SORITES

$$\boxed{\sigma\omega\rho\acute{\iota}\tau\eta}$$

Editor: Lorenzo Peña

Institute of Philosophy, CSIC
[Spanish Institute for Advanced Studies]
Madrid

Associate Editors:

Jesús Padilla-Gálvez
Johannes Kepler Universitaet
Linz, Austria
(also Reviews Editor)

Francisco J. Díez Ausín
CALIJ
(Centre for Logic and Juridical
Analysis, San Sebastian, Spain)

Manuel Liz
University of La Laguna
(Canary Islands, Spain)
(also Submissions Editor)

Issue #03 — November 1995

SORITES

SORITES

ISSN 1135-1349

Issue #03. November 1995

TABLE OF CONTENTS

ABSTRACTS OF THE PAPERS

REASONING WITH IMPERATIVES USING CLASSICAL LOGIC

Joseph S. Fulda

Traditionally, imperatives have been handled with deontic logics, not the logic of propositions which bear truth values. Yet, an imperative is issued by the speaker to cause (stay) actions which change the state of affairs, which is, in turn, described by propositions that bear truth values. Thus, ultimately, imperatives affect truth values. In this paper, we put forward an idea that allows us to reason with imperatives using classical logic by constructing a one-to-one correspondence between imperatives and a particular class of declaratives.

♣ ♣ ♣ ♣ ♣ ♣ ♣

A NAIVE VARIETY OF LOGICAL CONSEQUENCE

Enrique Alonso

The semantic analysis of logical consequence must obey a set of requisites which nowadays have acquired a dogmatic status. This situation prevents the development of other varieties of this fundamental relation. In this issue we try to define what we call a naive variety of logical consequence. The main feature of this relation is the way it depends on formulas in premises and conclusion: every sentence must contribute to the acceptability of an argument in a *significative* way. This circumstance can be of some interest for research programs demanding a logical apparatus sensitive to application context. We think of the logic LP developed by G. Priest — Priest [1979] — in relation to Gödel incompleteness theorems as a test for our points of view.

♣ ♣ ♣ ♣ ♣ ♣ ♣

Humor and Harm

Laurence Goldstein

For familiar reasons, stereotyping is believed to be irresponsible and offensive. Yet the use of stereotypes in humor is widespread. Particularly offensive are thought to be sexual and racial stereotypes, yet it is just these that figure particularly prominently in jokes. In certain circumstances it is unquestionably wrong to make jokes that employ such stereotypes. Some writers have made the much stronger claim that in all circumstances it is wrong to find such jokes funny; in other words that people who laugh at such jokes betray sexist/racist attitudes. This conclusion seems false. There is, as I shall argue, a thin dividing line between being properly sensitive to the rights and feelings of women and of racial groups different from our own, and being excessively sensitive to oversensitivity. Oversensitivity is, in this context, a kind of intolerance, and there is no reason why we should pander to that. One can be opposed to the unchecked dissemination of certain kinds of racist or sexist humor without oneself being a racist or sexist for finding such humor funny. The use of various stereotypes in humor serves the linguistic purpose of facilitating brevity and punch, the cultural purpose of preserving, in a sanitized form traditional rivalries and antipathies, and the psychological purpose of discharging fears. Blanket moral condemnation is inappropriate, though there will, of course, be circumstances under which the promulgation of certain types of humor, or even its enjoyment, ought to be condemned.

♣ ♣ ♣ ♣ ♣ ♣ ♣

What is Semantics? A Brief Note on a Huge Question

Newton C. A. da Costa, Otávio Bueno & Jean-Yves Béziau

After mentioning the cogent connection between pure semantics and the particular set theoretical framework in which it is formulated, some issues regarding the conceptual status of semantics itself, as well as its relationship to logic, are concisely raised.

♣ ♣ ♣ ♣ ♣ ♣ ♣

A Note on Truth, Deflationism and Irrealism

Pierluigi Miraglia

The paper deals with a problem about irrealist doctrines of content, according to which there are no real properties answering to content-attributing expressions. The central claim of the paper is that the distinction between factual and non-factual discourse (key to irrealism) is independent from particular conceptions of truth, and is thus compatible with a deflationary conception. This claim is sustained by an examination of what I take to be significant aspects of the deflationary conception. I argue therefore directly against Paul Boghossian's paper «The Status of Content», which attempted to show that irrealism about content is inconsistent.

REASONING WITH IMPERATIVES USING CLASSICAL LOGIC

Joseph S. Fulda

§1.— Translation

Consider the imperative issued on the streets of a city, «Give me ten dollars.» This is not a complete rendering of what the speaker intends: «Give me ten dollars or else....» is a more complete representation. The sanction that follows is indicated by an ellipsis, since the imperative mood encompasses both requests and commands of all sorts. Thus, if the man is a beggar, we might have «Give me ten dollars or I will starve,» while if the man is a dangerous criminal, we might have «Give me ten dollars or I will shoot.» Thus it is proposed to represent the imperative «Do x» with the declarative «If you don't do x, then sanction s,» where s may be a sanction applied against the speaker (the beggar case) or the listener (the criminal case). The sanction may also be applied against an innocent third party, as with the emperor who kills your family or the Klingons who kill innocent civilians. And, the sanction may be very weak: Mere displeasure is often the only sanction for ignoring a polite request. Regardless, however, of whether the sanction is severe or trivial, and against whom it is applied, it is always there, lurking in the background.

Very often «or else» indicates an exclusive disjunction, in which case the proposition corresponding to «Do x» is «If you do x, then sanction s will be avoided» & «If you don't do x, then sanction s.» In this case, the proposition corresponding to the imperative is a biconditional, rather than a simple conditional. Sometimes, the speaker intends not that the action be performed and the sanction avoided, but that the action not be performed and the sanction applied, as when an impossible action is demanded solely that the sanction can be applied or when the IRS requires reporting illegal income, only so that if the illegal activity is ultimately discovered, the list of charges in the indictment can include tax evasion.

Other times, the sanction will be null. «Excuse me!» someone says, but there is no sanction against him or the person accidentally bumped against if he is not excused. (There may be a sanction for the bumping itself, but not I think for the request to be excused.) In that case, we have «If you don't excuse me, then T» as the declarative corresponding to the imperative and it is a tautology. Another way to view such cases is to regard them as interjections, rather than (semantic) imperatives: Interjections do not change truth values, even through the mediation of actions (except when they are themselves actions of a sort).

Still other times, perhaps most times, the sanction will not be known to the listener and may not even be known to the speaker. This is entirely unproblematic: We require of a translation scheme only that it get what is being

translated as right as a human listener would get it, no more. So when a criminal approaches a man on the street with a demand for ten dollars and has no idea what he will do if the demand is refused, we have, simply, a propositional function, rather than a proposition — and that is entirely proper, since it represents accurately the intention behind the statement that has been uttered. Since most speakers use imperatives whose consequences are at times unclear to (at least) those to whom they are directed, it should disturb no one that the logician working on a theory of translation for imperatives should fare no better.

While a literature search has not revealed a detailed implementation of a sanction-based system, Anderson (1958), in his reduction of deontic logic to a form of modal logic, does mention the possibility of using a «penalty (reward) in a suitably broad sense» as a basis for understanding a system so reduced.

The present theory also makes sense out of contradictory imperatives, such as those of the wife who at times urges her husband to do his duty and go to war to defend his country and at other times urges him to stay home with her and their children and fulfill his domestic obligations. We translate the apparently contradictory «Go to war» & «Don't go to war» as «If you don't go to war, you'll be abdicating your duties as a citizen» & «If you do go to war, you'll be abdicating your responsibilities to me and our children.» Since the sanctions are different, we don't have a contradiction at all, but rather an instance of the Constructive Dilemma.

One might, of course, have the following discourse:

(1) If you don't go to war, you'll be abdicating your duties as a citizen.

(2) If you go to war, you'll be abdicating your responsibilities as a husband and father.

Therefore, (3) Don't go to war.

But this, too, is no contradiction, for the sanction in (3) is (at least) the displeasure of the speaker having considered both (1) and (2), which is very probably different from (arguably stronger, arguably weaker) the sanction in (2) alone. Finally, Kant's categorical imperative — an imperative which is often said to exclude the possibility of an «or else...» — does not, in fact, present a problem for our translation scheme: The sanction is simply blighting one's soul or something such.

The real test of a theory of translation, however, is how it fares when used to evaluate the validity of arguments, an enterprise to which we now turn.

§2.— Arguments

Consider first the argument scheme:

(1) Don't let x happen.

(2) If you do y, then x.

Therefore,

(3) Don't do y.

This is a clearly valid scheme containing a mixture of declaratives and imperatives, which reduces to:

(1') x→s (2') y→x Therefore, (3') y→s

A typical instance of this is:

(1'') Don't let the cat escape.

(2'') If you open the front door, the cat will escape.

Therefore,

(3'') Don't open the front door.

The second argument scheme is similar:

(4) Do x. (5) In order to do x, you must do y. Therefore, (6) Do y.

This is also a clearly valid scheme containing a mixture of declaratives and imperatives, and it reduces to:

(4') ~x→s (5') x→y Therefore, (6') ~y→s

A typical instance of this is the inference from «Clean up your room» to «Hang up your coat.»

A third valid argument scheme, but one which is far more complex, is:

(7) Do p or q. (8) If you do p, then do r. (9) If you do q, then do s. Therefore, (10) Do r or s.

It may not be presumed here that the sanctions for not doing p, q, r, and s are the same; they may or may not be. Hence this argument scheme reduces to:

(7') ~(p∨q)→x

(8') p→(~r→y)

(9') q→(~s→z)

Therefore,

(10') ~(r∨s)→(x∨y∨z), where x is at least y∨z.

A typical instance of this scheme is:

(7'') Either tell her nothing or tell her everything.

(8'') If you decide to tell her nothing, feign ignorance of the entire matter.

(9'') If you decide to tell her everything, tell it in a way that coheres credibly.

Therefore,

(10'')　If you neither feign ignorance of the entire matter nor tell her the whole story so that it coheres credibly, you'll be suspected of leaving something out or of lying.

I leave an appropriate context to the imagination of the reader. Notice that our choice of x was simply y∨z, but it could well be something stronger (y & z comes to mind, as well as more complicated propositions that entail y, z, or both).

However, the validity of the three schemata above depends critically on the assumption of a common context in which the premises and conclusion are asserted or else we have the usual problem with indexicals. Indeed, without this assumption, even the following argument scheme is invalid:

(11) Do x. Therefore, (12) Do x.

Since it does not follow from «Give me ten dollars or I will starve» that «Give me ten dollars or I will shoot» and since the sanction is, in both the premise and the conclusion, left implicit, the notion of validity for arguments with imperatives read as material (bi)conditionals depends critically on a common context. As another example, consider the imperative «Tell me who did it!» Asked by a curious friend, the sanction for not answering is mere displeasure; asked in a court of law, the sanction for not answering is being found in civil contempt and incarcerated — quite different!

Besides the problem of context, a more subtle problem arises if it is not clear whether the sanction will be applied only if the imperative is ignored or at least when the imperative is ignored, i.e. when it is not clear whether the conditional is a simple conditional or a biconditional. Thus it may appear that the following argument is surely valid:

(13) Give me ten dollars. Therefore, (14) Give me at least five dollars.

However, if the conditional corresponding to (13) is a biconditional, i.e. the sanction will be avoided if the command is obeyed (e.g., the criminal will not shoot if he is paid off), (14) may simply not be sufficient to avoid the sanction.

Verily, our ability to analyze arguments is hampered by lack of knowledge of context, intention, and the like, and this is the situation for declaratives just as for imperatives. If it seems like it is more troublesome for the latter, that is, indeed, the case, since often when imperatives are issued as commands (as opposed to requests) they are an abrogation of the rights and will of others, in which case the context is such that the intentions are necessarily less clear than when two people are having a (consensual) conversation: The situation of the starving beggar can be resolved more easily than the situation of the street criminal, i.e. it is surely a simpler matter to ascertain whether something less than ten dollars will do to satisfy the man's hunger than it is to ascertain what the man with his hand on the trigger will do if he is given less than he demands.

Acknowledgments

REFERENCE

Alan Ross Anderson, «A Reduction of Deontic Logic to Alethic Modal Logic,» *Mind* 67 (1958): 100-103.

Joseph S. Fulda

701 West 177th Street, #21

New York, NY 10033, USA

E-mail: kcla@csulb.edu

A

NAIVE VARIETY OF LOGICAL CONSEQUENCE

Enrique Alonso

§1. Two dogmas

This issue argues for a revision of some of the conditions traditionally imposed on any definition of logical consequence. These conditions could be summed up in two dogmas:

[1] Any precise definition of consequence relation on a formal language can be carried out by means of two kinds of resources, syntactic and semantic. There is not a genuine logical system for which only derivability (entailment) could be formally set up.

[2] Derivability for a formal system can be alternatively defined by means of a variety of syntactic resources: axiomatic systems, natural deduction, etc. In contrast with this situation, entailment presents a relatively stable and universal definition: truth-preservation with respect to some class of models.

It is obvious that our first dogma does not say that every formal system must have equivalent proof-theoretic and model-theoretic definitions of logical consequence. Apparently it only affirms that every genuine formal system can be alternatively analyzed in terms of proof-theoretic notions and model-theoretic ones. Nevertheless I think this dogma depends on a deeper thesis, that is, the thesis that *present formulations* of derivability and entailment respond to some frontier inside human mathematical intuition. There are not mixed-definitions inhabiting the space between derivability and entailment, there are not techniques combining proof-theoretic methods with model-theoretic ones to produce new consequence definitions.

This context justifies the importance conferred to soundness and completeness results. To prove the extensional equivalence of two relations defined by means of very different tools is always a matter of some interest and many times it yields positive mathematical knowledge.

Nevertheless, I think that nothing justifies the blank between derivability and entailment. The imaginary frontier dividing these fields could be — for some elementary cases — more a matter of convention than a genuine mathematical

fact. This is the part of the first dogma I do not accept. I think that it should be possible to define new relevant varieties of consequence not obeying this traditional distinction and making use of semantic techniques as well as syntactic ones.

The second dogma listed above constitutes properly the subject of this issue. In this dogma I mention a relatively stable an universal definition for semantic consequence whose format I offer now:

[3] $\Gamma \vdash \beta$ iff $\forall I \in I_v[(\forall \gamma_i \in \Gamma \; I(\gamma_i) \in D^+) \Rightarrow I(\beta) \in D^+]$, where,

1. I ranges over I_v,

2. I_v is the set of every admissible valuations, and

3. D^+ is a proper subset in the range of valuation functions I. I call this subset the set of designated values.

The element in this format which can be modified to give place to almost every imaginable semantic consequence relation is the set I_v. We can consider as admissible valuations over a formal language a great variety of mathematical objets. In fact, it could be a hard task to impose any limits whatsoever to what is admissible at this point. However, the other components in [3] do not offer a comparable level of variation. In fact, it is difficult to imagine any alternative to [3] different from it in some relevant aspect other than I_v. Universal quantification over valuations in some set, universal quantification over formulas in the set of premises, and material conditional between premises and conclusion are features which seem to be intrinsically related to our basic intuitions about semantic consequence.

The variety of consequence I try to define departs from tradition in one of these fundamental features. I have said that the set I_v can be illustrated by different mathematical objects. However, we always have a completely defined set of admissible valuations settled by a precise definition. If we change our set I_v we automatically change our logic. Is it possible to consider a family of sets I_v where we usually take only one? Can we have a suitable semantic definition of a consequence relation based on a substantive family of sets of admissible valuations?

I agree with some deviant schools — relevant and paraconsistent logicians — that some of the conditions imposed by classical logic over sets I_v to be admissible are overrestrictive. Nevertheless I do not think that the solution is merely to liberalize these conditions looking for more permissive ones. This strategy does not differ from tradition in one fundamental aspect: it is always necessary to have some precise criteria in order to define the correct extension of I_v.

What seem to be wrong this time is the strong dependence of a given criteria to define an unique set I_v, a set which remains constant with total independence of the content of those arguments whose validity allows to judge. If we consider a set of criteria sensitive to information codified by standard

propositional language we obtain a family of sets I_v depending on different admissibility conditions, conditions which will be settled by linguistic information codified by arguments in our language.

This suggestion could seem to be paradoxical at a first glance. How is then possible that information codified in arguments could determine admissibility criteria over I_v if all the information we can see codified into a formula is the subset of I_v which satisfies it? Nevertheless I are not alone defending new kinds of relations among language, information, admissible valuations and logical validity.

I think of the Heterogeneous Logic of I. Humberstone — Humberstone [1988] — as a first instance along this line. The novelty supplied by Humberstone is the consideration of two sets of admissible valuations, one for evaluating formulas in the set of premises of a given argument and the other for evaluating the conclusion. Reflexivity, monotonicity and other abstract properties predicable of consequence relations can be recovered by means of conditions relating assignments over sentential variables in premises with those over variables in conclusion. Underlying to this development we can find a very remarkable suggestion: information codified by premises via valuations can contribute in a different way to validity that information supplied by conclusion.

In fact, this is the point which serves to H. Marraud — Marraud [1994] — to elaborate his own suggestion. Under this logic, the set of premises plays a new role with respect to argumental validity. Formulas in premises determine admissibility conditions for those valuations relevant to judge the argument. It would take some time to give a more complete description of this issue so that we omit the details here.

Humberstone and Marraud offer good instances of what can be taken as a new line of research. One which considers that some of the information codified in an argument can have a definitive influence over the aspect and properties of those mathematical objects relevant to set up its validity.

§2. Avoidable commitments

The second dogma described above imposes two kinds of conditions over any suitable semantic definition of a relation of logical consequence. First of all, logical meaning, in the sense of those set of valuations which satisfies a given formula or set of formulas, has to explained in terms of a set belonging to another set fixed from the beginning. In other words, to assign a meaning to a formula we proceed to determine a subset of a previously fixed set — I_v — following appropriated instructions. Secondly, we are supposed to assume as true those conditions implicitly or explicitly followed in the inductive definition of the set I_v.

I do not think these two commitments are of the same importance for our investigation. In fact, I only mention the second one by historical reasons. Nevertheless, it could be of some help in what follows to analyze briefly this point.

R. Routley — Routley [1979] — points out the existence of an ontological commitment lying down classical logic. An examination of relevance failures in classical logic shows that part of responsibility for these failures is

owned to admissibility conditions over valuations. These conditions do not respond, following Routley, to considerations about logical structure, in fact there are not reasons of a purely logical character which could explain some of the requisites classical logic imposes over valuations. The conditions mentioned by Routley are those referred to assignments over variables, that is, those which establishes as a matter of pure logic that the only way to assign a value to a sentential variable is by means of a function whose domain is the set of sentential variables and whose range is {t,f}.

Tradition considers truth functions as a natural basis for assignments over variables. Routley's argumentation shows that truth functions are the resource that classically minded logicians employ to retain some ontological thesis referred to truth and falsity. The world would be furnished in such a way that sentences always have a truth value and never have more than exactly one. To put it otherwise, the world — at least the idealized world logicians consider in focus — only admits consistent and complete state-descriptions.

Till now the argumentation sustained by Routley is, from our point of view, basically correct and highly suggesting. Nevertheless I do not consider his solution an effective way to avoid the ontological commitment just identified. It is true that the problem, so posed, seems to offer an immediate solution. If we admit incomplete and inconsistent assignments over variables the ontological commitment vanishes. It only rests to identify a mathematical resource capable to do the task truth functions execute in classical logic.

Every beginner in non-classical logics can enumerate a list of mathematical techniques developed to do the job. I mention only three: 1) the inclusion of two defective truth values corresponding to non-standard assignments, that is, corresponding to gaps and gluts, 2) the relevant semantics developed by Australian relevant school based on an involution operator «*» inside a Pseudo-Kripkean semantics, and finally, 3) to design a more general resource to assign value to variables, that is, to consider relations of a certain kind where we made use of truth functions.

The first strategy, sometimes used for technical reasons, is the worst response one could afford to solve the problem of the ontological commitment. The inclusion of new values in the set {t,f} only suggests a change of ontology to the effect of liberalize overrestrictive conditions formerly sustained.

I think that Routley's position should be defended on a very different basis. If classical truth functions have to be presented as a subtle way to introduce some ontological thesis in logical machinery, we think that it should be of some interest to find out some mathematical resource of a more fundamental character. Relations $R \subseteq L \times \{t,f\}$ seem to offer the desired tool. Let us note that these relations allow at the same time to codify inconsistent and incomplete assignments, to consider one of these situations independently from the other and finally, to recover classical valuations as a very special case of relations, viz., as total functions.

