DIVISION OF COMPUTER SCIENCE

Temporal Logic and "Reverse Semantics"

J Broido

Technical Report No.233

October 1995

Broido, J , *Temporal Logic and "Reverse Semantics"* Abstract

This paper examines the relationships between certain types of linear frames and their alleged "temporal" Logics, S4.3 and S4.3.1, and shows how abnormal these relations are in comparison with common Modal Logics and the frame-types normally associated with them. Inasmuch as one refuses to regard as temporal structures any but one directional linear frames, the paper will disqualify such or other propositional logics from being strictly "temporal".

The paper defines plausible senses in which a Modal Logic can be said to *capture* a pre-selected subset of its models. Beyond completeness, this requires that every *nominalistically acceptable* model be "composed" of models of the chosen type (such modes of composition being termed *intensional unions*). While the Logics T,S4,S5 and B capture the frames traditionally used to characterise them, and while S4.3 captures weakly-linear frames (but not the linear!), S4.3.1 fails to capture any frame type. Specifically, it captures neither its discrete-linear models, nor the set of ω-frames—albeit semantically characteristic. On the other hand, S4.3.1 captures those *entropy-maximisers* [end-extensions of open-ended linear structures, with an appended S5-cluster], in which every *cut* is either *quasi-deterministic* or *modally continuous*. Quasi-Determinism means that any formula with a constant limiting value in shrinking "immediate futures" of the cut, must be so with respect to "recent pasts"; whereas Modal Continuity means—for a cut with an *open past*—that a formula is uniformly possible in every recent past only if it so in some immediate future. It is shown that such models can be elementarily embedded in *cofinal, modally continuous, order complete* extensions.

The results are significant independently of how one explicates the idea of *a Logic "capturing" a subset of its models.* In particular, no alternative explication can salvage the alleged status of S4.3.1(or any other propositional alternative) as a Logic of an open-ended, discrete-linear (or, at least, of ω-framed) temporality, since it is proved (Theorem 6.2.I) that *every model conforming to this characterisation is elementarily embeddable in a non-trivial entropy-maximiser,* which in general is neither linear nor discrete. This, together with the fact that some entropy maximising models of S4.3.1 have no elementarily equivalent linear submodel, demonstrates why the mere completeness of a logic vis à vis a models with a certain type of frame does not automatically make it into *the Logic of that frame-type.* In particular, the above results indicate why there can be no true propositional logic of one directional linear time!

The attempt to retain S4.3 and S4.3.1 as full-fledged temporal logics may connect better, according to this paper, with insights into the general—not quite linear—nature of physical time, that are suggested by considering the causal evolution of finitely quantised scalar fields—themselves treated as *cellular automata* operating in finite-dimensional state-spaces. The possible *basins of attraction* of such automata must always be intensional unions of Entropy Maximisers.

TEMPORAL LOGICS AND "REVERSE SEMANTICS"

Jonathan Broido

0. Introduction.

0.1. The Idiosyncrasy of Temporal (and Spatio-Temporal) Logics. The inspiration behind the development of Temporal Logic in the last 40 years sets it apart from most other Modal and Intensional Logics. S4.3, for example, which both Kripke and Hintikka came to regard as the primal logic of linear temporality, had innocently enough originated in the attempt to codify intuitions about Diodoran modalities (e.g.,*p is and will always be true*); yet from the outset the codification was constrained or even dictated by pre-conceptions about the *structure of temporality itself.* This applies to most of the work done in temporal and spatio-temporal logic, which reversed the usual direction for Intensional Logics in general: *Rather than seek a Semantics to elegantly characterise a given Logic, one looked around for a Logical Theory that would admit of a specific, preconceived, Semantic Infrastructure.*

Leading logicians[1] in the late 50's and the early 60's were thus quite happy to manufacture many different temporal and tense logics, as long as their semantics could be framed in an ordering of one kind or another of the "possible" temporal states, and such a preoccupation with particular underlying *topologies* was quite overpowering. Even the interest in Diodoran modalities was quickly replaced by an interest in ω-frames and other "discrete" temporal structures. A propositional logic S4.3.1 (Prior's D) of such"discrete" temporality was believed to have been realised by adding to S4.3 a single schema (such as **AD:**$L(L(\alpha \supset L\alpha) \supset \alpha) \supset L(ML\alpha \supset \alpha)$ in §4.0 below). When its discrete linear models were shown by Kripke (and algebraically, in [Bull, 65]) to provide it with completeness, it was concluded—*rather rashly, as we shall argue* —that this must indeed be the sought for "Diodoran" logic of discrete linear temporality. Similar completeness results seemed to have sealed the "Logical Capture" of other temporal and spatio-temporal structures.

Despite all this, it is our contention that some of the most famous temporal propositional logics can not be viewed as being specifically about the kinds of temporality or temporal structures that they are alleged to capture. We claim that what they do succeed in capturing is at times quite different, and even unspecifiable in pure topological terms. Yet, since an allusion to a "Logic of (such and such) temporality" presupposes, or at least suggests, that one has been successful in specifically "capturing" the temporal structure invoked, our claims here will have also challenged the underlying dogma that ordinary semantic characterisation (consistency & completeness) automatically warrants such a "capture" in the opposite direction—a capture of the characterising models by the characterised Logic.

0.2. **"Reverse Semantical Capture".**[2] To show that a Logical Theory specifically *captures* a certain favourite class of its models (which we might then choose to regard as *paradigmatic*), it is clearly not enough to prove completeness with respect to this class. For there might exist *other legitimate models* of the theory, not in the favourite class, which, *by this very completeness , are ineliminable* by any added schemata (that is, without also eliminating favourite models). Therefore, in claiming that a given Logic specifically captures a certain kind of "favoured" structure, one must somehow show, *in addition to completeness with respect to such structures, that all its other models are essentially reducible, by some standard decompositional means, to the preferred models.* Furthermore, such a reduction must be reasonably strong, in the sense that it retains enough information about the original reduced models.

Talking of *The Logic of ...(such and such structures)*, and thereby alluding only to a proper subset of all the conceivable models, therefore commits us—besides completeness— to a strong reductive relationship between all legitimate models and our preferred subset. When this is not available, we can not claim any priviledged status for our favourite models. This, indeed, is what happens with the Logic S4.3.1: The relationship between all its models and the subset of the discrete linear ones is so tenuous that we can no longer say that it captures this latter subset. This is no condemnation of S4.3.1 itself; we should only take greater care in choosing its "paradigmatic models", or what we want to regard it as a logic "of" !

In this respect things deteriorate in moving from "Classical Modal" to "Temporal" logics. S4, S5, T, and B all capture ideally their well known standard *accessibility frames* : every model outside the appointed frames can be rejected *a-priori* on strong conceptual grounds (e.g., because internally indistinguishable worlds have different accessibility ranges). However, the choice of models standardly associated with the logic at hand is misleading for S4.3 ("Logic of Linear Temporality"), and is totally ill-begotten for S4.3.1("The Logic of Discrete Linear Temporality"). S4.3 fails to capture linear frames, although it does capture *weakly linear* ones: every *strongly normal* model (where inaccessibility is reflected in the difference in values of some modal formula) is in a strong sense *composed* (by means of what we term a *strong intensional union*) of *weakly linear* structures. By contrast there exist models of S4.3.1 that *can not be composed*—even in a much looser sense—of discrete linear structures.

The discrete linear models of S4.3.1, and in particular ω-framed structures, constitute only the odd lot out of a class otherwise made of entirely different kinds of structures, which are *neither discrete nor linear!* The models in this paradigmatic class have to do with the semantic properties of *Quasi-Determinism* and *Modal Continuity* of the *cuts* in their maximal linear part—of which they may be a (non-linear) *end-extension.* Quasi-Determinism means that any formula *indeterminate* in "arbitrarily small pasts" (*that is,* in any *recent past*) of the *cut* will remain so in "arbitrarily small futures" (*that is,* in *immediate futures*) of the same cut; whereas Modal Continuity—relevant only to cuts with non-terminating past—means that any formula, possible in every recent past of a cut, will remain so in some immediate future. Furthermore, all these paradigmatic models can be elementarily embedded in *order complete* models which are *modally continuous in their non-discrete cuts* (THEOREM **5.3.5**).

The paradigmatic models of S4.3.1 use two different types of frames: (**A**) *Linear structures without end* ; and (**B**) *Strucutres in which a linear part is followed by a single final mutual accessibilty cluster (of cardinality≥1)*. A structure of either type will become an **S4.3.1** model if and only if *every cut in it is quasi-deterministic or modally continuous.* Either type suffices, by itself, for completeness, and neither suggests the other on pure *topological* grounds. The differences between the above frame-types are therefore "invisible" to first order intensional logic. This blindness is nonetheless significant from a logical point of view, and is highlighted by the fact (THEOREM **6.2.I**) that *every* **A** *type structure can be elementarily embedded in an end-extension structure of type* **B**—a fact correlated with the syntactic meta-theorem $S5 \vdash \alpha$ iff $S4 \vdash ML\alpha$. On the other hand, a type **B** model is not always elementarily equivalent to some linear submodel. Ironically, then, it is shown that S4.3.1—the alleged logic of discrete temporality—can be characterised altogether without its discrete linear models, and even more informatively at that! Furthermore, the paradigmatic non-linear B-type models may begat a new "temporal" significance in the context of causal-deterministic structures that explain the *"eventual" destruction* of linear temporality! (§7.2 below).

0.3. **Styles of Semantics and Reverse Semantic Reflection.** In the case of Modal and Intensional logics, in general, one needs to avoid the creation of unfavoured or philosophically spurious models merely as an accident of *wasteful semantic styles*—where the relation between framing and basic semantic valuation is too loose! It is trivial to create some S4 models that are not transitive, or S4.3 models that branch "toward the future"—simply by replicating the same valuations at certain nominally different positions of the accessibility-frame. The usual response to this problem consists, again, in nothing more than showing that we do not need such wasteful models in order to achieve completeness. Philosophically, however, we believe that a more rigorous response should trimline *a-priori* the notion of a model and demand that *framing-relations between possible worlds should be reflected in the relations between and the properties of the valuations of formulae within these worlds.*

Making the convenient separation between frames and models, and regarding the "reverse semantical" direction as one which attempts to elicit a logic from a class of structures, the complete enterprise of "reverse semantics" can be represented by the diagram :
frames- --->***semantic structures*** (candidate-models)---->*logic* . This seems to us to require that the framing assumptions should be reflected in the added semantic component of the structures—i.e., in their valuation schemes—while the structures themselves are reflected in the logic, in the sense that they characterise it while allowing us to "reconstruct" even those models whose structure does not fit the chosen frames.

"Reverse semantical" reflection and a somewhat more nominalistic style of semantics therefore go hand in hand. This can be expressed in various "reverse reflection" principles which spell out how different aspects of the framing infra-structure are to be explicated by features of the valuation schemes. The following are few examples, used in the present paper [where W_M, R_M, V_M stand for the set of possible *cases ("worlds")*, the *"accessibility"* relation, and the *valuation-function*, respectively, in a "model" M]:

(i) ***Model-Normality*** Any two *semantically equivalent* cases, w, w' [i.e., where $(\forall\alpha)(V_M(\alpha,w)=V_M(\alpha,w'))$] are interchangeable in any accessibility context. (*This allows a projection into semantical-equivalence-classes without loss of information*).

(ii) ***Strong-Model-Normality*** .[3] If in a model M, all the possible cases in $W_1 \subseteq W_M$ are accessible to w_0, whereas w_1 is not, then there must be a formula true everywhere in W_1, but false in w_1. This principle is equivalent to

$<w_0,w_1> \notin R_M \rightarrow \exists\beta(V_M(L\beta,w_0)=1 \,\&\, V_M(\beta,w_1)=0)$ which, for Logics containing S4, is equivalent to $<w_0,w_1> \in R_M \leftrightarrow \forall\beta(V_M(L\beta,w_0)=1 \rightarrow V_M(L\beta,w_1)=1)$

(iii) ***Modal Continuity in Linear Models*** .[4] If R_M is a Linear ordering of W_M, then the supremum (infimum) of any subset $W_1 \subseteq W_M$ —if it exists— should retain any Possibility (Necessity) true everywhere in W_1. This is equivalent to

If, in a linear Model, w *is the first (last) case in which* $L\beta$ ($M\beta$) *is true, then* w *can not be a right-(left-) accumulation point.*

The results of this paper show, however, that even a sufficiently nominalistic style of semantics—in which the model's framing features are reflected in valuation schemes and relations between them— does not solve by itself all the problems of unwanted models. While for some Modal Logics mere adherence to the above reflection principles ensures that *only desirable models are left*, for other Logics—such as S4.3—it can only guarantee that a (nearly) desirable type of models is "captured", in the sense indicated above and made precise below (§1.2). In still other cases a logic—such as S4.3.1—may *fail* to capture a prescribed class of models (e.g., based on finite linear and ω-frames) that characterises it. Thus the Temporal Logics S4.3 and S4.3.1 illustrate, respectively, a non-trivial partial success and a complete failure of a meaningful reverse semantical enterprise.

