Title: Theory Confirmation in One World and its Failure in Many

Date: Oct 02, 2008 10:00 AM

URL: http://pirsa.org/08100048

Abstract: I discuss how we can give a satisfactory account of theory confirmation for theories with random data, such as Copenhagen quantum theory, despite the lack of a completely satisfactory definition of probabilistic theories of nature. I also explain why neither this nor any other proposed account of scientific confirmation works for many-worlds theories

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

Theory Confirmation ~~in One World and its~~* Failure in Many Worlds

Adrian Kent

DAMTP, University of Cambridge
and
Perimeter Institute

talk at "The Clock and the Quantum"

10002809 2008@ PI

* speaker was trying to do too much in one talk. Apologies

Motivation for this talk:

If we could make sense of Everettian many-worlds quantum theory, and if it reproduced all the scientific successes of Copenhagen quantum theory, then we should take it seriously. So ... if we think we can't, or it doesn't, then we should try to explain why.

Quite a lot hangs on this. Mathematically elegant, universally applicable, scientifically adequate versions of quantum theory that give a realist ontology are not exactly thick on the ground. If (as I think) Everett fails, it's very plausible that the problem is a limitation of quantum theory itself, and the ultimate solution is new physics.

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

want to engage with -- and, to be frank, try to refute -- some
ery ingenious and interesting proposals due to Wallace and
ireaves-Myrvold, who try to explain how to make sense of
robability in many-worlds theories and then explain how to test
nd confirm them.

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

want to engage with -- and, to be frank, try to refute -- some ery ingenious and interesting proposals due to Wallace and Greaves-Myrvold, who try to explain how to make sense of robability in many-worlds theories and then explain how to test nd confirm them.

nd then I want to describe a slightly non-standard way of inking about one-world probabilistic theories, which allows us to ake and confirm or refute predictions without running into roblems in interpreting probability.

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

want to engage with -- and, to be frank, try to refute -- some very genious and interesting proposals due to Wallace and Greaves-lyrvold, who try to explain how to make sense of probability in many-orlds theories and then explain how to test and confirm them.

n another occasion (see e.g. forthcoming Everett book) I plan to escribe a slightly non-standard way of thinking about one-world robabilistic theories, which allows us to make and confirm or refute redictions without running into problems in interpreting probability.

or details of Wallace's, Greaves-Myrvold's arguments, critical esponses by Price, Albert, A.K. and others, and many other papers, ee forthcoming O.U.P. volume "Many Worlds?" (Saunders, Barrett, .K., Wallace, eds), which includes extended versions of papers given t last summer's PI "Many Worlds at 50" and Oxford "Everett at 50" onferences.

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

The arguments I want to make here aren't specific to quantum theory*: as Myrvold and Greaves stress, if we're going to take any many-worlds theories seriously then we need a way of testing and confirming or refuting general many-worlds theories, just as we do for one-worlds theories.

It's also much simpler to run all the arguments in toy many-worlds theories, so I will.

(* So, Everettian arguments that rely on the specific structure of quantum theory won't be covered in this talk. But they are elsewhere.)

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

The arguments I want to make here aren't specific to quantum theory*: as Myrvold and Greaves stress, if we're going to take any many-worlds theories seriously then we need a way of testing and confirming or refuting general many-worlds theories, just as we do for one-worlds theories.

It's also much simpler to run all the arguments in toy many-worlds theories, so I will.

(* So, Everettian arguments that rely on the specific structure of quantum theory won't be covered in this talk. But they are elsewhere.)

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

Importance: the Many-Worlds Analogue of Probability

To make sense of Everett and other many-worlds heories, we need to allow that different future branches may ave different "importance weights" which play a role nalogous to probabilities in one-world theories.

Importance: the Many-Worlds Analogue of Probability

To make sense of Everett and other many-worlds heories, we need to allow that different future branches may ave different "importance weights" which play a role nalogous to probabilities in one-world theories.

All future branches will be realised, but for the sake of the iscussion we allow the possibility that physics implies that ome are more important to us than others. (Without this, verettians can't explain why they worry about events in high orn weight future branches more than low weight ones.)

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Importance: the Many-Worlds Analogue of Probability

To make sense of Everett and other many-worlds heories, we need to allow that different future branches may ave different "importance weights" which play a role analogous to probabilities in one-world theories.

All future branches will be realised, but for the sake of the iscussion we allow the possibility that physics implies that ome are more important to us than others. (Without this, Everettians can't explain why they worry about events in high Born weight future branches more than low weight ones.)

We leave open for now whether importance weight is a hysical postulate (bizarre though *that* seems) or something xplained by the structure of the many-worlds theory.

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

Toy Many-Worlds Theories

'hese toy theories might initially seem a little intellectually nsettling -- but they turn out to be quite illuminating.

Theory Confirmation in One World and its - Windows Journal

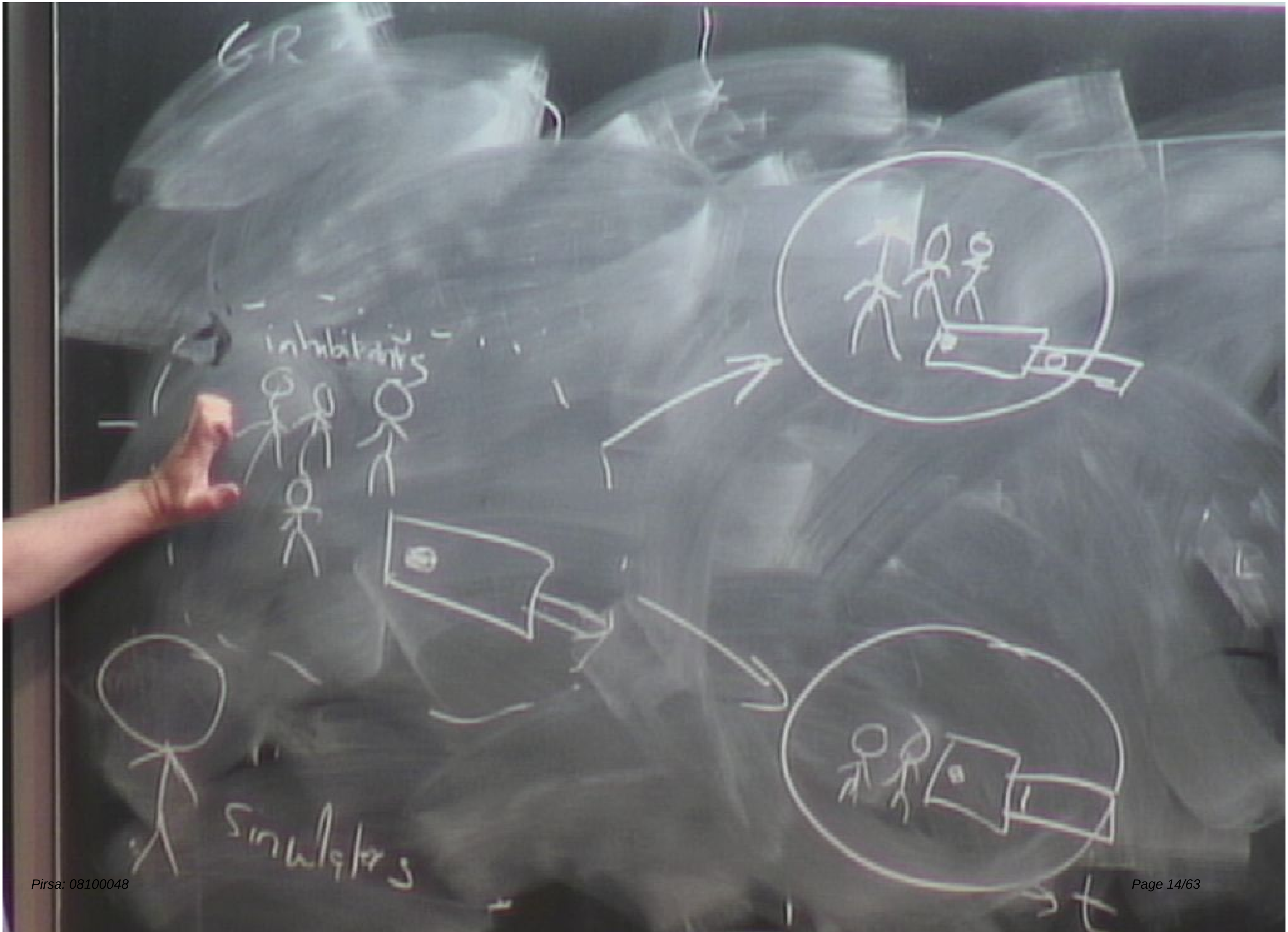
e Edit View Insert Actions Tools Help

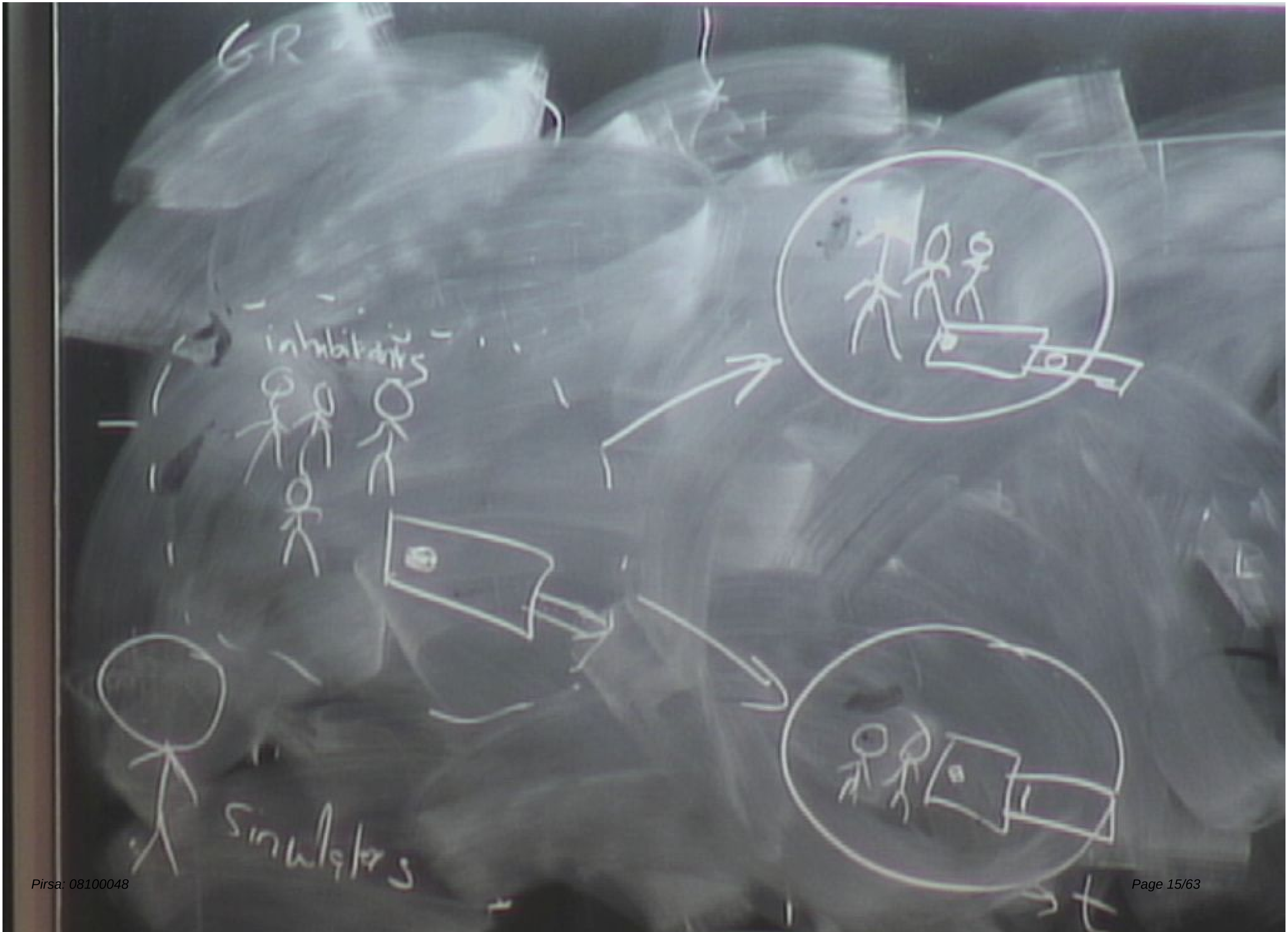
Reading View

Toy Many-Worlds Theories

'hese toy theories might initially seem a little intellectually nsettling -- but they turn out to be quite illuminating.

