Journal of Philosophical Logic (2005) 34: 581–605
DOI: 10.1007/s10992-004-8539-5

HORACIO ARLÓ COSTA

NON-ADJUNCTIVE INFERENCE AND CLASSICAL MODALITIES

ABSTRACT. The article focuses on representing different forms of non-adjunctive inference as sub-Kripkean systems of *classical modal logic*, where the inference from $\Box A$ and $\Box B$ to $\Box A \wedge B$ fails. In particular we prove a completeness result showing that the modal system that Schotch and Jennings derive from a form of non-adjunctive inference in (Schotch and Jennings, 1980) is a classical system strictly stronger than **EMN** and weaker than **K** (following the notation for classical modalities presented in Chellas, 1980). The unified semantical characterization in terms of neighborhoods permits comparisons between different forms of non-adjunctive inference. For example, we show that the non-adjunctive logic proposed in (Schotch and Jennings, 1980) is not adequate in general for representing the logic of high probability operators. An alternative interpretation of the forcing relation of Schotch and Jennings is derived from the proposed unified semantics and utilized in order to propose a more fine-grained measure of epistemic coherence than the one presented in (Schotch and Jennings, 1980). Finally we propose a syntactic translation of the purely implicative part of Jaśkowski's system D_2 into a classical system preserving all the theorems (and non-theorems) explicilty mentioned in (Jaśkowski, 1969). The translation method can be used in order to develop epistemic semantics for a larger class of non-adjunctive (discursive) logics than the ones historically investigated by Jaśkowski.

KEY WORDS: classical modal logic, epistemic logic, high probability operators, paraconsistent logic, non-adjunctive logic

1. INTRODUCTION

Non-Adjunctive logical systems are those where the inference from A and B to $A \wedge B$ fails. As is indicated in (Priest and Tanaka, 2000) the first of these systems to be produced was also the first formal *paraconsistent logic*. This was Stanislaw Jaśkowski's *discussive* (or *discursive*) logic (Jaśkowski, 1969). The central idea in discussive logic is to formalize the process of reasoning from the pooled views of various rational agents, who might nevertheless disagree about the truth of various facts. Most applications of non-adjunctive inference are epistemically motivated.

Another salient, and perhaps better studied, example is the logic of monadic operators of high probability. Many authors have suggested that formalizing the logic of such operators requires the use of non-adjunctive inference (or the use of some form of paraconsistent formalism). Views pro and con are discussed in (Kyburg, 1995). A related, but slightly different argument proposes that 'it is highly probable that' should be formalized as an epistemic modal operator. It is quite obvious that $A \wedge B$ might fail to be highly probable, even when A and B are highly probable. Under this construal what fails is not the inference from A and B to $A \wedge B$, but the inference from $\Box A$ and $\Box B$ to $\Box A \wedge B$, where '$\Box$' is the monadic operator of high probability.

Formalizing this latter account possesses also significant challenges. In fact, the weakest of the system of modal logics endowed with a relational semantics, the system **K**, satisfies the schema:

(C) $(\Box A \wedge \Box B) \rightarrow \Box(A \wedge B)$

So, this suggests that studying a logical system where failures of this modal form of Adjunction occur requires a generalization of the standard Kripke semantics for modal operators. This generalization, even though less carefully studied than its relational counterpart, has indeed been proposed (independently) by Dana Scott (1970) and Richard Montague (1970). A systematic presentation of this semantics and of the systems of *classical modal logic* that correspond to them is offered in Part III of (Chellas, 1980) – this includes work first presented in (Segerberg, 1971). Brian Chellas calls this generalization of relational semantics *minimal models*. They are otherwise known as *neighborhood models*, and this will be the terminology adopted here. (Arló-Costa, 2002) proposes the use of the family of sub-relational *classical* modal logics in order to formalize epistemic operators where different failures of logical omniscience occur. In particular it is suggested the possibility of using some of these systems in order to model monadic operators of high probability. The logical focus of (Arló-Costa, 2002) is to study a first order extension of some of the classical modal systems weaker than **K**. Kyburg and Teng (2002) have focused on the propositional level and on applications considering high probability. They identify the logical system **EMN** as the one involved in representing high probability operators.[1] Classical systems can be introduced succinctly as follows (we will provide more background below):

DEFINITION 1.1. A system of modal logic is *classical* if and only if it contains the axiom $\Diamond A \leftrightarrow \neg\Box\neg A$, and is closed under the rule of inference **RE**, according to which $\Box A \leftrightarrow \Box B$ should be inferred from $A \leftrightarrow B$.

In addition **EMN** satisfies the axioms:

(M) $\Box(A \wedge B) \rightarrow (\Box A \wedge \Box B)$

as well as:

(N) $\Box$**true**

The weakest Kripkean system, the system **K**, is equivalent to **EMCN**, so **EMN** is one of the classical systems which fail to satisfy the modal counterpart of adjunction. It is not difficult to see that the remaining axioms and rules are naturally motivated for an operator of high probability.

One of the virtues of the previous account is the natural intuitiveness of the use of a modal operator in order to represent qualitative probability. Another virtue is the fact that the analysis can be carried out by using an extension of classical logic, without modifying the underlying notion of logical consequence. Of course, the analysis still requires using a generalization of relational semantics in order to understand the nature of the modal operator '$\Box$'.

There is, nevertheless, an intimate connection between non-adjunctive inference and relational modal operators (where (C) fails), which we intend to study in detail here. As has happened in other areas of philosophical logic, connections between non-standard logical systems and extensions of classical logic illuminate the nature of both (an example is given by the connections between intuitioninsm and the modal system S4).

Schotch and Jennings have offered in (Schotch and Jennings, 1980) one of the standard contemporary systems of non-adjunctive inference, and in the process of doing so, they also derived a modal system from the non-adjunctive notion of consequence (the *forcing* relation) used in their analysis. Nevertheless, the nature of this modal operator and its potential relationships with the epistemic $\Box$ axiomatized by **EMN** is not immediate. They offer a semantics, which they see as a generalization of Kripke semantics. One of my goals here is to show that their modal operator has neighborhood models of the type proposed by Scott and Montague. I shall provide closure conditions on neighborhoods that completely characterize Schotch and Jennings' modal operator. The semantics allows us to locate Schotch and Jennings' modal system as a classical system of modal logic stronger than **EMN** and weaker than **K**. The system in question has not been independently studied by modal logicians. I shall also study a natural strengthening of their logic, also weaker than **K**.

I shall show via examples that the resulting classical modal system is not adequate to represent operators of high probability. This, in turn, sheds some light concerning the nature of the forcing relation proposed by Schotch and Jennings. Even when the notion in question admits an epistemic interpretation, I shall argue that the interpretation in question is very different from the one required for monadic operators of high probability.

The nature of non-adjunctive inference in Schotch and Jennings' system, as well as various strengthenings studied here, seem to be more naturally related to the first systems developed by Jaśkowski.

I shall proceed as follows. First, I shall introduce the forcing relation of Schotch and Jennings as well as their derived modal operator. Then I shall enter the details of neighborhood semantics and I shall prove that the axiomatization of Schotch and Jennings' is complete with respect to the proposed neighborhood model. Once this is done we will use the models in question in order to show that operators of high probability need not meet the constraints on neighborhoods needed for Schotch and Jennings' operators (with the exception of some limit cases). Finally a strengthening of Schotch and Jennings' modal logic will be considered. I shall close with some philosophical remarks concerning the epistemic nature of the forcing relation and I shall utilize them in order to motivate a new measure of coherence of information.

