Many Worlds and the Born Rule

Taha Dawoodbhoy

In non-deterministic Wavefunction-collapse theories like the Copenhagen Interpretation, probabilities are built-in by axiom, and set the frequency of the Wavefunction collapsing around some position. In hidden-variable theories like Pilot Wave Theory, probabilities emerge from our lack of knowledge of the hidden variables, and set the frequency of particles being located around some position. The Many Worlds Interpretation (MWI), on the other hand, is both deterministic and absent of any hidden information, so how can we use it to ascribe probabilities to measurement outcomes? How do probabilities enter the theory at all? A question like

> "What is the probability that I will find this particle located near position $\boldsymbol{x}$?"

is not actually well-defined in MWI, because that result is *guaranteed* to happen in some branch(es) of the Wavefunction, as are all other possible results – there are versions of myself that find that result *and* versions that do not. So how do we reconcile this picture with the fact that some measurement outcomes appear to occur more often than others, with probabilities given by the Born Rule? Well, to start, we should construct a 'probability question' that actually *is* well-defined in MWI. Rather than asking about the probability of obtaining a particular measurement outcome (which is indefinite), the natural choice is to ask about the probability of *being in a particular branch of the Wavefunction*.

Let us examine a simple case by considering the following Wavefunction for the measurement of a "quantum coin" at a time shortly after the measurement takes place. In this scenario, the coin has just become entangled with the apparatus used to 'measure' its state, and interactions between the apparatus and its environment have just led to some parts of the environment becoming entangled, as well:

$$|\Psi\rangle = \alpha \left|\boxed{\mathrm{H}}\right\rangle \otimes |\mathcal{E}_{\mathrm{H}}\rangle \otimes |\text{🤔}\rangle + \beta \left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_{\mathrm{T}}\rangle \otimes |\text{🤔}\rangle \tag{1}$$

The first factor in each term (with the boxed H or T) represents the state of the coin (heads or tails) and the measurement apparatus. The second factor (the $\mathcal{E}$ states) represents the entangled environment – stray electrons, photons, atoms, etc. in the room that have interacted with the apparatus, and whose states are entangled with the coin's. These first two factors will make this Wavefunction decoherent, so we can already think of it as two independent worlds. However, the cascade of entangling interactions has not yet reached the point that the observers in these two worlds are aware of the measurement outcome. They are included in the third factor, which represents the parts of the Wavefunction whose states are not yet entangled with the coin's, because the cascade of interactions has not yet reached them. These observers can then sensibly ask the question

> "What is the probability that I am in the world in which the apparatus measures the coin to be in the 'heads' state (or 'tails' state)?"

Thus, probabilities in MWI arise from the uncertainty that observers in branched Wavefunctions will initially have about which world they are actually in; they are the *credences* that the observers have of being in a particular world. Probabilities in MWI are, therefore, necessarily *Bayesian*, rather than *frequentist*.

But how do we actually determine what these Bayesian probabilities should be? Ultimately, we will need to show that they are consistent with the Born Rule, but it is worth noting that this is largely a conceptual problem, not a technical one. The square-amplitudes of the Wavefunction are a natural choice for probabilities in quantum mechanics, because they are always positive numbers that add up to 1. Therefore, essentially any attempt to uncover a reasonable calculation of probabilities based on the structure of quantum theory will lead to the Born Rule. That being said, a conceptual problem is still a problem, nonetheless. The challenge is to derive a definitive line of reasoning that would lead observers to assign credences based on the structure of quantum theory, so as to recover the Born Rule.

Unfortunately, there is a rather natural line of reasoning, often referred to as "branch counting", that is actually *inconsistent* with the Born Rule. The rationale for branch counting can be stated in a few different but related ways. For instance, considering the 2-branch Wavefunction in Eq. 1, the observers might reason along the lines of:

"I know that there are two observers in two different branches, and I know that I am one of them, so the probability that I am either one in particular should be 1 out of 2, or 50%."