Ve imagine a simulated classical universe, containing a nachine with a red button and a tape. Each time the nhabitants press the button, the simulation is deleted, and wo more are created, identical to the original except that the ape now has an extra 0 in one simulation and 1 in the other


GR
inhabitants
Simulators


GR
inhabitants
simulators

Toy Many-Worlds Theories

hese toy theories might initially seem a little intellectually nsettling -- but they turn out to be quite illuminating.

Ve imagine a simulated classical universe, containing a nachine with a red button and a tape. Each time the nhabitants press the button, the simulation is deleted, and wo more are created, identical to the original except that the ape now has an extra 0 in one simulation and 1 in the other.

here may or may not be something in the physics which could give the inhabitants some reason to assign different mportance weights to the two outcome branches. They correctly) believe they are living in a branching universe, and re interested in trying to develop and test theories about the mportance weight.

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

Toy Many-Worlds Theories

hese toy theories might initially seem a little intellectually nsettling -- but they turn out to be quite illuminating.

Ve imagine a simulated classical universe, containing a nachine with a red button and a tape. Each time the nhabitants press the button, the simulation is deleted, and wo more are created, identical to the original except that the ape now has an extra 0 in one simulation and 1 in the other.

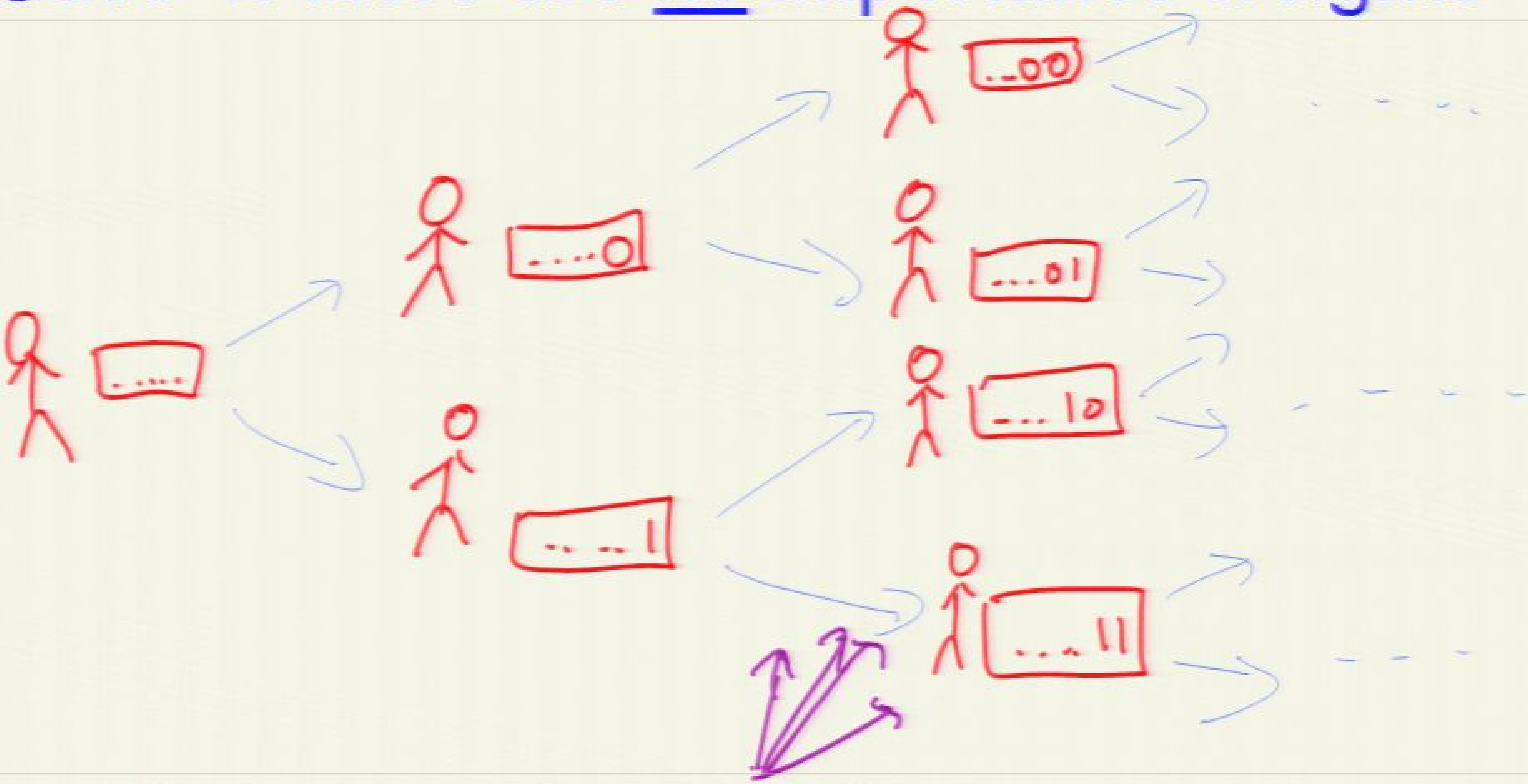
here may or may not be something in the physics which could give the inhabitants some reason to assign different mportance weights to the two outcome branches. They correctly) believe they are living in a branching universe, and re interested in trying to develop and test theories about the mportance weight.

heory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

Case 1: there are no importance weights



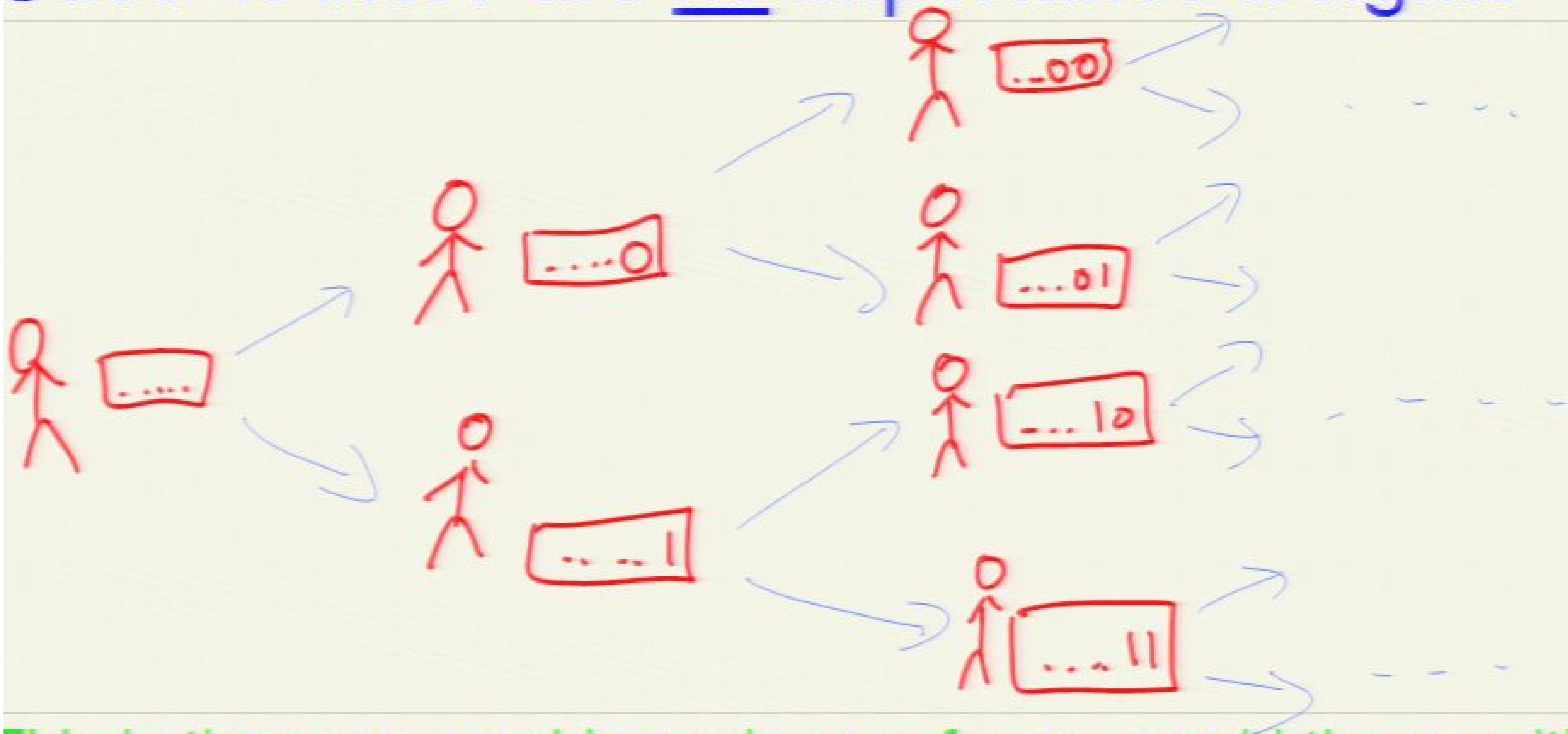
nhabitants developing and testing theories about
mportance weights (which as it happens don't exist)

