A Hoard of Hidden Assumptions

David Miller

1 Introduction

Such is our fondness for deductive validity that we are at times tempted, when confronted with an invalid argument, to postulate some missing premise, or hidden assumption, that, if adopted, would render the argument valid after all. A familiar philosophical example is Hume's contention that 'all our experimental conclusions proceed upon the supposition that the future will be conformable to the past' [3, § IV, Part II, p. 35]. Much can be said, and has been said, about the usefulness of a metaphysical principle of uniformity that is anything like as general as the supposition that Hume flirted with, but what interests me in this paper is the implicit logical thesis that, in most instances of deductive failure, it is possible to identify, more or less uniquely, a statement to plug the gap in the defective argument. In this vein, some who call themselves deductivists, such as Musgrave [8], recommend that all allegedly inductive inferences (but not necessarily all invalid inferences) are best treated as enthymemes calling for systematic deductive rehabilitation.

The usual rules governing the material conditional imply that if the argument from A to C is invalid then $\mathsf{A} \to \mathsf{C}$ is the logically weakest additional premise that is strong enough to make it valid. That consideration would not have been congenial to Hume, even if it had occurred to him, since he evidently thought that there is some general statement underwriting all 'reasonings from experience' [2, p. 651]. Yet it can be proved ([6, Chapter 8, §§ 1f.]) that there does exist a logically weakest strictly universal statement that restores validity to the kinds of argument that Hume denounced, and it is perhaps possible that something like that statement was skulking, unformed, at the back of his mind.

2 Critical thinking

In the *critical thinking movement*, an approach to the teaching of argumentation (and sometimes also of logic) that has become startlingly popular in the last thirty years or so, the question of how invalid arguments are to be reinvented as valid is handled somewhat differently. It is accepted that, in most cases in which the argument from A to C is invalid, there are ever so many further premises B that, although sufficient to make the argument from A to C into a valid argument, cannot be identified with the premise that was missing. Indeed, it is generally acknowledged that B cannot qualify as the missing premise if it alone validly implies the conclusion C (in particular, B cannot be identified with C). These uncontroversial observations generate conditions **a** and **b** in § 4 below. What is distinctive in the treatment in some critical thinking texts of the essential 'thinking skill' called *assumption spotting* is the method deployed to identify a statement B not only as sufficient for the argument from A and B to C be valid, but also as necessary.

According to the authors of [1, § IV, p. 563], who are highly critical, the so-called *negative test*, or *reverse test*, requires additionally that 'the negation of the original conclusion is derivable from the addition of the negation of [the missing] assumption to the original [premises]'. In short, the statement B is the assumption missing from the invalid argument from A to C only if the argument from A and ¬B to the conclusion ¬C is valid. This generates condition **c** in § 4 below. It is noted further that 'those who advance the negative test appear to abide by ... [a] fourth requirement, even though it is not stated' ([1, § V, p. 567]), to the effect that 'the negation of the conclusion is not derivable from the negation of the [missing] assumption alone', in short that the argument from ¬B to the conclusion ¬C be not valid. This generates condition **d** in § 4 below.

These authors conclude their rather relaxed discussion of the two forms of the negative test by wondering whether 'there are invalid arguments for which no assumption satisfies all four requirements and other invalid arguments where multiple alternatives satisfy those requirements', and 'what the underlying rationale of the negative test actually is' (*ibidem*). In the spirit of critical rationalism ([4, 6]), which advocates criticism rather than attempted justification as the only proper way to evaluate contested hypotheses, I discharge myself from pursuing the second inquiry, but I do wish to pursue the first. It will be shown in the rest of the paper how far the negative test (in both variants) is from enabling us to identify the assumptions that are missing from invalid arguments.

For exposure, exposition, and criticism, of some of the epistemological shortcomings of the critical thinking movement, see [5] and [7, §5].

3 A few technicalities

To begin, let us be explicit about terminology and notation. Given the class of meaningful sentences of some unspecified language that incorporates at least elementary sentential logic, sentences that are logically equivalent will be identified and called *statements*. We shall (nonstandardly) use $\Vdash$ and $\vdash$, in analogy with $\leq$ and $<$; that is, $\mathsf{A} \Vdash \mathsf{B}$ (rather than the usual $\mathsf{A} \vdash \mathsf{B}$) signifies that B *is derivable from* A, while $\mathsf{A} \vdash \mathsf{B}$ means that B *is properly (or unilaterally) derivable from* A; that is, $\mathsf{A} \Vdash \mathsf{B}$ but $\mathsf{A} \neq \mathsf{B}$. As is customary, $\top$ is the *tautological* or *logically true* statement, and $\bot$ is the *inconsistent* or *logically false* statement.

