Abductive Conceptual Engineering

Christian J. Feldbacher-Escamilla

Spring 2020

Project Information

Publication(s):

- Feldbacher-Escamilla, Christian J. (under revision). "Conceptual Engineering and Abduction". In: *manuscript*.

Talk(s):

- Feldbacher-Escamilla, Christian J. (2020b-01-17). *Abductive Conceptual Engineering*. Research Seminar. Presentation (invited). Lunchtime Talk. University of Pittsburgh: Center for Philosophy of Science.

Project(s):

- Visiting fellowship at the *Center for Philosophy of Science, University of Pittsburgh*; Spring 2020;
- Research grant of the *Fritz Thyssen Foundation*; Spring 2020;
- DFG funded research unit *Inductive Metaphysics* (FOR 2495); subprojects: *Creative Abductive Inference and its Role for Inductive Metaphysics*.

Introduction

Conceptual engineering is about improving our representational devices.

Standard view: Carnap's approach was too narrow.

Problem: new approach is pretty much meta-theoretical, not so much about how the engineering is done.

E.g., Cappelen (2018, p.199): "I've given you a theory of conceptual engineering without concepts and without engineering."

This talk has two aims: first, to show that Carnap was basically on the right track; and second, to identify abductive concept formation as a tool of conceptual engineering

The first part is instrumental for the second.

Contents

Explication and Conceptual Engineering

Conceptual Engineering

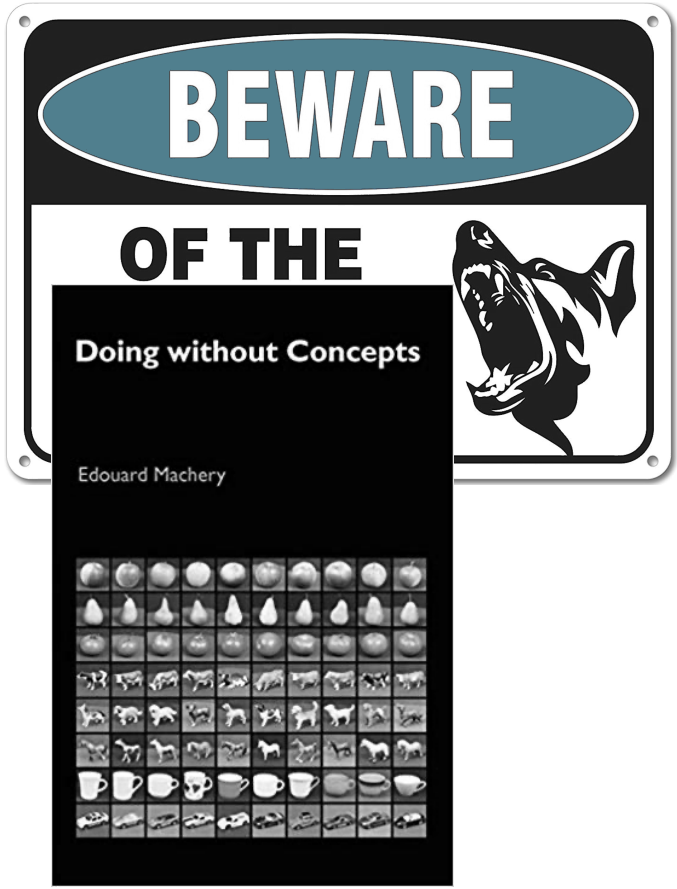
We distinguish:

- conceptual explanation
- modest explication
- full-blown explication
- conceptual engineering

We start with modest explication.

$\Rightarrow$ conceptual explanation as a particular "instance"

Concepts



Also Cappelen (2018) is very critical.

His austerity approach also dispenses concepts.