Theory Confirmation in One World and its - Windows Journal

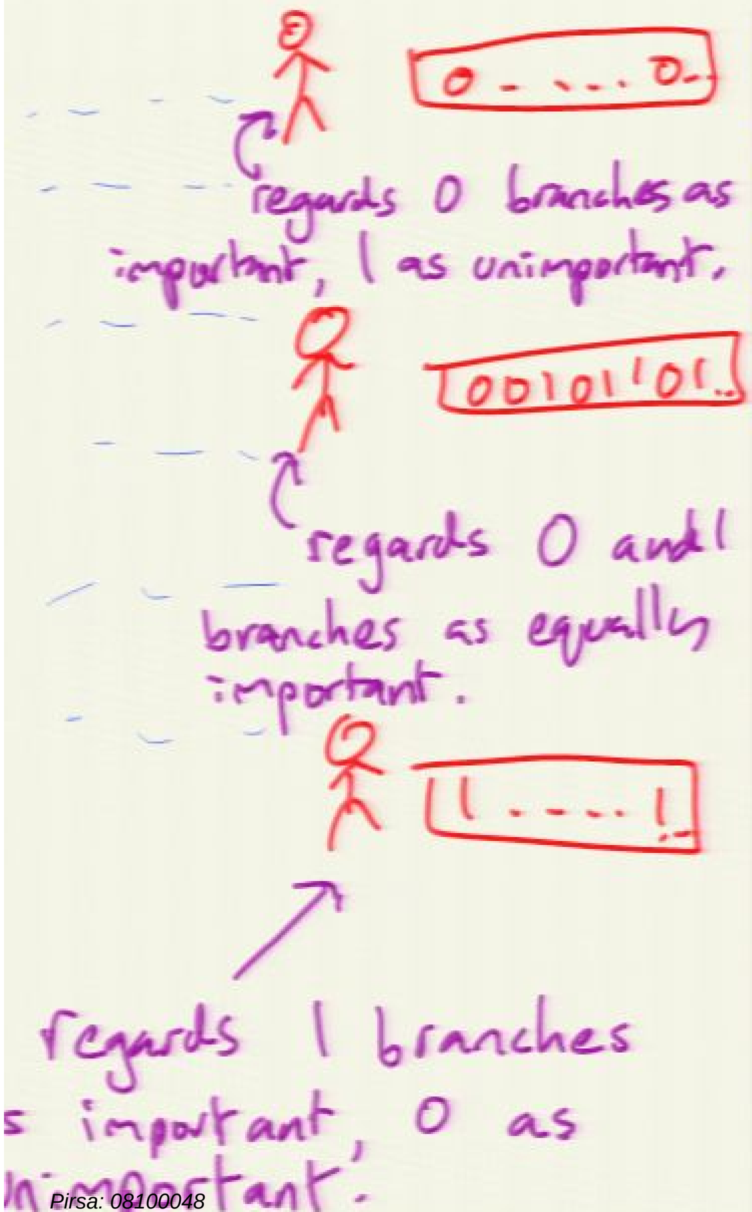
Edit View Insert Actions Tools Help

Reading View

Case 1: there are no importance weights

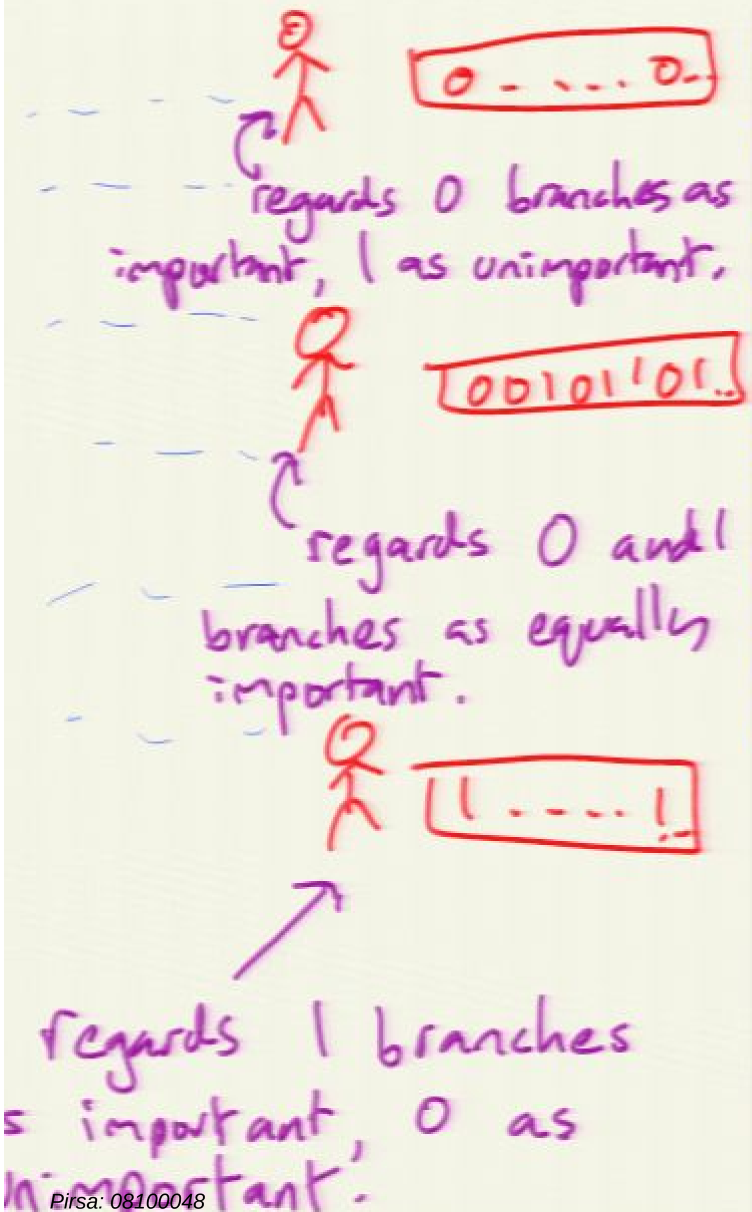


This is the many-worlds analogue of a one-world theory with a random tape for which there are no laws, neither deterministic nor probabilistic.



In the long run, eacn inhabitant tends to the theory that the importance weights equal the relative frequencies on the tape in their branch

Theory Confirmation in One World and its - Windows Journal

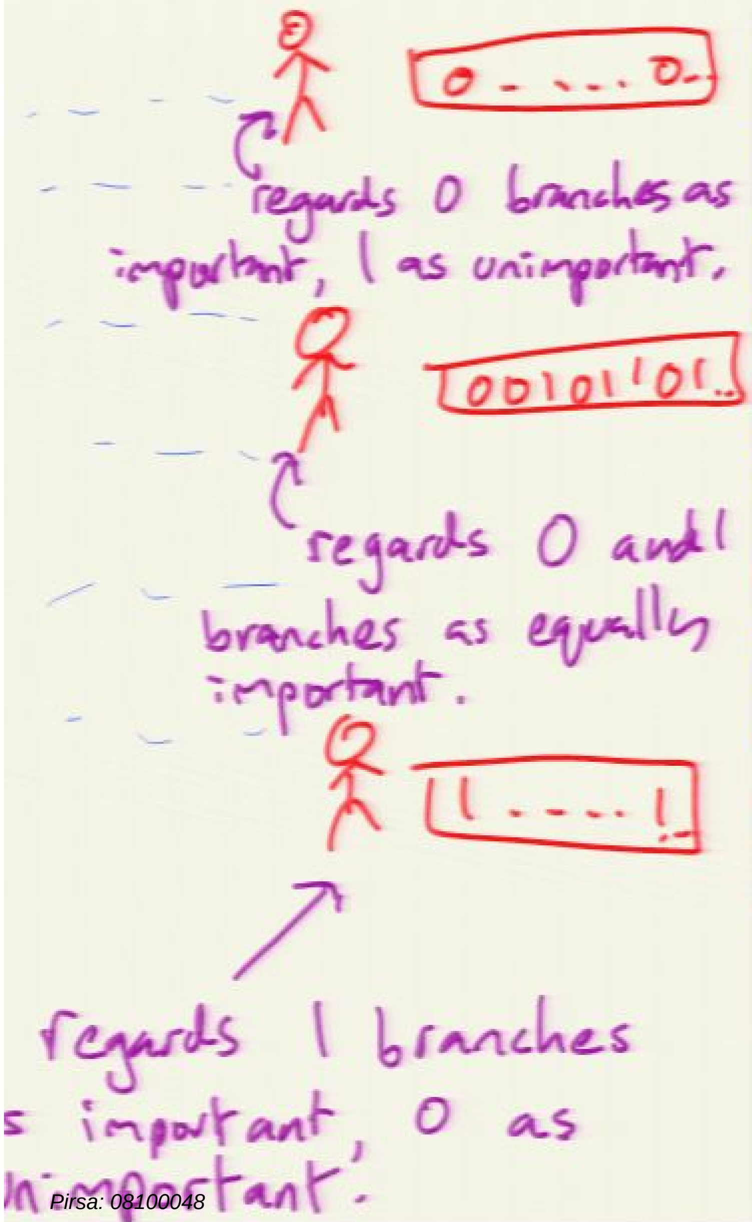


In the long run, eacn inhabitant tends to the theory that the importance weights equal the relative frequencies on the tape in their branch -- just as Everettians regard the observation of Born weight relative frequencies as confirming that Born weights define importance of future branches.

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

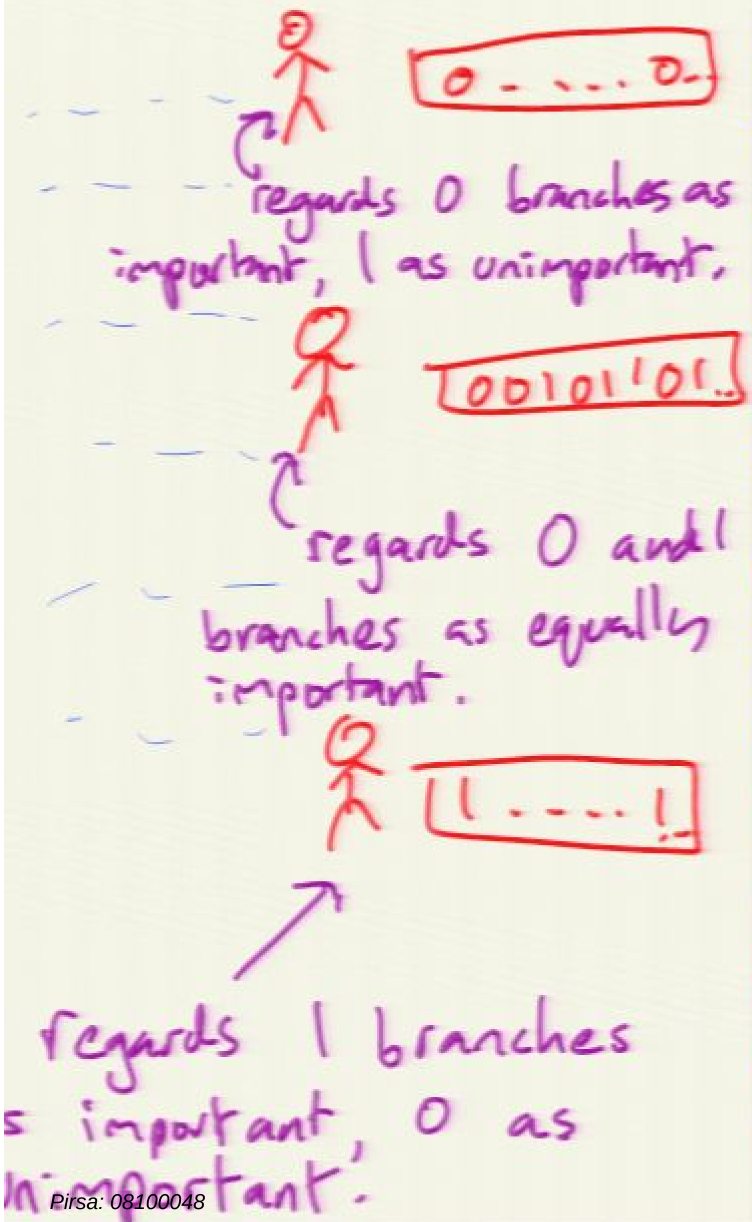


In the long run, eacn inhabitant tends to the theory that the importance weights equal the relative frequencies on the tape in their branch -- just as Everettians regard the observation of Born weight relative frequencies as confirming that Born weights define importance of future branches. We can derive this formally using Greaves-Myrvold's analysis.

