A MANUAL OF INTENSIONAL LOGIC

Second Edition
Revised and Expanded

Johan van Benthem

CSLI
CENTER FOR THE STUDY
OF LANGUAGE
AND INFORMATION

CSLI was founded early in 1983 by researchers from Stanford University, SRI International, and Xerox PARC to further research and development of integrated theories of language, information, and computation. CSLI headquarters and the publication offices are located at the Stanford site.

CSLI/SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

CSLI/Stanford
Ventura Hall
Stanford, CA 94305

CSLI/Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304

Printed in the United States

94 93 92 91 90 89 88 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Benthem, J. F. A. K. van, 1949–
A manual of intensional logic / Johan van Benthem. -- 2nd ed., rev. and expanded.
p. cm. -- (CSLI lecture notes ; no. 1)
Includes bibliographies and index.
1. Logic. I. Stanford University. Center for the Study of Language and Information. II. Title. III. Series.
BC71.B38 1988
160—dc 19 87-34935
CIP

ISBN 0-937073-30-X
ISBN 0-937073-29-6 (pbk.)

To Marie Virgine

Acknowledgments

It is a pleasure to acknowledge some debts incurred when writing this manual, both the original and its revision.

The course upon which this book is based was taught at Stanford in the spring of 1984, during a stay made possible by the support of the Department of Philosophy, CSLI, and the Sloan Foundation. John Perry was especially instrumental in organizing this visit of an exponent (and product) of the "ancien régime" in intensional logic.

As for editorial assistance, and indeed guidance in the production of the work, the help of Dikran Karagueuzian and his collaborators has been invaluable. Their combination of efficiency and friendliness has deeply impressed this veteran of the Dutch academic bureaucracy.

Contents

Introduction

These notes form the content of a graduate course in Intensional Logic taught at Stanford University in the Winter of 1984.

Intensional Logic as understood here is a research program based upon the broad presupposition that so-called "intensional contexts" in natural language can be explained semantically by the idea of *multiple reference.* Thus, temporal contexts require shifting patterns of ordinary Tarski-style denotations through time, modal contexts are related to varying denotations in some set of relevant worlds, or situations. Instead of developing one most general framework for this technique, we shall consider the above, as well as other examples.

The inspiration for the semantic theories to be presented here has been diverse, coming from linguistics, traditional philosophy, but also, for example, from the philosophy of science. On the intellectual map of a logician there is nothing strange to this combination, these being his natural neighbors. We shall see these various forms of inspiration recur throughout our examples.

From a technical point of view, the multiple reference move has the advantage that classical notions and arguments often remain applicable. On the other hand, this also makes intensional logic, on the whole, an area of application rather than innovation for mainstream logic. Even so, quite a few technical developments have occurred in the area with intrinsic logical interest (witness R. A. Bull's reviews in the *Journal of Symbolic Logic*, vol. 47:2, 1982, p. 440–445, and vol. 48:2, 1983, p. 488–495). In these notes, no such topics are explored in any depth, our aim being rather a survey of trends and questions in the area.

One conspicuous feature of intensional semantics, setting it apart from its classical ancestor, is the role played by *patterns* of classical situations. Thus, temporal instants come ordered by "before," possible worlds by "accessibility," "transition," or "similarity." We shall stress the intimate relation between types of inference in intensional logic and the kind of semantic pattern presupposed by these. In a sense, this is the most "intrinsic" question generated by our paradigm, concerning the nature of our semantic pictures, and hence of the conceptual enlightenment provided by this kind of semantics. It is precisely because these correspondences are often surprising that intensional semantics has been more than a systematic spelling out of the obvious.

The plan of these notes is as follows. First, some central examples of the more traditional phase of the subject are presented: the logic of tense, modality, conditionals, and their combinations. Then, what is perhaps the most conspicuous recent development is investigated: the shift from "total" to "partial" views of semantic entities and linguistic interpretation. Finally, a connection is made with another recent development in linguistic semantics: intensional notions will be studied as kinds of generalized quantifiers.

Obviously, a selection had to be made in all of this—and I have followed my predilections (and abilities). There is more of an emphasis on propositional than on predicate logic, there is a constant playing down of pretended exclusive rights to "the" philosophical interpretation of the formalism, and, finally, no comprehensive survey of the literature is attempted.

One cannot live in Stanford for long without realizing that "possible worlds semantics" is losing favor with the philosophical community. Evidently, fighting trends is a losing battle, in philosophy as much as in science or fashion. Nevertheless, this course is meant to give an overview of the above *research program*, as distinct from some of the philosophical *ideologies* commonly attached to it. The fund of questions and approaches gathered under the above umbrella, constantly renewing itself since the sixties, is something even the most advanced semanticists should come to terms with; I would hope, terms of endearment.

