May 15th, 9:00 AM - May 17th, 5:00 PM

Layered protocols in coalescent argumentation

Allan Randall
York University

Title: Layered Protocols in Coalescent Argumentation
Author: Allan F. Randall
Response to this paper by: Leo Groarke

---

1. Introduction

Gilbert (1997) has argued for an approach to argumentation theory in which the argument itself, *as* an argument, is not considered the sole object of study, but merely one piece of a larger context that includes much more than just what the argument is superficially about. Argumentation theory, he argues, must recognize the essentially goal-oriented nature of argument. In order to understand the argument itself, we must bring into the picture the whole background of beliefs and goals of the participants as crucial to what the argument is. To understand the argument *as* an argument requires that *more than* the argument be taken into consideration.

I will look at a simple (and purely hypothetical) example argument, in order to see what deeper understanding we can arrive at through hierarchical goal analysis. I will use Taylor's theory of hierarchical goals in communication, which he calls "layered protocols" (Taylor 1989; Farrell et al. 1999), as it is firmly rooted in basic principles of biology, in this case perceptual control theory (McFarland *et al.* 1957; Powers 1973; Taylor 1999; CSG). I will put forth some preliminary ideas towards the development of practical argumentation therapy, in which participants learn to argue better, based on Gilbert's notion of coalescent argumentation (Gilbert 1997).

In sections 2-7, I will give a brief summary of perceptual control theory. In sections 8-11, I will outline the theory of layered protocols (which is just perceptual control theory applied to communication). In sections 12-14, I will explain how layered protocols relate to argumentation in particular, and in sections 15-23, I will explain how this theory might be used in coalescent argumentation therapy. I will conclude with some general remarks and discussion of future work.

2. Perceptual Control Theory: Behaviour is the Control of Perception

To understand argument, one must understand communication in general, and to understand communication, one must view it as an activity of living organisms. So any theory of argument must be built on a proper theory of what a living organism is, and feedback control theory is the closest thing we currently have to a generally accepted foundation for biology. Taylor builds his theory of layered protocols (Taylor 1989; Farrell et al. 1999) on the control theory of Powers, a variation on standard control theory called "perceptual

control theory", or just "PCT" (McFarland *et al.* 1957; Powers 1973).

PCT, unlike standard control theory, does not presume that an organism is controlling something "out there" in the external world—instead, the organism controls only internal perceptions. The standard notion of feedback control, arising out of the work of Wiener (1948), views a living organism as a system that controls certain variables in its environment, so that when a cat catches a mouse, it is controlling the variable "mouse position", among others. Likewise, when a human drives a car down a road, they are controlling the position of the car relative to the centre of the road. Feedback control theory sees most of the complex behaviours of living organisms as arising out of complex organisations of control systems controlling various "environmental variables" that are important for the organism's survival.

These ideas have been enormously useful, but Powers, working on the ideas of Ashby (1956), saw that the standard formulation had seriously misplaced the role of perception. The basic idea of perceptual control dates all the way back to Aristotle (*c*. 335-322 BC), who broadly speaking could be called the father of PCT. Powers, who is generally considered the father of the modern, more technical version of the theory, follows a similar line of reasoning to that of his ancient Greek forefather, recognizing that the variable under control cannot be an external variable "out there" in the world, but must be an internal variable, namely a perception. (Some psychologists restrict the word "perception" to mental states that correspond in a straightforward and obvious way to direct, conscious sensory input, but the PCT community uses the word in a very general sense that covers just about any, if not all, mental states, whether conscious or not.)

Powers' system, like Aristotle's, is hierarchical. Perceptions are built up in a layered fashion, on top of other perceptions, which are built on top of other perceptions, etc., so we each perceive the world through a very different, complex system of input filters that produce in each of us entirely different perceptual variables. We cannot necessarily even perceive the variables that another person perceives and controls (even when that person perceives them perhaps quite effortlessly). We each divide the world up in our own way, filtering it and analyzing it in ways that can potentially seem well nigh unintelligible to another. Thus, another person's behaviour sometimes makes no sense at all to us when we try to find the external object out there in the world that they are controlling, that towards which their thought is directed. This is because each of us has analyzed the world into *different* world objects, or "percepts" or "complex environmental variables", as they are called in the PCT community.

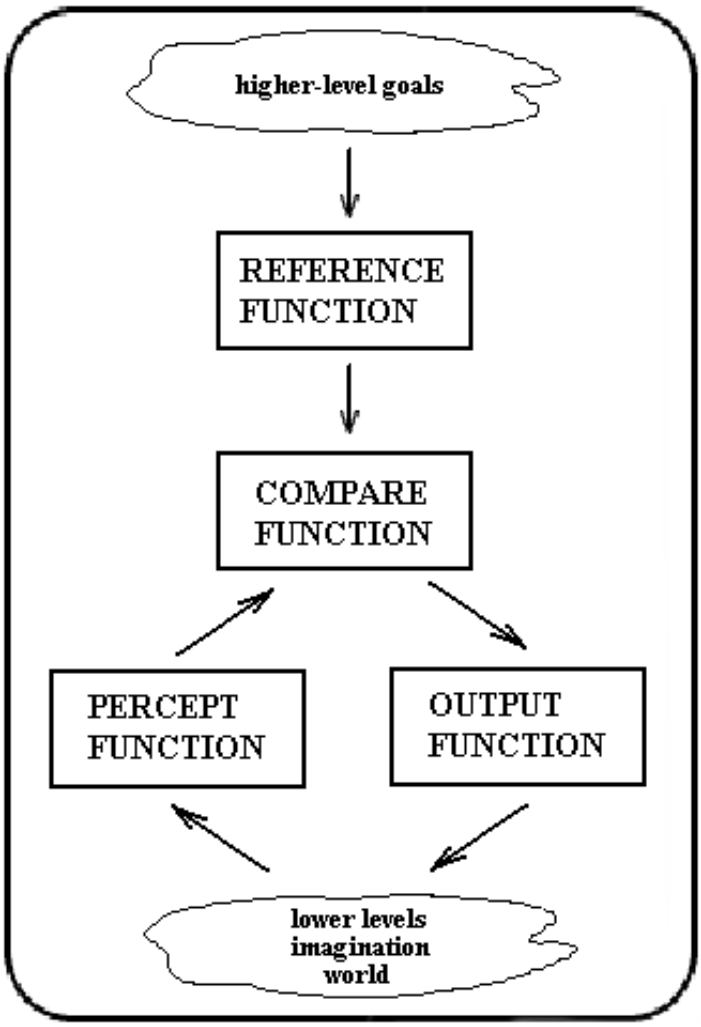
Lower level motor control seems to work pretty much the same in all of us; it is the higher level functions—such as choosing a job, a mate or a world-view, for example—that can sometimes involve the control of perceptions that seem almost impossible to fully grasp without actually being the person in question. It can even be very difficult for the person to articulate to others what they are

controlling for. Scientifically, we must view such a person *not* as controlling some external variable, but rather as controlling for having certain kinds of perceptions, feelings, etc. Behaviour is the control of perception, not of the external world (although the external world, as part of the feedback loop, will normally end up being altered in the process).

3. The Basic Controller

The basic notion of perceptual control is simple. Every perceptual variable that is controlled in the brain has a "comparator function", which takes some perceptual input (perhaps already highly processed) and compares it to a "reference" (what the system would like this variable to be). This produces an "error" signal (the amount by which the input falls short of the reference), which is then fed to an "output function", which decides how to act on the lower levels (or the world) in order to minimize error. In this manner, the system gets continual feedback on the result of what it is doing, forming a tight "closed loop" with its environment.

Figure 1 shows a schematic of the simplest type of control system. If we take the perceptual and output functions to be simple sums, and the comparator to be a subtraction, we get the simplest possible feedback control system. For various technical reasons, the actual functions are likely to be at least somewhat more complicated than this, but can still be quite simple (although they can theoretically also be quite complex). In many models, the output function is an integrator (i.e., an accumulator) with a decay rate, so that the value changes relatively smoothly, instead of in fits and starts. There is also usually a "gain" associated with the output function, which is simply a multiplicative constant that determines how hard the controller pushes on the environment (the lower levels) to get what it wants.



The hope of PCT is that our brain can be understood as a vast hierarchy of such basic controllers. At the bottom of the hierarchy are the motor outputs to the world and the sensory inputs, while at the highest level are the reference signals that represent the ultimate goals that the entire rest of the hierarchy is subservient to.

