Taylor & Francis
Taylor & Francis Group

# On a razor's edge: evaluating arguments from expert opinion

Douglas Walton*

*CRRAR, University of Windsor, 401 Sunset Ave., Windsor, ON, Canada N9B 3P4*

This paper takes an argumentation approach to find the place of trust in a method for evaluating arguments from expert opinion. The method uses the argumentation scheme for argument from expert opinion along with its matching set of critical questions. It shows how to use this scheme in three formal computational argumentation models that provide tools to analyse and evaluate instances of argument from expert opinion. The paper uses several examples to illustrate the use of these tools. A conclusion of the paper is that from an argumentation point of view, it is better to critically question arguments from expert opinion than to accept or reject them based solely on trust.

**Keywords:** nonmonotonic accounts of argument; automated argumentation reasoning systems; argument & automated reasoning; argument representation

## 1. Introduction

This paper offers solutions to key problems of how to apply argumentation tools to analyse and evaluate arguments from expert opinion. It is shown (1) how to structure the argumentation scheme for argument from expert opinion (2) how to apply it to real cases of argument from expert opinion, (3) how to set up the matching set of critical questions that go along with the scheme, (4) how to find the place of trust in configuring the schemes and critical questions, (5) how to use these tools to construct an argument diagram to represent pro and con arguments in a given argument from expert opinion, (6) how to evaluate the arguments and critical questions shown in the diagram, and (7) how to use this structure within a formal computational model to determine whether what the expert says is acceptable or not.

One of the critical questions raises the issue of trust, and a central problem is to determine how the other critical questions fit with this one. The paper studies how trust is related to argument from expert opinion in formal computational argumentation models.

Section 2 poses the problem to be solved by framing it within the growing and now very large literature on trusting experts. It is shown that there can be differing criteria for extending trust to experts depending on what you are trying to do. This section explains how the argumentation approach is distinctive in that its framework for analysing and evaluating arguments rests on an approach of critically questioning experts rather than trusting them. Argument from expert opinion has long been included in logic textbooks under the heading of the fallacy of appeal to authority, and even though this traditional approach of so strongly mistrusting authority has changed, generally the argumentation approach stresses the value of critical questioning. For example, if you are receiving advice from your doctor concerning a treatment that has been recommended, it is advocated that you should try not only to absorb the information she is communicating to you, but also try your best to ask intelligent questions about it, and in particular to critically question aspects you have doubts or reservations about. This policy is held to be consistent with rational

---

*Email: dwalton@uwindsor.ca

principles of informed and intelligent autonomous decision-making and critical evidence-based argumentation.

Sections 3 and 4 explain certain aspects of defeasible reasoning that are important for understanding arguments from expert opinion, and outline three formal computational systems for modelling arguments from expert opinion, ASPIC+, DefLog and the Carneades Argumentation System (CAS). Section 5 reviews and explains the argumentation scheme for argument from expert opinion and its matching set of critical questions.

Section 6 explains a basic difficulty in using critical questions as tools for argument evaluation within formal and computational systems for defeasible argumentation. Section 7 explains how the CAS overcomes this difficulty by distinguishing between two kinds of premises of the scheme called assumptions and exceptions. Based on this distinction, Section 7 shows how the scheme for argument for expert opinion, including representing the critical questions as assumptions and exceptions, is modelled in the CAS. Section 8 uses a simple example to show how Carneades has the capability for evaluating arguments from expert opinion by taking critical questions and counterarguments into account. Following the advice that real examples should be used to test any theory, Section 9 models some arguments from expert opinion in a real case discussing whether a valuable Greek statue (kouros) that appears to be from antiquity is genuine or not. Section 10 summarises the findings and draws some conclusions.

## 2. Arguments from expert opinion

Argument from expert opinion has always been a form of reasoning that is on a razor's edge. We often have to rely on it, but we also need to recognise that we can go badly wrong with it. Argument from expert opinion was traditionally taken to be a fallacious form of argument coming under the heading of appeal to authority in the logic textbooks. But research in studies on argumentation tended to show by an examination of many examples of argument from expert opinion that many of these arguments were not fallacious, and in fact they were reasonable but defeasible forms of argumentation. At one time, in a more positivistic era, it was accepted that argument from expert opinion is a subjective source of evidence or testimony that should always yield to empirical knowledge of the facts. However, it seems to be more generally acknowledged now that we do have to rely on experts, such as scientists, physicians, financial experts and so forth, and that such sources of evidence should be given at least some weight in deciding what to do or what to believe in practical matters. Thus the problem was posed of how to differentiate between the reasonable cases of argument from expert opinion and the fallacious instances of this type of argument. This problem has turned out to be a wicked one, and it has become more evident in recent years that solving it is a significant task with many practical applications.

The way towards a solution proposed in Walton (1997) was to formulate an argumentation scheme for argument from expert opinion along with a set of critical questions matching this scheme. The scheme and critical questions can be used in a number of ways to evaluate a given instance of argument from expert opinion. The scheme requires this type of argument to have certain premises articulated as special components of the scheme, and if the argument in question fails to have one or more of these premises, or otherwise does not fit the requirements of the scheme, then the argument can be analysed and even criticized on this basis. The missing premise might be merely an unstated premise or an incomplete argument of the kind traditionally called an enthymeme. Or in another more problematic kind of case, the expert source might not be named. This failure is in fact one of the most common problems with appeals to expert opinion found in everyday conversational arguments, such as political arguments and arguments put forward in newsmagazines. One premise of the given argument is that an expert says such and such,

or experts say such and such, without the expert being named, or the group of experts being identified with any institution or source that can be tracked down. In other instances, the error is more serious, as suggested by the fallacy literature (Hamblin, 1970). In some instances fallacies are simply errors, for example the error to name a source properly. However in other instances fallacies are much more serious, and can be identified with strategic errors that exploit common heuristics sometimes used to deceive an opponent in argumentation (Walton, 2010). Fallacies have been identified by Van Eemeren & Grootendorst (1992) as violations of the rules of a type of communicative argumentation structure called a critical discussion. Such implicit Gricean conversational rules require that participants in an argumentative exchange should cooperate by making their contributions to the exchange in a way that helps to move the argumentation forward (Grice, 1975). There is an element of trust presupposed by all parties in such a cooperative exchange.

Some might say that the problem is when to trust experts, and suggest that arguments from expert opinion become fallacious when the expert violates our trust in someone. Trust has become very important in distributed computational systems: a distributed system is a decentralised network consisting of a collection of autonomous computers that communicate with each other by exchanging messages (Li & Sighal, 2007, p. 45). Trust management systems aid automated multi-agent communications systems that put security policies in place that allow actions or messages from an unknown agent if that agent can furnish accredited credentials.

Haynes et al. (2012) reported data from interviews in which Australian civil servants, ministers and ministerial advisors tried to find and evaluate researchers with whom they wished to consult. The search was described as one of finding trustworthy experts, and for this reason it might easily be thought that the attributes found to be best for this purpose would have implications for studying the argument from expert opinion of the kind often featured in logic textbooks. In the study by Haynes et al. (2012, p. 1) evaluating three factors was seen as key to reaching a determination of trustworthiness: (1) *competence* (described as "an exemplary academic reputation complemented by pragmatism, understanding of government processes, and effective collaboration and communication skills"; (2) *integrity* (described as "independence, authenticity, and faithful reporting of research"); and (3) *benevolence* (described as "commitment to the policy reform agenda"). The aim of this study was to facilitate political policy discussions by locating suitable trustworthy experts who could be brought in to provide the factual data needed to make such discussions intelligent and informed.

Hence there are many areas where it is important to use criteria for trustworthiness of an expert, but this paper takes a different approach of working towards developing and improving arguments based on an appeal to expert opinion. This paper takes an argumentation approach, motivated by the need to teach students informal logic skills by helping them to be able to apply argumentation tools for the identification, analysis and evaluation of arguments. Argument from expert opinion has long been covered in logic textbooks, mainly in the section on informal fallacies in such a book, where the student is tutored on how to take a critical approach. A critical approach requires asking the right questions when the arguer is a layperson who is confronted by an argument that relies on expert opinion.

