n+1]])(e') = 0 + (r * [[e|n]])(e') + \alpha$$

$$
\begin{aligned}
&= \beta^r_\alpha(e|n, e') + \alpha \text{ (by the induction hypothesis)} \\
&= \beta^r_\alpha(e|n, e') + \beta_\alpha(e|n+1, e'(n)) \text{ (by the case hypotheses)} \\
&= \beta^r_\alpha(e|n+1, e').
\end{aligned}
$$

Case 2.B: $n \notin \Gamma(e_0, e)$. Hence, $e[n] \in [n, e(n)]$. So by ii, we obtain: (vi) $\min\{(r * [[e|n]])(e'') : e'' \in \mathrm{dom}(r * [[e|n]]) \cap B - [n, e(n)]\} > 0$. Let $e'' = e[n] \ddagger n$. Then $e''|n = e|n$ and $\{i \leq n : i \in \Gamma(e_0, e'')\} = \{n\}$. So by i, we obtain: (vii) $(r * [[e|n]])(e'') = 1$. Since $e[n](n) = e(n)$, (viii) $e'' \notin [n, e(n)]$. Since $e'' \in G^\omega(e_0) = \mathrm{dom}(r)$, vi, vii, viii yield $\min\{(r * [[e|n]])(e'') : e'' \in \mathrm{dom}(r * [[e|n]]) \cap (B - [n, e(n)])\} = 1$. So

$$
\begin{aligned}
(r * [[e|n+1]])(e') &= (r * [[e|n]])(e') + \alpha - 1 \\
&= \beta^r_\alpha(e|n, e') + \alpha - 1 \text{ (by the induction hypothesis)} \\
&= \beta^r_\alpha(e|n, e') + \beta_\alpha(e|n+1, e'(n)) \text{ (by the case hypotheses)} \\
&= \beta^r_\alpha(e|n+1, e'). \dashv
\end{aligned}
$$

Proof of proposition 13 Let $* = *_{J,2}$. Recall that $(\neg e_0)(n) = \neg(e_0(n))$. Let $r \supseteq r^H_{e_0}|G^\omega_{\text{even}}$ and suppose for reductio that $(r, *_{J,2})$ identifies $G^\omega_{\text{even}}(e_0) \cup \{\neg e_0\}$. For each i, let $e_i = (\neg e_0) \ddagger i = (e_0 \ddagger 0) \ddagger i \in G^2_{\text{even}}(e_0)$.

Case A: $r(\neg e_0) \geq \omega$. Then for each $i, r(e_i) < r(\neg e_0)$, contradicting the isolation condition (proposition 22).

Case B: for some $n \in \omega$ that $r(\neg e_0) = n$. By the reductio hypothesis, there is a $k \in \omega$ such that (i) $(r * [[\neg e_0|k]])^{-1}(0) = \{\neg e_0\}$. Let $j = \max\{n+1, k\}$. Then (ii) $r(e_j) > n = r(\neg e_0)$ and $r(e_{j+1}) = r(e_j) + 1$ and (iii) $e_j|j = e_{j+1}|j = \neg e_0|j$ and $e_{j+1}|j+1 = \neg e_0|j+1$. By timidity and stubbornness and i, iii, for each j' such that $k \leq j' \leq j+1$, (iv) $(r * [[e_{j+1}|j']])^{-1}(0) = \{\neg e_0\}$. By iv, (v) $(r * [[e_{j+1}|j]])(e_j) > 0$. By positive invariance and ii, (vi) $(r * [[e_{j+1}|j]])(e_{j+1}) = (r * [[e_{j+1}|j]])(e_j) + 1$. By positive invariance, iv, and iii, (vii) $(r * [[e_{j+1}|j+1]])(e_{j+1}) = (r * [[e_{j+1}|j]])(e_{j+1})$. By iii and iv, $\min\{(r * [[e_{j+1}|j]])(e') : e' \in \mathrm{dom}(r) \cap (B - [j, e_{j+1}(j)])\} \geq 1$. So since $e_j(j) \neq e_{j+1}(j)$, the definition of $*$ yields

$$\begin{aligned}(r * [[e_{j+1}|j+1]])(e_j) &\leq -1 + (r * [[e_{j+1}|j]])(e_j) + 2 \\ &\leq (r * [[e_{j+1}|j]])(e_j) + 1 \\ &\leq (r * [[e_{j+1}|j]])(e_{j+1}) \text{ (by vi).}\end{aligned}$$

