

HOW TO REFUTE AN ARGUMENT USING ARTIFICIAL INTELLIGENCE

There is a family of terms in argumentation that are closely related to each other and that all refer to some way in which a given argument is attacked, rebutted, refuted, undercut, critically questioned or objected to, thereby defeating it or casting it into doubt. Proper understanding of this family of terms is fundamental to argumentation theory and to building argumentation technologies in artificial intelligence. This paper refines, clarifies and classifies them, using the Carneades Argumentation System. It begins with a simple example that illustrates two main ways of refuting an argument, and concludes with a seven-step procedure for seeking a refutation or objection.

This paper applies a computational model to examples of argument attack, challenge, critical questioning and rebuttal, in order to study techniques for refuting an argument. The aim is to improve our understanding of how to attack and refute an argument by clarifying a group of related terms including ‘attack’, ‘rebuttal’, ‘refutation’, ‘challenge’, ‘defeater’, ‘undercutting defeater’, ‘rebutting defeater’, ‘exception’ and ‘objection’ that are commonly used in the literature on argumentation and artificial intelligence. One special kind of objection that is studied is that of irrelevance. As shown in the paper, these terms are, at their present state of usage, not precise or consistent enough for us to helpfully differentiate their meanings in framing useful advice on how to attack and refute arguments. To help remedy this situation, a classification system comprising all these key terms is built and defended.

The term ‘rebuttal’ is often associated with the work of Toulmin (1958), while the terms ‘undercutting defeater’ and ‘rebutting defeater’ are associated with the work of Pollock (1995), and are commonly used in the AI literature. For this reason, section 1 and section 7 of this paper are given over to terminological discussions of these key terms. In section 2, however, we get down to the main job by presenting and working with an example meant to illustrate two fundamentally different ways of refuting an argument. One of the attack procedures is called an internal refutation and the other is called an external refutation. It is this distinction, and the example of its use, that provide the departure point for the rest of the paper. In section 3 argumentation schemes with matching sets of critical questions are introduced, using the example of the scheme for argument from expert opinion. In section 4 a computational system from artificial intelligence called Carneades is introduced, and in section 5 and 6 it is shown how Carneades models rebuttal and refutation. Carneades also has a working graphical user interface that is used to visualize arguments and refutations, and this tool is used to analyze the arguments, rebuttals and refutations that are studied. A summary of the seven-step practical procedure for attacking and refuting an argument is given at the end.

1. Questions about Attack, Rebuttal, Objection and Refutation

One finds it to be a widely held commonplace in writings on logic and artificial intelligence that there are three ways to attack an argument (Prakken, 2010, 169). One is to argue that a premise is false or insufficiently supported. Let’s call this premise attack. Another is to argue that the conclusion doesn’t follow from the set of premises that were presented as supporting it. This could be called an undercutting attack, as we will see below. The third is to argue that the conclusion is shown to be false by bringing forward a counter-argument opposed to the original argument. What the attacker needs to do in such a case is to put forward a second argument that is stronger than the original argument and that provides evidence for rejecting the conclusion of the original argument. Such an attack is sufficient to defeat the original argument, unless its proponent can give further reasons to support it.

The undercutting type of attack does not apply to deductively valid arguments. If an argument fits the form of a deductively valid argument, it is impossible for the premises to be true and the

conclusion false. Deductive reasoning is monotonic, meaning that a deductive argument always remains valid even if new premises are added. However there is a method of attack on defeasible arguments that is highly familiar in the recent research on nonmonotonic logics for defeasible reasoning. It is to argue that there is an exception to the rule, and that the given case falls under the category of this type of exception. This way of attacking an argument is very familiar in recent studies of defeasible reasoning, like the classical Tweety inference: birds fly; Tweety is a bird; therefore Tweety flies. This inference is based on the defeasible generalization that birds normally fly, or it could also be analyzed as being based a conditional rule to the effect that if something is a bird it flies. Such a conditional is open to exceptions, meaning that it may default in some cases. The argument can be attacked by pointing out the exception to the rule.

To attack an argument in the third way, it may be enough to simply question whether its conclusion is true, but if a given argument that is being attacked has a certain degree of strength, merely questioning its conclusion may not be sufficient. What the attacker needs to do in such a case is to put forward a second argument that is stronger than the original argument and that provides evidence for rejecting the conclusion of the original argument. Such an attack is sufficient to defeat the original argument, unless its proponent can give further reasons to support it (Dung, 1995). Still another way to attack an argument is to ask a critical question that casts the argument into doubt, and that may defeat the argument unless its proponent can make some suitable reply to the question. The form of attack will be taken up in section 4.

Even though the given argument may stand, having repelled all attacks of the first three kinds, it may still be defeated on other grounds. One of these is that the argument is irrelevant, even though it may be valid. What is presupposed by this fourth kind of attack is that the given argument is supposed to be used to resolve some unsettled issue in a discussion that is being carried on in the given case. To attack an argument in the fourth way, matters of how the argument was used for some purpose in a context of dialogue need to be taken into account. If an argument has no probative value as evidence to prove or disprove the ultimate *probandum* in this particular discussion, it may be dismissed as irrelevant. Discussions of argument attack and refutation in the literature tend to acknowledge the first three ways of attacking an argument but to overlook the fourth way. The reason could be that this fourth way is more contextual than the first three ways in that it more directly relates to the context of dialogue surrounding the given argument. It could be classified as a procedural objection rather than as an attack.

Still another way to attack an argument is to claim that it commits the fallacy of begging the question. A circular argument, like ‘Snow is white therefore snow is white’, may be deductively valid but still be open to attack on the grounds that it fails to prove its conclusion. The failure here relates to the requirement that the premises of an argument that is being used to prove a conclusion should carry more weight than the conclusion itself. Thus if one of the premises depends on the conclusion, and cannot be proved independently of the conclusion, it is useless to increase the probative weight of the conclusion. Such an argument may be valid, but it is open to the criticism that it is useless to prove the conclusion it is supposed to be proving.

Although there may be four basic ways to attack an argument, asking a critical question is a way of making an objection to an argument that may, or may not be seen as an attack on the argument. The notion of making an objection to an argument seems to be much broader than the notion of attacking an argument, for making an objection can be procedural in nature. We also need to be careful to note that there can be ways of making an objection to an argument that do not fall into any of these five categories of attack on an argument (Krabbe, 2007). Thus the task of defining the notion of an objection precisely, and the task of classifying the various types of objections that can be made to an argument, remain open questions for future work. Still, in this section we have made some progress towards this investigation by carefully describing four

basic ways to attack an argument, and by adding that asking a critical question may also often be seen as a way of attacking an argument by raising critical doubts about it. Argument attacks surely represent some of the central ways of raising an objection about an argument.

Perhaps the best known use of the term ‘rebuttal’ in argumentation theory is Toulmin’s use of it in his argument model, containing the elements datum, qualifier, claim, warrant, backing and rebuttal. In the model (Toulmin, 1958, 101), the datum is supported by a warrant that leads to a claim that is qualified by conditions of exception or rebuttal. For example (99), the claim that a man is a British subject might be supported by the datum that he was born in Bermuda, based on the warrant that a man born in Bermuda will be a British subject. The warrant appears to be similar to what is often called a generalization in logic. This example of an argument is defeasible, because the generalization is subject to exceptions, and hence the argument is subject to defeat if the information comes in showing that the particular case at issue is one where an exception holds. For example, although a man may have been born in Bermuda, he may have changed his nationality since birth (101). Toulmin uses the word ‘rebuttal’, but other words like ‘refutation’ or ‘defeater’ might also be used to apply to such a case.

The meaning term of the term ‘warrant’ in Toulmin’s argument layout has long been the subject of much controversy (Hitchcock and Verheij, 2006). A Toulmin warrant is in typical instances a general statement that acts as an inference license, in contrast to the datum and claim that tend to be specific statements. In logical terms, it could be described as a propositional function or open sentence of this form: if a person x was born in Bermuda, then generally that person x is a British subject.

A rebuttal, judging by Toulmin’s Bermuda example, is an exception to a rule (warrant, in Toulmin’s terms). However, according to Verheij (2008, 20), rebuttal is an ambiguous concept in Toulmin’s treatment, and five meanings of the term need to be distinguished. First, rebuttals are associated with “circumstances in which the general authority of the warrant would have to be set aside” (Toulmin, 1958, 101). Second, rebuttals are “exceptional circumstances which might be capable of defeating or rebutting the warranted conclusion” (Toulmin, 1958, 101). Third, rebuttals are associated with the non-applicability of a warrant (Toulmin, 1958, 102). But a warrant could also be an argument against the datum, a different sort of rebuttal from an argument against the warrant or the claim. In traditional logical terms, this would be an argument claiming that a premise of the inference being rebutted does not hold. Verheij also distinguishes between the warrant that acts as an evidential support of the conditional and the conditional that is one premise in the inference. On his analysis a rebuttal can attack the conditional or it can attack the warrant that supports the conditional as evidence.

