MODALITY, BISIMULATION AND INTERPOLATION IN INFINITARY LOGIC

Johan van Benthem

ILLC, University of Amsterdam

july 1997

Abstract

Rohit Parikh has been one of the pioneers at the interface of modal and dynamic logic. Bringing dynamic, process-oriented concerns into modal logic has been a major move, whose repercussions are still being felt today. In this contribution, I consider the bridge concept of bisimulation, which links modal logic with computational process theories. Our main results show how its model-theoretic invariance properties, first established within first-order model theory, may be lifted to infinitary logic, a natural generalized habitat for theories of programs and processes. The techniques used for this purpose also suggest some further proof-theoretic uses, going beyond purely modal languages.

to appear in

M. Fitting, R. Ramanujan, & K. Georgatos, eds.

"Festschrift for Rohit Parikh's 60th Birthday"

Annals of Pure and Applied Logic.

1 Bisimulation in Modal and Dynamic Logic

Processes can be represented by poly-modal Kripke models, giving their states and possible transitions. More precisely, a *labeled transition system* (LTS) is a set of states with a family of binary transition relations over these: $(S, \{R_a\}_{a \in A})$, plus an evaluation function for unary predicates over states. Different notions of process then correspond to classes of LTSs, closed under suitable semantic relations of 'process equivalence'. In theories of computation, 'bisimulation' between is a central measure for this purpose.

<u>Definition</u> A *bisimulation* between two labeled transition systems is a binary relation C between their state sets satisfying 'atomic harmony' at C–corresponding states, plus zigzag clauses allowing 'mutual tracing' of the process for any action a . More precisely,

(i) if $s\,C\,s'$, then s, s' validate the same atomic propositions

(iia) if $s\,C\,s'$ and $s\,R_a\,t$, then there is some t' with $s'\,R_a\,t'$ and $t\,C\,t'$

(iib) and vice versa.

An assertion $\phi(x)$ (in any formal language) is called *invariant for bisimulation* if, whenever C is a bisimulation between $\mathbf{M}, \mathbf{M'}$ with $s\,C\,s'$, $\mathbf{M}, s \models \phi$ iff $\mathbf{M'}, s' \models \phi$.

Bisimulation first emerged in modal logic. It is the key semantic invariance for the modal propositional language describing LTSs, which has the usual Boolean operators, as well as indexed modalities $\langle a \rangle$ for each atomic action $a \in A$. By an easy induction, whenever C is a bisimulation between two models $\mathbf{M}, \mathbf{M'}$ with $s\,C\,s'$, we have

$$\mathbf{M}, s \models \phi \quad \text{iff} \quad \mathbf{M'}, s' \models \phi \qquad\qquad \text{for all modal formulas } \phi$$

Atomic harmony provides the induction base, and the two zigzag clauses are just what is needed here to push the induction through for the existential modalities. This observation can be reversed. Bisimulation invariance is indeed the defining semantic characteristic of the modal formalism. This the content of a model-theoretic preservation result (van Benthem 1976) – which views the modal language as a fragment of the obvious first-order description language for Kripke models (by the well-known 'standard translation').

<u>Modal Invariance Theorem</u> A first-order formula $\phi(x)$ over labeled transition systems is invariant for bisimulation iff it is definable by means of a modal formula.

A general proof via ω–saturated models is in van Benthem 1996. By now, many proofs exist for the MIT, as well as its variations and extensions. The general format measures the syntactic expressive power associated with any semantic process equivalence.

The analysis of assertions in the Modal Invariance Theorem extends to programs in *propositional dynamic logic*. Consider the following notion of invariance for program operations – which comes up naturally in an inductive proof of bisimulation invariance for the class of all PDL-formulas, which recurse toward regular programs:

Definition An operation $O(R_1, ..., R_n)$ on programs is *safe for bisimulation* if, whenever C is a relation of bisimulation between two models for their transition relations $R_1, ..., R_n$, then it is also a bisimulation for the defined relation $O(R_1, ..., R_n)$. In a derived sense, we call any formula $\pi(R, S, ..., x, y)$ defining an operation on binary relations $R, S, ...$ safe for bisimulation when its denotation is safe in the preceding sense.

It is easy to show that the regular operations of relational composition **;** and choice $\cup$ (Boolean union) have this property, and so do test relations (ϕ)**?** for modal formulas ϕ. Typically non-safe operations are program intersection and Boolean complement. But the following negation operation is safe: $\sim(R) = \{ (x, y) \mid x=y$ *and for no* $z : x R z \}$. All these operations are first-order definable in an obvious language over LTSs. Indeed, we have this counterpart to the above Modal Invariance Theorem (van Benthem 1993):

Modal Safety Theorem A first-order operation $O(R_1, ..., R_n)$ is safe for bisimulation iff it can be defined using atomic relations $R_a xy$ and atomic tests (q)**?** for propositional atoms q in our models, using the three operations **;** , $\sim$ and $\cup$.

