Theories of Something, Everything & Nothing.

Marcus Hutter
Canberra, ACT, 0200, Australia
http://www.hutter1.net/
THE AUSTRALIAN NATIONAL UNIVERSITY

Contents

Abstract

The progression of theories suggested for our world, from ego- to geo- to helio-centric models to universe and multiverse theories and beyond, shows one tendency: The size of the described worlds increases, with humans being expelled from their center to ever more remote and random locations. If pushed too far, a potential theory of everything (ToE) is actually more a theory of nothing (ToN). Indeed such theories have already been developed. I show that including observer localization into such theories is necessary and sufficient to avoid this problem. Ockham's razor is used to develop a quantitative recipe to identify ToEs and distinguish them from ToNs and theories in-between. This precisely shows what the problem is with some recently suggested universal ToEs. The suggested principle is extended to more practical (partial, approximate, probabilistic, parametric) world models (rather than ToEs). Finally, I provide a justification of Ockham's razor.

Topic of This Talk

- information-theoretic and computational approach for addressing the philosophical problem of judging theories (of everything) in physics.
- Slides (over)simplify & focus on the core problem and solution idea.
- Classical models in physics are essentially differential equations describing the time-evolution of some aspects of the world.
- A Theory of Everything (ToE) models the whole universe or multiverse, which should include initial conditions.
- I will argue, it can be crucial to also localize the observer, i.e. to augment the ToE with a model of the properties of the observer, even for non-quantum-mechanical phenomena.
- I call a ToE with observer localization, a Complete ToE (CToE).

THEORIES OF SOMETHING, EVERYTHING & NOTHING

- A number of models have been suggested for our world.
- They range from generally accepted to increasingly speculative to apparently bogus.
- For the purpose of this work it doesn't matter where you personally draw the line.
- The following (in)sane models will help to make clear the necessity of observer localization.

What is a Theory or Model

- By theory I mean a model which can explain $\approx$ describe $\approx$ predict $\approx$ compress our observations.

- deterministic theory/model + initial conditions = compact representation of observation sequence = bit string.

- Example: Newton mechanics maps initial planet positions+velocities into a time-series of planet positions.

- Stochastic model = probability distribution over observation strings.

Egocentric Model

A young child believes it is the center of the world.

+ Localization is trivial. It's always at coordinate (0,0,0).

– Cannot explain similarity of self and other humans.

Geocentric Model

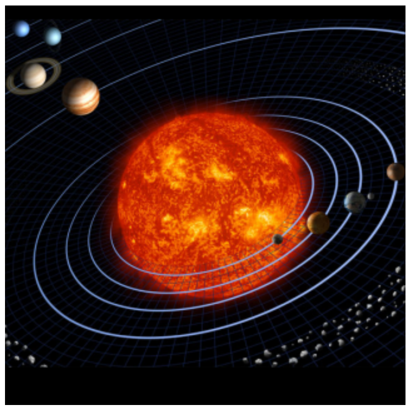
Human race and Earth is at the center of the universe.

+ Leads to understanding & well-functioning society.

– Why am I this particular person and not any other?

– Complex epicycle model for planets.

Heliocentric Model



Sun is at the center of the solar system / universe.

+ Simpler and better model of celestial motions.

– Why are we on planet 3 ?

Our Observable Universe

described by standard model + general relativity

+ Describes all known phenomena in our universe.
– Does not explain why are we in this solar system in this galaxy.

Large Universes & Multiverse Theories

Many theories (can be argued to) imply a multitude of essentially disconnected universes (in the conventional sense), often each with their own (quite different) characteristics.

- String theory: Different compactifications lead to different universes.
- Inflation: Universe much larger than visible part. Regions differ. Like the infinite fantasia land from the NeverEnding Story, where everything happens somewhere.
- Oscillating universe (Wheeler): a new big bang follows the assumed big crunch, and this repeats indefinitely.
- Baby universes (Smolin) Every black hole recursively produces new universes on the “other side” with quite different properties.
- Quantum universes (Everett): many-worlds interpretation of quantum mechanics postulates that the wave function doesn't collapse but the universe splits (decoheres) into different branches, one for each possible outcome of a measurement.

The Universal Universe

General recipe: If theory X contains some unexplained elements Y (quantum or compactification or parameter or other indeterminism), postulate that every realization of Y results in its own universe, and we just happen to live in one of them.

Often the anthropic principle is used in some hand-waving way to argue why we are in this and not that universe.

