Proceedings of the Conference on

HYPOTHETICAL REASONING

23–24 August 2014, Tübingen

Edited by

Thomas Piecha
and
Peter Schroeder-Heister

University of Tübingen
2015

Contents

Left to right, back to front:

Sergey Melikhov
Ernst Zimmermann Koji Mineshima
Youyou Cong Arnon Avron Guillaume Schlaepfer
Luca Tranchini · David Binder
Marine Gaudefroy-Bergmann Paul Egré
Reinhard Kahle · René Gazzari
Grigory Olkhovikov · Paolo Pistone
Erdinç Sayan · Nissim Francez
Andrzej Indrzejczak · Alberto Naibo
Kosta Došen
Peter Schroeder-Heister Francesca Poggiolesi
Torben Braüner

Preface

Hypothetical reasoning or reasoning under assumptions is a key concept of logic, philosophy of science and mathematics. The *Conference on Hypothetical Reasoning* focussed on its logical aspects, such as

- assumption-based calculi and their proof theory,
- logical consequence from a proof-theoretic or model-theoretic point of view,
- logics of conditionals,
- proof systems,
- structure of assumption-based proofs,
- hypotheses in proof-theoretic semantics,
- notions of implication,
- substructural logics,
- hypotheses in categorial logic,
- logical aspects of scientific explanation,
- hypothetical reasoning in mathematics,
- reasoning from definitions and axioms.

The conference took place 23–24 August, 2014 in Tübingen at the Department of Philosophy, in conjunction with ESSLLI 2014. The proceedings collect abstracts, slides and papers of the presentations given.

The conference and its proceedings were supported by the French-German ANR-DFG project "Hypothetical Reasoning: Its Proof-Theoretic Analysis" (HYPOTHESES), DFG grant Schr 275/16-2.

Thomas Piecha
Peter Schroeder-Heister

Programme

Saturday, 23 August

8.30–9.10	*Registration*
9.10–9.15	*Opening*
9.15–10.15	Francesca Poggiolesi: *Counterfactual logics: natural deduction calculi and sequent calculi*
10.15–10.45	*Coffee break*
10.45–11.15	Sergey Melikhov: *A joint logic of problems and propositions, a modified BHK-interpretation and proof-relevant topological models of intuitionistic logic*
11.15–11.45	Torben Braüner: *Seligman-style deduction for hybrid modal logic*
11.45–12.15	Grigory Olkhovikov: *Truth-value gaps and paradoxes of material implication*
12.15–14.15	*Lunch break* (Wirtshaus Lichtenstein)
14.15–15.15	Paul Egré: *Negating indicative conditionals*
15.15–15.45	*Coffee break*
15.45–16.15	Michael Cohen: *Explanatory Justice: The Case of Disjunctive Explanations*
16.15–17.15	Zoran Petrić: *Cuts and Graphs*
17.15–18.15	Kosta Došen: *An introduction to deduction*
20.00	*Conference dinner* (Wirtshaus Casino am Neckar)

Sunday, 24 August

9.00–10.00	Michel Bourdeau: *Comte's ≪ Théorie fondamentale des hypothèses ≫*
10.00–10.30	*Coffee break*
10.30–11.00	Erdinç Sayan: *How do vacuous truths become laws?*
11.00–11.30	Guillaume Schlaepfer: *Scientific modeling: a two layer based hypothetical reasoning*
11.30–12.00	Reinhard Kahle: *Axioms as Hypotheses*
12.00–14.00	*Lunch break* (Restaurant Mauganeschtle)
14.00–15.00	Andrzej Indrzejczak: *Hypersequents and Linear Time*
15.00–15.30	*Coffee break*
15.30–16.00	Nissim Francez: *On a dogma of Proof-Theoretic Semantics: generalising canonicity of derivations*
16.00–16.30	Paolo Pistone: *Rules, types and the transcendence of second order logic*
16.30–17.30	Arnon Avron: *Using Assumptions in Gentzen-type Systems*

Arnon Avron
School of Computer Science Department
Tel Aviv University

Using Assumptions in Gentzen-type Systems

Conference on Hypothetical Reasoning, August 2014

What is a Propositional Logic?

This is a pair $\langle \mathcal{L}, \vdash \rangle$, where $\mathcal{L}$ is a propositional language, and $\vdash$ is a relation between sets of formulas of $\mathcal{L}$ and formulas of $\mathcal{L}$ that satisfies:

Reflexivity: if $\varphi \in T$ then $T \vdash \varphi$.
Monotonicity: if $T \vdash \varphi$ and $T \subseteq T'$ then $T' \vdash \varphi$.
Transitivity: if $T \vdash \psi$ and $T, \psi \vdash \varphi$ then $T \vdash \varphi$.
Structurality: $T \vdash \varphi$ then $\sigma(T) \vdash \sigma(\varphi)$
Consistency $p \nvdash q$

Gentzen-style Proof Systems

- Hilbert-style systems operate on $\mathcal{L}$-formulas. Gentzen-style systems operate on *sequents*.
- A sequent: an expression of the form $\Gamma \Rightarrow \Delta$, where Γ, Δ are finite sets of $\mathcal{L}$-formulas.
- A standard Gentzen-type system for $\mathcal{L}$ consists of:
  1. Standard axioms: $\psi \Rightarrow \psi$.
  2. Structural Weakening and Cut rules:

$$\frac{\Gamma \Rightarrow \Delta}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\ (Weakening) \qquad \frac{\Gamma, \psi \Rightarrow \Delta \quad \Gamma \Rightarrow \Delta, \psi}{\Gamma \Rightarrow \Delta}\ (Cut)$$

  3. Logical introduction rules for the connectives of $\mathcal{L}$.

The Associated Consequence Relations

- Let Θ be a set of sequents, and $\Gamma \Rightarrow \Delta$ a sequent. $\Theta \vdash_G \Gamma \Rightarrow \Delta$ if there is a finite list of sequents whose last sequent is $\Gamma \Rightarrow \Delta$, and each sequent in this list is either an axiom of $\mathbf{G}$, or a member of Θ, or is obtained from previous sequents in the sequence by applying some rule of G.

The Associated Consequence Relations

- Let Θ be a set of sequents, and $\Gamma \Rightarrow \Delta$ a sequent.
  $\Theta \vdash_G \Gamma \Rightarrow \Delta$ if there is a finite list of sequents whose last sequent is $\Gamma \Rightarrow \Delta$, and each sequent in this list is either an axiom of $\mathbf{G}$, or a member of Θ, or is obtained from previous sequents in the sequence by applying some rule of G.
- $\mathbf{T} \vdash_G^t \psi$ if there is some finite $\Gamma \subseteq \mathbf{T}$, such that $\Gamma \Rightarrow \psi$ is provable in G.
- $\mathbf{T} \vdash_G^v \psi$ if there is some finite $\Gamma \subseteq \mathbf{T}$, such that $\{\Rightarrow \varphi \mid \varphi \in \Gamma\} \vdash_G \Rightarrow \psi$.

An Example: Modal Logics

In all Gentzen-type systems for the important modal logics:

$$\surd \; \frac{\Rightarrow \psi}{\Rightarrow \Box\psi} \qquad\qquad X \;\; \psi \Rightarrow \Box\psi$$

$$\downarrow \qquad\qquad\qquad \downarrow$$

$$\psi \vdash^v \Box\psi \qquad\qquad \psi \nvdash^t \Box\psi$$

Hence in all of them:

$$\vdash^v \qquad \neq \qquad \vdash^t$$

Sources and Processor: Belnap's Model

- A processor collects and processes information from a set of sources. Each source may provide the processor with information about atomic formulas. The information has the form of a truth-value in $\{1, 0, \mathbf{I}\}$.
- The processor assigns to an atom p a subset $d(p)$ of $\{0, 1\}$:
  - $1 \in d(p)$ iff some source has assigned 1 to p
  - $0 \in d(p)$ iff some source has assigned 0 to p

$\mathbf{t} = \{1\}$	- told to be true but not told to be false
$\mathbf{f} = \{0\}$	- told to be false but not told to be true
$\top = \{0, 1\}$	- told to be true and told to be false
$\bot = \emptyset$	- not told to be true and not told to be false

Belnap's Model (Continued)

Let $\mathcal{F} = \mathcal{F}_{\{\neg,\vee,\wedge\}}$.

The processor's valuation is extended to $\mathcal{F}$ as follows:

(db1) $0 \in d(\neg\varphi)$ iff $1 \in d(\varphi)$;

(db2) $1 \in d(\neg\varphi)$ iff $0 \in d(\varphi)$;

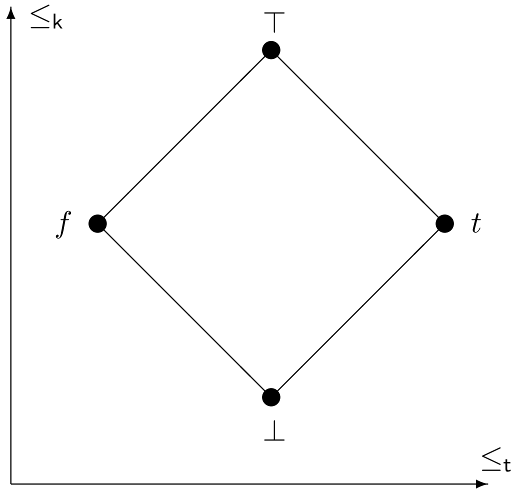
(db3) $1 \in d(\varphi \vee \psi)$ iff $1 \in d(\varphi)$ or $1 \in d(\psi)$;

(db4) $0 \in d(\varphi \vee \psi)$ iff $0 \in d(\varphi)$ and $0 \in d(\psi)$

(db5) $1 \in d(\varphi \wedge \psi)$ iff $1 \in d(\varphi)$ and $1 \in d(\psi)$;

(db6) $0 \in d(\varphi \wedge \psi)$ iff $0 \in d(\varphi)$ or $0 \in d(\psi)$.

The Basic Bilattice $\mathcal{FOUR}$



More General Model

- A source may provide information about any formula.
- The processor valuation d is the minimal function from $\mathcal{F}$ to $\mathcal{P}(\{0, 1\})$ which satisfies:

(d0) For $x \in \{0, 1\}$, $x \in d(\varphi)$ if $s(\varphi) = x$ for some source s.

(d1) $0 \in d(\neg\varphi)$ iff $1 \in d(\varphi)$;

(d2) $1 \in d(\neg\varphi)$ iff $0 \in d(\varphi)$;

(d3) $1 \in d(\varphi \vee \psi)$ if $1 \in d(\varphi)$ or $1 \in d(\psi)$;

(d4) $0 \in d(\varphi \vee \psi)$ iff $0 \in d(\varphi)$ and $0 \in d(\psi)$;

(d5) $1 \in d(\varphi \wedge \psi)$ iff $1 \in d(\varphi)$ and $1 \in d(\psi)$;

(d6) $0 \in d(\varphi \wedge \psi)$ if $0 \in d(\varphi)$ or $0 \in d(\psi)$.

Non-deterministic Matrices

A non-deterministic matrix (Nmatrix) for $\mathcal{L}$ is a tuple $\mathcal{M} = \langle \mathcal{V}, \mathcal{D}, \mathcal{O} \rangle$:

- $\mathcal{V}$ - the set of truth-values,
- $\mathcal{D}$ - the set of designated truth-values,
- $\mathcal{O}$ - contains an interpretation function $\tilde{\diamond} : \mathcal{V}^n \to P^+(\mathcal{V})$ for every n-ary connective $\diamond$ of $\mathcal{L}$.

A (legal) valuation in an Nmatrix $\mathcal{M} = \langle \mathcal{V}, \mathcal{D}, \mathcal{O} \rangle$ is any function $v : \mathcal{F}_{\mathcal{L}} \to \mathcal{V}$ which satisfies:

$$v(\diamond(\psi_1, ..., \psi_n)) \in \tilde{\diamond}(v(\psi_1), ..., v(\psi_n))$$

In ordinary matrices: each $\tilde{\diamond}$ returns only singletons.

Logics Induced by Nmatrices

- A valuation v in an Nmatrix $\mathcal{M}$ is a model of:
  - a formula ψ ($v \models^{\mathcal{M}} \psi$) if $v(\psi) \in \mathcal{D}$.
  - a theory $T \subseteq \mathcal{F}_{\mathcal{L}}$ ($v \models^{\mathcal{M}} T$) if $v \models^{\mathcal{M}} \psi$ for all $\psi \in T$.
- The formula consequence relation induced by $\mathcal{M}$ is the relation $\vdash_{\mathcal{M}}$ on $\mathcal{P}(\mathcal{F}_{\mathcal{L}}) \times \mathcal{F}_{\mathcal{L}}$ such that $T \vdash_{\mathcal{M}} \varphi$ if every model of T in $\mathcal{M}$ is also a model of φ.

THEOREM: For any finite Nmatrix $\mathcal{M}$ for $\mathcal{L}$, $\mathbf{L}_{\mathcal{M}} = \langle \mathcal{L}, \vdash_{\mathcal{M}} \rangle$ is a decidable and finitary propositional logic.

Characterization of a Reasonable Processor

A function $d : \mathcal{F} \to \{\mathbf{f}, \bot, \top, \mathbf{t}\}$ is an ESP processor valuation iff it is an $\mathcal{M}_I^4$-valuation, where $\mathcal{M}_I^4 = \langle \mathcal{V}, \mathcal{D}, \mathcal{O} \rangle$ is the Nmatrix:

$$\mathcal{V} = \{\mathbf{f}, \bot, \top, \mathbf{t}\}, \mathcal{D} = \{\top, \mathbf{t}\}, \mathcal{O} = \{\tilde{\neg}, \tilde{\vee}, \tilde{\wedge}\}$$

where the interpretations of the connectives are given by:

$\tilde{\neg}$	
$\mathbf{f}$	$\{\mathbf{t}\}$
$\bot$	$\{\bot\}$
$\top$	$\{\top\}$
$\mathbf{t}$	$\{\mathbf{f}\}$

$\tilde{\vee}$	$\mathbf{f}$	$\bot$	$\top$	$\mathbf{t}$
$\mathbf{f}$	$\{\mathbf{f}, \top\}$	$\{\mathbf{t}, \bot\}$	$\{\top\}$	$\{\mathbf{t}\}$
$\bot$	$\{\mathbf{t}, \bot\}$	$\{\mathbf{t}, \bot\}$	$\{\mathbf{t}\}$	$\{\mathbf{t}\}$
$\top$	$\{\top\}$	$\{\mathbf{t}\}$	$\{\top\}$	$\{\mathbf{t}\}$
$\mathbf{t}$	$\{\mathbf{t}\}$	$\{\mathbf{t}\}$	$\{\mathbf{t}\}$	$\{\mathbf{t}\}$

$\tilde{\wedge}$	$\mathbf{f}$	$\bot$	$\top$	$\mathbf{t}$
$\mathbf{f}$	$\{\mathbf{f}\}$	$\{\mathbf{f}\}$	$\{\mathbf{f}\}$	$\{\mathbf{f}\}$
$\bot$	$\{\mathbf{f}\}$	$\{\mathbf{f}, \bot\}$	$\{\mathbf{f}\}$	$\{\mathbf{f}, \bot\}$
$\top$	$\{\mathbf{f}\}$	$\{\mathbf{f}\}$	$\{\top\}$	$\{\top\}$
$\mathbf{t}$	$\{\mathbf{f}\}$	$\{\mathbf{f}, \bot\}$	$\{\top\}$	$\{\mathbf{t}, \top\}$

The Need for Sequents

The expressive power of the formulas in $\{\neg, \vee, \wedge\}$ is too weak:

- We cannot fully express negative information (e.g., that $1 \notin d(\varphi)$).
- We cannot fully express disjunctive information (e.g., that either $1 \in d(\varphi)$ or $1 \in d(\psi)$).
- The language provides no implication connective corresponding to the intended consequence relation.

