

Title: Everett and Evolution

Date: Sep 22, 2007 04:30 PM

URL: <http://pirsa.org/07090065>

Abstract: A fundamental question for Everettians is whether they can formulate a many-worlds interpretation of quantum theory which explains why, amongst all possible types of intelligent creature with all possible types of evolutionary and experimental history, we find ourselves among those whose histories apparently confirm Copenhagen quantum mechanics. Since the theory clearly allows that we could have found ourselves otherwise, the answer has to be probabilistic. Everettians then need to supply some account of how probability is or can be attached to an apparently deterministic theory. I argue that it cannot arise through any of the various notions of subjective uncertainty advocated by Saunders, Wallace, Vaidman and Greaves, among others, since none of these notions are valid. Nor could an adequate notion of probability be inferred from any account of the “caring weights” that we, as hypothetical rational Everettian agents, should use when considering the welfare of our Everettian successors – even if a unique rational strategy were to exist, which I argue is not the case. A proposition of the desired form could simply be postulated, but at the price of reducing the entire interpretation – probability postulate, preferred basis, and interpretation of basis states – to unsupported and maybe untestable hypotheses about consciousness. On the brighter side, I describe a new proposal for solving the measurement problem by proposing a definite mathematical structure for possible branching worlds, which appears to have the advantages claimed for Everettian ideas (in particular, respect of Lorentz invariance and of conservation laws), without the interpretational and scientific difficulties. I also note one subtle distinction between Everettian and standard accounts of evolution, which implies that (in principle, albeit not necessarily in practice) it could be possible to distinguish between the theories.



Taking branching seriously: "real world" interpretations (AK, arXiv:0708.3710)

A new proposal for solving the measurement problem.

Respects Lorentz invariance and conservation laws.

No immediately evident "tails problem".

Adds minimal extra structure to q.m., namely a preferred factorisation $H_A \otimes H_B$ and a preferred (possibly hypersurface dependent) projective decomposition on H_B .



branches separating over finite time interval

The branching structure which emerges is one in which individual branches are defined by mixed states, which become perfectly orthogonal (and pure) only as $t \rightarrow \infty$.

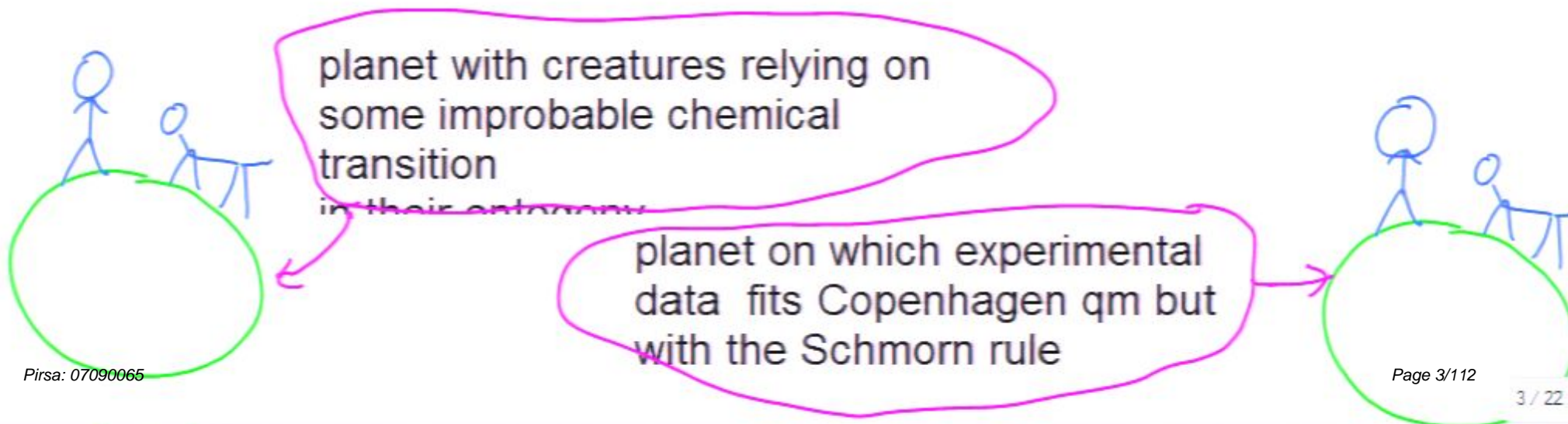
Measurement processes lead to outcome branches which become approximately orthogonal over finite time.



I want to consider evolution in two senses:

1) the Darwinian sense, applied to an Everett universe

An Everettian cosmology predicts the existence of multiple intelligent species - including multiple versions of humanity - with histories differing in various ways from our own. Many of these have evolutionary and experimental histories that would be enormously improbable according to non-Everettian qm. What can and should we, as scientists, say about this account of life in the universe?

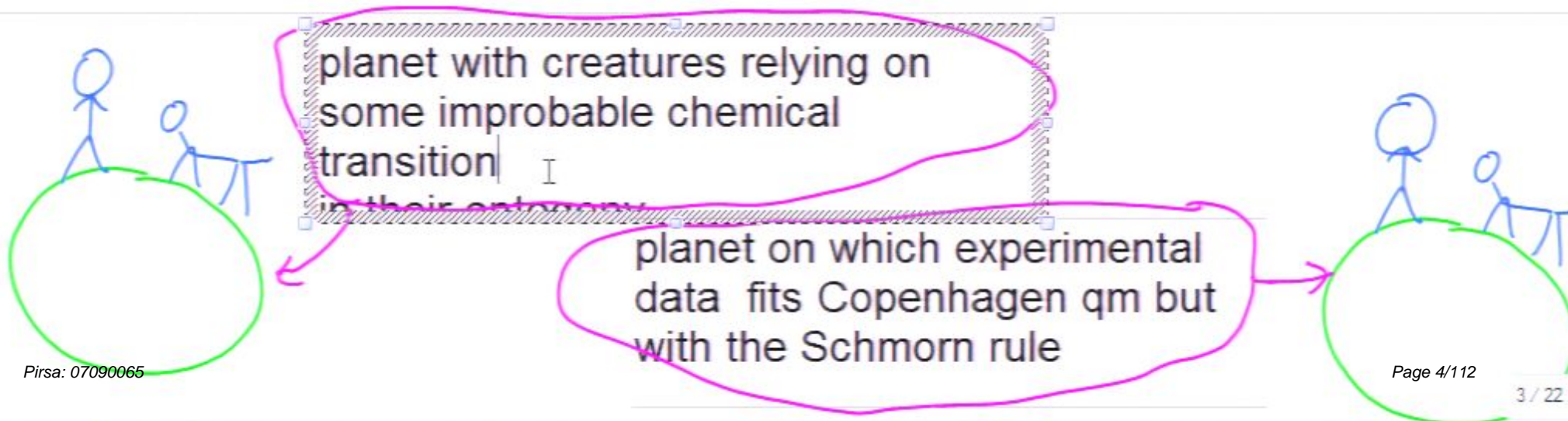




I want to consider evolution in two senses:

1) the Darwinian sense, applied to an Everett universe

An Everettian cosmology predicts the existence of multiple intelligent species - including multiple versions of humanity - with histories differing in various ways from our own. Many of these have evolutionary and experimental histories that would be enormously improbable according to non-Everettian qm. What can and should we, as scientists, say about this account of life in the universe?

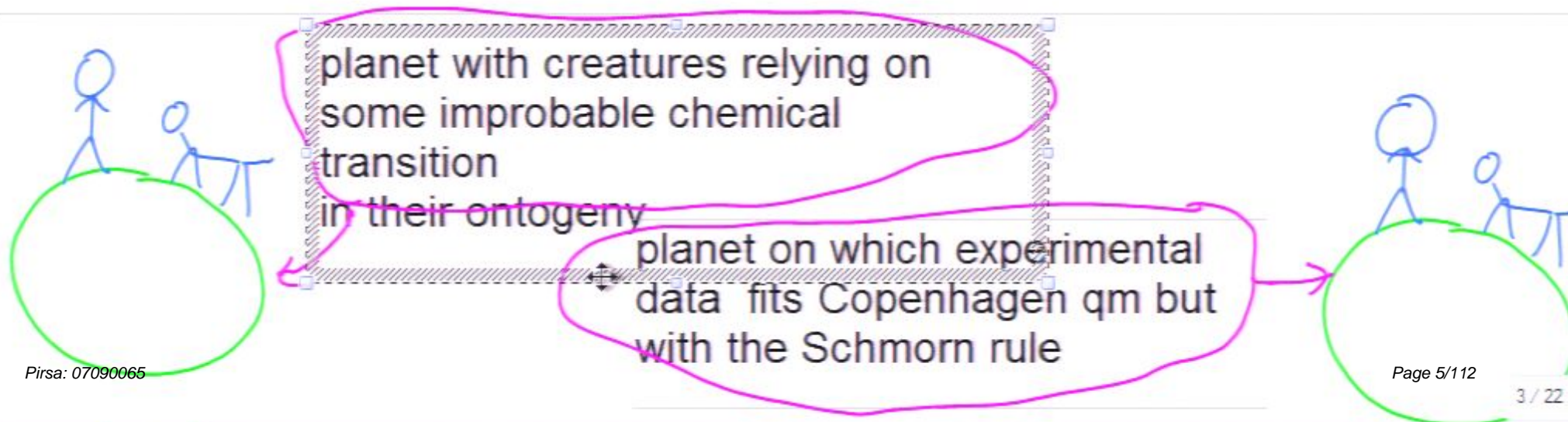




I want to consider evolution in two senses:

1) the Darwinian sense, applied to an Everett universe

An Everettian cosmology predicts the existence of multiple intelligent species - including multiple versions of humanity - with histories differing in various ways from our own. Many of these have evolutionary and experimental histories that would be enormously improbable according to non-Everettian qm. What can and should we, as scientists, say about this account of life in the universe?

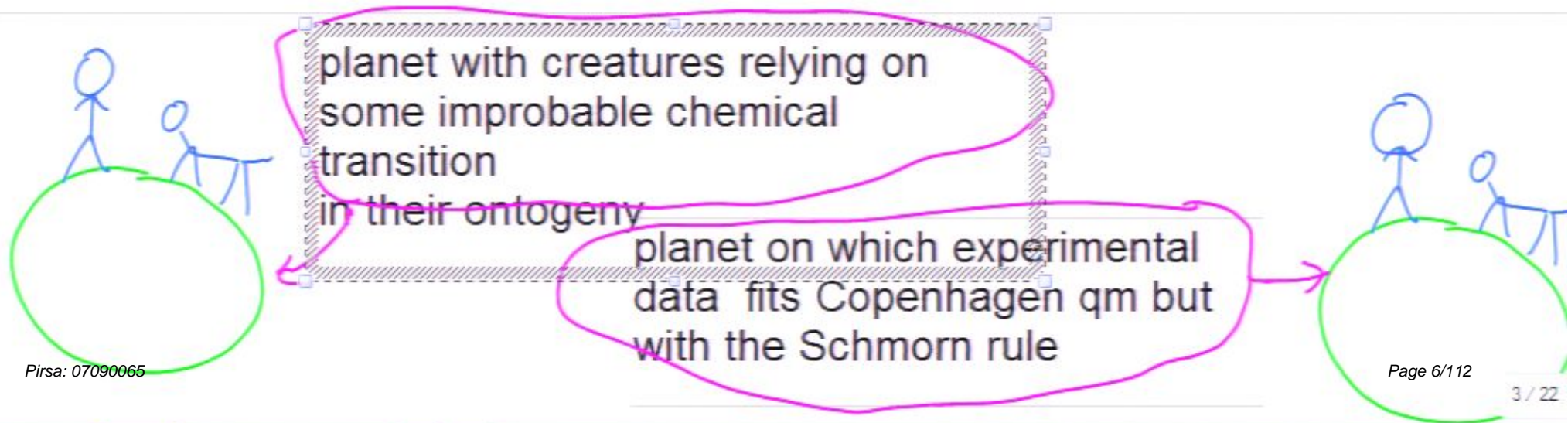




I want to consider evolution in two senses:

1) the Darwinian sense, applied to an Everett universe

An Everettian cosmology predicts the existence of multiple intelligent species - including multiple versions of humanity - with histories differing in various ways from our own. Many of these have evolutionary and experimental histories that would be enormously improbable according to non-Everettian qm. What can and should we, as scientists, say about this account of life in the universe?

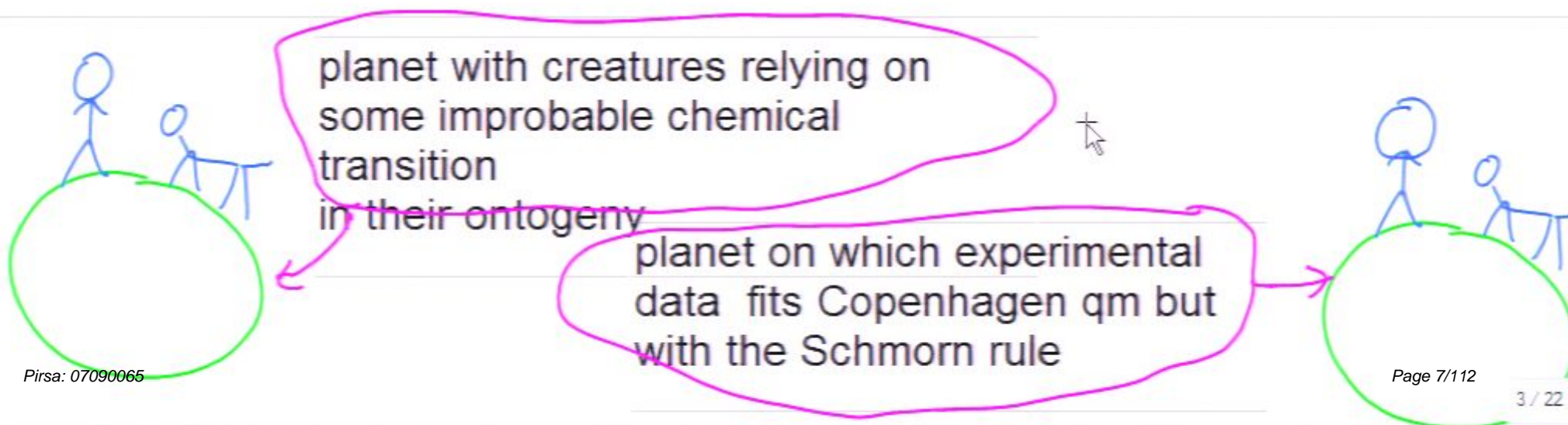




I want to consider evolution in two senses:

1) the Darwinian sense, applied to an Everett universe

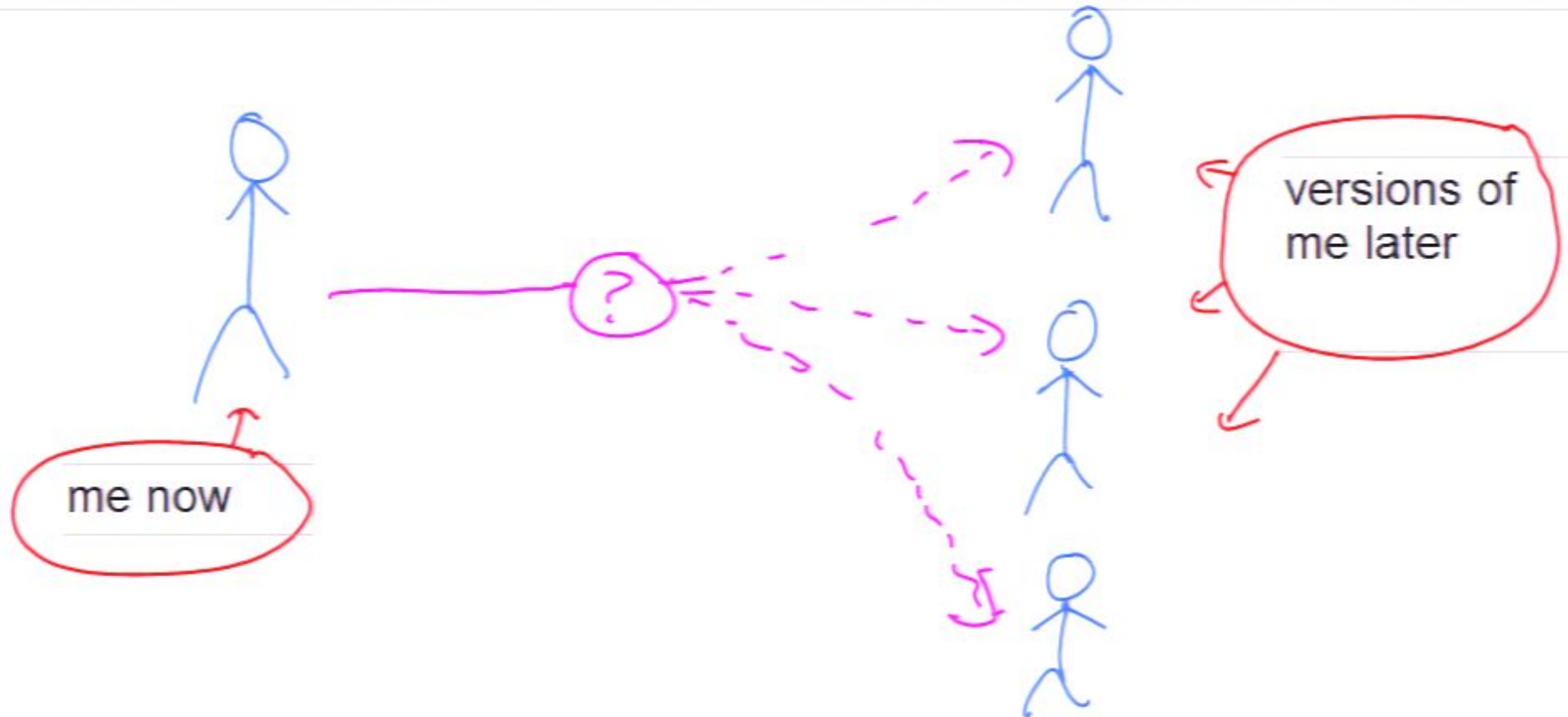
An Everettian cosmology predicts the existence of multiple intelligent species - including multiple versions of humanity - with histories differing in various ways from our own. Many of these have evolutionary and experimental histories that would be enormously improbable according to non-Everettian qm. What can and should we, as scientists, say about this account of life in the universe?