Take this to the extreme (Schmidhuber, Tegmark):
Universal Universe consists of every computable/mathematical universe.

Since our universe seems computable/mathematical, then it is contained in the universal universe, so we have a ToE already in our hands !

Epistemology: Bit-String Ontology

- All observations can be coded as a bit-string,
  e.g. camera image in robots or optic nerve signal in humans.

- Classical epistemology operates on a much higher conceptual level and therefore requires stronger (and hence more disputable) philosophical presuppositions.

- We assume
  a temporal bit-string of increasing length is the only observation;

- all higher ontologies are constructed from it
  and are pure "imagination".

All-a-Carte Models

are even simpler ways of obtaining ToEs

- Discretize our observable space-time universe at e.g. Planck level, and code it into a huge finite bit string o.
- Think of a digital high resolution 3D movie of the universe from the big bang to the big crunch.
- Now define infinite bit string:

  $u :=$ Infinite sequence of random bits (fair coin tosses), or

  $u :=$ Champernowne's number $= 0.1\,10\,11\,100\,101\,110\,111\ldots$, or

  $u := \boxed{\sqrt{2}} = \boxed{1.0110101000001001111001100110011111110011\ldots}$

String u contains o (actually infinitely often) $\Longrightarrow$ u is a perfect ToE.

[Reminiscent of Boltzmann's idea: in a sufficiently large random universe,]
[there exist low entropy regions that resemble our own universe.]

... but something doesn't seem right here ...

PREDICTIVE POWER & OBSERVER LOCALIZATION

- Some models seem bogus, others solid, and some are borderline.
- Many now accepted theories have once been regarded as insane.

$\Longrightarrow$ scientific community or general public as a judge is problematic and can lead to endless discussions.

- Examples: Historic geo$\leftrightarrow$heliocentric battle.
  Ongoing discussion of whether string theory is a ToE or more a ToN.
- Problem: Line of sanity differs for different people and different historical times.
- Standard (pseudo)justifications: Authority, Bible, Popper, Anthropic
- This talk: rational criterion whether a ToE makes sense or not.

Intuitive Sanity Status of Some Models

- Moving the Earth out of the center of the Universe was (and for some even still is) insane.
- The Standard Model + Gravity is accepted by nearly all physicists as the closest approximation to a ToE so far.
- Only outside physics, often by opponents of reductionism, this view has been criticized.
- Some respectable researchers including Nobel Laureates go further and take String Theory and even some Multiverse Theories serious.
- Universal ToE also has a few serious proponents.
- All-a-Carte Models seem clearly bogus.

Trend: Size of Worlds Increases

The progression of theories suggested for our world, from ego- to geo- to helio-centric models to universe and multiverse theories and beyond, shows one trend:

– The size of the described worlds increases, with humans being expelled from their center to ever more remote and random locations.

+ More accurate and comprehensive models of the world.

o First, larger model is ridiculed, later accepted. Can this go on forever? Will Multiverse, Universal Universe & All-a-Carte become accepted?

Rigorous scientific justification?

Predictive Power of Multiverse Theories

- Multiverse models: explain existence of our universe, but have reduced predictive power:

- Need to know in which Universe we are to make testable predictions
  Inflation: where, String theory: which compactification, SM+G: 20 parameters, baby universes: which, ...

- Anthropic arguments are not convincing! (Smolin 2004)

- Universal Universe: perfect ToE, but need to know which program generates **our** observable Universe to make testable predictions.

- All-a-Carte Models: Useless ToE, except o in u has been localized. But localization of o in u requires specification of whole o itself.

Predictive Power of 'Small' Historic Models

- The loss of predictive power when enlarging a Universe to a Multiverse model has nothing to do with Multiverses per se.
- Indeed, distinction between Universe and Multiverse is not absolute.
- Egocentric models can be used directly for prediction.
- Geocentric model: Need to localize yourself out of 10^{10} humans.
- Heliocentric model: Need also to know on which planet we are in order to predict celestial movement.
- (Assume deterministic) Universe model:
  need to know which is our Sun out of 10^{22} stars.

Conclusion

- We lose something (some predictive power) when progressing to too large Universe and Multiverse models.
- Localizing yourself can be important to make predictions. If pushed to the extreme, ToE becomes trivial but localization infeasible.

$\Longrightarrow$ A Complete ToE requires model of universe & observer.

- Example: If and only if we know we were in the center of universe $u = 001011011$, we can predict that we will 'see' $o = 1011$ when 'looking' to the right.
- Cf. Egocentric model $u = 1011$ needs no extra specification.