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View



In the long run, eacn inhabitant tends to the theory that the importance weights equal the relative frequencies on the tape in their branch -- just as Everettians regard the observation of Born weight relative frequencies as confirming that Born weights define importance of future branches. We can derive this formally using Greaves-Myrvold's analysis. (But it's a priori pretty clear it must work this way if at all here: the only data the inhabitants have are the observed relative frequencies.)

An inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the importance weights are also (p,1-p).

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

An inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the importance weights are also (p,1-p). She knows that there will be inhabitants who see all possible relative frequencies.

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

An inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the importance weights are also (p,1-p). She knows that there will be inhabitants who see all possible relative frequencies. But this, if she takes Everettian reasoning seriously, isn't supposed to shake her faith in the theory, because if the importance weights are (p,1-p) then the importance of a branch with r zeroes and (N-r) ones is p^r (1-p)^(N-r), and the branches with relative frequencies close to (p,1-p) dominate according to the importance measure.

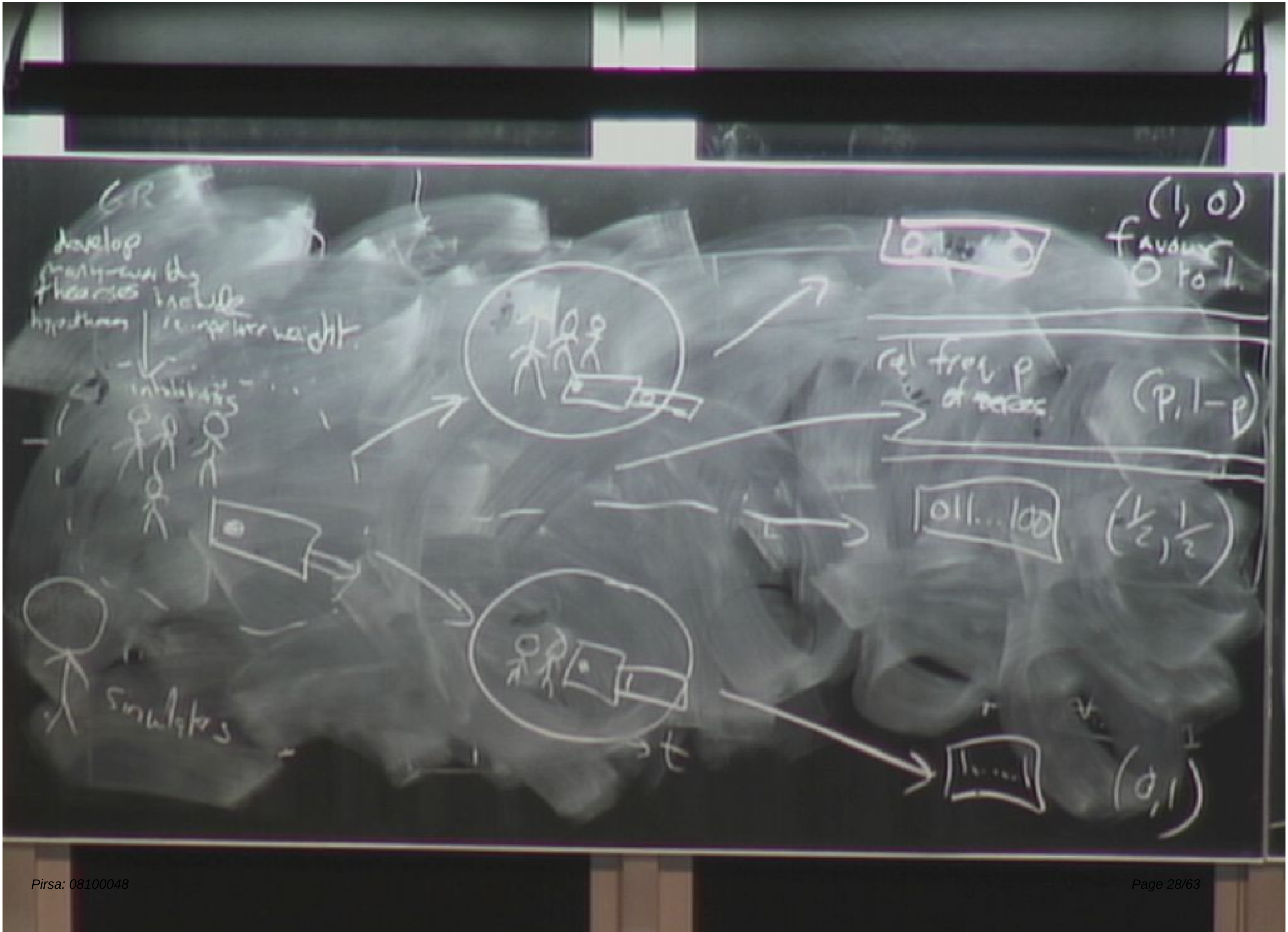
$(p, 1-p)$ wt. believe

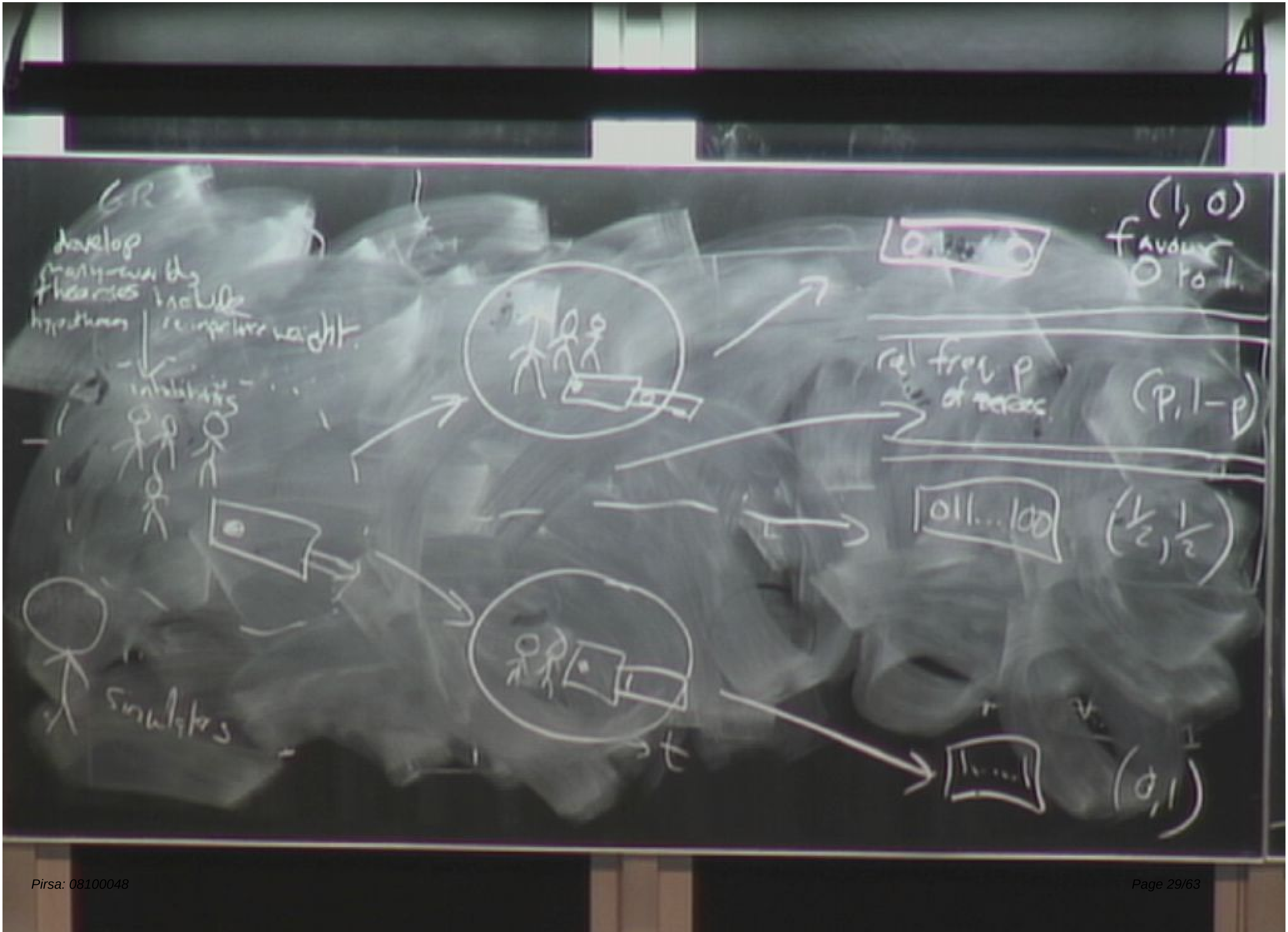
reasoning:

importance of branch

with r zeros, (N-r) ones

$\sim p^r (1-p)^{N-r}$


GR
develop
theories
include
simulators
(1, 0)
favour
0 to 1
rel freq p
of theories
(p, 1−p)
011...100
(1/2, 1/2)
(0, 1)


GR
develop
theories
include
inhabitants
simulators
(1, 0)
favour
0 to 1
(p, 1-p)
011...100
(1/2, 1/2)
(0, 1)

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

An inhabitant who sees relative frequencies (p,1-p) of zeroes and
ones after N branches tends to the theory that the importance
weights are also (p,1-p). She knows that there will be
inhabitants who see all possible relative frequencies. But this, if
she takes Everettian reasoning seriously, isn't supposed to
shake her faith in the theory, because if the importance weights
are (p,1-p) then the importance of a branch with r zeroes and
(N-r) ones is p^r (1-p)^(N-r), and the branches with relative
frequencies close to (p,1-p) dominate according to the
importance measure.

Remember, this is true for all values of p. Everyone thinks their
theory of importance weights is confirmed. Everyone thinks
they, and observers who see relative frequencies similar to
theirs, are the important branches, and the rest are negligible.

