EPISTEMOLOGY, COMPUTATION AND THE LAWS OF PHYSICS

David H. Wolpert

NASA Ames Research Center

David.H.Wolpert@nasa.gov

http://ti.arc.nasa.gov/people/dhw/

Walter Read

Mathematics Department, CS Fresno

NASA-ARC-03-0667

ROADMAP

1) ***Shortcomings of current impossibility results concerning laws of physics***

↓

2) ***Knowledge operators and their shortcoming***

↓

3) ***Formalize mathematical structure shared by observation and prediction:*** **inference devices**

↓

4) ***Elementary properties of inference devices***

COMPUTATIONAL IMPOSSIBILITY IN PHYSICS

1) Impossibility results of Moore, Pour-El and Richards, etc., rely on uncountable number of states of universe.

- *What if universe is countable, or even finite?*
- *What if there exist oracles, so Halting theorem (the basis of those results) is irrelevant?*

2) Impossibility results of Lloyd rely on current model of laws of physics (e.g., no superluminal travel).

- *What if laws are actually different?*

COMPUTATIONAL IMPOSSIBILITY IN PHYSICS

3) To apply Godel's incompleteness theorem presumes physical laws are "written in predicate logic"

- ***Barrow: What if universe "written" in different lang.?***
- ***What if there are no "laws" at all, just a huge list of events, which just happen to appear to have patterns?***
- ***What if Godel-style intuitionism is correct?***

What If Our Models Are Wrong???

COMPUTATIONAL IMPOSSIBILITY IN PHYSICS

3) To apply Godel's incompleteness theorem presumes physical laws are "written in predicate logic"

- ***Barrow: What if universe "written" in different lang.?***
- ***What if there are no "laws" at all, just a huge list of events, which just happen to appear to have patterns?***
- ***What if Godel-style intuitionism is correct?***

What If Our Models Are Wrong???

Is there some more fundamental model, almost indisputable, that we can analyze?

ROADMAP

1) ***Shortcomings of current impossibility results concerning laws of physics***

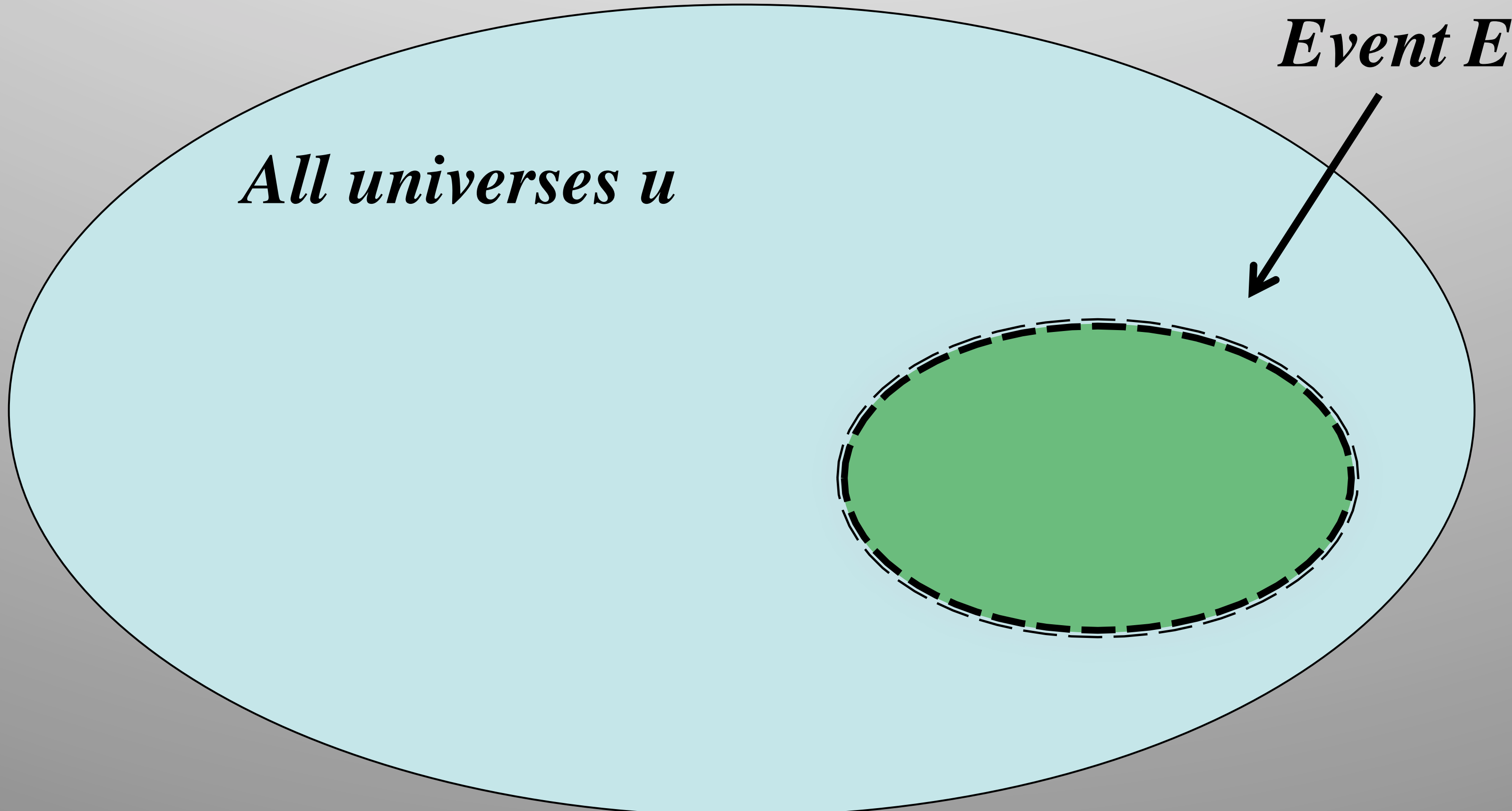
2) ***Knowledge operators and their shortcoming***

3) ***Formalize mathematical structure shared by observation and prediction:*** **inference devices**

4) ***Elementary properties of inference devices***

KNOWLEDGE OPERATORS

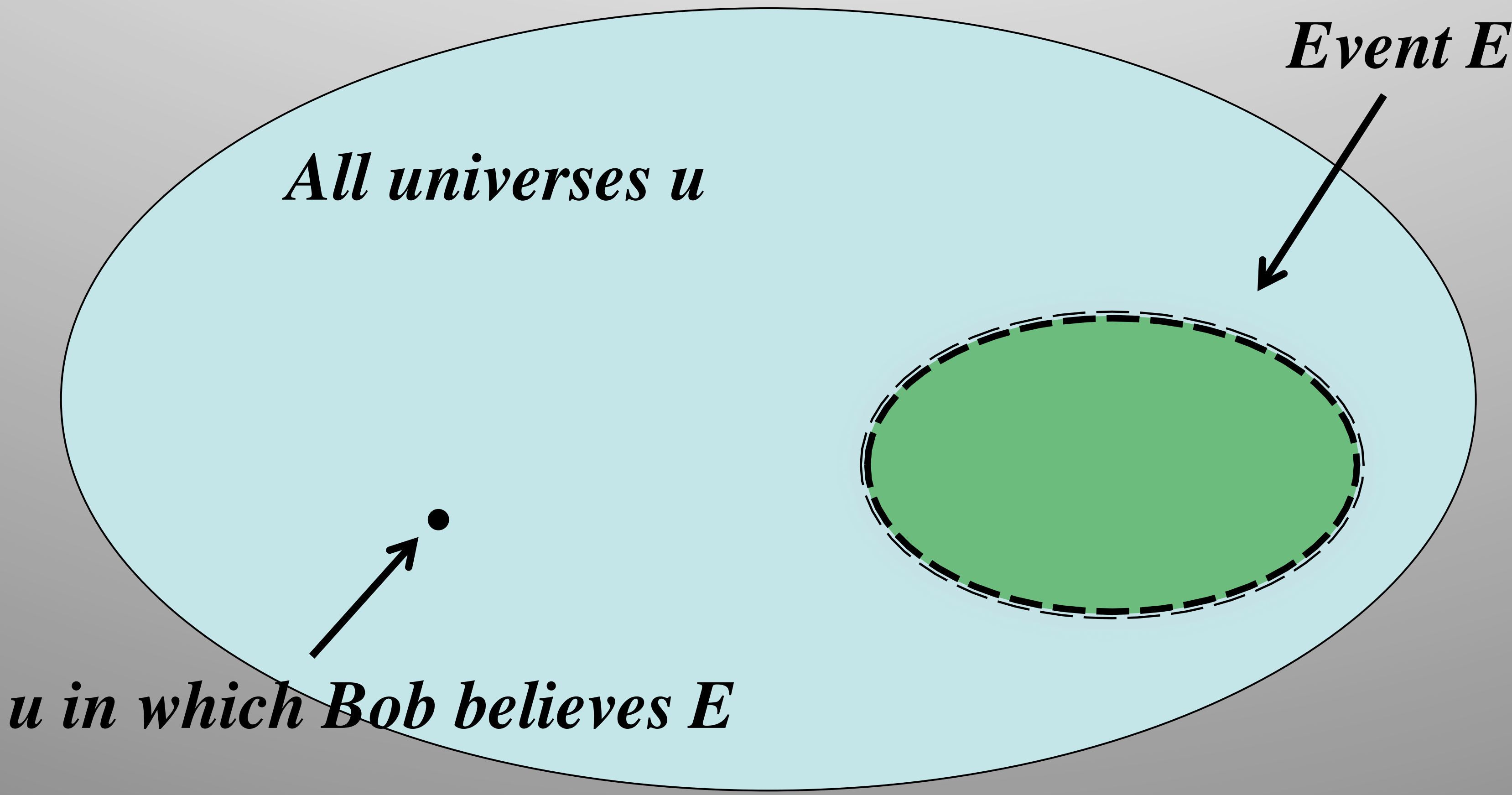
1) *To Bob, laws are patterns among "events Bob knows"*
2) *But what does it mean for Bob to "know" an event?*



KNOWLEDGE OPERATORS

1) At indicated u, Bob "believes" that he's in E.

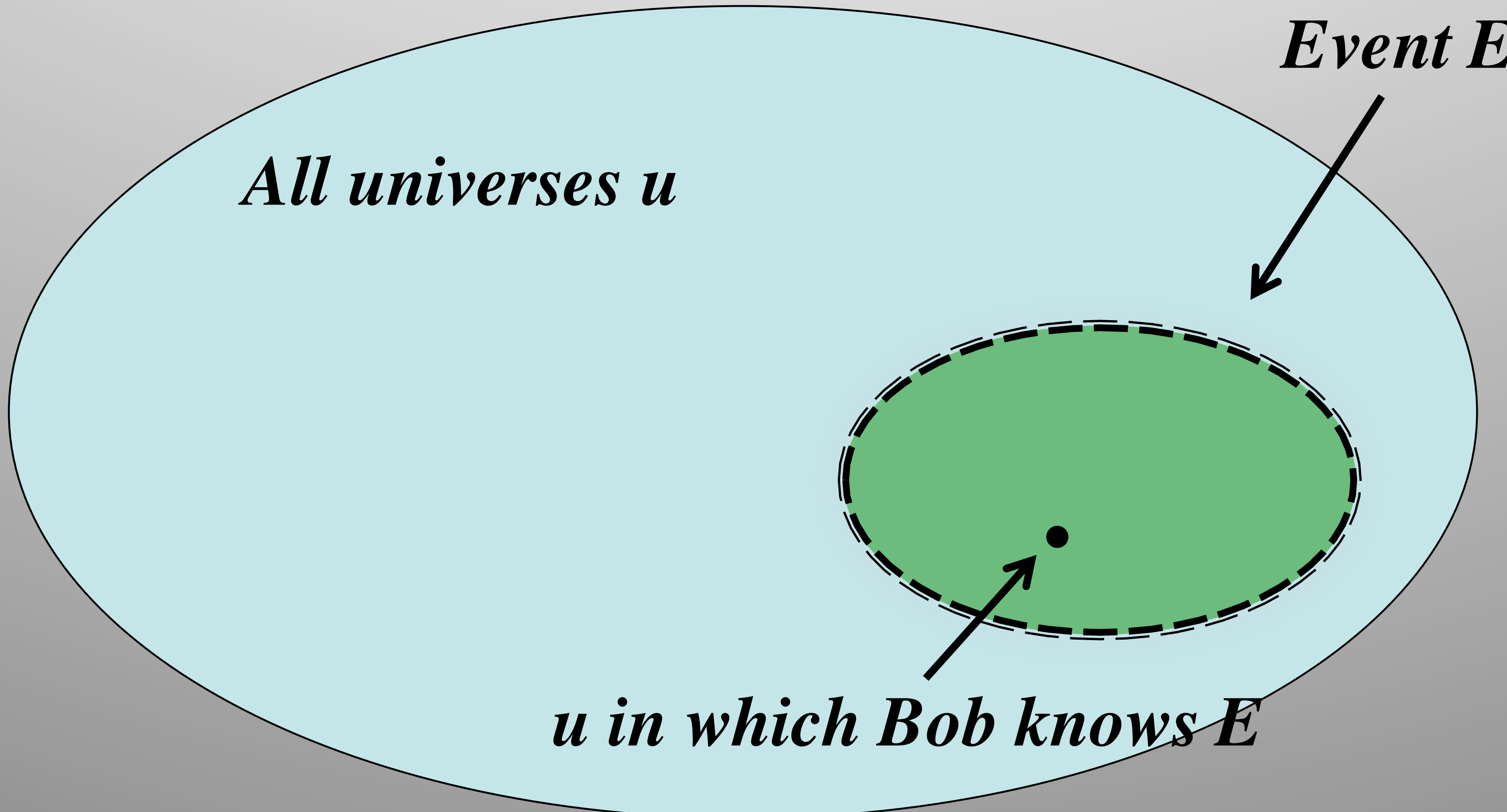
2) So belief is a function from u to subsets of {u}