4. The Control Hierarchy

Lower level goals can change over time, since ultimately they are simply functioning in service of the higher level goals. But the highest level goals of all are physiological and hard-wired into us by evolution; they can never be changed. All behaviour is some kind of attempt by this vast hierarchy to correct certain error signals to bring its perception in line with its goal or reference (behaviour is the control of perception). Higher level controllers tell lower level controllers how to behave in order to achieve the higher-level goals. These lower level controllers are free to come up with their own subgoals in order to achieve this. These subgoals, while they may be pursued vigorously by the organism, are still subordinate to the higher goals they serve, and if the organism perceives at any time a better way to achieve the higher goal, it will readily abandon the current lower-level goal.

Imagine yourself driving your car to work. You are continually controlling for perceiving yourself near the centre of the road by acting on your steering wheel percept. But why are you doing that? You might not be consciously thinking of it at the time, but examine yourself carefully, and you will discover that you are in fact controlling for getting to work, only on a somewhat longer time scale. On an even longer time scale, you are controlling for making money, and ultimately for being happy. "Happiness" in fact is the usual translation of Aristotle's word for the highest level goal in the hierarchy. In the modern theory, the highest level has a variety of different references, corresponding to physiological states, such as the levels of certain chemicals in the bloodstream, hormones, certain kinds of neural excitation, etc. In general, one can think of the highest-level goals as the basic drives and needs predetermined genetically (or perhaps environmentally if they are determined early enough in life, while the brain is still in development, so as to become effectively hard-wired).

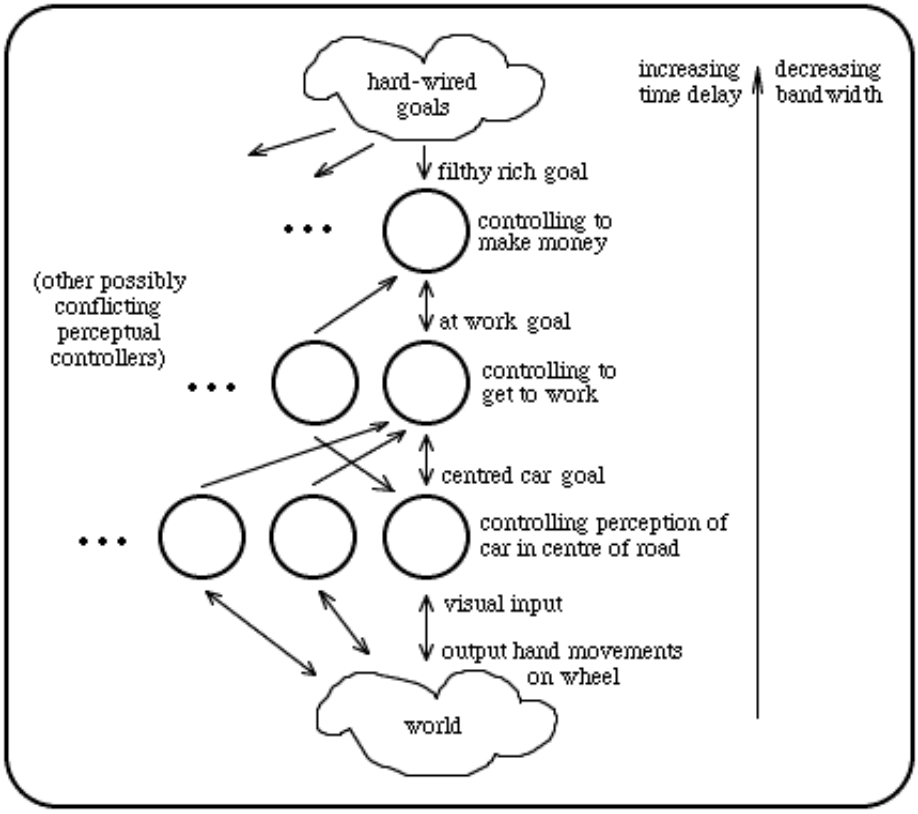


Figure 2 depicts part of a control hierarchy for driving to work. At the highest level shown (which is still far from being the highest level of all), the system is controlling for making money. This money-controller is telling lower-level controllers to drive the car, which are telling lower-level controllers to keep the center of the road in the center of the visual field, which are telling still lower-level controllers to move the steering wheel this way and that. The referent at each lower level is supplied by the behaviour of the higher level. The referent for the staying-on-the-road subsystem is supplied by the behaviour of the getting-to-work subsystem, for instance. The perceptual input for each higher level system is taken from the percepts of multiple lower level subsystems. Not only do higher level goals get controlled for on a longer time scale, but as we go up the hierarchy, there are increasingly long delays between acting and

getting feedback about the result. Similarly, there is less ability to process great quantities of information. Communication theorists would say that the higher level controllers are processing input at a lower "bandwidth", or information throughput.

With all these different controllers acting on the shared pool of lower level subsystems, there will naturally be conflicts between different controllers attempting to set lower level goals differently—hence the feelings of tension that we experience all the time, such as "I want to go to work, and make money, but at the same time I want to stay in bed and sleep."

5. Learning and Reorganisation

Humans do not have a fixed, unchanging control network to tell them how to behave (although some animals *do* have such a system). Humans can change their system radically in light of experience. So an accurate model of human cognition in terms of PCT must include some kind of learning algorithm, although there is no particular set of such algorithms written in stone within the PCT community.

Some neural network algorithms can allow for incremental changes to a control system's functions, so that it gradually learns to control better and better. However, when error in the system is high for an extended period of time, something more dramatic than small incremental changes may be called for. In such a situation, the system is signaled to reorganize, at which time radical, possibly random, changes are made, as the system attempts to overhaul its strategy in a major way.

Reorganization could occur on the local level of a single controller, if that particular controller has been failing to control its percept properly. Or, if the whole system has been stuck in a rut, and error throughout the control network has been too high, global reorganisation could be triggered, and fundamental changes will be made throughout the system in an attempt to hit upon a superior strategy for meeting the highest-level goals.

6. Imagination

Crucial to Powers' system is the concept of imagination. Note that in figure 1 the lower levels are alternatively labelled "world" or "imagination". Instead of outputting to lower level controllers that eventually output to the real world, control can also taken place internally, in imagination. Here, the feedback loop is short-circuited and never passes through the environment at all, but instead works through an internal world model. The organism interacts not with the world but with its own internal imaginary world.

Why should an organism be involved in such a delusional activity? Simply to test out methods of control before trying them for real. Once a method of control has been executed on the real world, it may be too late to change the result. Imagination is crucial in any control system sophisticated enough to learn from its experience. Many possible techniques can be tried in imagination, and the results observed, all in much less time than it takes to try just one technique out for real, and without any potentially negative consequences. Techniques (i.e., combinations of percept and output functions) that work well in imagination can then be fixed as the ones to use for real. This enables the organism to use its model of the world to make more correct decisions about the best course of action to take to meet its goal.

7. The Test for the Controlled Variable

One of the fundamental principles of perceptual control theory can be summed up by the slogan: "You can't tell what someone is doing by watching what they are doing." Of course, at times you can, but in general you cannot. It is entirely possible to have two organisms behaving identically, while in fact they are not really controlling for the same perceptual variables at all (or even particularly similar ones). This is as true for rats in a lab as it is for people. The way a scientist tests a rat to see what is "really going on" when it runs a maze, eats cheese and presses buttons, is not just to watch and record observations, but to repeatedly, and in various ways, *disturb* the organism in its attempt to control. This is called the "test for the controlled variable".

To test if an apparent control for variable X is actually subservient to some other dominant variable Y, we must disturb the (hypothesized) control of Y so that control of the lower-level X variable no longer aids the organism in controlling for the higher-level Y variable. If the organism immediately and dramatically stops controlling for X, then there is strong evidence that Y is a dominant control variable. If my boss calls me up on my wireless phone and fires me while I'm on my way to work, I may suddenly lose my motivation for driving to work.

But what if my boss calls, not to fire me, but to say that my salary has been cut dramatically. If my main goal in working at that particular job was to make lots of money, I may immediately quit, and turn the car around. If, on the other hand, the main goal I was satisfying by working there was personal fulfillment, and I love my job, I may be relatively unaffected by the news. You cannot determine which of these situations is the case by watching me drive to work. My behaviour is identical in both cases, even though what I'm *really* doing, on the longer time scale, is quite different.

