L. FLORIDI

ON THE LOGICAL UNSOLVABILITY OF THE GETTIER PROBLEM

ABSTRACT. The tripartite account of propositional, fallibilist knowledge that *p* as justified true belief can become adequate only if it can solve the Gettier Problem. However, the latter can be solved only if the problem of a successful coordination of the resources (at least truth and justification) necessary and sufficient to deliver propositional, fallibilist knowledge that *p* can be solved. In this paper, the coordination problem is proved to be unsolvable by showing that it is equivalent to the "coordinated attack" problem, which is demonstrably unsolvable in epistemic logic. It follows that the tripartite account is not merely inadequate as it stands, as proved by Gettier-type counterexamples, but demonstrably irreparable in principle, so that efforts to improve it can never succeed.

1. INTRODUCTION

According to the tripartite account of propositional and fallibilist knowledge that *p*, an epistemic agent *S* knows that *p* if and only if

(i) *p* is true,
(ii) *S* believes that *p*, and
(iii) *S* is justified in believing that *p*.

However, well-known Gettier-type counterexamples prove that this version of the tripartite account is inadequate. Conditions (i)–(iii) may be necessary but are certainly insufficient to define propositional knowledge, since they fail to ensure *S* against mere epistemic luck (Gettier 1963).

Most epistemologists agree that Gettier-type counterexamples pose a genuine challenge (see for example the reviews offered by Dancy 1985; Dancy and Sosa 1992; Everitt and Fisher 1995; Hetherington 1996; Steup 1996; and Greco and Sosa 1999). Many hope that the challenge may be met by revising the tripartite account to avoid the counterexamples without incurring new difficulties. There are two interpretations of this strategy. One, incorrect, argues that if the counterexamples are avoidable then the account can become adequate, that is:

LEMMA 1*. (i) if Gettier-type counterexamples are avoidable, at least in principle, then the tripartite account can become adequate, at least in principle; (ii) Gettier-type counterexamples are avoidable, at least in principle; therefore (iii) the tripartite account can become adequate, at least in principle.

Lemma 1* begs the question. If one could prove (ii) that the counterexamples are indeed avoidable, one would have proved that the tripartite account can become adequate, thanks to (i), but (i) is acceptable only if one already assumes (iii), that is, only if one already believes that the tripartite approach is a step in the right direction, but this is precisely the point in question. Thus, according to Lemma 1*, proving that Gettier-type counterexamples are unavoidable in principle would not affect the potential adequacy of the account, a clear *non sequitur*.

The correct interpretation argues that if the tripartite account can become adequate, at least in principle, then Gettier-type counterexamples must be avoidable, at least in principle, and hence that a successful strategy must prove that they are not demonstrably unavoidable:

LEMMA 1. (i) if the tripartite account can become adequate, at least in principle, then Gettier-type counterexamples are avoidable, at least in principle; but (ii) Gettier-type counterexamples are not avoidable, even in principle, therefore (iii) the tripartite account is irreparably inadequate in principle.

Correctly, Lemma 1 does not presuppose the adequacy of the tripartite account. This is the lemma that will be taken into consideration.

The crucial point in lemma 1 is to try to show that (ii) is not the case. Now this may be attempted by revising the tripartite account in only three ways:

(a) by strengthening/modifying the only flexible feature of the account, namely the justification condition (iii) (Chisholm 1989); or
(b) by adding at least one more condition that would prevent the Gettierization of the required justified true beliefs or, alternatively, allow their de-Gettierization; or
(c) by combining (a) and (b).

No other general strategies are available, in the sense that anything more radical than (a)–(c) amounts to a *de facto* rejection of the tripartite account. Plato, for example, departs from it, after having considered its viability in the *Thaetetus*.

Let me recapitulate. If there is any chance that the tripartite definition of knowledge may ever become adequate, it must somehow be possible to avoid or overcome Gettier-type counterexamples, at least in theory. In order to show that the counterexamples are avoidable, one may try to revise the definition in three ways.

Each of the three strategies has been probed and applied in various ways, usually following the reasonable maxim of keeping changes to a minimum. Yet four decades of relentless effort have yielded no ultimate solution or even a point of convergence (Griffiths 1967; Roth and Galis 1970; Pappas and Swain 1978; Pappas 1979, Shope 1983, Plantinga 1993b, Floridi 1996, Steup 2001). This raises circumstantial doubts as to whether Gettier-type counterexamples may indeed be avoidable at all (Kirkham 1984; Schreiber 1987; Zagzebski 1994) without abandoning the tripartite approach. This paper sets out to demonstrate that they are not avoidable and, most importantly, to explain why they can never be, no matter how the tripartite account is revised, improved or expanded. It follows that the tripartite account is not merely inadequate as it is, but demonstrably irreparable in principle. We should stop trying to fix it and start looking for a different approach.

2. WHY THE GETTIER PROBLEM IS UNSOLVABLE IN PRINCIPLE

To prove that the tripartite account is irreparably inadequate in principle, it is sufficient to prove three more lemmas:

LEMMA 2. All Gettier-type counterexamples are instances of a single Gettier Problem (GP).

LEMMA 3. (i) If the tripartite account can become adequate, at least in principle, then GP can be solved, at least in principle; but (ii) GP is not solvable, even in principle, therefore (iii) the tripartite account is irreparably inadequate, even in principle.

