

## INFORMATION, ACTION AND PERSONS

By John Perry

[This essay is Chapter 9 of *Identity, Personal Identity and the Self*. Indianapolis: Hackett Publishing, 2002]

### 1. INTRODUCTION

In Essay 8 I put forward an account of personal identity that goes as follows. There is a theory, the person theory, we have about human beings, and that we use to explain human thought and action, including language, memory and intention. This theory allows us to explain and predict what people will do in the future based on the present and past. We think that there is some causal connection among the stages of human beings that accounts for this theory working as well as it does. The unity relations for persons, that is the relation that obtains between human stages that are stages of the same person, is a matter of being causally connected in this way, whatever it is.

What I had in mind by "the person theory" is what David Lewis called "folk-psychology," the set of common sense principles that we apply to explain and predict human behavior. This includes our doxastic concepts: belief, recognition, knowledge, know-how, habit, and so forth; our volitional concepts: decision, deliberation, choice, intention, and the like; the various pro-attitudes: want, desire, preference, lust, the like our catalog of emotions, and much else. It seems to me that this body of common sense principles is an amazing intellectual accomplishment, and perhaps, even with all that science has brought us, it is still , together with common sense physics, the most useful knowledge we have for navigating through an ordinary day.

Still, it is not a scientific theory, and it hasn't brought us scientific knowledge. As Lisa Hall puts it,

Contrary to the Popperian ideal, it is a shallow and loosely-organized system with extremely vague conditions of application...these

features severely limit the value of common sense psychology as a tool in scientific research. However, they do not give us reason to question the system's empirical adequacy. On the contrary, the very qualities that make common sense psychology unsuitable for scientific purposes provide grounds for thinking that the theses to which it commits us are probably true (Hall, 1993. P. 41).

For our common sense psychology to serve the purpose I have in mind within a theory of persons and personal identity, it does not need to be a scientific theory, or a theory of any sort, in the usual sense. The practices of explanation must, however, be capable of providing coherent and naturalistically plausible explanations and predictions to a sufficient degree to suggest an underlying causal principle of its success (see Hall, 42ff.). In want to address one central puzzle about how this can be so, which I'll call the circumstantial nature of the attitudes. My approach may seem to stray from Hall's vision of the common sense system as a sprawling, loosely connected set of principles encircled by special cases and hedges. Instead I simplify the underlying principle of explanation to a simple caricature:

belief + desire motivates action.

My defense for dealing with this caricature, aside from intellectual limitations, is that by showing how this caricature could be coherent while at the same time make sense from a naturalistic point of view, I hope to provide a springboard for further work that will produce a picture more adequate to the sprawling reality.

## 2. *HOW CAN CIRCUMSTANTIAL ATTITUDES EXPLAIN?*

The view that our cognitive states are both causes of our action, and provide reasons for it, is central to common-sense theory of persons. If a belief and a desire *motivate* an action, then the belief state and desire state *cause* the action, and their contents *rationalize* the action. A belief and a desire rationalize an action if the action will promote the satisfaction of the desire if the belief is true.

The causal role and contents of our attitudes must *mesh*. Beliefs and desires must cause actions they rationalize.

Beliefs, desires, hopes, fears and other attitudes seem to be located in the heads of the people that have them. Our attitudes are accessible to us through introspection. For example, Vice-President Cheney can tell that he believes Bush to be a Republican just by examining the "the contents of his own mind"; he doesn't need to investigate the world around him. We think of beliefs and desires as being caused at certain times by events that impinge on the subject's body, specifically by perceptual events, such as reading a newspaper or seeing a picture of an ice cream cone or having someone tell you, "I'm a Republican, just like my dad." These attitudes can in turn cause changes in other mental phenomena, and eventually in the dispositions to behave and the observable behavior of the subject. Seeing the picture of an ice cream cone leads to my desire for one, which leads me to forget the meeting I am supposed to attend and walk to the ice cream shop instead. All of this seems to require that attitudes be states and activities that are localized in the subject where they do their causal work.