Need to balance model complexity & observer localization complexity ...

COMPLETE TOES (CTOES)

need specification of

(i)	initial conditions	}	deterministic ToE
(e)	state evolution		in conventional sense
(l)	localization of observer	}	required for CToE
(n)	random noise	}	for stochastic models
(o)	perception ability of observer	}	explained later

We will ignore noise (n) and perception ability (o) in the following and resume to these issues later.

Next we need a way to compare ToEs ...

Predictive Power and Falsifiability

- Whatever the intermediary guiding principles for designing theories/models (elegance, symmetries, tractability, consistency), the ultimate judge is predictive success.

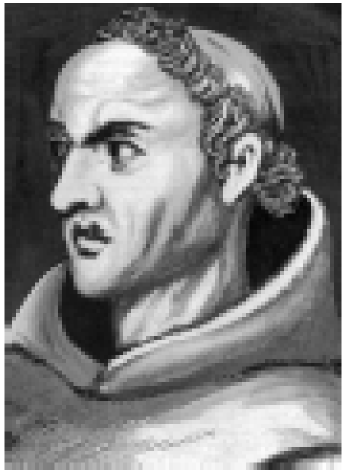
- Unfortunately we can never be sure whether a given theory makes correct predictions in the future.

- Example: Grue Emerald Paradox:
  Theory 1: All emeralds are green.
  Theory 2: All emeralds found till y2020 are green & thereafter blue.

- Both theories are equally consistent with the observations.
  Popper's falsifiability principle doesn't help.

- Solution: Ockham's razor: take simplest theory consistent with data.

Ockham's Razor



- ... tells us to choose the simpler among two otherwise equally good theories.
- ... is the most important principle in science
- ... maybe is even the definition of science

- One can show that simpler theories indeed more likely lead to correct predictions.

- For a discussion in the context of theories in physics, see Gell-Mann's (1994) book.

Quantification of Simplicity/Complexity

- Roughly, the complexity of a theory can be defined as the number of symbols one needs to write down the theory.
- More precisely, write down a program for the state evolution together with the initial conditions which reproduces the observation/data, and define the complexity of the theory as the size in bits of the file that contains the program.
- Identify theories with programs and write Length(q) for the length=complexity of program=theory q.
- Keywords: Kolmogorov complexity & Solomonoff induction, Minimum Description/Message Length principle, Overfitting & regularization in statistics (bias↔variance trade-off).

CToE Selection Principle – Informally

Among two CToEs, select the one that has shorter overall length

$$\text{Length}(i) + \text{Length}(e) + \text{Length}(l)$$

- Length/Complexity of Theory:
  All-a-Carte $<$ Universal $<$ Multiverse $<$ Universe.

- Length/Complexity of Localizing Observer:
  All-a-Carte $\gg$ Universal $>$ Multiverse $>$ Universe.

$\Longrightarrow$ All-a-Carte Model does not minimize above expression.

- Universal Universe is nearly as good as Multiverse.

Localization Within our Universe

- So far we have only localized our Universe in the Multiverse, but not ourselves in the Universe.
- Assume the $\sim 10^{11} \times 10^{11}$ stars in our Universe are somehow indexed. In order to localize our Sun we only need its index, which can be coded in about $\log_2(10^{11} \times 10^{11}) \approx 73$ bits.
- To localize earth among the 8 planets needs 3 bits.
- To localize yourself among 7 billion humans needs 33 bits.
- These localization penalties are tiny compared to the difference in predictive power of the various theories (ego/geo/helio/cosmo).
- This explains and justifies theories of large universes in which we occupy a random location.

(C)ToE – Formalization

- UTM = Universal Turing Machine = general-purpose computer.

- UTM takes a program coded as a finite binary string $q \in \{0,1\}^*$, executes it and outputs binary string $u \in \{0,1\}^\infty$.

$$\boxed{\text{UTM}(q) = u_1^q u_2^q u_3^q ... =: u_{1:\infty}^q}$$

- Formal ToE: $u_{1:\infty}^q$ will be the Universe (or Multiverse) generated by ToE candidate q. (high-resolution 3D movie of the whole Universe from big bang to big crunch)

$\Longrightarrow$ q incorporates initial condition (i) and state evolution (e).

Observational Process & Complete ToE

- Consider human in Universe u observing $o \equiv o_{1:\infty}$=parts of the world
  Observation is direct and classical. Think of
  a video camera mounted on a robot recording o.

