Explanation in *Science*

James A. Overton (james@overton.ca)

PSA, Chicago – November 7, 2014

Common Questions

- what is a scientific explanation?
- what is explained?
- what does the explaining?

Diverse Answers

- deductive-nomological (D-N) (Hempel and Oppenheim 1948; Hempel 1965)
- statistical relevance (Salmon 1971)
- unification (Friedman 1974; Kitcher 1989)
- pragmatic (van Fraassen 1980)
- causal-mechanical (Salmon 1984; Dowe 2000)
- intervention (Woodward 2003)
- mechanistic (Machamer, Darden, and Craver 2000; Bechtel and Abrahamsen 2005)
- asymptotic (Batterman 2002)
- model-based (Bokulich 2009)

Science Dataset

- 781 articles from one year of the journal *Science*
- large set of small case studies, randomly sampled
  - Sample A: 25 “explain” sentences
  - Sample B: 100 sentences
  - Sample C: 25 abstracts
- I use Sample A to build my account and the others to test it

Previous Work

- "Explain" in scientific discourse, *Synthese* 190(8):1383–1405, 2013.
- explanation is a goal of scientific practise
- explanation is important for understanding scientific practise
- explanation is general, across sciences

- justifies, at least in part, the diversity of philosophical accounts
- is there a unity to scientific explanation?

Current Work

- a general philosophical account of scientific explanation

Case A10: Quotation

No clear theoretical predictions for a star with parameters similar to those for HIP 13044 exist, hence it is possible that some high-order oscillations can explain the 1.4- or 3.5-day signal.

Setiawan, J., R.J. Klement, T. Henning, H.W. Rix, B. Rochau, J. Rodmann, and T. Schulze-Hartung. 2010. A giant planet around a metal-poor star of extragalactic origin. *Science* 330(6011):1642.

Case A10: Gloss

High-order oscillations *of luminance* in the theoretical predictions for *the stellar dynamics* of stars with parameters similar to HIP 13044 CAN POSSIBLY EXPLAIN the length of the signals *in the luminance* of HIP 13044.

Case A10: Normal Form

can possibly explain

The high-order
of oscillations
of luminance
in the theory
of stellar dynamics
in models of stars
with parameters similar
to HIP 13044

the length
of the signals
in the luminance
of HIP 13044
which is a star
with parameters similar
to HIP 13044.

Case A10: Patterns

phrase	can possibly explain	
top	The high-order	the length
core	of oscillations *of luminance* in the theory *of stellar dynamics* *in models*	of the measurements of the oscillations *of luminance* of HIP 13044
base	of stars with parameters similar to HIP 13044	*which is a star* *with parameters similar* *to HIP 13044.*

Case A10: Base

phrase	can possibly explain	
top	The high-order	the length
core	of oscillations *of luminance* in the theory *of stellar dynamics* *in models*	of the measurements of the oscillations *of luminance* of HIP 13044
base	**of stars** **with parameters similar to HIP 13044**	***which is a star*** ***with parameters similar to HIP 13044.***

Case A10: Qualities

phrase	can possibly explain	
top	**The high-order**	**the length**
core	of oscillations *of luminance* in the theory *of stellar dynamics* *in models*	of the measurements of the oscillations *of luminance* of HIP 13044
base	of stars with parameters similar to HIP 13044	*which is a star* *with parameters similar* *to HIP 13044.*

Case A10: Core

phrase	can possibly explain	
top	The high-order	the length
core	**of oscillations**	**of the measurements**
	of luminance	**of the oscillations**
	in the theory	***of luminance***
	of stellar dynamics	**of HIP 13044**
	in models	
base	of stars	*which is a star*
	with parameters similar	*with parameters similar*
	to HIP 13044	*to HIP 13044.*

Case A4: Quotation

Physiological concentrations of ADP [adenosine diphosphate] inhibit kinase activity in the oscillator, and a mathematical model constrained by data shows that this effect is sufficient to quantitatively explain entrainment of the cyanobacterial circadian clock.

