Comments on Elga and Rayo, "Fragmentation and Logical Omniscience"
Eastern APA 2021
Andy Egan, Rutgers University
andyegan@rutgers.edu

0. Intro:
    a. I really enjoyed, and really appreciate, this paper. It's super helpful and interesting.
    b. Glad to be invited to be in this session – thanks!
    c. Structure:
        i. do some stage setting and comprehension checking, say back what I take to be some of the central threads of the paper and give authors a chance to correct me.
        ii. Some appreciation of and drawing attention to cool stuff that happens in the paper.
        iii. A couple of "whydja do it that way" questions, both of which do in fact have the sort of "wouldn't it have been better to do it this other way?" objections that you might expect lurking in the background.
        iv. Draw attention to some other nearby cans of worms, a mix of "it'd be great to hear more about this" ones, and "I'm worried this is going to be a problem" ones.

1. Stage setting & comprehension check
    a. The Watson case – obviously he should look at the note!
    b. But not according to standard decision theory.
    c. (Because standard decision theory doesn't use a representation of belief that allows for failures of logical omniscience.)
    d. So: We move to a fragmented model of doxastic states, where doxastic states are modeled with access tables – these are or might as well be sets of pairs of choice conditions and credence functons.
    e. Also go to a decision rule built to take fragmented doxastic state representors as inputs: Rational action in choice situation c is the one that maximizes expected utility given subject's c-associated credence function.
    f. (They also drop a footnote about fragmenting value.)
    g. A selling point of the fragmented model: It lets us capture logical competence as something different from, and less demanding than, logical omniscience.
    h. Another: It lets us give a nice model of what's happening in logical thought, where we make cognitive progress by working out logical relations.

2. Appreciation of cool stuff
    a. Cool result about the impossibility of imposing just a "respects obvious entailments" requirement on incoherent credence functions, without imposing a logical omniscience requirement.

b. Cool way of assessing fragmented states for a certain kind of rationality, that lets us distinguish *respecting obvious entailments* from *being logically omniscient* in a fragmented framework.
    i. Comprehension check:
        1. The reason why the argument that there's no way to require respecting obvious entailments without requiring logical omniscience doesn't go through on the AE/AR model is that different sentences are salient in different choice situations (and maybe different ways of assembling them into complex sentences count as *direct*?) , so we can require that in c where S and W are salient but *S&W* and *~S or V* aren't, we have to respect the entailment from S&W to S, but not the one from (S&W)&(~S or V) to V.
        2. Then I count as respecting obvious entailments, and therefore being basically logically competent, iff for all c, my c-associated credence function respects the obvious entailments between sentences relevant in c and sentences directly constructed from them.
        3. Is that right?

c. Cool model of logical thought.
    i. Start off with a fragmented state, in which logical information (information about entailments between sentences) is scattered across fragmented states, but not uniformly available.
    ii. What happens in logical thinking is that some new logical information becomes uniformly available – every probability function associated with any choice condition gets updated with it.
    iii. That's how you make logical progress, and that's what achieving some logical insight amounts to: Adding to the range of entailments your currently active credence function respects, and updating all your other fragments to incorporate the same logical information.

3. "Whydja do it that way" questions
    a. Here's a way of characterizing the project and motivation:
        i. There's this problem with standard ways of modeling rational belief and rational decisions, which is that when you want to evaluate believers for logical competence, the only test the models lets us construct is a super-demanding binary (pass/fail) test that everybody fails.
        ii. So what we're going to do is to replace that with a less demanding binary test.
        iii. I have two worries about that approach:
            1. Even though it's less demanding, it's still demanding enough that everybody I know still fails it. (And I gotta say, I know some pretty logically sophisticated people.)

            a. So maybe still not great for dividing actual agents into the ones who are doing a good job, logic-wise, and the ones who aren't.
        2. Not at all clear that what we really want, to do the evaluative work in the neighborhood, is a binary test rather than something degreed. (Maybe something more complicated than a simple scale, even.)
    iv. So – Whydja do it that way?

b. Logical thought and logical insight modeled with (a) increasing range of entailments your current fragment respects, (b) updating all your other fragments so they also respect the same additional entailments.
    i. That's super demanding!
    ii. Why require (b) as well as (a)?
    iii. (a) is enough to explain Watson, right?
    iv. And (b) is a very demanding binary condition – might want to just insist on (a) for doing a good job in logical reasoning, and leave (b) for a different, cross-cutting kind of doxastic assessment. Or might want to have some graded assessment, where you do better the more widely the updates spread, but it's not just that you pass if it's everywhere and fail if it's not.

4. Other interesting nearby cans of worms
    a. I have a big "yes" written in the margins by the spot where you bring up fragmentation of value. Invite you to say more about how you think that's gonna work!
        i. What might the models look like – same deal, access tables with value functions?
        ii. What kinds of rational assessment, if any, are fragmented value states gonna be subject to?
    b. Another point, independently relevant but I think more pressing when you've also got fragmented value:
        i. How do you keep the project one of providing some kind of interesting, theoretically valuable characterization of a subject's mental states, the kind of thing that could potentially explain a subject's behavior, rather than a complicated redescription of the subject's behavioral dispositions?
        ii. Always gonna be a way of pairing information states and choice situations that matches up with what the subject does, or is disposed to do.
        iii. (Gonna be a bunch of ways – especially once we've got value functions as another independently movable part)
        iv. Need some extra constraints – what are those going to be?
    c. Rationalizing action – the way you do it, what's rational in c is what going to be what maximizes expected utility given my c-credences and c-values.

i. That seems like it captures one kind of evaluation, one way in which I can do a bad job – I can fail to do what's most likely to have the best results, given the information and values that I've in fact brought to bear on the decision.
ii. Another way of going wrong that it misses though: I don't bring to bear the right information or values.
    1. Value case: I just bring to bear my deliciousness-based values, and not my health-related ones, when I decide to eat the cake. It really does maximize EU given the c and v functions that get activated in that choice situation, but that was the wrong v function to activate.
        a. (Or I only bring to bear my short-term rather than long-term values, or…)
    2. Belief case: I make the rational choice given the information I *do* bring to bear, but not the rational choice given the information I *should have* brought to bear.
        a. Given that I'm not bringing to bear the info that I promised to pick the kids up from school, I'm maximizing EU by working on my comments. But I ought to be bringing that info to bear.

d. Two different motivations for fragmented pictures that I think could stand to be pulled apart more clearly:
    i. Lots of people can't be modeled with existing tools.
        1. Unified models of belief and value just aren't well-suited to model the actual states of actual agents, and so there's a market for fragmented models.
    ii. Existing tools tell us everybody in the world gets the only kind of rational thumbs down that we know how to give, puts everybody in the "fail" bucket.
        1. With unified models we have to model everybody who's not logically omniscient with an incoherent credence function
        2. all the incoherent credence functions get failing grades on the "is it coherent?" test for logical competence, which is the only logical-competence-ish test we know how to formulate in unified models.
    iii. Those are different – the "can't be modeled" complaint and the "doesn't allow for the right kinds of normative evaluation" complaint. Would be useful to hear a bit more about which bits of the AEAR picture are motivated by which thing.

To dos:

Replacing a binary test everybody fails with one only some people fail

Also: The “can’t model agents like us in the framework” motivation vs the “everybody gets the only thumbs down the framework knows how to issue” motivation, and different bits of story responsive to each

Overdemandingness of universal percolation

Say something about fragmented value

Say something about the collapse to complicated redescription worry