LEMMA 4. GP is logically equivalent to the so-called "coordinated attack" problem.

As we shall see, a group of important theorems in epistemic logic proves that the "coordinated attack" problem and some of its variations are logically unsolvable. Thus, proving Lemma 4 means proving that the Gettier problem too is logically insolvable (one of the first to draw attention to the connection between Gettier's analysis and epistemic logic is Lenzen 1978). Once this is established, it is simple to see that Lemma 3 shows that the tripartite account is irreparably inadequate in principle. Of course, all this applies only given the constraints posed by the tripartite account itself. The proviso is crucial, and more will be said on it in the conclusion. For the moment, consider the problem of squaring the circle. The problem is not constructing a square equal in area to a circle, but doing so by using only algebraic means (straight-edge and compass). Once Lindemann proved that π is transcendental (not an algebraic number of any degree), it became clear that solving the problem and satisfying its constraints were mutually exclusive. We shall see that the same holds true of GP. Given the conditions set up by the tripartite account of knowledge, Gettier-type counterexamples are unavoidable in principle, no matter what new strategies are then adopted to improve the account. It seems that Plato's view is vindicated: the very idea of defining knowledge on a doxastic basis, in terms of true justified belief, proves to be misguided. Let us now turn to the proof.

Lemma 3 simply follows from Lemmas 1 and 2 and Lemma 2 is trivial. Epistemologists agree (see for example Steup 1996) that there are countless Gettier-type counterexamples but only one logical problem, namely a *lack of successful coordination* (more on this in Section 3) between the truth of p and the reasons that justify S in holding that p. A Gettier-type counterexample arises because the truth and the justification of p happen to be not only independent (as they should be, since in this context we are dealing with fallibilist knowledge) but also opaquely unrelated, that is, they happen to fail to converge or to agree on the same propositional content p in a relevant and significant way, without S realising it (*Gettierization*). Once this feature is grasped, anyone can produce her own favourite counterexample. Thus, Shope (1983) lists 98 examples in the literature and Zagzebski (1994) provides an elegant recipe for cooking up your own. Yet nobody would expect each counterexample to require its own specific solution.

If you are not already convinced, consider the following argument. Suppose one argues that Gettier-type counterexamples can be avoided by making sure that the truth and the justification of p are successfully coordinated, and hence that all one needs to add to the tripartite account is a fourth clause specifying that:

(iv) the relationship between the truth of p and S' justification for p is successfully coordinated.

Adding (iv) would be begging the question because it would be equivalent to adding a clause specifying that

(iv*) the relationship between the truth of p and S' justification for p is not Gettierisable.

And clause (iv*) is precisely what the revised version of the tripartite account needs to achieve, in order to qualify as an adequate account of propositional knowledge, not something that can merely be decreed by *fiat*. Now the fact that (iv) and (iv*) are logically equivalent shows that Gettier-type counterexamples are caused by a lack of successful coordination between the truth and the justification of p, namely GP.

The demonstration of the logical unsolvability of GP really rests on the possibility of proving Lemma 4, the only one which is not trivial. The proof can be introduced by considering a familiar Gettier scenario. Note that, in view of further discussion in Section 3, it will be useful to pay attention to the fallibilist nature of the types of knowledge discussed.

John Smith has dental problems. Two molar teeth in his right mandible have initial interproximal caries (known as IIC). His dentist, Tracy (in the following analysis she will stand for the truth resource, hence the T for Tracy), suspects that *John's teeth have IIC* (call the sentence in italics p), but she is unable to detect its presence by clinical observation. Her true belief that p is therefore an unsubstantiated intuition, a lucky guess. However, Tracy knows that visual detection of IIC is often difficult, so she refers John to a dental radiologist, Jane (in the following analysis she will stand for the justification resource, hence the J for Jane), for a CDR (Computer Digital Radiography). Taking a CDR is usually a reliable procedure to diagnose IIC, though of course it is entirely fallible. Suppose the CDR shows that John's molar teeth are affected by IIC. Jane has very strong evidence in favour of p. However, unaided by Tracy, she too cannot correctly claim to know that p. Interpreting a CDR is a procedure that requires some expertise, so there is a chance that Jane is mistaken and holds a false belief. At this point, Tracy by herself does not yet know that p because she might be merely lucky, and Jane does not yet know that p because she might be wrong. GP is going to affect Tracy, while sceptical problems affect Jane (but this is another story). The hope is that Jane and Tracy may be individually necessary and jointly sufficient. Clearly, they need to coordinate their efforts. Jane emails the CDR to Tracy. Unfortunately, she sends the CDR of a homonymous patient, who also suffers from IIC. Were Tracy to rely on this piece of evidence, she would be justified in believing that p, yet she would still not know that p, for she would be merely lucky. As it happens, Tracy notices a number of inconsistencies between the CDR and what she knows about her patient. She concludes that there must have been a mistake and that the allegedly supporting evidence is in fact irrelevant. So, she asks Jane to make sure that she sends the CDR of the right John Smith. Jane sends a new email, this time with the relevant CDR. Unfortunately, the actual traces of IIC in John's molar teeth have been transposed during the imaging process. This is unlikely but possible. John's mandible was not optimally positioned and now it looks as if John has two molar teeth with IIC in the right *maxilla*. Again, were Tracy to rely on this second CDR, she would have a true and justified belief that p, yet still fail to know that p. A less experienced dentist might be fooled, but not Tracy. Noticing some anomalies in the shape and granularity of the CDR, she asks Jane to re-process the image. Finally, a correct CDR for the right John Smith reaches Tracy. Unfortunately, it fails to show the caries because of their very early stage of development. At the same time, the CDR can be interpreted as showing that John has IIC in two molar teeth of the right mandible, due to the presence of tartar and some noise in the data. At this point, there are two scenarios. Tracy may no longer trust her source of justification and so suspend her epistemic commitment. She does not claim to know that p, but opts for some epistemically weaker attitude, for example she says the she suspects that p or that she is quite confident that p. Alternatively, Tracy may rely completely on the evidence provided by the radiography, concluding correctly, but only by chance, that John suffers from IIC, a typical Gettier case. If she is epistemically cautious, she will not immediately operate on John, thus making a mistake, although arguably a small and recoverable one. If she makes an epistemic mistake and misreads the CDR, she will operate on John immediately, thus succeeding in her duties, to the advantage of John's health, yet merely by chance. In either case, she does not know that p.

