Getting to the Truth through Conceptual Revolutions

Kevin T. Kelly; Clark Glymour

PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol. 1990, Volume One: Contributed Papers (1990), 89-96.

Stable URL:
http://links.jstor.org/sici?sici=0270-8647%281990%291990%3C89%3AGTTTTC%3E2.0.CO%3B2-3

PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association is currently published by The University of Chicago Press.

Getting to the Truth Through Conceptual Revolutions[1]

Kevin T. Kelly and Clark Glymour

Carnegie Mellon University

> [I]t would be absurd for us to hope that we can know more of any object than belongs to the possible experience of it or lay claim to the least knowledge of how anything not assumed to be an object of possible experience is determined according to the constitution that it has in itself.
>
> * * *
>
> It would be... a still greater absurdity if we conceded no things in themselves or declared our experience to be the only possible mode of knowing things....
>
> [Kant, Prolegomena to Any Future Metaphysics]

1. Theory-Laden Data and the Aims of Inquiry

A certain line of skepticism about normative epistemology has become more or less standard in contemporary philosophy of science. It runs like this.

> Scientific method is a matter of choosing among theories on the basis of evidence.
>
> But in "conceptual revolutions", meaning, truth, and even what counts as observable can all be theory-relative.
>
> So since the data changes depending upon which theory one views it through, it is possible that a scientist who views the world through theory T1 will see the evidence as supporting T1 over T2, but a scientist who views the world through theory T2 will see the evidence as supporting T2 over T1.
>
> So there is no objective stance from which the choice of the first theory is more rational than the choice of the second.
>
> So theory choice is arbitrary, or if it is not arbitrary, it is based up on political or other "extra-rational" considerations (e.g. which of the competing theorists is more popular).
>
> Since science is just a sequence of non-rational theory choices, it is a non-rational activity.

PSA 1990, Volume 1, pp. 89-96

"Edifying" critics like Rorty (1979), who see the history of epistemology as a sustained attempt to legislate a fixed, Archimedian point to anchor inter-cultural debate, seize upon such arguments in the philosophy of science as part of their case that the quest by "systematic" epistemology for such an objective standpoint is hopeless. Opponents of this criticism argue that in real scientific cases sufficient data was objectively shared to make the actual choices rational. But both sides agree in their focus on convergence to agreement as a necessary element of the rationality of scientific inquiry.

We explore a different approach. With the "edifying" critics, we entertain the possibility of radically theory-relative evidence and grant the sceptics that if evidence is radically theory-relative, then one should not expect or demand eventual convergence of opinion. But unlike the critics, we do not infer from this fact that normative epistemology is hopeless or pointless, for it has other traditional goals that still make sense given theory-laden data.

According to *convergent realists*, a major goal of inquiry is the eventual approach to a true, informative theory of the world one lives in. This goal can be recognized in some form in the work of such philosophers as Plato, Aristotle, Bacon, Descartes, Mill, Whewell, Peirce, Reichenbach, and (early) Putnam. Each of these philosophers saw normative scientific method as a means toward truth, and each of them rejected the idea that methodology is a descriptive science.

We develop a formal theory of achieving truth in the limit, in which syntax, semantics, logical truth, and the observation/theory distinction may all be radically relative to one's conceptual scheme (neo-Kantianism), to semantical rules (logical positivism), to one's overall inferential dispositions (holism), to the way one dresses at parties (social pragmatism), or to anything else that tickles the philosophical fancy.

Then we extend insights from formal learning theory to our relativistic setting to prove structural necessary and sufficient conditions for the existence of a method that can reliably converge to the semantic status (true, false, or ill-formed) of a given string by means of conceptual revolutions. Reliable convergence is taken in the following, simple sense: no matter which possible truth-dependency is actual, and no matter which order data true of the scientist's current conceptual schemes arrives in, there must arise a time after which his guesses about the semantic status of the hypothesis under investigation are all correct for the conceptual schemes he elects to visit (see the technical section for a precise definition).

The necessity side of the argument is a sequential generalization of Blum and Blum's locking sequence lemma (c.f. Osherson, Stob and Weinstein 1986, p. 25) which embodies a relativistic generalization of the Putnam-Gold diagonal arguments (Putnam 1975) (Gold 1967).