To compensate for this, we reasons with sequents. Given an ESP structure $\langle S, d \rangle$, a sequent $\varphi_1, \ldots, \varphi_n \Rightarrow \psi_1, \ldots, \psi_k$ expresses the information that either $1 \notin d(\varphi_1)$, or $1 \notin d(\varphi_2)$, or $\ldots$ or $1 \notin d(\varphi_n)$, or $1 \in d(\psi_1)$, or $\ldots$ or $1 \in d(\psi_k)$.

Sequents and (N)matrices

- A valuation v in an Nmatrix $\mathcal{M}$ is a model of a sequent $\Gamma \Rightarrow \Delta$ ($v \models^{\mathcal{M}} \Gamma \Rightarrow \Delta$) iff either $v \models^{\mathcal{M}} \psi$ for some $\psi \in \Delta$, or $v \not\models^{\mathcal{M}} \psi$ for some $\psi \in \Gamma$.
- The sequent consequence relation induced by $\mathcal{M}$ is the relation $\vdash^{s}_{\mathcal{M}}$ on $\mathcal{P}(Seq_{\mathcal{L}}) \times Seq_{\mathcal{L}}$ defined by: $\Theta \vdash^{s}_{\mathcal{M}} \Gamma \Rightarrow \Delta$ iff $v \models^{\mathcal{M}} \Gamma \Rightarrow \Delta$ whenever $v \models^{\mathcal{M}} \Theta$.
- $T \vdash_{\mathcal{M}} \varphi$ iff $\{\Rightarrow \psi \mid \psi \in T\} \vdash^{s}_{\mathcal{M}} (\Rightarrow \varphi)$.
  If T is finite, then $T \vdash_{\mathcal{M}} \varphi$ iff $\vdash^{s}_{\mathcal{M}} (T \Rightarrow \varphi)$.

Systems For Reasoning with Sources

[¬¬⇒] $\dfrac{\Gamma, \psi \Rightarrow \Delta}{\Gamma, \neg\neg\psi \Rightarrow \Delta}$ [⇒¬¬] $\dfrac{\Gamma \Rightarrow \Delta, \psi}{\Gamma \Rightarrow \Delta, \neg\neg\psi}$

[∧⇒] $\dfrac{\Gamma, \psi, \varphi \Rightarrow \Delta}{\Gamma, \psi \wedge \varphi \Rightarrow \Delta}$ [⇒∧] $\dfrac{\Gamma \Rightarrow \Delta, \psi \quad \Gamma \Rightarrow \Delta, \varphi}{\Gamma \Rightarrow \Delta, \psi \wedge \varphi}$

[¬∧⇒] $\dfrac{\Gamma, \neg\psi \Rightarrow \Delta \quad \Gamma, \neg\varphi \Rightarrow \Delta}{\Gamma, \neg(\psi \wedge \varphi) \Rightarrow \Delta}$ [⇒¬∧] $\dfrac{\Gamma \Rightarrow \Delta, \neg\psi, \neg\varphi}{\Gamma \Rightarrow \Delta, \neg(\psi \wedge \varphi)}$

[∨⇒] $\dfrac{\Gamma, \psi \Rightarrow \Delta \quad \Gamma, \varphi \Rightarrow \Delta}{\Gamma, \psi \vee \varphi \Rightarrow \Delta}$ [⇒∨] $\dfrac{\Gamma \Rightarrow \Delta, \psi, \varphi}{\Gamma \Rightarrow \Delta, \psi \vee \varphi}$

[¬∨⇒] $\dfrac{\Gamma, \neg\psi, \neg\varphi \Rightarrow \Delta}{\Gamma, \neg(\psi \vee \varphi) \Rightarrow \Delta}$ [⇒¬∨] $\dfrac{\Gamma \Rightarrow \Delta, \neg\psi \quad \Gamma \Rightarrow \Delta, \neg\varphi}{\Gamma \Rightarrow \Delta, \neg(\psi \vee \varphi)}$

What is a Canonical Rule?

- An "ideal" logical rule: an introduction rule for *exactly one connective*, on *exactly one side of a sequent*.
- In its formulation: *exactly one occurrence* of the introduced connective, no other occurrences of other connectives.
- The rule should also be *pure* (i.e. context-independent): no side conditions limiting its application.
- Its active formulas: *immediate subformulas* of its principal formula.

What is a Canonical Rule? (Continued)

Stage 1.

$$\frac{\Gamma, \psi, \varphi \Rightarrow \Delta}{\Gamma, \psi \wedge \varphi \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow \Delta, \psi \quad \Gamma \Rightarrow \Delta, \varphi}{\Gamma \Rightarrow \Delta, \psi \wedge \varphi}$$

Stage 2.

$$\frac{\psi, \varphi \Rightarrow}{\psi \wedge \varphi \Rightarrow} \qquad \frac{\Rightarrow \psi \quad \Rightarrow \varphi}{\Rightarrow \psi \wedge \varphi}$$

Stage 3.

$$\{p_1, p_2 \Rightarrow\}/p_1 \wedge p_2 \Rightarrow \qquad \{\Rightarrow p_1 \ ; \ \Rightarrow p_2\}/ \Rightarrow p_1 \wedge p_2$$

Canonical Systems

- *A clause*: a sequent consisting of atomic formulas.
- A *canonical rule* has one of the forms:

$$\{\Pi_i \Rightarrow \Sigma_i\}_{1 \leq i \leq m} / \diamond (p_1, ..., p_n) \Rightarrow$$

$$\{\Pi_i \Rightarrow \Sigma_i\}_{1 \leq i \leq m} / \Rightarrow \diamond (p_1, ..., p_n)$$

where $m \geq 0$ *and for all* $1 \leq i \leq m$*:* $\Pi_i \Rightarrow \Sigma_i$ *is a clause over* $\{p_1, ..., p_n\}$.

Applications of Canonical Rules

The form of an additive application of the canonical rule $\{\Pi_i \Rightarrow \Sigma_i\}_{1 \leq i \leq m} / \diamond (p_1, ..., p_n) \Rightarrow$:

$$\frac{\{\Gamma, \Pi_i^* \Rightarrow \Delta, \Sigma_i^*\}_{1 \leq i \leq m}}{\Gamma, \diamond(\psi_1, ..., \psi_n) \Rightarrow \Delta}$$

Here Π_i^* and Σ_i^* are obtained from Π_i and Σ_i (respectively) by substituting ψ_j for p_j for all $1 \leq j \leq n$, and Γ, Δ are any finite sets of formulas (the context).

Example 1

Implication rules:

$$\{p_1 \Rightarrow p_2\} \;/\; \Rightarrow p_1 \supset p_2 \qquad \{\Rightarrow p_1 \;;\; p_2 \Rightarrow\} \;/\; p_1 \supset p_2 \Rightarrow$$

Their applications:

$$\frac{\Gamma, \psi \Rightarrow \Delta, \varphi}{\Gamma \Rightarrow \Delta, \psi \supset \varphi} \qquad \frac{\Gamma \Rightarrow \Delta, \psi \quad \Gamma, \varphi \Rightarrow \Delta}{\Gamma, \psi \supset \varphi \Rightarrow \Delta}$$

Example 2

Quasi-implication rules:

$$\{\Rightarrow p_1 \;;\; p_2 \Rightarrow\} \;/\; p_1 \rightsquigarrow p_2 \Rightarrow \qquad \{\Rightarrow p_2\} \;/\; \Rightarrow p_1 \rightsquigarrow p_2$$

Their applications:

$$\frac{\Gamma \Rightarrow \Delta, \psi \quad \Gamma, \varphi \Rightarrow \Delta}{\Gamma, \psi \rightsquigarrow \varphi \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow \Delta, \varphi}{\Gamma \Rightarrow \Delta, \psi \rightsquigarrow \varphi}$$

Example 3

"Tonk" rules:

$$\{p_2 \Rightarrow\} \;/\; p_1 T p_2 \Rightarrow \qquad \{\Rightarrow p_1\} \;/\; \Rightarrow p_1 T p_2$$

Their applications:

$$\frac{\Gamma, \psi \Rightarrow \Delta}{\Gamma, \varphi T \psi \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow \Delta, \varphi}{\Gamma \Rightarrow \Delta, \varphi T \psi}$$

What Sets of Rules are Acceptable?

- A standard Gentzen-type system is canonical if each of its logical (i.e. non-structural) rules is canonical.
- If G is a canonical system, then $\vdash_G$ is a structural and finitary tcr. But is it a logic? i.e., is it also consistent?

Coherence

- A canonical calculus G is *coherent* if for every pair of rules $\Theta_1 / \Rightarrow \diamond(p_1, ..., p_n)$ and $\Theta_2 / \diamond(p_1, ..., p_n) \Rightarrow$, the set of clauses $\Theta_1 \cup \Theta_2$ is classically unsatisfiable (and so inconsistent, i.e., the empty sequent can be derived from it using only cuts)
- For a canonical calculus G, $\vdash_G$ is a logic iff G is coherent.

Coherent Calculi:

$$\{p_1 \Rightarrow\} \ / \ \Rightarrow \neg p_1 \qquad \{\Rightarrow p_1\} \ / \ \neg p_1 \Rightarrow$$

$$\{p_1, p_2 \Rightarrow\} \ / \ p_1 \wedge p_2 \Rightarrow \qquad \{\Rightarrow p_1 \ ; \ \Rightarrow p_2\} \ / \ \Rightarrow p_1 \wedge p_2$$

$$\{p_1 \Rightarrow p_2\} \ / \ \Rightarrow p_1 \supset p_2 \qquad \{\Rightarrow p_1 \ ; \ p_2 \Rightarrow\} \ / \ p_1 \supset p_2 \Rightarrow$$

$$\{\Rightarrow p_1 \ ; \ p_2 \Rightarrow\} \ / \ p_1 \leadsto p_2 \Rightarrow \qquad \{\Rightarrow p_2\} \ / \ \Rightarrow p_1 \leadsto p_2$$

Non-coherent: “Tonk”!

$$\{p_2 \Rightarrow\} \ / \ p_1 T p_2 \Rightarrow \qquad \{\Rightarrow p_1\} \ / \ \Rightarrow p_1 T p_2$$

From these rules, we can derive $p \Rightarrow q$ for any p, q:

$$\dfrac{\dfrac{p \Rightarrow p}{p \Rightarrow pTq} \qquad \dfrac{q \Rightarrow q}{pTq \Rightarrow q}}{p \Rightarrow q}$$

Exact Correspondence

Theorem: If G is a canonical calculus, then the following statements are equivalent:

1. $\vdash_G$ is consistent (and so it induces a logic).
2. G is coherent.
3. G has a characteristic two-valued Nmatrix.
4. G admits cut-elimination.
5. G admits strong cut-elimination.

Strong Cut Elimination (Propositional Case)

A Gentzen-type system G admits Strong cut-elimination if has the following property:

$\{\Gamma_i \Rightarrow \Delta_i \mid i \in I\} \vdash_G \Gamma \Rightarrow \Delta$ iff there is a proof of $\Gamma \Rightarrow \Delta$ from $\{\Gamma_i \Rightarrow \Delta_i \mid i \in I\}$ in which all cuts are made on formulas in $\bigcup_{i \in I}(\Gamma_i \cup \Delta_i)$

(In particular: $\vdash_G \Gamma \Rightarrow \Delta$ iff $\Gamma \Rightarrow \Delta$ has a cut-free proof in G).

Usually, if a system admits cut-elimination, it admits also strong cut-elimination.

Applications: subformula property, decidability, proof search.

How Can We Prove Strong Cut Elimination?

- Prove ordinary cut-elimination. Then prove the strong cut-elimination by induction on the number of premises.

  This works fine if the system is Pure (and closed under weakening).

- Use some version of Gentzen's syntactic proof for LK and LJ.

- Use semantic methods.

An Example: The Provability Logic GL

H_{GL}**:** The Hilbert System for K (or $K4$) together with:

$$\mathrm{L} \qquad \Box(\Box\varphi \to \varphi) \to \Box\varphi$$

G_{GL}**:** Gentzen System for propositional classical logic and:

$$\frac{\Box\varphi, \Box\Gamma', \Gamma \Rightarrow \varphi}{\Box\Gamma', \Box\Gamma \Rightarrow \Box\varphi} \qquad \left(\frac{\Box\varphi, \Box\Gamma, \Gamma \Rightarrow \varphi}{\Box\Gamma \Rightarrow \Box\varphi}\right)$$

Semantics: Irreflexive, transitive and finite Kripke frames.

Semantics of Propositional GL

$F = \langle W_F, R, V\rangle$

- W_F – A nonempty finite set
- R – An irreflexive and transitive binary relation on W_F.
- $V : W_F \times \mathcal{F}_{\mathcal{L}} \to \{t, f\}$

  (i) $V(a, \varphi \to \psi) = t$ iff $v(a, \varphi) = f$ or $v(a, \psi) = t$

  (ii) $V(a, \bot) = f$

  (iii) $V(a, \Box\varphi) = t$ iff $\forall b(aRb \Rightarrow v(b, \varphi) = t)$

φ is true in $a \in W_F$ iff $v(a, \varphi) = t$

φ is valid in F iff φ is true in every $a \in W_F$

Semantics of Propositional GL (Continued)

- A $t-$model of φ is a pair $\langle F, a\rangle$ s.t. F is as above, $a \in W_F$ and φ is true in a.
- A $v-$model of φ is a triple F as above in which φ is valid.
- A $t-$model of $\Gamma \Rightarrow \Delta$ is a pair $\langle F, a\rangle$ as above s.t. ψ is false in a for some $\psi \in \Gamma$ or φ is true in a for some $\varphi \in \Delta$.
- A $v-$model of $\Gamma \Rightarrow \Delta$ is a triple F as above s.t. $\langle F, a\rangle$ is a $t-$model of $\Gamma \Rightarrow \Delta$ for every $a \in W_F$

Main Results about G_{GL}

1. G_{GL} admits strong cut-elimination.
2. Let Θ be a set of sequents, and $\Gamma \Rightarrow \Delta$ a sequent.
   $\Theta \vdash_{\mathrm{G}_{GL}} \Gamma \Rightarrow \Delta$ iff every $v-$model of Θ is a $v-$model of $\Gamma \Rightarrow \Delta$.

Corollaries:

- $\mathbf{T} \vdash^t_{\mathrm{G}_{GL}} \psi$ iff every $t-$model of $\mathbf{T}$ is a $t-$model of ψ.
- $\mathbf{T} \vdash^v_{\mathrm{G}_{GL}} \psi$ iff every $v-$model of $\mathbf{T}$ is a $v-$model of ψ.

Proof of the Main Results about G_{GL}

Let $\Theta = \{\Gamma_1 \Rightarrow \Delta_1, \ldots, \Gamma_n \Rightarrow \Delta_n\}$,
$S_\Theta = \Gamma_1 \cup \Gamma_2 \cup \cdots \cup \Gamma_n \cup \Delta_1 \cup \cdots \Delta_n$.

A good proof of $\Gamma \Rightarrow \Delta$ from Θ is a proof in which all cuts are on formula in S_Θ.

A sequent $\Gamma' \Rightarrow \Delta'$ is called saturated if:

(1) It consists of subformulas of $S_\Theta \cup \Gamma \cup \Delta$

(2) $\Gamma' \Rightarrow \Delta'$ has no good proof.

(3) If $\varphi \to \psi \in \Delta'$ then $\varphi \in \Gamma'$ and $\psi \in \Delta'$

(4) If $\varphi \to \psi \in \Gamma'$ then $\varphi \in \Delta'$ or $\psi \in \Gamma'$
$\ldots$

(*) If $\varphi \in S_\Theta$ then $\varphi \in \Gamma' \cup \Delta'$

The Counter-model

Suppose that $\Gamma \Rightarrow \Delta$ has no good proof from Θ.