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

An inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the importance weights are also (p,1-p). She knows that there will be inhabitants who see all possible relative frequencies. But this, if she takes Everettian reasoning seriously, isn't supposed to shake her faith in the theory, because if the importance weights are (p,1-p) then the importance of a branch with r zeroes and (N-r) ones is p^r (1-p)^(N-r), and the branches with relative frequencies close to (p,1-p) dominate according to the importance measure.

Remember, this is true for all values of p. Everyone thinks their theory of importance weights is confirmed. Everyone thinks they, and observers who see relative frequencies similar to theirs, are the important branches, and the rest are negligible.

Theory Confirmation in One World and its - Windows Journal

File Edit View Insert Actions Tools Help

Reading View

An inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the importance weights are also (p,1-p). She knows that there will be inhabitants who see all possible relative frequencies. But this, if she takes Everettian reasoning seriously, isn't supposed to shake her faith in the theory, because if the importance weights are (p,1-p) then the importance of a branch with r zeroes and (N-r) ones is p^r (1-p)^(N-r), and the branches with relative frequencies close to (p,1-p) dominate according to the importance measure.

Remember, this is true for all values of p. Everyone thinks their theory of importance weights is confirmed. Everyone thinks they, and observers who see relative frequencies similar to theirs, are the important branches, and the rest are negligible. (And recall, by hypothesis, they're all wrong!)

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

This phenomenon of spurious confirmation is already rather worryng for many-worlders. If we're bound to find a coherent theory confrimed even in a multiverse in which no such theory applies, how seriously can we take any confirmation?

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

Note: of course we'll get exactly the same phenomenon even if there are importance weights attached to the branches: the only difference is that a subset of the inhabitants will end up with the "right" answer.

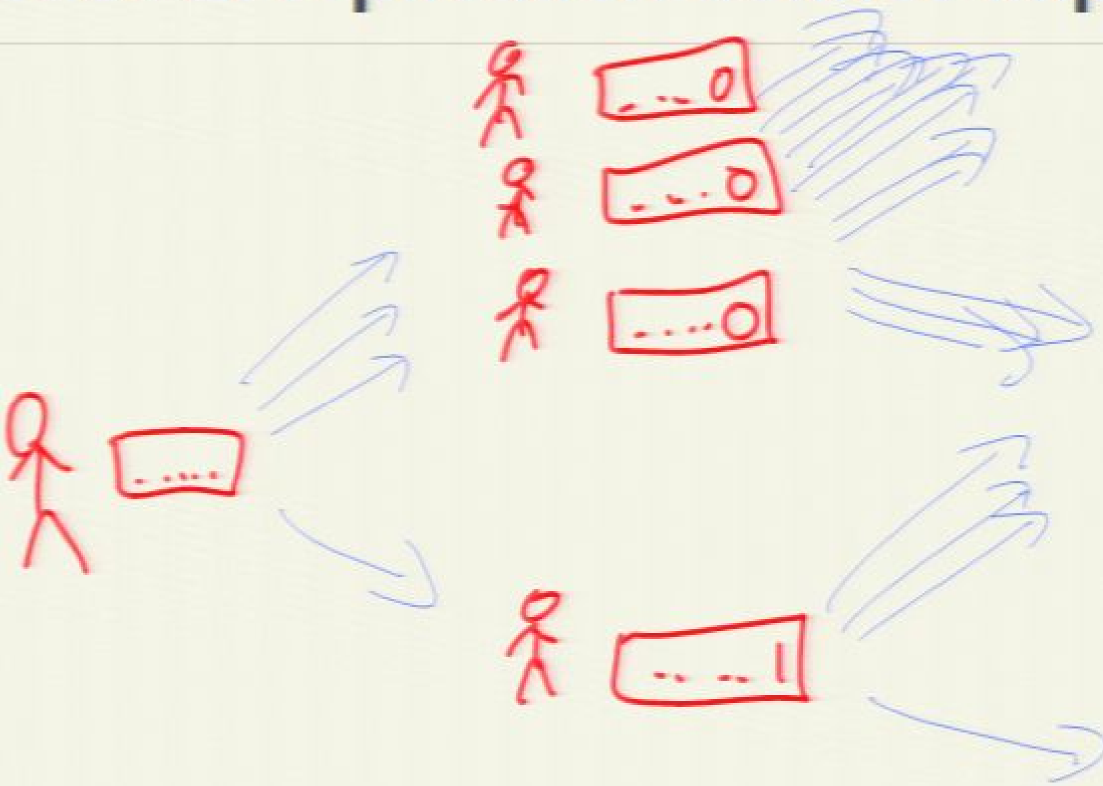
Let's now look into this more closely....

Theory Confirmation in One World and its - Windows Journal

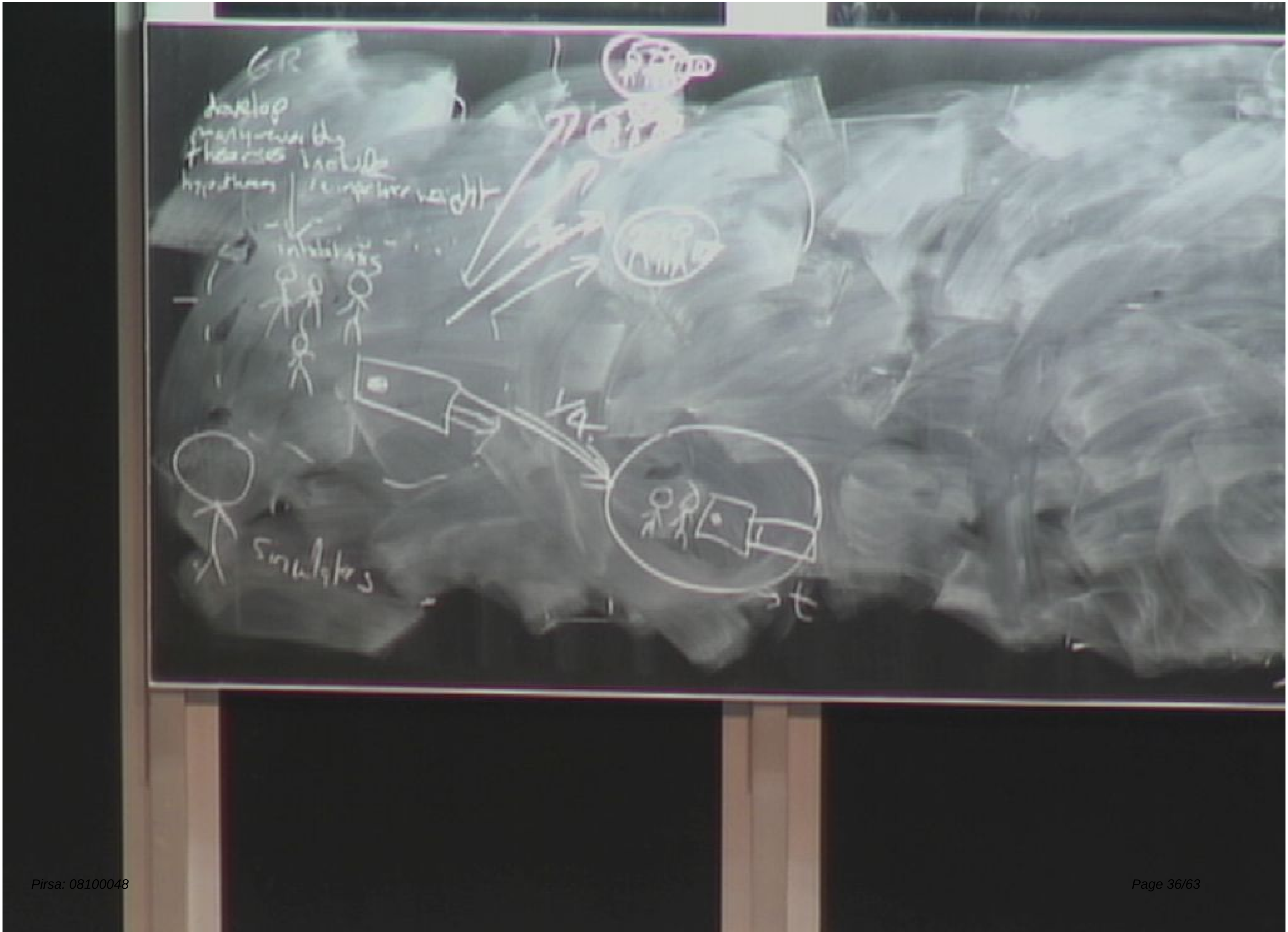
e Edit View Insert Actions Tools Help

Reading View

Case 2: replication defines importance weights



his time, the simulators make three new identical universes with ach 0 outcome, and as before just one with 1 outcome. The habitants don't know this -- but arguably, if they did, they should ow assign importance weights (3/4,1/4) to (0,1).


GR
develop
Simulates

Again, an inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the importance weights are also (p,1-p). In particular, inhabitants who see relative frequencies close to (3/4,1/4) tend to the theory that the importance weights are close to (3/4,1/4).

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

Again, an inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the importance weights are also (p,1-p). In particular, inhabitants who see relative frequencies close to (3/4,1/4) tend to the theory that the importance weights are close to (3/4,1/4).