1. Intensional Unions and the Notion of Capturing.

The success in capturing (by a theory) of certain privileged Models should depend on the amount of "information" about any acceptable model, in general, that can be retrieved, by means of an agreed mode of "composition", from the knowledge of these privileged models. Such compositional modes are entitled here *Intensional unions* and are defined non-constructively below. We will have *strong, regular* and *weak intensional unions,* corresponding to diminishing degrees of "information retrievability".

1.1. **Basic Definitions.** We normally understand the term *sub-model* in the loosest sense: M' is a sub-model of a given modal-model $M=\langle W_M, R_M, V_M\rangle$, iff $W_{M'} \subseteq W_M$, $R_{M'} \subseteq R_M$, and V_M agrees with $V_{M'}$ on $VAR \times W_{M'}$, where VAR is the set of propositional variables.
Let R be any binary relation on W (i.e, $R \subseteq W \times W$). $R''S =_{df} \{y \mid (S \times \{y\}) \cap R \neq \emptyset\}$.
For $w \in W$ define

$[w]_0^R = \{w\}$; $[w]_{n+1}^R = [w]_n^R \cup R''[w]_n^R$; $[w]^R = \bigcup\{[w]_n^R \mid n<\omega\}$=the R-closure of $\{w\}$.

Given two such binary relations R, R' $\subseteq W \times W$, say that R' is *exact for* R *at* $w \in W$ iff for every $n<\omega$, $[w]_n^{R'} = [w]_n^R$.

A model M' *mimics* M *at* $w \in W_{M'} \cap W_M$ iff $V_M(\alpha,w) = V_{M'}(\alpha,w)$, for every α.
Notice that
Lemma 1.1.1. *Any submodel* M' *of* M, *whose accessibility-relation* R' *is exact for* R_M *at every* $w' \in [w]^{R'} = [w]^{R_M}$ *must mimic* M *at any such world*. (Proof by Induction on Modal depth).

A sub-model M' of M is a *casewise elementary submodel* iff it mimics M in every world of its own.
For any $w \in W_M$ define $M|_w$ as the sub-model obtained by the restriction of the possible cases to $[w]^{R_M}$. (so called *the submodel generated by* w). The following are elementary consequences:
Lemma1.1.2. *The accessibility relation of* $M|_w$, $(R_M \cap [w]^R \times [w]^R)$, *is exact for* R_M *anywhere in* $[w]^R$;
Lemma 1.1.3. $M|_w$ *is a casewise elementary submodel of* M. (This is Theorem 5.6 in [Hughes and Cresswell, 84]: *A Companion To Modal Logic*).
$M|_w$ also inherits from M any nice reverse semantical reflection features (such as strong normality or modal continuity).

1.2. **Intensional Unions.** M is a *regular Intensional Union* of a set S of Modal models iff

(1) Every member of S is a sub-model of M; (2) $W_M = \bigcup\{W_{M'} \mid M' \in S\}$; and

(3) For every $w \in W_M$, $(\exists M' \in S)(R_{M'}$ *is exact for* R_M *at* w). [Note: (3) entails $R_M = \bigcup\{R_{M'} \mid M' \in S\}$]

M is a *strong Intensional Union* of a set S of Modal models iff, in addition to being a regular intensional union of S, every member of S is a case-wise elementary submodel of M.
Note: Any model M is such a union of its generated submodels $\{M|_w \mid w \in M\}$. (Lemma 2.1.4 below).

M is a *weak Intensional Union* of a set S of Modal models iff
(1) Every member of S is a submodel of M; and
(2) $(\forall w \in W_M)(\exists M' \in S)$(M' mimics M at w).

1.3. **Capturing.** We are ready now to define the concept of "capturing":

A Logic L *captures (regularly, strongly, weakly)* a class C of semantic structures iff
(1) C is an ordinary *semantic characteristic* of L (i.e., L is consistent and complete with respect to C); and
(2) Every strongly normal model of L is an intensional union (regular, strong, weak) of a subset of C.

For the sake of terminological completeness we might add to these "capturing" terms a special term for the case where every strongly normal model of the Logic L is itself in a characteristic class C. In such a case we can say that L *captures ideally* C.

Say now that a Logic captures (in any of the above senses) the frames defined by a property of the accessibility relation, whenever it captures (in this sense) the class of all modal semantic structures whose accessibility satisfies this property.

It is easy to prove
THEOREM 1. I. *The common modal Logics* T,S4,S5,and B, *capture ideally the frames characterised by the following properties, respectively: (1) reflexivity; (2) reflexivity+ transitivity, (3) reflexivity+transitivity+symmetry ; and (4) reflexivity+symmetry.*
The proof is left to the reader[5].

On the other hand, as is showed below (theorem 2.I), S4.3 captures *strongly,* but *not ideally,* the frames characterised by reflexivity+transitivity+connectivity (weak-linearity). S4.3.1 fails to capture *in any sense* any set of structures that can be defined solely by a property of their frame. To the extent that we can talk of it as "capturing" any type of model, this will involve an intertwining of frame and valuation properties.

1.4. **Definitions for Frames that are Reflexive, Transitive and Connective**. Let '$R^{\smile}$' denote the converse of R, so that for a model M, connectivity can be represented by
$R_M \cup R_M^{\smile} = W_M X W_M$, and weak anti-symmetry by $R_M \cap R_M^{\smile}$ =the identity on $W_M X W_M$.
For R which is reflexive, transitive and connective, define $<_R$ as $R - R^{\smile}$. If R is weakly anti-symmetric this must be a proper linear ordering. For any model M, a possible case $w \in W_M$ is a *right accumulation point* (case) iff, in some maximally linear subset of W_M (relative to $<_{R_M}$) $(\forall y)[y <_{R_M} w \rightarrow (\exists z)(z <_{R_M} w \;\&\; y <_{R_M} z)]$. Likewise, a *left accumulation point* w will be defined by the condition $(\forall y)[w <_{R_M} y \rightarrow (\exists z)(w <_{R_M} z \;\&\; z <_{R_M} y)]$.

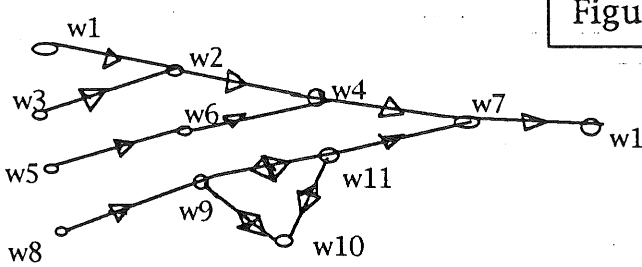
In the context of Modal models and subsets of their cases, we shall use the term *linearity* only with respect to $<_{R_M}$ .We shall call w an *accumulation point* when it is either a right, or a left, accumulation point (or both). R_M is *discrete* when there are no such points, and if $<_{R_M}$ is a linear ordering, in addition, and we can talk of M as a *linear discrete model.* [In some sources—in particular in connection with the logic S4.3.1—*discreteness* seems to have been used as merely *precluding left accumulation points*[6]—which should be better described as *discreteness-on-the-right* . This, however, will not affect our contention that S4.3.1 is not about discrete linear models. It is equally true that it is not about right-discrete linear models (or about left-discrete models).]

A model, M, will be called *boundedly compact* when every linearly ordered subset of W_M , with an upper(lower) bound in W_M, has a supremum (infimum) in W_M. This property is also known as *order completeness*.[7] Any well ordered model is obviously boundedly compact, but there are many boundedly compact models that are not well-ordered. Thus both $\omega+\omega$ and $\omega^{-}+\omega$ frames are boundedly compact but the latter is not well ordered. On the other hand an $\omega+\omega^{-}$ frame is not boundedly compact but is fully discrete! The importance of these definitions and observations to S4.3.1 stems from particular and general results established later (sections 4. and 5). For example: One can use an $\omega+\omega^{-}$ frame the construct a particular model refuting S4.3.1, although such frames are*fully discrete* . Any modally continuous boundedly compact model validates S4.3.1, and any S4.3.1 model can be elementarily embedded in such a model. Bounded compactness (or order completeness) is therefore much more relevant to this logic than discreteness.

2. Models of S4.3.

2.0. S4.3 can be represented as $S4+\{L(L\alpha\supset\beta)\vee L(L\beta\supset\alpha)\}$[8].

Figure 2.A



A "river confluence" frame (the arrow stands for a transitive and reflexive accessibility. w9,w10 and w11 form a mutual accessibility "cluster")

Any "multiple river confluence" (with permitted "river loops"), as in 2.A to the left, fails connectivity, and is therefore neither weakly nor properly linear. But such frames are perfectly legitimate as S4.3 frames, and can serve as an infrastructure for strongly normal models.

The significance of such non-standard models is usually dismissed because all their generated subframes are of the standard weakly linear type, and because for ordinary completeness one can take as an alternative characteristic class the set of all generated submodels of members of a characteristic class[9]. This approach works, from a reverse semantical point of view, only when those generated submodels are of the the type that is desirable to us—as would be the case with respect to S4.3 and weakly linear models as the chosen standard.

2.1. We want to start our investigation by reproducing, from a *reverse semantical* point of view, a well-known characterisation theorem for S4.3 . In particular we want to prove

THEOREM 2.I. *Every strongly normal model of* S4.3 *is a strong intensional union of weakly linear models—that is, models that are reflexive, transitive and connective.*

As a preliminary we need only the following definition and easily proved lemmas:

Definition 2.1.2. A model has *the lower bound property* iff for any its two possible cases there is a possible case to which they are both accessible.
Lemma 2.1.3. *If* M *is transitive and reflexive then, for any* $w \in W_M$, $M|_w$ *has the lower bound property;*
Lemma 2.1.4. M *is a strong intensional union of its generated submodels* $\{ M|_w \mid w \in W_M \}$;
Lemma 2.1.5. *If* M *is strongly normal so is* $M|_w$ *, for any* $w \in W_M$.

Combining the above with the fact (theorem 1.I) that any strongly normal model of S4 is reflexive and transitive, and with lemma 1.13 ($M|_w$ mimics M in all its worlds), we get
Corollary 2.1.6. If M is a strongly normal model of S4.3 (or any extension S_α of S4) then it is a strong intensional union of strongly normal models of S4.3 (S_α) which are reflexive, transitive and have the lower bound property, and which are all casewise elementary submodels.

To prove theorem 2.I then it is enough to prove the following

Lemma 2.1.7. *If* M *is a strongly normal model of* S4.3 *with the lower bound property then it is connective* ($R_M \cup \breve{R}_M = W_M X W_M$).
Proof. Let w_1 and w_2 be any two possible cases of M. Assume that neither of them is accessible to the other. Let w be a common lower bound, so that in this case $w \neq w_1$ and $w \neq w_2$, by assumption. Since $<w_1,w_2> \notin R_M$ we have by strong normality that $\exists\gamma(V_M(L\gamma,w_1)=1$ & $V_M(\gamma,w_2)=0)$, and likewise, since $<w_2,w_1> \notin R_M$, that $\exists\delta(V_M(L\delta,w_2)=1$ & $V_M(\delta,w_1)=0)$. It follows that $L\gamma \supset \delta$ is false in w_1 while $L\delta \supset \gamma$ is false in w_2. Since both w_1 and w_2 are accessible to w, both $L(L\gamma \supset \delta)$ and $L(L\delta \supset \gamma)$ are false in w and so is $L(L\gamma \supset \delta) \vee L(L\delta \supset \gamma)$, contradicting the validity of the S4.3 schema in M ■
The results of this section can be summarised then by the following statement:

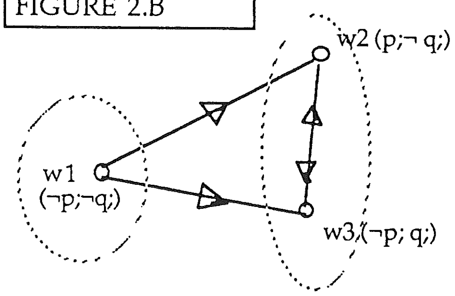
Corollary 2.1.8. *Weakly linear models are strongly captured by* S4.3 .