So the branch counting rule is that, for a Wavefunction containing N independent branches, the probability of being in a particular branch is simply $1/N$. Of course, we know this is wrong *a posteriori*, because it is inconsistent with the Born Rule, and therefore with experiments. Nonetheless, the rationale that leads to branch counting is quite intuitive, so it behooves us to discuss what exactly is wrong with this intuition, and how exactly it should be corrected. The tacit assumption built-in to the rationale of branch counting is sometimes referred to as the "Principle of Indifference", which is, in essence, the statement that we are *indifferent* towards the various possibilities, and therefore assign equal probability to each. To make this explicit, let us restate the rationale as:

"I know that there are two observers in two different branches, I know that I am one of them, *and I am indifferent towards which one, because I have no reason to believe that either one is more likely than the other.* So, given that I believe both possibilities are equally likely, and that their combined probability should total to 1, the probability of each should be 0.5."

Put this way, it should be clear that this an assumption worthy of some skepticism. Do we really have *no* reason to believe that one is more likely than the other? Let us try to determine the conditions under which this assumption of indifference can be justified by starting with a classical analogy.

Classical Indifference

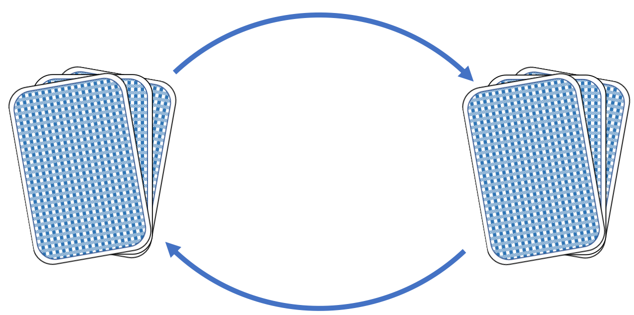

Figure 1

A dealer places two stacks of three cards face down on a table in front of you – one on the left and one on the right – and tells you that one of the cards is a joker. What is the probability (i.e. what is your credence) that the joker is in the stack on the left? You may have an intuition that the Principle of Indifference can be applied here, so that it is equally likely for the joker to be in either stack. But why? Suppose you swapped the positions of the two stacks, so that the stack that was on the left is now on the right, and vice versa, and reconsider the question: What is the probability that the joker is in the stack on the left? From your point-of-view, nothing has changed; as far as you can tell, the setup looks the same as it did before you swapped the stacks; there are still two stacks of three cards face down on the table, one on the left and one on the right (see Fig. 1). In that case, you must insist that your credence about whether the joker is in the stack on the left or the right is unchanged by this swap, so the probability of the joker being in the stack on the left *before* the swap is the same as the probability of it being in the stack on the left *after* the swap:

$$P(L, \text{before}) = P(L, \text{after}) \tag{2}$$

Furthermore, you must insist that the probability of the joker being in the stack on the *left after* the swap is the same as the probability of it being in the stack on the *right before* the swap, simply because the stack on the left after the swap *is* the stack on the right before the swap:

$$P(L, \text{after}) = P(R, \text{before}) \tag{3}$$

Combining these two equations, you find

$$P(L, \text{before}) = P(R, \text{before}) \tag{4}$$

which is the Principle of Indifference! Thus, we have justified the assumption for this scenario. Combining this with

$$P(L, \text{before}) + P(R, \text{before}) = 1 \tag{5}$$

you obtain the probabilities

$$P(L, \text{before}) = 1/2 \tag{6}$$

$$P(R, \text{before}) = 1/2 \tag{7}$$

We might call this result "stack counting": given a set of N stacks, the probability of the joker being in a particular stack is simply $1/N$.