KNOWLEDGE OPERATORS

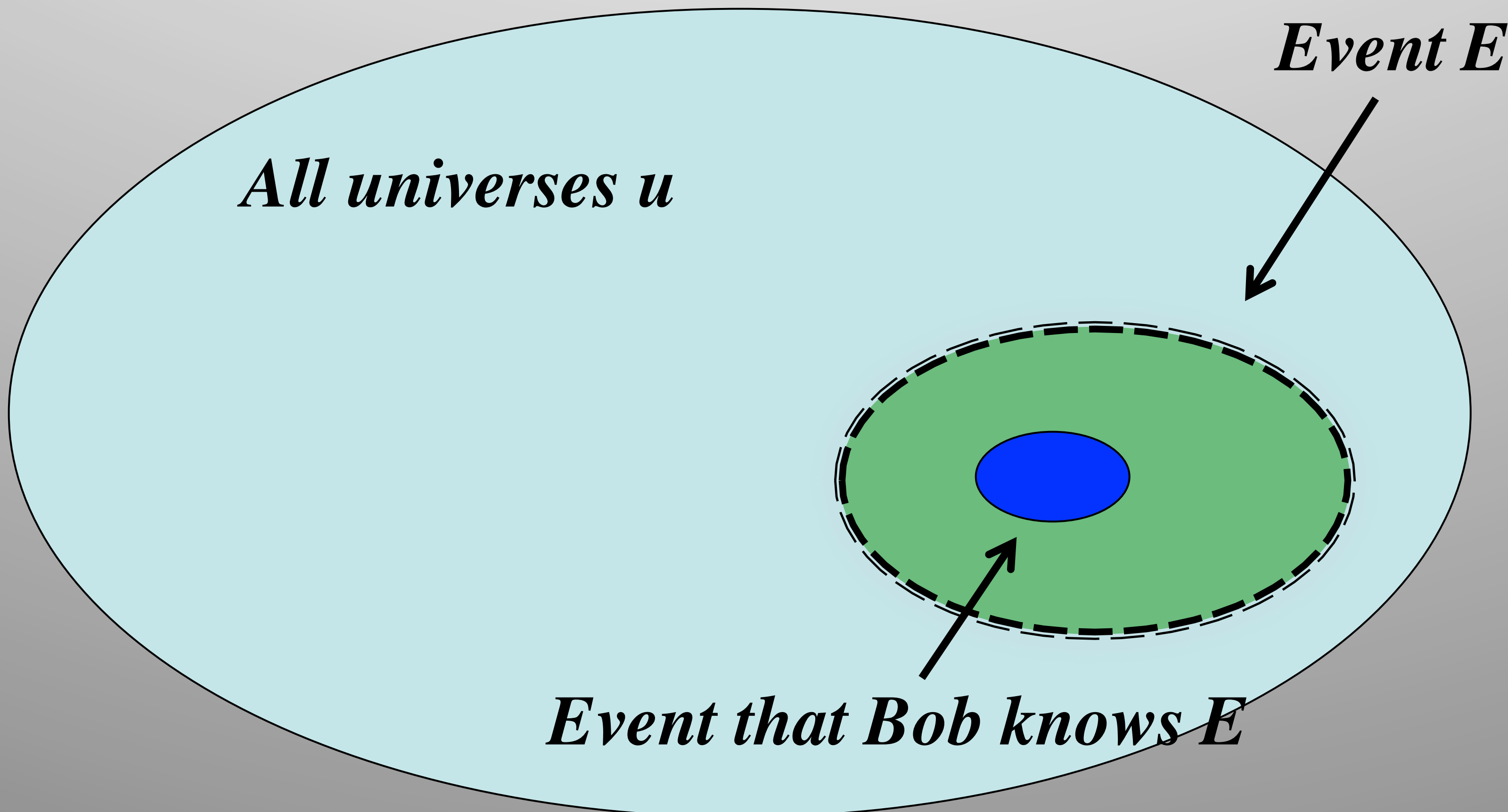
1) At indicated u, Bob <u>knows</u> he's in E.

2) A knowledge function is any belief function where the image of u contains u



KNOWLEDGE OPERATORS

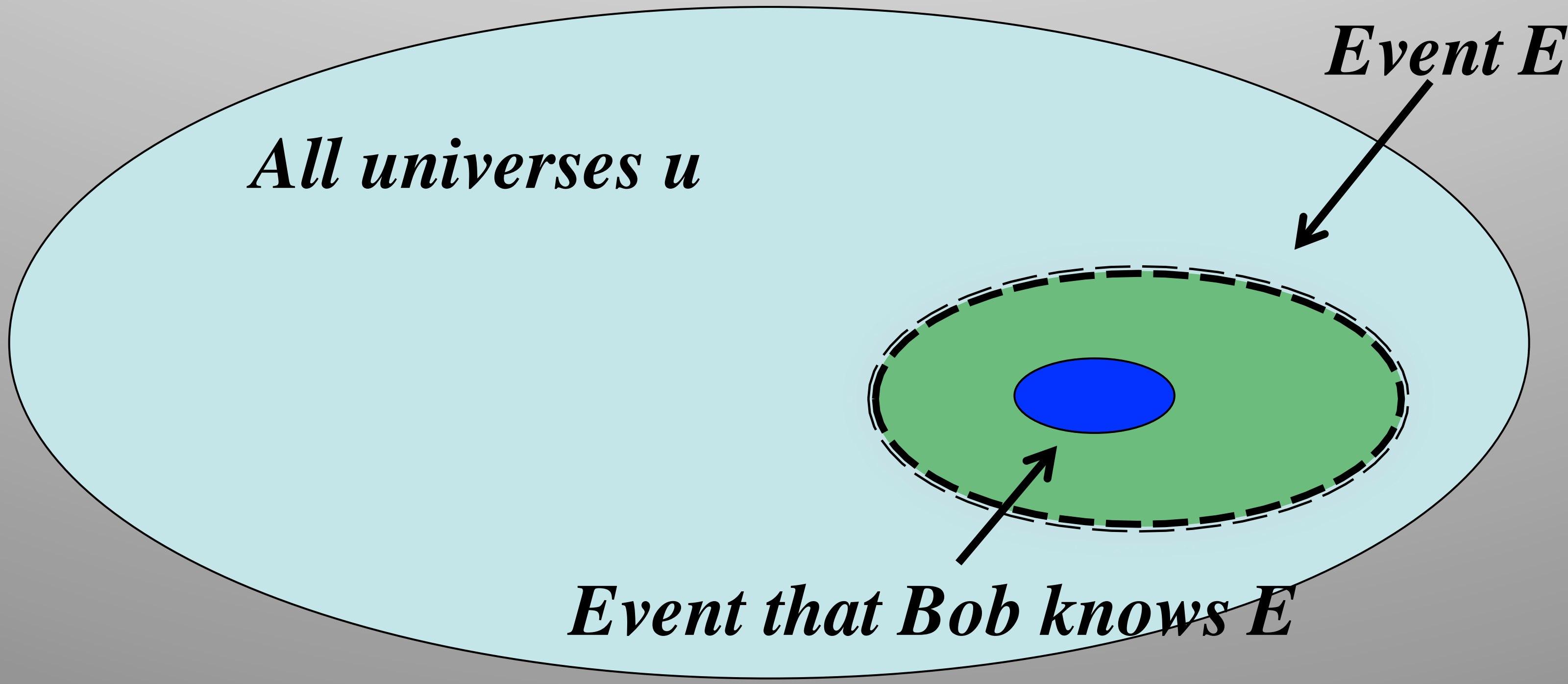
1) Throughout blue region, Bob knows he's in E.
(At other u in E, either Bob knows some E' that overlaps E, or Bob knows nothing.)



KNOWLEDGE OPERATORS

1) If Bob's belief / knowledge has any physical meaning, there must be a **physical manifestation** *of it we can see.*

2) So Bob must be able to **physically answer** *questions about what he knows (either implicitly or explicitly).*



KNOWLEDGE OPERATORS

1) *Formally, what does it mean for Bob to "be able to physically answer questions about what he knows"?*

2) *To answer this, analyze physical phenomena where Bob knows an event.*

3) *These are phenomena where information* <u>outside</u> *Bob gets* <u>inside</u> *Bob.*

4) *Examples:*
    - *Observation*
    - *Prediction*
    - *Memory*
    - *Control*

ROADMAP

1) ***Shortcomings of current impossibility results concerning laws of Physics***

2) ***Knowledge operators and their shortcoming***

3) ***Formalize mathematical structure shared by observation and prediction:*** **inference devices**

4) ***Elementary properties of inference devices***

EXAMPLE OF PHYSICAL KNOWLDEGE: OBSERVATION

1)Present a stylized example of observation.

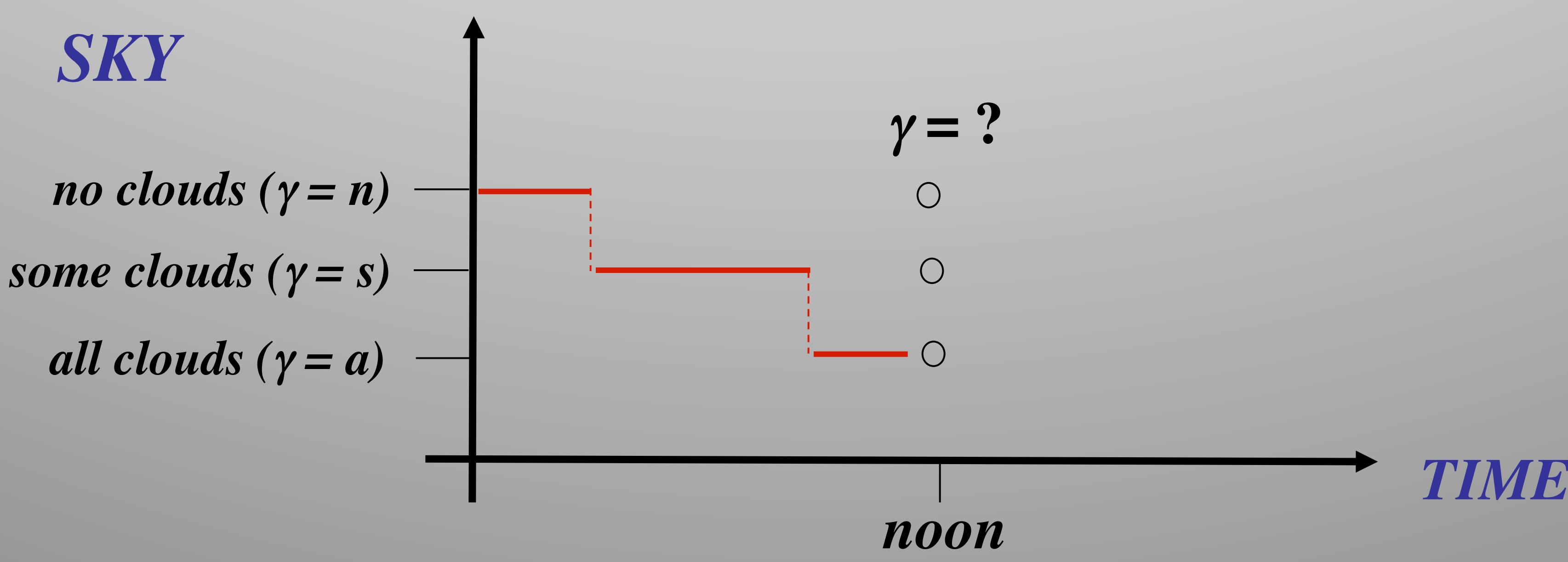
2)Emphasize features of that example found in all "observations"