Concerning the Second Edition (Summer 1987)

One of the surprising things about Intensional Logic is its ability to inspire new applications, even after its reputed demise. Especially in Computer Science, there is a flourishing these days of notions and techniques from possible worlds semantics. A new chapter has been added to describe some of the main developments here. This computational connection also has beneficial effects on Intensional Logic itself: many traditional topics, for instance in the logic of knowledge, have acquired a new impetus.

In addition, the main text has been corrected and expanded, for instance in the discussion of temporality and partiality. At the end, a new chapter has been added on intensional types, reflecting current interest in flexible type theories for logical semantics.

I Classical Theories

Many different types of intensional phenomena have become the subject of special branches of Intensional Logic. Perhaps the most central example is modal logic, but there is also epistemic logic (knowledge and belief), deontic logic (obligation and permission), tense logic, as well as their offspring. For instance, modal logic has served as a model for such diverse theories as the provability logic of arithmetic and the "dynamic logic" of computation (and action in general). Moreover, various combined systems have been investigated. The strategy in these notes has been, not to find a greatest common denominator for all of this, but rather to present the three specific examples which seem to have been richest as research programs: tense, modality and conditionals. We begin with the topic which is generally felt to be the most concrete of the three.

1 Tense and Time

Literature

Prior, A. 1967. *Past, Present and Future.* Oxford: Clarendon Press.
Van Benthem, J. 1983. *The Logic of Time.* Dordrecht: Reidel.

Motivation

The subject of "tense logic" has had a dual motivation. From a philosophical and logical point of view, there is an interest to the calculus of temporal reasoning, breaking with the traditional view that reasoning can only involve timeless eternal propositions. Of the many temporal indicators in natural language, tenses turn out to be reasonably universal and stable. From a linguistic point of view, a logical description of such a ubiquitous and important phenomenon as tense will obviously also be quite welcome.

Some simple examples of temporal reasoning will serve as a point of departure. There is a synonymy between *Dahlia will sob or snore* and *Dahlia will sob or Dahlia will snore*. This ought to be explained by our semantics. Opinions differ as to the synonymy between *Dahlia had lied* and *Dahlia lied*. Our semantics ought to explain the temporal option involved here. And finally, patent non-inferences demand explanation as well, for instance the one from *Dahlia will laugh* and *Dahlia will cry* to *Dahlia will laugh-and-cry*.

Our first step (by no means a negligible discovery) is the introduction of a suitable notation:

$F\varphi$: φ will be the case (at least once),
$P\varphi$: φ was the case (at least once).

The above examples then become:

$$F(q \vee r) \leftrightarrow Fq \vee Fr, \quad PPq \leftrightarrow Pq?, \quad Fq \wedge Fr \not\rightarrow F(q \wedge r).$$

Notice that the exact derivation of these *logical forms* from the earlier actual *sentences* may already involve some manipulation. Ever since Montague's pioneering work, logicians have become aware that the establishment of such links, traditionally thought of as an art rather than a science, may actually have a logic of its own, that ought to be an explicit part of the semantic enterprise. This point of view seems valid, but we shall continue to sin in these notes.

There are more profound examples of temporal reasoning in the philosophical tradition, where actually something is at stake, such as McTaggart's famous proof of the "Unreality of Time." Many of these have a *modal* flavor as well, but an example seems pertinent, even at this stage.

In the *Master Argument* of Diodorus Cronos (as reconstructed by M. White), Diodorus is reported to have proved the metaphysical principle of "Plenitude" (all possibilities are actual), in the following form:

$$\Diamond\varphi \rightarrow \varphi \vee F\varphi.$$

That is, whatever is possible ($\Diamond$) is or will be the case! In a historical analysis, White finds the following premises involved in the argument:

a. $FF\varphi \rightarrow F\varphi$,

b. $F\varphi \rightarrow G(F\varphi \vee \varphi \vee P\varphi)$ (with "G" standing for "it is always going to be the case": $G\varphi \leftrightarrow \neg F \neg \varphi$),

c. P *true*,

as well as a principle expressing Necessity ($\Box$) of the past:

d. $Pq \rightarrow \Box Pq$.

(The Necessity of the Past and Modal-Tense Logic Incompleteness, *Notre Dame Journal of Formal Logic*, 25:1, 1984, pp. 59–71.)

