

# Responsibility Without Causation - Probability Raising as an Independent Desert Base

Sarah Hiller

*Institute for Mathematics, Free University Berlin, Arnimallee 3,  
14195 Berlin, Germany*

Current crises such as climate change, mass extinction or the Covid pandemic and the related questions regarding collective decision making have exemplified the need for a formal quantification of responsibility in complex interactive situations. Our focus lies on an objective and normative assessment.

In the formal ethical literature it goes undisputed that actual causation is one of the preconditions of responsibility. That is, both cause and effect must actually obtain and they must stand in the relevant causal relation to each other. How exactly this causal relation is implemented may vary from causal models [4] to NESS causation [2] or the very strict variant of *seeing to it that* [6], but the presupposition of actual causation for responsibility ascription remains intact.

However, there are definitely cases in which we want to assign responsibility while actual causation is absent. For example, in legal contexts *attempts* are well-known cases of non-actualized causal relations. One way this has been integrated into the above-mentioned formal representations is via an explicit addition of action failure to the existing structure [1]. On a more general view, in situations in which we wish to assign moral responsibility without being able to rely on actual causation because an unsuccessful attempt was made to reach a certain outcome, researchers have noted that the action in question increased the chance of the unwanted outcome occurring [3]. This, we believe, has added to an existing confusion regarding the relation between causation and probability raising. Namely, the differences between these two concepts have been overlooked and it was attempted to reduce causation to probability raising [7, 9]. However, we know that there may be causation without probability raising and probability raising without causation [5], so this reduction seems implausible.

We argue, instead, that probability raising may serve as an independent desert base for responsibility ascription, as was suggested previously in a legal context e.g. by Moore [8]. Each desert base can maximally define a lower

bound for responsibility quantification; an absence of probability increase does not imply an absence of responsibility, as there may still be a causal connection.

We formalise the decision situations in which we wish to model responsibility via extensive-form game trees equipped with additional deontic and epistemic features. These represent decisions of agents in a temporal progression and can be seen as a finite fragment of the infinite *stit* frames, or as an extension giving a temporal dimension to the single-moment normal-form games used by Braham and van Hees [2]. In addition to nodes in which a single agent has a choice we also include nodes representing uncertainty, both probabilistic and non-probabilistic, as well as an agent-specific indistinguishability equivalence relation between decision nodes. Outcome nodes - the leaves of the finite tree - are partitioned into ethically desirable and ethically undesirable ones. A graphical representation of an example scenario within our model is depicted in Figure 1.

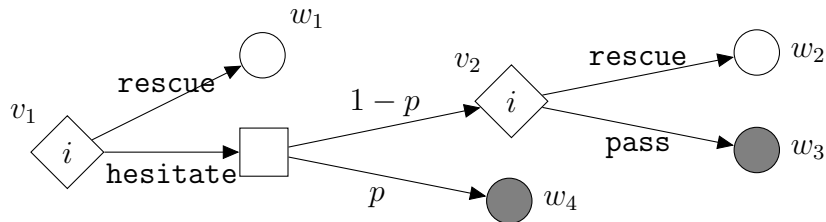


Figure 1: Graphic representation of a decision situation within our model. Agent  $i$  has the choice to either **rescue** an innocent stranger from drowning immediately or to **hesitate**, in which case there is some probability  $p$  that the stranger will drown ( $w_4$ ), and with another probability  $1 - p$  that agent  $i$  will get a second chance ( $v_2$ ) for the same decision. The stranger drowning is the ethically undesirable outcome and shaded in gray.

Subsequently to formally representing the decision situations we want to compute responsibility within this representation as an increase in the probability of an unwanted outcome. To this end we need to determine the baseline which serves as the point of comparison from which an increase or decrease is observed. It turns out that this selection is not as straightforward as it might initially seem. Specifically, with the help of two paradigmatic example scenarios we are able to show that neither a comparison with a single best nor a single worst scenario adequately captures all aspects of responsibility ascription. The scenarios in question are depicted in Figure 2.