The example presents GP in a distributed system scenario. The two resources, truth and justification, are introduced as agents interacting to achieve a common goal. This feature is not very common in the literature but it is useful to add further generality and clarity to the present analysis. Interpreting the epistemic subject S as a stand-alone, single agent is only a case limit (the Cartesian subject) and not even the most interesting. On the contrary, looking at the problem from a multi-agent, distributed system perspective, we can more easily identify GP as a problem of coordination between resources, in the following way.

Consider our two agents in more general and abstract terms, as parts of a simple, multi-agent, distributed system (for an introduction to agents and distributed systems see Wooldridge 2002). When speaking of agents in this sense it is vital that we are clear that these are not knowing subjects like S. On the contrary, like the generals in the "coordinated attack" problem (see below), Tracy and Jane are (clusters of) *resources* that have to be co-ordinated to deliver a product, in this case knowledge. More specifically, Tracy is any truth-producing oracle T, consisting of whatever resources are sufficient to generate $n \geq 1$ true propositions $\{p_1, p_2, \ldots, pn\}$. Jane is any justification-producing reasoner J, consisting of whatever resources are sufficient to justify T's true propositions. Their shared goal is to deliver propositional knowledge that p for $p \in \{p_1, p_2, \ldots, p_n\}$. One can picture this as the goal of defeating a third agent, Charles, a propositional-knowledge challenger C consisting of whatever resources are sufficient to prove that no propositional knowledge that p has been delivered. Let us assume the most favourable case in which

(a) T and J are non-faulty (they never fail to behave according to their specifications). Note that this condition is not essential, but just a matter of convenience. In this context we shall deal with the case that is most favourable to the tripartite account. If the agents can be faulty, scepticism arises, and one has the less favourable case represented by the untrustworthy agents known as the "Byzantine generals", see Pease et al. (1980) and Fagin et al. (1995);
(b) the communication medium between T and J is reliable and fault-tolerant but (provably) not fault-free. This is equivalent to saying that the case of knowledge in question is fallible;
(c) T and J deal with the same p;
(d) T and J are individually necessary to produce propositional knowledge that p (i.e., to defeat C);
(e) if T and J can coordinate their efforts successfully, then they are also jointly sufficient to produce propositional knowledge that p (i.e., to defeat C).
(f) T and J are *non-strategic agents*. Strategic agents act in their own interests, whereas *non-strategic* agents follow rules given to them; this assumption makes more precise the intuitive view that there is some sort of harmony between the justification and the truth of p. Again, the condition is assumed for the sake of simplicity.

Enquiring whether the tripartite account can be revised, so that it provides an adequate analysis of propositional knowledge in terms of necessary and sufficient conditions, means enquiring whether one can ensure that the two *agents* T and J can defeat the third *agent* C. The trivial answer is that, in order to ensure their victory, it is sufficient to ensure that T and J succeed in coordinating their efforts. This prompts the interesting question whether, given conditions (a)–(f), there is indeed a way in which the two agents can interact through some communication protocol that guarantees that they succeed in coordinating their efforts. This is equivalent to saying that the tripartite account can become adequate only if GP can be solved, and that GP can be solved only if the problem of a successful coordination between the two agents T and J (the truth and the justification of p) can be solved. And this is Lemma 4.

Although not in the same terms, the general point made by lemma 4 is often stressed by some of the best and most influential analyses of GP, such as Goldman (1967) and Nozick (1981). We have seen that possible strategies to achieve indefeasible coordination (not *indefeasible knowledge*) comprise a modification of the nature of clause (iii) in the tripartite account, and/or an addition of at least a fourth condition. Unfortunately, the coordination problem is demonstrably insolvable (Gray 1978 and Halpern and Moses 1990 provide full coverage of the proof; the reader will find in the Appendix a brief summary of the relevant theorems). This is so no matter which strategy is adopted (here the proviso discussed above in connection with squaring the circle applies). Let us see why.

