Simplifying Simplicity

Christian J. Feldbacher-Escamilla

Summer 2019

Project Information

Publication(s):

- Feldbacher-Escamilla, Christian J. (submitted). "Simplifying Simplicity". In: *manuscript*.

Talk(s):

- Feldbacher-Escamilla, Christian J. (2019b-08-05/2019-08-10). *Simplicity in Abductive Inference*. Conference. Presentation (contributed). 16th Congress of Logic, Methodology and Philosophy of Science (CLMPST16). University of Prague: Division of Logic, Methodology, Philosophy of Science, and Technology (DLMPST).
- Feldbacher-Escamilla, Christian J. (2019a-06-03/2019-06-04). *Defective Information and Abduction*. Conference. Presentation (contributed). Understanding Defectiveness in the Sciences. Universidad Nacional Autónoma de México: Instituto de Investigaciones Filosóficas (UNAM).
- Feldbacher-Escamilla, Christian J. (2019d-05-23/2019-05-25). *Simplifying Simplicity*. Conference. Presentation (contributed). Simplicities and Complexities. University of Bonn: The Epistemology of the LHC.
- Feldbacher-Escamilla, Christian J. (2019c-02-25/2019-02-27). *Simplicity in Abductive Inference*. Conference. Presentation (contributed). GWP.2019. University of Cologne: The German Society for Philosophy of Science.

Project(s):

- DFG funded research unit *Inductive Metaphysics* (FOR 2495); subproject: *Creative Abductive Inference and its Role for Inductive Metaphysics in Comparison to Other Metaphysical Methods*.

Simplifying Simplicity

There is a plurality of simplicity constraints.

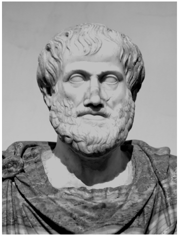
Questions:

- How to characterise simplicity?
- In particular: What is the epistemic role of it?
- How to account for the different forms of simplicity?

Contents

The Many Faces of Simplicity

The Tradition of Simplicity: Aristotle



Aristotle:

> *"Let that demonstration be better which, other things being equal, depends on fewer postulates or suppositions or propositions. For if they are equally familiar, knowing will come about more quickly in this way; and that is preferable."*
> *(Aristotle 1995, Posterior Analytics, 86ª30ff, p.322)*

Fewer propositions or axioms $\Rightarrow$ Better dynamics of knowledge acquisition

$=_{def}$ Axiomatic Simplicity

The Tradition of Simplicity: Ockham

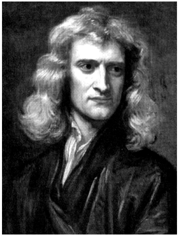
William of Ockham:

"Pluralitas non est ponenda sine necessitate"
Plurality must never be posited without necessity.
(cf. Gauch 2003, p.272)

Libertus Fromondus in 1649:

"Novaculam [...] Occami:
Non sunt multiplicanda entia sine necessitate[.]"
Ockham's razor:
Entities are not to be multiplied without necessity.
(cited after Hübener 1983, p.84)

The Tradition of Simplicity: Newton

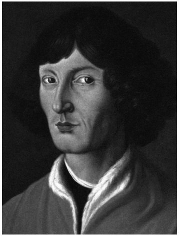


Sir Isaac Newton:

> *"Rule 1: No more causes of natural things should be admitted than are both true and sufficient to explain their phenomena.*
>
> *As the philosophers say: Nature does nothing in vain, and more causes are in vain when fewer suffice. For nature is simple and does not indulge in the luxury of superfluous causes."*
> *(Newton 1726(E3)/1999, Regulæ Philosophandi, Regula I, p.794)*

Simplicity of domain of our theories $\Rightarrow$ Maps to simplicity of nature $=_{def}$ Ontic Simplicity

The Tradition of Simplicity: Copernicus



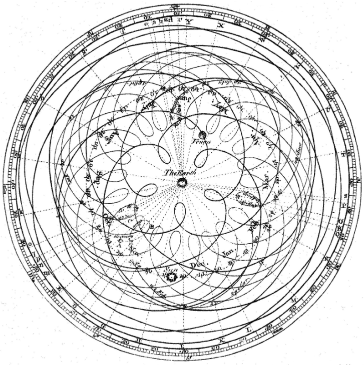
Nicolaus Copernicus:

> *"Alongside the ancient hypotheses, which are no more probable, let us permit these new hypotheses also to become known, especially since they are admirable as well as simple and bring with them a huge treasure of very skillful observations."*
> *(Copernicus 1543/1992, pp.XX, my emphasis)*

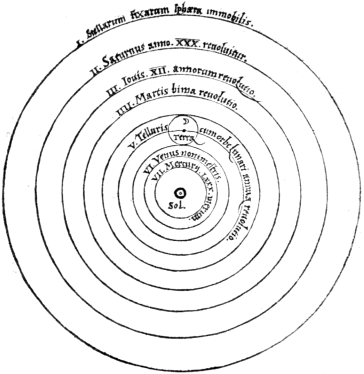
Simplicity of model $\Rightarrow$ Instrumental for predictive success

$=_{def}$ Parametric Simplicity

The Tradition of Simplicity: Copernicus



(James Ferguson, based on diagrams of Giovanni Cassini, in: A Society of Gentlemen in Scotland 1771, p.448, plate XL)



(Copernicus 1543/1992, chpt.10)

The Many Faces of Simplicity

- **Axiomatic Simplicity**: T_S is axiomatic simpler than T_C iff the cardinality of the set of *non-conjuctive and independent* axioms of T_S is smaller than that of T_C (when comparing against a common interpretation and background logic).

- **Ontic Simplicity**: T_S is ontic simpler than T_C iff the smallest cardinality of domains of models (interpretations) of T_S is smaller than that of T_C.

- **Parametric Simplicity**: T_S is parametric simpler than T_C iff T_S and T_C are polynomials and T_S has a lower degree than T_C.

How do these notions of simplicity relate to truth?

How to epistemically justify principles of simplicity?

The Epistemic Value of Parametric Simplicity

Truth-Conduciveness of Simplicity: Main Idea

E: Evidence: data points

H: Hypothesis/theories/models: functions (polynomials)

E might contain error.

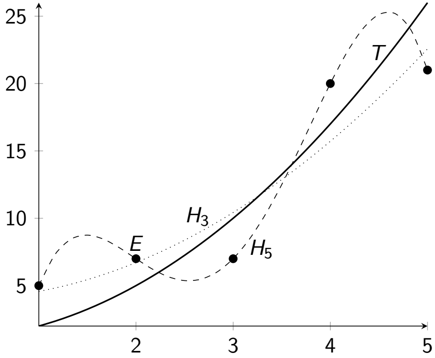
Task: H ought to fit, but not overfit E.

Complex models are prone to overfit.

Truth-Conduciveness of Simplicity: Argument

1. Data E might be noisy and involve *error*. *Error*
2. An accurate fit of a model H to the data E fits also error.
   $Error \Rightarrow (Accuracy \Rightarrow Falsehood)$
3. Whereas a less accurate fit of H to E may depart from error.
   $Error \Rightarrow (Inaccuracy \Rightarrow PosTruth)$
4. Fact: The more parameter, the more prone to overfit.
   $Complexity \Rightarrow Accuracy\ \&\ Simplicity \Rightarrow Inaccuracy$
5. Hence: Simplicity (having less parameters) may account for inaccuracy w.r.t. data E in order to achieve accuracy w.r.t. the truth.
   $Complexity \Rightarrow Falsehood\ \&\ Simplicity \Rightarrow PosTruth$

Truth-Conduciveness of Simplicity: Example



Curve fitting with a polynomial of degree 4 with 5 parameters H_5 and a polynomial of degree 2 with 3 parameters H_3. H_5 perfectly fits data set E, whereas H_3 deviates from E. However, H_5 has more distance from the truth T, whereas H_3 approximates T.

Truth-Conduciveness of Simplicity: Theory

The estimated predictive accuracy of the family of a model H given some data E (*Akaike information criterion* $AIC(H, E)$) is determined by (cf. Forster and Sober 1994, p.10):

$$AIC(H,E) = \frac{1}{|E|} \cdot (\log(Pr(E|H')) - c(H)) \qquad \text{(AIC)}$$

$c(H)$: number of parameters of H
H': most accurately parametrised model of family H regarding E

E.g.: Given equal accuracy: $Pr(E|H'_1) = Pr(E|H'_2)$
It follows: $c(H_1) < c(H_2) \Rightarrow AIC(H_1, E) > AIC(H_2, E)$

Upshot: Simplicity matters for estimated predictive accuracy.
$\Rightarrow$ epistemic value

Back to the Plurality of Simplicity

There are several characterisations of "simplicity". Most common are:

- number of axioms of a theory (axiomatic)
- number of presupposed entities (ontic)
- number of parameters of a model (parametric)

The argument for the epistemic value of simplicity from above is about the parametric notion.

