Logique & Analyse **181** (2003), 7–12

THE CONJUNCTION FALLACY

JACK COPELAND AND DIANE PROUDFOOT

> My own view about vagueness is that a meaning assignment to a λ-categorial language represents a possible way of making a vague sentence precise. Because of the indeterminacy of translation [Quine 1960: 27] there will be many different ways of doing this. Lakoff [1973] and others have proposed 'fuzzy logic' as the answer. I am adopting a 'wait and see' attitude on this. I suspect it will not be the whole answer. For myself I prefer a more holistic view of vagueness, for I think that what we can do by way of making the language precise in one area may well depend on what compensatory adjustments we are prepared to make elsewhere. It may be that English turns out to be best captured by a 'fuzzy set' of value assignments.
>
> Max Cresswell (1985: 36)

We too adopt a wait-and-see attitude concerning the overall utility of fuzzy logic in the semantics of natural language. We believe that fuzzy semantics not only has a role to play in the analysis of vagueness and indeterminacy, but also offers insight elsewhere in the investigation of natural language and of natural reasoning. In this note we add a stitch to the fabric of fuzzy semantics. We discuss psychological experiments alleged to demonstrate an almost universal tendency to commit a fundamental logical fallacy, the 'conjunction fallacy'. Using concepts from fuzzy semantics we outline a new approach to natural language conjunction. If the relevant competence of speakers involves (what we call) *non-minimizing* conjunction then in fact no fallacy is committed.

1. *Violating the Conjunction Rule?*

In probability theory the *conjunction rule* states that a conjunction cannot be more probable than one of its conjuncts. That is, where P is the probability measure and A, B are any statements:

$$P(A \,\&\, B) \leq P(A) \quad \text{and} \quad P(A \,\&\, B) \leq P(B).$$

The *conjunction fallacy* is the violation of the conjunction rule (Tversky and Kahneman 1983).

In a famous series of experiments designed to assess natural probability judgements, subjects were asked to assign probabilities (or frequencies) to a conjunction and to its conjuncts (care being taken to eliminate deviant interpretations of the conjunction and its conjuncts, and of the meaning of 'probability'). The subject-matters and formats of the tests, and the degree of statistical sophistication of the (more than 3,000) subjects, varied. One notorious example (tested on undergraduates at the University of British Columbia) is as follows. The subjects were told

> Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. (*ibid.*, p. 297)

and were asked which of the following is more probable:

> Linda is a bank teller.
>
> Linda is a bank teller and is active in the feminist movement. (*ibid.*, p. 299)

85% of subjects rated the conjunction more probable than the conjunct, 'in a flagrant violation of the conjunction rule' (*ibid.*, p. 299). This result is robust across different tests, despite 'increasingly desperate manipulations designed to induce subjects to obey the conjunction rule' (*ibid.*, p. 299).[1]

The 'Linda' experiment is designed to elicit the conjunction fallacy by exploiting the differing degrees of 'representativeness' (of a stereotypical Linda) of the conjunction and the conjunct (when Linda was 'identified merely as a "31-year-old woman", almost all respondents obeyed the conjunction rule' (*ibid.*, p. 305)). The same effect is seen when the conjunction expresses a causal or motivational connection. Tversky and Kahneman diagnose the 'stubborn failures' (*ibid.*, p. 297) to conform to the conjunction rule as follows: ordinary probability judgements are conflated with, or biased by, other judgements — of representativeness and causal connection — that the tasks naturally elicit. The inability, given concrete examples, to see the conjunction rule as decisive is, they suggest, analogous to a child's inability,

[1] Conformity to the conjunction rule increases in gambling tasks, 'which promot[e] extensional reasoning by emphasizing the conditions under which the bets will pay off' (*ibid.*, p. 304). In a gambling situation, the relevant outcomes are treated bivalently.

at a certain developmental stage, to see the conservation of volume rule as decisive (*ibid.*, p. 300). Tversky and Kahneman's work has, it is claimed, 'bleak implications for human rationality' (Nisbett & Borgida 1975, p. 935).