T and J are any two agents/resources that are individually necessary to achieve a particular goal – in our case, achieving propositional and fallible knowledge that p by defeating C – but need to be successfully coordinated (i.e., need to interact in a certain way, which can be left unspecified here, for reasons given in Section 3) to become jointly sufficient as a dynamic system. The coordination problem arises because the two resources T and J are not only logically but also empirically independent, so they do not yet deliver knowledge (let alone indefeasible knowledge), but need to rely for their communication/coordination on some empirical interaction, which cannot be assumed to be completely fault-free. Now, this system can be elegantly modelled in terms of a distributed, asynchronous message-passing system – like Tracy and Jane, or two divisions of an army attacking a common enemy, or three or more Byzantine (i.e., unreliable) generals, or T and J playing against C. Or it can be modelled by a synchronous message-passing system, in which message delivery is as reliable as one might wish but still not fault-free, that is, a system in which a message can take an arbitrarily long time to arrive (on this distinction see Halpern 1995). Since the tripartite account aims at establishing necessary and sufficient conditions for propositional knowledge, the question whether GP is solvable in principle is equivalent to the question whether there can be a time t at which the n (for $n \geq 2$) agents involved are successfully coordinated with respect to p. In the case of a message-passing system, the latter question is modelled as the question whether there is a communication protocol that can guarantee coordination between the n agents at a certain time in the future with respect to p. No protocol satisfies these requirements. This is proved in terms of a *regressus ad infinitum*. T and J are separate and independent agents playing against a third agent C. It is clear that if both agents play against C simultaneously they will defeat him, while if only one agent plays against C she will be defeated. The agents do not have pre-established strategies (this means that we are dealing with fallible knowledge) and T wishes to coordinate a simultaneous move against C at some time t. Neither agent will play unless she is sure that the other will play at the same time. In particular, an agent will not play if she receives no messages. The agents can communicate by means of messages. It takes a message some time $t \geq 1$ to get from the sender to the receiver. However, the message may get lost or corrupted. How long will it take them to coordinate a move against C? Suppose T sends a message m to J saying "Let's play move s against C at time t". J receives m. Now J is informed that m (in standard notation: $K_J m$), but will J play move s? Of course not, since the channel of communication is not fault-free, and T cannot be sure that she (J) received the message she (T) sent, and thus may not play. So J replies with an acknowledgement. Suppose the message reaches T. Now $K_T K_J m$ holds. So will T play move s? No, because now J does not have the information that T received the message, so J thinks that T may think that she (J) did not receive the original message, and thus not play. The next acknowledgment brings us to $K_J K_T K_J m$, and so forth. The regressus *ad infinitum* is obvious. Each time a message is received, the depth of the agents' information increases by one, yet there is no stage at which they are both informed that they are both informed that ... they will play move s at time t. The agents never attain common information (basically in the technical sense of "common knowledge", see Fagin et al. (1995)) that the move s is to be played at time t, because there is no protocol of communication that allows the distributed system to reach the established *fixed point*. As long as there is a possibility that the message may be lost or corrupted – and this possibility is guaranteed by the empirical and hence fallible nature of the interaction between the agents – common information is unattainable, even if the message is in fact delivered.

To summarise: successful coordination is a prerequisite for guaranteeing a successful game move, but common information is a prerequisite for guaranteeing successful coordination, and common information is unattainable in any distributed system in which there is any doubt at all about message delivery time. Such doubt is inevitable if the agents are at least logically independent and must interact through empirical protocols. The tripartite account sets up exactly such a distributed system. T and J are logically separate resources in need of empirical coordination to be able to deliver knowledge that p. We now know that there is no communication protocol that guarantees that they will be successfully coordinated to produce knowledge that p. Of course, this does not mean that they cannot be coordinated sometimes or even often, or that sub-optimal strategies cannot be devised (Halpern and Tuttle 1993; Morris and Shin 1997), but it does prove that counterexamples are inevitable in principle. The epistemic agent

S may know that p but has no way of ascertaining that the truth of *p*, i.e., the resource *T* to which *S* has access, and the justification for *p*, i.e., the resource *J* that is also available to *S*, are sufficient to provide *S* with propositional knowledge that *p*, unless *S* can be sure that they are indeed successfully coordinated with respect to *p* (a similar point is emphasized by Apel 1975 and Alston 1986). But the latter condition is unachievable in principle given the system set up by the tripartite account.

3. COMMENT

The proof can be further clarified by considering some potential objections.

One might ask whether the argument presented above implicitly presupposes an interpretation of knowledge as *indefeasible*, *certain* or *infallible*. All these positions are at least controversial and would be utterly inadequate in the context of a discussion of the tripartite account of knowledge, which, as we have seen, explicitly addresses empirical knowledge of a *fallibilist* kind.