2. Measures of Coherence and Non-Adjunctive Inference

The central idea behind Schotch and Jennings' notion of forcing is their proposal for measuring the coherence of a set of sentences. Their *coherence function c* is a function having as its domain the set of all finite sets of sentences and as a codomain the set $Nat \cup \{w\}$, where *Nat* is the set of natural numbers.

DEFINITION 2.1. For **false** $\notin \Gamma$, $c(\Gamma) = m$ if and only if m is the least integer such that there are sets

$a_1, \ldots, a_m$, with $a_i \nvdash$ **false** $(1 \leq i \leq m)$
and $\bigcup_{i=1}^{m} a_i = \Gamma$

where $\vdash$ is the classical notion of consequence and where $c(\Gamma) = w$ by convention if **false** $\in \Gamma$.

Now we can define a notion of derivability in terms of this notion of levels of coherence. The *forcing relation* $[\!\vdash$ is characterized as a relation between finite sets of sentences and sentences and defined as follows:

DEFINITION 2.2. For $c(\Gamma) = n(w)$, $\Gamma [\!\vdash A$ if and only if for every n-fold (w-fold) decomposition $a_1, \ldots, a_n$, of Γ, there is some i such that $a_i \vdash A$ $(1 \leq i \leq n(w))$.

The forcing relation obeys the following structural rules (as the classical notion of consequence $\vdash$):

(*Ref*) $A \in \Gamma \Rightarrow \Gamma \,[\!\vdash A$
(*Mon*) $\Gamma \,[\!\vdash A \Rightarrow \Gamma \cup \Delta \,[\!\vdash A$, when $c(\Gamma \cup \Delta) = c(\Gamma)$.
(*Trans*) $\Gamma \cup \{A\} \,[\!\vdash B$ and $\Gamma \,[\!\vdash A \rightarrow \Gamma \,[\!\vdash B$

We need an additional structural rule as well, which depends on the previous concept of m-cluster. For a any finite set we say that $C \subseteq 2^a$ is an *m-cluster* if and only if $m \in Nat$ and:

For all $f \in m^a$, there is $x \in C$, and there is $y \leq m$: $x \subseteq f^{-1}[y]$.

In words: for any way of dividing a into m subsets there is a member x of C such that x is included in at least one of the m subsets into which a has been divided.

(*Clus*) If $C = \{c_1, \ldots, c_n\}$ is an m-cluster constructed out of $A_1, \ldots, A_k \in \Gamma$ and $c(\Gamma) = m$, then $(c_1 \,[\!\vdash B, c_2 \,[\!\vdash B, \ldots, c_n \,[\!\vdash B) \Rightarrow \Gamma \,[\!\vdash B$

In addition we have the usual rules for introducing and eliminating connectives, with the notable exception that the rule for introducing conjunction only holds for sets Γ, such that $c(\Gamma) = 1$. In other words, from $\Gamma \,[\!\vdash A$ and $\Gamma \,[\!\vdash B$ it no longer follows that $\Gamma \,[\!\vdash (A \wedge B)$, unless $c(\Gamma) = 1$.

We can now introduce a generalization of the standard (logical) notion of *theory*. The most immediate definition is an obvious generalization of the classical notion:

DEFINITION 2.3. Δ is a m-theory if and only if $c(\Delta) = m$ and $\Delta \,[\!\vdash A$ entails $A \in \Delta$.

The notion of m-theory can also be expressed via two closure conditions, without appealing to a direct use of '$[\!\vdash$'.

DEFINITION 2.4. Δ is a m-theory if and only if

(a) $A \in \Delta$ and $\vdash A \rightarrow B$, entail that $B \in \Delta$
(b) If $\{c_1, \ldots, c_k\} \subseteq 2^\Delta$ is an m-cluster, then $\bigvee_1^k\{\wedge c_1, \ldots, \wedge c_k\} \in \Delta$, where $\wedge c_i$ denotes '$\wedge A_{i1}, \ldots, \wedge A_{ij}$' for $c_i = \{A_{i1}, \ldots, A_{ij}\}$.

This notion of m-theory, which generalizes the standard notion of theory, will be useful in order to introduce the necessity operator that Schotch and Jennings derive from the forcing relation. This derivation will be the focus of the next section.

3. NECESSITY DERIVED FROM THE FORCING RELATION

The first step in the derivation is to enlarge the language of the propositional calculus PC with a new connective $\Box$. A *pre-model* $\mathcal{B}$ for the enlarged language $PC(\Box)$ is a standard model (U, P) of a non-empty set U and a valuation P mapping the atoms of the language to events in 2^U. P is extended (uniquely) to a function $\|\cdot\|^{\mathcal{B}}$ which evaluates all sentences of the language by means of classical truth conditions. So, we have $\|A \wedge B\|^{\mathcal{B}} = \|A\|^{\mathcal{B}} \cap \|B\|^{\mathcal{B}}$, etc.

No specific truth conditions are introduced for '$\Box$' aside from stipulating that $\|\Box A\|^{\mathcal{B}} \in 2^U$, for all A. So, '$\Box$' is not really behaving as a logical constant at this stage. Still the $\Box$-operator can be used in order to 'guard' inconsistent formulae. So, even when $\|\{A, \neg A\}\|^{\mathcal{B}} = \emptyset$, we have $\|\{\Box A, \Box\neg A\}\|^{\mathcal{B}} \neq \emptyset$.

The second step in the derivation of a model for the $\Box$-operator will be to restrict the class of pre-models to a special subclass called *full pre-models*.

DEFINITION 3.1. $\mathcal{B}$ is a full $PC(\Box)$ pre-model if and only if, $\mathcal{B}$ is a pre-model and for all u such that $\Box(u)^{\mathcal{B}} = \{A \mid \models_u^{\mathcal{B}} \Box A\}$ is an n-theory, $a \subseteq \Box(u)^{\mathcal{B}}$, such that $a \nvdash$ **false**, $a \subseteq b$ and $b \nvdash$ **false**, then $\|b\|^{\mathcal{B}} \neq \emptyset$.

Now as a final step of the construction we derive the underlying structure of the desired model and the truth conditions for $\Box A$ form the restrictions imposed by $[\vdash$. This is done in two steps, the first of which is to define a n-natural relation.

DEFINITION 3.2. Let $\mathcal{B}$ be a full $PC(\Box)$ pre-model. For each $n \in Nat$ let r be a function $r : \{x \mid c(\Box(x)^{\mathcal{B}}) = n\} \to U^n$. Let u be an element of U such that $c(\Box(u)^{\mathcal{B}}) = n$. Further let $\Delta(u) = \{\delta \mid \delta : \Box(u)^{\mathcal{B}} \to n\}$ be the set of non-trivial n-fold decompositions of $\Box(u)^{\mathcal{B}}$.

Then $r(u) = \{\langle x_1, \ldots, x_n\rangle \mid x_i \in \|\delta^{-1}[i]\|^{\mathcal{B}}$ $(1 \leq i \leq n)$ for some $\delta \in \Delta(u)\}$. Finally if $\langle x_1, \ldots, x_n\rangle \in r(u)$ we write $uRx_1, \ldots, x_n$ and call R the n-natural relation of u.

Now we can prove the following theorem stating the desired truth conditions for the derived modal operator:

THEOREM 3.1 (Schotch and Jennings). *If $\mathcal{B}$ is a full $PC(\Box)$ pre-model and $\Box(u)^{\mathcal{B}}$ is an n-theory and R the n-natural relation, then $\models_u^{\mathcal{B}} \Box A$ if and only if for all $x_1, \ldots, x_n$, if $uRx_1, \ldots, x_n$, then $\models_{x_1}^{\mathcal{B}} A$ or ... or $\models_{x_n}^{\mathcal{B}} A$.*

Schotch and Jennings comment that their semantics is a generalization of Kripke semantics. Nevertheless, the presentation is non-standard and dependent of the notion of pre-model. It would be nice to see whether the proposed semantics can be classified in terms of some of the well-known generalizations of relational semantics. I shall focus on this topic in Section 4.

