Reasoning with Partial Knowledge

László Pólos and Michael T. Hannan

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2000-30-ORG
Publication status / version	draft / version June 2000
Number of pages	44
Email address first author	Lpolos@fbk.eur.nl
Address	Erasmus Research Institute of Management (ERIM) Rotterdam School of Management / Faculteit Bedrijfskunde Erasmus Universiteit Rotterdam PoBox 1738 3000 DR Rotterdam, The Netherlands Phone: # 31-(0) 10-408 1182 Fax: # 31-(0) 10-408 9020 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	We investigate how sociological argumentation differs from the classical first-order logic. We focus on theories about age dependence of organizational mortality. The overall pattern of argument does not comply with the classical monotonicity principle: adding premises does not overturn conclusions in an argument. The cause of nonmonotonicity is the need to derive conclusions from partial knowledge. We identify meta-principles that appear to guide the observed sociological argumentation patterns, and we formalize a semantics to represent them. This semantics yields a new kind of logical consequence relation. We demonstrate that this new logic can reproduce the results of informal sociological theorizing and lead to new insights. It allows us to unify existing theory fragments and paves the way towards a complete classical theory.	
Library of Congress Classification (LCC)	5001-6182	Business
	5546-5548.6 5548.7-5548.85	Office Organization and Management Industrial Psychology
	6951	Industrial Sociology
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	M 10 L 2	Business Administration: general Firm Objectives, Organization and Behaviour
	L 29	Firm Objectives, Organization and Behaviour : Other
European Business Schools Library Group (EBSLG)	85 A	Business General
	100B 240 B	Organization Theory (general) Information Systems Management
	80 J	Industrial Sociology
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen
	85.05 85.08	Management organisatie: algemeen Organisatiesociologie, organisatiepsychologie
	85.08	Organisatiesociologie, organisatiepsychologie
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Organisatieleer, informatietechnologie, prestatiebeoordeling	
	Organisatiesociologie, Logica	
Free keywords	Logic of theory building, Non-monotonicity, Social Science Methodology, Organizational Mortality Respectively	
Other information		

Reasoning with Partial Knowledge[1]

László Pólos
Eötvös Loránd University, Budapest
and
Erasmus University, Rotterdam

Michael T. Hannan
Stanford University

June 9, 2000

[1]We appreciate the support of the Centre for Formal Studies in the Social Sciences at Eötvös Loránd University, ERIM Institute of Erasmus University, and the Stanford Graduate School of Business.

Abstract

We investigate how sociological argumentation differs from the classical first-order logic. We focus on theories about age dependence of organizational mortality. The overall pattern of argument does not comply with the classical monotonicity principle: adding premises does not overturn conclusions in an argument. The cause of nonmonotonicity is the need to derive conclusions from partial knowledge. We identify meta-principles that appear to guide the observed sociological argumentation patterns, and we formalize a semantics to represent them. This semantics yields a new kind of logical consequence relation. We demonstrate that this new logic can reproduce the results of informal sociological theorizing and lead to new insights. It allows us to unify existing theory fragments and paves the way towards a complete classical theory.

Reasoning with Partial Knowledge

> *Observed inferential patterns which seem 'wrong' according to one notion of inference might just as well signal that the speaker is engaged in correct execution of another style of reasoning.* (Johan van Benthem [1996])

1 Introduction

When instances of sociological theorizing are examined from the perspective of formal logic, a confusing picture emerges. The argumentation seems to be erroneous; the sets of explanatory principles used in different parts of theories seem inconsistent. This sort of impression easily leads formally minded people to conclude that the sociological theories are unsystematic and unreliable.

Several papers have recently challenged this view. Each selected relatively small fragments of theories of organizations, formalized them in classical first-order logic, and showed that the basic conclusions were logically sound. These formalization attempts might engender two objections. Some did not stick to the informal, verbal theory; these added new assumptions to derive the formal counterparts of the theorems. Others re-interpreted some theorems and assumptions so that the (classical) first-order consequence relation holds for the claimed theorems. Second, they examined only fragments—not complete theories. Nonetheless, this research shows that some fragments of sociological theory are consistent if their stated premises are augmented with explicit representations of certain (more or less justified) hidden assumptions. Furthermore, contradictions can be removed by careful limitations of the validity domains of the explanatory principles; and consistent and sound formalizations of the theories can be provided.

A key difficulty facing sociology originates from incompleteness. Typical theories are in flux. Carefully constrained explanatory principles are not (yet) available. Nonetheless, the argumentation still follows certain logical principles. We want to identify some of these principles and show how they operate in the normal routines of sociological theorizing.

Studying these questions in full generality entails a vast—perhaps impossible—task. Instead, we narrow the focus considerably. We examine theories in organization sociology concerning age dependence in mortality processes. Empirical research provided three different tendencies to be explained. Historically these facts were explained by separate theory fragments, each selected to explain (a subset of) the empirical findings.

Certain populations of organizations exhibit a negative relationship between organizational age and the hazard of mortality. Theory fragments were developed under the label of liability-of-newness theories to explain this relationship. In other populations, organizations appeared to be most vulnerable not at founding, but somewhat latter. A fragment, called liability-of-adolescence theory, was developed to account for this pattern, especially the initial rise of mortality. Finally, research on some other populations finds that older organizations have the highest hazard. Theory fragments that explain these finding are called theories of liability of obsolescence, and liability of senescence, and network saturation.

No one claims that the various theory fragments hold simultaneously. But, what should be expected of a not-yet-studied population of organizations? First-order logic does not give enough room to keep all of these theory fragments on board; at most one of them can be true. Hannan (1998) formalized these theory fragments in classical first-order logic. This effort was only partly successful. The resulting formulation integrated two—but not all three—of the fragments. In other words, the formalization yielded two internally consistent fragments that seemingly cannot be reconciled in first-order logic. Making sense of this situation and making headway in refining the theories would seem to demand a different approach. Perhaps this entails use of a logic that imposes less stringent constraints. We argue that nonmonotonic logic suits this situation well.

Use of nonmonotonic logic in theory building constitutes a substantial departure from long-standing practice in sociology. When sociologists examine the logic of argumentation in their field, they invariably employ (often informally) propositional logic or, increasingly, first-order logic. We argue that the predictive power of available knowledge increases if we assume that sociological argumentation in general—or, at least in this particular case—follows a nonmonotonic logic. In this context, monotonicity means that the set of conclusions that follow from a set of premises grows monotonically as premises are added. In other words, monotonicity means that adding premises does not overturn conclusions that follow from the original (smaller) set of premises. In contrast, nonmonotonic logics allow the addition of new premises (reflecting new knowledge) to overturn existing conclusions. In such nonmonotonic logics, introduction of premises that would result in contradictions according to a classical first-order logic do not necessarily create inconsistency. Switches between explanatory principles, without further field-specific stipulations, follow the generic guideline of nonmonotonic logic: (1) when different principles give conflicting results, base inferences on the most specific principles that apply; and (2) when conflicting principles do not differ in specificity, do not make any inference.[1]

To identify how nonmonotonicity arises in argumentation, we need to learn more about the mechanisms that activate and de-activate certain premises. We follow the following strategy:

- Take a close look at the arguments. Design the syntax of a formal language to represent their premises and the conclusions.

[1]Standard technical references on the subject include: McCarty (1980), Makinson (1994), and Veltman (1996); Brewka, Dix, and Konolige (1997) provide an accessible overview of the field of nonmonotonic logic.

- Identify the logical principles used in this type of argumentation and generalize them.
- Formalize these principles in a semantics and a semantically defined system of inference for the sentences of this language.
- Verify that application of the inference system yields the desired results and provides logically sound explanations for the empirical findings.
- Identify some novel implications of the explanatory principles. In other words, check whether the theory built in this new logic has increased explanatory power.

We leave to future research two tasks:

- Checking whether new results can be supported empirically.
- Explore the relation between this semantically defined inference system and other nonmonotonic logics.

2 Theories of Age Dependence

2.1 Liability of Newness

We pick up the line of history of relevant arguments with the famous paper of Stinchcombe (1965), which makes four distinct arguments for a liability of newness (LoN, for short).

Arguments in Stinchcombe (1965)

First, new organizations normally lack the technical and social requirements for smooth functioning: appropriate skills of members, appropriate decision criteria, sensible divisions of responsibilities, development of loyalty, and learning what can go wrong. In old organizations, the members can have learned the relevant specialized knowledge and can have developed loyalty to the organization. Even when members get replaced, incumbents can convey the necessary knowledge to newer entrants. In contrast, new organizations have to get by with the generalized skills produced outside the organization, which normally fit the organizational context less than perfectly, or they have to invest in training. Second, new organizations must invent roles and role relationships and structure rewards and sanctions. The need for much learning by doing lowers performance in young organizations, leading to higher hazards of mortality. Third, in new organizations, most social relations are relations among strangers. Thus new organizations pose considerably higher levels of uncertainty than old organizations. The need to compensate for such uncertainties takes away vital resources from the young organizations, and that, in turn, lowers their survival chances. Fourth, young organizations normally lack strong ties to external constituencies. This makes it harder to mobilize resources and ward off attacks.

Stinchcombe makes clear that this story applies best to the beginning of the industrial age or to the beginnings of the histories of particular industries and that it might not apply in contemporary societies. He mentions explicitly that the growing availability of general skills, socially induced capacities to learn, and ease of recruitment might eliminate this effect. So the LoN might be only an historical observation that need not (in principle) be supported by facts of the contemporary world. Nonetheless, researchers exposed the LoN theorem to testing in many populations of organizations. Much early research found evidence of a LoN (Carroll and Hannan 2000). Such evidence made relevant those arguments that, from one premise-set or another, pointed towards negative age dependence.

The empirical validity of the LoN story (at least with size not taken into account—see below) presents a theoretical dilemma. What makes the claim an empirically justified theorem? We see at least two possibilities. First, some research that covers the whole histories of organizational populations brings into the picture organizations that were founded in the premodern period that (in accordance with Stinchcombe's argument) were exposed to the LoN. Furthermore, the tendency toward industrial concentration gives extra weight to these numerous but small and vulnerable organizations in the early history of the population. According to this interpretation, the LoN is indeed a premodern phenomenon; but fossil records of the early history of organizational populations preserved it for us. Alternatively, the findings might not be limited to the pre-modern phase. In this case, the theorem requires a different explanation, one that applies to modern/contemporary organizations.

An Argument in Hannan and Freeman (1984)

One argument put forward by Hannan and Freeman (1984) concerns the continuing relevance of the LoN. It can be recapitulated as follows:

1. Selection favors reliable and accountable organizations.
2. Reliability and accountability require highly reproducible structures
3. Structural reproducibility rises with age.
4. To connect selection to organizational mortality it seems natural to assume that whenever particular type (T_1) is preferred over an other type (T_2) then the mortality hazard of organizations of type (T_1) is lower than that of organizations of the type (T_2)

These assumptions imply the LoN theorem (Péli, Pólos, and Hannan 2000). For the purposes of the present paper, this argument simply replaces one, more or less historical argument with one that applies to contemporary organizations as well.

2.2 Liability of Adolescence

After the LoN had been demonstrated empirically and successfully integrated into theory, the empirical picture became clouded. Researchers found evidence that the hazard starts low, then rises to a peak, and finally declines again as organizations age (Carroll and Hannan 2000). Some analysts called this pattern a liability of adolescence (LoA, for short). One plausible explanation for the pattern works as follows. New organizations are endowed; they begin with a given amount of capital (financial and social) and a position in a more or less given network of ties. Surviving the initial period does not depend on smooth functioning, because endowments provide (partial) immunity. As endowments get spent down, performance matters more and the hazard increases. The mortality hazard in a population peaks when most organizations have exhausted their endowments. From then on, the normal process takes over, and the hazard declines as organizations accumulate reliability, accountability, and organization-specific human capital.

2.3 Aging, Obsolescence, Senescence, and Network Saturation

Empirical work on organizational mortality began to show a more serious divergence during the past decade, depending upon how researchers handled organizational size. Freeman, Carroll, and Hannan (1983) noted that negative age dependence might reflect the operation of unobserved heterogeneity rather than aging and that organizational size, which was not measured in early studies, could serve as the heterogeneous force. When researchers found data sets that allowed them to control for initial sizes and subsequent growth/decline, more than half of the studies produced evidence of positive age dependence (Carroll and Hannan 2000). This change in empirical findings motivated consideration of processes that might generate positive age dependence.

Barron, West, and Hannan (1994) offered several explanatory schema. One emphasizes the alignment of organizations with environments. If organizations tend to be aligned with their environments at founding and intensifying structural inertia makes it less and less likely that they can keep pace with changes in the environment, then the quality of alignment worsens with passage of time—aging, in the metric used in standard analyses. Relatively poor alignment with environments elevates the hazard for old organizations—there is a liability of obsolescence (LoO).

A second story focuses on daily (more or less routine) procedures. As an organization ages, friction accumulates and procedures become slower and perhaps also less accurate. Without maintenance, routines tend to deteriorate; and the accumulated inertia makes routine maintenance both more difficult and more expensive to complete. Old organizations loose on speed and efficiency, and, after a while, they vanish. This pattern has been called a liability of senescence (LoS).