) the sense of personal evolution:

How could my present self be related to the multiple versions of my future self in an Everettian universe?





Everett-Darwinian evolution: the key question

Why, amongst all the possible types of intelligent creature with all the possible types of evolutionary and experimental history, do I find myself one of those whose history confirms*

Copenhagen q.m.?

(apparently, and within the theory's domain of validity)

A clearly inadequate answer: this is **consistent** with an Everettian account (true) and hence tends to **confirm** it scientifically (not true if our criterion is that a successor should **explain** the validity of Copenhagen q.m.).

"Your history might take any form" isn't an interesting result to emerge from an account of cosmology and evolution.



Everett-Darwinian evolution: the key question

Why, amongst all the possible types of intelligent creature with all the possible types of evolutionary and experimental history, do I find myself one of those whose history confirms*

Copenhagen q.m.?

(apparently, and within the theory's domain of validity)

A clearly inadequate answer: this is **consistent** with an Everettian account (true) and hence tends to **confirm** it scientifically (not true if our criterion is that a successor should **explain** the validity of Copenhagen q.m.).

"Your history might take any form" isn't an interesting result to emerge from an account of cosmology and evolution.



Everett-Darwinian evolution: the key question

Why, amongst all the possible types of intelligent creature with all the possible types of evolutionary and experimental history, do I find myself one of those whose history confirms*

Copenhagen q.m.?

(apparently, and within the theory's domain of validity)

A clearly inadequate answer: this is **consistent** with an Everettian account (true) and hence tends to **confirm** it scientifically (not true if our criterion is that a successor should **explain** the validity of Copenhagen q.m.).

"Your history might take any form" isn't an interesting result to emerge from an account of cosmology and evolution.



Everett-Darwinian evolution: the key question

Why, amongst all the possible types of intelligent creature with all the possible types of evolutionary and experimental history, do I find myself one of those whose history confirms*

Copenhagen q.m.?

(apparently, and within the theory's domain of validity)

A clearly inadequate answer: this is **consistent** with an Everettian account (true) and hence tends to **confirm** it scientifically (not true if our criterion is that a successor should **explain** the validity of Copenhagen q.m.).

"Your history might take any form" isn't an interesting result to emerge from an account of cosmology and evolution.



I take Everett to agree that his programme succeeds iff it answers the key question:

"The aim is not to deny or contradict the conventional formulation of quantum theory... but rather to supply a new, more general and complete formulation, from which the conventional interpretation can be deduced." (Everett, 1957)

Now clearly Everettians won't say that the theory implies that we'll **certainly** find our history to be Copenhagen-esque. They can only hope to show it **probably** will. Fine: but *how can a probabilistic conclusion emerge from this deterministic theory?*



Why is our history Copenhagen-esque?

To explain why we find our evolutionary and experimental history apparently in accord with Copenhagen qm, Everettians have to produce a rule which says something very like the following:

If each branch i with wavefunction ψ_i contains N_i creatures with property P , the probability that one finds oneself in branch j , given that one has property P , is

$$\frac{N_j |\psi_j|^2}{\sum_i N_i |\psi_i|^2}$$

In principle this could either be offered as a fundamental postulate or derived from others.



Why is our history Copenhagen-esque?

To explain why we find our evolutionary and experimental history apparently in accord with Copenhagen qm, Everettians have to produce a rule which says something very like the following:

If each branch i with wavefunction ψ_i contains N_i creatures with property P , the probability that one finds oneself in branch j , given that one has property P , is

$$\frac{N_j |\psi_j|^2}{\sum_i N_i |\psi_i|^2}$$

In principle this could either be offered as a fundamental postulate or derived from others.



Why is our history Copenhagen-esque?

To explain why we find our evolutionary and experimental history apparently in accord with Copenhagen qm, Everettians have to produce a rule which says something very like the following:

If each branch i with wavefunction ψ_i contains N_i creatures with property P , the probability that one finds oneself in branch j , j given that one has property P , is

$$\frac{N_j |\psi_j|^2}{\sum_i N_i |\psi_i|^2}$$

In principle this could either be offered as a fundamental postulate or derived from others.



Why is our history Copenhagen-esque?

To explain why we find our evolutionary and experimental history apparently in accord with Copenhagen qm, Everettians have to produce a rule which says something very like the following:

If each branch i with wavefunction ψ_i contains N_i creatures with property P , the probability that one finds oneself in branch j , given that one has property P , is

$$\frac{N_j |\psi_j|^2}{\sum_i N_i |\psi_i|^2}$$

In principle this could either be offered as a fundamental postulate or derived from others.



How could probability arise in an Everettian universe at all, given that the dynamics are deterministic?

One popular line of thought argues for some notion of **subjective uncertainty** in an Everettian universe. According to Saunders & Wallace, Everett plus functionalism implies such a notion: **I will become** one of the two possible future versions of myself after observing a Stern-Gerlach experiment, **but I don't know which**. I think (as Vaidman and Greaves have argued) that this is simply a mistake:





How could probability arise in an Everettian universe at all, given that the dynamics are deterministic?

One popular line of thought argues for some notion of **subjective uncertainty** in an Everettian universe. According to Saunders & Wallace, Everett plus functionalism implies such a notion: **I will become** one of the two possible future versions of myself after observing a Stern-Gerlach experiment, **but I don't know which**. I think (as Vaidman and Greaves have argued) that this is simply a mistake:



Select Page	
Paste	Ctrl+V
Insert Text Box	
Insert/Remove Space	
Insert Picture...	
Insert Flag	





How could probability arise in an Everettian universe at all, given that the dynamics are deterministic?

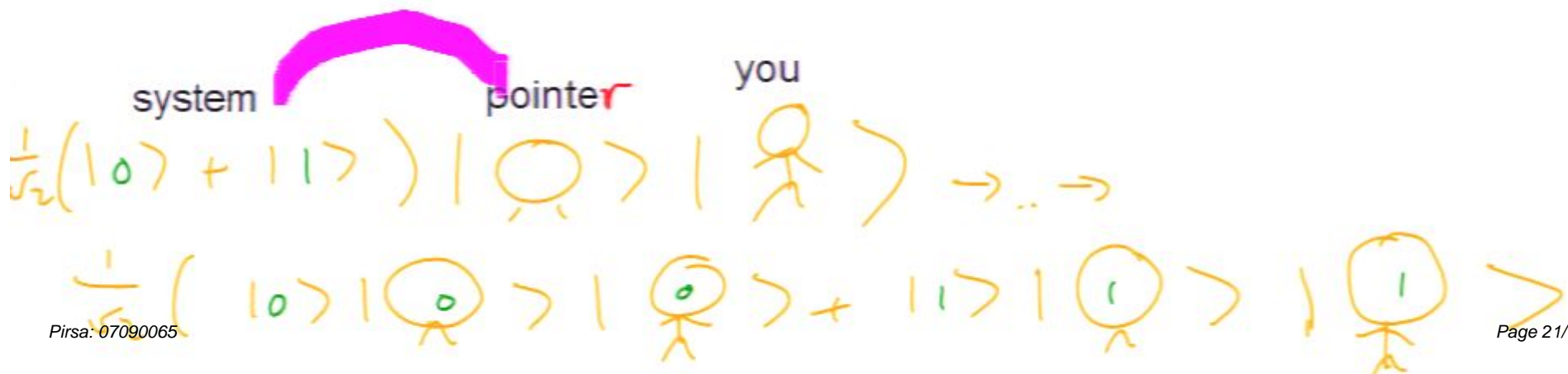
One popular line of thought argues for some notion of **subjective uncertainty** in an Everettian universe. According to Saunders & Wallace, Everett plus functionalism implies such a notion: **I will become** one of the two possible future versions of myself after observing a Stern-Gerlach experiment, **but I don't know which**. I think (as Vaidman and Greaves have argued) that this is simply a mistake:





How could probability arise in an Everettian universe at all, given that the dynamics are deterministic?

One popular line of thought argues for some notion of **subjective uncertainty** in an Everettian universe. According to Saunders & Wallace, Everett plus functionalism implies such a notion: **I will become** one of the two possible future versions of myself after observing a Stern-Gerlach experiment, **but I don't know which**. I think (as Vaidman and Greaves have argued) that this is simply a mistake:





How could probability arise in an Everettian universe at all, given that the dynamics are deterministic?

One popular line of thought argues for some notion of **subjective uncertainty** in an Everettian universe. According to Saunders & Wallace, Everett plus functionalism implies such a notion: **I will become** one of the two possible future versions of myself after observing a Stern-Gerlach experiment, **but I don't know which**. I think (as Vaidman and Greaves have argued) that this is simply a mistake:





Wallace (2002) summarizes Saunders thus:

"There are only three logical possibilities:

- 1) I should expect to become both future selves.*
- 2) I should expect to become one or the other.*
- 3) I should expect nothing: oblivion.*

Of these (3) seems absurd. (1) is at least coherent- we could imagine some telepathic link- but ... this link will have to supervene on some physical interaction between the two copies which is not in fact present. This leaves (2) as the only option."

I find this mystifying. (3) is coherent, and might be justifiable: if identity turns out to fail to propagate forwards in time then indeed "I" would cease to exist.

In any case, (1) makes perfect sense and requires no telepathy. (2) is unsupported by the dynamics and wave function ontology, and so must be rejected.



Wallace (2002) summarizes Saunders thus:

"There are only three logical possibilities:

- 1) I should expect to become both future selves.*
- 2) I should expect to become one or the other.*
- 3) I should expect nothing: oblivion.*

Of these (3) seems absurd. (1) is at least coherent- we could imagine some telepathic link- but ... this link will have to supervene on some physical interaction between the two copies which is not in fact present. This leaves (2) as the only option."

I find this mystifying. (3) is coherent, and might be justifiable: if identity turns out to fail to propagate forwards in time then indeed "I" would cease to exist.

In any case, (1) makes perfect sense and requires no telepathy. (2) is unsupported by the dynamics and wave function ontology, and so must be rejected.



Wallace (2002) summarizes Saunders thus:

"There are only three logical possibilities:

- 1) I should expect to become both future selves.*
- 2) I should expect to become one or the other.*
- 3) I should expect nothing: oblivion.*

Of these (3) seems absurd. (1) is at least coherent- we could imagine some telepathic link- but ... this link will have to supervene on some physical interaction between the two copies which is not in fact present. This leaves (2) as the only option."

I find this mystifying. (3) is coherent, and might be justifiable: if identity turns out to fail to propagate forwards in time then indeed "I" would cease to exist.

In any case, (1) makes perfect sense and requires no telepathy. (2) is unsupported by the dynamics and wave function ontology, and so must be rejected.



Wallace (2002) summarizes Saunders thus:

"There are only three logical possibilities:

- 1) I should expect to become both future selves.*
- 2) I should expect to become one or the other.*
- 3) I should expect nothing: oblivion.*

Of these (3) seems absurd. (1) is at least coherent- we could imagine some telepathic link- but ... this link will have to supervene on some physical interaction between the two copies which is not in fact present. This leaves (2) as the only option."

I find this mystifying. (3) is coherent, and might be justifiable: if identity turns out to fail to propagate forwards in time then indeed "I" would cease to exist.

In any case, (1) makes perfect sense and requires no telepathy. (2) is unsupported by the dynamics and wave function ontology, and so must be rejected.



Wallace (2002) summarizes Saunders thus:

"There are only three logical possibilities:

- 1) I should expect to become both future selves.*
- 2) I should expect to become one or the other.*
- 3) I should expect nothing: oblivion.*

Of these (3) seems absurd. (1) is at least coherent- we could imagine some telepathic link- but ... this link will have to supervene on some physical interaction between the two copies which is not in fact present. This leaves (2) as the only option."

I find this mystifying. (3) is coherent, and might be justifiable: if identity turns out to fail to propagate forwards in time then indeed "I" would cease to exist.

In any case, (1) makes perfect sense and requires no telepathy. (2) is unsupported by the dynamics and wave function ontology, and so must be rejected.



Wallace (2002) summarizes Saunders thus:

"There are only three logical possibilities:

- 1) I should expect to become both future selves.*
- 2) I should expect to become one or the other.*
- 3) I should expect nothing: oblivion.*

Of these (3) seems absurd. (1) is at least coherent- we could imagine some telepathic link- but ... this link will have to supervene on some physical interaction between the two copies which is not in fact present. This leaves (2) as the only option."

I find this mystifying. (3) is coherent, and might be justifiable: if identity turns out to fail to propagate forwards in time then indeed "I" would cease to exist.

In any case, (1) makes perfect sense and requires no telepathy. (2) is unsupported by the dynamics and wave function ontology, and so must be rejected.



Wallace (2002) summarizes Saunders thus:

"There are only three logical possibilities:

- 1) I should expect to become both future selves.*
- 2) I should expect to become one or the other.*
- 3) I should expect nothing: oblivion.*

Of these (3) seems absurd. (1) is at least coherent- we could imagine some telepathic link- but ... this link will have to supervene on some physical interaction between the two copies which is not in fact present. This leaves (2) as the only option."

I find this mystifying. (3) is coherent, and might be justifiable: if identity turns out to fail to propagate forwards in time then indeed "I" would cease to exist.

In any case, (1) makes perfect sense and requires no telepathy. (2) is unsupported by the dynamics and wave function ontology, and so must be rejected.



if it's sensible to say, as a form of shorthand, that you become **anything** in this process, it's surely the combination of your **successors**. Your physical system becomes the combination of their physical systems. And, if a functionalist view of Everett says anything intelligible about what your **mind** becomes, it's that it becomes a combination of their minds (which are disconnected and noninteracting).

This - pace Saunders & Wallace - makes perfect sense and requires no mysterious telepathic connection.



Vaidman rejects **pre-experiment** uncertainty but claims there is a legitimate form of **post hoc** uncertainty: after the experiment, but before I know the outcome, I will be in one of two possible branches, but I don't know which.

Vaidman compares this to one of two anaesthetized and separated successors waking up in one of two possible locations (he knows he is somewhere definite but not where.)

But as Saunders, Wallace, Greaves et al. stress, **there is no well-defined way of counting branches**. In realistic models, quantum events are frequent, near-ubiquitous, and impossible to keep systematic account of. **There is no well-defined set of events defining "your branch"** before you perceive them: the only well-defined state corresponding to "you" is the vector defining your internal state.

(cf. Papineau)



Vaidman rejects **pre-experiment** uncertainty but claims there is a legitimate form of **post hoc** uncertainty: after the experiment, but before I know the outcome, I will be in one of two possible branches, but I don't know which.

Vaidman compares this to one of two anaesthetized and separated successors waking up in one of two possible locations (he knows he is somewhere definite but not where.)

But as Saunders, Wallace, Greaves et al. stress, **there is no well-defined way of counting branches**. In realistic models, quantum events are frequent, near-ubiquitous, and impossible to keep systematic account of. **There is no well-defined set of events defining "your branch"** before you perceive them: the only well-defined state corresponding to "you" is the vector defining your internal state.

(cf. Papineau)



It's tempting to try to interpret

system pointer you

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) | \text{circle} \rangle | \text{stick} \rangle \rightarrow$$

$$\frac{1}{\sqrt{2}} (|0\rangle | \text{circle with dot} \rangle | \text{stick} \rangle + |1\rangle | \text{circle with dot} \rangle | \text{circle} \rangle)$$

to mean

you in 0 branch

you in 1 branch

but in a realistic model this decomposition is in no way canonical: the only well-defined representative of your current state is the unbranched vector

$$| \text{stick} \rangle$$

(There is no unique complete list of possible quasiclassical worlds in which you could currently be located.)



It's tempting to try to interpret

system pointer you

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) | \text{circle} \rangle | \text{stick} \rangle \rightarrow$$

$$\frac{1}{\sqrt{2}} (|0\rangle | \text{circle with dot} \rangle | \text{stick} \rangle + |1\rangle | \text{circle} \rangle | \text{stick} \rangle)$$

to mean

you in 0 branch

you in 1 branch

but in a realistic model this decomposition is in no way canonical: the only well-defined representative of your current state is the unbranched vector

$$| \text{stick} \rangle$$

(There is no unique complete list of possible quasiclassical worlds in which you could currently be located.)



It's tempting to try to interpret

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) | \text{you in } 0 \text{ branch} \rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle | \text{you in } 0 \text{ branch} \rangle + |1\rangle | \text{you in } 1 \text{ branch} \rangle)$$

to mean

you in 0 branch

you in 1 branch

but in a realistic model this decomposition is in no way canonical: the only well-defined representative of your current state is the unbranched vector

$$| \text{you in } 0 \text{ branch} \rangle$$

(There is no unique complete list of possible quasiclassical worlds in which you could currently be located.)