Now $(j+1|e_j) = (j+1|e_{j+1})$, so by positive invariance, for all $k' \geq j+1$, $(r*[[e_{j+1}|k']])(e_j) \leq (r * [[e_{j+1}|k']])(e_{j+1})$. Hence, for all such k', $(r * [[e_{j+1}|k']])^{-1}(0) \neq \{e_{j+1}\}$. $\dashv$

0.15 Appendix IV: Restrictiveness Proofs

Proof of proposition 16 (1) Case: $G^1(e_0)$. Let $R = (G^1(e_0)), \leq)$ be defined so that $b(R) = \{e_0\}$ and for each $k, k' > k \in \omega$, $e_0 \ddagger k' < e_0 < e_0 \ddagger k$ (note that this condition induces an infinite descending chain in R). It is easy to see that R is an epistemic state and that $(R, *_M)$ succeeds.

Case: $G^1_{\text{even}}(e_0)$. Let $R = (G^1_{\text{even}}(e_0), \leq)$ be defined so that $b(R) = \{e_0\}$ and if $|S| = |S'| = 2$ then $e_0 \ddagger S \leq e_0 \ddagger S'$ just in case $\min(S') \leq \min(S)$ (this condition also induces an infinite descending chain). R is an epistemic state and $(R, *_M)$ succeeds.

(2) Case: $G^1(e_0)$. Let $* = *_M$, let $R = (D, <_R)$, and suppose for reductio that $(R, *)$ identifies $G^2(e_0)$. Then for some k, (i) $b(R*[[e_0|k]]) = \{e_0\}$. Define $e = e_0 \ddagger k \in G^2(e_0)$. By proposition 23, we can find $k' > k$ such that, letting $e' = e \ddagger k' \in G^2(e_0)$, (ii) $e' >_{(R*[[e_0|k]])} e$. But i, ii, and proposition 21 contradict the reductio hypothesis.

Case: $G^1_{\text{even}}(e_0)$. Let $* = *_M$ and suppose for reductio that $R = (D, *)$ identifies $G^2_{\text{even}}(e_0)$. Then for some k, (i) $b(R * [[e_0|k]]) = \{e_0\}$. Define $e = e_0 \ddagger \{k, k+1\} \in$

$G^2_{\text{even}}(e_0)$. By proposition 25, we can find $k' > k+1$ such that, letting $e' = e \ddagger \{k', k'+1\} \in G^2_{\text{even}}(e_0)$, (ii) $e' >_{(R*[[e_0|k]])} e$. Let $e'' = e_0 \ddagger \{k', k'+1\} \in G^2_{\text{even}}(e_0)$. By proposition 21 and i, (iii) $e' \geq_{(R*[[e_0|k]])} e''$. Note that: (iv) $e_0|k = e'|k$, (v) $e(k) = e'(k) \neq e''|k$, and (vi) $k+1|e'' = k+1|e'$. By ii, iv, (vii) $e' \notin \min((R * [[e'|k]]), [k, e'(k)])$. By v, (viii) $e'' \notin \min((R * [[e'|k]]), [k, e'(k)])$. So by iii, v, vii, viii, and clause (2) of the definition of $*_M$: (ix) $e' \geq_{(R*[[e'|k+1]])} e''$. So by vi and positive order-invariance (proposition 7), for all $k' \geq k+1$, $e' \geq_{(R*[[e'|k']])} e''$, contradicting the reductio hypothesis. $\dashv$

Definition 18 *e is propped up at n in r just in case for each $e' \in [e]^{<}_{r}, e'(n) \neq e(n)$. e is propped up in r just in case there exists an n such that e is propped up at n in r.*

Proposition 21 (propping condition for $*_M$) *If $(r, *_M)$ identifies dom(r) then for each $e \in$ dom(r), for each m, there is an $m' \geq m$ such that e is propped up in $(r *_M [[e|m]])$ at m'; so in particular, e is propped up in r.*

Proof: Suppose that for all $m' \geq m$, $e \in \text{dom}(r*_M[[e|m]])$ is not propped up in $(r*_M[[e|m]])$ at m'. Using the definition of $*_M$, show by a straightforward induction on $k - m$ that for all $k \geq m, k' \geq k$, e is not propped up in $(r *_M [[e|k]])$ at k'. Hence, for all $k \geq m$, $(r *_M [[e|k]])(e) > 0$, so $(r, *_M)$ does not identify dom(r). $\dashv$

The following definition generalizes the notion of isolated points to the case in which there is sufficient information after a given position n to distinguish e from all other points in S. Observe that k-isolation is more stringent than isolation when $k > 0$. For example, 0^∞ is isolated but not 1-isolated in $\{10^n 1^\infty : n \in \omega\}$.