This result expresses functional completeness for dynamic counterparts of the Boolean primitives $\wedge$, $\neg$, $\vee$. For a proof, cf. van Benthem 1996, Chapter 5. New proofs are found in Hollenberg 1997, which studies safety over several broader notions of process equivalence, as well as extended languages including monadic second-order logic (following the MIT for the modal μ–calculus found in Janin & Walukiewicz 1996).

2 Modal Invariance in Infinitary Logic

It is natural to seek *infinitary* versions of MIT and MST. For, the usual regular program operations include Kleene iteration – and many further natural programming constructs are infinitary as well. Barwise & Moss 1996 show how infinitary modal logic ties in with non-well-founded set theory, and the infinitary first-order logic of bisimulation. So, consider an infinitary first-order language $L_{\infty\omega}$ over possible worlds models **M** – in the appropriate similarity type $\{R_a^2, R_b^2, ..., =, P^1, Q^1, ...\}$ – which allows arbitrary set conjunctions $\bigwedge_{i\in I} \phi_i$ and disjunctions $\bigvee_{i\in I} \phi_i$, but quantifiers only one at a time. Moreover, in this construction, we only allow formulas with a finite number of free variables – with one exception in Section 4, which we will indicate separately. The infinitary modal language to be used here extends the basic one in the same manner. Infinitary modal formulas are invariant for bisimulation, as their infinitary conjunctions and disjunctions fall within the obvious inductive argument. This time, a folklore result states a straightforward converse. Any two models **M**, x, **N**, y sharing the same infinitary modal theory have a bisimulation C connecting x to y. This does not yield an MIT-style preservation result in any obvious way. But we do have the latter, too – even though first-order proofs based on compactness and saturation fail for $L_{\infty\omega}$.

Theorem An infinitary first-order formula is invariant for bisimulation iff it is definable by means of an infinitary modal formula.

Proof This result is proved in van Benthem 1996, Ch. 10, using modified 'consistency families' to circumvent compactness. Another proof is in Barwise & van Benthem 1996. We will use techniques from the latter to also extend the Modal Safety Theorem to $L_{\infty\omega}$. Therefore, we give a brief sketch of the relevant argument. It involves crucial use of the following remnant of compactness retained by the infinitary language:

Boundedness Theorem for $L_{\infty\omega}$

Let $\psi(<)$ be a formula of $L_{\infty\omega}$ with models whose domains can be well-orders $<$ of any size. Then ψ also has a model where $<$ is not a well-order.

Now, suppose that ϕ is invariant for bisimulation. We prove that

\# There exists an ordinal κ such that for all models **M**, s $\models \phi$ and all models **M'**, s' having the same modal theory as **M**, s *up to modal operator depth* κ, **M'**, s' $\models \phi$

Modal operator depth is measured here in the usual way. (With infinitary combinations, it can run up to arbitrarily high ordinals.) The crucial property of this notion is this

Modal Karp Lemma

Two models **M**, s, **M'**, s' share the same modal theory up to depth κ iff there exists a descending chain of sets of 'partial bisimulations' between them of length κ, with zigzag conditions holding downward from levels $\beta+1$ to β.

More precisely, this description says that there exists a sequence $(I_\beta)_{\beta\leq\kappa}$ of non-empty sets of partial bisimulations between **M** and **N** such that $I_\beta \subseteq I_\gamma$ when $\gamma\leq\beta$, and such that, if $E\in I_{\beta+1}$ contains a tuple (a, b) and aRc for some $c\in|\mathbf{M}|$, then there is an object $e\in|\mathbf{N}|$ with cRe and some F in I_β with $E\subseteq F$ and $(c, e)\in F$. Moreover, the same requirement must hold in the opposite direction. (Similar results involving Ehrenfeucht games for standard first-order logic may be found in Doets 1996):

From # , the desired (infinitary) modal definability of ϕ follows quite easily. Consider the set (!) of all complete modal descriptions up to depth κ of all models **M**, s for ϕ . Then ϕ is equivalent to the $L_{\infty\omega}$–disjunction of all of these. (By construction, ϕ implies the disjunction. Conversely, if some disjunct holds in a model **M'**, s' , then, by definition, there is a model **M**, s sharing the same modal theory up to modal operator depth κ where ϕ holds. Property # then ensures the transfer of ϕ to **M'**, s' .)

