Categorical versus graded beliefs

Franz Dietrich

Abstract

This essay discusses the difficulty to reconcile two paradigms about beliefs: the binary or categorical paradigm of yes/no beliefs and the probabilistic paradigm of degrees of belief. The possibility for someone to hold both types of belief simultaneously is challenged by the lottery paradox, and more recently by a general impossibility theorem. The nature, relevance, and implications of the tension are explained and assessed. A more technical elaboration can be found in Dietrich and List (2018, 2021).

1   Two types of belief and their potential coexistence

Rational-choice theory and logic have very different concepts of belief, each of which enjoys significant appeal and wide applications. Rational-choice theory takes agents to have graded beliefs of the form of subjective probability assignments. One might believe that it rains with subjective probability 2/3, or that one will stay healthy with subjective probability 3/4. By contrast, logic takes agents to have categorical beliefs, of the form of 'yes' or 'no' (or abstention). One might believe that it rains, or that one will stay healthy, in a categorical rather than graded sense. Believing something categorically should not be confused with complete certainty, i.e., with maximal graded belief: otherwise one would hardly ever believe anything in the categorical sense.

The advantage for rational choice theory of assuming probabilistic beliefs is considerable: it opens to the door to the classic notion of a rational agent seeking to maximise expected utilities, since expected utilities are the result of combining the probabilistic model of beliefs with the utility-based model of goals, values, or desires. As such, probabilistic beliefs form an intrinsic part of the classic homo oeconomicus. By contrast, logicians are less interested in decision making, and hence do not need to combine beliefs with goals, values, or desires. Instead, they typically focus on beliefs alone, which they usually take to be truth-oriented, logically consistent, and deductively closed, and to evolve via reasoning and revision.

Of course, rational-choice theory has its own theory of belief revision: a highly unified Bayesian theory, in which probabilistic beliefs undergo Bayesian updating as new information arrives. But it is questionable whether Bayesianism yields a theory of reasoning, and more generally whether probabilistic beliefs lend themselves to any

form of reasoning at all. Indeed, reasoning is not revision, as it does not draw on new information but rather on inference from existing beliefs. For logicians, reasoning happens in language, and is a process of drawing conclusions from initially believed premises. The categorical notion of beliefs lends itself much better to reasoning.

Rational choice theorists and logicians are both right in some sense, since both models of belief fulfills the purpose set by the respective discipline. But can both concepts of belief coexist in the same agent? Such an agent would for instance simultaneously believe that it rains with subjective probability 2/3 and believe that it rains simpliciter. More generally, for any relevant proposition p, the agent would hold some subjective probability of p and hold some yes/no belief about p. Depending on the context, the agent might draw either on their categorical beliefs or on their graded beliefs. In some contexts, the agent might reason logically with categorical beliefs by drawing inferences from existing beliefs, thereby forming new beliefs. When learning information, the agent might on the one hand logically revise categorical beliefs, and on the other hand Bayes-revise graded beliefs. In decision-making contexts, the agent might either use a simple heuristic that draws on categorical beliefs, or use a more sophisticated decision rule (possibly the expected-utility rule) that draws on graded beliefs. In short, each type of belief would play a different functional role. Neither type would be redundant, since each type is tailored to its own role, and each type outperforms the other type in its own area of application. Under this attractive division-of-labour picture, each belief type would be a legitimate component of the agent's psychology.

But this picture can only be maintained if the two belief types are mutually compatible in some sense, i.e., if the agent can maintain categorical and graded beliefs in a mutually coherent way. What exactly 'coherence' amounts to is a question on its own, but roughly speaking one would expect the agent to categorically believe propositions in which they have high degree of belief, and to categorically disbelieve propositions in which they have low degree of belief.

The question of whether and how one can coherently hold both types of belief has recently received renewed attention. See for instance contributions by Hawthorne and Bovens (1999), Douven and Williamson (2006), Lin and Kelly (2012a, 2012b), Leitgeb (2014), and Dietrich and List (2018, 2021).

2 A general impossibility theorem about coexistence of both belief types

Our notion of 'can coexist' is normative, not positive. That is, we do not describe real agents, but we ask whether an idealised agent – perhaps to be called a 'rational' agent – can coherently hold both belief types.