A third explanation builds on the saturation of possible ties (Barron 1992).

If extending the web of ties is vital for an organization (say, for the recruitment of new customers and key staff) saturation of the space of possible ties lowers vitality. The older the organization, the more likely that its own space of new possible ties has become saturated. Therefore, older organizations have higher mortality hazards.

3 Nonmonotonicity in Theory Building

We distilled some guidelines for both syntax and semantics from these examples. Theory building (at least in this case) conforms to the principle of *informational monotonicity:* explanatory principles are not withdrawn, even when their first-order consequences are falsified. Instead, they are maintained; and their effects are controlled by more specific arguments. Thus, explanatory principles clearly differ from classical first-order (universal) generalizations, because they show an informational stability. Our approach conforms to the spirit of Frank Veltman's (1996) language for update semantics. Like Veltman, we need three operators: two for the predictions (expectations and derived generalization) of the theory and one to indicate the genericity of explanatory principles.[2]

The details of our approach differ from previous developments in nonmonotonic logic in how we define the specificity orderings of arguments. As we explain below, we build formal models of arguments involving empirical generalizations in terms of sequences of intensions of open formulæ.

What do we expect from a suitable nonmonotonic logic? To develop some intuition, we begin with a simple inference pattern that shows a failure of monotonicity.[3]

Premise 1. Birds fly.

Premise 2. Tweety is a bird.

Proposition 1. Tweety flies.

Now consider the following:

Premise 1. Birds fly.

Premise 2. Tweety is a bird.

Premise 3. All penguins are birds.

Premise 4. Penguins do not fly.

Premise 5. Tweety is a penguin.

Proposition 2. Tweety flies.

[2]We also follow Veltman in assuming that these operators cannot be applied recursively; we require that they be the outermost operators in all formulæ in which they appear.

[3]This stylized example and the Nixon Diamond (discussed below) are ubiquitous in the technical literature.

Although we kept the original premises unchanged, we added three new ones. What happens to the conclusion that we found justified based on the first two premises alone? Not only do we find this conclusion unjustified, but we are tempted to derive the *opposite* conclusion:

Proposition 2. Tweety does not fly.*

What is going on here? Compare the two possible arguments about flying. One builds on a premise about birds; the other builds on a premise about penguins. Tweety is both a bird and a penguin; so both premises apply. But, the premise about penguins seems to be more relevant for Tweety than the premise about birds. Why? Given that all penguins are birds but not all birds are penguins, the premise about penguins is more specific than the premise about birds. This difference in the specificity of the premises accounts for the difference in the relevance of the arguments. We want to use the most relevant arguments available. So, we go for the conclusion: "Tweety does not fly."

Where does this (implicit) specificity ordering come from? We argue that it comes from the third premise "All penguins are birds." One might object that the presence/absence of this premise should not matter much, because our common background knowledge holds that penguins are birds. To see that common background knowledge does not always clarify inferences, consider the famous (in logic) Nixon Diamond.

Premise 1. Quakers are doves.

Premises 2. Republicans are hawks.

Premise 3. Dick is a republican.

Premise 4. Dick is a Quaker.

Proposition 1. ???

Our background knowledge does not inform us about the specificity of certain premises. Lacking a dependable specificity order, we cannot conclude either that "Dick is a dove" or that "Dick is a hawk."

Seeing all this, one might go back to the previous example and decide that the original conclusion was, perhaps, unjustified and that we were wrong to conclude that "Tweety flies."[4] "Birds fly" is a rule—a rule with exceptions; and we did not know whether Tweety was an exception. Although we should have waited until we learned something about this, we just jumped to the conclusion.

Well, sometimes we must draw conclusions before all relevant facts are known—there is an urgency of action (Descartes 1897–1913). We almost always face this kind of urgency in theory building. We cannot wait until we know all relevant things about the subject. We often take what we have and draw conclusions. In technical terms, this attitude is called the *closed-world assumption:* we argue from a set of premises as if these premises represent all that is known to be true. In fact, this very feature of theorizing accounts for the nonmonotonic nature of much argumentation in sociology and other fields.

[4]Without doubt, this would be the reaction of some classically minded logicians.

4 A Language for Sociological Theorizing

A logic is a formal definition of a consequence relation. In the case of theorizing, the relevant consequence relation holds among sentences of a language. So it is natural to start the formal characterization of the logic with a definition of the language. Let us consider first what kind of language we desire.

We are focusing on nonmonotonicity, the uncertainty that our conclusions really follow from the theory as such, even though they appear to be conclusions of a given stage of the theory. This possibility is frightening, and we would like to avoid it if we could. If these dangerous beasts are here to stay, we want to mark them clearly. We do so by defining separate languages for the well-behaved part of the theory, where the conclusions are safe (and, as such, re-usable in further derivations), and the part that misbehaves in the sense of showing signs of nonmonotonicity.

Two kinds of sentences generate failures of monotonicity:

- empirical generalizations (rules with possible exceptions), and
- provisional theorems (theorems derived from generalizations).

The difference between them becomes clear when we consider their roles in theorizing. Serious empirical generalizations capture some relevant (and valid) insight. These are the sentences that yield the "a ha!" feeling that good theories can provide. Neither their validity nor their relevance gets automatically undermined by any accidental counter-example. In fact, the meaning of empirical generalizations cannot be properly reproduced by their truth conditions, expressed in term of the number or proportion of (positive) instantiations. They do not talk about what *is* the case. Rather, their main semantic contribution lies in shaping our expectations.

We need extra tools to express expectations. Because they provide insight, generalizations should be treated as *informationally stable* in the sense that extensions of the theory keep them intact. All that might happen to generalizations as a theory develops is that new insights restrict their domains of applicability. But, this happens only when knowledge about exceptions develops. Even new knowledge should not led us to update the empirical generalizations. New rules get added to the body of knowledge; and new, more specific rules can—and occasionally do—override older, more general rules.

Provisional theorems (propositions derived from generalizations), on the other hand, have a haphazard existence. New knowledge might wipe them out without a trace. They belong to particular stages of a developing theory—they represent the predictions that can be sensibly formed at that stage.

These considerations set a methodological agenda. The "official" empirical generalizations of a theory should be restricted to those that the theorist regards as dependable and insightful enough to be accepted as permanent assumptions of the theory. If there are doubts about the future acceptability of a generalization, then it does not deserve the status of (official) empirical generalization.

In the context of an evolving theory, part of the theory can normally be expressed in the language of first-order logic (FoL), which follows monotonicity. For instance, definitions and strict generalizations (universally quantified sentences) are expressed in this way. Assumptions of this sort provide the firm foundation for a theory. Any sentence that is logically derivable (according to the rules of classical FoL) from these assumptions counts as a dependable proposition of a theory. In formal terms, the deductive closure of the assumptions is part of a theory.

We add a new logical form to represent "serious" empirical generalizations. Study of the linguistic forms of the sentences that express empirical generalizations shows that it unlikely that they can be expressed in the language of first-order logic. *Generic* sentences are general, but not universal. Although sentences such as "Birds fly" are true or false, their truth conditions cannot be expressed in terms of truth and falsity about the flying ability of individual birds or bird species. Even if most birds do not fly, the sentence would still be an acceptable generalization, provided that this is a justifiable expectation of a creature with only one known property: it is a bird. Empirical generalizations cannot be expressed adequately in first-order logic, where generality means universal quantification, because any exception falsifies a universally quantified sentence.

In the strategy we propose, statements of (official) empirical generalizations become part of a theory. However, their role differs from that of classical (first-order) sentences. Although the consequences of first-order sentences are necessarily part of a theory, the consequences of empirical generalizations are not.

To learn what a theory predicts while still in progress, certain tests should be run on the (actual stage of the) theory. In formal terms, these tests are nonmonotonic inferences. Formalizing such inference requires a different logical form. Nonmonotonic tests concern predictions (or expectations) based on a particular stage of a theory. Such predictions come in two varieties. They can express that an object, or a sequence of objects, *normally* has a certain (perhaps-complex) property. Alternatively, they can state that one of two (possibly complex) properties *presumably implies* the other.

4.1 Syntax

We define the language (specifically, its syntax) to specify the well-formed sentences (or formulæ). Next, we assign meaning to the well-formed sentences. Finally, we spell out the consequence notion that fits this semantics. We assume familiarity with the language of first-order logic ($\mathcal{L}_F$, for short) at the level of such standard texts as Barwise and Etchemendy (1993). (Complete details of all of the formal structures are spelled out in Pólos and Hannan (2000a,b). Mastery of these technical details is not needed for understanding the general method and its applications. The technical issues are of interest mainly to those who might want to design alternative logics for the kinds of sociological applications we consider or who want to contrast our scheme with other applications in nonmonotonic logic.)

It is important to distinguish between open and closed sentences. To fix ideas, we give a small example. Suppose we define a one-place predicate $Org(\cdot)$, where $\cdot$ is the placeholder in the argument slot. $Org(x)$ is a sentence, which reads as "some object x is an organization." Such a sentence holds that an object possesses a property. As stated, the truth of this sentence cannot be established, because we have not identified the object in question. In technical terms, the variable x is free (not bound), and the sentence $Org(x)$ is an open sentence (or formula). In $\mathcal{L}_F$, we can get a closed sentence from an open formula in several ways. We can replace the variable with the proper name of an object (or individual constant), e.g., $Org(Intel)$. In this case, we check the truth of the sentence by examining the named object to ascertain whether it is in fact an organization. Alternatively, we can quantify over the variable x (thereby binding it) by forming sentences such as $\exists x[Org(x)]$ which reads as "some object (in the universe of discourse) possesses the property of being an organization" or $\forall x[Org(x)]$, which reads "every object (in the universe of discourse) has the property of being an organization." In each of these three revised sentences, all variables are bound and the sentences are said to be closed.

In sociological applications, we are usually interested in connections between predicates. Suppose that we have in hand the predicates $Young(\cdot)$ and $Rel/Acc(\cdot)$ which state the properties of "being young" and "having reliability/accountability." Many relevant sociological propositions assert a relation of material implication: if A then B. In formal terms, we use the non-logical constant $\rightarrow$ to express material implication: $A \rightarrow B$ (or, alternatively, $\neg[A \wedge \neg B]$). A typical open sentence expressing a relation of implication might be

$$Young(x) \wedge Org(x) \rightarrow \neg Rel/Acc(x).$$

A closed version can be gained by, for example, universal quantification:

$$\forall x[Young(x) \wedge Org(x) \rightarrow \neg Rel/Acc(x)],$$

The revised sentence states that it is the case for every object (in the universe of discourse) that if the object is both young and an organization then it does not possess the property of reliability/accountability.

In defining syntax (and, later, semantics), we refer to sentences (or formulæ) with shorthand expressions such as ϕ and ψ, where ϕ and ψ refer to some particular sentences. For example, we might express the last mentioned sentence as

$$\forall x[\phi \rightarrow \psi],$$

where ϕ stands for $Young(x) \wedge Org(x)$ and ψ stands for $\neg Rel/Acc(x)$.

The Language of Working Theory ($\mathcal{L}_W$)

We construct the language of working theories ($\mathcal{L}_W$, for short) as an extension of $\mathcal{L}_F$. We introduce new types of sentences to express empirical generalizations by adding (1) a nonlogical constant: a generic quantifier, which we denote as $\mathcal{Q}$, and (2) a new type of clause to the definition of sentences. The key idea is that we define generic sentences: sentences formed by applying the generic quantifier to link two open sentences that share all of their free variables.

Definition 1 (Language of Working Theory) $\mathcal{L}_W$ is the smallest extension of $\mathcal{L}_F$ such that, for any pair of open sentences ϕ, ψ, the language contains a new type of closed formula $\mathcal{Q}\bar{x}[\phi \rightarrow \psi]$, where $\bar{x}$ denotes the set of free variables that ϕ and ψ share.[5] This novel kind of formula can be read as "ϕs are ψs," or, with an emphasis on the nonmonotonicity, "ϕs normally are ψs" or "ϕ normally implies ψ."

The quantifier that creates logical forms for empirical generalizations must always be the outmost operator of formulæ. This means that the generalization-forming operation is not recursive. Empirical generalizations pertain to the world, not to (other) empirical generalizations.

The Language of Theory Testing ($\mathcal{L}_T$)

$\mathcal{L}_T$ is also defined as an extension of $\mathcal{L}_F$. We are interested in two kinds of nonmonotonic conclusions: derived generalizations and nonmonotonic conclusions, predictions for the objects (or for sequences of objects) in the domain of the theory. We add two logical constants (Λ and   ) and two new types of clauses to the definition of well-formed sentences. The logical constant Λ serves as the "is likely to be" connective, and   is the "presumably implies" connective.

Definition 2 (The language of theory testing) $\mathcal{L}_T$ is the smallest extension of $\mathcal{L}_F$ that satisfies the following conditions:

1. If ϕ is a first-order sentence with n free variables and $a_1, \ldots, a_n$ are individual constants (proper names), then $\Lambda\,[\phi(\langle a_1, \ldots, a_n\rangle)]$ is a closed formula that reads as "the sequence $\langle a_1, \ldots, a_n\rangle$ is likely to satisfy the open formula ϕ." We use open formulæ to express (sometimes-complex) properties. A looser, but intuitive, reading is "$\langle a_1, \ldots, a_n\rangle$ is likely to have the property ϕ."