It's tempting to try to interpret

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) | \text{circle} \rangle | \text{stick} \rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle | \text{circle} \rangle | \text{stick} \rangle + |1\rangle | \text{circle} \rangle | \text{stick} \rangle)$$

to mean

you in 0 branch

you in 1 branch

but in a realistic model this decomposition is in no way canonical: the only well-defined representative of your current state is the unbranched vector

$$| \text{stick} \rangle$$

(There is no unique complete list of possible quasiclassical worlds in which you could currently be located.)



It's tempting to try to interpret

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) | \text{circle} \rangle | \text{stick} \rangle \rightarrow$$

$$\frac{1}{\sqrt{2}} (|0\rangle | \text{circle} \rangle | \text{stick} \rangle + |1\rangle | \text{circle} \rangle | \text{stick} \rangle)$$

to mean

you in 0 branch

you in 1 branch

but in a realistic model this decomposition is in no way canonical: the only well-defined representative of your current state is the unbranched vector

$$| \text{stick} \rangle$$

(There is no unique complete list of possible quasiclassical worlds in which you could currently be located.)



It's tempting to try to interpret

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) | \text{circle} \rangle | \text{stick} \rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle | \text{circle with dot} \rangle | \text{stick} \rangle + |1\rangle | \text{circle with dot} \rangle | \text{stick} \rangle)$$

to mean

you in 0 branch

you in 1 branch

but in a realistic model this decomposition is in no way canonical: the only well-defined representative of your current state is the unbranched vector

$$| \text{stick} \rangle$$

Conclusion: Everett simply does not admit any notion of subjective uncertainty. We need to look elsewhere if we want an account of probability in Everett.



Probability from "caring weights"?

It's reasonable to suggest one should care about the welfare of successors in an Everett universe, even though there is no way to identify yourself with any particular one of them.

Deutsch and Wallace argue that, given assumptions which are claimed to be reasonable, if one believes in Everettian q.m., then there is a unique rational way to act, effectively assigning a "caring weight" of 2^{-i} to a successor whose state is a factor of 2^i .

I think there are several problems with this line of argument. But even taking the conclusion at face value, it (per se) wouldn't suffice: telling us how to treat our successors if we believe in Everett q. m. doesn't (per se) explain why we found ourselves in the position to arrive at Everett q.m. in the first place.



Probability from "caring weights"?

It's reasonable to suggest one should care about the welfare of successors in an Everett universe, even though there is no way to identify yourself with any particular one of them.

Deutsch and Wallace argue that, given assumptions which are claimed to be reasonable, if one believes in Everettian q.m., then there is a unique rational way to act, effectively assigning a "caring weight" of 2^{-i} to a successor whose state is a factor of 2^i .

I think there are several problems with this line of argument. But even taking the conclusion at face value, it (per se) wouldn't suffice: telling us how to treat our successors if we believe in Everett q. m. doesn't (per se) explain why we found ourselves in the position to arrive at Everett q.m. in the first place.



Probability from "caring weights"?

It's reasonable to suggest one should care about the welfare of successors in an Everett universe, even though there is no way to identify yourself with any particular one of them.

Deutsch and Wallace argue that, given assumptions which are claimed to be reasonable, if one believes in Everettian q.m., then there is a unique rational way to act, effectively assigning a "caring weight" of $|24:1|^2$ to a successor whose state is a factor of 24 .

I think there are several problems with this line of argument. But even taking the conclusion at face value, it (per se) wouldn't suffice: telling us how to treat our successors if we believe in Everett q. m. doesn't (per se) explain why we found ourselves in the position to arrive at Everett q.m. in the first place.



Probability from "caring weights"?

It's reasonable to suggest one should care about the welfare of successors in an Everett universe, even though there is no way to identify yourself with any particular one of them.

Deutsch and Wallace argue that, given assumptions which are claimed to be reasonable, if one believes in Everettian q.m., then there is a unique rational way to act, effectively assigning a "caring weight" of $1/4:1^2$ to a successor whose state is a factor of $1/4$.



I think there are several problems with this line of argument. But even taking the conclusion at face value, it (per se) wouldn't suffice: telling us how to treat our successors if we believe in Everett q. m. doesn't (per se) explain why we found ourselves in the position to arrive at Everett q.m. in the first place.



Probability from "caring weights"?

It's reasonable to suggest one should care about the welfare of successors in an Everett universe, even though there is no way to identify yourself with any particular one of them.

Deutsch and Wallace argue that, given assumptions which are claimed to be reasonable, if one believes in Everettian q.m., then there is a unique rational way to act, effectively assigning a "caring weight" of $1/4:1^2$ to a successor whose


state is a factor of 

I think there are several problems with this line of argument. But even taking the conclusion at face value, it (per se) wouldn't suffice: telling us how to treat our successors if we believe in Everett q. m. doesn't (per se) explain why we found ourselves in the position to arrive at Everett q.m. in the first place.



Probability from "caring weights"?

It's reasonable to suggest one should care about the welfare of successors in an Everett universe, even though there is no way to identify yourself with any particular one of them.

Deutsch and Wallace argue that, given assumptions which are claimed to be reasonable, if one believes in Everettian q.m., then there is a unique rational way to act, effectively assigning a "caring weight" of $1/4:1^2$ to a successor whose state is a factor of  I

I think there are several problems with this line of argument. But even taking the conclusion at face value, it (per se) wouldn't suffice: telling us how to treat our successors if we believe in Everett q. m. doesn't (per se) explain why we found ourselves in the position to arrive at Everett q.m. in the first place.



Probability from "caring weights"?

It's reasonable to suggest one should care about the welfare of successors in an Everett universe, even though there is no way to identify yourself with any particular one of them.

Deutsch and Wallace argue that, given assumptions which are claimed to be reasonable, if one believes in Everettian q.m., then there is a unique rational way to act, effectively assigning a "caring weight" of $|24:1|^2$ to a successor whose state is a factor of $24:1$.

I think there are several problems with this line of argument. But even taking the conclusion at face value, it (per se) wouldn't suffice: telling us how to treat our successors if we believe in Everett q. m. doesn't (per se) explain why we found ourselves in the position to arrive at Everett q.m. in the first place.



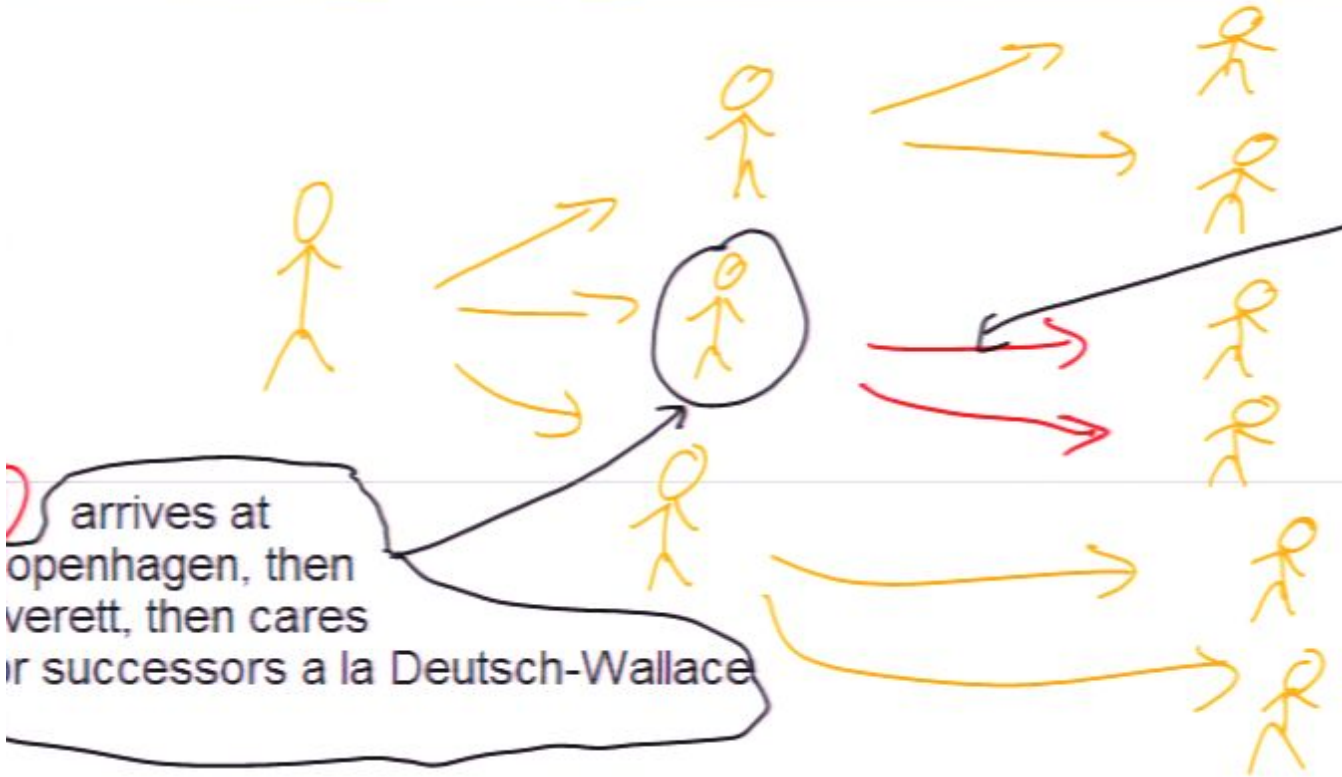
Probability from "caring weights"?

It's reasonable to suggest one should care about the welfare of successors in an Everett universe, even though there is no way to identify yourself with any particular one of them.

Deutsch and Wallace argue that, given assumptions which are claimed to be reasonable, if one believes in Everettian q.m., then there is a unique rational way to act, effectively assigning a "caring weight" of $1/2^i$ to a successor whose state is a factor of 2^i .



I think there are several problems with this line of argument. But even taking the conclusion at face value, it (per se) wouldn't suffice: telling us how to treat our successors if we believe in Everett q. m. doesn't (per se) explain why we found ourselves in the position to arrive at Everett q.m. in the first place.



flow of "proper" caring

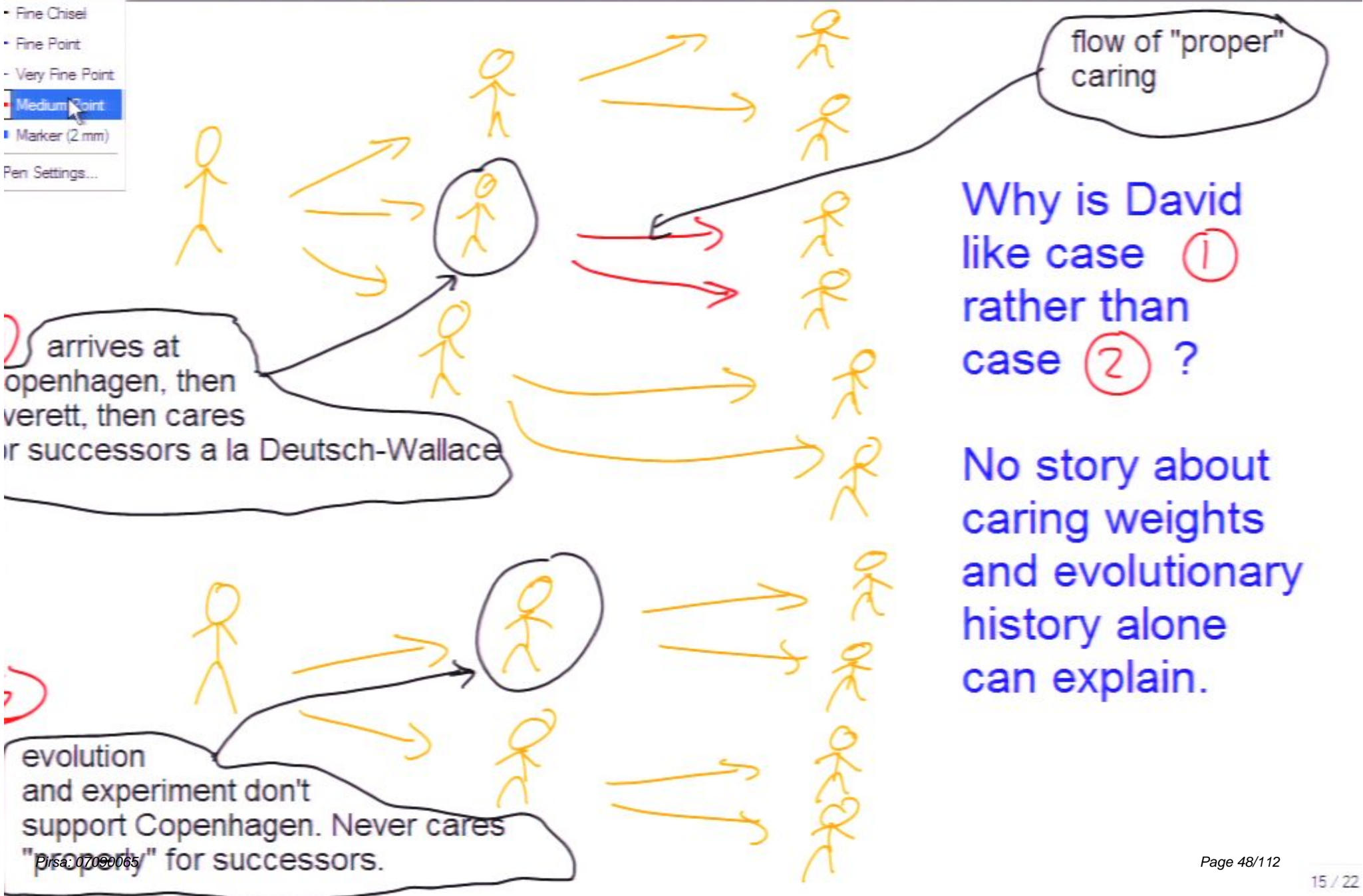
Why is David like case ① rather than case ② ?

No story about caring weights and evolutionary history alone can explain.





- Fine Chisel
- Fine Point
- Very Fine Point
- Medium Point
- Marker (2 mm)
- Pen Settings...



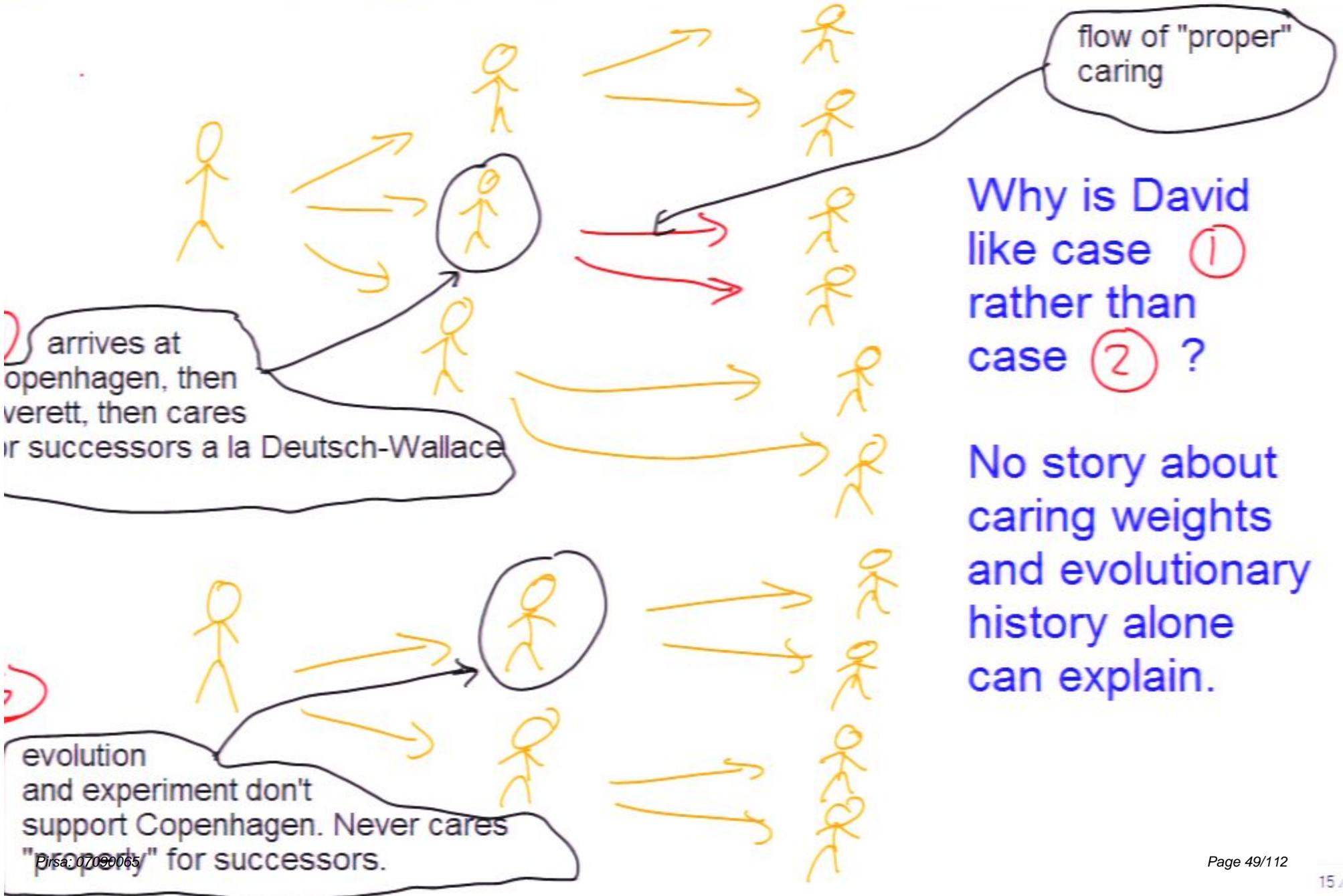
arrives at
openhagen, then
verett, then cares
for successors a la Deutsch-Wallace