However, what all need/use is:

- expression
- intension
- extension

Concept: $\langle expression, intension, extension \rangle$

From time to time we will add uses, governing principles, subjects/topics

Concepts: Convenient Symbolism

Since explication is about revision, we have an initial concept (explicandum):

$$C^A = \langle expression^A,\ intension^A,\ \underbrace{extension^A}_{E^A},\ use^A,\ principles^A,\ topics^A \rangle$$

And a resulting concept (explicatum):

$$C^\Omega = \langle expression^\Omega,\ intension^\Omega,\ \underbrace{extension^\Omega}_{E^\Omega},\ use^\Omega,\ principles^\Omega,\ topics^\Omega \rangle$$

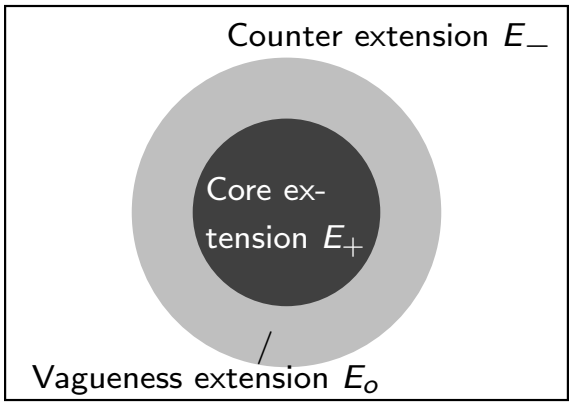
Fruitful, independently of a particular explicative interpretation of A, Ω. E.g. linguistic relation of homonymy $(expression^A = expression^\Omega,\ intension^A \neq intension^\Omega)$

$\Rightarrow$ Chalmers' subscript gambit (verbal disputes)

Or regarding concept abandonment: $use^\Omega = \emptyset$

Concepts: Convenient Symbolism

Regarding the extension E of a concept we differentiate: E_+, E_-, E_o



Concept *fish*: trouts $\subset E_+$, humans $\subset E_-$, whales $\subset E_o$

Modest Explication

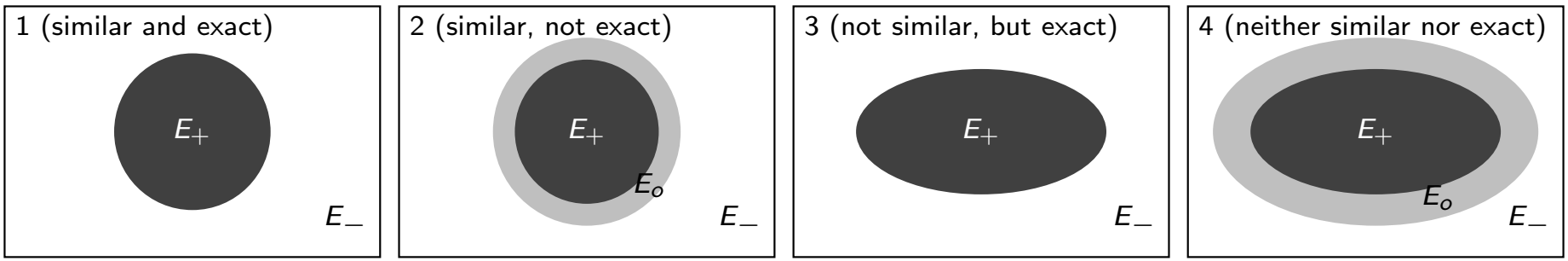
Carnap on explication:

> *"The task of explication may be characterized as follows. If a concept is given as explicandum [C^A], the task consists in finding another concept as its explicatum [C^Ω] which fulfils the following requirements to a sufficient degree.*
>
> 1. *[**Similarity:**] The explicatum is to be similar to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted.*
> 2. *[**Exactness:**] The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an exact form, so as to introduce the explicatum into a well-connected system of scientific concepts.*
> 3. *[**Fruitfulness:**] The explicatum is to be a fruitful concept, that is, useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept).*
> 4. *[**Simplicity:**] The explicatum should be as simple as possible; this means as simple as the more important requirements (1), (2), and (3) permit."*

Modest Explication

Interpretation in terms of modest explication:

$$\underbrace{E_+^A \subseteq E_+^\Omega}_{\text{core-extensional correctness}} \quad \& \quad \underbrace{E_-^A \subseteq E_-^\Omega}_{\text{counter-extensional correctness}} \quad \& \quad \underbrace{E_o^\Omega \subset E_o^A}_{\text{reduced extensional vagueness}}$$



Main idea: An explication is adequate only, if the conceptual core (E_+, E_-) is preserved and vagueness reduced (E_o); modest, because *modest* revision

Examples: 1 is not prime, Frege-style treatment of: 'The present king of France is bald.'