If an agent is unsure about which context they find themselves in, and it might be that their action differentiates between a good and a bad outcome, but it might also be that they have no influence over the outcome because it will be undesirable no matter what they do, then the comparison to the worst case scenario will result in a lack of responsibility. Similarly, in the parallel example where it might be that the agent has no agency because the outcome might be desirable no matter what they do, a comparison with a best

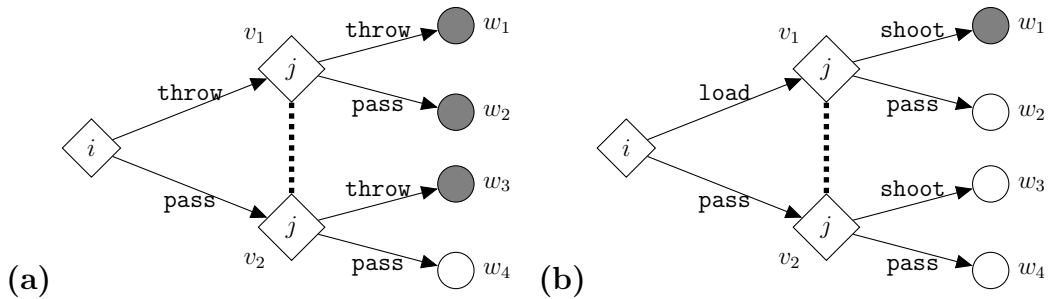


Figure 2: (a) Agent  $i$  can decide to throw a rock at a bottle and thus certainly shatter it or to refrain from throwing. A moment later agent  $j$  is faced with the same choice, not knowing what agent  $i$  selected. The bottle shattering is the undesirable outcome. (b) Agent  $i$  has the choice to either load a gun that they pass to agent  $j$  or to not load the gun. Not knowing which action  $i$  selected, agent  $j$  has to decide whether or not to shoot at a bottle. Again, the bottle shattering is undesirable.

case scenario leads to a lack of responsibility. We argue, however, that in both cases the agent must assume that their action might be the one determining the outcome, so that we do want to assign responsibility in both cases. Thus, we need to incorporate a comparison with a range of possible alternative scenarios, rather than a simple single best or worst case. We will present the details of this comparison and show how we are able to resolve further paradigmatic scenarios with this approach.

**Keywords.** Formal Ethics, Responsibility Quantification, Causation, Probability Raising

## References

- [1] Alexandru Baltag, Ilaria Canavotto, and Sonja Smets. “Causal Agency and Responsibility: A Refinement of STIT Logic”. 2020.
- [2] Matthew Braham and Martin van Hees. “An Anatomy of Moral Responsibility”. In: *Mind* 121.483 (July 2012), pp. 601–634.
- [3] Jan Broersen. “Modeling Attempt and Action Failure in Probabilistic STIT Logic”. In: *Proceedings of Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI 2011)*. Ed. by Toby Walsh. IJCAI, 2011, pp. 792–797.
- [4] Hana Chockler and Joseph Y. Halpern. “Responsibility and Blame: A Structural-Model Approach”. In: *Journal of Artificial Intelligence Research* 22 (2004), pp. 93–115.

- [5] C. Hitchcock. “Do All and Only Causes Raise the Probabilities of Effects?” In: *Causation and Counterfactuals*. Ed. by J. Collins, E. J. Hall, and L. A. Paul. MIT Press, 2004.
- [6] John F. Horty. *Agency and Deontic Logic*. Oxford University Press, 2001.
- [7] Igal Kwart. “Probabilistic Cause and the Thirsty Traveller”. In: *Journal of Philosophical Logic* 31 (2002), pp. 139–179.
- [8] Michael S. Moore. “Causing, Aiding, and the Superfluity of Accomplice Liability”. In: *University of Pennsylvania Law Review* 156.2 (2007), pp. 395–452.
- [9] Peter Vallentyne. “Brute Luck and Responsibility”. In: *Politics, Philosophy and Economics* 7.1 (2008), pp. 57–80.