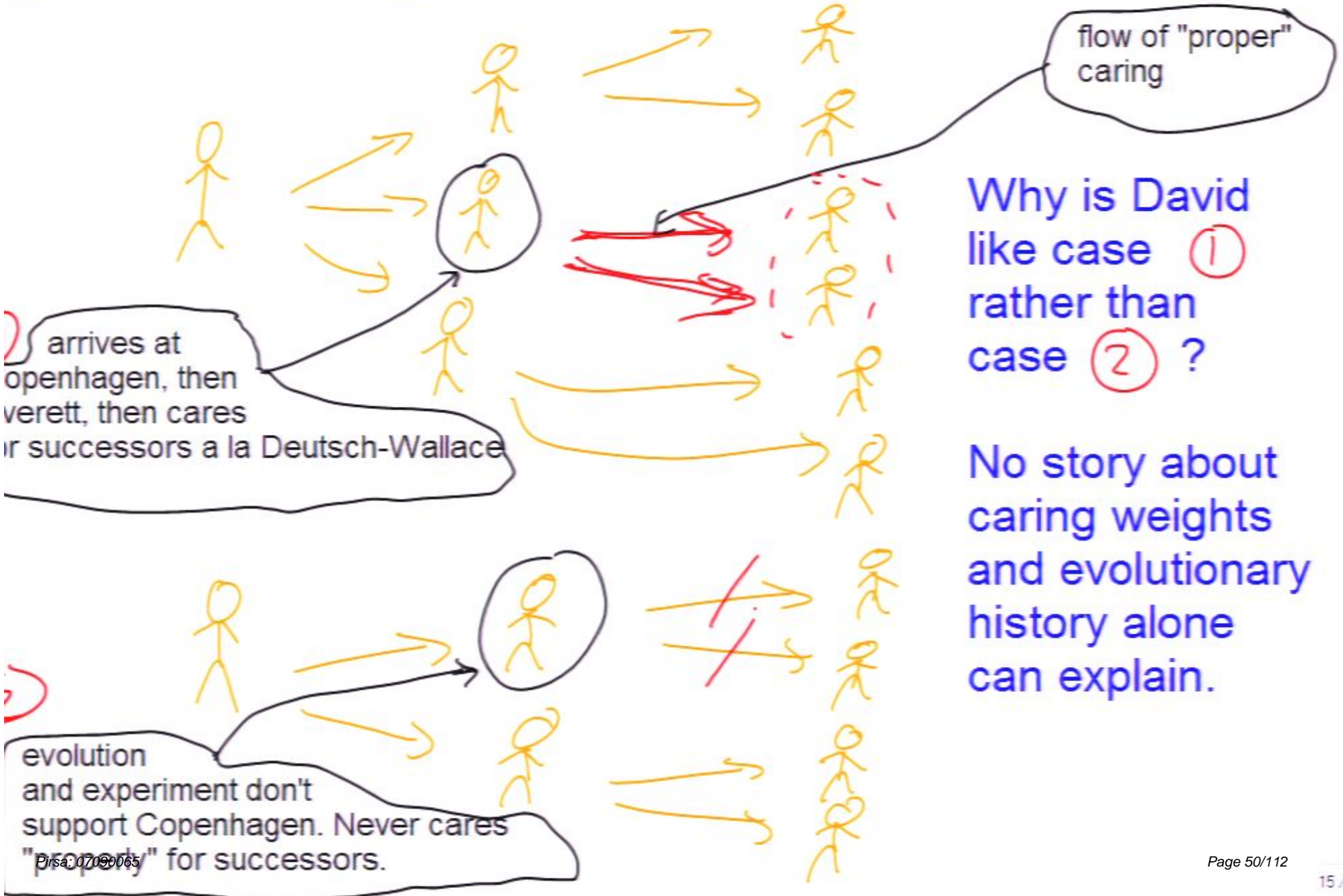
evolution
and experiment don't
support Copenhagen. Never cares
"property" for successors.

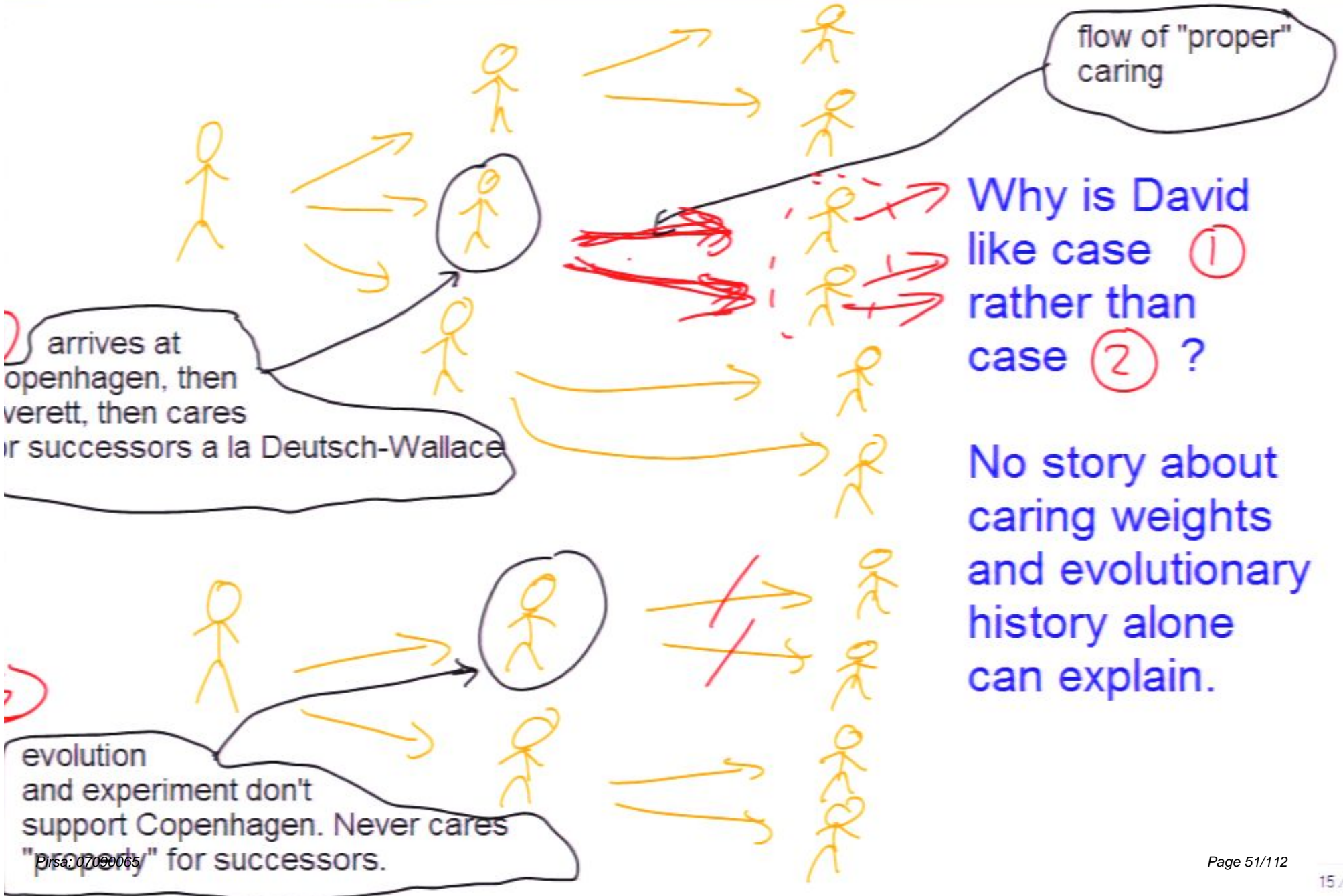
flow of "proper"
caring

Why is David
like case ①
rather than
case ② ?

No story about
caring weights
and evolutionary
history alone
can explain.







arrives at openhagen, then verett, then cares for successors a la Deutsch-Wallace

flow of "proper" caring

Why is David like case ① rather than case ② ?

No story about caring weights and evolutionary history alone can explain.

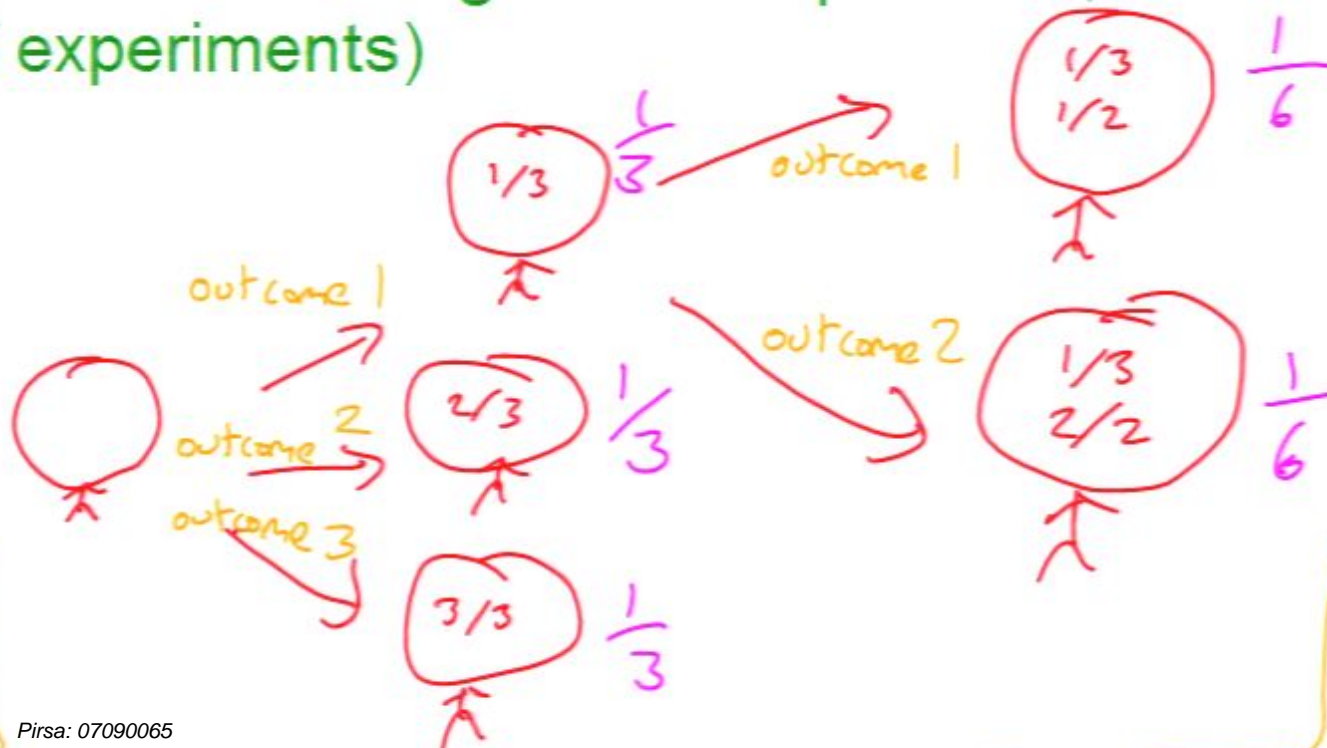
evolution and experiment don't support Copenhagen. Never cares "property" for successors.



Some problems with the Deutsch-Wallace argument

(i) There are other well-defined strategies, which it seems defensible to term rational, for example:

(i) weighting successors according to their memory patterns (branch-counting made respectable, for well-defined sequences of experiments)



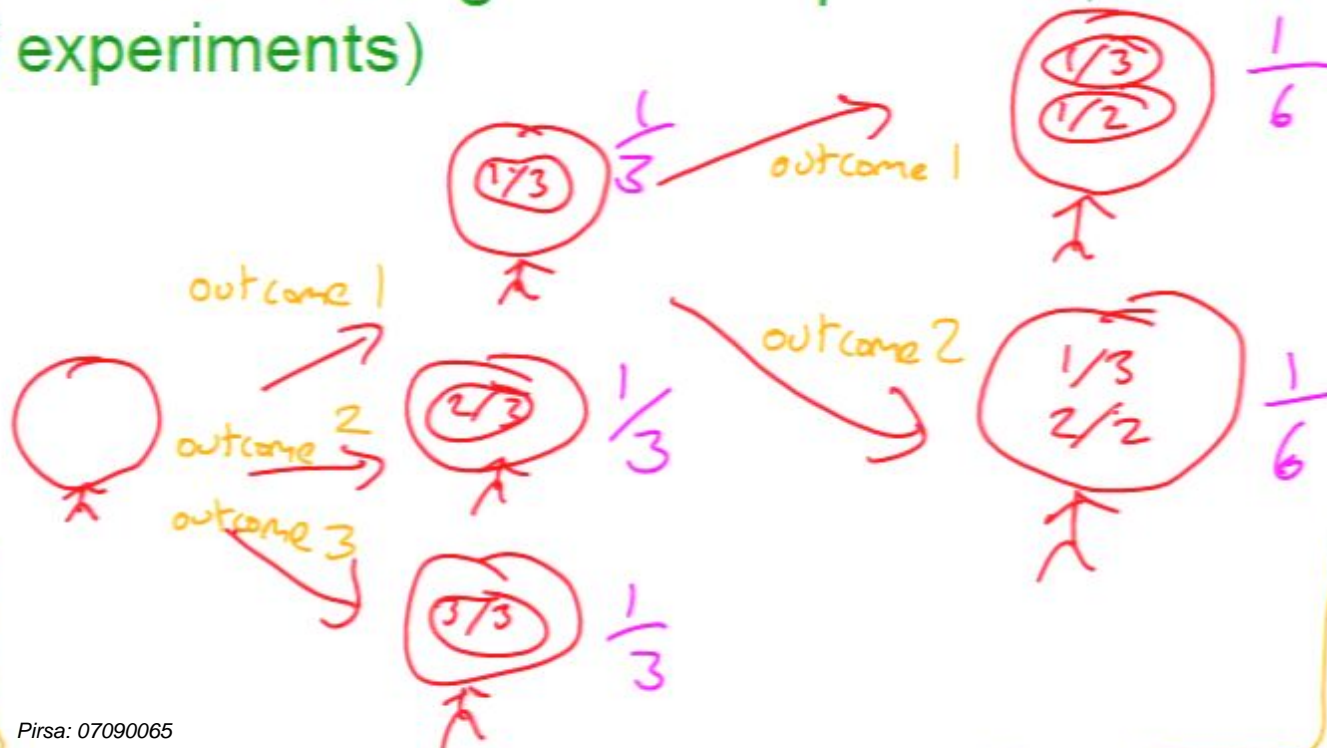
(ii) weighting by a post-selected final state
 (iii) branch-counting using a preferred consistent set



Some problems with the Deutsch-Wallace argument

(i) There are other well-defined strategies, which it seems defensible to term rational, for example:

(i) weighting successors according to their memory patterns (branch-counting made respectable, for well-defined sequences of experiments)



(ii) weighting by a post-selected final state

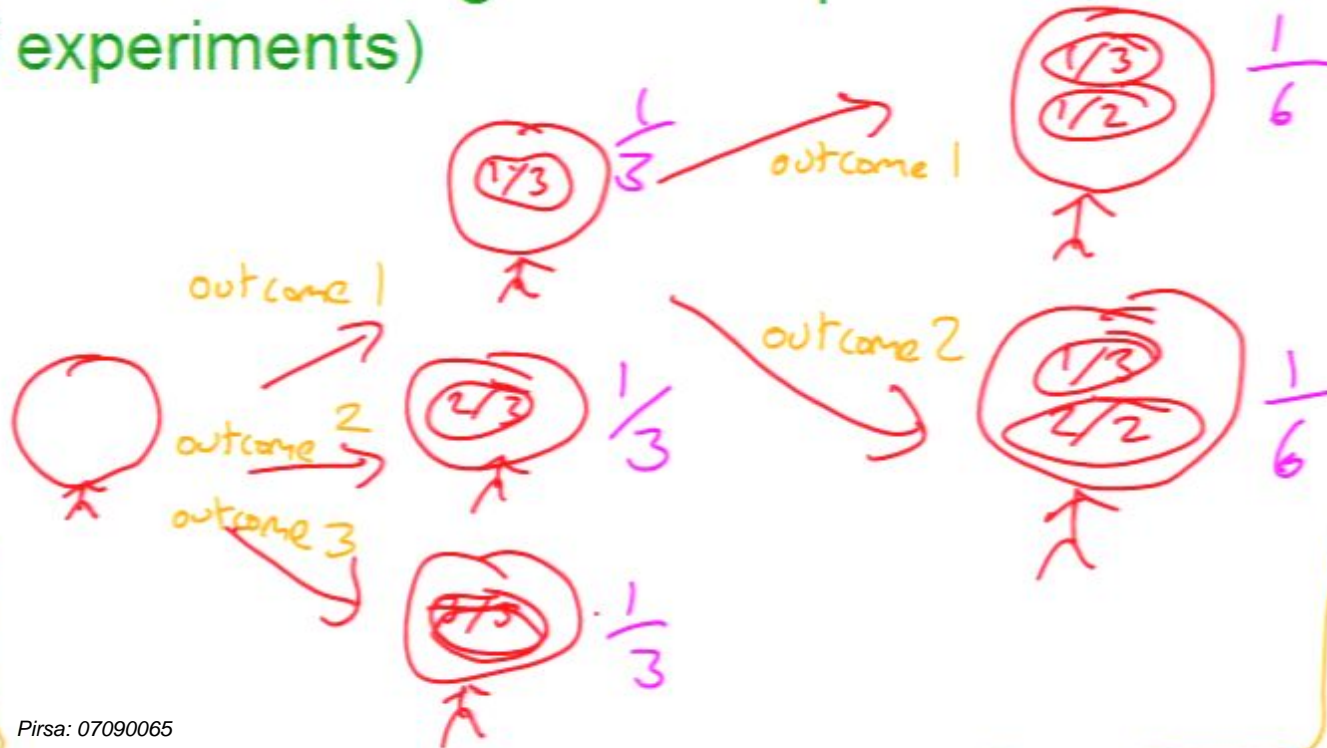
(iii) branch-counting using a preferred consistent set