Conceptual Explanation

Making things explicit

Main idea of conceptual explanations: show more or less hidden implications.

Important: $C^A = C^\Omega$

However, our grasp of them is different (improved from C^A to C^Ω)

Our "model" of modest explications fits also such explanations:

Re-interpret E_+, E_-, E_o not as extensions, but as our *mastering* of them.

In this interpretation E_o is governed by "hidden" implications.

Full-Blown Explication

As is well known, Carnap allowed for more, even shifts in the conceptual core: from E^A_- to E^Ω_+ and from E^A_+ to E^Ω_-

He suggested two steps:

1. conceptual clarification (explanation and modest explication)
2. revision

Important constraint: not any such shifts, but only fruitful ones.

However, there is an important tension: what keeps changing concepts still aligned with the same topics?

Strawson's Critique

Strawson:

> *"To offer formal explanations of key terms of scientific theories to one who seeks philosophical illumination of essential concepts of non-scientific discourse, is to do something utterly irrelevant—is a sheer misunderstanding, like offering a textbook on physiology to someone who says (with a sigh) that he wished he understood the workings of the human heart. [...] To do this last is not to solve the typical philosophical problem, but to change the subject."* (Strawson 1963, p.505)

So, the worry is that by changing *intension*, *extension*, one also changes *topic*:

$$topic^{A} \neq topic^{\Omega}$$

This is, what makes the similarity requirement so important.

Similarity

Similarity underwent some changes:

Conceptual explanation: $E^A = E^\Omega$

Modest explication: $E^A_{+/-} \subseteq E^\Omega_{+/-}$

Full-blown explication:

1. $\#E^A_{+/-} \cap E^\Omega_{+/-}$ is large enough
2. $\#E^A_{+/-} \cap E^\Omega_{+/-} > 0$ (overlap, cf. Carnap 1963)
   Problem of $\nexists$ domains: set-theoretical reconstruction of numbers/points
3. E^A *maps*-1:1-*to* E^Ω (extensional isomorph., cf. Goodman 1951/1977)
   Problem: only a cardinality constraint

Similarity: Extensional Isomorphism

C^A: *the current president of the United States*

C^Ω: *the current president of Russia*

$$\#E_+^A = \#E_+^\Omega$$

Conceptual Engineering

Conceptual engineering is about widening *explication* in all dimensions.

$$\mathit{explication}^{A} \quad \Rightarrow \quad \underbrace{\mathit{explication}^{\Omega}}_{\mathit{conceptual\ \ engineering}}$$

So-to-say: fullest-blown explication

Dimensions:

- arity of the relation: binary: revision; unary: abandonment or elimination; but also introduction;
- relata: concepts, sets of concepts, representational devices in general, principles, theories, methodologies
- conditions of adequacy: anything which leads to an improvement

Important: not *any* improvement

Conceptual Engineering: Ad improvement

Not any improvement counts:

Rather, improvement in form of change of extension *via* change of intension

Conceptual Engineering: Extension

Proponents:

- Georg Brun . . . classically oriented equilibrium-thinking
- Alexis Burgess and David Plunkett . . . *conceptual ethics*
- Herman Cappelen . . . austerity approach and stressing Strawson
- Sally Haslanger . . . activism: *revisionary projects*, *ameliorating projects*
- Edouard Machery . . . *naturalised conceptual analysis* (vs. problem of method of cases, rather: X-Phi)

Ad *explication*: a species of conceptual engineering

> *"Carnap's notion of explication, however, is narrower than the activity I'm interested in. He recognizes only one kind of deficiency, 'inexactness'. Similarly, improvements for Carnap are also of a specific kind [namely his four conditions]." (cf. Cappelen 2018, sect.6.1)*