Given all these observations, to finish our proof of the MIT, it remains to discuss the

<u>Subproof of #</u> Suppose that this transfer property fails. This means that, for each ordinal κ , there exist models **M**, s and **M'**, s' with

(i) **M**, s $\models \phi$ (ii) not **M'**, s' $\models \phi$
(iii) **M**, s and **M'**, s' have the same modal theory up to depth κ

By the above lemma, **M** and **M'** have a descending 'κ–tower' of partial bisimulations. Now, this situation may be coded up by a single infinitary first-order formula $\Phi(<)$. (This syntactic trick comes from a well-known proof of Lindström's Theorem.) Using fresh predicate letters A, B, C^k, I, <, one states that

> $(\phi)^A$, $(\neg\phi)^B$, $<$ is a linear order, while $C^k\ i\ \mathbf{x}\ \mathbf{y}$ is a (1+2k)-ary predicate defining a partial bisimulation of size k between matched members in the sequences $\mathbf{x}$, $\mathbf{y}$.

Here, the variable i runs over an index set I ordered by < , and we can also state the key zigzag properties. E.g., if $C^k\ (i+1)\ \mathbf{x}\ \mathbf{y}$ and Au, $R_a(\mathbf{x})_j u$, then there exists some v with Bv, $R_a(\mathbf{y})_j v$ such that $C^{k+1}\ i\ \mathbf{x}u\ \mathbf{y}v$. Now, the Boundedness Theorem says that $\Phi(<)$ has a model in which $<$ is not a well-order. Therefore, that model must have at least one countably descending chain of indices. Collecting all finite partial bisimulations along its stages, we get a true bisimulation, without any bound on its zigzag properties. But then, we have found two models A, B connected by a bisimulation which disagree on ϕ . And this refutes invariance for bisimulation. ■

3 Modal Safety in Infinitary Logic

By similar reasoning with the Boundedness Theorem, we now derive our main result.

<u>Theorem</u> A relational operation $\pi\ (R_1, ..., R_n)$ in $L_{\infty\omega}$ is safe for bisimulation iff it can be defined using atomic relations $R_a xy$ plus atomic tests (q)**?** , using only three operations **;** , $\bigcup$ and **~** , where the unions may now be *infinitary.*

Proof We recall the main features of the proof for the finitary first-order case (cf. van Benthem 1996, Ch. 5), while identifying those steps where a new subroutine is needed. The top-level argument remains the same, up to one important module.

I Specifying the relevant vocabularies, if a relational operation defined by $\pi = \pi(x, y)$ is safe for L-bisimulations (where L contains the relevant relational arguments and atomic predicates), then the $L_{\infty\omega}$–formula $\exists y\, (\pi(x, y) \wedge Qy)$ is invariant for (L+Q)-bisimulations, where Q is a new unary predicate letter. By the above infinitary Modal Invariance Theorem, there is then an equivalent infinitary modal formula $\phi(q)$.

II Due to the simple occurrence of Q , the latter has a further strong semantic property. Call $\phi(q)$ *distributive* in the proposition letter q if the following equivalence holds in each model (stated for convenience with some benign abuse of notation):

for each family of subsets $\{P_i\}_{i\in I}$, $\phi(\bigcup_{i\in I} P_i) \leftrightarrow \bigvee_{i\in I} \phi(P_i)$

From right to left, this amounts to semantic *monotonicity* whose syntactic correlate is obligatory positive occurrence for q – but the other half excludes a lot more. We want a syntactic preservation theorem for distributive infinitary modal formulas. This can be done – and the resulting normal forms are described in a separate theorem below.

III From these forms, one extracts the following explicit information. Any safe relation $\pi(x, y)$ may be defined as an infinitary union of finite sequential compositions of successive atomic actions $R_a xy$ plus tests (α)**?** for infinitary modal formulas α .

IV Finally, the latter possibly complex tests may be unpacked to combinations of atomic tests, using the following batch of valid equivalences:

$$(\bigvee_{i\in I} \phi_i)? = \bigcup_{i\in I} (\phi_i)? \qquad (\neg \phi)? = \sim (\phi)? \qquad (<a>\phi)? = \sim\sim (a\, ;\, (\phi)?)$$

■

In module **II**, we must prove the following independent model-theoretic preservation theorem – which is the technical core of our contribution. It also has an independent interest (perhaps, even surpassing that of the Modal Safety Theorem for some readers):

Theorem Up to logical equivalence, the q–distributive infinitary modal formulas $\phi(q)$ are just those definable by infinitary disjunctions of 'existential forms' $\alpha_0 \wedge q$, $\alpha_0 \wedge <a_1> (\alpha_1 \wedge q)$, $\alpha_0 \wedge <a_1> (\alpha_1 \wedge <a_2> (\alpha_2 \wedge q))$, etc., where all the infinitary modal formulas α_i are q–free.

As said before, from the syntactic format in the preservation theorem for distributivity, one easily extracts the desired normal form for operations that are safe for bisimulation.