2.2. What about properly linear frames? It is well known that S4.3 is semantically characterised by linear frames, but this does not mean that S4.3 is "specifically about them", unless we can prove that every strongly normal weakly linear model is an intensional union (at least a weak one) of legitimate linear models .

Unfortunately there are some very simple counter-examples. A weakly linear model which is not properly linear must have some mutual accessibility equivalence class containing more than one member. Now, it is easy to construct such a model in which every world w *satisfies uniquely* *some classical formula* ϕ_w, as is the case in the simple three world model $\mathcal{M}$ depicted in figure 2.B below.

w_2 and w_3 are mutually accessible while w_1 is an "earlier" case to which they are both accessible (but not vice versa). Given that classical formulae maintain their value at u in submodels containing u, and that $M\phi_u$ has to be kept true in any world up to u but made false later, any linear submodel of $\mathcal{M}$ mimicking $\mathcal{M}$ at w_1 must arrange in a definite linear order all the worlds of $\mathcal{M}$ but must retain the original order between mutual accessibility equivalence classes! Thus we have only $<w_1<w_2<w_3>$ or $<w_1<w_3<w_2>$ as possible frames for such a linear submodel mimicking the original model at w_1. Yet neither will do , since the first frame makes MLq true at w_1 while the second makes MLp true at w_1—neither of which are true at w_1 in $\mathcal{M}$.

FIGURE 2.B



A weakly linear model with a final non-trivial "cluster"

(arrow stands for transitive& reflexive accessibility)

This model is what we term later an 'Entropy Maximiser'. The proof above establishes, then, that some such models can not be decomposed into properly linear components without losing too much valuable information—such as LMp&LM¬p as a valid formula.

Contrasting this example with Segerberg's "bulldozing technique", which can be used for the transition from weakly linear to linear frames, when one wants to prove the completeness of S4.3 vis à vis linear models[10], we can now guage better the philosophical damage incurred by going for completeness at any price. The "bulldosing" of the only"cluster" (non-trivial mutual accessibility equivalence class)in $\mathcal{M}$ above requires $\aleph_0$ copies of $w_2<w_3$ or of $w_3<w_2$—leading to a redundant and anti-nominalistic model. The philosophical irony of such situations is that *the additional infra-structural simplicity desired can be achieved only by losing the ability of our models to reflect in their semantical make-up the very kind of frame desired!* Linearity has to be reflected, in this sense, in *new necessary truths emerging with each true successor moment,* but this is lost if we match S4.3 with properly linear frames. Weak linearity then must be the natural frame-correlate here. In the case of S4.3.1, we show later, we shall describe a way of getting rid of all "clusters" except the last which—as far as "capturing" goes—is sometimes ineliminable.

3. Digression: Simplified completeness of S4.3 and other extensions of S4.

The intensional decomposition into desired models of an extension of S4 can be sometimes used to obtain a completeness proof vis à vis such models. Thus we have

THEOREM 3.0. *Let* S_α *be any consistent extension of* S4, *closed under* S4-*deduction, which satisfies the following conditions, with respect to a class,* C, *of structures:*

(1) *Every strongly normal model of* S_α *is a (strong, weak, regular) intensional union of structures in* C.
(2) *Every member of* C *is an* S_α*-model (consistency).*

Then S_α *is also* **complete** *with respect to* C, *and therefore captures its structures in the corresponding sense (strong, weak, regular).*

To prove this theorem, we need to recapitulate a certain strong completeness result for S4:

THEOREM 3.I. *There exists a strongly normal "canonical" model of* S4, $M^U = \langle W_U, R_U, V_U \rangle$, *in which every* S4- *consistent set of formulae is satisfied in some possible case.*

Foregoing the well known type of proof, we only describe the make-up of such a model. Define the set W_U of possible cases (worlds) as the set of all maximally consistent sets (of Wffs) with respect to S4-provability. Define R_U by $\langle S, S_1 \rangle \in R_U \leftrightarrow (\forall \beta)$ ['$L\beta$'$\in S \rightarrow$'$L\beta$'$\in S_1$], and, finally, for any propositional variable υ, define $V_U(\upsilon, S)$ as 1 or 0 according to whether $\upsilon \in S$ or not, respectively. Prove, first, by induction,

Lemma 3.1.1. For any Wff α, $V_U(\alpha, S)=1$ iff $\alpha \in S$, and then, trivially

Lemma 3.1.2. M^U is strongly normal.

Then one proceeds to

Proof of Theorem 3.0. Let S^α be the set {'$L\theta$' | θ is a theorem of S_α}. The deductive closure of S^α is the same as that of S_α. Let S^* be *any* maximally consistent set containing S^α. This defines a case in the universal S4 model M^U. Let $M^U|_{S^*}$ be the submodel of M^U, defined as in §1.1, by taking the R_U-closure of {S^*}. By the S4 axioms, all its cases satisfy S^α, and it is therefore a model of S_α. If Γ is any set of formulae consistent with respect to S_α, then it is consistent with respect to S^α. Choose S^* ,then, as containing $S^\alpha \cup \Gamma$ from the outset. $M^U|_{S^*}$ becomes then, by lemmas 3.1.2 and 2.1.4, a strongly normal model of S_α satisfying Γ at its "root"-case.

Given such a strongly normal model, $\mathcal{M}$, of S_α satisfying Γ at one of its cases, w, we have, by condition (1) of theorem 3.0, that it is an intensional union of models in C (which, by condition (2) are all models of S_α). One of these component C-models must mimic $\mathcal{M}$ at w (true of any union concept!) and therefore must satisfy Γ there. ■

We see then that for an extension of S4 completeness vis à vis a desired type of model is guaranteed by the mere "intensional" decomposability of any proper model of such an extension. In particular it follows from theorem 2.I that S4.3 is complete with respect to weakly linear structures, and that we did not have to cite earlier proofs of this completeness result in order to obtain our "capturing" corollary 2.1.8.

On the other hand, theorem 3.0 entails that incompleteness vis à vis a class of models means also that such a class is insufficient to compose, even by weak intensional union, every strongly normal model.

4. Prior's D (S4.3.1) and Discrete Models, or What S4.3.1 is not about.

4.0. The Logic S4.3.1 can be represented as S4.3+{$L(L(\alpha \supset L\alpha) \supset \alpha) \supset L(ML\alpha \supset \alpha)$}. To refute in a linear frame the schema added to S4.3 (AD) is tantamount to finding for some formula α two possible cases w_0 and w_1, where $w_0 <_R w_1$, such that α, is false at w_0, is necessarily true at w_1, but changes its truth value twice more (and therefore infinitely often) after any case $\geq w_0$ at which it is false.

It is easy to construct a perfectly discrete model $\mathcal{M}^*$ which falsifies S4.3.1. Consider the frame provided by the set $\{-1/n \mid 1 \leq n < \omega\} \cup \{1/n \mid 1 \leq n < \omega\}$ together with the usual real order. Define the valuation V by (i) $V(p_1, w_r)=1$ if and only if $r>0$ or $|1/r|$ is even; and (ii) for $i>1$, p_i is true($V=1$) at $w_{(-1)^i/[i/2]}$ and nowhere else ['[x]' denotes the integral part of x]. Thus p_1 will be necessarily true at any positively indexed world, will be false at w_{-1}, but will alternate infinitely many times for the negatively indexed worlds following the latter. The frame is of order type $\omega+\omega^-$ which has no accumulation point of any sort and is therefore perfectly discrete.

In well ordered frames we have *discreteness on the right*, which some people considered (wrongly) the relevant frame property for S4.3.1. But well-orderedness beyond ω does not guarantee the validity of this logic either. It is trivial to construct an $\omega+1$ framed model refuting S4.3.1: we just have to see to it that some formula will change its truth value infinitely many times. Either it or its negation will be necessarily true at the last world(w_ω) and will therefore serve to refute **AD** at any naturally-indexed world in which it is false.

4.1. **S4.3.1 and Modal Continuity—Preliminary remarks.** It is easy to show that the only well-ordered S4.3.1 *frames* (that is, *independently of valuation*) are either finite or ω frames. Yet this is to some extent misleading. There are infinitely many strongly normal linear models of S4.3.1 of other order types(the non-linear models will be discussed later). In particular we have well ordered models of S4.3.1 for every denumerable ordinal(lemma 4.1.3 below).

Many of these models are *modally continuous* (see definition (iii), §0.3 above)—which satisfy for any α: *If it exists, the supremum (infimum) of a set of cases (worlds) where* α *is everywhere possibly(necessarily) true must retain this possibility (necessity).* The intimate connection between modal continuity and the Logic S4.3.1, for some order types, is revealed by pinpointing the S4.3.1-failure in such $\omega+1$ framed counter-models as described in 4.0 above. *Any* such $\omega+1$ framed counter-model to S4.3.1 must be modally discontinuous!

Although this suggests the correct insight that modal continuity is enough to ensure S4.3.1-validity for *well ordered* models, it does not reveal the exact root cause of S4.3.1-failure in other types of linear counter-models, such as the $\omega+\omega^-$ framed "discrete" model $\mathcal{M}^*$, described above in 4.0. In that case, for instance, there were no accumulation points but the model failed to be *boundedly compact (order complete)* . It was this that allowed a"discrete" linear counter model to S4.3.1.

Since every well ordered model is boundedly compact, we will now generalise the above insight about well ordered models and state

Lemma 4.1.1. *Every Modally continuous boundedly compact linear model validates* S4.3.1.
We leave this as an excercise to the reader

Corollary 4.1.2. *Every modally continuous well ordered model validates* S4.3.1
It is equally easy to prove

Lemma 4.1.3. *There exist well ordered strongly normal and modally continuous models of* S4.3.1 *corresponding to every denumerable ordinal* ν.

Proof. Each possible case will be indexed by a member of ν (a lesser ordinal). Let $W_D(\nu)$ be the subset of perfectly discrete cases (i.e, the set of non-limit members of ν). Let f be a 1-1 mapping of ω *on* this subset. Define a valuation V by $V(p_i,\mu)=1$ iff $f(i)\leq\mu$. Each propositional variable thus becomes true for the first time at some discrete non-limit ordinal and remains so thereafter. It is easy to prove, then, by induction on formula-structure, that for each formula there is a first ordinal$<\nu$ at which it becomes non-contingent (necesarily true or necessarily false) and that this ordinal must be a *non-limit* one (i.e, either the first or a successor to another ordinal). Given any α which is possibly true at all the ordinals below a given limit ordinal κ, this possibility can not terminate at κ—or else we would contradict the assertion above, producing a limit ordinal as the place where $M\alpha$ becomes false (forever) for the first time.

It is likewise easy to prove that the above model is strongly normal. For each $\gamma\in W_D(\nu)$, '$Lpf^{-1}(\gamma)$' will be true at or above γ, but will be false before γ. If λ is a limit ordinal$<\nu$, and μ is not accessible to λ then $\mu<\lambda$ and there exists a discrete non-limit ordinal δ between μ and λ and we can then use $Lpf^{-1}(\delta)$' as a formula true at or above λ but false at μ.■

4.2. **S4.3.1-models and Intensional Unions.** We will now prove

THEOREM 4.2.2. *The exists a modally continuous, strongly normal, well-ordered model of* S4.3.1 *which is not a weak intensional union of discrete linear submodels.*

Proof. Define a model M_2 as follows: The frame is of order type $\omega+\omega$ with a single accumulation point at w_ω. Assume that the first case is w_1. The valuation on the variables is defined by $V(p_{2i-1},w_\mu)=1$ iff $\mu\geq i$ $(\mu<\omega+\omega)$; and $V(p_{2i},w_\mu)=1$ iff $\mu\geq\omega+i$ $(\mu<\omega+\omega)$.
Now assume that this model is a weak intensional union of some discrete submodels. Verify first that each discrete case (w_i or $w_{\omega+i}$,where i is natural≥1) has a classical formula which is uniquely satisfied in it. For w_i ($w_{\omega+i}$) ,where i is natural≥1, it will be '$p_{2i-1}\&\neg p_{2i+1}$' ('$p_{2i}\&\neg p_{2i+2}$'). Thus, the formulae $\{M(p_i\&\neg p_{i+2}) \mid 1\leq i<\omega\}$ must all be all true at the case w_1. Since the weak intensional union must include a submodel N_1 mimicking M at w_1, such N_1 will have to contain, for each i, a case accessible to w_1, at which $p_i\&\neg p_{i+2}$ is true. Yet since classical formulae retain their values in any submodel, N_1 will have to contain *all* the discrete cases of M_2. Given that all components of the union are linear, N_1 will have to retain the entire M_2 order structure amongst the discrete cases. But a submodel of M_2 containing all its discrete cases and the order amongst them can not be discrete, since if it removes w_ω, it will have made $w_{\omega+1}$ into a new accumulation point. Contradiction. The Modal continuity and strong normality of M_2 are easy to establish■

Since strong and regular intensional unions are *a fortiori* weak ones, it follows that M_2 is not decomposable into discrete linear models in any of the intensional-union-modes suggested in this paper.