The coexistence of both belief types is challenged by the well-known *lottery paradox* (cf. the above-cited literature). The lottery paradox starts from a natural hy-

pothesis about the relationship between both belief types: one believes a proposition categorically if and only if one has a high enough degree of belief in it. The lottery paradox is meant to illustrate that this hypothesis runs into a problem: even if graded beliefs are perfectly rational, i.e., obey probability theory, the corresponding set of categorical beliefs, formed in accordance with the mentioned hypothesis, can be irrational, i.e., neither consistent nor deductively closed.

Why? In the lottery paradox, you are given a book of 100 pages. You know that exactly one page is black and all others are white. You have no idea about which page is black. So for each page you have a subjective probability of 99/100 that it is white. This subjective probability is high enough to make you (categorically) believe that the page is white. Meanwhile you have a subjective probability of 1 that not all pages are white. This maximal subjective probability is of course high enough to make you (categorically) believe that not all pages are white. The problem is: you believe that the first page is white, that the second page is white, and so one; but you fail to believe an implication of these 100 beliefs, namely that all pages are white – a violation of deductive closure. Worse, you believe the opposite of this implication, namely that *not* all pages are white – a violation of logical consistency.

Although the lottery paradox seems special in its setup, the underlying problem is very general. This fact has long been recognized informally, and has recently been established formally through an impossibility theorem derived in different versions by Dietrich and List (2018, 2021). According to this theorem, the two belief types cannot generally coexist subject to respecting certain initially plausible conditions. What are these axiomatic conditions, informally?

Three conditions pertain just to one belief type. One of them requires the agent to only ever hold a categorical belief set that is consistent and deductively closed. Another one requires the agent to only ever hold degrees of belief that are probabilistically coherent (so that, for instance, the probability of 'rain or snow' is the sum of the probabilities of 'rain' and 'snow'). The remaining one pertains again to graded beliefs and requires that any (probabilistically coherent) degrees of belief are allowed, i.e., can be held jointly with some categorical beliefs.

Three further conditions pertain to the relationship between the two belief types. One of them requires to categorically believe any proposition that has maximal subjective probability 1. Another one requires that the two belief types impose at least some non-trivial constraints on one another, rather than being essentially disconnected. The third condition requires that all relationships (i.e., mutual constraints) between the two belief types take a 'local' rather than 'global' form, in a sense defined below.

The impossibility theorem says that these axiomatic conditions are mutually inconsistent. So, it is strictly impossible to hold beliefs of both types in accordance with these axiomatic conditions.

To be a little more precise, let me sketch the formal setup. Consider a set X of

propositions (or events) of interest; in the lottery paradox, X contains propositions about which page(s) are white.[1] The agent's graded beliefs are represented by a *degree-of-belief function* Pr that assigns to each proposition $p \in X$ a subjective probability $Pr(p) \in [0, 1]$. The agent's categorical beliefs are represented by a *belief set* $B \subseteq X$, containing the (categorically) believed propositions. In different contexts the agent may hold different combinations (Pr, B) of a degree-of-belief function and a belief set. But not every combination counts as coherent: some combinations are coherent, the others are incoherent. Formally, coherence (or co-tenability) defines a binary relation between degree-of-belief functions Pr and belief sets B.

What is the structure of the coherence relation? The theorem assumes that the relation satisfies six conditions. They were stated informally above. Here are more formal re-statements:

- Categorical beliefs are logically coherent: all permissible belief sets B are logically consistent and deductively closed. "Permissible" means that B is coherent with at least one degree-of-belief function Pr.

- Graded beliefs are probabilistically coherent: any permissible degree-of-belief function Pr obeys the laws of probability. "Permissible" means that Pr is coherent with at least one belief set B.

- No coherent graded beliefs are ruled out: every probabilistically coherent degree-of-belief function Pr is permissible. "Permissible" was just defined.

- Completely certain propositions are categorically believed: for any coherent combination (Pr, B) and any proposition $p \in X$, if $Pr(p) = 1$ then $p \in B$.

- The two belief types are non-trivially connected: at least one (permissible) degree-of-belief function Pr requires to believe some proposition $p \in X$ that is not completely certain, i.e., satisfies $Pr(p) \neq 1$. This rules out that only completely certain propositions are ever required to be believed. Technically, a degree-of-belief function Pr is said to "require" to believe a proposition p if p is contained in all belief sets coherent with Pr.

- The last condition, "locality", is informally defined below.