Relations of the type just described present classical functions as an elaborate tool obtained by successive addition of extra criteria. Moreover, we can obtain this family of semantic resources without removing classical features of sentential connectives, that is, connectives are not responsible for any changes.

If we were to accept Routley's thesis, the strategy which allows relations where we formerly put total functions seems to be the most satisfactory one. It is quite difficult to imagine some other alternative capable to fulfill the requisites demanded by Routley's thesis. The central point of this thesis states that ontological commitment is a consequence of unjustified restrictions over admissible classes of models. Those restrictions are not referred in this case to connectives but to assignments over variables. Relations so stated, provide a starting point for the valuations demanded by Routley, and therefore, they should avoid the acquirement of any kind of ontological commitments.

Routley's thesis depends hardly on an higher level supposition: «there are mathematical tools — or ways to deal with mathematical resources — which do not determine how things are when used to define a suitable semantic». It is true that relations in $L \times \{t,f\}$ show that total functions could be an overrestrictive starting point to do semantics, but it does not mean that we could find out mathematically neutral devices to assign values to sentential variables. In our opinion such ideal starting point is an illusion which results inconsistent with the existence of a previously defined set I_v of admissible valuations and outside of which nothing count as a suitable interpretation. I think that ontological commitment has to do with the existence of this set such as it is conceived by tradition and expressed in [3].

Once we have an assignment over variables which respond to admissibility conditions previously and rigidly stated, we can obtain ontological conclusions with respect to the way language represents facts in the world. Admissibility conditions, however permissive they are, stated at start as definitory conditions for a logic only can be explained through discourses about how things are — or about how we think things are, and so on. Our thesis is that ontology cannot be completely avoided without a deep revision of the way language acquires meaning through semantic machinery.

§3. Admissible valuations and significance

The basic claim sustained by Routley referred to the convenience of a logic independent of considerations about ontology is highly valuable. We agree with Routley in the necessity of intensive research around this problem. Nevertheless we do not think his strategy or those strategies devoted to liberalize admissibility conditions for I_v could achieve an effective solution.

We have conjectured that the solution for this problem has to do with the way in which admissible interpretations are introduced to define semantic consequence. Lines before we claimed for a new relation between arguments and valuations relevant to judge their validity. This relation has to do with information codified by language independently from that obtained from satisfaction with respect to some set of admissible valuations or models.

In what follows we are going to explain the apparent paradox contained in our suggestion. We want to describe some procedure which allows to obtain information from sentences in an argument to the effect of determining the conditions which have to obey the valuations relevant for the validity of this argument. If we have success in the task just enunciated we could arrive at a definition of consequence independent of any set of admissible valuations fixed

from the beginning. In fact the valuations relevant in each case would depend on formulas in the argument analyzed in each moment.

Classical propositional calculus makes of sentences such as $\beta\&\neg\beta$ and $\beta\vee\neg\beta$ very special cases relatively to the set of admissible valuations stated for the matter. Classical tautologies and antilogies have the salient feature to express in the object language conditions imposed in the metalanguage over admissible valuations. Their special status is owned to the fact that those admissibility conditions have been settled from the beginning and remain constant. The logical meaning of a tautology is, under these considerations, the set I_v of every admissible valuations meanwhile the meaning of an antilogy is the empty set. The weight that sentences of this kind have over the validity of a given argument under a definition of consequence like [3] is therefore a very particular one and very different from the weight that contingent sentences have.

We'll say that every sentence whose logical meaning corresponds to the entire space of valuations, or alternatively to the empty set, present a «conflict of significance». We'll make extensive this term to sentences not incurring in that situation but containing subformulas which present such conflict.
Nevertheless, it is quite easy to find contexts where tautologies and antilogies are used in a significative way. They are not used to make mention to admissibility conditions — to some informal mate — but to genuine information. This fact takes place, for instance, when we discover, perhaps with some surprise, that some sentence β and its negation are both true. We do not want to say that β is therefore paradoxical, we only want to express what we have said and the way to do that is by means of a true — not paradoxical — contradiction $\beta\&\neg\beta$. Something very similar could be said with respect to the falsity of tautologies. We can think of situations which make false β and its negation as giving place to a false — not undefined — tautology.

These real life considerations can be found amongst the motivations of many partial and paraconsistent logics. The most significative developments in these areas adopt technical devices which generalize anomalies — gaps and gluts — equating them with standard valuations — true and false. I do not think situations like those we have described are the norm but the exception. Priest has defended this same position about paraconsistency in many places but his logic LP does not comply with this intuitive principle.

The pool of strategies developed to capture inconsistent (partial) situations via valuations generalizes this possibility allowing assignments which attribute the value «paradoxical» («undefined») to every sentential variable. This is an immediate consequence of liberalizing the admissibility conditions associated to the set I_v and therefore an effect of the standard definition of semantic consequence relation.

Our task will be to define a consequence relation whose validity criteria includes amongst other things, the condition that every formula present in an argument occurs significatively in the context supplied by that argument.

A significative occurrence of a formula in an argument is an informal notion which requires additional comments. Nevertheless, we cannot say very much in this moment. We know that a significative lecture of a formula can make true a classical antilogy declaring true and false some sentential variable and that

some parallel situation can stated for false tautolgies. We also know that this movement should not be predicated of sentential variables whatsoever, we do not accept inconsistencies or incomplete information without an explicit and concrete reason to proceed in that way. We can conclude therefore that significative lectures of formulas cannot be obtained merely modifying the acceptability conditions for valuations. This operation would affect to the entire language and this is a possibility we explicitly reject. If we analyze an argument looking for significance conflicts, it is possible that some sentential variables should be interpreted as allowing gaps (gluts) meanwhile the rest retain a perfectly classical behavior or it is still possible that more complex alternatives have to be considered.

The procedure to be developed will take as starting point the context supplied by an argument — eventually a set of formulas — locating significance conflicts which affect to sentential variables occurring in that argument. Once we have identified these sets of variables we have to determine the conditions to be satisfied by relevant valuations for the validity of the argument, leaving apart all those sentential variables which does not present any conflict in that argument. It would be a lost of time to delay the formal translation of these considerations.

§4. A significative variety of consequence

In the sequel we are going to adopt a formal expression for our fundamental notions and concepts. Anyhow we avoid detailed proofs which many times do not offer extra information to the reader and prevent an illuminating comprehension of the main ideas.

In what follows we are going to deal only with finite sets of formulas. As we shall see this is a restriction associated to some essential features of the procedure developed to establish significative lectures of formulas in arguments.

> ***Definition 1***: Let Γ be a finite set of formulas and let γ be its characteristic formula, i.e., $\gamma = \bigwedge \beta_i$ for β_i in Γ. We shall call μ a **bearer subformula** in Γ (in γ) iff μ is a classical tautology or antilogy and no subformula of μ has this property.

To identify bearers in Γ — or in formulas whatsoever — is the first step to obtain a significative lecture of formulas in some set Γ. Let us note that the notion of a bearer relative to some set Γ goes beyond the notion of tautological (antilogical) sentence. We can have sets whose respective characteristic formulas are not tautologies (antilogies) and have bearers in Γ pointing out inner conflicts of significance in that set. Let us note that a bearer in Γ goes beyond classical tautlogies (antilogies) also in another sense: only innermost classical tautologies (antilogies) are bearers. We proceed to identify first the smaller pieces of classical structure which can result responsible for conflicts of significance.

To find out bearers in Γ is a process which has to be associated to some effective method. We are going to adopt a procedure based on analytic tableau calculus — TA. We define the **positive (negative) tableau for** β, $T^+(\beta)$ $(T^-(\beta))$ in symbols as those tableau whose top is given by β ($\neg\beta$). We make use of

positive tableau to look for antilogies and negative tableau to look for antilogies. Once we obtain a closed tableau we have to analyze if the generating formula contains subformulas whose positive or negative tableau result to be closed. If this is the case, the generating formula is not a bearer. As we can see the procedure is a bit tedious, anyhow it is not difficult to realize that is an effective decision method.

>   ***Definition 2*:** We say that a formula μ is **completely analyzed** iff each occurrence of a bearer in μ has been labeled with an auxiliary symbol — put «*».

A completely analyzed formula shows by direct inspection which are the innermost and smaller pieces of its structure which presents conflicts of significance. Anyhow, we are supposed to obtain from these conflicts information about admissible valuations for the set in which they occurs, and we are supposed to do that by means of some conditions referred to sentential variables. It is of fundamental importance to realize that a bearer is associated at most to a closed tableau, positive or negative. The sentential variables responsible for the closure of this tableau can be founded inspectioning paths in the tableau. In what follows we shall speak of the tableau associated to a bearer to mention the tableau which identifies that subformula as a bearer.

>   ***Definition 3*:** Let $T(\psi)$ the tableau associated to a bearer subformula ψ. By the set $L(\psi)$ of **bearer atoms** of ψ we understand the closure under unions of the set whose members are the sets of atoms responsible in each path of $T(\psi)$ for the closure of this path.

This definition could seem more complex than expected. Nevertheless, it is justified by the impossibility of determining an unique set of atoms associated to the closure of a tableau. Let us consider the formula $\psi=(A\lor B)\&\neg(A\lor B)$. The completely analyzed formula generated by it is $[(A\lor B)\&\neg(A\lor B)]^*$, and its characteristic — positive, in this case — tableau $T(\psi)$ consists of two paths, one closed by the presence of $\{A,\neg A\}$ and another closed by $\{B,\neg B\}$. If we consider the ways to avoid the closure of $T(\psi)$ we find that the sets $\{A\}$, $\{B\}$ and $\{A,B\}$ are equally responsible for the closure we try to block. This example makes clear the reason which carry us to adopt such a strange definition for the set $L(\psi)$ of **bearer atoms** of ψ.

>   ***Definition 4*:** By the set $\Sigma(\beta)$ of **conflicting atoms** of formula β we understand the union of all those sets $L(\psi)$ where ψ is a bearer in β.

We now introduce a definition decisive in what follows.

Definition 5: We shall say that a formula μ is a **truth-specification** of ψ iff it is of the form ψ_s where *s* is a finite sequence in the Cartesian n-product $\{0,1\}^n$, for n finite.

We are supposed to give an intended interpretation for truth-specificators, one which allows to read formulas of the form ψ_s. Mimicking a recursive definition we shall read ψ_{s1} as saying that «ψ_s is true» and ψ_{s0} as saying that «ψ_s is false». We do not pretend to offer anything different from a technical device, but it has to be recognized that truth-specificators can play a role important by philosophical reasons. For instance, we think that a true contradiction says exactly $\psi_{s1}\&\psi_{s0}$ and something similar could be said with respect to false tautologies. Truth-specificators seem to be of some utility when we have to express facts which differs form what is considered usual with respect to truth. Truth-specificators seem to be a way out for the norm in matters where truth plays a fundamental role.

So much for philosophy. The notions of bearer subformula, conflicting atoms, and finally, truth-specification of some formula allows to define a procedure to deal with conflicts of significance. We are going to solve the conflicts identified in a completely analyzed formula by means of its bearers making use of truth-specificators affecting to conflicting atoms. This solution suggests to make use of a translation function from standard classical propositional language to a propositional language accepting truth-specificators over sentential variables. It is pointless to say that truth-specifications over conflicting atoms yield some information about admissibility conditions for those valuations which makes significative the formula analyzed.

Our translation have to be defined in terms of a composition of two translation functions, t/Δ and g/Δ, both relative to a set Δ of atoms contained in $\Sigma(\beta)$ for a certain formula β.

Translation function t/Δ:

c0) If $\beta_s \in Var$, then:

    i) if $\beta_s \in \Delta$, then $[(\beta_s)_i]^t = (\beta)_i$, where $i \in \{0,1\}$

    ii) if $\beta_s \notin \Delta$, then:

$$\text{a) } [(\beta_s)_1]^t = \beta$$
$$\text{b) } [(\beta_s)_0]^t = \neg\beta,$$

where Δ is a finite set of atoms.

c1) If $\beta = \neg\alpha$, then:

    i) $[(\neg\alpha)_1]^t = \neg[(\alpha)_0]^t$

    ii) $[(\neg\alpha)_0]^t = \neg[(\alpha)_1]^t$

c2) If $\beta = (\alpha \vee \psi)$, then:

> i) $[(\alpha \vee \psi)_1]^t = [(\alpha)_1]^t \vee [(\psi)_1]^t$
>
> ii) $[(\alpha \vee \psi)_0]^t = [(\alpha)_0]^t \& [(\psi)_0]^t$

c3) If $\beta = (\alpha \& \psi)$, then:

> i) $[(\alpha \& \psi)_1]^t = [(\alpha)_1]^t \& [(\psi)_1]^t$
>
> ii) $[(\alpha \& \psi)_0]^t = [(\alpha)_0]^t \vee [(\psi)_0]^t$

c4) If $\beta = (\alpha \rightarrow \psi)$, then:

> i) $[(\alpha \rightarrow \psi)_1]^t = [(\alpha)]^t \vee [(\psi)_1]^t$
>
> ii) $[(\alpha \rightarrow \psi)_0]^t = [(\alpha)_1]^t \& [(\psi)_0]^t$.

Translation function g/Δ:

c0) If β is of the form $(\alpha)^*$, then $[(\alpha)^*]^{g/\Delta} = [(\alpha)_1]^{t/\Delta}$

c1) If β contains some subformula of the form $(\alpha)^*$, α different from β, then:

> i) $\beta = \neg\psi$, $[\neg\psi]^{g/\Delta} = \neg([\psi]^{g/\Delta})$
>
> ii) $\psi = (\alpha \circ \psi)$, $[(\alpha \circ \psi)]^{g/\Delta} = ([\alpha]^{g/\Delta} \circ [\psi]^{g/\Delta})$, where $\circ \in \{\vee, \&, \rightarrow\}$

c2) If β does not contain any formula of the form $(\alpha)^*$, then $[\beta]^{g/\Delta} = [(\beta)_1]^{t/\Delta}$,
Δ being a finite set of atoms.

The way in which t/Δ and g/Δ are related can be deduced from clauses c0) and c2) for g/Δ.

Now we can define one of the most fundamental notions in this issue:

Resolution of β: By **a resolution of a sentence** β we understand the following formula:

$\vee_{\Delta \in \Sigma(\beta)} (\beta)^{g/\Delta}$.

We think that some detailed example could clarify the procedure just defined.

Example: Resolution of $\Gamma = \{\neg(p\&\neg p), p, \neg q\}$

1. $\neg(p\&\neg p)\&p\&\neg q$	Characteristic formula of Γ.
2. $\neg[(p\&\neg p)]^*\&p\&\neg q$	by the routine for bearer subformulas.
3. $\Sigma(\beta)=\{\{p\}\}$	by the routine for conflicting atoms.
4. $[\neg[(p\&\neg p)]^*\&p\&\neg q]^{g/\{p\}}$	first step of the translation routine.
5. $[\neg[(p\&\neg p)]^*]^{g/\{p\}}\&[p]^{g/\{p\}}\&[\neg q]^{g/\{p\}}$	by c1)ii of g/Δ
6. $\neg[(p\&\neg p)]^{*g/\{p\}}\&[p]^{g/\{p\}}\&[\neg q]^{g/\{p\}}$	by c1)i of g/Δ
7. $\neg[(p\&\neg p)]^{g/\{p\}}\&[p]^{g/\{p\}}\&[\neg q]^{g/\{p\}}$	by c0) of g/Δ
8. $\neg(p\&\neg p)_1^t\&p_1^t\&(\neg q)_1^t$, [mód.$\{p\}$]	by c2)
9. $\neg(p_1^t\&(\neg p)_1^t)\&p_1^t\&q_2^t$	by c3) and c1)ii of t/Δ
10. $\neg(p_1^t\&p_2^t)\&p_1^t\&q_2^t$	by c1)ii of t/Δ
11. $\neg(p_1\&p_2)\&p_1\&\neg q$	by c0)i and ii of t/Δ.

Once again we take finite sets and its characteristic formulas as interchangeable notions when needed.

I mention some other examples without going into details. Let Γ be the set $\{(p\vee q), \neg(p\vee q)\}$. The set of conflicting atoms associated to its characteristic subformula β is $\Sigma(\beta)=\{\{p\},\{q\},\{p,q\}\}$ what yields a resolution consisting in:

$[(p_1\vee q)\&p_2\&\neg q]\vee[(p\vee q_1)\&\neg p\&q_2]\vee[(p_1\vee q_1)\&p_2\&q_2].$

One of the most salient features of the way we are dealing with conflicts of significance is that it proceeds stepwise. We first localize the innermost bearer subformula of a given formula — finite set — and then we determine a resolution for this formula accordingly to the appropriate definition. Nothing prevents that a resolution of a formula could contain itself conflicts of significance of an higher order. An example easy to understand is given by $\Gamma=\{(p\&\neg p), \neg(p\&\neg p)\}$. Its resolution yields the formula $(p_1\&p_2)\&\neg(p_1\&p_2)$ which is not free of significance conflicts. This time our conflicts affect to truth-specifications of standard atoms which can be taken as new atoms if is necessary.

Anyhow, the conflict showed by this formula can be solved iterating the entire process once again. This time the resolution will be the formula $(p_{11}\&p_{21})\&(p_{12}\vee p_{22})$.

> ***Definition 6***: By the **last resolution** of a formula β, β^σ in symbols, we mean the resolution free of bearer subformulas.

It is quite obvious that last resolutions of formulas constitute the basic elements to define the variety of consequence we were looking for. Nevertheless, the process of resolution just defined does not make mention to arguments, it only deals with formulas and finite sets of formulas. But an standard argument is nothing different form an ordered pair $<\Gamma, \beta>$ of sets of formulas and formulas respectively, so that some relation can be expected. If we ignore order and limit ourselves to finite sets of premises, we obtain a set $\Gamma \cup \{\beta\}$ which seems to offer the natural context to execute our resolution procedure. The relevant information supplied by this set is that referred to conflicting atoms, and once we have a set $\Sigma(\Gamma \cup \{\beta\})$ generated by some argument $\Gamma \vdash \beta$, for Γ finite, we can look for successive resolution of the formulas in the argument. Eventually we can reach a resolution for premises and conclusion which satisfy conditions imposed by definition 6. The resulting argument is the translation of the original argument in a sentential language allowing specifications of atoms and constitutes a lecture free of conflicts of significance.

***Naive consequence*:** Let Γ be a finite set of formulas and let β be a formula, then,

$\Gamma \Vdash \beta$ iff $\Gamma^{\rho} \vdash_{cpc} \psi^{\rho}$, with respect to the set $\Sigma(\Gamma \cup \{\beta\})$, being Γ^{ρ} and ψ^{ρ} the last resolutions of the characteristic formula of Γ and β respectively.

I omit the proof that the resolution method is effective in the sense that it stops reaching an argument formed by the last resolutions of all those formulas in the argument under examination.

§5. Some comments about naive consequence and its applications

I do not justify the utility of this variety of consequence appealing to some successful research program outside the main topics in Logic. On the contrary, I think that naive consequence can bear its fruites once traditional problems in logic are revisited under new perspectives.

Naive consequence offers a partial and paraconsistent variety of a consequence relation preserving classical inferences when possible, that is, if there is not explicit significance conflicts.

Priest in Priest [1979] introduces the notions of «valid» and «quasi-valid inference» in the context of a revision of Gödel's incompleteness Theorems from the point of view of paraconsistent logic. The argument given by Priest shows that a paraconsistent interpretation of Gödel's incompleteness Theorems can be carried out without going into triviality. Now Gödel sentence is not independent from Peano Arithmetic, it results to be paradoxical and therefore, following Priest, an acceptable consequence of PA axioms. Unprobability of PA-consistency in PA offers another example of paradoxical sentence therefore removing a considerable amount of classical orthodoxy.

The success of this paraconsistent threat to arithmetical orthodoxy hangs on the adequacy of the elementary logic supplied in the place of CPC. Nevertheless, the Logic LP developed by Priest fails in some relevant aspects. For instance, this logic does not recognize the intended difference between valid and quasi-valid inferences. LP is a paraconsistent logic obtained by means of a revision of admissibility conditions over assignments an therefore, LP is paraconsistent everywhere. Its characteristic consequence relation cannot distinguish those inferences acceptable only in consistent contexts of deduction — where no paradox is present — from those inferences valid everywhere. LP rejects MP and DS because paradoxes can occur in premises allowing to conclude something false. Let us take as instance of DS the argument: A, $\neg A \lor B \vdash B$. The reason to reject its validity is that nothing prevents, in a paraconsistent Logic as LP, that A could be true and false at the same time, what destroy the basis to conclude B. From the point of view of naive consequence an instance of DS such as A, $\neg A \lor B \vdash B$ remains valid. No conflict of significance is present, or at least, no such conflict has been made explicit, so that we are not supposed to proceed under considerations not supplied by the argument under examination. However, if we look for the last resolution of a variant for DS such as $\neg A$, A, $\neg A \lor B \vdash B$ we see that it is naively rejectable. Now we know that A has been taken as a true and false sentence and this is enough to validate Priest argument. If we compare both instances of DS we see that naive consequence does not obey monotonicity but fortunately this is not a very popular property nowadays, at least if we consider the relative success of non-monotonic logics.