Describing rebuttal as citing an exception to a rule of inference on which an argument was based sounds similar to what is called undercutting in the literature on defeasibility (Pollock, 1995). Pollock’s distinction between two kinds of counter-arguments called rebutting defeaters and undercutting defeaters (often referred to as rebutters versus undercutters) is drawn as follows. A rebutting defeater gives a reason for denying a claim by arguing that the claim is a false previously held belief (Pollock, 1995, p. 40). An undercutting defeater attacks the inferential link between the claim and the reason supporting it by weakening or removing the reason that supported the claim. The way Pollock uses these terms, a rebutter gives a reason to show the conclusion is false, whereas an undercutter merely raises doubt whether the inference supporting the conclusion holds. It does not show that the conclusion is false. The classic example is the Tweety argument: Birds fly, Tweety is a bird; therefore Tweety flies. If new information comes in telling us that Tweety is a penguin, the original Tweety argument is undercut. Generally speaking, the argument still holds. Generally birds fly, and hence, given that Tweety is a bird, it follows that Tweety flies. But in this particular case, we have found out that

Tweety is a penguin. Hence in this particular case, since we know that Tweety is type of bird that does not fly, we can no longer use the former inference to draw the conclusion that Tweety flies.

Pollock has another example (1995, 41) that illustrates a defeasible argument that could be called argument from perception.

For instance, suppose x looks red to me, but I know that x is illuminated by red lights and red lights can make objects look red when they are not. Knowing this defeats the *prima facie* reason, but it is not a reason for thinking that x is *not* red. After all, red objects look red in red light too. This is an *undercutting defeater* (Pollock's italics in both instances).

To show how the red light example has the defining characteristics of a species of rebuttal, we can analyze it as an initial (given) argument and a counter-argument posed against it. The original argument says: when an object looks red, then (normally, but subject to exceptions) it is red, and this object looks red to me, therefore this object is red. The rebuttal of the original acts as a counter-argument that attacks the original argument: this object is illuminated by a red light, and when an object is illuminated by a red light, this can make it look red even though it is not, therefore the original argument (the *prima facie* reason for concluding that this object is red expressed by the original argument) no longer holds. According to Pollock (1995, 41) the counter-argument should be classified as an undercutter rather than a rebutter because red objects look red in red light too. Even given the attacking argument, the object may be red, for all we know. Thus in Pollock's terms it would not be right to say that the attacking argument is a rebutting defeater that shows that the conclusion of the original argument is false. What it shows is that because of the new information about the red light, the counter-argument, built on this new information, casts doubt on the conclusion of the original argument. As an undercutter it acts like a critical question that casts an argument into doubt.

Pollock's distinction between rebutters and undercutters is clearly a fundamental to any understanding of defeasible reasoning, but from a practical point of view, it leaves a number of questions open. Is an undercutter a particular instance that makes a defeasible generalization fail in a specific case? Or is an undercutter a special type of counterargument that attacks a prior defeasible argument and acts as a rebuttal to it? Is there a special characteristic of the logical structure of defeasible arguments that leaves them open to an undercutter type of attack, and if so how can we identify this characteristic so that we can learn when making an undercutter type of attack is appropriate? These are all practical questions that seek guidance that might be helpful in telling a participant in argumentation, or a critic of an argument, how to attack that argument or critically question it by finding some sort of standard rebuttal that applies to it. There are also some terminological questions about how to classify the terms 'attack', 'rebuttal' and 'refutation'. Pollock's terminology can be somewhat confusing when we try to apply it to giving practical advice on how to attack, rebut, critically question or refute a given argument, because undercutting does not sound all that different from rebutting. If I find an exception to a rule that defeats the defeasible argument, as in the red light example, surely it is reasonable to say that I have rebutted the original argument. We could even take a step further and use a stronger word, saying that I have refuted the argument. How is rebuttal different from refutation, a term often used in logic textbooks and writings on logic over the centuries? To approach these questions it is best to begin with a practical, and apparently simple example, in which advice is given on how to refute an argument, or perhaps as we might say, to rebut or attack an argument.

2. Internal and External Refutation

Goodwin (2010) presented a methodical procedure to her students on how to refute an argument that contrasts two strategies. The first strategy is that of focusing on the argument's conclusion and arguing for the opposite. She offered the following example. If one side argues

that video games lead to violence, the other side can argue that video games do not lead to violence. This can be recognized as a strategy often called rebuttal or refutation. It is the strategy when confronted with a target argument to present a new argument that has the opposite (negation) of the target argument as its conclusion. Although conceding that this is an important and often effective strategy, she suggests another one that may be even better. Instead of just looking at the conclusion of the other argument, this second strategy is to examine the reasons the other side is giving to support its argument, and see if these reasons hold up under questioning. Among the questions she proposed as ways of attacking the other argument are (1) to ask whether the other side is relying on a biased source, (2) to ask whether the evidence the other side is citing is relevant, or (3) to ask whether the analogy put forward by the other side is really similar.

What is suggested by this advice is that there are basically two ways of attacking an argument. One way, generally called refutation, is to present a new argument that has as its conclusion the negation of the original argument.¹ The other way, generally called asking critical questions, or casting doubt on an argument, is to ask questions that relate to the particular form of the original argument. For example, if the original argument was based on a source, like witness testimony or expert testimony, one could ask the critical question of whether that source is biased. Or if the original argument has the form of an argument from analogy, one could ask the critical question of whether the two cases at issue are really similar. Goodwin states that although attacking the other side's reasons by asking critical questions involves more strategy and paying attention to what the other side says, it can often be more effective because it attacks the opposed argument internally, nicely causing it to fall down.

This practical advice on how to refute an argument is generally very interesting from the point of view of argumentation theory, because it suggests there are two distinctive strategies, refutation and critical questioning, as each might be called, that need to be separated, and that each calls for a different approach. She has shown that each type of argument strategy has a distinctively different structure from the other. This is an important distinction for argumentation theory. Hamblin (1970, 162) distinguished between a weaker and a stronger sense of the term 'refutation'. The weaker he describes as "destruction of an opponent's proof" and the stronger as "construction of the proof of a contrary thesis". It would be nice to have some terminology to make this important distinction between these two meanings of the term 'refutation'. Let us call destruction of an opponent's proof internal refutation, because, as Goodwin has described it, this strategy is to examine the reasons the other side is giving to support its argument, and see if these reasons hold up under questioning. It is an internal attack on the argumentation offered by the other side. Let us call the construction of the proof of a contrary thesis external refutation because it goes outside the original argument to present a new argument that has as its conclusion the negation of the original argument. Attacks can be internal or external.

An example she gives to illustrate the technique of internal refutation is quoted below (with some parts deleted).

The other side said that Dr. Smith's study clearly shows that video games do not lead to violence. But Dr. Smith is biased. His research is entirely funded by the video game industry. That's what the 2001 investigation by the Parent's Defense League demonstrates. So you can see that the other side has no credible evidence linking video games to violence.

¹ Below we will challenge this generally accepted meaning of the term 'refutation' on the grounds that it is too broad. The problem is that we often have cases where a new argument has as its conclusion the negation of an original argument, but the new argument might still be weaker than the original argument. In such case it is questionable whether the new argument is a refutation of the original one. For the moment, however, we accept the broad conventional meaning of the term 'refutation' as a point of departure.

In the example one can see the components of a refutation. First, there are two parties that are presenting arguments on opposed sides of the disputed issue. The issue is whether or not video games lead to violence. The first side has argued that video games do not lead to violence, and has supported its claim by bringing forward the evidence that Dr. Smith's study shows that this claim is true. The opposed side then presents a counterargument, but this counterargument is not an external refutation, a new argument that supports the claim that video games do lead to violence. Instead, it attacks the original argument internally by making the claim that Dr. Smith is biased, and supports it with the reason that his research is entirely funded by the video game industry. So this is a counterargument, but not a refutation in the sense defined above. It is something else. It corresponds to the other technique of attacking an argument that Goodwin described as attacking the reasons the other side is giving by asking critical questions.

We can even analyze this internal type of attacking strategy more deeply by pointing out that the original argument took a particular form. It appears to be an argument from expert opinion that cites a study by someone called Dr. Smith that supposedly showed that video games do not lead to violence. The field of expertise of Dr. Smith is not stated, but it appears we are meant to assume that Dr. Smith is an expert in some field that includes the study of whether video games lead to violence or not. If we can make this assumption, the form of the original argument can then be identified as that of argument from expert opinion. Given this assumption we can understand a little more about the structure of the internal attack used against this argument. The attack makes the claim that Dr. Smith is biased, and this particular type of attack undercuts the argument by finding a weak point in its structure that, once pointed out and supported by evidence, subjects the argument to doubt in such a way that it no longer holds up as a way of supporting its conclusion that video games do not lead to violence. To understand more about how defeasible arguments can have different forms we need to examine the notion of an argumentation scheme.

3. Argumentation Schemes and Critical Questions

Pollock's red light example can be fitted to an argumentation schemes that has been called argument from appearance (Walton, 2006). Although Pollock did not employ the concept of an argumentation scheme with matching critical questions, the pattern of inference of the red light example can be called argument from perception (Walton, Reed and Macagno, 2008, 345).

PREMISE 1: Person P has a ϕ image (an image of a perceptible property).

PREMISE 2: To have a ϕ image (an image of a perceptible property) is a *prima facie* reason to believe that the circumstances exemplify ϕ .