5. What is S4.3.1 about? Temporal Cuts from a Modal Perspective.

5.0 The previous section makes it hard to view S4.3.1 as a "propositional logical reflection" of discrete linearity. ω-framed models (which are necessarily discrete) constitute a traditional semantic characteristic, but do not give us a clue as to S4.3.1 models, in general. Even worse, they are misleading, since they can't even suggest those model-features which characterise many other well-ordered models of S4.3.1. We will therefore try to obtain a better characterisation of strongly normal models of S4.3.1.

5.1. **Cuts in Weakly Linear Frames.** Given a frame $<W,R>$, where R is a weakly linear relation, define a *cut* , as usual, as an ordered pair $<W_1,W_2>$, satisfying the conditions: (1)$W=W_1 \cup W_2$; (2)$W_1 \cap W_2=\emptyset$; and (3) $\emptyset \neq W_1 \times W_2 \subseteq R-\breve{R}$. We shall add $<\emptyset,W>$ and $<W,\emptyset>$ as two trivial cuts.

Let '$\equiv_R$' stand for mutual accessibility ($R \cap \breve{R}$). Given a cut $<W_1,W_2>$, we will refer to W_1 as its *total past* and to W_2 as its *total future*. Since our frames are assumed weakly linear, R must induce a proper linear ordering of the equivalence classes modulo $\equiv_R$.
Any cut also induces a classical Dedekind cut in the induced linear ordering, since by condition (3) above two members of the same equivalence class cannot be on opposite sides of the cut. The cuts themselves are linearly ordered by the relation of (strict) incluson between their total pasts.
A cut is called *discrete* when its total past has a *last* member and its total future has a first member. It is called *open (bilaterally)* when neither of these conditions is satisfied. If exactly one of these conditions holds we might call the cut *real*. Thus, all the cuts are discrete in the Integers, while in the Rationals no cut is discrete, but some cuts are open. In the Reals every cut is real. A linear frame is boundedly compact when no cut is bilaterally open —that is, when every cut is discrete or real. Notice that although discreteness of all non-trivial cuts implies discreteness of frame, the converse is not true[11] ! A weakly linear frame is an *Entropy-Maximiser* when it has a cut whose total past is properly linear while its total future is a single cluster of mutually accessible cases (a universal S5 frame)[12].

5.1.1. **Recent Pasts and Immediate Futures; Determinism in Cuts.** A *recent past* of a cut is any non-empty intersection of its total past with the total future of some other cut (below it), while an *immediate future* of a cut is analogously any non-empty intersection of its total future with the total past of some other cut (above it). The semantic character of cuts in a model depends on the limiting behaviour of formula valuation in such "neighbourhood" pasts and futures.

Given any formula, α, we will say that it *recently (immediately) necessary* in a given cut iff it is true everywhere in some recent (immediate) past (future) of this cut. In particular we will be interested in *recent* or *immediate determinacy* in a cut, which means that the formula at hand has a constant value in some recent past, or immediate future, respectively, in the given cut.
A cut is called *quasi-deterministic* iff every formula which is immediately determinate in it is also recently determinate. A cut is *stoically determinate* iff every formula is both immediately and recently determinate in it. In particular every discrete cut must be stoic. Finally, a cut is *strictly deterministic* iff every formula or its negation is both recently and immediately necessary in the cut. Models will be called quasi, or stoically deterministic when all their non-trivial cuts are such.

5.1.2. Synthetic and Analytic Clusters in Strongly Normal Models. Let the term *cluster* stand for an equivalence classe modulo $\equiv_R$ (mutual accessibility). When a cluster is a singleton it will also be regarded as *trivial.* It is clear that a weakly linear model is properly linear iff all clusters are trivial, and is an Entropy Maximiser iff all clusters but the last are trivial.

We refer to a cluster as *analytic* iff every formula attains only one value within it. Otherwise it is to be called *synthetic.* A non-trivial analytic cluster obviously exhibits total redundancy, since it is semantically representable by any single case (world) therein; yet such multiplicity *solo numero* is not ruled out by strong normality!

A weak linear accessibility induces a proper linear relation between the clusters. As noted above, any cut in a weakly linear frame can *eo ipso* be regardedas a cut in the induced linear structure of its clusters.

Given any set of cases, S, denote by S^L the set of all necessities true everywhere in S (call it the *necessity set* of S). Any non-empty subset of the same cluster (in a reflexive and transitive model) has the same necessity set. The strongly normality of an S4 model is tantamount to an *isomorphism* between the induced clusters' frame and the partial inclusion-order amongst the necessity sets of these clusters. For a weakly linear model strong normality is entailed by different clusters having different necessity sets.

Lemma 5.1.2.1. *In a strongly normal* S4.3.1 *model no synthetic cluster has an immediate successor cluster, and furthermore, if such a cluster,* C, *is not terminal then* $C^L = \bigcap_{C'>C} C'^L$.

proof. Let α be both falsified and satisfied in C. If C has an immediate successor cluster, C_+, then clearly $C^L \subset C_+^L$ (proper inclusion). Let $L\beta$ be true in C_+ and false in C.

Then $\gamma=\alpha \vee L\beta$ is both satisfied and falsified in C, while it is true every where in C_+ and above. The cut between C and C_+ clearly falsifies S4.3.1: the special S4.3.1 schema (AD) is falsified in any case of C in which γ is false.

Thus C has no immediate successor; furthermore, if it has any clusters above it, it must satisfy any necessity true in all of them (or else derive a contradiction by same type of argument as above—omitting its second sentence and reading 'C_+' as 'everything above C') ■

From the above one can easily deduce

Corollary 5.1.2.2. *If a formula is satisfiable in a strongly normal weakly linear* S4.3.1 *model, then it is either satisfiable above every non-terminal cluster, or it is satisfiable everywhere in some cluster. (notice that this is utterly trivial for entropy maximisers and linear models).*

5.2. S4.3.1 Models as Weak Intensional Unions of Entropy Maximisers. We now prove

THEOREM 5.2.1. *Let* $\mathcal{M}$ *be any strongly normal weakly linear* S4.3.1 *model and let* $\mathfrak{f}$ *be any choice function which maps every non-terminal cluster on one of its own singleton subsets or on the empty set* $\varnothing$. *Let* 'C_{fin}' *denote the final cluster, if it exists, and* $\varnothing$ *otherwise. Let* $\mathcal{M}_\mathfrak{f}$ *be the submodel of* $\mathcal{M}$ *obtained by taking the cases in* $C_{fin} \cup \bigcup_{C \neq C_{fin}} \mathfrak{f}(C)$ *as the set of possible cases.Then (1)* $\mathcal{M}$ *is a weak intensional union of some or all those submodels of type* $\mathcal{M}_\mathfrak{f}$ *that are also casewise elementary submodels; and (2) if* $C_{fin} \neq \varnothing$ *we can take the union to be made exactly of those submodels of type* $\mathcal{M}_\mathfrak{f}$ *for which* $\mathfrak{f}(C) \neq \varnothing$, *for every non-terminal* C.

Proof. We separate the argument into two parts:

(a) $C_{fin} \neq \emptyset$. In this case it suffices to show that when $f(C) \neq \emptyset$ for every non-terminal C, $\mathcal{M}_f$ is a casewise elementary submodel of $\mathcal{M}$. The fact that $\mathcal{M}$ is a weak intensional union of these models will follow merely by observing that for each $w \in W_{\mathcal{M}} - C_{fin}$ there is an appropriate choice function picking w for the cluster containing it.

Let $f(C) \neq \emptyset$ for every non-terminal C. To prove that $\mathcal{M}_f$ is a casewise elementary submodel, it suffices to show (the Modal equivalent of Vaught's test) that if $V_{\mathcal{M}}(M\alpha,w)=1$ for w of $\mathcal{M}_f$, then $V_{\mathcal{M}}(\alpha,w')=1$ in some accessible case w' of $\mathcal{M}_f$.

If $w \in C_{fin}$ or if α is true in w itself or any where in C_{fin}, the claim is trivial. Assume then that $M\alpha$ and $\neg\alpha$ are both true in $w \notin C_{fin}$ but that α is true nowhere in C_{fin}.

Thus in the S4.3.1 model $\mathcal{M}|_w$ (casewise elementary submodel of $\mathcal{M}$) α is satisfiable but not in the final cluster. If there is a penultimate cluster, then by the above corollary (5.1.2.2) α would have to be satisfied everywhere in it. On the other hand, if there is no penultimate cluster, then by the same lemma either some non-terminal cluster would satisfy α everywhere, or else α would be both satisfied and falsified in infinitely many non-final clusters (while necessarily true in C_{fin}). Since the last alternative is barred by the fact that $\mathcal{M}|_w$ is a weakly linear S4.3.1 model, it follows that α is true everywhere in some cluster C above w, and in particular at $f(C)$.

(b) $\mathcal{M}$ has no terminal cluster. Let $\mathcal{F}$ be the set of all wffs and let $<\alpha_1,\alpha_2,...>$ be an enumeration thereof. For any case w let C_w be the cluster containing it. Define a function ψ from $\mathcal{F}$ into $R_{\mathcal{M}}$"{w} as follows: If $M\alpha_i$ is false in w then $\psi(\alpha_i)=w$. If $M\alpha_i$ is true in w, let w'_i be the highest case in the set $S_i=\{w,\psi(\alpha_1),\psi(\alpha_2),...,\psi(\alpha_{i-1})\}$ in which $M\alpha_i$ is still true.

α_i is satisfiable in $R_{\mathcal{M}}$"$\{w'_i\}$. If it is satisfiable above $C_w \cup C_{\psi(\alpha_1)} \cup ... \cup C_{\psi(\alpha_{i-1})}$ let $\psi(\alpha_i)$ be such a case. (clearly then $w'_i <_{R_{\mathcal{M}}} \psi(\alpha_i)$); otherwise α_i is satisfiable in $C_{w'_i}$ but not above! By corollary 5.1.2.2. above it must be satisfiable everywhere in this latter non-terminal cluster. We can choose then $\psi(\alpha_i)$ as w'_i.

The set $\Psi(w)=\{w\} \cup \{\psi(\alpha_i) \mid i<\omega\}$ contains at most one element per each cluster of $\mathcal{M}$. Define $f(C)=C \cap \Psi(w)$. $\mathcal{M}_f$ has then $\Psi(w)$ as its set of cases. Assume $M\alpha$ is true in $\mathcal{M}$ at a case $w' \in \Psi(w)$. Clearly w'=w or $w'=\psi(\alpha_n)$ for some $n \geq 1$>. Since the enumeration of wffs above is complete, there must be infinitely many values of k for which '$\alpha \equiv \alpha_{n+k}$' is a substitution instance of a propositional tautology. By definition $\psi(\alpha_{n+k}) \geq \psi(\alpha_n)$, where $\psi(\alpha_{n+k})$ satisfies α_{n+k} and therefore α as well. ■

The above theorem yields several weak intensional decomposition results which, combined with theorem 2.I, and with the (easily proved) fact that an intensional union (of any sort) of weak intensional unions (of models of a given type) is itself a weak intensional union of such models, produce the overall result that strongly normal S4.3.1 models are decomposable into (strongly normal) entropy maximisers (we forego detailed proofs). The semantic characteristics of these S4.3.1 entropy maximisers will be fully discussed in the next section. The immediate consequences of 5.2.1 are the following (remembering that the term *cluster* includes the trivial case of a singleton):

Corollary 5.2.2. *Any (strongly normal) weakly linear model of* S4.3.1 *without any final cluster is a weak intensional union of (strongly normal) linear* ω*-framed models.*

If, on the other hand, we start with a weakly linear S4.3.1 model that has a final cluster, the above tells us that it must be a weak intensional union of entropy maximisers—all of which end with the original model's final cluster. Since it is easily seen that any analytic cluster in any (casewise) elementary submodel can be replaced by any of its members without loss of "mimicry" ability in the reduced submodel (relative to remaining cases), an entropy maximiser with an analytic end-cluster must itself be a weak intensional union of linear models with a final case. Using the "transitivity" of intensional union concepts we get

Corollary 5.2.3. *Any (strongly normal) weakly linear model of* S4.3.1 *with an analytic final cluster is a weak intensional union of (strongly normal) linear models with a final case.*

Corollary 5.2.4. *Any (strongly normal) weakly linear model of* $E=S4.3.1+\{M(L\alpha \vee L\neg\alpha)\}$[13] *is a weak intensional union of (strongly normal) proper linear models.*

A third straightforward consequence is
Corollary 5.2.5. *Any (strongly normal) weakly linear model of* S4.3.1 *with a synthetic final cluster is a weak intensional union of (strongly normal) entropy maximisers with a single end cluster. It is not such a union of strongly normal linear submodels.*

That this is the best we can do in in this case should be apparent from the following argument (related to the one used in the counter-example of §2.2) :
A final *synthetic* cluster C_{fin} in a strongly normal weakly linear model, $\mathcal{M}$, would entail the existence of some α for which $L(M\alpha \wedge M\neg\alpha)$ is true everywhere in C_{fin}.Thus any linear submodel of $\mathcal{M}$ mimicking it at any world, w, of the final cluster would have to be *infinite,* if it is to satisfy the above formula at w; yet we can easily choose its final synthetic cluster to be finite[14].