Propositional attitudes, however, seem essentially relational in nature. They are "directed at" propositions and at the objects those propositions are about. These objects are the subject matter of the belief or desires, the objects facts about which determine if the belief is true, or the desire satisfied. They may be quite remote from the minds of subject. An attitude seems to be individuated by the agent, the type of attitude (belief, desire, etc.) and the proposition at which it is directed. It seems essential to the attitude reported by (1),

(1) Cheney believes that Bush is a Republican

for example, that it is directed towards the proposition *that Bush is a Republican*. And it seems essential to this proposition that it is about Bush. And it is facts about Bush that will determine whether Cheney's belief is true---the very same facts that determine whether (2)

(2) Bush is a Republican

is true. But how can a local mental state or activity of a person essentially involve some other individual in this way?

The view I advocate is a two-tiered view. One must distinguish between cognitive states and attitude properties. Cognitive states are local, causally interact with one another, and explain how muscles and bodies move. Attitude properties, are circumstantial, reported by attitude reports, explain actions by revealing their rationale and are individuated by propositions---by the objects they are about, and what their truth requires of those objects.

Cheney's belief that Bush is a Republican is a particular cognitive state he is in, a brain state. We can model it as a sentence of mentalese written in a box, or, to use the vocabulary I prefer, an association in his mind between ideas. It makes sense to ask when he came to be in that state, and about its causes and effects.

The way Cheney and those ideas are set into the world, the causes of the formation of the ideas, ways they are related to perceptions and these perceptions are related to object and properties, the way the ideas cause bodily movements, and the effects those movements have on situations in the world, that determine what the ideas are *of* and the content of the belief they comprise. That is, it is not only the internal state and its structure, but the wider circumstances of the belief, which determine the proposition that Cheney believes. Believing the proposition that Bush is a Republican involves Cheney's brain states being related to a certain person, Bush, and a certain property, being a Republican. It is a property of Cheney's that he believes Bush is a Republican, but not a "state" in the normal sense of the term, which connotes something internal and local. It is conceivable that Cheney could be in the very state he is in, but have a different attitude property in virtue of being in that state. This is the lesson we learn from various "twin" cases.

Paralleling the distinction between cognitive states and attitude properties is one between two kinds of actions: *executing movements* and *bringing about results*. I'll call these "executions" and "accomplishments." The execution of a movement is a basic action, something that can be caused by internal brain states affecting central nervous system, causing certain muscles to contract, etc. The philosophical paradigm is the movement of a finger. An accomplishment is something one does *by* executing a movement in certain circumstances, as one might bring it about that an elevator comes to the first floor, by moving one's finger in the right circumstances, while it was just in front of the elevator button. One can also decide to think or calculate or daydream, and do so without the train of causation going outside, so we'll include those as executions, although they don't feel very much like movements---movements of the mind perhaps, but that seems merely a metaphor.

On the two-tiered, view, then, we have cognitive states causing executions, and propositional attitudes rationalizing actions. But why should these two different levels, the local, causal states, and the circumstantial, rationalizing attitude properties mesh? What coordinates them? There must be some sort of "pre-established harmony" between these levels, for our cognitive life to make sense. To take an extreme case of non-coordination, suppose Cheney's belief that Bush was a Republican, and his desire to have Bush elected, causes him to give a million dollars to the Democratic Party. If this happened, of course, our belief in coordination would lead us to try to close the gap between motivation and action. We would look for further beliefs. Perhaps Cheney believes the more money the Democratic campaign has, the worse it will do. Or further desires. Perhaps Cheney wants to have Bush win a very close race. We would not take the situation at face value, thinking, "Oh well, this is just one of those cases when the internal causal states weren't quite aware of their content." Common sense psychology is deeply committed to harmony between cognitive states and attitude properties; they must somehow mesh.

### 3. MESHING

Here is a first shot at the meshing principle:

- (1) If believing  $P$  and desiring  $Q$  cause  $A$ , and then  $A$  promotes  $Q$ , given  $P$ .

This needs to be sorted out, in light of the two-tiered strategy. (1) has the attitude properties in the antecedent, causing action, and propositions in the consequent, rationalizing it. We need states in the antecedent causing, and circumstantially determined propositions in the consequent, doing the rationalizing. Here is an improved version.  $B$  and  $D$  are internal cognitive states,  $M$  is a type of movement,  $P$ ,  $Q$ , and  $R$  are propositions:

- (2) If belief states  $B$  and desire states  $D$  cause executions of movement  $M$ , and, on a specific occasion one believes  $P$  in virtue of being in  $B$ , desires  $Q$  in virtue of being in  $D$  and brings it about that  $R$  in virtue of executing  $M$ , then on that occasion,  $R$  promotes  $Q$  given  $P$ .