8. Layered Protocols: Communication is the Control of Belief

Just as in general human behaviour is the control of perception, communication in particular, according to Taylor, is the control of belief (Taylor

1989; Farrell *et al.* 1999), including the control of one's beliefs about the beliefs of others. For Taylor, a "belief" is a kind of coherent system of perceptions brought together in one's internal model of the world, required to transform error signals into the appropriate output to achieve control.

"Control of belief" may be a bit too general as a slogan for "communication", as the word is generally understood, since it would cover some situations where one is not really communicating in the traditional sense, such as sitting in an armchair and ruminating about the world. Here, one is trying to control certain aspects of one's belief system, yet no communicating partner is involved. One tries, within oneself, to bring one's internal belief system into line with some goal (logical consistency, for instance, or perhaps the ability of the belief system to make one feel good). This could, of course, still in a sense be thought of as communication, if we conceive it as a kind of "talking to oneself". Unless specifically noted otherwise, I will use the word "communication" in the more general and paradigmatic sense of information flowing between different persons, rather than between internal components of a single person.

This is not to say that in communication, we are *directly* controlling another person's belief system. In PCT, the other person's belief is never controlled directly, but only indirectly through the control of one's own beliefs *about* that person's beliefs. This process may or may not be that relevant to what the other person's belief actually is, since our perception of their belief is filtered through many layers of the control hierarchy.

9. Felicity Conditions for Communication

Taylor lists three fundamental propositions that can be used to express the basic "felicity conditions" that determine in general whether a communication will take place and when it will terminate (although any particular situation may yield a larger number of felicity conditions specific to that particular communication):

P1: the receiving partner has an understanding of the message.

P2: the receiving partner's understanding of the message is adequate.

P3: it is no longer worth trying to achieve (P1 & P2).

A communication is terminated when both partners come to believe strongly in P3. The goal of the communication, at least for either or both partners that are cooperating in the attempted communication, is to come to believe (P1 & P2), *and* to believe that the partner believes (P1 & P2) as well. From this, belief in P3 normally follows (since if P1 and P2 are already achieved in both partners, there is obviously no longer any point in trying to achieve them). If only P3 is believed, however, the communication is also terminated, but with the goal of the communication left unfulfilled.

If both partners believe all three propositions, then the communication has ended successfully. If, for example, Bill was sending a message of some kind to Susan, then Bill now believes that Susan has some belief/perception in her mind that she believes corresponds to the belief/perception that Bill has in his mind that he was trying to communicate (i.e. he believes P1). Susan believes she has a perception/belief in her mind that corresponds to the one Bill has in his mind, so she also believes P1. Bill also believes that Susan believes P1—he believes that Susan believes that he believes she has received the message. And Susan believes that Bill believes P1—she believes that he believes that she believes she has received the message. If a similar situation of reciprocal belief exists for the adequacy of the message, then both partners also believe P2, and hence P3, and the communication terminates.

So by satisfying the conditions for successful communication, Bill must do more than just believe that Susan perceives his perception/belief accurately. He must do more than believe that she believes that she perceives it accurately. He must *also* believe that she believes that he believes that she perceives it accurately.

An infinite regress suggests itself here. Why not demand that Bill also believe that Susan believe that he believes that she believes that she perceives the message correctly, and so on to as many levels as desired? Taylor's answer is simply that further levels provide no further useful information that allows either partner to meet their goals, and thus has no effect on actual behaviour. In actual dialogue, it is enough for each partner to not only believe the communication was successful, but to believe that their partner is likewise satisfied. This general requirement yields the number of reciprocal belief levels described above, no more and no less.

10. Belief

I have been calling Bill's message a "belief/perception" to avoid committing to whether Bill is communicating a belief-perception or a non-belief-perception. Although the boundary line between the two is fuzzy, I will define a belief as a perception or complex of perceptions that has become so embedded in a person's internal world model that it exerts a persistent influence over time on how their control system transforms its goal and current perception into action (behaviour). This means that a belief need not be at all propositional. It simply needs to be a perceptual pattern that one persistently uses in understanding the world.

11. Layered Messages and Protocols

Just as a regular perceptual control network is layered, with long-time-frame higher levels telling short-time-frame lower levels what to do, a higher level

controller that is attempting to send a message to a recipient may need to initiate the sending of many lower-level submessages in order to communicate the main message (what Taylor calls the "prime message").

Sending a message is not just a matter of a stream of letters or sounds leaving the sender and arriving at the recipient. As the message is being sent, the sender receives constant feedback from the other person as to how they are doing, how close they are getting to producing in the other person the perception they want. This requires that many "protocol" submessages be sent, such as "I'm ready to receive", "yes, I got that" or "huh, say again?" The standard protocols we tend to use most often in human communication have been formalized by Taylor in his General Protocol Grammar (GPG), although the full technical details will not concern us here.

12. Argumentation and Layered Protocols

In order to apply layered protocols to argumentation, we must first determine what makes argument different from regular discourse. I will characterise argumentative discourse as follows:

(1) The communication is two-way, each partner acting as both a sender *and* a receiver.

(2) The messages being sent are not merely perceptions, but beliefs.

(3) Each partner believes, at some level in their hierarchy, that they cannot believe both their own message *and* their partner's message, seeing the incoming message as incompatible with their existing belief system.

If any of these conditions are absent, then we are not really dealing with an argument, or at least not a mutual argument (although it is possible, for instance, for one partner to see the communication as an argument, where the other sees it as a nonargumentative discussion, and the dividing line between the two will naturally be fuzzy). In argumentation, we are trying to *convince* the other to not only *perceive* some idea or thought, but to *believe* it. That means more than merely inducing in your partner some perception or other; your partner must make this perception a reasonably permanent part of the internal world model they use to control their perceptions in general. Furthermore, both you and your partner must each believe, at some level of your control hierarchies, that the other's message is incompatible with your own—that it is not possible to believe both.

Note that I said "at *some* level of the hierarchy". This qualification is necessary, since it is possible that, at a higher level of the hierarchy, one of the arguing partners does not really believe in the truth of their message at all, nor even in its incompatibility with the incoming message. In cases like this, such as insincerity or devil's advocate, the arguer *is* nonetheless undeniably taking on

the beliefs required in (3) above, even if at some subordinate level of control. According to our technical definition, this still qualifies as belief, since it affects control persistently over time. But it is belief that is induced by a higher, dominant goal level where the belief is not actually held. The higher level is, in a sense, *simulating* the belief at the lower level.

If one was to use the broader definition of communication mentioned earlier, where we allow communication to be within an organism between its internal components, we would say that the higher level has *communicated* a belief to the lower level, one which it does not itself hold. This might even take the form of an internal argument, if the lower level had to be "convinced". This does not result in the same degree of internally felt tension, however, as conflicting goals at the same level, since the "pretend" goal is clearly subordinate to the higher level goal. If I am truly conflicted, really not able to decide whether it means more to me to get out of bed and go to work or to stay in bed and sleep, then my internal "argument" is most likely between goals at the same level of the hierarchy.

We should avoid the temptation to say that the incompatible incoming and outgoing messages are contradictory, or that the goal is to arrive at the truth. What makes the discussion argumentative is the *belief* that the two messages are incompatible in some way or another. This may or may not be an explicit belief in their logical inconsistency. If Bill tells Susan, "I think your dress is too flashy," and Susan responds, "Well, it's what I'm in the mood for today", Susan does not believe her desire to wear the dress is logically inconsistent with its being too flashy. This simply is not solely a matter of inconsistency, nor is it strictly a matter of propositions. Susan has a perception of the dress, at least on this particular day given her mood, that has enough permanence and influence on her behaviour that we call it a belief. But it is not properly understood as a proposition. It is made up of, as least partially, her perception of how the dress fits into what she feels satisfies her goals for this occasion, whatever those may be. This "belief" may be completely prelinguistic, and Susan may even be incapable at this point of fully expressing it in words. Susan is not trying to convince Bill that her mood is inconsistent with the dress being too flashy, she is simply trying to say that right now, her mood is what is more important to her, and she just wants Bill to see that. She sees Bill's message as conflicting with her belief in the current wearability of the dress, were she to adopt it *on the same level* as her belief in her mood. So she cannot believe both *on that level*.

