

## Value-Based Practical Reasoning

Value-based Practical Reasoning, *From Knowledge Representation to Argumentation in AI, Law and Policy Making: a Festschrift in Honour of Trevor Bench-Capon*, ed. K. Atkinson, H. Prakken and A. Wyner. London: College Publications, 2013, 259-282.

Douglas Walton

*CRRAR*

Surely one of the most important contributions of Trevor Bench-Capon to computer science, and especially to artificial intelligence and law, and for that matter to argumentation theory and cognitive science generally, is the model of value-based practical reasoning he has given us. He introduced value-based argumentation frameworks (VAF's) in (Bench-Capon, 2002; Bench-Capon, 2002a; Bench-Capon, 2003). He then applied this framework to practical reasoning, yielding the argumentation scheme for value-based practical reasoning, in a series of papers with his colleagues Katie Atkinson and Peter McBurney. Greenwood was the unmarried name of Katie Atkinson. The first one appears to be (Greenwood, Bench-Capon and McBurney, 2003). Before 2003 there had been several books and numerous articles written on the subject of practical reasoning, and of course the notion originally derived from Aristotle.

The literature in philosophy was almost exclusively concentrated on what could be called instrumental practical reasoning, a notion that did not take values explicitly into account, even though some philosophers recognized the importance of values in practical reasoning. But the only argumentation model of practical reasoning previously put forward (Walton, 1992) concentrated exclusively on instrumental practical reasoning, the assumption being that the values-laden variant contained so many complexities of its own that getting a basic notion of practical reasoning without considering the role of values in it was the best place to start. Previous to 2002, there was simply no idea of how a value-based model of practical reasoning could be added on to the existing argumentation model. Much less was there any idea of how such a notion of value-based practical reasoning could be formally modeled or implemented computationally. So what Trevor and his colleagues have given us is a tool of no small importance for many fields. The significance of the potential applications and philosophical implications of it should not be taken lightly.

Having a value-based version of the argumentation scheme for practical reasoning will no doubt prove to be extremely useful, not only in the field of artificial intelligence and law, but in many applications of argumentation theory and methods, including law itself. But there are many questions that arise concerning the relationship between the new value-based version of practical reasoning and the old simpler model of practical reasoning that does not take values into account. Also, there are more critical questions matching the value-based version of the scheme, and therefore questions are raised generally about how both kinds of critical questions should be used in evaluating practical reasoning, and how they fit together. Another basic question is how the value-based version and the older and simpler version of the argumentation scheme for practical reasoning fit together. Is one a special instance of the other? Is the value-based version the more general scheme so that the other scheme fits into it as a special instance? Or is the other scheme the more general one, so that the value-based scheme is an extension of the more basic scheme?

There have long been questions about the argumentation scheme and critical questions for practical reasoning that can prove troublesome to those trying to apply this scheme to argumentation in natural language discourse. Some users of this scheme have found it difficult to

distinguish in particular cases between instances of practical reasoning and instances of argument from negative consequences. Argument from negative consequences is a particularly common type of argument, perhaps even the most common type of argument in political discourse, as recent studies at University of Windsor (Hansen and Walton, 2013) have suggested. But asking about side effects of a projected course of action that could be used to fulfill a goal in practical reasoning is in fact one of the critical questions matching the basic scheme for practical reasoning. Is argument from negative consequences a distinctive type of counterattacking argument in its own right, or is it better seen as a critical question matching the scheme for practical reasoning? Or can it fulfill both these roles?

Another question raised is that since there appear to be around sixteen critical questions matching the value-based scheme for practical reasoning, could there be some way to make this list of schemes more manageable? For example, perhaps critical questions could be classified into subcategories in a tree structure so that it would be easier for a user to keep them in mind. At any rate, these are the questions that will be explored in this paper. Finally, good answers are hard to come by quickly in this new and vital area of investigation in argumentation, but still, it is worth moving ahead by trying to provide answers to them since practical reasoning is so important as a foundational mode of reasoning in argumentation studies.

## 1. Simple Practical Reasoning and its Extensions

The simplest and most basic kind of practical inference that is readily familiar to all of us can be represented in the following scheme according to the account of it given in (Walton, 2006, 300-301). The first-person pronoun 'I' represents an agent. More correctly, it could be called a rational agent of the kind described by Woodridge (2000), an entity that has goals, some (though possibly incomplete) knowledge of its circumstances, and the capability of acting to alter those circumstances and to perceive (some of) the consequences of so acting.

Goal Premise: I have a goal, *G*.

Means Premise: Carrying out this action *A* is a means to realize *G*.

Conclusion: I ought (practically speaking) to carry out this action *A*.

This simplest form of practical reasoning can be represented by the decision-making structure in visualized in figure 1.