We shall evaluate this argument in the semantics of Chapter 4—with the outcome that Diodorus' argument is valid for a certain metaphysical treatment of possibility and necessity, but not for plausible temporalized accounts of these notions.

Linguistically inclined students were struck by some pleasant analogies between the above temporal sentence operators and tenses in natural language, as tabulated below:

Dahlia lies	q	*Dahlia had lied*	PPq
Dahlia will lie	Fq	*Dahlia will have lied*	FPq
Dahlia lied	Pq	*Dahlia would lie*	PFq

correspondences have been suggested, such as PFP—*would have*, FF—*will be going to*. But eventually, the infinite number of formal combinations is going to outrun the actual (finite) number of tenses: there is "too much." There is also "too little": tenses such as the present perfect and progressive remain unaccounted for. But then, the formalism can be extended in due course to include further operators.

The Basic Formal Semantics

The above notation may be thought of as a formal language, in the simplest case, a propositional language of the usual kind, with added operators F, P, G (as above), and also H ("always in the past," $H = \neg P \neg$).

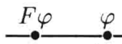
Models $M = \langle T, <, V \rangle$ consist of a "flow of time" $\langle T, < \rangle$, (i.e., a set of "moments" ordered by "earlier than" or "before") with a "valuation" V giving, for each proposition letter q, the set $V(q)$ of times when q is true. Alternatively, V provides a family of ordinary propositional valuations ("snapshots"), one at each moment in time. One can think here of various pictures of time: linear sequences, or perhaps branching trees, with changing events.

As usual, a recursive truth definition fixes the interpretation of the symbolism from the proposition letters upward:

$$M \models \varphi[t] \qquad (\text{“}\varphi \textit{ is true in } M \textit{ at } t\text{”}).$$

The clauses are as follows:

$M \models p[t]$ iff $t \in V(p)$,
$M \models \neg\varphi\,[t]$ iff *not* $M \models \varphi[t]$,
(and likewise for $\wedge$/*and*, $\vee$/*or*, $\rightarrow$/*if then*, etc.),

$M \models F\varphi\,[t]$ iff *for some* $t' > t$, $M \models \varphi[t']$ 

,
$M \models P\varphi\,[t]$ iff *for some* $t' < t$, $M \models \varphi[t']$ 

,
(and likewise for G and H).

If one wants to speak here of "the denotation" of a sentence φ in a model M, the obvious candidate is not one truth value, but a family of these: one for each world (Montague-style). Equivalently, one could say that the denotation of φ is the set of all points in time where φ is true (the common practice in technical tense logic). Notice that this view does not automatically commit one to the converse: that every set of points in time is an admissible "proposition." Indeed, models $\langle T, < \rangle$ with restricted ranges of denotations for sentences have been used extensively in technical research.

Digression. it may be instructive to realize that the view of sentences as denoting a *truth value* occurs in only one tiny corner of classical logic, viz. "local" propositional semantics. But even in propositional logic already, a sentence φ is often taken to denote all valuations making it true (if you wish, all "situations" where it holds). Moving to predicate logic, a proper Fregean compositional setup makes it imperative to let formulas denote sets of variable assignments. (Properly viewed, this is also what we have done in the above tense-logical case.) When special care is needed with complexity (as in set-theoretic meta-arguments), these will even be partial rather than total assignments. In other words, many monolithic philosophical views of "denotation" have been far removed from the realities (and flexibility) of logical practice.

When the temporal operators are added to a predicate logic, the above picture will become more detailed, locating, at each point $t \in T$, a whole Tarskian structure consisting of a domain of discourse and interpretations for basic predicates. In particular, this picture allows variation in domains over time, as well as changes in the behavior of individual objects. Such variation is needed to account for the various readings of, say,

Every man will do his duty: $\forall x(man(x) \rightarrow F\,duty(x))$,
$F\forall x(man(x) \rightarrow duty(x))$;

Every boy will have heard Dahlia: $\forall x(boy(x) \rightarrow FPhear(x,d))$,
$F\forall x(boy(x) \rightarrow Phear(x,d))$,
$FP\forall x(boy(x) \rightarrow hear(x,d))$.

A more radical way of "temporalizing" our ontology would be to dismiss the above stable objects, and take individuals to be "lifelines" assigning to each moment the corresponding manifestation of the individuals. (Compare Montague's "individual concepts.") This view, though intriguing, has never become predominant.

Evidently, the semantic scheme presented here can be used to interpret many other temporal expressions. The above is just the simplest system in existence, be it a fairly typical one.