5.3 Quasi-Determinism , Modal Continuity and S4.3.1. In an Entropy Maximiser every non-trivial cut has a linear past and an Entropy maximising future. The last non-trivial cut, with a single cluster as its only immediate & total future, will be called *the special cut.*
We start by extending the definition of Modal Continuity to cuts:

A cut $<W_1,W_2>$ is *modally continuous* iff $\bigcup_{w\in W_1} w^L = \bigcap_{w\in W_2} w^L$. Notice that a *model* is modally continuous, according to our *previous* definition, iff every *real* cut is such!
We now have
Lemma 5.3.1. *In a weakly linear model of* S4.3.1 *every cut is either quasi-deterministic or modally continuous.*

Proof. If a cut $<W_1,W_2>$ is not modally continuous, then (since $\bigcup_{w\in W_1} w^L \subseteq \bigcap_{w\in W_2} w^L$) let $L\beta$ be true everywhere in W_2 but false everywhere in W_1. If the cut is not quasi-deterministic, there is a formula α which is not recently determinate in it. The formula $\alpha \vee L\beta$ will serve then to refute the S4.3.1 special axiom. at any possible case where α is false.

The converse is true as well for entropy maximisers

Lemma 5.3.2. *An entropy maximiser in which every cut is modally continuous or quasi-deterministic is an* S4.3.1 *model.*

Proof. Let α be a formula refuting the special schema of S4.3.1 (AD), in an entropy maximiser, as described for a linear model at the opening comments of §4.0 above. Consider the two sets of cases: S_1, all these cases above which α changes its value at least twice; and S_2, all these cases in which α is necessarily true. The refutation of S4.3.1 requires only that (i) S_1 and S_2 be both non-empty, while (ii) any α-falsifying case belong to S_1. To prove 5.3.2 it is enough to show that $<S_1,S_2>$ constitutes a proper cut, for then it will either violate modal continuity or fail quasi-determinism, respectively, according to whether S_2 has a minimum or not. (If S_2 is the final cluster, we can treat the whole cluster as its own minimum). Since $S_1<S_2$, $<S_1,S_2>$ must be a proper cut if the set S_3 of cases above which α changes only once is shown to be empty! Since α becomes necessarily true it must remain so to the end. Therefore, at any member of S_3, α must be false (one change backwards from eventual truth). But then by (ii) above it must be a member of S_1.

We can be more precise about cuts that fail quasi-determinism in models of S4.3.1 .

Lemma 5.3.3. *If a cut* $<W_1,W_2>$, *in a strongly normal* S4.3.1-*entropy maximiser, is not quasi-deterministic, then it is a real cut in which* W_2 *has a minimum case* w* *which is the infimum of* $W_2-\{w^*\}$. *Furthermore, removing* w* *from the given model would produce a casewise elementary submodel, with* $<W_1,W_2-\{w^*\}>$ *as quasi-deterministic.*

Proof. Suppose first that some formula, α, would be immediatey determinate while failing recent determinacy in the special cut. Since this cut has only one immediate future—the final cluster—immmediate determinacy means that α is not contingent in the final cluster. Since this cut must also be, by lemma 5.3.1, modally continuous, either Lα or L¬α must be true somewhere in its past, implying recent determinacy. (contrad.)

Let us abbreviate *quasi-deterministic* as *q.D.* We now prove the following

claim. *A (bilaterally) open cut in a model of the assumed type must be quasi-deterministic.*

Suppose a non-trivial regular cut $<W_1,W_2>$is not q.D. Let γ be immediatey determinate while failing recent determinacy in it. W_1 must be infinite without maximum. If W_2 has no minimum , then any immediate future ΔF in which γ has fixed value (such must exist) has infinitely many cases. Let w'<w" be two such cases which are not in the final cluster (this is possible because the cut is not special). By strong normality there is a δ such that Lδ is true at w" but not in w'. Either γvLδ or ¬γvLδ would then refute S4.3.1 in our model (in much the same way as in the $\omega+\omega^-$ model $\mathcal{M}^*$ of 4.0).This proves the claim.

A non q.D cut as above could not be open, then, and its future W_2 must have a minimum—say w*. But this w* must also be the infimum of W_2–{w*], since a proper successor case w^{*+} entails a contradiction to S4.3.1, by strong normality, using the same technique as above: Let η be γ if γ is true at w* and ¬γ otherwise, for a γ manifesting the non-q.D nature of the cut, and let Lξ be true at w^{*+}, but not before. Then ηvLξ is the refuting formula.

Suppose Lζ is true everywhere in W_2–{w*], but no-where in W_1. Let η be as above. ηvLζ would then constitute a counter-example to S4.3.1. We thus gather that

$\bigcup_{w \in W_1} w^L = w^{*L} = \bigcap_{w \in W_2 - \{w^*\}} w^L$ and therefore, that the truth of any formula at w* must imply its satisfiability above. This is enough in order to prove, at the crucial induction step, that removing w* leaves the infrastructure for a casewise elementary submodel .
To prove that the cut $<W_1, W_2 - \{w^*\}>$ is q.D in this submodel , it is enough, by the proved claim above, to note that it is open and that this submodel is itself an S4.3.1model.■

The above lemma suggests that one could stop fussing about the topological nature of S4.3.1 models and try to characterise them solely by the deterministic character of their "cuts". It is indeed possible to use the above to prove[15]

THEOREM 5.3.4. *Let* $\mathcal{M}$ *be a strongly normal boundedly compact* S4.3.1– *entropy maximiser, Then* $\mathcal{M}$ *is a casewise elementary extension of the quasi-deterministic submodel* $\mathcal{M}^Q$ *obtained by removing all the cases in the linear part which do not have a quasi-deterministic cut immediately below them. Furthermore, if* $\mathcal{M}$ *is dense in the linear part, then* $\mathcal{M}^Q$ *is dense in* $\mathcal{M}$: *there is a case of* $\mathcal{M}^Q$ *between any two cases of the linear part of* $\mathcal{M}$.

Just as importantly, we can reverse the direction of this theorem. Inasmuch as a strongly normal S4.3.1 entropy maximiser has (like the submodel above) some cuts that are either open or fail modal continuity (all of which must be quasi-deterministic) we can plug each of these cuts with a new "reifying" case—so as to create a boundedly compact and modally continuous elementary extension (both casewise and globally).

THEOREM 5.3.5. *Any strongly normal denumerable* S4.3.1–*Entropy Maximiser is casewise elementarily embeddable in an (elementarily equivalent) cofinal entropy maximiser which is boundedly compact (order complete) and modally continuous.*

Proof. If a strongly normal S4.3.1–Entropy Maximiser, $\mathcal{M}$, is not boundedly compact , it must have some open cuts (none of which is special), and all such cuts must be quasi-deterministic. Likewise, if $\mathcal{M}$ is not modally continuous, there must be a real cut with a non-ending past, which is not modally continuous, but is also quasi-deterministic (lemma 5.3.1). In either case the problem cuts have an infinite past without a maximal element. For each such cut $C=<W^C_1, W^C_2>$ let $\Delta(W^C_1)$ denote the set of all *recent* pasts. $\Delta(W^C_1)$ has the finite intersection property. We can therefore extend it to an ultrafilter, $D(W^C_1)$, over W^C_1, which , for such cuts, must contain the *Fréchet filter* over W^C_1 , $Fr(W^C_1)= \{X \mid X \subseteq W^C_1 \ \& \ W^C_1 - X$ is finite$\}$—since every member of $Fr(W^C_1)$ contains members of $\Delta(W^C_1)$. Also, since every member of $\Delta(W^C_1)$ is included in some member of $Fr(W^C_1)$ it follows that an ultra-filter over W^C_1 contains $\Delta(W^C_1)$ if and only if it extends $Fr(W^C_1)$. Finally we notice that every member of such an ultra-filter must be cofinal with W^C_1, or else it would have a null intersection with some recent past of C.
Define now a *new case* w_C—where the value of a variable v is set by
$V(v,w_C)=1$ iff $\{w \mid w \in W^C_1 \& V(v,w)=1\} \in D(W^C_1)$
Each such w_C is placed "inside" the corresponding cut (i.e $W^C_1 < w_C < W^C_2$). Let M* be any model obtained in this manner; we will show that this operation can not create new problem cuts but must eliminate the "treated" ones!
Prove by induction the following <u>two claims (together)</u>:

(a) For any plugged C as above $V_{\mathcal{M}^*}(\alpha,w_C)=1$ **iff** $\{w \mid w\in W^C{}_1 \& V_{\mathcal{M}}(\alpha,w)=1\}\in D(W^C{}_1)$, for any formula α—implying that

<u>recent necessity (impossibility) of α in C $\Rightarrow$ α is true (false) in w_C</u> — and

(b) For any wff α, $V_{\mathcal{M}^*}(\alpha,w)=V_{\mathcal{M}}(\alpha,w)$ at any *old* case $w\in W_{\mathcal{M}}$.

The basis claim and the induction step for truth functions are simple to prove, and depend only on universal properties of ultra-filters (completely analogous steps are carried out in proving Theorem 6.2.I in appendix I below).

If $V_{\mathcal{M}^*}(L\alpha,w_C)=1$ then

$(\forall w')\,(w' \geq_{\mathcal{M}^*} w_C \rightarrow V_{\mathcal{M}^*}(\alpha,w')=1)$, which, by the Induction Hypothesis—using both (a) and (b)—is true iff

both $(\forall w')\,(w' \in W^C{}_2 \rightarrow V_{\mathcal{M}}(\alpha,w')=1)$ and

$(\forall C')\,(C' \geq C \,\&$ C' is a "problem cut" in $\mathcal{M} \rightarrow \{w \mid w\in W^{C'}{}_1 \& V_{\mathcal{M}}(\alpha,w)=1\}\in D'(W^{C'}{}_1))$.

where the inequality in the last expression refer to the natural order amongst cuts in $\mathcal{M}$. The last conjunct clearly implies that $S_1 =_{Df} \{w \mid w\in W^C{}_1 \& V_{\mathcal{M}}(\alpha,w)=1\}\in D(W^C{}_1)$, while the first conjunct implies that $\alpha \in (W^C{}_2)^L$. Since $\mathcal{M}$ is <u>quasi-deterministic</u>, it follows that α is <u>recently determinate</u> in C. This, together with the cofinality of S_1 with $W^C{}_1$, implies that $\{w \mid w\in W^C{}_1 \& V_{\mathcal{M}}(L\alpha,w)=1\}$ includes a recent past of C and therefore must belong to $D(W^C{}_1)$.

Conversely, if $\{w \mid w\in W^C{}_1 \& V_{\mathcal{M}}(L\alpha,w)=1\}$ belongs to $D(W^C{}_1)$, it must be non-empty, and α must be <u>recently necessary</u> in C, and in every cut above it, and true in every old case of $W^C{}_2$. It then follows from the induction hypothesis that $V_{\mathcal{M}^*}(L\alpha,w_C)=1$.

A straighforward application of the induction hypothesis shows that $L\beta$ must remain false at old cases in $\mathcal{M}^*$ when it was so in $\mathcal{M}$. All the above shows not only the case wise elementary equivalence of $\mathcal{M}^*$ and $\mathcal{M}$, but also their ordinary elementary equivalence. That the new model is boundedly compact follows from the mere plugging of all open cuts. New open cuts can not be created.

