Wall vs Fire

by Sven Nilsen, 2019

Wall vs Fire means two different ways of weighting a Path Reference Class under constrained conditions. A Wall adds more weight to remaining choices such that their sum equals a constant that defines what it means to preserve the observer under the single-world interpretation. A Fire does not add more weight to remaining choices, but reduces the weight for a specific choice to zero, such that the sum is less than a constant that defines what it means to preserve the observer under the single-world interpretation.

A normalized Path Reference Class changes weights into probabilities. Under the Path Reference Class Duality, this has a many-worlds interpretation and a single-world interpretation. In the single-world interpretation, probabilities mean how likely the observer is to be in a single cell from the chance of choices that the observer makes. In the many-worlds interpretation, probabilities mean the fraction of observers that exists in a single cell.

Since normalization usually happens after the update of values, one must be careful to pick rules for updating such that the model predicts what our intuition indicates is correct about observers. Our intuition depends on how we interpret probabilities.

A Fire destroys the observer under the single-world interpretation. However, in the many-worlds interpretation, there is a possibility that the branching of observers is reduced instead. This produces an identical prediction of probability that corresponds to the observer being destroyed in the fire, although no observer actually die. This means that Fire is used to talk about probabilities under the many-worlds interpretation which might not mean terminating observer states, but since there is no natural intuition for how this happens across many-worlds, one borrows the term from the single-world interpretation.

By mistake, it is easy to use the Fire method while intending to use the Wall method. This can lead to weird predictions, such as the probabilities vanishing inside a closed room simply by existing alongside a slightly bigger closed room in the same world. This happens because when the weights propagate to the Fire, they are either ignored or removed, resulting in less weight “reflecting” back.

A Wall is a way of adding more weight to remaining choices such that the probabilities do not vanish relative to the unconstrained conditions. For example, if the observer has 4 choices under unconstrained conditions and 1 choice is a wall, then each remaining choice need to be updated with 1/3 instead of 1/4.

However, since it is common to define updates as a function that returns the new value for each cell, it can be somewhat harder to program this correct, e.g. in the following case where yellow means a wall:

The sum of the weights in the unconstrained case is `14`, but since `7` weights are walls, the remaining weights should be divided by `14-7` instead of `14`.