Definition 19 *e is k-isolated in $S \Leftrightarrow$ there exists an $n \geq k$ such that $[(k|e)|n] \cap S \subseteq \{e\}$.*

Proposition 22 (isolation condition) *If $*$ is positively order-invariant and $(r, *)$ identifies dom(r) then for each $e \in$ dom(r), for all k, e is k-isolated in $[e]^{\leq}_{(r*[[e|k]])}$; so in particular, e is isolated in $[e]^{\leq}_{r}$.*

Proof: Suppose e is not k-isolated in $[e]^{\leq}_{(r*[[e|k]])}$. Then for each $n \geq k$, there is an $e_n \neq e$ such that $(r * [[e|k]])(e_n) \leq (r * [[e|k]])(e)$ and $(k|e_n)|n = (k|e)|n$. So by positive order-invariance, for each n, if $(r * [[e|n]])(e) = 0$ then $(r * [[e_n|n]])(e') = 0$. Hence, $(r, *)$ does not identify $\text{dom}(r)$. ⊣

Proposition 23 *If $*$ is positively order-invariant and $(r, *)$ identifies $G^{n+1}(e_0) \subseteq dom(r)$ and $e \in G^n(e_0)$ then for all k, for all but finitely many j, $(r*[[e_0|k]])(e) < (r*[[e_0|k]])(e \ddagger j)$.*

Proof: Let $e \in G^n(e_0)$. Then for each j, $e \ddagger j \in \text{dom}(r)$. Suppose that for some k there are infinitely many distinct j such that $(r * [[e|k]])(e) \geq (r * [[e|k]])(e \ddagger j)$. Then e is not k-isolated in $[e]^{\leq}_{(r*[[e_0|k]])}$. Apply proposition 22. ⊣

Proposition 24 (stacking lemma) *For all $k, n, n' \leq n$, if $*$ is positively order-invariant and $(r, *)$ identifies $G^n(e_0) \subseteq dom(r)$ and $(r * [[e_0|k]])(e_0) = 0$ then there exists an $e_{n'}$ such that*

1. *$e_{n'} \in g^{n'}(e_0)$,*
2. *$e_0|(k + 1) = e_{n'}|(k + 1)$ and*
3. *$(r * [[e_0|k]])(e_{n'}) \geq n'$.*

Proof: Assume the antecedent. Let n, k be given. We show the consequent by induction on $n' \leq n$. When $n' = 0$, (1-3) are trivially satisfied by e_0. Now suppose that $n' + 1 \leq n$ and that there exist $e_0, \ldots e_{n'}$ satisfying (1-3). Since $n' + 1 \leq n$, $(r, *)$ identifies $G^{n'+1}(e_0)$. So by proposition 23, we may choose j sufficiently large so that (i) $(r * [[e_0|k]])(e_{n'}) < (r * [[e_0|k]])(e_{n'} \ddagger j)$ and (ii) $j > \max(\Gamma(e_0, e_{n'}) \cup \{k + 1\})$. Now set $e_{n'+1} = e_{n'} \ddagger j$. $e_{n'+1}$ satisfies (1, 2) because of ii and the fact that $e_{n'}$ does. $e_{n'+1}$ satisfies (3) at $n' + 1$ because of statement i and the fact that $e_{n'}$ does. ⊣

Proposition 25 *If $*$ is positively order-invariant and $(r, *)$ identifies $G^{n+1}_{even}(e_0) \subseteq dom(r)$ and $e \in G^n_{even}(e_0)$ then for all k, for all but finitely many j, for all $m > 0$, $(r*[[e_0|k]])(e) < (r * [[e_0|k]])((e \ddagger j) \ddagger j + m)$.*

Proof: Similar to the proof of proposition 23. $\dashv$

Proposition 26 (even stacking lemma) *Proposition 24 continues to hold when G^n, g^n are replaced with G^n_{even}, g^n_{even}.*

Proof: Similar to the proof of proposition 24, using proposition 25. $\dashv$

Proposition 27 (with Oliver Schulte) *For all $e_0 \in B$, For all $j \geq 2$, $G^j(e_0)$ is identifiable using just j retractions, but is not identifiable by $*_{R,1}$.*