3.1. *The Logic K_n*

The goal of Schotch and Jennings is to show that the class of structures generated in the previous section determine a modal logic extending the non-adjunctive logic presented above. The logic in question is obtained by supplementing the axioms and rules constraining the $[\vdash$ relation with the following rule:

(RK_n) If $c(\Gamma) = n$, and $\Gamma [\vdash B$, then $\Box[\Gamma] \vdash \Box B$, where $\Box[\Gamma] = \{\Box A \mid A \in \Gamma\}$.

Schotch and Jennings built a canonical model for the resulting logic K_n, and they prove that the model is a full $PC(\Box)$ pre-model, satisfying the closure restriction used in the theorem presented above.

In the following sections I shall proceed as follows. First I shall provide some background about neighborhood models of modalities. Then I shall introduce a constraint on neighborhoods, called *clustering*, and I shall show that this constraint is a 'natural' semantic counterpart of the notion of m-theory. This introductory result might help connecting the new neighborhood structures with the ones built up by Schotch and Jennings. Then I shall prove a general representation result for K_n in terms of neighborhood models, which does not require using the full $PC(\Box)$ pre-models of Schotch and Jennings.

4. Neighborhood Models for Modalities

We will introduce here the basis of the so-called *neighborhood semantics* for propositional modal logics. We will follow the standard presentation given in Part III of (Chellas, 1980).

DEFINITION 4.1. $\mathcal{M} = \langle W, N, P\rangle$ is a neighborhood model if and only if:

(1) W is a set,
(2) N is a mapping from W to sets of subsets of W,
(3) P is a standard valuation mapping the atoms of the language to subsets of W.

Of course the pair $\mathcal{F} = \langle W, N\rangle$ is a *neighborhood frame*. The following definition makes precise the notion of truth in a model.

DEFINITION 4.2 (Truth in a neighborhood model). Let u be a world in a model $\mathcal{M} = \langle W, N, P\rangle$. P is extended (uniquely) to a relation $\models_u$ (where $\mathcal{M} \models_u$ A states that A is true in the model $\mathcal{M}$ at world u). The extension is standard for Boolean connectives. Then the following clauses are added in order to determine truth conditions for modal operators.

(1) $\mathcal{M} \models_u \Box A$ if and only if $\|A\|^{\mathcal{M}} \in N(u)$
(2) $\mathcal{M} \models_u \Diamond A$ if and only if $\|\neg A\|^{\mathcal{M}} \notin N(u)$

where, $\|A\|^{\mathcal{M}} = \{u \in W : \mathcal{M}, \models_u A\}$.

$\|A\|^{\mathcal{M}}$ is called A's *truth set*. Intuitively $N(u)$ yields the propositions that are necessary at u. Then $\Box A$ is true at u if and only if the 'truth set' of A (i.e. the set of all worlds where A is true) is in $N(u)$. If the intended interpretation is epistemic $N(u)$ contains a set of propositions understood as epistemically necessary. This can be made more precise by determining the exact nature of the epistemic attitude we are considering. $N(u)$ can contain the known propositions, or the believed propositions, or the propositions that are considered highly likely, etc. Then the set $P = \{\|A\|^{\mathcal{M}} \in 2^W : \models_u \Diamond A\}$ determines the space of epistemic possibilities with respect to the chosen modality – knowledge, likelihood, etc.

Clause (2) forces the duality of possibility with respect to necessity. It just says that $\Diamond A$ is true at u if the denial of the proposition expressed by A (i.e. the complement of A's true set) is not necessary at u. $N(u)$ is called the *neighborhood* of Γ.

4.1. *Augmentation*

The following conditions on the function N in a neighborhood model $\mathcal{M} = \langle W, N, P\rangle$ are of interest. For every world u in $\mathcal{M}$ and every proposition (set of worlds) X, Y in $\mathcal{M}$:

(m) If $X \cap Y \in N(u)$, then $X \in N(u)$, and $Y \in N(u)$.
(c) If $X \in N(u)$, and $Y \in N(u)$, then $X \cap Y \in N(u)$.
(n) $W \in N(u)$.

When the function N in a neighborhood model satisfies conditions (m), (c) or (n), we say that the model is *supplemented*, is *closed under intersections*, or *contains the unit* respectively. If a model satisfies (m) and (c) we say that it is a *quasi-filter*. If all three conditions are met it is a *filter*.

Notice that filters can also be characterized as non-empty quasi-filters – non-empty in the sense that for all worlds u in the model $N(u) \neq \emptyset$.

DEFINITION 4.3. A neighborhood model $\mathcal{M} = \langle W, N, P \rangle$ is *augmented* if and only if it is supplemented and, for every world u in it:

$$\bigcap N(u) \in N(u).$$

Now we can present an observation (established in Chellas, 1980, Section 7.4), which will be of heuristic interest in the coming section.

OBSERVATION 4.1. *$\mathcal{M}$ is augmented just in case for every world u and set of worlds X in the model:* (a) *$X \in N(u)$ if and only if $\bigcap N(u) \subseteq X$.*

It easy to see that every augmented model is a filter: supplemented, closed under intersections and possessed of the unit. Moreover, every finite filter is augmented. This suggests a tight relationships between neighborhood and Kripke models: a Kripke model is essentially an augmented neighborhood model.

4.2. *Epistemic Interpretation of Augmentation*

In recent work in epistemic logic it is quite usual to represent agents by *acceptance sets* or *belief sets*, obeying certain rationality constraints. If the representation is linguistic the agent is represented by a logically closed set of sentences. If the representation is done in a σ-field or relative to a universe of possible worlds, the agent is represented by a set of points such that all propositions accepted (believed) by the agent are supersets of this set of points. Adopting either representation is tantamount to imposing logical omniscience as a rationality constraint.

When a neighborhood frame is augmented we have the guarantee that, for every world u, its neighborhood $N(u)$ contains a smallest proposition, composed of the worlds that are members of every proposition in $N(u)$. In other words, for every u we know that $N(u)$ always contains $\bigcap N(u)$ and every superset thereof (including W).

We will propose to see the intersection of the neighborhood of a world as an acceptance set for that world, obeying the rationality constraints required by logical omniscience. The following results help to make this idea more clear.

OBSERVATION 4.2 *If $\mathcal{M}$ is augmented, then for every u in the model: $\models_u \Box A$ iff and only if $\bigcap N(u) \subseteq \|A\|^{\mathcal{M}}$, and $\models_u \neg\Box A$ iff and only if $\bigcap N(u) \not\subseteq \|A\|^{\mathcal{M}}$.*

Epistemic possibility is, in this setting, understood in terms of compatibility with the belief set $\bigcap N(u)$. In other words $\models_u \Diamond A$ if and only if $\|A\|^{\mathcal{M}} \cap (\bigcap N(u)) \neq \emptyset$. This in turn means that, when the model $\mathcal{M}$ i s augmented, $\models_u \Diamond A$ holds whenever $\|A\|^{\mathcal{M}}$ is logically compatible with *every* epistemically necessary proposition in the neighborhood.