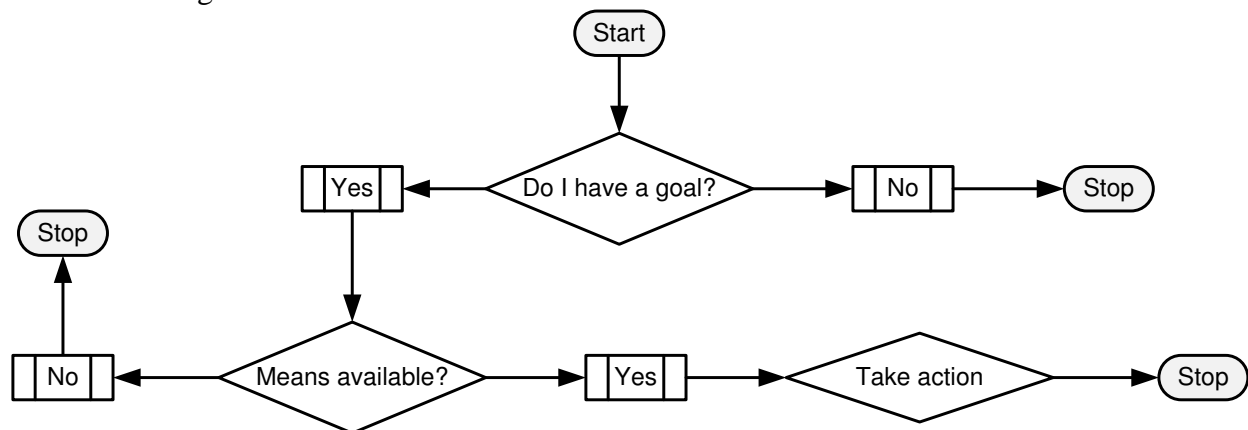


Figure 1: Simple Practical Reasoning

This form of practical inference is much too simple to represent realistic cases of the use of practical reasoning. An agent may have many goals, they may conflict, a chain of actions may be required, and there may be many ways to carry out a goal. Because practical reasoning is generally a defeasible form of argumentation, the basic evaluation procedure recommended in (Walton, 2006, 301) is to begin by asking one or more of a set of basic critical questions. Six of the most basic critical questions are listed below.

CQ1: Are there alternative actions available other than action *A*?

CQ2: Is *G* a possible (realistic) goal?

CQ3: Are there other goals that might conflict with *G*?

CQ4: Are there negative consequences of bringing about *A* that should be considered?

CQ5: Is *A* the best or most acceptable of the alternatives?

CQ6: Are intervening actions required to get from *A* to *G*?

To get some idea of how these critical questions work, let's start with a relatively simple example from (Walton, 1990, 89-90). Bob wants to balance his budget. He observes that the heating bill for his house has a large item in his budget. Let's suppose Bob already possesses a wood stove, a chainsaw and a trailer to carry wood. He is now in a position to take action, except he must have some source of wood to cut. Let's suppose he has a permit to cut wood in a woodland area designated by the government. So now, in order to carry out his goal, he hitches up his trailer to his car, and puts the chainsaw in the car. Next, he has to drive to the woodland area, which involves a lengthy collection of intervening actions, stopping at stop lights, putting his foot on the accelerator or the brake, and so forth. When he gets to the woodland area, he has to find some suitable trees, select one to start with, and then get his equipment unloaded and ready to start cutting a tree. He then has to go through a sequence of actions that would be familiar to him if he is very experienced with cutting trees with his chain saw.

If Bob's chain saw is new, or a rental unit that he is not familiar with, he might need to the read instructions given for starting his chain saw (Walton, 1997, 89).

1. Switch on the ignition (push the stop switch to the left so that the 1 is visible).
2. Pull out the choke control.
3. Push down the throttle safety catch.
4. Open the throttle fully.
5. Push the starting throttle ratchet backwards. Now all controls are in starting position and the chain saw is ready to be started.
6. Put your right foot on the plate beneath the rear handle.
7. Grasp the front handle with your left hand and press the saw against the ground.
8. Grasp the starter handle with your right hand and pull out the starting cord slowly until the starter pawls engage.
9. Give the starting cord a short sharp tug.

When analyzing what appears to be a simple teleological action like getting some wood for the fireplace, studying even what appears to be an obvious case that can be summed up as one action typically leads to the need to break it down into a sequence of actions that might be quite lengthy. We often don't think about these intervening actions because they are routine. We do them habitually without thinking about them. But in studying any realistic case, it can be seen that they involve some sequences of actions that are comparable to computer programs.

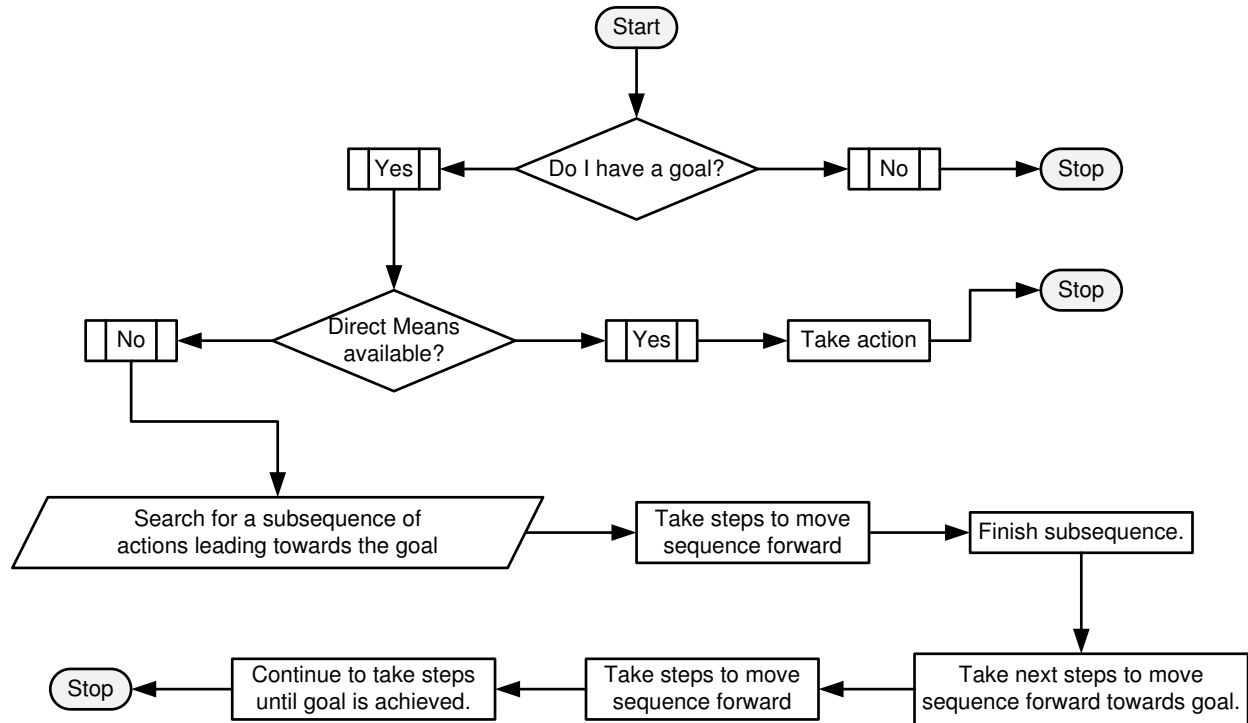


Figure 2: Practical Reasoning with Intervening Actions

There is a first step, and intervening sequence of steps, and finally a last step that represents the outcome. In such cases, the model of simple practical reasoning shown in figure 1 will not work. We have to make the model more complex by taking the need for intervening subsequences of actions into account. A rough attempt at a more adequate model is shown in figure 2.