CONCLUSION: It is reasonable to believe that ϕ is the case.

Walton Reed and Macagno (2008, 345) list this form of argument as an argumentation scheme with the following critical question matching it: are the circumstances such that having a ϕ image is not a reliable indicator of ϕ ?

Consider another example (Prakken, 2003): if something looks like an affidavit, it is an affidavit; this object looks like an affidavit; therefore it is an affidavit. This inference might fail if we are taking part in a TV series about a trial in which props are used. A document on a desk might look like an affidavit, but after all this is a TV series. It might not be an affidavit, but merely a prop made to look like one. In the context, the original argument fails to support the conclusion that the document in question is an affidavit. But maybe it is a real affidavit. An easy way to get such a prop for the TV series would be to ask someone who has access to real affidavits to get one for use in the TV series. This example has the same scheme as the red light example.

The scheme representing argument from expert opinion was formulated in (Walton, 1997, 210), with some minor notational changes, as shown below, with two premises and a conclusion. *E* is an autonomous agent of a kind that can possess knowledge in some subject domain. The domain of knowledge, or subject domain, is represented by the variable *F* for field of knowledge. It is assumed that the domain of knowledge contains a set of propositions.

Major Premise: Source *E* is an expert in field *F* containing proposition *A*.

Minor Premise: *E* asserts that proposition *A* (in field *F*) is true (false).

 Conclusion: *A* may plausibly be taken to be true (false).

As shown in (Walton, 1997) any given instance of an argument from expert opinion needs to be evaluated in a dialogue where an opponent (respondent) can ask critical questions. This form of inference is defeasible, provided we take it to be based on a defeasible generalization to the effect that if an expert says *A*, and *A* is in the right field for the expert, then *A* may plausibly be taken to be acceptable as true (subject to exceptions). What kinds of exceptions need to be taken into account corresponding to critical questions matching a scheme? The six basic critical questions matching the appeal to expert opinion (Walton, 1997, 223) are the following.

CQ₁: *Expertise Question*: How knowledgeable is *E* as an expert source?

CQ₂: *Field Question*: Is *E* an expert in the field *F* that *A* is in?

CQ₃: *Opinion Question*: What did *E* assert that implies *A*?

CQ₄: *Trustworthiness Question*: Is *E* personally reliable as a source?

CQ₅: *Consistency Question*: Is *A* consistent with what other experts assert?

CQ₆: *Evidence Question*: Is *E*'s assertion based on evidence?

CQ₁ refers to the expert's level of mastery of the field *F*. CQ₄ refers to the expert's trustworthiness. For example, if the expert has a history of lying, or is known to have something to lose or gain by saying *A* is true or false, these factors would suggest that the expert may not be personally reliable. The assumption made in (Walton, 1997) was that if the respondent asks one of the six critical questions, the initiative shifts back to the proponent's side to respond to the question appropriately. The asking of the critical question defeats the argument temporarily until the critical question has been answered successfully. This approach was a first pass to solving the problem of how to evaluate an argument from expert opinion. More specifically, it was designed to offer students in courses on critical argumentation some direction on how to react when confronted with an argument from expert opinion. Although the critical questions stated in (Walton, 1997) were meant to be practically useful for this purpose, they are also open to formulation in a more precise manner that might make theoretical refinement possible.

The study of attacks, rebuttals and refutations would be aided considerably if some structure could be brought to bear that would enable us to anticipate in a particular case what sort of attack an argument is susceptible to. Here the critical questions matching a scheme can be very useful. For example if the argument is an appeal to expert opinion, we can see already from examining the critical questions matching scheme for argument from expert opinion that this argument will tend to be open to certain types of attack. For example, it will be open to an attack on the grounds that the expert is not a trustworthy source. One of the standard ways of arguing that an expert is not a trustworthy source is to allege that the expert is biased because she has something financially to gain by making the claim. However, it has been shown that critical questions differ in their force. In some instances, merely asking a critical question makes the original argument default, while in other instances, asking the critical question does not make the argument default unless the question asker can offer evidence to back up the question (Walton and Godden, 2005). There are differences between the critical questions on how strongly or weakly asking the question produces such a shift of initiative. Such observations have led to two theories about

requirements for initiative shifting when critical questions matching the argument from expert opinion are asked (Walton and Godden, 2005). According to one theory, in a case where the respondent asks any one of these critical questions, the burden of proof automatically shifts back to the proponent's side to provide an answer, and if she fails to do so, the argument defaults (is defeated). On this theory, only if the proponent does provide an appropriate answer is the plausibility of the original argument from expert opinion restored. According to the other theory, asking a critical question should not be enough to make the original argument default. The question, if questioned, needs to be backed up with some evidence before it can shift any burden that would defeat the argument.

Recent advances in artificial intelligence have developed formal systems to model argumentation, and argument visualization tools that can be used to represent not only reasons given to support an argument, but also attacks on it. Some of these formal systems with visualization tools can also accommodate argumentation schemes. One such system, called Carneades after the Greek skeptical philosopher, can use heuristic strategies to search a space of arguments induced by argumentation schemes (Gordon, 2010). Argumentation schemes in the Carneades model function as heuristic search procedures that apply statements from a database to find arguments pro or con a claim at issue. The arguments that turn up in the resulting stream are alternative ways that can be used to prove the claim. Carneades provides an integrated dialectical framework enabling a variety of legal argumentation schemes, such as argument from expert opinion reasoning, to be used in a comprehensive system supporting both argument construction and argument evaluation tasks.

4. The Carneades System

Part of the definition of a rebuttal is that it is an attack on an argument, and a rebuttal itself would normally seem to be an argument. In order to define the notion of a rebuttal, we also need to have some clear notion of what an argument is. There is not much agreement in argumentation theory on how to define an argument, however. To cope with this problem, it is best to begin with a minimalist account of the structure of an argument. According to this account, an argument is composed of three things: a set of premises, a conclusion, and an inference that leads from the premises to the conclusion. The conclusion is generally taken to be a claim that has been made, and the premises are propositions that are put forward in support of the claim. Beyond this minimal account, it will prove to be useful to have a formal model to represent the notion of an argument, preferably one that would enable us to visualize the premises and conclusion of an argument in a clear way to represent examples of attacks rebuttals and refutations. For example, if we could represent Goodwin's example of an internal refutation, this capability could be extremely helpful. There are many such argumentation visualization tools available at the present time, but it is especially helpful to use one that provides not only a formal model of argumentation, but also an argument visualization tool that fits the model.

Carneades is a mathematical model consisting of definitions of mathematical structures and functions on these structures (Gordon, Prakken and Walton, 2007), and a computational model, meaning that all the functions of the model are computable (Gordon and Walton, 2009). Carneades has been implemented using a functional programming language, and has a graphical user interface. (<https://github.com/carneades/carneades>). Argumentation is modeled by Carneades in a tree structure where the nodes are text boxes containing premises and conclusions of an argument (Gordon, 2010). The premises are connected to the conclusion in the normal way in an argument with an arrow pointing to the conclusion. An argument that supports a conclusion is indicated by a circle containing a + sign. The premise is an exception is joined to a circle by a

dashed line. How Carneades displays the structure of the argument in Pollock's red light example is shown in figure 1.

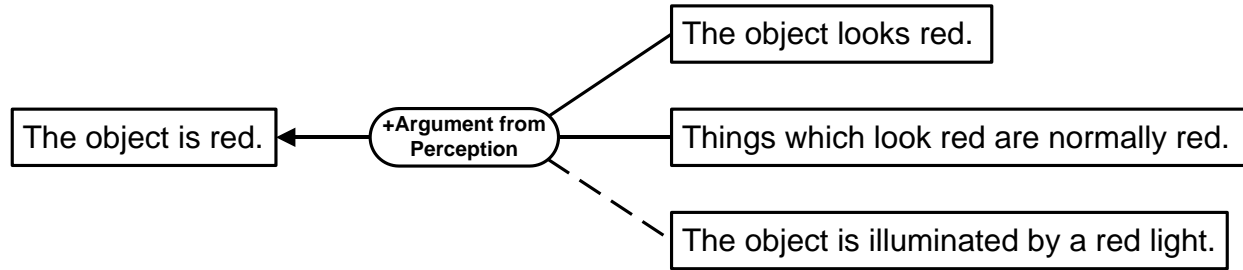


Figure 1: Exception in Pollock's Red Light Example

As shown in figure 1, the statement at the bottom right is an exception, and so the argument as a whole represents a Pollock-style undercutter. In the Carneades model, this argument is represented as a typical defeasible argument that has two normal premises, displayed as the top two boxes on the right in figure 1. But this argument is subject to an exception, and in Carneades the exception is represented as an additional premise of a special kind that can defeat the original argument. Carneades can also be used for evaluating arguments, and how the procedure of evaluation can be illustrated using the case of Pollock's red light example is shown in figure 2. As shown by the checkmark text box at the bottom, the statement that the object is illuminated by red light has been accepted. Once the statement has been accepted, even though the two premises above it would normally enable the conclusion to be accepted provided these two premises are accepted, in this situation, since the exception applies, the conclusion is cast into doubt. The status of the conclusion is represented by the question mark appearing in its text box. What this analysis visualizes is a situation in which the conclusion is rendered questionable, and hence not acceptable. It does not tell us, however, that the conclusion is false or unacceptable.

