From the *Aufbau* to the Canberra Plan

David J. Chalmers

Carnap's *Aufbau*

- Rudolf Carnap (1928) *Der Logische Aufbau der Welt (*The Logical Structure of the World)

- Aims for a characterization of the world in terms of a minimal vocabulary, from which all truths about the world can be derived.

The Vocabulary

- Carnap has one non-logical primitive:
  - The relation of recollected phenomenal similarity (between elementary experiences).
- The world-description can be given using an expression for this relation, and first-order logical expressions.
- In principle the relation can be eliminated, giving a purely logical description of the world.

The Derivation Relation

- All truths are held to be derivable from the world-description plus definitional sentences for non-basic vocabulary.
  - Definitional sentences give explicit definitions
- Guiding idea: Non-basic truths are analytically entailed by basic truths
  - Aiming for an epistemological and semantic reduction
  - Although: extensional criterion of adequacy for definitions?

Problems for the *Aufbau*

(1) Goodman's critique (construction of the visual field)
(2) Quine's critique (definition of spatiotemporal location)
(3) Doubts about phenomenal reduction
(4) Doubts about analyticity
(5) Doubts about definitional analysis
(6) Newman's problem for structuralism

The Canberra Plan

- The "Canberra Plan": A program for semantic/epistemological/ metaphysical reduction
  - Grounded in the Ramsey-Carnap-Lewis method for the analysis of theoretical terms
  - But extended to concepts and expressions of all sorts
- Regiment, Ramsify, and rigidify where necessary!
- Q: Might the Canberra plan be used to vindicate Carnap?
  - A minimal world-description that analytically/a priori entails all truths?
  - N.B. Concentrate on prospects for epistemological/semantic entailment, not modal/metaphysical entailment.

Regimentation

- Applying the method to e.g. ‘charge’:
- First, regiment one’s theory of the role charge plays
  - Charge is a quantitative property that can take positive/negative values
  - Entities with opposite charge attract (in such-and-such way)
  - Entities with same-sign charge repel (in such-and-such way)
  - …
- The result can be put in the form P(charge), for some complex predicate P
  - The expressions used in P are the “O-terms”
  - This regimentation is supposed to capture our understanding of ‘charge’
  - Idea: it is a conceptual truth that a property ϕ is charge iff P(ϕ)

Ramsification and Rigidification

- Then we can analyze the sentence ‘x has charge’ as
  - $\exists\phi$ (P(ϕ) & ϕ(x)) [or $\exists\phi$ (P(ϕ) & instantiates (x, ϕ))]
  - A Ramsey sentence for ‘charge’
- Likewise for other sentences involving ‘charge’
  - Analyzed via Ramsey sentences including just logical expressions and O-terms
  - All ‘charge’ truths derivable from complete enough truth in the O-vocabulary.
- Rigidification (where necessary)
  - $\exists\phi$ ϕ (x) & actually P(ϕ)
- Charge is whatever (actually) plays the charge role.

Repeated Ramsification

- One can regiment/Ramsify multiple expressions one at a time, yielding Ramsey sentences with O-terms excluding those expressions
  - Then all truths in the full vocabulary will be derivable from truths in the O-vocabulary
- Canberra Plan: Apply this method not just to theoretical terms in science, but to expressions of all sorts
  - Free will is what plays the free will role
  - Water is what (actually) plays the water role
  - Gödel is whoever (actually) plays the Gödel role
  - And so on

Definitions and A Priori Entailment

- Complication: There are reasonable doubts about the availability of explicit finite definitions: e.g. *knowledge* = *such-and-such*

- But for the current project, one doesn't need finite definitions, just a priori entailments
  - 'Knowledge'-truths a priori entailed by truths in a more basic vocabulary
  - T-truths a priori entailed by non-T truths [C&J 2001]
  - E.g. a priori entailed by Ramsey sentence involving O-terms

- Repeated application of this method will yield some limited vocabulary V such that all truths are a priori entailed by V-truths
  - There will be a V-sentence D such that for all truths T, 'D ⊃ T' is a priori

Global Ramsification?

- Thought: repeated Ramsification might eventually yield a basic sentence describing the world
  - E.g. A true sentence of the form 'there exist entities and properties that stand in such-and-such relations'.
- This sentence might play the role of Carnap's basic world-description: all truths derivable from it, via logic plus (Ramseyan) definitions, or by a priori entailment.
- Q: What might such a sentence look like?
  - Extreme version: a purely logical sentence (all O-terms are Ramsified away).
  - Less extreme version: a sentence involving some primitive O-terms (that are not Ramsified away).

