Law and Philosophy Library

José Juan Moreso

Legal Indeterminacy and Constitutional Interpretation

Springer-Science+Business Media, B.V.

$$(1)\, f(OP) = \wedge p$$

and

$$(2)\, f(Pp) = \vee p.$$

Once we have this function, we can introduce the resulting declarative sentence, *p'*, in the theory of sense and analyze it with the help of convention-T.

This approach also makes it possible to ascribe propositional attitudes to the person uttering the sentences. The standard ascription is that if someone says that *p*, then he believes that *p*; and if someone prescribes that *p*, then the standard ascription is that he wishes that *p* (Platts 1979, 63). In any case, propositional attitudes concern the place of intention in a theory of meaning. That is what I will turn to now.

b) Intentions and meanings

Some 25 years ago, Strawson (1971, 170-189) already examined two contending approaches in philosophical theories of meaning: theories of communicative intention, on the one hand, and theories of formal semantics, on the other. He presented that conflict in the following terms:

„A struggle on what seems to be such a central issue in philosophy should have something of a Homeric quality; and a Homeric struggle calls for gods and heroes. I can at least, though tentatively, name some living captains and benevolent shades: on the one side, say, Grice, Austin, and the later Wittgenstein; on the other, Chomsky, Frege, and the earlier Wittgenstein." (Strawson 1971, 172)

Strawson himself advocates an approach in the line of the theorists of communicative action; the main target of his criticism are theorists of formal semantics *à la* Davidson.

Theorists of communicative action usually present their approach in the following way: They form a primitive concept of *communication* (or communicative intention) such that it does not presuppose the concept of *linguistic meaning*, and they then show that the latter can and must be explicated in terms of the former.

The most elaborate conception of this kind is that of Grice (1989). Grice's crucial distinction is the one between the notions of *sentence-meaning* and *utterer's meaning*, i. e., between 'sentence *s* means that *p*' and 'by uttering sentence *s*, *x* means that *p*'. The theory then defines *utterer's meaning* in terms of the notion of the utterer's intentions, and finally attempts to define *sentence-meaning* in terms of the previously defined utterer's meaning.

Here, I do not wish to analyze Grice's several attempts to refine his definition of *utterer's meaning*,[11] as a reaction to the criticisms it provoked (cf. Grice 1989, 93-104).

[11] Grice's standard definition of 'utterer's meaning' is as follows:
„'*U* meant something by uttering *x*' is true iff, for some audience *A*, *U* uttered *x* intending:
(1) *A* to produce a particular response *r*
(2) *A* to think (recognize) that *U* intends (1)
(3) *A* to fulfill (1) on the basis of his fulfilment of (2)." (Grice 1989, 92)

I will only try to show that *none* of the approaches to meaning by theorists of communicative action can substitute the theory of sense in terms of truth-conditions.

According to Grice, sentence-meaning is defined in terms of the intentions with which a sentence is uttered, perhaps together with the receivers' response habitually guaranteed by that utterance. Now, Platts (1979, 89 f.) argues, that explication fails as an explication of the meaning of sentences in natural languages. Most of those sentences (of which there is an infinite number in natural languages) will never be uttered. Therefore, neither will they be uttered with any intention, nor will their being uttered induce any response in some audience. Thus, Platts continues,

„The obvious move, perhaps the only possible one, is to hold the meanings of such sentences to be definable in terms of the intentions with which they would be uttered were they to be uttered and the responses they would then induce".

Now, Grice confronts a dilemma: Either those intentions and responses are somehow restricted, or they aren't. In the latter case, the meanings of those unuttered sentences will be completely undetermined; in fact, they will not mean anything at all. And in the former case, that restriction can only come from the meaning of the sentence itself. Therefore, the attempt to define the meanings of unuttered sentences through hypothetical intentions and responses fails; a notion of sentence-meaning must be presupposed. And such a notion we obtain from the theory of sense in terms of truth-conditions.

But then, what is it that makes the approach of the theorists of communicative action so attractive? Perhaps, Platts (1979, 92 f.) suggests, that it does not distinguish between the following theses:

„(T.1) A system of utterances would not be a language unless intentions were ascribed to the utterers; more specifically, an utterance is a piece of *linguistic* behaviour only if it is intentional.
(T.2) The notion of sentence-meaning can be defined in terms of the notion of utterers' intentions.
(T.3) The meaning of any particular sentence in a language can be determined by reference to the intentions with which it is uttered."

While we can assume the truth of T.1, it in no way implies T.2 or T.3. But the truth of T.2 and T.3 is precisely what an approach to meaning like that of communicative action requires.

Acceptance of T.1 does not refute the theory of sense I have presented. Intentions must even be given a central place in our theory of linguistic behaviour. That place, however, is not in the theory of sense, but in the theory of force; more precisely, in the component of the theory of force by which we identify the mood of the utterance of a sentence. As we have seen, it is the theory of force that shows us the propositional attitude (the direction of fit) with which a sentence is uttered.

That means that the 'Homeric struggle' between theorists of communicative action and theorists of formal semantics must be concluded by a hybrid theory of linguistic behaviour which gives room to intentions (in the theory of force), but which also preserves an approach in terms of truth-conditions (for the theory of sense) (cf. Rumfitt 1995 on the need for such a hybrid theory).

There is, however, another idea of Grice's (1989, 22-40) which, I think, deserves closer attention. It is the idea that in any declarative utterance we can distinguish that which is asserted from that which is *implied* by what is asserted. Something similar can be said of imperative utterances, where we can distinguish that which is explicitly prescribed from that which is implied by what is prescribed (it is precisely this distinction which enables us to speak of a normative system as a deductive system). Now, one could think that the classical notion of *logical consequence* is all that is needed here. If someone says 'There was a table and a chair in the room', he has also (implicitly, as a consequence of his explicit assertion) said 'There was a chair in the room'. Now, since from that last expression we logically derive 'There was a chair or a dog in the room', he would also implicitly have asserted the proposition expressed by that last sentence. Part of Grice's work on so-called „conversational implicatures" attempts to restrict the possibility of obtaining such counterintuitive conclusions.

In the next section, I will propose a notion of *relevant conclusion* that can mitigate those counterintuitive conclusions. In order to do that, I will take as a paradigm a certain conception of the justificaton of judicial decisions.

c) Relevance and implicit meaning

In the recent literature on the justification of judicial decisions, we usually find a distinction being made between *external* and *internal* justifications (Wroblewski 1971, 411 f.; Alexy 1978, 119 f.; Aarnio 1987, 119 f.). Internal justification concerns the validity of the inference from some premises to a judicial decision, taken as their conclusion. The function of external justification, in contrast, is centered on the control of the premises.

A judicial decision is *internally* justified if and only if it follows from the norms applicable to the case, and from the sentence(s) describing the facts of the case. On that conception, a judicial argument *(JA)* can take the following form:

(1) If x does A, x must be punished with sanction S.
(2) x does A.

Therefore, *(3)* x must be punished with sanction S.

(3) is justified since it is a deductive consequence of *(1)* and *(2)*. Despite the numerous criticisms it has evoked, it seems to be widely agreed that deductive arguments play a central role in internal justification (e. g., MacCormick 1994, ix). Thus, internal justification can be reconstructed as a logical inference.

Most criticisms of deductive logic, however, relate to the insufficiency of the instruments of formal logic for analyzing the question of the adequacy of the premises (i. e., external justification). And the main problems of justification in the law are often said to concern external justification (Aarnio 1987, 119 f.)

