



## Essay Review

### Has Foundationalism Failed?

By Majid Amini

A critical review of *Coherence in Thought and Action*. By Paul Thagard.  
Cambridge (Mass.): MIT Press, 2000/2002.

Foundationalism and coherentism in epistemology, like so many other polar philosophical positions, have been going through periodic bouts of trials and tribulations to the delight of each other's opponent. In the recent cycle of this ongoing saga, coherentism has suffered a severe blow from the defection of Laurence Bonjour, as one of its most prominent and persistent defenders, to the opposition camp.<sup>1</sup> Indeed, in his new foundationalist incarnation, Bonjour concludes one of his papers by the stark statement that 'coherentism is pretty obviously untenable, indeed hopeless'.<sup>2</sup> However, Paul Thagard in his latest book, *Coherence in Thought and Action*, attempts not only to defend coherentism from its detractors but also to demonstrate that foundationalism does not have a chance of competing with coherentism in an account of human consciousness with its multifarious manifestations.<sup>3</sup> What essentially Thagard does is to tender *coherence* as the holy grail of human inference and thinking. This is epitomised by his repeated observation that 'the foundational search for certainty was pointless, and that what mattered was the growth of

knowledge, not its foundations.' (p. 90) Thagard contends that the key to the growth of knowledge and to its very understanding is nothing other than coherence, where in a similar epistemic exercise foundationalism 'has undoubtedly failed'. (p. 8) The purpose of this critical review is, therefore, to look at the technical details and finesses of Thagard's conception of coherence to ascertain whether its alleged extensive explanatory power renders foundationalism obsolete.<sup>4</sup>

To set the scene and also to have a better appreciation of the significance of Thagard's contention to ensconce coherence at the centre of philosophy, a brief background to the debate may not be amiss. Generally, coherence has been a recurrent theme in philosophy; indeed, it does not take long to find a philosopher commenting on certain competing accounts to commend the more coherent scientific or ethical theory, or the more coherent plan, or the more coherent theory of something else. Yet, there has not been much of an account of what exactly coherence itself is. The problem was not that of coming by synonymous, or semi-

synonymous, words or phrases like one's beliefs cohere if they "hang together" or one's goals make up a coherent plan if they "fit well with one another." The problem was not even that of itemising the ingredients of coherence such as appropriateness of means to ends, logical consistency, and so on. The problem was the lack of any specification of how these ingredients were to be calculated and combined with respect to a set of propositions. That is, how one was supposed to work out the degree each item on the list would contribute to the overall coherence of a theory or strategy. In Thagard's own words: 'no insight on how to achieve it.' (p. 6) *Practically speaking*, there was no decision procedure for determining *comparative coherence*.

However, in the absence of such an account, counselling to choose the most coherent cluster of thoughts or course of action would be no more than empty words. In fact, the received opinion was that an account of coherence sufficiently detailed and definite to be turned into a computer programme was beyond reach. Coherence is not something for which we have an algorithm, but is something that we ultimately judge, in Hilary Putnam's "memorable" expression, 'by "seat of the pants" feel'.<sup>5</sup> It is, therefore, against this background that Thagard's attempt to characterise coherence 'as mathematically precise as the tools of deductive logic and probability theory' (p. 16) becomes interesting and challenging.

Formally, Thagard defines coherence in terms of the notion of a *coherence problem* as follows:

Let  $E$  be a finite set of elements  $\{e_i\}$  and  $C$  be a set of constraints on  $E$  understood as a set  $\{(e_i, e_j)\}$  of pairs of elements of  $E$ .  $C$  divides into  $C^+$ , the positive constraints on  $E$ , and  $C^-$ , the negative constraints on  $E$ . With each constraint is associated a number  $w$ , which is the weight (strength) of the constraint. The problem is to partition  $E$  into two sets,  $A$  and  $R$ , in

a way that maximizes compliance with the following two *coherence conditions*:

If  $(e_i, e_j)$  is in  $C^+$ , then  $e_i$  is in  $A$  if and only if  $e_j$  is in  $A$ .

If  $(e_i, e_j)$  is in  $C^-$ , then  $e_i$  is in  $A$  if and only if  $e_j$  is in  $R$ .

Let  $W$  be the weight of the partition, that is, the sum of the weights of the satisfied constraints. The coherence problem is then to partition  $E$  into  $A$  and  $R$  in a way that maximizes  $W$ . (p. 18)

In a non-technical language, this simply means that: *given* (1) a number of elements like propositions or objectives, (2) two sets of positive and negative constraints such that if a certain element is in, then another one should also be in, and if a certain element is in, another one should be out, and (3) weights for the constraints that specify how important satisfying a constraint is, *then* one needs to find a way of dividing up the set of elements into an accepted ( $A$ ) and a rejected ( $R$ ) set which satisfy as many constraints as possible. Thus, the higher the summed weights of the satisfied constraints, the more coherent the solution to the coherence problem would be in a particular case.