The sufficiency side of the argument amounts to the construction of a universal, relativistic method for getting to the truth about a given hypothesis. The method is universal in the following sense. If the relativistic philosopher will tell the method which dependencies are *possible* according to his theory, then the method will reliably get to the truth about a given string *if it is possible for any method to do so*. Actually, our result is somewhat weaker. The conditions are necessary and sufficient only under the hypothesis that there are at most finitely many conceptual schemes. Some analogous result holds when this condition is not met, but we haven't figured out what it is yet.

Our universal method directs changes during the course of inquiry in the parameters that its metaphysical theory deems relevant to truth, observability, and well-formedness.

So the method achieves its universal power by directing the adoption of "conceptual revolutions". To put it another way, the method actively adopts new formal systems, rather than remaining stuck in a single system throughout the course of inquiry.

We also prove that no method that eventually settles upon a given conceptual scheme is as powerful as our universal method, which is unrestricted in the number of revolutions it may have.

Aside from their technical interest to inductive theorists, these results have some more general, philosophical significance. First of all, they show that Kuhn's famous relativistic attack upon the intersubjectivity of particular theory choices in science does not add up, as Rorty would like, to the general demise of normative epistemology. Universal power confers some normative force to a method, and the investigation and recommendation of such methods is an eminently suitable task for normative epistemologists. Moreover, this task does not fall within the purview of empirical historiography, Rorty's heir apparent to normative epistemology (c.f. Rorty 1979 p. 272)..

Another philosophical consequence of our results is that a broadly logical attitude can be taken toward inductive methodology without begging any questions against relativism and holism. Positivism and methodological clarity are not the same thing. Our analysis of convergent relativism is not committed to *a priori* truth of any kind, to analyticity, to a fixed observation-theory distinction, to exhaustion of theoretical meaning by observation, to compositional semantics, or even to theory-neutral rules of syntax for the hypothesis language. Even its formal relation of "truth" can be construed in as "edifying" a manner as empirical adequacy or the community's disposition to concur.

A third consequence is that relativism does not trivialize the problem of induction. Radical subjectivists (e.g. Levi 1980 p. 8) hold that whatever you believe is true because you believe it, and this position certainly does trivialize the goal of getting to the truth in the limit. But radical subjectivism is just one, extreme form of relativism. Positivism, holism, social pragmatism, and the semantics of quantum logic are all theories of truth-dependency that can support inductive problems for which there is no possible reliable solution.

A fourth consequence is to lend partial support to P. K Feyerabend's (1975) vague dictum that "anything goes" in science. Since any method that is forced to converge to a particular conceptual scheme is less powerful than our universal method, which is not so restricted, there should be no enforcement of a particular regimen of "normal science" upon the scientific community. On the other hand, it does not follow that "any method goes", for some methods are more powerful than others, and *ceteris paribus*, it would make little sense to choose the less powerful method over the more powerful. Nor does it follow that "any revolution goes at any time". During the course of inquiry directed by a powerful method, one should have the revolutions the method directs, when it directs them.

Our extension of the theory of reliable inductive inference to a relativistic framework also provides a blueprint for extending the reliabilist theory of knowledge to relativistic settings. Indeed, if one believes that reliable belief formation is a necessary condition on knowledge, then our necessary conditions on reliable inductive inference amount to a kind of relativized transcendental deduction.

The thesis that data is theory laden has often been raised in a negative way, to criticize formal studies in epistemology and to encourage mass desertions of philosophers to historiography. But once one takes the structure of the issue seriously, it is less a reason to abandon formal methodology than an invitation to explore new and exciting territory. We can start to winnow out the solvable, relativistic inductive problems from the unsolvable ones. We can begin to look at which combinations of intuitive techniques suffice to yield universal inductive procedures. We can look at how difficult inductive inference becomes under different, standard theories of reference. We can look at inductive inference in non-standard logics, such as quantum logic, in which truth depends upon which experimental regimen one adopts. We can obtain a precise account of how conceptual revolutions assist convergence to the truth. And we can look at how standard methodological directives and theories of confirmation interact with reliability.

Finally, the approach of this paper has wider ramifications than relativism. Meaning shifts are just one way in which truth can depend upon what the experimenter does. In the social sciences, one often encounters systems in which a scientist's announced conclusion causes changes in the system he attempts to predict. For an extreme example of such a case, consider the effect of the federal reserve board chairman predicting a major stock market crash by the end of the week. Similar cases arise in such fields as clinical psychology, in which the presuppositions of earlier questions on a questionnaire can actually mould the personality the psychologist wishes to profile. Even though meaning shift and causal feedback are two very different ways to arrive at a relativized notion of truth, the approach taken in this paper provides a unified account of both types of inductive problems.