3)Why those features are always found in observations:

Without those features, the observation conveys no semantic information

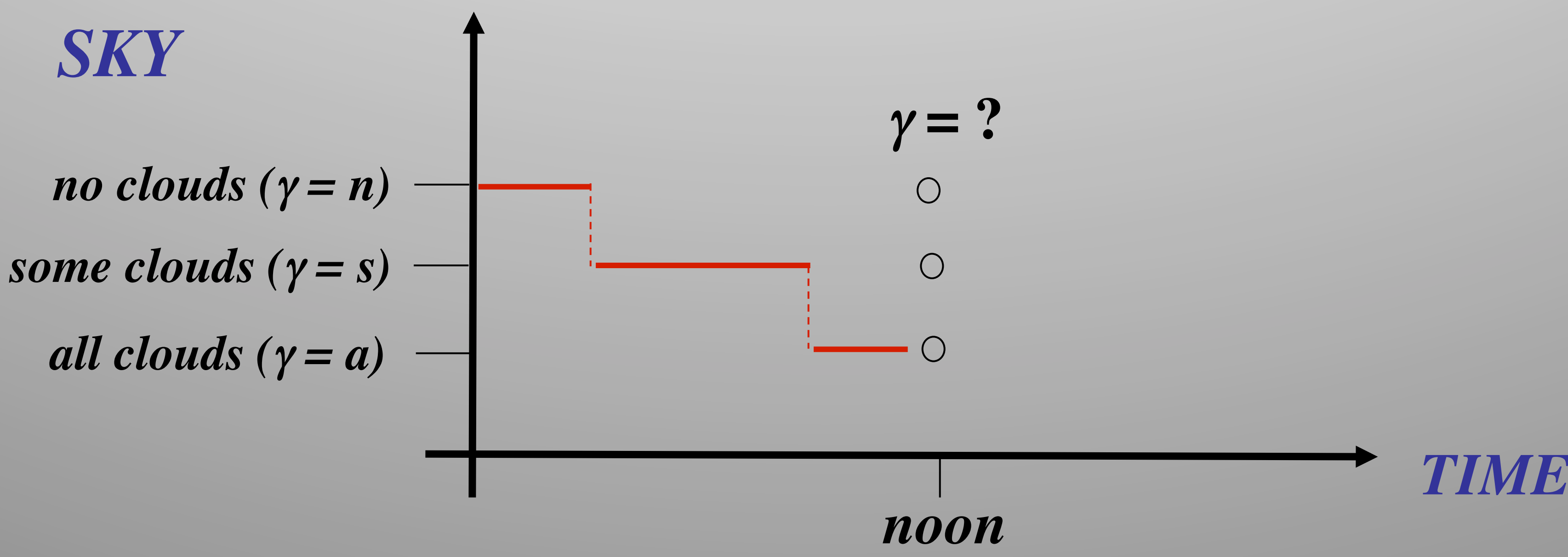
OBSERVATION

- ***Want to observe γ, state of sky at noon tomorrow***



OBSERVATION

- ***Bob claims to be able to make that observation***



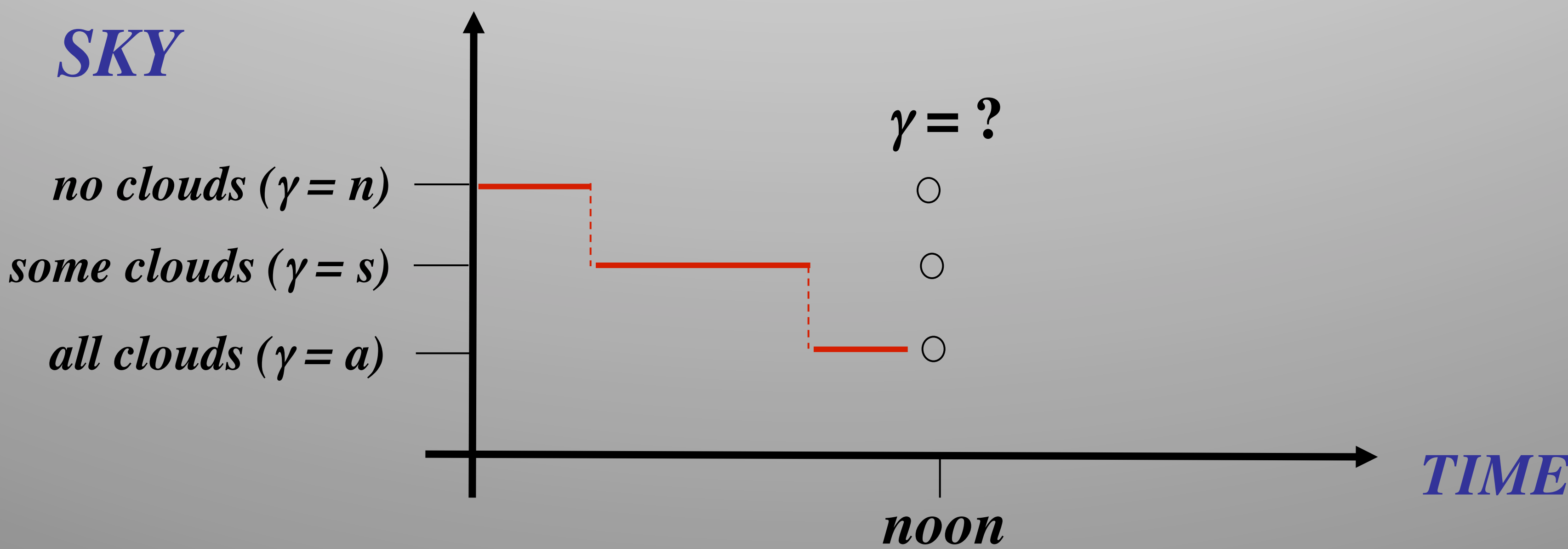
OBSERVATION

If Bob's claim is true, he will be able to correctly answer three questions that could be posed to him:

i) Does $\gamma =$ 'n'? (Yes / no)

ii) Does $\gamma =$'s'? (Yes / no)

iii) Does $\gamma =$ 'a'? (Yes / no)



OBSERVATION

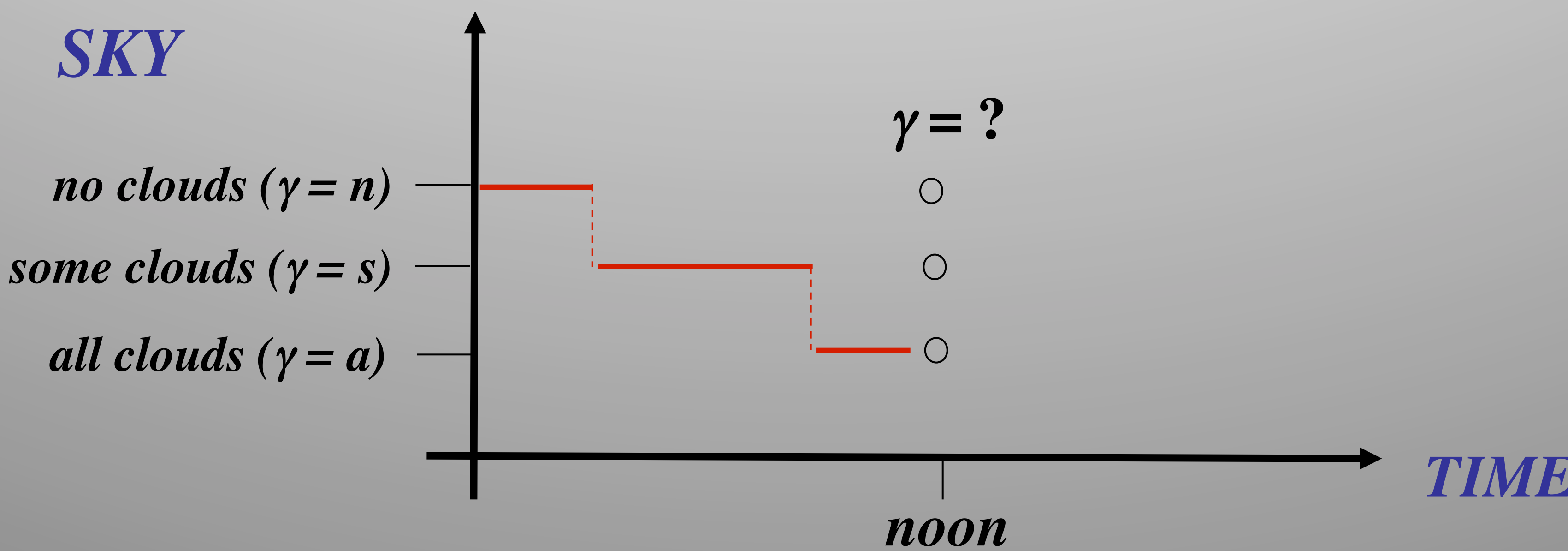
If Bob's claim is true, he will be able to correctly answer three questions that could be posed to him:

i) Does $\gamma =$ 'n'? (Yes / no)

ii) Does $\gamma =$'s'? (Yes / no)

iii) Does $\gamma =$ 'a'? (Yes / no)

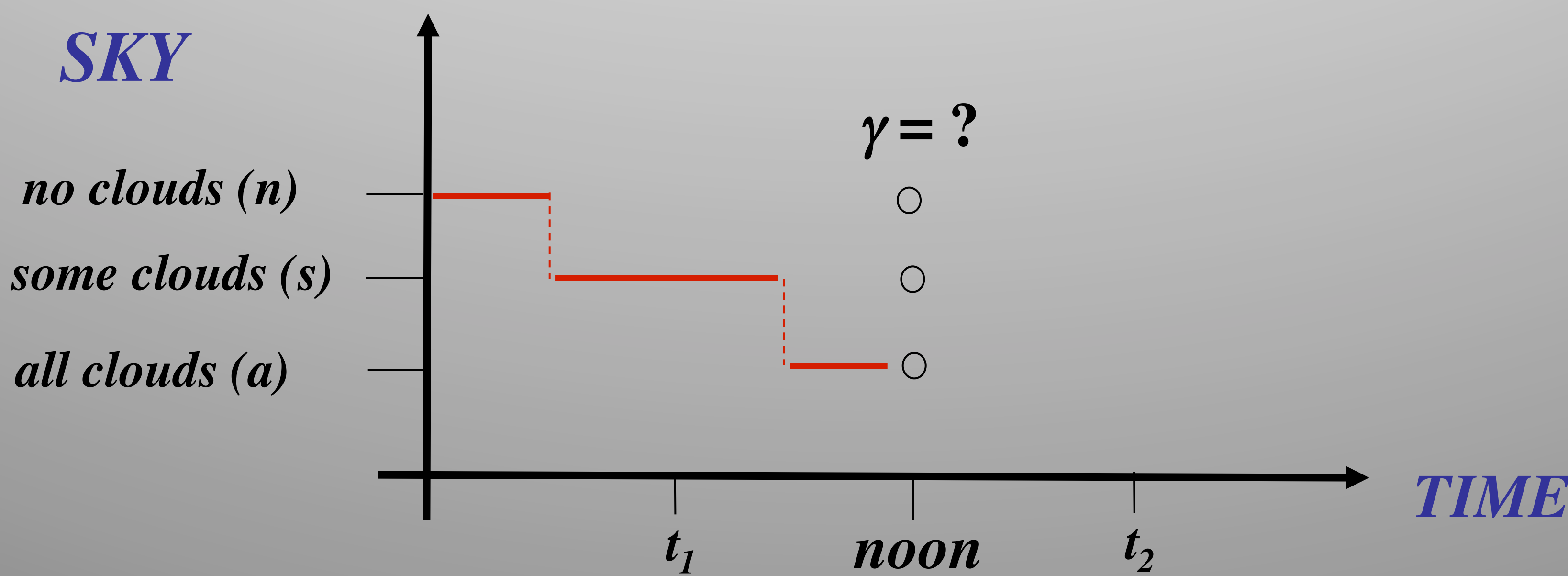
How formalize this?