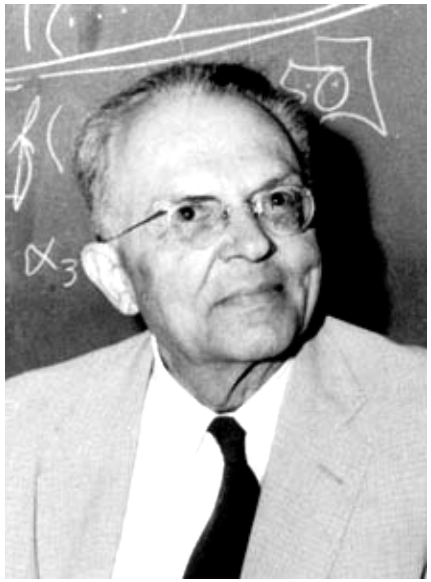
Whatever serves as a standard for improvement, serves also for adequacy.

Conceptual Engineering: Back to the Roots

Conceptual engineering: no discussion of general conditions of adequacy.

Suggestion: wider reading of Carnap provides some general structure:

- **Exactness:** CP "The characterization of $[C^{\Omega}]$ is to be given in an *exact* form."
- **Simplicity:** CP C^{Ω} "should be as *simple* as possible;"
- **Fruitfulness:** C^{Ω} "is to be a *fruitful* concept, that is, useful for the [purpose at hand]."
- **Similarity:** "The explicatum $[C^{\Omega}]$ is to be *similar to the explicandum* $[C^{A}]$ in such a way that $[C^{A}$ and C^{Ω} are about the same subject/topic]."



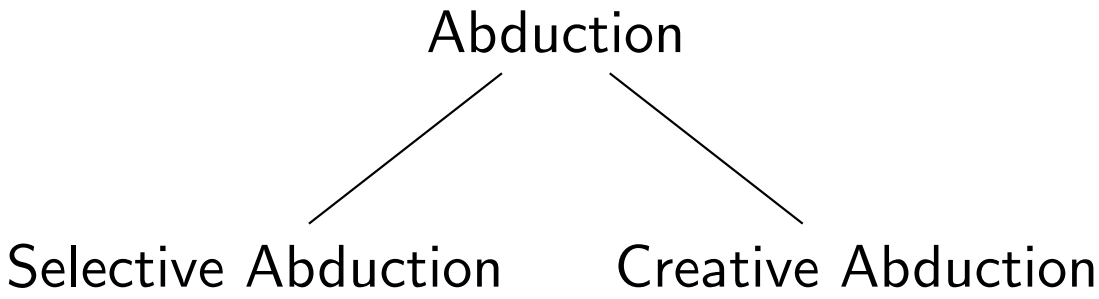
Particularly fruitfulness (allowing for change) and similarity (upholding subject-relatedness) are to be balanced (tension)

Carnap *brought home* . . .

Creative Abduction

Different Forms of Abduction

Abduction is an important inference method in science.



- **Selective Abduction (IBE):** aims at determining the best hypothesis from a set of available candidates (Lipton 2004; Niiniluoto 1999)
- **Creative Abduction:** inference method for generating hypotheses featuring new theoretical concepts on the basis of empirical phenomena (Douven 2018; Schurz 2008)

IBE: balancing of accuracy (fit) and simplicity

A similar heuristics applies also to creative abduction.

Creative Abduction: Main Idea

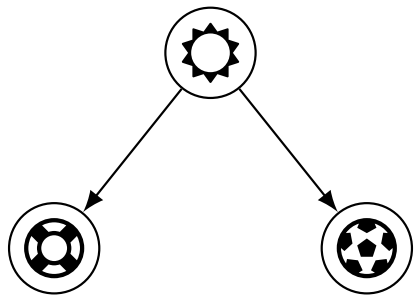
Assume you have two random variables ⚽ and ⚽ and you observe a correlation among them:

$$P(\text{⚽}|\text{⚽}) > P(\text{⚽})$$

Now, assume you cannot account for one by help of the other:

$$\text{Neither } \text{⚽} \longrightarrow \text{⚽}, \text{ nor } \text{⚽} \longrightarrow \text{⚽}$$

The way to go is to assume that there is something causing both: some X



This is in accordance with Reichenbach's principle of the common cause.