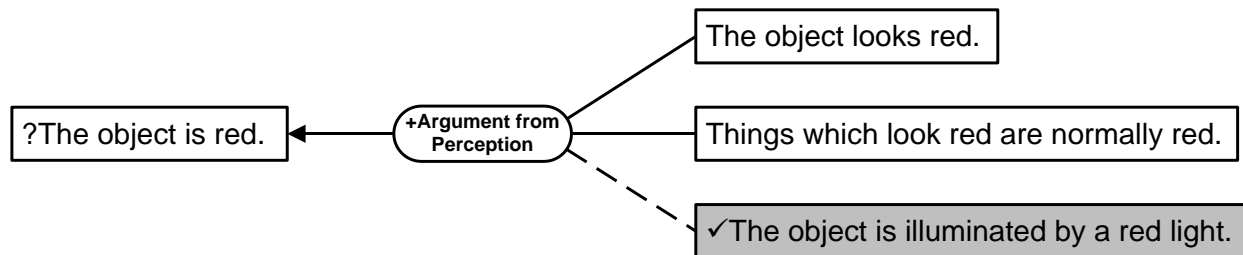


Figure 2: Undercutter in Pollock's Red light Example

Carneades defines formal properties that are used to identify, analyze, construct, visualize and evaluate arguments (Gordon and Walton, 2006). Part of the definition of a rebuttal is that it is an attack on an argument, and a rebuttal itself is also an argument. It follows that in order to define the notion of a rebuttal, we surely also need to have some clear notion of what an argument is. As noted just above in this section, an argument is taken to have three basic components: a set of premises, a conclusion, and an inference that leads from the premises to the conclusion.

Figures 1 and 2 show how these three components are related. In the following formal definition of an argument in Carneades (Gordon and Walton, 2009), a distinction is drawn between two types of opposition. One is negation, represented in the same way as in classical propositional logic where a proposition p is true if and only if its negation is false. The negation

of a proposition, in other words, has the opposite truth value of the original proposition. The other is complement. The complement of a set is the set of things outside that set (Gordon and Walton, 2009, 242-243).

Let L be a propositional language. An *argument* is a tuple $\langle P, E, c \rangle$ where $P \subset L$ are its *premises*, $E \subset L$ are its *exceptions* and $c \in L$ is its *conclusion*. For simplicity, c and all members of P and E must be literals, i.e. either an atomic proposition or a negated atomic proposition. Let p be a literal. If p is c , then the argument is an argument *pro p*. If p is the complement of c , the argument is an argument *con p*.

According to this definition we can understand the notions of an argument *pro* a proposition p and argument *con* a proposition p as follows. If p is the conclusion of the argument, the argument is said to be *pro p*, whereas if some proposition other than p is the conclusion of the argument, the argument is said to be *con p*. Defeaters (rebuttals) are modeled as arguments in the opposite direction for the same conclusion. If one argument is *pro* the conclusion, its rebuttal would be another argument *con* the same conclusion. Premise defeat is modeled by an argument *con* an ordinary premise or an assumption, or *pro* an exception (Gordon, 2005, 56). In the Carneades system, critical questions matching an argument are classified into three categories: ordinary premises, assumptions or exceptions. External refutations are modeled as arguments in the opposite direction for the same conclusion. If one argument is *pro* the conclusion, its refutation would be another argument *con* the same conclusion. Premise defeat is modeled by an argument *con* an ordinary premise or an assumption, or *pro* an exception (Gordon, 2005, 56). See how Carneades models the distinction between internal and external refutation, we show how this distinction works in the case of argument from expert opinion.

Let's begin with the notion of external refutation to see how it works generally in cases of argument from expert opinion. In a case of external refutation, as shown in figure 3, we have one argument from expert opinion in which the premise is that expert 1 says that some proposition A is true and the conclusion is the proposition that A is true. This is the argument shown at the top in figure 3, and it is a *pro*-argument, as shown by the + in the circle representing the argument. Beneath it is the second argument that attacks the first argument, based on the premise that there is another expert who says that the opposite of A is true. The second argument is an external refutation of the first one, because it is a separate opposed argument that has the opposite conclusion of the first argument.

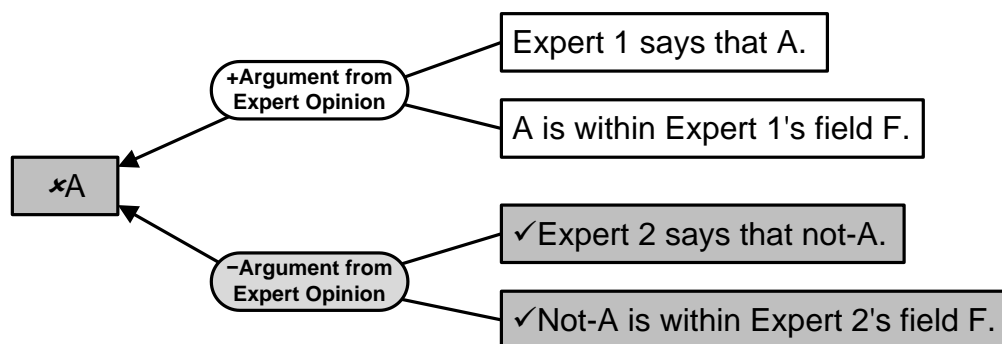


Figure 3: How Carneades Models External Refutation

But if the second argument is merely a rebuttal of the first argument can it properly be called a refutation? Certainly it fits the definition of an external refutation of the kind attributed to Hamblin above, but there is more to say about it. Notice that in figure 3, the premises shown at the bottom appear in darkened textboxes and have check marks in front of them, indicating that

this premise has been accepted. Notice that the premises shown at the top appear in undarkened text boxes with no check mark in front, indicating that each premise has merely been stated but has not been accepted. What will happen automatically in Carneades is that the bottom argument will be taken as refuting the top one. Since it has two accepted premises, when both premises are considered together, the conclusion *A* comes out as rejected (indicated by ✖).

In such a case, we can say that the first argument is refuted by the second one in a strong sense of the term ‘refutation’ meaning not only that the second argument goes to the opposite conclusion of the first one, but it does so in such a way that it overwhelms the first argument, providing a reason to infer that the conclusion of the first argument is no longer acceptable. We could say that in this strong sense of refutation, the second argument successfully refutes the first argument. Or perhaps we could draw the distinction in a different way by saying that second argument not only rebuts first argument but also refutes it. The terminology remains uncertain here but we will clarify it later.

No matter how we describe what has happened in this example in terms of the distinction between rebuttal and refutation, we can see why it illustrates how Carneades models the notion of an external refutation. In an external refutation, we have two separate arguments, and one attacks the other externally by providing an independent line of argument that goes to the opposite of the conclusion of the first argument. Carneades models the notion of an internal refutation in a completely different way by focusing on the critical questions matching the argumentation scheme, and goes into considerations of different ways these critical questions can be used to attack the original argument. One of the main features of Carneades is that it enables critical questions to be represented on argument diagrams of the kinds shown in the figures 1, 2 and 3 above (Walton and Gordon, 2005). In the standard argument diagrams, the text boxes (nodes in the tree) contain propositions that are premises and conclusions of arguments, but there is no obvious way that critical questions can be represented on such a diagram. Carneades solves this problem by enabling a distinction to be drawn between two ways an argument from expert opinion should be critically questioned, and thus enables the critical questions to be represented as implicit premises of an argumentation scheme on an argument diagram. The two assumptions that (1) the expert is not trustworthy and (2) that what she says is not consistent with what other experts say, are assumed to be false. It is assumed, in other words, that (1) and (2) are false until new evidence comes in to show that they are true. The two assumptions that (1) the expert is credible as an expert and that (2) what she says is based on evidence, are assumed to be true, until such time as new evidence comes in showing they are false. Also assumed as true are the ordinary premises that (1) the expert really is an expert, (2) she is an expert in the subject domain of the claim, (3) she asserts the claim in question, and (4) the claim is in the subject domain in which she is an expert.

Now let’s look once again at the expertise question, to see how it could be classified. It is about *E*’s depth of knowledge in the field *F* that the proposition at issue lies in. As noted above, the expertise question seems to ask for a comparative rating. What if the proponent fails to answer by specifying some degree of expertise, like “very credible” or “only slightly credible”? As noted above it seems hard to decide what the effect on the original argument should be. Should it be defeated or merely undercut? It seems like it should only be undercut, because even if we don’t know how strong the argument from expert opinion is, it might still have some strength. It might even be very strong, for all we know.

The field and opinion questions can be modeled as ordinary premises of the arguments from expert opinion scheme in Carneades. Now let’s look back at the trustworthiness question, which refers to the reliability of the expert as a source who can be trusted. If the expert was shown to be biased or a liar, that would presumably be a defeater. It would be an *ad hominem* argument used

to attack the original argument, and if strong, would defeat it. But unless there is some evidence of ethical misconduct, as noted above, the proponent could simply answer ‘yes’, and that would seem to be enough to answer the question appropriately. As noted above, to make such a charge stick, the questioner should be held to supporting the allegation by producing evidence of bias or dishonesty.