OBSERVATION

- ***U = {all universe-histories consistent with physics, in which: Bob and the sky exist; at t_1 Bob considers a q; he observes γ; he gives honest answer to that q at t_2}***

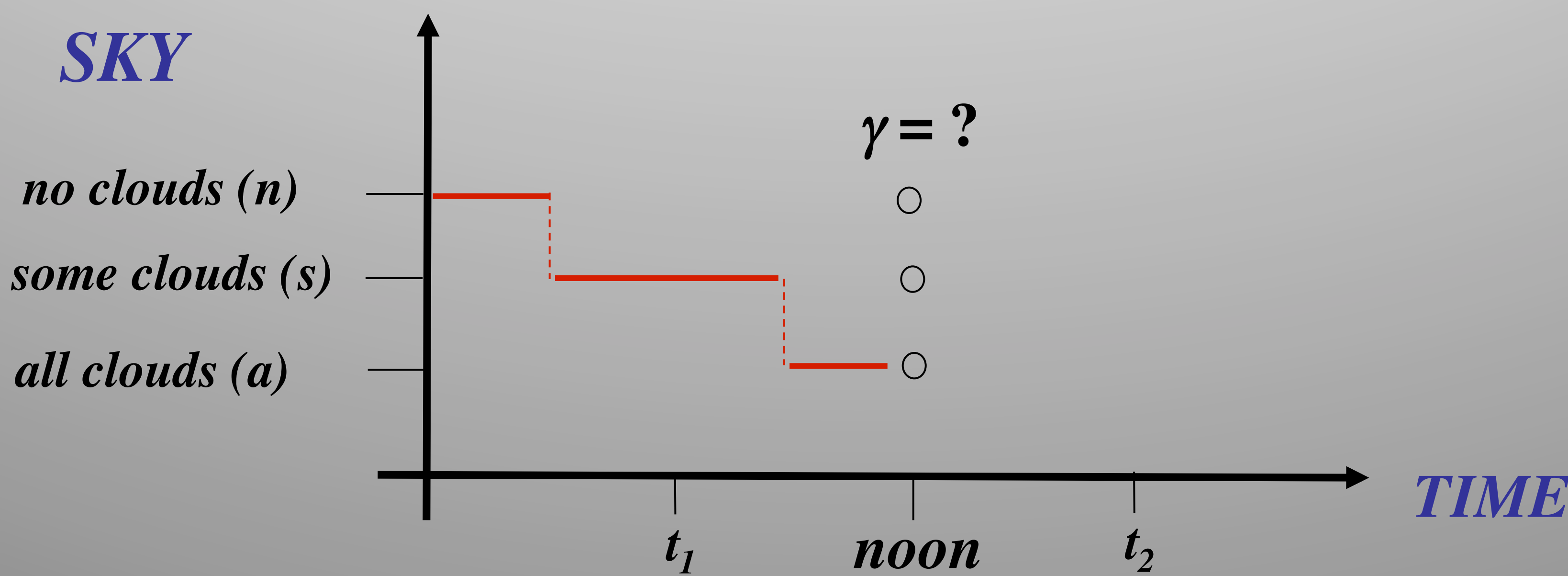
- ***State of sky at noon is fixed by $u \in U$, the actual universe-history. So $\gamma = \Gamma(u)$ for some function Γ***



OBSERVATION

- *U = {all universe-histories consistent with physics, in which: Bob and the sky exist; at t_1 Bob considers a q; he observes γ; he gives honest answer to that q at t_2}*

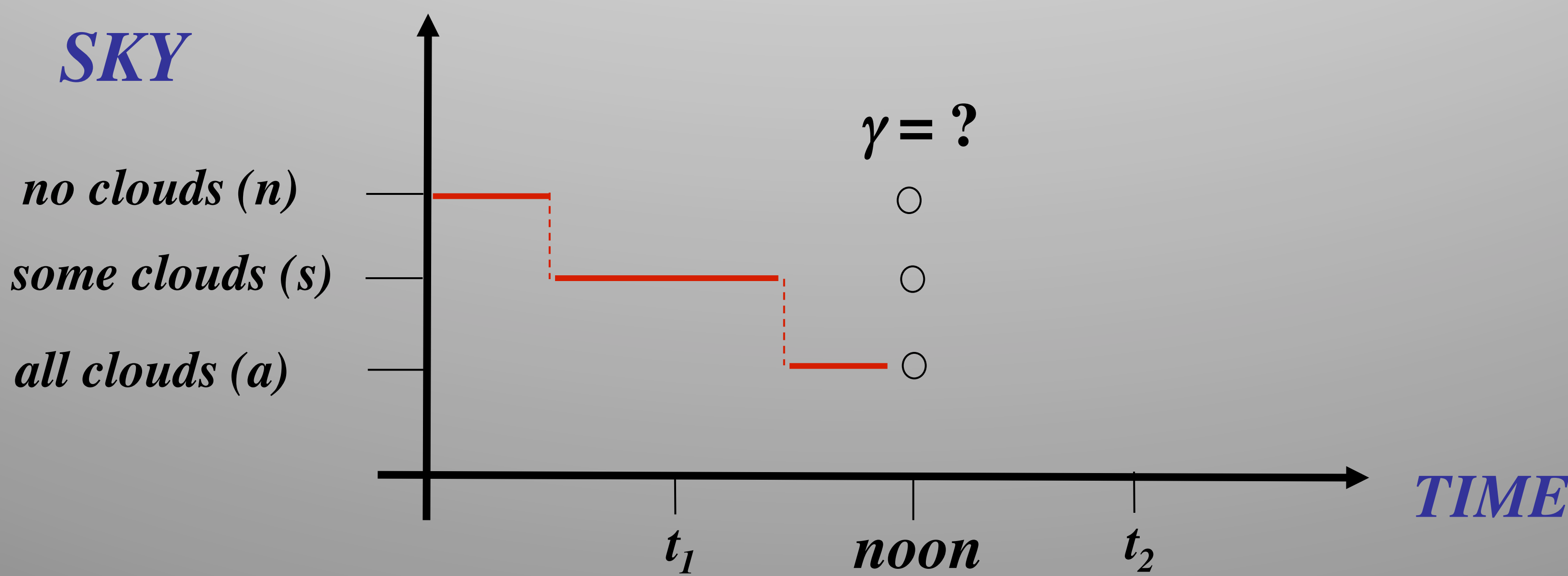
- *The question Bob considers at t_1 is set by actual universe-history $u \in U$: Bob considers $x = X(u)$ for some func. X*



OBSERVATION

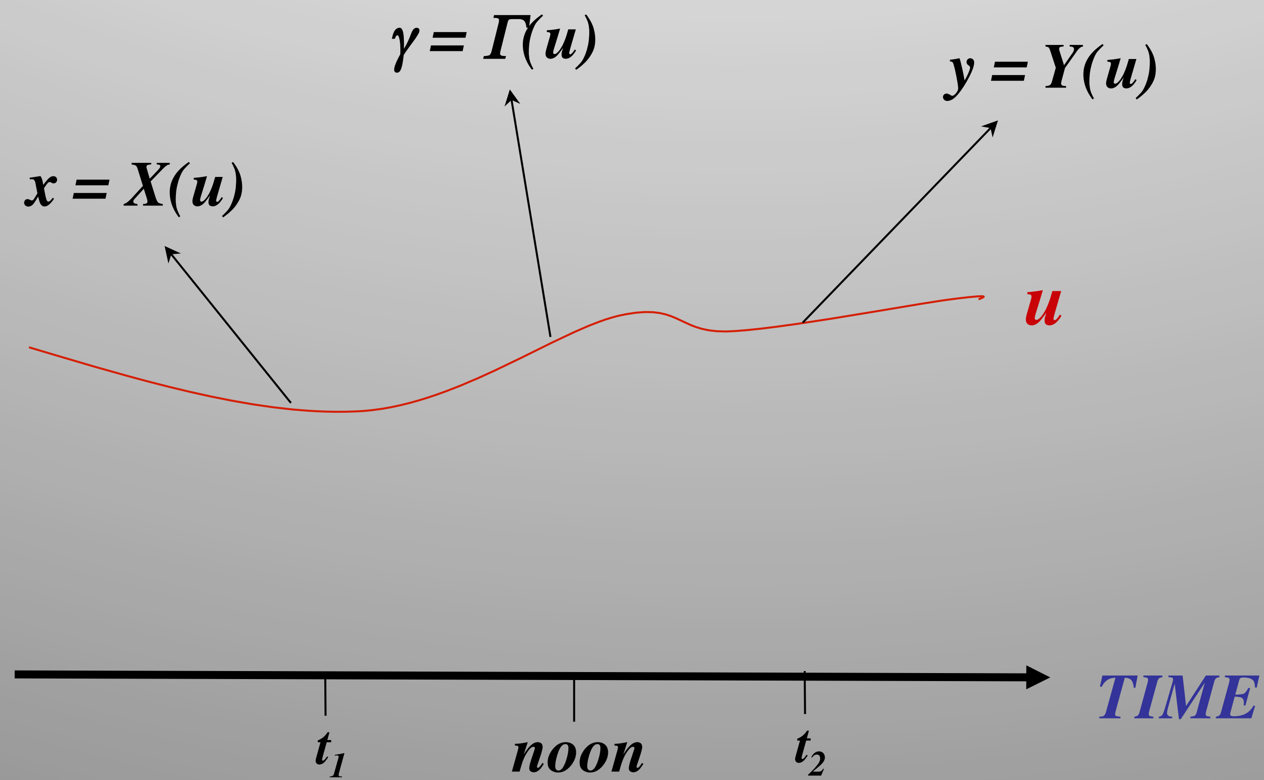
- ***U = {all universe-histories consistent with physics, in which: Bob and the sky exist; at t_1 Bob considers a q; he observes γ; he gives honest answer to that q at t_2}***

- ***Bob's answer at t_2 is given by actual universe-history $u \in U$: binary answer $y = Y(u)$ for some func. Y***



OBSERVATION

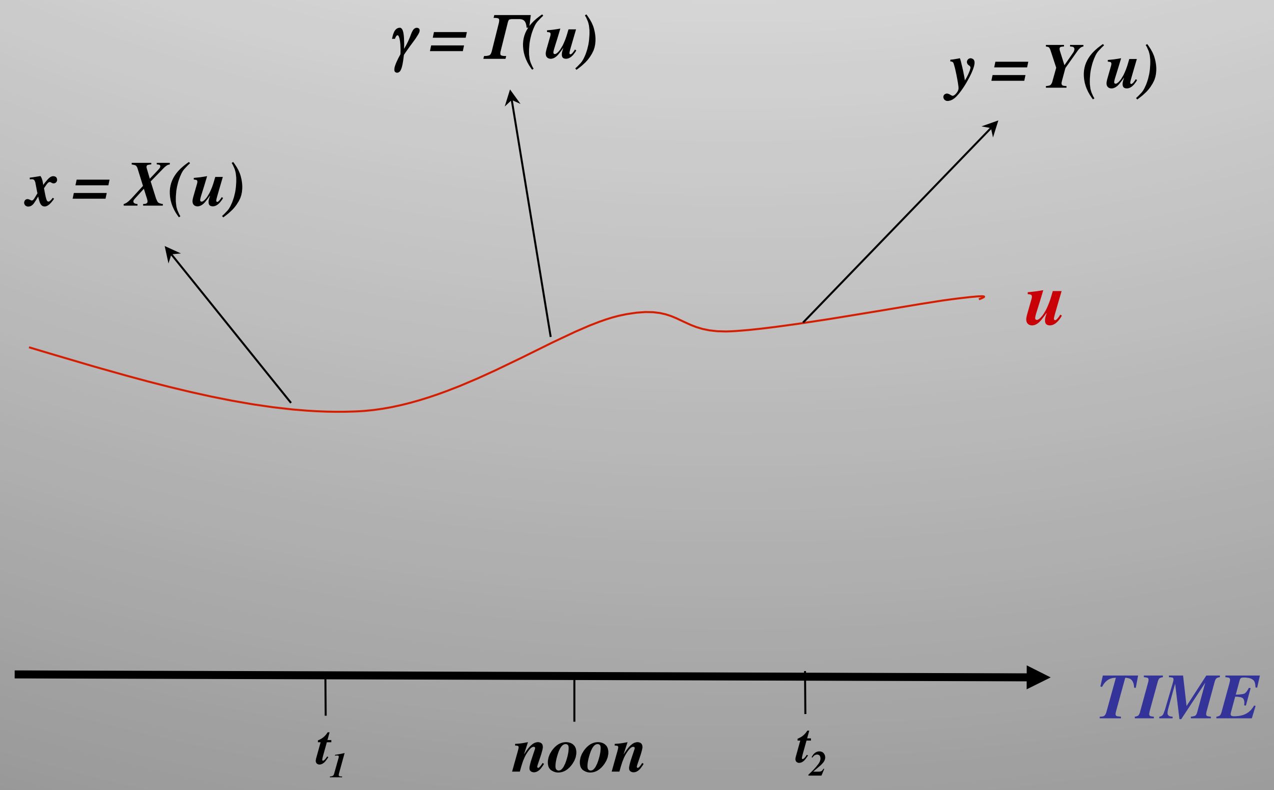
- $\gamma \in \{$ 'n', 's', 'a'$\}$ = sky at noon = $\Gamma(u)$
- x = (what q Bob considers at t_1) = $X(u)$
- y = (Bob's answer at t_2) = $Y(u)$