<u>To prove Modal Continuity</u>, let S be any bounded-above-set of cases (possible worlds) in $\mathcal{M}^*$, which has no maximal (latest) member, but which satisfies $M\alpha$ everywhere. Suppose it has a supremum w^S in $\mathcal{M}^*$. Let C* be the cut immediately below this case, i.e, $<\{\text{cases}<w^S\},\{\text{cases}\geq w^S\}>$. C*'s past is cofinal both with S and with the past of C*'s old counterpart in $\mathcal{M}$, C ($\{\text{old-cases}<w^S\},\{\text{old-cases}\geq w^S\}$). Thus $M\alpha$ is recently necessary both in C* and in C. If w^S is a new case, it satisfies $M\alpha$ by claim (a) above. If w^S is an *old case*, then C*=C, where C as a real cut must have been originally modally continuous. ■

6. Strongly normal S4 Models and the Idiosyncracy of ω-frames.

6.0. The above results are enough to show that S4.3.1 is not about discrete linearity and that its discrete well-ordered models (finite and ω-based) are exceptional, in a sense, despite their sufficiency for an ordinary completeness result. Yet the question lingers: *If the ω–frames are so idiosyncratic, what is their relationship to the more general type of model?*

6.1. Alluring Topological Intuitions. Consider strongly normal S4 models. Within their necessarily transitive frames there may occur end points and "bubbles", which constitute "local" S5 submodels. A seeming "semantical eye-catching" topological predicate of such frames has to do with such end features. An*end cluster* —which may be a singleton—has only its own elements as accessible to any of its worlds. Now, there is a topological difference between models/ frames in which every possible-case has in sight some accessible end-cluster, on one hand, and those, on the other, in which there may be a"branch" or "open ended section" with no such clusters. One is tempted to correlate the former structures with the validity of the set of statements asserting the *possible necessity of S5 theorems.*

6.2. But Misleading. This however is a topological mirage for logicians! The truth is that *there is nothing formulable as an extension of S4 that is exclusively about open ended linearities such as* ω. For every ω-based model there is an elementary extension which is an Entropy Maximiser, and for some ω-based models such extensions are necessarily non-linear.

THEOREM 6.2.I. *Let* $\mathcal{M}$ *be an* ω*-model, i.e,* $\mathcal{M} = \langle\omega, \leq, V\rangle$. *Then there exists an Elementary End Extension of* $\mathcal{M}$, $\mathcal{M}^* = \langle W^*, R^*, V^*\rangle$, *satisfying:*

(1) $\mathcal{M}^*$ *is an Entropy-Maximiser with* $\mathcal{M}$ *as a maximal initial linear model, and, therefore*

$\overline{M} = \mathcal{M}^* - \mathcal{M} = \langle W^* - \omega, R^* - \leq, V^* \restriction_{\text{Formulae} \times W^* - \omega}\rangle$, *is a S5-model;*

(2) $\mathcal{M}$ *is a case-wise elementary submodel of* $\mathcal{M}^*$*; and*

(3) *For any* α, $\mathcal{M} \vdash ML\alpha \iff \overline{M} \vdash L\alpha$.

This is proved in Appendix I below, where the cases of $\overline{M}$ correspond to all the ultrafilters over ω containing the Fréchet filter ($= \{X \mid X \subseteq \omega \ \&\ \omega - X \text{ is finite}\}$). Notice that if $\mathcal{M}$ has some necessarily contingent formula, $\mathcal{M}^*$ will be irreducibly non-linear.

From the proof of this theorem, and its generalisation to models with any linear order type without end, one can see that discarding models with open ended linear "branches" is not going to affect the ordinary characterisation of any S4 extension logic (this does not mean that any ω-framed model is reducible, or dispensible from a reverse semantical point of view).

6.3. Syntactic correlate of Theorem 6.2.I. The syntactic relationship deducible from the above theorem therefore amounts to

THEOREM 6.3.I. *If* $S5 \vdash \alpha$ *then* $S4 \vdash ML\alpha$

which can be proved directly by straightforward *syntactic induction*. The converse is trivial.

7. Concluding remarks.

7.1. The concept of (reverse semantical) "capture". We recognise that our explications of this concept seem to contain an arbitrarily circumscribed move—that of trying to represent every (nominalistically acceptable) model of a logic $\mathcal{L}$ as an intensional union —*in one of the senses used in this paper*—of the models we want to capture. After all, it might be claimed, one could have presented a *different* mode of "intensional" composition, which might have given rise to different relationships between reductandum-models and candidate reductans-models.

One must stress, however, that the required reduction is subject to some essential constraints, first and foremost of which is the natural reductive requirement that the components of the reduction should "worldwise" retain between them all the semantic value information of the original (strongly normal) model. But this, of course, does not always allow a perfect reconstruction of the original accessibility-frame. Since such accessibility frames, in general, can take many mathematical shapes and forms, we used a universal mode of composition, in which frame-overlaps among components could be as large as needed to retain essential frame properties. This corresponds to what is required in classical model theory in order to form unions of models (out of chains or directed families)[16].

A critic of the restriction to "simplistic" intensional unions might argue that our very proof of theorem 6.2.I shows how an entropy maximising elementary extension of an ω-model can be obtained from this very ω-model by using "composition" by ultra-filters, and may go on to suggest that any legitimate entropy-maximising model of S4.3.1 is related to its initial linear-segment in a similar manner. Our remarks in §2.2 suggest however that, in analogy to the case of finite entropy maximisers, this is not so for infinite ones, either. Indeed, one can easily construct infinite entropy maximising models of S4.3.1, with an initial ω-segment, that are not elementarily equivalent to any of their ω-submodels.

More important, however, is the fact that most of our results, concerning different types of S4.3.1 models, are significant independently of the concept of intensional union, and can be interpreted conventionally as stating that this logic has many (nominalistically) legitimate-models, which are very different from its linear-discrete models and are seemingly irreducible to them by any standard device that is structurally intelligible independently of specific semantic information. Furthermore, as indicated above, while linear models can always be embedded in non-linear entropy-maximisers, the latter do not always contain (casewise) elementarily equivalent linear submodels. The fact that each entropy-maximiser can be made to agree in a world of its linear part with some infinite (well-ordered) linear extension of its linear part (as in Segerberg's constructions) is not *philosophically* significant and certainly does not "reduce" that entropy maximiser to linearity—even if we ignore the fact that such linear extensions violate any canon of (nominalistic) semantic parsimony. Such constructions are mere artefacts of completeness proofs.

Equally important is the result that modally-continuous, order-complete entropy-maximising (and linear) models are a better choice as *paradigmatic* models of S4.3.1, which match more perspicuously its special axiom. One may argue that entropy maximisers are quite linear (although to our mind the "linearity" of such weakly linear models diminishes when the cardinality of their linear part is small compared with that of its (final) S5-cluster); but one can hardly claim that order-complete, dense, sets (like the Reals) are quite discrete! Even if we were to choose *all modally continuous, order complete, strictly linear structures* as our preferred semantically characteristic set for S4.3.1, in the *classical* sense, we would still be putting ourselves far beyond the set of ω-framed models—which would then constitute a very special subset of our favoured linear models.

More generally we can say that the weak-linearity of certain classes of S4.3.1 models that suffice for completeness is inherited from S4.3. Its special axiom. however, does not dictate an additional frame property, but a quasi-determinism and/or modal continuity of the valuation function, within the same frame types. The completeness-sufficiency of ω-frames for S4.3.1 is thus interesting but not model-informative! All these results are valid independently of the significance one attaches to the enterprise of "reverse semantics".

However, it is only through some "reverse semantical" concept of "capturing" that one can elaborate a coherent notion of a *paradigmatic type* of model (of a Logic). As the greek term suggests, a pardigm should be be fully instrumental to understanding the totality of beings it exemplifies. Thus, for example, it is not enough to choose the *prime* models of a theory as paradigmatic, unless we have a standard way of reconstructing from them the other legitimate models[17]. We should not regard the ω-framed models of S4.3.1 as paradigmatic any more than we should regard the "standard" model of Arithmetic as paradigmatic of Peano Axioms. Indeed the partial-similarity between S4.3.1 and Peano-Axioms with respect to the relationships between "standard" and "non-standard" models, and inter alia amongst the latter , using cofinal and end-extensions, is quite striking[18].

7.2. Entropy Maximisers, Deterministic-Causal mechanisms and The Nature of Time.
Results such as theorem 6.2.I show that no extension of S4 can capture "exactly" the concept of a strictly linear time, *in its full generality* (results such as 5.2.3 and 5.2.4 are connected only with those linear frames, with a terminal "moment", that can be derived from *intensionally-decomposable* entropy maximisers; they do not apply to *open-ended* linearity). *Yet can't we say that logics such as* S4.3.1 *are about an alternative concept of time* ? One reason for raising this possibility here is that entropy-maximisers are strongly indicated, as a minimally liberalised version of strictly linear time, by two very different conceptual schemes—both of which are intimately linked to Physics.

First there is the old idea (Clausius and Boltzmann) that the one-directionality of time is connected with the overall increase of entropy (in a closed system). This suggests that when entropy reaches a maximum, the directionality of time breaks down in the sense that allowed physical transitions are reversible and there is no criterion to distinguish very recent past from some "future". In a finite universe with a certain kind of physics a maximal entropy is inevitable. It seems to us, then, that the appropriate global physical structures that make time possible for such a universe must be those of *non-linear entropy maximisers* (or intensional unions thereof). Linear time then will have been explained by a structure which *eo ipso* explains also its ultimate destruction (notwithstanding the seeming paradox behind this very statement!).

A second and perhaps more fascinating indication come from a different quarter, which connects well with Field theory. Consider a "quantised" scalar field Φ in space-time obeying a deterministic-causal law of evolution. For simplicity assume that Φ is restricted to a finite range of values. The causal action of the field on itself may be restricted by some universal upper limit (e.g., the speed of light). In a discrete space and time, this would be represented by a rule of the type $\Phi(x, t+1)=F(\{\Phi(x_i, t) \mid x_i \in \mathcal{N}(x)\})$, (where $\mathcal{N}(x)$ is an appropriate "neighborhood" of space-cell x, that can be made "commensurate in size" with the maximal speed of propagation of causal action). Such a rule describes what is known as a *cellular automaton* and is itself alluded to as its *transition function* or *transition-rule* . The behaviour of such automata can be adequately modelled by using a two-valued scalar function.

Now there is an obvious way in which each such binary cellular automaton will provide us with a finitary model of discrete temporal logic. Let the concurrent (binary) scalar field values at the different "spatial" cells correspond to the valuation of (finitely many) different "basic" propositions, so that a *global state configuration* (GSC) corresponds to a finitary "possible world" or a "moment", and let $F^k(C)$ be the GSC obtained by applying this cellular automaton (defined by F) k times to the global state configuration, C; Then α is *necessarily true* in C *iff* α is true in $F^k(C)$ for all $k \geq 0$. Inasmuch as α is a modal function of propositions whose values for each GSC are the scalar values in predetermined cells therein, the truth-value of α is computable. Remembering, now, that such mechanisms are used here to model a deterministic universe, we can say that the corresponding *time-like structure* is determined by the way the automaton *partitions* the state-space (the set of all *a-priori* possible GSC's) into paths of deterministic evolution.

Yet here again we end up with nothing but Entropy Maximisers! Such cellular automata will always partition the state-space into separate *basins of attraction*, each of which consists of an *attractor-cycle* (which could be a fixed-point) with any possible number (including 0) of *transient lines* leading into this cycle (transients may merge like rivers before reaching an attractor cycle, but they can not split-up). It is clear that each such basin of attraction is a (strong) *intensional union* of entropy maximisers, and that we can regard the whole induced global partition structure as such a union as well. We notice that there is no possibility of an infinite-length transient, in the above universe, unless the state-space itself is infinite. Since we assumed the number of scalar values to be finite, a space of *finitely many* "cells" would preclude any infinite (ω) transient. However an ω-time line can be easily generated by a cellular automaton operating on infinite state configuration.

Analysing the *global* temporal structures that such mechanisms induce in the state-space—structures that just began to be seriously mapped (e.g, by [Wuensche and Lesser, 92]) without the seductive analogy of "Chaos Theoretical" attributes of continuous non-linear dynamic behaviour—opens up a whole new field of fascinating study into the relationship between the fundamental transition function that determines the evolution of a discrete scalar field and the overall structure of the intensional-unions of entropy maximising "time lines" that describe this evolution. Although these *hyper-temporal* structures depend both on the transition function and on the innate topology of the "spatial" field (e.g., for cellular automata, on the exact total number of cells and on what constitutes a neighbourhood), there are some important features that can be linked directly, for most spaces, to the nature of the transition-function itself. For instance, it is possible to deduce high symmetries in such hyper-temporal structures—isomorphism amongst most of the maximal entropy maximising sub-frames thereof—in the case of transition functions that are relatively good at partially preserving the (semantic) information encoded in global-state-configurations.