According to the discussion above, only the consistency and backup evidence questions need some evidence to back them up before the mere asking of the question defeats the original argument. Hence only these two of the critical questions are treated as exceptions. The results of how the critical questions should be classified as premise on the Carneades model can be summed up as follows.

Premise: *E* is an expert.

Premise: *E* asserts that *A*.

Premise: *A* is within *F*.

Assumption: It is assumed to be true that *E* is a knowledgeable expert.

Assumption: It is assumed to be true that what *E* says is based on evidence in field *F*.

Exception: *E* is not trustworthy.

Exception: What *E* asserts is not consistent with what other experts in field *F* say.

Conclusion: *A* is true.

It is shown in figure 4 how argument from expert opinion is visually represented in the Carneades graphical user interface, and how each premise is represented. A normal premise is represented by a solid line, an exception is represented by a dashed line, and assumption is represented by a dotted line.

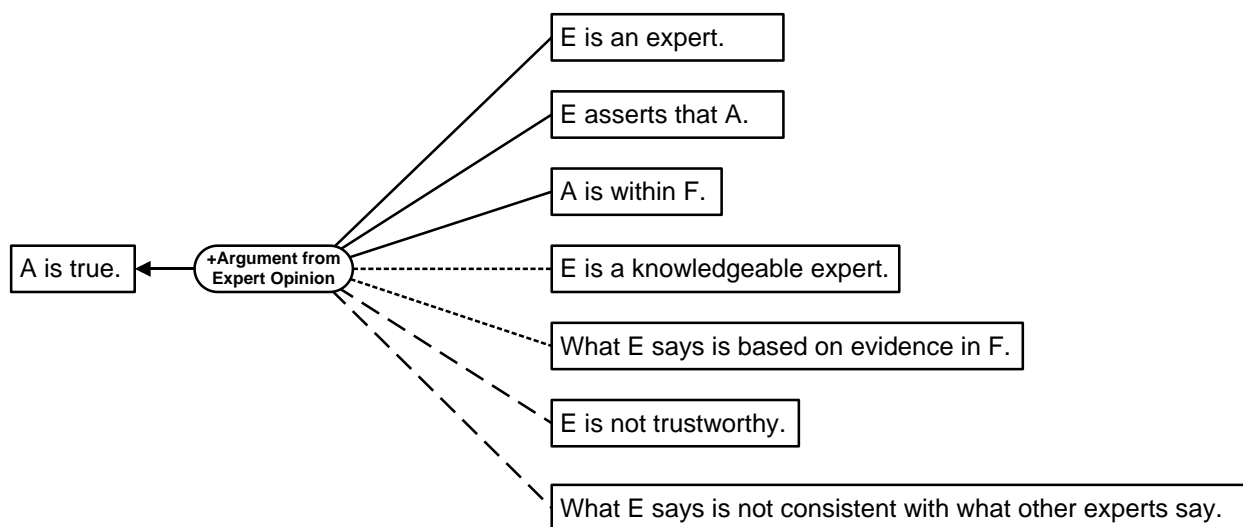


Figure 4: Visualization of Argument from Expert Opinion in the Carneades Interface

As figure 4 shows, the critical questions are represented as additional premises alongside the ordinary premises in the scheme for argument from expert opinion. This means that, as far as Carneades is concerned, attacking the argument by asking anyone critical questions can be classified as a premise attack argument. According to the Carneades model, the ordinary premises are stated, whereas the other premises expressing critical questions are either assumptions or exceptions.

If we are using the Carneades graphical user interface to help us devise a strategy to refute an argument we are confronted with, we can look over the evidence available in the case, or that could possibly be collected in the case, in order to decide which of the critical questions would be the best one to pose. Posing a critical question of the assumption type requires no evidence to back it up in order to defeat the original argument. These would be the first premises to look at. Goodwin described the strategy as one of examining the reasons the other side is giving to support its argument to see if these reasons hold up under critical questioning. However, if there is evidence that could be used to back up one of the critical questions, that would be the question to pose. As we see in the case of Dr. Smith, there is evidence that could be used to back up the claim that he is biased. Hence Carneades can automatically point to the trustworthiness question, represented as an exception in the argument visualization, and indicate that the best strategy is to ask this question.

5. How Carneades Models Attacks and Rebuttal

Not only are schemes classified under other schemes, but critical questions also have a classification structure as well. For example, although argument from bias is a specific type of argument in its own right with its distinctive argumentation scheme, asking a critical question about bias is so common in responding to arguments from expert opinion that it needs to be identified as a specific critical question in its own right with respect to the scheme for argument from expert opinion. In (Walton 1997, 213-217) the bias critical question is treated as a sub-question of the trustworthiness question. In other words, questioning whether an expert is biased is treated as a special case of questioning whether the expert is personally reliable as a source. The reason is that questioning on grounds of bias is a way of questioning the trustworthiness of an expert source. A biased expert need not be completely untrustworthy, but if there are grounds for suspecting a bias, that is a good reason for having reservations about the strength or even the acceptability of an argument from expert opinion.

Let's go back to the example Goodwin gave to illustrate the technique of attacking the reasons the other side has put forward in its argument. In this example, the attack alleges that Dr. Smith is biased, because his research is entirely funded by the video game industry. Next, evidence to support this claim of bias is put forward. It is claimed that the 2001 investigation by the Parent's Defense League constitutes evidence to support bias. The structure of this argument from expert opinion is shown in figure 5. The three ordinary premises of the argument from expert opinion are shown at the top in the three darkened boxes. In each case, a check mark appears before the proposition in the box. The checkmark redundantly shows, along with the darkened box, that these three propositions have been accepted. Carneades would automatically darken the box for the conclusion and put a checkmark before the proposition that *A* may be taken to be true. This is the normal evaluation procedure that Carneades is set up to automatically carry out. However in this instance, the box in the middle at the bottom containing the proposition that *E* is not trustworthy also has a checkmark in front of it. Moreover, this proposition is supported by evidence of *E*'s bias, and we can take it that this evidence is strong enough to be accepted. Since the premise that *E* is not trustworthy is an exception, the argument from expert opinion to the conclusion that *A* is true is cast into doubt. Hence we see that this proposition has a question mark in front of it in figure 5. Figure 5 shows generally how the trustworthiness and bias critical questions are modeled by Carneades, and how a finding of bias functions as support for a trustworthiness exception of the kind that can cast an argument from expert opinion into doubt by rebutting it.

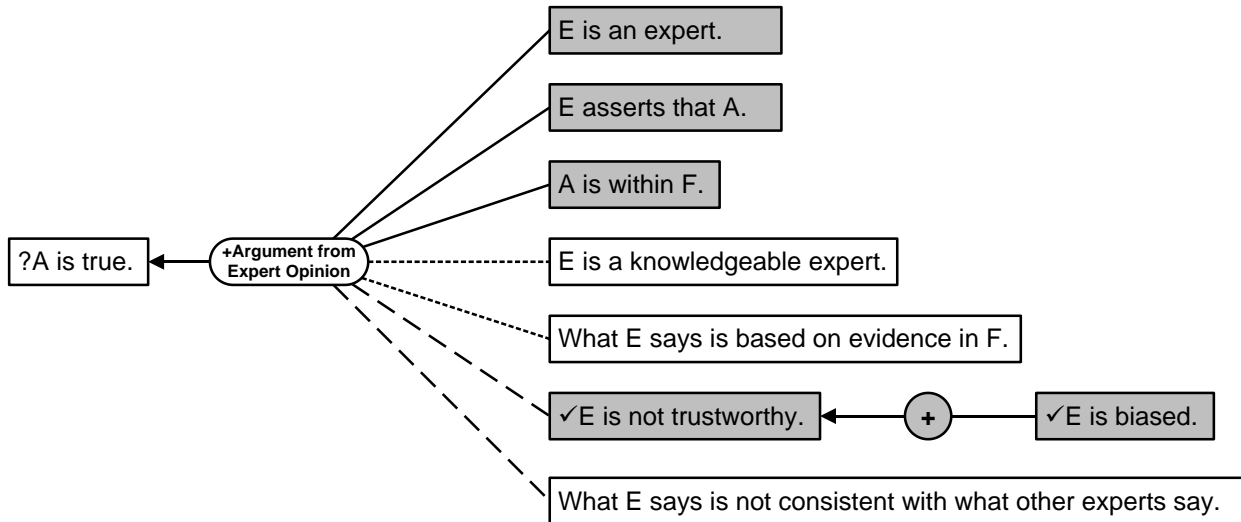


Figure 5: Argument Undercut by Bias Attack

Next let's examine how Carneades represents the example of the second strategy of refutation described by Goodwin. In this example, it is argued that Dr. Smith's study can be attacked internally by arguing that it was paid for by the video industry. The basic argument is shown in figure 6.