But I will not go into these problems here. Therefore, in what follows, 'justified legal decision' must simply be taken as a synonym for '*internally justified* legal decision'.[12] I will also not treat the problem of whether the insufficiency of deductive logic may be due to the complexity of some particular external justification.[13] Instead, I will be concerned with the question of whether deductive logic (and internal justification) does not actually prove too much. It should not be forgotten that any set of premises logically entails an infinite number of consequences and, therefore, an infinite number of sentences would be justified with any judicial argument. In *JA*, from *(1)* and *(2)* we cannot only deduce *(3)*, but also

(3') x must be punished with sanction S or given award A.
(3'') If it rains tomorrow, x must be punished with sanction S.

and (infinitely) many more.[14]

(3') and *(3'')* are logical consequences of *(1)* and *(2)*, but intuitively they seem to be paradoxical consequences. No jurist would regard *(3')* or *(3'')* as adequate ('justified') sentences grounded in *(1)* and *(2)*. But if the only criterion of justification for judicial decisions is that of logical deduction, then we must try to find out why these conclusions seem paradoxical, and to find possible solutions. The paradox perhaps lies in the lack of a connection of *(3')* and *(3'')* with the premises of the argument, in — as we may say — the *irrelevance* of those conclusions.

The notion of relevance, however, eludes precise conceptual reconstruction and, what's even more important, an adequate logical treatment.[15]

---

[12] Lyons (1993, 119-140) prefers to reserve the expression 'justified judicial decision' for morally adequate judicial decisions. Such decisions are not necessarily the same as the decisions that derive from legal norms. But here I will not go into this topic either.

[13] In Moreso/Navarro/Redondo 1992 we have treated that question under the name of the 'thesis of insufficiency'.

[14] *JA* could be represented as follows:

$$\begin{array}{l} 1)\ p \rightarrow Oq \\ 2)\ p \\ \hline \vdash 3)\ Oq \end{array}$$

I will assume that expressions like '$p \rightarrow Oq$' are adequate representations of conditional norms. As is well-known, there is no agreement on this in the literature on deontic logic. Von Wright, to mention an undisputed authority, has taken a vacillating stand on this central point of the logical representation of norms. Thus, in 1980 he wrote: „It should be noted that this standard system differs in two important respects from the system proposed in my (1951). First, it admits 'mixed' formulae formed of propositional variables, sentential connectives, and deontic expressions. For example: $p \rightarrow Oq$, is a well-formed formula of the standard system ..." (von Wright 1980, 403). But only a few years later, he observed: „On the interpretation of Oq as a prescription (norm), the expression $p \rightarrow Oq$ would consist of a descriptive and a prescriptive ingredient joined by a sentential connective. Does such an expression 'make sense'? The question is not really easy to answer. It seems to me clear, however, that if the standard connective in question is a *truth-connective*, then this 'linguistic hybrid' is a monster with no place in meaningful discourse." (von Wright 1983a, 151)

[15] In Appendix B, I will discuss some of the notions of relevance to be found in the philosophical literature, one of which (the one that seems to me the most adequate) will be applied to the question of the justification

To reason is to show that from a given set of premises *{P}*, some conclusion *C* can be deduced. In classical logic, that *C* can be deduced from *{P}* means (in terms of classical semantics) that if the premises *{P}* are true, then *C* is necessarily true, i. e., that *{P} entails C*. The notion of entailment is regarded to be the converse of the notion of 'following from' or 'being deduced from'. Apparently, G. E. Moore (1922, 291; cf. also von Wright 1957, 166) was the first one to introduce this notion:

„We require, first of all, some term to express the converse of the relation which we assert to hold between a particular proposition *p* and a particular proposition *q*, when we assert that *q follows from* or *is deducible from p*. Let us use the term 'entails' to express the converse of this relation." (G. E. Moore 1922, 291)

Thus, if we have two premises

$(1)\, p \rightarrow q$ $\qquad\qquad$ $(2)\, p$

we can deduce

$\vdash q,$

but we can also deduce

$\vdash q \vee r,$
$\vdash r \rightarrow q,$
$\vdash \neg q \rightarrow r,$

and many other conclusions. For instance, from the premises 'If you come with me, I go to the movie' and 'You come with me', we can deduce 'I go to the movie', but also 'I go to the movie or I stay at home', 'If you insult me, I go to the movie' or 'If I don't go to the movie, you will insult me'.[16] The paradoxical note of these conclusions (which are necessarily true if the premises are true) is due to the fact that there is no connection between them and the starting point. Apparently, an argument is irrelevant if there is no connection between the premises and the conclusion; and the connection of logical entailment is not enough to produce relevant arguments.

If to this we add the fact that many logicians mistrust a formal treatment of the notion of *relevance*,[17] it seems that a theory of reasoning based on the formal-logical notion of deduction is highly deficient.

---

of judicial decisions, followed by some reflections, related to that notion of relevance, on the concept of a legal system as a normative system.

[16] The paradoxical overtone may become manifest in the following way: I say to a friend 'If you come with me, I'll go to the movie'; he answers 'I'll go with you'; and I say 'If you insult me, I'll go to the movie'. The contexts of human communication in which the use of practical reasoning takes place show that this kind of reasoning is inadequate (for common purposes).

[17] Thus, Suppes (1957, 8) writes that „... the notion of connection or dependence being appealed to here is too vague to be a formal concept of logic"; and Susan Haack (1978, 16 f.) holds that „Considerations of rele-

Among the philosophers interested in reasoning, this leads to three possible attitudes:

I) A (possibly partial) rejection of formal logic as a criterion for controlling our arguments. This thesis has two corollaries:

I.a) There are valid arguments that do not follow the patterns of formal logic.

I.b) There are invalid arguments that do follow the patterns of formal logic.

II) A construction of a new, deviant, non-classical logic that does not produce the paradoxes of irrelevance. In other words, the elaboration of a (deviant) notion of logical consequence which incorporates the notion of relevance.

III) A restriction of classical logic with additional criteria of relevance. For this conception, there can be no valid arguments that do not follow the patterns of classical logic; but there are formally valid arguments that are irrelevant. For this conception, reasoning consists in logic *plus* relevance.

In *Appendix B*, the reader will find a detailed analysis of these three approaches. Suffice it here to point out that, in my view, in order to show that a set of normative and factual premises *justifies* a conclusion, the conclusion must not only derive logically from the premises; it must also be shown that it is a *relevant* consequence of that set of premises. For this reason (among others), in a normative system it is important to distinguish the derived norms which are relevant consequences of that system from those which are irrelevant consequences.

4. Conclusions

(1) The law can be seen as a set of prescriptive and conceptual rules containing all their logical consequences, i. e., as a *normative system.*

(2) In such a set, it is important to distinguish the consequences that are *relevant* from those that are *irrelevant.*

(3) In order to adopt this approach, we need to *attribute meaning* to norm-formulations, i. e., to *interpret* those formulations, which are the result of certain speech-acts by certain norm-authorities.

(4) Such an attribution of meaning presupposes the attribution of propositional attitudes to the authorities. For this, we need a theory of *illocutionary force* that can account for the intentions of such norm-authorities.

(5) In order to identify the normative content prescribed by authorities, we need a theory of *sense*. Such a theory is provided by *convention-E*, which is parasitic on *convention-T* and, therefore, uses an approach in terms of truth-conditions.

---

vance are apt to be relegated to the rhetorical rather than the logical dimension of assessment of arguments". This attitude is neatly summarized — if only to submit it to the most important critical revision in logico-philosophical literature so far — by Anderson and Belnap (1975, XXI): „The difficulty of treating relevance with the same degree of mathematical sophistication and exactness characteristic of treatments of extensional logic led many influential philosopher-logicians to believe that it was *impossible* to find a satisfactory treatment of the topic."