I bring these comments to a close by pointing out some problems we shall be faced with in further developments. Naive consequence has been set up in an indirect way. No semantics has been defined nor inference rules or axioms have been offered. From the point of view of orthodoxy, such as it is resumed by the dogmas listed above, we have not a genuine mathematical interpretation of any consequence relation. At most we have a translation which defines some secondary mathematical object probably having to do with consequence.

The definition of naive consequence depends on the resolution procedure and classical consequence relation. Resolution is a syntactical tool based on finitary considerations. Once we have obtained the last resolution of an argument in does not matter if we take classical derivability or classical entailment to establish naive consequence in terms of the resolution just mentioned. The resolution has to be effective in a sense that makes of naive consequence a syntactic notion whose semantic mate could not be easy to define. Nevertheless, a resolution of an argument is nothing different from another argument in a language allowing truth-specificators for atoms. But this notion of truth-specificators has a semantic flavor I do not want to deny. We should think of a naive analysis of an argument as a procedure which has to determine — in an effective way — first the admissibility conditions for relevant valuations and then everything goes classically. We can say that naive consequence changes the logic used to analyze arguments always that the context given by an argument requires such a change. To make clear the point, we can think of naive consequence as a relation which goes along the entire hierarchy $C_0, ... C_n$, $n < \omega$ of Arruda and da-Costa looking for the most convenient system for the occasion. Technically we can say that naive consequence is like C_ω a non-finitely trivializable logic, but it does retain like C_n for $n < \omega$ an strong negation.

The intended relation between naive semantics and the hierarchy $C_0,...C_n$ goes beyond paraconsitency including also the partial behavior exhibited by naive consequence. Anyhow this relation only can be taken as a metaphor, in the limit it does not work.

We can think of resolution method as a procedure which serves to identify the logic that in a chain of the kind of Arruda's results relevant to judge an argument. The maximal length of truth-specificators in the last resolution of formulas in an argument can be taken as an index for the position that this logic takes in the ordered chain. This suggest an inductive process which can be of a fundamental utility to extend resolution method to non-finite sets. If we define resolution process inside an inductive procedure we can expect for a fixed-point theorem. We are supposed then to extend the resolution method allowing for infinite truth-specifications for atoms and prove that any set can reach a point where no further conflicts of significance occur. Anyhow we can think of sets so defined, that conflict of significance always persist, for instance,

Example: Put $\mu_0^+=\beta$, and $\mu_0^-=\neg\beta$. Now we define inductively $\mu_{i+1}^+=\mu_i^+\&\neg\mu_i^-$, and $\mu_{i+1}^-=\neg\mu_{i+1}^+$. We obtain the sets $\Gamma_j=\{\mu_i^+/0\leq i<j\}\cup\{\mu_i^- /0\leq i<j\}$, and finally let Γ_ω be the union of each Γ_i. It is quite obvious that the resolution process for this set does not seem to reach a fixed-point.

All these comments show that naive consequence does not fit well into the framework of standard semantics and, in general, of standard definition strategies for logical consequence. However, I think that it is not the last word to say about naive consequence, but the effect of a significative departure from tradition. I guess that some effort can yield genuine semantics and proof-theoretic tools to deal with naive consequence, and so to prove its utility in partial and paraconsistent revision programs for fundamental paths of contemporary mathematics.

References

da Costa, N.C.A. & Marconi, D. [1989]: «An Overview of Paraconsistent Logic in the 80's», *The Journal of Non-Classical Logic*, vol.6, nº1, pp. 5-32.

Humberstone, I.L. [1988]: «Heterogeneus Logic», *Erkenntnis*, 29, pp.395-435.

Marraud, H. [1994]: *Curso de Doctorado*, Dept. de Lingüística, Lenguas Modernas, Lógica y Filosofía de la Ciencia. U.A.M.

Priest, G. [1979]: «The Logic of Paradox», *Journal of Philosophical Logic*, 9, pp.415-435.

Priest, G. [1984]: «The Logic of Paradox Revisited», *Journal of Philosophical Logic*, 13, pp.219-235.

Routley, R. [1979]: «The Choice of Logical Foundations: Non-Classical Choices and Ultralogical Choices», *Studia Logica*, vol.XXXIX, nº1, pp.77-98.

Enrique Alonso

Unirsidad Autónoma de Madrid

E — 28071 Madrid, Spain

Humo

r and Harm

Laurence Goldstein

It is important, in attempting to combat sexism and racism, that the sources and manifestations of these vices not be mislocated. The reason why this is important is purely pragmatic: if the battles are fought on the wrong fronts then the fighting of well-meaning people will be ridiculed, their efforts will be counterproductive and the war will be lost. This worry has surfaced in a debate that has been raging for some time about the propriety of a writer's gender-neutering his or her pronouns. On another front, some sanctimonious philosophers have been arguing of late that certain *jokes* are not funny, or, at least, that anyone who finds them funny betrays racist or sexist attitudes.

Presumably if we ourselves find such jokes funny then, if we are persuaded by the argument and condemn racism and sexism, we shall educate ourselves to so abhor what we condemn that we no longer laugh and we will censure and perhaps seek to censor such forms of humor. However, censoring humor and getting people not to laugh is a pretty serious business and I somehow doubt that this is a front on which scholars should be fighting. I question the claim that to laugh at a joke which employs sexual or racial stereotypes is in general, an indicator of sexist or racist attitudes.

At the opposite extreme from those whom I have tendentiously dubbed sanctimonious, are those whom, equally tendentiously, I shall call callous, who claim that all joking is *only* joking, and therefore does not raise any serious moral or social problems. For example, the author of an extensive survey of ethnic humor world-wide concludes that «jokes... are not thermostats regulating and shaping human behavior, but they are social thermometers that measure, record and indicate what is going on. To become angry about jokes and to seek to censor them because they impinge on sensitive issues is about as sensible as smashing a thermometer because it reveals how hot it is. Those who do so deserve all the extra derision they then incur, for they are fools indeed.»[1]

There has also been a popular backlash against what is perceived as thought-control. The *New York Times Magazine* carried a series of letters in response to an article about the quadriplegic sick cartoonist John Callahan. One correspondent wrote: «John Callahan's work is a welcome antidote to the intellectual poison of the so-called politically correct movement, one of the

---

[1] Christie Davies, *Ethnic Humor Around the World: A Comparative Analysis* (Bloomington, Ind.: Indiana University Press, 1990), p.326

devitalizing forces in American culture and language. The P.C. movement has elevated euphemism from social palliative to socio-political fetish.»[2]

This position is simplistic. A Callahan cartoon called «Alzheimer Hoedown», which depicted confused couples at a dance, scratching their heads, unable to follow the instruction «return to the girl that you just left», not surprisingly upset people suffering from Alzheimer's, and it is doubtful whether taking into account the feelings of those people ought to be regarded as a socio-political fetish. Also, one cannot simply dismiss the worry that the attitudes of young children are to some extent shaped by the prevailing humorous norms. However, there is no incompatibility between an individual's feeling uneasy about the uncontrolled dissemination of certain types of joke and his finding those jokes funny. The claim that I am disputing is that there is something morally wrong with a person who does find them funny.

One might think that there was nothing particularly philosophical about the disputed claim; that the explanation of the relation between our attitudes and what we find funny is a matter for psychology to discover.[3] That this thought is incorrect may be seen by considering how we it could be possible to empirically *refute* the claim: one would present a set of such jokes to a subject about whom one had no reason to believe that he or she was a sexist or a racist; if the subject laughed at any of the jokes, then the claim would be falsified and my view would be vindicated.

It takes but a moment's reflection to realize that such an experiment would be worthless. For an alternative account of the result of the experiment might be that, despite all the evidence we had accumulated about the subject's attitudes, the subject's response proved that he or she was sexist/racist after all. Moreover (so the alternative story might continue), we should not be too surprised at this outcome since most of us, despite what we profess and believe about ourselves are *viscerally* sexist or racist.[4] We can even have or, as a result of exposure to such jokes, come to have, negative attitudes towards groups of which we ourselves are members. One could envisage a refinement of the experiment in which subjects antecedently ranked for sexism or racism were presented with a batch of jokes, and the differing extents of their amused reactions recorded.[5]

---

[2]   Ted Blumberg, «Defiantly Incorrect,» *The New York Times Magazine* (28 June, 1992).

[3]   There is quite a large body of literature on such psychological experimentation. See, for example, Dolf Zillmann and Jennings Bryant, «Retaliatory Equity as a Factor in Humor Appreciation,» *Journal of Experimental Social Psychology* 10 (1974): 480-488.

[4]   See Irving Thalberg, «Visceral Racism,» in *Today's Moral Problems*, ed. Richard Wasserstrom (New York: Macmillan Publishing Co. Inc., 1975), pp.187-204.

[5]   As was done in Zillmann and Bryant, *op. cit.*.

However, it should be clear that the kinds of measurement required for this experiment lie well outside the limits of reliable quantitative methods, and, besides, one could not detect or reliably ascertain what the subjects were laughing *at* nor whether, say, a reaction of mild amusement was the behavioral manifestation of being mildly amused or was the result of successfully supressing an overt indication of intense amusement.

Experimentation alone, then, will not suffice to settle the claim one way or the other. And perhaps experiment is beside the point if there is not a contingent but a necessary connection between finding jokes about Xs amusing, and having certain beliefs about Xs. That the connection is, indeed necessary has been argued by Ronald de Sousa.[6] De Sousa considers the following example:

N. [a woman in the public eye famous for her alliances] goes to visit the hockey team. When she emerges, she complains that she has been gang-raped. Wishful thinking.

I'm inclined to agree with de Sousa, that the joke, in precisely this form, is a malicious one; that is to say, anyone who finds it funny is likely to bear malice towards N.. But my reason for thinking this (which differs from de Sousa's), is that the joke has so few humorous features. It is not sonorous — it cannot be delivered well and the punch line lacks vitality — and although the idea of someone's having a voracious sexual appetite is a common humorous theme, the very fact that it is so frequently employed means that most of us will find it stale and wearisome unless given a new twist — which the present joke does not provide. But, more important, the non-idiomatic term «gang-raped» is used instead of the more colloquial, fun expression «gang-banged». For an uninhibited woman, to be gang banged is not inconsistent with her having a good time (my informant made the obvious proviso: that she like and be sexually interested in every member of the gang), but to be raped is to have sexual intercourse *against one's wishes*, so to characterize a non-masochistic woman's thought of being raped as wishful thinking is just a stupid contradiction.

A lot of humor depends on perceiving lurking contradictions, so it is important to distinguish the subtle from the stupid. I would count the following extract from a newspaper report as containing a subtle (and therefore amusing) contradiction:

«The dead man was white, in his mid-thirties and spoke with an Irish accent.»

---

[6]   Ronald de Sousa, *The Rationality of Emotion* (Cambridge, MA: MIT Press, 1987), esp. pp.275-299; a version of the relevant secion is included in *The Philosophy of Laughter and Humor* ed. John Morreall (Albany: State University of New York Press, 1987), pp.226-249. De Sousa, incidentally, has a footnote (p.292 = p.249 in the Moreall) whimsically canvassing some difficulties inherent in psychological experimentation on mirth.

The contradiction here (a dead man speaking) arises from failure of substitutivity of identicals abetted by the ambiguous temporal reference of «spoke». It takes a certain amount of perceptiveness to realize how the contradiction has come about (this typically wouldn't involve being able to *formulate* it in the way that I've just done) and there is a great deal of pleasure associated with this exercise of the intellect. Nothing comparable is present in the joke about N.. I conclude that, in most cases, a man who finds that joke funny derives malicious pleasure at the thought of the abasement of a *particular person* (N.) whom he may despise or envy. Whereas de Sousa sees laughing at this joke as evidence of sexism — as evidence of the belief that generally women's sexual desires are indiscriminate — I think it is clear that if one substituted «a woman» for «N» in the joke, then the class of people who laugh at the revised version would not include all those who laugh at the original.

In practice, of course, it is frequently by no means clear what aspect of a joke someone finds funny. Consider, for example

Q.: If you keep your beer in beer cellars, and your wine in wine cellars, where do you keep your knives?

A.: In Monica Seles.

Monica Seles is a hugely successful and truculent tennis star who was stabbed during a tennis tournament by a crazed fan of her main adversary. Undoubtedly some who envy her success or who dislike her attitude to the game will laugh at the joke for reasons that make the rest of us feel rather uncomfortable. But the rest of us may well laugh at the joke because of its ingenious play on sounds or because the idea of using a human body as a convenient medium for storing knives, like the idea of using buttocks for parking bicycles, is compelling and absurd.

The point of the above discussion is not to suggest that there are no sexist jokes directed against women in general, but to warn against the danger of seeing sexism or misogyny when they are not present. Quite clearly, in the joke about N., the same form of joke could have as its subject a well-known stud visiting a women's hockey team, and the humor (such as it is) would be preserved. The same is true in the Monica Seles case, which could take as its subject any controversial celebrity who had been stabbed and whose name rhymes with that of a receptacle. This interchangeability of the subject (or subject-group) of a joke is a clear indication that one can find the joke amusing for reasons other than its being targetted on that subject or subject-group. It is therefore wrong to condemn the following as an example of sexist humor:

A husband says to his wife «Women always take everything so personally,» and the wife replies indignantly «*I* don't».

Again, *formulating* what the laugher is laughing at is quite difficult, and that is some indication that the laughter is an expression of intellectual pleasure. The wife is taking her husband's remark personally in the very act of denying that she does this. Perhaps the realization of this is sufficient to make us laugh. But it

is also likely that we instantaneously fill in a bit of background. E.g., the husband's remark occurs in the middle of a row. He thinks that his wife is taking something personally and wants to criticize her on this score, but in order to defuse the situation, he expresses his criticism in a general, impersonal way. Yet her reply, with the stress on «I», indicates that she is still taking things personally.

Of course, some will find the use of a stereotype (the nagging, niggling wife) an added humorous element, and, for this reason, it may be right to claim that such jokes are bad because they cause harm to women. But the joke works (though less well) without this element. For example, we might tell a story about two men having an argument, in the course of which one says to the other «Some guys take things so personally»; and the other replies «*I* don't». The original joke is not funny *just in virtue of* having a stereotyped subject; the subject could be changed and the humor would not completely disappear. It is important to notice, though, that the employment of a stereotype does make *some* difference. To describe someone's face as looking like a bag of nails is funny, but it's funnier when the person so described is one's mother-in-law. The mother-in-law in question is not, of course, one's own but is a representative of the stereotype *mother-in-law*, just as it is the stereotypical woman, not all actual women, who always take things personally.

Even when subject-interchangeability is not possible, it may still be the case that finding a joke funny does not amount to holding contemptuous and contemptible attitudes towards its subject. The following rape joke, I wish to suggest, can be enjoyed with a clear conscience:

A woman, returns home late one night and reports to her flatmate «I've been graped». The flatmate replies «Don't you mean 'raped'?» «No, there was a whole bunch of them.»

The subject of this joke is a woman (any woman), so that one might be inclined to think that those who enjoy it reveal sexist attitudes. However, unlike our first example, of rape humor, this one has very many redeeming features. Apart from the obvious play on words («bunch of grapes»), the other elements that contribute to making this joke funny are (i) the incongruity of an alarming outburst being turned into a philological discussion, (ii) the satisfying appropriateness of the prefix «g» before «raped» which has connotations of *g*reat, *g*igantic magnitude (as in «g-forces», «giga-bytes») and which can be thought of as abbreviating «gang», and (iii) the phonetic and orthographic similarity of the newly minted verb with «to grope», a verb quite frequently used for another sexual offence, which produces a punning element. Why features like these are mirth-making is a difficult question to which nobody yet has a satisfactory answer. But *that* they are is beyond dispute since they figure in all manner of jokes which are inoffensive by anyone's lights. The only reason why the joke alludes to women is simply that it is they who are usually the victims of rape. That aside, the reference to *women* is inessential to the humor. If someone finds this joke amusing because a woman is the fall-guy, one can only say that he is laughing at the wrong thing, and the same would be true of someone who laughed just because he found the word «rape» funny.

Two further points of some importance are, first, that this joke is 'impersonal' — one is not laughing at the misfortune of a real rape victim.

Second, it may be true that to laugh when this joke is told implies that one is not focusing on such facts as that a high proportion of women are raped, that the assumption that women enjoy rape has made legal redress difficult for raped women to obtain, that rape occurs in the same social context in which female children are molested and adult women are battered by their partners etc... But is it any more morally reprehensible to temporarily put such things to the back of one's mind than it is to put to the back of one's mind the fact that Frege and Wagner were fascists when one is enjoying reading the beautiful *Die Grundlagen der Arithmetik* or listening to the great *Ring* cycle? One occasionally enjoys a good meal knowing full well that concurrently two thirds of the world's population is going hungry. It would be morally reprehensible for an educated adult not to be acquainted with facts about starvation, discrimination and oppression; we should dwell on such facts frequently and educate our children to dwell on them too. But morality does not demand that we do so all of the time.

Similar remarks can be made about many allegedly racist jokes. I have heard children retailing the following riddle:

Q.: How do you stop a black man jumping up and down on your bed?

A.: Put velcro on the ceiling.

Are we really to say that those who find this funny are racists? Well, first, the joke passes the subject-interchangeability test. Children find almost as funny this variant:

Q.: How do you stop a robot jumping up and down on your bed?

A.: Put a magnet on the ceiling.

Second, in either of its forms, this has the hallmark of a good riddle: The question sets the mind racing on a flurry of wild goose chases; the solution is unexpected and punchy. The reason that the variant is not quite as good as the original is that, while the adhesion of metal to magnets is common and of no particular interest, the idea of a velcro-to-frizzy-hair bond is vivid and imaginative — an idea that gives pleasure because of its ingenuity. Of course, someone who *tells* this joke may intend simply to denigrate blacks, perhaps by suggesting that their behavior is too uninhibited so that the resulting punishment of being rendered ludicrously helpless perfectly fits the crime (compare mediaeval stocks). But surely someone could be criticized for *laughing* at the joke only if his laughter stemmed from similarly hostile attitudes.

I would claim that, even in jokes where allusion is *essentially* made to sexual or racial stereotypes, the fact that someone finds them funny does not necessarily indicate that he holds sexist or racist views. Another riddle may illustrate the point.

Q.: What does a Jew do with his old razor blades?

A.: He shaves with them.

Here the question sets our minds racing in the direction of circumcision, but the resolution has to do with something quite different — meanness — which, in humorous contexts, is stereotypically attributed to Jews as also in the following which combines that with the stereotype of Jewish aversion to sports.

Q.: What does a Jew think is the point of American Football?

A.: To get the quarterback

Now, there is a danger, and one that should not be underestimated, of creating such stereotypes, because not everyone can sharply separate caricature from character. Hence sexist and racist humor can instill bad attitudes, and can foster poor self-images among members of the caricatured groups. So a case could be made for refraining from telling such jokes. But, when you hear the above riddle for the first time, then, although you may disapprove of the person telling it, and although you may have absolutely nothing against Jews, you will probably laugh. Why? The reason is not that you temporarily adopt the perspective of the bigot[7] (could you adopt the perspective of Hitler and his followers, past and present, and laugh at holocaust jokes?) but because the joke is short and deft and could only be so if it relied on a shared background of make-believe assumptions. One does not need to hate lawyers or to «adopt the perspective» of lawyer-haters to enjoy this:

Q.: What do you have if you have a lawyer buried up to his neck in sand?

A.: Not enough sand.

A good joke, like a valid argument, can often rest on assumptions known to be false.

The proof that we can enjoy sexual or racial humor about certain groups without holding unfavorable beliefs about those groups is that, without holding *any* antecedent beliefs about a certain group, we can make it the target of such humor. For example, I have never held the belief that Cornish people are incestuous and I do not hold it at present. Nor do most people. But now, having sown the seed in your mind, I can tell the following story of a young Cornish man disconsolately reporting to his father that, having asked his girl friend to marry him, the proposal was rejected. His father asks, «Was she a virgin, son?», and the boy answers «Yes». «Don't worry, then, son», says the father, «if she's not good enough for her own family, she's not good enough for ours».