OBSERVATION

So:

- *For each of three binary questions* $q_\gamma : \Gamma(U) \rightarrow \mathrm{B}$
- $\exists\, x$ *such that*
- $X(u) = x \;\; \Rightarrow \;\; Y(u) = q_\gamma(\Gamma(u))$

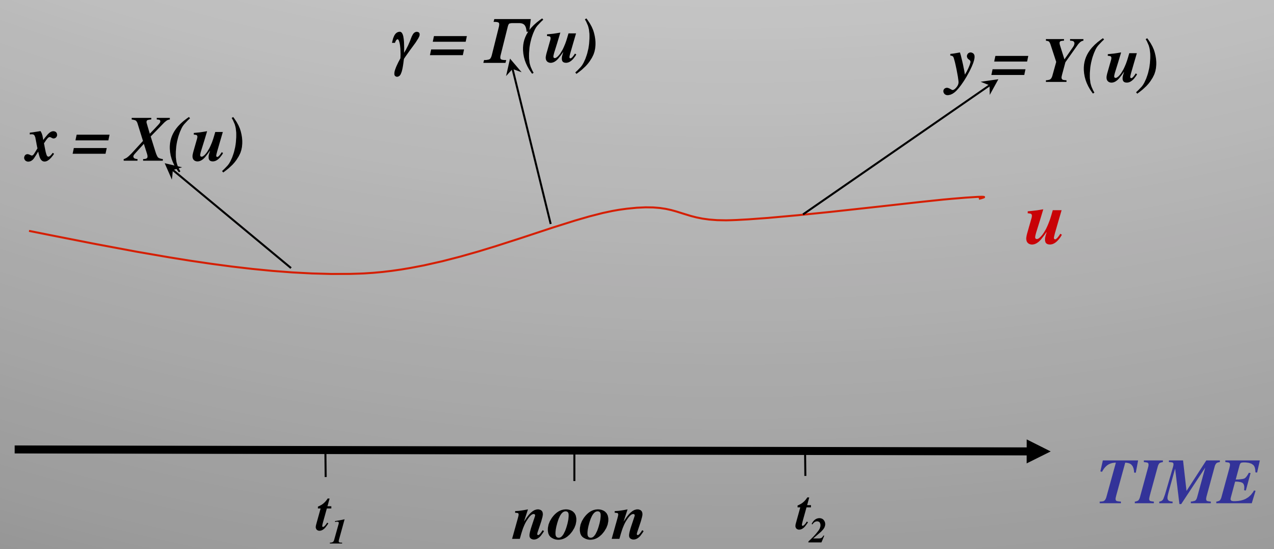


OBSERVATION

- *Nothing about observation* **process;** *all about what it means to* **successfully observe.**
- *The 'what' of observation, not the 'how'.*

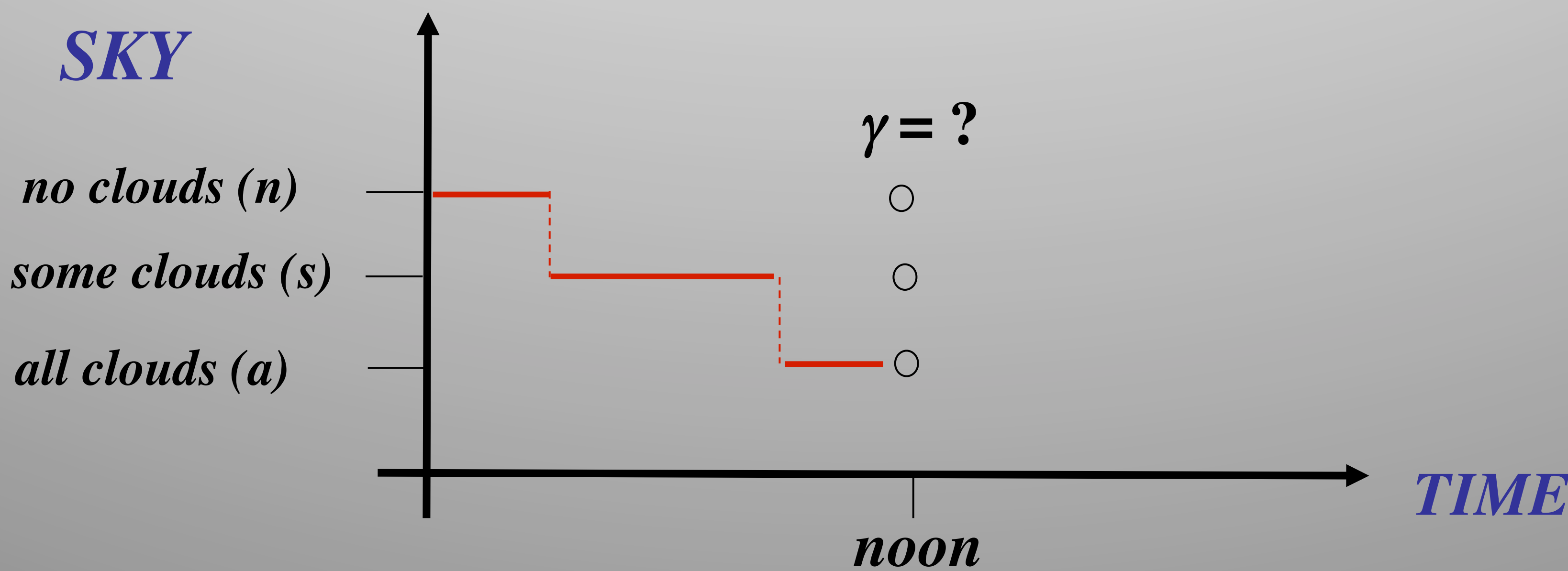
For each of three binary-valued questions q_γ,

$$\exists\, x \text{ such that } X(u) = x \quad \Rightarrow \quad Y(u) = q_\gamma(\Gamma(u))$$



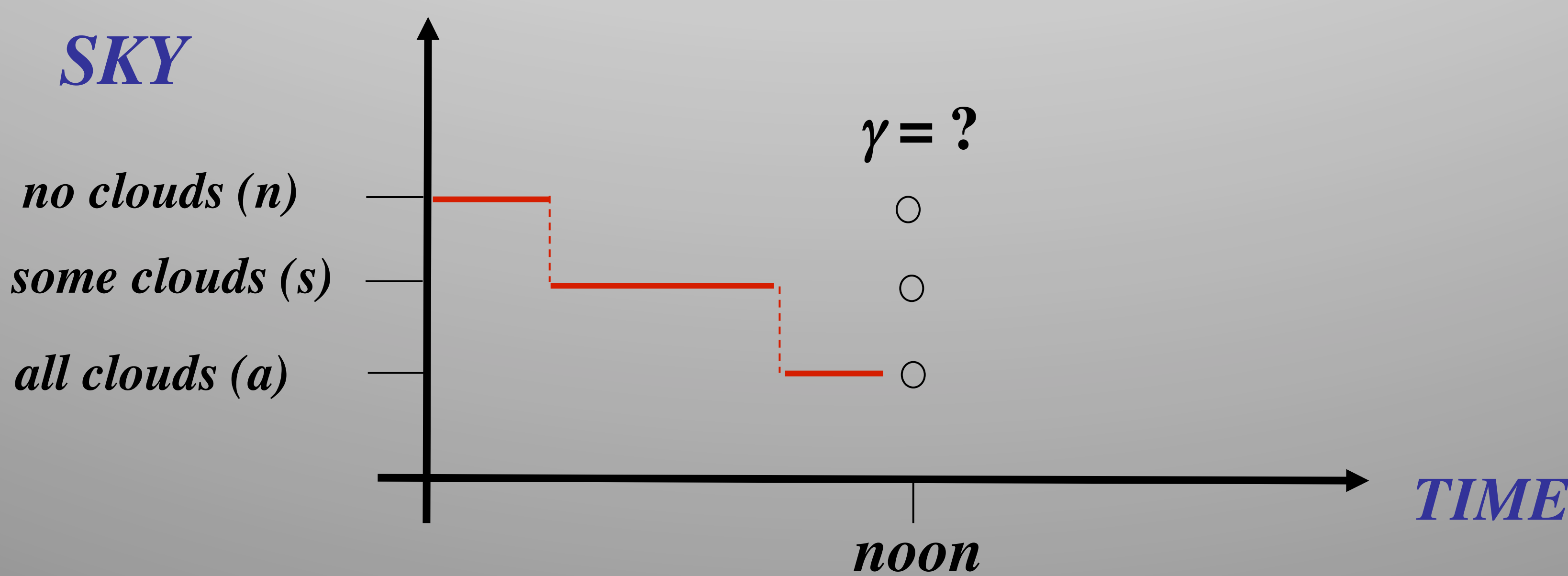
PREDICTION

- ***Want to predict γ, state of sky at noon tomorrow***



PREDICTION

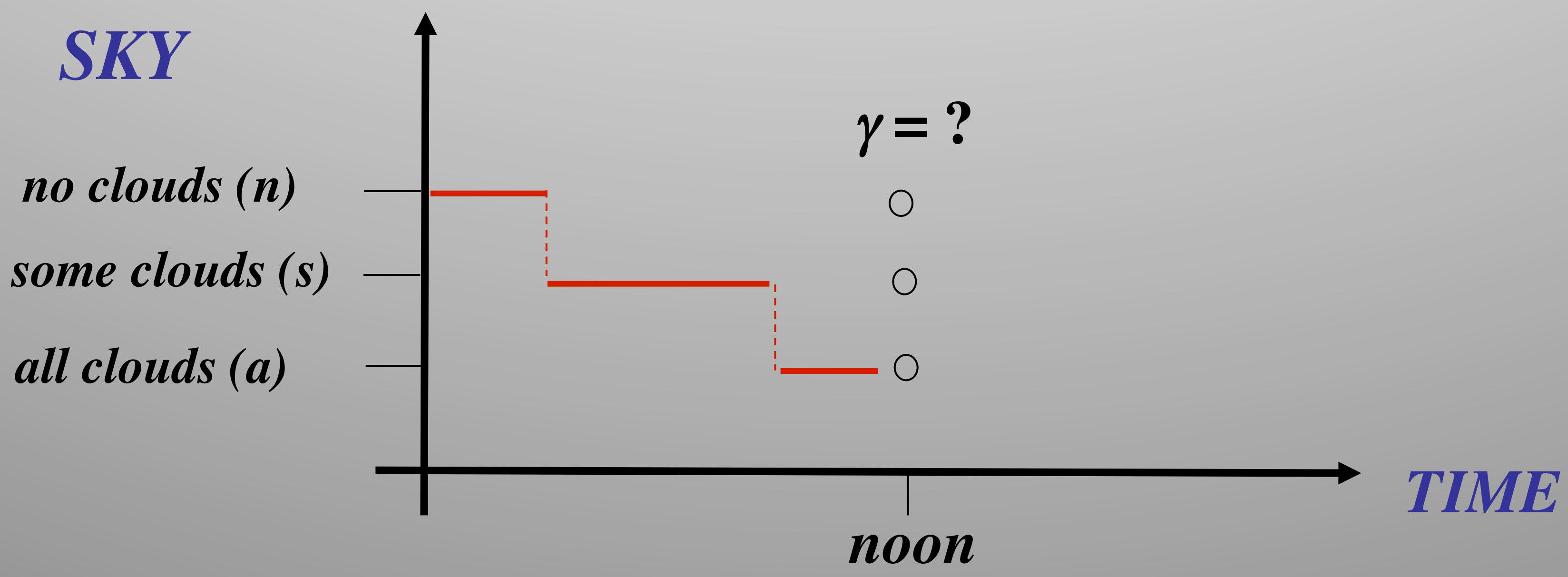
- ***Bob claims to have a laptop that he can program to make that prediction***