Proof All forms described are evidently distributive w.r.t. the proposition letter q . The hard part is the converse. Let us first analyse the models **M**, s where a distributive formula $\phi(q)$ holds. Any denotation for q can be written as a union of singletons, and so, by distributivity, ϕ will hold with q true in only one world t . (In case q's denotation is empty, monotonicity will make it true for any singleton denotation $\{t\}$.) Moreover, we may assume that this single q–world lies at some finite successor distance from s , since we also have ϕ true at the submodel generated from the root. Thus, there is some finite sequence $s = x_1, ..., x_n = t$ of states via atomic actions leading from s to the unique world where q holds. Call a model **M'**, s' a κ–*relative of* **M**, s if it has a corresponding sequence $x_1', ..., x_n'$ leading to a q–world $t' = s_n'$, such that matched worlds x_i, x_i' satisfy the same infinitary q–free modal formulas up to operator depth κ . (As in the above, we will refer to the relevant vocabulary as language L .) We prove a

Lemma There exists an ordinal κ such that, if **M**, s $\models \phi$ and **M'**, s' is a κ–relative of **M**, s , then **M'**, s' $\models \phi$.

From this, the required definition for ϕ arises as a disjunction of all modal descriptions up to depth κ of finite q–paths in models **M**, s for ϕ as described just now. (This is a set, because of the restriction to fixed modal depth.) The justification for this definition is similar to an earlier one. Clearly, by our construction, ϕ implies this disjunction. But also conversely, whenever some disjunct holds, we are in a model which is a κ–relative of some such **M**, s , and the Lemma tells us that ϕ must hold there as well.

Proof of the Lemma The argument starts like in the earlier proof of the infinitary Modal Invariance Theorem. Assume, for each ordinal κ , that there exist two models **M**, s $\models \phi$ and **M'**, s' with κ–corresponding finite branches as above, such that ϕ fails in **M'**, s' . Now, code up this situation in one infinitary formula $\Psi(A, B, \mathbf{x}, \mathbf{y}, C^k, I, <)$ which describes, in particular, the existence of some $<$–descending sequence (as measured along the ordered index set I) of partial L-bisimulations. The sequence starts with the simulation sending the states $(\mathbf{x})_i$ to the corresponding $(\mathbf{y})_i$ at the top. Also, we can state explicitly that in A, there is just one q–world. This formula then has models with well-orders of arbitrarily high cardinality for $<$. By the Boundedness Theorem, it must have a model where $<$ is not a well-order on I . Using a countable descending chain of indices in I as before, such a model yields the following situation:

- a model $\mathbf{M}$, s $\models \phi$ with a finite sequence $s=s_1, ..., s_n=t$ of atomic steps leading from s to its only q–world t
- a model $\mathbf{M'}$, s' where ϕ fails, with a similar sequence $s_1', ..., s_n'$ leading from s' to some q–world $t'=s_n'$
- an L–bisimulation C between $\mathbf{M}$ and $\mathbf{M'}$ with s_i C s_i' $(1 \leq i \leq n)$

The remainder of the argument is as for the finitary Safety Theorem (van Benthem 1996, Ch. 5), without any need for infinitary modifications. Here is an outline. Given two models as described, one can perform a number of semantic constructions that do not affect truth of the relevant infinitary modal formulas. In particular, in the usual way, one can *unravel* the two models $\mathbf{M}$, $\mathbf{M'}$ to *trees* $\mathbf{T}$, $\mathbf{T'}$ with branches of length at most ω . Next, using successive (L+q)–*bisimulation-preserving* moves of copying subtrees and re-attachment of nodes on these trees, one can even improve the above linkage to an L–bisimulation between tree models $\mathbf{N}$, $\mathbf{N'}$ with the following extra:

> in the L-bisimulation C , the links between corresponding nodes in the distinguished branches are *unique*: these nodes do not attach to any others

Moreover, we can perform the relevant 'gardening moves' in such a way that q will still be true at only one world in $\mathbf{N}$. As all moves involved are (L+q)–bisimulations, $\mathbf{N}$, s $\models \phi(q)$. Now consider the model $\mathbf{N}^*$ which is $\mathbf{N'}$ with the following difference: q is true only in t' . Clearly, our L-bisimulation is even an (L+q)–bisimulation between $\mathbf{N}$, s and $\mathbf{N}^*$, s' . (This is precisely why we wanted the uniqueness of the match between t and t' .) Then we can argue as follows. Since the modal formula $\phi(q)$ holds at $\mathbf{N}$, s , it also holds at $\mathbf{N}^*$, s' (as we now even have an (L+q)–bisimulation.) But then, by monotonicity, $\phi(q)$ also holds in the model $\mathbf{N'}$, s' (whose denotation for q can only be larger). But then, again, as the gardening and unraveling moves were all (L+q)–bisimulations, we would also have that $\mathbf{T'}$, s' $\models \phi(q)$, and hence $\mathbf{M'}$, s' $\models \phi(q)$. At this point, we have derived a contradiction with the given failure of ϕ at $\mathbf{M'}$, s' . ■

Hollenberg 1997 has some new proofs of modal safety for non-first-order languages, which might simplify the above laborious 'argument from surgery'.