But now what about the other critical questions? We need an even more complex model that not only takes subsequences of chained actions into account, but also accounts for the realistic possibilities that there may be a choice of several different actions each of which by itself represents a means to achieving the goal. In order to factor in these other critical questions, the model of practical reasoning as a sequence of steps in the thinking of a rational agent as it goes through thinking required to move towards its goal starts to get much more complex. Complex sequences of practical reasoning of these kinds have already been widely studied in the field of AI and planning. But here our aim is only to get a model of the schema for practical reasoning along with a set of matching critical questions.

## 2. The BDI and Commitment Models

The dominant approach is the belief-desire-intention (BDI) model, which uses ‘intention’ (or other variants like ‘want’ or ‘desire’) in place of the term ‘goal’ in the first premise. In this model the cognitive architecture of an agent is made up of a set of beliefs as well as a set of desires or intentions. In particular, an agent is said to have beliefs about its circumstances such that carrying out some particular action or set of action in those circumstances could be a means to realizing its intention. Clarke (1985, 17) posited a simple model of this sort in which  $E$  is a want or intention,  $M$  is a means (an action) that might potentially lead to the satisfaction of this want, and  $C$  is a set of circumstances of the agent’s choice situation.

I want *E*  
 My doing *M* is a means to attaining *E* if *C* obtains  
 Hence I should (ought to) do *M*

Clarke (1985, 3) cited the following example as fitting his version of the scheme.

I want to keep dry.  
 Taking this umbrella is a means of keeping dry if it rains [because it will rain].  
 Hence I should take this umbrella.

This version expressing the major premise as a want fits the BDI (belief-desire intention) model.

The BDI model has been widely adopted in cognitive science and computing, especially in work on artificial intelligence (Bratman, 1987; Bratman, Israel and Pollack, 1988; Audi, 1989; Pollock, 1995; Wooldridge, 2002; Paglieri and Castelfranchi, 2005,). On this model of how a cognitive agent reasons, the agent possesses a set of beliefs that are continually being updated by sensory input coming in from its environment as it moves forward taking actions, and a set of desires (wants) that form its intentions. On the BDI model, a distinction can be made between goals and intentions. There seems to be little firm agreement on how to draw this distinction precisely, but one way is that an agent's intentions can be defined as a subset of its goals - persistent goals that are stable over time and are not easily given up as the agent continues to act.

The BDI accounts which represent the dominant model of practical reasoning in philosophy appear to be directed to modeling the behavior of a person who is trying to decide what to do based on information about the circumstances that might be changing as the situation evolves. Hence the model is based on the notion of an agent taking action based on its beliefs and intentions, desires and wants. It is possible to take a different approach to practical reasoning, however, where the object is to model the reasoning of a robot that is built to act autonomously in accord with goals that have been programmed into it. Such a robot acts by continually moving through a set of circumstances recorded in its memory and perceived by its sensors as it navigates through the circumstances.

Such a rational agent would need to have several important characteristics. First obviously it needs to have goals and the capability to carry out actions that might realize those goals, or in some instances might steer away from the realization of those goals. It must have some knowledge contained in its memory representing information about its circumstances and about the external world generally. As it steers its way through the circumstances it must have the capability for feedback, requiring the capability to observe the consequences of its actions and the capability to steer them towards its goals. It must have the capability to mesh abstract goals with particular actions, and hence it needs to have some resources for dealing with abstraction. It must have the capability to look forward into the future by constructing hypotheses about possible future consequences of its actions. It must have the capability to comparatively examine different possible alternative courses of action to determine which course might better steer towards fulfillment of its goals, or might lead to entrance of fulfillment of its goals. It must have a memory that has the capability of keeping track of its actions so that it can organize them, and so it can tell which actions have had certain consequences in the past.

These are the basic characteristics it must have, but at a higher level it might have four other characteristics, if it is to be an especially intelligent agent. One is plasticity, meaning the

characteristic to act in a flexible way moving forward and adapting to information about new circumstances. A second one is the characteristic of persistence, referring to trying an alternative course of action if one action being considered turns out to be impossible or not useful towards reaching a goal. A third one is planning capability, built on the agent's capability of constructing hypotheses about possible future consequences of its actions and its capability of estimating the most likely consequences of its actions in normal situations of the kind it is familiar with, or has had some experience with. A fourth one is criticism, referring to the ability to see defects in its own action plans, including inconsistencies and practical difficulties.

An agent of this sort can be programmed with goals in the form of propositions that can be more concrete or more abstract. For example if my goal is health, that goal is quite abstract. It may be nontrivial to operationalize it in a specific case where I have to make a decision about my health on the basis of what I know or don't know. Or my goal might be to lose weight. This is a more specific goal, but still quite general. It might be much more specific. For example my goal may be to lose ten pounds within the next six months. A goal in this sense of the word can be represented as a proposition that is inserted into the memory of the agent to represent the commitment that the agent has adopted. For example, if an agent says it is his goal to lose ten pounds within the next six months, or that he intends to do so, or uses some comparable expression of the sort, it can be inferred that he has taken on a commitment to do whatever he can to lose ten pounds within the next six months. Presumably, in the case of a nonhuman agent like a robot, it has a goal because it has been programmed to act in such a way as to try to bring about some proposition in its commitment store if it has the resources to do so, and if the possibility of doing so becomes possible in the circumstances that it confronts.