2. If ϕ and ϕ are first-order wffs with exactly the same free variables, then $(\phi \hookrightarrow \psi)$ is a closed formula of $\mathcal{L}_T$. $(\phi \hookrightarrow \psi)$ reads as "ϕ presumably implies ψ" or "ϕs are presumably ψs."

Our use of the expressions "normally implies," "is likely to," and "presumably implies" reflects a substantial influence of Veltman's (1996) brilliant "Defaults in Update Semantics," which offers formal semantics for the expressions "normally implies," "presumably," and "presumably implies." Even though our research questions led us to a different formal semantics, the credit for developing some basic, dependable intuitions about a domain, which is normally recognized as very slippery, should go to Veltman.

[5] We assume that there are no free variables that they do not share and that there is at least one free variable that they do share.

4.2 Semantics

Classical First-Order Semantics

Assigning meaning to sentences in $\mathcal{L}_F$ requires specifying: what objects the individual constants refer to, what the variables stand for, and what properties/relations the predicates denote.[6] Individual constants refer to—and variables stand for—elements of the universe of discourse ($\mathcal{U}$). One-place predicates refer to properties that can be represented as subsets of the universe of discourse. Two-place predicates refer to relations represented by subsets of pairs of objects/elements of $\mathcal{U}$, three-place predicates to relations represented by subsets of triplets of objects/elements, and so forth.

We can define the semantics in a systematic manner in terms of interpretation functions.

Definition 3 (Interpretation functions) A function defined on the set of predicates and names, ρ, is an interpretation if and only if (iff) it maps

1. individual constants to elements of the universe of discourse, and
2. n-place predicates (elements of P^n) to sets of n-long sequences of elements of the universe of discourse. (The n-long sequences of the elements of $\mathcal{U}$ are elements of the n-th power of that universe, $\mathcal{U}^n$.)

We define the language (syntax) to specify the well-formed sentences (formulæ). Next we assign meaning to them. Finally, we spell out what kind of consequence notion this semantics implies.

For example, *Org* is a one-place predicate. The interpretation function assigns to it the set of objects that—according to an interpretation—qualify as organizations. This set is called the *extension* of the predicate. Extensions obviously depend upon interpretations. *Acquired* is a two-place predicate; and the interpretation function assigns to it a set of pairs of objects, those pairs for which the first component of the pair acquired the second component (according to the interpretation). This example shows that the order of the components matters, $\langle BMW, Rover\rangle$ might be in the extension of the predicate *Acquired* even when $\langle Rover, BMW\rangle$ is not in the extension, in the natural interpretation, at the present.

Among the nonlogical expressions, only the variables do not get interpreted. Just as pronouns get their denotations from ostensive actions, variables have to be valuated.

Given an interpretation function and a valuation of variables, we can calculate truth values of sentences. (We use 1 for truth and 0 for falsity.)

Definition 4 (Truth values of the sentences of $\mathcal{L}_F$)

[6] Valuable background material on the semantics we propose can be found in Dowty, Wall, and Peters (1980) and Gamut (1991).

1. The value of a term is given either by an interpretation function or by a valuation. (If a is an individual constant, then it is interpreted by ρ. If it is a variable, then it is valuated by v.)

2. An identity statement is true iff both terms refer to the very same element of the universe of discourse.

3. The sentence stating that a sequence of objects possesses a property, e.g., "Org (BMW)", is true iff the objects actually do possess the property, i.e., BMW is indeed an organization.

4. The negation of a sentence is true iff the sentence is false.

5. The conjunction of two sentences is true iff both conjuncts are true.

6. The disjunction of two sentences is true iff at least one of the disjuncts is true.

7. A material implication is false iff the antecedent is true and the consequent is false.

8. A universally quantified sentence is false iff there is a counter-example to it.

9. An existentially quantified sentence is true iff the embedded sentence has a positive instance.

According to this definition, truth/falsity depends upon the choice of interpretation. But, not all interpretations are equally useful; some might have very little to do with the real world. We need factual knowledge to tell which sets the predicates denote. It is not logic that tells which creatures are organizations or which creatures have acquired which other creatures. One has to go into the world and find out. If no factual information is available, then all denotations are equally possible. In other words, we have the whole set of different interpretations—possible worlds, in the usual formal language—but we know nothing about which of these possible worlds is the *actual* world. When we learn some relevant facts, the set of (still) possible worlds is made smaller.

A Possible-Worlds Semantics for $\mathcal{L}_F$

The next step is to tell what sets the predicates denote in the various possible worlds. A generalized interpretation of the language assigns references to individual constants and denotation to all predicates in *every* possible world. We can use the interpretation function to characterize the semantics of a set of the possible worlds.

Definition 5 (Interpretation functions on possible worlds) The interpretation function for possible worlds is a function defined on the set of nonlogical constants (predicates and names) that defines the extensions of the relevant predicates in every possible world. (We assume that each individual constant denotes the same element of the universe in all possible worlds.)

The concept of *intension* plays a very important role in possible-world semantics. The intension of a predicate is the function that tells the extension of the predicate (the set of those objects for which the predicate is true) in every possible world. According to the conditions given above, the interpretation function assigns intensions to predicates. Given an interpretation, we can assign truth-values to all first-order sentences in every possible world, according to one valuation or another.

We defined empirical generalizations as generic sentences. As we noted above, a generic sentence is either true or false; but, its truth/falsity generally cannot usefully be expressed in terms of the proportion of positive and negative instances in a world (the "inductivist" view). We consider a generic sentence to be true if the regularity it expresses is present in the world. Carlson (1988: 33) calls this view the "rules-and-regulations" approach:

> According to this approach, generic sentences depend for their truth and falsity upon whether or not there is a corresponding structure in the world, structures being not the episodic instances but rather the causal forces behind those instances.

To provide a formal model for the presence of such regularities, we take advantage of linguistic knowledge. The linguistic research concluded that the underlying structure of generic sentences contains two open formulæ with identical free variables (Carlson 1977; Diesing 1988; Kratzner 1988). Our strategy is based on the view that the ordered pair of the intensions of (two) open formulæ provides a useful semantic representation of the regularity that makes an empirical generalization true.

We have to define the intensions whose composition yields a generalization. We do so by extending the definition of intensions to apply to open sentences (as well as predicates):

Definition 6 (Intensions of open sentences) The intension of an open sentence is the function that gives, for every possible world, the set of those sequences of objects for which this particular (open) sentence is true in that world.

Definition 7 (Rules and regularities) We express empirical rules (generalizations) as generically quantified sentences. These sentences are composed of two open formulæ with the same free variables. We model the regularity expressed by the generalization as the ordered pair of intensions of these open formulæ.

At a given stage of theory building, all the representations of its empirical rules (generalizations)—and nothing else—are collected into the set $\mathcal{G}$.

Theory building reflects learning. Such a process depends upon our becoming aware of more universal rules and empirical rules (generalizations). Adding universal rules reduces the possibilities about how the "real" world might be, the set of still-possible worlds shrinks. Learning a new empirical generalizations does not have the same effect. Following Carlson's rules-and-regulations approach, we assume that learning a new empirical generalization means only that we add a new item to the stock of empirical-rule representations.

Semantics for $\mathcal{L}_W$

With these preliminaries in hand, we can represent a stage of a theory.

Definition 8 (Stage of a theory) The stage of a theory is a pair of sets: a set of (still) possible worlds $\mathcal{X}$, the worlds that can be the real world as far as the (first-order part of the) theory is concerned, and the set $\mathcal{G}$ representing empirical generalizations.

The idea that a stage of a theory has both of these components corresponds to the view that both first-order premises[7] and empirical generalizations matter. We consider generalizations to be just as real as objects, properties, or relations.

Sentences in $\mathcal{L}_W$ will be evaluated in such stages. The sentences of the first-order part of the theory (the classical premises) can be true, false, or undefined in any stage. Intuitively speaking, a first-order sentence is true in a stage if it is true in all of the still possible worlds, false if it is false in all still possible worlds, and undefined otherwise.

As we pointed out above, an empirical rule or generalization ought to be modeled semantically by the presence of the corresponding regularity, because its truth or falsity cannot be characterized in terms of a true (universally quantified) proposition about the world. Instead of focusing on truth, we treat sentences expressing generalizations as true if they represent explicit generalizations of the theory and as false otherwise. Their logical impact will be characterized by the role that they play in inference, i.e., in terms of the conclusions they (together with first-order premises, of course) support.

Definition 9 (Truth conditions for empirical generalizations) For empirical generalizations (sentences of $\mathcal{L}_W$), $\mathcal{Q}\ \bar{x}[\phi \rightarrow \psi]$ is true in a stage of a theory if the pair of intensions of ϕ and ψ are in $\mathcal{G}$ in that stage, and it is false otherwise.

As we construct the situation, an empirical generalization is unambiguously either true or false at any stage of a theory. It is important to realize that universal rules and empirical generalizations are stable elements of the theory. If they are true in a given stage of the theory, then they remain true after any extensions of the theory. But, as we pointed out earlier, their consequences can change from stage to stage.

Semantics for $\mathcal{L}_T$

Because we lack a rich semantics for empirical generalizations, we leave it to the semantics of the sentences of $\mathcal{L}_T$ to characterize the basic intuitions about inference from generalizations. Loosely speaking, a nonmonotonic test will succeed if we can construct a tentative—but convincing—argument from the generalizations. We call an argument based on generalizations a *rule chain.* A test will

[7] It is worth noting that so-called meta-considerations can also be expressed as first-order premises. Such considerations include background information such as rules of arithmetic, set theory, and so forth.

succeed if we have such a tentative argument that is more specific than all of the (tentative) counter-arguments.

Rules in the chains represent empirical rules or universal rules (first-order premises, definitions, and meta-considerations). Because strict comparisons of rule chains are needed to test arguments, we have to define carefully the proper construction of the chain: which rule can follow which other rule in the chain.

As a preparatory step, we need to define a transitive and reflexive specificity relation on intensions. We want to relativize the notion of specificity to *information states.* Things a theory tells about the real world include information about the specificity of the empirical generalizations. The semantics of the theory represent the empirical generalizations as pairs of intensions (γ, γ'), with the first component being the intension of the antecedent and the second component being the intension of the consequent.

Think of the following empirical generalization from Barron, West and Hannan (1994).

Old and small organizations are vulnerable.

The formal counterpart of such a generalization is

$$\mathcal{Q}x[Org(x) \wedge Old(x) \wedge Small(x) \rightarrow Vulnerable(x)]$$

In this case, the antecedent is the formula $Org(x) \wedge Old(x) \wedge Small(x)$. The intension of this formula is the function that tells for every possible world which objects are the old and small organizations in that world.

Suppose we have more information about a set of organizations that we believe to be an exception to this rule, say handcraft producers (in some industry).

All craft producers are old and small organizations:
$\forall x\,[Craft(x) \rightarrow Org(x) \wedge Old(x) \wedge Small(x)]$;

Craft producers are normally not vulnerable:
$\mathcal{Q}x\,[Craft(x) \rightarrow \neg Vulnerable(x)]$.

What should we think of craft producers based on the new information state? If this is all that we know, then we cannot conclude that the extension of the predicate 'Craft (producer)' is smaller than that of the complex predicate 'old and small organization' in the actual world. It might be smaller, or it might be equal. Yet, we are convinced that the 'craft producer' rule is more specific for any craft producer than the 'old and small organization' rule. If so, the relation between the two rules is this. In all still-possible worlds, the extension of the 'craft producer' predicate is smaller than or equal to the extension of the 'old and small organization' predicate; and the extension of the 'craft' predicate is smaller in some worlds (where one sees small, old liberal arts colleges, say). Exactly this relation is what we capture by the definition of the specificity relation, which we now discuss.

The specificity relation for pairs of intensions, denoted by $a \sqsubseteq_{\mathcal{X}} b$, is relativized to the set of (still) possible worlds at a stage of a theory, $\mathcal{X}$, such that $\mathcal{X} \subseteq \mathcal{W}$, as follows.

Definition 10 (The specificity relation for (still) possible worlds) $a \sqsubseteq_{\mathcal{X}} b$ iff $(a \rightarrow b)$ is true in all still-possible worlds, the elements of $\mathcal{X}$.

To define rule chains, we start with a given (sequence of) object(s) or a (perhaps complex) predicate. The first component of the chain, γ_1, identifies the subject of the argument. If the argument concerns a particular individual (or sequence of individuals), then $\langle a_1, \ldots, a_n \rangle$ γ_1 is the singleton whose only element is the reference of that individual (or the sequence of the reference of individuals) $\{\langle a_1, \ldots, a_n \rangle\}$. If the subject is a type of individual, then γ_1 is the set of objects of that type. Comparison of first elements in different chains tells only whether the arguments concern the same subject. Unless all of the chains have identical first elements, it does not make sense to proceed. If they are the same, then we want to begin comparisons of rule chains by comparing the specificity of their *second* elements (γ_2).