---

[7]   *Pace* Robert Roberts, «Humor and the Virtues,» *Inquiry* 31 (1988): 127-149.

Now, although it may be true that some people are disgusted that anyone laughs at this joke, the source of their disgust can hardly be that the laughter betrays a keeness on incest, nor that it is directed against Cornishmen. We surely wouldn't specially refrain from telling such jokes in the presence of Cornishmen, and we would just regard a Cornishman as touchy if he took offence. Similarly, it's hard to imagine a Jewish American girl taking offence at many of the JAP (Jewish American Princess) jokes. E.g.:

Q.: What is the difference between a Jewish American Princess and Russia?

A.: Russia sucks.

The reason why this would not give offence is that nobody would identify herself as a Jewish American Princess so the category is treated as fictional — even though some real girls probably do satisfy its qualifying characteristics — Jewish, American, always spoilt by her father rather than her mother and not wildly sexy.[8]

A stereotype is simply a bundle of fictional attributions usually related only tenuously to fact. The use of stereotypes becomes dangerous and unacceptable in humor when those involved in the joke-exchange do not recognize that the attributions are false, or are encouraged by the joke to believe that they are true. The stereotypes used in racist and sexist humor are the product of a simple process. A certain group (in European countries, this group tends to consist of white heterosexual males) are marked as the Norm, and stereotypes are created by figuratively taking other groups that are different from the Norm and accentuating and distorting those differences. This is done by treating each «Other» group as homogeneous (all mothers-in-law are the same) and as possessing, to an extreme degree, characteristics conceived not to be present in the Norm group.

Being able to rely on shared knowledge of such stereotypes is useful for the humorist. Quite frequently in humor, these stereotypes incorporate elements of sex, stupidity, dirtiness, cowardice, toilets and bodily discomfort — things which, in our culture and for reasons unknown, are sources of amusement. So, just as we invent characters such as Santa Claus and the Man in the Moon around which to build stories to amuse children, so likewise we have created fictions (e.g. that black men have big penises, that Poles are stupid) to feed our need for laughter. Having the stereotypes spares us the trouble of spelling out joke scenarios at tedious length; the hearer is assumed to be able to fill in the necessary background. The use of these stereotypes may be dangerous when it helps foster false beliefs or bad attitudes — but people susceptible to that kind of influence will generally be those who have difficulty in distinguishing fact from fiction.

---

[8] See Victor Raskin, *Semantic Mechanisms of Humor*. (Dordrecht: Reidel, 1985)

I have encountered a weaker thesis than de Sousa's defended in the following way: Given that you can't joke about things that are sufficiently horrible, when you do joke about something you imply that it's not all that horrible. The thesis is that a laugher, while not cruel or vicious, may be thoughtless and insensitive. With this view I am in only partial agreement. We *do* laugh at torture jokes — e.g. the picture of Christ hanging on the cross, with the cartoonist's bubble coming out of his mouth enclosing the words «What a way to spend Easter». I would claim, that we can joke (with some people) about Christ's crucifixion *without* implying «that it's not all that horrible». And suppose we heard this joke from a child:

The judge says «Attempting to blow up Parliament is a very serious offence, Mr. Fawkes. I'm going to send you down for a long stretch.»

We, and the child might be rather pleased about him understanding the pun, and (I think) we wouldn't interpret his laughter as an endorsement of barbaric forms of punishment. That's very different from joking about torture under General Pinochet. What we can't joke about are things that we are currently very upset about, or take terribly seriously for personal reasons. That is because we cannot detach ourselves from our emotional involvement in them. Now, some people can joke about things that we find upsetting, because although intellectually they see that what is being joked about is horrible, they are not sufficiently involved emotionally to let it worry them. In some circumstances, such people should be condemned for not making our concerns theirs.

I well remember, as a six year old boy, coming home from school and retailing this slogan (based on an advertising campaign for the Gas Board) to my father: «Go gas — six million Jews can't be wrong». He was deeply ashamed that his son thought this amusing, and I now think he was fully justified, even though the joke does have what I have called «redeeming features», viz. the ambiguity of «go» and the implication that the Jews had a consumers' choice in the means of their slaughter. But my father was justified in feeling ashamed not only because I was too insensitive to realize that he couldn't emotionally detach himself from horrifying events that occurred less than a decade before, but also because *I* should have learned enough about the recent persecution of Jews to realize that, for him, it was no laughing matter, and a cause for concern that his son was associating with people who almost certainly were using such jokes to spread anti-semitism. Similarly, one might argue that it is culpable to laugh about (say) blacks knowing that they themselves wouldn't find the jokes funny. If you say that it's all right to laugh behind blacks' backs then, if you are not black, reflect how *you* would like it if you suspected that people, in your absence, were unfavorably caricaturing members of a minority group to which *you* belong. Isn't it *callous* for whites not to be emotionally involved in matters that they know make blacks very upset?

This question needs to be handled with some care. We might invoke what Daniel Dennett calls an intuition pump — an extreme example designed to coax a person's intuitions in a certain direction. Switch from humans to animals. Even those who most ardently respect the rights of animals would not object to the following joke at the expense of cows:

Two cows were grazing in a field. One says to the other «This mad cow's disease, it's terrible isn't it?». And the other replies «It doesn't bother me — I'm a sheep.»

Obviously, a cow could not be *offended* by this; we cannot be offended by something of which we are unaware. But neither is the cow demeaned — nobody would be caused by this joke to regard cows in a less favorable light or to lose respect for them. Now is a member of a minority (say) demeaned by jokes about that minority? Sensible people (such as readers of this journal) would no more come to regard minorities unfavorably as a result of hearing those jokes as they would adopt bad attitudes to cows as a result of hearing cow jokes. This claim is likely to propel philosophers into counterexample mode, but the inclination can be resisted simply by asking yourself: «Has my friendship for any woman or for any minority person *ever* been affected by jokes about these groups?».

Unfortunately, however, not everyone is as sensible as readers of this journal. Suppose, while joking around, that we just have a faint suspicion that one member of the audience is laughing because his feelings of superiority are being stoked by the humor, or that one member is feeling mildly uncomfortable at the jokes. We wouldn't laugh so easily. But, even when we are confident that no such person is present, don't, or shouldn't we feel guilty that our joking must be confined to this clandestine coterie? In principle the answer should be «No», since acts of humor by consenting adults are not subject to moral censure provided they are performed in private and nobody is harmed. But, in practice, we can seldom guarantee that there would be no bad effects of our «only joking». Given how little we know about *why* people laugh at jokes and exactly *what* it is they find funny, we have no right to feel in the least confident that joke-telling about minority groups will not instill, reinforce or legitimize bad attitudes towards these minorities except under very special circumstances (e.g. telling jokes about blacks, purely for scholarly purposes, to an audience which consists exclusively of black sociologists).[9]

Ethnic groups in America are now voicing the kinds of concern adumbrated above, and this has created what Mahadev Apte has called «an American sociocultural dilemma» — a tension between two core cultural values: sense of humor *versus* the recognition of cultural and ethnic pluralism.[10] Evidence shows that hostile or degrading wit is the most popular form of American humor,[11] so that it can be regarded as part of the American way of life to which minority groups should accommodate themselves if they wish to be regarded as thoroughly assimilated. This «social manipulation» view is endorsed by Charles

---

[9]   I borrow this example from Michael Philips, «Racist Acts and Racist Humor,» *Canadian Journal of Philosophy* 14 (1984): 75-96.

[10]   Mahadev Apte, «Ethnic Humor Versus 'Sense of Humor',» *American Behavioral Scientist* 30 (1987): 27-41.

[11]   See Harvey Mindess, «The Panorama of Humor and the Meaning of Life,» *American Behavioral Scientist* 30 (1987): 82-95 on the Antioch Humor Test.

Schutz: «Ethnic humor as public humor serves the larger society by implicitly upholding the dominant standards and way of life» thereby facilitating «most pleasurably the ethnic adjustments and assimilation necessary to the social whole»[12] Apte shows that this position is open to serious question: «With the growing emphasis on cultural pluralism during the last twenty five years and the positive self-image that many ethnic groups began to emphasize, intergroup interactions and attitudes have changed. For one thing, members of various ethnic groups no longer seem to believe that they have to internalize so-called American cultural values that were, by and large, imposed by the dominant white Anglo-Saxon Protestant segment of the population in order to acquire a 'true' American identity.» (pp.32-33)

What is striking, however, about Apte's findings is that an ethnic group's deeming unamusing ethnic humor directed against itself does not translate into that group's taking a stance against the use of ethnic stereotypes *in general*. Indeed, his research seems to show that a minority group readily accepts deprecatory humor directed at *itself*, just so long as the joke-teller is a member of that group. This implies that the humor is regarded as valuable, and only as obnoxious when it is seen to pose a threat or to constitute an attack. In the perceived absence of any such danger members of minority groups seem content to accept the use of stereotypes, and, if my analysis is correct, this is because such stereotypes are regarded as mere fictions which are part of a tradition, just like folk tales and nursery rhymes, or are convenient pegs on which to hang a joke.

The lesson seems to be that we should not be too prim. It may be true, as Philips (*op. cit.*) implies, that philosophers of Polish descent feel badly knowing that their colleagues get pleasure from telling each other innocent jokes that feign to ridicule Poles. But suppose that the jokes were genuinely innocent (none of the jokers believe or could be brought to believe that Poles are stupid, and each knows this about the other) and that philosophers are sufficiently courteous to ensure that their Polish colleagues who mind about such things do *not* know of the traffic in Polish jokes. Would this be so different from my refraining from discussing, in a tutorial group, an example involving death, knowing that a close relative of one member of the group had just died? Outside of such special circumstances, there is nothing wrong in telling jokes about death; we don't thereby demean the dead. The stereotype Polack is a figment; jokes about this abstract entity do not demean any real Polish persons, and we should refrain from telling such jokes only in the company of those who, rightly or wrongly, get upset by them, or in the company of those who really are stupid enough to become (more) bigoted.

A similar point can be made about sexist jokes. The objection might be raised that although the objects of sexist jokes could be male, it is no accident that they tend not to be. It is no accident, so the objection continues, that these jokes tend to be told by men about women, and explaining why it is no accident goes something like this: Such jokes are part of a tradition, a tradition of male thinking about women (as domestic creatures, as sex objects, as less intelligent

---

[12]   Charles Schutz, «The Sociability of Ethnic Jokes,» *Humor* 2 (1989): 165-177. See p.176.

and reasonable than men, say) and telling such jokes often, perhaps usually, serves to legitimate and confirm such assumptions. These assumptions had a crucial rôle in the subjugation of women and continue to prevent women from attaining full equality.[13] The reply is to concede that sexist jokes did indeed serve this function in the past and that, insofar as they continue to do so, we should strenuously attempt to prevent their being so used. But manacles were used in the past to prevent slaves from escaping. We now detest such grotesque forms of maltreatment, and the whole system of which it was a part. Yet manacles can now be used for entirely different purposes (say, as an aid to lovemaking) and, so long as it is clear to the users that the purpose is indeed entirely different, there is no need to feel guilty because of the historical association of these devices with practises we now despise. We should surely not wish to forbid the use of manacles by consenting lovers and would regard as absurd the suggestion that, by using them, they were somehow endorsing slavery.

Suppose that we have made our case that to laugh at racial and sexual humor is not necessarily to embrace racist and sexist attitudes and that the stereotypes employed serve as a convenient foil for the humorous exercise of our linguistic skills. Suppose too that we accept the empirical evidence that, under the right circumstances, the retailing of such humor does not give offence. Should we nevertheless refrain from such humor and train ourselves to be affronted by it because, after all, enjoying this kind of humor is not a particularly valuable value to ourselves, and may be of considerable harm to others unless we are extremely vigilant? This is not a simple problem to resolve. What is needed is a solid defence of the claim that ethnic and sexist humor really does have significant positive value. As we have seen, some of the arguments put forward to date in support of this conclusion have been fairly feeble. I should like to adduce two new considerations.

Within a single day of the disastrous explosion of the space shuttle *Challenger* on January 28, 1986, sick jokes were beginning to circulate, most of them centring on Christa McAuliffe, the sole female member of the crew. (More recent examples centre on David Koresh, the «toast of Texas».) Now, it is not unlikely that, before the death of their daughter, the parents of Christa McAuliffe, like many Americans, found this brand of sick humor funny. What is certain is that they do not find it funny now. And yet others who saw the live TV transmission, and who watched horrified as the spacecraft disintegrated, were able to laugh at the sick jokes a few hours after the event. A plausible explanation of this is that one function of particular kinds of humor is to relieve fear. But sick jokes can no longer serve this purpose for the parents of Christa McAuliffe, since their worst fears have now been realized.[14]

---

[13]   I'm grateful to Andrew Jack for the formulation of this objection.

[14]   In his review, «Reason, Love and Laughter,» *Dialogue* 28 (1989): 499-507 of de Sousa's book, Steven Burns claims, less plausibly I think, that the rôle of such jokes is to help cope with grief. It is very doubtful whether the joke-tellers were *grieving* about the fate of the astronauts, and it is certain that these jokes would not have helped alleviate the grief of the astronauts' close relatives.

Surely a similar account can be given for the appeal of ethnic humor. One cause of (say) whites finding jokes against blacks funny is that there is a deep-rooted fear of losing one's membership in the majority, in a society where minorities are discriminated against. An easy way to silence whites who protest that «affirmative action» programs involve unfair discrimination against white people is to ask whether, in order to reap the benefits of such programs, they themselves would want to be black. A white person forced to consider (perhaps for the first time) what it is like for a black to be black soon becomes aware of the cultural institutionalization of a white aesthetic in which black bodies, including hair, are regarded as deviant and ugly, a culture in which white-looking blacks receive more privileges than a black person with classically African features. Curiously, an analogous riposte is less effective where men are protesting against affirmative action in favor of women. I suspect that the reason for this is that it is far easier for a white to envisage the humiliations and deprivations heaped on members of ethnic minorities than it is for a man to see the world as a woman sees it — to understand what it is like to be a woman.[15] Another reason is that the situation of women is perceived as not being so bad since, after all, women are not a *minority*. Historically it has been minorities — the mentally handicapped, the deformed, the speech defective — who have been the victims of mockery and abuse. Along the Appian way outside Rome one can still see the remains of cave-like apertures where such unfortunates were caged for the amusement of passing travellers. This is commonly explained by pointing out that such handicaps were regarded as divine punishment for sins, so that, by mocking the afflicted , one was endorsing the acts of the gods, thereby insuring oneself against a similar fate. Perhaps today's ethnic humor has its roots in primitive fear and ancient superstition.

If I am right, then humor which makes use of ethnic sterotypes serves the therapeutic purpose of alleviating fear. It might be argued that these fears are unreasonable, yet, if they are *real*, one can hardly condemn people for the psychological barriers they erect against them. This is where we came in, for our original concern was not with whether it was morally right to *tell* racist and sexist jokes, but with whether a person is morally culpable for laughing at such jokes. My answer was «No», and I have now offered a supplementary reason in terms of the therapeutic value of a psychological defence mechanism the output of which is laughter. It is a consequence of my position that as (if) society becomes more integrated, with opportunities equalized across races and sexes, so the attractiveness of this kind of humor will diminish (but not disappear). Whether I am right about this, only time (a very long time, I fear) will tell.

My second consideration in favor of sexist and racist humor is related to the pleasure associated with the exercise of linguistic skills. Suppose we find the sight of a madman funny. How would we convey our pleasure to a third person not lucky enough to witness the mad antics himself? A picture might help, but a

---

[15]   For an intelligent discussion of the possibility of grasping what it is like to be something other than what one is, see Thomas Nagel, *The View from Nowhere* (New York: Oxford University Press, 1986), esp. chap.11. I hasten to add that Nagel, in personal conversation, has resisted extending his argument about the interspecies inaccessibility of 'point of view' in the way that I have suggested.

thousand words would not necessarily, unless they were very well chosen. Certain obvious principles would guide that choice. The words themselves might have a mad sound — «loopy», «schlemiel», «oaf», «dolt», «goof», «nincompoop» — or might engage some apposite semantic connections — «lunatic», «dumbell», «turkey», «blockhead» — or we might make use of similes — «mad as a hatter, as a march hare» — or avoid the pedestrian literal by giving some metaphors a free run — «not playing with a full deck», «lost his marbles», «two sandwiches short of a picnic», «a roo loose in the top paddock» etc.. Some of these associations, particularly the odd but peculiarly apposite ones, are funny. Notice that I'm not saying here that ingenious wordplay makes what would be a sad subject humorous; on the contrary, the playful words are a vehicle for what is antecedently thought to merit just such a form of transport.

Humans (especially young ones) take a primitive delight in conflict and violence, or at least in their graphic depiction. What is especially entertaining about cartoon depictions of the «Tom and Jerry» sort is that in cartoons we can outdo the real world. The cat can drop a thousand feet off a mountain, bounce back up and straight into a mincer, emerging out of the other end dishevelled but good for a further string of catastrophes. And words can outdo pictures not only because we can have verbal descriptions of what is pictorially impossible, but also because the *thoughts* attributed to an individual generally cannot be put into pictures and because the scope for amusing verbal ambiguity is far greater than that for pictorial ambiguity. The force of an explosion can be seen to turn Tom inside out, but when we ask what was the last thing to go through Christa McAuliffe's head (answer: her ass) we are engaging a dimension of humorous representation beyond the four available to visual depiction. Here there is not just the «Tom and Jerry»-type image, but a verbal play in which, to our surprise and amusement, we are sharply brought back to the literal meanings of the words when we were assuming a metaphorical sense — thoughts are not things that literally go through the head. Exactly the same switch occurs in this reflection of Groucho Marx's: «Outside of a dog, man's best friend is a book; inside of a dog, it's too dark to read».

Even pure wordplay is amusing. Puns are funny — there need be no connection with anything beyond the words. A friend of mine, writing a review of a book on Scepticism in which the name of Peter Unger, one of the main players in this particular field, did not appear, entitled his piece «Book Lacking Unger». As puns go, this is not particularly great, and there is no real connection between the philosophical theory and Osborne's drama, yet, *despite lacking any point*, the exercise of verbal ingenuity makes us laugh (or at least chuckle). So, in talking about what we already find amusing — sex, disaster befalling others, modes of speech or behavior with which we are not familiar — it is natural to employ verbal dexterity of the sorts that we have been discussing; the two are made for each other. There thus arises a tradition (or, set of traditions) in which prominent aspects of the human condition are described, embellished and caricatured in the medium of playful language — puns, ambiguities, amphibolies etc.. As these traditions develop, so certain norms and stereotypes become entrenched, though, like most other aspects of social life, the traditions of humor are dynamic.

One highly significant feature of verbal humor is that sophistication in the production of it is acquired at a fairly late age — evidence in fact shows that, at a very early age, the linguistic element of a joke is so insignificant to children that they laugh just as much when the punch line is exchanged for a completion which, by adult lights, is not funny at all. And most of the humor produced by adolescents is, well, adolescent. Verbal agility and the ability to devise new variations on well-worn themes are the mature products of a highly developed sense of language and of immersion in the humorous milieu. A joke which denigrates blacks or Irish or women may appeal to a child because hostility to outside groups is part of the child's world of fun. But, for mature adults, the cleverness of a joke is its most important feature, so a non-clever joke which denigrates some group will seem merely embarrassing — most of the jokes that young children tell us are just that.

Seen in this way, racist and sexist humor become critical tests of wit. If this is correct, such humor is an indication of good rather than evil moral values, for what the promulgation and the appreciation of racist and sexist jokes acknowledges is the adult recognition that mere unadorned prejudice is unworthy, unfunny and unpleasant. For a joke to work in this territory it must have novelty, perhaps a peculiar association of ideas, or vivacious imagery or linguistic subtlety or some combination of these. In other words, it must have the characteristics we associate with sharp humor; anything less will fall flat or fall foul. Similarly with sick humor: think how nauseating a *clumsy* joke about Christa McAuliffe would be. We can grant to De Sousa the very limited thesis that there is something disturbing about adults who find *bad* jokes funny.