Proof: The positive claim is from proposition 11. For the negative claim, suppose for reductio that there is an IA r such that $(r, *)$ identifies $G^j(e_0)$, where $* = *_{R,1}$ and $j \geq 2$. Then since $e_0 \in G^0(e_0)$ and $0 < j$, there exists a least n such that (i) $(r, *)(e_0|n) = \{e_0\}$. Then there exists a least $k > n$ such that (ii) $(r, *)((e_0 \ddagger n)|k) = \{e_0 \ddagger n\}$, since $e_0 \ddagger n \in G^1(e_0)$ and $1 < j$. Define $R = \{e'' \in B : |\Gamma(e_0, e'')| \text{ is odd}\}$. Since $|\Gamma(e_0, (e_0 \ddagger n))| = 1$ is odd, we have by statements i and ii that there is a least $k' > n$ such that (iii) $b(r * [[(e_0 \ddagger n)|k']]) \cap G^j(e_0) \subseteq R$. Since k' is least, there exists an e such that (iv.a) $(r*[[(e_0 \ddagger n)|(k'-1)]])(e) = 0$, (iv.b) $(r * [[(e_0 \ddagger n)|k']])(e) > 0$, and (iv.c) $e \in G^j(e_0) - R$. Since $* = *_{R,1}$, we also have[31] (iv.d) $(r * [[(e_0 \ddagger n)|k']])(e) = 1$.

Case 1: $k' = n + 1$. Then we may choose e to be e_0, by i, iv.a. Define $e' = (e_0 \ddagger n) \ddagger n + 1$. Hence, (v.a) $e'|n + 1 = (e_0 \ddagger n)|n + 1$, (v.b) $n + 1|e' = n + 1|e_0$, and (v.c) $e' \in G^j(e_0) - R$, since $|\Gamma(e_0, e')| = 2$ and $j \geq 2$.[32] By v.a, v.c and the reductio hypothesis,

[31]This is where the value $\alpha = 1$ enters the negative argument.

[32]This is where $j \geq 2$ enters the argument.

$e' \in \mathrm{dom}(r * [[(e_0 \ddagger n)|n+1]])$, else $(r, *)$ fails to identify $e' \in G^j(e_0)$. So by iii, v.a, v.c, and the case hypothesis, (vi) $(r * [[e'|n+1]])(e') \geq 1$. By ii, iv.d, v.a, the case hypothesis, and positive invariance, (vii) $(r * [[e'|n+1]])(e) = 1$. By positive invariance and v.b, vi, vii, we have that for each $m \geq n+1$, $(r * [[e'|m]])(e') \geq (r * [[e'|m]])(e_0)$, contradicting the reductio hypothesis.

Case 2: $k' \geq n+2$. Then by the definition of * and iv.a, iv.b, we have (viii.a) $e(k'-2) = (e_0 \ddagger n)(k'-2) = \neg e_0(k'-2)$ and (viii.b) $e(k'-1) = \neg(e_0 \ddagger n)(k'-1) = e_0(k'-1)$. Let e' be defined so that: (ix.a) $e'|k' = (e_0 \ddagger n)|k'$, and (ix.b) $k'|e' = k'|e$. By viii.a, there exists some $j \leq k'-2$ such that $j \in \Gamma(e_0, e)$. By viii.a,b, $k'-1 \in \Gamma(e_0, e)$. So $|\{j \leq k' : j \in \Gamma(e_0, e)\}| \geq 2$. But by ix.a,b we also have $|\{j \leq k' : j \in \Gamma(e_0, e')\}| \leq 2$. So by ix.b, $\Gamma(e', e_0) \leq \Gamma(e, e_0)$. So by iv.c, (x) $e' \in G^j(e_0)$. So by the reductio hypothesis and ix.b, (xi) $e' \in \mathrm{dom}(r * [[(e \ddagger n)|k']])$, else $(r, *)$ does not identify $e' \in G^j_{\text{even}}(e_0)$. By iv.c, $|\Gamma(e_0, e)|$ is even. Hence, e agrees almost everywhere with e_0. By ix.b, e' agrees almost everywhere with e and hence with e_0. So, $|\Gamma(e_0, e')|$ is even. So $e' \notin R$. Thus, by iii, xi, $(r * [[(e_0 \ddagger n)|k']])(e') \geq 1$. So by positive invariance, iv.d, ix.a, ix.b, we have that for all $m \geq k'$, $(r * [[e'|m]])(e') \geq (r * [[e'|m]])(e_0)$, contradicting the reductio hypothesis. $\dashv$