Creative Abduction: Bayesian Networks

This idea can be generalised to network structures: Bayesian networks.

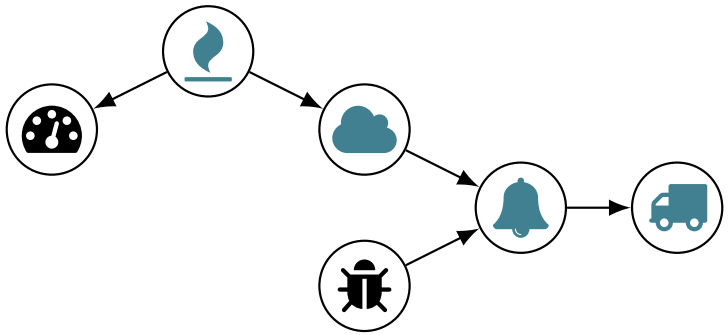
A Bayesian network consists of:

- a set of nodes/variables: $V_1, \ldots, V_n$
- edges between some of the variables: $\longrightarrow$
- and a probability distribution over the set: P

Most important condition: generalisation of the Reichenbach principle:

Markov Condition

Variables are independent of their non-descendants, conditional on their parents.

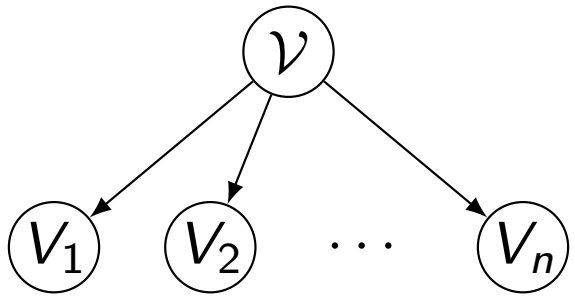


Creative Abduction: Necessary Conditions

Assume, $V_1, \ldots, V_n$ are correlated.

Assume there are no direct causal relations between them.

Creative abduction: introduce a common cause $\mathcal{V}$:



Two conditions are necessary for "probability flow" (cf. Feldbacher-Escamilla and Gebharter 2019; Dardashti et al. 2019):

i. P over $\mathcal{V}$ is non-extreme: $0 < P(\mathcal{V}) < 1$

ii. Positive dependence from $\mathcal{V}$ to the V_is: $P(V_i|\mathcal{V}) > P(V_i)$

Creative Abduction: Virtue

From i and ii one can deduce the correlations.

Now, observe the performance: Given n correlated V_is:

- # unified statements: $\binom{n}{2} = n \cdot (n-1)/2$ possibilities to combine the V_is
- # unifying statements: $n+1$ (conditions i and ii)

Suggested measure for unification:

$$u(n) = \frac{\#\ unified}{\#\ unifying} - 1$$

intro of new concept is $\boxplus$ vs. $\square$ vs. $\boxminus$ is intro of new concept

Regarding our example:

$$u(n) = \frac{n \cdot (n-1)}{2 \cdot (n+1)} - 1 \approx \ linear\ growth$$

Taking into consideration full information (e.g. cond. dependencies: $P(V_1|V_2, V_3) > P(V_1|V_2)$)
$\Rightarrow$ exponential growth