A goal in this sense does not have to be an intention, a want or desire of the agent. It is often described that way, but such a description is somewhat misleading, or even anthropomorphic. A goal can be understood as a state of the world that an agent is aiming at. This world state can be represented in the agent's knowledge base (or in its commitment set, in an argumentation model) as one or more propositions. To call this (or these) propositions a goal for some agent is simply to say that the agent is committed to bringing about this world state, or to arrange for its bringing about, and acts accordingly. Typically, an agent will maintain this objective for an extended period of time, which is to say that the agent is committed to the achievement of the world state.

We tend to think of a rational agent of this sort as a solitary decision maker who looks at both sides of a choice and examines the pros and cons of each course of action it contemplates undertaking. But in a multiagent context, such a decision maker would engage in group deliberations in cases where the group shares common goals, to some degree, common knowledge of a set of circumstances in which a choice that affects all of them has to be made, and can put forward and criticize proposals for action (Atkinson, Bench-Capon and McBurney, 2004). The basic structure of a single agent, or the way a group of agents collaborates in deliberation to make a decision on action, is basically the same in that in both instances the argumentation and the outcome of the decision-making procedure is based on practical reasoning.

The problem with using the concept of intention as the primitive constant in the major premise of the scheme for practical reasoning is the difficulty of knowing what the intention of an agent really is in a given case. We can all easily recognize that it is extremely difficult very often in everyday life to try and figure out what the intention of another agent really is. Indeed, one might add that it may involve considerable speculative guesswork to try to figure out even what one's own intention is in some situation where one is making a decision to act, or has made

such a decision. Both intentions and beliefs, the other key primitive notion in the BDI model, are private psychological notions internal to an agent, while commitments are statements externally and verbally accepted by an agent in a communicative context. The BDI model is clearly necessary and important in cognitive science in the fields of psychology and law, because the notions of motive and intention are fundamentally important in these fields. However, precisely because of the richness of these notions, and the difficulty of determining them in particular cases, it introduces of complexity in using them in the simplest form of the argumentation scheme for practical reasoning.

An alternative approach is to take goals as statements that the agent has expressed or formulated in a public way, based on what he has said and committed himself to, either through as speech or actions. On this approach, a goal is simply a proposition stated or accepted by an agent. On the commitment model of argumentation, there is a dialogue structure in which each contributes speech acts and each party has a commitment set containing the propositions it has accepted, judging by its speech acts in the previous dialogue (Walton and Krabbe, 1995). In such a formal dialogue model, as each move is made, propositions are inserted into or retracted from an agent's commitment store according to protocols (rules governing the moves) (Hamblin, 1970; 1971). The commitment model represents a weaker logic-based approach, as opposed to the richer psychological approach of the BDI model. The two models are certainly related to each other, and one approach to investigating both of them is to begin with the commitment model, which is simpler and more open to logical analysis, and then move from there to the investigation of the BDI model. On the commitment model, practical reasoning is modeled in a dialogue format which utilizes the device of a set of critical questions matching an argumentation scheme in order to evaluate the application of the scheme to a case.

### 3. Argument from Consequences

Suppose that Bob has now gotten his chain saw started, and has selected the first tree he wants to cut down. He steps towards the tree, but then realizes that if he makes the cut from the front, the way he originally planned, there is a very good chance that the tree will fall back towards him and cause him serious injury. Therefore, using practical reasoning, he moves around to a different angle and makes the cut in such a way that the tree will fall away from him. Here Bob has used argument from negative consequences as follows.

If I cut the tree from the front, it will have consequences for me.  
 These consequences are highly negative.  
 Therefore I will not cut the tree from the front.

Because of these negative consequences, Bob searches around for an alternative course of action that might lead to his goal of cutting down a tree without the negative consequences of injuring him. He sees that if he cuts the tree from a different angle, it will fall away from him, and so the cutting of the tree will not have these negative consequences. So he chooses the alternative course of action. This conclusion assumes however that the alternative course of action does not also have some negative consequences, especially ones that might be even worse than the original set of negative consequences.

To model this kind of reasoning, we have to construct a decision-making flowchart that takes negative consequences of a proposed course of action into account. This model needs to be

capable of considering the question of whether the negative consequences are bad enough that they outweigh the goal. In other words, the negative consequences have to be weighed against the goal. This process of weighing the negative value of the consequences against the positive value of the goal might suggest that values are now involved. In other words, it suggests that we are now already doing value-based practical reasoning and not just simple instrumental practical reasoning of the kind that does not involve values. This observation exposes a general problem. Could it be that because the side effects critical question entails classifying consequences as good or bad, values are inextricably bound up with even the simpler forms of practical reasoning. This in turn suggests that perhaps there is no such thing as simple instrumental practical reasoning, because all practical reasoning involves values. Here we have a central problem, assuming we want to draw a distinction between two kinds of practical reasoning, instrumental practical reasoning versus value-based practical reasoning.

Practical reasoning is commonly used in advertising, for example in health product ads (Walton, 2010). Typical examples cite positive consequences of taking the medication, but will also offer a long list of side effects. For example an ad for the pain reliever medication Cymbalta (*Newsweek*, December 31, 2012, 12-14) tells us that Cymbalta can help to significantly reduce chronic low back pain. This good consequence can be seen as representing the goal of someone who takes this medication. However, the ad also states a long list of side effects included under the heading of important safety information. This list includes itching, upper belly pain, headache, weakness or feeling unsteady, confusion, problems in concentrating, high fever, confusion, stiff muscles, muscle twitching, racing heart rate, abnormal bleeding, serious possibly life-threatening skin reactions, abnormal mood (mania), racing thoughts, talking more or faster than usual, reckless behaviour, and seizures or convulsions. It may be presumed that all of these side effects may be taken to be bad consequences. They may be taken to be reasons to at least consider not taking the drug, or to stop taking it if the side effects are so bad as to be intolerable.