W: The set of saturated sequents

R: $(\Sigma_1 \Rightarrow \Pi_1) R (\Sigma_2 \Rightarrow \Pi_2)$ if:

(i) If $\Box\varphi \in \Sigma_1$ then $\{\Box\varphi, \varphi\} \subseteq \Sigma_2$

(ii) There is at least one sentence of the form $\Box\psi$ in $\Sigma_2 - \Sigma_1$

V: $V(\Gamma' \Rightarrow \Delta', p) = \begin{cases} t & p \in \Gamma' \\ f & \text{otherwise} \end{cases}$

Main Lemmas about the Counter-model

(1) If $\varphi \in \Gamma'$ then $v(\Gamma' \Rightarrow \Delta', \varphi) = t$
If $\varphi \in \Delta'$ then $v(\Gamma' \Rightarrow \Delta', \varphi) = f$

(2) R is transitive and irreflexive

(3) W is finite

(4) $\Gamma \Rightarrow \Delta$ is not valid in $\langle W, R, V \rangle$

(*) $\Gamma_i \Rightarrow \Delta_i$ $(i = 1, \ldots, n)$ are valid in $\langle W, R, V \rangle$

Proof of (*): If $\Gamma' \Rightarrow \Delta' \in W$ then $\Gamma_i \cup \Delta_i \subseteq \Gamma' \cup \Delta'$. It is impossible that both $\Gamma_i \subseteq \Gamma'$ and $\Delta_i \subseteq \Delta'$. Hence (*) follows from (1).

FOL: The Case of LK

$$(\forall \Rightarrow) \quad \frac{\Gamma, \varphi(t/x) \Rightarrow \Delta}{\Gamma, \forall x\varphi \Rightarrow \Delta} \qquad {}^{(*)} \frac{\Gamma \Rightarrow \Delta, \varphi}{\Gamma \Rightarrow \Delta \quad \forall y\varphi(y/x)} \quad (\Rightarrow \forall)$$

$$^{(*)}(\exists \Rightarrow) \quad \frac{\Gamma, \varphi \Rightarrow \Delta}{\Gamma, \exists y\varphi(y/x) \Rightarrow \Delta} \qquad \frac{\Gamma \Rightarrow \Delta, \varphi(t/x)}{\Gamma \Rightarrow \Delta, \exists x\varphi} \quad (\Rightarrow \exists)$$

The starred rules are impure. Therefore $\vdash^v \neq \vdash^t$

$$\surd \ \frac{\Rightarrow \psi}{\Rightarrow \forall x\psi} \qquad\qquad X \ \ \psi \Rightarrow \forall x\psi$$

$$\downarrow \qquad\qquad\qquad \downarrow$$

$$\psi \vdash^v \forall x\psi \qquad\qquad \psi \not\vdash^t \forall x\psi$$

FOL: The Case of LK (Continued)

The set $\{(\Rightarrow p(x)), (p(a) \Rightarrow)\}$ is inconsistent:

$$(\text{Cut})\dfrac{(\Rightarrow \forall)\dfrac{\Rightarrow p(x)}{\Rightarrow \forall x p(x)} \quad (\forall \Rightarrow)\dfrac{p(a) \Rightarrow}{\forall x p(x) \Rightarrow}}{\Rightarrow}$$

However, there is no way to derive $\Rightarrow$ from this inconsistent set without the logical rules for the quantifiers.

The strong cut-elimination theorem for LK

Let LKS be LK+ the rule of substitution of terms for free variables:

$$\frac{\Gamma \Rightarrow \Delta}{\Gamma\{t/x\} \Rightarrow \Delta\{t/x\}}$$

Let S be a set of sequents and suppose $S \vdash_{\mathrm{LKS}} \Gamma \Rightarrow \Delta$.

Then there is a proof in LKS of $\Gamma \Rightarrow \Delta$ from S in which the substitution rule is applied only to sequents in S, and all cuts are on instances of formulae which occur in sequents of S.

Outline of a Syntactic Proof

First we use Gentzen's method to show that all cuts which are not on instances of formulae which occur in S are eliminable.

The substitution rule is needed at this stage when the cut formula begins with a quantifier and we want to apply the induction hypothesis to immediate subformulae of it.

Then we use induction on length of such proofs to show that all applications of the substitution rule can be done on sequents in S.

The Basis of Resolution

Let S be a set of clauses and s a clause such that $S \vdash s$. Let S' be a subset of S with the following properties:

- (1) S' contains no tautology.
- (2) Each element in S is subsumed by some element of S'.
- (3) No element in S' is subsumed by another element of S'.

Then there exists a clause s' which subsumed s so that s' can be inferred from S' using only substitutions and cuts.

Corollary: If S is an inconsistent set of clauses then there is a finite set S' of instances of clauses in S from which $\Rightarrow$ can be derived using only cuts.

La théorie positive des hypothèses

Michel Bourdeau

Institut d'histoire et de philosophie des sciences et des techniques
(CNRS, Paris 1, ENS)

Abstract

In spite of its importance (it triggered the interest for the topic in the nineteenth century), Comte's theory of hypothesis has not received the attention it deserves, and there has been no in-depth study of it since a paper by Laudan in 1971.

The argument takes its starting point in the fact that there is more in our ways of reasoning than induction and deduction. Science could not progress without scientists resorting to hypothesis, that is, moving away from data, assuming some phenomenon, drawing consequences from it and asking if they agree or not with facts. Then, the question is: what kind of hypothesis is admissible, what kind is not?

Comte's answer is often dismissed as verificationist but there is much more to say about it. When he spoke about unverifiable hypothesis, he had in mind instances like phlogiston, calorific or "éther luminifère", which were still quite common in his time, and he relied on the pioneer thermodynamical works of Fourier, who precisely began by rejecting this very kind of hypothesis; but the theory has also to be understood in relation to his anti-metaphysical stance. The influence of Comte's theory can be seen in the work of Duhem but mainly in Peirce's theory of abduction.

Au début du dix-neuvième siècle, l'idée que les sciences expérimentales seraient des sciences inductives était encore très répandue. La théorie des hypothèses exposée dans la vingt-huitième leçon du *Cours de philosophie positive* (1830-1842)[1] constitue donc une contribution novatrice, puisque son point de départ se trouve dans le refus d'une telle position : l'induction ne suffit pas, il est impossible de s'en tenir aux seules données et le savant n'a d'autre issue que de prendre les devants. Bien plus, dans l'économie générale de la philosophie comtienne des sciences, l'exposé canonique de cette théorie, qui occupe la fin de la vingt-huitième leçon, constitue un dispositif central, polémique et problématique[2].

[1] Pour les œuvres de Comte, les abréviations suivantes seront utilisées :

- *C* *Cours de philosophie positive* (1830-1842). On renvoie à l'édition Hermann, 2 volumes, Paris, 1975 (1er vol., leçons 1-45 ; 2e vol., leçons 46-60) ; pour les leçons 46-51, nous renvoyons à la nouvelle édition en cours, Paris, Hermann, 2012.
- *TPAP* *Traité philosophique d'astronomie populaire* (1844), Paris, Fayard, 1985.
- *S* *Système de politique positive* (1851-1854), Paris, L. Mathias, 4 volumes. On donne le volume puis la page.

[2] Parmi les études consacrées à la question, il en est deux de remarquables, que le lecteur est instamment invité à consulter :

- Laudan, L. : ≪ Towards a Reassessment of Comte's "Méthode Positive." ≫, *Philosophy of Science* 38 (1971), 35-53 ; repris dans *Science and Hypothesis: Historical Essays on Scientific Methodology*. Dordrecht, D. Reidel 1981 ;
- Bachelard, G. : *Etude sur l'évolution d'un problème de physique*, Paris, Vrin 1973 (la première édition est de 1927), chap. 4, 55-72.

Le dispositif est central pour les sciences expérimentales, le recours à l'hypothèse y jouant un rôle capital ; mais il est aussi central pour la philosophie positive, ce qui ne doit pas surprendre, les sciences expérimentales servant en quelque sorte de modèle de la positivité. En ce sens, la théorie des hypothèses est inséparable d'une réflexion plus générale sur la nature de la positivité, ce qui ne va pas sans créer d'assez sérieuses difficultés.

Le caractère polémique de la théorie résulte directement de cette situation. Si Comte parle des hypothèses, c'est en raison de leur fécondité : leur usage est rigoureusement indispensable en science. Mais cet aspect pourtant primordial passe vite au second plan. Une fois la porte ouverte, il convient de ne pas laisser passer n'importe quoi ; aussi le plus clair de la discussion consistera à déterminer dans quelles conditions une hypothèse peut être tenue pour recevable, c'est-à-dire positive. La théorie prend ainsi la forme d'une critique de certains types d'hypothèses, et les partisans de ces dernières ne se sont pas privés de contre attaquer, accusant Comte de poser des limitations arbitraires à l'activité du savant. Une bonne partie des objections adressées à la philosophie positiviste des sciences s'est ainsi cristallisée autour de la théorie des hypothèses, présentée comme une des meilleures marques de son étroitesse d'esprit). Pour les adversaires du positivisme, tout se passe comme si critiquer la métaphysique revenait *ipso facto* à être scientiste, et la question est de savoir s'ils ont raison de nous enfermer dans une telle alternative.

Si l'on cherche maintenant à se faire une première idée de ce que dit cette théorie, il suffira à ce stade de retenir qu'elle présente deux versants, positif et négatif. Le premier consiste à tenir l'hypothèse comme une simple anticipation sur l'expérience future. C'est un artifice essentiellement provisoire et à ce titre éliminable. Le versant négatif consiste alors à exclure les hypothèses qui ne satisfont pas à ces exigences, et elles sont nombreuses. Les hypothèses métaphysiques, bien sûr ; mais Comte estime que, sur ce point, la bataille est à peu près gagnée. Si donc la théorie est l'occasion d'un approfondissement de la notion de positivité, c'est qu'elle conduit à exclure également les succédanés de métaphysique qu'a produit ce que Comte appelle un positivisme incomplet ou un positivisme bâtard.

Comte lui-même reconnaissait que la fonction fondamentale des hypothèses en physique était « difficile à analyser » (*C*, 28^{e} l., 456). Des indications qu'il a fournies, il n'est pas toujours aisé de dégager un sens clair et univoque et le jugement porté variera beaucoup avec l'interprétation retenue. Il est assez sûr toutefois que l'histoire ne s'est pas engagée dans la voie qu'il souhaitait et que le *positivisme bâtard* qu'il condamnait compte encore beaucoup d'adeptes. Parallèlement, le travail accompli en philosophie des sciences nous permet d'avoir aujourd'hui une vue plus précise et plus exacte du rôle et de la nature des hypothèses dans les sciences expérimentales. En conclure que Comte n'aurait plus rien à nous apprendre serait toutefois bien hâtif et des esprits aussi éminents que Peirce ou Duhem ont pris cette doctrine assez au sérieux pour entreprendre de la développer. La difficulté se concentre autour de la théorie exposé dans le *Cours* mais, avant de l'examiner en détail, il convient de la placer dans un contexte plus général.

La spontanéité de l'esprit et les divers usages d'*hypothèse*

Signe de l'importance qu'il accorde de façon générale au raisonnement hypothétique, Comte est revenu sur le sujet à diverses reprises, notamment dans les *Conclusions générales* du *Cours*, puis dans le *Système de politique positive* (1851-1854) et, si l'usage qu'il fait du terme ne se conforme pas toujours aux règles prescrites dans la vingt-huitième leçon, c'est que celles-ci ne sont destinées à s'appliquer que dans un contexte bien spécifique.

Le premier point qui se dégage est que, à travers la théorie des hypothèses, c'est « le genre de liberté resté facultatif pour notre intelligence » qui est en jeu (*C*, 58e l., 735). Le primat accordé à l'observation sur l'imagination ne signifie en aucune façon une adhésion à l'empirisme de la table rase. L'imagination n'est pas moins indispensable en science qu'en art et l'adoption de la méthode subjective après 1848 ne fera que donner une part croissante à cette liberté spéculative. L'hypothèse est un procédé qui trouve son origine non hors de nous, mais dans la spontanéité d'un esprit qui « n'est jamais passif dans ses relations avec le monde » (*S*, III, 19). — Celle-ci se manifeste encore sous d'autres formes et, aux côtés de l'hypothèse, la méthode positive fait également une place à des fictions scientifiques, comme ces « organismes fictifs, artificiellement imaginés », qu'elle propose d'intercaler entre les organismes connus « de manière à faciliter leurs comparaison, en rendant la série biologique plus homogène et plus continue, en un mot plus régulière »[3]. — Comte est ainsi conduit à faire un usage assez varié du terme *hypothèse*. La dynamique sociale, par exemple, repose sur « l'hypothèse d'un peuple unique ». Encore qualifiée de « fiction rationnelle » (*C*, 48e l., 171) car il ne s'agit pas d'en déduire des conséquences, l'idée empruntée à Condorcet est décrite comme une idéalisation destinée à ordonner les données et simplifier le récit en introduisant une continuité non attestée dans l'histoire. De la même façon, le tome deux du *Système* nous invite à considérer l'hypothèse d'un monde sans lois, ou celle d'un pays de cocagne (respectivement *S*, II, 28-29 et 141-149). Ce sont là des expériences de pensée et non plus des idéalisations : dans les deux cas, la situation en question est reconnue comme contrefactuelle, et il s'agit d'en explorer les conséquences, ou les facettes, pour ensuite la comparer avec la réalité, dans le seul but de mieux caractériser celle-ci.

La cinquante huitième leçon avait déjà, sinon assoupli les conditions formulées dans la vingt-huitième leçon, du moins élargi le cadre théorique dans lequel elles s'inscrivent. Y est introduite en effet une distinction jusqu'alors absente, qui étend la sphère de la positivité. Une fois exclues les causes, et à s'en tenir à la seule recherche des lois, deux types de question peuvent en effet se présenter : certaines sont simplement prématurées, alors que d'autres portent sur des sujets « indéfiniment inaccessibles, quoique de nature positive » (*C*, 58e l., 735). Le premier cas rentre directement dans le cadre de la théorie des hypothèses. Le second revient à admettre l'existence de questions « que l'esprit humain ne saurait certainement résoudre jamais, et qui méritent cependant d'être qualifiées de positives, parce qu'on peut concevoir qu'elles deviendraient accessibles à une intelligence mieux organisée »[4]. En dépit de cette profonde différence, le principe utilisé dans le premier cas, mais plus libéral que celui de 1835, s'applique également : « former les suppositions les plus propres à faciliter notre marche mentale, sous la double

[3] *C*, 40e l., 728. Comte dit emprunter l'idée aux mathématiques où « on a souvent trouvé de grands avantages à imaginer directement une suite quelconque de cas hypothétiques, dont la considération, quoique simplement artificielle, peut faciliter beaucoup, soit l'éclaircissement plus parfait du sujet naturel des recherches, soit même son élaboration fondamentale. Un tel artifice diffère essentiellement de celui des hypothèses proprement dites, avec lesquelles il a été toujours confondu jusqu'ici par les plus profonds philosophes. Dans ce dernier cas, la fiction ne porte que sur la seule solution du problème ; tandis que, dans l'autre, le problème lui-même est radicalement idéal, sa solution pouvant être, d'ailleurs, entièrement régulière ». Voir encore 60e l., 786.