The reason Bill's statement and Susan's reply is an "argument", and not simply a discussion, is that Bill is trying to get Susan to change her belief (or at least he perceives that he is doing this—he may be mistaken about her beliefs). But Susan does not believe it possible to hold her existing belief about her dress and its role in her current control, and also to hold the one Bill is trying to induce in her, *at the same level* in her hierarchy. So she argues back, trying to change Bill's belief in turn. In this case, she is being mostly defensive, and may only want to get Bill to believe that there is no point in bugging her anymore, or

perhaps a bit more strongly she may want him to believe that it is a good and reasonable thing for her to be positive about her dress. In any case, she wants to change his belief in *some* way, or else there would only be a one-way argument.

13. Argument Settlement and Resolution

It is possible for Bill and Susan to be at complete odds with respect to the messages they think are being sent and received. Bill might think he is trying to convince Susan that red is a tacky colour, while Susan believes that it is the pattern that is perhaps a little flamboyant but perfect for this particular occasion, and so this must be what Bill is talking about. The point is that they mutually agree that there *is* an incompatibility in the reciprocal messages, not that they need have any agreement as to *what* those messages are. Indeed, misunderstandings as to what messages are being sent are quite normal, since much of real argument consists in the participants struggling to belief (P1 & P2), and until they do they could be completely mistaken as to the nature of the messages. If they both understood each other's message perfectly, then quite likely they would either come to an agreement or believe P3 without believing (P1 & P2). If all three propositions are believed, then we will say that the argument is "resolved". If only P3 is believed, then the argument is "settled", but not resolved.

There are in general four beliefs to be concerned about in argumentation. There is (1) the belief Bill is trying to transmit to Susan and (2) the belief Bill believes Susan is trying to transmit to him. These may or may not be the same as (3) the belief Susan believes Bill is trying to transmit to her and (4) the one she is trying to transmit to him. Bill's side of the communication qualifies as argumentative if at some level affecting the discussion he believes that he cannot believe both messages as he perceives them. Likewise for Susan, who sees her discussion with Bill as argumentative because she does not believe she can believe that the dress is too tacky and *also* believe in her desire to wear the dress enough to actually put it on and go out. If she does find herself believing both of these, at the same level of the hierarchy, while still seeing her discussion with Bill as an argument, then she is in a classic case of conflict, and will conduct an "internalized argument" of some kind to try and resolve the issue.

Since the two partners start out believing that they *cannot,* at the same level of control, believe both messages, a resolution (as opposed to a settlement) necessarily implies that either one or both participants have:

(1) changed the message they were sending,

(2) changed the message they perceived the other person to be sending, or

(3) ceased believing they could not believe both messages.

In (1) the person has become convinced of the other's position. In (2), the person has changed their mind about what they thought the other person was claiming, resulting in case (1). In (3), they have not changed their mind about the nature of the messages at all, but have decided that *both* can be believed simultaneously.

The attempt to achieve (P1 & P2), so crucial to any argument, is rarely the result of a straightforward transmission of a message from A to B and then another from B to A. Recall that we are starting out with neither partner particularly understanding what beliefs are even at issue. Therefore, much time is spent exchanging protocol messages, to indicate things like "yes I understand", or "huh? what's that?" (which tells the other person that you do not understand and they must put off believing P2; it also gives them a clue as to how to go about successfully transmitting the message).

14. An Argument is More Than Just an Argument

If an argument is to be properly analyzed, it is not enough to analyze the hierarchical structure of the argument *per se* (*as* an argument). The reason is that, just as a lower level subargument cannot be fully separated from the dominant argument it serves without misunderstanding it, so the entire argument as such cannot be separated from the rest of the control hierarchy without misunderstanding it. What people do when they argue is to control, or seek to control, their internal beliefs about what others believe. Beliefs are just a kind of perception, so argumentation is simply one kind of general feedback control, which uses the protocols of cultural communication for feedback.

This process is inevitably hierarchical. There are subgoals and subarguments that are spawned in order to support the more general argument. This corresponds in a precise manner to a control hierarchy (argument being just one type of feedback control activity a living organism can engage in). The transmission of a message is modified according to feedback from the communicating partner (which is largely the "world" for the purposes of analyzing the argument). In order to achieve this, a subargument may need to be argued for and resolved first. This subargument, with its own possibly quite different incoming and outgoing messages, is not the dominant argument, however, as it merely serves the main argument. The main argument takes place, as for any higher level control, on a longer time scale, with longer delays, and less bandwidth.

In order to convince Susan that her dress is tacky, Bill may get sidetracked into convincing her first that the social function they are attending is semi-formal, which has nothing directly to do with her dress. This argument, as a subargument, gets resolved or settled before the main argument. The main argument has longer delays, because it must wait for the subargument to be settled (which may take hours) before it can "resume". (Although it is of course an abstraction to view the higher level as "suspended" at all, since the

discussion about the semi-formal issue *is* the conducting of the tacky-dress argument, which in a way has not really been suspended at all.)

What may be surprising is that much (although not all) of what is left implicit and unspoken in an argument is actually higher up in the hierarchy, and so *dominant* over the more explicit and obvious aspects of the argument. The explicit messages act in service to the implicit messages. Moreover, the entire argument is itself a mere part of a larger control hierarchy, which is not in itself argumentative at all.

15. Layered Goals and Beliefs In Coalescent Argumentation

Gilbert argues that, rather than focussing on an idealized, logic-based formulation of an argument as a search for the truth, we should seek to analyze the goal structure of the participants. More importantly, the participants themselves can use this kind of analysis to improve their chances of reaching a resolution. Instead of concentrating narrowly on the "issue at hand" (i.e., the explicit lower subordinate levels), the arguers can work at constructing proper models of each other's goal hierarchies, attempting to uncover as much of the hierarchy as possible. In particular they should be concerned with uncovering goals that are relevant to the dynamics of the argument, but previously unrecognized.

This approach will allow the participants to avoid just "going for the win". In uncovering the dominant, upper layers of the goal hierarchy, as well as unrecognized goals on the same level as the subordinate argument, the participants can concentrate on achieving their goals in *some other fashion* than that assumed in the subordinate argument. In this way, seemingly intractable arguments can often be resolved into agreement.

I will begin with an overview of goals in argumentation as presented by Gilbert (1997: 67-74), developed within the paradigm of layered protocols. Gilbert's view of a belief system is more of a "web" than a hierarchy, but the strict hierarchical nature of PCT and layered protocols can easily be modified to accommodate something like this (although by its very nature, layered protocols is best suited to at least a loosely hierarchical structure).

According to the logic-based approach to argumentation theory, the goal of an argument is to convince the opponent that one's claim is true. According to Gilbert, while this may be *a* goal of an argument, it is not necessarily *the* goal. There may be many goals, acting on different levels (these different levels we will take to be roughly equivalent to the layered protocol goal hierarchy). Arguers must negotiate between many possibly conflicting goals, of three basic types, which Gilbert organizes into the following ordered (i.e. prioritized) sets:

Motives: M = {m1, m2, …}

These determine the other goals in a broad, general way. They are thus roughly equivalent to the highest level goals in the PCT hierarchy (those that are above the level of the argument *per se*).

Task goals: T = {t1, t2, …}

These goals are the obvious focus of the argument, and are thus roughly equivalent to the lower level goals in the PCT hierarchy.

Face goals: F = {f1, f2, …}

These are goals having to do with the personal relationship between the arguers. Thus, they can also be either task goals *or* motives. In the PCT hierarchy, these might be found high up in the hierarchy, or lower down.

All these goals are delimited by a set of procedures, P = {p1, p2, …}. These procedures are roughly equivalent to the protocols used in the PCT hierarchy to exchange information, although we could interpret them more broadly to include any rule of operation of the control hierarchy, whether directly related to communication or not. For instance, the protocol of nodding one's head every once and a while when listening, indicating understanding, would be one such procedure, and is also clearly a communication protocol. But we could also include the general procedure used in all control systems of taking one's goal (or reference signal), comparing it to the perceptual input, and thereby producing an error signal that is used to produce behaviour. This general procedure is, of course, not special to communication or argumentation, but is simply used in all behaviour for any living organism. Other less general delimiting procedures would include any particular way of translating error signals into behaviour that a particular individual has adopted from experience. For instance, an individual might have a tendency (conscious or unconscious) to always check out whether a person is fidgeting while they talk, as they tend not to trust someone who fidgets. This is a procedure used to translate error into behaviour, although not a universally adopted one. Another kind of procedure that we will also include would be the way that a person builds their perception up in the first place, before comparing it with their goal (as mentioned earlier, we all do this differently).

Thus, we see that so far there are (at least) four different types of delimiting procedures:

(1) general rules for the operation of control systems.

(2) specific rules for translating error into behaviour.

(3) specific rules for translating sensory data into perception.

