Updating with Confidence

1 Introduction

Non sequitur. If we don't have absolute certainty, we would be certain to some degree, but that's different from confidence.

~~Seldom do we have the luxury of absolute certainty. This raises a question:~~ what does it mean to have a moderate degree of confidence in a piece of information ϕ? Probability gives a perfectly coherent answer: to have high confidence in ϕ is to find it highly probable, while to have low confidence is to find it improbable. In a world where probabillity is so dominant, this reading of "confidence" seems to have shadowed another closely related notion.

For us, confidence is a measure of trust, rather than likelihood. Like probability, confidence is a continuum between two extremes, and can be represented as a real number $\chi \in [0, 1]$. While probability ranges from untenable to undeniable, confidence ranges from completely untrustworthy to fully trusted. This paper explores on how confidence works in the context of updating beliefs. In this setting, we start with some belief state θ, and then recieve an observation ϕ with some degree of confidence χ, which we take into account to arrive at a new belief state θ'. When we have no confidence in ϕ (i.e., $\chi = 0$), we argue that it is appropriate to ignore it; if we have full confidence in ϕ (i.e., $\chi = 1$), then we ought to fully incorporate it into our beliefs.

What about intermediate values of confidence? As one might expect, they represent a smooth interpolation between ignoring and fully incorporating new information. So concretely, $\chi = 0.7$ indicates that one should move one's beliefs "70% of the way" towards fully incorporating the information. For example, if θ is a probability distribution μ over W, and the observation is an event $A \subset W$, one natural interpretation of "70%" is via convex combination, in which case the result of an update is

$$\theta' = (1 - \chi)\, \mu + (\chi)\, \mu|A.$$

This is not the only reasonable path between μ and the conditioned distribution $\mu|A$.[1] That said, some paths have nicer properties than others, and for this reason we will devote a significant amount of energy to characterizing them, so as to nail down some shared properties of confidence. In the example above, the full-confidence endpoint (the conditioned distribution) is obvious, and an interperetation of confidence amounts to a suitable choice of path between them. In some cases, the endpoint is unclear upfront, and rather defined implicitly as the ultimate destination of the appropriate path. For example, suppose our belief state θ specifies the weights of a neural network $f_\theta : X \to \Delta Y$, and the observation is a sample (x, y). Now, given a loss function, gradient descent[2] defines a path in belief space from the initial weights θ to some θ' such that $f_{\theta'}(x)$ is as close as possible placing all mass on y. We would like to apply the same notion of confidence—the same scale between ignoring and fully incorporating information—to both examples.

Intermediate levels of confidence come with some inherrent difficulties. For one, we suddenly have to worry about independence of our observations. For instance, observing A with confidence $\chi = 0.5$, and then again with confidence $\chi = 0.5$, moves us 50% of the way towards being certain in A, and then 50% of the remaining way, in total moving us 75% of the way towards certainty. This seems inappropriate if the second observation is a copy of the first. By contrast, dealing only with certainties, conditioning on A twice is no different from conditioning on A once, and so there's no need for such careful accounting. So what are we going to do to deal with this issue?

Probability and Confidence. Superficially, probability and confidence play quite different roles. A confidence measurement is an additional (possibly externally supplied) piece of information that we can use in updating our beliefs, while a probability is usually a function of our current belief state. Simply put, confidence is typically an input, while a probability is an output.

This point doesn't make sense to me If we replace confidence by probability, that would be an input, even if you had a belief state..

[1] another is given by $\mu'(x) \propto \mu(x)(1 - \chi)^{-\mathbf{1}[x \in A]}$

[2] techically gradient flow, which is the limit of gradient descent for small step size

I don't understand this at all. Why should the derivative be relevant? Where is the rest coming from? I strongly suggest you cut it. If not, you have to explain it much better.

I would cut this; it's a red herring.

In principle, there's no reason a probability couldn't be an input, in principle. So, what about a probabilistic view of confidence, in which χ is the probability that the information is trustworthy?

High confidence is in many ways like high probability: if we really trust a statement ϕ, we should fully incorporate it into our beliefs, and thereby come to believe it with high probability. Similarly, it only makes sense to be extremely confident in a ϕ if you believe that ϕ is extremely likely to be true. Low confidence, on the other hand, is quite different from low probability. If we have little trust in ϕ, we should *ignore* ϕ, rather than coming to believe that ϕ is unlikely. For example, if an adversary tells you something that you happen to already believe, you might say you have low confidence in their statement, but nevertheless ascribe it high probability.