University of Hertfordshire
May 1995

Appendix I: Elementary End Extensions of ω-Models

THEOREM: Let $\mathcal{M}$ be an ω-model, i.e, $\mathcal{M} = <\omega, \leq, V>$. Then there exists an Elementary End Extension of $\mathcal{M}$, $\mathcal{M}^* = <W^*, R^*, V^*>$, satisfying:

(1) $\mathcal{M}^*$ is an Entropy-Maximiser with $\mathcal{M}$ as a maximal initial linear model, and, therefore

$\overline{M} = \mathcal{M}^* - \mathcal{M} = <W^* - \omega, R^* - \leq, V^* \restriction \text{Formulae} \times W^* - \omega>$, is a S5-model;

(2) $\mathcal{M}$ is a case-wise elementary submodel of $\mathcal{M}^*$; and

(3) For any α, $\mathcal{M} \models ML\alpha \iff \overline{M} \models L\alpha$.

Proof: Let **Fr** be the *Fréchet filter* over ω, i.e, $\mathbf{Fr} = \{X \mid X \subseteq \omega \ \&\ \omega - X \text{ is finite}\}$.
Let 1 be the value corresponding to truth, while 0 corresponds to falsehood.
Let 'TV(s)' denote the truth-value of a statement s.
Let UF(ω) be the set of all ultra-filters over ω, and let W_{end} and V_{end} be defined by

$W_{end} = \{D \mid D \in UF(\omega) \ \&\ D \supset Fr\}$ and, for any propositional variable, u,

$V_{end}(u,D) = \text{truth-value}(\{i \mid V(u,i) = 1\} \in D)$,

together with the usual recursive valuation clauses.

Let $R_{end} = W_{end} \times W_{end}$ and set $\overline{M}$ as $<W_{end}, R_{end}, V_{end}>$ so that

$W^* = \omega \cup W_{end}$; $R^* = \leq_\omega \cup W^* \times W_{end}$; $V^* \restriction Var \times \omega = V$ and $V^* \restriction Var \times W_{end} = V_{end}$.

Lemma 1. For any formula α, $V_{end}(\alpha,D) = \text{truth-value}(\{i \mid V(\alpha,i) = 1\} \in D)$.

Proof (by induction on construction of α). For propositional variables this is true by definition. If the main connective is negation($\neg$) then

$$
\begin{aligned}
V_{end}(\neg\alpha, D) &=_{Df} 1 - V_{end}(\alpha,D) \\
&= 1 - TV(\{i \mid V(\alpha,i) = 1\} \in D) \quad \text{(by Induction Hyp.)} \\
&= 1 - TV(\omega - \{i \mid V(\alpha,i) = 1\} \notin D) \quad \text{(by Ultra-filter properties)} \\
&= 1 - TV(\{i \mid V(\alpha,i) = 0\} \notin D) \\
&= 1 - TV(\{i \mid V(\neg\alpha,i) = 1\} \notin D) \\
&= TV \neg(\{i \mid V(\neg\alpha,i) = 1\} \notin D) \\
&= TV(\{i \mid V(\neg\alpha,i) = 1\} \in D).
\end{aligned}
$$

If the main connective is conjunction ($\wedge$) then

$$
\begin{aligned}
V_{end}(\alpha \wedge \beta, D) &= V_{end}(\alpha,D) \times V_{end}(\beta,D) \\
&= TV(\{i \mid V(\alpha,i) = 1\} \in D) \times TV(\{i \mid V(\beta,i) = 1\} \in D) \quad \text{(by Induction Hyp.)} \\
&= TV(\{i \mid V(\alpha,i) = 1\} \in D \ \&\ \{i \mid V(\beta,i) = 1\} \in D)
\end{aligned}
$$

but since *for any ultra-filter* D, $A \in D \ \&\ B \in D$ iff $A \cap B \in D$ the above is

$$
\begin{aligned}
&= TV(\{i \mid V(\alpha,i) = 1\} \cap \{i \mid V(\beta,i) = 1\} \in D) \\
&= TV(\{i \mid V(\alpha,i) = V(\beta,i) = 1\} \in D) \\
&= TV(\{i \mid V(\alpha \wedge \beta, i) = 1\} \in D).
\end{aligned}
$$

For other truth-functional connectives the proof is similar, relying on the inductive hypothesis about the operands and on the properties of any ultrafilter.
The crucial part of proof is for the case when the main operator (connective) is *necessity* (L). Then

$V_{end}(L\alpha,D) = 1$ iff

$(\forall D')(D' \in W_{end} \rightarrow V(\alpha,D') = 1)$, which, by the Induction Hypothesis, is true

iff $(\forall D')(D' \in W_{end} \rightarrow \{i \mid V(\alpha,i) = 1\} \in D')$, or equivalently

iff $\{i \mid V(\alpha,i) = 1\} \in \cap \{D' \mid D \in W_{end}\}$

But it is easy to prove that *the intersection of all ultra-filters over ω which include a proper filter F over ω must be F itself* [19]; hence

$$V_{end}(L\alpha,D) = 1 \text{ iff } \{i \mid V(\alpha,i) = 1\} \in \mathbf{Fr}. \qquad (\Delta)$$

Now suppose

$\{i \mid V(L\alpha,i) = 1\} \in D$

This is the case iff (by definition)

$\{i \mid (\forall j)(j \geq i \rightarrow V(\alpha,j) = 1)\} \in D$

Since D is a proper Ultrafilter, this implies that

$\{i \mid (\forall j)(j \geq i \to V(\alpha,j)=1)\} \neq \varnothing$ (the empty set)

which means that

$(\exists i)(\forall j)(j \geq i \to V(\alpha,j)=1)$

Let i_0 be the minimal value of i satisfying the matrix of the existential quantifier. Then the set $\{i \mid i \geq i_0\}$ is exactly the set $\{i \mid (\forall j)(j \geq i \to V(\alpha,j)=1)\}$ and it is obviously a member of the Fréchet Filter, Fr. Thus by the equivalence (Δ) above, $V_{end}(L\alpha,D)=1$.

Conversely: Suppose $V_{end}(L\alpha,D)=1$, for any $D \in W_{end}$. Then, by (Δ)

$\{i \mid V(\alpha,i)=1\} \in \mathbf{Fr}$

$\leftrightarrow \omega - \{i \mid V(\alpha,i)=1\}$ is finite

$\leftrightarrow (\exists i)(\forall j)(j \geq i \to V(\alpha,j)=1)$

$\leftrightarrow (\exists i)(\forall k)[k \geq i \to (\forall j)(j \geq k \to V(\alpha,j)=1)]$

$\leftrightarrow (\exists i)(\forall k)[k \geq i \to V(L\alpha,k)=1)]$

$\leftrightarrow \{k \mid V(L\alpha,k)=1\} \supseteq \{i \mid i \geq i_0\}$, where i_0 is any value of i satisfying the existential quantifier above. But since the smaller set belongs to **Fr**, so does the superset $\{k \mid V(L\alpha,k)=1\}$.

Since **Fr** is the intersection of all the members of W_{end}, $\{k \mid V(L\alpha,k)=1\} \in D$, for any $D \in W_{end}$.

We can now prove the rest of thr above theorem. We prove first [(3) above] that for any α, $\mathcal{M} \vdash ML\alpha \Leftrightarrow \overline{M} \vdash L\alpha$. This will follow from

Lemma 2.

$(\mathcal{M} \vdash ML\alpha) \to (\forall D)(D \in W_{end} \to V_{end}(\alpha,D)=1)$;

$(\mathcal{M} \vdash ML\neg\alpha) \to (\forall D)(D \in W_{end} \to V_{end}(\alpha,D)=0)$; and

$(\mathcal{M} \vdash LM\alpha \wedge LM\neg\alpha) \to (\exists D_1)(D_1 \in W_{end} \,\&\, V_{end}(\alpha,D_1)=1) \,\&\, (\exists D_2)(D_2 \in W_{end} \,\&\, V_{end}(\alpha,D)=0)$.

Proof.

$(\mathcal{M} \vdash ML\alpha) \to (\exists i)(\forall j)(j \geq i \to V(\alpha,j)=1)$

$\to (\exists i)[\{j \mid j \geq i\} \subseteq \{k \mid V(\alpha,k)=1\}$

$\to \{k \mid V(\alpha,k)=1\} \in \mathbf{Fr}$

$\to (\forall D)(D \in W_{end} \to \{k \mid V(\alpha,k)=1\} \in D)$, since $(\forall D)(D \in W_{end} \to D \supset \mathbf{Fr})$

$\to (\forall D)(D \in W_{end} \to V_{end}(\alpha,D)=1)$

$(\mathcal{M} \vdash ML\neg\alpha) \to (\exists i)(\forall j)(j \geq i \to V(\neg\alpha,j)=1)$

$\to (\exists i)[\{j \mid j \geq i\} \subseteq \{k \mid V(\alpha,k)=0\}$

$\to \{k \mid V(\alpha,k)=0\} \in \mathbf{Fr}$

$\to (\forall D)(D \in W_{end} \to \{k \mid V(\alpha,k)=0\} \in D)$, since $(\forall D)(D \in W_{end} \to D \supset \mathbf{Fr})$

$\to (\forall D)(D \in W_{end} \to V_{end}(\alpha,D)=0)$

If $(\mathcal{M} \vdash LM\alpha \wedge LM\neg\alpha)$ then both $\{k \mid V(\alpha,k)=1\}$ and $\{k \mid V(\alpha,k)=0\} = \omega - \{k \mid V(\alpha,k)=1\}$ must be infinite. So neither of them belong to **Fr.** We now use the following basic lemma[20]. *If F is a proper filter over ω and a subset Y of ω is such that neither it nor its ω complement, $(\omega - Y)$, belong to F, then there is a proper ultrafilter over ω containing $F \cup \{Y\}$.*

We can apply this lemma twice to show the existence of ultrafilters D_1 and D_2 containing $Fr \cup \{k \mid V(\alpha,k)=1\}$ and $Fr \cup (\omega - \{k \mid V(\alpha,k)=1\}$, respectively, which by definition implies $(\exists D_1)(D_1 \in W_{end} \,\&\, V_{end}(\alpha,D_1)=1) \,\&\, (\exists D_2)(D_2 \in W_{end} \,\&\, V_{end}(\alpha,D)=0)$.

A corollary of this proof is (3) For any α, $\mathcal{M} \vdash ML\alpha \Leftrightarrow \overline{M} \vdash L\alpha$.

That $\mathcal{M}^*$ as defined is an Entropy Maximiser follows straightforwardly from definitions. To prove that $\mathcal{M}$ is a casewise elementary submodel of $\mathcal{M}^*$, we show that if $(\forall i<\omega)[V(\alpha,i)=V^*(\alpha,i)]$ then $(\forall i<\omega)[V(L\alpha,i)=1 \Leftrightarrow V^*(L\alpha,i)=1]$.

But $V^*(L\alpha,i)=1 \Leftrightarrow_{df} (\forall D)(D \in W_{end} \to V_{end}(\alpha,D)=1) \,\&\, (\forall j)(j \geq i \to V(\alpha,j)=1)$. The second conjunct on the right is definitionally equivalent to $V(L\alpha,i)=1$ which implies $\mathcal{M} \vdash ML\alpha$, which implies, by the above, $\overline{M} \vdash L\alpha$, which is equivalent by definition to the first conjunct.

Thus $V^*(L\alpha,i)=1 \Leftrightarrow (\forall j)(j \geq i \to V(\alpha,j)=1) \Leftrightarrow V(L\alpha,i)=1$.

This is sufficient to prove by induction that $\mathcal{M}$ is a casewise elementary submodel of $\mathcal{M}^*$.

Appendix II: Syntactic Proof that If S5 $\vdash \alpha$ then S4 $\vdash ML\alpha$.

In the following proof we shall use the notations of and references from Huges and Creswell's *An Introduction to Modal Logic* [henceforth abbreviated as IML]. 'PC' stands for classical propositional calculus.

Derived rules of Inference

DR1: If $\vdash \alpha\supset\beta$ then $\vdash L\alpha\supset L\beta$ [IML p.33]
valid in $K\subset T\subset S4$.

THEOREM. $\vdash L(\alpha\supset\beta)\supset(M\alpha\supset M\beta)$ (**T8** in IML, p.37, valid in $K\subset T\subset S4$.)

DR3: If $\vdash \alpha\supset\beta$ then $\vdash M\alpha\supset M\beta$ [IML p.37]
Follows from above theorem via Necessitation of hypothesis. Valid in $K\subset T\subset S4$.