What follows from the above line of argument is the paradoxical-sounding conclusion that commerce in racist humor is of positive value in a multicultural setting. While, as a theoretical proposition, the conclusion may seem paradoxical, in practise it is a fairly obvious truth. Those who enjoy a multicultural social circle, while being sensitive to the fact that, in certain cultures, certain types of humor have not got a foothold, are equally aware that racial differences and stereotypes are widely accepted as grist for the mill. Trading insults and telling jokes are ways of enjoying that vitality of language and of life on which we are right to place a high premium. Those who do not enjoy a diverse social circle and are unable to speak at first hand could study the behavior of older children in well integrated, multicultural schools. Although racist humor is rife, it is regarded as part of the rich pattern of school life; it does not impair friendships, nor does it breed hostility. On the contrary, insulting is often a way of cementing friendship, (so long as the insults are «insults»). Not to see this is to confuse mention and use.

Laurence Goldstein

University of Hong Kong, Department of Philosophy

LAURENCE@hkucc.hku.hk

What is Semantics?

A Brief Note on a Huge Question

Newton C. A. da Costa

Otávio Bueno

Jean-Yves Béziau

> Thus mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we say is true.
>
> Bertrand Russell

INTRODUCTION

In what follows, we present, in a rather rough and preliminary way, some general remarks on a quite delicate issue: semantics. To some extent, as will be clear anyway as we proceed, we are here concerned with formulating and spelling out some questions, problems and ideas on this topic, rather than considering their possible solutions. Our basic aim thus consists in just pointing out to some problems that, as far as we see, deserve to be considered and examined — a project, in fact, for a series of works. This explains, or so we hope, the rather concise style adopted throughout the piece.

After some general considerations, made in section 1, we shall briefly present, in section 2, nine thesis on semantics.

Before continuing, however, we wish to add a last introductory remark. It consists in stressing the considerable departure found today between the original sense of this term («semantics») and its current, rather multiple uses. This fact, however, by no means reduces itself to a matter of words. Underlying this meaning variance, it is possible to identify, as far as we can evaluate, a strange shift on the main direction of the semantic analysis of a formal system. As a result, it seems to us, some very important conceptual questions are not correctly spelled out — or not even perceived. Our main purpose now is to call the attention upon them.

1. LOGIC, SEMANTICS, SET THEORY

When first proposed in the fields of logic and formal sciences, the term «semantics» used to present a clear sense. It was supposed to denote that part of an analysis of a language concerned with the determination of the meanings of its (well formed) expressions. (On this regard, see the interesting comments as well as the references presented by Church, in a section dedicated to semantics, at the end of his introduction to the celebrated Church [1946].) More recently, however, faced with an enormous variety of alternative meanings, it is no longer possible to specify an exact sense to this word. Indeed, the process of stretching its meaning has reached such a point that even a *semantic* conception of theories, within the philosophy of science, has recently been advanced!

In intimate connection to this point, our first remark stresses the fact that, as far as the earlier sense of semantics is concerned, Tarski's set theoretical semantics is *not*, in a strict sense, a semantics: it just represents an extensional association between, on the one hand, terms and predicates of a language to, respectively, particular objects and classes of objects of a fixed domain, on the other (this point, indeed, was already noticed by Church himself). By no means the meaning of these terms and predicates is established this way: no intensional factors are taken into account!

More importantly, however, on this regard, is perhaps to note that a set theoretical semantics for a non-classical logic (e.g., relevant or paraconsistent logics) — besides not being, strictly speaking, a semantics —, being constructed within classical set theory, it reveals itself, from a philosophical perspective, completely unsatisfactory. One reintroduces, so to speak, by the backdoors, exactly what was intended to be left on the entrance!

That is the reason why one of the authors (Newton da Costa), when first developed his paraconsistent systems, presented them through a syntactical approach. At that time (1954), not having yet a paraconsistent set theory at his disposal, it would not be possible to articulate a reasonable (set theoretical) semantics for that logic.

(It should be noted, and we shall return to this point later, that in order to have a logic minimally developed, at least *three* conditions must be met: besides the formulation of a propositional calculus, a quantificational theory is to be advanced; furthermore, the same shall be stated for a set theory. Thus, Smiley and some other reputed forerunners of paraconsistent logic, despite the undeniable relevance of their work, have not elaborated, strictly speaking, such a logic: more should had been done.)

More generally, the usual set theoretical semantics, given the way it is articulated at present, depends on its underlying set theory: if one changes such a theory, the semantics itself is, *ipso facto*, changed. In particular, the same is the case for Tarski's definition of truth.

A last word. As some recent researches within valuation theory has shown (see Grana [1990]), every logic admits a two valued semantics. The question then naturally results: in what sense can we talk of a sound *semantics*?

2. SOME THESES ON SEMANTICS

Based on some of the previous remarks (though not only on them), we shall now concisely present nine distinct, but interconnected, thesis on semantics.

1. There is no radical semantics — in the sense of a presuppositionless one. As we have already remarked, Tarski's semantics, for instance, depends on the particular set theoretical setting within which it is formulated. However, there seems to be a kind of «intuitive semantics» underlying our standard semantical constructions. Nevertheless, it is employed just on heuristic, and by no means justificationist, grounds. Its role consists in supplying some theoretical guidelines in order to help us in obtaining our semantical results.

2. As far as we understand, and returning to an earlier point, in order to have strictly speaking a *logic*, one has to present definitions of the notions of *demonstration* and *thesis* adequate to (1) a propositional calculus, (2) a quantificational theory (with identity), and eventually (3) to a set theory. This indeed is an important constraint, given that one of the roles of a logic consists in supplying some tools in order to assist us in the development of conceptual systems. Frequently, however, in various domains, such a development depends on the adoption of particular set theories — and here comes the need of them. It is plain that such a remark is undeniably straightforward as far as scientific contexts, both in formal as well as in empirical domains, are concerned. If one intends to develop mathematics, physics or some further scientific field, set theory, in some or other form, as it is obvious, is probably to be employed. Thus, if your most beloved logic is to be of any use within this process (and it seems fairly reasonable to suppose so, or at least to intend that), then the best you can do is to have it developed up to a set theoretical level.

Given these remarks, we may conclude that, strictly speaking, there might not be a relevant logic. Indeed, at least as far as our current knowldge is concerned, it is not possible to develop a strictly *relevant* set theory. (As is known, it is not even possible to demonstrate, based on this logic, the unicity of the empty set, for such a proof depends on the fact that $A, \neg A \vdash B$.) Furthermore, in connection to this point, and granting that relevant logic is a logic, given the non existence of a relevant set theory, we wonder if, from a philosophical perspective, it is legitimate simply to adopt, as it is usually done, in order to formulate a semantics for this logic, classical set theory, which, as we know, is constructed based on classical, and not on relevant, logic. Unfortunately, this move seems to be rather puzzling, given the relevant theorist's rejection of classical logic.

3. A convenient logical system, as far as contemporary science is concerned, should somehow contain classical logic and its semantics — or, at least a considerable portion of it. Otherwise, on the one hand, some basic scientific applications would not be developed, nor, on the other, some aspects involved in the construction of mathematics will be possible. This point, indeed, was already noticed by Hilbert himself. In fact, though perhaps being a bit hasty in his generalization regarding the role of Aristotelian laws of logic in the construction of mathematics, from his viewpoint:

[...] we cannot relinquish the use either of the principle of excluded middle or of any other law of Aristotelian logic expressed in our axioms, since the construction of analysis is impossible without them. (Hilbert [1927], p. 471).

4. Classical logic, just as relevant logic, is based on a certain kind of semantic atomism: under certain contexts, a particular proposition is true or false independently of any other. (Wittgenstein appears to have adopted such an assumption in his *Tractatus*.) Physics, in this sense, seems to be committed to this kind of atomism. This, anyway, would be a limitation to classical logic.

On this regard, how to apply relevant logic to dialectics, if the latter assimilates everything to a «fluid»? And to the coherence theory of truth? In these cases, are classical semantical construction sufficient?

5. We can imagine a logic in which every proposition depends, for the determination of its truth-value, on the propositions in its neighborhood. So, we would have a «neighborhood» semantics, quite different from the classical one.

6. Given that classical logic is contained within some paraconsistent logic, the former can be employed in order to supply a semantics for the latter. Furthermore, it is possible to construct *paraconsistent* set theories. Thus the argument just presented agains relevant semantics can not be directed to paraconsistent logic.

7. Paraphrasing Einstein's celebrated remark, we can say that the propositions of logic, as far as they are true, do not apply to reality, and as far as they apply to reality, they are not true. This is the case when logic is not conceived in an absolutist way, but as a theory among further theories. There are at least four arguments for such a claim: (1) the plurality of logics; (2) the apparent fuzziness of reality; (3) the fact that formal sciences (in particular, logic) are human constructions; and (4) the opposition between logical rigour and the fuzziness of reality. A good semantics has to cope with all of these issues.

8. How to choose between alternative logical and semantical systems? To put it in a nutshell: through an examination of its consequences. Obviously, some pragmatic aspects are also to be considered here. Being more specific, to some extent, we make our choices based on various considerations that can however be divided into two classes: (1) *formal* requirements, and (2) *material* conditions. Regarding (1), we find the usual formal constraints on a logical system: its soundness and completeness, its relative consistency and so on. (The failing of some of these conditions may, to some extent, present some negative evidence against the system.)

Concerning the material conditions, one may present some criteria of choice roughly based on the following grounds: (2.1) *heuristic aspects* of the system (as far as its deductive power is concerned, for instance), (2.2) its *problem-solving resources*, (2.3) its *adequacy* in order to make sense of some scientific applications and some scientific strategies of reasoning.

9. Someone may present the following question: what is the usefulness of paraconsistent logic and its semantics? To such a question we may reply with a further question: what is the usefulness of classical logic, if we have paraconsistent logic which (at least in some systems) contains classical logic?

3. **CONCLUDING REMARK**

As it might be easily noted, it seems to us that if the preceding theses are nearly correct, Russell's comment on mathematics, presented above, with obvious changes can also be rather naturally applied, in particular, to *pure semantics* itself.

REFERENCES

Church, A. [1946]: *Mathematical Logic.* Princeton: Princeton University Press.

Grana, N. [1990]: *Sulla Teoria delle Valutazioni di N. C. A. da Costa.* Napoli: Liguori Editore.

Hilbert, D. [1927]: «The Foundations of Mathematics», English translation of the original 1927 German paper reprinted in van Heijenoort (ed.) [1967], pp. 464-479.

van Heijenoort, J. (ed.) [1967]: *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931.* Cambridge, Mass.: Harvard University Press.

Newton C. A. da Costa [ncacosta@usp.br]

Otávio Bueno [oasbueno@usp.br]

Jean-Yves Béziau [jyb@alpha.lncc.br]

Department of Philosophy

University of São Paulo

São Paulo-SP, Brazil

A NOTE ON TRUTH, DEFLATIONISM AND IRREALISM

Pierluigi Miraglia[1]

§1.—.

It seems reasonable to require of a naturalistic account of a given region of scientific or ordinary discourse that it construe the reference of expressions central to that region in terms of naturalistically acceptable entities (ultimately, physical objects or states). In many areas of discourse, however, this project stumbles upon notorious difficulties, not all attributable to what appear to be contingent gaps in the current state of scientific knowledge. It is appealing in such cases to regard the utterances and predicates of the given area of discourse as playing a different, non-descriptive role — that of evincing the speaker's stance, for example, or expressing an attitude. The thought is that this move offsets the need to specify physicalist denotata of the predicates in question, for the «point» of uttering judgments in the given area of discourse would not be, in effect, to *denote* anything. Such is the thrust of an *irrealist* approach to an area of discourse. Of course, irrealists immediately face a further question: if the predicates under examination are non-descriptive and non-denoting, can they still play a legitimate role in our conceptual ecosystem, or is it simply a mistake to go on using them? A choice must then be made: should we preserve the discourse in question or should we rather «quine» it, possibly consigning it to extinction? The conservationist option requires, at a minimum, that we explain how discourse involving predicates that fail to denote may still support some standards of cognitive legitimacy, justificatory procedures and so on. This strategy characterizes the conservationist variety of irrealism known as *non-factualism*. The alternative strategy — to acknowledge the mistake intrinsic to non-denoting discourse — is distinctive of an *error-theoretical* approach.

An important issue in irrealism is: how far can one go? Paul Boghossian's «The Status of Content» (1990a; SOC henceforth) lays out a powerful argument against irrealist conceptions of *content*: he believes, in a nutshell, that no irrealist about a certain area of discourse (different from content) can be irrealist about content. Let *irrealism about an area of discourse F* be the doctrine according to which there are no real properties corresponding to (or denoted by) the predicates

---

of F. Irrealism about content consequently holds that such characteristic semantic predicates as «has truth-conditions p» or «means that p» (which apply to sentences or utterances) do not denote real properties. SOC aims at showing that irrealism about content is an «unstable», intrinsically incoherent doctrine. I shall counter this claim by showing that Boghossian's arguments are based on a distorted view of the commitments attendant upon irrealist views. But our scope will of necessity be broader. The interest of the SOC argument resides in its forcing us to confront us some deep and far-reaching issues in metaphysics and semantics. It involves a detailed discussion of a *deflationary* conception of truth, which is alleged to have dire consequences for irrealist views. Now, that issues surrounding truth and semantics are central to irrealist projects in general hardly needs emphasizing:[2] the point urged by Boghossian is therefore a crucial one and, if sustained, would have significant consequences for the very possibility of irrealism. I hold out hope that, by clarifying the relations between irrealist (and specifically *non-factualist*) projects and philosophical views about truth, the present discussion may also serve the broader purpose of deepening our understanding of irrealism.

Now for a preview of things to come: section 2 presents a reconstruction of the central argument of SOC, stated around p. 175; § 3 investigates aspects of the deflationary conception of truth which seem to me essential for the purpose at hand; IV lays out the main objection to the SOC argument. Much of what Boghossian says about content, irrealism and related matters will be granted without argument. I take no issue with his contention that irrealist views about content apply as much to mental content as to linguistic meaning, so I shall assume in what follows that we may restrict our attention to the linguistic case. Similarly, I assume that «the essential core of the ordinary notion of content does consist simply in the idea of a truth condition» (1990a, p. 173). Thus, such content-ascribing locutions as «S has truth condition p» (where S is the name of a sentence), «S is true» and the like will be the focus. The viability of irrealism about more comprehensive notions of content (if any exist) will not be discussed.

Furthermore, I briefly hinted at the fact that irrealist doctrines about any subject matter F come in two varieties: error theories — according to which the predicates of F purport to be genuinely referential but are in fact systematically empty — and non-factualist theories. But my discussion will concern only the latter; nothing substantial is said in this paper about the arguments in SOC that deal with the alleged difficulties of error theories. My concern is with the objections against non-factualist theories of content. In the remainder of this paper, «non-factualism» and «irrealism» are used interchangeably.

§2.—

According to Boghossian, *non-factualism* about a predicate P belonging to a given area of discourse F is characterized by adherence to the following two claims:

(i) the predicate P does not denote a property;

---

[2]   See for example chapters 7 and 8 in S. Blackburn 1984, a standard bearer of contemporary irrealism. Blackburn articulates the irrealist strategy that we shall be most concerned with, namely non-factualism.

> (ii) a declarative sentence containing P (such as the atomic
> sentence «x is P») is not truth-conditional, i.e. does not have a
> truth-condition.

It is a belief in (ii) that sets non-factualism apart from error theories about P, for those are committed to the view that statements containing P are to be evaluated as if they had a «genuine» truth value, although their truth-conditions are never satisfied (i.e. they are always false). Non-factualism about content, then, will be characterized by the following two theses about the content-ascribing predicate «has truth conditions p» (given the modest construal of content, having truth conditions p amounts to having content p):

> (1) The predicate «has truth condition p» does not denote a property;
>
> (2) «S has truth condition p» (where S is an appropriate name for a given sentence) is not truth-conditional.

The central argument against non-factualism about content unfolds in two stages, separately developing the conflicting implications of (1) and (2). More precisely, the trouble with non-factualism is that (1) presupposes a *deflationary* view of truth conditions, while only a non-deflationary, correspondence conception can justify acceptance of (2). We shall consider deflationism more closely in the next section. Let us now examine Boghossian's argument in some detail.[3]

> (A) If «true» does not refer to a property, then any declarative
> sentence is (trivially) truth-conditional.

That «true» does not refer to a property is, for Boghossian, the central tenet of deflationism. The important point is that

> there is no more to a sentence's being truth-conditional — genuinely apt for (deflationary) truth or falsity — than its being a significant sentence possessing the appropriate syntactic potentialities (SOC, p. 164).

What Boghossian seems to have in mind is a disquotational or «homophonic» construal of (a sentence's having) truth conditions. On a disquotational reading the truth condition for any given declarative sentence is that expressed by the sentence itself. In effect, nothing more is needed in order for S to be able to express a truth condition than just that S be assertible at all. S must be merely a candidate for assertion in order to be truth-conditional. From a deflationist standpoint the qualifying requisite for such candidacy must be, according to Boghossian, extremely weak: just being «meaningful» (a syntactic property) and «declarative» should suffice. Hence:

---

[3]   The reader will notice that Boghossian takes a deflationary conception of truth to be defined by the thesis that the predicate «true» does not refer to a property. I shall for the time being respect this use, although I think that deflationism is more perspicuosly stated in a slightly different manner (see the definition [DEFL] below). A further terminological caveat: one often sees the dispute between deflationary and non-deflationary conceptions of truth characterized in terms of «robustness», a notion which tries hard to legitimate a certain metaphysical picture of the dispute. In these terms, the opposite of a deflationary notion of truth would be a «robust» notion of truth. I find the picture misleading, so I avoid this terminology. The proper opponent of deflationary truth is not robust truth, but correspondence truth (more on this in sect. III). But if I am right, the latter is no more robust than the former.

> Any meaningful declarative sentence would be (at a minimum) a *candidate* for assertion... Any such sentence would count, therefore, as truth-conditional in a deflationary sense (SOC, p. 165).

(B) By contraposition of (A), (2) implies that «true» does refer to a property. Were the non-factualist to adopt a deflationist reading of locutions such as «expresses a truth condition», she would have no room to *deny* that any meaningful declarative sentence is truth-conditional — but that is precisely what non-factualists deny.

The idea is to show that, while (2) rules out deflationism for the reasons just given, (1) entails it. The case is condensed in the following passage:

> For the truth value of a sentence is fully determined by its truth condition and the relevant worldly facts. There is no way, then, that a sentence's possessing a truth value could be a thoroughly factual matter («true» does express a property) if there is non-factuality in one of its determinants («has truth condition p» does not express a property). (1990a, p. 175)

This can be reconstructed as follows:

(C) The predicate «true» denotes a property if and only if «a sentence's possessing a truth value» is a «thoroughly factual matter».

Strictly speaking, the argument requires only the left-to-right direction. *Prima facie*, though, the right-to-left direction fails on a disquotational notion of truth: I shall argue in the next section that a disquotationalist should maintain that a sentence's truth value can be a perfectly factual matter without conceding that «true» denotes a property. On the other hand, if the deflationist envisioned by Boghossian adheres to a view of truth as some kind of pragmatic «virtue» of statements, (C) might well be in order: for this kind of deflationist, the truth of an assertion is indeed a thoroughly non-factual matter. One senses here the need to take a closer look at deflationism. At any rate, since the equivalence is not necessary to the SOC argument, these are for the time being definitional quibbles.

(D) (1), the claim that the predicate «has truth condition p» does not denote a property, entails that a sentence's possessing a truth value is not a thoroughly factual matter.

(E) Therefore (by the left-to-right direction of (C)), (1) entails that «true» does not refer to a property, i.e. a deflationary conception of truth. By (B), we obtain a contradiction. Hence, irrealism (i.e. non factualism) about content is inconsistent.

§3.—

If standard usage is to offer any guidance in semantics, it appears that the predicate «true» (or «is true») as applied to sentences in the indicative mood satisfies equivalences of this form:

(T) «Snow is white» is true if and only if snow is white.