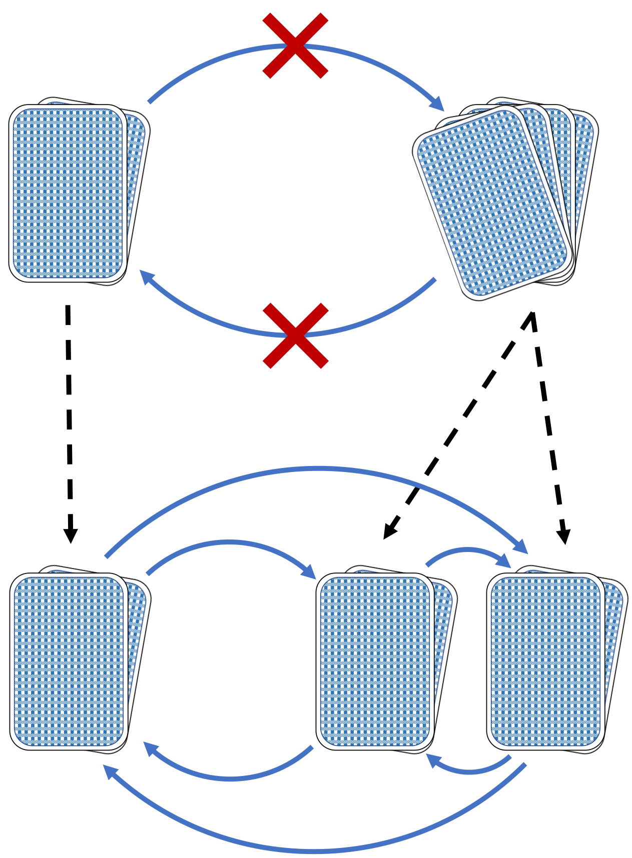

Figure 2

Now suppose the dealer takes one of the cards from the stack on the left and moves it to the stack on the right, so that the left stack has 2 cards and the right stack has 4. Reconsidering the question – What is the probability that the joker is in the stack on the left? – you may have an intuition that the Principle of Indifference *cannot* be applied here. But why? If you swap the stacks as you did before, you will find that, in this case, the setup does *not* look the same as it did before the swap. The stack on the left after the swap is now apparently different from the stack on the left before the swap – the former contains 4 cards while the latter contained 2 (see Fig. 2, top). Therefore, Eq. 2 does not follow, and you cannot justify the Principle of Indifference in this case. So what can you do instead? Well, imagine splitting the stack of 4 cards on the right into two stacks of 2. Let us call these imaginary stacks "pseudo-stacks". Now you have three (pseudo-)stacks of 2, one on the left and two on the right, and you can swap any two of them without changing the setup at all (see Fig. 2, bottom). Therefore, following the same reasoning as before, you must insist that the Principle of Indifference applies between these (pseudo-)stacks, i.e.

$$P(L, \text{before}) = P(R_1, \text{before}) = P(R_2, \text{before}) \tag{8}$$

where R_1 and R_2 refer to the two pseudo-stacks on the right. From this, it directly follows that

$$P(L, \text{before}) = 1/3 \tag{9}$$

$$P(R_1, \text{before}) + P(R_2, \text{before}) = P(R, \text{before}) = 2/3 \tag{10}$$

In this case, despite the fact that the Principle of Indifference was invalid at the outset, we were still able to leverage it in order to obtain the correct probabilities through a clever use of "pseudo-stacks".

Quantum Indifference

The situation in quantum mechanics and MWI is slightly more complicated, but we can refer to many of the ideas outlined in the above 'classical' analogy to guide us through it. Let us start again with Eq. 1, which I'll call $|\Psi_1\rangle$:

$$|\Psi_1\rangle = \alpha \left|\boxed{\mathrm{H}}\right\rangle \otimes |\mathcal{E}_\mathrm{H}\rangle \otimes |\text{🤔}\rangle + \beta \left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_\mathrm{T}\rangle \otimes |\text{🤔}\rangle \tag{11}$$

Recall the question at hand: What is the probability that you (an observer) are in the world (i.e. branch of the Wavefunction) in which the apparatus measures the coin to be in the 'heads' state? To see if the Principle of Indifference can apply here, let us follow the classical analogy and 'swap' the environment states between the two branches:

$$|\Psi_2\rangle = \alpha \left|\boxed{\mathrm{H}}\right\rangle \otimes |\mathcal{E}_\mathrm{T}\rangle \otimes |\text{🤔}\rangle + \beta \left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_\mathrm{H}\rangle \otimes |\text{🤔}\rangle \tag{12}$$