Each side effect in the list may be taken as an instance of the side effects critical question. But should any of these side effects actually be experienced by the person taking the drug, the critical question is transformed into an argument from negative consequences. Now it is an argument against continuing to take the drug. However, the person will have to weigh the negative consequences against the positive benefits of the drug in helping with back pain. It is a question of balancing how bad the back pain is against how bad the side effects are.

#### 4. Value-based Practical Reasoning

Atkinson, Bench-Capon and McBurney, (2006, 166) - and before that in Atkinson, Bench-Capon and McBurney 2003) - formulated their model of value-based practical reasoning said to be an extension of Walton's instrumental practical reasoning scheme as follows.

In the circumstances  $R$   
 we should perform action  $A$   
 to achieve new circumstances  $S$   
 which will realize some goal  $G$   
 which will promote some value  $V$



The first version of this scheme that I found was in (Greenwood, Bench-Capon and McBurney, 2003, 22). Matching this scheme Atkinson, Bench-Capon and McBurney (2006, 165-167) cited ten critical questions 165-167, labeled as follows.

- CQ1a: Are there alternative ways of realizing the same consequences?
- CQ1b: Are there alternative ways of realizing the same goal?
- CQ1c: Are there alternative ways of promoting the same value?
- CQ3a: Would doing action *A* promote some other value?
- CQ3b: Does doing action *A* preclude some other action which would promote some other value?
- CQ4a: Does doing action *A* have a side effect which demotes the value *V*?
- CQ4b: Does doing action *A* have a side effect which demotes some other value?
- CQ5: Are the circumstances such that doing action *A* will bring about goal *G*?
- CQ6: Does goal *G* promote value *V*?
- CQ7: Is goal *G* possible?

In this list, all the critical questions concern values, except for CQ1a, CQ1b, CQ5 and CQ7.

In addition to this main list of ten critical questions, Atkinson, Bench-Capon and McBurney (2006) stated that some of these critical questions need to address eight other issues. First they state (166-167) that that an answer to CQ5 needs to address four issues:

- (a) Whether the believed circumstances *R* are possible.
- (b) Whether the believed circumstances *R* are true.
- (c) Assuming both of these, whether the action *A* has the stated consequences *S*.
- (d) Assuming all of these, whether the action *A* will bring about the desired goal *G*.

Second, they state (167) CQ6 needs to address two issues:

- (a) Whether goal *G* does realize the value intended; and
- (b) Whether the value proposed is indeed a legitimate value.

Third, they state (167) that CQ7 needs to address two additional issues:

- (a) Whether the situation *S* believed by agent *a* to result from doing action *A* is a possible state of affairs.
- (b) Whether the particular aspects of situation *S* represented by *G* are possible.

These issues can be reformulated as questions, and indeed they do seem to represent additional critical questions that can be subsumes under the original seven. This extended reformulation would give us a grand total of eighteen critical questions. Only the issues in the second group concern values. The comments of Atkinson, Bench-Capon and McBurney suggest that subsets of the eight issues identified above can be fitted under the concerns of specific critical questions. In light of these matters, it is interesting to compare the later list of sixteen critical questions offered where it is said that Walton's original four critical questions associated with his (instrumental) scheme have been extended to address the elements in the value-based scheme, producing an expanded list with eighteen critical questions. In this new list of critical question (Atkinson and Bench-Capon, 2008, 141), has been reduced to the following sixteen.

- CQ1: Are the believed circumstances true?
- CQ2: Assuming the circumstances, does the action have the stated consequences?
- CQ3: Assuming the circumstances and that the action has the stated consequences, will the action bring about the desired goal?

- CQ4: Does the goal realise the value stated?
- CQ5: Are there alternative ways of realising the same consequences?
- CQ6: Are there alternative ways of realising the same goal?
- CQ7: Are there alternative ways of promoting the same value?
- CQ8: Does doing the action have a side effect which demotes the value?
- CQ9: Does doing the action have a side effect which demotes some other value?
- CQ10: Does doing the action promote some other value?
- CQ11: Does doing the action preclude some other action which would promote some other value?
- CQ12: Are the circumstances as described possible?
- CQ13: Is the action possible?
- CQ14: Are the consequences as described possible?
- CQ15: Can the desired goal be realised?
- CQ16: Is the value indeed a legitimate value?

The two that appear to be missing from the original list are these ones: (1) Are the circumstances such that doing action *A* will bring about goal *G*? (2) Are the particular aspects of situation *S* represented by *G* possible?

Citing Atkinson, Bench-Capon and McBurney, (2004), the technical report (Atkinson, Bench-Capon and McBurney, 2005a, 3) tells us that there are sixteen critical questions matching the scheme for value-based practical reasoning. The latter source is especially interesting for our purposes in that it proposes (2005a, 3) that the list of critical questions can be subdivided into three distinct categories relating to the nature of the attack: “issues relating to the belief as to what is the case; issues relating to desires as to what should be the case; and issues relating to representation concerning the language being used and the logic being deployed in the argument”.

However, another way of classifying the sixteen critical questions that seems to me easier to identify and apply is set out below, classifying them into seven categories.

1. Belief about circumstances: CQ1, CQ12
2. Consequences: CQ2, CQ3, CQ8, CQ9, CQ14
3. The action being considered (not just its consequences): CQ 10, CQ11, CQ13
4. Goal: CQ4, CQ15,
5. Alternatives: CQ5, CQ6, CQ7
6. CQ16 is only about values.
7. All are about values except CQ1, CQ2, CQ3, CQ5, CQ6, CQ12, CQ13, CQ14.