Question: What is the epistemic rationale of an *axiomatic* and *ontic* simplicity constraint?

The Epistemic Value of Axiomatic and Ontic Simplicity

A Reductionist Approach

Main idea: Try to reduce all simplicity considerations to that of the parametric notion.

If this succeeds, then the model-theoretic argument for the epistemic value of simplicity can be employed in the other cases too.

How to reduce the axiomatic and ontic approach to the model-theoretic approach?

Regarding the ontic notion, the main idea can be found already in (Forster and Sober 1994, sect.4).

We aim at elaborating on this and expand it also to the axiomatic notion.

Ontic Simplicity

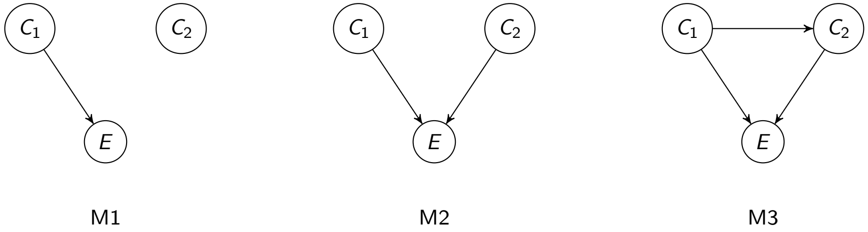
Recall Newton's causal reasoning constraint:

> *"No more causes of natural things should be admitted than are both true and sufficient to explain their phenomena." (cf. Newton 1726(E3)/1999, pp.794–796)*

Epistemic rationale (cf. Forster and Sober 1994, sect.4): Assume E is to be explained by causes $C_1, C_2, \ldots$:

Ontic Simplicity: Options

There are several explanatorial options available:



Probabilistic modelling of causal relations:

$Pr(E\|\cdot)$	C_1	$\overline{C_1}$
C_2	$c_0 + c_1 + c_2 + c_{1,2}$	$c_0 + c_2$
$\overline{C_2}$	$c_0 + c_1$	c_0

The c_is are the parameters of the models.

Ontic Simplicity: Models

The respective models are:

- M1 $Pr(E|C_1, C_2) = c_0 + c_1 \cdot val(C_1)$ (single cause)
- M2 $Pr(E|C_1, C_2) = c_0 + c_1 \cdot val(C_1) + c_2 \cdot val(C_2)$ (non-interactive causes)
- M3 $Pr(E|C_1, C_2) = c_0 + c_1 \cdot val(C_1) + c_2 \cdot val(C_2) + c_{1,2} \cdot val(C_1) \cdot val(C_2)$ (causes with interactions)

By employing reasoning of model selection:

- Accuracy is typically better in M3>M2>M1
  $\Rightarrow$ on average more fitting of error in data, i.e. overfitting.
- Simplicity is better in M1>M2>M3
  $\Rightarrow$ on average less fitting of error in data.

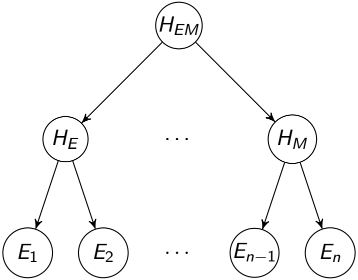
One can reduce the ontic to the parametric parlance about simplicity.

Axiomatic Simplicity

E.g. different patterns of abduction (cf. Feldbacher-Escamilla and Gebharter 2019; Schurz 2008).

One might aim at explaining empirical correlations by help of dispositions.

For Example:



Axiomatic Simplicity: Unification

Performance is measured via unification
(in $[-1, \infty]$, positive: there is unification):

$$u(E|H_i) = \frac{\#\text{ of statements to be unified } (E)}{\underbrace{\#\text{ of unifying statements}}_{c \text{ ... complexity } c(H_i)}} - 1$$

Why choosing that H_i which needs least $\#$ of unifying statements?

Answer: Also via structural equations as before.

Axiomatic Simplicity: More Details

Let Pr be the probability distribution provided by the axioms of the theory.