Now reconsider the question at hand: What is the probability that you are in the world in which the apparatus measures the coin to be in the 'heads' state? From your point-of-view, nothing has changed, because the apparatus states are exactly the same as they were before the swap; only the environment states changed, and they aren't relevant to the question at hand; we only care about the apparatus states at the moment. Therefore, you must insist that your credences are unchanged by the swap, so you can reasonably assert that the probability of being in the branch of $|\Psi_1\rangle$ in which the apparatus measures 'heads' is the same as that for $|\Psi_2\rangle$:

$$P\left(\boxed{\mathrm{H}}, \Psi_1\right) = P\left(\boxed{\mathrm{H}}, \Psi_2\right) \tag{13}$$

Now, let us note that the probability of being in the branch of $|\Psi_2\rangle$ in which the apparatus is in the state $\left|\boxed{\mathrm{H}}\right\rangle$ is the same as the probability of being in the branch of $|\Psi_2\rangle$ in which the *environment* is in the state $|\mathcal{E}_\mathrm{T}\rangle$:

$$P\left(\boxed{\mathrm{H}}, \Psi_2\right) = P\left(\mathcal{E}_\mathrm{T}, \Psi_2\right) \tag{14}$$

This is because, for $|\Psi_2\rangle$, the branch in which the apparatus is in the state $\left|\boxed{\mathrm{H}}\right\rangle$ *is* the branch in which the environment is in the state $|\mathcal{E}_\mathrm{T}\rangle$.

Next, we can swap the *apparatus states* of the two $|\Psi_2\rangle$ branches, to obtain a third Wavefunction:

$$|\Psi_3\rangle = \alpha \left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_\mathrm{T}\rangle \otimes |\text{🤔}\rangle + \beta \left|\boxed{\mathrm{H}}\right\rangle \otimes |\mathcal{E}_\mathrm{H}\rangle \otimes |\text{🤔}\rangle \tag{15}$$

Following the same reasoning as before, you must insist that the probability of being in the branch in which the *environment* is in the state $|\mathcal{E}_\mathrm{T}\rangle$ is unchanged by this swap, because the environment states were unaffected (only the apparatus states changed):

$$P(\mathcal{E}_\mathrm{T}, \Psi_2) = P(\mathcal{E}_\mathrm{T}, \Psi_3) \tag{16}$$

And, again, since the $|\Psi_3\rangle$ branch in which the environment state is $|\mathcal{E}_\mathrm{T}\rangle$ is the same as the $|\Psi_3\rangle$ branch in which the apparatus state is $\left|\boxed{\mathrm{T}}\right\rangle$, we have

$$P(\mathcal{E}_\mathrm{T}, \Psi_3) = P\left(\boxed{\mathrm{T}}, \Psi_3\right) \tag{17}$$

Putting this all together,

$$P\left(\boxed{\mathrm{H}}, \Psi_1\right) = P\left(\boxed{\mathrm{H}}, \Psi_2\right) = P(\mathcal{E}_\mathrm{T}, \Psi_2) = P(\mathcal{E}_\mathrm{T}, \Psi_3) = P\left(\boxed{\mathrm{T}}, \Psi_3\right) \tag{18}$$

Finally, notice that when the branch amplitudes are equal ($\alpha = \beta$), $|\Psi_3\rangle$ is exactly equal to $|\Psi_1\rangle$, so we can replace Ψ_3 with Ψ_1 in the last part of the equation above, which gives us

$$P\left(\boxed{\mathrm{H}}, \Psi_1\right) = P\left(\boxed{\mathrm{T}}, \Psi_1\right) \quad \text{when} \quad \alpha = \beta \tag{19}$$