- Let $s \in \{0,1\}^*$ be program that extracts obs. o^{sq} from universe u^q:

$$\text{UTM}(s, u^q_{1:\infty}) = o^{sq}_{1:\infty}$$

- Program s contains all information about the location and
  orientation and perception abilities of the observer/camera,

$\Longrightarrow$ q specifies not only item (l) but also item (o).

A *Complete* ToE (CToE) consists of a specification of a (ToE,Subject) pair (q, s). Since it includes s it is a *Subjective* ToE.

CToE Selection Principle – Formally

- Let $o_{1:t}^{true}$ be true past observation (whole life experience).
- The observation sequence $o_{1:\infty}^{sq}$ generated by a correct CToE must be consistent with the true observations $o_{1:t}^{true}$

$\Longrightarrow$ Among a given set of perfect $o_{1:t}^{sq} = o_{1:t}^{true}$ CToEs $\{(q,s)\}$ select the one of smallest Length(q) + Length(s). Formally ...

$$(q^*, s^*) := \arg\min_{q,s}\{\text{Length}(q) + \text{Length}(s) : o_{1:t}^{sq} = o_{1:t}^{true}\}$$
where $o_{1:\infty}^{sq} = \text{UTM}(s, \text{UTM}(q))$.

The selected CToE (q^*, s^*) can and should then be used for forecasting future observations via ...$o_{t+1:\infty}^{forecast} = \text{UTM}(s^*, u_{1:\infty}^{q^*})$.

Universal ToE - Formalization

generates all computable universes

q	UTM(q)					
ϵ	u_1^ϵ	u_2^ϵ	u_3^ϵ	u_4^ϵ	u_5^ϵ	$\cdots$
0	u_1^0	u_2^0	u_3^0	u_4^0	$\cdots$	$\cdots$
1	u_1^1	u_2^1	u_3^1	$\cdots$	$\cdots$	
00	u_1^{00}	u_2^{00}	$\cdots$	$\cdots$		
$\vdots$	$\vdots$	$\vdots$	$\vdots$			

Each row corresponds to one universe.

(Schmidhuber, 2000)

Linearize by dovetailing in diagonal serpentines:

$$\breve{u}_{1:\infty} := u_1^\epsilon u_1^0 u_2^\epsilon u_3^\epsilon u_2^0 u_1^1 u_1^{00} u_2^1 u_3^0 u_4^\epsilon u_5^\epsilon u_4^0 u_3^1 u_2^{00} ...$$

Not hard to construct an explicit program $\breve{q}$ for UTM that computes $\breve{u}_{1:\infty} = u_{1:\infty}^{\breve{q}} = \text{UTM}(\breve{q})$.

Extensions to More Realistic Models

Partial&approximate theories: E.g. Newton only predicts planetary positions approximately, but not phenomena involving light.
Solution: Add Length(b), where b are bits that are not or wrongly predicted.

Probabilistic theories: (E.g. QM) Replace programs (q, s) by probability distributions $(s(o|u), q(u))$, and (Shannon) code noise in $\log 1/p(o)$ bits.
$p(o) := \sum_u s(o|u)q(u)$. Cf. two-part MDL.

Theories with parameters: Code parameters to suitable finite accuracy.
For smooth parametric models, a parameter accuracy of $O(1/\sqrt{n})$ is needed, which requires $\frac{1}{2}\log n + O(1)$ bits per parameter.

Infinite/continuous universes: *(a) All separable spaces* have a countable characterization, e.g. rational points in $I\!R^4$. *(b) Loewenheim-Skolem* theorem (an apparent paradox) implies that Zermelo-Fraenkel set theory (ZFC) has a countable model. – And all physics is separable and formalizable in ZFC.

JUSTIFICATION OF OCKHAM'S RAZOR

Proof of Ockham's Razor

- Ockham's razor: Select the simplest:=shortest theory consistent with the observations.
- Universal Self-Sampling Assumption (USSA):
  A priori it is equally likely to be in any of the universes u^q generated by some program $q \in \{0,1\}^*$.
- Theorem: We are most likely in a universe that is (equivalent to) the simplest universe consistent with our past observations.
- Corollary: Among all considered theories, the one selected by Ockham's razor is the one that most likely leads to correct predictions.