An S4 Derived rule NCN: If S4 $\vdash L\alpha\supset\beta$ then S4 $\vdash L\alpha\supset L\beta$
Proof: apply DR1, then use S4 $\vdash L\alpha\supset LL\alpha$,and Hyp. Syll.

An S4 Derived rule MLdist: If S4 $\vdash ML(\alpha\supset\beta)$ then S4 $\vdash ML\alpha\supset ML\beta$
Proof:
(0) S4 $\vdash ML(\alpha\supset\beta)$ [Hypothesis]
(1) S4 $\vdash \alpha\supset(\alpha\supset\beta\supset\beta)$ [PC]
(2) S4 $\vdash L\alpha\supset L(\alpha\supset\beta\supset\beta)$ [**DR1** (1)]
(3) S4 $\vdash L(\alpha\supset\beta\supset\beta)\supset(L(\alpha\supset\beta)\supset L\beta)$ ['L'-distribution over '$\supset$' or **A6** in **IML**]
(4) S4 $\vdash L\alpha\supset(L(\alpha\supset\beta)\supset L\beta)$ [Hyp. Syll., (2)&(3)]
(5) S4 $\vdash L\alpha\supset L(L(\alpha\supset\beta)\supset L\beta)$ [NCN (4)]
(6) S4 $\vdash L(L(\alpha\supset\beta)\supset L\beta)\supset(ML(\alpha\supset\beta)\supset ML\beta)$ [Subs. inst. of T8, IML p.37]
(7) S4 $\vdash L\alpha\supset(ML(\alpha\supset\beta)\supset ML\beta)$ [Hyp. Syll., (5)&(6)]
(8) S4 $\vdash ML(\alpha\supset\beta)\supset(L\alpha\supset ML\beta)$ [PC—Permutation.,(7)]
(9) S4 $\vdash L\alpha\supset ML\beta$ [MP, (8)&(0)]
(10) S4 $\vdash ML\alpha\supset MML\beta$ [**DR3**, (9)]
(11) S4 $\vdash MML\beta\supset ML\beta$ [**T18** ,IML p.46]
(12) S4 $\vdash ML\alpha\supset ML\beta$ [Hyp. Syll., (10)&(11)]

An S4 Derived rule MLMP: If S4 $\vdash ML(\alpha\supset\beta)$ and S4 $\vdash ML\alpha$, then S4 $\vdash ML\beta$
Proof: apply MLdist and MP.

Lemma1. If S4 $\vdash\alpha$, then S4 $\vdash ML\alpha$
Proof: trivial.

THEOREM: S4 $\vdash ML(ML\alpha\supset L\alpha)$
Proof:
(1) $L\alpha\supset(ML\alpha\supset L\alpha)$ [Subst. inst. of PC]
(2) $LL\alpha\supset L(ML\alpha\supset L\alpha)$ [**DR1**, (1)]
(3) $L\alpha\supset LL\alpha$ [S4 schema]
(4) $L\alpha\supset L(ML\alpha\supset L\alpha)$ [Hyp. Syll., (3)&(2)]
(5) $\neg ML\alpha\supset(ML\alpha\supset L\alpha)$ [Subst. inst. of PC]
(6) $LM\neg\alpha\supset\neg ML\alpha$ [M definition, K and PC]
(7) $LM\neg\alpha\supset(ML\alpha\supset L\alpha)$ [H.Syll, (6)&(5)]
(8) $LLM\neg\alpha\supset L(ML\alpha\supset L\alpha)$ [**DR1**, (7)]
(9) $LM\neg\alpha\supset LLM\neg\alpha$ [S4 axiom]
(10) $LM\neg\alpha\supset L(ML\alpha\supset L\alpha)$ [H.Syll, (9)&(8)]
(11) $\neg ML\alpha\supset LM\neg\alpha$ [M definition, K and classical logic]
(12) $\neg ML\alpha\supset L(ML\alpha\supset L\alpha)$ [H.Syll, (11)&(10)]
(13) $L(ML\alpha\supset L\alpha)\supset ML(ML\alpha\supset L\alpha)$ [instance of $L\alpha\supset M\alpha\in D\subset T\subset S4$]
(14) $\neg ML\alpha\supset ML(ML\alpha\supset L\alpha)$ [H.Syll, (12)&(13)]
(15) $ML\alpha\supset ML(ML\alpha\supset L\alpha)$ [**DR3**, (4)]
(16) $\neg ML\alpha \vee ML\alpha$ [subst. inst. of exc. middle]
(17) $ML(ML\alpha\supset L\alpha)$ [Constructive dilemma,(14)&(15)&(16)]

MAIN THEOREM: If S5 $\vdash \alpha$ then S4 $\vdash ML\alpha$.

We prove by induction on the length of the minimal proof $\Pi(\alpha)$ of α in S5 that applying ML to all the formulae in $\Pi(\alpha)$ yields a sequence that can be enriched by interpolation to an S4-proof . If the length=1, then α is an S5 axiom. If it is also an S4 axiom, we can derive $ML\alpha$ in S4 according to Lemma 1. If it is an instance of the S5 Axiom Schema $ML\delta \supset L\delta$, then the proof is given by the last theorem provedabove.
Suppose the length of the minimal proof of α in S5 is n+1, where $n \geq 1$. Then α is not an axiom of S5. It must follow from previous formulae either by Modus Ponens(MP) or by
N(Necessitation). Apply ML to all the formulae of the proof . An MP transition will correspond then S4-transition by **MLMP** above. By Ind. Hyp. the ML-transformed premises are provable in S4. Therefore the ML-transformed conclusion must provable in S4.
A Necessitation inference (from S5 $\vdash \beta$ to S5 $\vdash L\beta$.) is transformed into (from S4 $\vdash ML\beta$ to S4 $\vdash MLL\beta$.) which can be enriched by interpolation to ($\vdash ML\beta$, $\vdash L\beta \supset LL\beta$, $\vdash ML\beta \supset MLL\beta$, $\vdash MLL\beta$). By Induction Hyp. S4 $\vdash ML\beta$, while the second and the third provability statements are an S4-axiom and an application to it of **DR3** (see above). S4 $\vdash MLL\beta$ follows from these by Modus Ponens.
Since the converse of the main theorem is trivially proved through the S5 Axiom Schema $ML\alpha \supset L\alpha$, and the fact that S5 contains S4, we have
COROLLARY: S5 $\vdash \alpha$ if and only if S4 $\vdash ML\alpha$.

Footnotes.

1. Prominently, Prior, Kripke, Hintikka, Lemmon, Cocchiarella and Dummet. [Cf. the appendices to Prior's *Past,Present and Future* (1967)]

2. Terms such as 'reverse semantics' are used tongue in cheek : We do not really believe that there should be any privileged direction in proceeding between formal-theories, on one hand, and structures in which we wish to anchor their semantics, on the other. It is interesting however that a strong consciousness of the reverse semantical direction, which is so germane to mathematics, occurs almost as an after-thought to some Modal Logicians. Thus, in their book *A Companion to Modal Logic* (84)—most of which is devoted to the relationship between Modal theories and frames—the idea of looking at the whole semantical enterprise from the opposite point of view occurs to the authors, G.E.Hughes and M.J.Cresswell, only on a single occasion(ch.3, p.47), in a section entitled **conditions not corresponding to any axiom.** It is only there that we find the question posed: *"But what, we may wonder, is the position about the reverse direction?"*. The ensuing discussion is scanty.

3. Model-Normality follows from *distinguishability* (see [Hughes & Cresswell,84, p.75]), and is also entailed by Strong Normality (It is equivalent to strong normality for finite models).

4. The principle of Modal continuity entails also the following: *In a Linear Model every discrete case is semantically distinguished from all others by a uniquely true formula.*

5. The proofs follows along the same lines as those which establish the relevant frame properties of Canonical models of said modal logics.

6. Cf. [Hughes G.E& Cresswell M.J , *An Introducton to Modal Logic*, 1968] p.289.

7. Cf. Kelly, J.L., *General Topology*, p.14.

8. Cf. [Hughes G.E& Cresswell M.J, *A Companion to Modal Logic*, 1984] p.30

9. Op.Cit. Corollary 5.10, p.81. This, however, does not work for S4.3.1.

10. Cf. Segerberg, K, An Essay in Classical Modal Logic, Uppsala 1971, pp.78-81; and also [Hughes and Cresswell, 84, p.84].

11. An $\omega+\omega^-$ frame, such as in $\mathcal{M}^*$ above, is discrete although it has one open cut.

12. Formally <W,R> is an *entropy maximiseing frame* iff $W=W_1 \cup W_2$, $W_1 \cap W_2 = \varnothing$, $(W_1 \cup (W_2 x W_2) \subset R$, and $<W_1, R\cap(W_1 x W_1)>$ is <u>linear</u>. (Notice that $<W_2, R\cap(W_2 x W_2)>$ is then a universal S5 frame).

13. This is related to sobociński's K systems. The E schema above is equivalent to the schema $Kb=\{LM\alpha \supset ML\alpha)\}$ in [Hughes & Cresswell, 68] p.265. The system E would seem to be K4.3.1

14. Essentially this shows that a non-trivial S5 model without"Occamian" redundancy can not be described as a union of linear submodels. Segerberg's method of bulldosing clusters (and replacing them with infinite linear stretches) is, in fact, an intuitive result of the above argument.

15. The proof is straightforward, using the above lemma and the techniques used in proving it. The denseness of the submodel follows from the fact (due to strong normality) that for any (topologically)closed interval [w,w'] in the original model there is a formula satisfied only within it. One uses then the elementary submodel property to show that the submodel could satisfy it only within the interval.

16. c.f [Chang& Keisler's *Model Theory* 73] pp. 113-121, and especially exercise 3.19

17. A prime set of models of T is a minimal one with repect to the property that every model of T contains a submodel isomorphic to one of its members.

18. Cf. [Gaifman,H.,1972, pp.128-144.] where it is proved that every extension of a non-standard model of Arithmetic can be obtained by taking a cofinal elementary extension and then an end extension. A very similar theorem to this theorem of Gaifman can be shown by our results to apply to models of S4.3.1.

19. Suppose $S=[\bigcap \{D \mid D \in UF(\omega) \ \& \ D \supset \mathbf{F}\}] - \mathbf{F}$, where **F** is a proper filter, is not empty. Let X be a member of S. Clearly X does not belong to **F**. Suppose $\omega - X$ belongs to **F** . Then it belongs to every ultra-filter that contains **F** , and therefore X does not belong to any of them, and *a fortiori* is not a member of their intersection and could not belong to S. (contradiction, derived fron the fact that for any ultrafilter over ω contains a subset of ω iff it does not contain its ω-complement). Thus neither X nor $\omega - X$ belong to **F** .
On the other hand, we must have a non-empty intersection $Y \cap (\omega - X)$ for every member Y of **F**, since otherwise we would have $Y \subseteq X$, for some Y in **F**, which would imply—by the proper filter properties—that X belongs to **F**. Thus the set $F \cup \{Y \cap (\omega - X) \mid Y \in F\}$ must have the *finite intersection property* and is extendible to an unltrafilter D^* over ω. Since $X \in D^*$ as a member of S, and since $(\omega - X) = \omega \cap (\omega - X) \in D^*$ by its construction, we get $\varnothing = X \cap (\omega - X) \in D^*$ (contradiction).

20. For any X in F, $Y \cap X \neq \varnothing$ or else $X \subseteq \omega - Y$ and by filter properties $\omega - Y$ belongs to F (contrad.) Hence prove that $F \cup \{Y \cap X \mid X \in F\}$ has the finite intersection property and can therefore be extended to a proper ultrafilter.

BIBLIOGRAPHY

Bull,R.A. *An Algebraic Study of Diodoran Modal Systems*, JSL,Vol 30, 1,pp.58-64, 1965

Chang,C.C.&Keisler,H.J. *Model Theory*, North Holland 1973

Gaifman, H., *A Note on Models and Submodels of Arithmetic*, pp. 128-144 in Conference in Mathematical Logic—London 1970, Springer Verlag 1972

Hughes,G.E & Cresswell, M.J, *An Introduction to Modal Logic*, Methuen 1968

Hughes,G.E & Cresswell, M.J, *A Companion to Modal Logic*, Methuen 1984

Kelly, J.L, *General Topology*, Van Nostrand 1955

Prior, A.N., *Past, Present and Future*, Oxford 1967

Segerberg, K, *An Essay in Classical Modal Logic*, Uppsala 1971

Wuensche,A. & Lesser, M., *The Global Dynamics of Cellular Automata*, Santa Fe Studies in the Science of Complexity, Addison-Wesley 1992