In our view, it is far from the case that Tversky, Kahneman, and their supporters have established that the judgements elicited in the 'Linda' experiment are instances of the conjunction fallacy. Indeed these judgements are not necessarily fallacious.

2. *Fuzzy Logic*

Fuzzy logic aims to extend ordinary deductive methods to situations in which the information available may be only partly true. In fuzzy logic each statement has some numerical *degree* of truth or falsity. These degrees are the truth-values of the logic, and for full generality continuum-many values are usually permitted. Every statement under consideration is taken to have a value lying in the closed interval $[0, 1]$ of the real numbers ('closed' meaning simply that the end-points 0 and 1 are included in the interval). 0 represents complete or determinate falsehood, 1 complete or determinate truth.[2]

The three-valued approach (True, False, and Neither), with its one non-classical value, has little to recommend it by comparison. With only three values to hand, all statements that are neither completely true nor completely false have to be thrown indiscriminately into the Neither box. This eliminates much significant information, for example that A is truer than B, or that C is nearly but not completely true, or that drawing a valid inference will produce a conclusion that is no less true than the premiss. The three-valued approach cannot generate a helpful extension of the classical theory of inference, since the effect of assigning a statement the catch-all value Neither is simply to prevent us from usefully employing it as a premiss in inference. It is of no practical assistance to be told that if one or the other premiss of an application of modus ponens has the value Neither, then the conclusion is either True or Neither (as in Łukasiewicz 1920: 88 and 1930: 166).

The standard truth conditions in fuzzy semantics for propositional compounds are as follows.

(∼) $v(\sim A) = 1 - v(A)$
($-$ is subtraction)

(&) $v(A \mathbin{\&} B) = min(v(A), v(B))$

[2] (Copeland 1997) defends fuzzy logic from a number of popular but misguided objections, for example that fuzzy logic is encumbered by 'excess precision', and that fuzzy logic is unable to cope with 'higher-order' vagueness.

($min(i,j)$ is the smaller of the two numbers i and j)

($\vee$) $v(A \vee B) = max(v(A), v(B))$
($max(i,j)$ is the larger of i and j)

($\rightarrow$) $v(A \rightarrow B) = 1$ if $v(B) \geq v(A)$; $v(A \rightarrow B) = 1 - (v(A) - v(B))$ if $v(B) < v(A)$.

The above semantics has its origins in (Łukasiewicz 1922), where Łukasiewicz introduced the now familiar plethora of values while retaining what are in effect the classical rules for evaluating compounds. Łukasiewicz interpreted the values of his calculus as degrees of probability (1922: 130). Zadeh took over Łukasiewicz's formal apparatus and interpreted the values of the calculus as degrees of truth (Zadeh 1975). Is there any reason, however, to expect that Łukasiewicz's rules — and in particular the conjunction rule (&) — transfer satisfactorily from a domain of degrees of probability to one of degrees of truth?

3. *Non-Minimizing Conjunction*

Williamson (1994: 115) shows that (&) follows from three principles which 'seem plausible':

($\&_1$) $v(A) \leq v(A \,\&\, A)$

($\&_2$) $v(A \,\&\, B) \leq v(A)$ and $v(A \,\&\, B) \leq v(B)$

($\&_3$) If $v(A') \leq v(A)$ and $v(B') \leq v(B)$ then $v(A' \,\&\, B') \leq v(A \,\&\, B)$.