A rule chain can be extended beyond its first element in any of the following ways:

- The next two elements in the chain are the two components of the semantic representation of an empirical generalization, such that the first element of this component stands in the $\sqsubseteq_{\mathcal{X}}$ relation with the element that precedes it.
- The last element in the chain stands in the $\sqsubseteq_{\mathcal{X}}$ relation with the element that precedes it.
- The last element of the chain is the intension of the (perhaps complex) predicate that we try to test on the first element of the chain.

Positive chains correspond to (tentative) arguments, negative chains to (tentative) counter arguments.

Definition 11 (Minimal rule chains) A rule chain of length k, $\langle \gamma_1, \ldots, \gamma_i, \ldots, \gamma_k \rangle$, is minimal if $\langle \gamma_1, \ldots, \gamma_{i-1}, \gamma_{i+1}, \ldots, \gamma_k \rangle$ is not a rule chain for all $1 < i < k$.

Definition 12 (More-specific rule chains) The rule chain $\langle \gamma_1, \gamma_2, \ldots, \gamma_k \rangle$ is more specific than $\langle \gamma'_1, \gamma'_2, \ldots, \gamma'_l \rangle$ if $\gamma_2 \quad_{\mathcal{X}} \gamma'_2$ and $\neg(\gamma'_2 \quad_{\mathcal{X}} \gamma_2)$ or if $k < 3$.

The machinery of FoL allows all sorts of lengthening of arguments (simply by using tautologies). Adding such irrelevant material to a line of argumentation would destroy the possibility of comparing arguments sensibly with respect to their specificity as measured by the length of the chain of argument. We defined minimal rule chains such that the irrelevant (though sound) logical part is cut out. All FoL arguments can be reduced to two-element chains in all stages of a theory. If the argument is first-order, then there is a $\sqsubseteq_{\mathcal{X}}$ relation between the first and the last element of the chain for any $\mathcal{X}$ set of possible worlds.

Now we have all of the ingredients needed to define the semantics for the language of working theories. We want to express the semantics for Λ (the "likely to possess" relation) and $\hookrightarrow$ (the "presumably implies" connective).

Definition 13 (Semantics for $\mathcal{L}_T$)

1. $\Lambda\,[\phi(\langle a_1, \ldots, a_n\rangle)]$ is true in a stage of the theory iff:

   (a) there exist minimal positive two-element $a_1, \ldots, a_n$–ϕ rule chains, or

   (b) there exist minimal positive $a_1, \ldots, a_n (n > 2)$–ϕ rule chains, and

   (c) if there also exist minimal negative $a_1, \ldots, a_n (n > 2)$–ϕ chains, then all minimal positive chains must be more specific than all minimal negative chains.

   $\Lambda\,[\phi(\langle a_1, \ldots, a_n\rangle)]$ is false otherwise.

2. $(\phi(x_1, \ldots, x_n) \hookrightarrow \psi(x_1, \ldots, x_n))$ is true in a stage of the theory if:

   (a) there exist minimal positive two-element ϕ–ψ chains, or

   (b) there exist minimal positive ϕ–ψ chains, and

   (c) if there also exist minimal negative ϕ–ψ chains, then all minimal positive chains must be more specific than all minimal negative chains.

   $\phi(x_1, \ldots, x_n) \hookrightarrow \psi$ is false otherwise.

We can always construct two-element chains for statements that are logically true, making Λ and $\hookrightarrow$ tests succeed for them.[8] Such a first-order argument (two-element chains) will always overrule tentative arguments (minimal chains having at least three elements).

Figure 1 illustrates a pair of rule chains that give rise to a Nixon Diamond. The first element in each chain is the singleton γ_1. The second elements are the sets γ_2 and γ_2', and the third elements are γ_3 and γ_3'. The "funnels" connecting the sets represent the regularities that connect the complex properties, and the heavy arrows represent the conditionalities in the arguments. Here the two rule chains differ in their implications: the upper one leads to the conclusion $\neg\psi$ and the lower one to ψ. The two chains have the same length and their specificity cannot be compared (neither γ_2 nor γ_2' is a proper subset of the other). Therefore, our strategy of inference does not generate a conclusion in this case.

Figure 2 illustrates the Penguin Principle. It shows a case in which the inference strategy does produce a conclusion even though two lines of argument disagree. The situation is just as in figure 1 except that γ_2 is a proper subset of γ_2'; in other words, the argument represented in the lower chain is more specific than the opposing argument in the upper chain. Therefore, we draw the conclusion ψ.

[8] A sentence is logically true if there is no interpretation that would make it false in any possible world.

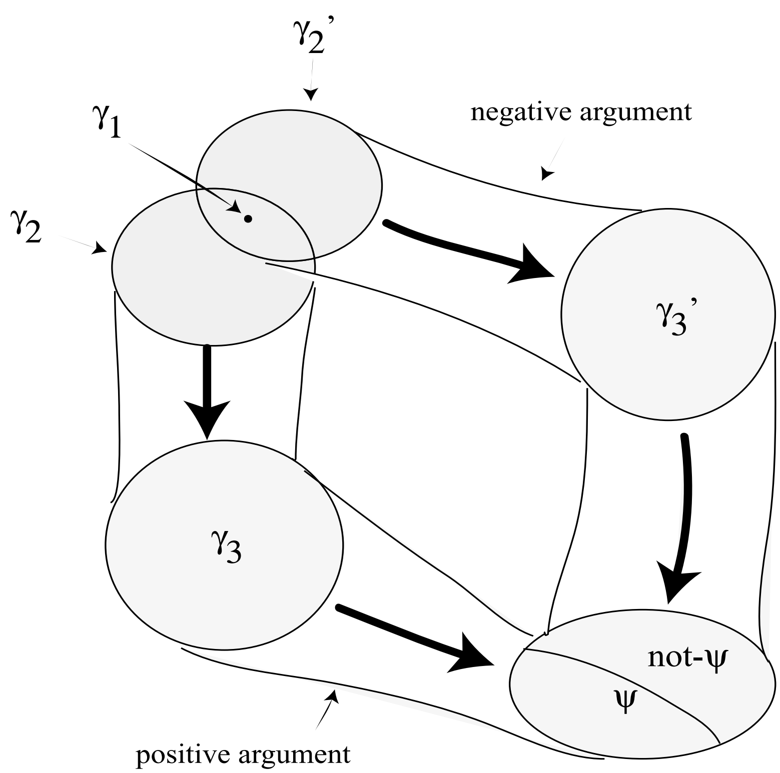

Figure 1: Illustration of the rule chains for an argument containing a Nixon Diamond

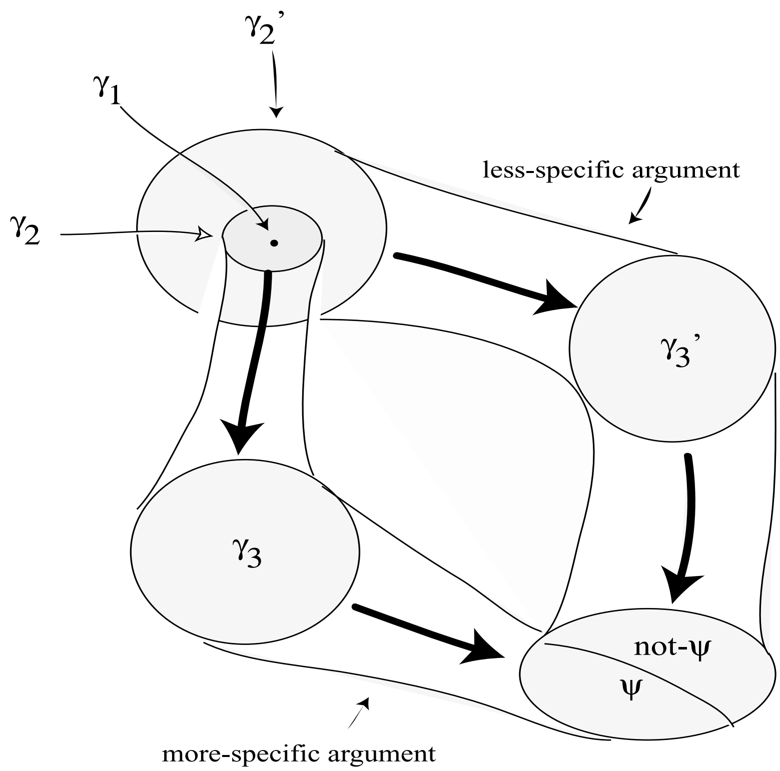

Figure 2: Illustration of rule chains for an argument subject to the Penguin Principle

Inference Within Working Theories

The inferencing that seems most relevant to sociological work starts with premises expressed in $\mathcal{L}_F$ and $\mathcal{L}_W$ and builds conclusions expressed in $\mathcal{L}_T$. To make sense of such argumentation, we need to define the semantic consequence relation.

Definition 14 (Rule-minimal stage of a theory) Let Π_F be a set of premises expressing universal (first-order) sentences and Π_W be the premises expressing empirical generalizations, i.e., sentences of $\mathcal{L}_W$. A $\langle \mathcal{X}, \mathcal{G} \rangle$ stage of the $\Pi_F \bigcup \Pi_W$ theory is a rule-minimal stage if $\mathcal{G}$ is the intersection of all those representations of empirical generalizations that make all elements of Π_W true.

Definition 15 (Implication relation for working theories) Let ϕ be a well-formed formula.
ϕ is a consequence of $\Pi_F \bigcup \Pi_W$ iff all the rule-minimal stages of the theory $\Pi_F \bigcup \Pi_W$ make ϕ true.

For a formula ϕ in $\mathcal{L}_F$, logical implication depends only on the first-order premises (giving the classical logical implications). In contrast, an empirical generalization ($\phi \in \mathcal{L}_W$) is derivable only if it is already present in the premise set.

Examples[9]

The first example illustrates that conclusions drawn from empirical generalizations are "weaker" than conclusions drawn from universally quantified sentences.

> Young organizations face a LoN:
> $\mathcal{Q}x[Org(x) \wedge Young(x) \rightarrow LoN(x)]$;
>
> HiTech is a young organization:
> $(Org(HiTech) \wedge Young(HiTech))$.

This pair of assumptions does not imply that 'HiTech faces the liability of newness', i.e., $LoN(HiTech)$. But, it does imply that 'HiTech is likely to face the liability of newness,' i.e., $\Lambda\, [LoN(HiTech)]$.

In this case, the present stage of the "theory" consists of one fact: 'HiTech is a young organization;' and the rule-minimal version has only one empirical generalization: 'Young organizations face the liability of newness.' In this situation, no two-long chain connects $HiTech$ and LoN. But, there is one three-long chain, which starts with the singleton $\{HiTech\}$. The second element is the intension of the formulæ $(Org(x) \wedge Young(x))$, $(Org(x) \wedge Young(x))$, and the third and last element in the chain is the intension of $LoN(x)$. This is the only minimal chain, and it is positive. So the test with $\Lambda\, [LoN(HiTech)]$ succeeds.

The second example shows that facts are also likely things. Consider the following argument, based loosely on Sørensen and Stuart (2000).

[9] Jaap Kamps helped us by working out similar examples for a related paper (Pólos and Kamps 1998).

Young organizations are innovative:
$\mathcal{Q}x[(Org(x) \wedge Young(x)) \rightarrow Innov(x)]$;

HiTech is a young organization:
$Org(HiTech) \wedge Young(HiTech)$;

HiTech is not innovative:
$\neg Innov(HiTech)$.

This trio of assumptions implies that

HiTech is not innovative:
$\neg Innov(HiTech)$;

HiTech is likely not to be innovative:
$\Lambda[\neg Innov(HiTech)]$.

Now we can construct a two-long chain from *HiTech* to *Innov*, and this chain is positive. Obviously, any two-long chain is more specific than any longer chain. So the test with $\Lambda[\neg Innov(HiTech)]$ succeeds.

The next example (based loosely on Baron, Hannan, and Burton (2000)) returns to the Penguin Principle.

Young organizations are innovative:
$\mathcal{Q}x[(Org(x) \wedge Young(x)) \rightarrow Innov(x)]$;

Young autocracy-based organizations are not innovative:
$\mathcal{Q}x[\mathit{Autocracy\text{-}based}(x) \wedge Young(x)) \rightarrow \neg Innov(x)]$;

All autocracy-based organizations are organizations:
$\forall x[\mathit{Autocracy\text{-}based}(x) \rightarrow Org(x)]$;

HiTech is a young autocracy-based organizations:
$\mathit{Autocracy\text{-}based}(HiTech) \wedge Young(HiTech)$.

This set of assumptions implies that

HiTech is likely not to be innovative:
$\Lambda[\neg Innov(HiTech)]$.

There are two minimal HiTech–innovative chains. The one whose second element is the intension of the formula $(\mathit{Autocracy\text{-}based}(x) \wedge Young(x)$ is more specific than the one whose second element is the intension of $(Org(x) \wedge Young(x)$. The more-specific minimal chain is positive, therefore the test with $\Lambda[\neg Innov(HiTech)]$ succeeds.

Finally, we consider a Nixon Diamond:

Young organizations are innovative:
$\mathcal{Q}x[(Org(x) \wedge Young(x)) \rightarrow Innov(x)]$;

HiTech is an autocracy-based organization:
$Autocracy\text{-}based(HiTech)$.