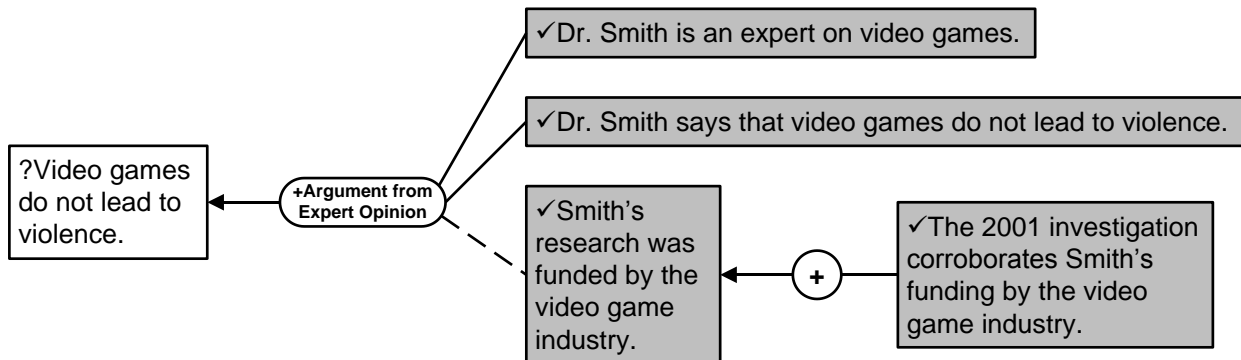


Figure 6: Defeasible Structure of the Video Example Visualized by Carneades

When it comes to evaluating the argument to see how a rebuttal works, we could look at the Carneades visualization of it in figure 6, where the two normal premises at the top are accepted. The remaining premises, as shown in figure 5, are not shown in figure 6. In figure 6 it is also shown that the exception at the bottom, the proposition that Smith was paid by the video industry, is accepted on the grounds that it is supported by the evidence of the 2001 investigation. Although the two premises at the top would normally be enough to support acceptance of the conclusion on the Carneades model, in this instance the conclusion is not accepted. It is shown as questioned. The reason is that the premise at the bottom, the proposition that Smith was paid by the video industry, is an exception, and moreover it is an exception that has been accepted, based on the evidence of the investigation. Thus what figure 6 shows is that the argument from expert opinion has been defeated. It has been undercut by giving evidence to show that an exception applies. Goodwin described the strategy as one of examining the reasons

the other side is giving to support its argument to see if these reasons hold up under critical questioning. In this example, it is fair to say that the argument did not hold up under critical questioning. But the question is: has this argument been refuted, or has the conclusion merely been cast into doubt?

The notion of an attack is another concept that needs to be fitted into this system of classification. In the Carneades system, a proposition can be stated, questioned, assumed or accepted. In Carneades one argument can attack another in basically four ways.

1. It can attack one or more of the premises of the prior argument and show that one or more of them is questionable.
2. It can attack one of these premises and show that one or more of them is not acceptable.
3. It can attack the conclusion by posing a counterargument that shows that the conclusion is questionable.
4. It can attack the conclusion by posing a counterargument that shows that the conclusion is not acceptable.

Is an attack the same thing as a rebuttal? At first, it seems that it is, because an attack on an argument is designed to show that the argument is questionable, that it is not supported by the evidence, or even that the evidence shows that it is untenable. On the other hand, it would seem that it is not, because asking a critical question could perhaps be classified as an attack on an argument, it would not seem quite right to say that asking such a critical question is a rebuttal.

This classification may be borderline, however. Asking a critical question casts doubt on an argument, but is casting doubt on an argument rebutting it? What Carneades has shown is that critical questions matching argumentation schemes are of two different kinds in this regard (Walton and Gordon, 2005). Some critical questions act as rebuttals when they are asked, because unless the proponent of the argument replies appropriately to the question, the argument is defeated. Asking other critical questions does not defeat the original argument unless the question is backed up by some evidence. In this kind of case it does not really seem quite right to describe the asking of the critical question as a rebuttal. The word 'rebuttal' also implies that the attacking is being done by posing another argument, and not merely by asking a question about the original argument, even if it is a critical question that casts doubt on the argument.

In addition to the three basic ways of attacking an argument listed in section 1, we also considered some other ways. One of these ways is to argue that the given argument is not relevant to the ultimate conclusion to be proved in the case at issue. To attack an argument in the fourth way, matters of how the argument was used for some purpose in a context of dialogue needs to be taken into account. Even though the given argument may stand, having repelled all attacks of the first three kinds, its force as argument may be nullified if it is irrelevant. But is this kind of charge a rebuttal? It is not, if it is not an attack on the argument itself, but rather a charge that the argument is not useful for some purpose. A charge of irrelevance is best seen as a procedural objection to the effect that the argument is not useful to resolve the ultimate issue under discussion. To model this kind of procedural objection, we have to look at argumentation as a process, after the manner of Carneades.

6. How Carneades Models Relevance

The Carneades system can be used to assist an agent preparing a case by constructing arguments used to prove a claim in a situation where there is an information service that continually provides new information that might be useful for this purpose (Ballnat and Gordon, 2010). The agent only presents his case once the resources provided by the information service have been exhausted. If that has not happened, the agent tries to make his case by asking

questions and searching for new information to construct arguments. Then he selects which arguments to put forward in order to prove the goal thesis that he wants to prove. In this system there is a continuous loop as the agent keeps collecting new information from the information service and uses that information to construct new arguments. A simplified version of this process comparable to the figure in (Ballnat and Gordon, 2010, 52) is shown in figure 7.

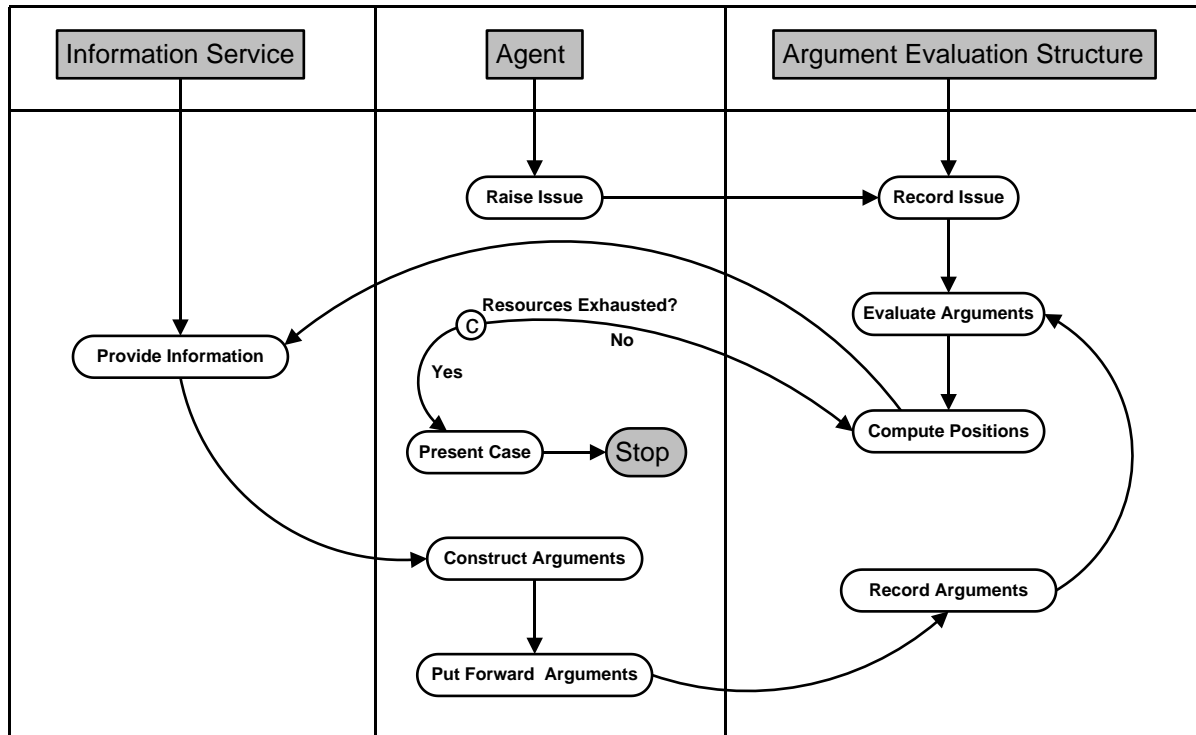


Figure 7: An Argumentation Process

Only once these information and argument construction resources are exhausted does the agent either prove his thesis or find that there are insufficient resources to do so. As the agent proceeds through this argumentation process, he tries to find alternative positions to support his argument.

Suppose I want to prove my claim that proposition A is true. What should I do? Should I make a further argument *pro* A ? Or should I make another argument *con* B , where B is some proposition that is being used by the opposition to refute A ? Or should I put forward arguments supporting some premise of one of my previous arguments that were put forward in support of A ? In other words, what should be my next goal, where a goal is a proposition that a party searches around for to work on next, by looking for arguments *pro* or *con* the proposition he ultimately wants to prove in the dialogue. Carneades is being used here as a device to find which arguments are relevant by telling him which propositions he should choose to work on next, given the information he already has.