What would be helpful is to divide the critical questions into categories so that we could have made critical questions and then sub-questions under these main critical questions. That way we could have a tree structure of critical questions where the user starts with the basic critical questions and then probes into the given argument in a more detailed way by asking further sub questions. Of course the sub questions can also have some questions, and so the problem of the completeness of practical reasoning as a defeasible form of argumentation is posed. Basically, the answer to this question is that practical reasoning is inherently defeasible when employed under conditions of changing circumstances and incomplete information about all the circumstances. The closure problem can only be solved by viewing practical reasoning is a form

of argumentation set into a dialogue structure with an opening and a closing stage. This is not an impossible project to carry out because Atkinson, Bench-Capon and McBurney view practical reasoning this way as a form of argument that needs to be set into a framework of dialogue.

So what I have proposed above as a way of classifying the critical questions, based on the comments of Atkinson Bench-Capon and McBurney, might serve as a beginning point for such a classification project. However, the problem I have is to make sense of some of the critical questions in the list of sixteen. For example consider critical questions CQ3, which asks whether the action will bring about the desired goal? What does this question made? Does it mean that the action that is being considered will be sufficient all by itself to carry out the goal? Or does it mean that the action, taken along with a lot of other sub-actions along the way in a sequence of the kind described above in the chain saw example, will bring about the desired goal? Or does it mean that the action, although it may not be sufficient by itself, is a necessary condition for bringing about the desired goal that can be sufficient in the circumstances, perhaps taken along with some other actions that are also required?

## 5. Critical Questions and Premises

In a book published fifteen years ago (Walton, 1990) I set out an argumentation scheme for a form of goal-directed, knowledge-based reasoning that concludes in an action, called practical reasoning. In this analysis, two argumentation schemes for practical reasoning were postulated: a necessary condition scheme and a sufficient condition scheme. Here is the necessary condition scheme (Walton, 1990, p. 48).

(N1) My goal is to bring about A (*Goal Premise*).

(N2) I reasonably consider on the given information that bringing about at least one of  $[B_0, B_1, \dots, B_n]$  is necessary to bring about A (*Alternatives Premise*).

(N3) I have selected one member  $B_i$  as an acceptable, or as the most acceptable necessary condition for A (*Selection Premise*).

(N4) Nothing unchangeable prevents me from bringing about  $B_i$  as far as I know (*Practicality Premise*).

(N5) Bringing about A is more acceptable to me than not bringing about  $B_i$  (*Side Effects Premise*).

Therefore, it is required that I bring about  $B_i$  (*Conclusion*).

The sufficient condition scheme is the same (1990), except that in its premises (N2) and (N3), the expression ‘sufficient condition’ must be substituted for ‘necessary condition’. This way of framing the scheme is interesting because it takes the approach of incorporating the critical questions into the scheme itself by treating them as premises (in some cases additional premises).

In a table Atkinson, Bench-Capon and McBurney, (2006, 174) reconfigure the critical questions for their scheme for value-based practical reasoning as set of sixteen attacks representing what they call “the nature of conflict for critical questions”. The reader can also see by scanning over the list of these attacks given below how each attack is said to be associated with a specific critical question. The label of the critical question is given to the right of each attack.

1. Disagree with the description of the current situation CQ5b
2. Disagree with the consequences of the proposed action CQ5c

3. Disagree that the desired features are part of the consequences CQ5d
4. Disagree that these features promote the desired value CQ6a
5. Believe that the consequences can be realized by some alternative action CQ1a
6. Believe that the desired features can be realized through some alternative action CQ1b
7. Believe that the desired value can be realized in an alternative way CQ1c
8. Believe that the action has undesirable side effects which demote the desired value CQ4a
9. Believe that the action has undesirable side effects which demote some other desired value CQ4b
10. Agree that the action should be performed, but for different reasons CQ3a
11. Believe that the action will preclude some more desirable action CQ3b
12. Believe that the action is impossible CQ2
13. Believe that the circumstances as described are not possible CQ5a
14. Believe that the consequences as described are not possible CQ7a
15. Believe that the desired features cannot be realized CQ7b
16. Disagree that the desired value is worth promoting CQ6b

The association of the prior list of critical questions with a list of what are taken to be typical types of argument attacks on practical reasoning is interesting from the point of view of argumentation theory for a number of reasons. First, it raises the question of whether each of the critical questions could be reconfigured as a positive assertion or statement of some sort that expresses the content of the critical question. This issue of whether the questions can be reformulated as statements has been discussed in the literature on argumentation (Walton and Godden, 2005). Second, it presents quite an expanded list of critical questions matching scheme for practical reasoning, one might almost say a proliferation of critical questions. Third, it suggests the possibility that some of the critical questions matching the scheme for practical reasoning could be classified under others, suggesting a way of organizing the critical questions in some way that might prove useful to users of the scheme.

This approach has now been systematically worked out in the Carneades Argumentation System, which manages critical questions matching a scheme by treating them as premises of the scheme and classifying them as assumptions or exceptions (Gordon, Prakken and Walton, 2007). The ordinary premises are the ones explicitly stated in the argumentation scheme are classified as assumptions. But there are also assumptions in the form of additional premises that are assumed to hold, just like the ordinary premises, but if questioned automatically fail to hold unless the proponent of the argument gives some evidence to support the premise. Exceptions are also additional premises, but they are assumed not to hold as exceptions unless evidence to back them up is given by the critical questioner. They do not defeat the argument unless the questioner gives backup evidence to support the question.

The following entry for ‘argument from practical reasoning’, citing (Atkinson, Bench-Capon and McBurney, 2007) as the source, can be found in the catalogue of schemes in the Carneades Argumentation System (<http://carneades.github.com/>).

### **Argument from Practical Reasoning**

**id:** practical-reasoning

**strict:** false

**direction:** pro

**conclusion:** Action *A* should be performed.

**premises:**

- *S1* is the case in the current circumstances.
- Performing *A* in *S1* would bring about *S2*.
- *G* would be realized in *S2*.
- Achieving the goal *G* would promote the value *V*.