It is routinely presumed, after Tarski, that a theory of truth for a given language L should be able to derive all equivalences of this form, one for each declarative sentence of L. In other words, it is an adequacy constraint on the definition of a truth predicate for L that the schema resulting by writing a sentential variable in place of the sentence «snow is white» in (T) above be validated by such a predicate. Tarski's Convention T is of course such a schema. This being the case, it might be tempting to conceive of Convention T as telling

us in effect *all* we need to know (and all we can expect to know) about the semantic role of the predicate «true»: this conception I call *deflationism* about truth.[4]

> (DEFL) for any (declarative) sentence S, the assertoric content of «S is true» is the same as the content of S; i.e., to say that S is true is to say no more and no less than what is expressed by asserting S.[5]

This thesis puts the deflationist at variance with what is usually labeled as the «correspondence theory» of truth, according to which to say of a sentence S that it is true amounts to saying that there is a special relation of «correspondence» between S and some parcel of ontology — a state of affairs, a fact, a combination of objects, etc. This conception then opens a gap between what I have called the assertoric content of «S is true» and the content of S itself, for on the face of it the latter contains no reference to «correspondence». One might retort perhaps that to say of S that it corresponds in the appropriate way with the facts is *just* to say that S (putting aside the fact that such a claim runs counter to the intuitive judgement of most speakers), i.e. that correspondence is itself a «disappearing» or redundant property. This might have been Ramsey's view when he said:

> We can, if we like, say that [the proposition aRb] is true if there exists a corresponding fact that a has R to b, but this is essentially not an analysis but a periphrasis, for «The fact that a has R to b exists» is no different from «a has R to b». (1927, p. 39)

---

[4]   There is also the temptation to read this statement as the deflationist's *definition* of truth. One might be well-advised to resist such a temptation. Anil Gupta has recently challenged this basic deflationary idea: in his view, it is simply false that sentences like (T) provide a definition of truth, in any plausible sense (although they certainly are an important consequence of a proper definition of truth). Furthermore, there is the «inflationary argument» advanced by Crispin Wright 1992, ch. 1, which purports to demonstrate the incoherence of a similar version of deflationism. However, I shall not discuss Gupta's ideas, which rest upon a different set of considerations from those central to this paper. I plan to discuss Wright's argument elsewhere («A brief against the inflationary argument», in preparation), but it seems to me that the outcome of that dispute is independent from the treatment presented in this paper: here my primary concern is to show the compatibility of what Wright would call «minimalism» (which is at least coherent, in his view) and non-factualism about content.

[5]   The assertoric content of a sentence (roughly, the meaning conveyed by an assertive utterance of that sentence) must be understood to exclude certain pragmatic factors that make, in particular contexts, an assertion of, e.g., «That is true» not quite the same as an assertion of whatever utterance is referred to by «That» (a similar problem may arise for utterances containing indexicals). Such factors are discussed and dealt with by D. Grover, J.L. Camp and N. Belnap 1975, see especially pp. 79 and ff.

But it seems fair to say that a correspondence theory couched up along these lines is no longer a conception of truth *as* correspondence: it is just the old sheep in wolf's clothing, that is, the deflationary conception repackaged.

My critique of the SOC argument depends on a number of more or less elementary facts about deflationism, which I now proceed to explain. A vindication of the deflationary conception of truth, worthwhile as it is, is clearly beyond the scope of this paper. Yet certain basic aspects of the deflationary conception are made very short shrift of in Boghossian's SOC, as well as in several of the critical papers it generated. By taking such aspects into account, we shall have made some progress toward a clearer understanding of the issue.

Firstly: The most straightforward embodiment of the deflationary conception is the so-called disquotationalist theory of truth; I take disquotational truth to be the paradigmatic example of a deflationary notion of truth, and I shall use disquotationalism to show that irrealism about content is ultimately immune to Boghossian's objections. This is in contrast with most other critics of Boghossian's views, who have generally conceded his point about disquotationalism even while disputing his assessment of deflationism. According to the disquotationalist theory, «S is true» is equivalent to a (typically infinite) conjunction of the form

> if S is «snow is white», then snow is white; if S is «the cat is on the mat», then the cat is on the mat; ...

(where, as usual, S stands for any declarative sentence). Thus, disquotationalism about truth involves in effect little more than strict adherence to DEFL.[6] By contrast, another view of the notion of truth, also classified by many as deflationary, offers a quite different outlook. According to a *pragmatist* conception of truth, truth is a special compliment paid to a select class of sentences — perhaps those that we are prepared to «defend against all comers», or those that we deem «explanatorily indispensable», and so on.[7] Such a conception is clearly incompatible with the correspondence theory. But is it a deflationary conception? If I *say* of «snow is white» that it deserves to be treated by me with the respect I reserve for beliefs that are indispensable to my world picture (which is roughly what I would imply, according to the pragmatist, by asserting that «snow is white» is true), I do not seem to be saying just that snow is white. On the other hand, a careful pragmatist could reply that the compliment paid to the sentences in the select class does not add to the *content* of these sentences, so the contents of the two relevant assertion — that S and that S is true — remains the same. At any rate, I mentioned a pragmatist conception of the

---

[6]   See H. Field 1986, p. 58. The prosentential theory of truth (see Grover, Camp, Belnap, op. cit.) seems to be a particularly sophisticated version of what I (following Field) have called «disquotationalism» here.

[7]   The first formulation of the compliment is by R. Rorty, *Philosophy and the Mirror of Nature* (Oxford: Blackwell, 1980), p. 308; the second I owe to R. Kraut, and is articulated throughout his (1993). Obviously, my sketchy account of coherentism and pragmatism about truth has no pretense to do justice to either doctrine.

truth predicate to evince a certain contrast that might emerge in relation to my next point.

The point is, *secondly*, that a deflationist about truth may be an irrealist with respect to (the predicates of) a given area of discourse, but she *need not* be one. Deflationism is perfectly compatible with realism. The disquotationalist holds that to assert that «snow is white» is true is the same as to assert that snow is white. But must she be an irrealist about the color of snow? Trivially, no. I conjecture that the reason why this fact may on occasion appear less than completely trivial is that it tends to be obscured when the deflationist view one has in mind is a pragmatist one. Suppose someone asserts that it is true that there are electrons in my kitchen. If I say, along with a certain kind of pragmatist, that the speaker is thereby expressing his conviction that the sentence asserted is explanatorily indispensable, I may easily slip into making the *existence* of electrons in my kitchen a matter of explanatory expediency — and this would seem to be close to an irrealist conception *of electrons*. Thus, if this kind of pragmatist conception of truth has intuitive appeal, it often is on the score of a more general anti-realist project which recommends it. Whatever the matter with pragmatism, however, deflationism in general is independent from irrealism (or realism). In the same case, a disquotationalist would have no need to finesse her stance with respect to the existence of electrons in my kitchen; she simply asserts the *identity* of such a fact with the fact that «there are electrons in my kitchen» is true — assuming, of course, that there is such a fact. In other words, the disquotationalist need not import any *sui generis* conception of the truthmakers of such a sentence. Thus, her account of truth is, to a large extent, metaphysically flexible. How flexible, of course, is precisely the question addressed by Boghossian's paper: an essential part of the argument there is that deflationism is incompatible with *mixed* accounts (irrealism about some areas of discourse but not others), because drawing the boundary line between factual and non-factual areas requires a non-deflationary conception of truth.

Thirdly, and lastly: The deflationary conception faces no serious or special difficulty in explaining the concept of truth-conditions of a sentence. To put things in terms of Tarski's theory of truth: the truth-conditions of a sentence in the first sense are simply given by the respective T-sentence, a biconditional having the same form as (T) above. The truth-conditions of the sentence named on the lefthand side of the biconditional are just the righthand side of the biconditional. So, the truth-conditions of «snow is white», as well as «la neve è bianca», are *that* snow is white. That, of course, is precisely what the deflationist (disquotationalist) theory will predict. And yet there is a sense that some important element is missing from deflationary truth-conditions, something that «real» truth-conditions should have. I shall proceed here with a bit more caution.

A brief reflection on Tarski's contribution may help here. Tarski notoriously claims, in the opening paragraphs of his (1935), to be concerned exclusively with the «classical» notion of truth, i.e. (in his words) truth as correspondence. But the condition for the «material adequacy» of a definition of truth he stipulates (i.e., Convention T) merely prescribes that the theory entail all

T-sentences;[8] truth as defined by the disquotationalist would satisfy Tarski's criterion. This should be no surprise, since what the disquotationalist does is, in effect, to define truth with the help of a homophonic translation of the object language into the metalanguage, and that is pretty much what Tarski did.[9] Still, one might object at this point that the analysis of truth conditions in the Tarskian approach is not exhausted by mere consideration of the T-sentences; it also matters how the T-sentences are *derived* from the theory, and the base axioms used in such derivation will be (in compliance with compositional requirements) statements about the denotation of terms and predicates. Can the deflationary conception accommodate such denotation axioms? Isn»t denotation «intrinsically» a correspondence notion? Well, the theory of denotation that can be extracted from Tarski's work is something like the following (in the case of, e.g., proper names; other cases are defined accordingly):

> To say that the name N denotes a given object *a* is the same as to stipulate that either *a* is France and N is «France», or ... or *a* is Germany and N is «Germany» (Field (1972), p. 365).

In the same passage Field continues:

> This is Tarski's account of denotations for English proper names...[Such theories of denotation as the above] satisfy criteria of adequacy exactly analogous to the criteria of adequacy that Tarski accepted for theories of truth.

Now, a theory of denotation couched in these terms is a theory that will be perfectly congenial to a deflationist: we could call such a theory a disquotational or homophonic theory of *denotation*. We may raise all sorts of grievances against such a theory. For example, we may wonder, as Field does, «what a real explication of denotation in nonsemantic terms would be like» (ibid.), in which case we should look at a different theory to supply that. Nevertheless, the point is clear: to the extent that talk of truth conditions is cast in terms of a Tarskian truth theory (i.e. as talk about the righthand side of the T-sentences), truth-conditions pose no special threat to the deflationist. But now a problem arises.

Talk of truth-conditions is often perceived as ambiguous: on the one hand, we seem to have in mind the righthand side of Tarskian biconditionals — that snow is white; on the other hand, many seem to expect and demand something more «robust», or at least as robust as real snow. Yet attempts to clarify the

---

[8]   Convention T requires that an «adequate definition of truth» have as consequences:

> all sentences which are obtained from the expression «x ∈ Tr if and only if p» by substituting for the symbol «x» a structural-descriptive name of any sentence of the language in question and for the symbol «p» the expression which forms the translation of this sentence into the metalanguage;...(A. Tarski 1935, p. 188)

Clearly, the disquotational definition does entail all such sentences, i.e. sentences like «'snow is white' is true iff snow is white», for it results in fact from an infinite conjunction of them.

[9]   For a thorough examination of this aspect of Tarski's theory, see H. Field 1972, especially pp. 354 and ff. (To be sure, this is not quite to say that the notion that Tarski defined was actually disquotational truth; the point is rather that Tarski's contributions to the theory of truth are compatible with the deflationary conception, and are consequently available to the deflationist. For a review of some technical and philosophical differences between the disquotationalist and Tarski's notion of truth, see Field 1986, cit., pp. 64-65 and footnotes.)

perceived ambiguity and to strengthen *the expression of* truth-conditions in one way or another tend to be vacuous.[10] What could the robust truth-conditions of «snow is white» be like? Perhaps *the fact that* snow is white? Certainly we could (and should) say that the truth-conditions of «snow is white» are the fact that snow is white, but this would in no way mark a difference between us and a deflationist: the described fact is precisely that expressed by the righthand side of the relevant biconditional. Could we then protest, with M. Devitt, that on the deflationary picture of truth-conditions truth or falsity «do not apply to sentences partly in virtue of contingent properties of sentences determined by facts about language» (1990, 254)? This will not take us very far: it may well be in virtue of a causal connection (a primeval baptism, or whatnot) between the word «snow» and snow that the truth-conditions of «snow is white» are what they are, but the baptism itself is not the truth-conditions of the sentence, although it explains why the truth-conditions are in fact that snow is white. But, more generally, it is not as though talk of baptism, causal connections or other «contingent properties» will scare off the deflationist: it is not part of the deflationary conception that sentences acquire their truth-conditions by magic or by fiat. Well, if all this is right it will be difficult to embarrass the deflationist by producing something (the «robust» truth-conditions) which she *cannot* express or refer to. But we still have the opposite problem to deal with: she may trivialize the notion by being too prodigal, by expressing *too many* truth-conditions. In fact, this is the charge levelled by Boghossian: that a deflationist will not be in a position to deny that any sentence is truth-conditional.

§4.—

   Thus, the serious problem turns out to be this: a deflationary conception of truth precludes an understanding of what it is for a sentence to be *truth-conditional*, because a deflationist cannot make sense of a sentence's *failing* to be such. For a deflationist must classify any sentence as truth-conditional (see steps (A) and (B) in our reconstruction of SOC). But we will se that, whatever the merit of this claim, the problem for the deflationist is much overblown. My response to the allegation will be that, to the extent that the deflationist has trouble with the notion of truth-conditionality in question, so does everybody else; and to the extent that a correspondence theorist can make sense of the claim that some sentences are *not* truth-conditional, so can the deflationist. A consequence of this thesis will be that claims about what sentences are not truth-conditional have little, if anything, to do with one's conception of truth. This is a mildly surprising conclusion, because it has been customary in irrealist quarters to label non-factual discourse as «non-truth-conditional». Yet I submit that what matters is not the label, but the features of sentences and utterances that the label is meant to designate — and I hold that these features have in fact little to do with a

―――――――――

[10]   A similar point is very carefully worked out in ch. 2 and 3 of Wright (1992), cit. The question there is whether one or the other «platitude» about truth — the kind of pronouncement which is typically intended to mark the robustness of truth — proves impervious to the minimalist (the Wrightean deflationist), and can thus be used to demarcate minimalist truth from robust conceptions. Wright shows that no such platitude emerges.

particular conception of truth. So, non-factualists can be accused of loose talk, but not — if I am right — of incoherence.

What do people mean when they declare a sentence or utterance to be truth-conditional? Or, perhaps better: what *can* they mean? The safe, if trivial, answer is of course: that the sentence in question has truth conditions. By which in turn it is meant, without straying too far from triviality, that the sentence expresses a proposition that at least sometimes is true, and thus occupies the right side of the biconditional in the relevant T-sentence.

So far, so good for the deflationist: we have not exceeded the resources available under the deflationary conception. As noted in §3, this notion of truth conditions is perfectly compatible with deflationism. But the notion of truth conditions so easily wrung out of Tarskian biconditionals is also quite weak, or so it seems. It applies to any sentence for which the theory of truth is to yield a corresponding T-sentence. Here lies the problem, or so the criticism goes. For we may want to *deny* that a sentence is truth-conditional even though that sentence «has truth conditions» in the weak sense palatable to the deflationist. It must then be the case that such denials, if we are to make sense of them, involve a *different* notion of truth conditions and being truth-conditional. Thus Boghossian:

> declarative sentences cannot fail to possess truth conditions except against the background of a robust [i.e. correspondence] notion of truth (1990, p. 166).

I agree that there is a non-trivial notion of being truth-conditional at play here, but I deny that a correspondence theorist is any better equipped to capture it than a deflationist. The non-trivial notion that we may want to capture is that of a sentence or a class of sentences being *fact-stating*, or part of fact-stating discourse. We may suspect (correctly, I believe) that the utterances of certain areas of discourse perform a different function than that of asserting that a given state of affairs obtains; such utterances would not be fact-stating. As we saw previously, to hold such a view of the sentences belonging to a given region of discourse is precisely to be an irrealist (and a non-factualist) about that region. Now, it seems to me essential that we understand what is going on when someone takes a non-factualist stance with respect to a certain area of discourse, that is, *what* is being asserted by declaring that some sentences are not truth-conditional.

When non-factualists and irrealists in general claim that sentences belonging to a given area of discourse are not truth-conditional, they seem to mean, first, that no *fact* (state of affairs, etc.) answers to the sentences of that area. This strikes me as a characteristically metaphysical thesis, i.e. one that has to do with ontology, with objects and combination of objects and their presence or absence in the world we inhabit. There is, however, a further step that one needs to have taken before being in a position to declare a sentence *non*-truth-conditional. For the simple absence of a fact to make a sentence true might just be taken to indicate that that sentence is *false*. Now, false sentences are seemingly truth-conditional. So the non-factualist claim about the non-truth-conditionality of the area of discourse in question must contain a further element: the non-factualist must mean that the communicative *function* of sentences in that area of discourse is radically different from that of sentences that are to be evaluated for truth or

falsity. This is more of a semantic claim,[11] having to do with the proper way to interpret certain sentences or utterances: it is a claim, for example, about people *do* when they utter judgments of a certain type. Notice that what is proprietary to non-factualists is the second (semantic) claim, for the first (metaphysical) claim they share with the error theorists. The error theorists (about an area of discourse F) also hold that there are no things answering to the predicates of F, but from this realization they conclude, roughly, that everything said in F is false. Non-factualists want to escape precisely that conclusion. So they *need* to make both points, if non-factualism is to be at all viable.

A question then naturally arises as to the precise connection between the ontological claim and the semantic claim. The question could be stated this way: how does the thesis about what facts or objects there are translate into a view about the proper applicability of truth and falsity to an area of discourse? Or if you prefer: does the ontological non-factualist thesis *entail* the semantic non-factualist thesis? If it did, then the realization that the world does not contain facts about F, say, would be sufficient to conclude that F-discourse is non-factual, and therefore not truth-conditional. Consequently: *If* the correspondence theorist had a better grip on the ontology (i.e. the world) than the deflationist, as those dissatisfied by deflationism seem to suggest, then he (the correspondence theorist) could at least make clear sense of a F-sentence's not being truth-conditional: the things or properties asserted by F to exist are not there. But if the ontological thesis does not imply the semantic thesis, then *even if* the correspondence theorist has a better grip on the ontology, he has no significant advantage over a non-correspondence theorist — simply because the ontology in his grip is by itself *inconclusive as to the semantic status* (truth-conditionality) of F-sentences. I argue that, at least prima facie, there is no direct implication between the ontological claim and the semantic claim.

One may well protest the vagueness of such expressions as «having a better grip on ontology», by which I have characterized the correspondence theorist. I am sympathetic: I am not sure what it means. But the burden of clarification falls on the correspondence theorist. The strategy is this: the fans of correspondence complain that deflationary truth really does not establish some needed connection discourse-ontology. Let us grant that, and let us grant that a correspondence theorist will have privileged, unfettered access to such connection; even though we have really no idea of what connection we are talking about here, this is the kind of things one hears from the fans of correspondence, and this we concede: we are stacking the deck in favor of correspondence. Then let us show that even under such assumption the correspondence theory will offer no significant advantage over the competing deflationary notion.

I ask the reader to imagine a correspondence theorist as singularly well-equipped to discover the real properties and objects to which predicates and terms refer in the given language. He is an ontological detector: if there is a property answering to a certain predicate, this theorist will see that there is, and

---

[11]   I use the attribute «semantic» here in the broad sense of pertaining to interpretation. This may not entirely correspond to the narrow sense familiar from contemporary philosophy of language, where «semantics» is almost synonymous with truth theory.

viceversa.[12] The question is whether this peculiar ability provides him with a criterion to discriminate factual (or «truth-conditional») from non-factual discourse. Now let us imagine this theorist confronted with the realization that there is no property of wrongness and therefore no fact about, say, the wrongness of telling lies. In accordance with his view of truth, he would presumably regard such a fact as the truth conditions of «lying is wrong». But he cannot take the lack of fact about the wrongness of lying as a direct indication that a sentence like «lying is wrong» is not truth-conditional. Why not say, more economically (at least from a semantic point of view), that the sentence is plainly false? Perhaps our correspondence theorist might reply that the situation is peculiar in that the lack of a fact making «lying is wrong» true does not seem to translate into a fact making «lying is not wrong» true, as we might expect for sentences and areas of discourse of the ordinary sort. But this is no good reason to relegate sentences about the wrongness of lying in a class by themselves, since the same peculiar situation obtains for very many sentences about, e.g., the present king of France (or perhaps the future king of England).[13] But those sentences surely seem as truth-conditional as any. Nor would it matter very much to establish that the truth conditions of (i.e. the fact that) «lying is wrong» *can never* be realized (assuming this could be established), since that is also the plight of a lot of necessarily false sentences (such as «water is not H_2O»). But impossible sentences also are truth-conditional. Anyway, we have supposed that there is no object in the world around us with the property of wrongness. But there is also no object with the property of being a winged horse, nor could there be. By assumption, the correspondence theorist is aware of this situation; but then how will he be able to determine that utterances concerning winged horses are to be counted as false, while utterances about the wrongness of lying are to be put in the non-factual limbo? What *fact* about the world could justify this disparity of treatment?

---

[12]   A word of caution on the way «ontology» is used here: I am assuming a somewhat parsimonious, naturalistic world, one in which there are no Platonic universals hanging around with no instances. For some, this may be unduly restrictive, but I am prepared to face the risk.