PREDICTION

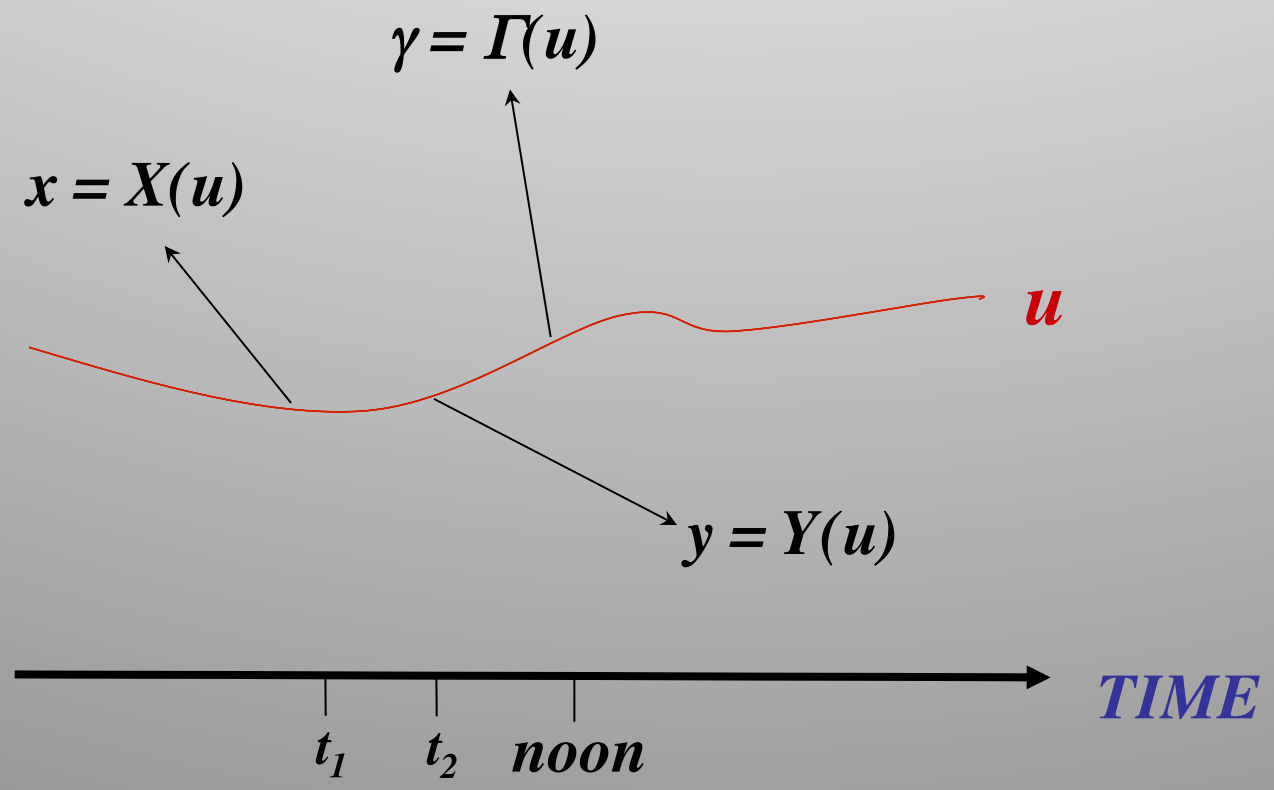
If Bob's claim is true, he will be able to correctly answer three questions:

- ***i) Does $\gamma =$ 'n'? (Yes / no)***
- ***ii) Does $\gamma =$'s'? (Yes / no)***
- ***iii) Does $\gamma =$ 'a'? (Yes / no)***



PREDICTION

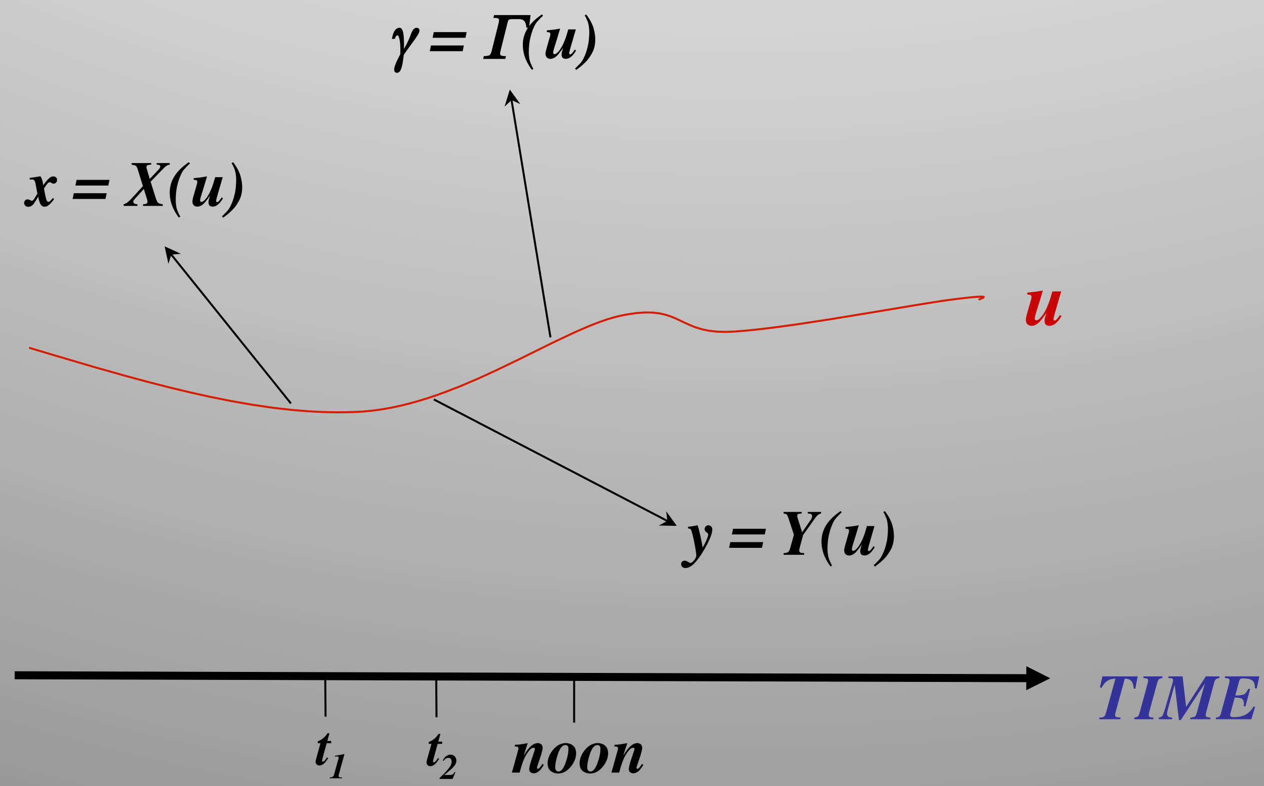
- $\gamma \in \{$ 'n', 's', 'a'$\}$ = sky at noon = $\Gamma(u)$
- x = (laptop program at t_1) = $X(u)$
- y = (Bob's answer at t_2) = $Y(u)$



PREDICTION

So:

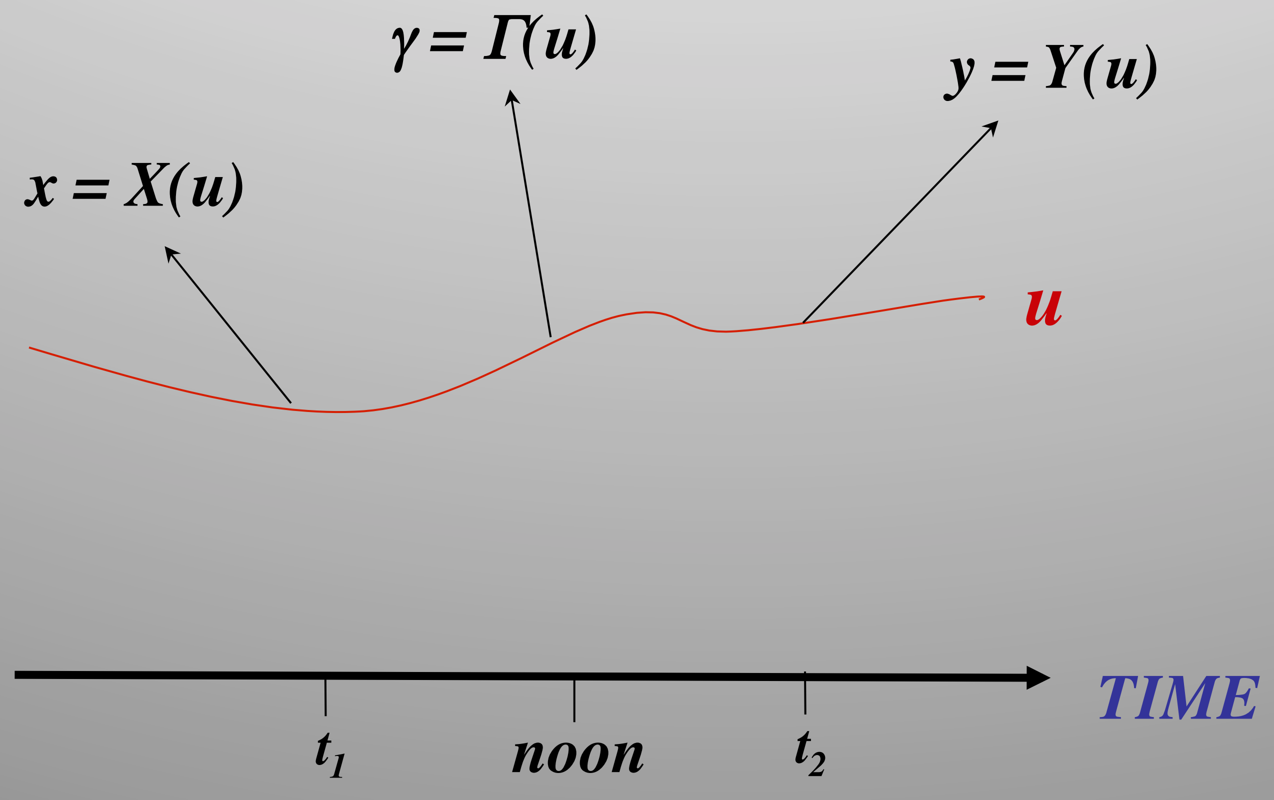
- *For each of three binary questions* q_γ,
- $\exists\, x$ *such that*
- $X(u) = x \;\;\Rightarrow\;\; Y(u) = q_\gamma(\Gamma(u))$



OBSERVATION

So:

- ***For each of three binary questions q_γ,***
- ***$\exists\, x$ such that***
- ***$X(u) = x \;\Rightarrow\; Y(u) = q_\gamma(\Gamma(u))$***



KNOWLEDGE

- *More generally, if at some time, “Bob knows the state of the sky at noon”, γ, then he can answer three questions:*
  - *i) Does $\gamma =$ ‘n’? (Yes / no)*
  - *ii) Does $\gamma =$ ‘s’? (Yes / no)*
  - *iii) Does $\gamma =$ ‘a’? (Yes / no)*

- *Note no chronological ordering. Just three functions:*
  X (what question Bob considers),
  Y (his answer),
  Γ (the sky’s actual state at noon),
  - *all three are functions of $u \in U$*

INFERENCE DEVICES

- *An* <u>inference device</u> *is any two functions (X, Y) over U, where range of Y is binary.*

An inference device (X, Y) (<u>weakly) infers</u> *a function Γ over U iff*

$\forall \gamma$ *in* Γ*'s range,*

$$\exists x \text{ such that } \quad X(u) = x \quad \Rightarrow \quad Y(u) = q_\gamma(\Gamma(u))$$

- *A necessary condition to say that (X, Y) "observes", "predicts", or "knows" Γ is that (X, Y) weakly infers Γ.*

- *No claims of sufficiency; observation, prediction, knowledge, etc. involve much more than just weak inference.*

- *But even requiring weak inference restricts observation, prediction, and knowledge.*

INFERENCE DEVICES

- *An* **inference device** *is any two functions (X, Y) over U, where range of Y is binary.*

*An inference device (X, Y) (***weakly) infers*** *a function Γ over U iff*

$$\forall \gamma \text{ in } \Gamma\text{'s range},$$

$$\exists\, x \text{ such that } \quad X(u) = x \;\Rightarrow\; Y(u) = q_\gamma(\Gamma(u))$$

- *N.b., use of counterfactual questions (possibility of Bob asking if $\Gamma(u)$ has a value that it does not have). Similar to using intervention to define causality in Bayes nets.*

- *Contrast inference with Aumann-style "knowledge operators"*

INFERENCE DEVICES TERMINOLOGY

1) **Setup** ***function X over U***

2) **Conclusion** ***binary-valued function Y over U***

3) ***"$(X, Y) > \Gamma$" means (X, Y) weakly infers Γ***

INFERENCE DEVICES AND THE LAWS OF PHYSICS

1) A reality is a space U, a set of devices defined over U, and a set of functions the devices might infer.

2) So a reality is a triple, $(U, \{X_j, Y_j\}, \{\Gamma_i\})$.