APPENDIX A: A FORMAL THEORY OF SENSE FOR PRESCRIPTIONS

a) Language LN

In philosophy of language, the question what entities are to be considered to be true or false (propositions, type-sentences, inscription-sentences, etc.) has been much discussed. In the same way, one can discuss what entities can be said to have the property of being 'effective'. In this book, I will assume that *norms* are the vehicles of efficacy. I will also assume that norms can be represented in an artificial language. Therefore, I will hold that the vehicles of efficacy are a subset of the well-formed expressions of an artificial language *LN*. *LN* is a language of first-order predicate deontic logic.

Primitive symbols:

a) Constants: a, b, c, a', b', c', a'', ...
b) Predicates: F, G, H, F', G', H', F'' ...
c) Connectives: $\neg, \wedge, \vee, \rightarrow$
d) Parenthesis: (,)
e) Variables: x, y, z, x', y', z', x'' ...
f) Quantifiers: $\forall, \exists$
g) Deontic operators: O, Ph, P[18]

The well-formed expressions of *LN* can be characterized recursively as well-formed expressions of a declarative *(d-formulae)*, imperative *(i-formulae)*, or mixed *(m-formulae)* nature (where the latter are actually of an imperative nature too).

Definition of a d-formula:

1) An n-adic predicate letter, followed by n constants is an atomic d-formula.
2) If A is a d-formula, then $\neg(A)$ is a d-formula.
3) If A and B are d-formulae, $(A \wedge B)$ is a d-formula.
4) If A and B are d-formulae, $(A \vee B)$ is a d-formula.
5) If A and B are d-formulae, $(A \rightarrow B)$ is a d-formula.
6) If A is a d-formula, and α is the result of the substitution in A of a constant by a variable, then the expression formed by the universal quantifier ('$\forall$') followed by the variable, followed in turn by α also is a d-formula.

[18] As we know, the operators (to be read as 'obligatory', 'prohibited', and 'permitted'), are interdefinable:

$P = \neg O \neg.$

$P = \neg Ph.$

There is also the operator 'facultative' *(F)* which can be defined with the help of 'permitted' and therefore is not needed here:

$F = P \wedge P\neg.$

7) If A is a d-formula, and α is the result of the substitution in A of a constant by a variable, then the expression formed by the existential quantifier ('$\exists$') followed by the variable, followed in turn by α also is a d-formula.

Definition of an i-formula:

1') If A is an atomic d-formula, then a deontic operator followed by (A) is an atomic i-formula.
2') If A is a i-formula, then $\neg(A)$ is an i-formula.
3') If A and B are i-formulae, $(A \wedge B)$ is an i-formula.
4') If A and B are i-formulae, $(A \vee B)$ is an i-formula.
5') If A is an i-formula, and α is the result of the substitution in A of a constant by a variable, then the expression formed by the universal quantifier ('$\forall$') followed by the variable, followed in turn by α also is an i-formula.
6') If A is an i-formula, and α is the result of the substitution in A of a constant by a variable, then the expression formed by the existential quantifier ('$\exists$') followed by the variable, followed in turn by α also is an i-formula.

Definition of an m-formula:

1'') If A is a d-formula and B is an i-formula, then $(A \rightarrow B)$ is an m-formula.
2'') If A is a d-formula and B is an m-formula, then $(A \rightarrow B)$ is an m-formula.
3'') If A is an m-formula, and α is the result of the substitution in A of a constant by a variable, then the expression formed by the universal quantifier ('$\forall$') followed by the variable, followed in turn by α also is an m-formula.
4'') If A is an m-formula, and α is the result of the substitution in A of a constant by a variable, then the expression formed by the existential quantifier ('$\exists$') followed by the variable, followed in turn by α also is an m-formula.

Comments:

The main objective of language *LN* is to provide an adequate base for the analysis of the sense of imperatives. In contrast to other artificial languages, the construction of *LN* is not aimed at the solution of problems of deontic logic; instead, it is intended to facilitate the investigation of the semantics of directive discourse.

Therefore, it may be helpful to clarify some of the restrictions implied by language *LN*:

(a) Language *LN* allows the use of all the logical rules of a natural deductive calculus for first-order predicate logic, provided the expression obtained is a well-formed expression (a formula) of *LN*.

(b) A conditional with a deontic operator in the antecedent, as in '$O(Fa) \rightarrow O(Ga)$', is not a well-formed expression of *LN*. The only conditionals admitted are mixed expressions the antecedent of which is a d-formula and the consequent of which

is an i-formula or an m-formula. Although syntactically such expressions are mixed, semantically they are of an imperative nature. The restriction arises from pragmatic considerations. In imperative language, it does not seem possible to use conditionals with an imperative antecedent. For instance, the imperative sentence 'If you go to the market, buy oranges!' is a well-formed expression, whereas the sentence 'If, go to the market!, buy oranges!' makes no sense.[19] A normative proposition, e. g., a sentence describing the existence of a norm, however, can be the antecedent of a conditional. For instance, 'If you must go to the market, then buy oranges!' has meaning; but the antecedent must be represented as a d-formula instead of an i-formula.

For the same reason, some transformations authorized by the rules of logic are invalid in *LN*. Thus, in *LN* it is impossible to go from '$(Fa) \rightarrow O(Ga)$' to '$\neg O(Ga) \rightarrow \neg(Fa)$', since the second expression is not a formula of *LN*.

A certain inconvenience of expressions like '$O((Fa) \rightarrow (Ga))$', which is either equivalent to '$(Fa) \rightarrow O(Ga)$' or incomprehensible, is the reason why *LN* authorizes deontic operators only in front of atomic d-formulae.

(c) In *LN*, mixed expressions of a conjunctive or disjunctive nature are not permitted, for similar reasons as those pointed out by Smart (where '*Ap*' is an *assertive* or *declarative* sentence, and '*Ip*' an *imperative* sentence):

„I am inclined to think that mixtures of imperatives occur in colloquial language only in the form Ap → Iq (and perhaps in a rather strained way in the form Ap∨q which comes to the same as A(¬p) → Iq. We do not get Ap∧q though we do get Ap and Iq asserted as separate premisses. I do not think we ever get Ip → Aq, even though this would be a way of saying A(¬q) → I(¬p).“ (Smart 1984, 16)

(d) General norms, e. g., 'Persons of age must vote' are expressed in *LN* through m-formulae and have the following canonical formulation:

$$\text{'}\forall x\,(Fx \rightarrow O(Gx))\text{'}.$$

(e) Note that the predicates of *LN*, when modalized by a deontic operator, must be interpreted as properties attributed to human beings, concerning the performance of certain actions. *LN* is a language that permits quantification only over actors, not over actions (Ziemba 1976, 383 ff.; Makinson 1981, 87-91; Hernández Marín 1984, 113 ff.). For instance, 'The chairs must recite a Shakespeare sonett' is an expression that makes no sense. This is analogous to the situation generated by expressions such as 'Prime numbers are blue' (cf. Grant 1968, 189 f.).

b) Open Sentences

The expressions which in the definitions above were designed by α, i. e., expressions with free variables, we will call *open* d-sentences, i-sentences, or m-sentences, respec-

[19] Not all scholars of deontic logic share this view. Weinberger (1984, 1991), e. g., has insisted that expressions of the form '$p \rightarrow Oq$' are unsatisfactory reconstructions of normative conditionals.

tively. Examples of such open sentences are: '$(Fx \rightarrow Gx)$', '$(O(Gx) \wedge P(Hx))$', or '$(Fy \rightarrow Ph(Hy))$', etc.

Open d-sentences are neither true nor false; rather, they are satisfied or not satisfied by certain objects (or pairs of objects, or triplets of objects, etc.). Thus, if '*F*' is the predicate 'being a philosopher', the open d-sentence '*Fx*' is satisfied by Aristotle, but (perhaps) not by Alexander the Great. If '*G*' is 'being the lover of', the open d-sentence '*Gxy*' is satisfied by the pair <Marcus Antonius, Cleopatra>, but (perhaps) not by the pair <Caesar, Brutus>.