[4] *C*, 58e l., 735 ; Comte pense sans doute ici à l'être omniscient dont parle Laplace dans l'introduction de son *Essai philosophique sur les probabilités*.

condition permanente de ne choquer aucune notion antérieure, et être toujours disposé à modifier ces artifices, aussitôt que l'observation viendrait à l'exiger ≫ (*Ibid.*). Comte donne aussitôt deux exemples : ≪ l'hypothèse, spontanément adoptée en physique, sur la constitution moléculaire des corps, pourvu toutefois qu'on ne lui attribue jamais une vicieuse réalité ≫ et ≪ l'artifice fondamental du dualisme ≫ en chimie[5]. Dans la trente sixième leçon du *Cours*, il avait en effet proposé de considérer toute composition chimique comme résultant d'une suite de combinaisons, et donc ultimement comme binaire ; en prenant soin d'ajouter : ≪ je ne propose point le dualisme universel et invariable comme une loi réelle de la nature, que nous ne pourrions jamais avoir aucun moyen de constater ; mais je le proclame un artifice fondamental de la vraie philosophie chimique, destiné à simplifier toutes nos conceptions élémentaires, en usant judicieusement du genre spécial de liberté resté facultatif pour notre intelligence ≫[6]. Dès lors qu'on voit dans un tel procédé un simple artifice logique et non une loi naturelle, qu'on ne confond pas le subjectif et l'objectif, rien ne s'oppose à son emploi.

La théorie de la vingt-huitième leçon

Ce cadre une fois fixé, il devient possible d'entrer plus en détail dans l'examen de la théorie exposée à la fin de la vingt-huitième leçon du *Cours*. Si ces pages ne constituent pas la vue définitive de Comte sur le sujet, les développements ultérieurs ne forment que de brèves remarques, alors qu'on a affaire ici à un exposé détaillé et c'est pourquoi, depuis Mill, c'est sur lui que s'est concentrée la discussion.

Après avoir brièvement constaté l'impossibilité à rendre compte de la démarche expérimentale au moyen des seuls procédés reconnus d'ordinaire par les logiciens, à savoir la déduction et l'induction, et la nécessité qui s'ensuit de recourir à un autre type de raisonnement consistant à ≪ anticiper sur les résultats, en faisant une supposition provisoire, d'abord essentiellement conjecturale, quant à quelques-unes des notions mêmes qui constituent l'objet final de la recherche ≫, Comte s'empresse d'ajouter :

> Mais, l'emploi de ce puissant artifice doit être constamment assujetti à une condition fondamentale, à défaut de laquelle il tendrait nécessairement, au contraire, à entraver le développement de nos vraies connaissances. Cette condition, jusqu'ici vaguement analysée, consiste à ne jamais imaginer que des hypothèses susceptibles, par leur nature, d'une vérification positive, plus ou moins éloignée, mais toujours clairement inévitable, et dont le degré de précision soit exactement en harmonie avec celui que comporte l'étude des phénomènes correspondants. En d'autres termes, les hypothèses vraiment philosophiques doivent constamment présenter le caractère de simples anticipations sur ce que l'expérience et le raisonnement auraient pu dévoiler immédiatement, si les circonstances du problème eussent été plus favorables[7].

Qui cherche à comprendre ce que Comte demande d'une bonne hypothèse et ce

[5]Nous parlerions aujourd'hui d'hypothèse atomique ; la restriction explique qu'à la fin du dix-neuvième siècle, les positivistes aient été hostiles à l'atomisme.

[6]*C*, 36e l., 602 ; voir encore *S*, I, 553-554. L'hypothèse a vite été critiquée par Laurent, qui lui reprochait de multiplier inutilement les niveaux d'analyse.

[7]Dans ce qui suit, il ne sera pas question de cet aspect de la condition qui porte sur le degré de précision requis des hypothèses. Sur ce point, on se reportera à l'ouvrage de Bachelard cité n. 1.

qu'il rejette se heurte vite à toute sorte de difficultés, que deux principes d'interprétation aideront à démêler. Tout d'abord, l'analyse doit être à la fois conceptuelle et historique, une des conclusions de l'exposé étant précisément que « la philosophie des sciences ne saurait être convenablement étudiée séparément de leur histoire » (*C*, 28e l., 464 ; cf. 49e l., 237, qui renvoie à ce passage). De plus, l'analyse se meut tour à tour au plan descriptif et au plan normatif. On ne comprendrait rien à la théorie fondamentale des hypothèses si on ne voyait qu'elle entend répondre à la situation que l'auteur avait sous les yeux. Le positivisme bâtard est d'abord un fait, de l'ordre du constat. Mais, dans la pratique des savants de son temps, Comte croit nécessaire de faire le partage entre la bonne pratique, telle qu'illustrée par Fourier, et la mauvaise.

La place accordée à l'astronomie illustre bien ce jeu du descriptif et du normatif. La pratique des astronomes est en effet donnée en modèle aux physiciens, confirmant une fois de plus le caractère exemplaire de la plus ancienne des sciences naturelles.

> Tel fait est encore peu connu, ou telle loi est ignorée : on forme alors à cet égard une hypothèse, le plus possible en harmonie avec l'ensemble des données déjà acquises ; et la science, pouvant ainsi se développer librement, finit toujours par conduire à de nouvelles conséquences observables, susceptibles de confirmer ou d'infirmer, sans aucune équivoque, la supposition primitive (*C*, 28e l., 458).

Les leçons d'astronomie ou le *Traité* de 1844 contiennent ainsi de nombreux exemples d'un usage irréprochable de l'hypothèse. Ainsi, ce que l'humanité a longtemps observé dans le ciel, c'était uniquement la constance des diverses configurations d'étoiles, et leur déplacement régulier d'est en ouest ; mais cela ne suffisait pas à l'astronome pour rendre compte des apparences et prévoir la position future des astres. Aussi l'astronome a-t-il posé l'existence d'une voûte céleste, d'une sphère des fixes ayant pour centre la terre. Telle est, ajoute Comte, « la première grande conception scientifique que l'esprit humain ait dû former. Sans doute, nous n'y attachons plus le même sens que les anciens, qui y voyaient l'expression absolue de la réalité ; mais à titre d'artifice astronomique, elle comportera toujours la même efficacité habituelle »[8].

Le fait que l'examen d'une question aussi importante que celle qui nous occupe ait été reporté aux leçons de physique, alors pourtant que c'est l'astronomie qui nous enseigne le bon usage des hypothèses, indique le sens polémique ou, si l'on préfère, la fonction préventive que Comte accorde à cette théorie. Les astronomes se conformant spontanément à la condition fondamentale énoncée plus haut, il n'était pas nécessaire de la formuler explicitement. Tel n'est toutefois plus le cas quand on passe d'une science à l'autre :

> Les diverses hypothèses employées aujourd'hui par les physiciens doivent être soigneusement distinguées en deux classes : les unes, jusqu'ici peu multipliées, sont simplement relatives aux lois des phénomènes ; les autres, dont le rôle actuel est beaucoup plus étendu, concernent la détermination des agents généraux auxquels on rapporte les différents genres d'effets naturels. (*C*, 28e l., 458).

[8] *TPAP*, 122. La détermination de la figure de la terre fournit un autre exemple de la façon dont l'esprit humain a été amené à forger une suite d'hypothèses de plus en plus satisfaisantes. Voir *TPAP*, 155-156 et les leçons correspondantes du *Cours*.

Il est même possible de préciser davantage la cible visée. Deux des branches de la physique, la barologie ou théorie de la gravitation, la thermologie ou théorie de la chaleur, ont en effet suivi le modèle de l'astronomie et se sont débarrassées des hypothèses du second type. Bachelard a notamment montré de façon convaincante tout ce que, sur le point qui nous occupe, Comte doit à Fourier, un des deux dédicataires du *Cours*[9]. Reste donc seulement l'étude de la lumière et de l'électricité. Alors que Fourier nous a appris comment faire l'économie du calorique, dans ces deux derniers cas, on continue à admettre l'existence d'un fluide électrique ou d'un éther luminifère. Or, objecte Comte, l'existence de ces prétendues entités n'est

> pas plus susceptible de négation que d'affirmation, puisque, d'après la constitution qui leur est soigneusement attribuée, ils échappent nécessairement à tout contrôle positif. Quelle argumentation sérieuse pourrait-on instituer pour ou contre des corps ou des milieux dont le caractère fondamental est de n'en avoir aucun ? Ils sont expressément imaginés comme invisibles, intangibles, impondérables même, et d'ailleurs inséparables des substances qu'ils animent : notre raison ne saurait donc avoir sur eux la moindre prise (*C*, 28ᵉ l., 459).

La lumière indépendante du corps lumineux, ou l'électricité séparée du corps électrique ne diffèrent des entités scolastiques que par une corporéité ≪ fort équivoque, puisqu'on leur ôte expressément, par leur définition fondamentale, toutes les qualités susceptibles de caractériser une matière quelconque ≫ (*C*, 28ᵉ l., 461).

Si ce type d'hypothèse n'a donc pas sa place en science, reste à expliquer pourquoi la décision de les exclure se heurte à une telle résistance. La réponse proposée renvoie à la structure de l'esprit humain, telle qu'elle ressort de la loi des trois états. ≪ Quoique la métaphysique ne constitue elle-même [. . .] qu'une grande transition générale de la théologie à la science réelle, une transition secondaire, et, par là, beaucoup plus rapide, devient ensuite nécessaire entre les conceptions métaphysiques et les conceptions vraiment positives ≫ (*Ibid.*). Cet état de la physique, qualifié pour cette raison de *positivisme incomplet* ou *bâtard*, constitue donc un intermédiaire historiquement indispensable, mais voué à disparaître dès lors qu'il a rempli sa fonction, comme le montre l'exemple de la théorie cartésienne des tourbillons : historiquement considérée, en introduisant l'idée d'un mécanisme quelconque, elle représentait un incontestable progrès par rapport aux explications proposées un peu plus tôt par Kepler mais, ce service une fois rendu, il a bien fallu se résoudre à l'abandonner.

La théorie des hypothèses constitue bien une des pièces maîtresses de l'épistémologie positiviste. Le recours à l'hypothèse étant rigoureusement indispensable en science, on voit mal comment l'épistémologue pourrait ne pas l'étudier. Preuve de la distance qui sépare la pensée de Comte des doctrines qui lui sont d'ordinaire attribuées, l'hypothèse est d'abord pour lui une libre création de l'esprit. A ce titre, elle est semblable aux fictions qu'il n'hésitait pas à introduire en science pour faciliter la marche du savant.

[9]Voir le texte de Bachelard, cité n. 1, qui conclut en ces termes : ≪ Ainsi, on peut justifier, à bien des points de vue, la prudence scientifique de Comte. L'intransigeance de sa réaction contre l'esprit métaphysique était elle-même nécessaire dans une période où la science prétendait assurer ses fondations. On ne peut pas refuser à Comte une claire vision des conditions scientifiques de son époque, et surtout la compréhension exacte de l'organisation et de la discipline qui sont indispensables pour faire travailler à plein rendement la société savante. ≫

Une comparaison avec cet autre type d'artifice logique permet de mieux en saisir la spécificité. Tout d'abord, l'hypothèse est un instrument heuristique, le point de départ d'un raisonnement qui nous permet d'augmenter nos connaissances, tandis que la fiction n'est pas là pour qu'on en déduise quoi que ce soit. La fécondité de l'hypothèse est liée également à son caractère transitoire d'anticipation. Si Comte insiste tant sur son éliminabilité, c'est qu'il y voit le gage de sa positivité : l'hypothèse est destinée, non pas à garder indéfiniment le statut d'hypothèse, mais à devenir l'énoncé d'un fait ou d'une loi, comme la monnaie est destinée à être échangée pour une marchandise[10]. A la différence de la fiction, elle émet la prétention à décrire ≪ par provision ≫ la réalité et, une fois connu le verdict de l'expérience, elle sera acceptée, ou rejetée.

La théorie comtienne des hypothèses combine les approches conceptuelle et historique. Il convient de la lire dans le même esprit et de la replacer dans son contexte. On en comprend mieux alors le caractère hautement polémique, qui se réclamait à juste titre des travaux de Fourier sur la chaleur. Il y avait de bonnes raisons de vouloir se débarrasser des fluides ou de l'éther et autres entités qui encombraient la théorie physique. Duhem a cherché à construire une optique conforme aux canons positivistes et la théorie de l'abduction de Peirce est donnée explicitement comme la reprise du projet comtien[11].

[10]En dépit des apparences, Comte n'est pas vérificationiste. Il savait très bien que les lois scientifiques ≪ ne sont jamais que des hypothèses assez confirmées par l'observation ≫ (*S*, II, 33). Parmi les hypothèses disponibles à un moment donné, on choisit celle qui s'accorde le mieux avec les données expérimentales. La question de la figure de la terre donne un bon exemple d'une suite d'hypothèses de plus en plus satisfaisantes. Voir également *C*, 24e l., 391, où Comte reconnaît que le refus des notions absolues conduit à admettre que même les lois qui nous paraissent les mieux établies, comme à son époque les lois de Newton, ne sont pas à l'abri de réfutation.

[11]Voir respectivement : M. Blay : ≪ Comte et Duhem ou la construction d'une optique positive ≫, *Revue philosophique*, 2007-4, 493-504 et Peirce, pour qui ≪ the true maxim of abduction is that which Comte endeavored to formulate when he said that any hypothesis might be admissible if and only if it was verifiable ≫ (in : *Harvard Lectures on Pragmatism* (1903 ; Lecture VI: the Nature of Meaning) ; in Houser, N. and Kloesel, Ch. (eds) : *The Essential Peirce*, Bloomington, Indiana U. P., 1992, vol. 2, 225.

Seligman-style deduction for hybrid modal logic

Patrick Blackburn, Torben Braüner*, Klaus Frovin Jørgensen
Roskilde University, Denmark

and

Thomas Bolander
Technical University of Denmark

Abstract. A number of different sorts of proof-systems for modal logics are available: Major examples are labelled systems, display logics, multiple sequents, and hybrid-logical systems. It is not the purpose of this talk to discuss the pros and cons of all these systems, but instead to focus on one particular sort of proof-system based on hybrid modal logic, which is an extension of ordinary modal logic allowing explicit reference to individual points in a Kripke model (the points stand for times, locations, states, possible worlds, or something else).

This sort of proof-system was put forward by Jerry Seligman in the late 1990s [4]. One particular feature of this sort of system is the possibility to jump to a hypothetical time (or whatever the points stand for), do some reasoning, and then jump back to the present time again. In natural deduction versions [2] the hypothetical reasoning is kept track of using machinery called explicit substitutions, similar to "proof-boxes" in the style of linear logic. Such a hybrid-logical proof-box encapsulates hypothetical reasoning taking place at one particular time. Within the proof-box, information depending on the hypothetical time can be dealt with, but only non-indexical information, that is, statements whose truth-values do not depend on the time, can flow in and out of the box.

In my talk I'll present Seligman-style natural deduction, and I'll in particular discuss the above mentioned machinery enabling hypothetical reasoning. I'll also briefly describe two other lines of work in the area of Seligman-style reasoning:
1. Seligman-style natural deduction has turned out to be useful for formalizing so-called false-belief tests in cognitive psychology [3].
2. Recently developed Seligman-style tableau systems with desirable proof-theoretic properties [1].

References

[1] P. Blackburn, T. Braüner, T. Bolander, and K.F. Jørgensen. A Seligman-style tableau system. In K. McMillan, A. Middeldorp, and A. Voronkov, editors, *Logic for Programming, Artificial Intelligence, and Reasoning*, volume 8312 of *Lecture Notes in Computer Science*, pages 147–163. Springer Publishing Company, 2013.

[2] T. Braüner. *Hybrid Logic and its Proof-Theory*, volume 37 of *Applied Logic Series*. Springer, 2011.

[3] T. Braüner. Hybrid-logical reasoning in false-belief tasks. In B. C. Schipper, editor, *Proceedings of Fourteenth Conference on Theoretical Aspects of Rationality and*

*The talk was given by Torben Braüner, torben@ruc.dk. Torben Braüner acknowledges the financial support received from The Velux Foundation as funding for the project Hybrid-Logical Proofs at Work in Cognitive Psychology (VELUX 33305).