4 Infinitary Interpolation

The Modal Invariance Theorem may also be generalized to an *interpolation property*. Barwise & van Benthem 1996 propose a generalization of the Craig Interpolation Theorem which also applies to infinitary first-order logic, as well as other logical formalisms which lack the standard version of interpolation. (Examples where this strategy works are finite-variable fragments of first-order logic. Cf. Andréka, van Benthem & Németi 1996.) Their general strategy is the replacement of ordinary Tarski consequence by a more general notion of consequence along any model relation.

Definition Let R be any semantic relation between models. We say that ϕ *implies* ψ *along* R if, whenever $\mathbf{M} \models \phi$ and $\mathbf{M}\, R\, \mathbf{N}$, then $\mathbf{N} \models \psi$.

Standard consequence is the case of the identity relation. Another important case has R as 'potential L-isomorphism': the existence of a family of finite partial L-isomorphisms between $\mathbf{M}, \mathbf{N}$ with the usual back and forth properties. Note that this notion of consequence does not have all the structural properties of standard consequence. E.g., the Cut rule will only hold if the relation R is transitive. We state the main result here, which generalizes both interpolation and preservation theorems.

Theorem For $L_{\infty\omega}$–formulas $\phi(x), \psi(x)$, the following are equivalent:

(i) there is an $\alpha \in L_\phi \cap L_\psi$ such that $\phi \models \alpha \models \psi$

(ii) ϕ implies ψ along potential $L_\phi \cap L_\psi$ –isomorphism.

Barwise & van Benthem 1996 give a proof for this theorem based on the Boundedness Theorem for $L_{\infty\omega}$. They show how to specialize the argument to get an equivalence between consequence along *bisimulation* and the existence of some *modal* interpolant. The Modal Invariance Theorem then follows by setting $\psi = \phi$. In this Section, we present a more traditional Henkin-style proof for the infinitary Interpolation Theorem. Its route is more laborious – but in return, it provides suggestive additional information.

A New Proof From (i) to (ii), this is an immediate consequence of the simple fact that potential isomorphism in a similarity type L preserves truth of all L–formulas. For the direction from (ii) to (i), assume that ϕ, ψ have no interpolant in $L = L_\phi \cap L_\psi$. We are going to construct a counterexample to (ii), using '*good triples*' (E, Σ, Δ), where the idea is that Σ describes a model for ϕ over some domain of constants A, Δ one for $\neg\psi$ over constants B, and E a potential L–isomorphism between A, B, all 'in statu nascendi'. We start with some preliminaries. First, set $\mu =_{def} \max(\aleph_0, |subformulas(\phi)|, |subformulas(\psi)|)$. Next, choose two disjoint sets of constants A, B of size μ^+, the first regular cardinal greater than μ . For convenience, in what follows, we shall be working with formulas in *normal form*, constructed from atoms and their negations using both quantifiers, as well as arbitrary set conjunctions and disjunctions. Moreover, throughout, formulas will only contain a finite number of constants.

<u>Definition</u> A *good triple* (E, Σ, Δ) satisfies the following requirements:

(1) E is a set of tuples **a**, **b** ($\mathbf{a} \subseteq A$, $\mathbf{b} \subseteq B$) with length (**a**) = length (**b**) where all elements of **a** are distinct from each other, and similarly for **b**

(2) Σ is a set of subformulas of ϕ made into sentences by plugging in constants from A; and likewise for Δ w.r.t. subformulas of $\neg\psi$ and constants from B

(3) the cardinalities $|E|, |\Sigma|, |\Delta|$ are all smaller than μ^+

(4) Σ, Δ are *L-inseparable via* E . That is, there is no set $\{ (\mathbf{a}^s, \mathbf{b}^s) \mid s \in S \}$ of tuples in E , each with a corresponding L-formula $\beta(\mathbf{x})$, such that for some infinitary $\bigvee$ –combination α of the formulas $\beta(\mathbf{x})$, (i) $\Sigma \models \alpha\ [\mathbf{x}^s{:=}\mathbf{a}^s]$, while (ii) $\Delta \models \neg\alpha\ [\mathbf{x}^s{:=}\mathbf{b}^s]$

<u>Note</u> These $\vee$ –combinations genuinely extend $L_{\infty\omega}$, but they are still invariant under potential L–isomorphism, in an obvious sense. (Allowing existential quantifiers over infinite conjunctions, like in $\exists x\ \wedge_n Rxa_n$, *would* give problems with invariance.) In what follows, we shall call these combinations 'extended modal formulas'.