There is a minimal logic of the above semantic scheme, consisting of all those formulas which are valid in all models at all points in time. One example is the earlier equivalence $F(q \vee r) \leftrightarrow Fq \vee Fr$. (Basically, the existential quantifier in the semantic clause for F distributes over disjunction.) On the other hand, for example, $Fq \wedge Fr \rightarrow F(q \wedge r)$ has obvious counterexamples. The *minimal tense logic* thus obtained can be axiomatized as follows:

a. all propositional tautologies;

b. the definitions:

$F\varphi \leftrightarrow \neg G \neg \varphi$,
$P\varphi \leftrightarrow \neg H \neg \varphi$;

c. the tense-logical axioms:

$G(\varphi \rightarrow \psi) \rightarrow (G\varphi \rightarrow G\psi)$,
$H(\varphi \rightarrow \psi) \rightarrow (H\varphi \rightarrow H\psi)$,
$\varphi \rightarrow GP\varphi, \quad \varphi \rightarrow HF\varphi$;

d. the rules of inference:

$\varphi, \varphi \rightarrow \psi / \psi$ (Modus Ponens),
$\varphi / G\varphi, \ \varphi / H\varphi$ ("Eternity").

But, further principles become valid as soon as one imposes reasonable constraints on the temporal order. Indeed, precise *correspondences* arise between tense-logical axioms and ordering conditions. To state these, the following notion of truth is needed, depending only on the ordering pattern, ignoring accidental features of any particular valuation:

- $\mathbf{T}$ $(= \langle T, < \rangle) \models \varphi[t]$ if $\langle T, <, V \rangle \models \varphi[t]$ for all valuations V,
- $\mathbf{T} \models \varphi$ if $\mathbf{T} \models \varphi[t]$ for all $t \in T$.

An example is provided by the earlier putative principle $PPq \leftrightarrow Pq$. We have:

$$\begin{array}{llll} \mathbf{T} \models PPq \rightarrow Pq\,[t] & \text{iff} & \forall x < t\; \forall y < x\colon\; y < t & (\textit{transitivity}), \\ \mathbf{T} \models Pq \rightarrow PPq\,[t] & \text{iff} & \forall x < t\; \exists y < t\colon\; x < y & (\textit{density}). \end{array}$$

Thus, we see which consequences endorsing the above equivalence of simple past and past perfect would have for our picture of time.

Other examples occur in the mentioned version of Diodorus' Master Argument:

- $FFq \rightarrow Fq$ defines *transitivity* as well,
- $Fq \rightarrow G(Fq \vee q \vee Pq)$ defines *right-linearity* (i.e., $\forall t\, \forall x > t\; \forall y > t\; (y < x \;\vee\; y = x \;\vee\; x < y)$), while
- P *true* defines *left-succession* $(\forall t\, \exists x\colon x < t)$.

Conversely, interesting tense-logical axioms have also been discovered in attempts to match existing conditions on the temporal order. For instance, in order to describe *discreteness*

$$\begin{array}{l} \forall t\colon\; \exists x > t\; \forall y < x\; (y = t \;\vee\; y < t), \\ \forall t\colon\; \exists x < t\; \forall y > x\; (y = t \;\vee\; y > t), \end{array}$$

"Hamblin's Axiom" was invented:

$$(q \wedge Hq) \rightarrow FHq, \quad (q \wedge Gq) \rightarrow PGq.$$

For a survey of this area, see J. van Benthem, Correspondence Theory, in D. Gabbay and F. Guenthner, eds., *Handbook of Philosophical Logic*, Vol. II, Reidel, Dordrecht, 1984.

Thus, through a study of actually proposed "valid" patterns of inference, one can form a conception of their presupposed picture of Time. This is the direction of modeling various temporal logics. Conversely, one may start with some preferred picture of Time, say, that of the *integer* or *real* line, and ask for all valid inferences there. For most well-known structures, such axiomatizations had been found by 1970. Actually, there is no general logical reason why all these structures should have axiomatizable theories at all. And, at least for temporal *predicate* logic, there is a well-known warning example (discovered independently by

Lindström and Scott): the tense predicate logic of the real time line is not effectively axiomatizable.

Of course, questions of correspondence and axiomatization are not the only queries arising in this semantics. For instance, an obvious question is, on any given temporal structure, what happens to the potentially infinite number of "tenses," that is, sequences of operators F, P, G, H in our language. A good exercise to become familiar with the peculiarities of our formalism is to prove Hamblin's *Fourteen Tenses Theorem*:

> On the real time axis, there are only 14 logically distinct sequences of operators.