This epistemic interpretation of augmentation can be extended to the case of neighborhoods that are not augmented. The basic idea is to extend the previous account even for inconsistent neighborhoods with empty intersection:

DEFINITION 4.4 (Poss). $\mathcal{M} \models_u \Diamond A$ if and only if for every X in $N(u)$, $\|A\|^{\mathcal{M}} \cap X \neq \emptyset$.

The central idea being that an (unclosed) inconsistent set of statements can be used in order to establish what is possible as follows: if a statement contradicts *a member* of the set, then it is not possible. For certain standard of quality control, that one of the inspected pieces is not OK is not a serious possibility. And at the same time, it is not a serious possibility that all the inspected pieces are OK. A detailed analysis of the logical consequences of adopting this extended notion of epistemic possibility for first order languages, as well as some consequences concerning the lottery paradox, is presented in (Arló-Costa, 2002).

4.3. *The Level of Coherence of Neighborhoods*

The previous remarks bring us directly to the fact that most of the non-normal classical models will contain inconsistent neighborhood models. We can measure the level of coherence of these neighborhoods as we can measure the level of coherence of a set of sentences.

DEFINITION 4.5. A set of propositions N has level of coherence m if and only if m is the least integer such that there is a sequence of sets of propositions $X_1, \ldots, X_m$ where each of these sets is in N, such that $\emptyset \neq \bigcap X_i$ and $\bigcup_i X_i = N$. Each of the sequences $X_1, \ldots, X_m$ will be called an m-decomposition of N.

This is a straightforward adaptation of the ideas of Schotch and Jennings presented above. We can add a bit of useful notation here:

DEFINITION 4.6. If a set of propositions N has degree of coherence m, and $X_1, \ldots, X_m$ is an m-decomposition δ of N, then the sets $G_1 = \bigcap X_1$, $\ldots$, $G_m = \bigcap X_m$, are called a set of m-generators of $N\delta$.

Now the following closure condition on neighborhoods, which we can call *clustering* is of interest:

DEFINITION 4.7. A neighborhood model $\mathcal{M} = \langle W, N, P\rangle$ is *m-clustered* if and only if for every $u \in W$, and for every $X \subseteq 2^W$, if $N(u)$ has level of coherence m.

$X \in N(u)$ if and only if for all generators $G_1, \ldots, G_m$ for $N(u)$, either $G_1 \subseteq X$, or ..., or $G_m \subseteq X$.

We will say that a neighborhood of level of coherence m in a clustered model is m-clustered. It is not difficult to see that a clustered neighborhood is supplemented and possesses the unit, even though it need not be closed under intersections. The next section will be devoted to show that clustering is the counterpart for neighborhoods of the syntactic notion of m-theory.

5. From Pre-Models to Neighborhood Models

Let $\mathcal{N} = \langle U, N, P\rangle$ be a neighborhood model. We can then prove the following result about clustering:

THEOREM 5.1. *Let $\mathcal{N}$ be a neighborhood model and let $\Box(u)^{\mathcal{N}} = \{A \mid \mathcal{N} \models_u \Box A\}$. $\Box(u)^{\mathcal{N}}$ is an m-theory if and only if $N(u)$ is m-clustered.*

Proof. Assume that $X = \|A\|^{\mathcal{N}} \in N(u)$. Then, by the truth conditions of the $\Box$-operator, $A \in \Box(u)^{\mathcal{N}}$. Therefore, $\Box(u)^{\mathcal{N}}$ $[\vdash A$ (by [*Ref*]). So, for all m-decompositions δ of $\Box(u)^{\mathcal{N}}$, there is i, $\|\delta^{-1}[i]\|^{\mathcal{N}} \subseteq \|A\|^{\mathcal{N}}$.

Since $\Box(u)^{\mathcal{N}}$ is an m-theory, $N(u)$ has level of coherence m and all m-generators of $N(u)$ are given by the sets $\|\delta^{-1}[1]\|^{\mathcal{N}}, \ldots,]\|\delta^{-1}[m]\|^{\mathcal{N}}$ for each δ. So, we have that for all sets of m-generators $G_1, \ldots, G_m$ for the neighborhood, there is G_i in the set entailing X, and this is enough to establish the LTR part of the proof.

For the RTL part of the proof assume that $X = \|A\|^{\mathcal{N}} \notin N(u)$. Then we have that $A \notin \Box(u)^{\mathcal{N}}$ and since $\Box(u)^{\mathcal{N}}$ is, by hypothesis, a m-theory, $\Box(u)^{\mathcal{N}}$ $[\nvdash A$. Therefore there is $\delta \in \Delta(u)$: for all i $(1 \leq i \leq n)$ $\|\delta^{-1}[i]\|^{\mathcal{N}} \not\subseteq \|A\|^{\mathcal{N}}$. For the reasons invoked in the first part of the proof, this means that there is a set of m-generators of $N(u)$, $\|\delta^{-1}[1]\|^{\mathcal{N}}, \ldots, \|\delta^{-1}[n]\|^{\mathcal{N}}$ such that neither of them entails $\|A\|^{\mathcal{N}}$. This completes the proof. □

The proof is more direct than a similar proof in terms of the m-natural relation, offered by Schotch and Jennings. The reason for this directness is the fact that m-clustering seems to be a more 'natural' semantic counterpart for the syntactic notion of m-theory.

6. Completeness in Terms of Canonical Neighborhood Models

A direct completeness proof in terms of canonical neighborhood models proceeds as follows. The construction of the canonical model is standard. Let Σ be a system of classical modal logic. Then let $Max_\Sigma\, \Gamma$ denote a maximal and consistent set of sentences of Σ. Let p_n denote the atoms of the language. In addition we have the notation $|A|_\Sigma = \{Max_\Sigma \mid A \in \Gamma\}$, where $|A|_\Sigma$ is A's *proof set* for the system Γ. The canonical model $\mathcal{N} = \langle W, N, P\rangle$ is built up as follows:

(1) $W = \{\Gamma \mid Max_\Sigma\, \Gamma\}$.
(2) For all $u \in \mathcal{N}$, $\Box A \in u$ if and only if $|A|_\Sigma \in N(u)$.
(3) $P_n = |p_n|_\Sigma$, for $n = 0, 1, \ldots$.

I shall not repeat here the main results about canonical models for classical systems, which can be found in (Chellas, 1980). The result that interest us in order to prove a determination result for the logic K_n is the following one:

THEOREM 6.1. *Let $\mathcal{N} = \langle W, N, P\rangle$ be the smallest canonical neighborhood model for the classical system containing the rule RK_n. Then for every u in $\mathcal{N}$ $N(u)$ has degree of coherence n if and only if $N(u)$ is n-clustered.*

Proof. Let Σ be a system of classical logic containing the rule RK_n and let $\mathcal{N}$ be the smallest canonical neighborhood model for Σ, i.e. a model such that $N(u) = \{|A|_\Sigma \mid \Box A \in u\}$. Assume for arbitrary u that $N(u)$ has degree of coherence n. Assume in addition that $X \in N(u)$. Then, $X = |A|_\Sigma$, for $\Box A \in u$.

Since $\Box A \in u$ we have that $A \in \Box(u)^{\mathcal{N}} = \{A \mid \Box A \in u\}$. This indicates, by Reflexivity of $[\vdash$, that $\Box(u)^{\mathcal{N}} [\vdash A$. So, for all n-decompositions δ of $\Box(u)^{\mathcal{N}}$, there is i, $\delta^{-1}[i] \vdash A$. Now, all the n-generators of $N(u)$ are the sets:

$|\delta^{-1}[1]|_\Sigma, \ldots, |\delta^{-1}[n]|_\Sigma$ for some decomposition δ of $\Box(u)^{\mathcal{N}}$.