Goldman (2001, p. 85) frames the problem to be discussed as one of evaluating the testimony of experts to "decide which of two or more rival experts is most credible". Goldman defines expertise in terms of authority, and defines the notion of authority as follows: "Person *A* is an authority in subject *S* if and only if *A* knows more propositions in *S*, or has a higher degree of knowledge of propositions in *S*, than almost anybody else" (Goldman, 1999, p. 268). This does not seem to be a very helpful definition of the notion of an expert, because it implies the consequence that if you have two experts, and one knows more than the other, then the second cannot be an expert. The good thing about the definition is that it defines expertise in a subject, in relation to the

knowledge that the person who is claimed to be an expert has in that subject. But a dubious aspect of it from an argumentation point of view is that it differentiates between experts and nonexperts on the basis of the number of propositions known by the person who is claimed to be an expert, resting on a numerical comparison. Another questionable aspect of the definition is that it appears to include being an authority under the more general category of being an expert. This is backwards from an argumentation point of view, where it is important to clearly distinguish between the more general notion of an authority and the subsumed notion of an expert (Walton, 1997).

In a compelling and influential book, Freedman (2010) argued that experts, including scientific experts, are generally wrong with respect to claims that they make. Freedman supported his conclusions with many well-documented instances where expert opinions were wrong. He concluded that approximately two-thirds of the research findings published in leading medical journals turned out to be wrong (Freedman, 2010, p. 6). In an appendix to the book (231–238), he presented a number of interesting examples of wrong expert opinions. These include arguments from expert opinion in fields as widely ranging as physics, economics, sports and child-raising. Freedman went so far as to write (6) that he could fill his entire book, and several more, with examples of pronouncements of experts that turned out to be incorrect. His general conclusion is worth quoting: "The fact is, expert wisdom usually turns out to be at best highly contested and ephemeral, and at worst flat-out wrong" (Freedman, 2010). The implications of Freedman's reports of such findings are highly significant for argumentation studies on the argument from expert opinion as a defeasible form of reasoning.

Mizrahi (2013) argues that arguments from expert opinion are inherently weak, in the sense that even if the premises are true, they provide either weak support or no support at all for the conclusion. He takes the view that the argumentation scheme for argument from expert opinion is best represented by its simplest form, "Expert *E* says that *A*, therefore *A*". To support his claim he cites a body of empirical evidence showing that experts are only slightly more accurate than chance (2013, p. 58), and are therefore wrong more often than one might expect (p. 63). He even goes so far as to claim (p. 58) that "we do argue fallaciously when we argue that [proposition] *p* on the ground that an expert says that *p*". He refuses to countenance the possibility that other premises of the form of the argument from expert opinion need to be taken into account.

From an argumentation point of view, this approach does not provide a solution to the problem, because from that point of view what is most vital is to critically question the argument from expert opinion that one has been confronted with, rather than deciding to go along with the argument or not on the basis of whether to trust the expert or not. One could say that from an argumentation point of view of the kind associated with the study of fallacies, it is part of one's starting point to generally be somewhat critical about arguments from expert opinion, in order to ask the right questions needed to properly evaluate the argument as strong or weak. Nevertheless, as will be shown below, trust is partly involved in this critical endeavour, and Freedman's findings about expert opinions being shown to be wrong in so many instances are important.

One purpose of this paper is to teach students informal logic skills using argumentation tools. Another purpose is to show that the work is of value to researchers in artificial intelligence who are interested in building systems that can perform automated reasoning using computational argumentation. Argumentation is helpful to computing because it provides concepts and methods used to build software tools for designing, implementing and analysing sophisticated forms of reasoning and interaction among rational agents. Recent successes include argumentation-based models of evidential relations and legal processes of examination and evaluation of evidence. Argument mapping has proved to be useful for designing better products and services and for improving the quality of communication in social media by making deliberation dialogues more efficient. Throughout many of its areas, artificial intelligence has seen a prolific growth in uses of argumentation,

including agent system negotiation protocols, argumentation-based models of evidential reasoning in law, design and implementation of protocols for multi-agent action and communication, the application of theories of argument and rhetoric in natural language processing, and the use of argument-based structures for autonomous reasoning in artificial intelligence.

The way forward advocated in the present paper is to use formal computational argumentation systems that (1) can apply argumentation schemes (2) that are to be used along with argument diagramming tools and (3) that distinguish between Pollock-style rebutters and undercutters (Pollock, 1995). On this approach, the problem is reframed as one of how laypersons should evaluate the testimony of experts based on an analysis or examination of the argument from expert opinion and probe into it by distinguishing different factors that call for critical questions to be asked. On this approach, a distinction is drawn between the expertise critical question and the reliability critical question. Credibility could ambiguously refer to either one of these factors or both.

From an argumentation point of view, dealing with the traditional informal fallacy of the *argumentum ad verecundiam* (literally, argument from modesty) requires carefully examining lots of examples of this type of strategic manoeuvering for the purpose of deception. This project was carried forward in Walton (1997) and brought out common elements in some of the most serious instances of the fallacy. In such cases it was found that it is hard for a layperson in a field of knowledge to critically question an expert or the opinion of an expert brought forward by a third party, because we normally tend to defer to experts. To some extent this is reasonable. For example in law, expert witnesses are given special privileges to express opinions and draw inferences in ways stronger than a nonexpert witness is allowed to. In other instances, however, because an expert is treated as an authority, and since as we know from psychological studies there is a halo effect surrounding the pronouncements of an authority, we tend to give too much credit to the expert opinion and are reluctant to critically question it. It may be hard, or even appear inappropriate, for a questioner to raise doubts about an opinion that is privy to experts in the field of knowledge if one is not oneself an expert in this field. Thus the clever sophist can easily appeal to argument from expert opinion in a forceful way that takes advantage of our deference to experts by making anyone who questions the expert appear to be presumptuous and to be on dubious grounds. In this paper, however, the view is defended that argument from expert opinion should be regarded as an essentially defeasible form of argument that should always be open to critical questioning.

## 3. Formal computational systems for modelling arguments from expert opinion

There are formal argumentation systems that have been computationally implemented that can be used to model arguments from expert opinion and to evaluate them when they are nested within related arguments in a larger body of evidence (Prakken, 2011). The most important properties of these systems for our purposes here are that they represent argument from expert opinion as a form of argument that is inherently defeasible, and they formally model the conditions under which such an argument can be either supported or defeated by the related arguments in a case.

One such system is ASPIC+ (Prakken, 2010). It is built on a logical language containing a set of strict inference rules as well as a set of defeasible inference rules. Although it would normally model argument from expert opinion as a defeasible form of argument, it also has the capability of modelling it as a deductively valid form of argument, should this be required in some instances, for example when a knowledge base is assumed to be closed. ASPIC+ is based on a Dung-style abstract argumentation framework that determines the success of argument attacks and that compares conflicts in arguments at the points where they conflict (Dung, 1995). ASPIC+ is built around the notion of defeasibility attributed to (Pollock, 1995). Pollock drew a distinction between two kinds of refutations he called rebutting defeaters, or rebutters, and undercutting defeaters, or undercutters (Pollock, 1995, p. 40). A rebutter gives a reason for denying a claim. We could say, to

use different language, that it refutes the claim by showing it is false. An undercutter casts doubt on whether the claim holds by attacking the inferential link between the claim and the reason supporting it.

Pollock (1995, p. 41) used a famous example to illustrate his distinction. In this example, if I see an object that looks red to me, that is a reason for my thinking it is red. But suppose I know that the object is illuminated by a red light. This new information works as an undercutter in Pollock's sense, because red objects look red in red light too. It does not defeat (rebut, in Pollock's sense of the term) the claim that the object is red, because it might be red for all I know. In his terminology, it undercuts the argument that it is red. We could say that an undercutter acts like a critical question that casts an argument into doubt rather than strongly refuting it.

The logical system DefLog (Verheij, 2003, 2005) has been computationally implemented and has an accompanying argument diagramming tool called ArguMed that can be used to analyse and evaluate defeasible argumentation. ArguMed is available free on the Internet: (http://www.ai.rug.nl/~verheij/aaa/argumed3.htm) and it can be used to model arguments from expert opinion. The logical system is built around two connectives called primitive implication, represented by the symbol ~>, and dialectical negation, represented by X.

There is only one rule of inference supported by primitive implication. It is the rule often called *modus non excipiens* by Verheij (2003) but more widely called defeasible modus ponens (DMP).

> $A \sim> B$.
> A.
> Therefore $B$.

The propositions in DefLog are assumptions that can either be positively evaluated as justified or negatively evaluated as defeated. The system may be contrasted with that of deductive logic in which propositions are said to be true or false, and there is no way to challenge the validity of an inference. The only way to challenge a deductively valid argument is to attack one of its premises or pose a counterargument showing that the conclusion is false. No undercutting, in Pollock's sense, is allowed.