Proposition 28 *Let $e_0 \in B$.*

1. *$G^1(e_0)$ is identifiable by $*_{R,1}$.*

2. *For all $j \geq 1$, $G^j(e_0)$ is not identifiable by $*_M, *_{J,1}, *_{A,1}$.*

Proof of (1). Let $* = *_{R,1}$. Define $r^{-1}(0) = g^0(e_0) = \{e_0\}$ and $r^{-1}(1) = g^1(e_0) = \{e_0 \ddagger k : k \in \omega\}$. Then $\mathrm{dom}(r) = G^1(e_0)$. Let $e \in G^1(e_0)$. Case: $e = e_0$. Then by timidity and stubbornness, we have that for each $k, b(r * [[e|k]]) = \{e\}$. Case: for some n, $e = e_0 \ddagger n$. By timidity and stubbornness: (i.a) $b(r * [[e|n]]) = \{e_0\}$. So by positive invariance, we have that for all $n' \geq n$, (i.b) $(r * [[e|n]])(e_0 \ddagger n') = 1$. So by proposition 6, we have that for each $n'' < n$, (i.c) $(r * [[e|n]])(e_0 \ddagger n'') \geq (n - n'') > 1$. On data $e|n+1$, e_0 is refuted and moves up one level along with all data streams of form $e_0 \ddagger n'$, where $n' > n$. By i.a,b,c, e is the lowest data stream consistent with the data, so e drops to level 0. All data streams of form $e \ddagger n'$ such that $n' < n$ also drop one level with e, but fortunately, by i.c they all end up above level 0. So $b(r * [[e|n+1]]) = \{e\}$. By timidity and stubbornness, e remains uniquely at level 0 forever after.

Proof of (2). Case: $* = *_{A,1}$. Instance of proposition 29. Case: $* = *_{J,1}, *_M$. Suppose for reductio that there is an IA r such that $(r, *)$ identifies $G^1(e_0)$. Then there exists a least n such that (i) $b(r * [[e_0|n]]) = \{e_0\}$. Furthermore, (ii) $\exists k \geq n$ such that $\forall k' > k$, $(r * [[e_0|n]])(e_0 \ddagger k') \geq (r * [[e_0|n]])(e_0 \ddagger k)$; for otherwise, there would exist an infinite descending chain of ordinals in the range of $(r * [[e_0|n]])$. By ii, there exists a $k \geq n$ such that (iii) $(r * [[e_0|n]])(e_0 \ddagger k+1) \geq (r * [[e_0|n]])(e_0 \ddagger k)$. Observe that: (iv) $(e_0 \ddagger k)|k = (e_0 \ddagger k+1)|k = e_0|k$ and (v) $(e_0 \ddagger k+1)(k) = e_0(k) \neq (e_0 \ddagger k)(k)$. By timidity and stubbornness and i, iv, v, (vi) $\forall n', n \leq n' \leq k+1 \Rightarrow b(r * [[(e_0 \ddagger k+1)|n']]) = \{e_0\}$. By iii, iv, vi and positive order-invariance (proposition 7), (vii) $\forall n', n \leq n' \leq k \Rightarrow (r * [[(e_0 \ddagger k+1)|n']])(e_0 \ddagger k+1) \geq (r * [[(e_0 \ddagger k+1)|n']])(e_0 \ddagger k) > 0$. Now it is claimed as well that: (viii) $(r * [[(e_0 \ddagger k+1)|k+1]])(e_0 \ddagger k+1) \geq (r * [[(e_0 \ddagger k+1)|k+1]])(e_0 \ddagger k)$. For consider the case of $*_M$. By v, vi and the definition of $*_M$, $(r *_M [[(e_0 \ddagger k+1)|k+1]])(e_0 \ddagger k) = (r *_M [[(e_0 \ddagger k+1)|k]])(e_0 \ddagger k)+1$ and $(r *_M [[(e_0 \ddagger k+1)|k+1]])(e_0 \ddagger k+1) = (r *_M [[(e_0 \ddagger k+1)|k]])(e_0 \ddagger k+1)+1$. So by vii, we have viii for $*_M$.