The following not unexpected results concerning interpolation will be useful. Suppose that $\mathsf{Z} \vdash \mathsf{X}$, so that $\mathsf{X} \wedge \neg\mathsf{Z}$ is consistent. Provided that the theory $\mathsf{X} \wedge \neg\mathsf{Z}$ is not maximal (that is, negation complete), there exists a *proper interpolant* between Z and X, that is to say, a statement Y such that $\mathsf{Z} \vdash \mathsf{Y} \vdash \mathsf{X}$. Such an interpolant Y may be constructed as follows. Since $\mathsf{X} \wedge \neg\mathsf{Z}$ is consistent and not maximal, there exists at least one statement U that is undecided by $\mathsf{X} \wedge \neg\mathsf{Z}$; that is, neither $\mathsf{X} \wedge \neg\mathsf{Z} \Vdash \mathsf{U}$ nor $\mathsf{X} \wedge \neg\mathsf{Z} \Vdash \neg\mathsf{U}$. For any such undecided U, let $\mathsf{Y} = \mathsf{Z} \vee (\mathsf{X} \wedge \mathsf{U})$. Since $\mathsf{Z} \vdash \mathsf{X}$, we have $\mathsf{Z} \Vdash \mathsf{Y} \Vdash \mathsf{X}$. If $\mathsf{Y} = \mathsf{Z}$, then $\mathsf{X} \wedge \mathsf{U} \Vdash \mathsf{Z}$, and hence $\mathsf{X} \wedge \neg\mathsf{Z} \Vdash \neg\mathsf{U}$; and if $\mathsf{X} = \mathsf{Y}$ then $\mathsf{X} \wedge \neg\mathsf{Z} \Vdash \mathsf{U}$. Since U (and $\neg\mathsf{U}$) are undecided by $\mathsf{X} \wedge \neg\mathsf{Z}$, we have shown that $\mathsf{X} \neq \mathsf{Y} \neq \mathsf{Z}$. In other words, $\mathsf{Z} \vdash \mathsf{Y} \vdash \mathsf{X}$; that is, Y is a proper interpolant between Z and X. So too, of course, is $\mathsf{Z} \vee (\mathsf{X} \wedge \neg\mathsf{U})$. It is easy to see that, by the distributive law, these two interpolants can be expressed as $\mathsf{X} \wedge (\mathsf{Z} \vee \mathsf{U})$ and $\mathsf{X} \wedge (\mathsf{Z} \vee \neg\mathsf{U})$.

The converse holds too: if $\mathsf{Z} \vdash \mathsf{Y} \vdash \mathsf{X}$ for some Y, then $\mathsf{X} \wedge \neg\mathsf{Z}$ is consistent and not maximal, and indeed, Y is undecided by $\mathsf{X} \wedge \neg\mathsf{Z}$. For if $\mathsf{X} \wedge \neg\mathsf{Z} \Vdash \mathsf{Y}$ then, since $\mathsf{Z} \vdash \mathsf{Y}$, we may conclude that $\mathsf{X} \Vdash \mathsf{Y}$, contrary to assumption. Likewise, if $\mathsf{X} \wedge \neg\mathsf{Z} \Vdash \neg\mathsf{Y}$, then $\mathsf{X} \wedge \mathsf{Y} \Vdash \mathsf{Z}$; and since $\mathsf{Y} \vdash \mathsf{X}$, we may conclude that $\mathsf{Y} \Vdash \mathsf{Z}$, again contrary to assumption. Since $\mathsf{Y} = \mathsf{Z} \vee (\mathsf{X} \wedge \mathsf{Y})$, it follows that every proper interpolant between Z and X has the form $\mathsf{Z} \vee (\mathsf{X} \wedge \mathsf{U})$ where U is a statement undecided by $\mathsf{X} \wedge \neg\mathsf{Z}$.

4 Statement of the problem

Given statements A, C such that the conclusion C cannot be validly derived from the premise or assumption (sometimes called the reason) A,

$$♠ \qquad \mathsf{A} \nVdash \mathsf{C},$$

the problem is to characterize all those *missing premises* or *hidden assumptions* B that provide solutions to the following set of conditions:

$$\begin{array}{llll} \textbf{a} & \mathsf{A}, \mathsf{B} \Vdash \mathsf{C} & \textbf{c} & \mathsf{A}, \neg\mathsf{B} \Vdash \neg\mathsf{C} \\ \textbf{b} & \mathsf{B} \nVdash \mathsf{C} & \textbf{d} & \neg\mathsf{B} \nVdash \neg\mathsf{C}. \end{array}$$

By contraposition, these conditions may be more succinctly rewritten:

$$\begin{array}{llll} \mathbf{a} & \mathsf{A}, \mathsf{B} \Vdash \mathsf{C} & \mathbf{c} & \mathsf{A}, \mathsf{C} \Vdash \mathsf{B} \\ \mathbf{b} & \mathsf{B} \nVdash \mathsf{C} & \mathbf{d} & \mathsf{C} \nVdash \mathsf{B}. \end{array}$$

Interest will be restricted to the conditions **a–c** and the conditions **a–d**.