Rust, M.J., S.S. Golden, and E.K. O'Shea. 2011. Light-driven changes in energy metabolism directly entrain the cyanobacterial circadian oscillator. *Science* 331(6014):220.

Case A4: Patterns

phrase		is sufficient to quantitatively explain
top	The physiological concentration	*the rate*
core	of ADP in the mathematical *model* of the activity of kinase	of entrainment
base	in circadian clocks of cyanobacteria	in circadian clocks of cyanobacteria.

Case A12: Quotation

In summary, changes in water mass formation processes are not necessarily required to explain the high GNAIW [Glacial North Atlantic Intermediate Water] end-member $\delta^{13}C$ values.

Olsen, A., and U. Ninnemann. 2010. Large $\delta^{13}C$ gradients in the preindustrial North Atlantic revealed. *Science* 330(6004):658.

Case A12: Patterns

phrase	are not necessarily required to explain	
top	Changes	the large size
core	in the processes of formation	of measurements of $\delta^{13}C$ in end-members of GNAIW
base	of water masses	*which is a water mass.*

Five Categories

- data
- entity
- kind
- model
- theory

Data

- a statement about an entity
- Types:
  - measurements
  - observations
  - images
- Examples:
  - the luminosity measurements of HIP 13044 (A10)
  - the measurements of rates of entrainment of circadian clocks (A4)
  - the measurements of end-member δ^{13}C values (A12)
  - the observations of the severity of the Fog phenotype in *C. elegans* (B58)

Entity

- a concrete particular thing or process
- Types:
  - stars
  - samples
  - specimens
- Examples:
  - star HIP 13044 (A10)
  - GNAIW's formation process (A12)
  - the Tagish Lake meteorite (A19)
  - the sample of carbon monoxide extracted from ice core D47 in (A15)

Kind

- an abstract universal class of entities
- Types:
  - natural kinds
  - species
  - universals
- Examples:
  - stars with parameters similar to HIP 13044 (A10)
  - circadian clocks (A4)
  - ADP (A4)
  - water masses (A12)
  - lithium (A1)
  - *E. coli* (A8)
  - Mn_4CaO_2 (B68)

Model

- an abstract description of the relationships that hold between kinds and their qualities
- Types:
  - sets of differential equations
  - mechanisms
  - flow charts
- Examples:
  - models of stellar dynamics (A10)
  - mathematical models of kinase activity in circadian clocks (A4)
  - Brownian random walks modelling foraging behaviour (B49)
  - reaction-diffusion equations modelling spatially periodic biological structures (B45)
  - a hierarchical model of stem cell crypts (A14)

Theory

- a principle, set of principles, or a formal system that is a building block for models.
- Types:
  - laws
  - empirical generalizations
  - mathematical formalisms
- Examples:
  - the theory of stellar dynamics (A10)
  - the theory of chromosomal supercoiling (B21)
  - universal hydrodynamics (B2)
  - the defensive function of sabre teeth (B54)
  - the mathematical theory of differential equations (B45)