So, we know that for any arbitrary set of n-generators $G_1, \ldots, G_n$ for $N(u)$ where:

$G_1 = |\delta^{-1}[1]|_\Sigma, \ldots, G_n = |\delta^{-1}[n]|_\Sigma$ for some decomposition δ of $\Box(u)^{\mathcal{N}}$

there should be i, such that $G_i = |\delta^{-1}[i]|_\Sigma \subseteq |A|_\Sigma$.

On the other hand, if we assume that for all n-generators $G_1, \ldots, G_n$ of $N(u)$ there is at least one G_i entailing $|A|_\Sigma$; this is tantamount to assume that $\Box(u)^{\mathcal{N}}$ [$\vdash A$. So, by the rule RK_n, $\Gamma' = \{\Box B \mid \Box B \in u\} \vdash \Box A$. Therefore $\Box A \in u$, which entails that $|A|_\Sigma \in N(u)$, as needed. □

7. High Probability Neighborhoods and Clustering

Let $\mathcal{N}_P = \langle U, N_P, V\rangle$ be a high n-probability model, where U is the universe, V a valuation and P a probability function defined on a Boolean sub-algebra of the power set of U. In addition, N_P is defined as follows:

DEFINITION 7.1. $N_P(u) = \{X \mid P(X) \geq n\}$.

As we reported above, it is clear from the work of Kyburg and Teng (2002) that high probability models are supplemented and possess the unit, and they are not closed under intersections. In addition, we can apply Schotch and Jennings' ideas here by measuring the coherence of high probability neighborhoods.

Some salient cases are immediate. High n-probability neighborhoods which contain a point $w \in U$ such that $P(\{w\}) \geq n$ are augmented with $\bigcap N(u) = \{w\}$. Nevertheless, high probability neighborhoods are not in general clustered:

EXAMPLE 7.1. Consider a .6-probability neighborhood where U contains four points, and let $P(\{w_1\}) = .5$, $P(\{w_2\}) = .1$, $P(\{w_3\}) = .3$, $P(\{w_1\}) = .1$. This neighborhood has level of coherence 1 with only one generator in $\{w_1\}$. Nevertheless, the neighborhood is not clustered, because $\{w_1\} \notin N(u)$.

The example gives some hints about the nature of the notion of clustering itself and about the nature of the non-adjunctive logics based on it. There are, of course, a variety of possible epistemic interpretations for the notion. I am interested here in the fact that clustering admits an interpretation, which is compatible with endorsing the most demanding standards of epistemic rationality in terms of logical closure. Here is the basis of this account of clustering (comparisons with the logic of high probability will flow naturally from the interpretation itself).

The gist of the idea is to see clustering as an account of epistemic indeterminacy prompted by data, which can be incoherent. An alternative example will help motivating the idea.

EXAMPLE 7.2 Consider a .6-probability neighborhood where U contains four points, and let $P(\{w_1\}) = .4$, $P(\{w_2\}) = .3$, $P(\{w_3\}) = .2$, $P(\{w_1\}) = .1$. This neighborhood has level of coherence 2. Possible generators include $G_1 = \{w_2\}$, $G_2 = \{w_3\}$; $G_1 = \{w_1\}$, $G_2 = \{w_2, w_3\}$; $G_1 = \{w_1\}$, $G_2 = \{w_4, w_2, w_3\}$; $G_1 = \{w_1\}$, $G_2 = \{w_2\}$; $G_1 = \{w_1\}$, $G_2 = \{w_3\}$.

A *cluster* in the sense of Schotch and Jennings is: $C_1 = \{w_1, w_2\}$, $C_2 = \{w_1, w_3\}$, $C_3 = \{w_4, w_2, w_3\}$. So we have that $N(u) = \{X \mid \text{ either } C_1 \subseteq X, \text{ or } C_2 \subseteq X, \text{ or } C_3 \subseteq X\}$.

Even when the last example was generated by utilizing high probability, I ask the reader to abstract from that fact and to just consider the data in the neighborhood as a possible data set independently of its origin. An agent facing the set of possible 2-decompositions of the data can be seen as being in doubt between various ways of articulating the data as the pooled knowledge of two consistent, but unclosed views. So the idea of forcing can be articulated as a form of cautious inference, where one should be committed to infer something from the data as long as it follows from any of the possible manners of articulating the data. So, for example, one of the generators will be in the neighborhood as long as it is a conclusion inferable from all possible articulations of the data.

So, clustering can be seen as a condition which requires the maximum degree of logical perfection as is compatible with the degree of indeterminacy represented in the neighborhood. If, as in the first example, there is no degree of indeterminacy, and the level of coherence is one, then the agent should be logically omniscient, i.e. the neighborhood should be augmented. Of course, this is a requirement which clashes with the logic of high probability, which permits augmentation only in some limit cases, but that establishes its own standard of rationality, not necessarily coherent with logical closure (seen as an ideal of rationality).

8. A More Fine-Grained Measure of Coherence?

The previous epistemic account of clustering suggests, in turn, that Schotch and Jennings' measure of coherence might not be fine-grained enough to reflect an intuitive notion of coherence in terms of epistemic determinacy. It seems that there are cases where neighborhoods, which intuitively bear different degree of coherence, receive nevertheless the same measure. For example, consider the following class of neighborhoods of level of coherence n:

DEFINITION 8.1. A neighborhood $N(u)$ is *closed under a set of m-generators* $G_1, \ldots, G_m$ if and only if $N(u)$ can be represented as $\{X \subseteq 2^U \mid G_1 \subseteq X\} \cup \cdots \cup \{X \subseteq 2^U \mid G_m \subseteq X\}$.

It is clear that neighborhoods closed under m-generators have level of coherence m. Nevertheless, of two neighborhoods of level of coherence m it seems that if one is closed under generators and the other is not, the one which is closed is epistemically more determinate (and intuitively more coherent) than the one which is unclosed. For example, consider the following modification of our second example:

EXAMPLE 8.1 Consider a neighborhood $N(u)$ which is closed under the generators $G_1 = \{w_2\}$, $G_2 = \{w_3\}$ in Example 7.2.

Certainly there are grounds here in order to make a logical distinction. The models whose neighborhoods are closed under generators are a strict subclass of the clustered models. Syntactically the requirement can be expressed by constraining further the forcing relation:

(CG) If $a_1, \ldots, a_n$ is an n-decomposition of Γ, then Γ [$\vdash \wedge a_i$, for all i, $1 \leq i \leq n$.

A possible improvement on the measure of coherence we have been using can be to define:

DEFINITION 8.2. For consistent Γ, $c(\Gamma) = m.n$ if and only if m is the least integer such that there are sets

$$a_1, \ldots, a_m, a_i \nvdash \textbf{false} \ (1 \leq i \leq m)$$

$$\text{and } \bigcup_{i=1}^{m} a_i = \Gamma$$

where $\vdash$ is the classical notion of consequence, $c(\Gamma) = w$ if **false** $\in \Gamma$, and where n is the number of possible m-decompositions of Γ.

So, according to this proposal, the coherence of the neighborhood in Example 7.1 is 1; the coherence of the neighborhood in Example 7.2 is 10; and the coherence of the neighborhood in Example 8.1 is 2.

9. CLOSURE UNDER GENERATORS

Neighborhood semantics is a generalization of relational semantics. So, we can obtain the entire normal family of modal systems by adding appropriate constraints on neighborhoods. Here is a list of those constraints for the standard schemas D, T, B, 4 and 5 (in epistemic logic 4 is called KK or positive introspection, etc.). The constraints are conditions on a model $\mathcal{M} = \langle U, N, P \rangle$, for every world u and proposition X in $\mathcal{M}$.