HiTech is a young organization:
$Org(HiTech) \wedge Young(HiTech)$.

Autocracy-based organizations are not innovative:
$Qx[Autocracy\text{-}based(x) \rightarrow \neg Innov(x)]$.

This set of propositions does *not* imply either that

HiTech is likely not to be innovative:
$\Lambda[\neg Innov(HiTech)]$, or

HiTech is likely to be innovative:
$\Lambda[Innov(HiTech)]$.

There are two minimal 'HiTech'–'innovative' chains. The one whose second element is the intension $(Autocracy\text{-}based(x)$ is not more specific than the one whose second element is the intension of $(Org(x) \wedge Young(x)$, nor the other way around. Therefore, neither the test with $\Lambda[\neg Innov(HiTech)]$ nor the test with $\Lambda[Innov(HiTech)]$ succeeds.

5 Back to Sociology

This formal nonmonotonic logic for theory building might be interesting for its own sake. But, the really important test is this: Does this novel formal approach contribute to our understanding of the sociological phenomena? Here we try to show that it does, by extending Hannan's (1998) formalization of a series of alternative formulations of age dependence. The key extensions involve unifications of the various fragments.

As we pointed out above, use of this kind of default reasoning relies heavily on considerations of specificity of the relevant arguments. We use an intuitive notion of specificity. The antecedents of the premises (partially) describe positions of the (a_1, a_2) interval on the scale of possible age: say the real line. Suppose that the antecedent of one generalization says, for example, that a_2 has to be in a given set, and the antecedent of another empirical generalization says that a_2 has to be within a *proper subset* of the same given set. Then arguments that use the second empirical generalization are more specific than those that use the first. Formalizing what is required to show that this intuition is correct would not add much to the understanding of the issues we are addressing here. Therefore, we simply assume here that these intuitions provide sufficient guidelines to decide specificity issues.

5.1 Liability of Newness

In what follows, we treat theory fragments that claim the LoN as one theory, even though there are at least six clearly distinguishable arguments. We do so because logical analyzes using each yields similar results.

5.1.1 Reliability/Accountability (Fragment 1)

One way to get the LoN claim uses premises concerning reliability/accountability (Hannan and Freeman 1984):

Premise $\mathbf{F}_1$.1 Reliability/accountability normally increases with age at all ages:

$$\mathcal{Q}x, a_1, a_2[(A(x, a_1) < A(x, a_2)) \rightarrow RA(x, a_1) < RA(x, a_2)].$$

Premise $\mathbf{F}_1$.2 Increases in reliability/accountability normally lower the hazard:

$$\mathcal{Q}x, a_1, a_2[RA(x, a_1) < RA(x, a_2) \rightarrow M(x, a_1) > M(x, a_2)].$$

These premises do indeed imply the strong-form version of the LoN claim in the following format

Proposition $\mathbf{F}_1$.1 The hazard presumably declines with age at all ages:

$$(a_1 < a_2) \hookrightarrow M(x, a_1) > M(x, a_2).$$

Proof of $\mathbf{F}_1$.1 To show that this theorem goes through in our new logic one has to consider all the rule chains that start with the intension of the (open) formula $A(x, a_1) < A(x, a_2)$ and end in the intension of $M(x, a_1) > M(x, a_2)$. At this stage of the theory, there is only one such rule chain; and this rule chain is positive. ■

We treat these premises and the proposition that follows as the default theory. In other words, this is the first stage of the theory. These premises will be included in every subsequent stage. Notice that, because Proposition F_1.1 applies to any pair of ages, it is extremely non-specific with respect to the scope of applicability. It will turn out that this default is usually overridden by more specific premises in the more developed theories.

5.1.2 Endowment (Fragment 2)

The next development introduced the notion of endowments. According to the endowment-based argumentation, there is a period of given length, call it ϵ, in which an organization enjoys high survival chances due to the protection offered by some endowment. We treat the period of normal endowment as a population characteristic and ϵ as a population parameter.[10] Some populations face resource-poor environments or intense competition for resources with the result that newly founded organizations are unlikely to be endowed (Carroll and Hannan 1989). We can represent this case by setting $\epsilon = 0$. The other extreme, infinitely persisting endowment, does not seem reasonable. So we will assume that the normal period of endowment is finite.

[10] Of course, empirical research might be able to specify the sizes of endowments at the organizational level. In such a case, the more specific information ought to overrule the population-level default.

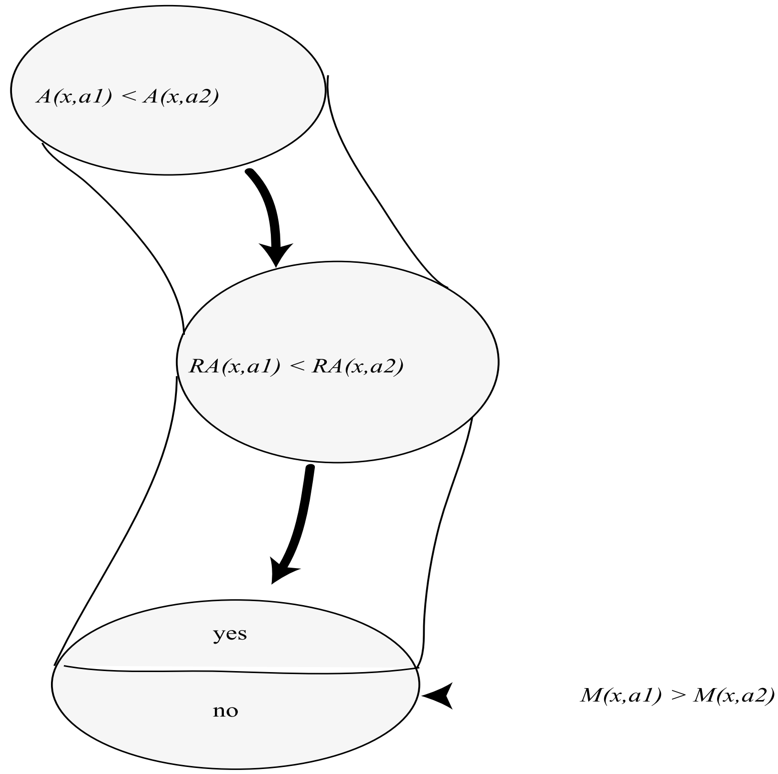

Figure 3: Rule chain in the default theory: Fragment 1

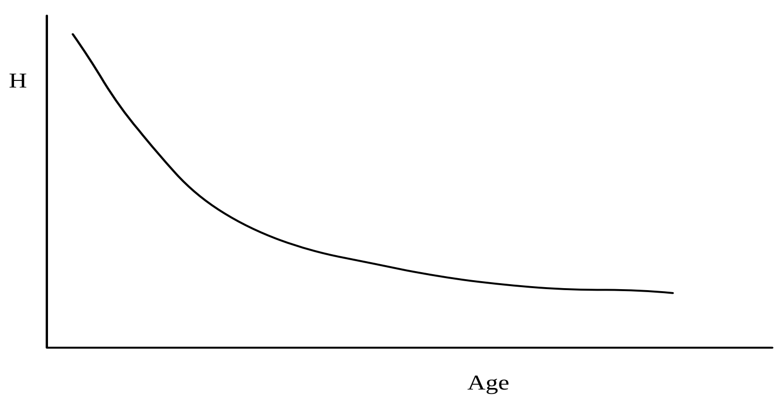

Figure 4: The pattern of age dependence of organizational mortality according to the default theory: liability of newness

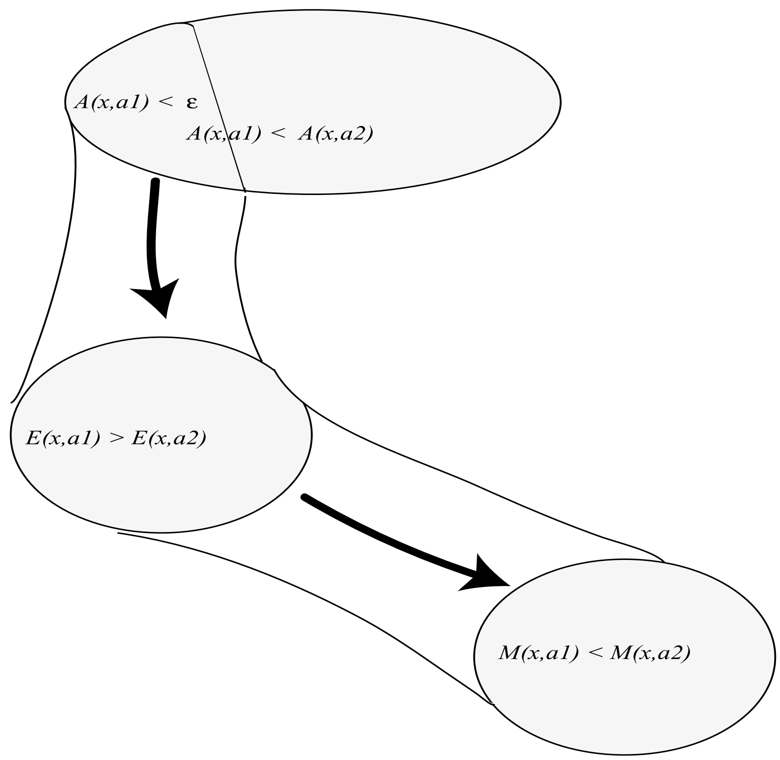

Figure 5: Rule chains in the second theory fragment: liability of adolescence

Figure 6: Pattern of age dependence of mortality of an endowed population of organizations according to the second theory fragment: liability of adolescence

Premise F_2.1 Before age ϵ, stocks of initial endowment are normally positive and they normally decline with age:

$$\mathcal{Q}x, a_1, a_2[A(x, a_1) < A(x, a_2) \wedge A(x, a_1) < \epsilon) \rightarrow E(x, a_1) > E(x, a_2)].$$

Premise F_2.2 A larger endowment normally lowers the hazard:

$$\mathcal{Q}x, a_1, a_2[E(x, a_1) < E(x, a_2) \rightarrow M(x, a_1) > M(x, a_2)].$$

These premises indeed imply that the hazard increases with age during an initial period in which endowments get spent down:

Proposition F_2.1 The hazard presumably increases with age within the period of initial endowment:

$$(A(x, a_1) < \epsilon \wedge A(x, a_1) < A(x, a_2)) \hookrightarrow M(x, a_1) < M(x, a_2).$$

Proof of F_2.1 The proof of this proposition follows the line of the previous proof. It is concerned with the rule-chains that connect the intension of $(A(x, a_1) < \epsilon \wedge A(x, a_1) < A(x, a_2))$ and that of $M(x, a_1) < M(x, a_2)$. Again, there is only one such chain; and this chain is positive. ■

5.1.3 First Unification Attempt (U_1)

Now consider the first attempt at unification: the theory that uses all four of these premises. We can regard this construction as the second stage of the theory.

Premises: F_1.1–2 and F_2.1–2.

At this stage of the theory, we get the following theorems:

Theorem U_1.1 The hazard presumably increases with age during the period of initial endowment:

$$(A(x, a_1) < \epsilon \wedge A(x, a_1 < A(x, a_2)) \hookrightarrow M(x, a_1) < M(x, a_2).$$

Theorem U_1.2 The hazard presumably decreases with age after the period of initial endowment:

$$(\epsilon < (A(x, a_1) \wedge A(x, a_1 < A(x, a_2)) \hookrightarrow M(x, a_1) > M(x, a_2).$$

Proofs U_1.1, U_1.2 Now we have to consider the most specific rule chains that connect $(A(x, a_1) < \epsilon \wedge A(x, a_1 < A(x, a_2))$ and $M(x, a_1) < M(x, a_2)$, and $(\epsilon < (A(x, a_1) \wedge A(x, a_1 < A(x, a_2))$ and $M(x, a_1) > M(x, a_2)$, respectively. Both two-step chains are positive. ■

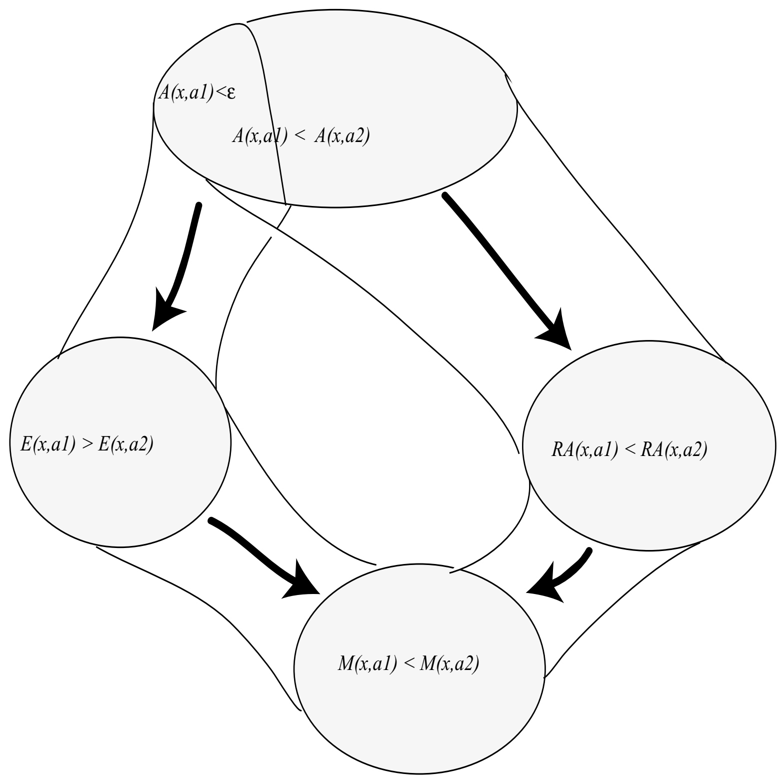

Figure 7: Rule chains in the first unification: combining the first and second fragments