(4) communication protocols (of type 2 and 3).

(5) argumentation protocols (of type 2 and 3).

"Goals alone," says Gilbert, "are insufficient to predict or delimit action." There are multiple ways to achieve any given goal. What behaviour will be chosen depends on many other variables, such as the procedures and protocols adopted by the participants, *and* of course the actual physical situation in which they find themselves.

Note that Gilbert's goal sets are prioritized, meaning that some are more important than others. When two goals conflict, the more important one will be acted on, and the other neglected. For Gilbert, the goal of an argument is less to persuade and more to maximize one's "overall degree of [goal] satisfaction". This is not a simple ratio of number of successes over failures, since the more important goals will have more weight. In terms of PCT, since we have a huge hierarchy of goals seeking fulfillment simultaneously, the "overall satisfaction" is roughly the inverse of the total error at the top level we are concerned with, weighted by the gain on each subsystem within that level (recall that the gain on each individual controller is how hard that particular subsystem is trying to achieve its goal).

Higher priority goals will thus have higher gains than less important goals. The gain is usually multiplied by the error to produce the actual output of the controller. So if we double the gain on a "get to work" controller that is telling a lower level controller to "stay on the road", the system will be twice as likely to act, or will act twice as forcefully (depending on the exact nature of the output function).

A crucial aspect of coalescent argumentation, as compared to more logic-based approaches, is the increased role played by face goals and motives, as opposed to the more obvious task goals (although the role of task goals is by no means downplayed). In most real argumentation, task goals, while important, are only a small piece of the picture. Face goals may be more important, and the participants usually have broader motives not obvious from the actual argumentation.

16. Constructing Goal Hierarchies: Test for Controlled Variables

Goal analysis (the actual construction of a PCT goal/belief hierarchy) concerns us in two respects: (1) a third-party argument analysis will begin with an analysis of the goal structures, or control hierarchies, of the participants, and (2) the participants themselves will need to perform something like this analysis, explicitly or implicitly, on both their opponent and themselves if they are to effectively argue.

Constructing such an analysis will require using the test for the controlled variable. While much of a person's hierarchy may be uncovered by passive observation, without interacting with the person and disturbing their control, much of the hierarchy will remain a mystery. Application of the test also illustrates why goals that are above the level of the argument *per se* are so

important in an analysis of the argument, since they can have dramatic impact on the way an argument will be affected by new and unexpected disturbances.

Arguing partners, of course, have only a limited ability to use the test for the controlled variable, and only limited ability to model their own or their opponent's hierarchies. We cannot, during a real argument, fire someone or cut their salary just to see how they will react. However, we *can* simulate the disturbance by suggesting it, getting the person to have one level of their control hierarchy simulate the salary cut at a subordinate level, and then we can observe the result of *that.*

17. An Example: Bill and Susan

Bill is arguing with Susan, who is a coworker. To a casual observer, it appears that Bill and Susan are simply squabbling over whether company X is or is not good at making widgets. A non-goal-oriented, logic-based approach might analyze the validity of their arguments with respect to that issue, and ignore their larger motives and face goals. This does not mean the logic-based approach would necessarily restrict itself to formal, deductive methods, but it would generally insist on looking only at the validity of the arguments put forward concerning company X's widget making ability. Other factors would be allowed only if they were relevant to justifying these positions.

But Gilbert's approach says no, while it may be important and significant to look at the logical validity of Bill's and Susan's arguments, if we are to really help Bill and Susan argue more effectively, we need to consider the broader picture of what these two people are *actually doing*. In analyzing their goal hierarchies, we do not wish to go looking so high in their hierarchies that we just end up looking at the same physiological goals everybody shares. But neither do we want to stay so close to the argument *per se* that we miss what is really important.

It may turn out that Bill is really more concerned with getting a raise than "finding out the truth" about company X's widget expertise. To test for this, we must do more than simply observe Bill in action and look at whether his arguments are valid, we must actively interfere with the system, as with the rat in the maze, poking and prodding Bill in various ways to see if he has significant higher level goals and what effect (if any) they have on the current argument. This is not only something we do as third-party observers attempting to describe and explain the interaction, but it is something Susan will also try to do, should she choose to argue with Bill's position (especially if she decides to use our brand of coalescent argumentation therapy). It is also something Bill himself ought to do (self-analysis), especially if he is not fully aware of his own motivations for badgering Susan.

As it turns out (let us presume), Bill and Susan's company, in Bill's judgement, would greatly benefit from a relationship with a company skilled in widget

making, and this would very likely result in raises for everyone in Bill's division. As a result, he is trying very hard to get Susan, the resident widget-expert, to forge new ties with company X. He tries to convince her that this will result in good things for the company and a raise for her (although what he is really after in all this is a raise for himself). But he finds her resistant. She perceives company X through very different eyes than Bill, and cannot really see why he thinks they are good at making widgets. To her, their widgets are obviously totally lame. She is tempted to dismiss Bill's arguments, but instead decides to delve into the argument, since it is always *possible* Bill might have a point, and since she *would* like a raise, it is worth pursuing up to a point.

Susan may not have a pre-existing model of Bill she is satisfied with, and so will apply the test for the controlled variable, trying to figure out what makes Bill tick. Ideally she should actually change the conditions for what she is hypothesizing might be Bill's higher level goal, for instance by actually removing the circumstances within the company that might potentially lead Bill to think that ties with the other company would get him a raise. But, as is often the case, this kind of radical disturbance is not practical, so Susan will be somewhat subtle and simulate such disturbances, by suggestion. She will try to induce in Bill a simulation of the desired belief at a lower level, under control of a higher level that very well knows that such is not the case. In other words, she gets Bill to *imagine* that their company would *not* respond with raises for everyone in Bill's division. We do this sort of thing all the time in argumentation and it is very important. To build a model of the opponent's hierarchical goal structure, we suggest disturbances, and see if we get some indication of the reaction that a real disturbance would have elicited.

Not only are higher level goals, or motives, important to argument analysis, but face goals must also be included. A face goal is a goal that has to do with the maintenance of the participants' personal relationship. Whenever people communicate, they will (almost always) be highly constrained by their desire to maintain or establish certain kinds of relationships with others. Perhaps Susan really would have just dismissed Bill's argument, uninterested in continuing, were it not for the fact that she has a crush on him, and is interested in him sexually. Sex is a powerful motivator, even when we try so hard to pretend it is irrelevant. Were we to apply the test for the controlled variable to Susan, by informing her that Bill is gay for instance, she might realize she has no chance, and lose interest in the argument, as the only thing now contributing to her goal of understanding Bill is her desire for a raise, and she really does not take his argument seriously enough to pursue the discussion on that basis alone.

Bill, of course, no doubt has face goals as well. He may also be controlling for spending more time with Susan. But perhaps Bill is doing it with a different motive. Whereas Susan has a romantic crush on Bill, Bill simply has a general, overarching principle of wanting to work with any colleagues he has not yet spent much time with. As a result, he also has a face goal to spend more time with Carol down the hall. We could, again, test for this (at least in theory) using the test for the controlled variable.

Note that while task goals are generally lower-level and higher-bandwidth, they are not necessarily at the bottom of the hierarchy. They may themselves need to further invoke lower level controllers (for hand-eye coordination, perhaps) that are needed to carry out the communication, but are not themselves the obvious focus of the argument.

While the main task goal is the most apparent goal to an observer of the argument, the actual arguer might see a higher-level goal as more the focus of what s/he is doing. Bill, for instance, is to a casual observer arguing with Susan about company X's widget expertise. But he himself may be very aware throughout the argument of his more crucial goal of getting Susan to forge ties with company X. Thus, it is this goal that Gilbert calls the "apparent strategic goal", which may or may not be the main task goal, depending on the circumstances.