Let us turn now to the case of $*_{J,1}$. By v, vi, $\min\{(r * [[(e_0 \ddagger k+1)|k]])(e') : e' \in \mathrm{dom}(r * [[(e_0 \ddagger k+1)|k]]) \cap [k, (e_0 \ddagger k+1)(k)]\} = 0$ and $\min\{(r * [[(e_0 \ddagger k+1)|k]])(e') : e' \in \mathrm{dom}(r * [[(e_0 \ddagger k+1)|k]]) \cap (B - [k, (e_0 \ddagger k+1)(k)])\} > 0$. So by v and the definition of $*_{J,1}$, $(r *_{J,1} [[(e_0 \ddagger k+1)|k+1]])(e_0 \ddagger k+1) = -0 + (r *_{J,1} [[(e_0 \ddagger k+1)|k]])(e_0 \ddagger k+1)$ and $(r *_{J,1} [[(e_0 \ddagger k+1)|k+1]])(e_0 \ddagger k) \geq -1 + (r *_{J,1} [[(e_0 \ddagger k+1)|k]])(e_0 \ddagger k) + 1$. So again by vii we have viii for $*_{J,1}$.

Finally, since $(k+1)|(e_0 \ddagger k) = (k+1)|(e_0 \ddagger k+1)$, we have by viii and positive order-invariance that for all $k' \geq k+1$, $(r*[[(e_0 \ddagger k+1)|k']])(e_0 \ddagger k+1) \geq (r*[[(e_0 \ddagger k+1)|k']])(e_0 \ddagger k)$, contradicting the reductio hypothesis. $\dashv$

Proposition 29 (restrictiveness of $*_{A,n}$) *Let $e_0 \in B$.*

1. *$G^0(e_0)$ is identifiable by $*_{A,0}$.*

2. *for all n, $G^{n+1}(e_0)$ is identifiable by $*_{A,n+2}$.*

3. *for all $m > n+1, G^m(e_0)$ is not identifiable by $*_{A,n+2}$.*

Proof of (1). Let $\mathrm{dom}(r) = \{e_0\}$ and let $r(e_0) = 0$. Then for all k, $(r, *_{J,0})(e_0|k) = \{e_0\}$.

Proof of (2). By propositions 9 and 11.

Proof of (3). Suppose for reductio that $(r, *_{A,n+1})$ identifies $G^{m+1}(e_0)$, with $m \geq n$. Then for some j, (i) $b(r *_{A,n+1} [[e_0|j]]) = \{e_0\}$. So by positive invariance and since $G^{m+1}(e_0) \subset \mathrm{dom}(r)$ by the reductio hypothesis, proposition 24 yields (ii) there exists an $e \in G^{n+1}(e_0) - G^n(e_0)$ such that $e_0|j+1 = e|j+1$ and $(r * [[e_0|j]])(e) \geq n+1$. $e \neq e_0$, so let z be least such that $e(z) \neq e_0(z)$. So, (iii) $z > j$. So since $z > 0$, we may define e' to be just like e except that $e'(z-1) = \neg(e_0(z-1))$. Hence, (iv) $e'(z-1) \neq e(z-1) = e_0(z-1)$. Also (v) $e|z = e_0|z$ and (vi) $z|e' = z|e$. By i, v, and the timidity of $*_{A,n+1}$, (vii) for all x such that $j \leq x \leq z$, $(r *_{A,n+1} [[e|x]])(e_0) = 0$. By positive invariance and ii, v, vii, (viii) $(r *_{A,n+1} [[e|z]])(e) \geq n+1$. By iv and the definition of $*_{A,n+1}$, (ix) $(r *_{A,n+1} [[e|z]])(e') = n+1$. By vi, viii, ix, and positive invariance, (ix) for all $k' \geq z$, $(r *_{A,n+1} [[e|k']])(e) \geq (r *_{A,n+1} [[e|k']])(e')$. Hence, $(r, *_{A,n+1})$ does not identify $G^{n+1}(e_0)$, contradicting the reductio hypothesis. $\dashv$

Proposition 30 *Let $e_0 \in B$.*

1. $G^{\omega}_{even}(e_0)$ *is identifiable by* $*_{J,1}$, $*_{R,1}$.

2. $\forall m \geq 1, G^{m}_{even}(e_0)$ *is not identifiable by* $*_M$.

3. $G^{0}_{even}(e_0)$ *is identifiable by* $*_{A,0}$.

4. $G^{n}_{even}(e_0)$ *is identifiable by* $*_{A,n+1}$.

5. $\forall m \geq n, G^{m+1}_{even}(e_0)$ *is not identifiable by* $*_{A,n+1}$.