Pairs of Categories

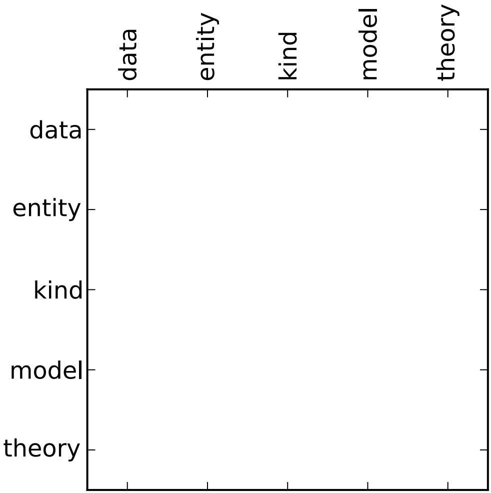

Figure 1: Pairs of categories

Sample A: Variety

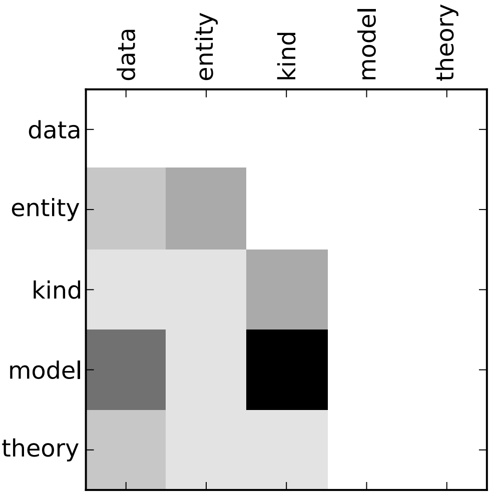

Figure 2: Sample A heatmap

Relations Between Categories

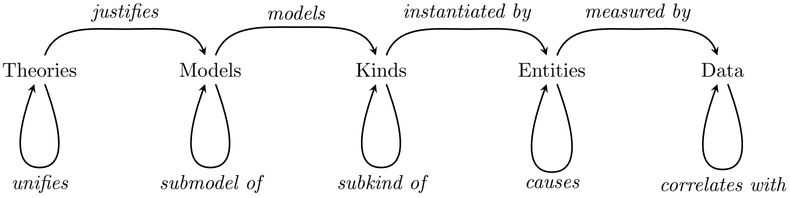

Figure 3: Some relations between instances of categories

Structure of an Explanation: Basic

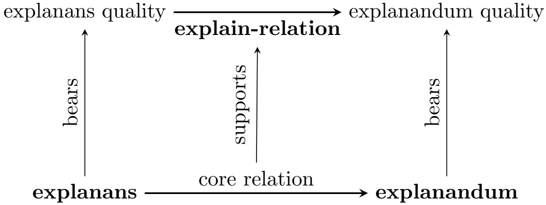

Figure 4: Basic structure of a scientific explanation.

Scientific Explanation

1. explanans:
    - a quality/property/characteristic
    - of a data/entity/kind/model/theory
    - at least as *general* as the explanandum
2. explanandum
    - a quality/property/characteristic
    - of a data/entity/kind/model/theory
    - at least as *specific* as the explanans
3. explain-relation:
    - expresses the counterfactual dependence of the explanandum quality on the explanans quality
    - answers: What if things had been different?
    - supported by a core relation:
        - connects explanans to explanandum
        - counterfactual supporting

Structure of an Explanation: Theory-Data

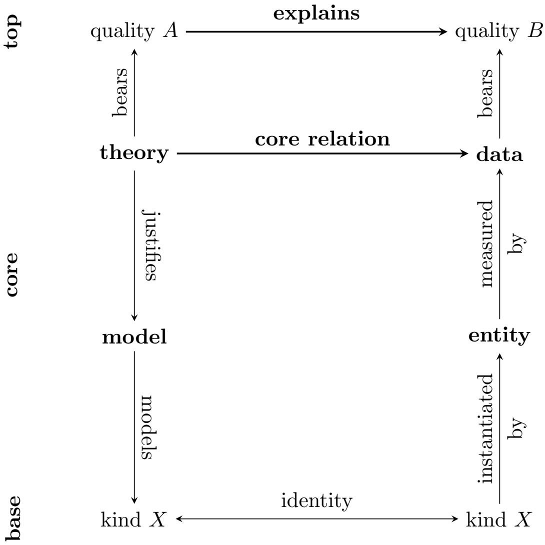

Figure 5: General structure of a theory-data explanation.

Evidence and Explanation

- in Samples B and C, order can be reversed
- explanation
  - general to specific
  - lower-left triangle
- evidence
  - specific to general
  - upper-right triangle
- otherwise the same structure