Here I use "desire" and "belief" loosely and broadly. "Pro-attitude" and "doxastic pro-attitude" might be better. In our example, the total explanation of Cheney's making the movements necessary to make out the check to the Democratic Party and send it would include bits of know-how (he knows how to sign his name), recognition (he thinks *that* is his checkbook), attunement to background constraints (he doesn't act like he was in a space-ship, but rather as if he were in a normal earthly gravitational environment) and so forth. I don't think, however, that we need assume that the totality of Cheney's beliefs or cognitive states are involved in the causation and rationalization of each act. This is all discussed in more detail, although probably still not enough, in (Israel, Perry, Tutiya, 1993).

(2) helps to make the problem clearer. Since what a person believes in virtue of being in  $B$ , desires in virtue of being in  $D$ , and accomplishes in virtue of

executing  $M$  will depend on the circumstances one is in, it would seem like a bit of a miracle, for all that is said in (2), if the antecedent guaranteed the consequent. To see that it need not be a miracle, we need to get inside the circumstances that determine the contents and results on a specific occasion to see how they might interact in a coordinated way.

Let  $Y_b$  be a function that assigns contents to belief states, considered in the widest sense, relative to agents, times and circumstances. If

$$Y_b(a,t,F,B) = P$$

then if  $a$  at  $t$  is in belief state  $B$  and in circumstance  $F$  then  $a$  believes proposition  $P$  at  $t$ . Similarly for  $Y_d$  and desires in the broadest sense.

Let  $P$  be a relation between agents, times, circumstances movements and results. If

$$P(a,t,F,M,R)$$

then if  $a$  executes  $M$  at  $t$  in circumstances  $F$ ,  $a$  at  $t$  brings it about that  $R$ . Then I propose:

- (3) If belief state  $B$  and desire state  $D$  cause movement  $M$ , then:  
 Whenever there is an agent  $a$ , time  $t$ , circumstance  $F$ , and propositions  $P$  and  $Q$  such that;
- (i)  $a$  at  $t$  is in states  $B$  and  $D$  and circumstances  $F$ , and
  - (ii)  $Y_b(a,t,F,B) = P$ , and
  - (iii)  $Y_d(a,t,F,D) = Q$
- there are also circumstances  $F'$  and a result  $R$  such that
- (i)  $P(a,t,F',M,R)$ , and
  - (ii)  $a$ 's bringing it about that  $R$  at  $t$  in  $F$  promotes the satisfaction of  $Q$

Let's see how this works in a simple case. I see a glass  $c$  before me which I take to have water in it. Next to it is another glass,  $c$ ; which I take to be empty. I want to quench my thirst. I reach out, pick up  $c$ , and drink from it. Here my

perception causes me to go into an internal state *B*. The content of the state is *P*: that *c* is a certain distance and direction from me, and has water in it. That state combines with the desire *D*, the content of which is *Q*: that my thirst be quenched. If my belief about the direction and distance of *c* from me are true, the result of these movements will be that the cup is brought to my lips and the liquid from it poured into my mouth. If my belief that *c* has water in it is true, my thirst will be relieved.

Now suppose *c'* had been in front of me instead, looking just the same as *c* did, and *c* is off to the side and looking empty, just like *c'* was in the original case. In that case, I would have believed that *c'* was a certain distance and direction from me, that *c'* had water in it and that *c* did not. Same internal states, different propositions believed. Since the internal states are the same, the movement will be the same. But here is where the pre-established harmony kicks in. If the a glass is a certain distance and direction from me, and is the glass I see, and the glass that gives rise to my belief state, it will be the glass my belief is about; but if it is that distance and direction from (and if I am properly attuned to the length of my arms and the relevant gravitational forces and the like), that movement will bring the very glass I see and formed the belief about to my lips. The facts that produce this result are not exactly the same ones that determine which glass my belief is about (hence the difference between *F* and *F'* in the conditions), but if the *F* facts obtain, so will the *F'* facts. The contents of my beliefs are different in the second case, since they are about *c'* instead of *c*. But the results of my action are also different, since I pick up *c'* instead of *c*. The attitudes and the states mesh, even though we kept the states fixed while the attitude properties changed with the circumstances.