Some problems with the Deutsch-Wallace argument

(i) There are other well-defined strategies, which it seems defensible to term rational, for example:

(i) weighting successors according to their memory patterns (branch-counting made respectable, for well-defined sequences of experiments)



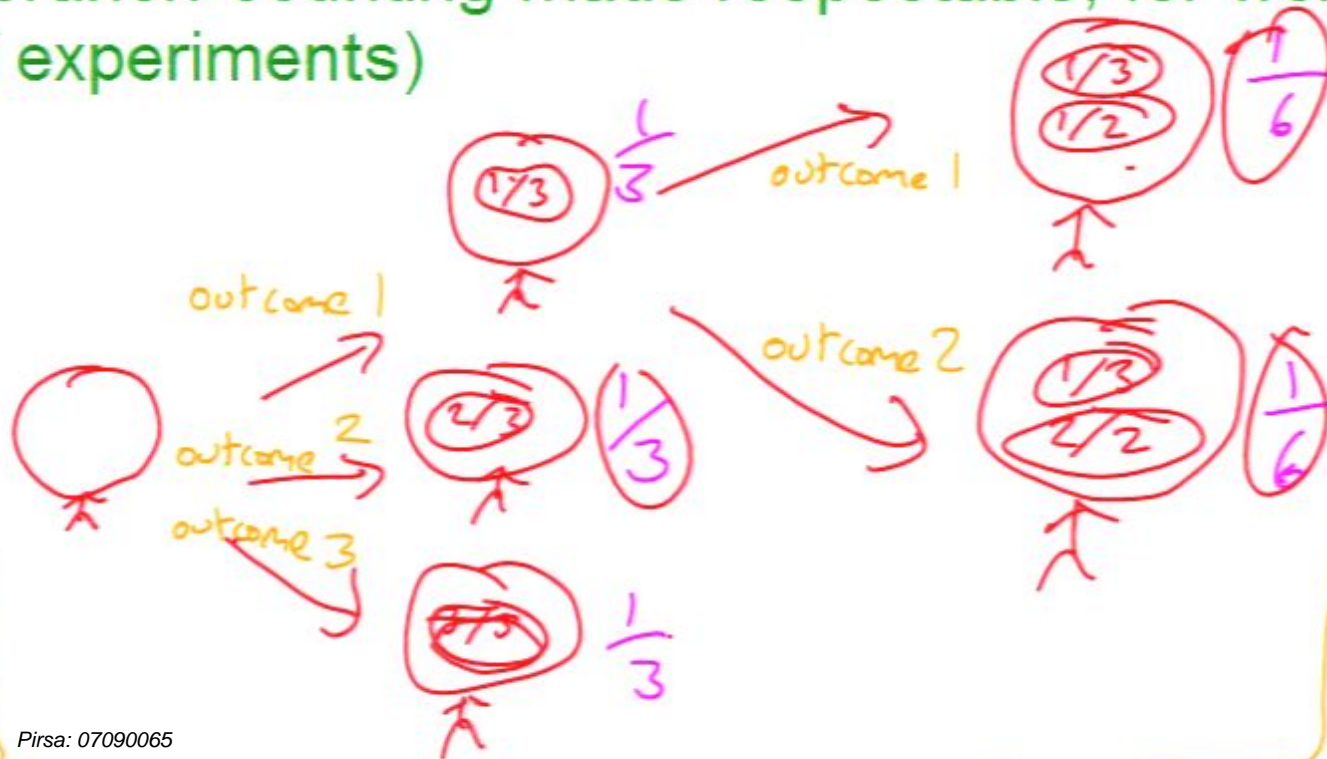
(ii) weighting by a post-selected final state
 (iii) branch-counting using a preferred consistent set



Some problems with the Deutsch-Wallace argument

There are other well-defined strategies, which it seems defensible to term rational, for example:

(i) weighting successors according to their memory patterns (branch-counting made respectable, for well-defined sequences of experiments)



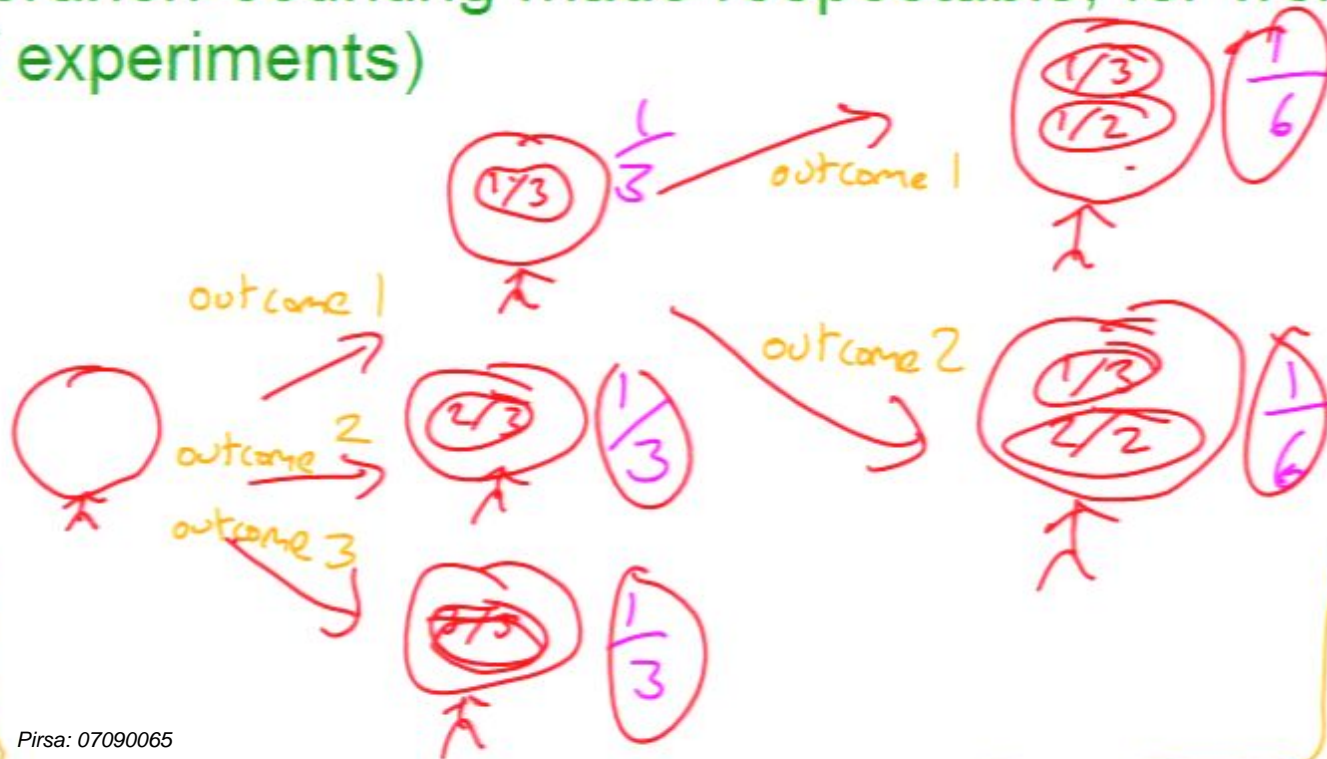
(ii) weighting by a post-selected final state
 (iii) branch-counting using a preferred consistent set



Some problems with the Deutsch-Wallace argument

(i) There are other well-defined strategies, which it seems defensible to term rational, for example:

(ii) weighting successors according to their memory patterns (branch-counting made respectable, for well-defined sequences of experiments)



(ii) weighting by a post-selected final state

(iii) branch-counting using a preferred consistent set



Some problems with the Deutsch-Wallace argument

There are other well-defined strategies, which it seems defensible to term rational, for example:

(i) weighting successors according to their memory patterns (branch-counting made respectable, for well-defined sequences of experiments)



(ii) weighting by a post-selected final state

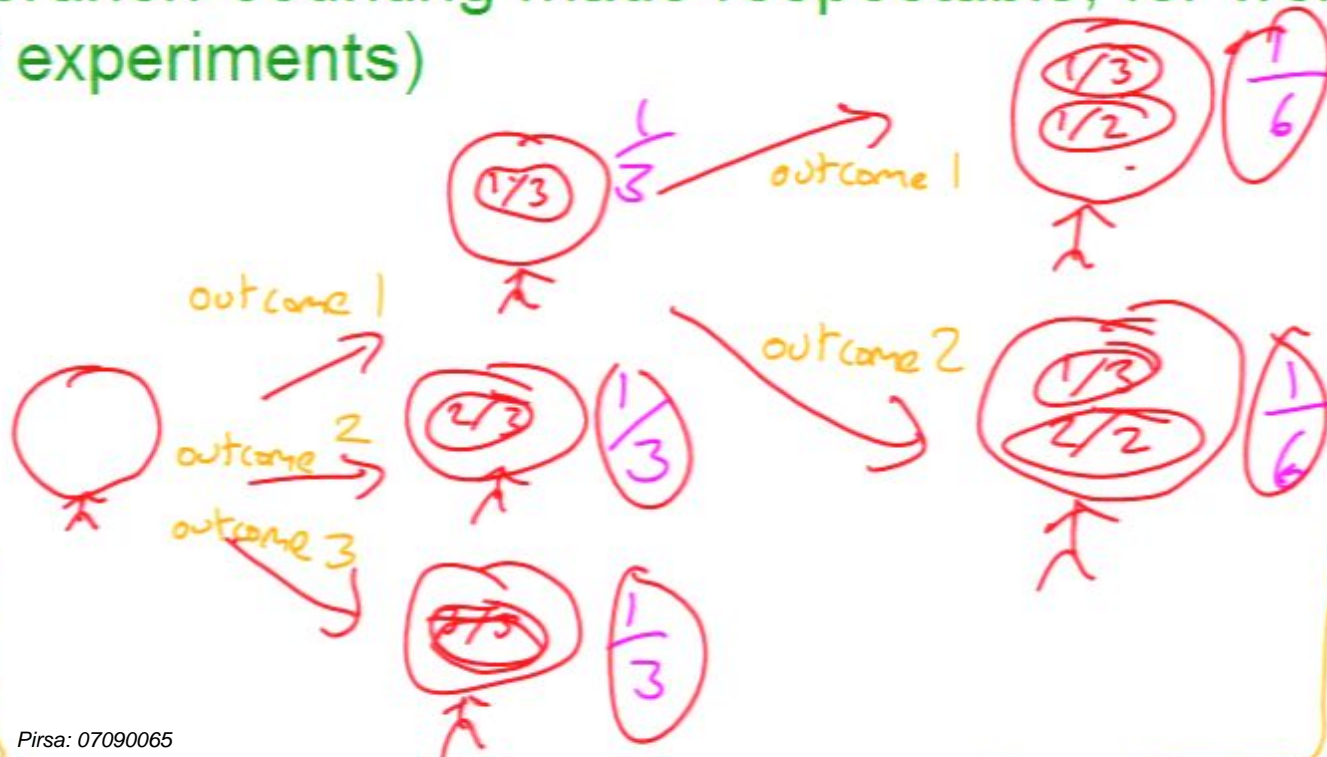
(iii) branch-counting using a preferred consistent set



Some problems with the Deutsch-Wallace argument

There are other well-defined strategies, which it seems defensible to term rational, for example:

(i) weighting successors according to their memory patterns (branch-counting made respectable, for well-defined sequences of experiments)



(ii) weighting by a post-selected final state

(iii) branch-counting using a preferred consistent set



Some problems with the Deutsch-Wallace argument

In any case, a unique well-defined decision strategy doesn't necessarily deserve the title "rational" or shed any insight into the structure of a universe: it could be unique and still apparently pointless.

Think of a powerful universe creator and destroyer who writes numbers in the sky that behave like probabilities under splitting. Treating them as probabilities is a well-defined caring strategy - perhaps the only possible such - but they're still just numbers in the sky.



Some problems with the Deutsch-Wallace argument

In any case, a unique well-defined decision strategy doesn't necessarily deserve the title "rational" or shed any insight into the structure of a universe: it could be unique and still apparently pointless.

Think of a powerful universe creator and destroyer who writes numbers in the sky that behave like probabilities under splitting. Treating them as probabilities is a well-defined caring strategy - perhaps the only possible such - but they're still just numbers in the sky.



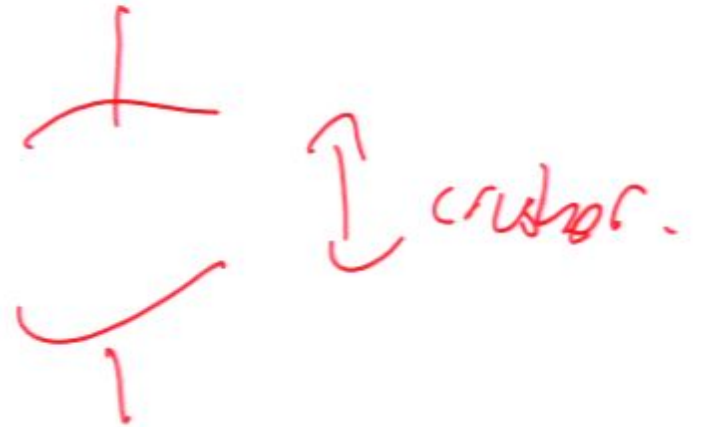
Some problems with the Deutsch-Wallace argument

In any case, a unique well-defined decision strategy doesn't necessarily deserve the title "rational" or shed any insight into the structure of a universe: it could be unique and still apparently pointless.

Think of a powerful universe creator and destroyer who writes numbers in the sky that behave like probabilities under splitting. Treating them as probabilities is a well-defined caring strategy - perhaps the only possible such - but they're still just numbers in the sky.

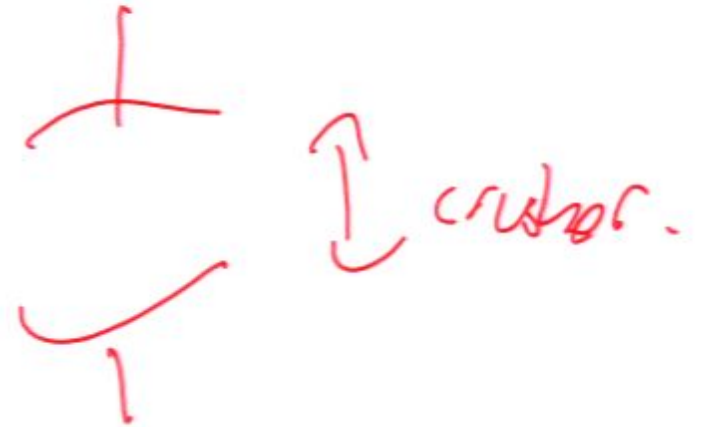








You create
un. verses

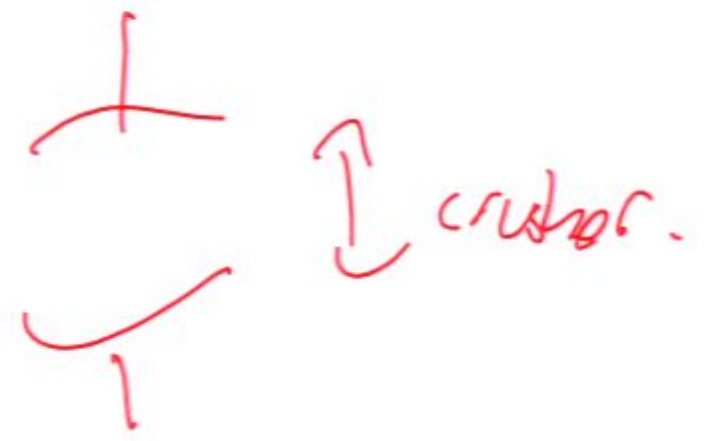




You create
universes



with conscious
beings and forms

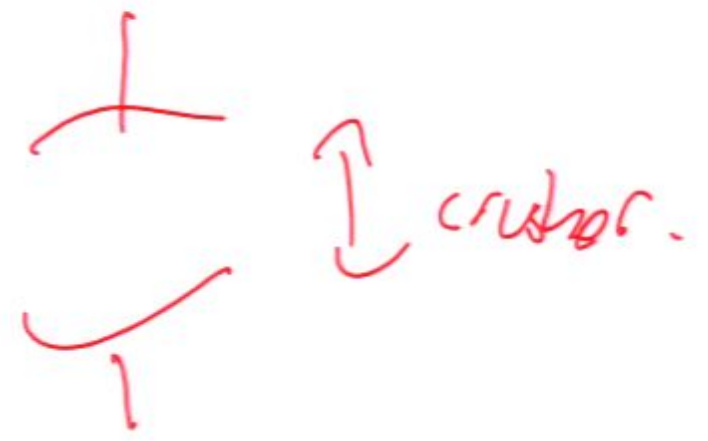




You create
universes



with conscious
beings and (from within universe)
appa

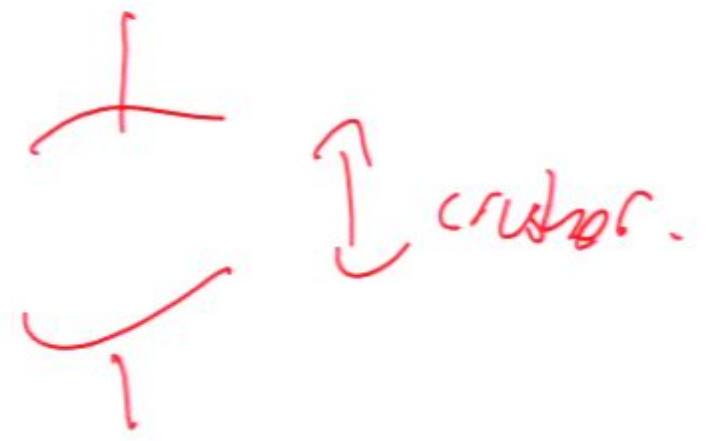


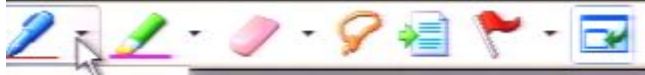


You create
universes



with conscious
beings and (from within universe)
apparently random events.