Open sentences are sentential functions (it is easily seen that the so-called open d-sentences are not d-formulae). The notion of *satisfaction* can be explicated informally in the following way. An open d-sentence is satisfied by certain objects if the expression becomes a *true* d-formula on substituting the free variables by the names of those objects. Thus, we can say that the pair <Marcus Antonius, Cleopatra> satisfies the open sentence '*x* is the lover of *y*', since 'Marcus Antonius is the lover of Cleopatra' is a true sentence. But, since — according to Tarski — the definition of truth requires the notion of satisfaction, we cannot (without circularity) use the predicate 'true' for the definition of 'satisfaction'.[20]

In any case, this seems an adequate way of informally introducing the notion of *compliance* for open i-sentences and open m-sentences. An expression like '*x* should be a movie actress' is complied with by Julia Roberts, but not by Margaret Thatcher, and '*x* should kill *y*' is complied with by the pair <Brutus, Caesar>, but not by many other pairs. Therefore, we can sustain that an open i-sentence (or m-sentence) is complied with by some individuals if on substituting the names of those individuals for the free variables that expression becomes an *effective* i-formula (or m-formula). Thus, the pair <Kennedy, Johnson> complies with the expression 'If *x* dies, *y* should succeed him', since 'If Kennedy dies, Johnson should succeed him' is an effective sentence.

In view of this presentation, it should be clear that efficacy is parasitic on truth: An i-formula or m-formula is effective if and only if its corresponding d-formula is true.[21] Let's take the simplest case: that of atomic i-formulae, with expressions such as '$O(Ga)$'. The corresponding d-formula is 'Ga'.

In order not to complicate things unnecessarily, I will choose a subset of the i-formulae and m-formulae of *LN*. Based on them, I will show what their corresponding

20 In the words of Grayling (1982, 160): „Thus, for example, snow (not the name 'snow' but the actual stuff) satisfies 'x is white' because the sentence 'snow is white' is true. However this is a *merely* heuristic way of explaining satisfaction, for 'true' is being *used* here; and because we wish to define 'true', we must seek for an account not involving 'true'."

21 Thus, Ross (1941, 60) writes: „... an imperative I is said to be satisfied, when the corresponding indicative sentence S describing the theme of demand, is true, and non-satisfied, when that sentence is false". Williams (1973, 187 f.; cf. also Hart 1983, 325) formulated a similar suggestion: „... that corresponding to any imperative 'do x', there is an indicative statement, 'x is done', which might be called its 'obedience statement'; and we may say that two imperatives are inconsistent if their obedience statements are inconsistent." A critique of this way of analysis can be found in Kelsen (1979, 163 f.; 173-176).

d-formulae are. For that subset of formulae, the notion of compliance will be defined recursively, and we will thus arrive at a definition of efficacy.

c) Satisfaction and Compliance

i) A subset of LN

Let us take a subset of *LN* in which all predicates of the language are monadic, and let us restrict the connectives to negation and conditional. The negation only precedes atomic formulae, and the conditional links only atomic formulae, or formulae preceded by the negation. Finally, in i-formulae and m-formulae, we will use only the universal quantifier. In this subset, all expressions are of the following form:

d-formulae:	*i-formulae:*
Fa	$O(Fa)$, $Ph(Fa)$, $P(Fa)$
$\neg(Fa)$	$\neg(O(Fa))$, ...
$(Fa \rightarrow \neg Ga)$	$\forall x\,(O(Fx))$
$\forall x\,(Fx \rightarrow \neg Gx)$	$\forall x\,(P(Fx))$
$\exists x\,(Fx \wedge Gx)$	

m-formulae:

$(Fa \rightarrow O(Ga))$	$\forall x\,(Fx \rightarrow O(Gx))$
$(Fa \rightarrow P(Ga))$	$\forall x\,(Fx \rightarrow P(Gx))$

ii) Corresponding d-formulae

I will now show what the d-formulae corresponding to that subset of i-formulae and m-formulae are. The basic intuition is the following: The d-formulae corresponding to i-formulae and m-formulae are true descriptions with respect to some possible worlds. These are deontically perfect worlds with respect to those i-formulae and m-formulae. In those worlds, norms are always effective.

The d-formula corresponding to an atomic i-formula of obligation or permission X_i is obtained by eliminating the deontic operator from X_i. The d-formula corresponding to an atomic i-formula of prohibition X_i is obtained by eliminating the deontic operator and putting the sign of negation in front of X_i.

Let us now look at i-formulae with quantifiers. An i-formula like '$\forall x\,(O(Fx))$' is effective if it is complied with by all its addressees. Therefore, we can say that $\forall x\,(Fx)$ is its corresponding d-formula. One can argue, however, that this is asking too much, and that the compliance of a norm *N* by some majoritarian percentage of its addressees is sufficient for the efficacy of *N*. I will comment on that question later.

Now, when can we say that an i-formula like '$\forall x\,(P(Fx))$' is effective? Following von Wright (1983a, 139), I will use the following convention: A permissive norm is complied with if and only if it is used by its addressees on at least one occasion; in other words, if the state of affairs permitted by the norm is produced at least once. Von Wright asserts that it is rather reasonable to expect that the states of affairs permitted by an authority will obtain on some occasion, adding:

„I shall say that a permissive norm is satisfiable if, and only if, it is possible that the permitted state of affairs obtains at *some* time in the history of the norm. And it is satisfied if, and only if, at some time in its history that which it permits actually is also the case." (von Wright 1983a, 139)

A permission given with a certain condition of application, e. g., 'Smoking is permitted during dinner', is complied with (and therefore effective) if and only if some addressee of that permission smokes during dinner.[22] If one accepts that convention, then the i-formula '$\forall x\ (P(Fx))$' is effective if and only if there is some x that is F. Therefore, we will say that the d-formula corresponding to that i-formula is '$\exists x\ (Fx)$'.[23]

Let us now analyze the negations of atomic i-formulae. According to an old, but controversial intuition of von Wright (1963a, 138), the negation of a norm is a norm just as the negation of a proposition is a proposition. In view of the interdefinability of the operators, the negation of an i-formula *O-* or *Ph-* is an i-formula *P-*, and the negation of an i-formula *P-* is an i-formula *O-* or *Ph-* Thus, the i-formula '$\neg(O(Fa))$' is effective if and only if '$\neg(Fa)$' is true, and this then is the d-formula corresponding to the above i-formula.

An m-formula like '$(Fa \rightarrow O(Ga))$' has as its corresponding d-formula '$(Fa \rightarrow (Ga)$', and the d-formula corresponding to '$\forall x\ (Fx \rightarrow O(Gx))$' is '$\forall x\ (Fx \rightarrow (Gx)$'.[24]

[22] This reconstruction has a counterintuitive consequence. If norms making an action optional (facultative) are defined as the conjunction of the permissions to do and to omit that action, then in this language making an act-individual facultative at a given time would always be ineffective (no-one can do and omit an action at the same time). Of course, this was not the only way of approaching the question of permissions. An alternative way was to assume that, since permissions in principle cannot be complied with, they do not really express prescriptions in the same way as obligations and prohibitions and that they are perhaps expressions of acts of rejecting (previous) prohibitions or acts of rejecting them in advance. This is a way explored by Alchourrón and Bulygin (1981). It would lead to an elimination of our *P*-expressions from language *LN*.

[23] Incidentally, this intuition corresponds to the analogy some deontic logicians (cf. Kalinowski 1975, 40 f.) establish between the square of opposition representing the relations of Aristotelian predicate logic (where universal statements imply their corresponding individual statements through a relation of subalternation) and the square representing the relations between deontic expressions (where expressions of obligation entail permissive expressions). On the origin of that analogy in the work of Bentham, cf. Moreso 1992, 148 f.