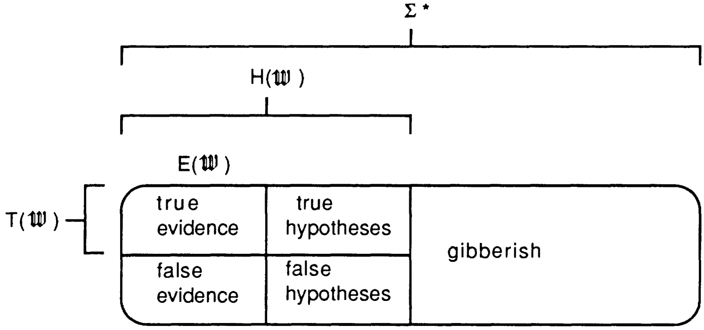
2. Appendix: Definitions and Results

Sequence Operations

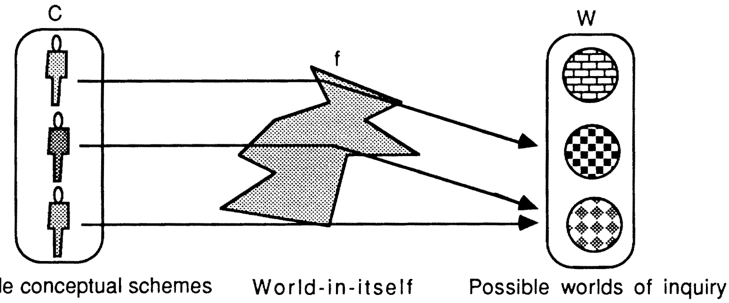
last(τ) = the object occurring at the end of finite sequence τ.
τ_n = the item occurring in position n of sequence τ.
$\tau[n]$ = the initial segment of sequence τ of length n.
τ- = the result of deleting the last item in finite sequence τ.
$\sigma * \tau$ is the concatenation of sequence τ and finite sequence σ.
rng(τ) = the set of all objects occurring in τ.

Metaphysics

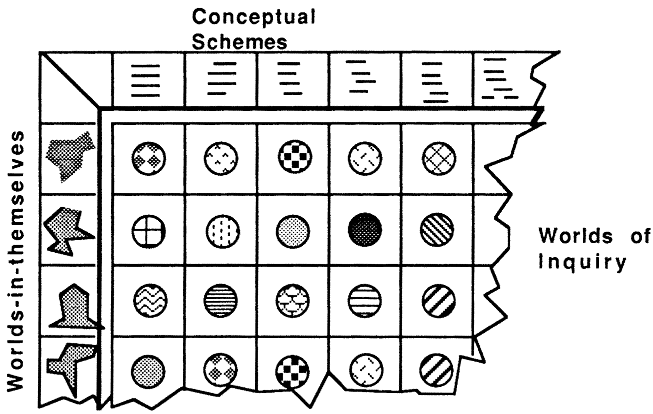
Let Σ be a countable alphabet
Let Σ^* = the set of all finite strings on A (including the empty string)
An *hypothesis* is some string $s \in \Sigma^*$.
A world-of-investigation $\mathfrak{W}$ is a triple of sets <H, E, T> such that T, E $\subseteq$ H.
H($\mathfrak{W}$) represents the *well-formed* hypotheses in $\mathfrak{W}$.
Note that a conjectured hypothesis need not be well-formed.
E($\mathfrak{W}$) represents the "evidence language" subset of H in $\mathfrak{W}$.
T($\mathfrak{W}$) represents the subset of H that is true in $\mathfrak{W}$.



W is the set of all worlds-of-investigation.
A *conceptual scheme* is an arbitrary point.
C = the set of all conceptual schemes.
A *world-in-itself* is a total map f: C $\rightarrow$ W.



A *relativistic system* is just some set F of worlds-in-themselves.
A relativistic system is *countable* $\Leftrightarrow$
F is countable, C is countable and W is countable.



The *truth values* are {T, F, U}.
Define tv: Σ^* X W $\rightarrow$ {T, F, U} as follows:
tv(s, $\mathfrak{W}$) = T if s $\in$ H($\mathfrak{W}$) $\cap$ T($\mathfrak{W}$).
tv(s, $\mathfrak{W}$) = F if s $\in$ H($\mathfrak{W}$) - T($\mathfrak{W}$).
tv(s, $\mathfrak{W}$) = U if s $\notin$ H($\mathfrak{W}$).
T means "true", F means "false" and U means "no truth value".