This concern is reasonable, and its roots may be traced to the technical vocabulary required by the analysis of the coordination problem. The reader will recall, for example, that it has been argued that the question whether GP may be solvable in principle is equivalent to the question whether there can be a time *t* at which the resources involved are successfully coordinated with respect to *p*. However, this concern can be allayed once we clearly distinguish between (i) how one qualifies the kind of coordination required between the resources that are necessary and sufficient for a successful delivery of fallible knowledge and (ii) how one qualifies the knowledge delivered. The medical example chosen in the discussion and the careful constraint placed on any sceptical drift were meant to facilitate and support this distinction, but an analogy may make it sharper. Suppose that two independent processes, say packaging and handling, are individually necessary and jointly sufficient to deliver a box of fresh eggs unbroken to your house, as long as they are successfully coordinated. One may qualify this condition – successful coordination – in several ways. Suppose that, unless the coordination is 100% successful, there is no guarantee that the eggs will be delivered unbroken. Of course none of this is going to affect the intrinsic fragility of the eggs. Nor is it equivalent to saying that the eggs cannot be delivered unbroken when coordination is less than fully successful. The delivery may be successful just by luck. Let us now assume that the shop will not send you the eggs unless there is 100% successful coordination between the two processes. The proof offered in the previous section shows that such a level of coordination is unattainable in principle, no matter how one modifies the processes involved or how far one extends their number and scope. Of course, the proof may still be questioned – but not the fact that the attainability (or unattainability) of coordination is independent of the "fragility" of the specific case of fallibilist propositional knowledge taken into consideration. To use the terminology of the "coordinated attack" problem, unless the coordination is guaranteed, only a risky attack can be launched, but this begs the question, since the problem requires the launching of a safe attack. The necessary resources are also *jointly* sufficient only in a "well-coordinated" sense of "jointly". Obviously, most of the time one can have knowledge that p by having true and justified beliefs that p, but this is not what we are looking for. We are seeking instead a definition of knowledge in terms of necessary and sufficient conditions, such that, if obtained, they successfully deliver "fragile" knowledge "unbroken".

A different but related concern, because it still addresses the correct focus of the argument, can be phrased thus: suppose the argument does prove that Gettier counterexamples are inevitable in principle, isn't this a solution to Gettier's challenge? Yet the challenge was not (a) for us never to be Gettierized, but it was (b) for us to understand what it is not to be Gettierized. So there seems to be a worrying level-confusion afflicting the argument.

This new concern can be resolved in two steps. First, the argument does provide a solution to the Gettier Problem, but a negative one. The real achievement would not be to show that the tripartite account of knowledge is inadequate - this is precisely what the Gettier Problem is about - but to prove that the Gettier Problem is insolvable *no matter how one tries to revise and improve the original account*, and hence that the inadequacy of the account cannot be remedied. This clarification leads to the second step, which concerns the alleged confusion between (a) and (b). It is certainly important to distinguish between understanding a challenge, as stressed in (b), and meeting it successfully, as specified in (a). But it is also equally important to understand that, in a negative proof, (a) and (b) are strictly connected. Recall the comparison with the squaring of the circle: a better understanding of the mathematical nature of π leads to a proof that it is impossible. Now the argument proceeds exactly along the same lines. It shows that the Gettier Problem is a special case of the "coordinated attack" problem by showing that it is a problem about the coordination between whatever resources are deemed necessary and sufficient to deliver knowledge. This amounts to understanding what it is for us not to be Gettierized, that is, (b). But once this is clear, it becomes equally clear that the problem is unsolvable. And this amounts to proving that it is impossible never to be Gettierized, given the preconditions set by the standard account, that is, (a). If the challenge is properly understood, it becomes clear that it cannot be met.

Once the focus of the argument is fully vindicated, one may still have reservations about its formulation. One may suspect that the argument presented above depends on some equivocation regarding the crucial concept of "coordination". For (1) "coordination" between T and J is presumably meant to correspond to the Generals' *common knowledge* that they will both attack. But (2) there is no warrant for assuming that the reason why knowledge fails in Gettier-type counterexamples is lack of common knowledge of justified true belief, or anything sufficiently formally analogous to it to satisfy the conditions for a generalized version of the negative findings concerning a coordinated attack. (3) Any sense of "coordination" that allows epistemologists to agree that lack of coordination between truth and justification is the key to the Gettier problem (as claimed above) is an extremely vague and unspecific one, and little more than a label for the problem. (4) In particular, it is not agreed that coordination is an epistemic relation rather than one of, e.g., causation or counterfactual dependence. In the one-person case, having common knowledge that one has justified true belief entails knowing that one knows that the belief is true; but since many epistemologists reject the KK principle ("positive introspection"), common knowledge cannot be assumed to be necessary for knowledge itself. Similar considerations apply in the many-person case. But if this is the case, then (5) no plausible *prima facie* case for lemma 4 has been made, which is essential to the argument of the paper.

The objection highlights some important features of the argument. Regarding (1), the suggestion is plausible but mistaken. The confusion concerns the level of analysis. Coordination between T and J does not correspond to the generals' *common knowledge* about the correct message m containing the time at which they will attack but to their coordination, interpretable as the synchronization of their actions. The generals' synchronization can only be guaranteed by, but is clearly different from, their common knowledge of m (for example, the generals may be lucky and be coordinated even without common knowledge). Their synchronization in turn guarantees, but is also different from, the safety of their action and hence the successfulness of their attack. The attack never takes place because (a) the generals attack only if "failure is not an option" and (b) their common knowledge of m requires perfectly fault-free communication, which is unobtainable in the given circumstances.

