Contradiction, Rationality, and Belief Change

Graham Priest

June 16, 2004

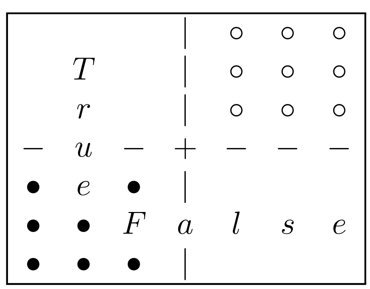
- *Main Thesis of Talk*: It may be perfectly rational to fully believe a contradiction.

1 Contradiction

A contradiction is something of the form $A\&\neg A$, where negation has the truth conditions:

$\neg A$ is true if and only if A is false

The relationship between true sentences and false sentences can be depicted as follows:



$\circ$ = *neither true nor false*. $\bullet$ = *both true and false*. If A is both true and false A and $\neg A$ are both true. According to classical logic, the quadrants marked $\circ$ and $\bullet$ are both empty, but this is not entirely obvious. Possible examples of *neither*: future contingents ('there will be a sea battle tomorrow'). Possible examples of *both*: paradoxes of self-reference ('this sentence is false').

2 Rationality and Contradiction

Two major arguments as to why it is irrational to believe contradictions:

1. Contradictions entail everything (*Explosion*: $A \& \neg A \vdash B$). Rational belief is closed under entailment. It is not rational to believe everything.

2. Contradictions cannot be true (*Law of Non-Contradiction, LNC*). This is obviously true. It is irrational to believe something that is obviously not true.

Re 1: It is not obvious that rational belief is closed under entailment. More importantly, Explosion is valid only if it is impossible for something to be both true and false. (An inference is valid if it is possible that its premises are true and its conclusion is not true. If it is possible for $A \& \neg A$ to be true, it is possible for $A \& \neg A$ to be true and B to be false.) Hence this argument collapses into Argument 2.

Re 2: There are *prima facie* counter examples. Why should one suppose the LNC to hold? There is only one sustained defence of this in Western Philosophy: Aristotle's in *Metaphysics*, Γ. There is one major argument in this, it is tangled and contorted. It is hard to see that it works; and even if it does, it hardly makes the LNC *obvious*. The other arguments are short and aim at establishing that not *all* contradictions are true (or even that one cannot *believe* that all contradictions are true). This is quite compatible with *some* contradictions being true.

Hence, there seem to be no good reasons as to why it is irrational to believe a contradiction.

3 A Model of Rationality

Even if consistency is a constraint on rationality, it is a relatively weak one. (E.g., the belief that the earth is flat can be held quite consistently.) There must be other constraints. Traditional epistemology identifies a number of criteria which speak in favour of the rational acceptability of a set of beliefs (whilst their opposites speak against). These include:

- adequacy to the data
- simplicity
- unity (few *ad hoc* hypotheses)
- explanatory power
- parsimony (no unnecessary entities)
- consistency

All of these may come by degrees; none is an absolute constraint (even consistency as I argued); and some may be more important than others. But most importantly, the criteria may pull in different directions. E.g., Copernican astronomy was simpler than Ptolemaic, but it required many more *ad hoc* hypotheses to cover the fact that it was dynamically impossible, given te received (Aristotelian) dynamics.

Given rival sets of beliefs (theories), it seems plausible to suppose that:

- *one theory is rationally preferable to the others if it is sufficiently better on sufficiently many of these criteria.*

This is somewhat vague, but it at least makes clear how an inconsistent theory may be rationally acceptable. Though it is inconsistent, it out-performs each competitor on many of the other criteria.

One may make the account more precise as follows. We suppose that for any criterion, c, and any set of beliefs, K, we may assign it a real number, $\mu_c(K)$, which measures how good K is according to constraint c. Suppose that the criteria are c_1, ..., c_n; and suppose that the relative importance of these is given by the weights (real numbers), w_1, ..., w_n. Then the *rationality index* of K, $\rho(K)$, can be defined by:

$$\rho(K) = w_1.\mu_{c_1}(K) + ... + w_n.\mu_{c_n}(K)$$

Given a collection of possible belief sets, K_1, ..., K_m, the one that it is rational to believe is the one with the greatest rationally index. (If there is more than one, believers have free choice.) This one may be inconsistent.

4 Belief Revision

An important objection to the claim that one can rationally believe contradictions is as follows: if one could accept a contradiction, one would never have to *revise* one's beliefs; for given any evidence that contradicts one's beliefs, one could simply *add* it, even though the result is inconsistent.

We can now see that this objection fails. Given new information that is inconsistent with one's beliefs, there are many possible responses, such as:

- simply add it
- say that it was mistaken
- add it, and throw out something to maintain consistency (this may be possible in several different ways)

- reject the whole conceptual framework in which one is working and replace it with a set of beliefs in another famework (in the way that phlogiston chemistry was replaced by oxygen chemistry in the light of information about weights)

These responses will produce a whole collectiont of possible belief sets. According to the account of the previous section, one should believe whichever of these has the greatest rationality index. This may well not be the one obtained by simply adding the in formation (which may produce a failure of simplicity, unity, etc.).

- *Conclusions*:
  - Consistency is not a necessary condition for rationality. (It may be rational to be inconsistent.)
  - The rational set of beliefs (in any given context) is the one with the highest rationality index, i.e., that performs best *overall* on the rationality criteria.
  - Even though inconsistencies may be rationally acceptable, new evidence that is inconsistent with our beliefs may well occasion rationally dropping beliefs.