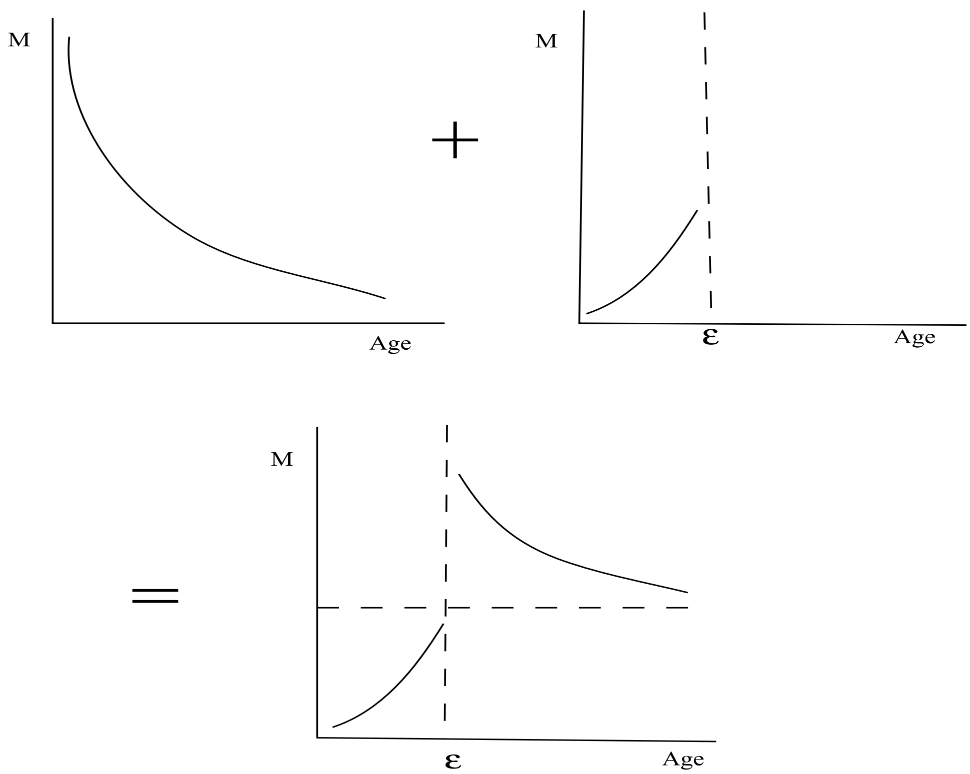

Figure 8: The pattern of age dependence of mortality according the first unification

A somewhat surprising corollary also follows from these premises. Even though we see a period of monotonic increasing hazards and a period of monotonically decreasing hazards, the machinery implies that the hazard in the second period cannot fall below the highest value of the hazard in the first period. In other words, there is an overall tendency toward positive age dependence.

Theorem:U$_1$.3 The hazard presumably increases over all intervals that span ϵ:

$$(A(x, a_1) < \epsilon < A(x, a_2)) \hookrightarrow M(x, a_1) < M(x, a_2).$$

Theorems U$_1$.2 and U$_1$.3 together imply a nice result about the global behavior of the process:

Theorem U$_1$.4 The hazard presumably jumps at to its highest level at ϵ.

Proof: So far we have derived:

Theorem U$_1$.1: $(A(x, a_1) < A(x, a_2)) < \epsilon) \hookrightarrow M(x, a_1) < M(x, a_2)$

Theorem U$_1$.2: $(\epsilon < A(x, a_1) < A(x, a_2)) \hookrightarrow M(x, a_2) < M(x, a_1)$

Theorem U$_1$.3: $(A(x, a_1) < \epsilon < A(x, a_2)) \hookrightarrow M(x, a_1) < M(x, a_2)$

According to Theorem U$_1$.1, $A(x, a_2)$ can be arbitrarily close to ϵ, and then $M(x, a_2)$ will likely approximate the supremum (least upper bound) of all of the values of the hazard before ϵ. Similarly, in Theorem U$_1$.2, $A(x, a_2)$ can be arbitrarily far above ϵ; and, then, $M(x, a_2)$ will likely approximate the infinum (greatest lower bound) of all the values of the hazard after ϵ. According to Theorem U$_1$.3, it is likely that $M(x, a_1) < M(x, a_2)$, i.e., the infinum of the values after ϵ is likely to either equal or exceed the supremum of the values before ϵ.

Now from Theorem U$_1$.2, we know that the value of the hazard above, but infinitely close, to ϵ is likely to approximate the supremum of the values after ϵ. Furthermore, we know that the set of the values of the hazard after ϵ has several elements, and therefore the infinum of this set is strictly smaller then the supremum.

If we put all these consideration together, then we can conclude that the hazard likely jumps at ϵ and the height of the jump is likely to be at least as big as the difference between the infinum and the supremum of the set of hazard values after ϵ. These considerations also show that the hazard reaches its maximum at ϵ. ■

5.2 Superimposing a Liability of Obsolescence

At least three different lines of argumentation support the claim of a liability of oldness: (1) environmental drift, inertia, and the need for being aligned with the environment: the LoO; the need for smooth operation, and the accumulation of organizational friction: the LoS; and (3) the need to expand the network of organizational ties, and the saturation of the space of (possible) ties.

Environmental Drift and Alignment (Fragment 3)

We assume that the quality of the alignment between an organization and its environment affects its mortality hazard. We also assume that organizations are relatively inert, that, in the long run, their structures cannot follow environmental changes. So, after a period of given length, they will normally no longer be aligned with the environment. As in Hannan (1998), we parameterize this process in terms of environmental drift. Specifically, organizations are normally best aligned with the environments in which they first appear; but environments drift over time. Due to inertia, drift causes alignment to decline steadily over time (age, from the perspective of an organization). The drift is such that within a period of length σ, the quality of alignment does not change so much from the founding conditions that if affects the hazard. However, beyond σ, the environment has normally drifted far enough as to drive the quality of alignment below a threshold that affects the hazard. Further drift, after σ, continually degrades alignment.

The parameter σ presumably varies among populations. At one extreme, the environment is so volatile that even new organizations cannot match their structures, strategies, and routines to the environment. In this case, we can reasonably set σ to zero. At the other extreme, environmental change is so slow that the threshold of poor alignment will normally not be reached in an organization's lifetime, that is, $\sigma = \infty$.

Premise F$_3$.1 After age σ, organizations are normally less and less well aligned with their environments:

$$\mathcal{Q}x, a_1, a_2[(A(x, a_1) < A(x, a_2) \wedge A(x, a_2) > \sigma) \rightarrow AL(x, a_1) > AL(x, a_2)].$$

Premise F$_3$.2 Superior alignment with the environment normally lowers the hazard:

$$\mathcal{Q}x, a_1, a_2[AL(x, a_1) < AL(x, a_2) \rightarrow M(x, a_1) > M(x, a_2)].$$

These two premises yield the following theorem:

Proposition F$_3$.1 The hazard of organizational mortality presumably increases with age after the initial period of alignment:

$$(A(x, a_1) < A(x, a_2) \wedge \sigma < A(x, a_2)) \hookrightarrow M(x, a_1) < M(x, a_2).$$

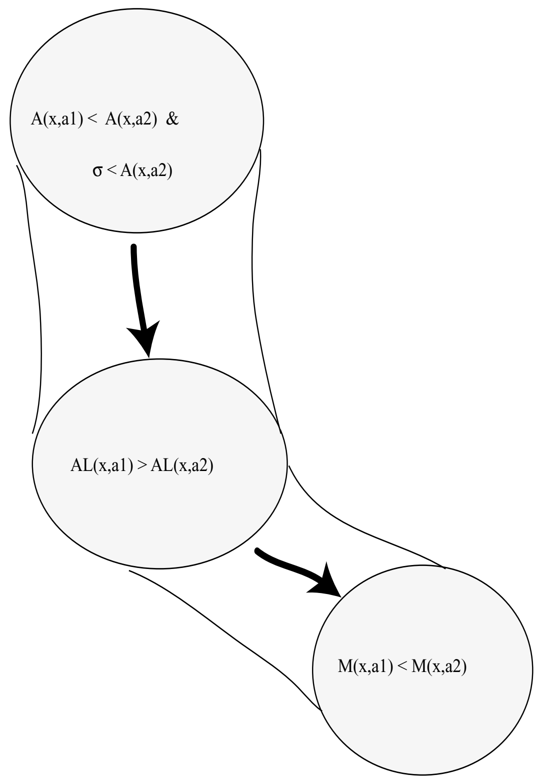

Figure 9: Rule chains in the third theory fragment: liability of obsolescence

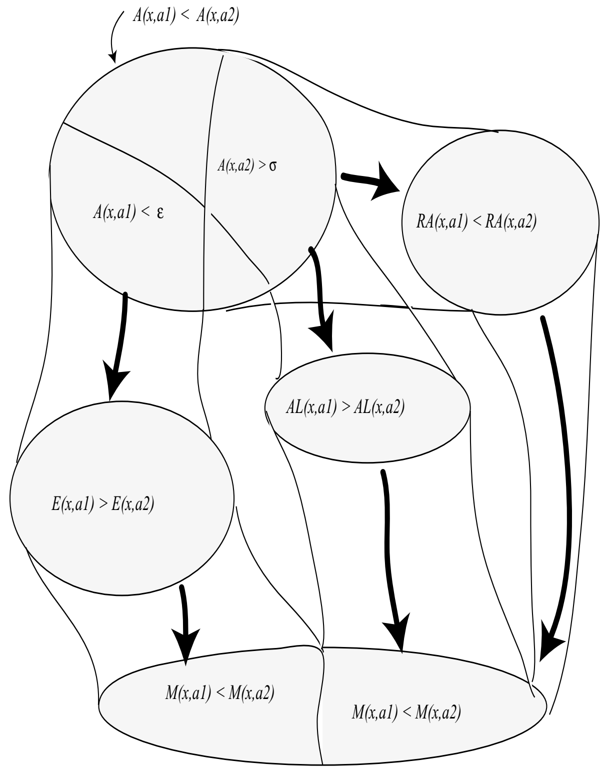

Figure 10: Rule chains in the second unification

Proof F_3.1 In this fragment there is only one rule chain that connects the intension of $A(x, a_1) < A(x, a_2)$ and the intension of $M(x, a_1) < M(x, a_2)$. This rule chain is positive. ■

5.2.1 Second Unification Attempt (U_2)

Now we consider the theory that uses all seven premises and one additional factual premise.

Premises: Premises F_1.1–2, F_2.1–2, F_3.1–3, and one fact: endowment is exhausted before the organization experiences obsolescence: $\epsilon < \sigma$ (It will be discussed later what happens in populations of organizations which does not provide empirical justification for this consideration.)

In this unification, the third stage of the theory, the first three theorems from the first unification attempt are still valid:

Theorem U_2.1 (same as U_1.1) The hazard presumably increases with age

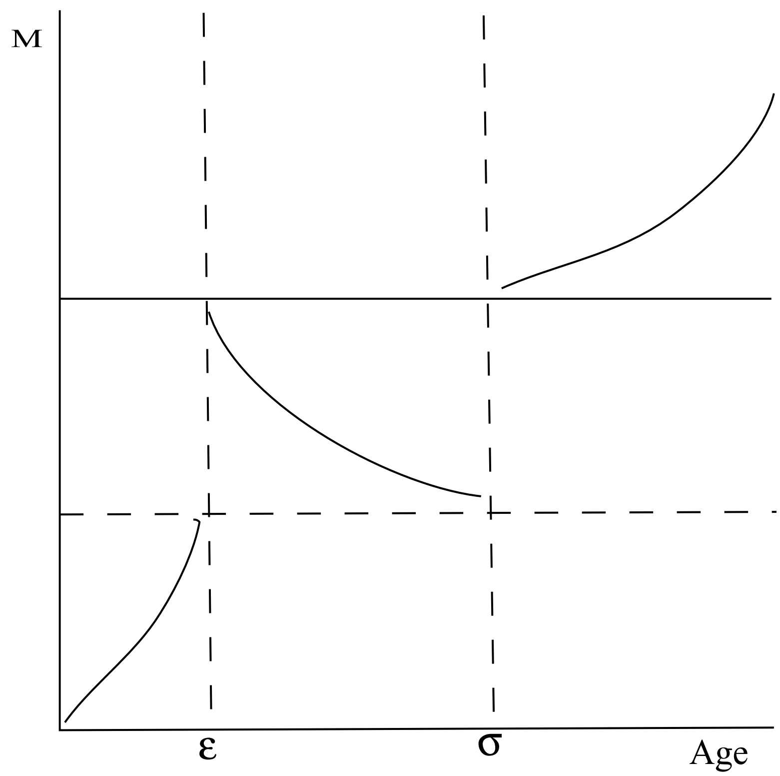

Figure 11: Age dependence of mortality according to the second unification

in the period of initial endowment and testing:

$$(A(x, a_{)} < A(x, a_2) < \epsilon) \hookrightarrow M(x, a_1) < M(x, a_2).$$

Theorem U_2.2 (same as U_1.2) The hazard presumably decreases with age after the initial period of endowment and testing but before the ending of alignment:

$$(\epsilon < A(x, a_{)} < A(x, a_2) \leq \sigma) \hookrightarrow M(x, a_1) > M(x, a_2).$$

Theorem U_2.3 (same as U_1.3) The hazard increases over intervals that span ϵ:

$$(A(x, a_{)} < \epsilon < A(x, a_2)) \hookrightarrow M(x, a_1) < M(x, a_2).$$

Theorem U_1.4 is now limited to the period before the onset of obsolescence:

Theorem U_2.4 The hazard presumably jumps to its highest level in the age interval $[0, \sigma)$ at ϵ.