(d) If $X \in N(u)$, then $X^c \notin N(u)$.
(t) If $X \in N(u)$, then $u \in N(u)$.
(b) If $u \in N(u)$, then $\{w \in \mathcal{M} \mid X^c \notin N(w)\} \in N(u)$.
(iv) If $X \in N(u)$, then $\{w \in \mathcal{M} \mid X \in N(w)\} \in N(u)$.
(iv) If $X \notin N(u)$, $\{w \in \mathcal{M} \mid X \notin N(w)\} \in N(u)$.

Now consider the following constraint on neighborhoods:

DEFINITION 9.1 (J-Σ). For every $u \in U$, $Y \in 2^W$, $N_Y(u) = \bigcup\{N(w) \mid w \in Y \subseteq U$ and $N(w)$ is a neighborhood in a model of a classical system $\Sigma\}$.

When Σ is any normal Kripkean system, (J-Σ) seems to model constraints inspired by Jaśkowski's ideas (the next section provides an introduction to some of these ideas).

(J-Σ)-neighborhoods represent the sum of the beliefs of n participants in a 'discussion' (where n is the cardinality of Y). The set of beliefs of each participant $w \in Y$ is represented by the propositions in N_w. The standards of rationality of participants are given by the constraints imposed by Σ.

Notice that as long as Σ is any normal system, a model closed under (J-Σ) is clustered, closed under generators and of level of coherence n, where n is the number of participants. In fact, since in this case we have $\mathbf{K} \subseteq \Sigma$, the corresponding model of Σ is augmented, with $\emptyset \neq \bigcap N(w)$ for each $N(w) \in N_Y(u)$. Therefore each $\bigcap N(w)$ is a generator.

There is a hierarchy of models closed under generators which obey (J-Σ) for the various possible normal Σ. These systems are interesting per se independently of their eventual connection with the systems historically proposed by Jaśkowski's (a concrete connection with one of Jaśkowski's systems is studied below).

10. Jaśkowski's System D_2

The first non-adjunctive system in the literature was proposed by Jaśkowski in an article published in 1948 (Jaśkowski, 1948), later reprinted in 1969 in *Studia Logica* (Jaśkowski, 1969). Jaśkowski was interested in studying *contradictory* deductive systems including at least two theses that contradict each other. He was aware that any contradictory system based on a two-value logic is trivial or, as he preferred to call it: over-complete, in the sense that every meaningful formula is a thesis of the system. This is so in virtue of the following law;

(Over Completeness) $A \to (\neg A \to B)$

So, Jaśkowski presented the problem of studying contradictory systems in the following terms:

The task is to find a system of the sentential calculus which: (1) when applied to the contradictory systems would not always entail their over-completeness, (2) would be rich enough to enable practical inference, (3) would have an intuitive justification (Jaśkowski, 1969, page 145).

Kolmogorov's system (Kolmogorov, 1924) based on the four Hilbert's axioms of positive logic plus:

(K9) $(A \to B) \to ((A \to \neg B) \to \neg A)$

was one of the known systems in the literature accessible to Jaśkowski which failed to satisfy Over-completeness, but Kolmogorov's system when applied to a contradictory system produced undesirable results (every formula beginning with the symbol of negation is a thesis). So, Jaśkowski developed a new formalism whose intended interpretation diverged from systems like Kolmogorov's (or other multi-valued systems proposed at the time).

The main idea of a discursive system is simple. Jaśkowski invites the reader to consider a deductive system pooling the theses advanced by several participants in a discourse. So, Jaśkowski proposes: 'Let such a system which cannot be said to include theses that express opinions in agreement with one another, be termed a *discursive* system' (Jaśkowski, 1969, page 149). Moreover he immediately remarks:

To bring out the nature of the theses of such a system [a discursive system] it would be proper to precede each thesis by the reservation: "in accordance with the opinion of one of the participants of the discourse" [...] Hence the joining of a thesis to a discursive system has a different intuitive meaning than has assertion in an ordinary system (Jaśkowski, 1969, page 149).

And he proposes to encode this epistemic attitude via the modality usually reserved to encode the notion of possibility (metaphysical or epistemic):

Discursive assertion includes an implicit reservation of the kind specified above, which – out of the functions so far introduced in this paper – has its equivalent in possibility *Pos*. Accordingly, if a thesis α is recorded in a discursive system, its intuitive sense ought to be interpreted so as if it were preceded by the symbol *Pos*, that is, the sense "it is possible that α" (Jaśkowski, 1969, page 149).

It is clear that Jaśkowski proposes this representation of discursive assertion for instrumental reasons even when it reverses the usual practice in epistemic logic of reserving $\Box$ for the representation of the main attitude (knowledge, belief, etc.) and $\Diamond$ for the encoding of possibility. His discussion of *discursive implication* shows in an even more clear manner that he is trying to utilize known modal notions in order to represent a notion of implication that does not obey Over-completeness. He considers and discards various modal translations of a discursive connective $\rightarrow_d$. For example the translation of $A \rightarrow_d B$ as $\Diamond(A \rightarrow B)$, which he discards for not obeying elemental patterns of inference like modus ponens. Finally he notices that the translation of $A \rightarrow_d B$ as $\Diamond A \rightarrow B$, preserves patterns of inference that he considers desirable.

There seems to be a tension in Jaśkowski's writings between intuitive and instrumental adequacy. In fact, as he notices, the straightforward translation of $A \rightarrow_d B$, in accordance with the proposed epistemic reading of $\Diamond$ would be: "if anyone states that A, then B". But it is unclear how B should be read in this interpretation of $A \rightarrow_d B$. The literal translation would be: "if anyone states that A, then B is the case". Nevertheless this interpretation should be attentive to the fact that each assertion (including conditional assertions) are preceded by $\Diamond$. So, for example, modus ponens works because for a (Kripkean) modality the following is a theorem:

(MP) $\Diamond((\Diamond A \rightarrow B) \wedge \Diamond A) \rightarrow \Diamond B)$

So, the assertion of $A \rightarrow_d B$ should be intuitively interpreted along the following lines: "In accordance with the opinion of one of the participants of the discourse, if anyone states that A, then B," which nests various epistemic modalities.

Max Urchs presents in (Urchs, 1970) a characterization of Jaśkowski's system D_2 along lines suggested by Jaśkowski' in (Jaśkowski, 1969).[2] Let FOR_d be the set of formulas freely generated from a denumerable set of atomic propositions by means of some boolean complete set of connectives and two additional connectives: the discussive conjunction $\wedge_d$ and the discussive implication $\rightarrow_d$. Then Urchs provides a syntactic translation $t : FOR_d \rightarrow FOR_m$ from FOR_d to propositional modal language FOR_m with identical set of atoms AT. Let $\Diamond$ be S5-possibility. The translation proceeds as follows:

(1) $t(H) =_{df} H$, for $H \in AT$.

(2) $t(\neg H) =_{df} \neg t(H)$.

(3) $t(H \wedge G) =_{df} t(H) \wedge t(G)$.

(4) $t(H \wedge_d G) =_{df} t(H) \wedge \Diamond t(G)$.

(5) $t(H \rightarrow_d G) =_{df} \Diamond t(H) \rightarrow t(G)$.

Urchs, commenting on the translation, says that 'it may appear somewhat bewildering'. In fact, the role of the translation is purely instrumental.

DEFINITION 10.1. $D_2 =_{df} \{H \in FOR_d : \Diamond t(H) \in S5\}$.