The impossibility theorem states that these conditions are mutually incompatible.[2]

---

[1] 'Propositions' could for instance be modelled as sets of possible worlds, i.e., subsets of some fixed underlying set Ω of possible worlds. This 'semantic' or 'set-theoretic' notion of proposition is common in rational-choice theory and probability theory, where propositions are usually called 'events'.

[2] The theorem assumes that the set X of propositions under consideration contains enough interconnections, of the sort that some propositions in X entail or are inconsistent with some other propositions in X. In the absence of any interconnections, no impossibility could possibly occur, since all belief sets $B \subseteq X$ would then be trivially consistent and deductively closed.

A special kind of coherence relation deserves being mentioned: so-called *functional* or *deterministic* relations, in which the graded beliefs always fully determine the categorical beliefs. Formally, functionally means that each permissible degree-of-belief function Pr is coherent with exactly one belief set B. Such a functional relation can be captured by a *binarization function* f which maps any (permissible) degree-of-belief function Pr to the corresponding belief set $B = f(Pr)$.

The mentioned impossibility result was initially stated under the assumption of functionality, hence as a theorem about the inexistence of any binarization *function* satisfying certain conditions (Dietrich and List 2018). To our later surprise, the impossibility extends to the much more general case without functionality assumption (Dietrich and List 2021). The non-functional case allows one's categorical beliefs to be related only very loosely to one's graded beliefs: one's degrees of belief could impose almost no constraints on categorical beliefs, thereby leaving much freedom in what to believe categorically. Despite such freedom, it remains impossible to hold both belief types in accordance with the theorem's plausible conditions.

3 What to make of the formal impossibility?

Different reactions to the impossibility are imaginable. Either one takes rational agents to have only graded belief, no categorical beliefs – against the logical paradigm. Or one takes rational agents to have only categorical beliefs, no graded beliefs – against the rational-choice-theoretic paradigm. Or one maintains both belief types, but gives up some of the conditions assumed in the incompatibility theorem. As a matter of fact, most conditions in the theorem seem normatively inescapable. But there are two important exceptions:

- One can give up locality. Locality demands that any implications of graded beliefs for categorical beliefs are 'local', i.e., proposition-by-proposition. More precisely, whether one's graded beliefs *require* to believe a given proposition only depends on the graded belief in *this* proposition (where "require to believe" was defined above). For instance, if the graded beliefs require to believe in rain, then changing the degree of belief in sunshine without changing the degree of belief in rain does not lift the requirement to believe in rain. An example of a local relation between both belief types is the mentioned hypothesis whereby one believes a proposition categorically if and only if its subjective probability is high enough. Locality is a strong demand. Many interesting non-local ('holistic') ways to relate both belief types are presented in Dietrich and List (2018, 2021).

- More radically, one could turn to a different theory of graded beliefs, by giving up probabilistic beliefs in favour of some other notion of graded belief. Multi-valued logic and ranking theory provide alternative kinds of graded belief. Interestingly, ranking-theoretic beliefs (Spohn 2012) would escape the im-

possibility and allow for a viable coexistence of graded and categorical beliefs. Needless to say, orthodox rational-choice theory would be reluctant to replace "their" probabilistic paradigm by an altogether different, albeit graded, notion of belief.

References

Dietrich, Franz, and Christian List (2018) From degrees of belief to binary beliefs: Lessons from judgment-aggregation theory, *Journal of Philosophy* 115: 225-270

Dietrich, Franz, and Christian List (2021) The relation between degrees of belief and binary beliefs: a general impossibility theorem. In: I. Douven (ed.) *Lotteries, Knowledge, and Rational Belief. Essays on the Lottery Paradox*, Cambridge University Press, 2021, pp. 223-54

Hawthorne, James, and Luc Bovens (1999) The Preface, the Lottery, and the Logic of Belief, *Mind* 108(430): 241–264

Leitgeb, Hannes (2014) The Stability Theory of Belief, *Philosophical Review* 123(2): 131–171

Lin, Hanti, and Kevin T. Kelly (2012a) "Propositional Reasoning that Tracks Probabilistic Reasoning", Journal of Philosophical Logic 41(6): 957–981

Lin, Hanti, and Kevin T. Kelly (2012b) A Geo-logical Solution to the Lottery Paradox, *Synthese* 186(2): 531–575

Spohn, Wolfgang (2012) *The Laws of Belief: Ranking Theory and Its Philosophical Applications*, Oxford: Oxford University Press