This is the Principle of Indifference! Combining this with

$$P\left(\boxed{\mathrm{H}}, \Psi_1\right) + P\left(\boxed{\mathrm{T}}, \Psi_1\right) = 1 \tag{20}$$

we obtain the branch counting result

$$P\left(\boxed{\mathrm{H}}, \Psi_1\right) = 1/2 \quad \text{when} \quad \alpha = \beta \tag{21}$$

$$P\left(\boxed{\mathrm{T}}, \Psi_1\right) = 1/2 \quad \text{when} \quad \alpha = \beta \tag{22}$$

Thus, we have demonstrated that the Principle of Indifference and branch counting are valid and justified *when branches have equal amplitudes*, which is exactly the condition under which branch counting is consistent with the Born Rule. So far, so good.

When the amplitudes are unequal, however, the justification for the Principle of Indifference and branch counting breaks down (because $|\Psi_3\rangle$ will no longer be equal to $|\Psi_1\rangle$), so we have more work to do to derive the rest of the Born Rule. Following the classical analogy, we can still try to leverage indifference by imagining separating the Wavefunction into some number of *"pseudo-branches"* with equal amplitudes, and then applying the Principle of Indifference to those. To see how this goes, let us assume that the square-amplitudes of the two branches, $|\alpha|^2$ and $|\beta|^2$, are rational numbers; probabilities for Wavefunctions with irrational square-amplitudes can be interpolated from the former, if needed. Then, we can write the Wavefunction as

$$|\Psi\rangle = \sqrt{\frac{a}{c}}\left|\boxed{\mathrm{H}}\right\rangle \otimes |\mathcal{E}_\mathrm{H}\rangle \otimes |\text{🤔}\rangle + \sqrt{\frac{b}{c}}\left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_\mathrm{T}\rangle \otimes |\text{🤔}\rangle \tag{23}$$

where $|a|$, $|b|$, and $|c| = |a| + |b|$ are positive integers. Now we can separate the two branches into equal-amplitude pseudo-branches. To make the amplitudes equal, we can multiply the $\left|\boxed{\mathrm{H}}\right\rangle$-branch amplitudes by $1/\sqrt{a}$, and the $\left|\boxed{\mathrm{T}}\right\rangle$-branch amplitudes by $1/\sqrt{b}$. In that case, each pseudo-branch will have an amplitude of $1/\sqrt{c}$. By the Pythagorean Theorem, this requires that we split the $\left|\boxed{\mathrm{H}}\right\rangle$ branch into $|a|$ pseudo-branches, and the $\left|\boxed{\mathrm{T}}\right\rangle$ branch into $|b|$ pseudo-branches, so that the vector addition of the pseudo-branches returns the original branch that they came from (with its original length). In the classical analogy, this is like imposing that the combined total number of cards in the pseudo-stacks be equal to the number of cards in the original stack. For example, we can perform the splitting on the environment states as follows:

$$|\mathcal{E}_\mathrm{H}\rangle = \frac{1}{\sqrt{a}}(|\mathcal{E}_{\mathrm{H},1}\rangle + |\mathcal{E}_{\mathrm{H},2}\rangle + ... + |\mathcal{E}_{\mathrm{H},|a|}\rangle) \tag{24}$$

$$|\mathcal{E}_\mathrm{T}\rangle = \frac{1}{\sqrt{b}}(|\mathcal{E}_{\mathrm{T},1}\rangle + |\mathcal{E}_{\mathrm{T},2}\rangle + ... + |\mathcal{E}_{\mathrm{T},|b|}\rangle) \tag{25}$$