Abductive Conceptual Engineering

Engineering Tools

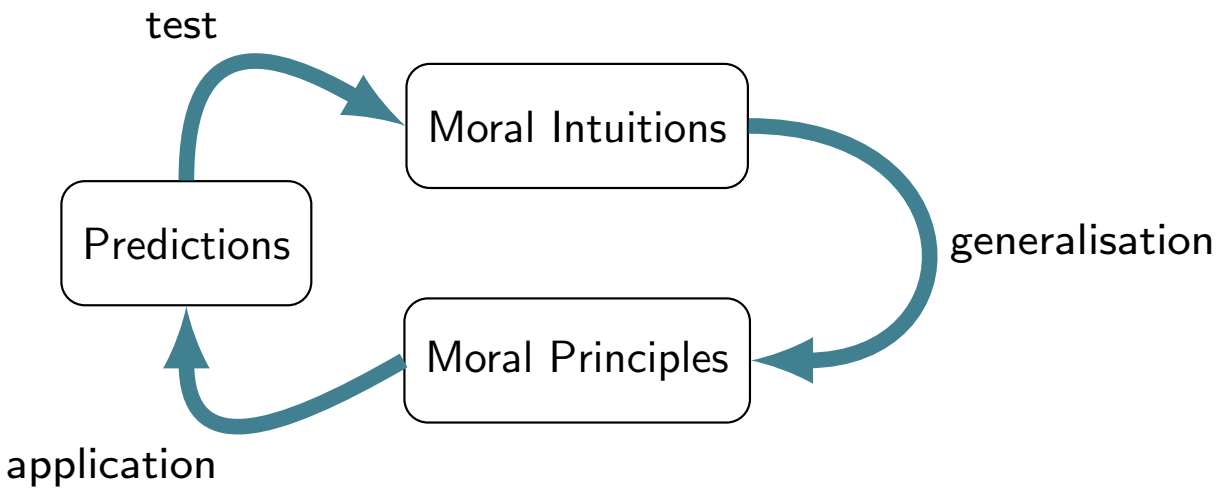
Let us briefly consider classical engineering tools and check whether they relate to the widened Carnapian conditions.

This serves as another indicator in favour of the claim that these conditions broadly cover conceptual engineering.

We then argue that abductive concept formation is also aligned with them, and, hence, a conceptual engineering tool.

Reflective Equilibirium

Rawls style engineering:



Already before: Goodman style engineering: Inductive . . .

Goodman's extensional isomorphism ⇒ cardinality ⇒ consequence:

- holism (set of notions)
- dynamics vs. Carnapian linear structure from step 1 to step 2.

Ad Carnapian conditions: similarity by 🔃; fruitfulness by *generalisations*

Definitions

Reflective equilibrium is a method for re-engineering: interplay: C^A–C^Ω

Another important (and most classical) engineering tool: definitions:

> *"A good definition of a word must include all entities which are always denoted and must exclude all entities which are never denoted by the word. [...] A good definition should extend the use of the word by dealing with objects not known or not dealt with in ordinary language." (Menger 1943, p.4)*

Also Frege (1979, p.33): "fruitfulness is the acid test of concepts"

Definitions serve both ends, revision and introduction.

However, as a general engineering tool they are too demanding (necessary and sufficient conditions).

Ad Carnapian conditions: obvious

Creative Abduction

We propose creative abduction as another important engineering tool.

Recall, the Bayesian network treatment brings about interesting unifications.

As we will see, the holistic factor widening the explicatory approach to conceptual engineering plays an important role for interpreting creative abduction within this framework.

Let us see some examples!

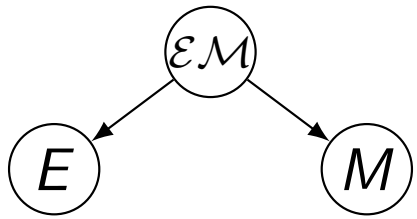
Creative Abduction: Electromagnetism

Assume that an empirical correlation between two dispositions is found:

E ... producing electricity when moved along a wire

M ... attracting iron

According to creative abduction, one accounts for it by assuming a common cause/feature: $\mathcal{EM}$ ... electromagnetism



Introducing $\mathcal{EM} \Rightarrow$ unification $\Rightarrow$ improvement of the system of concepts

Creative Abduction: Electric Dipole Structure

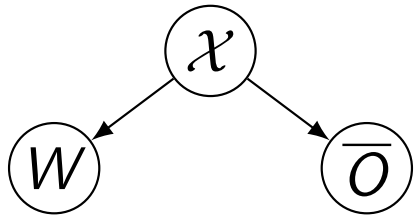
Another example stems from chemistry (cf. Schurz 2008, p.224).