The idea is to characterize something less familiar (D_2) in terms of something more familiar, namely the system S5. Since then recent work in paraconsistent logic has followed Jaśkowski's practice of characterizing D_2 in terms of Kripkean modalities (da Costa and Dubikajtis, 1977), in particular via the appeal to the system S5.

At this point it should be clear to the reader why D_2 is non-adjunctive. The reason is that the following formula is not a thesis of D_2:

(Ad) $A \rightarrow_d (B \rightarrow_d (A \wedge B))$

in virtue of the fact that the corresponding translated formula utilizing S5-possibility fails to be a theorem in S5. The final part of this article will be devoted to show that there are alternative translations of D_2 into classical systems of modal logic. I think that this alternative translation reveals in a more clear manner the epistemic nature of Jaśkowski's ideas. In a nutshell, I think that the translation to non-normal classical systems has not only instrumental but also foundational value, by suggesting possible epistemic reconstructions of non-adjunctive inference. And, of course, representing D_2 in this way is part of my attempt here of showing that various well-known non-adjunctive systems have counterparts in corresponding systems of classical modalities which are non-normal.

I shall focus on the purely implicative part of D_2 and therefore I shall only present a sketch of a complete translation (which should include as well at least discursive equivalence or discursive conjunction). The emphasis here is on the nature of the required translation, which can also be used (by adjusting parameters) to produce other discursive systems (in the broad sense of the Jaśkowskian notion) other than the ones historically proposed by Jaśkowski.

We need first some notational distinctions about the underlying language. Let L be the language obtained by adding a discursive implication to the purely Boolean language L_0. Let, in addition, L_B be $L - L_0$, and finally let $L_{\rightarrow}$ be the purely implicative fragment of the language, containing formulae $A \rightarrow_d B$, where A and B are either Boolean or conditionals. In order to represent a discussion it would be useful in addition to have a set agents I participating in the discussion and the duals operators $\Box_i$, $\Diamond i$, for $i \in I$. $\Box_i A$ carries the intuitive meaning: "agent i stated A" or "Agent i believes A". We also need two 'group' operators:

DEFINITION 10.2.

$\Box A =_{df} \exists i \Box_i A$
$\Diamond A =_{df} \forall i \Diamond_i A$

The definition intends to capture the intuitive reading of $\Box$ proposed by Jaśkowski: "A has been stated by someone in the discussion" (or the simpler and weaker reading that will be proposed below: "there is an agent in a group that believes that A"). The dual modal notion of epistemic possibility states that A is possible as long as it is epistemically possible for each agent in the discussion. Now we can utilize these operators in order to propose a modal translation for $A \rightarrow_d B$. The idea is to translate the discursive conditional as $\Box_i A \rightarrow \Diamond_i B$, i.e. "$B$ is epistemically possible for some agent i, given that he believes A". This basic idea can be made more precise via the following translation:

(1) $t(H) =_{df} \Box_i H$, for $H \in AT$.
(2) $t(H \wedge G) =_{df} t(H) \wedge t(G)$.
(3) $t(\neg H) =_{df} \Box_i \neg H$, for $H \in AT$.
(4) $t(\neg H) =_{df} \neg t(H)$, for $H \in L - AT$.
(5) $t(H \rightarrow_d G) =_{df} \mathbf{t}(H) \rightarrow T(G)$.

where we have the following additional definitions:

$\mathbf{t}(A) = t(A)$ if $A \in L_B$ and $\mathbf{t}(A) = \Box_i A$ otherwise.
$T(B) = t(B)$ if $B \in L_B$ and $T(B) = \Diamond_i B$ otherwise.

Let's consider some examples in order to see how the translation works.

$(A \rightarrow_d B) \wedge \neg B) \rightarrow_d \neg A$

The translation of this formula is:

$((\Box_i A \rightarrow \Diamond i B) \wedge \Box_i \neg B) \rightarrow \Diamond_i \neg A$

Intuitively the translated formula says that if agent i believes $\neg B$ but B is possible for i if A is believed, then i considers that $\neg A$ is possible. The central thesis $A \rightarrow_d A$ follows as long as our operators $\Box_i$ are constrained by:

(D) $\Box_i A \rightarrow \Diamond_i A$

In other words, the neighborhoods for these operators are *consistent* in the sense that we should have:

(d) If $X \in N_i(u)$, then $X^c \notin N_i(u)$

The D_2 theorem $(A \wedge B) \rightarrow_d A$, as well as the law of importation follow as long as the operator $\Box_i$ obeys:

(M) $\Box_i(A \wedge B) \rightarrow (\Box_i A \wedge \Box_i B)$

Some D_2 theses follow straightforwardly in virtue of the definitions and the duality of each modal operator involved in the translation. Examples of theses of this sort considered by Jaśkowski are:

$(\neg A \rightarrow_d A) \rightarrow_d A$
$(A \rightarrow_d \neg A) \rightarrow_d \neg A$

The following D_2 theorem imposes another important constraint on the underlying modal logics for agents:

$(\neg A \rightarrow_d (B \wedge \neg B)) \rightarrow_d A$

Here the crucial constraint is:

(N) $\Box_i$**true**

The translation offered so far is still insufficient to meet some of the basic conditions of adequacy imposed by Jaśkowski. For example, modus ponens does not hold. Nevertheless the use of additional constraints which apparently were also required in Jaśkowski's system[3] guarantee not only modus ponens but also the following additional theses of D_2:

$(A \rightarrow_d B \wedge A \rightarrow_d \neg B) \rightarrow_d \neg A$
$(\neg A \rightarrow_d B \wedge \neg A \rightarrow_d \neg B) \rightarrow_d A$

These theses hold as long as the epistemic modalities for each agent are constrained by the following *saturation* condition:

(S) $\Diamond_i A \rightarrow \Box_i A$

The logic imposed on the $\Box_i$ operators (and their duals) is still rather weak. In particular condition (C) need not be satisfied, and therefore the

EMNDS system does not have relational models. In order to see this just consider a neighborhood model with a finite domain $U = \{w_1, w_2, w_3\}$ and such that there is a neighborhood containing exactly the set U and all the propositions $\{x, z\}$ where $x \neq z$ and x, z range over the given universe. Condition (C) is violated although all axioms in *EMNDS* are satisfied.

This previous analysis suggests the interests of focusing on the following translation:

DEFINITION 10.3. $D =_{df} \{H \in L : t(H) \in EMNDS\}$.

D contains all the purely implicational theorems mentioned by Jaśkowski in (Jaśkowski, 1969). By the same token all the non-theorems explicitly mentioned in (Jaśkowski, 1969) do not belong to D (here the failure of (C) is important):

$((A \rightarrow_d B) \wedge (A \rightarrow_d \neg B)) \rightarrow_d C$
$(A \rightarrow_d (\neg A \rightarrow_d B)$
$(A \rightarrow_d (\neg A \rightarrow_d \neg\neg A)) \rightarrow_d B$
$(A \rightarrow_d (B \rightarrow_d (A \wedge B))$

It is easy to see that the fact that the $\Box_i$ operators obey the laws of *EMNDS* entails that $\Box$ obeys the laws of *EMNS* – the condition D for the group operators requires that $\Box A \rightarrow \Diamond A$, which is not entailed by $\Box_i A \rightarrow \Diamond_i A$.[4]

This also indicates that we can translate formulae of D_2 directly to the modal logic *EMNS* containing only the operators $\Box$ and $\Diamond$, rather than to a multi-modal logic. The idea, in virtue of the previously stated facts is to translate $A \rightarrow_d B$ directly to $\Box A \rightarrow \Box B$.