Since (1) is not the case, (2) is correct but does not apply. The argument is not that knowledge fails in Gettier-type counterexamples because the epistemic subject S lacks common knowledge of justified true belief. If this were the case, the argument would indeed have to be rejected. The argument is that knowledge fails in Gettier-type counterexamples because there are cases in which, although T and J are both available to S, one can still show that there is no coordination between T and J (or better: Gettier-type counterexamples prove that it is impossible to guarantee that an epistemic commitment by the system $T + J$ will be safe and hence successful in delivering propositional knowledge that p). This possible lack of coordination cannot be overcome because it is caused by a lack of common knowledge, not by S, but between the two agents T and J, and not of justified true belief, but of the relevant circumstances, that is, coordination, in which the system can make a safe epistemic commitment. Hence S' lack of common knowledge of justified true belief is irrelevant here. Compare the generals' common knowledge of m as against the generals' actual launch of the attack. The lack of common knowledge between T and J is caused by the absence of fault-free communication, which is unobtainable given the specified constraints. In other words, $T + J$ cannot guarantee delivery of propositional knowledge. Thus, the point of the argument lies in modelling T and J as two agents/resources and their interaction as a message-passing procedure. In this way, one can appreciate the fact that Gettier-type counterexamples show that, no matter how many times the two agents T and J check and double-check that they are properly coordinated, no fixed point can ever be reached. It is always possible that T and J may in the end fail to be coordinated.

Regarding (3), the objection is correct in pointing out the unspecified nature of the relation of coordination between T and J. However, it is *useful* to keep the relation unspecified precisely because this makes the result applicable to any interpretation of it (see points (1)–(3) in Section 4). And it is *not necessary* to specify the nature of the coordination relation. This is so because the failure to deliver propositional knowledge does not depend on a particular interpretation of it, but on the fact that, whichever way the relation is interpreted, it requires common knowledge, which in turn requires communication between the n agents involved, and the communication, being fallible, can never eliminate the possibility of Gettier-type counterexamples.

Regarding (4), we have seen that the argument does not equate successful coordination and common knowledge, contrary to the assumption in objection (1). Nor does it imply any other epistemic interpretation of coordination, for it leaves it unspecified. As for the acceptance of the KK thesis, the objection is correct in pointing out that KK is controversial, but this is also as far as the objection can go, for two reasons. First, suppose the argument did entail KK; this might still be taken as a reason in favour of the popularity of epistemic systems such as S4 and S5, rather than a *reductio*, so the objection may be answered on a purely logical ground. Second, and most importantly, the argument does not entail KK in any way that facilitates the objection in the first place, so there is no problem. According to the argument, for the *epistemic commitment* of the system $T + J$ to be safe, T and J need to be successfully coordinated. For this to be the case, T and J need to achieve common knowledge of the relevant circumstances in which they can make a safe epistemic commitment, i.e., the message m. It is important to clarify what the claim is here. Common knowledge between T and J is not *necessary* for knowledge itself – we have seen in Section 2 that the agents may decide to adopt sub-optimal strategies, which can still deliver a result sometimes – but it is necessary for any epistemic commitment that needs to be safe, i.e., that guarantees the fulfilment of necessary and sufficient conditions for the delivery of propositional and fallible knowledge that p. Recall that the original theorem proves that the generals never attack, given the constraints in place; it does not prove that they could not win, should they decide to attack anyway. The generals can win the battle without having common knowledge, but they should not commit themselves to attacking the enemy in the first place, given the constraints they share. Likewise, T and J may deliver prepositional knowledge without sharing common knowledge (so $K_S p$ does not require or entail $K_S K_S p$), but no epistemic commitment (no delivery of p) is guaranteed to be successful without their coordination, which requires them to achieve common knowledge. This common knowledge requires fault-free communication, which is not achievable because of the constraints posed by the tripartite account itself: $T + J$ cannot claim $K_S p$ with total certainty without T and J being successfully coordinated, something unachievable given the tripartite account.

4. CONCLUSIONS

By way of conclusion, it is now worth stressing the generality and robustness of the result.

First, it would be a mistake to interpret the coordination problem as a mere message-passing issue. For the latter is not the difficulty itself but an elegant way of modelling the dynamic interactions between $n \geq 2$ agents (resources, processes, conditions, etc.) to prove that the goal of ensuring successful coordination in a distributed system, such as the tripartite account, is insurmountable. The real difficulty is that, if T and J are independent (as they should be, given the fact that we are speaking of empirical, fallible knowledge), the logical possibility of a lack of coordination is inevitable and there is no way of making sure that they will deliver knowledge in a Gettier-proof way. In the absence of coordination, the agents will play only at the risk of defeat, or they will not play, if the cost of a defeat is too high. Likewise, since it is one of the tripartite account's constraints that T and J are not pre-coordinated, the system $T + J$ will at best be able to claim to know that p only defeasibly (Quine *docet*), or will not commit itself epistemically (this is the initial Cartesian option in the *Meditations*, while waiting for the discovery of the *Cogito*). The possibility of GP can not be eliminated. Recall Tracy the dentist. In the end, either she trusts the CDR, inadvertently running the risk of not knowing that p, or she suspends her commitment, admitting that she may still not know that p, even if she is absolutely right about p. Either way, she never reaches the time t at which one (or she) can say for sure that she knows that p. The conclusion is that GP is logically unsolvable and now we know why. The tripartite account asks us to find a way to coordinate T and J successfully while satisfying constraints – empirical interaction between T and J as two independent resources – that, by their very nature, make it impossible to achieve the set goal, exactly like squaring the circle.