To see how primitive implication works, consider Pollock's red light example. Verheij (2003, p. 324) represents this example in DefLog by taking the conditional "If an object looks red, it is red" as a primitive implication. The reasoning in the first stage of Pollock's example where the observer sees the object is red and therefore concludes that it is red is modelled in DefLog as the following DMP argument.

> looks_red.
> looks_red ~> is_red.
> Therefore is_red.

The reasoning in the second stage of Pollock's example is modelled as follows.

> looks_red.
> illuminated by a red light.
> looks_red ~> X(looks_red ~> is_red).
> Therefore X(is_red).

The third premise is a nested defeasible primitive implication containing a defeasible negation. It states that if the object looks red under the circumstances of its being illuminated by red light it cannot be inferred that it is red simply because it looks red. The conclusion is that it cannot be
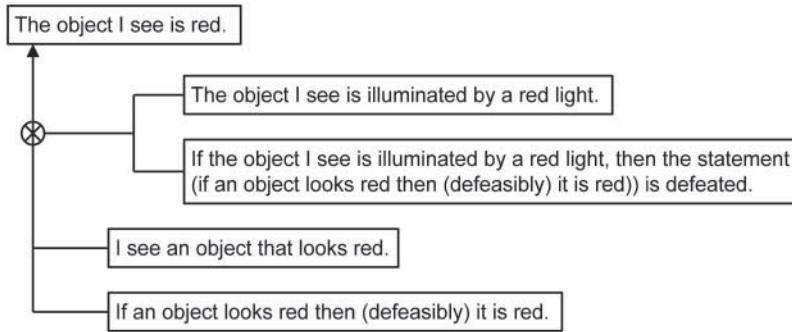
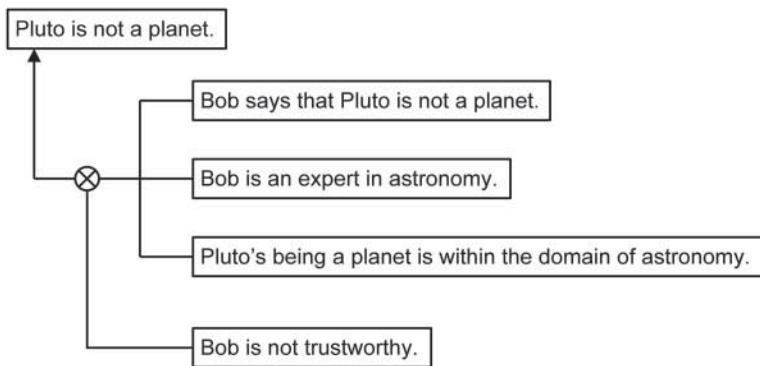Figure 1. Pollock's red light example modelled in DefLog.



Figure 2. Argument from expert opinion as a defeasible argument in DefLog.

concluded from the three premises of the argument that the object is red. Of course it might be red, but that is not a justifiable reason for accepting the conclusion that it is red.

How the red light argument above is visually represented in Verheij's argument diagramming system ArguMed can be shown using Figure 1.

The first stage of the reasoning in Pollock's example is shown by the argument at the bottom of Figure 1. It has two premises, and these premises go together in a linked argument format to support the conclusion that the object I see is red. Above these two premises we see the undercutting argument, which itself has two premises forming a second linked argument. This second linked argument undercuts the first one, as shown by the line from the second argument to the X appearing on the line leading from the first argument to the conclusion. So the top argument is shown as undercutting the bottom argument, in a way that visually displays the two stages of the reasoning in Pollock's example.

Next it is shown how an argument from expert opinion is modelled as a defeasible argument in DefLog by displaying a simple example in Figure 2. The argumentation scheme on which the argument represented in Figure 2 is based will be presented in Section 5. Even though this form has not yet been stated explicitly the reader can easily see at this point that in the example shown in Figure 2 a particular form of argument from expert opinion is being used.

In this example the argument from expert opinion is shown with its three premises in the top part of Figure 2. The proposition at the bottom, the statement that Bob is not trustworthy, corresponds to one of the critical questions matching this scheme for argument from expert opinion. Let us

say that when a critic puts forward this statement, it undercuts the argument from expert opinion based on Bob's being an expert in astronomy. The reason is that if Bob is not trustworthy, a doubt is raised on whether we should accept the argument based on his testimony. More will be shown about how to model trustworthiness in another system below.

## 4. The Carneades argumentation system

Carneades is a formal and computational system (Gordon, 2010) that also has a visualisation tool that is available at http://carneades.github.com. The CAS formally models argumentation as a graph, a structure made up of nodes that represent premises or conclusions of an argument, and arrows representing arguments joining points to other points (Gordon, 2010). An argument evaluation structure is defined in CAS as a tuple ⟨state, audience, standard⟩, where a proof standard is a function mapping tuples of the form ⟨issue, state, audience⟩ to the Boolean values true and false, where an issue is a proposition to be proved or disproved in L, a state is a point the sequence of argumentation is in and an audience is the respondent to whom the argument was directed in a dialogue. The audience determines whether a premise has been accepted or not, and argumentation schemes determine where the conclusion of an argument should be accepted given the status of its premises (accepted, not accepted or rejected). A proposition in an argument evaluation structure is acceptable if and only if it meets its standard of proof when put forward at a particular state according to the evaluation placed on it by the audience (Gordon & Walton, 2009).

Four standards were formally modelled in CAS (Gordon & Walton, 2009). They range in order of strictness from the weakest shown at the top to the highest shown at the bottom.

The Scintilla of Evidence Standard:

- There is at least one applicable argument.

The preponderance of Evidence Standard:

- The scintilla of evidence standard is satisfied and
- the maximum weight assigned to an applicable pro argument is greater than the maximum weight of an applicable con argument.

The Clear and Convincing Evidence Standard:

- The preponderance of evidence standard is satisfied,
- the maximum weight of applicable pro arguments exceeds some threshold $\alpha$ and
- the difference between the maximum weight of the applicable pro arguments and the maximum weight of the applicable con arguments exceeds some threshold $\beta$.

The Beyond Reasonable Doubt Standard:

- The clear and convincing evidence standard is satisfied and
- the maximum weight of the applicable con arguments is less than some threshold $\gamma$.

The threshold $\gamma$ is not given a fixed numerical value. It is left open to be specified by the contextual application and is meant to be specified by the user.

The visualisation tool for the CAS is still under development. The argument map drawn with CAS shown in Figure 3 indicates how a typical argument diagram appears to the user in the most recent version (1.0.2). The statements making up the premises and conclusions in the argument are inserted in a menu at the left of the screen, and then they appear in an argument diagram of the kind displayed in Figure 3. The default standard of proof is preponderance of the evidence, but
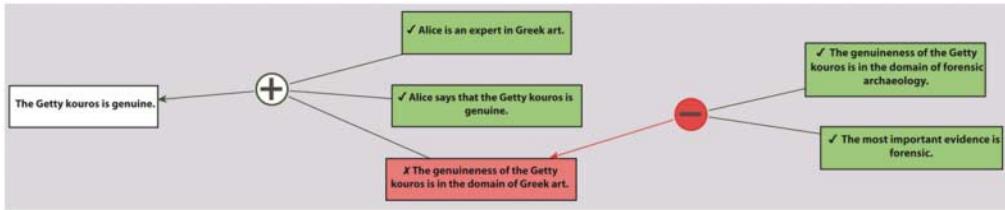
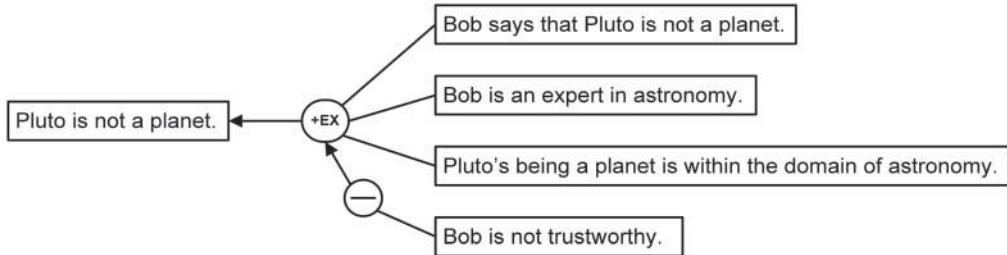Figure 3. An example argument visualised with Carneades.