Knowledge (TARK), pages 186–195, 2013. ISBN 978-0-615-74716-3, available at http://tark.org.

[4] J. Seligman. The logic of correct description. In M. de Rijke, editor, *Advances in Intensional Logic*, volume 7 of *Applied Logic Series*, pages 107–135. Kluwer, 1997.

Explanatory Justice: The Case of Disjunctive Explanations

Michael Cohen
München

Abstract. In recent years there has been an effort to explicate the concept of explanatory power in a Bayesian framework, by constructing explanatory measures. It has been argued that those measures should not violate the principle of explanatory justice, which states that explanatory power cannot be extended 'for free'. I argue, by formal means, that one recent measure that was claimed to be immune from explanatory injustice fails to be so. Thereafter, I propose a conceptual way of avoiding the counterintuitive side effects of the discussed formal results. I argue for another way of understanding the notion of 'extending explanatory success', for the cases of negative explanatory power. A consequence of this interpretation is that the original explanatory justice criticism disappears. Furthermore, some of the formal results derived from explanatory measures can be made conceptually clear under this kind of understanding.

An introduction to deduction

Kosta Došen
Mathematical Institute, SANU

Abstract. To determine what deductions are it does not seem sufficient to know that the premises and conclusions are propositions, or something in the field of propositions, like commands and questions. It seems equally, if not more, important to know that deductions make structures, which in mathematics we find in categories, multicategories and polycategories. It seems also important to know that deductions should be members of particular kinds of families, which is what is meant by their being in accordance with rules.

Negating indicative conditionals

Paul Egré
CNRS, Institut Jean-Nicod, Paris

Abstract. Negation is often viewed as a "litmus test" for theories of indicative conditionals in natural language (see Handley, Evans and Thompson 2006). Three main families of semantic theories compete when it comes to understanding the interaction of negation with conditional sentences, namely i) suppositional analyses inspired by the Ramsey Test, according to which the negation of "if P then Q" is of the form "If P then not Q"; ii) the two-valued truth-functional analysis according to which the negation should be a conjunction of the form "P and not Q" and iii) strict conditional analyses predicting a weak negation of the form "possibly P and not Q". We present the results of several experimental studies concerning the denial of indicative conditional sentences in dialogues, intended to show that all three forms of negations can be retrieved from the weak negation, depending on the additional assumptions available to the denier of the conditional.

(Joint work with Guy Politzer.)

On a dogma of Proof-Theoretic Semantics: generalising canonicity of derivations

Nissim Francez
Computer Science dept., the Technion-IIT, Haifa, Israel
(francez@cs.technion.ac.il)
and
Julien Murzi
School of European Culture and Languages,
University of Kent, UK
and Munich centre for Mathematical Philosophy
(j.murzi@kent.ac.uk)

Introduction

– A dogma within the PTS community: I-rules of a meaning-conferring ND-system fully determine meanings.
– As a result, there emerged a conception of a *canonical proof* ends with an I-rule.
– Recently, canonicity was extended to arbitrary derivations from open assumptions, again "essentially" ending with an I-rule.
– A canonical derivation is the *most direct* way of concluding a formula from some assumptions which are taken as *grounds for assertion* of the conclusion.
– This dogma has some far-reaching consequences:
- the notion of *harmony*, central to PTS, requires a balance between the I-rules and E-rules (only).
- One way in which harmony was formalized is by the existence of *derivation reductions* removing maximal formulas, formulas which are the conclusion of an I-rule and major premises of an E-rule.

Our goal

– Harmony is taken as a necessary condition for the proof-theoretic justification of a logic, the heart of the PTS programme.
- Consequently, many PTS-adherers reject Classical Logic, where in its standard presentation the E-rule for negation, the *double negation elimination* (DNE) lacks harmony, and prefer Intuitionistic Logic, the standard ND-system of which NJ does enjoy harmony.
– We claim that the focus of PTS has to be on *direct grounds* (for assertion), which need not necessarily coincide with I-rules.
- Sometimes, notably for *disjunctions*, such grounds may arise differently.
- In particular, we suggest that in ND-systems that have primitive *structural rules* (S-rules) allowing the conclusion of formulas, not just the manipulation of assumptions, a derivation ending with such a rule should also count as canonical, also providing grounds for assertion of its conclusion.

Justification

– Classicality by proof-theoretic kosher means is typically gotten by strengthening the structural resources of the meaning-conferring system, e.g.:
- the coordination principles in bilateral systems, or multiple conclusions in both sequent calculi and ND-systems.
- Since these resources allow us to assert new theorems, such as Peirce's Law, LEM and more, it may be justly thought that they are meaning-determining;
- they license new uses of disjunction, implication etc.
- If those meanings are to be accounted for in terms of canonical grounds, the notion of a canonical ground needs to be enriched and expanded accordingly.
– Furthermore, an S-rule can also eliminate a formula, similarly to an E-rule.
This is particularly significant for the pragmatist account of PTS, according to which the E-rules fully determine meaning.

The consequence

– As a consequence of this suggestion, there emerges a need for a revision of the definition of harmony.
– A balance is required between ways of introduction on the one hand (here applications of both I-rules and S-rules), and ways of elimination on the other hand (here applications of E-rules and, again, of S-rules).
- The establishment of this balance leads to the consideration of additional derivation reductions, on top of the usual detour elimination reductions, only confronting I-rules with E-rules.
The conception of *self-justifying*, meaning-conferring rules is thereby extended.

Classifying Structural rules

– S-rules are further divided as follows:
- S_a-rules: These are S-rules that only manipulate assumptions, like Gentzen's original rules (for single conclusion) *Weakening*, *Contraction* and *Exchange*. Typically in ND-systems, S_a-rules are not taken as primitive; rather, they are absorbed into the identity axiom.
- S_i-rules: Those are structural rules that allow the introduction of a conclusion, like some *coordination rules* in the bilateral presentation of Classical Logic (below), or right weakening in themultiple-concusions presentation of classical logic (also below).
- S_e-rules: Those are structural rules that allow the elimination of a conclusion, again like the *coordination rules* in the bilateral presentation of Classical Logic, or right contraction in the multiple-confusions presentation of classical logic.
- We assume the usual notion of a tree-like derivation, ranged over by $\mathcal{D}$.

S-canonicity, grounds, bridge formula

Definition: A derivation $\mathcal{D}$ is *S-canonical* iff its last step is one of the following:
1. An application of an "essential" I-rule.
2. An application of an S_i-rule.
- Denote S-canonical derivability by $\vdash^{sc}$.

Definition: Γ is a *grounds for assertion* for φ iff $\vdash^{sc} \Gamma : \varphi$.

Definition: A formula φ occurring as a node in a derivation $\mathcal{D}$ is a *bridge formula* iff it is either a conclusion of an application of an I-rule or of an S_i-rule, and at the same time the major premise of an application of an E-rule or a premise of an application of S_e-rule.
- The notion of a *maximal formula* is a special case of a bridge formula, when the formula is introduced via an I-rule and immediately afterwards eliminated via an E-rule.

Local-soundness

– We now cast harmony, known also as local-soundness, in terms of reducing derivations containing a bridge formula.

Definition: An ND-system is *S-locally-sound* iff every derivation containing a bridge formula can be reduced to an equivalent one (having the same conclusion and the same (or fewer) open assumptions) where that bridge formula does not occur.
– This definition gives rise to four kinds of reductions:
- I/E-reductions: These are the traditional detour-removing reductions, eliminating occurrences of maximal formulas.
- I/S_e-reductions: These reductions eliminate bridge formulas introduced via an I-rule and eliminated via an S_e-rule.
- S_i/E-reductions: These are reductions eliminating a bridge formula introduced via an S_i-rule and eliminated via an E-rule.
- S_i/S_e-reductions: These are purely structural reductions, eliminating bridge formulas introduced via an S_i-rule and eliminated via an S_e-rule.

A bilateral justification of Classical Logic

$$\frac{+\varphi \qquad +\psi}{+(\varphi \wedge \psi)}\ (\wedge^+ I)$$

$$\frac{+(\varphi \wedge \psi) \qquad \begin{array}{c}[+\varphi]_i, [+\psi]_j \\ \vdots \\ +\xi\end{array}}{+\xi}\ (\wedge^+ E^{i,j})$$

$$\frac{+\varphi}{+(\varphi \vee \psi)}\ (\vee^+ I_1) \qquad \frac{+\psi}{+(\varphi \vee \psi)}\ (\vee^+ I_2)$$

$$\frac{+(\varphi \vee \psi) \qquad \begin{array}{c}[+\varphi]_i \\ \vdots \\ +\xi\end{array} \qquad \begin{array}{c}[+\psi]_j \\ \vdots \\ +\xi\end{array}}{+\xi}\ (\vee^+ E^{i,j})$$

$$\frac{\begin{array}{c}[+\varphi]_i \\ \vdots \\ +\psi\end{array}}{+(\varphi \to \psi)}\ (\to^+ I^i) \qquad \frac{+(\varphi \to \psi) \qquad +\varphi \qquad \begin{array}{c}[+\psi]_i \\ \vdots \\ +\xi\end{array}}{+\xi}\ (\to^+ E^i)$$

$$\frac{-\varphi}{-(\varphi \wedge \psi)}\ (\wedge^- I_1) \qquad \frac{-\psi}{-(\varphi \wedge \psi)}\ (\wedge^- I_2) \qquad \frac{-(\varphi \wedge \psi) \qquad \begin{array}{c}[-\varphi]_i \\ \vdots \\ \xi\end{array} \qquad \begin{array}{c}[-\psi]_j \\ \vdots \\ \xi\end{array}}{\xi}\ (\wedge^- EI, j)$$

$$\frac{-\varphi \qquad -\psi}{-(\varphi \vee \psi)}\ (\vee^- I) \qquad \frac{-(\varphi \vee \psi) \qquad \begin{array}{c}[-\varphi]_i, [-\psi]_j \\ \vdots \\ \xi\end{array}}{\xi}\ (\vee^- E^{i,j})$$

$$\frac{+\varphi \qquad -\psi}{-(\varphi \to \psi)}\ (\to^- I) \qquad \frac{-(\varphi \to \psi) \qquad \begin{array}{c}[+\varphi]_i, [-\psi]_j \\ \vdots \\ \xi\end{array}}{\xi}\ (\to^- E^{i,j})$$

Defining negation bilaterally

$$\frac{-\varphi}{+\neg\varphi}\ (\neg^+ I) \qquad \frac{+\neg\varphi}{-\varphi}\ (\neg^+ E)$$

$$\frac{+\varphi}{-\neg\varphi}\ (\neg^- I) \qquad \frac{-\neg\varphi}{+\varphi}\ (\neg^- E)$$

– The definition of negation is based on the approach that asserting $\neg\varphi$ is warranted by denying φ, and denying $\neg\varphi$ is warranted by denying $\neg\varphi$.

Coordination rules

– The classical strength effect of the two forces of assertion and denial is established by means of appropriate *structural rules*, under the name of *coordination rules*, extending the meaning of the underlying derivability relation.
- For the expression of those rules, $\perp$ is used, viewed not as a propositional constant but as a logical punctuation sign. Consequently it does not occur with a force marker attached.

$$\Gamma, +\varphi, -\varphi : \sigma \ (INC) \qquad \frac{\Gamma : \sigma \quad \Gamma : \overline{\sigma}}{\Gamma : \perp} \ (LNC)$$

$$\frac{\Gamma, \sigma : \perp}{\Gamma : \overline{\sigma}} \ (RED^i)$$

- (INC) establishes the incoherence of n context, that cannot be coherently used in deductions. (*explosion*).
- (LNC) and (RED) allow for a *reversal* of the judgement (in both directions), once a contradiction is reached.

Some derived rules

– A useful *derived* bilateral structural rule is (Co), expressing *contraposition* displayed in its logistic presentation with explicit contexts:

$$\frac{\Gamma, \beta : \alpha}{\Gamma, \overline{\alpha} : \overline{\beta}} \ (Co)$$

- Contraposition leads to another useful structural rule, called here $(Cases)$, to be used in the reductions.

$$\frac{\Gamma, +\varphi : \xi \quad \Gamma, -\varphi : \xi}{\Gamma : \xi} \ (Cases)$$

- Notably, if $\perp$ *is* considered a propositional constant, the rule (RED) is impure. An alternative, pure form of a bilateral Reduction structural rule, evading the use of $\perp$, was proposed by Smiley:

$$\frac{\begin{matrix}[\sigma]_i \\ \vdots \\ \alpha\end{matrix} \quad \begin{matrix}[\sigma]_j \\ \vdots \\ \overline{\alpha}\end{matrix}}{\overline{\sigma}} \ (SR^{i,j})$$

S-harmony of the bilateral system

– Only I/E-reductions and S_i/E-reductions have to be considered, as there are no S_e-rules.
- The I/E-reductions were presented elsewhere and skipped here.
–Sample S_i/E-reductions are presented, showing there is no need to eliminate after inferring via the S_i-rule (RED).
- For each connective, there are separate reductions for the two force markers.

Implication

Assertion:

$$\dfrac{\begin{array}{c}\mathcal{D}_2\\ +\varphi\end{array} \quad \dfrac{\begin{array}{c}[-(\varphi\to\psi)]_i\\ \mathcal{D}_1\\ \bot\end{array}}{+(\varphi\to\psi)}\,(RED^i)}{+\psi}\,(\to^+ E) \quad \leadsto_r \quad \dfrac{\begin{array}{c}\dfrac{\begin{array}{c}\mathcal{D}_2\\ +\varphi\end{array} \quad [-\psi]_j}{-(\varphi\to\psi)}\,(\to^- I)\\ \mathcal{D}_1\\ \bot\end{array}}{+\psi}\,(RED^j)$$

Denial: Two negative E-rules.

$$\dfrac{\dfrac{\begin{array}{c}[+(\varphi\to\psi)]_i\\ \mathcal{D}_1\\ \bot\end{array}}{-(\varphi\to\psi)}\,(RED_1)}{+\varphi}\,(\to^- E_i) \quad \leadsto_r \quad \dfrac{\begin{array}{c}\dfrac{[-\varphi]_i}{+(\varphi\to\psi)}\,(*)\\ \mathcal{D}_1\\ \bot\end{array}}{+\varphi}\,(RED_i)$$

$(*)$:

$$\dfrac{\dfrac{\dfrac{[\varphi]_i \quad [+\varphi]_j \quad [-\psi]_k}{\bot}\,(LNC)}{+\psi}\,(RED^k)}{+(\varphi\to\psi)}\,(\to^+ I^j)$$

$$\dfrac{\dfrac{\begin{array}{c}[+(\varphi\to\psi)]_i\\ \mathcal{D}_1\\ \bot\end{array}}{-(\varphi\to\psi)}\,(RED^i)}{-\psi}\,(\to^- E_2) \quad \leadsto_r \quad \dfrac{\begin{array}{c}\dfrac{[+\psi]_i}{+(\varphi\to\psi)}\,(\to^+ I^l)\\ \mathcal{D}_1\\ \bot\end{array}}{-\psi}\,(RED^i)$$

Conjunction

Assertion:

$$\dfrac{\dfrac{\begin{matrix}[-(\varphi\wedge\psi)]_i\\ \mathcal{D}_1\\ \bot\end{matrix}}{+(\varphi\wedge\psi)}\,(RED^i)}{+\varphi}\,(\wedge^+E_1) \quad \rightsquigarrow_r \quad \dfrac{\begin{matrix}\dfrac{[-\varphi]_j}{-(\varphi\wedge\psi)}\,(\wedge^-I_1)\\ \mathcal{D}_1\\ \bot\end{matrix}}{+\varphi}\,(RED^j)$$