I conjecture that the translation sketched above (or some slight modification of it to accommodate other connectives) is enough from a logical point of view for the purposes of individuating a classical modal system (without Kripkean counterpart) corresponding to D_2. The general idea of the translation can be extended in order to capture different discussive logics. For example one might be interested in representing a discussion between *EMN* agents that are not only consistent and saturated (individually) but that eventually also obey individually (C). Moreover the agents in question can be opinionated (as in many reconstructions of Jaśkowski's ideas) in the sense that for every world w, the set $\{A : w \models \Box_i A\}$ is a maximal and consistent set of sentences. Cases of this sort can be represented via neighborhoods closed under generators – where the group operators used in the translation are still classical and non-Kripkean (even when the agent operators might be Kripkean). I focused here, nevertheless, on what seems the weakest translation available where not only the group operators but also the agent operators are non-Kripkean.

ACKNOWLEDGEMENTS

I would like to thank Rohit Parikh for various interesting conversations on the topic of this paper. Thanks are also due to an anonymous referee especially for various useful remarks concerning Jaśkowski's system D_2.

NOTES

[1] Moshe Vardi considered in (Vardi, 1986) the use of classical modal systems in order to represent both failures of logical omniscience and high probability operators. He also stated in passing that the proposal can be used in order to circumvent the lottery paradox. Nevertheless, after considering the use of classical modal systems, Vardi discarded them without exploring their logical power. For example, the main development in (Arló-Costa, 2002) is to utilize first order extensions of classical systems in order to obtain insights about the lottery paradox. In fact, propositional classical modalities are not expressive enough to encode what is logically interesting about the paradox. Vardi gave two reasons for not utilizing systems of classical modalities (which he dubs *intensional logic* following Montague's terminology). The central reason is that this approach "leaves the notion of a possible world as a primitive notion [...]. While this might be seen as an advantage by the logician whose interest is in epistemic logic, it is a disadvantage for the "user" of epistemic logic whose interest is mostly in using the framework to model belief states (page 297)." Vardi proceeds instead to establish that: "a world consists of a truth assignment to the atomic propositions and a collection of sets of worlds. This is, of course, a circular definition...". Vardi proceeds in ways that are almost completely orthogonal to the main theoretical considerations first introduced in (Arló-Costa, 2002) and re-iterated (more briefly) here. Even when circular definitions and 'composite' worlds can be made mathematically coherent in various ways (among them by abandoning the axioms of foundation in set theory (Barwise, 1988)) this is precisely what I intend to avoid by utilizing classical modal systems in order to represent 'bounded' epistemic operators. Rather than having composite worlds consisting of truth assignments and a collection of sets of worlds, it seems preferable to leave the worlds as primitives and *associating* them with a collection of sets of worlds (their neighborhoods). Once this foundational point is seen, one can go further and proceed to study interesting phenomena via appropriate constraints on the neighborhood function or by studying first order extensions, apparently completely unexplored until the publication of (Arló-Costa, 2002). So, the main point of Vardi's article, in spite of mentioning classical systems in passing, is to defend a position orthogonal with the one presented here and in (Arló-Costa, 2002). His main idea is to abandon the idea (common in epistemic semantics) that worlds should be treated as primitives. Exploring the technical complexities implicit in embracing this research program (which might include the use of circular definitions) might be interesting per se (Barwise and Moss, 1988), but such complexities do not seem to be needed for purely foundational reasons related to the application of epistemic logic to cases where logical

omniscience fails. The modal logics proposed by Scott and Montague, I would like to argue, offer a rich representational tool compatible with the use of standard set theory. Both for applied and for pure applications. In particular I intend to argue here that various types of paraconsistent systems can be parametrically classified by determining their 'classical' counterparts. Moreover, this classification offers simultaneously a manner of developing epistemic interpretations for these paraconsistent systems as well as a heuristic tool for the investigation of the family of non-Kripkean classical modalities itself.

[2] I am indebted here to an anonymous referee who suggested the potential interest of Urchs's work in his review.

[3] The following informal description of Jaśkowski's systems indicates that he might have required a stronger constraint than saturation:

In a discourse, each participant puts forward some information, beliefs, or opinions. What is true in a discourse is the sum of opinions given by participants. Each participant's opinions are taken to be self-consistent, but may be inconsistent with those of others. To formalise this idea, take an interpretation, I, to be one for a standard modal logic, say **S5**. Each participant's belief set is the set of sentences true in a possible world in I. Thus, A holds in I iff A holds at *some* world in I. Clearly, one may have both A and $\neg A$ (but not $A \wedge \neg A$) holding in an interpretation (Priest and Tanaka, 2000).

[4] There is a slight abuse of notation here. We are using the same letter (D) to characterize the agent axiom ($\Box_i A \rightarrow \Diamond_i A$) and the group-axiom ($\Box A \rightarrow \Diamond A$).

REFERENCES

Arló-Costa, H. (2002) First order extensions of classical systems of modal logic: The role of the Barcan schemas, *Studia Logica* **71**, 87–118.

Barwise, J. (1988) Three views of common knowledge, *TARK* **II**, 365–379.

Barwise, J. and Moss, L. (1996) *Vicious Circles: On the Mathematics of Non-Wellfounded Phenomena*, CSLI Publications, February.

Chellas, B. F. (1980) *Modal Logic: An Introduction*, Cambridge University Press, Cambridge.

da Costa, N. C. A. and Dubikajtis, L. (1977) On Jaśkowski's discussive logic, in A. I. Arruda, N. C. A. da Costa and R. Chuaqui (eds.), *Non-Classical Logics, Model Theory and Computability*, North-Holland Publishing Company.

Halpern, J. Y. and Rabin, M. O. (1987) A logic to reason about likelihood, *Artificial Intelligence* **32**, 379–405.

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

Jaśkowski, S. (1948) Rachunek zdań dla systemów dedukcyjnych sprzecznych, *Studia Societatis Scientiarum Torunensis, Section A* **1**(5), 57–77.

Jaśkowski, S. (1969) Propositional calculus for contradictory deductive systems, *Studia Logica* **XXIV**, 143–157.

Kyburg, H. E., Jr. (1961) *Probability and the Logic of Rational Belief*, Wesleyan University Press, Middletown.

Kyburg, H. E., Jr. (1997) The rule of Adjunction and reasonable inference, *J. Philosophy*, March.

Kyburg, H. E., Jr. and Teng, C. M. (2002) The logic of risky knowledge, in *Proceedings of WoLLIC*, Brazil.

Kolmogorov, A. N. (1924) O principie tertium non datur, *Mat. Sbornik* **32**.

Montague, R. (1970) Universal grammar, *Theoria* **36**, 373–398.

Priest, G. and Tanaka, K. (2000) *Paraconsistent Logic*, Stanford Encyclopedia of Philosophy, `http://plato.stanford.edu/entries/logic-paraconsistent/`, December 6.

Schotch, P. K. and Jennings, R.E. (1980) Inference and necessity, *J. Philos. Logic* **9**, 327–340.

Segerberg, K. (1971) *An Essay in Classical Modal Logic*, Stockholm.

Scott, D. (1970) Advice in modal logic, in Lambert, K. (ed.), *Philosophical Problems in Logic*, Reidel, Dordrecht, pp. 143–173.

Urchs, M. (1995) Discursive logic. Towards a logic of rational discourse, *Studia Logica* **54**, 231–249.

Vardi, M. (1986) On epistemic logic and logical omniscience, in Halpern, Y. (ed.), *Theoretical Aspects of Reasoning about Knowledge. Proceedings of the 1986 Conference*, Morgan Kaufmann, Los Altos.

Carnegie Mellon University
e-mail: hcosta@andrew.cmu.edu