Here, $|\mathcal{E}_{\mathrm{H}}\rangle$ is split into a superposition of a number of "pseudo-environments", each with an amplitude of $1/\sqrt{a}$. By the Pythagorean Theorem, there must be exactly $|a|$ such pseudo-environments in the superposition. Similarly, $|\mathcal{E}_{\mathrm{T}}\rangle$ is split into $|b|$ pseudo-environments, each with an amplitude of $1/\sqrt{b}$. Substituting these into the Wavefunction, we obtain

$$\begin{aligned} |\Psi\rangle =& \sqrt{\frac{1}{c}}\left|\boxed{\mathrm{H}}\right\rangle \otimes (|\mathcal{E}_{\mathrm{H},1}\rangle + |\mathcal{E}_{\mathrm{H},2}\rangle + ... + |\mathcal{E}_{\mathrm{H},|a|}\rangle) \otimes |\text{🤔}\rangle \\ &+ \sqrt{\frac{1}{c}}\left|\boxed{\mathrm{T}}\right\rangle \otimes (|\mathcal{E}_{\mathrm{T},1}\rangle + |\mathcal{E}_{\mathrm{T},2}\rangle + ... + |\mathcal{E}_{\mathrm{T},|b|}\rangle) \otimes |\text{🤔}\rangle \end{aligned} \tag{26}$$

So, we now have $|c|$ pseudo-branches of equal amplitude, and by following the previous reasoning, the Principle of Indifference and branch counting can be applied to them: the probability of being in a particular pseudo-branch should be $1/|c|$. Therefore, since the apparatus measures the coin's state to be 'heads' in $|a|$ out of $|c|$ of these pseudo-branches, the probability of being in such a pseudo-branch is

$$P\left(\boxed{\mathrm{H}}, \Psi\right) = \frac{|a|}{|c|} \tag{27}$$

Similarly,

$$P\left(\boxed{\mathrm{T}}, \Psi\right) = \frac{|b|}{|c|} \tag{28}$$

As expected, this is the Born Rule!

For the sake of clarity, consider the following example, with $a = 1$, $b = 2$, and $c = 3$.

$$|\Psi_1\rangle = \sqrt{\frac{1}{3}}\left|\boxed{\mathrm{H}}\right\rangle \otimes |\mathcal{E}_{\mathrm{H}}\rangle \otimes |\text{🤔}\rangle + \sqrt{\frac{2}{3}}\left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_{\mathrm{T}}\rangle \otimes |\text{🤔}\rangle \tag{29}$$

To make the amplitudes equal, we will need to multiply the amplitude of the second term by $1/\sqrt{2}$, which means that we will need to split it into 2 pseudo-branches. We can do this by identifying 2 pseudo-environments as:

$$|\mathcal{E}_{\mathrm{T}}\rangle = \frac{1}{\sqrt{2}}(|\mathcal{E}_{\mathrm{T},1}\rangle + |\mathcal{E}_{\mathrm{T},2}\rangle) \tag{30}$$

Plugging this into $|\Psi_1\rangle$, we obtain

$$|\Psi_1\rangle = \sqrt{\frac{1}{3}}\left|\boxed{\mathrm{H}}\right\rangle \otimes |\mathcal{E}_{\mathrm{H}}\rangle \otimes |\text{🤔}\rangle + \sqrt{\frac{1}{3}}\left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_{\mathrm{T},1}\rangle \otimes |\text{🤔}\rangle + \sqrt{\frac{1}{3}}\left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_{\mathrm{T},2}\rangle \otimes |\text{🤔}\rangle \tag{31}$$

Now we can swap any two of the environment states to carry out the justification for the Principle of Indifference. For example, swapping $|\mathcal{E}_{\mathrm{H}}\rangle$ and $|\mathcal{E}_{\mathrm{T},1}\rangle$, we obtain

$$|\Psi_2\rangle = \sqrt{\frac{1}{3}}\left|\boxed{\mathrm{H}}\right\rangle \otimes |\mathcal{E}_{\mathrm{T},1}\rangle \otimes |\text{🤔}\rangle + \sqrt{\frac{1}{3}}\left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_{\mathrm{H}}\rangle \otimes |\text{🤔}\rangle + \sqrt{\frac{1}{3}}\left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_{\mathrm{T},2}\rangle \otimes |\text{🤔}\rangle \tag{32}$$