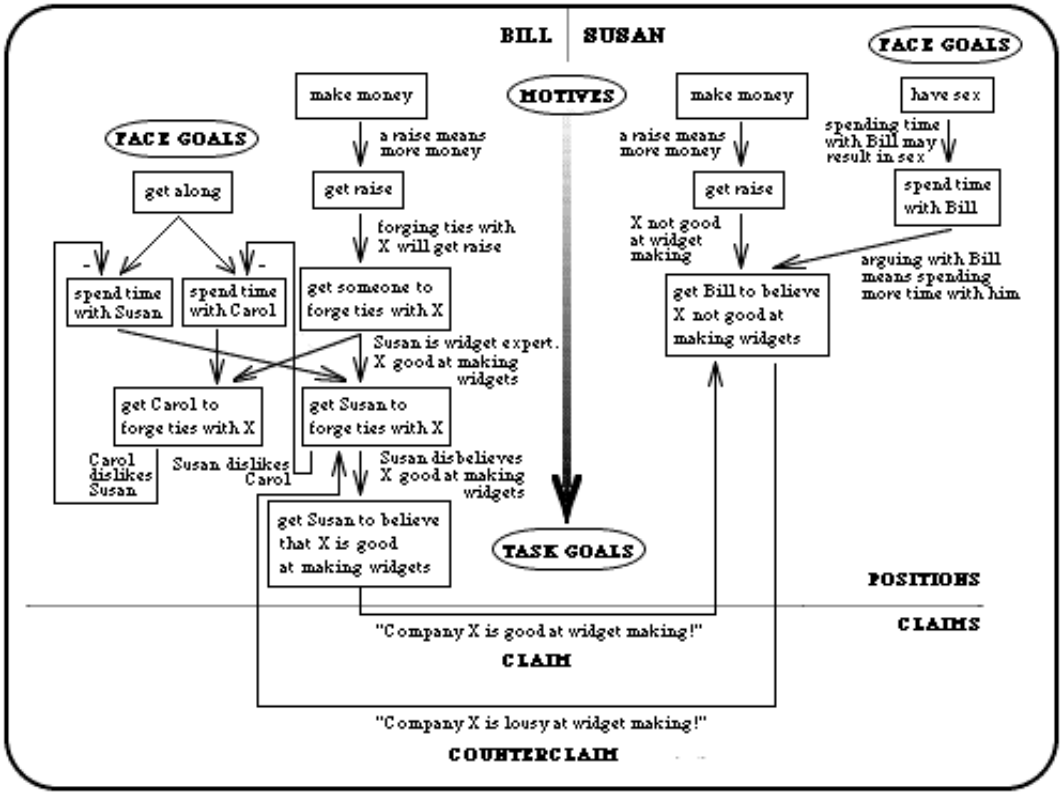


Figure 3 shows a simplified analysis of Bill's and Susan's control hierarchies (some parts of the diagram have yet to be explained). This is a greatly simplified diagram, of course, showing only some of the major connections between Bill and Susan. Each box is a subcontroller, controlled by one or more higher-level controllers. The arrows between the boxes, whether providing perceptual input or behavioural output, are labelled with beliefs (although the reader should keep in mind that we are using the word "belief" in the broad sense of any procedure for converting signals into behaviour or perception that persists in an individual over time).

There is a line in the diagram that separates the "claims" from the "positions". A claim is the outwardly apparent assertion made by one of the partners in the argument. Gilbert takes the term "position", however, as a deeper term

referring to the whole complex of beliefs and goals behind the claim. A claim is just an outward label for the much richer and more complex position. Thus, the claim side of the line in figure 3 is in the external world, while the position side is in the brain. Although I have shown the external world in figure 3 as only containing the claims, there is much more going on than that, of course. The world is the environment that links the two partners, and information can be transformed and constrained in all kinds of ways by it.

Indeed, only a small part of the actual control hierarchy is shown here, and many links that no doubt exist between the boxes are not drawn in. This is just to keep things simple. In reality, for example, the outputs that Bill and Susan trade back and forth are immensely more complicated than the simple claims in the diagram, and they input into many more of the other partner's subcontrollers than what is shown. I have simply depicted a few key points where the perceptual inputs particularly matter. Note, too, that exactly how this perceptual input gets processed is up to the organization of the perceiver's network. Susan cannot attach to her speech acts markers that tell Bill's brain how to perceive what she is saying. Each of us builds the world in our heads afresh and in our own unique way, so there is no guarantee as to how others will reconstruct what we have tried to communicate.

18. Loosening the Hierarchy

While PCT networks are generally perceived as hierarchical in nature, and W.T. Powers (who is arguably the father of the field) views them as *strictly* hierarchical, it is not necessary that this be so. Beliefs may be interconnected in something more like a web than a hierarchy. A web structure would make the system less predictable, since the effects of a radical change in one belief could not be traced so easily through the system if "lower-level" goals that are supposedly subordinate end up looping back up and changing the higher level goals that helped determine them in the first place! This could make analysis much more difficult, but it is perhaps more in keeping with the way human beings work.

Take, as an example, Bill's face goals. So far, we have seen that he has a face goal to spend more time with both Susan and Carol. Now there is obviously at least *some* conflict involved here, in the sense that he (probably) cannot easily spend more time with both of them. He must spend less time with each, or decide which one he wants to concentrate on getting to know better. This is not a serious conflict within his own web/hierarchy, since the source of the conflict is simply the constraints imposed on his control by the outside world.

However, let us consider a more serious *internal* conflict for Bill. Let us presume that Carol and Susan despise each other, so much so that they both tend to avoid anyone who chums around with the other. Now Bill knows this (or more accurately, he believes it, so it is a constraint on his action, whether it is really true or not), so he realizes that if he tries too hard to convince Susan to

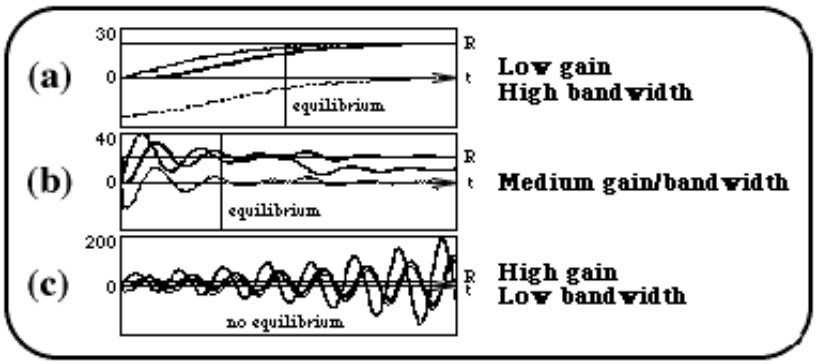
forge ties with company X, he may lose any chance of trying the same with Carol, and vice-versa.

In terms of his face goals, Bill has been torn for some time between Susan and Carol. Perhaps in his set of delimiting procedures, he has an overarching principle of leaving things alone in such cases, and he has been avoiding both Carol and Susan for quite a while. Perhaps someone else would have just flipped a coin. In figure 3, however, we see that Bill's task goal has caused him to go with Susan, because she is the widget expert. On the other hand, Bill could just as easily become conflicted again. For instance, he could be pushed back to Carol when he realizes that Susan is really very resistant, or perhaps when he realizes she is interested in him romantically.

Bill's goal/belief hierarchy is not a strict hierarchy at all in this case. The two subordinate subsystems looping back and mutually attempting to negate each other's parent control systems cannot be redrawn as a true hierarchy (the fact that they are attempting to negate a goal is indicated by a minus sign in the diagram). The more ubiquitous are these tangled loop-backs, the less sensible is it to speak of a strict hierarchy. However, I think it still makes sense to talk of the overall system as a hierarchy of sorts, as there will still be a general tendency for goals to be layered, with long time delays and low bandwidth for the more crucial goals, and high bandwidth, short time delays for the subordinate goals. However, I will refer to such a system as a "loose hierarchy", since the principle of hierarchical organization will be everywhere violated, with only a loose tendency for an overall layering into motives and task goals. Obviously, in such a complex system, the distinction between a motive, a task goal and a face goal can become very slippery indeed.

19. Thrashing

A common problem in control systems is the phenomenon of "thrashing". This occurs when the output gain on a controller is too high. Feedback control requires that the actions we output to our environment (such as speech acts) be slow enough for change to occur gradually and smoothly, and for our reactions to likewise be gradual and smooth. Radical action in a short time can completely swamp our ability to deal with the results. The output gain is the degree to which we are "pushing" out on the environment (and hence, in argument, on our communicative partner). If this gain is large, we are trying *very hard* to get what we want, to meet our goals. One of the most important lessons from control theory, whether applied to an individual organism or to society as a whole, is that if one tries *too* hard, everything falls apart.