And, we have theorems specific to the obsolescence process:

Theorem U_2.5 The hazard presumably increases with age after alignment ends:

$$(A(x, a_{)} < A(x, a_2) \wedge A(x, a_{)} > \sigma) \hookrightarrow M(x, a_1) < M(x, a_2).$$

Theorem U_2.6 The hazard presumably increases over intervals that span σ:

$$(A(x, a_{)} < \sigma < A(x, a_2)) \hookrightarrow M(x, a_1) < M(x, a_2).$$

Again we can derive implications about jumps and maxima in the process.

Theorem U_2.5 The hazard presumably jumps above its highest level when an organization reaches the age of obsolescence.

Proofs : All the proofs can be read from the diagram in figure 10. ■

Under specific conditions, the general picture reproduces the patterns of age-dependence found in empirical research:

Positive age dependence. If the exhaustion of endowments is not completed within an observation period, all changes in organizational mortality will be due to decreasing levels of endowment. This yields positive age dependence.

If a population is such that organizations become obsolete before they exhaust their endowments, there is no time left for realizing improvements in life chances due to improvements in reliability and accountability. The first and the last periods of positive age-dependence are "pushed together" and the mortality hazard increases monotonically.

The full picture can also be interpreted as entailing global positive age dependence: There are three distinct periods: before ϵ, between ϵ and σ, and after σ. We established above that the mortality hazard is lower before ϵ than between ϵ and σ and that it is lower between ϵ and σ than after σ. So the hazard is getting higher in the long run regardless of the changes observable in the details of the picture.

Negative age dependence. If the population is such that endowments are low or nonexistent, and the population did not reach the expected age of obsolescence yet, then the population shows negative age dependence in mortality.

Inverted-U-shaped age dependence. High adolescent mortality shows up in this picture when the presence of endowment offers some protection at young ages, the exhaustion of endowments results in a jump in organizational mortality, and then increasing reliability and accountability lowers the hazard. Here we assume that the population is not observed long enough to see obsolescence.

6 Smoothing the Process

The mortality process, as modeled so far, has jumps at the end of endowment and the onset of obsolescence. One can imagine situations in which such jumps might actually occur, as when endowments reflect some fixed-length regulatory protection granted to young organizations. Yet, it seems natural to assume that the process would shift gradually from a regime of decreasing to increasing hazards near σ. So it seems interesting to explore how the logical machinery and additional sociological arguments can smooth the process in that region.

Consider the onset of obsolescence. We want to create a Nixon Diamond by finding another specific argument that applies before σ. Then comparisons of hazards before and after σ would not be predictable because two equally and opposing specific rules apply to such intervals.

6.1 Saturation of Organizational Networks (Fragment 4)

Here we need to elaborate on three assumptions:

> The possibility of extending the network of ties grows once the organization survived the initial period within which organizations would normally have exhausted their initial endowments and/or be subjected to testing.
>
> Organizations can saturate their space of possible network ties, i.e. reach a stage when hardly any new tie is possible. One way to explain such a possibility is to point out that network ties occasionally bring about network incompatibilities: ties to organizations A exclude the possibility of having ties with organization B.
>
> The more extensive is an organization's network of ties, the lower is its hazard of organizational mortality.

We formalize these assumptions as follows:

Premise F$_4$.1 An organization's network normally grows more extensive with age, after the period of initial endowment/testing and before saturation.

$$\mathcal{Q}x, a_1, a_2[\theta > a_2 > a_1 > \max(\delta, \epsilon) \rightarrow NWT(x, a_2) > NWT(x, a_1)].$$

Premise F$_4$.2 An organization's network normally does not grow more extensive with age after the saturation time (θ).

$$\mathcal{Q}x, a_1, a_2[a_2 > a_1 > \theta > \max(\delta, \epsilon) \rightarrow \neg(NWT(x, a_2) > NWT(x, a_1))].$$

Premise F$_4$.3 If an organization's network expands, then its hazard falls.

$$\mathcal{Q}x, a_1, a_2[(a_2 > a_1 \wedge NWT(x, a_2) > NWT(x, a_1)) \rightarrow M(x, a_2) < M(x, a_1)].$$

Premise F$_4$.4 If an organization's network does not expand, then its hazard rises.

$$\mathcal{Q}x \forall a[(a_2 \geq a > a_1 \wedge \neg(NWT(x, a) > NWT(x, a_1)) \rightarrow M(x, a_2) > M(x, a_1)].$$

We now get two propositions.

Proposition F$_4$.1 Before the normal time of saturation, an organization's hazard falls with age.

$$(\theta > a_2 > a_1 > \max(\delta, \epsilon)) \hookrightarrow M(x, a_2) < M(x, a_1).$$

Proposition F$_4$.2 After the normal time of saturation, an organization's hazard rises with age.

$$(a_2 > a_1 > \theta > \max(\delta, \epsilon)) \hookrightarrow M(x, a_2) > M(x, a_1).$$

Each proposition is more specific than the one based on the default; but neither is more specific than the other for max($a_1 < \theta < \sigma < a_2$ pairs. Therefore, we have a Nixon Diamond here.

6.2 Third Unification Attempt (U_3)

With these additional assumptions, the hazard can be smooth near σ if the population is such that the saturation of possible ties happens before obsolescence strikes, i.e., when $\theta < \sigma$.

Premises: Premises F$_1$.1–2, F$_2$.1–2, F$_3$.1–3, F$_4$.1–2, and F$_4$.1–3.

This fourth stage of the theory yields the following theorems:

Theorem U$_3$.1 The hazard of organizational mortality presumably increases with age in the initial period of endowment/testing:

$$(a_1 < a_2 < \max(\delta, \epsilon)) \hookrightarrow M(x, a_1) < M(x, a_2).$$

Theorem U$_3$.2 The hazard of organizational mortality presumably decreases with age after the initial period of endowment/testing but before the ending of alignment:

$$(\max(\delta, \epsilon) < a_1 < a_2 < \sigma) \hookrightarrow M(x, a_1) > M(x, a_2).$$

Theorem U$_3$.3 The hazard of organizational mortality presumably increases with age after alignment ends:

$$(a_1 < a_2 \wedge a_2 < \sigma) \hookrightarrow M(x, a_1) < M(x, a_2).$$

Jumps in the hazard are consequences of theorems of the following type: and Theorem U$_2$.4:

$$(a_1 < \theta \wedge \theta \leq \sigma \wedge \sigma < a_2) \hookrightarrow M(x, a_1) < M(x, a_2).$$

These theorems say that the hazard is higher after σ than before. If we expect "smooth" transitions between the phases (without knowledge of the actual functional forms of the component mortality functions), then the theory should not yield any prediction—i.e., it should produce a Nixon Diamond—for intervals overarching the $[\theta, \sigma)$ interval—and therefore rule out any predictions.

Furthermore we gain the following insight: If the rise in the mortality hazard starts with a sudden jump, then it is likely that obsolescence strikes; if the hazard starts to rise smoothly, then it is likely to be due to the saturation of the network of ties and/or senescence (perhaps acting in concert with obsolescence).

7 Reflections

Now after seeing how nonmonotonic logic applies in the context of theory construction, it is time to ask whether the new notion of the consequence relation changes our understanding of theory. It does not. Tarski's view of theory—a deductively closed system of sentences, where deductive closure means closure under the classical first-order notion of logical consequence—still applies.

The analysis presented in this paper did not deal with a theory in the Tarskian sense. Instead, it dealt with a theory in flux. And, just as a former dean is not a dean any longer, a theory in flux is not yet a theory. Instead, it is raw material for a future theory. Future research might provide evidence that would override its present conclusions.

How can we get a proper theory from this raw material? Suppose that enough high-quality empirical work clarified the conditions of occurrence of all of the patterns of age dependence in organizational mortality. Then one would likely feel tempted to declare that all of the relevant facts are in hand. We would then want to construct a complete (classical) theory.

The nonmonotonic formalization of the theory in flux can be used to construct the classical theory. We want a classical theory that is not contradicted by any of the classes of known tendencies. We can do this as follows:

- take all of the empirical generalizations and restrict them to those domains where they were not overridden and
- universally quantify them, with antecedents that provide the required restrictions.

With this revision, all those conclusions that were "likely" to be the case then become the case; and the derived expectations ("normally implies" statements) turn into universally quantified, though restricted, theorems.

For example, suppose that the theory in flux had been frozen into a restricted classical theory after accommodating considerations about the liability of newness and endowments. Then the former empirical generalizations:

Premise $\mathbf{F}_1$.1 Reliability/accountability normally grows with age at all ages:

$$\mathcal{Q}x, a_1, a_2[(a_1 < a_2) \rightarrow RA(x, a_1) < RA(x, a_2)].$$

Premise $\mathbf{F}_1$.2 Higher reliability/accountability normally lowers the hazard:

$$\mathcal{Q}x, a_1, a_2[RA(x, a_1) < RA(x, a_2) \rightarrow M(x, a_1) > M(x, a_2)].$$

Premise $\mathbf{F}_2$.1 Before age ϵ, stocks of initial endowment are normally positive and normally decline with age:

$$\mathcal{Q}x, a_1, a_2[(a_1 < a_2 \wedge a_1 < \epsilon) \rightarrow E(x, a_1) > E(x, a_2) > 0].$$

Premise $\mathbf{F}_2$.2 A larger endowment normally lowers the hazard:

$$\mathcal{Q}x, a_1, a_2[E(x, a_1) < E(x, a_2) \rightarrow M(x, a_1) > M(x, a_2)].$$

will be reformulated as (constrained) universal sentences as follows:

Premise $\mathbf{F}_1$.1′. Reliability/accountability grows with age after ϵ:

$$\forall x, a_1, a_2[(\epsilon < a_1 < a_2) \rightarrow RA(x, a_1) < RA(x, a_2)].$$

This formula differs from [Premise F_1.1] in two respects. Its variables are bound by a universal quantifier instead of the generic quantifier, and the applicability domain is restricted to periods being positioned after ϵ. The latter restriction follows the domain restriction to the portion of the age range in which the process is not overridden.

Premise $\mathbf{F}_1$.2′. Higher reliability/accountability lowers the hazard:

$$\forall x, a_1, a_2[RA(x, a_1) < RA(x, a_2) \rightarrow M(x, a_1) > M(x, a_2)].$$

This formula differs from [Premise F_1.2] in only one respect: its variables are bound by a universal quantifier instead of the generic quantifier.

Premise $\mathbf{F}_2$.1′. Before age ϵ, stocks of initial endowment are positive and decline with age:

$$\forall x, a_1, a_2[(a_1 < a_2, \wedge a_1 < \epsilon) \rightarrow E(x, a_1) > E(x, a_2) > 0].$$

This formula also differs from [Premise F_2.1] in one respect: its variables are bound by a universal quantifier instead of the generic quantifier.

Premise $\mathbf{F}_2$.2′. A larger endowment lowers the hazard:

$$\forall x, a_1, a_2[E(x, a_1) < E(x, a_2) \rightarrow M(x, a_1) > M(x, a_2)].$$

This formula also differs from [Premise F_2.2] in one respect: its variables are bound by a universal quantifier instead of the generic quantifier.

These reformulated premises imply the following theorems:

Theorem $\mathbf{F}_2$.1′. The hazard increases with age within the period of initial endowment:

$$\forall x, a_1, a_2[(a_1 < a_2 < \epsilon) \rightarrow M(x, a_1) < M(x, a_2)].$$

This proposition is a restricted version of Theorem F_2.1, but it is universal, without exceptions.

Theorem $\mathbf{F}_1$.1′. The hazard declines with age after the period of initial endowment:

$$\forall x, a_1, a_2[(\epsilon < a_1 < a_2) \rightarrow M(x, a_1) > M(x, a_2)].$$

This proposition is a also restricted, universal version of Theorem F_1.1.

Theorem $\mathbf{U}_1$.3′. The hazard increases over intervals that span ϵ:

$$\forall x, a_1, a_2[(a_1 < \epsilon < a_2) \rightarrow M(x, a_2) > M(x, a_1)].$$

This proposition is also a universal, but restricted, version of Theorem U_1.3.