Figure 4. Carneades version of the pluto example.

that can be changed in the menu. The user inputs which statements are accepted or rejected (or are undecided), and then CAS draws inferences from these premises to determine which conclusions need to be accepted or rejected (or to be declared undecided).

In the example shown in Figure 3, the ultimate conclusion, the statement that the Getty kouros is genuine, appears at the left. Supporting it is a pro argument from expert opinion with three premises. The bottom premise is attacked by a con argument. The two premises of the con argument are shown as accepted, indicated by the light grey background in both text boxes and a checkmark in front of each statement in each text box. The con argument is successful in defeating the bottom premise of the pro argument, and hence the bottom premise is shown in a darker text box with an X in front of the statement and the text box. This notation indicates that the statement in the text box is rejected. Because of the failure of one premise of the argument from expert opinion, the node with the plus sign in it is shown with a white background, indicating the argument is not applicable. Because of this the conclusion is also shown in a white text box, indicating that it is stated but not accepted (undecided). In short, the original argument is shown as refuted because of the attack on the one premise.

CAS can also use argumentation schemes to model defeasible arguments such as argument from expert opinion, argument from testimony, argument from cause and effect and so forth. If the scheme fits the argument chosen to be modelled, the scheme is judged to be applicable to the argument and the argument is taken to be "valid" (defeasibly).

The name of the argumentation scheme in Figure 4 is indicated in the node joining the three premises to the ultimate conclusion. EX stands for the argument from expert opinion and the plus sign in the node indicates that the argument from expert opinion is used as a pro argument. The statement "Bob is not trustworthy" is the only premise in a con argument, indicated by the minus sign in the node leading to the node containing the argument from expert opinion. This con argument is modelled by CAS as a Pollock-style undercutter. ASPIC+, DefLog and CAS all use undercutters and rebutters to model-defeasible argumentation, but the way that CAS does this in the case of argument from expert opinion is especially distinctive. This will be explained using an example in Section 7.

## 5.   The scheme and matching critical questions

There can be different ways of formulating the argumentation scheme for argument from expert opinion. The first formulation of the logical structure of this form of argument was given in Walton (1989, p. 193), where *A* is a proposition.

> *E* is an expert in domain *D*.
> *E* asserts that *A* is known to be true.
> *A* is within *D*.
> Therefore, *A* may plausibly be taken to be true.

Matching the original version of this scheme, six critical questions were informally presented (Walton, 1989, pp. 194–196). The first is whether the opinion put forward by the expert falls within his or her field of competence. The second is whether the source cited as an expert is really an expert, as opposed to being a source that was cited on grounds of popularity or celebrity status. The third is the question of how authoritative the expert should be taken to be. The fourth is whether there are other experts who disagree. The fifth is whether the expert's opinion is consistent with any objective evidence that may be available. The sixth is whether the pronouncement made by the expert has been correctly interpreted.

A more recent version of the scheme for argument from expert opinion was given (Walton, Reed, & Macagno, 2008, p. 310) as follows. This version of the scheme is closely comparable to the one given in (Walton, 1997, p. 210).

> Major Premise: Source *E* is an expert in subject domain *S* containing proposition *A*.
> Minor Premise: *E* asserts that proposition *A* is true (false).
> Conclusion: *A* is true (false).

The difference between this scheme and the earlier one is that the assumption that proposition *A* is within domain *D* is stated as a separate premise in the original version, whereas in the later version it is included as part of the major premise.

It has also been noted that the scheme can be formulated in a conditional version that makes it have the structure of DMP in DefLog. This conditional version can be formulated as follows (Reed & Walton, 2003, p. 201).

> Conditional Premise: If Source *E* is an expert in subject domain *S* containing proposition *A*, and *E* asserts that proposition *A* is true (false) then *A* is true (false).
> Major Premise: Source *E* is an expert in subject domain *S* containing proposition *A*.
> Minor Premise: *E* asserts that proposition *A* is true (false).
> Conclusion: *A* is true (false).

Part of Mizrahi's (2013, p. 68) argument is that the conditional premise in the expanded version of the scheme for argument from expert opinion is "implausible" because it makes the claim that the fact that an expert says that proposition *P* makes it significantly more likely that *P* is true. However, he holds this opinion because, like Goldman, he takes the traditional view that such a conditional can only be deductive in nature, like the strict material conditional of classical deductive logic, or an inductive conditional that is statistical in nature. Carneades offers a third possibility by admitting a form of *modus ponens* that is defeasible but not inductive in nature.

On this view the conditional version of the scheme has the following logical structure, where $P_1$, $P_2$ and $P_3$ are meta-variables for the premises in the scheme and *C* is a meta-variable for the conclusion.

If $P_1$, $P_2$ and $P_3$ then $C$.
$P_1$, $P_2$ and $P_3$.
Therefore $C$.

It is important to emphasise that the scheme has to be seen as defeasible in nature when taken as an instance of the form of inference DMP. The assumption behind configuring the scheme in this defeasible manner is that generally speaking it is not justifiable to take the word of an expert as infallible, even though it is also generally reasonable to presume what an expert says is right in the absence of evidence to the contrary. To accept what an expert says as having to be right absolutely, beyond all questioning or possibility of doubt makes that form of argument inherently fallacious. Exploiting the tendency of some participants in argumentation to take what an expert says as sacrosanct has been identified as a fallacious form of argument from authority in which an arguer tries to get the best of a speech partner unfairly (Walton, 1997). When you are trying to decide what to do in a given set of circumstances or what proposition to accept, you can do much better if you tentatively accept what an expert says unless you have reason not to accept it, so long as you are prepared to critically question the advice given by the expert. In Walton, 1977) it is shown that it is important not to be intimidated by expert opinions because of the powerful halo effect of an expert pronouncement. The original critical questions matching the original scheme have been reformulated in a more precise way to match the newer version of the scheme. This new way of formulating the six basic critical questions (Walton et al., 2008, p. 310) has a name for each question. This way of formulating the six basic critical questions comes from the earlier version of the scheme given in (Walton, 1997, p. 223).

> *Expertise Question*: How credible is $E$ as an expert source?
> *Field Question*: Is $E$ an expert in the field $F$ that $A$ is in?
> *Opinion Question*: What did $E$ assert that implies $A$?
> *Trustworthiness Question*: Is $E$ personally reliable as a source?
> *Consistency Question*: Is $A$ consistent with what other experts assert?
> *Backup Evidence Question*: Is $E$'s assertion based on evidence?

The important factor to stress once again is the defeasible nature of the argument. This defeasible aspect is brought out by seeing how the critical questions function as devices for evaluating an argument from expert opinion. If a respondent asks any one of the six critical questions, the original argument defaults (meaning that the conclusion can no longer be taken to be accepted given that the premises are accepted) unless the question is answered adequately. But once the question has been answered adequately, the argument tentatively stands until further critical questions are asked about it. As more critical questions matching the scheme are answered appropriately, the argument from expert opinion gets stronger and stronger, even though it may have been weak to begin with.

## 6. Critical questioning and burdens of proof

It is important to realise that the six basic critical questions are not the only ones matching the scheme for argument from expert opinion. Through research on argument from expert opinion and its corresponding fallacies, and through teaching students in courses and informal logic how to try to deal intelligently with arguments based on expert opinion, these basic six critical questions have been distilled out as the ones best suited to give guidance to students on how to critically and intelligently react to arguments from expert opinion. However, each of the basic critical questions has critical sub-questions beneath it (Walton, 1997).

Under the expertise critical question, there are three sub-questions (Walton, 1997, p. 217).

   (1)  Is *E* biased?
   (2)  Is *E* honest?
   (3)  Is *E* conscientious?


Classifying and framing such critical questions is a matter of analysing examples of fallacious arguments from expert opinion, to see where these erroneous arguments went wrong (Walton, 1997). Once the errors were classified in a systematic way, sets of critical questions designed to pinpoint and cope with them were also classified.

The possibility that critical questions can continually be asked in a dialogue that can go on continually between an arguer and a critical questioner poses problems for modelling a scheme like argument from expert opinion in a formal and computational argumentation system. Can the respondent go on and on forever asking such critical questions? Open-endedness is of course characteristic of defeasible arguments. They are nonmonotonic, meaning that new incoming information can make them fail in the future even though they hold tentatively for now. But on which side should the burden of proof lie on bringing in new evidence? Is merely asking a question enough to defeat the argument or does the question need to be backed up by evidence before it has this effect?