$\Longrightarrow$ **Ockham's razor is 'correct' (under USSA).**

Consistent Universes

- Assumption: $o_{1:\infty} = u_{1:\infty}$, and $u_{1:t}$ = observations so far.
- Counting consistent universes:
  $Q_L := \{q : \text{Length}(q) \leq L \text{ and } \text{UTM}(q) = u_{1:t}^{true}*\}$
- Shortest consistent q:
  $q_{min} := \arg\min_q \{\text{Length}(q) : q \in Q_L\}$ and $l := \text{Length}(q_{min})$
- Added (unread) "garbage" g after the end of a program q does not change its behavior $\Rightarrow q_{min}g \in Q_L \Rightarrow |Q_L| \gtrsim 2^{L-l}$.
- Deep result in AIT: If there are many long equivalent programs, then there must also be a short one $\Rightarrow |Q_L| \lesssim 2^{L-l}$

Probabilistic Prediction

Given observations $u_{1:t}^{true}$ we now determine the probability of being in a universe that continues with $u_{t+1:n}$. by counting universes:

$$\begin{aligned} Q_L^n &:= \{q : \text{Length}(q) \leq L \text{ and } \text{UTM}(q) = u_{1:t}^{true} u_{t+1:n}*\} \subset Q_L \\ q_{min}^n &:= \arg\min_q \{\text{Length}(q) : q \in Q_L^n\} \quad \text{and} \quad l_n := \text{Length}(q_{min}^n) \\ |Q_L^n| &\approx 2^{L-l_n} \end{aligned}$$

The probability of being in a universe with future $u_{t+1:n}$ given $u_{1:t}^{true}$ is determined by their relative number

$$P(u_{t+1:n}|u_{1:t}^{true}) = \frac{|Q_L^n|}{|Q_L|} \approx 2^{-(l_n - l)} \qquad \text{(by USSA)}$$

This Implies Ockham's Razor

- $P(u_{t+1:n}|u_{1:t}^{true}) \approx 2^{-(l_n - l)} \Longrightarrow$ the most likely continuation $\hat{u}_{t+1:n} := \arg\max_{u_{t+1:n}} P(u_{t+1:n}|u_{1:t}^{true})$ is (approximately) the one that minimizes l_n.
- By definition, q_{min} is the shortest program in $Q_L = \bigcup_{u_{t+1:n}} Q_L^n$.
- $\Rightarrow \quad P(\hat{u}_{t+1:n}|u_{1:t}^{true}) \approx P(u_{t+1:n}^{q_{min}}|u_{1:t}^{true})$
- In words: We are most likely in a universe that is (equivalent to) the simplest universe consistent with our past observations.

Universal Self-Sampling Assumption

- USSA has by itself no bias towards simple models.
- Indeed, most q in Q_L have length close to L, and since we sample uniformly from Q_L, this actually represents a huge bias towards large models for $L \to \infty$.
- Other non-universal self-sampling leads to similar conclusion.
- How reasonable is UToE ?
  Nearly as good as any other correct ToE, and it is a safe(r) bet.
- How reasonable is USSA? If somebody (but how and who?) would tell us that the universe is computable but nothing else, universal self-sampling seems like a reasonable a priori UToE.

Comparison to Anthropic Self-Sampling

The anthropic self-sampling assumption states that a priori you are equally likely any of the (human) observers on earth, this universe, or any alternative or parallel universe, while

1) USSA samples from any universe and any location (living or dead) in the multiverse and not only among human (or reasonably intelligent) observers.

2) USSA has no problem of what counts as a reasonable (human) observer.

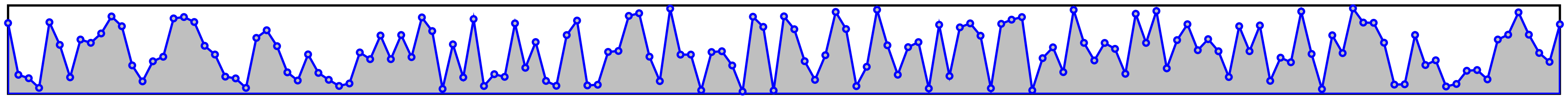
3) The USSA principle is completely formal.

Relation: both sample from the set of reasonable observers,
since $o_{1:t}^{true}$ includes snapshots of other (human) observers.

The No Free Lunch (NFL) Theorem/Myth

- Consider algorithms for finding the maximum of a function, and compare their performance uniformly averaged over all functions over some fixed finite domain.

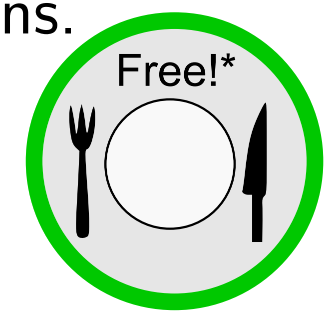
- Since sampling uniformly leads with (very) high probability to a totally random function (white noise), it is clear that on average no optimization algorithm can perform better than exhaustive search.