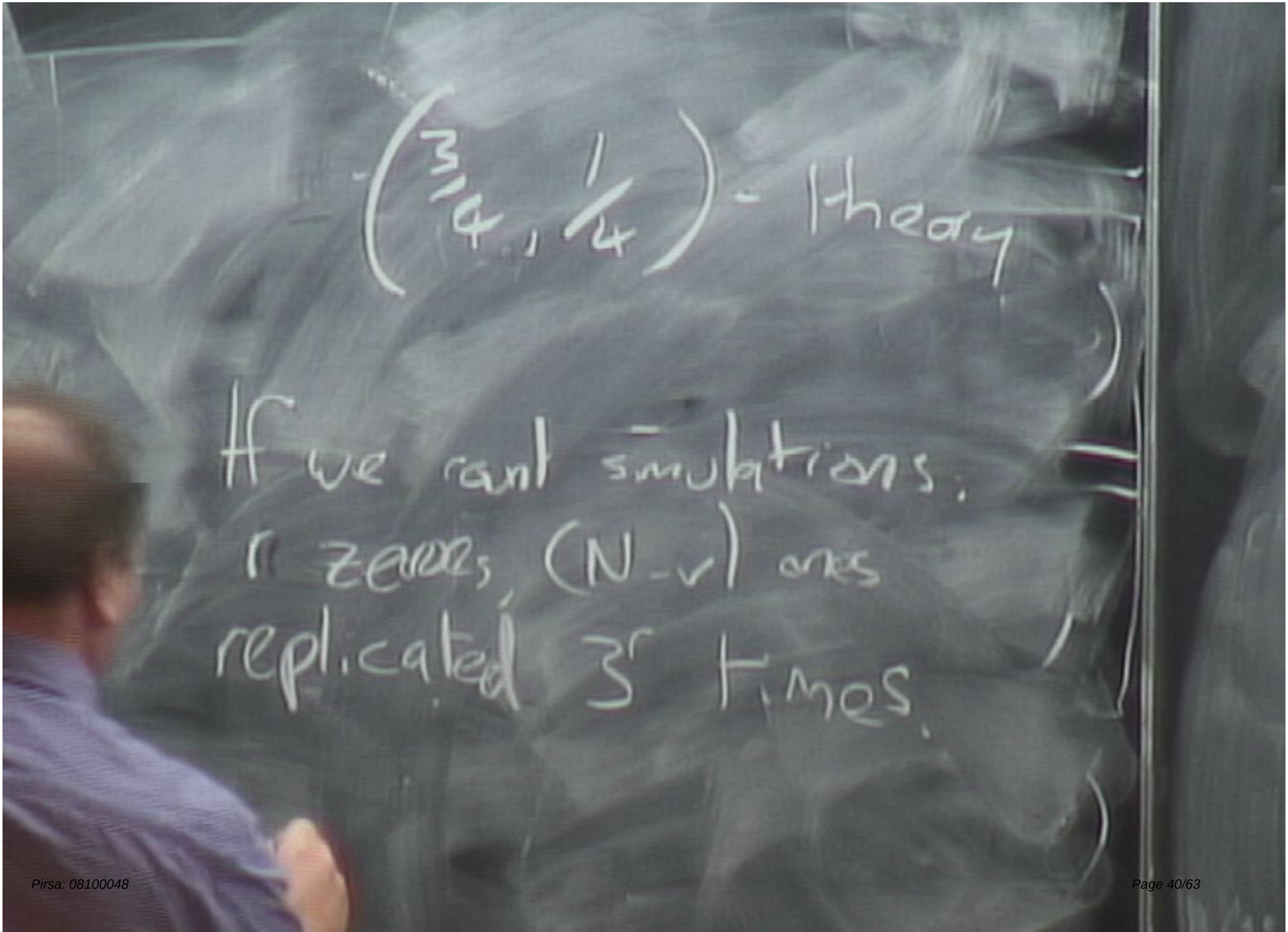
Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

Again, an inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the mportance weights are also (p,1-p). In particular, inhabitants who see relative frequencies close to (3/4,1/4) tend to the theory hat the importance weights are close to (3/4,1/4).

This time, if we count the simulations as equally important, these nhabitants are right -- and moreover, they dominate in the nultiverse, according to the counting measure. By that neasure, almost all inhabitants arrive at close to the (arguably) correct importance weights in the long run.


(3/4, 1/4) - theory
If we want simulations:
r zeros, (N-r) ones
replicated 3^r times.

$\left(\frac{3}{4}, \frac{1}{4}\right)$ - theory

If we want simulations,
r zeros, $(N-r)$ ones
replicated 3^r times.

$3^r \binom{N}{r}$ sims with $(r, N-r)$ $(0,1)$'s.

Again, an inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the importance weights are also (p,1-p). In particular, inhabitants who see relative frequencies close to (3/4,1/4) tend to the theory that the importance weights are close to (3/4,1/4).

This time, if we count the simulations as equally important, these inhabitants are <u>right</u> -- and moreover, they <u>dominate</u> in the multiverse, according to the counting measure. By that measure, almost all inhabitants arrive at close to the (arguably) correct importance weights in the long run.

Theory Confirmation in One World and its - Windows Journal

File Edit View Insert Actions Tools Help

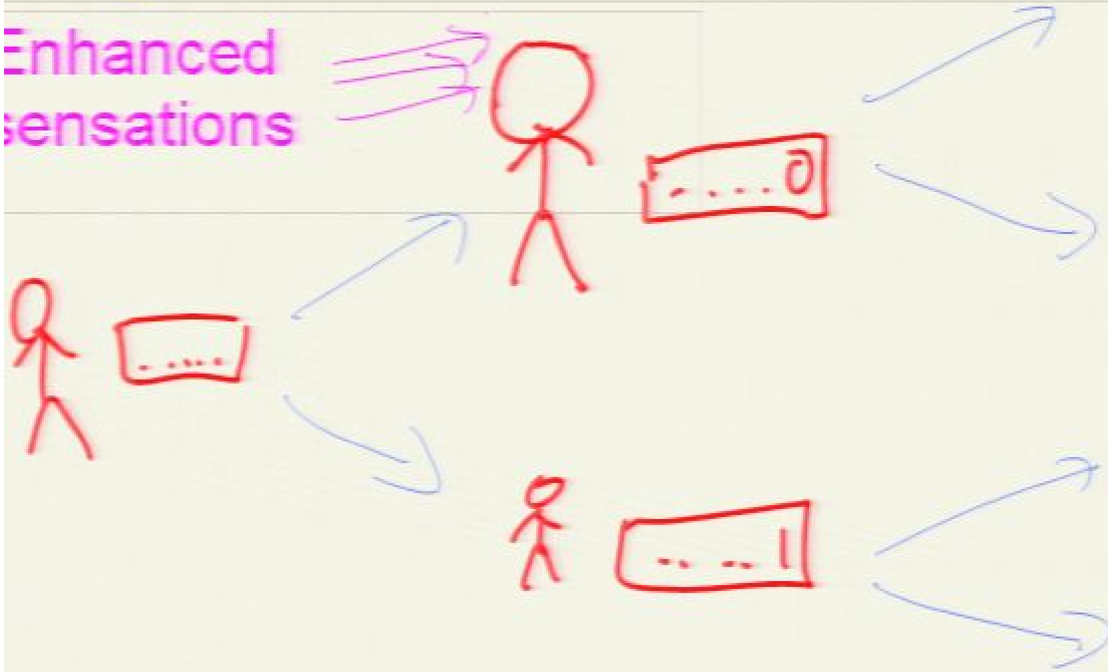
Reading View

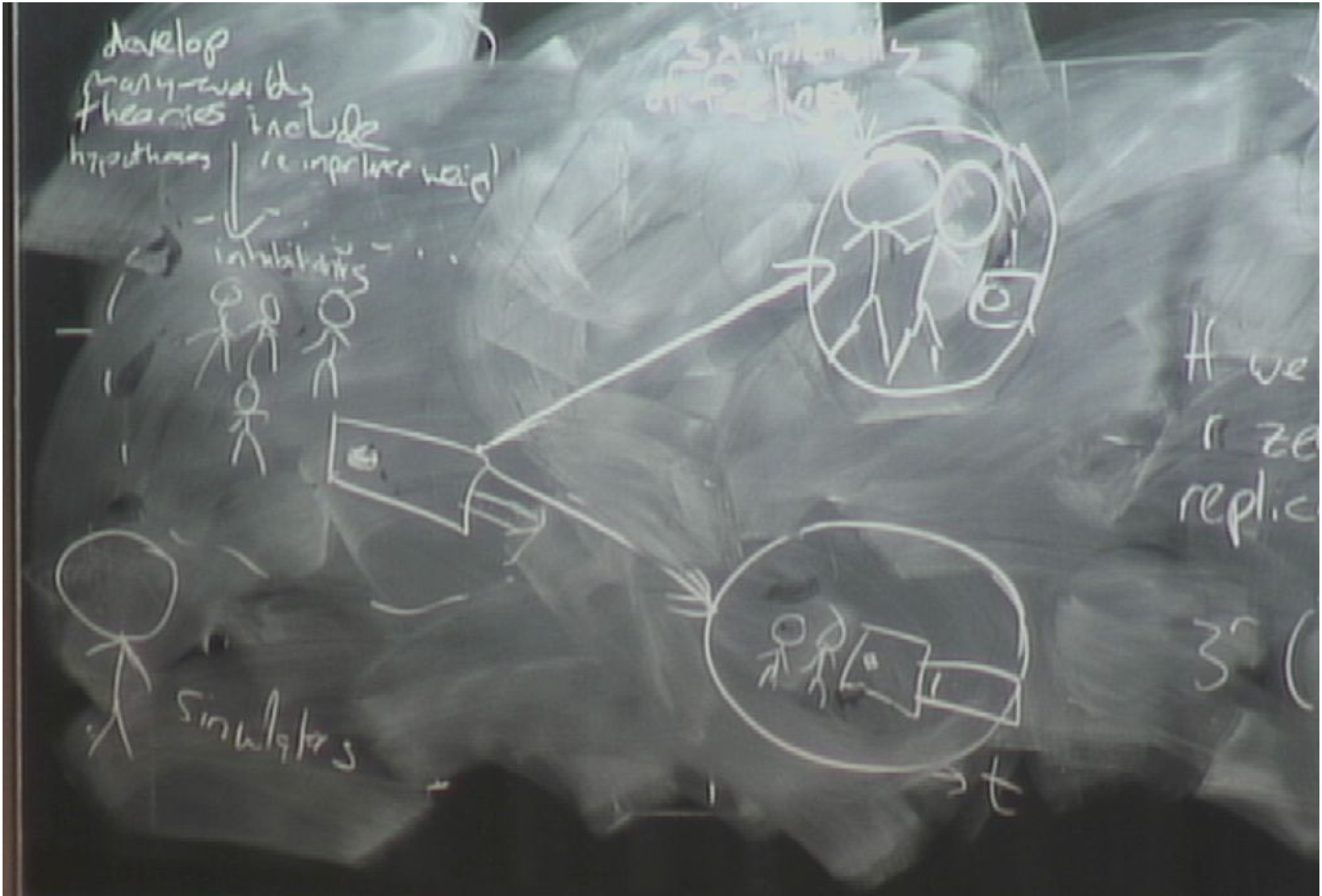
So, by the counting measure, almost all inhabitants arrive at close to (what are arguably) the right importance weights.

This still seems to me like a stage in an argument rather than the end of one. We seem now to need to say something anthropic -- that if you're an inhabitant of the multiverse, there's a very high chance of your being among the ones who finds the right weights and thus the right theory. I don't have anything new to say here about the puzzles of anthropic reasoning -- the intuition seems clear here, but seems hard to justify rigorously.

So, more discussion is needed, but still, we seem to have made interesting progress in this example. Perhaps many-worlds theory confirmation could sometimes work - or at least get somewhere closer to working - after all?