But, how can one compute this conception of coherence? That is, back to our earlier pragmatic problem of how to measure comparative coherence. Thagard moots the following five possible algorithms for calculating coherence:

An *exhaustive* search algorithm that considers all possible solutions

An *incremental* algorithm that considers elements in arbitrary order

A *connectionist* algorithm that uses an artificial neural network to assess coherence

A *greedy* algorithm that uses locally optimal choices to approximate a globally optimal solution

A *semidefinite programming* (SDP) algorithm that is guaranteed to satisfy a high proportion of the maximum satisfiable constraints (p. 26; original emphasis)

In the process of essaying each one in turn, he dismisses the first two for being of limited use but argues that the other three provide effective means of computing coherence. Thagard's favourites, however, are connectionist algorithms as, he claims, there is a 'natural alignment between coherence problems and connectionist networks' and they provide the 'most psychologically appealing models of coherence optimization'. (pp. 33, 40)

Now, in light of the foregoing outline of the formal and implementational components of Thagard's account of coherence, I would like to raise three sets of points. The first point to note about the formal characterisation of coherence is the use of the *biconditional* clause "if and only if" which is indicative of a larger issue about the insufficiency of coherence for constituting *truth*. The question is: how are we supposed to understand such clauses – in a coherence way on pain circularity or in a non-coherence manner? My problem is patently parochial but symptomatic of the global issue about the nature of truth. The significance of the question lies in the twist of Thagard's tale where he parts company with conventional coherentists by *not* defending 'a coherence theory of truth, since there are good reasons for preferring a correspondence theory'.<sup>6</sup> (p. 74) In fact, rather iconoclastically for a coherentist, he attempts to 'argue *against* a coherence theory of truth'. (p. 85; original emphasis) Nonetheless, in a spirit of conciliation, he says: 'truth is a matter also of correspondence, not coherence alone.' (p. 78) Thus, Thagard's eclectic approach allows him to parry perennial problems of coherentism such as *isolation objection* that

a set of beliefs may be internally coherent but not true – the case of illusory but consistent theories. But, obviously, his eclecticism not only fails to curry favour with hardline and puritanical coherentists but also highlights, rather self-defeatingly for Thagard, the persistence of non-coherentist constituents of cognitive architecture.

The second point to note is that Thagard's characterisation of coherence is reminiscent of a familiar problem in graph theory known as MAX CUT. Formally, Michael Garey and David Johnson express the problem thus:

INSTANCE: Graph  $G = (V, E)$ ,  
 "weigh"  $w(e) \in Z^+$  for each  $e \in E$ , positive integer  $K$ .  
 QUESTION: Can  $V$  be partitioned into two disjoint sets  $V_1$  and  $V_2$ , such that the sum of the weights of the edges from  $E$  that have one endpoint in each set is at least  $K$ ?<sup>7</sup>

Again, in plain language, the question is whether one can find a way of cutting a network into two parts such that the total capacity of the links crossing the cut is maximised. Now, the intriguing point here is that Thagard's characterisation of coherence is indeed a variation on MAX CUT, but MAX CUT is NP (Nondeterministic Polynomial)-complete. A problem is NP-complete when it is hard in principle: that is, no matter how large or fast a computer is, there are reasonably sized inputs for which there are *no* efficient (polynomial-time) procedures for solving the problem. In other words, like MAX CUT, Thagard's coherence is NP-complete and as such is *computationally intractable*.

However, Thagard himself is cognisant of these concerns and seems happy to settle for an *approximation* of optimal coherence: computing 'coherence is a matter of maximizing constraint satisfaction' which 'can be accomplished approximately'. (p. 40) That is, if the algorithms cannot be used to lasso the set of

elements with the maximum summed weights of coherence, one should perhaps opt for a set that comes close. But the problem with such approximations is that not only they fail to form the most coherent set but also fail to ensure that the chosen set is not dramatically different from the most coherent one. In other words, there is no guarantee that the *next most* coherent set is not drastically divergent from the *most* coherent one.

Nevertheless, for Thagard, there are still ways of shoring up coherence with varying degrees of vigour. Minimally, by taking the cue from the title of the book, one could concentrate on the *action* part, rather than *thought*, and emphasise the centrality of coherence in *conative* contexts. In planning tasks where the problem is not so much about truth or falsity but devising the most efficient way of reconciling various practical goals and objectives, approximations of most coherent plans are as good as the most coherent ones. Thus, from a practical perspective, coherence as a criterion of adequacy does play a principal part in our reasoning deliberations.