<u>Fact</u> Choose any starting tuple **a**, **b** for the free variables of ϕ , $\neg\psi$.
Then $(\{\langle \mathbf{a}, \mathbf{b}\rangle\}, \{\phi(\mathbf{a})\}, \{\neg\psi(\mathbf{b})\})$ is a good triple.

<u>Proof</u> The only non-trivial property to be checked is Non-Separation. But the above strong formulation reduces to the usual inseparability given by the negation of clause (ii) in our Theorem, in this special case where we only have one tuple **a**, **b** in E . ■

We check a bunch of extension principles for Σ (those for Δ are entirely similar), which are like the usual ones for 'consistency properties' in infinitary logic.

<u>Facts</u>

(i) If (E, Σ, Δ) is good, and $\wedge_i \phi_i \in \Sigma$, then $(E, \Sigma \cup \{\phi_i\}_i, \Delta)$ is good

(ii) If (E, Σ, Δ) is good, and $\vee_i \phi_i \in \Sigma$, then for some i, $(E, \Sigma \cup \{\phi_i\}, \Delta)$ is good

(iii) If (E, Σ, Δ) is good, and $\forall x\phi \in \Sigma$, then for all $a \in A$, $(E, \Sigma \cup \{\phi(a)\}, \Delta)$ is good

(iv) If (E, Σ, Δ) is good, and $\exists x\phi \in \Sigma$, then for any $a \in A$ that is *new* to Σ and E , $(E, \Sigma \cup \{\phi(a)\}, \Delta)$ is good

<u>Proof</u> (i) Adding all consequences ϕ_i of $\bigwedge_i \phi_i \in \Sigma$ does not affect (non-)separation. Moreover, the cardinality of the extended Σ stays below μ^+ . (ii) Here we need the extended class of infinitary $\bigvee$ –combinations. Suppose that all triples $(E, \Sigma \cup \{\phi_i\}, \Delta)$ *do* L–separate, say via extended formulas α_i . Then $\bigvee_i \alpha_i$ separates Σ, Δ via E : quod non. (iii) Again, adding the logical consequence $\phi(a)$ does not affect separation.

(iv) Adding $\phi(a)$ with a new constant a does not yield new separations. For, this move does not trigger new tuples in E, and then we have the usual valid inference from $\Sigma \cup \{\phi(a_{new})\} \models \alpha$ to $\Sigma \cup \{\exists x \phi\} \models \alpha$. (Note that the new a does not occur in α). ■

The new feature, as compared with consistency properties, are extension principles for the component E , that will create the required features of a potential L–isomorphism. As before, we suppress redundant symmetric versions.

Facts (Continued)

(v) If (E, Σ, Δ) is good, $\mathbf{a}, \mathbf{b} \in E$, $P \in L$, $P\mathbf{a} \in \Sigma$, then $(E, \Sigma, \Delta \cup (P\mathbf{b}\})$ is good

(vi) If (E, Σ, Δ) is good, $\mathbf{a}, \mathbf{b} \in E$, then for any $a \in A$ that is new to $\mathbf{a}$ and for any $b \in B$ that is *new* to Δ, E , $(E \cup (\mathbf{a}a, \mathbf{b}b\}, \Sigma, \Delta)$ is good

Proof (i) Suppose there were a separation, say by the formula α . Then we must have $\Sigma \models \alpha$ [*A–substitutions*] $\wedge$ $P\mathbf{a}$, and $\Delta, P\mathbf{b} \models \neg \alpha$ [*B–substitutions*]. The latter implies $\Delta \models \neg (\alpha \wedge P\mathbf{x})$ [*B–substitutions*] . This yields a separation for Σ, Δ via E after all. (Note that we can rewrite $\alpha \wedge P\mathbf{x}$ as an admissible infinitary disjunction.) (ii) Suppose that we get a separation via the new E–link. More precisely, $\Sigma \models \alpha$ [*A–substitutions*], and $\Delta \models \neg \alpha$ [*B–substitutions*], where α is an extended L–formula as before, now also involving L–subformulas $\beta(\mathbf{x}, y)$ associated with $\mathbf{a}a$ (for Σ) and $\mathbf{b}b$ (for Δ). This gives the following separation for the original case: $\Sigma \models \exists y \alpha$ [$\mathbf{a}$], $\Delta \models \neg \exists y \alpha$ [$\mathbf{b}$] (recall that b was *new*). Again, $\exists y \alpha$ is equivalent to an admissible extended formula, by moving the existential quantifier inside over disjunctions. The result is an ordinary L-formula with respect to the old pair $(\mathbf{a}, \mathbf{b})$. ■