3) As far as any device in a reality is concerned, *U is irrelevant*. It's only the inference graph relating the sets $\{X_j, Y_j\}$ and $\{\Gamma_i\}$ that matter:

The laws of Physics are patterns in the inference graph of a reality

ROADMAP

1) ***Shortcomings of current impossibility results concerning laws of Physics***

↓

2) ***Knowledge operators and their shortcoming***

↓

3) ***Formalize mathematical structure shared by observation and prediction:*** **inference devices**

↓

4) ***Elementary properties of inference devices***

ELEMENTARY PROPERTIES OF INFERENCE

1) **Inference need not be transitive:**

   $(X_1, Y_1) > Y_2$ and $(X_2, Y_2) > Y_3$ does not mean $(X_1, Y_1) > Y_3$

2) **For any Γ, ∃ a device that infers Γ.**

3) **For any device, ∃ a Γ it does not infer. (*Impossibility result*)**
   - ***Intuition*: X ~ initial configuration of a Turing machine.**

     Y (a bit) ~ whether Turing machine halts or not.

     So apply Halting theorem-style reasoning

IMPLICATIONS OF IMPOSSIBILITY RESULT

1) **For any simulator, there is always a prediction that cannot be guaranteed correct.**
    - ***Laplace was wrong.***
    - ***Impossibility results of Pour-El et al., Fredkin et al., Moore, etc. are far narrower than this result***

2) **For any observation apparatus, there is always an observation that cannot be guaranteed to be correct.**
    - ***Non-quantum mechanical "uncertainty principle"***

BREADTH OF IMPOSSIBILITY RESULT

1) **Holds even for a countable U (even for a finite one).**

2) **Holds even if current formulation of physics is wrong.**

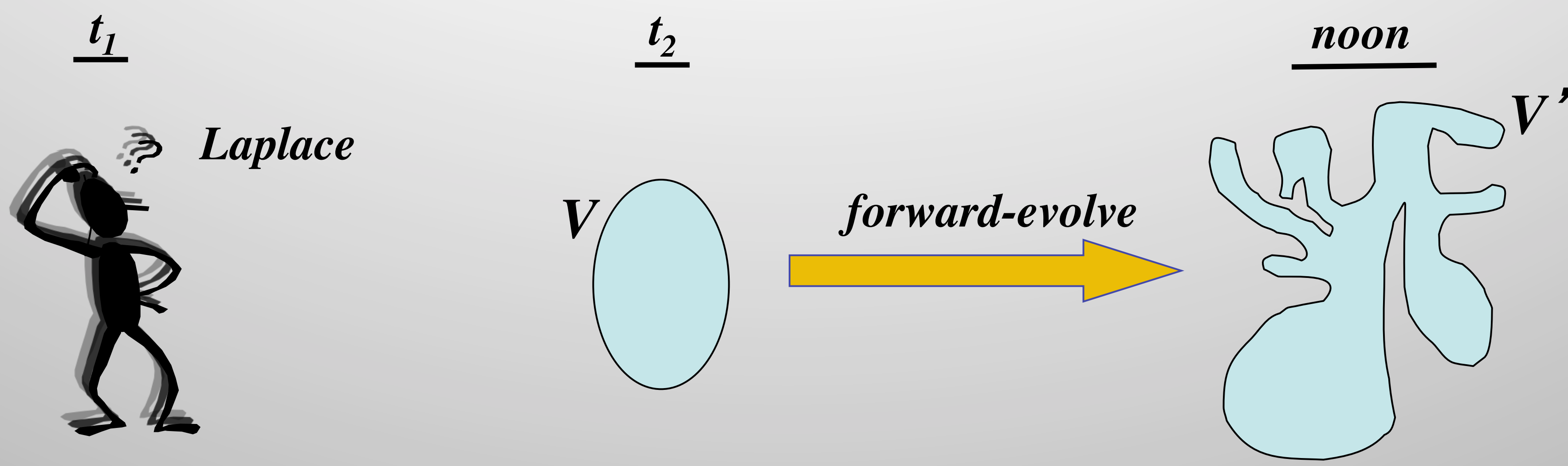
3) **Holds even if the device has Super-Turing capability**

4) **Holds even if laws of Physics are not written in predicate logic,**

   or intuitionism is correct,

   or even if there are no laws, just a huge list of events.

EXAMPLE: PREDICTION FAILURE



1. *V = {all time-t_2 universes where Laplace is answering "yes" to his t_1 question}*
2. *V' = V evolved forward to noon*
3. *At t_1, ask Laplace, "will universe be outside V' at noon?"*

Trivially, Laplace's answer is wrong

INFERENCE RELATIONS BETWEEN DEVICES

- **Often not interested in inference of arbitrary functions, but rather inference relation among a pre-fixed set of devices.**

I) *Two* devices (X, Y), (X', Y') are *pairwise distinguishable* iff every pair (x, x') occurs for some u

II) *A set* of devices $\{(X_i, Y_i)\}$ is *mutually distinguishable* iff every tuple $(x_1, x_2, \ldots)$ occurs for some u

- **Distinguishability can be seen as a formalization of "free will". (Compare to Conway's "free will" theorems.)**

INFERENCE RELATIONS BETWEEN DEVICES -2

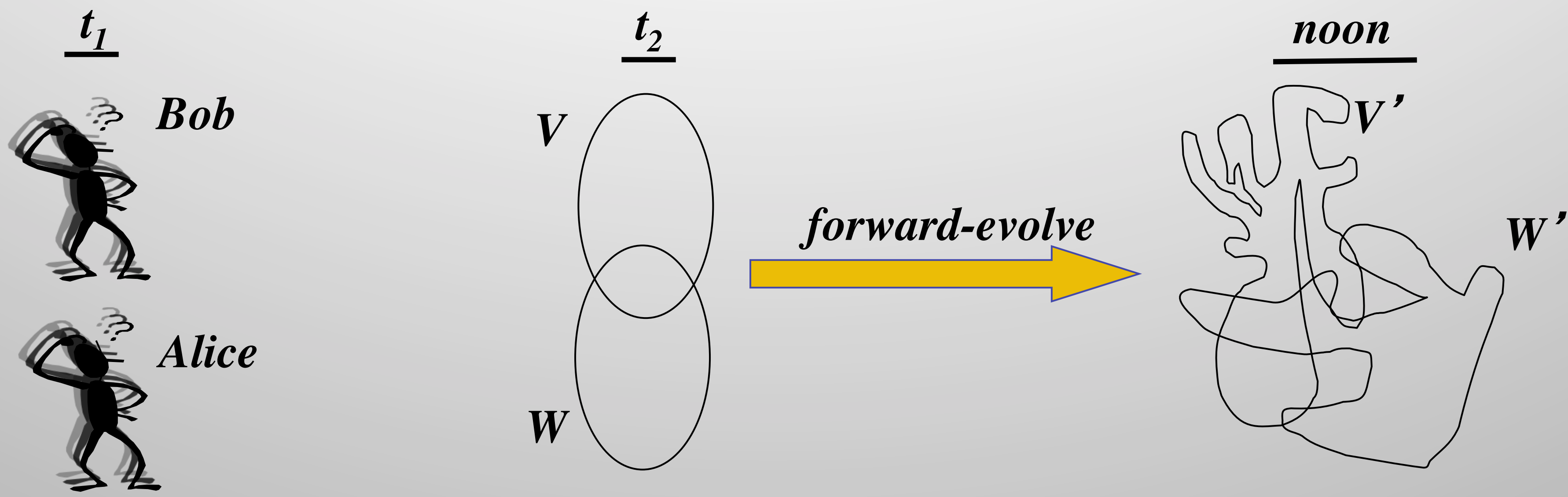
3) **If all pairs of devices from $\{C_i\}$ are pairwise distinguishable, $\exists$ at most one $k : C_k > C_j \; \forall j \neq k$. "*Monotheism*" theorem.**

- **N.b., control is a special type of inference.**

4) **If all pairs of devices from $\{C_i\}$ are pairwise distinguishable, can have $C_1 > C_2 > \ldots > C_1$.**

5) **If the set of devices $\{C_i\}$ is mutually distinguishable, can*<u>not</u>* have $C_1 > C_2 > \ldots > C_1$.**

MONOTHEISM EXAMPLE



- *V = {time-t_2 universes where Bob is answering 'yes' to his t_1 question}*
- *W = {time-t_2 universes where Alice is answering 'yes' to her t_1 question}*
- *V' = V evolved forward to noon*
- *W' = W evolved forward to noon*
- *At t_1, ask Bob, "will universe be in W' at noon?"*
- *At t_1, ask Alice, "will universe be outside of V' at noon?"*

Either Bob or Alice is wrong

INFERENCE KNOWLEDGE AND BOOLEAN ALGEBRA

Knowledge defined in terms of weak inference obeys many of the properties of Boolean algebra:

1) **(X, Y) may know A, or may know ~A, but not both.**

2) **If (X, Y) knows A $\Rightarrow$ B and (X, Y) knows B $\Rightarrow$ C, then (X, Y) knows A $\Rightarrow$ C.**

3) **If (X, Y) knows A, then (X, Y) knows event "(X, Y) knows A".**

4) **If (X, Y) knows event A, and knows event A $\Rightarrow$ B, then B is true.**
    - **However no implication that (X, Y) knows B; no problem of knowing all truths via deduction.**

STOCHASTIC INFERENCE

- **What changes if there is probability measure P over U?**

1) Given a function Γ and device C = (X, Y), C infers Γ with <u>*covariance accuracy*</u>

$$\varepsilon(C,\Gamma) = \frac{\sum q_L \max_x \left[E_P(Y q_L(\Gamma) \mid x) \right]}{|\Gamma(U)|}$$

2) Can't instead use mutual information; that only captures *syntactic* content of distributions, not *semantic* content.

EXAMPLE OF STOCHASTIC INFERENCE RESULT

1) For any probability distribution P over U,

$$\varepsilon((X,Y),\Gamma) \geq (2-n)\,\frac{\max_x\left[E_P(Y \mid x)\right]}{n}$$

where *n* = |Γ(U)|

2) For any probability distribution P over U, there exists two devices (X_1, Y_1), (X_2, Y_2) where X_1 and X_2 are distinguishable, but both $\varepsilon((X_1, Y_1), Y_2)$ and $\varepsilon((X_1, Y_1), Y_2)$ are arbitrarily close to one;

Second Laplace impossibility theorem is "barely true"

HOWEVER RELATED RESULTS ARE QUITE STRONG

1) Let C_1 and C_2 be two devices, where:
   i) Both $X_1(U)$ and $X_2(U)$ are the binaries;
   ii) $C_1 > C_2$ with accuracy ε_1, and $C_2 > C_1$ with accuracy ε_2.
   iii) $P(X_1 = -1) = \alpha$, and $P(X_2 = -1) = \beta$

2) Define H as the four-dimensional unit open hypercube, and
   i) $\forall z \in H$, $k(z) = z_1 + z_4 - z_2 - z_3$;
   ii) $\forall z \in H$, $m(z) = z_2 - z_4$;
   iii) $\forall z \in H$, $n(z) = z_3 - z_4$.

3) $\varepsilon_1\varepsilon_2 \leq \max_{z\in H} |\alpha\,\beta[k(z)]^2 + ak(z)m(z) + bk(z)n(z) + m(z)n(z)|$

4) E.g., for $\alpha = \beta = 1/2$, $\varepsilon_1\varepsilon_2 \leq 1/4$.

STRONG INFERENCE

- A universal Turing Machine T can emulate any other one, T'
- T does that by having its input be the program and input of T'

The analog with inference devices:

$C = (X, Y)$ ***strongly infers*** $C' = (X', Y')$ **iff:**

$$\forall \text{ questions } q_L \text{ of } Y'(U),\ \forall\, x',$$

$$\exists\, x \text{ s.t.}$$

$$X(u) = x \;\Rightarrow\; Y(u) = q_L(Y'[u]),\ X'(u) = x'$$

- "$C_1 >> C_2$" means C_1 strongly infers C_2

PROPERTIES OF STRONG INFERENCE

1) $C_1 >> C_2$ and $C_2 > \Gamma \Rightarrow C_1 > \Gamma$
   - Just like with UTM's and TM's (contrast weak inference)

2) $C_1 >> C_2$ and $C_2 >> C_3 \Rightarrow C_1 >> C_3$
   - Just like with UTM's (contrast weak inference)

3) For any C_1, $\exists$ C_2 that C_1 does not strongly infer

4) If $\forall x_1$, $|X_1^{-1}(x_1)| > 2$, then $\exists$ C_2 such that $C_2 >> C_1$

5) No two devices can strongly infer each other
   - Distinguishability irrelevant (contrast weak inference)
   - Holds even if C_1 and C_2 are same system just at different moments in time; "intelligent design theorem".

CONCLUSIONS

1) Previous work on the relation between computational impossibility and physics made strong assumptions about the computational model of the universe.

2) One can instead start at a more fundamental level, with a model of what it means to know a fact about the universe in which you are embedded.