Next, we can swap the apparatus states to obtain

$$\begin{aligned} |\Psi_3\rangle &= \sqrt{\frac{1}{3}}\left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_{\mathrm{T},1}\rangle \otimes |\text{🤔}\rangle + \sqrt{\frac{1}{3}}\left|\boxed{\mathrm{H}}\right\rangle \otimes |\mathcal{E}_{\mathrm{H}}\rangle \otimes |\text{🤔}\rangle + \sqrt{\frac{1}{3}}\left|\boxed{\mathrm{T}}\right\rangle \otimes |\mathcal{E}_{\mathrm{T},2}\rangle \otimes |\text{🤔}\rangle \\ &= |\Psi_1\rangle \end{aligned} \tag{33}$$

So, following the same reasoning as before, the credences satisfy

$$P\left(\boxed{\mathrm{H}}, \Psi_1\right) = P\left(\boxed{\mathrm{H}}, \Psi_2\right) = P(\mathcal{E}_{\mathrm{T},1}, \Psi_2) = P(\mathcal{E}_{\mathrm{T},1}, \Psi_3) = P(\mathcal{E}_{\mathrm{T},1}, \Psi_1) \tag{34}$$

Of course, we could go through the same procedure while having swapped $|\mathcal{E}_{\mathrm{H}}\rangle$ and $|\mathcal{E}_{\mathrm{T},2}\rangle$ initially, instead, in which case we would find

$$P\left(\boxed{\mathrm{H}}, \Psi_1\right) = P\left(\boxed{\mathrm{H}}, \Psi_2\right) = P(\mathcal{E}_{\mathrm{T},2}, \Psi_2) = P(\mathcal{E}_{\mathrm{T},2}, \Psi_3) = P(\mathcal{E}_{\mathrm{T},2}, \Psi_1) \tag{35}$$

Combining these results, we have

$$P\left(\boxed{\mathrm{H}}, \Psi_1\right) = P(\mathcal{E}_{\mathrm{T},1}, \Psi_1) = P(\mathcal{E}_{\mathrm{T},2}, \Psi_1) \tag{36}$$

which is the Principle of Indifference. Then, since probabilities must sum to 1,

$$P\left(\boxed{\mathrm{H}}, \Psi_1\right) + P(\mathcal{E}_{\mathrm{T},1}, \Psi_1) + P(\mathcal{E}_{\mathrm{T},2}, \Psi_1) = 1 \tag{37}$$

we obtain

$$P\left(\boxed{\mathrm{H}}, \Psi_1\right) = 1/3 \tag{38}$$

$$P(\mathcal{E}_{\mathrm{T},1}, \Psi_1) + P(\mathcal{E}_{\mathrm{T},2}, \Psi_1) = P\left(\boxed{\mathrm{T}}, \Psi_1\right) = 2/3 \tag{39}$$

which is consistent with the Born Rule. To emphasize the point, notice that the key to this result is really in the first step, where we separate the $\left|\boxed{\mathrm{T}}\right\rangle$ branch into 2 equal amplitude pseudo-branches. Again, the requirement that there be exactly 2 pseudo-branches in this case, in order to make the amplitudes equal, is a direct consequence of the Pythagorean Theorem, and, thus, the Born Rule is, as well.

I would like to emphasize that this work is largely derivative of the work of Zurek (2005) and Sebens & Carroll (2018). Many clarifying details are presented in these two papers, and I encourage the reader to consult them as needed. The full references are given below.

References

C. T. Sebens and S. M. Carroll. Self-locating uncertainty and the origin of probability in everettian quantum mechanics. *The British Journal for the Philosophy of Science*, 69(1):25–74, 2018. doi: 10.1093/bjps/axw004. URL https://arxiv.org/abs/1405.7577.

W. H. Zurek. Probabilities from entanglement, born's rule $p_k = |\ \psi_k\ |^2$ from envariance. *Phys. Rev. A*, 71:052105, 2005. doi: 10.1103/PhysRevA.71.052105. URL https://arxiv.org/abs/quant-ph/0405161.