Now we construct our models. We list all good triples in a sequence of length μ^+ , interspersed with all relevant formulas, and all constants. We make each item occur cofinally often, to ensure fair scheduling. This can be done, for cardinality reasons. The construction involves a (componentwise) growing sequence of good triples in an ordinal sequence $T_0, T_1, ..., T_\alpha, ...$ $(\alpha < \mu^+)$. Our steps follows the above decompositions, starting from the initial good triple $(\{<\mathbf{a}, \mathbf{b}>\}, \{\phi(\mathbf{a})\}, \{\neg \psi(\mathbf{b})\})$. Whenever a formula is scheduled, we check if it triggers a possible extension as listed in the above Facts, and then perform that – and the same with constants and E–zigzags. At limit ordinals, we take the union of our efforts so far, and continue. In the standard manner, this gives us two models – one based on A for $\cup_i \Sigma^i$, one based on B for $\cup_i \Delta^i$, while $\cup_i E^i$ describes a potential L–isomorphism between these two. ■

Here is the surplus in this proof. The core of the argument are the construction rules. These may also be viewed as *tableau rules* for a calculus of 'joint consistency' along potential L–isomorphism. The rules deviate from standard ones in their *ternary* format

$$\Sigma \text{ cons}_E \ \Delta$$

where E codes the relevant vocabulary and object links. The intended interpretation validates explicit equivalences like the following:

<u>Fact</u>

(i) $\Sigma + P\mathbf{a}$ $\text{cons}_{E + \mathbf{a}, \mathbf{b}}$ Δ iff $\Sigma + P\mathbf{a}$ $\text{cons}_{E + \mathbf{a}, \mathbf{b}}$ $\Delta + P\mathbf{b}$

(ii) likewise for negated atoms $\neg P\mathbf{a}$, $\neg P\mathbf{b}$

(iii) $\Sigma + \wedge_i \phi_i$ cons_E Δ iff $\Sigma + \wedge_i \phi_i + \{\phi_i\}_i$ cons_E Δ

(iv) $\Sigma + \vee_i \phi_i$ cons_E Δ iff *for some* i, $\Sigma + \phi_i$ cons_E Δ

(v) $\Sigma + \exists x \phi$ cons_E Δ iff *for some new* a, $\Sigma + \phi(a)$ cons_E Δ

(vi) Σ $\text{cons}_{E + \mathbf{a}, \mathbf{b}}$ Δ, $a \in \Sigma$ new to $\mathbf{a}$ only if *for some new* b, Σ $\text{cons}_{E + \mathbf{a}, \mathbf{b} + \mathbf{a}a, \mathbf{b}b}$ Δ

5 Discussion: Ternary Sequents Manipulating Vocabulary

The preceding analysis suggests an independent study of 'interpolation inferences'. One can recast the preceding principles as *inference rules* manipulating *ternary* sequents, with an additional argument recording relevant vocabulary:

$$\Sigma \Rightarrow_E \Delta$$

Historically, notions of inference keeping an explicit record of variable and fixed vocabulary occur as early as Bernard Bolzano's work in the early 19th century on styles of consequence (cf. van Benthem 1985). Working with such sequents may change familiar features of logical consequence. There are now three positions at which to formulate *structural rules*, and e.g., one can have Monotonicity or Additivity w.r.t. vocabulary. In this connection, recall that consequence along a model relation did not necessarily retain all usual structural rules. (In fact, what it does retain are strengthening and disjunction of antecedents, as well as weakening and conjunction of consequents.)

Indeed, a number of ternary inference notions occurred in the above. For convenience, disregard the complication of the *family of* links in E, with infinitary conjunctions and disjunctions over the associated formulas. The negation of Σ cons_E Δ then states the existence of some separating L-formula γ with $\mathbf{a}E\mathbf{b}$, $\gamma(\mathbf{a})$ implied by Σ and $\gamma(\mathbf{b})$ refuted by Δ. If we turn this into a positive statement, using negations of the formulas in Δ for convenience, then we get the existence of some L-formula with $\gamma(\mathbf{a})$ implied by Σ and $\gamma(\mathbf{b})$ implying the disjunction of Δ, as usual. This notion of 'interpolation consequence' implies our initial one of 'consequence along potential L-isomorphism'. But the latter may also, of course, be studied in its own right. (By the analysis of Barwise & van Benthem 1996, it is RE for first-order logic, and many of its variants.)