The defeasible nature of the argument from expert opinion can be brought out even further by seeing that evaluating an instance of the argument in any particular case rests on the setting in which there is a dialogue between the proponent of the argument and a respondent or critical questioner. The proponent originally puts forward the argument, and the respondent has the task of critically questioning it or putting forward counterarguments. Evaluating whether any particular instance of the argument from expert opinion holds or not in a given case depends on two factors. One is whether the given argument fits the structure of the scheme for argument from expert opinion. But if so, then evaluation depends on what happens in the dialogue, and in particular the balance between the moves of the proponent of the respondent. The evaluation of the argument depends on pro and contra moves made in the dialogue. It is possible to put this point in a different way by expressing it in terms of shifting of the burden of proof. Once a question has been asked and answered adequately, a burden of proof shifts back to the questioner to ask another question or accept the argument. But there is a general problem about how such a shift should be regulated and how arguments from expert opinion should be computationally modelled.

Chris Reed, when visiting at University of Arizona in 2001, asked a question. Is there any way the critical questions matching a scheme could be represented as statements of the kind represented on an argument diagram? I replied that I could not figure out a way to do it, because some critical questions defeat the argument merely by being asked, while others do not unless they are backed up by evidence. These observations led to two hypotheses (Walton & Godden, 2005) about what happens when the respondent asks a critical question: (1) When a critical question is asked, the burden shifts to the proponent to answer it and if no answer is given, the proponent's argument should fail. (2) To make the proponent's argument fail, the respondent needs to support the critical question with further argument.

Issues such as completeness of a set of critical questions are important from a computational perspective since they hold not only for the scheme for argument from expert opinion but for all schemes in general. But the question is not an easy one to resolve because context may play a role. For example an opinion expressed by an expert witness in court may have to be questioned in a different way from the case of an opinion being expressed in an informal setting or one put forward as a conclusion in a scientific paper. Wyner (2012) discusses problems of this sort that have arisen from attempts to provide formal representations of critical questions. In Parsons et al. (2012) argumentation schemes based on different forms of trust are set out. In particular there are

schemes for trust from expert opinion and trust from authority. These matters need to be explored further.

## 7. The Carneades version of the scheme and critical questions

The problem of having to choose between the two hypotheses led to the following insight that became a founding feature of the CAS: which hypothesis should be applied in any given case depends on the argumentation scheme (Walton & Gordon, 2005). In other words, the solution proposed was that a different hypothesis should be applied to each critical question of the scheme. This solution allows the burden of proof for answering critical questions to be assigned to either the proponent or the respondent, on a question-by-question basis for each argumentation scheme (Walton & Gordon, 2011).

The solution was essentially to model critical questions as premises of a scheme by expanding the premises in the scheme. The ordinary premises are the minor and major premises of the schemes. The assumptions represent critical questions to be answered by the proponent. The exceptions represent critical questions to be answered by the respondent. The two latter types of critical questions are modelled as additional premises. On this view whether a premise holds depends not only on its type but also the dialectical status of the premise during a sequence of dialogue. Shifts of burden take place as the argumentation proceeds in a case where the parties take turns making moves. They do not represent what is called burden of persuasion in law, but are more like what is called the burden of producing evidence or what is often called the evidential burden in law (Prakken & Sartor, 2009).

In the current version of the Carneades (https://github.com/carneades/carneades) there is a catalogue of schemes (http://localhost:8080/policymodellingtool/#/schemes). One of the schemes in the catalogue is that for argument from expert opinion, shown below.

> **id:** expert-opinion
> **strict:** false
> **direction:** pro
> **conclusion:** *A*
> **premises:**
>   - Source *E* is an expert in subject domain *S*.
>   - *A* is in domain *S*.
>   - *E* asserts that *A* is true.
> **assumptions**:
>   - The assertion *A* is based on evidence.
> **exceptions:**
>   - *E* is personally unreliable as a source.
>   - *A* is inconsistent with what other experts assert.

The trustworthiness critical question is represented by the statement that *E* is personally reliable as a source, classified as an exception. This means that if the respondent in the dialogue asks whether *E* is personally reliable as a source that can be trusted (the trustworthiness question), the proponent's argument from expert opinion will not be defeated unless the respondent backs up her allegation with some evidence. Otherwise the proponent is entitled to respond to the question by saying, "Of course the expert is personally reliable, and that holds unless you can provide evidence to the contrary".

In contrast, the backup evidence question is treated as an assumption. This means that if the respondent asks for backup evidence on which the experts can support her claim, the proponent
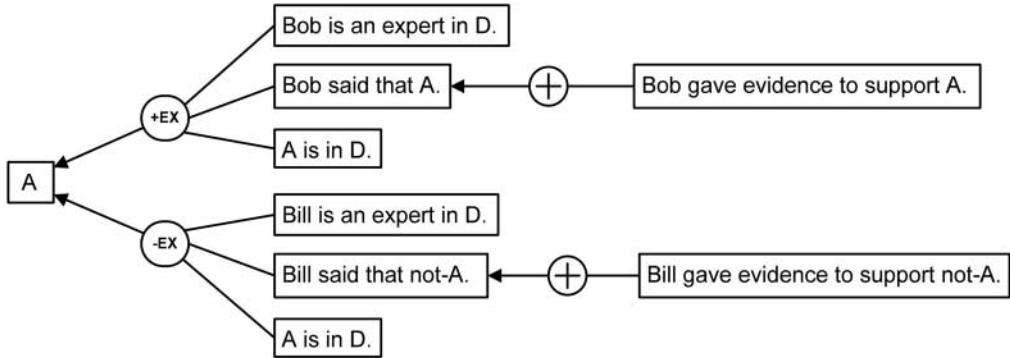
Figure 5. A case of the battle of the experts modelled in CAS.

is obliged to provide some evidence of this kind, else the argument from expert opinion fails. We reasonably expect experts to base their opinions on evidence, typically scientific evidence of some sort, and if this assumption is in doubt, an argument from expert opinion appears to be questionable. Once we have classified each critical question matching a scheme in this way, a standardised way of managing schemes in computational systems can be implemented.

## 8. An example of argument from expert opinion

To get some idea of how Carneades can model arguments from expert opinion, consider a typical case that involves some critical questions. The consistency critical question raises the issue of whether, in a case where an argument from expert opinion has been put forward, that opinion is consistent with opinions that may have been cited by other experts. The classic case, called the battle of the experts, occurs where one expert asserts proposition *A* and another expert asserts proposition not-*A*. To make the example more interesting let us consider a case of this sort that also involves the backup evidence question. In the example shown in Figure 5, Bob is an expert who has asserted proposition *A*, but Bill is an expert who has asserted proposition not-*A*. The ultimate conclusion of Bob's argument, namely proposition *A*, can be seen at the far left of Figure 5. As shown at the top, Bob's argument from expert opinion supports *A*. One of the three ordinary premises of the argument from expert opinion is supported by the proposition that Bob gave evidence to support *A*. This part of the argument diagram shows that the backup evidence critical question has been answered, possibly even before it has been raised by a critical questioner. Hence this example illustrates what is called a proleptic argument, an argument where the proponent responds to an objection even before the objection has been raised by the respondent. This strategy is a way of anticipating a criticism.

Below the pro argument, there is a con argument from expert opinion, based on what expert Bill claimed. Examining both arguments together in Figure 5, we can see that it represents a classic case of the battle of the experts, of a kind that is well known in legal trials where expert witnesses representing both sides are brought forward to offer evidence.

How should the argumentation in a case of a deadlock between two arguments from expert opinion of this sort be evaluated by CAS? A simplified version of how CAS evaluates such arguments can be presented to get some idea of how the procedure works. CAS evaluates arguments using a three-valued system, in which oppositions can be evaluated as accepted, not accepted or undecided. Here we will simplify this procedure by using a Dung-style labelling saying that a proposition is *in* if it is accepted and a proposition is *out* if it is either rejected, not accepted or
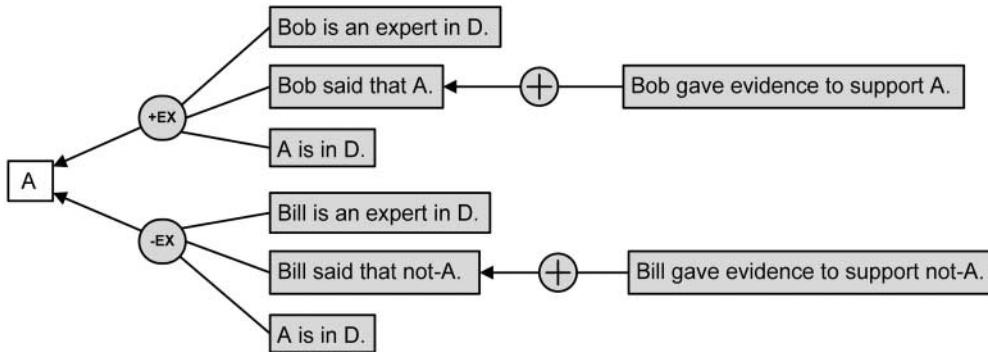
Figure 6. Extending the case by bringing in the audience.

undecided. These initial values of whether a proposition is in or out come from the audience. Once the initial values for the propositions are determined by this means, CAS calculates, using argumentation schemes and the structure of the argument, whether the conclusion is in or out.