An empirical correlation is established between:

W . . . solubility in water

$\overline{O}$ . . . non-solubility in oil

In accordance with creative abduction



It turned out that $\mathcal{X}$ can be described in detail: electric dipole structure

Creative Abduction: Metaphysics of/within Science

Abductive concept formation happens not only in natural science.

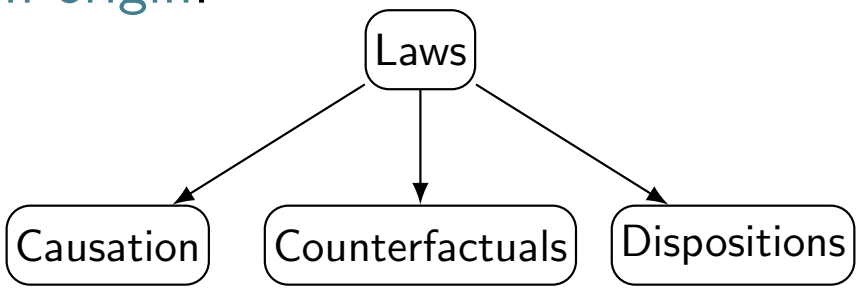
Something similar can be also observed in philosophy.

Example: Maudlin's account to the metaphysics of/within science, relating: laws, counterfactuals, dispositions, and causation (cf. Maudlin 2007).

Some "cor-"/relations between these notions:

- 'X causes Y.' relates to 'If X did not occur, Y would not occur.'
- 'x has the disposition of water-solubility.' relates to 'If x were put into water, x would dissolve.'

Maudlin: These "cor-"relations are best accounted for by taking laws as a fundamental common origin:



Carnapian Conditions: Exactness

How does creative abduction square with the widened Carnapian conditions?

Let us begin with exactness:

The structure of common causes/features is well explored.

Particularly within the wider framework of Bayesian networks.

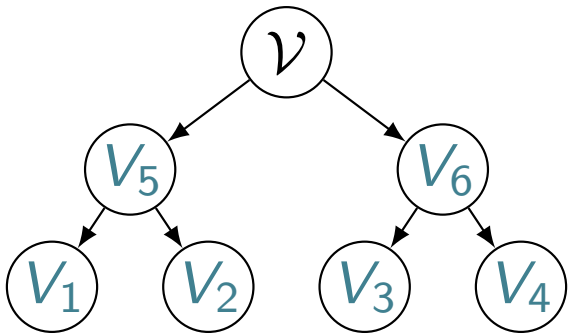
The framework is as exact as one can wish.

Carnapian Conditions: Simplicity

Common cause: simplest causal graph accounting for a correlation.

Sometimes not all correlations are clear from the beginning on.

This, then, leads to piecemeal linking. Example:



Intermediate causes decrease the degree of unification.

E.g., for pairwise linking:

$$u(n) = 0.25n - 0.25$$

This still brings linear growth with it, but it is less steep.

This is the default case due to practical reasons.

Carnapian Conditions: Fruitfulness

Fruitfulness seems easy to come by.

Given the task of explanation, common causes make up for paradigm cases of explanation.

Here is why: due to their screening off feature: $P(V_i|V_j) \overset{Why?}{>} P(V_i)$

Because of $\mathcal{V}$: correlation vanishes in light of $\mathcal{V}$: $P(V_i|V_j, \mathcal{V}) = P(V_i|\mathcal{V})$

The explanation is the better/the more fruitful, the more unified it is.

Carnapian Conditions: Similarity

Finally, a quick note on similarity.

Recall, similarity serves the purpose of guaranteeing subject-relatedness.

When introducing a concept, there is no respective C^A.

However, there is some sense in which such a constraint is also relevant for abductive coneptual engineering.

As is the case with IBE, also creative abduction balances simplicity with accuracy.

Due to the holistic nature of the approach, it can happen that by introducing a common cause/feature, some weak cor-/relations cannot be accounted for.