Data Presentations

a *data presentation* is an infinite sequence of strings in Σ^*.
SEQ = the set of all finite segments of data presentations.

Truth Detectors

A *truth detector* is a function δ: SEQ X $\Sigma^* \to$ C X {T, F, U}.

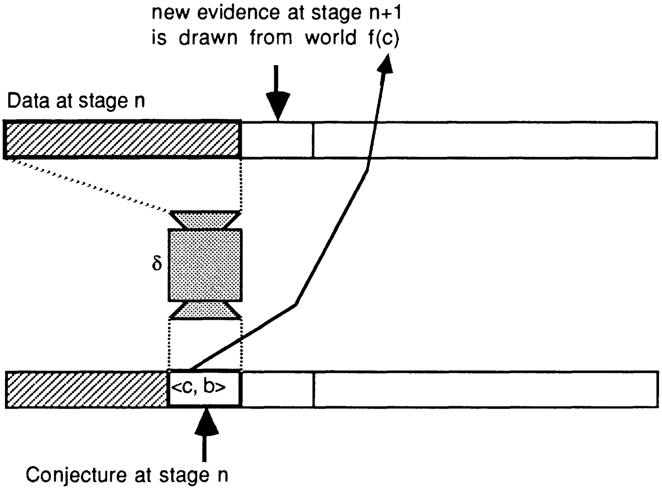
Data Presentations for Worlds and Detectors

$t[\delta, c, s]_n = t_k$ where k is the nth position in t such that $\delta(t[k-1], s)_1 = c$.
$t[\delta, c, s]$ = the sequence $<t[\delta, c, s]_1, t[\delta, c, s]_2, ..., t[\delta, c, s]_n, ...>$
$ev(\mathfrak{W}) = T(\mathfrak{W}) \cap E(\mathfrak{W})$
$\sigma \in$ SEQ is *sound* for δ, f, s $\Leftrightarrow \forall c$, $rng(\sigma[\delta, c, s]) \subseteq ev(f(c))$.
data presentation t is *sound for* δ, f, s $\Leftrightarrow \forall c \in C$ $rng(t[\delta, c, s]) \subseteq ev(f(c))$.

Note: soundness requires that the data read at stage n+1 be true with respect to the conceptual scheme produced by δ at stage n.

data presentation t is *complete for* δ, f, s $\Leftrightarrow$
$\forall c \in C$, $t[\delta, c]$ is infinite $\Rightarrow rng(t[\delta, c, s]) = ev(f(c))$.
data presentation t is *for* δ, f, s $\Leftrightarrow$ t is complete for δ, f, s and t is sound for δ, f, s.
PRES(f, δ, s) = {t: t is for δ, f, s}.

Relativistic Data Presentations



Reliable Detection

Let t $\in$ PRES(f, δ, s).
δ *converges* to c, b on t, s $\Leftrightarrow$
$\exists n\ \forall m > n\ \delta(t[m], s)_1 = c$ and $\delta(t[m], s)_2 = b$
δ *converges* to $\mathfrak{W}$, b on t, s $\Leftrightarrow$
$\exists n\ \forall m > n\ f(\delta(t[m], s)_1) = \mathfrak{W}$ and $\delta(t[m], s)_2 = b$.
δ *converges* to b on t, s $\Leftrightarrow$
$\exists n\ \forall m > n\ \delta(t[m], s)_2 = b$.
δ *detects* s on t $\Leftrightarrow$
$\exists n\ \forall m > n\ \delta(t[m], s)_2 = tv(s, f(\delta(t[m], s)_1))$.

Note: Detection requires that d produce a truth value that is correct for the conceptual scheme produced at the same moment as this truth value.

δ *detects* s in f $\Leftrightarrow$
  $\forall t \in$ PRES(δ, f), d [scheme-stably, world-stably, truth-stably] detects s on t.
δ is *reliable* over F $\Leftrightarrow$
  $\forall f \in$ F, δ [scheme-stably, world-stably, truth-stably] detects s in f.

Locking Sequence Lemma for Detection Simpliciter.