Let us reconsider the example shown in Figure 5, and say for the purposes of illustration that the audience has accepted all the propositions shown in a darkened box as in Figure 6. Moreover, let us say that in both instances, the requirements for the argumentation scheme for expert opinion have been met, as shown by the two darkened boxes of the nodes containing notation EX. If we were only to consider Bob's pro argument shown at the top of Figure 6, on this basis the text box containing proposition *A* would be darkened showing that the argument from expert opinion proves conclusion *A*. The argument proves the conclusion because it fits the requirements for the argumentation scheme for argument from expert opinion and all three of its ordinary premises are accepted. Moreover, one of them is even backed up by the supporting evidence given by the expert. But once we take Bill's con argument into account, the two arguments are deadlocked. One cancels the other.

There are various ways such a deadlock can be dealt with by CAS. One is to utilise the notion of standards of proof. Another is to utilise the presumption that the audience has a set of values that can be ordered in priority. By combining these two means or using them separately, one argument can be shown to be stronger than another. Both alternatives can be combined.

Next, let us see how the trustworthiness critical question might enter into consideration in the case of this sort. If we look at Bill's con argument shown at the bottom of Figure 7, we see that it has been undercut by the statement that Bill is not trustworthy. This is an instance of a critical question that is an exception being modelled as an undercutter. Because the trustworthiness question is an exception, it does not defeat the argument it is directed against unless some evidence is given to support the allegation.

As shown in Figure 7, the allegation that Bill is not trustworthy is supported by an argument that has a premise stating that Bill lied in the past. Because this premise provides a reason to support the allegation that Bill is not trustworthy, the asking of the critical question defeats the argument from expert opinion in this instance. As shown in Figure 7 the argument node containing the con argument from expert opinion is shown with a white background. Bill's argument from expert opinion is knocked out of contention, and so Bob's argument shows that *A* is now in.

This example has been merely a simple one made up for purposes of illustration so the reader can get a basic idea of how CAS models arguments, how it visually represents them using argument diagrams and how it evaluates them by using the notion of an audience. To get a better idea, as always in the field of argumentation studies, it is helpful to examine a real example.
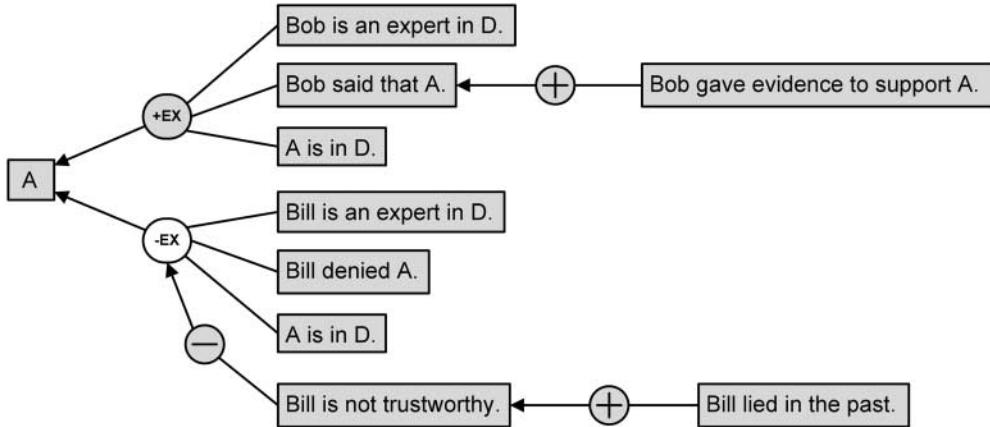
Figure 7. The trustworthiness critical question modelled as undercutter.

## 9.    The case of the Getty kouros

A kouros is an ancient Greek statue of a standing nude youth, typically standing with its left foot forward, arms at his sides, looking straight ahead. The so-called Getty kouros was bought by the J. Paul Getty Museum in Malibu California in 1985 for seven million dollars. Although originally thought to be authentic, experts have raised many doubts, and the label on the statue in the museum reads "Greek, about 530 BC, or modern forgery". Evidence concerning the provenance of the statue is weak. It was bought by the museum from a collector in Geneva who claimed he had bought it in 1930 from a Greek dealer. But there were no archaeological data tracing the statue to Greece. The documentary history of the statue appeared to be a hoax because a letter supposedly from the Swiss collector dated 1952 had a postcode on it that did not exist until 1972 (True, 1987). Figure 8 displays the structure of the two arguments from expert opinion and the argument from the provenance evidence.

As also shown in Figure 8, there was some evidence supporting the genuineness of the statue. It was made from a kind of marble found in Thrace. Norman Herz, a professor of geology at the University of Georgia determined with 90% probability that the source of the stone the statue was carved from was the island of Thasos. Stanley Margolis, a geology professor at the University of California at Davis, showed that the dolomite surface of the sculpture had undergone a process in which the magnesium content had leached out. He concluded that this process could only have occurred over the course of many centuries (Margolis, 1989). He stated that for these reasons the statue could not have been duplicated by a forger (Herz & Waelkens, 1988, p. 311).

CAS can be used to model the structure of these arguments using the standards of proof, the notion of audience as a basis for determining which premises of an argument are accepted, rejected or undecided and the other tools explained in Section 4. We begin by seeing how one argument from expert opinion attacks another. Whether the ultimate conclusion should be accepted or not depends in CAS on the standard of proof that is to be applied (Gordon, 2010). If the preponderance of the evidence standard is applied, the pro arguments for the genuineness of the kouros could win. If a higher standard is applied, such as clear and convincing evidence, or beyond a reasonable doubt, the pro argument might fail to prove the conclusion. On this view, the outcome depends on the standard of proof for the inquiry and on how acceptable the premises are to the audience that is to decide whether to accept the premises or not. In this case the standard of proof required to establish that the kouros is genuine is high, given the skepticism that is always present in such cases on the part of the experts due to the possibility of forgery, and the cleverness of forgers

exhibited in many comparable cases. The three main bodies of evidence required to meet this standard are (1) the geological evidence concerning the source of the stone statue is made of, (2) the judgement of experts concerning how close is the match between the artistic techniques exhibited in this statue and the comparable techniques exhibited in other statutes of the same kind known to be genuine and (3) the provenance evidence.

The case can be extended by introducing some evidence provided by a third expert as shown in Figure 9. In the 1990s a marine chemist named Miriam Kastner was able to artificially induce de-dolomitsation in the laboratory. Moreover, this result was confirmed by previous findings of Margolis (Kimmelman, 1991). These results showed that it is possible that the kouros was synthetically aged by a forger. This new evidence cast doubt on the claim made by Margolis that this process could only occur over the course of many centuries, weakening the argument based on the appeal to the expert opinion of Margolis by casting doubt on one of its premises.