Second, it should be clear that the logical insolvability of the Gettier Problem as a special case of the "coordinated attack" problem holds true independently of any modification in the nature of J (including its relation to T) and/or any addition of extra agents beside J and T. Indeed, other variables can play no useful role. The following is a list of those that have attracted most attention in the literature. The Gettier Problem/"coordinated attack" problem are logically unsolvable

1. whatever the nature of the coordination protocols relating T and J, including truth-tracking and a (non-instantaneous) synchronic, causal interaction; recall the case of the synchronous message-passing system outlined above;
2. whatever the degree of reliability satisfied by the coordination protocols;
3. whatever, and no matter how many relations (e.g., internalist, externalist) may occur between the truth and the justification of p;

(the previous three points further clarify why it is useful to leave the nature of the coordination unspecified);

4. whatever interpretation is offered of the concepts of truth and justification. The two agents, players, or divisions attacking a common enemy, can be as strong as one wishes and the two generals as smart as Alexander *T*he Great and *J*ulius Caesar;

5. whether the original concepts, especially justification, are replaced by other resources, e.g., well-foundedness or warrant (Plantinga 1993a). This is convincingly argued by Crisp (2000) and Pust (2000), even if in rather different terms;
6. whether any other agent (any other epistemic resource) is added, as long as all agents are still individually necessary and jointly sufficient only if successfully coordinated. Indeed, the analysis of the "Byzantine generals" is usually based on $n \geq 3$ agents, see Fagin et al. (1995);
7. whether one models the various components as $n \geq 2$ agents in a distributed system or as $n \geq 2$ processes in a single agent (even in the latter case, it is a matter of granularity).

All this holds true *provided* that the $n \geq 2$ epistemic resources in question are logically independent and need to be successfully coordinated through a less than fault-free communication protocol to achieve their common goal.

Third, the previous proviso is crucial, as it makes clear that many alleged "solutions" of the Gettier Problem/"coordinated attack" problem turn out to be fallacious. For they all presuppose or advocate some form of "pre-established harmony", such as pre-coordination, instantaneous synchronicity, reduction of the $n \geq 2$ agents T and J to 1, or some fault-free and simultaneous protocol of communication between T and J. Similar "Leibnizian" strategies do implement perfect coordination, and hence solve GP, but, apart from being unrealistic, they fail to respect the tripartite account's empirical constraints. Since p is an empirical and fallible proposition, its T and J cannot be assumed to be pre-coordinated a priori in a way that makes their lack of coordination logically impossible, so similar strategies beg the question, in that they are all equivalent to a mere "no-Gettierization" assumption.

Since GP is demonstrably unsolvable, it follows not only that the tripartite account is logically inadequate as it is, but also that it is irretrievably so in principle. GP is not a mere anomaly, requiring the rectification of an otherwise stable and acceptable account of propositional knowledge. It is proof that the core of the approach needs to be abandoned.

ACKNOWLEDGEMENTS

I would like to thank Joseph Halpern, Kia Nobre, Paul Oldfield, Gianluca Paronitti, Matteo Turilli, Jeff Sanders and the participants at the graduate seminar of the University of Bari for their useful comments and helpful discussions about previous versions of this work. I owe to an anonymous

colleague and to the two anonymous referees of the journal some of the clarifying comments discussed in the third section. They are responsible only for the improvements and not for any remaining mistakes.

APPENDIX

A full formalization of the results concerning the "coordinated attack" problem is provided in Halpern and Moses (1990) and in Fagin et al. (1995). To my knowledge, the best introduction to the problem and to other relevant results in epistemic logic is still Halpern (1995). In this appendix the relevant theorems are briefly recalled with a short commentary that connects them to the discussion in the main text.

THEOREM 1 (Halpern and Moses 1990, Theorem 3.1). In any run of a system that displays unbounded message delays, it can never be common knowledge that a message has been delivered.

Commentary: The *unbounded message delays* clause indicates that the communication protocol may be assumed to be synchronic and highly reliable, as long as it is not supposed to be unrealistically fault-free. The theorem says that no matter how many messages of any kind are successfully delivered, coordination between T and J of the kind required to launch the attack is unattainable.

COROLLARY 1 (Halpern and Moses 1990, Corollary 3.2). In any run of a system that displays unbounded message delays, it can never be common knowledge C among the generals that they are attacking; i.e., if G consists of two generals, then C_G(attack) never holds.

Commentary: The corollary says that, if the requirement is to attack only if victory is guaranteed, then the generals will never attack. In terms of the analysis of the Gettier Problem, this means that, given the system set up by the tripartite account, the requirement to provide necessary and sufficient conditions for the definition of propositional and fallibilist knowledge that p cannot be fulfilled.

THEOREM 2 (Halpern and Moses 1990, Theorem 3.3). In any system for coordinated attack, when the generals attack, it is common knowledge among the generals that they are attacking. Thus, if I is an interpreted system for coordinated attack, and G consists of the two generals, then at every point (r, m) of I, we have

$$(I, r, m) \models \text{attack} \rightarrow C_G(\text{attack}).$$

Commentary: This is the theorem that proves the necessity of common knowledge for coordinated attack. In epistemological terms, the resources T and J can deliver knowledge that p only if they are successfully coordinated.

COROLLARY 2 (Halpern and Moses 1990, Corollary 3.4). In any run of a system that displays unbounded message delays, the generals never attack.

Commentary: This corollary reminds us that, given the conditions set up by the tripartite account of knowledge, the resources T and J fail to deliver knowledge that p without running the risk of being Gettierized, that is, either they do not deliver knowledge, or they deliver it in a necessarily Gettierizable way.