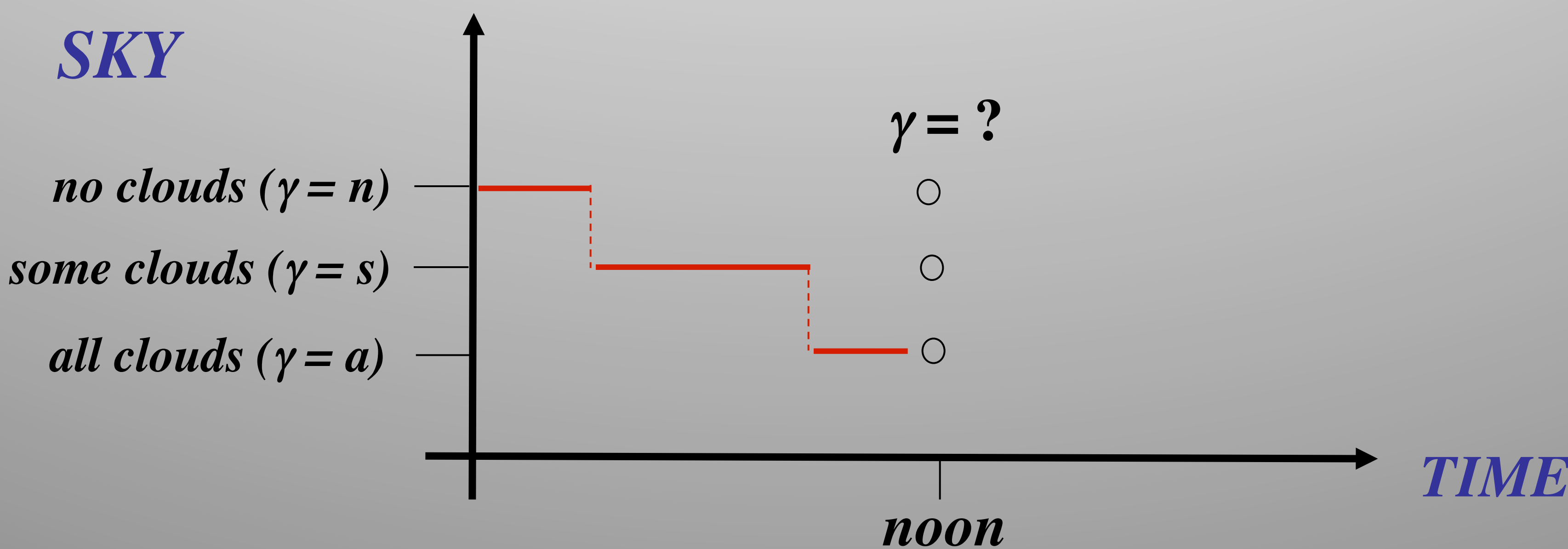
3) This model shares many attributes with computational models.

4) This model has additional impossibility results, some showing that Laplace was wrong, and some that are reminiscent of quantum mechanics.

OBSERVATION

What does this mean physically?

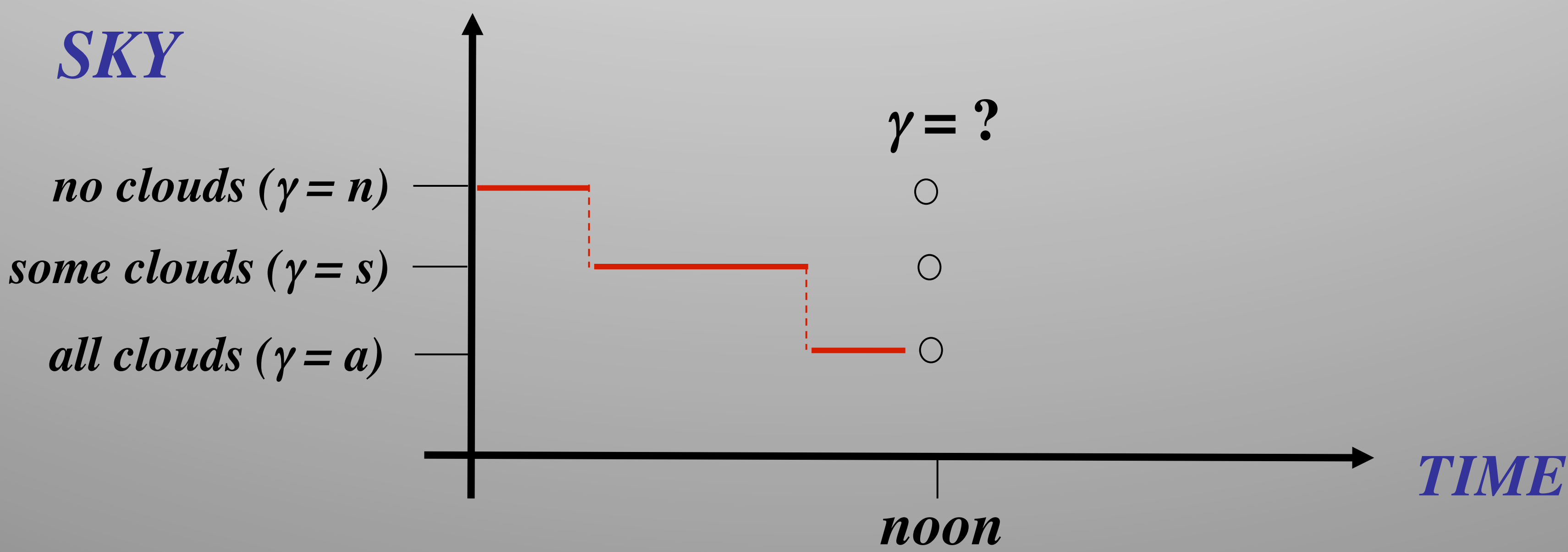
- ***Restrict attention to universes where Bob and the sky exist; Bob considers one of the three binary questions; observes γ; then gives his honest answer to that question.***



OBSERVATION

- ***Bob considers one of the three binary questions; observes γ; then gives his honest answer to that question.***

- ***Crucial point: what question Bob considers, the value γ, and what answer Bob gives, are all properties of the universe.***



OBSERVATION

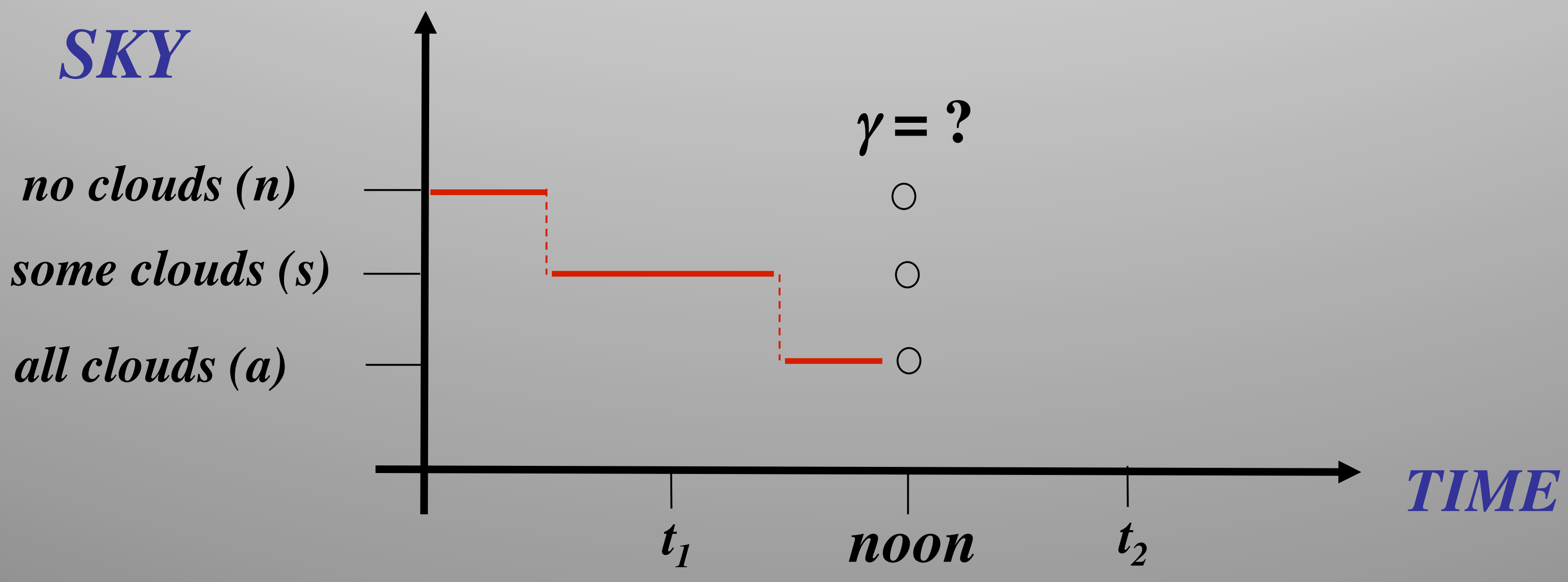
So Bob observes γ if for each of the three questions, q,

The universe having property x_q:

"At some t_1 Bob considers q"

$$\Longrightarrow$$

y, the binary answer Bob gives at some t_2 > noon, equals correct answer to q



PREDICTION

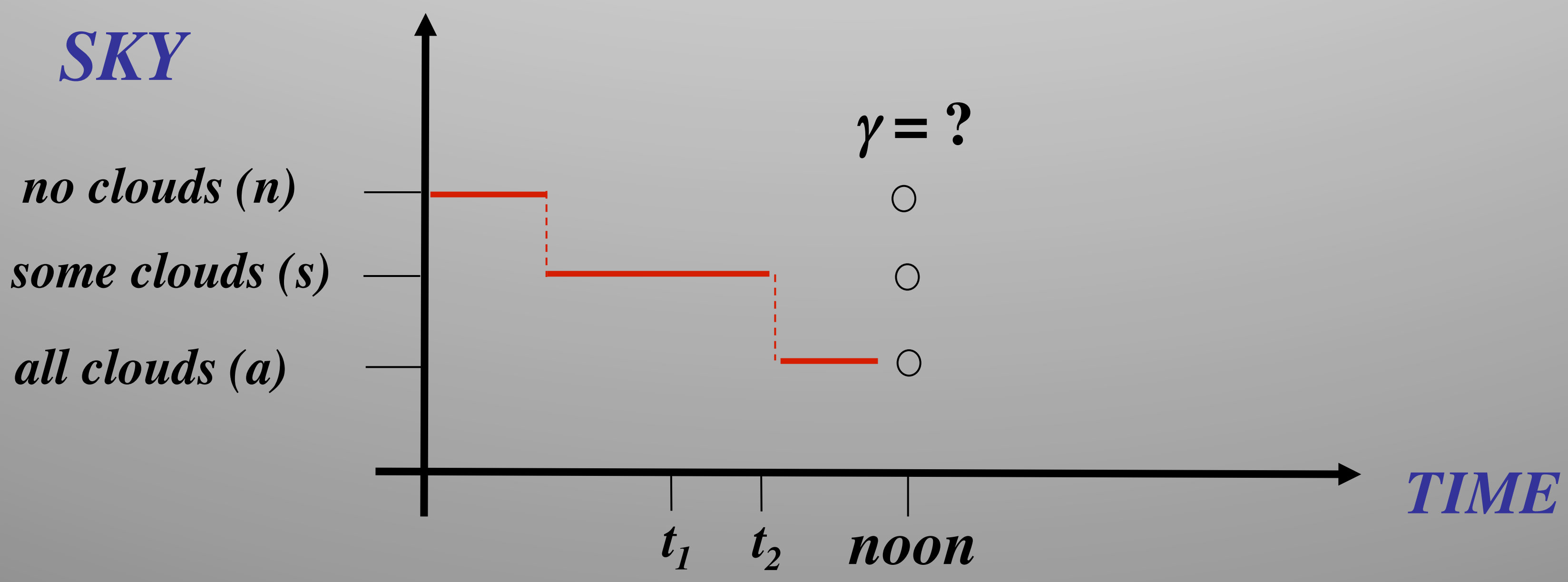
Bob can predict γ if for each of the three questions, q,

The universe has property x_q:

"At some t_1 < noon Bob programs the laptop to predict q"

$$\Longrightarrow$$

y, the binary answer Bob reads off at some t_2 < noon, equals correct answer to q



INFERENCE DEVICES

- ***Advantages of using binary questions:***

  i) Formalism doesn't change if range of Γ changes

  ii) Device never need give value $\Gamma(u)$, only confirm/reject suggested $\Gamma(u)$'s. (Cf. computational complexity)

  iii) Formalizes semantic information (contrast Shannon)