**assumptions:**

- *V* is a legitimate value.
- *G* is a worthy goal.
- Action *A* is possible.

**exceptions:**

- There exists an action that would bring about *S1* more effectively than *A*.
- There exists an action that would realize the goal *G* more effectively than *A*.
- There exists an action that would promote the value *V* more effectively than *A*.
- Performing *A* in *S1* would have side-effects which demote *V* or some other value.

What are called ‘premises’ in the list above are the so-called ordinary premises of the scheme for argument from practical reasoning.

Now the question is raised how the critical questions proposed by Atkinson, Bench-Capon and McBurney can be classified as assumptions or exceptions in order to deal with the problem of burden of proof posed by them. Some recommendations on this and related questions of how to configure the scheme for practical reasoning and its set of matching critical questions are given in the conclusions section.

## 6. Conclusions and Questions for Further Work

One problem with the current version of this scheme is that it may not take into account the kind of situation where there are conflicting goals for the agent. Consider the military planning example from Walton 1990, 64. In this case, a soldier wants to locate the enemy position, but in order for him to do that it is necessary for him to stand on a hill and make visual observations. However, if he does this he may reveal the location of his unit the enemy. So he has two goals. One is to locate the enemy position, while the other is not to reveal the location of this unit the enemy. The problem is that these two goals are in conflict.

Of course the purpose of value-based practical reasoning is precisely to resolve such goal conflicts by means of an ordering of values using value priorities. But even before a practical reasoning agent gets to this step, it may be important for it to recognize a conflict between two of its goals as a basis for proceeding further. Therefore, it was argued in Walton 1990 that one of the critical questions matching this scheme for practical reasoning should be the question of whether there is a conflict of goals that can be identified. At very least, the critical question should ask whether there are other goals that the agent should take into consideration, in addition to the goal *G* that is the primary goal being considered in the argumentation scheme. The recommendation is to add the premise ‘there is another goal *G*’ that is incompatible with *G* as an additional exception to the current list.

But also there is another critical question to be considered. This is the question of whether the goal *G* is possible, one of the questions that was taken into account in the version of the value-

based scheme put forward by Atkinson, Bench-Capon and McBurney. The recommendation suggested is that the premise ‘ $G$  is possible’ should be added under the list of exceptions.

Finally, there is a third additional critical question that ought to be taken into account. This is the question concerning intervening actions. As shown by the chain saw example, typically working towards a goal involves carrying out not just one action but a sequence of actions where each one is required (often a particular order) to be performed in order to move toward the goal. Thus is to ask whether there are intervening actions required to move from the action  $A$  to the goal  $G$ .

To deal with these issues, the following amended version of the Carneades scheme is put forward as the appropriate scheme for value-based reasoning practical reasoning with its matching set of critical questions is proposed.

### **Argument from Value-based Practical Reasoning**

**id:** practical-reasoning

**strict:** false

**direction:** pro

**conclusion:** Action  $A$  should be performed.

**premises:**

- $S1$  is the case in the current circumstances.
- Performing  $A$  in  $S1$  would bring about  $S2$ .
- $G$  would be realized in  $S2$ .
- Achieving the goal  $G$  would promote the value  $V$ .

**assumptions:**

- $V$  is a legitimate value.
- $G$  is possible.
- $G$  is a worthy goal.
- Action  $A$  is possible.

**exceptions:**

- There exists an action that would bring about  $S1$  more effectively than  $A$ .
- There exists an action that would realize the goal  $G$  more effectively than  $A$ .
- There are intervening actions required to move from the action  $A$  to the goal  $G$ .
- There exists an action that would promote the value  $V$  more effectively than  $A$ .
- Performing  $A$  in  $S1$  would have side-effects which demote  $V$  or some other value.
- There is another goal  $G'$  that is incompatible with  $G$ .

However, in addition, a scheme for Instrumental Practical Reasoning that contains no mention of values is put forward for consideration.

### **Argument from Instrumental Practical Reasoning**

**id:** practical-reasoning

**strict:** false

**direction:** pro

**conclusion:** Action  $A$  should be performed.

**premises:**

- *SI* is the case in the current circumstances.
- Performing *A* in *S1* would bring about *S2*.
- *G* would be realized in *S2*.

**assumptions:**

- *G* is possible.
- Action *A* is possible.

**exceptions:**

- There exists an action that would bring about *SI* more effectively than *A*.
- There exists an action that would realize the goal *G* more effectively than *A*.
- There are intervening actions required to move from the action *A* to the goal *G*.
- Performing *A* in *SI* would have side-effects that need to be taken into account.
- There is another goal *G'* that is incompatible with *G*.

The scheme for value-based practical reasoning, on this approach, is an extension of the scheme for instrumental practical reasoning. This is a useful way of designing a theory of practical reasoning as a form of argument that can be applied to real cases of argumentation because the simpler purely instrumental version of the scheme can be applied when values are not at issue.

There is one complication however. The side effects critical question is usually expressed as asking whether there are negative side consequences of the action *A*. The example of the side effect of the Cymbalta medication discussed in section 3 is the classic kind of case so often encountered. The question is whether the potential negative consequences of the medication are “negative” (bad) because they have a negative value for the agent. If so, then the side effects critical question cannot be purely instrumental in nature. It has to be a value-based consideration. If so, then all practical reasoning has to be value-based practical reasoning. How to separate out these schemes and distinguish between them in a useful manner also involved some related schemes.

Argument from negative consequences cites the consequences of a proposed course of action as a reason against taking that action. The following scheme is from (Walton, 1996, 75).

Premise: If *A* is brought about, negative consequences will plausibly occur.

Conclusion: *A* should not be brought about.

Of course there is also a matching scheme for argument from positive consequences.

Premise: If *A* is brought about, positive consequences will plausibly occur.

Conclusion: *A* should be brought about.