We reject ($\&_2$). Williamson says in support of ($\&_2$): 'if $[A]$ fails to be true to a certain degree, $[A \,\&\, B]$ also fails to be true to that degree, for the latter claim merely adds something to the former' (*ibid.*). We suggest, on the contrary, that when a new piece of information is added to a totality, the new totality may possess an overall degree of truth that is higher than that of the old. Suppose that $v(A) = 0$ and $v(B) = 1$. It seems harsh to say that the totality of information A, B is completely lacking in truth. If the totality is put forward as a representation of the world, there is *some* truth to what is asserted. Part of the representation is completely true, so the representation as a whole is partly true. And there seems no relevant difference between the joint assertion of the individual statements A and B and the assertion of their conjunction. Furthermore, since we advocate an account of entailment according to which a valid entailment preserves (or nearly preserves

(Copeland 1997)) the overall degree of truth of the premisses, the entailment $A, B \vdash (A \,\&\, B)$, which we endorse, guarantees that the conjunction is partly true if the premisses jointly are so. Thus we replace ($\&_2$) by the inequality

($\&_{\geq}$) $v(A \,\&\, B) \geq min(v(A), v(B))$.[3]

Non-minimizing conjunction operators are those that satisfy ($\&_{\geq}$).

This approach invites formal development, which we shall not pursue here. The important point is that subjects who endorse an instance of ($\&_{\geq}$) do not thereby evidence a departure from rationality. A subject who (given appropriate background information) agrees with the statement 'Linda is more likely to be a feminist bank teller than she is likely to be a bank teller' (Tversky and Kahneman 1983: 299) may indeed be behaving rationally if they mean to indicate that (given the background information) the degree of truth of 'Linda is a feminist bank teller' is more likely than not to exceed the degree of truth of 'Linda is a bank teller'.

Department of Philosophy
University of Canterbury
Private Bag 4800
Christchurch
New Zealand
E-mail: jack.copeland@canterbury.ac.nz
diane.proudfoot@canterbury.ac.nz

REFERENCES

Borkowski, L. (ed.) 1970 *Jan Łukasiewicz: Selected Works*, Amsterdam: North Holland.

[3] Edgington (1996) also questions (&) in the setting of fuzzy logic but proposes an account according to which

$$v(A \,\&\, B) \leq min(v(A), v(B)). \text{ (p. 306)}$$

Nevertheless, her small-red-ball example is instructive in the present connection.

> Suppose: $v(Ra) = 1$, $v(Sa) = 0.5$, [and] $v(Rb) = 0.5$, $v(Sb) = 0.5$ ... According to (&), $v(Ra \,\&\, Sa) = v(Rb \,\&\, Sb) = 0.5$. But it is plausible that a is a better case for 'red and small' than b: both are borderline in size, and a is clearly red while b is not. 'Bring me a ball which is red and small; if you can't find a clear case, bring the closest you can find.' Would not a — perfectly red and arguably small — be a better choice than b? (*ibid.*, p. 304)

Copeland, B.J. 1997 'Vague Identity and Fuzzy Logic', *Journal of Philosophy*, 94, 514–534.
Cresswell, M.J. 1985 *Adverbial Modification: Interval Semantics and Its Rivals*, Dordrecht: Reidel.
Edgington, D. 1996 'Vagueness By Degrees', in Keefe and Smith (1996).
Keefe, R., Smith, P. 1996 *Vagueness: A Reader*, Cambridge, Mass.: MIT Press.
Lakoff, G. 1973 'Hedges: a Study in Meaning Criteria and the Logic of Fuzzy Concepts', *Journal of Philosophical Logic*, 2, 458–506.
Łukasiewicz, J. 1920 'On Three-Valued Logic', in Borkowski (1970).
Łukasiewicz, J. 1922 'A Numerical Interpretation of the Theory of Propositions', in Borkowski (1970).
Łukasiewicz, J. 1930 'Philosophical Remarks on Many-Valued Systems', in Borkowski (1970).
Nisbett, R.E., Borgida, E. 1975 'Attribution and the Psychology of Prediction', *Journal of Personality and Social Psychology*, 32, 932–943.
Quine, W.V.O. 1960 *Word and Object*, Cambridge, Mass.: MIT Press.
Tversky, A., Kahneman, D. 1983 'Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgement', *Psychological Review*, 90, 293–315.
Williamson, T. 1994 *Vagueness*, Routledge: London.
Zadeh, L.A. 1975 'Fuzzy Logic and Approximate Reasoning', *Synthese*, 30, 407–28.