REFERENCES

Alston, W.: 1986, 'Epistemic Circularity', *Philosophy and Phenomenological Research* **47**, 1–30.

Apel, K. O.: 1975, 'The Problem of Philosophical Fundamental-Grounding in Light of a Transcendental Pragmatic of Language', *Man & World* **8**, 239–275.

Chisholm, R. M.: 1989, *Theory of Knowledge*, 3rd ed. Prentice-Hall, Englewood Cliffs.

Crisp, T. M.: 2000, 'Gettier and Plantinga's Revised Account of Warrant', *Analysis* **60**(265), 42–50.

Dancy, J.: 1985, *An Introduction to Contemporary Epistemology*, Blackwell, Oxford.

Dancy, J. and E. Sosa (eds): 1992, *A Companion to Epistemology*, Blackwell Reference, Oxford.

Everitt, N. and A. Fisher: 1995, *Modern Epistemology: A New Introduction*, McGraw-Hill, New York, London.

Fagin, R., J. Y. Halpern, Y. Moses, and M. Y. Vardi: 1995, *Reasoning About Knowledge*, MIT Press, Cambridge, MA, London.

Floridi, L.: 1996, *Scepticism and the Foundation of Epistemology: A Study in the Metalogical Fallacies*, Brill, Leiden.

Gettier, E.: 1963, 'Is Justified True Belief Knowledge?', *Analysis* **23**, 121–123.

Goldman, A.: 1967, 'A Causal Theory of Knowing', *Journal of Philosophy* **64**(12), 355-372.

Gray, J. N.: 1978, 'Notes on Database Operating Systems', in R. Bayer, R. Graham, and G. Seegmuller (eds), *Operating Systems: An Advanced Course*, Springer-Verlag, Berlin, pp. 393–481.

Greco, J. and E. Sosa (eds): 1999, *The Blackwell Guide to Epistemology*, Blackwell Publishers, Oxford.

Griffiths, A. P. (ed.): 1967, *Knowledge and Belief*, Oxford University Press, London.

Halpern, J. and B. Tuttle: 1993, 'Knowledge, Probability and Adversaries', *Journal of the Association for Computing Machinery* **40**, 917–962.

Halpern, J. Y.: 1995, 'Reasoning About Knowledge: A Survey', in D. Gabbay, C. J. Hogger, and J. A. Robinson (eds), *Handbook of Logic in Artificial Intelligence and Logic Programming*, Oxford University Press, Oxford, pp. 1–34. A version of the paper similar to the published version is available in postscript and pdf from http://www.cs.cornell.edu/home/halpern/abstract.html.

Halpern, J. Y. and Y. Moses: 1990, 'Knowledge and Common Knowledge in a Distributed Environment', *Journal of the Association for Computing Machinery* **37**(3), 549–587.

Hetherington, S. C.: 1996, *Knowledge Puzzles: An Introduction to Epistemology*, Westview Press, Boulder, CO, Oxford.

Kirkham, R. L. 1984, 'Does the Gettier Problem Rest on a Mistake?', *Mind* **93**, 501–513.

Lenzen, W.: 1978, *Recent Work in Epistemic Logic*, North-Holland, Amsterdam.

Morris, S. and H. S. Shin: 1997, 'Approximate Common Knowledge and Co-ordination: Recent Lessons from Game Theory', *Journal of Logic, Language and Information* **6**, 171–190.

Nozick, R: 1981, *Philosophical Explanation*, Harvard University Press, Cambridge, MA.

Pappas, G. S. (ed.): 1979, *Justification and Knowledge: New Studies in Epistemology*, Reidel, Dordrecht, Holland; Boston.

Pappas, G. S. and M. Swain (ed.): 1978, *Essays on Knowledge and Justification*, Cornell University Press, Ithaca, London.

Pease, M., R. Shostak, and L. Lamport: 1980, 'Reaching Agreement in the Presence of Faults', *Journal of the Association for Computing Machinery* **27**(2), 228–234.

Plantinga, A.: 1993a, *Warrant: The Current Debate*, Oxford University Press, New York, Oxford.

Plantinga, A.: 1993b, *Warrant and Proper Function*, Oxford University Press, Oxford.

Pust, J.: 2000, 'Warrant and Analysis', *Analysis* **60**(265), 51–57.

Roth, M. D. and L. Galis (ed.): 1970, *Knowing: Essays in the Analysis of Knowledge*, Random House, New York.

Schreiber, D. S. G.: 1987, 'The Illegitimacy of Gettier Examples', *Metaphilosophy* **18**, 49–54.

Shope, R. K.: 1983, *The Analysis of Knowing: A Decade of Research*, Princeton University Press, Princeton.

Steup, M.: 1996, *An Introduction to Contemporary Epistemology*, Prenctice Hall, Upper Saddle River.

Steup, M.: 2001, 'The Analysis of Knowledge', *Stanford Encyclopedia of Knowledge*, http://plato.stanford.edu/entries/knowledge-analysis/.

Wooldridge, M. J.: 2002, *An Introduction to Multiagent Systems*, J. Wiley, Chichester.

Zagzebski, L.: 1994, 'The Inescapability of Gettier Problems', *Philosophical Quarterly* **44**(174), 65–73.

Dipartimento di Scienze Filosofiche
Università degli Studi di Bari
70121 Bari
Italy
E-mail: luciano.floridi@philosophy.ox.ac.uk