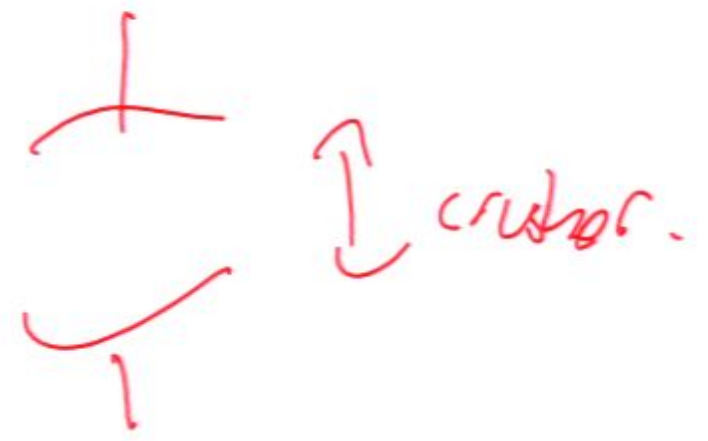
- Fine Chisel
- Fine Point
- Very Fine Point
- Medium Point
- Marker (2 mm)
- Pen Settings...

create



un. verses

with conscious
 beings and (from within universe)
 apparently random events.

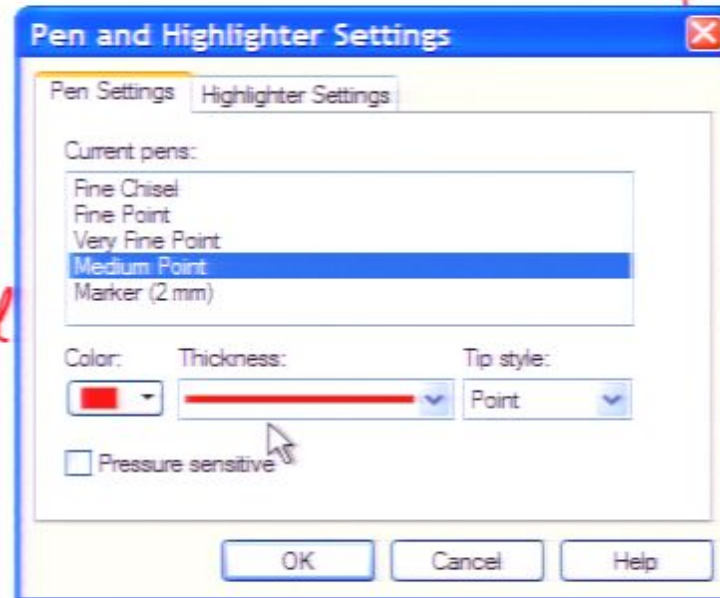




You create
un. verses



with conscious
beings and (from
apparently mad

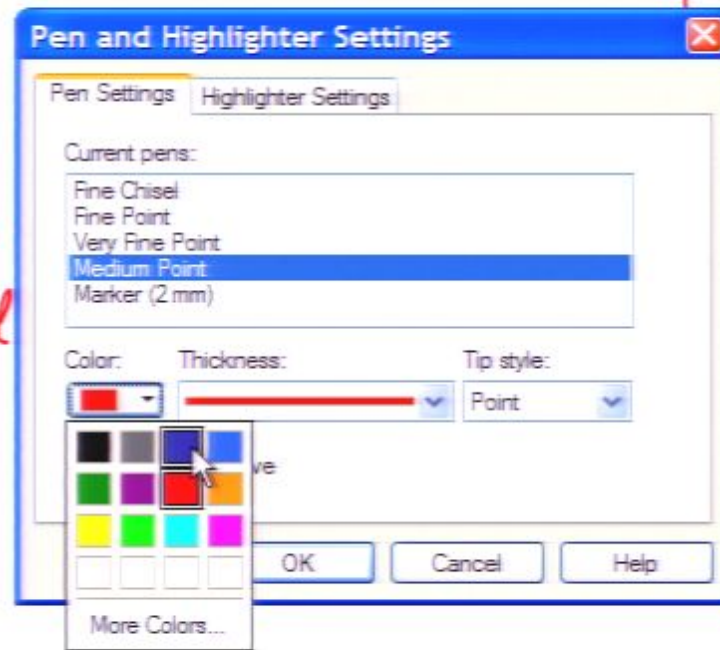
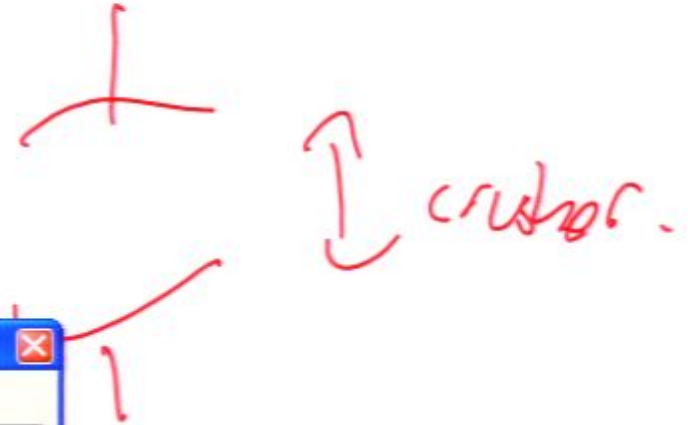




You create
universes



with conscious
beings and (from
apparently mad

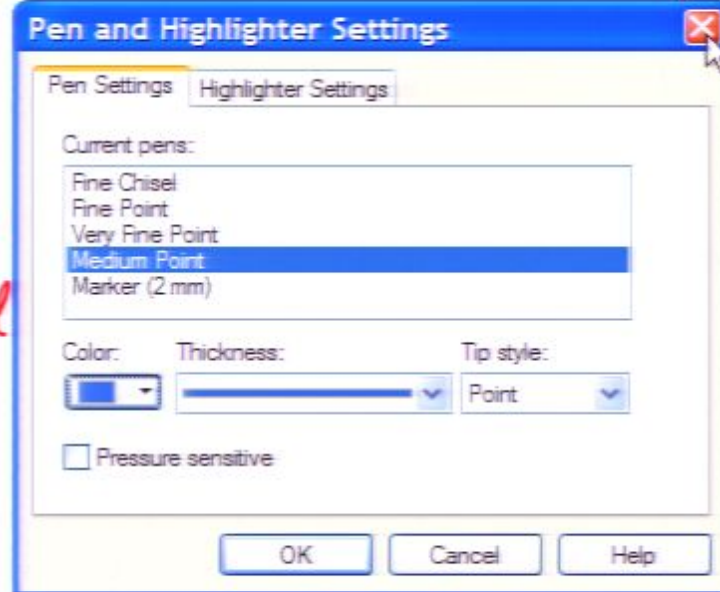
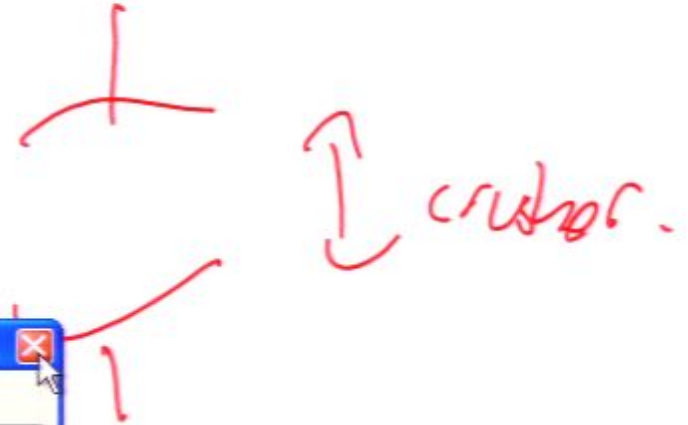




You create
universes



with conscious
beings and (from
apparently and



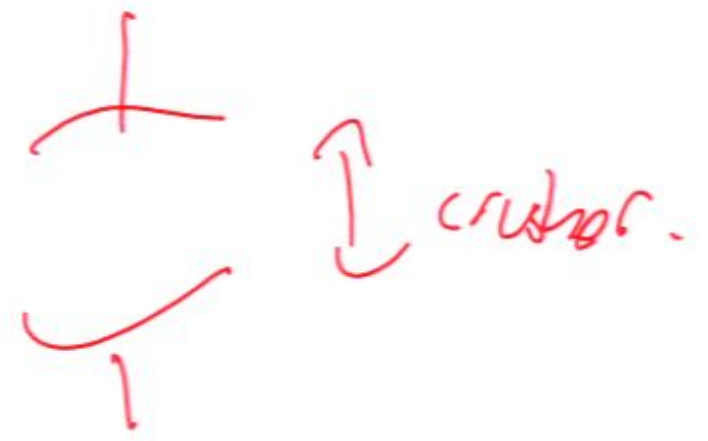


You create
universes



with conscious
beings and (from within universe)
apparently random events.

Or





- Fine Chisel
- Fine Point
- Very Fine Point
- Medium Point
- Marker (2 mm)
- Pen Settings

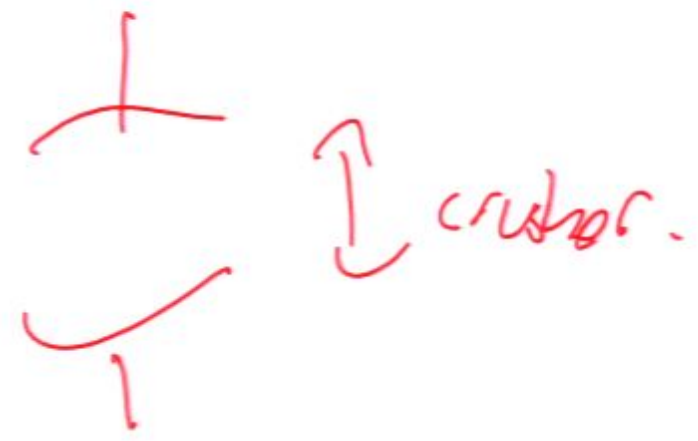
create



un. verses

with conscious
 beings and (from within universe)
 apparently random events.

Outcomes



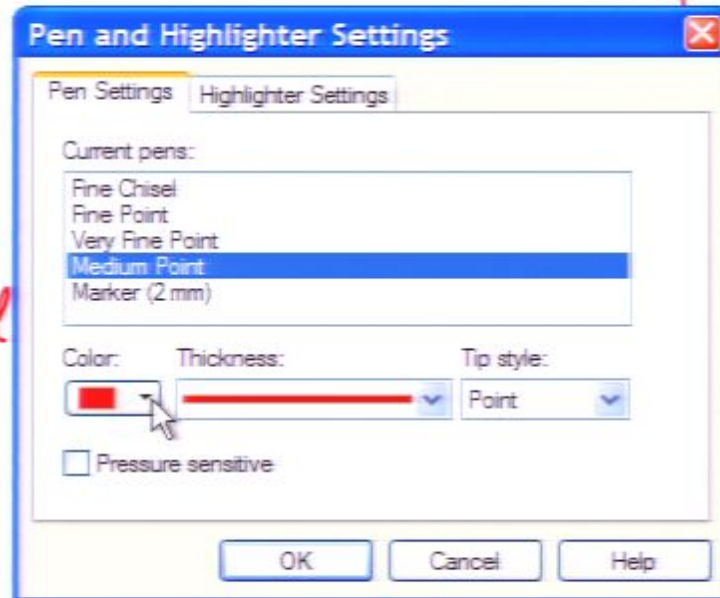
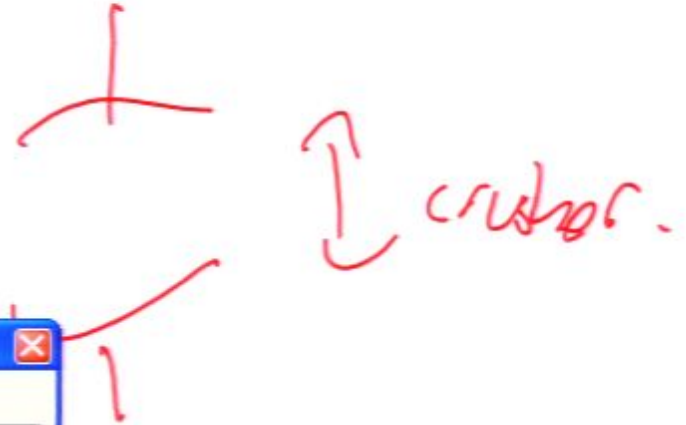


You create
un. verses



with conscious
beings and (from
apparently mad

Outcomes



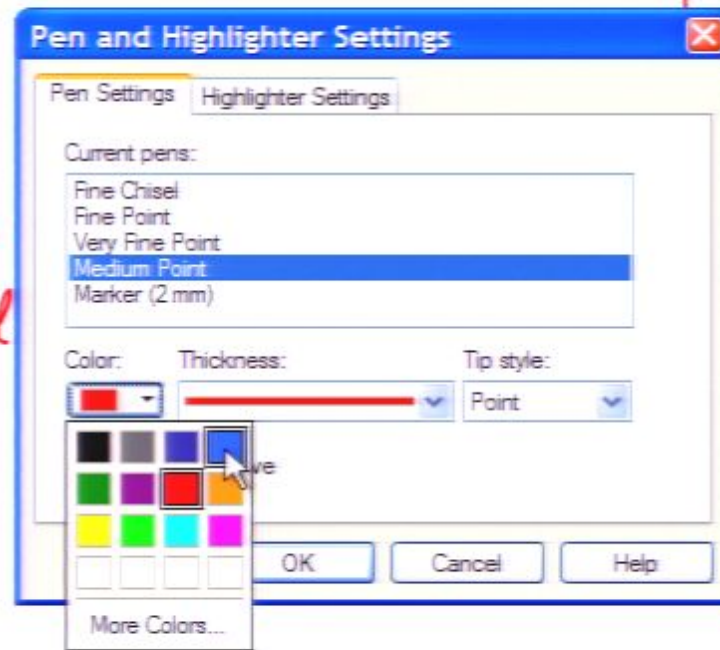
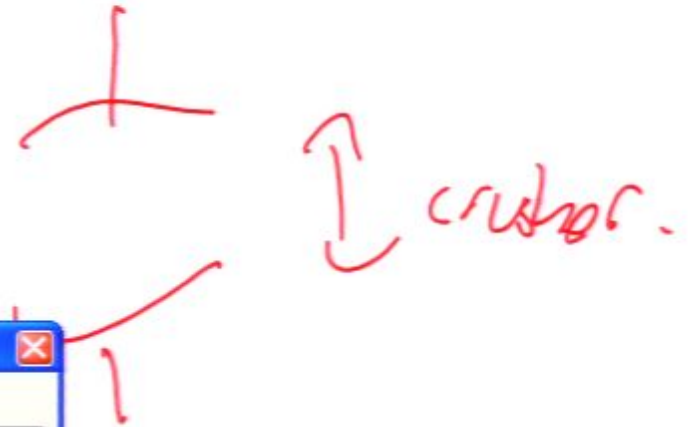


You create
universes



with conscious
beings and (from
apparently random

Outcomes



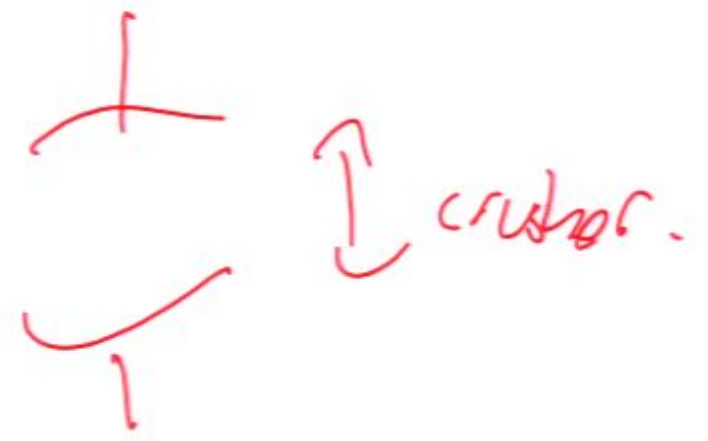


You create
universes



with conscious
beings and (from within universe)
apparently random events.