As well as providing a method for helping in arguer to determine which arguments are relevant, Carneades can also be used to help in arguer determine which arguments are not relevant. What is presupposed by a claim of relevance is that the given argument is supposed to be used to resolve some unsettled issue in a discussion that is being carried on in the given case. If an argument has no probative value as evidence to prove or disprove the thesis at issue in a

particular discussion, it may be dismissed as irrelevant. However, although this attack may knock the argument out of consideration, it is not, strictly speaking, a rebuttal. It should be classified as a procedural objection claiming that the argument under consideration is useless to prove some ultimate claim that the arguer is building a case to prove. On this analysis, the objection to an argument on grounds of relevance is different from the rebuttals and refutations we have been concerned with. Still, it is interesting to see that Carneades has the capability for dealing with claims of relevance and irrelevance because it can model argumentation as a process.

The procedure recommended for seeking some means of refuting or objecting to an argument broadly follows the line of investigation in the paper. It starts out by focusing on refutation in the narrower sense, referring to external and internal refutation, then goes on to means of attack and investigation of an argument offered by argumentation schemes and critical questions. From there, it looks more widely to other kinds of objections that may be procedural in nature, and that may not focus so narrowly on internal or external refutation. As it expands outwards, it takes into account the wider context of an argument, and can do so by viewing argumentation as a process using the Carneades system.

7. Classifying Objections, Rebuttals and Refutations

An objection does not necessarily have to be a counter-argument posed against an original argument. It could be merely the asking of a critical question. Even when an objection is a counter-argument posed against an original argument, it does not have to be an argument that the original argument is weak, unsupported or incorrect. It could be a procedural objection, not implying that the argument it is addressed against is incorrect, insufficiently supported by evidence, or even questionable as an argument in itself. Such a procedural objection could merely claim that the argument, even though it might be reasonable enough, or well enough supported in itself, is not appropriate for use in the context of the given discussion. In law for example, an argument might be objected to on the grounds that the evidence it purports to bring forward has been obtained illegally, even though that evidence might otherwise be quite convincing in itself as a rational argument. It follows that an objection is not necessarily a rebuttal or a refutation. The term ‘objection’ represents a wider category.

There may be a narrower sense of the word ‘objection’, however, that is used in logic. Govier (1999, 229) considers an objection to be an argument raised against a prior argument. Hence a question is not an objection: “On this view, a question purely considered as such does not itself constitute an objection”. On her account, an objection can be directed in one of two ways. The objection can claim that there is something wrong with the conclusion, or it can claim that there is something wrong with the argument. But these are not the only possibilities. She classifies five types of objections (231), depending on what the objection is specifically raised against: (1) against the conclusion, (2) against the argument in support of the conclusion, (3) against the arguer, (4) against the arguer’s qualifications, personal characteristics or circumstances, or (5) against the way the argument or conclusion was expressed. It is interesting to note that some of these categories of objection may correspond to or overlap with types of arguments associated with some of the traditional informal fallacies. The third category and two parts of the fourth may correspond to the *ad hominem* type of argument while the first part of the fourth may correspond to a common type of attack on arguments from expert opinion.

A different way of classifying objections to an argument has been put forward by Krabbe (2007, 55-57) who lists seven ways an opponent can critically react to a proponent’s expressed argument. (1) A request for clarification, explanation or elucidation may contain an implicit criticism that the argument was not clearly expressed to start with. (2) A challenge to an

argument comprises an expression of critical doubt about whether a reason supports the argument. (3) A bound challenge raises a more specific doubtful point that offers some reason for entertaining doubt. (4) An exposure of a flaw poses a negative evaluation of an argument and requests further amplification. (5) Rejection is a kind of critical reaction by an opponent who may not deny that the proponent's argument is reasonable, but takes up an opposite point of view. (6) A charge of fallacy criticizes the contribution of the proponent by claiming he or she has violated some rule of fair procedure. (7) A personal attack is a common kind of critical reaction that provides a means of defense against unreasonable moves by one's opponent. Krabbe (2007, 57) suggests that these critical reactions can properly be called objections, because they express dissatisfaction with an argument presented by a proponent. However, Krabbe (2007, 57) writes that to speak of a request for clarification or a pure challenge as an objection would be an overstatement, because objections presuppose a negative evaluation, whereas these other two types of reaction precede evaluation.

There are differences between these two views on what an objection is. Govier (1999, 229), requires that an objection be an argument when she wrote, "An objection is an argument, a consideration put forward, alleged to show either that there is something wrong with the conclusion in question or that there is something wrong with the argument put forward in its favor". Krabbe does hold the view that an objection has to be an argument. Ralph Johnson, in an unpublished manuscript shown to the author, has advocated the view an objection is a response to an argument that can be in the form of a question or a statement, and does not have to be an argument. I will take it that objection is a wider category than rebuttal, so that while putting forward a rebuttal is making an objection in some instances, there are also instances in which an objection to an argument should not be classified as a rebuttal.

The notion of a challenge is well known in argumentation. In his *Why-Because System with Questions*, Hamblin (1970, chapter 8), has a locution 'Why *A*?' that is a challenge or request made to the hearer to provide a justification (an argument) for the statement *A* queried. But what is a challenge to an argument (as opposed to a statement)? Most likely it would seem to be a critical question. But there could be other sorts of argument challenge, for example such a challenge could be a procedural objection that the argument is irrelevant.

Following the line of this paper, the notion of a rebuttal can be defined as follows. A rebuttal requires three things. First, it requires a prior argument that it is directed against. Second, the rebuttal itself is an argument that is directed against this prior argument. Third, it is directed against the prior argument in order to show that it is open to doubt or not acceptable.

A rebuttal is one of a pair of arguments, where the two arguments are ordered, logically rather than temporally, so that the one precedes the other, and so that the second one is directed against the first one. What does "directed against" mean? One argument can have another argument as its target. The one can be meant to support the other, or can be meant to attack the other, or the two arguments can be independent of each other. But something more is meant here. What seems to be implied is that a rebuttal is an argument directed against another argument to show that the first argument is somehow defective. To rebut an argument is to try to show that the argument is questionable, or not supported by the evidence, or even that the evidence shows that it is untenable.

Is a refutation the same as a rebuttal? One way to define the relationship between these two terms strongly suggested by our discussion of how Carneades handles the type of argument configuration shown in figure 3 would be to say that a refutation is a successful rebuttal. On this way of defining the two terms, a rebuttal is aimed to show that the argument it is directed against is questionable or untenable. A refutation is a rebuttal that is successful in carrying out its aim. A refutation is a counterargument that is not only posed against a prior argument, but weighs in

more strongly when evaluated against the prior argument so that it reverses the conclusion of the prior argument. So defined, the one term would seem to be a subspecies of the other. A refutation is a species of rebuttal that shows that the argument it is aimed at is untenable. When an argument you have put forward is refuted, it has to be given up. If the argument is confronted with a rebuttal, you may or may not have to give it up. Only if the rebuttal is a refutation do you have to give it up. The same point can be made about attack. Attack does not imply defeat.

The term challenge is widely used in formal dialogue systems. As noted above, Hamblin has a locution ‘Why A?’, called a challenge, in his Why-Because System with Questions. To respond appropriately the hearer is expected to provide premises that the challenger is committed to already, or can be brought to concede at future moves), and *A* is supposed to be a conclusion implied by these premises according to the inference rules in the system. A challenge, in this sense, is not an argument. It is a speech act that requests some evidence to support a claim made by the other party. As the distinction between assumptions and exceptions made in Carneades shows, some critical questions are merely challenges, whereas other critical questions, although they have the speech act format of a challenge, defeat the other party’s argument unless she comes forward with some evidence to support her argument.

The classification tree shown in figure 8 offers a way of clarifying these terms.