Sample A: 25 “explain” sentences

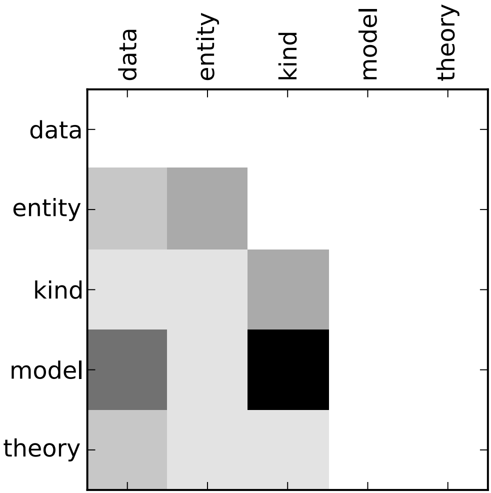

Figure 6: Sample A heatmap

Sample B: 100 sentences

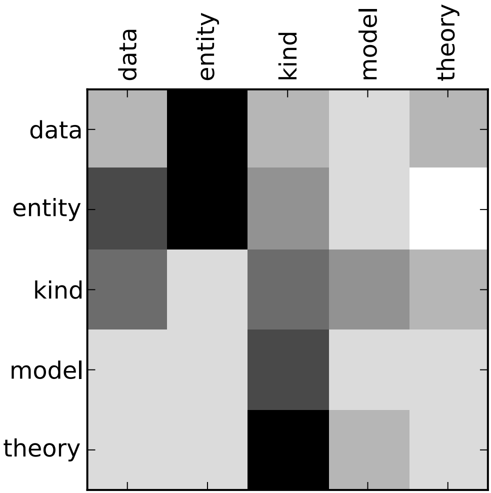

Figure 7: Sample B heatmap

Sample C: 25 abstracts

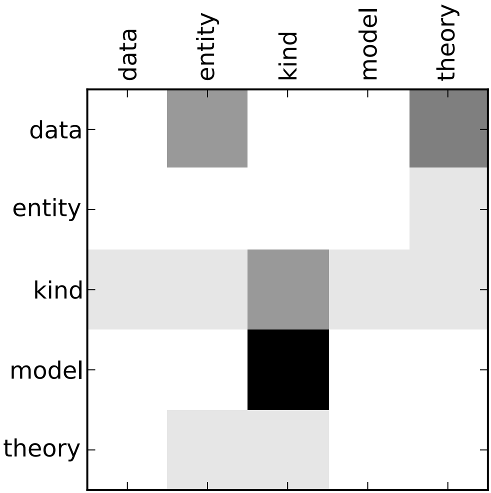

Figure 8: Sample C heatmap

Samples A, B, C

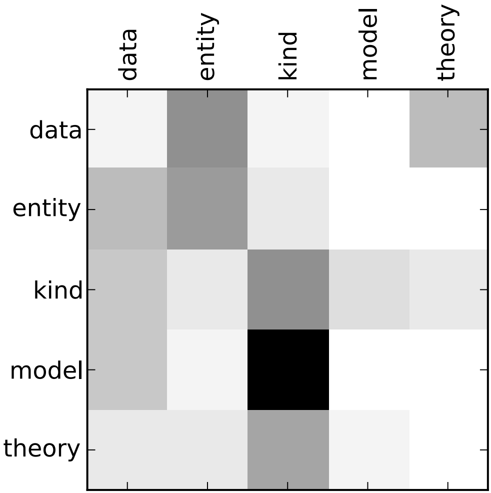

Figure 9: Samples A, B, C heatmap

Room for Other Accounts

- theory-data
  - deductive-nomological (D-N)
- theory/model
  - unification
  - asymptotic
- model/kind/entity
  - intervention
  - mechanistic
  - model-based
- entity/data
  - causal-mechanical
- no room
  - statistical relevance
  - pragmatic

Evidence for Other Accounts in the *Science* Dataset

- strong evidence
  - intervention
  - mechanistic
  - model-based
- very weak evidence
  - deductive-nomological
  - causal-mechanical (Salmon and Dowe)
- no evidence
  - statistical relevance
  - unification
  - asymptotic
- equivocal
  - pragmatic

Upshot

- one explain-relation, different core relations
- five categories for the explanans and explanandum
- pairs of categories determine core relation
- generalized counterfactual account
- many existing accounts fit the framework, but not everything goes

Appendix: Science Informatics

- data: measurements in databases, spreadsheets; images
- entity: subject IDs, barcodes, URIs
- kind: domain ontologies
- model: programs
- theory: software libraries