According to (Walton, 1996, pp. 76-77), there are three critical questions match either scheme.

CQ1. How strong is the probability or plausibility that these cited consequences will (may, might, must) occur?

CQ2. What evidence, if any, supported the claim that these consequences will (may, might, must) occur if *A* is brought about?

CQ3. Are there consequences of the opposite value that ought to be taken into account?

Both schemes appear to be species of value-based argumentation. But are they really? Bench-Capon takes as values abstract notions like ‘health’, ‘security’, ‘patriotism’, and so forth that support goals. But in arguments from positive and negative consequences, the term ‘values’ appears to be used in a different way, closer to what Stevenson (1944) called positive and negative values (Macagno and Walton, 2013).

It also needs to be taken into account that there are schemes for arguments from positive and negative values (Walton, Reed and Macagno, 2008, 321). In these schemes, a judgment of positive value is taken as a reason supporting commitment to a goal, while a judgment of negative value is taken as a reason going against commitment to a goal. So these schemes have to be factored in.

So future work on classification of argumentation schemes needs to much more carefully study how this cluster of schemes fits together so that some of them can be classified as subspecies of others. Until this work of scheme classification fits these schemes together into a unified system, the conclusions drawn here have to be tentative.

## References

- Atkinson, K. and Bench-Capon, T. (2007). Practical Reasoning as Presumptive Argumentation Using Action Based Alternating Transition Systems. *Artificial Intelligence* 171, 855–874.
- Atkinson, K. and Bench-Capon, T. (2008). Addressing Moral Problems through Practical Reasoning, *Journal of Applied Logic*, 6, 135-151.
- Atkinson, K., Bench-Capon, T. and McBurney, P., (2004). PARMENIDES: Facilitating Democratic Debate, *Electronic Government*, ed. R. Traummüller, Lecture Notes in Computer Science (LNCS), 3183.
- Atkinson, K., Bench-Capon, T. and McBurney, P. (2005). Agent Decision Making Using Argumentation About Actions, *Technical Report*, University of Liverpool.
- Atkinson, K., Bench-Capon, T. and McBurney, P. (2006). Computational Representation of Practical Argument. *Synthese*, 152 (2), 157-206.
- Audi, R. (1989). *Practical Reasoning*. London: Routledge.
- Bench-Capon, T. (2002). The Missing Link Revisited: The Role of Teleology in Representing Legal Argument, *Artificial Intelligence and Law*, 10 (2-3), 79-94.
- Bench-Capon, T. (2002a). Agreeing to Differ: Modelling Persuasive Dialogue Between Parties Without a Consensus About Values. *Informal Logic*, 22 (3), 231-45.
- Bench-Capon, T. (2003). Persuasion in Practical Argument Using Value-based Argumentation Frameworks, *Journal of Logic and Computation*, 13, 429-448.
- Bratman, M. (1987). *Intentions, Plans, and Practical Reason*, Cambridge, Mass., Harvard University Press.



- Bratman, M., Israel, D. and Pollack, M (1988). Plans and Resource-bounded Practical Reasoning, *Comput. Intelligence*, 4, 349-355.
- Clarke, D. S. (1985). *Practical Inferences*. London: Routledge.
- Gordon, T., Prakken, H. and Walton, D. (2007). The Carneades Model of Argument and Burden of Proof. *Artificial Intelligence* 171: 875-96.
- Greenwood, K., Bench-Capon, T. and McBurney, P. (2003). Towards a Computational Account of Persuasion in Law. In G. Sartor (editor): *Proceedings of the Ninth International Conference on AI and Law* (ICAIL 2003). Edinburgh: ACM Press, 22-31.
- Hamblin, C. L. (1970). *Fallacies*. London: Methuen, 1970.
- Hamblin, C. L. (1971). Mathematical Models of Dialogue. *Theoria* 37, 130-155.
- Hansen, H. and Walton, D. (2013). Kinds of Arguments Used in the Ontario Provincial Election, 2011, *Argumentation in Context*, 2(2), 226-258.
- Paglieri, F. and Castelfranchi, C. (2005). Arguments as Belief Structures, *The Uses of Argument: Proceedings of a Conference at McMaster University*, ed. D. Hitchcock and D. Farr, Ontario Society for the Study of Argumentation, 356-367.
- Macagno, F. and Walton, D. (2013). *Emotive Language in Argumentation*. Cambridge: Cambridge University Press, 2013.
- Pollock, J. L. (1995). *Cognitive Carpentry*. The MIT Press: Cambridge, MA.
- Stevenson, C. L. (1944). *Ethics and Language*. New Haven: Yale University Press.
- Walton, D. (1990). *Practical Reasoning: Goal-Driven, Knowledge-Based, Action-Guiding Argumentation*. Savage, Maryland: Rowman & Littlefield.
- Walton, D. (1996). *Argumentation Schemes for Presumptive Reasoning*. Mahwah, New Jersey.
- Walton, D. (1998). *The New Dialectic*. Toronto: University of Toronto Press.
- Walton, D. (2006). *Fundamentals of Critical Argumentation*. Cambridge: Cambridge University Press.
- Walton, D. (2007). Evaluating Practical Reasoning, *Synthese*, 157, 197-240.
- Walton, D. (2010). The Structure of Argumentation in Health Product Messages, *Argument and Computation*, 1(3), 179-198.

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

Walton, D. and Godden, D. (2005). The Nature and Status of Critical Questions in Argumentation Schemes, *The Uses of Argument: Proceedings of a Conference at McMaster University, 18-21 May 2005*, ed. D. Hitchcock. Hamilton, Ontario: OSSA, 476-484.

Walton, D. and Krabbe, E. (1995). *Commitment in Dialogue*. Albany: SUNY Press.

Walton, D. and Macagno, F. (2005). *Argumentation Schemes*. Cambridge: Cambridge University Press.

Wooldridge, M. (2002). *An Introduction to MultiAgent Systems*. Chichester: Wiley.