The graphs in figure 4 are adapted from computer simulations (Randall 1994, 1995) of control systems meant to mimic living organisms (although such simulations are, of course, only crude imitations of the real thing). The point of showing them here is not to make a rigorous technical point (I have not presented the technical details of control simulations enough in this paper to warrant such discussion), but rather just to suggest some general and very common phenomena in biology that may help to inform our search for the right way to conduct argument. Thus, for our current purposes, the reader should not worry too much about exactly what all the curves are that are graphed in figure 4. The important part of these diagrams for us is that over time (t) a perceptual signal is brought into line with a reference R (or goal), which is fixed. This is successfully achieved in 3(a) and 3(b), but not in 3(c). In complex real-world systems like human beings, the reference is not usually fixed like this, of course, but may fluctuate just like the perceptual signal, since any one subsystem will probably be subject to many other subsystems attempting to modify its goals. In a loose hierarchy these modifications may even come from "lower down". However, as a general rule, a reference signal will need to change more slowly than a perceptual signal if stable equilibrium is to be achieved.

In 3(a), a simple control system is given an output gain of 10, which is relatively low. The organism (or in this case, the computer simulation of an organism) is taking a relaxed, laid-back approach to getting what it wants (it isn't a "go-getter"), and it reaches stable equilibrium smoothly, but gradually (things just keep getting better and better, and only rarely get a little bit worse). However, it takes a long time to get what it wants. This would be analogous to an argument in which Bill did as little as possible to disturb Susan's belief system, taking his time to explain his position one step at a time, as slowly as possible (with Susan presumably doing the same, or else Bill's feedback would probably not be stable enough for this kind of control to work).

In 3(b), we see an organism with a higher gain of 50. Here, stable equilibrium is reached, but not without some thrashing about and instability first (things get worse before they get better, but they *do* get better). Here, Bill is not pushing Susan *too* hard, but he is disturbing her beliefs in an aggressive manner. He isn't "laid-back", but he's not trying to railroad her into agreeing with him either. Control is achieved more quickly, although the process is messier.

In 3(c), however, the output gain is 100. The system is in run-away oscillation,

thrashing wildly, getting further and further away from stability (things just keep getting worse and worse). If the gain were intermediate between 50 and 100, we might see more stable thrashing, where the system keeps oscillating without control being achieved, but where things at least don't get progressively worse. For instance, Bill and Susan may find themselves repeating the same mistakes over and over, going in circles, because neither is really listening to the other, and they are both just trying *too* hard.

While one possible explanation for thrashing is trying too hard, another possible reason is low bandwidth, meaning that the partners are not communicating enough information to each other to achieve stable control. Insufficient bandwidth causes very similar problems as trying too hard, and the two problems often go hand in hand. In fact, a result of trying too hard is usually that less information is transmitted. If Bill is pushing Susan too hard to understand his position right away, he is probably not taking his time to explain everything in detail. When we push too hard, we expect results right away, where what may be needed is a slow and steady approach with a lot of information being transmitted, and a more relaxed attitude towards getting results.

Sometimes, of course, low bandwidth can cause thrashing even though the partners are *not* trying too hard. The bandwidth can be low, for instance, simply due to physical constraints. Perhaps the partners are writing letters back and forth, and they simply need the higher bandwidth of a face-to-face meeting to achieve equilibrium. (You might want to argue that writing letters back and forth actually provides quite high bandwidth, since one is free to write very long letters packed full of information. However, what is important in control theory is how much information can be transmitted *before* feedback would normally be received back from the environment. Humans normally get feedback quite quickly in a face-to-face conversation; in letter writing, the feedback takes much longer, so there is relatively little bandwidth.)

If Bill, for example, uses a high-gain/low-bandwidth strategy, he will immediately change his method whenever he sees Susan not responding the way he wants (he is trying too hard). The right approach is often to give whatever method one is using time to work, and only change strategies if it fails over a prolonged period. Of course, this is not to say that low-gain/high bandwidth is *always* the way to go. Sometimes it really *is* best to try as hard as you can. But next time you find yourself "running around in circles", and not getting anywhere, remember figure 4, and try to think if perhaps a more relaxed approach to getting what you want might work better in the long run.

20. Coalescent Argumentation Therapy

I will divide coalescent argumentation therapy into three stages. The presumption here is that two arguers have come to us seeking therapy,

because their argument was going nowhere. Thus, they are predisposed to going through some trouble to fix things (like drawing out complex goal diagrams). However, the basic ideas could still work on a more informal level in everyday argumentation. The first step will be to draw the goal diagrams, after which the participants will uncover common goals and seek to merge their beliefs in a way acceptable to both parties.

Coalescent argumentation is more cooperative than other approaches where each partner is out to "win". Both partners must, for it to work most successfully, have coalescent argumentation as one of their goals. This means they must truly want to come to a mutually satisfactory conclusion. The back and forth of their argumentation will, if all goes well, result in the coalescing of their opposing views into a maximally satisfactory result. Figure 4 (a) and (b) showed diagrams of the kind of equilibrium that is sought in a living organism. The coalescing of two points of view that occurs in coalescent argumentation is a particular case of such biological equilibrium.

21. Coalescent Argumentation Therapy, Stage I: *Drawing the Goal Diagram*

Gilbert proposes a suite of tactics for achieving coalescence. One of the first things we must do, he says, is attempt to uncover as many hidden goals as possible. The set of Bill's goals of which he is aware will be called G(*Bill*) and the set of Susan's goals of which she is aware, G(*Susan*). Coalescent argumentation involves bringing into awareness (into G) as many of one's goals as possible. This "increases their likelihood of satisfaction" and "opens the possibilities of the identification of mutually held goals."

If Bill and Susan are both practicing coalescent argumentation, they will start by trying to understand their goals. Let us assume they have both been having this company X argument for a week now, and are getting nowhere—they are thrashing. They come to us for counseling in coalescent argumentation, of which they have heard great things. We start with Stage I, by giving them each the assignment of separately trying to draw figure 3. Since they perceive the world differently, they might come up with radically different diagrams. Already we have a meta-level argument about exactly *what* is going on in the primary argument in the first place. But note that it is unlikely that they will coalesce on a mutually agreeable conclusion if they proceed to argue each with an entirely different idea of what is going on in the argument. So once they have each attempted to draw figure 3, they must negotiate between them a working version that they can agree on.

Notice that "winning and losing" has entirely been put to the side, for now. We are not asking Bill and Susan to stop caring about winning and losing, just to put their concern with it a bit on the back-burner for the time being, and concentrate on understanding what each other's goals and beliefs are.

The focus of this step in the counselling is what Gilbert calls "dark-side" goals

and beliefs (Gilbert 1997: 105-106). These are of two main types: (i) *unrevealed*: goals or beliefs known to the arguer, but unrevealed for any of a variety of reasons, and (ii) *unknown*: goals or beliefs unknown to the arguer (or at least the arguer is not aware of their role in the argument). The more intense the disagreement, according to Gilbert, "the more crucial it is to uncover and explore the dark-side positional constituents."

The entire goal-belief network relevant to a particular argument claim is, recall, the position, for which the claim is merely an external manifestation. What we have been calling "belief", however, covers far more than the usual everyday use of the term, since we are including any persistent pattern of transforming perceptual signals into action. In Gilbert's terminology (Gilbert 1997: 75-88, 110), the traditional sense of "belief" is what we call such persistent patterns when they are in the "logical mode". When they are in the "emotional mode", we call them emotions. When in what Gilbert calls the "kisceral mode", we call them intuitions, and when in what he calls the "visceral mode", they correspond to aspects of the physical environment, which are of course also part of the control loop (although these are downplayed in figure 3, they should not be in general if they are relevant). Since the visceral components are external to the organism, however, I would tend not to say they are part of the "position". However, they are certainly part of the "positional analysis", but of course in a proper PCT analysis of an argument, positional, environmental and goal analysis cannot really be separated from each other. All are crucial components in the control feedback loop, the nature of which may completely change if only one is concentrated on and not the others. Hence, Gilbert's stages I and II (goal analysis and positional analysis respectively) are here merged into a single stage, since goals and beliefs/emotions/intuitions cannot be separated from each other in perceptual control theory.

If our counselling is to be practical, of course, we must avoid being too idealistic. We must not require absolutely that Bill and Susan uncover and put on the table *all* of their dark-side goals and beliefs. If Susan really wants to keep her crush on Bill a secret, she should not be required to reveal it. However, it is nonetheless important for her to try and draw as complete a version of figure 3 as she can, knowing that she need only show Bill that portion of it that she decides to make public. Once she shows her public diagram to Bill, keeping her private version hidden, they can compare notes, and proceed to modify each other's diagrams, attempting to come up with a mutually agreeable diagram. This may only be partially achievable, of course, and the participants may have to accept at some point that they have somewhat different views on what is happening and move on. But the more agreement can be achieved at this pre-argument stage, the more likely will the participants be able to succeed in the actual argumentation stage. By the time a diagram is actually agreed on, Bill and Susan might feasibly have come to understand each other's positions to such an extent that agreement on the actual argument will almost fall right out of their pre-argument agreement.