[24] This implies, of course, that a norm like 'All those with an income of more than 10 million dollars a year will pay an income tax of 50%' is effective if no-one has an income of more than 10 million dollars a year. This may seem paradoxical if efficacy is linked to the motivation of behaviour: in that case, such a norm apparently does not motivate any behaviour. But that paradoxical note is similar to that produced by conditional statements with a false antecedent, e. g., 'All centaurs are metaphysicists'. Such statements are a convenient way of increasing our set of truths without increasing our knowledge about the world. However, as has been noted (Dummett 1978, 9 f.), when the truth of the antecedent depends on an action that is in the power of the addressee of the imperative, then we cannot say that this kind of conditional imperatives does not motivate behaviour. Suppose a mother orders her son 'If you go out, wear your coat'. We cannot say that if the son does not go out it is as if the imperative had not been issued, since the reason why he does not go out could be, e. g., that he cannot find his coat.

The i-formula '$(Fa \rightarrow P(Ga))$' has as its corresponding d-formula '$(Fa \wedge Ga)$', while '$\forall x\,(Fx \rightarrow P(Gx))$' has the corresponding d-formula '$\exists x\,(Fx \wedge Gx)$'. Note that there is a certain asymmetry between expressions of obligation (or prohibition) and permissive expressions. Conditional formulae of obligation are effective even when the condition of application is absent, whereas conditional formulae of permission are not effective in that case. The analogy between universal sentences vs. singular sentences and expressions of obligation vs. expressions of permission is thus maintained.

In summary, we can now produce a list of some correspondences:

i-formulae	*corresponding d-formulae*
$O(Fa)$	Fa
$Ph(Fa)$	$\neg(Fa)$
$P(Fa)$	Fa
$\neg(O(Fa))$	$\neg(Fa)$
$\neg(Ph(Fa))$	Fa
$\neg(P(Fa))$	$\neg(Fa)$
$\forall x\,(O(Fx))$	$\forall x\,(Fx)$
$\forall x\,(Ph(Fx))$	$\forall x\,\neg(Fx)$
$\forall x\,(P(Fx))$	$\exists x\,(Fx)$
$\forall x\,\neg(O(Fx))$	$\exists x\,\neg(Fx)$
$\forall x\,\neg(Ph(Fx))$	$\exists x\,(Fx)$
$\forall x\,\neg(P(Fx))$	$\forall x\,\neg(Fx)$

m-formulae	*corresponding d-formulae*
$(Fa \rightarrow O(Ga))$	$(Fa \rightarrow (Ga))$
$(Fa \rightarrow P(Ga))$	$(Fa \wedge Ga)$
$\forall x\,(Fx \rightarrow O(Gx))$	$\forall x\,(Fx \rightarrow (Gx))$
$\forall x\,(Fx \rightarrow P(Gx))$	$\exists x\,(Fx \wedge Gx)$[25]

iii) Satisfaction and the truth of d-formulae

Now, let us recall that satisfaction is a *relation* between open sentences and ordered *n-tuples* of objects. According to Tarski, open sentences, like

$$F(x, x', x'', \ldots, x^n)$$

are satisfied by infinite sequences of objects like

$$<O, O', O'', \ldots, O^{n+1}, \ldots>.$$

25 I omit the combination with negations since, because of the equivalences between deontic operators, it follows from the application of the mechanism given for i-formulae.

Since such sentences are satisfied by the first *n* objects of the sequence, the rest of it can be ignored. Thus, we can say that the negation of an open d-sentence *A* is satisfied by all sequences that do not satisfy *A*. Open d-sentences of the form $A \rightarrow B$ are satisfied either if no sequence satisfies *A*, or by those sequences that satisfy *B*.

For any d-formula X_d, all members of any sequence are irrelevant for knowing whether the sequence satisfies X_d. Thus, a true d-formula in *LN* is satisfied by all sequences, and a false d-formula is not satisfied by any sequence. For example, '$\exists x\ (Fx)$' (where *F* is 'being a philosopher') is satisfied by any sequence, e. g. the sequence <Octavio, Augustus, ...>, since the open d-formula resulting from the elimination of the quantifier is '*Fx*', and that sentence is satisfied by some sequences, like <Aristotle, ...>.

Now we can offer a *recursive* or *inductive* definition of satisfaction for our subset of *LN*.[26] I will call '*var (i)*' the i-th variable of the vocabulary of *LN*, and x_i the i-th object of a sequence *X*. If we assume that '*A*' and '*B*' are monadic predicates of *LN*, the definition is as follows:

(1) For any *i* and for any *X*: *X* satisfies '*A*' followed by *var (i)* iff Ax_i.
(2) For any sequence *X* and for any d-formula *A*: *X* satisfies the negation of *A* iff *X* does not satisfy *A*.
(3) For any sequence *X*, for any d-formula *A* and for any d-formula *B*, *X* satisfies '$A \rightarrow B$' iff *X* does not satisfy *A* or *X* satisfies *B*.
(4) For any *X*, for any *A*, and for any *i*: *X* satisfies the universal quantification of *A* with respect to *var (i)* iff *A* is satisfied by any sequence *X'* such that $X_j = X_j'$, for all *j* with $j \neq i$.

With Quine, we can add:

„Taken altogether, the inductive definition tells us what it is for a sentence to satisfy a sentence of the object language. Incidentally it affords a definition also of truth, since, as lately noted, this just means being satisfied by all sequences.“ (Quine 1970, 42)

iv) Compliance and the efficacy of d-formulae and i-formulae

The inductive definition of *satisfaction*, together with the definition of a *d-formula corresponding* to an i-formula and an m-formula, allows us to define the notion of *compliance* as a relation between i-formulae or m-formulae and sequences of objects:

For any i-formula *A* and for any sequence *X*, *X complies* with *A* iff the d-formula *B* corresponding to *A* is satisfied by *X*.
For any m-formula *A* and for any sequence *X*, *X complies* with *A* iff the d-formula *B* corresponding to *A* is satisfied by *X*.

With this, we can now also define *efficacy*:

An i-formula is *effective* iff it is complied with by all sequences or, what amounts to the same, iff its corresponding d-formula is true.

---

[26] Since the above presentation follows Quine (1970), it should be added that proper names (the constants of *LN*) are eliminated in the following way: '*Fa*' is replaced by the expression '$\exists x\ (a = x \wedge Fx)$', where '$a =$' is a predicate. Cf. Quine (1970, 25 f.).

An m-formula is *effective* iff it is complied with by all sequences or, what amounts to the same, iff its corresponding d-formula is true.

Since these definitions of compliance and efficacy differ from the usual ones, the following points should be noted:

(a) Norms are not complied with by their addressees, but by ordered sequences of individuals. The addressees of norms are elements of those sequences.

(b) Compliance with d-formulae or with i-formulae depends on the sequences taken into consideration. Those sequences depend on some ascription of the elements of our language to a certain structure of the world in a given domain or universe (an *interpretation*).[27] Some sequences obviously satisfy certain i-formulae or d-formulae under one interpretation, but not under another.

(c) The idea underlying this formal definition can be better understood through the following examples:

The norm expressed by 'All persons of age should vote' can be symbolized in *LN* by the m-formula

$$(\alpha)\ \forall x\, (Fx \rightarrow O(Gx)).$$

The corresponding d-formula is:

$$(\beta)\ \forall x\, (Fx \rightarrow Gx)$$

(α) is complied with by all the sequences satisfying (β). A world in which no-one is of age, or a world in which all persons of age do vote is a world in which (β) is true and, therefore, (α) is effective.