Denial:

$$\dfrac{\dfrac{\begin{matrix}[+(\varphi\wedge\psi)]_i\\ \mathcal{D}_1\\ \bot\end{matrix}}{-(\varphi\wedge\psi)}\,(RED^i) \quad \begin{matrix}[-\varphi]_j\\ \mathcal{D}_2\\ \xi\end{matrix} \quad \begin{matrix}[-\psi]_j\\ \mathcal{D}_3\\ \xi\end{matrix}}{\xi}\,(\wedge^-E^{j,k})$$

$\rightsquigarrow_r$

$$\dfrac{\begin{matrix}\dfrac{\begin{matrix}\dfrac{[+\varphi]_i \quad [+\psi]_j}{+(\varphi\wedge\psi)}\,(\wedge^+I)\\ \mathcal{D}_1\\ \bot\end{matrix}}{-\varphi}\,(RED^i)\\ \mathcal{D}_2\\ \xi\end{matrix} \quad \begin{matrix}[-\psi]_k\\ \mathcal{D}_3\\ \xi\end{matrix}}{\xi}\,(Cases^{j,k})$$

A multiple-conclusions justification of Classical Logic

$$\dfrac{\varphi,\Delta \quad \psi,\Delta}{\varphi\wedge\psi,\Delta}\,(\wedge I) \qquad \dfrac{\varphi\wedge\psi,\Delta}{\varphi,\Delta}\,(\wedge_1 E) \qquad \dfrac{\varphi\wedge\psi,\Delta}{\psi,\Delta}\,(\wedge_2 E)$$

$$\dfrac{\varphi,\Delta}{\varphi\vee\psi,\Delta}\,(\vee_1 I) \qquad \dfrac{\psi,\Delta}{\varphi\vee\psi,\Delta}\,(\vee_2 I) \qquad \dfrac{\varphi\vee\psi,\Delta}{\varphi,\psi,\Delta}\,(\vee E)$$

$$\dfrac{\begin{matrix}[\varphi]_i\\ \mathcal{D}\\ \psi,\Delta\end{matrix}}{\varphi\to\psi,\Delta}\,(\to I^i) \qquad \dfrac{\varphi\to\psi,\Delta \quad \varphi,\Delta}{\psi,\Delta}\,(\to E)$$

$$\dfrac{\begin{matrix}[\varphi]_i\\ \mathcal{D}\\ \Delta\end{matrix}}{\neg\varphi,\Delta}\,(\neg I^i) \qquad \dfrac{\neg\varphi,\Delta \quad \varphi,\Delta}{\Delta}\,(\neg E)$$

$$\dfrac{\Delta}{\varphi,\Delta}\,(RW) \qquad \dfrac{\varphi,\varphi,\Delta}{\varphi,\Delta}\,(RC)$$

S-harmony of the MCND-system

– Note that right contraction (RC) is an S_e structural rule, while right weakening (RW) is an S_i rule.
- Thus, we have to show the S_i/E-reductions establishing local-soundness in reducing derivation in which a formula is concluded by weakening and immediately eliminated.
- In addition, we have to show the I/S_e-reductions, where the introduced formula φ is already in Δ, and is immediately eliminated by RC.
- Finally, there are also S_i/S_e-reductions, where a formula $\varphi \in \Delta$ is introduced by right weakening and immediately eliminated via right contraction.
- The I/E-reductions were presented elsewhere and are omitted.

Sample reductions

– S_i/E-reductions: The common structure of all the S_i/E-reductions is that weakening followed by elimination is reduced to weakening only (once or twice).

Implication:

$$\dfrac{\begin{matrix}\mathcal{D}'\\ \varphi,\Delta\end{matrix} \quad \dfrac{\begin{matrix}\mathcal{D}\\ \Delta\end{matrix}}{\varphi\to\psi,\Delta}\,(RW)}{\psi,\Delta}\,(\to E) \qquad \leadsto_r \qquad \dfrac{\begin{matrix}\mathcal{D}\\ \Delta\end{matrix}}{\psi,\Delta}\,(RW)$$

Conjunction:

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}\\ \Delta\end{matrix}}{\varphi\wedge\psi,\Delta}\,(RW)}{\varphi,\Delta}\,(\wedge E_1) \qquad \leadsto_r \qquad \dfrac{\begin{matrix}\mathcal{D}\\ \Delta\end{matrix}}{\varphi,\Delta}\,(RW)$$

Negation:

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}\\ \Delta\end{matrix}}{\neg\varphi,\Delta}\,(RW) \quad \begin{matrix}\mathcal{D}'\\ \varphi,\Delta\end{matrix}}{\Delta}\,(\neg E) \qquad \leadsto_r \qquad \begin{matrix}\mathcal{D}\\ \Delta\end{matrix}$$

S_i/S_e-reductions: Here there is only one reduction to consider.

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}\\ \varphi,\Delta\end{matrix}}{\varphi,\varphi,\Delta}\,(RW)}{\varphi,\Delta}\,(RC) \qquad \leadsto_r \qquad \begin{matrix}\mathcal{D}\\ \varphi,\Delta\end{matrix}$$

Hypersequents and Linear Time

Andrzej Indrzejczak
Institute of Philosophy UŁ, Łódź

Abstract. Hypersequent calculi (HC) are a generalised form of sequent calculi invented independently by Pottinger and Avron. Hypersequents are usually defined as finite sets or multisets of ordinary sequents. HC proved to be very useful in the field of nonclassical logics including several modal, many-valued, relevant, and paraconsistent logics. We consider the application of HC to temporal logics of linear time. The first approach based on Avron's rule for modelling linearity is sufficient for dealing with logics such as S4.3 or K4.3 where only future-looking operators are taken into account. In order to provide cut-free formalizations of bimodal logics such like Kt4.3 where past-looking operators are added we need a different solution. The second approach is based on hypersequents treated as finite sequences of sequents. This solution is more expressive in the sense that the symmetry of past and future in linear time may be modelled by suitable rules.

Axioms as Hypotheses

Reinhard Kahle

CENTRIA, CM, and DM, FCT
Universidade Nova de Lisboa
2829-516 Caparica, Portugal
kahle@mat.uc.pt

Hypothetical Reasoning
23–24 August 2014, Tübingen

Research supported by the Portuguese Science Foundation, FCT, through the projects *Hilbert's Legacy in the Philosophy of Mathematics*, PTDC/FIL-FCI/109991/2009 and *The Notion of Mathematical Proof*, PTDC/MHC-FIL/5363/2012.

The notion of Axiom

1. Traditional notion
   - True statement
   - Self-evident assumption
2. Modern notion
   - Starting point for formal reasoning
   - Arbitrary assumption

- Truth disappeared
- Evidence disappeared

Non-Euclidean Geometry

- If there are alternatives for the parallel axiom which give rise to a consistent description of a "geometric reality", in which *mathematical* sense one of the versions of the parallel axiom could be true, and the other could be false?
- Of course, the parallel axiom could still be *empirically* true (relativity theory suggests today that this is not the case).
- But empricial truth would, in any case, not be the truth mathematicians are after.
- And even if so, *self-evidence* would be gone for sure.

Side question:

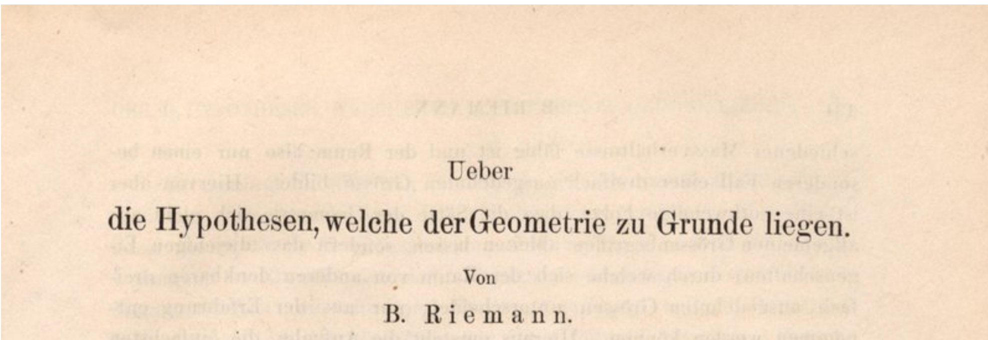
- The distinguished German mathematician Hans Grauert (1930–2011) tried to give additional evidence for Euclidean Geometry (by use of an interesting "homothety principle"), calling Euclidean Geometry **truer** than the other geometries.
- Is it philosophically possible to have a (non-trivial) comparative of "true"??

Bernhard Riemann

Riemann might have been the first(?) to change explicitely from "Axiom" to "Hypothesis":

Bernhard Riemann, 1854

These matters of fact are—like all matters of fact—not necessary, but only of empirical certainty; they are hypotheses.



Ueber

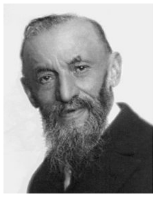
die Hypothesen, welche der Geometrie zu Grunde liegen.

Von

B. Riemann.

Hilbert's Axiomatic Method

- David Hilbert proposed the *Axiomatic Method* to set up (formal) mathematical theories to investigate all kinds of research areas.
- Doing so, the notion of axiom became increasingly disconnected from its traditional meaning.
- For Hilbert, the justification of an axiom system was now given by its raw consistency.
- Frege opposed to this, and see also Peano's comment on the next slide, but the historical development confirmed Hilbert.



Peano

Giuseppe Peano, 1906

But a proof that a system of postulates for arithmetic, or for geometry, does not involve a contradiction, is not, according to me, necessary. For we do not create postulates at will, but we assume as postulates the simplest propositions that are written, explicitly or implicitly, in every text of arithmetic or geometry. Our analysis of the principles of these sciences is a reduction of the ordinary affirmations to a minimum number, that which is necessary and sufficient. The system of postulates for arithmetic and geometry is satisfied by the ideas of number and point that every writer of arithmetic and geometry has. We think number, therefore number exists. A proof of consistency of a system of postulates can be useful, if the postulates are hypothetical and do not correspond to real facts.

(cited from Hubert C. Kennedy, Peano, Reidel, 1980, p. 118f)

- Thus, Peano still made a distinction between (tradtional) axioms and hypotheses.

Bourbaki

- As latest with the rise of modern abstract algebra, the notion of axiom has entirely shifted away from truth and evidence.
- Just think of the *axioms for a group*.

Nicholas Bourbaki, 1950

It goes without saying that there is no longer any connection between this interpretation of the word "axiom" and its traditional meaning of "evident truth".

(N. Bourbaki, The Architecture of Mathematics, The American Mathematical Monthly, 57(4), 1950, p. 225)

David Hilbert

- It is one thing to call axioms hypotheses (as Riemann did) ...
- it is, however, another thing to give hypotheses the status of axioms, as Hilbert suggeted:

David Hilbert, 1922

Nothing prevents us from taking as axioms propositions which are provable, or which we believe are provable. Indeed, as history shows, this procedure is perfectly in order: [...] Riemann's conjecture about the zeroes of $\zeta(s)$, [...]

Axioms as Hypothesis

- What is the difference of treating Riemann's Hypothesis (RH) as an axiom instead of as a "usual" hypothesis?
- A mathematical statement ϕ using RH as a hypothesis would be proven in a theory T as follows:
$$T \vdash RH \to \phi.$$
- Using RH as an axiom would change the situation to:
$$T \cup \{RH\} \vdash \phi.$$
- This seems to be an easy application of the *deduction theorem*.
- However, the situation is different, as $T \cup \{RH\}$—understood as a set of *axioms*—receives an autonomous status.

Side remark: Principia Mathematica

- An illustrating example of a comparable (in fact, inverse) situation can be found in Whitehead and Russell's *Principia Mathematica*.
- PM does **not** state the *axiom of infinite* (Inf); instead, when needed, a theorem is formulated in the form $\mathrm{Inf} \to \ldots$.
- PM has, however, the *axiom of reducibility* (Red).
- Avoiding Inf as axioms, allows to stay within the *logicist* paradigm.
- Why not doing the same with Red?

  Alasdair Urquhart conjectured that the different treatment of Inf and Red is simply due to the number of instances where the respective axiom is needed; in fact, Red is needed in PM much more often than Inf.

Potentially inconsistent theories

- What is the problem with $T \cup \{RH\}$ as a set of *axiom*?
- It could be inconsistent.
- Let assume, for the sake of the argument, that it would be inconsistent.
  - In the case of $T \vdash RH \rightarrow \phi$ we would just have studied a (now) trivially true sentence.
  - In the case of $T \cup \{RH\} \vdash \phi$ we would have studied an "empty" theory, one which doesn't even has any meaningful interpretation.
- Should we believe that Hilbert suggested to study "potentially meaningless theories"?

The challenge

- Can we attribute a meaningful *semantics* to a potentially inconsistent theory like $T \cup \{RH\}$, and to our work in such a theory?
- Any usual model-theoretic semantics has to fail here!
- Our proposal: *a proof-theoretic semantics*
- Or even: *a proof semantics*

Idea of a proof semantics

Thesis

The *semantics* of a *proof* (not only the the proven formula!) can be find in the concrete derivation given in the proof.

- In particular, if we have $T \vdash \phi$ the semantics should take into account "only" the subset $T_0 \subset T$ of axioms actually used in the derivation.
- In this way, the semantics is *local* with respect to the single derivation.
- The advantage should be clear: even, if T turns out to be inconsistent, T_0 might be consistent and the performed proof can savely carried over to another axiom system.
- Thus, even in an inconsistent theory, we might have two proofs $T \vdash \phi$ and $T \vdash \psi$ with incompatable ϕ and ψ. But the proofs $T_0 \vdash \phi$ and $T_1 \vdash \psi$ for different $T_0, T_1 \subset T$ might be meaningful in different consistent subsystems of T.

Historic examples

Frege's Grundgesetze

- As much as we can judge, the last situation holds for Frege's *Grundgesetze*.
- Frege, of course, does not use, anywhere, explicitly Russell's paradox and the ex-falso-quodlibet to obtain a formula.
- In contrast, all of his derivations seem to be "locally" meaningful.
- Modern work on *different* consistent subsystems of Frege's Grundgesetze supports this view.

Reinhardt cardinal

- A Reinhardt cardinal, which turned out to be inconsistent over ZFC (but unkown to be consistent or inconsistent over ZF only!), is another example.
- Work on a Reinhardt cardinal done at the time where it was not known to be incosistent does not seem to be "meaningless" work.

Conclusion

- Following Hilbert there should be no (philosophical) difference any longer between Axiom and Hypothesis.
- Treating axiom systems, in general, as "arbitrary" sets of assumptions (hypotheses), rises the question of potentially inconsistent axiom systems.
- Such systems (may) lack a model-theoretic semantics.
- Proof(-theoretic) semantics can step in.

A joint logic of problems and propositions, a modified BHK-interpretation and proof-relevant topological models of intuitionistic logic

Sergey Melikhov
Steklov Mathematical Institute

Abstract. In a 1985 commentary to his collected works, Kolmogorov remarked that his 1932 paper ("Zur Deutung der intuitionistischen Logik") "was written in hope that with time, the logic of solution of problems will become a permanent part of [a standard] course of logic. Creation of a unified logical apparatus dealing with objects of two types – propositions and problems – was intended." We describe such a formal system, QHC, which is a conservative extension of both the intuitionistic predicate calculus, QH, and the classical predicate calculus, QC. Moreover:

- The only new connectives ? and ! of QHC induce a Galois connection (i.e., a pair of adjoint functors) between the Lindenbaum algebras of QH and QC, regarded as posets.
- Kolmogorov's double negation translation of propositions into problems extends to a retraction of QHC onto QH.
- Gödel's provability translation of problems into modal propositions extends to a retraction of QHC onto its QC + (?!) fragment, which can be identified with the modal logic QS4.