[13]   According to the Russellian parsing of sentences such as «the king of France is bald», their English negation («the king of France is not bald») is ambiguous and generates two readings. If we take the negation as having wide or sentential scope, as in «It is not the case that the king of France is bald», then the result is true if there is no king of France. Is this an objection to the analogy I draw with the moral judgment «lying is wrong»? I am not sure. In the situation described, in which wrongness cannot be found, I would imagine that one may sensibly state that it is not the case that lying is wrong and that it is not the case that lying is right, either; but notice that the same would hold for the king of France sentences. It seems to me, however, that wide-scope negation in natural language is not all that clear (in fact, it is not even clear that there is such an operator). «It is the case that S» in ordinary language is almost synonymous with «It is true that S» — but then we are back to the starting point, for it certainly would be correct, in the situation envisioned in the text (no wrongness to be found), to assert that it is not true that lying is wrong. In fact, a correspondence theorist would be *committed* to asserting this.

The moral is clear enough, I think: the mere lack of a suitable fact or «truthmaker» for a sentence is not sufficient evidence for us to determine that the sentence is not truth-conditional *in the sense required by non-factualist claims*. But notice that we have worked from within the constraints of a correspondence conception of truth and truth-conditions. Thus we come to this realization: the special liaison between facts and sentences, to which the correspondence theorist claims privileged access, cannot be the criterion by which we single out truth-conditional or fact-stating discourse. That criterion must advert to issues of use, interpretation, conventions, etc. — to the function of an area of discourse in communication. But these issues do *not* require for their clarification a correspondence conception of truth. For that matter, they do not impose a deflationary conception either: the point is that the crux of this matter relatively independent from any particular conception of truth.[14]

Now one might still reply that, while the ontological component of the non-factualist thesis (the claim that there are no facts of a certain type) is not sufficient to ground the semantic component, yet it is necessary: and the correspondence theorist has the resources to make at least sense of the ontological thesis, while the deflationist lacks even those. This is in effect a version of the misconception we discussed in III, according to which a deflationary conception issues into generalized anti-realism, so that a grasp of facts in the world (of ontology) is quite out of its reach. But we have shown this prejudice to be mistaken. The deflationist (i.e., the disquotationalist about truth) can very well make sense of the absence or presence of certain facts. She can admit of some and reject others, so that her conception of *facts* is certainly non-trivial. Given this, she has as much of a right to the ontological thesis as the correspondence theorist does. Hence she can make as much sense of a non-factualist claim that some sentences are not truth-conditional as the correspondence theorist can, for the ontological thesis that there are no facts answering to those sentences is the *only* component of such a claim that can ever presuppose a correspondence conception.

The upshot is this, I believe: There is a certain ambiguity in the concept of truth-conditionality, as it applies to sentences or utterances (and perhaps in the very concept of truth conditions). There is a somewhat weak notion of being truth-conditional, which is associated with Tarskian biconditionals. Tarskian biconditionals are deflationary «in spirit», so to speak, but Boghossian is right in claiming that, on this notion, almost any sentence would be truth-conditional. So

---

[14]   A referee has pointed out to me that my argument here may be in concflict with the analysis of Jackson, Oppy and Smith 1994. (Unfortunately, I learned of their excellent paper too late to take it into account.) The conflict may stem from this: I may be seen here as suggesting a certain «minimalism» about truth-conditionality (what they call, much more felicitously, «truth-aptness»), while they claim that even a minimalist about truth (such as a disquotationalist might be) should not be confused into believing that truth-*aptness* itself can be construed minimally. But I find myself in agreement with this claim, and I think that it in fact goes in the direction I am gesturing toward: I do believe that truth-aptness is an important, non-minimal property of discourse; I do not believe that this property hinges upon a certain conception of truth. Obviously, much work here is left to be done.

this cannot be what people mean when they deny that sentences of a certain region of discourse are not truth-conditional. So, either non-factualism makes no sense as it is, or there also is a non-trivial reading of claims of truth-conditionality. But the non-trivial notion of truth-conditionality appealed to by non-factualist claims turns out, on closer inspection, to portend (i) a straightforward ontological thesis and (ii) a thesis about the role of certain sentences in communicative discourse. I have argued in this section that (i) does not imply (ii), and that, since (i) does not imply (ii), the correspondence theorist is in no better shape than the deflationist to account for the non-trivial notion of truth-conditionality. In effect, we might conclude that either non-factualism makes no sense at all (a stronger claim, which even Boghossian seems unwilling to underwrite), or the distinction between factual and non-factual discourse has less to do with a conception of truth than is often presumed.

Let us show now that Boghossian's argument is neutralized. To see this, let TC stand for the predicate «is truth-conditional». Thus, the expression TC(S) «says» that the sentence named by S is truth-conditional. Let $TC(Stc_p)$ be the sentence TC(«S has truth conditions p»). According to step (A) in the SOC argument, deflationism entails TC(S) for any declarative S; so in particular $TC(Stc_p)$. But the non-factualist holds (2), which is the negation of $TC(Stc_p)$. Therefore, (2) implies that deflationism is false. This is the SOC argument in a nutshell. The considerations developed in this and the previous section, however, show that there are two *distinct* predicates (or senses) of truth-conditionality available to the non-factualist: there is the «trivial» version borrowed from Tarski and the non-trivial one involving the «ontological thesis».

Let TC and TC* represent, respectively, the trivial and the non-trivial predicate, and return to the argument. On Boghossian's own account (2) involves a non-trivial notion of being truth-conditional: in other words, (2) is *not* the negation of $TC(Stc_p)$, but rather of $TC*(Stc_p)$. The negation of $TC*(Stc_p)$, however, does *not* contradict $TC(Stc_p)$, short of equivocation.[15] On the other hand, deflationism does *not* entail the non-trivial thesis, i.e. TC*(«S has truth conditions p»), for the reasons explained. So on the non-trivial interpretation step (A) is not a valid inference. Therefore we have two possible cases: Either (i) step (A) is valid, but its consequence is TC(S) and successively $TC(Stc_p)$; in which case there is no contradiction with (2), which denies $TC*(Stc_p)$. Or (ii) we take the conclusion of step (A) to be $TC*(Stc_p)$, thereby producing a contradiction with (2); in which case, though, step (A) is invalid, for deflationism does not entail that any sentence is truth-conditional in the non-trivial, ontologically significant sense.

The significance of these results may perhaps be better appreciated by a brief comparison with some of the critical literature generated by SOC. M. Devitt and G. Rey have generally conceded one of Boghossian's crucial points; they agree with him that «the idea that a sentence lacks truth conditions presupposes

---

[15]   My recognition of this point was greatly aided by N. Tennant's detailed parsing of Boghossian's argument in his (unpublished manuscript).

that the sense of truth in question is robust, not deflationary».[16] But I have shown in this section that «the idea that a sentence lacks truth conditions» either presupposes little about truth, or else is perfectly intelligible on a deflationary notion. On the other hand, Robert Kraut (1993) attacks the thesis that for a deflationist any sentence is trivially truth-conditional, i.e. step (A) in my reconstruction (notice that Devitt and Rey concede this, too). Thus his rejoinder to Boghossian's attack is parallel to that formulated in this paper. In order to provide a counterexample, however, Kraut adopts a pragmatist, not a disquotationalist, conception of truth — truth, on this view, is a special «compliment» paid to select utterances. Kraut is effective in characterizing a pragmatist conception immune to Boghossian's charge, but he admits that Boghossian's point is valid against disquotationalism and other «promiscuous» varieties of deflationism. If I am correct, however, either Boghossian's point about disquotationalism is incorrect, or else it does not mean what Boghossian takes it to mean. At any rate, I disagree with Kraut on the issue of promiscuity: the «non-promiscuity» that is necessary to formulate non-trivial irrealist and non-factualist theses about given areas of discourse has little to do, if I am correct, with truth itself.

REFERENCES

S. Blackburn 1984: *Spreading the Word* (Oxford: Clarendon Press).

P. Boghossian 1990a: «The Status of Content», *The Philosophical Review* XCIX, 159-184.

P. Boghossian 1990b: «The Status of Content Revisited», *Pacific Philosophical Quarterly* 71, 264-278.

M. Devitt 1990: «Transcendentalism about Content», *Pacific Philosophical Quarterly* 71, 247-263.

M. Devitt and G. Rey 1991: «Transcending Transcendentalism: a Response to Boghossian», *Pacific Philosophical Quarterly* 72, 87-100.

H. Field 1972: «Tarski's Theory of Truth», *The Journal of Philosophy* LXIX, 347-375.

H. Field 1986: «The Deflationary Conception of Truth», in *Fact, Science and Morality*, ed. by McDonald and Wright (Oxford: Blackwell). 55-117.

D. Grover, J.L. Camp, N. Belnap 1975: «A Prosentential Theory of Truth», *Philosophical Studies* 27, 73-125.

F. Jackson, G. Oppy, M. Smith 1994: «Minimalism and Truth Aptness», *Mind* 103, 287-302.

R. Kraut 1993: «Robust Deflationism», *The Philosophical Review* CII.

---

[16]   M. Devitt and G. Rey 1991, p. 94; M. Devitt 1990, pp. 252ff.; and Boghossian's reply in (1990b). Devitt, in fact, is even more critical than Boghossian toward deflationism, and claims that Tarski erred in calling his own the «semantic» conception of truth: for Devitt semantics in and of itself demands a correspondence conception, and evidently he agrees that Tarski's theory of truth is not substantially different from the deflationary conception. Of course, these issues are much too important in their own right to be fully discussed here.

F.P. Ramsey 1927: «Facts and Propositions», in *Philosophical Papers*, ed. by D.H. Mellor (Cambridge: Cambridge University Press, 1990).

A. Tarski 1935: «The Concept of Truth in Formalized Languages». English translation by J.H. Woodger, in *Logic, Semantic, Metamathematics*, ed. by J. Corcoran (Indianapolis: Hackett, 1983).

N. Tennant (unpublished), «On Two Transcendental Arguments about Truth and Truth-Conditionality».

C. Wright 1992: *Truth and Objectivity* (Cambridge, MA and London: Harvard University Press).

Pierluigi Miraglia
Department of Philosophy
The Ohio State University
350 University Hall
Columbus, OH 43210-1365, USA
E-mail: miraglia.1@osu.edu

A Theory of Textuality: The Logic and Epistemology. By Jorge J.E. Gracia. Albany: State University of New York Press, 1995. xxviii + 309 pages.

A Theory of Textuality is an important and welcome addition to the literature on texts and their interpretation. The value of Gracia's work lies in its extended treatment of the question: What is a text? Much of the literature on hermeneutics and theory of interpretation asks this question only in passing, if at all. The result has been a vague and imprecise conception of textuality which has swelled outside its proper domain. It has become common in continental circles, for example, to conceive of texts as occasions for interpretation. This conception allows dreams, historical occurrences, and other such phenomena to be considered texts. The problem with this conception is that it does not have a clear way to draw the line between text and nontext, for nearly everything is an occasion for interpretation of some sort. To borrow an example of Gracia's, even a car crash could be seen as a text under this conception. Something is clearly amiss with this understanding of textuality, and that cannot be ignored. Gracia sees this clearly and is of the conviction that we must come to terms with the logic of texts, including the definitional elements of their intension and extension, before considering the epistemological questions they pose.

Part one of *A Theory of Textuality* examines the logic of texts, and it is here that Gracia deals most directly with the question: What is a text? Chapter one is devoted to the intension of texts, chapter two to the extension of texts, and chapter three to a taxonomy of texts. In chapter one Gracia defines a text as, «a group of entities used as signs, which are selected, arranged, and intended by an author in a certain context to convey a specific meaning to an audience.»(4) Gracia justifies this definition through his characteristically rigorous argumentation, considering various counterexamples and counterarguments. Indeed, one of the virtues of the book is that the author takes his reader through his thought process and argumentation in scrupulous detail.

To return to the issue at hand, I find much of merit in Gracia's definition of 'text'. The definition, in fact, explains why we often use what seem to be incompatible predicates as applied to texts. For example, we use both physical predicates (such as heavy) and nonphysical predicates (such as incoherent) in speaking of texts. The distinction between signs and texts which is explicit in Gracia's definition (and which he argues for in detail) explains the apparent contradiction. A physical predicate refers to entities that can constitute texts (among them written signs as well as paper), and nonphysical predicates refer to the meaning of a text. This is a relatively simple and easily acceptable explanation. In fact it is so simple and acceptable that it is easy to overlook its weight and value. Much of Gracia's writing is similarly deceptive. It is often only after some reflection that the reader realizes that he has been mechanically nodding a yes as Gracia has made an important point on a controversial issue.

Gracia's example also explains why, for example, a car crash is not a text, although it may be an occasion for interpretation. Firstly, a car crash is not ordinarily composed of signs as a text must be. Secondly, even if it were composed of signs, these signs would not have been selected and arranged by someone with the intention of conveying meaning. Of course it is possible that one could orchestrate a textual car crash in which various parts of it would be intended as meaningful signs to some esoteric audience, but this is ordinarily not the case. Gracia's definition, then, does not restrict texts to the written, but it does draw sensible boundaries.

Chapter two deals with the extension of texts. Here Gracia is concerned with distinguishing texts from four entities with which they are sometimes confused: language, artifacts, art objects, and works. Texts are frequently composed in languages, but they are not themselves language. Texts do not have the flexibility and independence of authors and audiences that language per se does. A detailed discussion of the nature of artifacts concludes with the claim that texts are artifacts but not all artifacts are texts. An intriguing discussion of the distinction between the artistic and the aesthetic concludes that texts can be art objects or aesthetic objects but they need not be either.

Perhaps the most provocative discussion in chapter two focuses on the distinction between texts and works. The words 'text' and 'work' have often been seen as synonymous, largely due to their ordinary language use and conceptual overlap. Still, philosophers have sensed that there is a difference between 'work' and 'text', and have made attempts at clarifying the difference. Gracia considers all the most popular and viable attempts, in particular the idea that a work is a type of text. Nehamas has made this view attractive by speaking of works as interpreted texts. This definition does not work, as Gracia points out, because it does not account for translation. There is only one work *The Sun Also Rises*, but there are as many different texts which embody the work as there are translations of it. Translations are different texts because they are composed of different signs, but we speak of the work as remaining the same even when the text changes in translation.

Gracia proposes that we understand a work as being the meaning of a group of signs which is independent of that group inasmuch as other groups can be used to convey it. It is not the case, however, that the meaning of just any group of signs is a work. Gracia is not far from Nehamas's claim that only interpreted texts are works. Gracia does not hold, however, that all interpreted texts are works, but is somewhat more vague. As he says, «what makes the meaning of a particular text a work is that it fits a certain view of what a work is as developed by a culture at a particular point in history.»(67)

This is indeed an appropriate account because we do not speak of the meanings of all interpreted texts as works, but only of a select number. This is the first time in the book that Gracia makes use of a cultural explanation. He does so with impunity here largely because his aims are descriptive. As we shall see, in section two of his book Gracia at times leans too heavily on cultural explanations of normative and epistemological issues.

Chapter three which deals with the taxonomy of texts is the least interesting section of part one, although it contains at least one provocative discussion. Gracia's taxonomy offers both a modal and a functional classification of texts. The modalities of texts include actual texts, intended texts, and ideal

texts. The most intriguing and controversial claim of chapter three is that there is no such thing as an intended text. Gracia argues that this is so because a text is always a result of a process of production and does not precede such a process in any way.(72) Upon reflection it does seem true that before the composition of their texts authors have only more or less vague sets of ideas which they aim to realize and embody. They do not ordinarily have a text per se.

What Gracia can prove, however, is only the weaker claim that intended texts are quite uncommon especially when dealing with texts of any length and complexity. Although he addresses the most viable counterexamples, he does not do so successfully in all cases. We can, I submit, have intended texts particularly in the case of short and simple texts. For example, the nervous student who mentally rehearses the answer «Albany is the capital of New York.» only to say «Atlanta is the capital of New York.» surely had an intended text that was not realized in the spoken word.

It is true enough, as Gracia might object, that what the student had was an actual mental text. Still, that actual mental text also played the role of intended text for a spoken text that was imperfectly produced. There is no reason that a text cannot be actual in one medium, for example the mental, and intended for another medium, for example the spoken or written. Gracia suggests that we can explain slips of the tongue, such as in the example I have given, in terms of what we 'meant' and not what we 'intended'. Ordinarily we correct slips of the tongue by saying «That is not what I meant to say.» rather than «That is not what I intended to say.» This is very true, but the reason that it is true is that most slips of the tongue are not preceded by intended mental texts. Still some can be, as in the example I gave, and in those cases it would be proper to explain «That is not what I intended to say.» In the end, though, we owe a debt to Gracia for at least making us aware that intended texts are exceedingly scarce, especially for long and complex texts.

Part two of *A Theory of Textuality* deals with the epistemology of texts, specifically with questions of understanding, interpretation, and discernibility. Perhaps the greatest merit of chapter four is its clear and precise distinction between 'understanding' and 'meaning'. «Understanding is a kind of mental act whereby one grasps something which in the case of texts is their meaning.»(103) Meaning, on the other hand, is what is understood when one is said to understand a text.(108) Gracia has much of value to say about the nature of understanding itself, but his most important contribution here is this simple yet vital distinction between 'meaning' and 'understanding'. The two terms are frequently used as interchangeable in the literature, and clearly this practice is misguided.

The remainder of chapter four is devoted to a discussion of the limits of textual meaning and textual understanding. Noticeably absent from this section of the chapter is a discussion of textual significance. Gracia does deal with significance briefly in chapter one, saying that significance involves the relevance, importance, and consequences of a text.(18) Chapter four, however, is the place where a discussion of textual significance is most clearly needed. Ever since E.D. Hirsch first drew attention to the distinction between a text's meaning as opposed to its significance, this has become a distinction which demands attention. It is, then, disappointing and unfortunate to see it omitted here. Indeed, Gracia's discussion of the limits which authorial intention and understanding place on textual meaning suffers in the absence of a consideration of significance.

Gracia's account of the relationship between a text's «cultural function» and its meaning is also deficient in its neglect of significance. Gracia insists, as many would, that the cultural function of a text plays an important role in the determination of its meaning. I would argue that there are important consequences of a text's cultural function, but these are consequences affecting significance rather than meaning. Certainly this is a point of much debate, and I do not have the space to address it fully here.

Whether we attribute the relevance of a text's cultural function to its meaning or significance, we must look critically at what the cultural function demands. This is a normative issue which is inextricably bound to the epistemological issue, but one which Gracia neglects. We cannot afford to accept the status quo of textual cultural functions uncritically, just as we cannot afford to accept the cultural dictates of mores and ethics uncritically. Without critical reflection on and ethical analysis of the cultural functions of textual genres we slide down the slippery slope into interpretive relativism. Under Gracia's scheme a culture which took all works of literature to be the word of God would be justified in doing so. Certainly though, such an interpretive practice would demand more justification than mere cultural fiat.

Chapter five «Interpretation» provides an insightful account of the function of interpretations, and makes a keen distinction between two primary types of interpretation, textual and nontextual. Textual interpretations are those whose main or only purpose is to produce understandings of the meanings of texts and of the implications of those meanings. Nontextual interpretations are those whose primary aim is other than to produce such understandings. Nontextual understandings may be, for example, Freudian, Marxist, or feminist. Such interpretations are more concerned with a certain significance of the text than its meaning (although Gracia does not put it in quite these terms).

The distinction between textual and nontextual interpretations is important because, as Gracia makes clear, they are very different things and often interpreters are not themselves clear as to what type of interpretation they are offering. Gracia also cogently argues that there is nothing wrong with textual interpretation, as long as it is recognized as such and is built on understanding not misunderstanding.

Chapter six «Discernibility» addresses three questions: 1) How do I know that something is a text? 2) How do I learn the meaning of a text? 3) How can I be certain that I know the meaning of a text? To be clear, the certainty with which we can ever know that something is a text or know its meaning is not apodicitic certainty. Still, Gracia offers well-considered and valuable answers to these questions. The common sense analysis, that ultimately the basis on which we learn the meaning of a text is expected behavior in context, has definite appeal. Gracia claims that behavior is actually the key to breaking the hermeneutic circle. As he says, «to break the circle we need only one case in which we can have certainty that a text has been understood on the basis of something that is not a text.» (206) We should acknowledge that many continental philosophers do not see the hermeneutic circle as being a problem but as actually having an ontologically positive status. Still Gracia's proposal seems to me to be the beginning of a promising solution to one of the problems that has haunted interpretive theory since its inception.

A comprehensive conclusion to the book follows chapter six. The conclusion brings together all the major elements of Gracia's theory in a few short pages, and will serve as a good reference for the reader who wishes to reacquaint himself with the theory. A thirty-six page bibliography provides the reader with a detailed source of classical and contemporary hermeneutical studies, and a thorough index facilitates searches within the book.