Newman's Problem

- Pure structuralism (Russell, Carnap): The content of science can be captured in a purely structural description.

- A purely structural description of the world is a description of the form

  there exist relations R1, R2, …, and there exist entities x, y, z, …, such that …. [xR1y, ~xR2z, and so on]

- Newman (1928): Purely structural descriptions are near-vacuous.
  - They are satisfied by any set of the right cardinality.
  - Given such a set, we can always define up relations R1, R2, …, that satisfy the descriptions relative to members of the set

  - (Compare: Putnam's model-theoretic argument.)

Russell's Response

- Russell's response:
  - Newman is right about pure structuralism
  - Science delivers more than a purely structural description of the world
  - Its description involves a basic relation: the relation of "spatiotemporal copunctuality" between sense-data and physical objects.
  - We assume this relation R, and give an impure structural description:

    there exist entities x, y, z, [relations R1, R2, …, properties P1, P2, P3…] such that xRy, yRz [P1x, xR1y,…]

- The primitive relation R is such that we grasp it by understanding it (via Russellian acquaintance?).

Carnap's Response

- Carnap is initially a weak structuralist
  - His description D of the world invokes the primitive relation R, plus logical vocabulary.
- But he wants to be a pure structuralist, so he ultimately tries to drop R (sections 153-55).
  - i.e. "there exists a relation R such that D[R]"
- He then notices the threat of vacuity (Newman's problem!)
  - To avoid it, he stipulates that R is a "founded" ("natural", "experiencable") relation (cf. Lewis on Putnam)
  - Justifies this by claiming that "founded" is a basic logical concept!

Ramseyan Structuralism

- Extreme Global Ramsification is a form of pure structuralism, and is subject to Newman's problem.
  - Both Carnap's and Russell's response are available.
- Lewis gives a version of Carnap's response, appealing to 'natural' properties (though in the metasemantics, not in the Ramsey sentence)
- Alternatively, one can give a version of Russell's response, allowing other primitive O-terms that are not Ramsified away

The Appeal to Naturalness

- Newman: ‘If the world has cardinality C, then R’ is a priori, for Ramsey sentence R and appropriate cardinality C.
- Q: Does the appeal to naturalness affect the a priori truths?
- If no: it doesn’t help with Newman’s problem
- If yes: then naturalness is being smuggled into the ideology of the Ramsey sentence, as with Carnap
  - So the sentence in effect invokes a primitive concept of ‘natural property’
  - But then: why not other primitive concepts?

Other Primitive Concepts

- Everyone allows some primitive (unramsified) expressions
  - Logical expressions
  - Mathematical expressions (usually)
  - Naturalness (Carnap)
  - Experiential expressions (Putnam)
- So not every term needs to be Ramsified via a theoretical role
- The Ramsey sentence might contain some further primitives, e.g. expressing
  - Spatiotemporal concepts
  - Nomic/modal concepts
  - Mental concepts
- Then Newman's problem is avoided
- Q: What are the primitive O-terms?

Transparent Concepts

- *Transparent* concept: possessing the concept puts one in a position to know what its referent is
- In 2D terms, transparent concepts are epistemically rigid (constant primary intension)
  - Heuristic: Transparent expressions are not “Twin-Earthable”,
  - E.g. *friend* is arguably transparent, *water* is opaque
- Opaque concepts are Ramsified away
- Transparent concepts can be Ramsified, but need not be
  - So primitive O-terms may express transparent concepts

A Starting Point

- Chalmers and Jackson 2001: All truths are a priori entailed by PQTI
  - Conjunction of microphysical/phenomenal/indexical/that's-all truths
  - I.e. for all truths M, 'PQTI ⊃ M' is (ideally) knowable a priori
- PQTI is not plausibly a primitive basis
  - Microphysical terms (and phenomenal terms?) can be Ramsified
  - Microphysical concepts are arguably opaque
- But we can use PQTI as a starting point to narrow down the ultimate O-terms.

Spatiotemporal Structuralism

- Q: What might serve as ultimate O-terms for Lewis?
  - Physical terms are definable in terms of impact on observables
  - Observables are definable in terms of effect on experiences
  - Experiences are definable in terms of effect on behavior/processing
  - Cause/effect definable in terms of counterfactuals
  - Counterfactuals definable in terms of laws
  - Lawhood is definable in terms of spatiotemporal regularities

- Perhaps: Some spatiotemporal terms are O-terms, not theoretically definable
  - Cf. Lewis's Humean supervenience base, a distribution of properties across spacetime.
  - Truths about this base analytically entail all truths, but are themselves unanalyzable?