Modelling the example of the Getty kouros using CAS is useful for demonstrating a type of reasoning that the scheme for argument from expert opinion is intended to capture. It shows how one argument from expert opinion can be attacked by or supported by other arguments from expert opinion. However one subject that we will not deal with in exploring these examples is the role of accrual of arguments. We see that the pro argument from expert opinion based on the expertise of Herz in geology was supported by the corroborative pro argument from expert opinion based on the geological expertise of Margolis. It is implied that the first argument, while defeasible, must have had a certain degree of strength or plausibility to begin with, and then when the second argument based on the geological evidence came to be taken into consideration, the conclusion that the Getty kouros is genuine became even more plausible. But then, when the argument from expert opinion put forward by Margolis was attacked by the undermining argument based on an appeal to expert opinion from marine chemistry, the degree of acceptability of the conclusion must have gone down. These variations in the strength of the body of evidence supporting or attacking the ultimate conclusion that the Getty kouros is genuine suggest that some sort of mechanism of
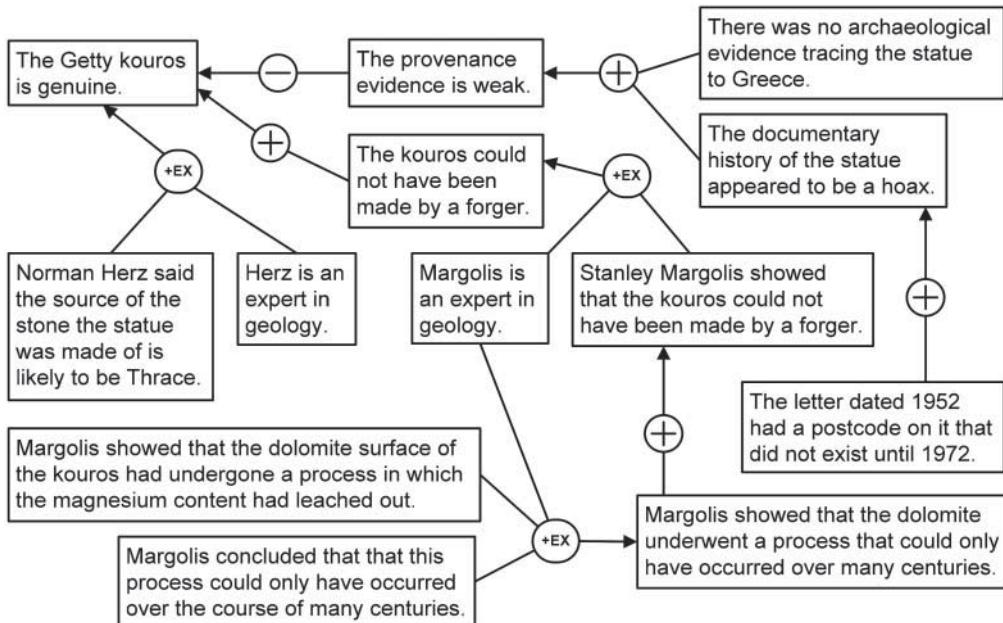


Figure 8. First two arguments from expert opinion in the Getty Kouros case.
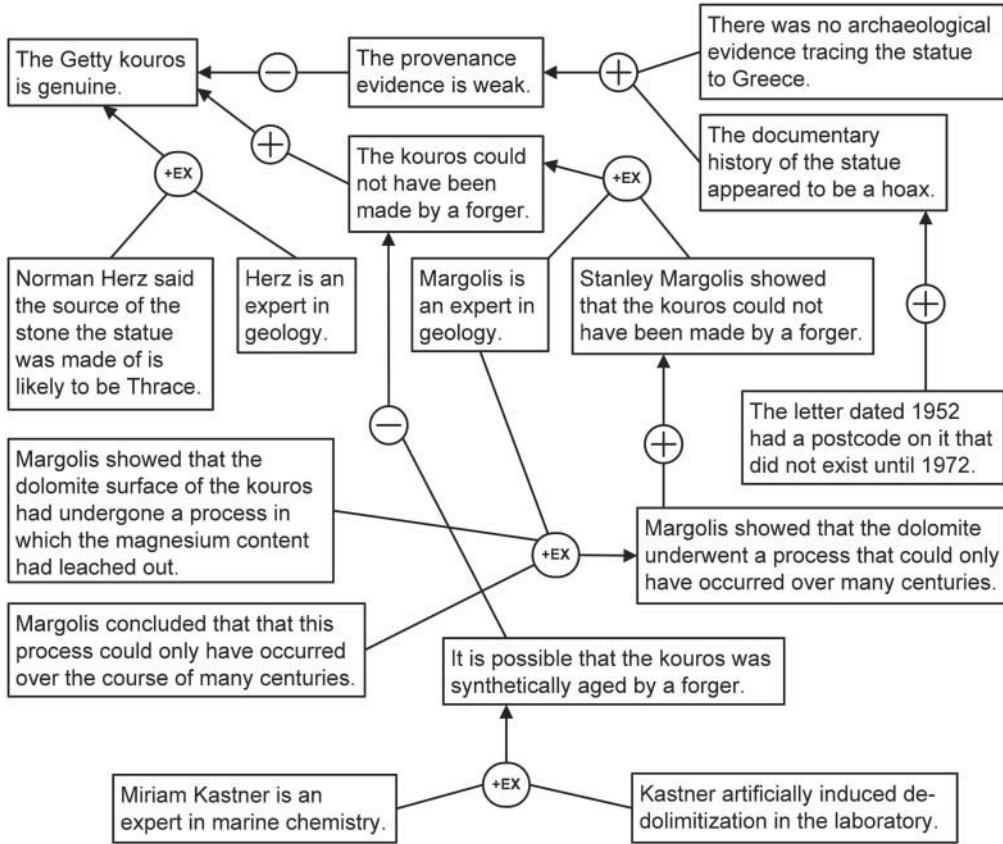
Figure 9. Third argument from expert opinion in the Getty case.

accrual of arguments is implicitly at work in how we evaluate the strength of support given by the evidence in this case. However it is known that accrual is a difficult issue to handle formally (Prakken, 2005).

The problem of how to model accrual of arguments by showing how the evidential weight of one argument can be raised or lowered by new evidence in the form of another argument that corroborates or attacks the first argument has not yet been solved for CAS. But even so, by revealing the structure of the group of related argument in the Getty kouros case as shown in Figure 9, we can still get help in taking the first necessary steps preliminary to evaluating the sequence of argumentation by some procedure of accrual of evidence.

Provenance evidence is especially important as a defeating factor even where the other two factors have been established by means of a strong body of supportive evidence. Looking at Figure 8, it can be seen that the geological evidence is fairly strong, because it is based on the concurring opinion of two independent experts. However, given the weakness of the provenance evidence, the standard of proof required to establish that the Getty kouros is genuine cannot reasonably be met. Also, the situation represented in Figure 8 represents a conflict, because the body of evidence under category 1 is conflicted with the body of evidence under category 3. What is missing is any consideration of the evidence under category 2.

If we were to take into account further evidence not modelled in Figure 8, the evaluation of the evidential situation might not turn out to be too much different, since there was a recurring

conflict of opinions on how close the match was between the Getty kouros and other statues of the same kind known to be genuine. Once we look at the further evidence shown in Figure 9, the geological evidence is weakened by the introduction of new evidence concerning Kastner's artificial aging of the stone in the laboratory by de-dolomitisation. In Figure 9, the new evidence based on the argument from expert opinion of Kastner is shown at the bottom of the argument diagram. This new argument attacks the conclusion of the Margolis expert opinion argument that the kouros could not have been made by a forger.

This new evidence brings the geological evidence even further from the possibility of meeting the standard of proof required to establish that the Getty kouros is genuine. If other experts independent of Margolis were to confirm Kastner's result, it would make the argument from geology stronger. However the fact that it was Margolis who confirmed Kastner's result is good as well, in a certain respect, because he was the original expert who claimed that the statue could not have been duplicated by a forger. Now it would seem that he would have to admit that this is possible. Although we do not have any evidence of his reaction, his confirmation of Kastner's result suggests that there is reason to think that he would have reason to retract his earlier claim that the statute could not have been duplicated by a forger.

## 10.  Conclusions

This paper concludes that (1) it is generally a mistake, from the argumentation point of view, to trust experts, (2) even though it is often necessary to rely on expert opinion evidence, but that (3) we can provisionally accept conclusions drawn from expert opinion on a presumptive basis subject to retraction. The paper showed how to evaluate an argument from expert opinion in a real case through a five-step procedure that proceeds by (1) identifying the parts of the argument, its premises and conclusion, using the argumentation scheme for argument from expert opinion (along with other schemes), (2) evaluating the argument by constructing an argument diagram that represents the mass of relevant evidence in the case, (3) taking the critical questions matching the scheme into account, (4) doing this by representing them as additional premises (assumptions and exceptions) of the scheme and (5) setting in place a system for showing the evidential relationships between the pro and con arguments preliminary to weighing the arguments both for and against the argument from expert opinion. It was shown that applying this procedure in a formal computational argumentation system is made possible by reconfiguring the critical questions by distinguishing three kinds of premises in the scheme called ordinary premises, assumptions and exceptions. Several examples were given showing how to carry out this general procedure, including the real example of the Getty kouros.