Case 3: sensation enhancement defines importance weights

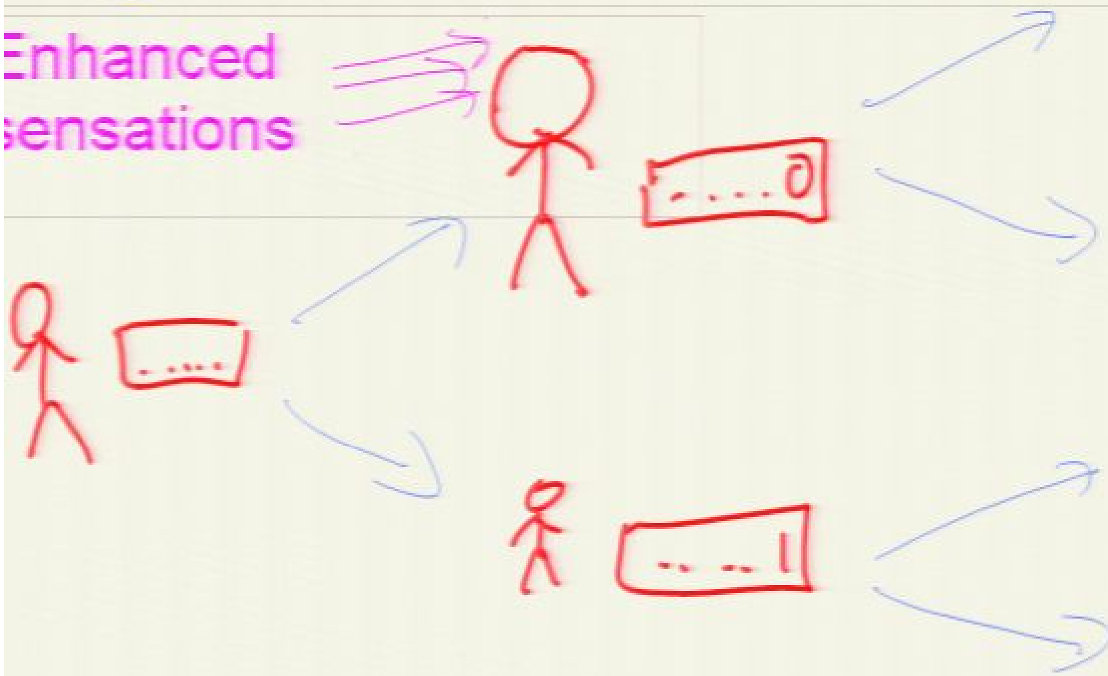



develop
theories include
hypotheses
inhabitants
simulators
replic

Edit View Insert Actions Tools Help

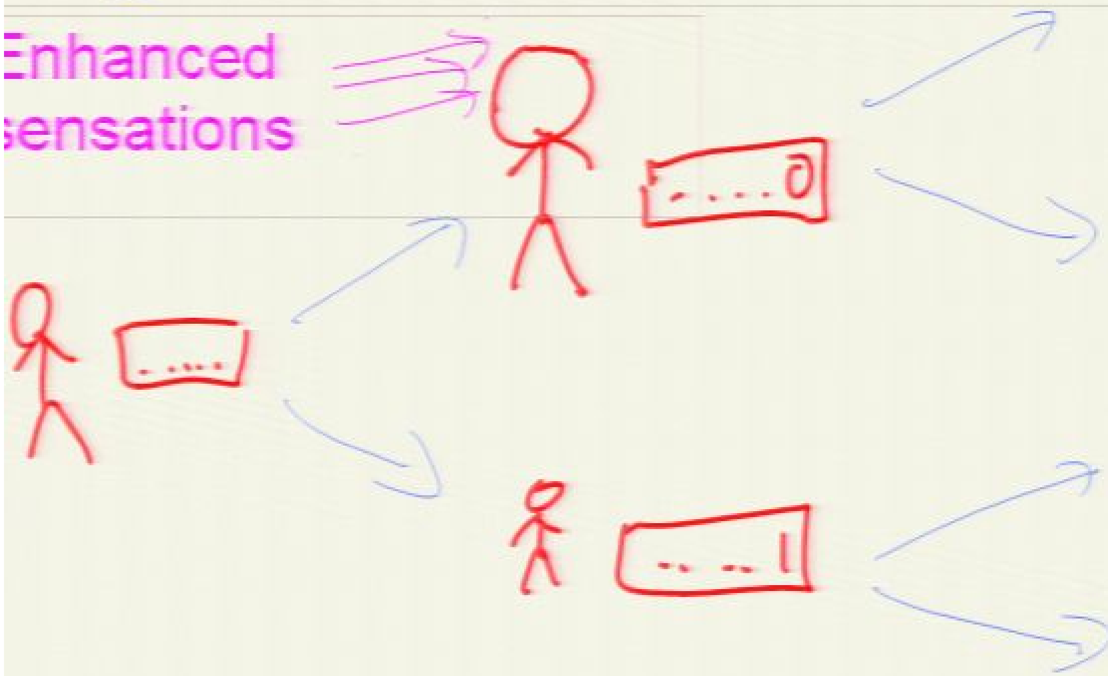
Reading View

Case 3: sensation enhancement defines importance weights



The simulators arrange things so that the inhabitants' sensations are three times as rich on the 0 branches as the 1 branches.

Case 3: sensation enhancement defines importance weights



The simulators arrange things so that the inhabitants' sensations are three times as rich on the 0 branches as the 1 branches.

he inhabitants don't know what's going on -- but again, rguably, if they did, they should assign importance weights 3/4,1/4) to (0,1) branches. Winning a donut brings three times ne pleasure on 0 branches as on 1 branches, and so on.

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

Once again, an inhabitant who sees relative frequencies (p,1-p) of eroes and ones after N branches tends to the theory that the nportance weights are also (p,1-p). In particular, inhabitants who ee relative frequencies close to (1/2,1/2) tend to the theory that the nportance weights are close to (1/2,1/2).

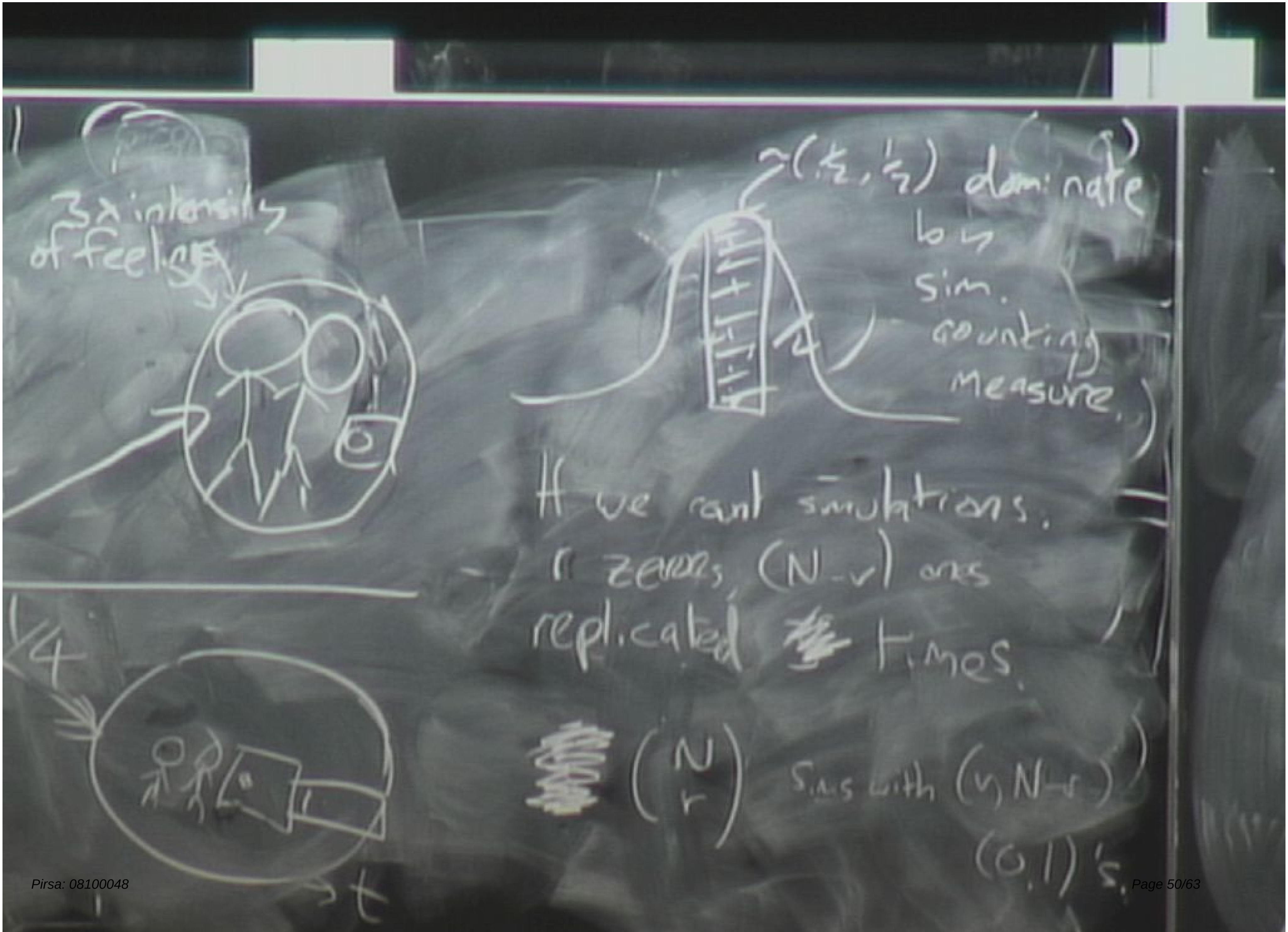
Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

Once again, an inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the importance weights are also (p,1-p). In particular, inhabitants who see relative frequencies close to (1/2,1/2) tend to the theory that the importance weights are close to (1/2,1/2).

This time, it's <u>these</u> inhabitants who dominate in the multiverse, according to the simulation counting measure. By that measure, almost all inhabitants tend to arrive at the <u>wrong</u> importance weights.



3λ intensity of feelings
~(½, ½) dominate
by sim. counting measure.
we want simulations,
r zeros, (N-r) ones
replicated times.
(N r)
sims with (r, N-r)
(0,1)'s.

)nce again, an inhabitant who sees relative frequencies (p,1-p) of eroes and ones after N branches tends to the theory that the nportance weights are also (p,1-p). In particular, inhabitants who ee relative frequencies close to (1/2,1/2) tend to the theory that the nportance weights are close to (1/2,1/2).

'his time, it's <u>these</u> inhabitants who dominate in the multiverse, ccording to the simulation counting measure. By that measure, lmost all inhabitants tend to arrive at the <u>wrong</u> importance weights.

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

)nce again, an inhabitant who sees relative frequencies (p,1-p) of eroes and ones after N branches tends to the theory that the nportance weights are also (p,1-p). In particular, inhabitants who ee relative frequencies close to (1/2,1/2) tend to the theory that the nportance weights are close to (1/2,1/2).

'his time, it's these inhabitants who dominate in the multiverse, ccording to the simulation counting measure. By that measure, lmost all inhabitants tend to arrive at the wrong importance weights.

low it's true we could instead count simulations weighted by intensity f sensation, and of course that would give a different answer: the 3/4,1/4) frequency inhabitants dominate again. But there's no really ompelling reason to use this rule: anthropic intuitions certainly don't orce it.

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

Once again, an inhabitant who sees relative frequencies (p,1-p) of zeroes and ones after N branches tends to the theory that the importance weights are also (p,1-p). In particular, inhabitants who see relative frequencies close to (1/2,1/2) tend to the theory that the importance weights are close to (1/2,1/2).

This time, it's these inhabitants who dominate in the multiverse, according to the simulation counting measure. By that measure, almost all inhabitants tend to arrive at the wrong importance weights.

Now it's true we could instead count simulations weighted by intensity of sensation, and of course that would give a different answer: the (3/4,1/4) frequency inhabitants dominate again. But there's no really compelling reason to use this rule: anthropic intuitions certainly don't force it.