This leads to a new paradigm that views intuitionistic logic not as an alternative to classical logic that criminalizes some of its principles, but as an extension package that upgrades classical logic without removing it. The purpose of the upgrade is proof-relevance, or "categorification"; thus, if the classical conjunction and disjunction are to be modeled by intersection and union (of sets), the intuitionistic conjunction and disjunction will be modeled by product and disjoint union (of sheaves of sets). More formally, we describe a family of sheaf-valued models of QH, inspired by dependent type theory (not to be confused with the well-known open set-valued sheaf models of QH), which can be regarded as proof-relevant analogues of classical topological models of Tarski et al., and which extend to models of QHC. We also give an interpretation of some of these models in terms of stable solutions of algebraic equations; this can be seen as a proof-relevant counterpart of Novikov's classical interpretation of some Tarski models in terms of weighting of masses.

The new paradigm also suggests a rethink of the Brouwer–Heyting–Kolmogorov interpretation of QH. Traditional ways to understand intuitionistic logic (semantically) have been rooted either in philosophy – with emphasis on the development of knowledge (Brouwer, Heyting, Kripke) or in computer science – with emphasis on effective operations (Kleene, Markov, Martin-Löf). Our "modified BHK interpretation" is rooted in the usual mathematical ideas of canonicity and stability; it emphasizes simply the order of quantifiers, developing Kolmogorov's original approach. This interpretation is compatible with two complete classes models of QH: (i) Tarski topological models and (ii) set-theoretic models of Medvedev–Skvortsov and Läuchli; as well as with (iii) our sheaf-valued models.

Truth-value gaps and paradoxes of material implication

Grigory Olkhovikov
Ural Federal University

Abstract. One of the uses of multi-valued logic is to overcome some restrictions of the classical two-valued one. A particularly well-known set of restrictions is connected with analysis of conditionals in two-valued logic, where one finds so called paradoxes of material implication. One of these paradoxes is basically as follows: a conditional with false antecedent is a truth-value gap from a commonsense point of view, but is true in two-valued logic.

Resources of three-valued logic provide an obvious solution to this paradox: that is to say, interpret sentences exhibiting truth-value gap as having the third (non-designated) truth-value and define implication accordingly. Strange enough, this way out was not seriously tried until recently. In our talk we recapitulate the main results obtained thus far about a three-valued logic which realizes this kind of definition for implication together with natural definitions for disjunction, conjunction and quantifiers. We briefly address complete and consistent Hilbert-style axiomatizations for this logic (for both propositional fragment and full first-order version) and its expressive power.

Cuts and Graphs

Zoran Petrić
Mathematical Institute, SANU

Abstract. Plural (or multiple-conclusion) cuts are inferences made by applying a structural rule introduced by Gentzen for his sequent formulation of classical logic. As singular (single-conclusion) cuts yield trees, which underlie ordinary natural deduction derivations, so plural cuts yield graphs of a more complicated kind, related to trees. The graphs of plural cuts are interesting in particular when the plural cuts are appropriate for sequent systems without the structural rule of permutation. The aim of this talk, which is based on a joint work with Kosta Došen, is to define these graphs both inductively and in a pure combinatorial manner.

Rules, types and the transcendence of second order logic

Paolo Pistone
Università di Roma Tre

Abstract. As soon as formulas of the form "there exists an X such that ..." or "for all integers n ..." are involved, the justification of logical rules by harmony (i.e. cut-elimination) looks like an empty shell: the intrinsic circularity of such arguments is indeed so harmful that we cannot, without accepting some preconceived assumptions, decide between "sound" and "paradoxical" versions of the Hauptsatz for higher order logics.

The aim of this talk is to argue that the acceptance of this fundamental form of ignorance (intimately related to the incompleteness theorems) makes room for a constructive understanding of second order logic.

Counterfactual logics: natural deduction calculi and sequent calculi

Francesca Poggiolesi
CNRS-CEPERC (UMR 7304), Aix-en-Provence

Abstract. Counterfactual logics, which have a long and venerable history [3, 1, 2], have been introduced to capture counterfactual sentences, i.e. conditionals of the form "if A were the case, then B would be the case", where A is false. If we interpret counterfactuals as material conditionals, we have that all counterfactuals are trivially true and this is an unpleasant conclusion. By means of counterfactual logics, on the other hand, we can give a different and meaningful interpretation of counterfactual sentences.

There are several different systems of counterfactual logics. Amongst them we focus on the system CK and its standard extensions, namely CK $+ \{\mathrm{ID}, \mathrm{MP}, \mathrm{CEM}\}$. These systems have a simple and useful semantics. One just needs to consider a set of possible worlds W, and a selection function f; for each world i and each formula A, f selects the set of worlds of W which are *closer* to i given the information A. Thus a counterfactual sentence $A > B$ is true at a world i if, and only if, B is true at all those worlds that are closer to i given the information A.

In this talk we aim at presenting a method for generating sequent calculi and natural deduction calculi for the system CK and its extensions. The method is based on and fully exploits the simple semantics interpretation of such systems. Sequent calculi and natural deduction calculi are proved to be equivalent; moreover, as for the natural deduction

calculi, we prove that the derivations normalize; while, as for the sequent calculi, we prove they are contraction-free, weakening-free and cut-free and that their logical rules are all invertible.

References

[1] Lewis, D., *Counterfactuals*. Basil Blackwell, 1973.

[2] Nute, D., *Topics in Conditional Logic*. Reidel, Dordrecht, 1980.

[3] Stalnaker, R. A theory of conditionals, *American Philosophical Quarterly, Monograph Series no. 2*, Blackwell, Oxford, 98–112, 1968.

How do vacuous truths become laws?

Erdinç Sayan

Department of Philosophy
Middle East Technical University
esayan@metu.edu.tr

The use of idealizations and approximations is pervasive in science. Idealizations and approximations are assumptions or hypotheses which are known to be patently false. They enter into a scientific analysis or explanation in basically two ways. First, they may be conjoined to a theory as extraneous hypotheses, mainly to make it easier to work with the theory. Second, they may be embodied in the very statement or formulation of laws and theories; I call such laws idealizational laws. Thus, for example, assuming that the universe contains only two bodies is an idealizing hypothesis that may be used in some contexts as input to the law of gravitation and Newton's second law of motion. On the other hand, insofar as Newton's second law is conceived as applying only to point masses, that law contains an idealization as part of its content. In this paper I confine myself to an examination of the problems of testing pertaining to idealizational laws.

Consider the simple conception that scientific laws are unrestricted universal statements of the form "$\forall x(Fx \rightarrow Gx)$", where the arrow sign is understood to denote material implication. If we assume the adequacy of this material-conditional interpretation of laws, the following problem arises. If no object instantiates the antecedent, as when, for example, F refers to the property of being a point mass or a frictionless plane, the law "$\forall x(Fx \rightarrow Gx)$" is true "vacuously." But of course we cannot be content with our laws being true vacuously, for in such laws the consequent can be any statement whatsoever and the "law" would be true. If we adopt the material-conditional model of laws, explaining the possibility of *falsifying* idealizational laws, which are *true* (vacuously), becomes an acute problem for Popperian falsificationism, for example.

There are difficulties for the falsificationist view even if we don't treat laws as material conditionals, however, as we shall see when we look at the hypothetico-deductivist theory of confirmation. (I shall hereafter use 'h-d' as short for "hypothetico-deductivist," "hypothetico-deductivism," or "hypothetico-deductive.") The usual h-d theory faces problems in accounting for confirmation and disconfirmation of idealizational hypotheses. According to one classical formulation of the h-d account we have the following schemata for confirmation and disconfirmation of an hypothesis H:

(HD 1)

Confirmation:		Disconfirmation:	
$H \vdash e \rightarrow f$	*true*	$H \vdash e \rightarrow f$	*true*
e	*true*	e	*true*
f	*true*	$\sim f$	*true*
$\therefore H$	*may be true*	$\therefore H$	*false*

Here H may be a law or theory, '$\vdash$' stands for logical entailment, and e and f are evidential statements.[1] Notice that (HD 1) does not presuppose H to be a material conditional. Now, let H be the idealizational law known as the ideal-gas law that says "All ideal gases obey $PV = kT$." This law entails the statement, "If g is an ideal gas, then g obeys $PV = kT$," for any given object g. So the confirmatory schema takes the following shape:

All ideal gases obey $PV = kT$ $\vdash$ If g is an ideal gas, then g obeys $PV = kT$	*true*
g is an ideal gas	*true*
g obeys $PV = kT$	*true*
$\therefore$ All ideal gases obey $PV = kT$	*may be true*

Since ideal gases do not exist, the second premise of this confirmatory argument cannot ever be satisfied. The same premise occurs also in the corresponding disconfirmatory argument. As a result, the ideal-gas law turns out to be neither confirmable nor disconfirmable on the h-d account; every object is irrelevant as far as being evidence for the law.[2] This is in contrast to falsificationism: every "rigorous attempt" to test the ideal-gas law by trying to find negative instances of it would be bound to fail, and of course we already know that. The same goes of course for any contrary of the ideal-gas law, such as "All ideal gases obey $P^2/V^7 = kT^3$" or "No ideal gas obeys $PV = kT$." Hence the falsificationist method, as it stands, seems inapplicable in the case of idealizational generalizations, because trying to corroborate them is hopeless.

Testability of idealizational generalizations turns out to be a problem also for another popular theory of confirmation: the Bayesian theory. According to this theory,

[1] Cf. Braithwaite 1983, pp. 45–46. For simplicity, these two schemata leave out the auxiliary hypotheses and initial conditions which need to be conjoined with H to secure entailment of $e \to f$ by H. Also, (HD 1) needs to be supplemented with a clause saying that e does not entail f without the help of H, i.e. it is not the case that $e \vdash f$. Otherwise we would be allowing any hypothesis to be confirmed in the event that e is true and entails f; for if $e \vdash f$, then $H \vdash e \to f$ for any H. For traditional logical positivists e and f must contain only observational terms and logical constants.

[2] The h-d model is sometimes given a cruder form. Let us look at how this alternative formulation confronts the problem of idealizational laws. On this version of h-d, we have these schemata:

(HD 2)	CONFIRMATION:		DISCONFIRMATION:	
	$H \vdash E$	*true*	$H \vdash E$	*true*
	E	*true*	$\sim E$	*true*
	$\therefore H$	*may be true*	$\therefore H$	*false*

E here represents the evidence statement. If we take E to be the conditional $e \to f$ of the previous version (HD 1), we get a problem: the falsehood of e would suffice to make E true, and hence to confirm H. To illustrate this in terms of our ideal-gas example, we obtain from (HD 2):

All ideal gases obey $PV = kT$ $\vdash$ If g is an ideal gas, then g obeys $PV = kT$	*true*
If g is an ideal gas, then g obeys $PV = kT$	*true*
$\therefore$ All ideal gases obey $PV = kT$	*may be true*

Thus every object g which is not an ideal gas would satisfy the second premise and hence its observation would confirm the ideal-gas law. Since in fact no object is an ideal gas, it follows that observation of *any* object constitutes confirming evidence for the ideal-gas law. This is in stark contrast to (HD 1), according to which *no* object has, as we have seen, any confirmatory or disconfirmatory relevance to that law.

incremental confirmation occurs just in case the posterior probability of a hypothesis is greater than the prior probability of that hypothesis on a given piece of evidence. So we have incremental confirmation of hypothesis H by evidence E if and only if $\Pr(H \mid E \,\&\, B) > \Pr(H \mid B)$, where B is the background knowledge (or conjunction of background assumptions). And incremental disconfirmation of H by E occurs on the Bayesian account if and only if $\Pr(H \mid E \,\&\, B) < \Pr(H \mid B)$.[3] Now, if our background knowledge B gives us strong reasons to believe that H is an idealizational hypothesis, i.e. a generalization that refers to nonexistent entities such as ideal gases, frictionless surfaces, perfect vacuum conditions, uniform electromagnetic fields and the like, then B tells us that we cannot find any actual evidence for H. That is to say, we know that we can never find ideal gases, perfect vacuums or perfectly elastic springs to serve as conditions or media to gather evidence for confirming or disconfirming H. It follows that testing of idealizational generalizations is not possible by the standard Bayesian criteria of incremental confirmation and disconfirmation. On the other hand, if we were to stick to the material-conditional model of laws, our background knowledge would entail that idealizational generalizations are (almost certainly) vacuously true. This means $\Pr(H \mid B) \approx 1$, which is to say that H is "absolutely confirmed" in Bayesian terminology. Hence, the Bayesian theory would grant every idealizational universal the status of very highly or absolutely confirmed law—whether the universal makes any sense or not. But of course some idealizational generalizations are confirmed while others are *dis*confirmed in science.

So how are idealizational generalizations tested? Let us consider this answer: They are tested by turning them into approximate generalizations. Here is a suggestion about how to do it. Consider once again our generic idealizational generalization[4]:

(IL) $\forall x(Fx \to Gx)$

Instead of interpreting our generalization as stating that everything that is F is also G, let us interpret it more loosely as asserting "Every x which is approximately F is also approximately G." And let us express this approximate law symbolically as:

(AL) $\forall x\,(\approx Fx \to \approx Gx)$

where '$\approx Fx$' stands for "x is approximately F," and similarly for '$\approx Gx$.' Even if there exist no objects in the world that are F, there may exist objects that are approximately F. Thus let us assume that there are no perfect vacuums in the world as a matter of nomological fact, and that therefore the following law of falling bodies belongs to the category of idealizational law:

(IFB) All bodies falling in perfect vacuum at the surface of the earth fall a height given by $h = \frac{1}{2}gt^2$.

g above is the gravitational acceleration at the surface of the earth and t is the duration of the fall. In accordance with (AL), we can interpret this law as asserting the following:

[3]There is neither confirmation nor disconfirmation of H by E if and only if $\Pr(H \mid E \,\&\, B) = \Pr(H \mid B)$.

[4]My use of the material conditional sign '$\to$' hereafter is only for expository purposes and does not reflect a commitment to the material-conditional interpretation of laws on my part. If one wants, one could read it as a stronger conditional than material conditional, say, as a "nomological conditional."

(AFB) All bodies falling in an approximate vacuum at the surface of the earth fall a height given by $h \approx \frac{1}{2}gt^2$.

We can then test this approximative version of the law, so the suggestion goes, in the usual way—whether in accordance with the h-d model, the falsificationist recipe, the Bayesian account or some other theory of confirmation.

It may be that a tacit assumption on the part of most philosophers that an idealizational generalization can easily be converted into an approximative hypothesis in this fashion is what has been responsible for those philosophers' neglect of a careful study of how idealizational laws and theories get confirmed. However, this simple way of converting an idealizational generalization into an approximative one is not unproblematic. An obvious problem is determining how approximate is approximate enough in a given context. How are we to decide what should count as an approximate vacuum, for example, and how much discrepancy between the values of h and $\frac{1}{2}gt^2$ should we allow to count $h \approx \frac{1}{2}gt^2$ as holding? In short, what should be the criteria for deciding when $\approx Fx$ and $\approx Gx$ are true and when they are false? In the case of the law of falling bodies, feathers and parachutes will *always* fall more slowly than stones, unless the medium of fall is a perfect vacuum and not a merely approximate one. But a perfect vacuum is not attainable (or so we are assuming). So, if we want the approximative version of the law of falling bodies (AFB) to be confirmed both by feathers and stones, we need some criteria that would qualify a medium as approximately a vacuum and at the same time ensure that $h \approx \frac{1}{2}gt^2$ is satisfied by both stones and feathers. That doesn't seem to be a feasible thing to do.