⇒ All reasonable optimization algorithms are equally good/bad on average.

- Conclusion correct, but obviously no practical implication, since nobody cares about the maximum of white noise functions.

- Uniform and universal sampling are both (non)assumptions, but only **universal sampling** makes sense and offers a free lunch.



*Subject to computation fees

Uniform(NFL)$\leftrightarrow$Universal(USSA) Sampling

USSA also samples uniformly, but over effective models=theories=programs.

A priori we assumed all programs to be equally likely, but the resulting universe distribution is far from uniform.

Phrased differently, we piped uniform noise (via M, see below) through a universal Turing machine.

Formally: We assume a universal Solomonoff distribution, rather than a uniform distribution.

Conclusion:

> The nearly vacuous assumption that the world has any effective structure breaks NFL down, and makes Ockham's razor work.

DISCUSSION

Miscellaneous

- Localizing the observer here has nothing to do with the quantum mechanical measuring process,
  although there may be some deeper yet to be explored connection.

- Schmidhuber (1997&2000): All apparent physical randomness may actually only be pseudo random.

- Hutter (2005): Believing in true random noise may be an unscientific position.

Assumptions

(i) Bit-string ontology: The observers' raw experience of the world can be cast into a single temporal binary sequence o. All other physical and epistemological concepts are derived.

(ii) Realism: There exists an objective world independent of any particular observer in it.

(iii) Computable universe: The world is computable, i.e. there exists an algorithm (a finite binary string) which, when executed, outputs the entire space-time universe.

(iv) Computable observer process:
The observer is a computable process within the objective world.

(v) Ockham's razor principle:
Choose the simplest theory consistent with the observations.

Conclusions

- Unlike falsificationism, quantified versions of Ockham's razor can serve as the foundation of science.

- A theory that perfectly describes our universe or multiverse, rather than being a Theory of Everything (ToE), might also be a theory of nothing.

- A predictively meaningful theory can be obtained if the theory is augmented by the localization of the observer.

- A truly Complete Theory of Everything (CToE) (q, s) consists of a conventional (objective) ToE q plus (subjective) observer process s.

- The bit-string ontology, realism, computability, subjectivism, and Ockham's razor quantified in terms of code-length minimization enable a scientifically meaningful and systematic quest for a theory of everything.

Conclusions (cntd)

- More precisely, the CToE Selection Principle allows a rigorous and quantitative comparison of CToEs and can even be used to select the "best" CToE (q^*, s^*).
- As a side result, this allows to separate objective knowledge q from subjective knowledge s.
- One might even argue that if q^* is non-trivial, this is evidence for the existence of an objective reality.
- Another side result is that there is no hard distinction between a universe and a multiverse; the difference is qualitative and semantic.
- Finally, since s traces the observer from birth till now, a CToE also makes predictions about the continuity and survival of the self and personal identity over time.

Summary

- Respectable researchers have dismissed and embraced each single model of the world discussed above
  – at different times in history and concurrently,
  – often based on unscientific arguments.

- I presented a more serious treatment of world model selection.

- I introduced and discussed the usefulness of a theory formally in terms of predictive power based on model *and* observer localization complexity.

- Outlook: Compute and compare complexities of concrete theories, e.g. compare SM+G with String Theory.

Thanks! Questions? Details:

- M. Hutter. *A complete theory of everything (will be subjective).* Algorithms, 3(4):329–350, 2010.
- M. Hutter. *The subjective computable universe.* A Computable Universe: Understanding and Exploring Nature as Computation, pages 399–416. World Scientific, 2012.
- J. Schmidhuber. *Algorithmic theories of everything.* arxiv.org/quant-ph/0011122, (2000).
- M. Gell-Mann. *The Quark and the Jaguar: Adventures in the Simple and the Complex*. W.H. Freeman & Company, (1994).
- S. Rathmanner and M. Hutter. *A philosophical treatise of universal induction.* Entropy, 13(6):1076–1136, 2011.
- A. Blumhofer and M. Hutter. *Family structure from periodic solutions of an improved gap equation.* Nuclear Physics, B484:80–96, 1997. Missing figures in B494 (1997) 485.
- M. Hutter. *One decade of universal artificial intelligence.* In Theoretical Foundations of Artificial General Intelligence, pages 67–88. Atlantis Press, 2012.