Outcomes

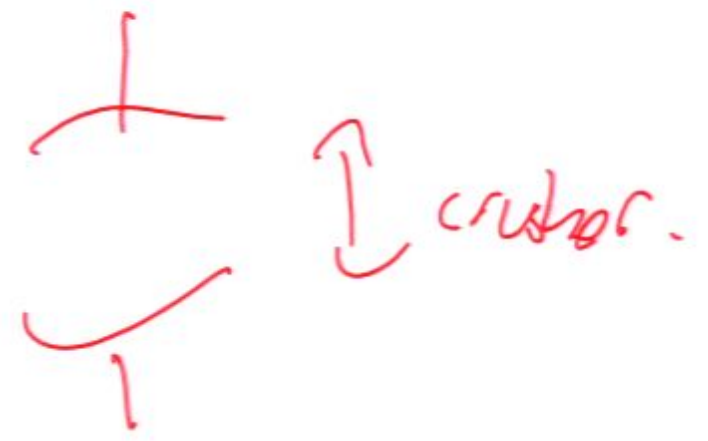




You create
universes



with conscious
beings and (from within universe)
apparently random events.

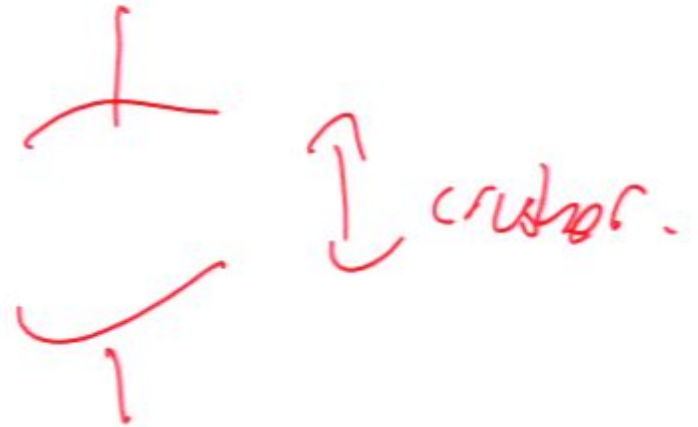




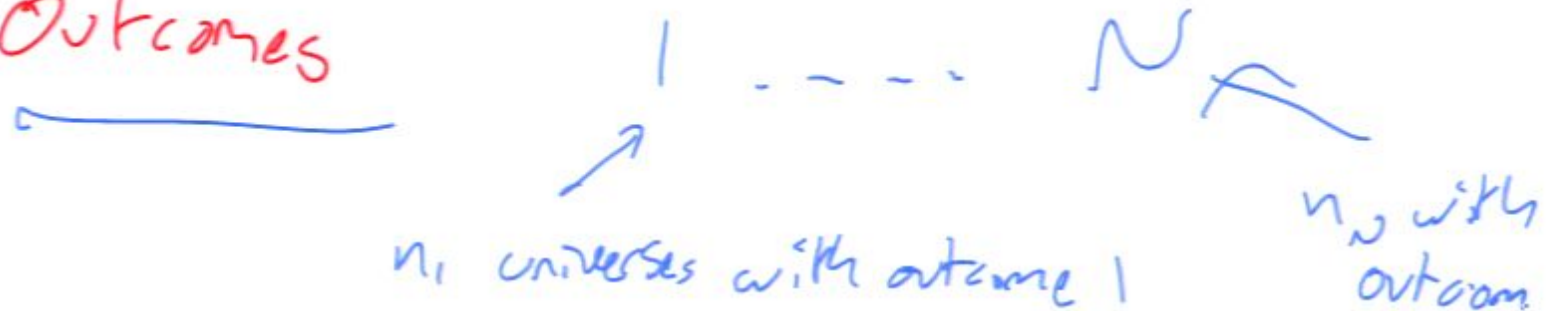
You create
universes



with conscious
beings and (from within universe)
apparently random events.

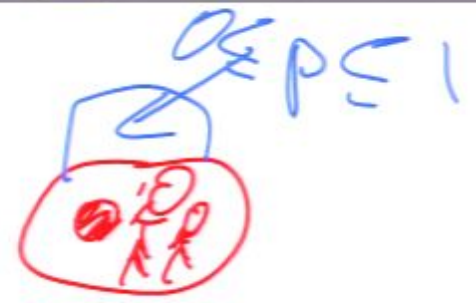


Outcomes

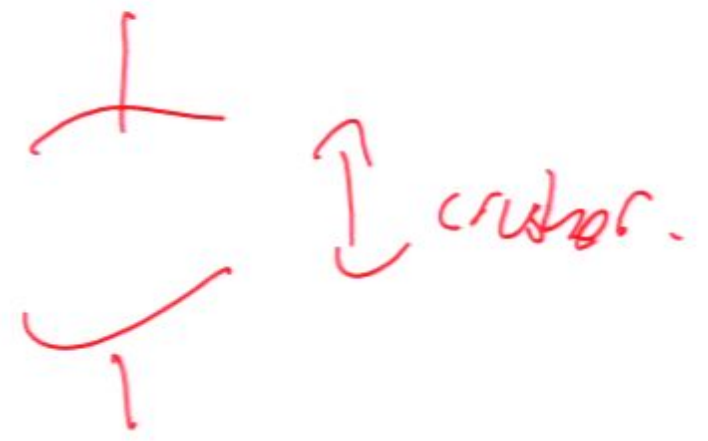




You create
universes

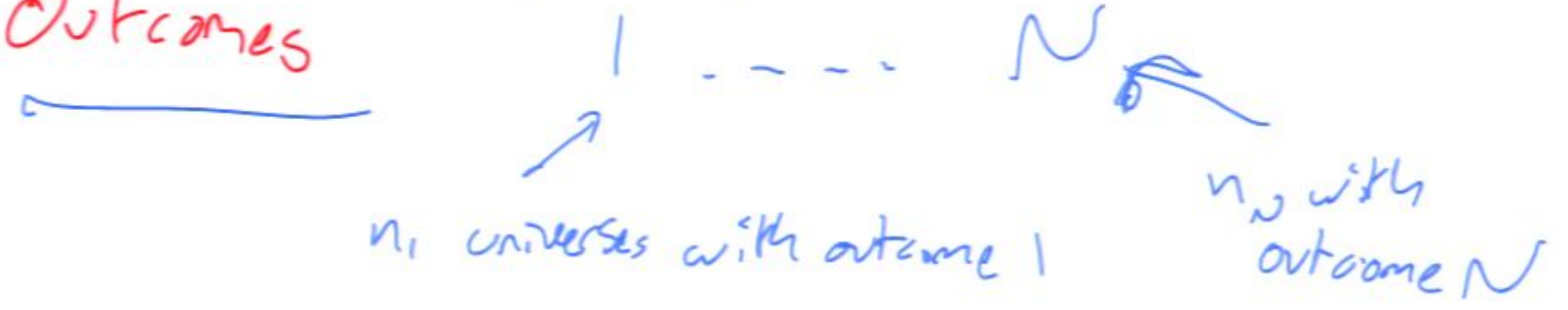


with conscious
beings and (from within universe)
apparently random events.



Outcomes

numbers in this list sum:

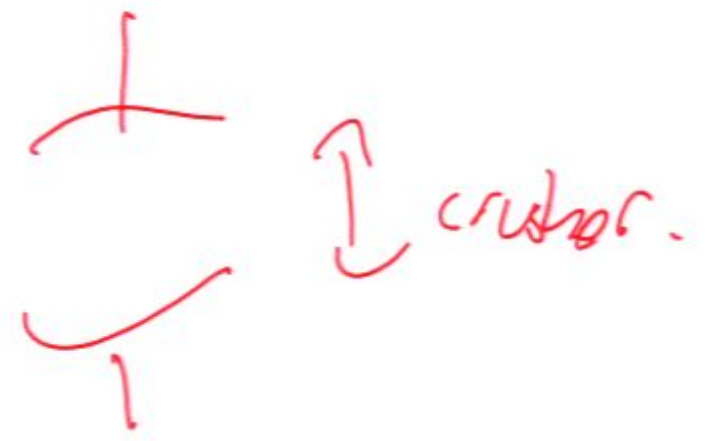




You create
universes

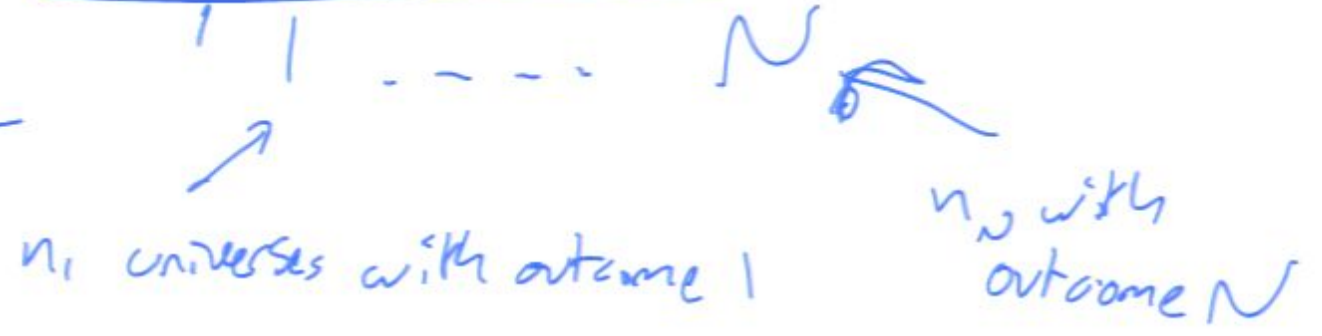


with conscious
beings and (from within universe)
apparently random events.



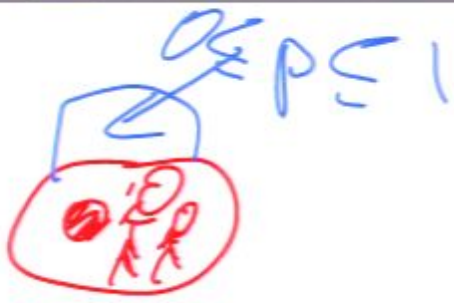
Outcomes

numbers in this they sum to P_j

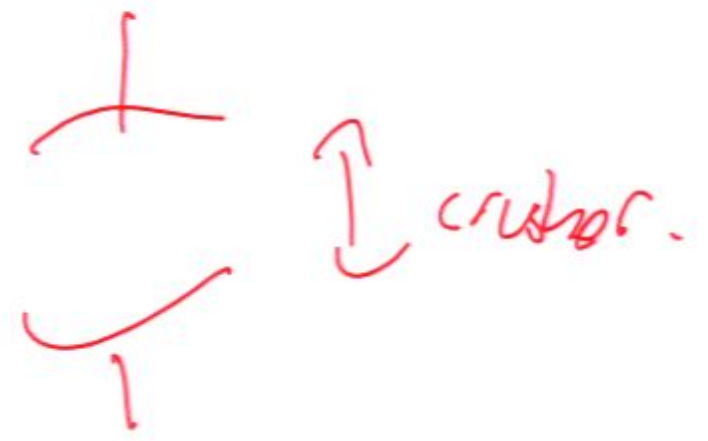




You create universes



with conscious beings and (from within universe) apparently random events.

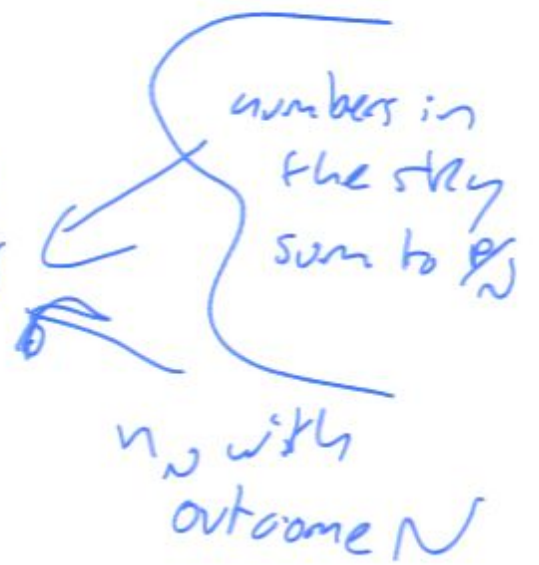


Outcomes

numbers in the n_i sum to P_i



n_i universes with outcome 1

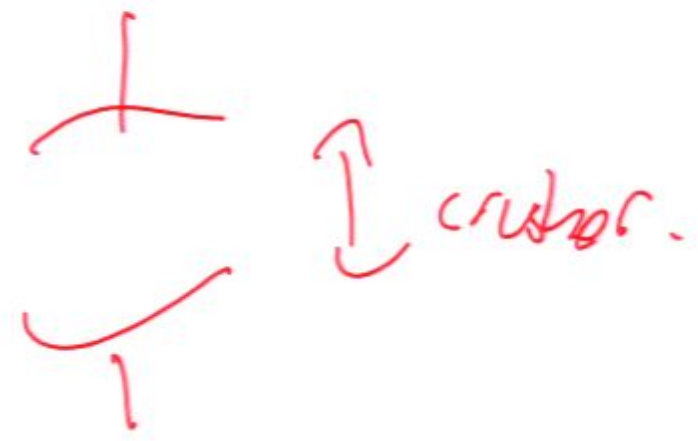




You create
universes



with conscious
beings and (from within universe)
apparently random events.

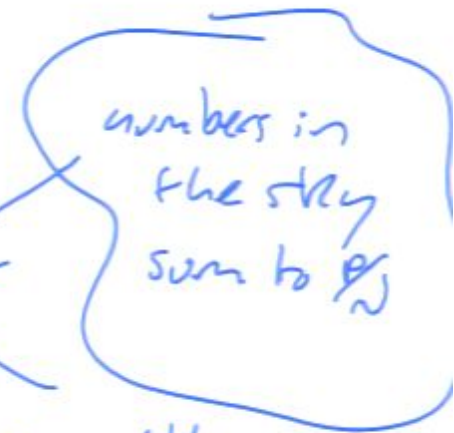


Outcomes

numbers in the fly sum to $\frac{P}{N}$



n_1 universes with outcome 1



n_N with
outcome N

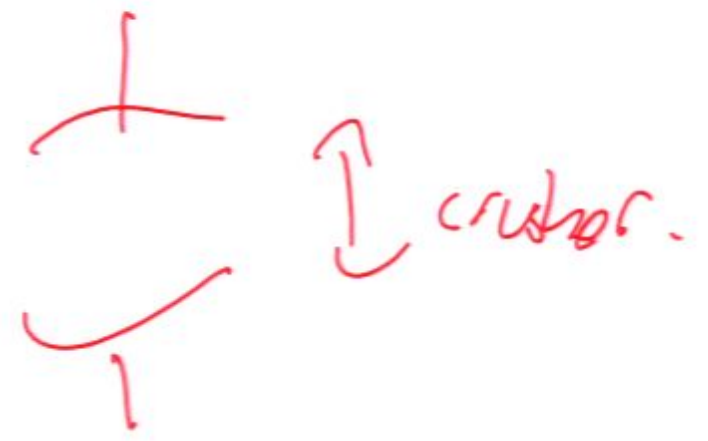
should
Case



You create universes



with conscious beings and (from within universe) apparently random events.



Outcomes

numbers in the P_i sum to P_i



numbers in the P_N sum to P_N

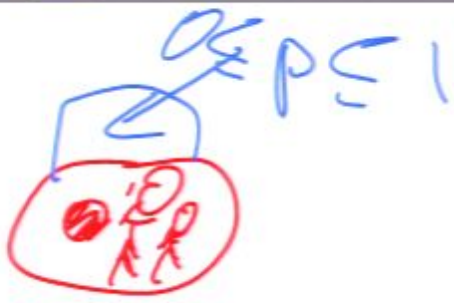
n_N with outcome N

should (only well-defined strategies) case

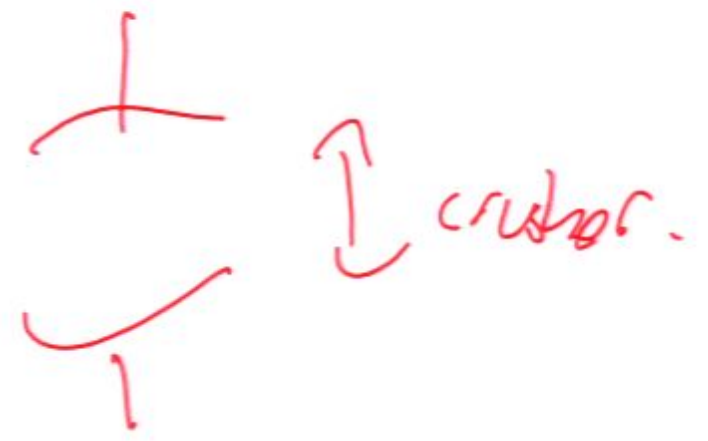
n_1 universes with outcome 1



You create
universes

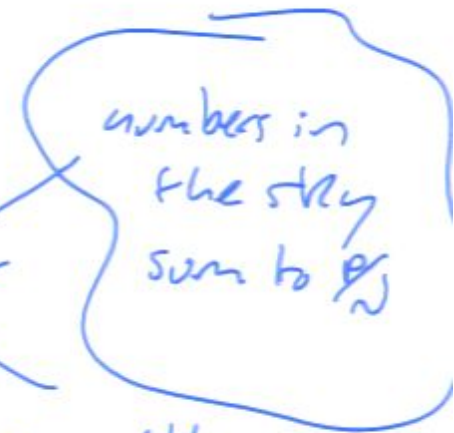


with conscious
beings and (from within universe)
apparently random events.