22. Coalescent Argumentation Therapy, Stage II: *Finding Mutual Goals and Beliefs*

Assuming Bill and Susan have now, relatively speaking, agreed on a belief/goal diagram, they can proceed to look for mutual goals to help them achieve maximum goal satisfaction. We will start them out by asking them to find goals on Bill's and Susan's side of the diagram that are mutual. The larger is the intersection of G(*Bill*) and G(*Susan*), the greater the chance of coalescence. "By becoming aware of the role goals play in argumentation, arguers can better focus on securing their own needs as well as attempting to satisfy those of their opposers. When the satisfaction of needs is maximized, the opportunity for a mutually agreeable conclusion is maximized as well." (Gilbert 1997: 74)

It should be noted that mutual goals fall into two general classes:

(1) *mutually achievable goals*, or goals that Susan and Bill can cooperatively achieve without conflict. These may or may not actually be that similar, so long as they can both be met simultaneously.

(2) *similar or like-minded goals*, which often go along with *similar beliefs*. This occurs, for instance, when Bill and Susan both belief something and have a related goal. For instance, they both have the goal of getting more money and believe getting a raise is the way to do it. Yet, these goals may not be mutually achievable, especially if there is competition for raises in the company, and Bill and Susan both know only one raise will be given in their division.

Both kinds of mutual goals can be important to uncover. The former is obviously important, since maximizing mutual goal satisfaction is a common goal of participants in coalescent argumentation. The latter is also important, though, since it provides the participants with a common base upon which to build an understanding of the parts of each other's hierarchies that are not so readily understood.

23. Coalescent Argumentation Therapy, Stage III: *Empathy and Merging*

Once Bill and Susan have agreed, to the extent possible, on the goal/belief network, and have isolated mutual beliefs and goals, they can move on to stage III, the attempt to actual merge their positions, to "coalesce". The groundwork is already laid in their mutual points of agreement, and on their pragmatic points of possible mutual satisfaction. But now comes the really hard part—finding points of coalescence between the areas of their networks that are *not* already compatible.

Stage III requires empathy. The goal here is to understand the other's goal/belief network well enough to almost feel it is one's own. Bill and Susan cannot possibly merge their conflicting views on widget making unless they can

truly come to understand each other's positions. This may well require a more detailed and careful analysis than I have provided in figure 3. They may have to analyze Bill's reasons for believing X's widgets are good, which will involve his beliefs about what a widget is good for, as well as his goals and motivations for being interested in widgets in the first place (some of which are face goals or broad financial motives). For Susan, who knows more about widgets than Bill, this may require recalling how it was when she was only a beginner at widget design. Or it may involve simulating in her own networks the processes going on in Bill's head, even if she has really never seen the world like that at all. Bill, on the other hand, will have to actually learn something about the more technical aspects of widget making.

Empathy is difficult, yet absolutely crucial to coalescent argumentation. If the conflicting position is not understood, it will not be possible to merge it with one's own. Gilbert describes this empathic stage as an act of sheer will. One must "project oneself into another's position". "The quality of listening and observation required for the empathic comprehension of another's position is no mean feat; it demands, at least temporarily, the suspension of the drive to persuade or convince, i.e., to win, in favor of the desire to agree." (Gilbert 1997: 111)

Of course, complete empathy will not always be possible, but then complete merging of the two positions will not usually be achieved either. We can only ask that Bill and Susan try their best, and they may be quite happy to settle for partial coalescence, such as a partial merging, or a failed attempt at merging that nonetheless results in better understanding between them. Gilbert points out as well (Gilbert 1997: 113) that even if no real merging materializes at all, the coalescent procedure might still eliminate many points of disagreement as not really central to the argument, providing the participants with a more focussed view of their disagreement.

24. Conclusion

In summary, there are three main stages to coalescent argumentation therapy:

(1) Model each other's goal-belief networks:

(a) design private versions.

(b) reveal public versions.

(c) negotiate a working version.

(2) Find compatible goals and beliefs:

(a) find mutually held beliefs and goals.

(b) find mutually satisfiable goals.

(3) Empathize and coalesce:

(a) project into remaining incompatible parts of your partner's network.

(b) find ways of merging these beliefs and/or goals with your own.

Of course, one can practice coalescent argumentation without explicitly drawing PCT diagrams, or any other technical diagrams for that matter. Still, using such diagrams may provide the structure needed to learn basic coalescent skills, before learning to practice the skills without the diagrams (since actually drawing large, complex diagrams is rarely practical, especially in typical day-to-day arguing). Also, the general principles of coalescent argumentation are independent of the particulars of perceptual control theory, so it could perhaps still be practiced even by those with a very anti-PCT view of beliefs and goals. Even so, I think PCT provides a very compatible framework in which to talk about coalescent argumentation.

In uncovering so much of the goal hierarchy, the arguing partners can concentrate, not on "defeating" their partner's attempts to achieve their goals, but instead on allowing the other participant to achieve their dominant goals in some other fashion than that assumed in the subordinate argument. Argument proceeds from some initial disagreement, to filling in of the other's goals, to mutual satisfaction of those goals "coalescing" into a happy resolution.

Of course, you cannot expect to resolve all arguments with this method. The point is not that disagreement is unreal and simply a byproduct of the misunderstanding of goals—if only we understand each other we will all agree. Rather, the point is that coalescent goal-centred argumentation will enable the participants in even the most intractable arguments to strip away many of the unnecessary obstacles to agreement and, even if agreement is never actually reached, develop a better understanding of each other's positions, and thus inevitably a better understanding of their own.

This work could perhaps be best extended by further developing the Bill-Susan example so that it is clear exactly how coalescence could be achieved by them. Beyond that, it would be quite useful to see how two actual arguers embroiled in a real argument react to being taught such a system of diagram drawing and modification as a way of improving their argumentation skill. Perhaps some would just scoff. Perhaps others would not be up to the challenge of having to empathize with someone they feel is irrational. The hope would be that most, however, would at least learn somewhat better communication skills, listening more often to their opponents, empathizing, slowing down and not trying so hard, and hopefully learning to argue by understanding instead of by confrontation.

References

Aristotle (*c*. 335?322 BC, 1941). "Nicomachean Ethics." In R. McKeon (Ed.), W.D. Ross (Trans.), *The Basic Works of Aristotle*. New York: Random House, 935-1126.

Ashby, W.R. (1956). *An Introduction to Cybernetics*. New York: John Wiley & Sons.

CSG (Control Systems Group). *http://www.ed.uiuc.edu/csg/*.

Farrell, P.S.E., J.G. Hollands, M.M. Taylor and H.D. Gamble (1999). "Perceptual Control and Layered Protocols in Interface Design." To be published in *International Journal of Human-Computer Studies (Special Issue on Perceptual Control Theory).*

Gilbert, Michael (1997). *Coalescent Argumentation.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

McFarland, R.L., W.T. Powers and R.K. Clark (1957). "A general feedback theory of human behavior." *University of Chicago Counseling Center Discussion Papers* 3(18).

Powers, William T. (1973). *Behavior: The control of perception.* Hawthorne, New York: Aldine Publishing.

Randall, A.F. (1994). *Perceptual Control Builder.* Toronto: NTT Systems (internal report).

Randall, A.F. and M. Martin Taylor (1995). *Reorganizing Control Networks in the Recognition of Syntax* (with M.M. Taylor), Toronto: Defence and Civil Institute of Environmental Medicine (internal report).

Taylor, M. Martin, *et al.* (1989). "Response timing in layered protocols: a cybernetic view of natural dialogue." In M. M. Taylor (Ed.), *The Structure of Multi-modal Dialogue.* Elsevier Science, North-Holland, 159-172.

Taylor, M. Martin, *et al.* (1999). "Editorial: Perceptual Control Theory and its application." To be published in *International Journal of Human-Computer Studies (Special Issue on Perceptual Control Theory).*

van Eemeren, Frans, *et al.* (1993). *Reconstructing Argumentative Discourse*. Tuscaloosa: The University of Alabama Press.

Wiener, Norbert (1948, 1961). *Cybernetics: or Control and Communication in the Animal and the Machine* (2nd Ed.). Cambridge, Massachusetts: MIT Press.