I think there is a more satisfactory way of confirming the idealizational law of falling bodies (IFB) while effectively disconfirming its contraries, than turning it into the approximative law of falling bodies (AFB). Our confidence in (IFB) actually derives from the fact that, for any given object, the *more* the conditions of its fall approximate to being in perfect vacuum, the *more closely* it obeys $h = \frac{1}{2}gt^2$. As we experimentally make the medium come closer and closer to a vacuum, we will observe that h for feathers as well as for stones will approach $\frac{1}{2}gt^2$. In the case of feathers, closing of the gap between h and $\frac{1}{2}gt^2$ will be more drastic than stones, since stones more closely obey $h = \frac{1}{2}gt^2$ to begin with. But both feathers and stones will converge to $h = \frac{1}{2}gt^2$ as a common limit. (And the difference in their rates of convergence can plausibly be accounted for in terms of the differences in the air resistance they experience.) If the opposite occurred, i.e. progressively closer approximations to perfect vacuum conditions did not result in progressively closer approximations of h to $\frac{1}{2}gt^2$ for some objects, this would prompt us to look for an explanation of this misbehavior. Should our attempts to explain the misbehavior fail, we come to regard the law as disconfirmed.

It seems that part of the meaning of an idealizational law (indeed of *most* laws, if not all) is that, as the property in its antecedent is more closely instantiated by an object, the property in its consequent is better approximated by that object. The counterpart of this is also implicit in what the law asserts: If closer instantiation of the antecedent-property of the law by any object *failed* to lead to closer approximation of its consequent-property by that object, then the law would fail, i.e. would be disconfirmed. Thus I propose the following modification on (AL), which gives us the "progressively approximative" interpretation of an idealizational law:

(PAL) $\forall x\,(\approx\uparrow Fx \rightarrow \approx\uparrow Gx)$

Here '$\approx\uparrow Fx$' stands for "x approximates progressively more to being an F" and '$\approx\uparrow Gx$' stands for "x approximates progressively more to being a G." Confirmatory instances of this reading of an idealizational law are those cases where making the experimental conditions more closely approximate to the idealizational antecedent F results in improved approximation of actual observations to G. And disconfirmatory instances of it are those where improved approximation to F does not make actual observations more closely approximate to G.

Sometimes, after observing an object in one medium, instead of letting it actually fall in another medium which more closely approximates to a vacuum than the first one, we may have reason to think that *had* we let it fall in this second medium, it *would have* come closer to obeying $h = \frac{1}{2}gt^2$. When such counterfactual reasoning is cogent, we tend to count the law as confirmed by the object.[5]

The relation between the counterfactual

> If an object had been let to fall in medium M_2, which more closely approximates to a vacuum than medium M_1, it would have come closer to obeying $h = \frac{1}{2}gt^2$

and the (PAL) reading of the idealizational law of falling bodies is noteworthy in the following respect. The truth of a relevant counterfactual is usually thought to testify to the lawlikeness of an ordinary, nonidealizational generalization. (We say that a lawlike generalization "supports" its relevant counterfactuals.) But the truth of the counterfactual above is not only an indication of the lawlikeness of (PAL), but also provides a confirmatory argument for it.

Here is how another idealizational law, the ideal-gas law, can be confirmed in accordance with its (PAL) form. Some of the properties of an ideal gas are that its molecules are point masses, and those molecules engage in totally elastic collisions with the walls of the vessel. Now take two real gases: radon and helium. We know that helium comes closer in its relevant properties to an ideal gas than radon does. The molecules of helium are closer to being point masses, for example, than the molecules of radon, which are much bigger than helium molecules. The collisions between the walls of the vessel and helium molecules are closer to being elastic collisions than are the collisions between the walls of the vessel and radon molecules. And so on. As a result, the equation of state for helium better approximates to $PV = kT$, the equation of state for an ideal gas, than does the equation of state for radon. The ideal-gas law can thus be confirmed by *real gases*! This is compatible with saying that this law is actually a false description of the behavior of real gases. For the ideal-gas law does not, strictly speaking, purport to correctly describe a real gas—it is about ideal gases.[6]

There is an intriguing question in connection with (PAL). Could it turn out for some laws that no object can even approximately—much less progressively-approximately—instantiate the antecedent of it? In other words, what if there exists no object which, for some F, can (progressively) approximate F? Now, if (progressive) approximation to F

[5]The analysis in the previous paragraph and this one, as well as some of the other ideas expounded in this paper, carry a family resemblance to Ronald Laymon's theory of confirmation in Laymon 1985. I am grateful to him for his comments and suggestions some years back on the issues discussed in this paper.

[6]The behavior of real gases confirms the ideal-gas law by another way also. When a real gas in a closed vessel becomes diluted after we expand the volume of the vessel, its equation of state comes closer to $PV = kT$. Diluting the gas decreases the forces of cohesion among its molecules, thus bringing it closer to an ideal gas, which is assumed to lack intermolecular forces of cohesion.

cannot be instantiated by any object in the world, this would mean that the domain of the law consists of purely fictional objects. Unlike idealized objects, fictional objects are the kind of things to which real objects cannot meaningfully be said to approximate.[7] Nonetheless, some universals which are about purely fictional entities appear to be of some scientific use. Laws about epicycles and deferents in Ptolemaic astronomy (which did have a measure of predictive power), or about virtual images in optics might be examples of such laws.[8] If there are genuine fictional laws, then, how are they confirmed? Conversion to their (PAL) versions won't do, since fictitious entities, as we described them, cannot be (progressively) approximated by real objects.

It seems to me that confirmation or disconfirmation of a fictional law will have to be in a different manner than the way idealizational generalizations are confirmed or disconfirmed. Clearly, a fictional law will have to be tested through the observational implications to be somehow elicited from it. But once a fictional law yields an observational consequence of the form $e \rightarrow f$, our progressively-approximative approach could pick up from there, and assert that the fictional law in question receives confirmation if better approximation of experimental conditions to e brings actual observations closer to f.

There is a complication that needs to be mentioned in the confirmation and disconfirmation procedures for idealizational laws à la the progressively-approximative approach we described. In many cases when we are testing an idealizational law, we employ a number of approximations and other idealizations than the one(s) that occurs in the antecedent of the idealizational law. Take the testing of the law of ideal (or simple) pendulum, for example, which states:

All ideal pendulums have a period of oscillation given by $T = 2\pi(l/g)^{1/2}$,

where l is the length of the suspension string. The assumptions that the pendulum bob is a point mass and that the suspension string is weightless are part of what it is to be an ideal pendulum, and hence are "internal" to the antecedent of the law. When experimentally testing the law, we would typically employ a number of "external" assumptions as well, such as that there is no air friction that would delay the swing; that g, the gravitational acceleration parameter, remains constant during the swing; that the ambient temperature does not affect the period; and so on. All these approximations and idealizations would introduce errors that might conspire in such ways that our efforts to bring real pendulums closer and closer to an ideal pendulum (by decreasing the weight of their strings or making their bobs smaller and smaller) do not result in their measured periods coming progressively closer to satisfy $T = 2\pi(l/g)^{1/2}$. In other words, increased agreement between our experimental pendulums and the ideal pendulum may not produce increasingly better agreement between our measurements of the periods of them with the consequent of the law. If such nonmonotonous and erratic behavior is only local and can convincingly be explained in terms of the conspiracy of

[7]Nancy Cartwright marks the distinction between idealizational and fictitious entities as follows:

> There are the obvious idealizations of physics—infinite potentials, zero time correlations, perfectly rigid rods, and frictionless planes. But it would be a mistake to think entirely in terms of idealizations—of properties which we conceive as limiting cases, to which we can approach closer and closer in reality. For some properties are not even approached in reality. They are pure fictions. (Cartwright 1983, p. 153)

[8]Virtual images are surrogates for real objects whose role is basically to simplify the mathematics involved (Laymon 1994, p. 36).

the errors introduced by the internal and external idealizations and approximations involved, the idealizational law would be regarded as confirmed. If the nonmonotonicty and nonconvergence in the experimental fit cannot be attributed to local conspiracies of errors, however, and can best be explained by the falsity of the idealizational law being tested, then the law will be regarded as disconfirmed.[9]

References

Braithwaite, R. B. (1983), "The Structure of a Scientific System," in Peter Achinstein (ed.), *The Concept of Evidence*, New York: Oxford U. P., pp. 44–52. (First published in R. B. Braithwaite, *Scientific Explanation*, Cambridge: The University Press, 1953, pp. 12–21.)

Cartwright, Nancy (1983), *How the Laws of Physics Lie*. Oxford: Clarendon Press.

Laymon, Ronald (1985), "Idealizations and the Testing of Theories by Experimentation," in Peter Achinstein and Owen Hannaway (eds.), *Observation, Experiment, and Hypothesis in Modern Physical Science*, Cambridge: The MIT Press, pp. 147–173.

Laymon, Ronald (1994), "Demonstrative Induction, Old and New Evidence and the Accuracy of the Electrostatic Inverse Square Law," *Synthese* 99, 23–58.

[9] I am grateful to some members of the audience at the Conference for their questions and comments.

Scientific modeling:
A two-layer based hypothetical reasoning

Guillaume Schlaepfer

Conference on Hypothetical Reasoning
August 23–24, 2014 in Tübingen, Germany

FACULTÉ DES LETTRES
DEPARTEMENT DE PHILOSOPHIE

UNIVERSITÉ
DE GENÈVE

Deductive-nomological explanation (DN model)

General laws

Particular circumstances

Particular facts or less general law

$$\left.\begin{array}{c} L_1, L_2, \ldots, L_r \\ \underline{C_1, C_2, \ldots, C_k} \end{array}\right\} \textbf{Explanans sentences}$$

$$E \qquad \textbf{Explanandum sentence}$$

Laws for Hempel:
regularity + additional criterion
Exceptionless generalizations

Hempel, C. G., 1965, Aspects of Scientific Explanation

Some criticisms against the DN model

Explanatory irrelevance:
(Salmon 1971, Kyburg 1965)

L: All males who take birth control pills regularly fail to get pregnant
C: John take birth control pills

E: John fails to get pregnant

→ DN model insufficient

Lawless explanations:

Special sciences (economics, psychology, biology…) have no universal laws

→ DN model unnecessary

Ceteris paribus laws (cp-laws)

Not so many universal laws in (special) sciences…

…lots of restricted laws:

> "ceteris paribus, doing physical activities increases life span"

Semantically obscure:
no explicit description of the ceteris paribus clause (cp-clause)

→ Hard to know when to apply them or to make predictions

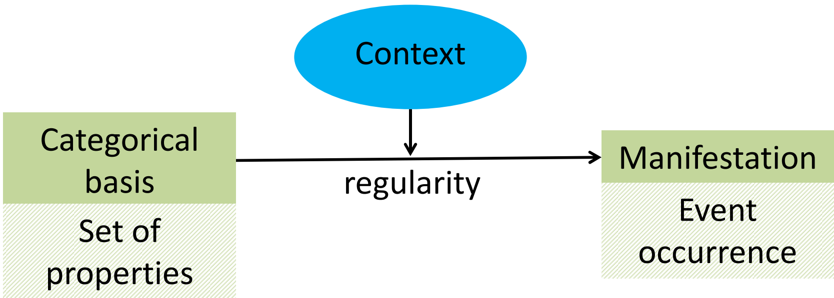
2 types of cp-clauses:

- Comparative: "other things being equal"
- Exclusive: "If no other relevant factor impinges"

Dispositional view of cp-laws

Cp-Law as manifestation of capacities/disposition (Cartwright 1989):

- Capacities underlie causal claims
- Causal factor acting across all context
- Not always manifested: depends on other capacities
- Explains why laws also explanatory in non-ideal context



"What is thought to be an exception to a principle is always some other and distinct principle cutting into the former" (Mill, 1836)

Dispositional view of cp-laws

2 degrees of generality/modality:

Causal law:

- supported by existence of a regularity
- relative to particular context

Capacity ascription:

- tendency associated with some properties/factor
- occurs in all context

Laws of nature and system laws

Schurz 2002 comes to a similar result for laws in physics:

Law of nature :
- unrestricted/universal law
- Total force law (2nd Newton's principle), special force laws
- Not applicable per se (would require specification of all active forces)

System law:
- Refer to specific system, need exclusive cp-clause
- Specification of all forces: boundary conditions
- Kepler law of planetary orbits, law of free fall etc.
- Often assume simplificatory assumptions
- More contingent, but still law-like generalizations (support counterfactuals)
- Typical law of special sciences

→ System laws akin to theoretical models

A two-layers based hypothetical reasoning

System laws/causal models: separable into two layers of laws

- Dispositional laws:
  - ➢ Exclusive cp-laws
  - ➢ Take real properties as argument
- Rules of composition
  - ➢ Take dispositional laws as argument

Paradigmatic case: Newtonian forces
- Dispositional laws: special force laws
- Rule of composition: vector addition

A two-layers based hypothetical reasoning

Dynamics of causal modeling:

- Choice of dispositions based on available properties (process under study and context) and explanatory goal
- Establishment of composition rule based on set of disposition
- Analysis of model's consequences etc…

Sources of error:

- False disposition (supposing a false inference from a set of properties)
- Irrelevant disposition (related to irrelevant properties)
- Missing disposition
- False composition rule

A two-layers based hypothetical reasoning

Problem of ad hoc hypotheses:

1. recognizing genuine capacities
2. Identifying minimal set of capacities (proper inferential role → 1.)

Some candidate criteria for genuine capacity:

- Preservation of explanatoriness over contextual change
- Production of new predictions
- Independent experimental manipulability
- Law resulting from agent based simulation
- Part of a minimal explanatorily exhaustive set of dispositions
- Theorems of second order logic? Non-monotonic logic?

Back to DN model

Explanatory irrelevance:
(Salmon 1971, Kyburg 1965)

L: All males who take birth control pills regularly fail to get pregnant
C: John take birth control pills

E: John fails to get pregnant

→ DN model insufficient

Lawless explanations:

Special sciences (economics, psychology, biology…) have no universal laws

→ DN model unnecessary

Back to DN model

Problem of lawfulness in special sciences:
What is special about physics then? Not that it does not offer knowledge about powers or capacities but rather that it has been able to establish other kinds of knowledge as well, knowledge that we can couple with our knowledge of capacities to make exact
predictions.
This additional knowledge is primarily of two kinds: 1) We know for the powers of physics when they will be exercised (e.g., a massive object always attracts other masses); and 2) we have rules for how to calculate what happens when different capacities operate together (e.g., the vector addition law for forces). This kind of knowledge is missing for many other subjects. That is why they cannot make exact predictions.

(Cartwright 2002, 6)

[…] non-physical sciences do not have laws of nature of their own, but they do have system laws of their own”

(Schurz 2002, 368)

Back to DN model

Explanation based on system laws

Problem of lack of lawfulness:
- Lawfull although more contingent/not exceptionless/idealizations
- Lawfulness supported by integration of universal laws/capacities
- Exclusive cp-laws might be defined theoretically (case in physics)

Problem of relevance:
System law should provide explanation only in situation displaying the relevant capacities and boundary conditions

Thanks for your attention!

Bibliographic references:

Cartwright, Nancy (1994, [1989]): Nature's Capacities and their Measurement.

Cartwright, Nancy (2002): In favor of laws that are not ceteris paribus after all. In *Erkenntnis* 57 (3), pp. 425–439.

Hempel, Carl Gustav (1966): Philosophy of natural science: Prentice-Hall Englewood Cliffs, NJ.

Kyburg, H.: Discussion: Salmon's Paper. In *PHILOS SCI* (32), pp. 147–151.

Salmon, Wesley C. (1971): Statistical Explanations. In Wesley C. Salmon, et al (Eds.): Statistical Explanation and Statistical Relevance, pp. 29–87.

Schurz, Gerhard (2002): Ceteris paribus laws: Classification and deconstruction. In *Erkenntnis* 57 (3), pp. 351–372.

guillaume.schlaepfer@unige.ch

FACULTÉ DES LETTRES
DEPARTEMENT DE PHILOSOPHIE

UNIVERSITÉ DE GENÈVE