Let Pr/the theory be about random variables $X_1, \ldots, X_n$

Then Pr/the theory basically just fixes independently of each other:

- $Pr(X_1)$
- $Pr(X_1|X_2), Pr(X_1|\overline{X_2})$
- $Pr(X_1|X_2, X_3), Pr(X_1|\overline{X_2}, X_3), Pr(X_1|X_2, \overline{X_3}), Pr(X_1|\overline{X_2}, \overline{X_3})$

$$\vdots$$

For n variables, the # of independently fixed probability statements is $2^n - 1$.

$\Rightarrow$ Structural equations as in the case of ontic simplicity (cf. Pearl 2000, sect.1.4.1).

$\Rightarrow$ We can count model parameters.

Axiomatic Simplicity: Tradition

Theory and hypothesis choice via Bayes' theorem:

$$Pr(H_i|E) = Pr(E|H_i) \cdot \frac{Pr(H_i)}{Pr(E)}$$

Two ingredients needed for comparison: $\underbrace{\text{likelihoods}}_{Pr(E|H_i)}$ and $\underbrace{\text{priors}}_{Pr(H_i)}$

There are different approaches for determining the priors.

E.g. principle of indifference—problems: Bertrand's paradox and infinity

Alternative: Ordering of priors via simplicity (cf.,e.g. Jeffreys 1939/2003): Simplicity-order of H_i; then assign $1/2, 1/4, \ldots$ $((1/2)^n)$ prior Pr.

So, priors are linked to $c(H_i)$.

Difference to the model-theoretical approach: Bayesianism provides no argument for simplicity (cf. Sober 2015, chpt.2).

Summary

- Simplicity comes in several forms:
  - axiomatic
  - ontic
  - parametric

- Regarding parametric simplicity: Akaike style measure of estimated predictive success; depends on:
  - Pr ... accuracy
  - c ... simplicity/complexity

- Rationale for minimising complexity c: avoiding error

- Reduction of axiomatic and ontic notions to parametric notion by help of structural equations

References I

A Society of Gentlemen in Scotland, ed. (1771). *Encyclopedia Britannica or A Dictionary of Arts and Sciences, Compiled Upon a New Plan. In which the different sciences and arts are digested into distinct treatises or systems; and the various technical terms, etc. are explained as they occur in the order of the alphabet. In three volumes.* Vol. Volume I. Edinburgh: A. Bell and C. Macfarquhar.

Aristotle (1995). *The Complete Works of Aristotle*. Ed. by Aristotle (ed. by Barnes, Jonathan). Vol. Volume One and Two. Princeton: Princeton University Press.

Copernicus, Nicholas (1543/1992). *On the Revolutions*. Translated by E. Rosen. Maryland: The Johns Hopkins University Press.

Feldbacher-Escamilla, Christian J. (submitted). "Simplifying Simplicity". In: *manuscript*.

Feldbacher-Escamilla, Christian J. and Gebharter, Alexander (2019-01). "Modeling Creative Abduction Bayesian Style". In: *European Journal for Philosophy of Science* 9.1, pp. 1–15. DOI: 10.1007/s13194-018-0234-4.

Forster, Malcolm R. and Sober, Elliott (1994). "How to Tell When Simpler, More Unified, or Less Ad Hoc Theories Will Provide More Accurate Predictions". In: *The British Journal for the Philosophy of Science* 45.1, pp. 1–35. DOI: 10.1093/bjps/45.1.1.

Gauch, Hugh G. (2003). *Scientific Method in Practice*. Cambridge: Cambridge University Press.

Hübener, Wolfgang (1983). "Occam's Razor Not Mysterious". In: *Archiv für Begriffsgeschichte* 27, pp. 73–92. DOI: 10.2307/24362877.

Jeffreys, Harold (1939/2003). *Theory of Probability*. Oxford: Clarendon Press.

References II

Newton, Isaac (1726(E3)/1999). *The Principia: Mathematical Principles of Natural Philosophy: A New Translation*. Ed. by Cohen, I. Bernard and Whitman, Anne. Berkeley: University of California Press.

Pearl, Judea (2000). *Causality. Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Schurz, Gerhard (2008). "Patterns of Abduction". English. In: *Synthese* 164.2, pp. 201–234. DOI: 10.1007/s11229-007-9223-4.

Sober, Elliott (2015). *Ockham's Razors. A User's Manual*. Cambridge: Cambridge University Press.