$\sigma \in$ SEQ is *locking* for δ, f, s $\Leftrightarrow$
  σ is sound for δ, f, s and
  $tv(s, f(\delta(s, \sigma)_1)) = \delta(s, \sigma)_2$ and
  $\forall \tau \in$ SEQ if
    (a) $\sigma \subseteq \tau$ and
    (b) τ is sound for δ, f, s
  then $tv(s, f(\delta(\tau, s)_1)) = \delta(\tau, s)_2$.

Locking Sequence Lemma:

If δ detects s in f then
  $\forall \tau$ sound for f, δ,
    $\exists s \in$ SEQ such that
    $\tau \subseteq \sigma$ and
    σ is locking for δ, f, s.

A Characterization of Truth Detection for Finite C and Countable Collections of Worlds-in-themselves

D is a *clue* $\Leftrightarrow$ D is a finite subset of C X Σ^*.
clue D is *sound for* f $\Leftrightarrow$ $\forall \langle c, e\rangle \in D, e \in ev(f(c))$.
clue D *involves* $c \in C \Leftrightarrow \exists e \in \Sigma^*$ such that $\langle c, e\rangle \in D$.
clue D is *contained in* s Œ SEQ with respect to d ¤
  $\forall \langle c, e\rangle \in D\ \exists n\ 1 \leq n \leq \text{length}(\sigma)$ and $\sigma_n = e$ and $c = \delta(\sigma[n], s)_1$.
Let F' $\subseteq$ F.
The *agreement zone* of F' mod s = $\{c: \forall f, g \in F'\ tv(s, f(c)) = tv(s, g(c))\}$.
$az_s(F)$ = the agreement zone of F mod s.
$az_s(\langle f_1, ..., f_n\rangle) = az_s(\{f_1, ..., f_n\})$.
Let f $\in$ F.
Let D be a clue.
D is a *clue for* f, s in F mod $\langle f_1, ..., f_n\rangle$ and $\langle D_1, ..., D_n\rangle \Leftrightarrow$
  (A) $D_1, ..., D_n$, D are sound for f and
  (B) $\forall f' \in F$
    if $D_1, D_2, ..., D_n$, D are sound for f'
    then
    (1)
    if $az_s(\langle f_1, ..., f_n, f, f'\rangle) = \varnothing$
    then $\exists c \in az_s(\langle f_1, ..., f_n, f\rangle)$
      $ev(f'(c)) - ev(f(c)) \neq \varnothing$ and
    (2)
    if
    (2. a) $az_s(\langle f_1, ..., f_n, f, f'\rangle) \neq \varnothing$ and
    (2. b) $\exists c \in az_s(\langle f_1, ..., f_n, f\rangle)$
      $tv(s, f(c)) \neq tv(s, f'(c))$ and

(2. c) $\forall c \in az_{\sigma}(<f_1, ..., f_n, f>)$, $ev(f'(c)) - ev(f(c)) = \varnothing$
then
$\exists D'$ such that D' is a clue for f' in F mod $<f_1, ..., f_n, f>$ and $<D_1, ..., D_n, D>$ and D' involves only schemes in $az_s(<f_1, ..., f_n, f>)$.

F is *sparse* mod s $\Leftrightarrow \forall f \in F$, f, s has a clue in F mod $<>$ and $<>$.

Theorem: If C is finite and F is countable then

s is detectable over F $\Leftrightarrow$ F is sparse mod s.

Corollary: The theorem holds for any set of truth-values of any cardinality.

Corollary: In the proof, a complete method for constructing scientists is given. This method has the property that if it is possible for some scientist to get to the relative truth in a given, relativistic system, the constructed scientist will do so.

References:

Gold, E.M. (1967), "Language Identification in the Limit". *Information and Control* 10: 447-474.

Kelly, K. and C. Glymour (1989), "Convergence to the Truth and Nothing but the Truth". *Philosophy of Science* 56:2.

Kuhn, T.S. (1970), *The Structure of Scientific Revolutions.* Chicago: Chicago University Press.

Levi, I. (1983), *The Enterprise of Knowledge.* Cambridge: MIT Press.

Osherson, D., Stob, M., and Weinstein, S. (1986), *Systems that Learn.* Cambridge: MIT Press.

Putnam, H. (1975), "Probability and Confirmation" in *Mathematics, Matter, and Method.* Cambridge: Cambridge University Press.

Rorty, R. (1979), *Philosophy and the Mirror of Nature.* Princeton: Princeton University Press.