Since this process freezes the theory in flux into a theory proper, these new versions of what had been derived generalizations can freely be used in further derivations. If there is nothing that would override the arguments that supported them, they are just as safe as the empirical generalizations. They are monotonic, their inferential behavior follows classical FoL.

Now, as far as the considerations of age dependence in organizational mortality are concerned, they could be combined by means of a nonmonotonic logic to provide interesting (in part unexpected) patterns, but extensive further research is needed before freezing the theory in flux into a theory becomes a tempting possibility for the research program.

To regain consistency after the unification, the reliability/accountability-based argument has been restricted to ages above the endowment-protected youth of organizations. It is not at all obvious that this restriction is well motivated substantively. Strong intuition suggests that quite the contrary; at very young age the gain in reliability and accountability is greater than in older ages.

To find a consistent way of unifying the arguments such that basic intuitions are not violated would require dependable knowledge of four functional forms for all populations of organizations:

- How (exactly) are aging and the growth of reliability and accountability related?
- How (exactly) does the growth of reliability and accountability influence the hazard?
- How (exactly) are aging and the level of endowment related?
- How (exactly) does the level of endowment influence the hazard?

The empirical research is not quite ready to deliver all these functional forms in the near future.

On the other hand, any publication commits its authors to a particular stage of the theory. From the point of view of the publication the theory in flux freezes when it gets published, and any classical first-order conclusion of that theory can and will be used against the authors.

8 Conclusion

We have proposed a new approach to unifying fragments of theory programs undergoing development. This approach requires the use of some logical tools not yet deployed in sociology, as well as some innovation in shaping the tools themselves.

We tried to illustrate the potential value of these developments by applying them to a well-studied yet still-recalcitrant problem: age dependence in organizational mortality processes. Thus we confront a situation in which several reasonably well developed theory fragments have withstood an attempt to bring make a coherent single theory expressed in standard form: first-order logic (Hannan 1998). We took it as a challenge to show that an appropriately formed nonmonotonic logic would succeed where the classical tools had not.

We claim that our approach has passed this test. We came up with a formalization that incorporates several theory fragments that could not be reconciled in formalizations based upon first-order logic. Moreover, we arrived at a consistent picture that surprised us. We worked hard at getting the local structure of the process right by concentrating on the specific arguments that applied to given phases of the organizational life course. What we did not expect to emerge from this effort is the strongly consistent global structure. When we shift from the ground-level view to a birdseye view, we see that the mortality hazard is higher in each later phase than in each earlier phase, despite the fact that the hazard does not rise monotonically within the phases. This unintended result seems pleasing, because much well designed empirical research has begun to find such a global pattern.

References

Baron, James N., Michael T. Hannan, and M. Diane Burton. 2000. "Labor Pains: Organizational Change and Employee Turnover in Young, High-Tech Firms." *American Journal of Sociology,* forthcoming.

Barron, David N. 1992. *An Ecological Analysis of the Dynamics of Financial Institutions in New York State, 1914–1934.* Unpublished Ph.D. Dissertation, Cornell University.

Barron, David N., Elizabeth West, and Michael T. Hannan. 1994. "A Time to Grow and a Time to Die: Growth and Mortality of Credit Unions in New York, 1914–1990." *American Journal of Sociology* 100:381–421.

Brewka, G., J. Dix, and K. Konolige. 1997. *Nonmonotonic Reasoning: An Overview.* Stanford: CSLI Publications.

Carlson, Gregory N. 1977. *Reference to Kinds in English.* Ph.D. dissertation, University of Massachusetts, Amherst.

——— 1988. "Truth-Conditions of Generic Sentences: Two Contrasting Views." Pp. 31–52 in *Genericity in Natural Language: Proceedings of the 1988 Tübingen Conference,* edited by M. Krifka. Universität Tübingen.

Carroll, Glenn R. and Michael T. Hannan. 1989. "Density Delay in the Evolution of Organizational Populations: A Model and Five Empirical Tests." *Administrative Science Quarterly* 34:411–30.

——— 2000. *The Demography of Corporations and Industries.* Princeton: Princeton University Press.

Descartes, René. 1897–1913. *Oeuvres de Descartes.* Edited by C. Adam and P. Tannery. Paris: J. Vrin.

Diesing, Molly. 1988. "Bare Plural Subjects and the Stage/Individual Contrast." Pp. 107–154 in *Genericity in Natural Language: Proceedings of the 1988 Tübingen Conference,* edited by M. Krifka. Universität Tübingen.

Dowty, David. R., Robert E. Wall, and Stanley Peters. 1980. *Introduction to Montague Semantics.* Dordrecht: Kluwer Academic.

Freeman, John, Glenn R. Carroll, and Michael T. Hannan. 1983. "The Liability of Newness: Age Dependence in Organizational Death Rates." *American Sociological Review* 48:692–710.

Gamut, L. T. F. 1991. "Logic, Language and Meaning." 2 vols. Chicago: University of Chicago Press.

Hannan, Michael T. 1998. "Rethinking Age Dependence in Organizational Mortality: Logical Formalizations." *American Journal of Sociology* 104:85–123.

Hannan, Michael T. and John Freeman. 1977. "The Population Ecology of Organizations." *American Journal of Sociology* 82:929–64.

——— 1984. "Structural Inertia and Organizational Change." *American Sociological Review* 49:149–64.

Kratzner, Angelika. 1988. "Stage-Level and Individual-Level Predicates." Pp. 247–284 in *Genericity in Natural Language: Proceedings of the 1988 Tübingen Conference,* edited by M. Krifka. Universität Tübingen.

McCarty, J. 1980 "Circumscription—a Form of Nonmonotonic Reasoning." *Artificial Intelligence and Logic Programming* 13:27–39.

Makinson, D. 1994 "General Nonmonotonic Logic." Pp. 35–110 in *Handbook of Logic in Artificial Intelligence and Logic Programming: Nonmonotonic Reasoning and Uncertain Reasoning.* Vol III. edited by D. M. Gabbay, C. J. Hogg, and J. A. Robinson. Oxford: Oxford University Press.

Péli, Gábor, László Pólos, and Michael T. Hannan. 2000. "Back to Inertia: Theoretical Implications of Alternative Styles of Logical Formalization." *Sociological Theory* 18:193–213.

Pólos, László and Michael T. Hannan. 2000a. "Nonmonotonicity in Theory Building." In *Simulating Organizational Worlds,* edited by A. Lomi and E. Larsen. Cambridge: MIT Press, in press.

Pólos, László and Michael T. Hannan. 2000b. "Reasoning with Partial Knowledge" Research Paper No. 1638, Graduate School of Business, Stanford University.

Pólos, László, Michael T. Hannan, and Jaap Kamps. 1999."Aging by Default." In *Proceedings of the 4th Dutch–German Workshop on Nonmonotonic Reasoning Techniques and Their Applications,* edited by H. Rott, C. Albert, G. Brewka, and C. Wittveen. Amsterdam: ILLC Scientific Publications.

Pólos, László and Jaap Kamps. 1998."Reasoning With Empirical Generalizations." CCSOM Technical Report 98–169. University of Amsterdam.

Schubert, Lenhardt and Francis J. Pelletier. 1988. "An Outlook on Generic Sentences." Pp. 357–372 in *Genericity in Natural Language: Proceedings of the 1988 Tübingen Conference,* edited by M. Krifka. Universität Tübingen.

Sørensen, Jesper B. and Toby E. Stuart. 2000. "Organizational Aging and Innovation." *Administrative Science Quarterly* 45:81–112.

Stinchcombe, Arthur L. 1965. "Social Structure and Organizations." Pp. 142–93 in *Handbook of Organizations,* edited by J. G. March. Chicago: Rand McNally.

Thomason, Richmond H. 1988. "Theories of Nonmonotonicity and Natural Language Generics." Pp. 395–406 in *Genericity in Natural Language: Proceedings of the 1988 Tübingen Conference,* edited by M. Krifka. Universität Tübingen.

van Benthem, Johan. 1996. "Logic and Argumentation Theory." In *Logic and Argumentation,* edited by J. van Benthem, S. van Eemereren, R. Grootendorst, and F. Veltman. Amsterdam: Royal Dutch Academy of Sciences.

Veltman, Frank. 1996. "Defaults in Update Semantics." *Journal of Philosophical Logic* 25:221–61.

ERASMUS **R**ESEARCH **I**NSTITUTE OF **M**ANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

Publications in the Report Series Research* in Management

Impact of the Employee Communication and Perceived External Prestige on Organizational Identification
Ale Smidts, Cees B.M. van Riel & Ad Th.H. Pruyn
ERS-2000-01-MKT

Critical Complexities, from marginal paradigms to learning networks
Slawomir Magala
ERS-2000-02-ORG

Forecasting Market Shares from Models for Sales
Dennis Fok & Philip Hans Franses
ERS-2000-03-MKT

A Greedy Heuristic for a Three-Level Multi-Period Single-Sourcing Problem
H. Edwin Romeijn & Dolores Romero Morales
ERS-2000-04-LIS

Integer Constraints for Train Series Connections
Rob A. Zuidwijk & Leo G. Kroon
ERS-2000-05-LIS

Competitive Exception Learning Using Fuzzy Frequency Distribution
W-M. van den Bergh & J. van den Berg
ERS-2000-06-LIS

Start-Up Capital: Differences Between Male and Female Entrepreneurs, 'Does Gender Matter?'
Ingrid Verheul & Roy Thurik
ERS-2000-07-STR

The Effect of Relational Constructs on Relationship Performance: Does Duration Matter?
Peter C. Verhoef, Philip Hans Franses & Janny C. Hoekstra
ERS-2000-08-MKT

Marketing Cooperatives and Financial Structure: a Transaction Costs Economics Analysis
George W.J. Hendrikse & Cees P. Veerman
ERS-2000-09-ORG

---

* ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Decision Making in Marketing Management
F&A Financial Decision Making and Accounting
STR Strategic Renewal and the Dynamics of Firms, Networks and Industries

A Marketing Co-operative as a System of Attributes: A case study of VTN/The Greenery International BV,
Jos Bijman, George Hendrikse & Cees Veerman
ERS-2000-10-ORG

Evaluating Style Analysis
Frans A. De Roon, Theo E. Nijman & Jenke R. Ter Horst
ERS-2000-11-F&A

From Skews to a Skewed-t: Modelling option-implied returns by a skewed Student-t
Cyriel de Jong & Ronald Huisman
ERS-2000-12-F&A

Marketing Co-operatives: An Incomplete Contracting Perspective
George W.J. Hendrikse & Cees P. Veerman
ERS-2000-13– ORG

Models and Algorithms for Integration of Vehicle and Crew Scheduling
Richard Freling, Dennis Huisman & Albert P.M. Wagelmans
ERS-2000-14-LIS

Ownership Structure in Agrifood Chains: The Marketing Cooperative
George W.J. Hendrikse & W.J.J. (Jos) Bijman
ERS-2000-15-ORG

Managing Knowledge in a Distributed Decision Making Context: The Way Forward for Decision Support Systems
Sajda Qureshi & Vlatka Hlupic
ERS-2000-16-LIS

Organizational Change and Vested Interests
George W.J. Hendrikse
ERS-2000-17-ORG

Strategies, Uncertainty and Performance of Small Business Startups
Marco van Gelderen, Michael Frese & Roy Thurik
ERS-2000-18-STR

Creation of Managerial Capabilities through Managerial Knowledge Integration: a Competence-Based Perspective
Frans A.J. van den Bosch & Raymond van Wijk
ERS-2000-19-STR

Adaptiveness in Virtual Teams: Organisational Challenges and Research Direction
Sajda Qureshi & Doug Vogel
ERS-2000-20-LIS

Currency Hedging for International Stock Portfolios: A General Approach
Frans A. de Roon, Theo E. Nijman & Bas J.M. Werker
ERS-2000-21-F&A

Transition Processes towards Internal Networks: Differential Paces of Change and Effects on Knowledge Flows at Rabobank Group
Raymond A. van Wijk & Frans A.J. van den Bosch
ERS-2000-22-STR

Assessment of Sustainable Development: a Novel Approach using Fuzzy Set Theory
A.M.G. Cornelissen, J. van den Berg, W.J. Koops, M. Grossman & H.M.J. Udo
ERS-2000-23-LIS

Creating the N-Form Corporation as a Managerial Competence
Raymond vanWijk & Frans A.J. van den Bosch
ERS-2000-24-STR

Competition and Market Dynamics on the Russian Deposits Market
Piet-Hein Admiraal & Martin A. Carree
ERS-2000-25-STR

Interest and Hazard Rates of Russian Saving Banks
Martin A. Carree
ERS-2000-26-STR

The Evolution of the Russian Saving Bank Sector during the Transition Era
Martin A. Carree
ERS-2000-27-STR

Is Polder-Type Governance Good for You? Laissez-Faire Intervention, Wage Restraint, And Dutch Steel
Hans Schenk
ERS-2000-28-ORG

Foundations of a Theory of Social Forms
László Pólos, Michael T. Hannan & Glenn R. Carroll
ERS-2000-29-ORG