(Hint: use simple collapsing principles such as "$PP = P$," "$GG = G$," as well as more exotic ones, such as "$FHF = F$," "$FHP = HP$." Moreover, cut down on calculations by exploiting symmetries.) Thus, the issue whether our formalism can match natural language tenses becomes more interesting, with possibly different answers in different semantic settings.

Finally, there is a logical interest to the study of flows of time $\langle T, < \rangle$, independently from any particular language being interpreted. For instance, many intuitions that we have concerning Time do not correspond to simple inferences, but are of a more global nature. One famous example is *Homogeneity*: "all points in Time are alike." It has the following technical formulation: "every point in T can be mapped onto any other one by some order-preserving automorphism of $\langle T, < \rangle$." Such intuitions impose more global constraints on our class of semantic models, with less "direct" effects than, say, the earlier transitivity or asymmetry. For instance, Homogeneity tells us (among other things) that the temporal order cannot change its behavior: for example, it is either dense or discrete throughout. Actually, the statement of such general intuitions, beyond what is needed for immediate "engineering" purposes, is a discernible recent trend in various areas of semantics.

Further Developments

The above paradigm has been extended repeatedly to account for further temporal phenomena. One notable direction here has been the introduction of auxiliary reference points. For instance, the permanence of "now," no matter how deeply embedded in a sentence, requires the continuing availability of the original moment of utterance, even while operators P and F are being unpacked, shifting the current point of evaluation. Thus, Kamp introduced "double indexing": φ is true in M at t, t_0; with a resetting function for "now." NOWφ is true in M at t, t_0 if and only if φ is true in M at t_0, t_0. Further moves in this direction have been proposed by Vlach, Åqvist and Guenthner, and by Gabbay—until a system arose whose temporal operators looked remarkably like the *Quine operators* for a variable-free predicate logic. By this time, the tense-logical formalism had become as strong as a full-blown two-sorted predicate logic, consisting of an ordinary predicate logic with an added domain of temporal items, freely allowing for quantification over the latter. (For instance, Fq corresponds to $\exists t > t_0\ Qt$,

$$\underline{F\,\forall x(\text{NOW}\,girl(x) \rightarrow woman(x))}$$

to

$$\exists t > t_0\ \forall x\ (girl(t_0, x) \rightarrow woman(t, x)),$$

while, for example, $\exists t < t_0\ \exists t' > t_0\ \forall s(t < s < t' \rightarrow rains(s))$ expresses the progressive *it is raining*.) A whole philosophical and methodological debate has raged about the relative virtues of these two formalisms.

The above two-sorted predicate logic with temporal parameters can be regarded as a kind of "descriptive limit" for temporal constructions. An early result is "Kamp's Theorem" showing that (modulo some assumptions on the temporal order) actually two constructions suffice for obtaining its full power:

$$\begin{array}{ll} \text{SINCE}\,\varphi\psi: & \exists t' < t_0(\varphi_{t'} \wedge \forall t'' \in (t', t_0)\psi_{t''}), \\ \text{UNTIL}\,\varphi\psi: & \exists t' > t_0(\varphi_{t'} \wedge \forall t'' \in (t_0, t')\psi_{t''}). \end{array}$$

Thus, in a sense, natural language attains maximal power of temporal expressibility using these two notions.

Finally, the perspective presented here has some obvious descriptive limitations. (The point, of course, is not that its students failed to notice these, but rather that they hoped to gain insights from a judicious idealization, as so often in science.) For instance, there are the "missing tenses," such as present perfect or progressive. Now, the latter can be given a new operator, but the former presents a more subtle problem. Its truth conditions seem to be identical to those for the simple past (P), and yet there is a difference. Reichenbach treated this distinction, already in the forties, using a so-called "reference point," introducing a perspective upon the event described: the simple past (*Dahlia sang*) has its perspective located at that past event, the present perfect (*Dahlia has sung*) may look at that same event from the present. Attempts to incorporate this idea into the earlier framework have not been quite successful. (We shall find a more recent attempt in the interval tense logic of Chap. 6.)

Another famous problem is the deictic use of tenses, as exemplified in Partee's *I did not turn off the stove*, which refers to some specific past time. Here, neither reading in the above formalism seems appropriate: $P\neg$ *turn off* is too weak and $\neg P$ *turn off* much too strong. (But perhaps, this just points at a "specific" reading of the existential quantifier in the $P\neg$-version—a phenomenon that has to be taken seriously even in ordinary predicate logic.)