5 Deducibility relations between A and C

It is easy to see that only some relations of deducibility can hold between A and C when **a–d**, or just **a–c**, hold. $\mathsf{A} \Vdash \mathsf{C}$ is at once ruled out by ♠. If $\neg\mathsf{A} \Vdash \mathsf{C}$, then $\neg\mathsf{A} \Vdash \mathsf{B} \to \mathsf{C}$, and by **a**, $\mathsf{A} \Vdash \mathsf{B} \to \mathsf{C}$, whence $\Vdash \mathsf{B} \to \mathsf{C}$, which contradicts **b**. If $\mathsf{C} \Vdash \mathsf{A}$ then by **c**, $\mathsf{C} \Vdash \mathsf{B}$, which contradicts **d**. If A and C are mutual contraries (that is to say, $\mathsf{A} \Vdash \neg\mathsf{C}$ or, equivalently, $\mathsf{C} \Vdash \neg\mathsf{A}$, or $\Vdash \neg\mathsf{A} \vee \neg\mathsf{C}$), then by **a**, A and B are also mutual contraries.

In brief, if all of the conditions **a–d** are to hold, then $\mathsf{C} \Vdash \neg\mathsf{A}$ is the only possible deducibility relation between A and C, while if only **a–c** are to hold, then $\mathsf{C} \Vdash \mathsf{A}$ and $\mathsf{C} \Vdash \neg\mathsf{A}$ are both possible. They are indeed possible simultaneously, since C may be the inconsistent statement $\bot$.

6 Truth-functional solutions

The conditions **a** and **c** together imply that $\mathsf{A} \wedge \mathsf{C} \Vdash \mathsf{B} \Vdash \mathsf{A} \to \mathsf{C}$. The only truth functions B of A and C that are in accordance with this restriction are (tritely) the conjunction $\mathsf{A} \wedge \mathsf{C}$ and the material conditional $\mathsf{A} \to \mathsf{C}$, and (hardly less tritely) the biconditional $\mathsf{A} \leftrightarrow \mathsf{C}$ and the statement C. But neither $\mathsf{A} \wedge \mathsf{C}$ nor C satisfies **b**, and $\mathsf{A} \to \mathsf{C}$ does not satisfy **d**.

In brief, $\mathsf{A} \to \mathsf{C}$ and $\mathsf{A} \leftrightarrow \mathsf{C}$ are the only solutions of **a–c** that are truth functions of A and C. These solutions coincide if & only if $\mathsf{C} \Vdash \mathsf{A}$. The biconditional $\mathsf{A} \leftrightarrow \mathsf{C}$ is the only truth-functional solution of **a–d**.

7 Necessary & sufficient conditions for solutions

As noted in § **5**, there are no solutions of **a–c** if $\neg\mathsf{A} \Vdash \mathsf{C}$, and there are no solutions to **a–d** if either $\neg\mathsf{A} \Vdash \mathsf{C}$ or $\mathsf{C} \Vdash \mathsf{A}$. Moreover, if $\neg\mathsf{A} \nVdash \mathsf{C}$, then neither $\mathsf{A} \to \mathsf{C} \Vdash \mathsf{C}$ nor $\mathsf{A} \leftrightarrow \mathsf{C} \Vdash \mathsf{C}$, and so both $\mathsf{A} \to \mathsf{C}$ and $\mathsf{A} \leftrightarrow \mathsf{C}$ satisfy **b**. But both $\mathsf{A} \to \mathsf{C}$ and $\mathsf{A} \leftrightarrow \mathsf{C}$ satisfy **a** and **c** for all A, C.

In brief, $\neg\mathsf{A} \nVdash \mathsf{C}$ is a necessary & sufficient condition (i) for **a–c** to have any solutions; (ii) for $\mathsf{A} \to \mathsf{C}$ to be a solution of **a–c**; and (iii) for $\mathsf{A} \leftrightarrow \mathsf{C}$ to be a solution of **a–c**. These solutions are different if & only if $\mathsf{C} \nVdash \mathsf{A}$. It is a necessary & sufficient condition for (i) **a–d** to have any solutions, and for (ii) $\mathsf{A} \leftrightarrow \mathsf{C}$ to be a solution of **a–d**, that $\neg\mathsf{A} \nVdash \mathsf{C} \nVdash \mathsf{A}$.