A Theory of Textuality is an excellent work of philosophy, and is essential reading for all those concerned with the study of texts and their interpretation. I believe both continental and analytic philosophers will find much of value in this book. Although the book is relatively free of struggles with specific historical and contemporary theorists, it reflects a broad reading and consideration of both analytic and continental philosophers. Gracia has continued in the spirit of rapprochement he called for in *Philosophy and Its History (1992)* by himself delving into much of the continental tradition. His consideration of the postmodernists, deconstructionists, and hermeneuticists is apparent, although these groups are unlikely to find his conclusions congenial. In conclusion, then, this reviewer highly recommends *A Theory of Textuality*, and wishes to draw the reader's attention to Gracia's forthcoming volume on the metaphysics of texts. This book will complete Gracia's study of textuality and address many of the important metaphysical and ontological issues which were outside the scope of the current volume.

William Irwin

State University of New York at Buffalo, Department of Philosophy
607 Baldy Hall45-38-25, Buffalo, New York 14260-1010
Fax 716)645-38-25

The Mind Association Research Fellowship 1996-97

The Association proposes to award from time to time, as funds permit, to individual scholars a sum of £ 5,000 to assist in their research activities. It is envisaged that the sum would be paid to the University at which the scholar is employed and that the funds would be used by the department in which the scholar works to release him or her from teaching and administrative duties, so that the scholar to whom the award is made can concentrate upon their research activities. The Fellowship is specifically designed to help academics who are unduly burdened with departmental commitments. The award will be competitive and scholars wishing to apply should write to the Secretary of the Association enclosing a CV and a short statement of the research to be carried out. A committee consisting of the Editor of Mind and two members of the Executive Committee of the Association will then make the award. The Secretary of the Association will communicate the result of the competition to the applicants. The deadline for applications for the academic year of 1996-7 is the 31st of December, 1995. Previous recipients of the award are Graham Macdonald, University of Bradford (1994-5), and Colin Howson, London School of Economics (1995-6).

♣ ♣ ♣ ♣ ♣ ♣ ♣

ISIS: Information, Statistics and Induction in Science
Melbourne, Australia, 20-23 August 1996
*** CALL FOR PAPERS ***
Conference Chair: David Dowe
Co-chairs: Kevin Korb and Jonathan Oliver

INVITED SPEAKERS:

Henry Kyburg, Jr. (University of Rochester, NY)

J. Ross Quinlan (Sydney University)

Jorma J. Rissanen (IBM Almaden Research, San Jose, California)

Ray Solomonoff (U.S.A.)

PROGRAM COMMITTEE:

Lloyd Allison, Mark Bedau, Hamparsum Bozdogan, Wray Buntine, Peter Cheeseman, Honghua Dai, David Dowe, Doug Fisher, Alex Gammerman, Clark Glymour, Randy Goebel, David Hand, Bill Harper, David Heckerman, Colin Howson, Lawrence Hunter, Frank Jackson, Max King, Kevin Korb, Henry Kyburg, Ming Li, Nozomu Matsubara, Aleksandar Milosavljevic, Richard Neapolitan, Jonathan Oliver, Michael Pazzani, J. Ross Quinlan, Glenn Shafer, Peter Slezak, Ray Solomonoff, Paul Thagard, Neil Thomason, Raul Valdes-Perez, Tim van Gelder, Paul Vitanyi, Chris Wallace, Geoff Webb, Xindong Wu, Jan Zytkow.

Inquiries to:

isis96@cs.monash.edu.au

David Dowe: dld@cs.monash.edu.au

Kevin Korb: korb@cs.monash.edu.au or

Jonathan Oliver: jono@cs.monash.edu.au

Information is available on the WWW at:

http://www.cs.monash.edu.au/~jono/ISIS/ISIS.html

This conference will explore the use of computational modelling to understand and emulate inductive processes in science. The problems involved in building and using such computer models reflect methodological and foundational concerns common to a variety of academic disciplines, especially statistics, artificial intelligence (AI) and the philosophy of science. This conference aims

to bring together researchers from these and related fields to present new computational technologies for supporting or analysing scientific inference and to engage in collegial debate over the merits and difficulties underlying the various approaches to automating inductive and statistical inference.

AREAS OF INTEREST.

The following streams/subject areas are of particular interest to the organisers:

Concept Formation and Classification.

Minimum Encoding Length Inference Methods.

Scientific Discovery.

Theory Revision.

Bayesian Methodology.

Foundations of Statistics.

Foundations of Social Science.

Foundations of AI.

CALL FOR PAPERS.

Prospective authors should mail five copies of their papers to Dr. David Dowe, ISIS chair. Alternatively, authors may submit by email to isis96@cs.monash.edu.au. Email submissions must be in LaTex (using the ISIS style guide [will be available at the ISIS WWW page]). Submitted papers should be in double-column format in 10 point font and not exceeding 10 pages. An additional page should display the title, author(s) and affiliation(s), abstract, keywords and identification of which of the eight areas of interest (see http://www.cs.monash.edu.au/~jono/ISIS/ISIS.Area.Interest.html) are most relevant to the paper. Refereeing will be blind; that is, this additional page will not be passed along to referees.

The proceedings will be published; details have not yet been settled with the prospective publisher. Accepted papers will have to be represented by at least one author in attendance to be published.

Papers should be sent to:

Dr David Dowe

ISIS chair

Department of Computer Science

Monash University

Clayton Victoria 3168

Australia

Phone: +61-3-9 905 5226

FAX: +61-3-9 905 5146

Email: isis96@cs.monash.edu.au

Submission (receipt) deadline: 11 March, 1996

Notification of acceptance: 10 June, 1996

Camera-ready copy (receipt) deadline: 15 July, 1996

CONFERENCE VENUE

ISIS will be held at the Old Melbourne Hotel, 5-17 Flemington Rd. North Melbourne.

The Old Melbourne Hotel is within easy walking distance of downtown Melborne, Melbourne University, many restaurants (on Lygon Street) and the Melbourne Zoo. It is about fifteen to twenty minutes drive from the airport.

REGISTRATION

A registration form will be available at the WWW site:

http://www.cs.monash.edu.au/~jono/ISIS/ISIS.shtml,

or by mail from the conference chair. Dates for registration will be considered to be met assuming that legible postmarks are on or before the dates and airmail is used. Student registrations will be available at a discount (but prices have not yet been fixed). Relevant dates are:

Early registration (at a discount): 3 June, 1996

Final registration: 1 July, 1996

SORITES

NOTES TO POTENTIAL CONTRIBUTORS

All submitted manuscripts will be refereed either by members of the Board of Advisors or by other specialists; as far as possible, each manuscript will be refereed by philosophers not unsympathetic to the paper's philosophical outlook or orientation.

No manuscript may be submitted if it is being considered for publication elsewhere.

Once accepted, papers may not be printed or displayed elsewhere or incorporated into a book, an anthology or any other publication of any sort, unless and until **SORITES** has accorded the author(s) permission to that effect — which in normal cases will be done routinely, provided **SORITES** is duly acknowledged as the primary source. By submitting a paper, the author agrees to the points, terms and conditions contained in the Copyright Notice included in each issue of **SORITES**.

All submitted papers must be written in English. The author's local variety of English (including the spelling) will be respected — be it Indian, Filipino, Australian, American, Western-African, British, Southern-African, Eastern-African, Jamaican, etc. All editorial material will be written in BBC English, which is the journal's «official» dialect.

There is no settled length limit for papers, but we expect our contributors to stand by usual editorial limitations. The editors may reject unreasonably long contributions.

We expect any submitted paper to be accompanied by a short abstract.

We welcome submissions of in-depth articles as well as discussion notes.

Ours is a journal granting a broad freedom of style to its contributors. Many ways of listing bibliographical items and referring to them seem to us acceptable, such as '[Moore, 1940]', or '[M:5]' or '[OQR]'. What alone we demand is clarity. (Thus, for instance, do not refer to '[SWT]' in the body of the article if no item in the bibliography collected at the end has a clear '[SWT]' in front of it, with the items sorted in the alphabetic order of the referring acronyms.) We prefer our contributors to refer to 'Alvin Goldman' rather than 'Goldman, A.', which is obviously ambiguous. We dislike implied anachronisms like [Hegel, 1989]' or '[Plato, 1861]' — but you are entitled to ignore our advice.

How to submit?

(1) We will be thankful to all contributors who submit their papers in the form of [I.B.M.-PC] WordPerfect 5.1 files. There are several convertors which can be used to turn docs from other word processor formats into WP5.1 format. (Notice that with WP5.1 you can write not only almost all diacritically marked characters of any language which uses the Latin script, but moreover all of Greek and virtually all symbols of mathematical logic and set theory.)

(2.1) In case a contributor can neither use WP5.1 nor have their doc converted into WP5.1 format, they can send us their file in its original format (be it a different version of WordPerfect or another sort of word-processor, such as MS-Word, MS-Word for Windows, WordStar, AmiPro, XyWrite, DisplayWrite, .rtf, etc). We'll try (and hopefully in most cases we'll manage) to convert those files from other formats into WordPerfect 5.1.[1]

---

[1] Unfortunately we cannot yet handle TeX or LaTeX files. The convertors we've tried have proved useless.

(2.2) When WP5.1 format is not available and we have been unable to use the original file, a good idea is for the author to have their doc converted to a .html file (there are lots of HTML editors and document-to-HTML converters from a great many formats — PC-Write, [La]TeX, MS-Word and Windows-Word etc). We expect HTML files to bear the extension '.htm'.[2]

(2.3) Anauthor solution is to use [stripped and extended] ASCII format, which means: text files (not binary ones) written using any printable ASCII characters of Code-page 437 (USA or default), i.e. any character except ASCII_00 through ASCII_31; with CRs (carriage returns) only between paragraphs — not as end-lines. Such files will here be called 'ASCII files'. We expect them to bear the extension '.ASC'.

(2.4) Another alternative (which is in itself worse, but which nevertheless may be more practical in certain cases) is to use the DOS text format, with no character outside the range from ASCII_32

---

[2]   **The following information is mainly due to Ian Graham. We have abridged some relevant parts of his document and added the item concerning LaTeX.**

HTML Writer HTML Writer is a Windows-based HTML editor. Additional information can be found at: http://www.et.byu.edu/~nosackk/html-writer/index.html.

HoTMetaL for Windows is a commercial HTML editor, but afree implementation is available via anonymous FTP. There may also be Mac (and other) versions. SoftQuad (who makes HoTMetaL) has their own Web server with up-to-date information. There are several anonymous ftp sources of the HoTMetaL executable. One is in gatekeeper.dec.com, while another is the NCSA ftp archive.

PC-Write-HTML-Editing-Macros, a package for editing HTML docs with the PC-Write editor, is available at: ftp://www.ucc.ie/pub/pcw4.zip.

HTML Assistant is an MS Windows text editor with extensions to assist in the creation of HTML hypertext docs to be viewed by World Wide Web browsers like Cello and Mosaic. FTP-available from ftp.cs.dal.ca/htmlasst/.

HyperEdit is a facility designed for MS-Windows users to aid in the creation of HTML docs. Version v0.2a is largely based on «A Beginners Guide to HTML» produced by NCSA. Author: Steve Hancock, s.hancock@info.curtin.edu.au.

CU_HTML.DOT CU_HTML.DOT is a Microsoft Word for Windows 2.0 doc template that allows users to create HTML docs inside Word in a WYSIWYG manner and generate a corresponding HTML file. The ZIP file is at ftp.cuhk.hk/ /pub/www/windows/util/CU_HTML.ZIP.

HTML for Word 2.0 by NICE technologies, France, creates a structured doc environment for Word 2.0. It creates doc instances that conform to ISO 8879 (SGML), and is available from the ftp.cica.indiana.edu FTP site or from its mirrors. Additional information: Eric van Herwijnen, NICE technologies, chemin des Hutins, Veraz, 01170 Gex, France. Tel (33)-50424940.

HTMLed, a customizable HTML editor with toolbars, can be obtained via anonymous ftp at pringle.mta.ca/pub/HTMLed. The file is htmed12.zip.

Rtftohtml, which converts Microsoft RichText Format to HTML, supports WORD 6.0 RTF files. This program is useful for MS-Word docs, as these use RTF format as the doc code. This code can be used on Macintoshes, PC's or on Unix boxes. RTFTOHTM-Tools is a distribution containing a conversion DLL and a doc template for WinWord 2.0. This sotware (rtftohtm.dll and html.dot) when used together, allow transparent conversion from WinWord to HTML, not only from RTF to HTML.

PSTOHTML (PostScript-to-HTML Converter)is a Perl-script package for converting postscript-to-html, and also for converting PostScript to plain text. If you have perl on you PC, then you can run this. Users of this code need a postscript interpreter, e.g.

LaTeX2HTML is a Perl program that converts documents written in LaTeX into the HTML format. It handles equations, tables, figures, footnotes, lists and bibliographies. It translates accented and special characters to the equivalent ISO-LATIN-1 character set whenever possible. The actual code is located at http://cbl.leeds.ac.uk/nikos/tex2html/latex2html.tar or http://cbl.leeds.ac.uk/nikos/tex2html/latex2html.tar-gz. The author is Nikos Drakos, <nikos@cbl.leeds.ac.uk>, http://cbl.leeds.ac.uk/nikos/personal.html.

through ASCII_126, no hyphenation, a CR at the end of each line and two CRs separating paragraphs. Such files will be here called 'text files'; we expect them to bear a '.txt' extension.

(3) In cases (2.2) and (2.4) the contributor can include their paper into an e_mail message sent to one of our editorial inbox (sorites@olmo.csic.es)

(4) Before sending us their file the contributor is advised to compress it — except in case they are sending us a text file through procedure (3) above. Compression reduces disk-storage and shortens transmission time. We can extract and expand files archived or compressed with Diet, ARJ (both warmly recommended), Tar, Arc, Zip (or PKZip), GZip, Compress (i.e. .Z files), LHA, Zoo, RaR, and some versions of the MAC archivers PackIT and StuffIT.

(5) The most expedient way for contributors to send us their submitted paper is through anonymous FTP. At your host's prompt, you enter 'FTP olmo.CSIC.es'; when you are prompted for your username, you answer 'FTP' or 'anonymous'; when you are next prompted for your password, you answer with your e_mail address; once connected, you enter 'cd pub/sorites/incoming', then 'binary', and then 'put xxx' — where xxx is the file containing your submitted paper and a covering letter. (If the file is an archive, the extension must reveal the archiving utility employed: '.gz', '.Arj', '.RAR', etc. (DIETed files needn't bear any special denomination or mark; they will always be automatically recognized by our reading software.)

(6) Whenever a paper is submitted, its author must send us a covering letter as an e_mail message addressed to one of our editorial inboxes.

(7) If a contributor cannot upload their file through anonymous FTP, they can avail themselves of one of the following alternatives.

(7.1) If the file is a '.htm' or a '.txt' file (i.e. in cases (2.2) and (2.4)), simply include it into a e_mail message.

(7.2) In other cases, an 8-to-7 bits converter has to be used, upon which the result can also be included into an e_mail message. 8-to-7 bits convertors «translate» any file (even a binary file) into a text file with short lines which can be e-mailed. There are several useful 8-to-7 convertors, the most popular one being UUenCODE, which is a public domain software available for many different operative systems (Unix, OS/2, DOS etc). Another extremely good such convertor, very easy to use, is Mike Albert's ASCIIZE.[3] We can also decode back into their binary original formats files encoded into an e-mailable ASCII format by other 8-to-7 bits convertors, such as: TxtBin, PopMail, NuPop, or University of Minnesota's BINHEX, which is available both for PC and for Macintosh computers. Whatever the 8-to-7 bits encoder used, large files had better be previously archived with Arj, Diet or any other compressor, the thus obtained archive becoming the input for an 8-to-7 bits convertor.[4]

(7.3) An alternative possibility for contributors whose submitted papers are WordPerfect 5.1 or WordPerfect 6 docs is for them to use a quite different 8-to-7 bits convertor, namely the one provided by the utility Convert.Exe included into the WordPerfect 5.1 package. (WordPerfect corporation also sells other enhanced versions of the convertor. WordPerfect 6.0 has incorporated

---

[3]   Mike Albert's address is P. O. Box 535, Bedford, MA 01730, USA.

[4]   For the time being, and as a service to our readers and contributors, we have a directory called 'soft' hanging from our directory sorites at the node olmo.csic.es. The directory contains some of the non-commercial software we are referring to, such as archivers or 8-to-7 encoders (or 7-to-8 decoders).

a powerful conversion utility.) A separate e_mail message is mandatory in this case informing us of the procedure. The result of such a conversion is a 'kermit-format' file.[5]

(8) You can also submit your manuscript in an electronic form mailing a diskette to one of the Submissions Editor (Prof. Manuel Liz, Facultad de Filosofia, Universidad de La Laguna, Tenerife, Canary Islands, Spain; Telephone Nr. +3422-603166; Fax Nr. +3422-603102). Diskettes will not be returned, and regular-mail correspondence will be kept to a minimum.

(9) Such submitted papers as are neither WordPerfect 5.1 files nor files in HTML format require some preparation.

(9.1) Ours is not a logic journal, but of course one of the glories of analytical philosophy is its rigour, which it partly owes to auxiliary use of symbolic notation in order to avoid ambiguities, make matters of scope clear or render arguments perspicuous. ASCII translations of symbolic notation are problematic, especially in cases of nonclassical logics, which may use sundry negations, disjunctions, conjunctions, conditionals, implications and also different universal and particular quantifiers (e.g. existentially and nonexistentially committed quantifiers, a familiar dichotomy in Meinongian circles). While using WordPerfect 5.1 you can represent a huge variety of such nuances, it is impossible to express them within the narrow framework of text or even ASCII files (i.e. even when the 224 printable [extended] ASCII characters can be used). Still, for some limited purposes, a translation of sorts can be attempted. You are free to choose your representation, but the following translation is — for the time being — a reasonable one: '(x)' for universal quantifier, '(Ex)' for existential quantifier; '&' for conjunction; 'V' for disjunction; '->' for implication (if needed — something stronger than the mere 'if ... then'); 'C' for conditional; '=>' for an alternative (still stronger?) implication; '_pos_' for a possibility operator; '_nec_' for a necessity operator.

(9.2) In ASCII or text files all notes must be end-notes, not foot-notes. Reference to them within the paper's body may be given in the form '\n/', where n is the note's number (the note itself beginning with '\n/', too, of course). No headings, footings, or page-breaks. In such files, bold or italic bust be replaced by underscores as follows: the italized phrase '*for that reason*' must be represented as '_for that reason_' (NOT: '_for_that_reason_'). A dash is represented by a sequence of a blanc space, two hyphens, and another blanc space.[6]

---

[5] In the case of WordPerfect 5.1, the procedure is as follows. Suppose you have a file called 'dilemmas.wp5' in your directory c:\articles, and you want to submit it to **SORITES**. At your DOS prompt you change to your directory c:\articles. We assume your WordPerfect files are in directory c:\WP51. At the DOS prompt you give the command '\wp51\convert'; when prompted you reply 'dilemmas.wp5' as your input file whatever you want as the output file — suppose your answer is 'dilemmas.ker'; when prompted for a kind of conversion you choose 1, then 6. Then you launch you communications program, log into your local host, upload your file c:\articles\dilemmas.ker using any available transmission protocol (such as Kermit, e.g.). And, last, you enter your e_mail service, start an e_mail to to sorites@olmo.csic.es and include your just uploaded dilemmas.ker file into the body of the message. (What command serves to that effect depends on the e_mail software available; consult your local host administrators.)

　　　With WordPerfect 6 the conversion to kermit format is simple and straightforward: you only have to save your paper as a 'kermit (7 bits transfer)' file.

[6] Those devices are temporary only. Later on we'll strongly advise and encourage those of our contributors who can use neither WordPerfect format nor one of the other word-processor formats our convertors can handle automatically to resort to HTML, with certain conventions in order to represent Greek characters as well as logical and set-theoretic symbols.

SORITES

An Electronic Quarterly of Analytical Philosophy

ISSN 1135-1349

---

[1]   The reader may find an excellent discussion of copyright-related issues in a FAQ paper (available for anonymous FTP from rtfm.mit.edu [18.70.0.209] /pub/usenet/news.answers/law/Copyright-FAQ). The paper is entitled «Frequently Asked Questions about Copyright (V. 1.1.3)», 1994, by Terry Carroll. We have borrowed a number of considerations from that helpful document.