Proof of (1). Case: $* = *_{R,1}$. Let r_e be r^H_e restricted to $G^{\omega}_{\text{even}}(e)$, so for each $e' \in G^{\omega}_{\text{even}}(e)$, $r_e(e') = \rho(e, e')$. Let e be given and let $e' \in G^{\omega}_{\text{even}}(e)$. Define e'_i so that (i.a) $e'_i|i = e'|i$ and (i.b) $i|e'_i = i|e$. Hence, $e = e'_0$. Recall that $G^{\omega}_{\text{even}}(e)$ is precisely the set of all finite variants of e, so $\Delta(e, e')$ is finite. Let $m = 1+ \max(\Delta(e, e'))$. Then (ii) for all $k \geq m, e'_k = e'$. I claim that $*$ satisfies the following symmetry conditions: for each $e', e'' \in G^{\omega}_{\text{even}}(e)$, (iii.a) $(r_e * [[e'|k]])(e'') \geq r_{e'_k}(e'')$, and (iii.b) if $e''|k = e'|k$ then $(r_e * [[e'|k]])(e'') = r_{e'_k}(e'')$. Then for each $e' \in G^{\omega}_{\text{even}}(e)$, for each $k \geq m$, $b(r_e * [[e'|k]]) = b(r_{e'_k}) = \{e'_k\} = \{e'\}$, by ii. Thus, $(r_e, *)$ identifies $G^{\omega}_{\text{even}}(e)$. So it remains only to establish iii.a, b. iii.a, b are immediate when $k = 0$. Now suppose iii.a,b hold at k. Then (iv) $b(r_e * [[e'|k]]) = \{e'_k\}$. Let $\text{var}(e'_k)$ be just like e'_k except that $\text{var}(e'_k)(k) = \neg(e'_k)(k) = \neg e(k)$. $\rho(\text{var}(e'_k), e'_k) = 1$ and $e'|k = \text{var}(e'_k)|k$ so iii.b of the induction hypothesis yields (v) $(r_e * [[e'|k]])(\text{var}(e'_k)) = 1$.

Case 1: $e'(k) \neq e(k)$. Then $\text{var}(e'_k)(k) = e'(k) \neq e'_k(k)$. So by iv, v, (vi) $\min\{(r * [[e'|k]])(e'') : e'' \in \text{dom}(r * [[e'|k]]) \cap [k, e'(k)]\} = 1$. Now let $e'' \in G^{\omega}_{\text{even}}(e)$.

Subcase: $e''(k) = e'(k)$. Then by the induction hypothesis and vi, $(r_e * [[e'|k+1]])(e'') = -1 + (r_e * [[e'|k]])(e'') \geq -1 + \rho(e_k, e'') = \rho(e_{k+1}, e'') = r_{e_{k+1}}(e'')$. When $e''|k+1 = e'|k+1$, the inequality just stated is strenghened to an equality by iii.b of the induction hypothesis, yielding $(r_e * [[e'|k]])(e'') = r_{e_{k+1}}(e'')$.

Subcase: $e''(k) \neq e'(k)$. Then by the induction hypothesis, $(r_e * [[e'|k+1]])(e'') = (r_e * [[e'|k]])(e'') + 1 \geq \rho(e_k, e'') + 1 = \rho(e_{k+1}, e'') = r_{e_{k+1}}(e'')$. Since $e''(k) \neq e'(k)$, $e'|k+1 \neq e''|k+1$ so iii.b holds trivially in this subcase.

Case 2: $e'(k) = e(k)$. Then (vii) $e'_k = e'_{k+1}$. Hence, $e'(k) = e'_k(k)$. So by iv, (viii) $\min\{(r * [[e'|k]])(e'') : e'' \in \mathrm{dom}(r * [[e'|k]]) \cap [k, e'(k)]\} = 0$. Now let $e'' \in G^{\omega}_{\text{even}}(e)$.

Subcase: $e''(k) = e'(k)$. Then by the induction hypothesis and vii, viii, $(r_e * [[e'|k+1]])(e'') = -0 + (r_e * [[e'|k]])(e'') \geq \rho(e_k, e'') = \rho(e_{k+1}, e'') = r_{e_{k+1}}(e'')$. When $e''|k+1 = e'|k+1$, the inequality is strenghened to an equality by iii.b of the induction hypothesis, yielding $(r_e * [[e'|k]])(e'') = r_{e_{k+1}}(e'')$.