8 Non-truth-functional solutions

It is not implied in § **6** that $\mathsf{A} \to \mathsf{C}$ and $\mathsf{A} \leftrightarrow \mathsf{C}$ are the only possible solutions of **a–c**, or that $\mathsf{A} \leftrightarrow \mathsf{C}$ is the only possible solution of **a–d**. What is implied is that no other solution is a truth function of A and C.

It is also implied in § **6** that any further solution B of **a–d** has to satisfy $A \wedge C \vdash B \vdash A \to C$. There are therefore three disjoint ranges in which such an assumption B could be located: $(\boldsymbol{\alpha})$ $A \wedge C \vdash B \vdash A \leftrightarrow C$; $(\boldsymbol{\beta})$ $A \leftrightarrow C \vdash B \vdash A \to C$; and $(\boldsymbol{\gamma})$ $A \wedge C \vdash B \vdash A \to C$, where B and $A \leftrightarrow C$ are logically incomparable (that is to say, $A \leftrightarrow C \nVdash B \nVdash A \to C$). We shall establish that, provided that A and C are neither too strong nor too weak, the ranges $(\boldsymbol{\alpha})$ and $(\boldsymbol{\beta})$ yield ample opportunities for new solutions to **a–c** and **a–d**. No similar result is known for the range $(\boldsymbol{\gamma})$.

It was shown in § **3** that, for there to exist an interpolant B between $A \wedge C$ and $A \to C$, it is necessary & sufficient that $(A \to C) \wedge \neg(A \wedge C)$, which is identical with $\neg A$, be consistent and not maximal. Now $A = \bot$ is ruled out by ♠, and hence $\neg A$ is consistent. That is, there exist solutions of **a–d** that are distinct from $A \leftrightarrow C$ if & only if $\neg A$ is not maximal (that is, if & only if A is not irreducible in the sense of [9, § 4]).

9 $(\boldsymbol{\alpha})$ Solutions properly between $A \wedge C$ and $A \leftrightarrow C$

It follows from $(\boldsymbol{\alpha})$ $A \wedge C \vdash B \vdash A \leftrightarrow C$ that $A, B \Vdash C$ and $A, C \Vdash B$; that is, **a** and **c** are satisfied. It is true also that if $B \Vdash C$ then $B \Vdash A$, and hence $B \Vdash A \wedge C$, contradicting $(\boldsymbol{\alpha})$; whence $B \nVdash C$, and **b** is satisfied. By § **3**, for there to be any B satisfying $(\boldsymbol{\alpha})$, it is necessary and sufficient that $(A \leftrightarrow C) \wedge \neg(A \wedge C)$, which is identical with $\neg A \wedge \neg C$, be consistent and not maximal. But the consistency of $\neg A \wedge \neg C$ is equivalent to $\neg A \nVdash C$, which by § **7**, is necessary and sufficient for **a–c** to have solutions.

By § **7** again, if **a–c** are satisfied, then **d** is satisfied if & only if $C \nVdash A$.

In brief, provided that $\neg A \wedge \neg C$ not a maximal theory, and that the conditions **a–c** have solutions, then there exists a statement B such that $A \wedge C \vdash B \vdash A \leftrightarrow C$, and every such statement B is a solution of **a–c**. For B to be a solution of **a–d**, it is necessary and sufficient that $C \nVdash A$.

10 $(\boldsymbol{\beta})$ Solutions properly between $A \leftrightarrow C$ and $A \to C$

It is clear that if the biconditional $A \leftrightarrow C$ and the conditional $A \to C$ are logically equivalent (that is, identical) the ranges $(\boldsymbol{\beta})$ and $(\boldsymbol{\gamma})$ identified in § **8** are empty, and there is nothing more to be said. Since the equivalence holds if & only if $C \Vdash A$, we now add to the assumption ♠ that $\neg A \nVdash C$ the assumption that $C \nVdash A$. The combined assumptions may be written

♣ $$A \wedge C \vdash C \vdash A \to C.$$

As shown in § **7**, ♣ is necessary & sufficient for **a–d** to have solutions.