Consequence along potential isomorphism has some interesting features, as compared with ordinary sequent calculi. We already mentioned the structural rules. But also, this calculus does not obey all the usual *logical rules*. E.g., the usual *conditionalization rule* fails for conditionals. To see this, let the infinitary formula $\phi = \phi\,(D, <, =)$ define the ordinal ω_0 categorically, with D interpreted as the whole domain. Likewsie, let the formula $\psi = \psi\,(D', <', =)$ define the ordinal ω_1 categorically, with D' equal to the whole domain. Evidently, $\phi\,(D, <, =), \psi\,(D', <', =) \Rightarrow_{\{=\}} \bot$. But this does not imply $\phi\,(D, <, =) \Rightarrow_{\{=\}} \psi\,(D', <', =) \to \bot$, since any two infinite domains admit of a $\{=\}$-potential isomorphism. Conditionalization does hold when we modify the E-argument. For, if Σ, A $\Rightarrow_E$ D, and L(A) is the vocabulary of A, then $\Sigma \Rightarrow_{E \cup \{\mathbf{L(A)}\}} A \to D$.

We conjecture that this ternary rule format captures consequence, even for deviant languages like *finite-variable* fragments, where no Gentzen system can ever axiomatize ordinary validity (cf. Andréka, van Benthem & Németi 1996). E.g., consider the following counter-example to interpolation inside the two-variable fragment (with =):

$$|A| \leq 1, |-A| \leq 1 \Rightarrow \neg\,(\exists x\,(Bx \wedge Cx) \wedge \exists x\,(Bx \wedge \neg Cx) \wedge \exists x\,(\neg Bx \wedge Cx))$$

There is no pure identity interpolant in two variables. Such formulas cannot distinguish between domains with 2 objects (where the antecedent may hold) and domains with 3 objects (where the consequent can be refuted). With our ternary inference, we do obtain

$$|A| \leq 1, |-A| \leq 1, \exists x\,(Bx \wedge Cx), \exists x\,(Bx \wedge \neg Cx), \exists x\,(\neg Bx \wedge Cx) \Rightarrow \bot$$
$$|A| \leq 1, |-A| \leq 1 \Rightarrow_{\{=, B, C\}} \neg\,(\exists x\,(Bx \wedge Cx) \wedge \exists x\,(Bx \wedge \neg Cx) \wedge \exists x\,(\neg Bx \wedge Cx))$$

The enforced E–registration of cross-over blocks the sequent for a standard interpolant:

$$|A| \leq 1, |-A| \leq 1 \Rightarrow_{\{=\}} \neg\,(\exists x\,(Bx \wedge Cx) \wedge \exists x\,(Bx \wedge \neg Cx) \wedge \exists x\,(\neg Bx \wedge Cx))$$

This analysis can be pushed still further, to probe where the classical proof of the preceding sequent must employ principles beyond the two–variable Gentzen format. Evidently, these are just speculations at this stage – that still have to be made good.

6 Further Directions

At this stage, several further directions are open to us. One can do a similar analysis of infinitary invariance, safety and interpolation for *other process equivalences*. There is also the matter of extending the above results to fixed-point logics, like the *μ-calculus*. And finally, *ternary interpolation calculi* seem a promising new proof format per se.

7 References

H. Andréka, J. van Benthem & I. Németi, 1996, 'Modal Logics and Bounded First-Order Fragments'. To appear in the *Journal of Philosophical Logic*.

J. Barwise & L. Moss, 1996, *Vicious Circles*, CSLI Publications, Stanford.

J. Barwise & J. van Benthem, 1996, 'Interpolation, Preservation, and Pebble Games', Research Report ML-96-12, Institute for Logic, Language and Computation, University of Amsterdam.

J. van Benthem, 1976, *Modal Correspondence Theory*, Ph.D. Thesis, Department of Mathematics, University of Amsterdam.

J. van Benthem, 1985, 'The Variety of Consequence, According to Bolzano', *Studia Logica* 44:4 (1985), 389-403.

J. van Benthem, 1996, *Exploring Logical Dynamics*, CSLI Publications, Stanford.

K. Doets, 1996, *Basic Model Theory*, CSLI Publications, Stanford.

M. Hollenberg, 1996A, 'Bisimulation Respecting First-Order Operations', Logic Group Preprint Series 156, Institute for Philosophy, Rijksuniversiteit Utrecht.

M. Hollenberg, 1996B, 'General Safety for Bisimulation', in P. Dekker & M. Stokhof, eds., *Proceedings 10th Amsterdam Colloquium*, Institute for Logic, Language and Computation, University of Amsterdam.

M. Hollenberg, 1997, *Modal Logic and Process Invariances*, forthcoming Ph.D. dissertation, Institute for Philosophy, Rijksuniversiteit Utrecht.

D. Janin & I. Walukiewicz, 1996, 'On the Expressive Completeness of the Propositional μ–Calculus with Respect to Monadic Second-Order Logic', Department of Mathematics and Informatics, University of Bordeaux & Department of Computer Science, Aarhus University.