So, some of the concepts and their interrelations which motivated the introduction of a new concept are no longer perfectly matching.

For subject-preservation, a similarity requirement is needed.

Summary

Widening Carnap's conditions of adequacy straightforwardly covers (the most) important features of conceptual engineering: change and subject-relatedness (*orthodoxy*)

That the widened Carnapian conditions of adequacy have some bearing on or even a match in virtues of creative abduction, makes creative abduction a tool of conceptual engineering.

References I

Brun, Georg (2016-12). "Explication as a Method of Conceptual Re-engineering". In: *Erkenntnis* 81.6, pp. 1211–1241. DOI: 10.1007/s10670-015-9791-5.

— (2017). "Conceptual Re-Engineering: From Explication to Reflective Equilibrium". In: *Synthese*. DOI: 10.1007/s11229-017-1596-4.

Burgess, Alexis and Plunkett, David (2013a). "Conceptual Ethics I". In: *Philosophy Compass* 8.12, pp. 1091–1101. DOI: 10.1111/phc3.12086.

— (2013b). "Conceptual Ethics II". In: *Philosophy Compass* 8.12, pp. 1102–1110. DOI: 10.1111/phc3.12085.

Cappelen, Herman (2018). *Fixing Language. An Essay on Conceptual Engineering*. Oxford: Oxford University Press.

Carnap, Rudolf (1963). "Replies and Systematic Exposition". In: *The Philosophy of Rudolf Carnap*. Ed. by Schilpp, Paul Arthur. La Salle: Open Court, pp. 858–1013.

Dardashti, Radin, Hartmann, Stephan, Thébault, Karim, and Winsberg, Eric (2019). "Hawking Radiation and Analogue Experiments: A Bayesian analysis". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*. DOI: https://doi.org/10.1016/j.shpsb.2019.04.004.

Douven, Igor (2018). "Abduction". In: *The Stanford Encyclopedia of Philosophy (Summer 2018 Edition)*. Ed. by Zalta, Edward N.

References II

Feldbacher-Escamilla, Christian J. (2020a). "A Complementary Approach to Aristotle's Account of Definition and Carnap's Account of Explication". In: *Logical Analysis and History of Philosophy* 22.1, pp. 19–40. DOI: 10.30965/9783957437310_003.

— (under revision). "Conceptual Engineering and Abduction". In: *manuscript*.

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

Frege, Gottlob (1979). "Boole's logical Calculus and the Concept-script [1880/81]". In: *Posthumous Writings*. Ed. by Hermes, Hans, Kambartel, Friedrich, and Kaulbach, Friedrich. Translated by Peter Long and Roger White. Oxford: Basil Blackwell, pp. 9–46.

Goodman, Nelson (1951/1977). *The Structure of Appearance*. Second Edition. Harvard: Harvard University Press.

Haslanger, Sally (2000). "Gender and Race: (What) Are They? (What) Do We Want Them To Be?" In: *Noûs* 34.1, pp. 31–55. DOI: 10.1111/0029-4624.00201.

Lipton, Peter (2004). *Inference to the Best Explanation*. 2nd Edition. London: Routledge.

Machery, Edouard (2017). *Philosophy Within Its Proper Bounds*. Oxford: Oxford University Press.

Maudlin, Tim (2007). *The Metaphysics Within Physics*. Oxford: Oxford University Press.

Menger, Karl (1943). "What is Dimension?" In: *The American Mathematical Monthly* 50.1, pp. 2–7. DOI: 10.2307/2303984.

References III

Niiniluoto, Ilkka (1999). "Defending Abduction". In: *Philosophy of Science* 66, S436–S451. DOI: 10.1086/392744.

Schilpp, Paul Arthur, ed. (1963). *The Philosophy of Rudolf Carnap*. La Salle: Open Court.

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

Strawson, Peter F. (1963). "Carnap's Views on Constructed Systems Versus Natural Languages in Analytic Philosophy". In: *The Philosophy of Rudolf Carnap*. Ed. by Schilpp, Paul Arthur. La Salle: Open Court, pp. 503–518.