It follows from § **7** that if **a–c** has any solutions then every B that satisfies $(\boldsymbol{\beta})$ $A \leftrightarrow C \vdash B \vdash A \to C$ is also a solution of **a–c**. For $A \to C$ satisfies **a**, and therefore any stronger B does; while $A \leftrightarrow C$ satisfies **b** and **c**, and therefore any weaker B does. By § **3**, for there to be any B satisfying ($\boldsymbol{\beta}$), it is necessary and sufficient that $(\mathsf{A} \to \mathsf{C}) \wedge \neg(\mathsf{A} \leftrightarrow \mathsf{C})$, which is identical with $\neg\mathsf{A} \wedge \mathsf{C}$, be consistent and not maximal. But for $\neg\mathsf{A} \wedge \mathsf{C}$ to be consistent it is necessary & sufficient that $\mathsf{C} \nvdash \mathsf{A}$. Since $\mathsf{C} = \mathsf{A}$ is ruled out by ♣, this condition can be strengthened to $\mathsf{C} \nVdash \mathsf{A}$.

Moreover, if $\mathsf{C} \Vdash \mathsf{B}$ then $\mathsf{C} \vee (\mathsf{A} \leftrightarrow \mathsf{C}) \Vdash \mathsf{B}$; that is, $\mathsf{A} \to \mathsf{C} \Vdash \mathsf{B}$, which ($\boldsymbol{\beta}$) declares impossible. In other words, **d** is satisfied by any B that lies in the open interval between $\mathsf{A} \leftrightarrow \mathsf{C}$ and $\mathsf{A} \to \mathsf{C}$ under consideration.

In brief, provided that $\neg\mathsf{A} \wedge \mathsf{C}$ is not a maximal theory, and that the conditions **a–c** have solutions, then there exists a statement B such that $\mathsf{A} \leftrightarrow \mathsf{C} \vdash \mathsf{B} \vdash \mathsf{A} \to \mathsf{C}$, and every such statement B is a solution of **a–d**.

11 Conclusion

It is not been possible to determine whether there are solutions to **a–c** or **a–d** that lie within the range (γ), but enough has been said to make it evident that, except in extreme circumstances, even the stronger set of conditions does not allow identification of the assumptions missing from invalid arguments. I doubt that those in the critical thinking movement will be much impressed by this technical result. Nevertheless, the status and significance of the negative test remain as obscure as they ever were.

References

[1] P. Gardner and S. Johnson. "Teaching the Pursuit of Assumptions". In: *Journal of Philosophy of Education* 49.4 (2015), pp. 558–570.

[2] D. Hume. *An Abstract of a Book Lately Published, Entitled A Treatise of Human Nature, Etc.* 1740. (Ed. by L. Selby-Bigge. Oxford: Clarendon Press 1888, 1978).

[3] D. Hume. *An Enquiry Concerning Human Understanding.* 1748. (Ed. by L. Selby-Bigge. Oxford: Clarendon Press 1893, 1975).

[4] D. W. Miller. *Critical Rationalism. A Restatement and Defence.* Chicago and La Salle IL: Open Court Publishing Company, 1994.

[5] D. W. Miller. "Do We Reason When We Think We Reason, or Do We Think?" In: *Learning for Democracy* 1.3 (2005), pp. 57–71.

[6] D. W. Miller. *Out of Error. Further Essays on Critical Rationalism.* Aldershot and Burlington VT: Ashgate Publishing Company, 2006. (now published by Routledge, Abingdon and New York).

[7] D. W. Miller. “Mashinnoye oogadivaniye, chast’ II”. In: *Voprosy Filosofii* 8 (2012), pp. 117–126. (Spanish translation, “Adivinación automática”, in G. Guerrero P. & L. M. Duque M. (eds.), *Filosofía de la ciencia: Problemas contemporáneos*, pp. 19–44, Programa Editorial, Universidad del Valle, Santiago de Cali, 2015. English version, `http://www.warwick.ac.uk/go/dwmiller/jsm.pdf`).

[8] A. E. Musgrave. “Popper and Hypothetico-Deductivism”. In: *Handbook of the History of Logic. Inductive Logic*. Ed. by D. M. Gabbay, S. Hartmann, and J. Woods. Vol. 10. Handbook of the History of Logic. Amsterdam: Elsevier B. V., 2011, pp. 205–234.

[9] A. Tarski. “Grundzüge des Systemenkalkül”. In: *Fundamenta Mathematicae* 25.4 (1935), pp. 342–383, 26.2 (1936) pp. 283–301. (English translation, “Foundations of the Calculus of Systems”, Chapter XII, of A. Tarski, *Logic, Semantics, Metamathematics*, Oxford: Clarendon Press, 1956, 2nd edition, Indianapolis: Hackett Publishing Company, 1983).