We have here two inequivalent branch measures.

What have we learnt?

This, I think -- there are two logically distinct notions of importance, not separated in the Everettian literature:

What have we learnt?

This, I claim -- there are two logically distinct notions of importance, not separated in the Everettian literature:

"caring weight" -- how much one should (given some assumptions) care about events on a given branch.

"confirmation weight" -- how much weight should (given some other assumptions) be attached to a branch when assessing the success or failure of theory testing and confirmation.

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

For many-worlds confirmation theory to work, we need our many-vorlds theories to be somehow equipped with a natural notion of ranch confirmation weight (not just caring weight).

But a novel rule for theory confirmation doesn't seem something one can be allowed just to postulate in a scientific theory. (Consider my new one-world theory which includes a postulate that the important people for confirming the theory are those who agree with my observations and my theoretical interpretation. It's self-consistent -- but it surely isn't science.)

However, to have any hope of deriving a confirmation weight, we can't appeal to a purported derivation of caring weight (even if one were accepted). We would need a separate justification: for instance, that underlying Everettian quantum theory is some "many-minds" model analogous to the replicant multiverse -- precisely the sort of ad hoc structure that all decent Everettians wish and purport to avoid.

e Edit View Insert Actions Tools Help

Reading View

Summary for Everettians:

Even if you can
(i) find a principled justification of $|\psi|^2$ as
caring weight - of "betting" in many-worlds as
though Born weights were probabilities

Summary for Everettians:

Even if you can
(i) find a principled justification of $|\psi|^2$ as caring weight - of "betting" in many-worlds as though Born weights were probabilities - you still need a separate argument

Theory Confirmation in One World and its - Windows Journal

e Edit View Insert Actions Tools Help

Reading View

Summary for Everettians:

Even if you can
(i) find a principled justification of $|\psi|^2$ as
caring weight - of "betting" in many-worlds as
though Born weights were probabilities - you
still need a separate argument for
(ii) treating $|\psi|^2$ as a confirmation weight - i.e.
neglecting atypical low $|\psi|^2$ branches on which
non-Born weight statistics are observed.

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

Summary for Everettians:

Even if you can
(i) find a principled justification of $|\psi|^2$ as
bearing weight - of "betting" in many-worlds as
though Born weights were probabilities - you
still need a separate argument for
(ii) treating $|\psi|^2$ as a confirmation weight - i.e.
neglecting atypical low $|\psi|^2$ branches on which
non-Born weight statistics are observed.
Few people believe even (i) is possible - but some do:
cf. Wallace, Deutsch.

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

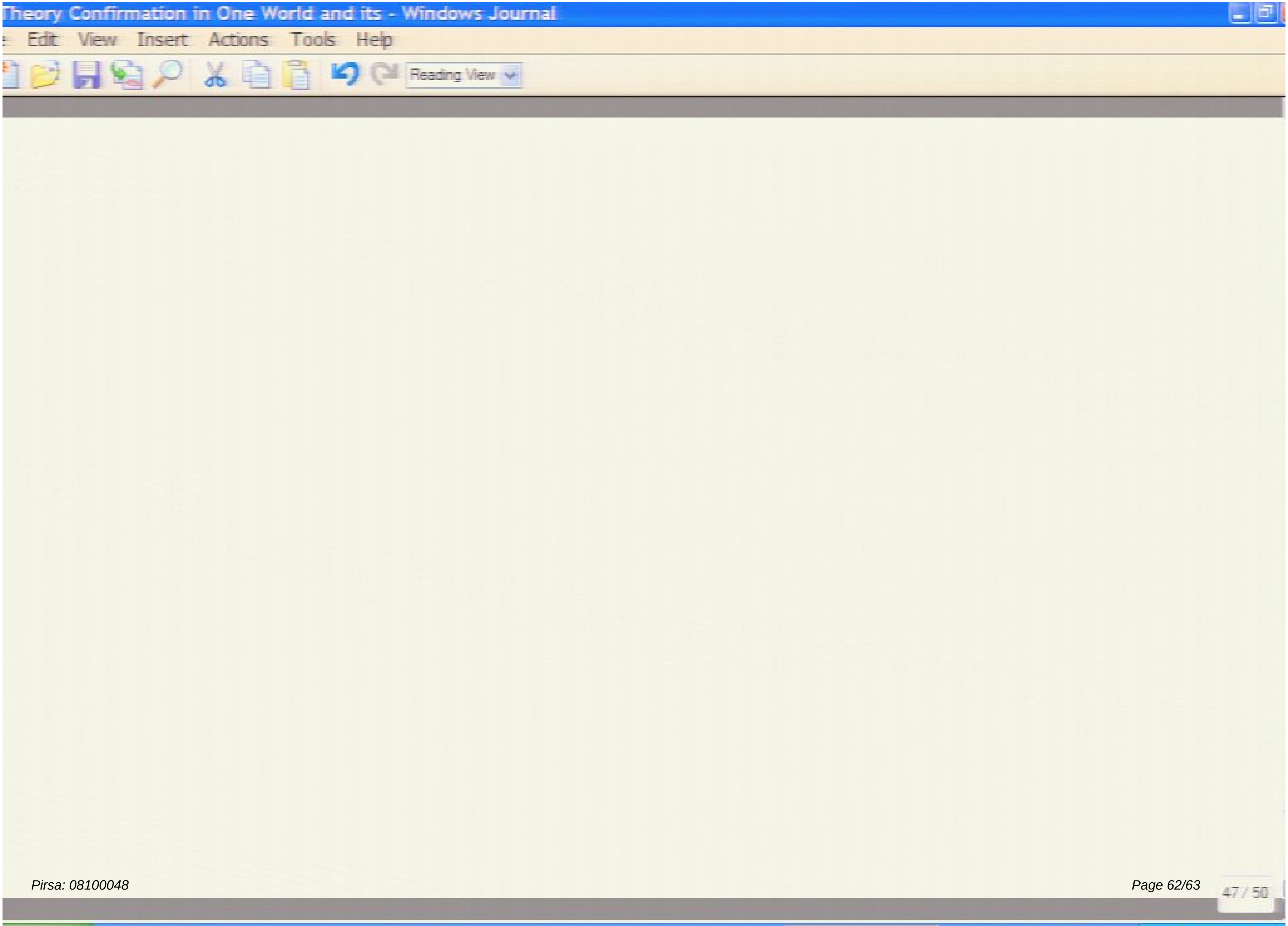
Reading View

Summary for Everettians:

Even if you can
(i) find a principled justification of $|\psi|^2$ as bearing weight - of "betting" in many-worlds as though Born weights were probabilities - you still need a separate argument for
(ii) treating $|\psi|^2$ as a confirmation weight - i.e. neglecting atypical low $|\psi|^2$ branches on which non-Born weight statistics are observed.

Few people believe even (i) is possible - but some do: cf. Wallace, Deutsch.

No one, as far as I am aware, has separately addressed (ii)

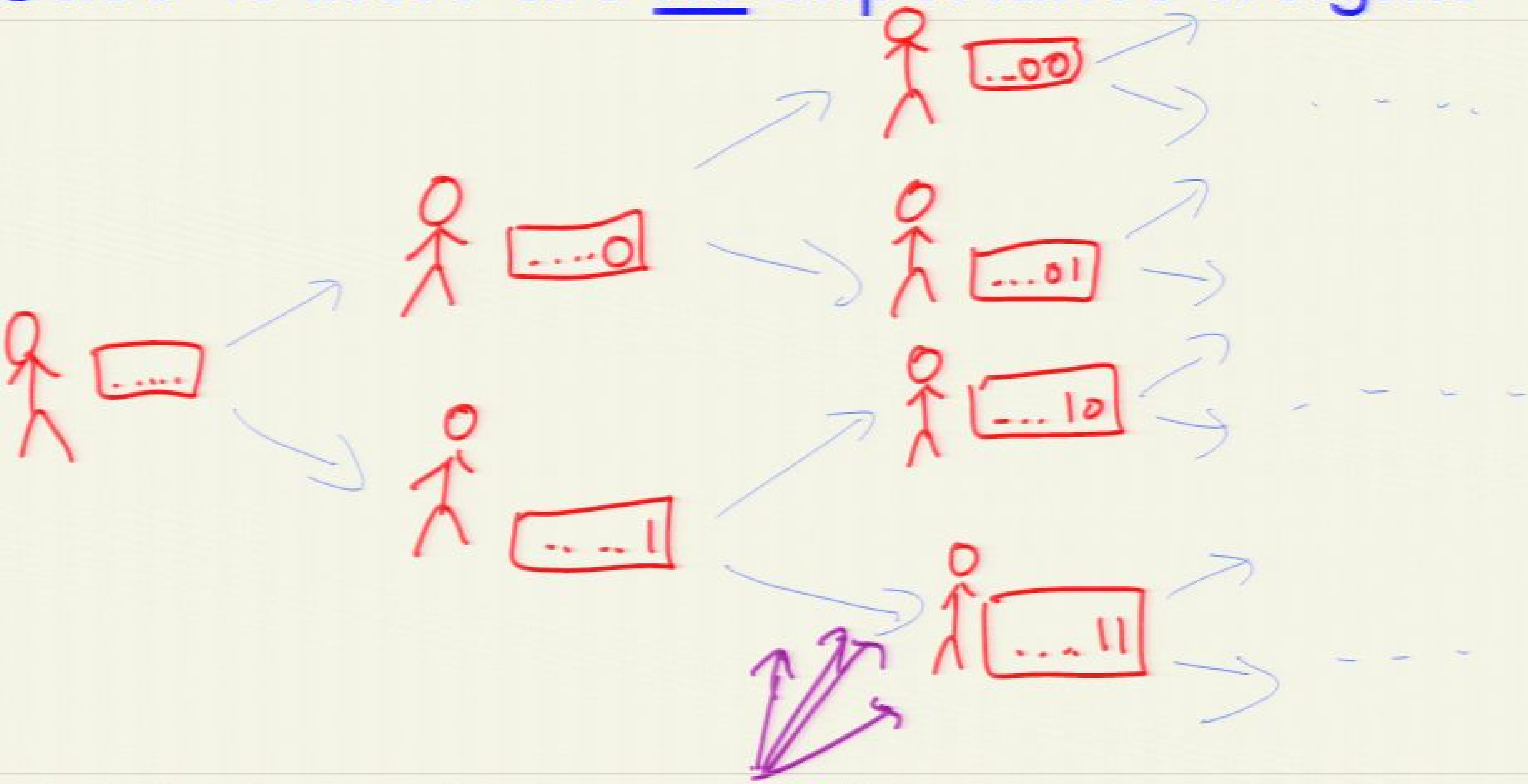

Theory Confirmation in One World and its - Windows Journal
Edit View Insert Actions Tools Help
Reading View
47 / 50

Theory Confirmation in One World and its - Windows Journal

Edit View Insert Actions Tools Help

Reading View

Case 1: there are no importance weights



nhabitants developing and testing theories about mportance weights (which as it happens don't exist)