It was shown in this paper how CAS applies this procedure because it uses a defeasible version of the scheme in its argument evaluation system based on acceptability of statements, burdens of proof and proof standards (Gordon, 2010, pp. 145–156). For these reasons CAS fits the epistemology of scientific evidence (ESE) model (Walton & Zhang, 2013). This model has been applied to the analysis and evaluation of expert testimony as evidence in law. It is specifically designed for the avoidance or minimisation of error and, like CAS, it is acceptance-based rather than being based on the veristic view of Goldman. In a veristic epistemology, knowledge deductively implies truth. On this view one agent is more expert than another if its knowledge base contains more true propositions than the other. The ESE is a flexible epistemology for dealing with defeasible reasoning in a setting where knowledge is a set of commitments of the scientists in a domain of scientific knowledge that is subject to retraction as new evidence comes in. It is not a set of true beliefs, nor is it based exclusively on deductive or inductive reasoning (at least the kind represented by standard probability theory).

Mizrahi's argument goes wrong because he uses the single-premised version of the argument from expert opinion as his version of the form of the argument in general. This is unfortunate because it is precisely when this simple version of the scheme is used to represent argument from expert opinion that the other critical questions are not taken into account. The simple version has heuristic value because it shows how we often leap from the single premise that somebody is an expert to the conclusion that what this person says is true. But the simple version also illustrates precisely why leaping to a conclusion in this way without considering the questions of whether the person cited is a real expert, whether he or she is an expert in the appropriate field, and so forth. It is precisely by overlooking its critical questions or, even worse, ignoring them or shielding them off from consideration that the *ad verecundiam* fallacy occurs. As shown in this paper, argument from expert opinion in its single-premised form is of no use for argument evaluation until the additional premises are taken into account. The single-premise version of the scheme has initial explanatory value for teaching students about the simplest essentials of arguments from expert opinion, but to get anywhere we need to realise that additional premises are involved. This is shown by the model of argument from expert opinion in the CAS.

Freedman is open to the criticism of having engaged in a circular form of reasoning because he quoted many experts in his book to prove his claim that many experts are wrong. However this form of circular reasoning does not commit the fallacy of begging the question, because Freedman's conclusion is based on empirical evidence showing how often experts have been wrong, and he is able to interpret this evidence and draw conclusions from it in an informed manner. His arguments about errors in expert reasoning, and his findings about why arguments from expert opinion reasoning so often go wrong, take place at a meta-level where it is not only important but necessary for users of expert opinion evidence to become aware of the errors in their own reasoning and correct them, or at least be aware of their weaknesses. But he does not draw the conclusion that arguments from expert opinion are worthless, and ought to be entirely discounted. He went so far in an interview (Experts and Studies: Not Always Trustworthy, *Time*, 29 June 2010) to say that discarding expertise altogether "would be reckless and dangerous" and that the key to dealing with arguments from expert opinion is to learn to distinguish the better ones from the worse ones. It has been an objective of this paper to find a systematic way to use argumentation tools to help accomplish this goal.

## References

Dung, P. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence, 77*(2), 321–357.

Freedman, D.H. (2010). *Wrong: Why experts keep failing us – and how to know when not to trust them*. New York: Little Brown.

Goldman, A. (2001). Experts: which ones should you trust? *Philosophy and Phenomenological Research, LXIII*(1), 85–110.

Gordon, T.F. (2010). An overview of the Carneades argumentation support system. In C. Reed & C.W. Tindale (Eds.), *Dialectics, dialogue and argumentation* (pp. 145–156). London: College.

Gordon, T.F., & Walton, D. (2009). A formal model of legal proof standards and burdens. In F.H. van Eemeren, R. Grootendorst, J.A. Blair, & C.A. Willard (Eds.), *Proceedings of the seventh international conference of the international society for the study of argumentation* (pp. 644–655). Amsterdam: SicSat.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 43–58). New York: Academic Press.

Hamblin, C.L. (1970). *Fallacies*. London: Methuen.

Haynes, A.S., Derrick, G.E., Redman, S., Hall, W.D., Gillespie, J.A., & Chapman, S. (2012). Identifying trustworthy experts: How do policymakers find and assess public health researchers worth consulting or

collaborating with? *PLoS ONE, 7*(3): e32665. doi:10.1371/journal.pone.0032665. Retrieved July 10, 2013, from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0032665.

Herz, N., & Waelkens, M. (1988). *Classical marble: Geochemistry, technology, trade*. Dordrecht: Kluwer Academic.

Li, H., & Sighal, M. (2007). Trust management in distributed systems. *Computer, 40*(2), 45–53.

Kimmelman, M. (1991, August 4). Absolutely real? Absolutely fake? *New York Times*. Retrieved from http://www.nytimes.com/1991/08/04/arts/art-absolutely-real-absolutely-fake.html

Margolis, S.T. (1989). Authenticating ancient marble sculpture. *Scientific American, 260*(6), 104–114.

Mizrahi, M. (2013). Why arguments from expert opinion are weak arguments. *Informal Logic, 33*(1), 57–79.

Parsons, S., Atkinson, K., Haigh, K., Levitt, K., McBurney, P., Rowe, J., Singh, M.P., & Sklar, E. (2012). Argument schemes for reasoning about trust. *Frontiers in artificial intelligence and applications, Volume 245: Computational models of argument*, pp. 430–441. Amsterdam: IOS Press.

Pollock, J.L. (1995). *Cognitive carpentry*. Cambridge, MA: The MIT Press.

Prakken, H. (2005). A study of accrual of arguments, with applications to evidential reasoning. *Proceedings of the tenth international conference on artificial intelligence and law*, *Bologna* (pp. 85–94). New York: ACM Press.

Prakken, H. (2010). An abstract framework for argumentation with structured arguments. *Argument and Computation, 1*(2), 93–124.

Prakken, H. (2011). An overview of formal models of argumentation and their application in philosophy. *Studies in Logic, 4*(1), 65–86.

Prakken, H., & Sartor, G. (2009). A logical analysis of burdens of proof. In H. Kaptein, H. Prakken, & B. Verheij (Eds.), *Legal evidence and proof: Statistics, stories, logic* (pp. 223–253). Farnham: Ashgate.

Reed, C., & Walton, D. (2003). Diagramming, argumentation schemes and critical questions. In F.H. van Eemeren, J.A. Blair, C.A. Willard, & A. Snoek Henkemans (Eds.), *Anyone who has a view: Theoretical contributions to the study of argumentation* (pp. 195–211). Dordrecht: Kluwer.

True, M. (1987, January). A kouros at the Getty museum. *The Burlington Magazine, 129*(1006), 3–11.

Van Eemeren, F.H., & Grootendorst, R. (1992). *Argumentation, communication and fallacies*. Mahwah, NJ: Erlbaum.

Verheij, B. (2003). DefLog: On the logical interpretation of prima facie justified assumptions. *Journal of Logic and Computation, 13*(3), 319–346.

Verheij, B. (2005). *Virtual arguments. On the design of argument assistants for lawyers and other arguers*. The Hague: TMC Asser Press.

Walton, D. (1989). *Informal Logic*. Cambridge: Cambridge University Press.

Walton, D. (1997). *Appeal to expert opinion*. University Park: Penn State Press.

Walton, D. (2010). Why fallacies appear to be better arguments than they are. *Informal Logic, 30*(2), 159–184.

Walton, D., & Godden, D. (2005). The nature and status of critical questions in argumentation schemes. In D. Hitchcock (Ed.), *The uses of argument: Proceedings of a conference at McMaster University 18–21 May, 2005* (pp. 476–484). Hamilton, Ontario: Ontario Society for the Study of Argumentation.

Walton, D., & Gordon, T.F. (2005). Critical questions in computational models of legal argument. In P.E. Dunne & T.J.M. Bench-Capon (Eds.), *Argumentation in artificial intelligence and law, IAAIL workshop series* (pp. 103–111). Nijmegen: Wolf Legal.

Walton, D., & Gordon, T.F. (2011). Modeling critical questions as additional premises. In F. Zenker (Ed.), *Argument cultures: Proceedings of the 8th international OSSA conference*. Windsor. Retrieved from http://www.dougwalton.ca/papers%20in%20pdf/11OSSA.pdf.

Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. Cambridge: Cambridge University Press.

Walton, D., & Zhang, N. (2013). The epistemology of scientific evidence. *Artificial Intelligence and Law, 21*(2), 173–219.

Wyner, A. (2012). Questions, arguments, and natural language semantics. *Proceedings of the 12th workshop on computational models of natural argumentation (CMNA 2012)*. Montpellier, France, pp. 16–20.