- Spatiotemporal structuralism: A fundamental world-description characterizing the distribution of certain (existentially specified) properties and relations in spacetime
  - Primitives: Spatiotemporal, logical/mathematical, categorical, indexical/totality?

Spatiotemporal Opacity

- Problem: Spatiotemporal concepts are arguably Twin-Earthable, and so opaque
  - They pick out relativistic properties in relativistic scenarios
  - Classical properties in classical scenarios
  - Computational properties in Matrix scenarios
- In effect: spatiotemporal concepts are concepts of that manifold of properties and relations that serves as the normal causal basis for our spatiotemporal experience.
- If so: spatiotemporal terms are not ultimate O-terms.
- So what are the ultimate O-terms?

Nomic/Phenomenal Structuralism

- Alternative package:
  - Physical terms analyzed in terms of effects on observables
  - Observables (inc spatiotemporal) defined in terms of effects on experience
  - Causation analyzed in terms of laws
- Ultimate O-terms include phenomenal terms and nomic terms
  - These show up ubiquitously in Ramseyan analyses of other terms.
  - Somewhat plausibly, phenomenal concepts are unanalyzable and transparent
  - Same for some nomic concepts (*law*, or *counterfactually depends,* or *cause*)
- Nomic/phenomenal structuralism: Ramsey sentence specifies a manifold of (existentially specified) properties and relations whose instances are nomically connected to each other and to experiences
  - Primitives: Nomic, phenomenal, logical/mathematical, categorical, indexical, totality?

Alternative Packages

- There are various available packages, depending on one's views about
  - Analyzing the nomic in terms of the non-nomic
  - Analyzing the experiential in terms of the non-experiential
  - Analyzing the spatiotemporal in terms of the non-spatiotemporal
- E.g. N, S, NE, SE, NSE
- But one had better not embrace all three analyses at once, at cost of Newman's problem
  - Also: one had better not ramsify away both nomic and spatiotemporal, at cost of a sort of phenomenalism.
  - One might also further analyze the experiential, e.g. in terms of relations to "Edenic" properties presented in perception.
- One could be pluralistic (cf. Carnap), allowing multiple minimal vocabularies

Ramseyan Humility?

- Ramsey sentence specifies basic physical properties existentially, via roles
  - Are there further truths about which properties these are?
- Answer 1: the properties are just numerically distinct (Lewis/Armstrong)
  - Then the Ramsey sentence (with that's-all) is epistemically complete
- Answer 2: the properties have a further ungraspable nature
  - Then the Ramsey sentence entails all graspable/expressible truths
- Answer 3: the properties have a further graspable nature
  - Graspable under transparent concepts -- e.g. phenomenal, Edenic, alien.
  - Then the Ramsey sentence must be supplanted: existential quantifiers for properties replaced by these transparent specifications
  - We will need primitive terms for these concepts, or a further analysis.

Scrutability and Meaning

- Scrutability: there is a limited vocabulary V such that all truths are a priori entailed by some V-truth
- Generalized scrutability: there is a limited vocabulary V such that all e-possible sentences are a priori entailed by some e-possible V-sentence.
  - S is e-possible when ~S [or ~det S] is not a priori
- Generalized scrutability allows a world-description for every e-possible scenario
  - With a vocabulary capturing the basic dimensions of epistemic space?
  - We can construct scenarios as maximal e-possible V-sentences
  - S is true at a scenario W iff 'D ⊃ S' is a priori, where D specifies W.
- One can then say that the intension of S is the set of scenarios at which S is true
  - Then 'S ≡ T' is a priori iff S and T have the same intension
  - A quasi-Fregean semantic value, vindicating Carnap's project in *Meaning and Necessity*?

Conclusion

- The Canberra plan, resting on the Ramsey-Carnap-Lewis method, offers some hope of vindicating Carnap's project in the *Aufbau.*
- Carnap's minimal vocabulary needs to be expanded, to include nomic (or perhaps spatiotemporal) vocabulary as well as phenomenal vocabulary.
- Carnap's derivation relation should be weakened from entailment via definition to a priori entailment.
- With these alterations, the project of the *Aufbau* is very much alive.