Similarly, the norm expressed by 'All persons of age may vote' can be symbolized in *LN* by the m-formula

$$(\alpha')\ \forall x\, (Fx \rightarrow P(Gx)).$$

The corresponding d-formula is:

$$(\beta')\ \exists x\, (Fx \wedge Gx)$$

(α') is effective if and only if (β') is true. The permissive norm expressed in (α') is effective only in the case that some of its addresses make use of the permission to vote. If

[27] For an alternative presentation of the notion of satisfaction using the notions of *structure* (as a set of objects referred to by the universal quantifier — a universe — and the denotation of the predicates of the language) and *interpretation* (a function correlating language with structure in such a way that one can say that if an interpretation I *satisfies* some set of formulae *{C}*, then I is a *model* of *{C}*), cf. Enderton (1972, 79-86) and Garrido (1981, 220-230).

no-one votes, then the permission is ineffective. That is the intuitive reason why (α) does not imply (α'), since in standard logic (β) does not imply (β').

(d) I will now analyze the possibility that the formulae corresponding to i-formulae and m-formulae of obligation are not d-formulae preceded by the universal quantifier. It will then be possible to account for the intuition that in order to predicate the efficacy of a norm we do not require that *all* its addressees comply with it *all* the time. I will present two alternative strategies for this.

On the one hand, one can use a *numerically defined quantifier* in the corresponding d-formulae (Quine 1972, 311-316). The idea is as follows: Suppose a norm N is effective if it is complied with by at least 75% of its addressees, and suppose also that N has ten addressees. The i-formula '(α) $\forall x\ (O(Fx))$' is effective if and only if '$\exists^{\geq 8} x\ (Fx)$' is true. This last expression would then be the d-formula corresponding to the previous i-formula.

In this way, d-formulae corresponding to universal i-formulae or m-formulae would be existential. And, what's more, all corresponding d-formulae in *LN* would be existential.

Now, as Quine has pointed out, this strategy fails when the number is indefinite, i. e., when the addressees of a norm are indefinite, and that obviously is what happens with many kinds of norms.

On the other hand, we can try to use so-called non-standard quantifiers (cf. Platts 1979, 100-106). Platts introduces the possibility of working with at least two quantifiers between 'all' and 'some', namely, 'most' and 'many'. This means that besides sentences like 'All philosophers are boring' and 'Some philosophers are boring', we could have 'Most philosophers are boring' and 'Many philosophers are boring'. I will present Platts's ideas about the 'most' quantifier — which I will call the 'majoritarian quantifier' ('*M*') — with the underlying intuition that the declarative statement corresponding to 'All citizens should vote' could be 'Most citizens vote'.

This idea will enable us to introduce a new clause in the definition of satisfaction:

(6) For any X, for any A and for any i: X satisfies the majoritarian quantification of A with respect to *var (i)* iff A is satisfied by most sequences X' such that $X_j = X_j'$, for all j with $j \neq i$.

Now, we need a canonical formulation for the kind of statements with the majoritarian quantifier. In principle, we have two options. We can either make it similar to a statement with the existential quantifier, or make it similar to a statement with the universal quantifier:

(∂) $Mx\ (Fx \wedge Gx)$

or

(γ) $Mx\ (Fx \rightarrow Gx)$

(∂) is clearly a bad candidate, since it does not mean that most philosophers are boring, but that most things are at the same time philosophers and boring. But (γ) does not seem to work any better, since because of the equivalence of '$Fx \rightarrow Gx$' and '$\neg Fx \vee Gx$' statement (γ) would have to be read as 'most things either are not philosophers or are boring'. Therefore, given the truth of the first disjunct, any statement beginning with 'Most philosophers ...' would be true. This too is an implausible conclusion, because it generates truth-conditions for the majoritarian quantifier that are very far away from our basic intuitions about its use in language.[28]

We must therefore give up our attempt to represent statements with the majoritarian quantifier through the use of connectives. This kind of sentences does not include connectives; they are rather *relational* sentences (similar to 'The moon is far from the sun'), but establishing a second-order relation. They do not establish a relation between individuals, but between predicates. The canonical formulation of 'Most philosophers are boring' thus could be:

(δ) $Mx\,(Fx, Gx)$,

which could be read as 'Most individuals satisfying predicate F also satisfy predicate G'. Therefore, the clause that should be added to our definition of satisfaction is another one:

(6') For any X, for any A and for any i: X satisfies the majoritarian quantification of A — say, '$Mx\,(Fx, Gx)$' — with respect to *var (i)* iff most individuals satisfying 'Fx' also satisfy 'Gx'.

This seems to give an adequate account of sentences with the majoritarian quantifier, but the formulation is too far from that of the classical quantifiers, and it makes logical relations between those quantifiers impossible. Therefore, it also makes impossible logical relations between the corresponding d-formulae which reflect the relationships between imperative statements.

[28] Perhaps an idea of von Wright (1984b, 46) about the possibility of restricting the scope of quantifiers could help to avoid this difficulty. But I will not analyze that possibility here.

APPENDIX B: RELEVANCE, ARGUMENTATION AND NORMATIVE SYSTEMS

a) Relevance Beyond Logic

In several relatively recent studies on pragmatics, the notion of relevance plays a central role. They follow up on a suggestion by Grice (1975, 46) who, when speaking about the general conditions of conversation, mentions as one of them (as the category of *relation*) the maxim 'Be relevant'. The question is what one must do in order to be relevant. If we want to answer that question we need an adequate conceptual reconstruction of the notion of relevance.

The work of Sperber and Wilson is one attempt in that direction. On the one hand, they assert that they „implicitly assumed that the process of inferential comprehension is non-demonstrative" (Sperber/Wilson 1986a, 65). On the other, they hold that „utterance comprehension involves a substantial inferential element" (Sperber/Wilson 1986b, 244). In their view, that inferential element has to do with logic, but

„there is good reason to think that the logic used in utterance comprehension is not a standard one. On the one hand, it must be much more extensive, providing rules for every concept that can play a role in the inferential processing of propositions, including many that are of no particular interest to logicians. On the other hand, it must be more restrictive in certain ways" (Sperber/Wilson 1986b, 247).

Here, I am interested in how these authors think logic should be restricted in order to account for relevance. To this end, they propose that one should accept only those rules of logic (of propositional logic, we must assume) that never lead to 'irrelevant' — in the sense of 'trivial' — conclusions.

They assume that the restriction should apply to systems of natural deduction within propositional logic. Such systems contain rules of introduction and of elimination for each connective.[29] The proposed restriction is that all rules are elimination rules (1986b, 248). They thus construct the notion of *non-trivial logical implication*:

---

[29] For a standard system of natural deduction like that of Gentzen (1934, 176-210), these rules are:

Rule for conjunction introduction ($\wedge I$):
$A, B \vdash A \wedge B$ and also $B \wedge A$
Rule for conjunction elimination ($\wedge E$):
$A \wedge B \vdash A$ and also B
Rule for disjunction introduction ($\vee I$):
$A \vdash A \vee B$ and also $B \vee A$
Rule for disjunction elimination ($\vee E$):
$A \vee B, A \ldots C, B \ldots C \vdash C$
Rule for conditional introduction (or *deduction theorem*) ($\rightarrow I$):
$A \ldots B \vdash A \rightarrow B$
Rule for conditional elimination (or *modus ponens*) ($\rightarrow E$):
$A \rightarrow B, A \vdash B$
Rule for negation introduction (or *reductio ad absurdum*) ($\neg I$):
$A \ldots B \wedge \neg B \vdash \neg A$
Rule for negation elimination ($\neg E$):

„A set of assumptions {P} *logically and non-trivially implies* an assumption Q if and only if, when {P} is the set of initial theses in a derivation involving only elimination rules, Q belongs to the set of final theses.“ (Sperber/Wilson 1986a, 97)

They then proceed to analyze under what conditions adding a new piece of information *{P}* to a given context *{C}* — a set of statements representing a store of information — allows one to gain *contextual implications*. Thus, the notion of contextual implication is defined in terms of the contextualization of *{P}* in *{C}* — i. e., of the union-set of *{P}* and *{C}*:

Contextual implication (Sperber/Wilson 1986a, 107)

„A set of assumptions {P} *contextually implies* an assumption Q in context {C} if and only if
(i) the union of {P} and {C} non-trivially implies Q,
(ii) {P} does not non-trivially imply Q, and
(iii) {C} does not non-trivially imply Q.“

Their approach to relevance, then, is bound up with contextual implication: The more contextual implications a proposition has in a given context, the more relevant it is (Sperber/Wilson 1986b, 249):

Relevance (Sperber/Wilson 1986a, 122)

„An assumption is relevant in a context if and only if it has some contextual effect in that context.“

Thus, we can say that the *relevant* consequences of an argument *A* are those that are contextual implications of some set of premises *{P}*. The advantage of this notion is that some of the consequences we had regarded as paradoxical no longer apply. From

$$(1)\ p \rightarrow q,$$
$$(2)\ p$$

we can now deduce neither '$q \vee r$' (since we have no rule like that of disjunction introduction) nor '$r \rightarrow q$' (since we have no rule for conditional introduction), nor many other conclusions which we intuitively regard as irrelevant.

The problems with this conception of relevance are, on the one hand, that the logical system these ideas give rise to is not rigourously spelled out (Sperber and Wilson could counter this[30] by saying that they want to offer a pragmatic characterization of

---

$$\neg\neg A \vdash A$$

Expressions of the kind '@ ... #' should be read as '# is obtained from @ (by rules of logic)'.

[30] Instead, they argue that their system is somewhat similar to Parry's system of analytic implication, as presented by Anderson/Belnap 1975, 430-432; but although *PAI* (Parry's analytic implication) does not accept in the system a rule that would enable one to go from 'p' to '$p \vee q$' (i. e., there is no rule for disjunction introduction, which is the cause of so many of the irrelevant conclusions we draw), it does have an explicit

the notion of relevance, rather than try to construct an alternative to the standard notion of logical consequence), and, on the other, that a system without introduction rules loses its inferential capacity.

Sperber and Wilson (1986b, 258 and n. 7; 1986a, 99-101) mention two aspects of their notion that may be problematic:

a) Since there are no introduction rules, the following inference is impossible (cf. n. 29):

(1) p
(2) q

$\vdash p \wedge q$

This is a case of the rule of conjunction introduction. Sperber and Wilson argue that the deductive capacity of human beings is imperfect, and that there are *valid* arguments they cannot directly grasp. Allegedly, only by showing that we cannot accept the premises and at the same time deny the conclusion can the validity of the above argument be shown. This assertion looks very much like an *ad hoc* hypothesis: It is implausible that the only way for us to accept the argument is by showing that it is inconsistent to assert at the same time the truth of the negation of the conclusion and the truth of the premises (this is a kind of a rule of *reductio ad absurdum* which, besides, in the calculi of natural deduction is also an introduction rule).

b) They also wonder how the following argument, which they regard as valid, can be represented in their terms:

(1) If the trains are on strike and the car has broken down, there is no way of getting to work.
(2) The trains are on strike.
(3) The car has broken down.

Therefore, there is no way of getting to work.

A standard system of logic would formalize and solve the case as follows:

(1) $(p \wedge q) \rightarrow r$
(2) p
(3) q

2, 3	*(4)* $p \wedge q$	$\wedge I$ 2, 3
1, 2, 3	*(5)* r	$\rightarrow E$ (MP) 1, 4

---

rule for conjunction introduction and thus violates Sperber/Wilson's criterion for non-trivial implication. For them, in contrast to *PAI*, one cannot obtain '$p \wedge q$' from 'p, q'.

But step (4) involves the use of an introduction rule. Therefore, they say that we should postulate another rule: the rule of *conjunctive modus ponens (CMP)*

$$(A \wedge B) \rightarrow C, A \vdash B \rightarrow C,$$

with which we can now show the validity of the argument as follows:

(1) $(p \wedge q) \rightarrow r$
(2) p
(3) q

1, 2	(4) $q \rightarrow r$	*CPM* 1, 2
1, 2, 3	(5) r	$\rightarrow E$ (MP) 3, 4

However, the rule of *conjunctive modus ponens* is nothing but a rule derived from any standard system of natural deduction presupposing the use of conjunction introduction. Actually, it is an *ad hoc* elimination rule which only apparently succeeds in renouncing the use of an introduction rule.[31]

Nor do their reflections on the psychological plausibility of those rules seem very convincing, since we do not seem to have much experimental evidence of the psychological accessibility of *conjunctive modus ponens*, which is accepted by the authors, whereas we do have evidence of the accessibility of the rule for conjunction introduction, which the authors reject (cf. Sperber/Wilson 1968a, 100 f.).

Besides these problems, the inferential capacity of a system of natural deduction that renounces the use of introduction rules is very limited. Without the rule for conditional introduction — the deduction theorem — the validity of the following argument (from deontic logic) cannot be shown:

(1) If a person is more than eighteen years old, she is of age.
(2) If a person is of age, then she has the obligation to vote.

Therefore, if a person is over eighteen years old, she has the obligation to vote.

[31] Most calculi of natural deduction would show that *conjunctive modus ponens* is a rule derived as follows:

(1) $(A \wedge B) \rightarrow C$
(2) A

3	(3) B	Assumption
2, 3	(4) $A \wedge B$	$\wedge I$ 2, 3
1, 2, 3	(5) C	$\rightarrow E$ (MP) 1, 4
1, 2	(6) $B \rightarrow C$	$\rightarrow I$ 3, 6

This proof uses two introduction rules: the rule for conjunction introduction (step 2), and the rule for conditional introduction (steps 3 and 6).
Something similar applies to what they call *disjunctive modus ponens*: $(A \vee B) \rightarrow C, A \vdash C$. This is a derived rule that presupposes the use of the rule for disjunction introduction.

In standard (deontic) logic, the validity of that argument can be shown as follows:

	(1) $p \rightarrow q$	
	(2) $q \rightarrow Or$	
3	*(3)* p	Assumption
1, 3	*(4)* q	$\rightarrow E$ (MP) 1, 3
1, 2, 3	*(5)* Or	$\rightarrow E$ (MP) 2, 4
1, 2	*(6)* $p \rightarrow Or$	$\rightarrow I$ 3,6

One could say that this deduction can also be done using another rule (the transitivity of the conditional), but again this would only apparently solve our problem, since that is a derived rule and presupposes the more basic rule for conditional introduction.

In summary, Sperber and Wilson's notion of relevance succeeds in avoiding irrelevant consequences only at the forbiddingly high price of a loss of inferential capacity, to the point of obstructing the way to consequences that are clearly relevant. In the previous example, a judge may need to show that the norm '$p \rightarrow Or$' is a norm derived from the system, and if we have no introduction rules, we cannot logically justify his decision.

b) The Logic of Relevance

That an argument is logically valid means that if the premises are true then the conclusion will also be true. This requires that the set of premises *{P} materially implies* conclusion *C*. Now, the notion of material implication has occupied a great number of logicians, because it gives rise to certain consequences regarded as paradoxical. Thus, Hunter (1993, 281) recently remarked:

„What distinguishes a decent logic of conditionals from the logic of material implication is the requirement that the antecedent of a conditional should be *relevant* to its consequent. Standard truth-functional logic does not meet that requirement ... A logic of conditionals must be a *relevance* logic of some kind."