Maximally, however, one may extend the debate to the level of *thought*. Thagard could pose the same problem of approximation to non-coherentist alternatives. For example, Bayesian probabilistic reasoning is similarly beset with computational intractability and as such it relies on approximations for computing posterior probabilities. Also probabilistic information updating leads to a combinatorial explosion, ‘since we need to know the probabilities of a set of conjunctions whose size grow exponentially with the number of propositions.’ (p. 250) Generally, and more importantly, it seems that any procedure sufficiently rich to be able to model everyday theory choices, whether scientific or otherwise, involves some measure of approximation.

The third point to note is the predilection that Thagard shows for connectionist algorithms in the implementation of coherence. Although Thagard is conscious of the computa-

tional limitations of connectionist algorithms, he capitalises on the encouraging empirical results from a number of such neural network models of coherence to propose them for their psychological appeal. As a matter of fact, Thagard says that his ‘characterization of coherence was abstracted’ from connectionist methods in the first place. (p. 15) Unfortunately, however, he does not engage with the criticisms of connectionism pressed by the classical computational theorists of mind and advocates of domain-specificity and modularity of brain, especially in point of the *psychological plausibility* of connectionist models of mind, which plainly leaves lacunae in his coherentist lattice of cognition.

In conclusion, through his eclecticism and approximation algorithms, Thagard seems able to tout a viable notion of coherence. However, this is only achieved by extensively curtailing the traditional claims of coherentism and by conceding, for example, that ‘the formation of elements such as propositions and concepts and the construction of constraint relations between elements depend on processes to which coherence is only indirectly relevant.’ (p. 24) Also, even though he may appear to have vaccinated coherentism against the virus of *isolation*, *i.e.*, there is no guarantee that the most coherent theory is also true, the susceptibility still remains the Achilles’ heel of coherentism. That is, approximating maximum coherence is not yet the same thing as approximating truth. Indeed, Thagard himself admits that for the coherentist project to succeed one needs ‘to see a much fuller account of the conditions under which progressively coherent theories can be said to approximate the truth.’ (p. 280) Therefore, in view of these pressing problems for coherentism, it seems rather too hasty to hypothesise, let alone to state categorically, that foundationalism “has undoubtedly failed”.

Majid Amini, Ph.D., Associate Professor in Philosophy, Department of History and Phi-

osophy, Virginia State University, P.O. Box 9070, VA 23806, USA. Email: [mamini@vsu.edu](mailto:mamini@vsu.edu).

## Notes

1. For the coherentist phase of Bonjour's work, see, for example, his *The Structure of Empirical Knowledge* (Harvard University Press, 1985) and his contribution to *The Current State of the Coherence Theory*, edited by John W. Bender (Kluwer, 1989). For his foundationalist conversion, see, for example, Bonjour's *In Defense of Pure Reason* (Cambridge University Press, 1998), 'The Dialectic of Foundationalism and Coherentism' in *The Blackwell Handbook of Epistemology*, edited by John Greco and Ernest Sosa (Blackwell, 1999), 'Foundationalism and the External World' in *Philosophical Perspectives: Volume 13*, edited by James Tomberlin (Blackwell, 1999), and 'Toward a Defense of Empirical Foundationalism' in *Resurrecting Old-Fashioned Foundationalism*, edited by Michael DePaul (Rowman and Littlefield, 2000).

2. 'The Dialectic of Foundationalism and Coherentism' in *The Blackwell Handbook of Epistemology*, edited by John Greco and Ernest Sosa (Blackwell, 1999) p.139.

3. Thagard has been publishing on the theme of coherence for the past fifteen years or so; see, for example his 'Explanatory Coherence', *Behavioral and Brain Sciences*, 12: 435-467,

1989. But, his latest book has the virtue of conveniently collecting all of his major publications on this issue with further elaborations and emendations in one single volume.

4. To evince the enormous explanatory efficacy of coherentism, six out of the nine chapters of the book are devoted to an exploration of the relevance and application of coherence to a wide variety of topics ranging from the nature of knowledge and reality to the philosophical and psychological problems in ethics and politics, the nature of emotions and how emotional coherence underpin beauty in science and art, and how coherentism fares significantly better than the rival probabilistic approach to the issues of theory choice. But, my concern in this paper will be with the details of Thagard's characterisation of coherence than its applicability or otherwise.

5. *Reason, Truth and History* (Cambridge University Press, 1982) p. 133.

6. Thagard's trouble with the coherentist theory of truth stems from his predilection for scientific realism. In general, he thinks that considerations of explanatory coherence strongly support the existence of an independent world, so that truth must be a matter of correspondence with this world.

7. *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, 1979) p. 87.