Outcomes

numbers in the sky sum to $\frac{P}{N}$



n_i with
outcome N

should
Care

(only well-defined strategies)

about successors in proportion in to P



Replacing "caring weights" by "consciousness weights"

If, indeed, stories about caring weights are inadequate to make Everett into a good scientific theory, what could do the trick?

We seem to need a "consciousness weight", defined precisely so as to allow us to infer statements about the probable form of our conscious existence. This doesn't seem to follow from anything simpler. Effectively we need, after all, to postulate:

For each branch i with wavefunction ψ_i contains N_i creatures with property P , the probability that one finds oneself in branch j , given that one has property P , is

$$\frac{N_j |\psi_j|^2}{\sum_i N_i |\psi_i|^2}$$



Replacing "caring weights" by "consciousness weights"

If, indeed, stories about caring weights are inadequate to make Everett into a good scientific theory, what could do the trick?

We seem to need a "consciousness weight", defined precisely so as to allow us to infer statements about the probable form of our conscious existence. This doesn't seem to follow from anything simpler. Effectively we need, after all, to postulate:

For each branch i with wavefunction ψ_i contains N_i creatures with property P , the probability that one finds oneself in branch j , given that one has property P , is

$$\frac{N_j |\psi_j|^2}{\sum_i N_i |\psi_i|^2}$$



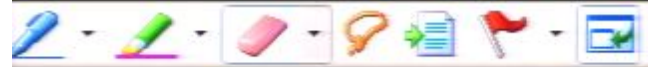
Replacing "caring weights" by "consciousness weights"

If, indeed, stories about caring weights are inadequate to make Everett into a good scientific theory, what could do the trick?

We seem to need a "consciousness weight", defined precisely so as to allow us to infer statements about the probable form of our conscious existence. This doesn't seem to follow from anything simpler. Effectively we need, after all, to postulate:

For each branch i with wavefunction ψ_i contains N_i creatures with property P , the probability that one finds oneself in branch j , given that one has property P , is

$$\frac{N_j |\psi_j|^2}{\sum_i N_i |\psi_i|^2}$$



Replacing "caring weights" by "consciousness weights"

If, indeed, stories about caring weights are inadequate to make Everett into a good scientific theory, what could do the trick?

We seem to need a "consciousness weight", defined precisely so as to allow us to infer statements about the probable form of our conscious existence. This doesn't seem to follow from anything simpler. Effectively we need, after all, to postulate:

For each branch i with wavefunction ψ_i contains N_i creatures with property P , the probability that one finds oneself in branch j , given that one has property P , is

$$\frac{N_j |\psi_j|^2}{\sum_i N_i |\psi_i|^2}$$



So why not postulate a consciousness weight?

.e. postulate this:

If each branch i with wavefunction ψ_i contains N_i creatures with property P , the probability that one finds oneself in branch j , given that one has property P , is

$$\frac{N_j |\psi_j|^2}{\sum_i N_i |\psi_i|^2}$$

One could. But note the consequences:

- 1) The whole Everett interpretation - probability postulate as well as preferred basis choice - has now been reduced to (dubious?) hypotheses about consciousness.
- 2) Since possible quantum events are continually affecting or failing to affect our brain state, we have no solid notion of identity moving forward in time, even moment to moment.



One could just postulate an existence weight. But if so note:

3) The case for functionalism, a theory of consciousness that's supposed to apply to creatures dynamically interacting with their environment, seems more than usually questionable if identities are merely instantaneous.

4) We seem to have arrived back at something very like the Bell-Barbour picture of Everett as a "marvellous moment" theory, with identity failing to propagate forward in time and memory having no necessary connection with a personal past. (Consistent, sure. But attractive?)

5) We're apparently inescapably committed to a form of anthropic principle in order to extract useful inferences from the theory.

6) $\sum N_i | \psi_i |^2$ surely isn't a unique probability measure over instantaneous mind-states.



A small point to please everyone*

*?

Everett as just outlined makes a subtly different prediction from "standard" one-branch q.m. Evolutionary success in Everett is measured by

$$\sum N_i |r_i|^2,$$

a measure insensitive to the risk variance of an evolutionary strategy. A very-high-extinction-risk-even-higher-reward strategy pays off handsomely in an Everett universe, while being near-certainly suicidal in a one-branch universe. (Gamblers with positive expectation don't need to worry about risk of ruin if they care only about the weighted sum of wealth over worlds.)

If we found evidence that our ancestors pursued such a strategy at some point, and that as a result we dominate the distribution of intelligent life, it would tend to support Everett.



A small point to please everyone*



Everett as just outlined makes a subtly different prediction from "standard" one-branch q.m. Evolutionary success in Everett is measured by

$$\sum N_i |r_i|^2,$$

a measure insensitive to the risk variance of an evolutionary strategy. A very-high-extinction-risk-even-higher-reward strategy pays off handsomely in an Everett universe, while being near-certainly suicidal in a one-branch universe. (Gamblers with positive expectation don't need to worry about risk of ruin if they care only about the weighted sum of wealth over worlds.)

If we found evidence that our ancestors pursued such a strategy at some point, and that as a result we dominate the distribution of intelligent life, it would tend to support Everett.



A small point to please everyone*

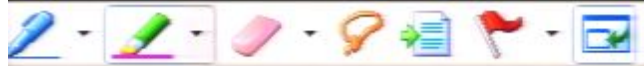


Everett as just outlined makes a subtly different prediction from "standard" one-branch q.m. Evolutionary success in Everett is measured by

$$\sum N_i | \psi_i |^2$$

a measure insensitive to the risk variance of an evolutionary strategy. A very-high-extinction-risk-even-higher-reward strategy pays off handsomely in an Everett universe, while being near-certainly suicidal in a one-branch universe. (Gamblers with positive expectation don't need to worry about risk of ruin if they care only about the weighted sum of wealth over worlds.)

If we found evidence that our ancestors pursued such a strategy at some point, and that as a result we dominate the distribution of intelligent life, it would tend to support Everett.



A small point to please everyone*



Everett as just outlined makes a subtly different prediction from "standard" one-branch q.m. Evolutionary success in Everett is measured by

$$\sum N_i | \psi_i |^2$$

a measure insensitive to the risk variance of an evolutionary strategy. A very-high-extinction-risk-even-higher-reward strategy pays off handsomely in an Everett universe, while being near-certainly suicidal in a one-branch universe. (Gamblers with positive expectation don't need to worry about risk of ruin if they are only about the weighted sum of wealth over worlds.)

If we found evidence that our ancestors pursued such a strategy at some point, and that as a result we dominate the distribution of intelligent life, it would tend to support Everett.



A small point to please everyone*



Everett as just outlined makes a subtly different prediction from "standard" one-branch q.m. Evolutionary success in Everett is measured by

$$\sum N_i | \psi_i |^2$$

measure insensitive to the risk variance of an evolutionary strategy. A very-high-extinction-risk-even-higher-reward strategy pays off handsomely in an Everett universe, while being near-certainly suicidal in a one-branch universe. (Gamblers with positive expectation don't need to worry about risk of ruin if they are only about the weighted sum of wealth over worlds.)

If we found evidence that our ancestors pursued such a strategy at some point, and that as a result we dominate the distribution of intelligent life, it would tend to support Everett.



Everett and Evolution

Note Title

15/09/2007

Adrian Kent

(DAMTP, University of Cambridge
and
Perimeter Institute)

"Many Worlds At 50" conference talk

²
163022092007@PI



Taking branching seriously: "real world" interpretations (AK, arXiv:0708.3710)

A new proposal for solving the measurement problem.

Respects Lorentz invariance and conservation laws.

No immediately evident "tails problem".

Adds minimal extra structure to q.m., namely a preferred factorisation $H_A \otimes H_B$ and a preferred (possibly hypersurface dependent) projective decomposition on H_B .



branches separating over finite time interval

The branching structure which emerges is one in which individual branches are defined by mixed states, which become perfectly orthogonal (and pure) only as $t \rightarrow \infty$.

Measurement processes lead to outcome branches which become approximately orthogonal over finite time.





Taking branching seriously: "real world" interpretations (AK, arXiv:0708.3710)

A new proposal for solving the measurement problem.

Respects Lorentz invariance and conservation laws.

No immediately evident "tails problem".

Adds minimal extra structure to q.m., namely a preferred factorisation $H_A \otimes H_B$ and a preferred (possibly hypersurface dependent) projective decomposition on H_B .



branches separating over finite time interval

The branching structure which emerges is one in which individual branches are defined by mixed states, which become perfectly orthogonal (and pure) only as $t \rightarrow \infty$.

Measurement processes lead to outcome branches which become approximately orthogonal over finite time.



Taking branching seriously: "real world" interpretations (AK, arXiv:0708.3710)

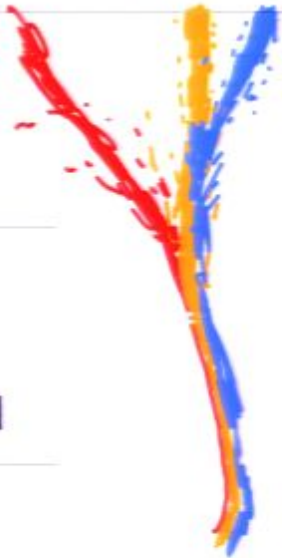
A new proposal for solving the measurement problem.

Respects Lorentz invariance and conservation laws.

No immediately evident "tails problem".

Adds minimal extra structure to q.m., namely a preferred

factorisation $H_A \otimes H_B$ and a preferred (possibly hypersurface dependent) projective decomposition on H_B .



branches separating over finite time interval

The branching structure which emerges is one in which individual branches are defined by

mixed states, which become perfectly orthogonal (and pure) only as $t \rightarrow \infty$.

Measurement processes lead to outcome branches which become approximately orthogonal over finite time.



Taking branching seriously: "real world" interpretations (AK, arXiv:0708.3710)

A new proposal for solving the measurement problem.

Respects Lorentz invariance and conservation laws.

No immediately evident "tails problem".

Adds minimal extra structure to q.m., namely a preferred factorisation $H_A \otimes H_B$ and a preferred (possibly hypersurface dependent) projective decomposition on H_B .



branches separating over finite time interval

The branching structure which emerges is one in which individual branches are defined by mixed states, which become perfectly orthogonal (and pure) only as $t \rightarrow \infty$.

Measurement processes lead to outcome branches which become approximately orthogonal over finite time.



A small point to please everyone*



Everett as just outlined makes a subtly different prediction from "standard" one-branch q.m. Evolutionary success in Everett is measured by

$$\sum N_i | \psi_i |^2$$

a measure insensitive to the risk variance of an evolutionary strategy. A very-high-extinction-risk-even-higher-reward strategy pays off handsomely in an Everett universe, while being near-certainly suicidal in a one-branch universe. (Gamblers with positive expectation don't need to worry about risk of ruin if they are only about the weighted sum of wealth over worlds.)

If we found evidence that our ancestors pursued such a strategy at some point, and that as a result we dominate the distribution of intelligent life, it would tend to support Everett.





How could probability arise in an Everettian universe at all, given that the dynamics are deterministic?

One popular line of thought argues for some notion of **subjective uncertainty** in an Everettian universe. According to Saunders & Wallace, Everett plus functionalism implies such a notion: **I will become** one of the two possible future versions of myself after observing a Stern-Gerlach experiment, **but I don't know which**. I think (as Vaidman and Greaves have argued) that this is simply a mistake:





Wallace (2002) summarizes Saunders thus:

"There are only three logical possibilities:

- 1) I should expect to become both future selves.*
- 2) I should expect to become one or the other.*
- 3) I should expect nothing: oblivion.*

Of these (3) seems absurd. (1) is at least coherent- we could imagine some telepathic link- but ... this link will have to supervene on some physical interaction between the two copies which is not in fact present. This leaves (2) as the only option."

I find this mystifying. (3) is coherent, and might be justifiable: if identity turns out to fail to propagate forwards in time then indeed "I" would cease to exist.

In any case, (1) makes perfect sense and requires no telepathy. (2) is unsupported by the dynamics and wave function ontology, and so must be rejected.





A small point to please everyone*



Everett as just outlined makes a subtly different prediction from "standard" one-branch q.m. Evolutionary success in Everett is measured by

$$\sum N_i |r_i|^2,$$

a measure insensitive to the risk variance of an evolutionary strategy. A very-high-extinction-risk-even-higher-reward strategy pays off handsomely in an Everett universe, while being near-certainly suicidal in a one-branch universe. (Gamblers with positive expectation don't need to worry about risk of ruin if they are only about the weighted sum of wealth over worlds.)

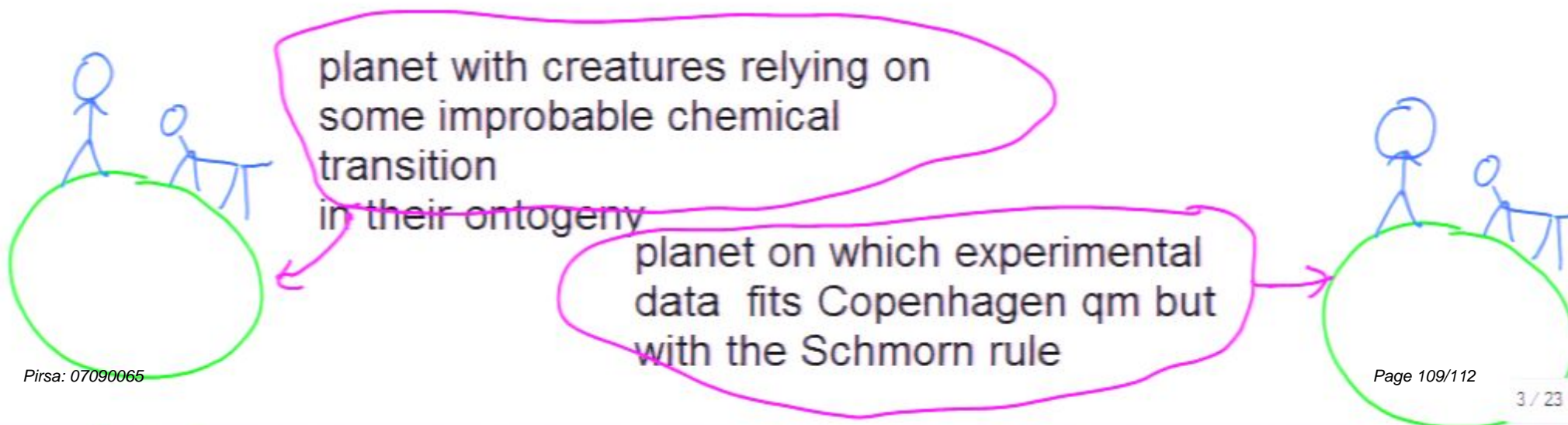
If we found evidence that our ancestors pursued such a strategy at some point, and that as a result we dominate the distribution of intelligent life, it would tend to support Everett.



I want to consider evolution in two senses:

1) the Darwinian sense, applied to an Everett universe

An Everettian cosmology predicts the existence of multiple intelligent species - including multiple versions of humanity - with histories differing in various ways from our own. Many of these have evolutionary and experimental histories that would be enormously improbable according to non-Everettian qm. What can and should we, as scientists, say about this account of life in the universe?





Taking branching seriously: "real world" interpretations (AK, arXiv:0708.3710)

A new proposal for solving the measurement problem.

Respects Lorentz invariance and conservation laws.

No immediately evident "tails problem".

Adds minimal extra structure to q.m., namely a preferred factorisation $H_A \otimes H_B$ and a preferred (possibly hypersurface dependent) projective decomposition on H_B .



branches separating over finite time interval

The branching structure which emerges is one in which individual branches are defined by mixed states, which become perfectly orthogonal (and pure) only as $t \rightarrow \infty$.

Measurement processes lead to outcome branches which become approximately orthogonal over finite time.



Everett-Darwinian evolution: the key question

Why, amongst all the possible types of intelligent creature with all the possible types of evolutionary and experimental history, do I find myself one of those whose history confirms*

Copenhagen q.m.?

(apparently, and within the theory's domain of validity)

A clearly inadequate answer: this is **consistent** with an Everettian account (true) and hence tends to **confirm** it scientifically (not true if our criteria is that a successor should **explain** the validity of Copenhagen q.m.).

"Your history might take any form" isn't an interesting result to merge from an account of cosmology and evolution.



I take Everett to agree that his programme succeeds iff it answers the key question:

"The aim is not to deny or contradict the conventional formulation of quantum theory... but rather to supply a new, more general and complete formulation, from which the conventional interpretation can be deduced." (Everett, 1957)

Now clearly Everettians won't say that the theory implies that we'll **certainly** find our history to be Copenhagen-esque. They can only hope to show it **probably** will. Fine: but *how can a probabilistic conclusion emerge from this deterministic theory?*