In fact, the development of modern modal logic received its impulse from C. I. Lewis's dissatisfaction (1912, 1914, 1918) with the notion of material implication as used by Frege in his *Begriffsschrift* (1879) and by Russell and Whitehead in *Principia Mathematica* (1910). Lewis observes that this notion enables one to show that the following 'paradoxical' formulae taken from sentential or propositional logic are valid theorems:

$$p \rightarrow (q \rightarrow p)$$
$$\neg p \rightarrow (p \rightarrow q)$$
$$(p \rightarrow q) \vee (q \rightarrow p)$$ [32]

[32] See the following proof (without premises) of the first one of these theorems. The reader can easily prove the other two (although in the third the use of derived rules is suggested: De Morgan's laws — which inter-

In the first case, the paradoxical nature lies in that if a statement is true then it is implied by any other statement; in the second, in that if a statement is false then it implies any other statement; and in the third, in that given any two statements, either the first implies the second, or the second implies the first.

Lewis proposes to avoid these paradoxical results with his notion of *strict* implication, where for 'p' strictly to imply 'q' it is not required that either 'p' is false or 'q' is true (as with material implication), but instead that 'p' cannot be true and 'q' false. Thus, „a strict implication is a material implication which is necessary" (von Wright 1957, 170).

To this end, Lewis introduces a new symbol for strict implication (which I will represent by '-<') and defines it with the help of the modal notion of necessity (where 'it is necessary that' is represented by 'L'):

$$p \text{ -}< \;=_{df} L(p \rightarrow q)$$ [33]

Lewis's concept, however, has been unable to avoid paradoxes altogether, because they could be reintroduced as paradoxes of strict implication. Thus, the following formulae are theorems of the various modal systems proposed by Lewis:

$$Lp \rightarrow (q \text{ -}< p)$$
$$L\neg p \rightarrow (p \text{ -}< q).$$

That means that a necessary statement is strictly implied by any statement, and an impossible statement strictly implies any statement. Lewis upheld these consequences and thought that those paradoxes are truths about the relation of deduction (cf. Haack 1978, 197 f.). For the second theorem, he proposed the following proof (a proof that consists in showing how an arbitrary conclusion can be deduced from an impossible premise: *ex falso quodlibet*):

	1) $p \wedge \neg p$	(impossible premise)
1	*2)* p	$\wedge E$ 1

---

define conjunction and disjunction — and the rules interdefining the conditional with conjunction and with disjunction):

1	*(1)* p	Assumption
2	*(2)* q	Assumption
1, 2	*(3)* $p \wedge q$	$\wedge I$ 1, 2
1, 2	*(4)* p	$\wedge E$ 3
1	*(5)* $q \rightarrow p$	$\rightarrow I$ 2,4
-	*(6)* $p \rightarrow (q \rightarrow p)$	$\rightarrow I$ 1,5

[33] If we add the modal operator of possibility (represented by M) and define it as $Mp =df \neg L\neg p$, that definition is equivalent to:

$$p \text{ -}< q =df \neg M(p \wedge \neg q).$$

1	*3)* $p \vee q$	$\vee I$ 2
1	*4)* $\neg p$	$\wedge E$ 1
1	*5)* q	*MTP* 3, 4[34]

If one accepts this form of argument as valid, then that means that totally arbitrary (and, in that sense, irrelevant) conclusions can be deduced from a set of contradictory (impossible) premises.

Therefore, for some logicians the relation of entailment expressed by the notion of strict implication is insufficient. For them, there must be some connection of meaning between the premises and the conclusion in order to say that the former imply the latter. The analysis of this idea has been the task of so-called *relevance logic*, in which the work by Anderson and Belnap (1975) stands out.[35]

My intention here is not to present Anderson and Belnap's conception or their ambitious logico-mathematical exposition in great detail. For my purpose, it is enough to sketch their main ideas, partly in an informal manner.

The logic of relevance is designed to guard against the paradoxes of both material and strict implication. According to Anderson/Belnap, for '$A \rightarrow B$' to be true (in the sense that A entails B), A must be *relevant* for obtaining B. That means, on the one hand, that it must be possible to use A in the deduction of B from A (Anderson/Belnap 1975, 18 and 30 f., construct a formal technique to this effect), and, on the other, that A and B must share some meaning content:

---

[34] This is, in fact, an example of the so-called disjunctive syllogism or *modus tollendo ponens* [$A \vee B$, $\neg A \vdash B$] which, in Gentzen's system of natural deduction presented in n. 29, is a rule derived from the rule for disjunction elimination and some other rules:

1) $A \vee B$
2) $\neg A$

3	*3)* A	Assumption
2, 3	*4)* $A \wedge \neg A$	$\wedge I$ 2, 3
5	*5)* $\neg B$	Assumption
2, 3, 5	*6)* $(A \wedge \neg A) \wedge \neg B$	$\wedge I$ 4, 5
2, 3, 5	*7)* $A \wedge \neg A$	$\wedge E$ 6
2, 3	*8)* $\neg\neg B$	$\neg I$ 5, 7
2, 3	*9)* B	$\neg E$ 8
10	*10)* B	Assumption
11	*11)* $\neg B$	Assumption
10, 11	*12)* $B \wedge \neg B$	$\wedge I$ 10, 11
10	*13)* $\neg\neg B$	$\neg I$ 11, 12
10	*14)* B	$\neg E$ 13
1, 2	*15)* B	$\vee E$ 1, 3-9, 10-14

[35] This work, which was co-ordinated by Anderson and Belnap but to which other prominent logicians also made contributions, was to appear in two volumes. Following Anderson's death in late 1973, the publication of the second volume was delayed until 1992.

„The second formal condition above is suggested by the consideration that informal discussions of implication or entailment have frequently demanded 'relevance' of A to B as a necessary condition for the truth of $A \rightarrow B$, where relevance is now construed as involving some 'meaning content' common to both A and B ...
A formal condition for 'common meaning content' becomes almost obvious once we note that commonality of meaning in propositional logic is carried by commonality of propositional variables. So we propose as a *necessary*, but by no means sufficient condition for the relevance of A to B in the pure calculus of entailment that A and B must share a variable." (Anderson/Belnap 1975, 32 f.)

Based on these ideas, they construct an axiomatic systems for relevant implication ($R\rightarrow$) (Anderson/Belnap 1975, 20):

$R1.\ A \rightarrow A$
$R2.\ (A \rightarrow B) \rightarrow ((C \rightarrow A) \rightarrow (C \rightarrow B))$
$R3.\ (A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))$
$R4.\ (A \rightarrow (A \rightarrow B) \rightarrow (A \rightarrow B)$

To this, they add other axioms concerning the other connectives, which allow them to obtain the complete system ($E\rightarrow$). Note especially that the complete system contains the following two axioms (8 and 9) about disjunction (Anderson/Belnap 1975, 231 f.):

$(E8)\ A \rightarrow (A \vee B)$
$(E9)\ B \rightarrow (A \vee B)$

I will now try to show the problems this conception raises with respect to the exclusion of the 'paradoxical' logical consequences presented earlier:

i) First, the logic of relevance is unable to eliminate one of the most paradoxical factors about relevance. In that logic, disjunction can be introduced, since '$A \rightarrow (A \vee B)$' is an axiom. Thus, with the addition of the deontic operators to the language of the logic of relevance, the following is a *valid* inference (in the sense in which Anderson and Belnap redefine the notion of logical validity):

$(1)\ p \rightarrow Oq$
$(2)\ p$

$\vdash O(q \vee r)$[36]

That this is a valid argument in the logic of relevance implies that that logic cannot rid us of the problems burdening the logical conception of the justification of judicial deci-

[36] Since from *(1)* and *(2)*, we can deduce 'Oq', and from 'Oq' we can deduce $O(q \vee r)$. This is one possible formulation of what is called Ross's Paradox (Ross 1941). One reason for Ross's mistrust in deontic logic was precisely the possibility to deduce the norm 'Mail the letter or burn it' from the premise 'Mail the letter'. Ross's Paradox, as is known, has given rise to an extensive debate in deontic logic.