Subcase: $e''(k) \neq e'(k)$. Then by the induction hypothesis and vii, $(r_e * [[e'|k+1]])(e'') = (r_e * [[e'|k]])(e'') + 1 \geq \rho(e_k, e'') = \rho(e_{k+1}, e'') = r_{e_{k+1}}(e'')$. Since $e''(k) \neq e'(k)$, $e'|k+1 \neq e''|k+1$ so iii.b holds trivially in this subcase.

Case: $* = *_{J,1}$. The argument is similar to the preceding one, except that the symmetry condition iii.a,b can be strengthened to: (iii) for each $e' \in G^{\omega}_{\text{even}}(e)$, $(r_e * [[e'|k]]) = r_{e'_k}$, which implies the success of $(r, *)$ as before.[33] Claim iii is immediate when $k = 0$.

In case 1, the induction hypothesis yields vi as well as (vi′) $\min\{(r * [[e'|k]])(e'') : e'' \in \mathrm{dom}(r * [[e'|k]]) \cap B - [k, e'(k)]\} = 0$. Subcase $e''(k) = e'(k)$, is as before, with an equality replacing the inequality. In subcase $e''(k) \neq e'(k)$, vi, vi′, yield: $(r_e * [[e'|k+1]])(e'') = -0 + (r_e * [[e'|k]])(e'') + 1 = \rho(e_k, e'') + 1 = \rho(e_{k+1}, e'') = r_{e_{k+1}}(e'')$.

In case 2, the induction hypothesis yields viii as well as (viii′) $\min\{(r * [[e'|k]])(e'') : e'' \in \mathrm{dom}(r * [[e'|k]]) \cap B - [k, e'(k)]\} = 1$. Subcase $e''(k) = e'(k)$ is as before, with an equality replacing the inequality. In subcase $e''(k) \neq e'(k)$, (viii, viii′) yield: $(r_e * [[e'|k+1]])(e'') = -1 + (r_e * [[e'|k]])(e'') + 1 = \rho(e_k, e'') = \rho(e_{k+1}, e'') = r_{e_{k+1}}(e'')$.

[33] Condition iii implies the hypercube rotation representation of the evolution of $(*, r)$, as depicted in fig. 3.

Proof of (2). Let $* = *_M$. Suppose for reductio that there is an IA r such that $(r, *)$ identifies $G^1_{\text{even}}(e_0)$. Then there exists a least n such that (i) $(r, *)(e_0|n) = \{e_0\}$. Furthermore, (ii) $\exists i \geq n, (r * [[e_0|n]])((e_0 \ddagger i) \ddagger i+2) \leq (r * [[e_0|n]])((e_0 \ddagger i+1) \ddagger i+3)$, else, there would exist an infinite descending chain of ordinals in the range of $(r * [[e_0|n]])$. Let $e = (e_0 \ddagger i) \ddagger i+2$, $e' = (e_0 \ddagger i+1) \ddagger i+3$, and $e'' = (e_0 \ddagger i) \ddagger i+3$.

Case 1: $(r * [[e_0|n]])(e'') > (r * [[e_0|n]])(e')$. Then by i, ii, e'' is not propped up in $(r * [[e_0|n]])$, contradicting the reductio hypothesis by proposition 21.

Case 2: $(r*[[e_0|n]])(e'') \leq (r*[[e_0|n]])(e')$. $e_0|i = e'|i = e''|i$, so by timidity and positive order- invariance (proposition 7), (iii) $(r*[[e'|i]](e_0) = 0 < (r*[[e'|i]])(e'') \leq (r*[[e'|i]])(e')$. So by timidity, stubbornness, iii and the fact that $e'|i+1 = e_0|i+1$, we have: (iv) $b((r * [[e'|i+1]])(.|[i, e'(i)])) = \{e_0\}$. So $e', e'' \notin b(r * [[e'|i+1]])(.|[i, e'(i)])$. So by the definition of $*_M$ and iii, $(r * [[e'|i+1]])(e'') = (r * [[e'|i]])(e'') + 1 \leq (r * [[e'|i]])(e') + 1 = (r * [[e'|i+1]])(e')$. So since $i+1|e' = i+1|e''$, positive order-invariance yields that for all $k \geq i+1, (r * [[e'|k]])(e'') \leq (r * [[e'|k]])(e')$, contradicting the reductio hypothesis.

Proof of (3). Immdediate.

Proof of (4). Immediate consequence of propositions 9 and 11.

Proof of (5). The argument is identical to the one provided for proposition 29 except that the appeal to proposition 24 is replaced with an appeal to proposition 26. $\dashv$