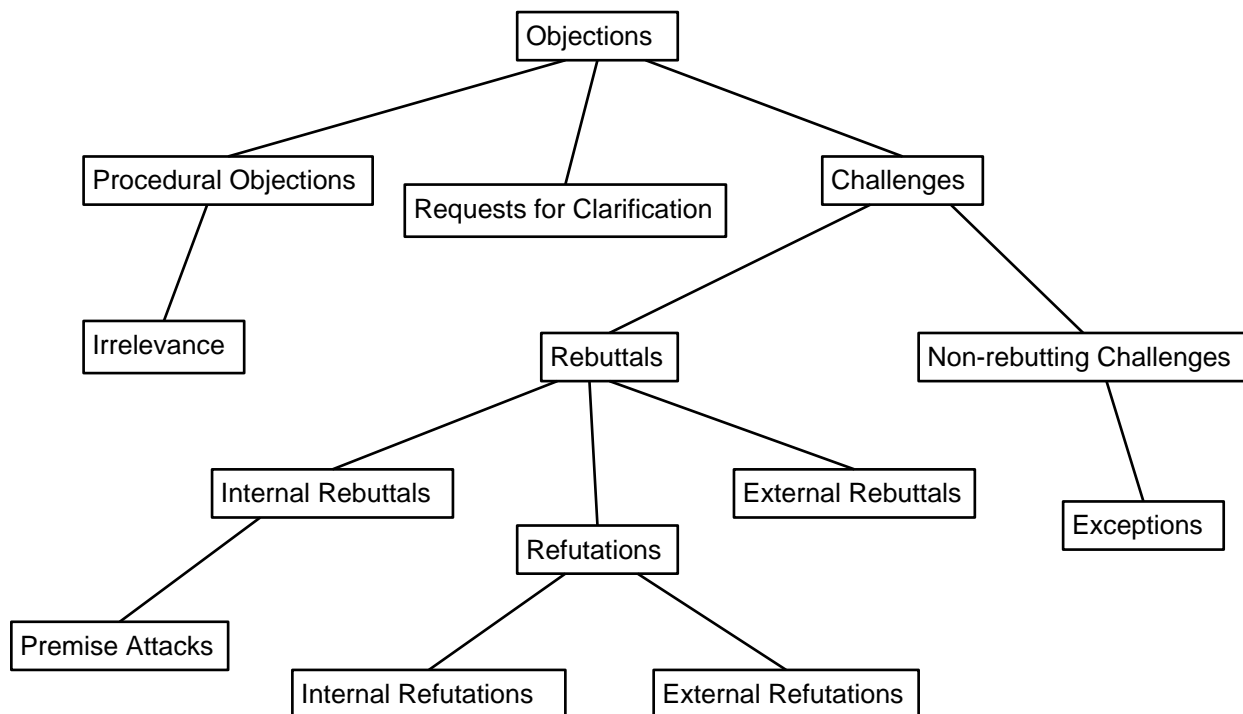


Figure 8: Classification Tree for Species of Objections

Objection is taken to be a wide category that includes procedural objections, and many kinds of attacks that should not, strictly speaking, be called rebuttals. An objection of irrelevance is shown as an example of a procedural objection. An objection does not have to be a rebuttal even though it is comparable to a rebuttal in that it assumes that there is something negative about an original argument, or move in argumentation, that needs to be responded to, called into question and corrected. The classification tree in figure 8 incorporates the notion of a challenge. A challenge is defined after the manner of Krabbe as a species of objection that comprises an

expression of critical doubt about whether a reason supports the argument that is challenged. However, this way of defining the notion of challenge makes it appear to be very close to a Pollock-style undercutter, a species of argument attack modeled as an exception in Carneades. Figure 8 clarifies the notion of the challenge by classifying the Pollock-style undercutter as an exception, using the term and its Carneades meaning. Exceptions are classified as critical questions that need to be backed up by evidence before they defeat the argument they are directed against. The classification tree shown in figure 8 also incorporates the distinction between an internal refutation or rebuttal and an external one. Hence it is a comprehensive classification scheme that includes all the species of objections analyzed in the paper.

A rebuttal is a species of objection. A refutation is a species of rebuttal that is successful in knocking down the argument it was directed against. A *rebuttal* is an argument directed against another argument to show that the first argument is somehow defective. An *attack*, in the sense of the word as used in the field of argumentation, is an argument directed against another argument to show that the first argument is somehow defective. In other words, for purposes of argumentation study, the words ‘rebuttal’ and ‘attack’ can be taken as equivalent.

To rebut an argument is to try to show that the argument is questionable, or that it is not supported by the evidence, or even that the evidence shows that it is untenable. A rebuttal can attack a premise of the original argument, it can attack the conclusion, or it can act as an undercutter that attacks the inference from the premises to the conclusion. How it does this, as illustrated by Pollock’s red light example and the Tweety example, is by finding an exception to a general rule that is the warrant of a defensible argument. A *refutation* is a species of rebuttal that shows that the argument it is aimed at is unacceptable. It could be called a knock-down counter-argument. When an argument you have put forward is confronted with a refutation, it has to be given up. Both rebuttals and refutations can be external or internal.

8. Conclusion

The practical argument attack and refutation procedure derived from the analysis in this paper has seven steps. The procedure can be applied using these seven steps

1. Look for a refutation in the sense described in section 2. If you have a counter-argument that can be used to prove the opposite of the conclusion claimed in the original argument, go for an external refutation.
2. Alternatively, if this seems to be a better route of attack, go for an internal refutation.
3. The first step in seeking a suitable internal refutation is to see if the argument you are trying to attack fits a known argumentation scheme. The list of the most basic types of arguments that have argumentation schemes are the following: argument from position to know, argument from witness testimony, argument from expert opinion, argument from analogy, argument from verbal classification, argument from rule, argument from precedent, practical reasoning, value-based practical reasoning, argument from appearances (perception), argument from ignorance, argument from consequences (positive or negative), argument from popular opinion, argument from commitment, direct *ad hominem* argument (personal attack), circumstantial *ad hominem* argument, argument from bias, argument from correlation to cause, argument from evidence to a hypothesis, abductive reasoning, argument from waste, and slippery slope argument.
4. If the argument fits a scheme that can be identified, look at the critical questions matching the scheme, and see which question is most appropriate.
5. In the Carneades model critical questions are represented as different kinds of premises. If the premise you choose to attack is either an ordinary premise or an assumption, simply question it.
6. If it is an exception, question it only if you have the evidence required to back it up.

7. If none of this procedure so far has come up with a good result, go on to look for some procedural objection, like questioning whether the argument is relevant.

Throughout the main part of paper the narrower concern has been with the concept of refutation illustrated by Goodwin's example that we began with. But later there was a move to considering other kinds of objections that, it was argued, do not fit this narrower model. The list of objections provided by Krabbe gives a good idea of what some of these objections are, but there is no reason to think that this list is complete. Some of the objections correspond to informal fallacies of the kind well known in the argumentation literature. Objecting to an argument on the grounds that it is circular and therefore begs the question is an example. The task of studying and classifying additional kinds of objections to an argument associated with fallacies is a project for future research.

References

Ballnat, S. and Gordon, T. F. (2010). Goal Selection in Argumentation Processes, *Computational Models of Argument: Proceedings of COMMA 2010*, ed. P. Baroni, F. Cerutti, M. Giacomin and G. R. Simari, Amsterdam: IOS Press, 51-62.

Dung, P. M. (1995). On the Acceptability of Arguments and Its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-person Games, *Artificial Intelligence*, 77, 321-357.

Goodwin, J. (2010). How to Refute an Argument, accessed November 26, 2010 at this site: <http://www.public.iastate.edu/~goodwin/spcom322/refute.pdf>

Gordon, Thomas F. (2010). An Overview of the Carneades Argumentation Support System, *Dialectics, Dialogue and Argumentation*, ed. C. Reed and C. W. Tindale, London: College Publications, 145-156.

Gordon Thomas F. and Walton, D. (2006). The Carneades Argumentation Framework, *Computational Models of Argument: Proceedings of COMMA 2006*, ed. P. E. Dunne and T. J. M. Bench-Capon. Amsterdam: IOS Press, 195-207.

Gordon, T. F., & Walton, D. (2009). Proof Burdens and Standards, ed I. Rahwan & G. Simari *Argumentation in Artificial Intelligence* Berlin, Germany: Springer-Verlag, 239-260.

Gordon, T. F., Prakken, H. and Walton, D. (2007). The Carneades Model of Argument and Burden of Proof, *Artificial Intelligence*, 171, 875-896.

Govier, T. (1999). *The Philosophy of Argument*. Newport News, Virginia: Vale Press.

Govier, T. (2006). *The Philosophy of Argument*, Newport News, Virginia, Vale Press, 1999.

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

Hitchcock, D. and B. Verheij (eds), *Arguing on the Toulmin Model: New Essays in Argument Analysis and Evaluation*. Dordrecht: Springer, 2006.

Krabbe, E. C. W. (2007). Nothing but Objections!, *Reason Reclaimed*, ed. H. V. Hansen and R. C. Pinto, Newport News, Virginia: Vale Press, 51-63.

Pollock, J. (1995). *Cognitive Carpentry*, Cambridge, Mass., MIT Press.

Prakken, H. (2003). Logical Dialectics: The Missing Link Between Deductivism and Pragmatic Dialectics, *Proceedings of the Fifth Conference of the International Society for the Study of Argumentation*, ed. Frans H. van Eemeren et al., Amsterdam: SicSat, 857-860.

Prakken, H. (2010). On the Nature of Argument Schemes, *Dialectics, Dialogue and Argumentation*, ed. C. Reed and C. W. Tindale, London, College Publications.

Toulmin, S. (1952). *The Uses of Argument*, Cambridge, Cambridge University Press.

Verheij, B. (2009). The Toulmin Argument Model in Artificial Intelligence, *Argumentation in Artificial Intelligence*, ed. I. Rahwan and G. Simari, Berlin: Springer, 2009, 219-238.

Walton, D. (1997). *Appeal to Expert Opinion*, University Park, Pa: Penn State Press.

Walton, D. (2006). Argument from Appearance: A New Argumentation Scheme, *Logique et Analyse*, 195, 2006, 319-340.

Walton, D. and Godden, D. M. (2005). The Nature and Status of Critical Questions in Argumentation Schemes, *The Uses of Argument: Proceedings of a Conference at McMaster University*, ed. D. Hitchcock, Hamilton, Ontario: OSSA, 476-484.

Walton, D. and Gordon, T. F. (2005). Critical Questions in Computational Models of Legal Argument. *Argumentation in Artificial Intelligence and Law*, IAAIL Workshop Series, ed. Dunne, P. E. and T. J. M. Bench-Capon, Nijmegen: Wolf Legal Publishers, 103-111.

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