One of the biggest shortcomings of probability is its inability to represent a truly neutral attitude towards a proposition—a transparent attitude, which defers entirely to prior belief. A value of $\frac{1}{2}$ may be equally far from zero as it is from one, but is by no means a neutral assessment in all cases: hearing that your favored candidate has a 50% chance of winning is big news if a win was previously thought to be inevitable. For this reason, telling someone the odds are 50/50 is quite different from saying you have no idea. By contrast, zero confidence represents something truly neutral: a statement made with zero confidence does not stake out a claim, and a statement recieved with zero confidence does not affect the recipient's beliefs. Why "transparent"?

Nevertheless, in some contexts, we will see that confidence corresponds to probability. To use a graphical metaphor, think of certainty as black or white. Probability describes shades of gray, while confidence describes opacity. If we are painting with black and start with a white canvas, there is a precise correspondence between the opacity and the resulting shade of gray.

What about connections to belief functions? Connections to variance?

Previous Versions

v0

There are many well-studied ways of representing and measuring uncertainty, probability chief among them. Indeed, an informal poll of our colleagues suggests that computer scientists view "confidence" as a synonym for probability. The enormous success of probability seems to have shadowed a closely related, but fundementally different, meaning of the word.

Probability is a numerical scale that ranges from untenable (0) to undeniable (1). No number on this scale is truly neutral. $\frac{1}{2}$ may split the difference between the extremes, but is by no means always a neutral assessment: learning that your favored candidate is likely to win with probabitliy $\frac{1}{2}$ is a big deal, if a win was previously thought to be inevitable. This shortcoming has perhaps been the primary selling point of many alternatives to probabiltiy, such as Dempster-Shafer Belief functions.

Confidence is also a scale between two extremes. It is a measure of trust on a scale of completley untrustworthy ($\perp$) to fully trusted $\top$. High confidence is quite like high probability: if we really trust a statement ϕ, we should fully incorporate it into our beliefs, and thereby come to believe it with high probability. Similarly, it only makes sense to be extremely confident in a ϕ if you believe that ϕ is extremely likely to be true. Low confidence, on the other hand, is quite different from low probability. If we have little trust in ϕ, we should *ignore* ϕ, rather than coming to believe that ϕ is unlikely. For example, if an adversary tells you something that you happen to already believe, you may have low confidence in their statement, but nevertheless ascribe it high probability. Importantly, zero confidence represents a truly neutral stance; a statement with zero confidence has no effect.

We will analyze this notion of confidence in the context of updating. In this setting, one has some belief state, and recieves inputs, which one might have some degree of confidence in, which is used to modify one's belief state. Confidence measures how seriously to take an input in updating beliefs.

v1

The ability articulate a "degree of confidence" is an important aspect of knowledge representation. Of course, there are many well-established ways of representing uncertainty, probability chief among them. Indeed, an informal poll of our colleagues suggests that most computer scientists view "confidence" as a synonym for probability. Although this use of the word is perfectly reasonable, it seems to have shadowed another conception of confidence—one that is fundementally different, if at first sublty so.

For us, confidence is a measure of trust. Like probability, it is a scale between two extremes. While probability ranges from untenable (0) to undeniable (1), confidence ranges from completely untrustworthy ($\perp$) to fully trusted ($\top$). This paper explores how confidence works in the context of learning. In this setting, one has some belief state, and recieves inputs, which one might have some degree of confidence in, which is used to modify one's belief state. Our use of confidence can be viewed as a measure of how seriously to take an input in updating our beliefs.

High confidence is in many ways like high probability: if we really trust a statement ϕ, we should fully incorporate it into our beliefs, and thereby come to believe it with high probability. Similarly, it only makes sense to be extremely confident in a ϕ if you believe that ϕ is extremely likely to be true. Low confidence, on the other hand, is quite different from low probability. If we have little trust in ϕ, we should *ignore* ϕ, rather than coming to believe that ϕ is unlikely. For example, if an adversary tells you something that you happen to already believe, you might say you have low confidence in their statement, but nevertheless ascribe it high probability.