Here is a twin-earth example. Twin-*a* is in the same internal cognitive states as *a*, and a pretty similar situation as *a*, except that he is a zillion miles away on twin-earth looking at glasses *d* and *d'*. They have twater instead of water on twin-earth (See Putnam, 1975), but they call it "water," since that's the

word for it in Twin-English. So, speaking in English, Twin-*a* is in the same state as *a*, but Twin-*a* believes that glass *d* has twater in it, while *a* believes that glass *c* has water in it. Why doesn't that cause a problem? Again, circumstances to the rescue. The reason Twin-*a* is thinking about twater instead of water is that when he grew up he saw and drank a lot of twater, used it to wash his face, and learned to recognize the difference between twater and tmud and torange-juice and tmilk and so forth. So he has an idea of twater, and internal state or aspect of an internal state that is *of* twater. This very idea would have been of water if he had been on earth. (The importance of water to our biology and the world in general may make some of these counterfactuals implausible, but I'll press on.) When he learned Twin-English, he associated the word "water" with the substance he had learned to recognize, drink, wash with, and so forth.

So, if Twin-*a*'s belief his true, given its circumstantially determined content, the glass in front of him will be full of twater. Although the movement his belief and desire cause are the same as those the same states cause in *a*, things still work out. He wants twater to his lips, and that's what he gets, if the glass is full of twater.

If condition (3) is met, the causal roles and the contents of the attitudes will mesh. But what exactly does (3) put conditions on? Basically on the Y functions, that assign contents to states, and the causal structure of the part of the world in which these states occur, which is reflected in the P relation.

A naturalistic account of cognitive states and their contents will have to explain how these conditions are met. The hypothesis is that cognitive states evolved as a system of harnessing information to control action, and that common sense psychology is built upon appreciation of this fact.

#### 4. *THE REFLEXIVE/CIRCUMSTANTIAL STRUCTURE OF INFORMATION*

I'll call any circumstance, thought of in terms of the information it carries, a signal. A signal is the fact that *A has  $\phi$* , for some object *A* and property  $\phi$ . *A* is the

*carrier* of information and  $\phi$  is the *indicating property*. Suppose you are at the dentist. The dentist holds an x-ray taken of some of your teeth. He points to a discoloration on a certain tooth. "That shows that you have a cavity," he says. The x-ray is the carrier, the pattern discoloration is the indicating property, having a cavity is the *indicated property* and the information is the proposition that your tooth has a cavity. Note the circumstantial structure; the signal is a local feature of the x-ray; the information is a proposition about your tooth, in virtue of the circumstance that the x-ray was taken of your tooth.

We get from the signal to the information in two steps. Information is basically what one part or aspect of the universe (the signal) shows about some other part or aspect (the subject matter, in this case, your tooth and the property of having a cavity). This is possible only because events are constrained by laws of nature, or as I prefer, because of its more liberal, common-sense, loose, and non-reductive connotations, by the way the things happen. The information carried by a signal is what *else* things have to be like, for the signal to have occurred, given the way things happen. Given the way things happen, an x-ray has a discoloration like that only if the tooth to which the relevant part of the x-ray was exposed has a cavity.

This means that a signal will carry the information that  $P$  if there is *some* constraint, some principle of how things work, such that given that constraint, the  $P$  has to be the case for the signal to occur. It is useful, however, to have an explicitly relative concept:

$S$  has the informational content that  $P$  relative to constraint  $C$  if given  $C$ , it has to be the case that  $P$  for  $S$  to occur.

One advantage of the concept of informational content is that it allows us to consider false or non-factual constraints as well as true ones, which is often useful, as we will see below. That is, we'll think of constraints as states of affairs, some of which are facts (or propositions, some of which are true, if one prefers).

Let's return to our example. Can we say that the spot on the x-ray shows that your tooth has a cavity, relative to a constraint *C*, where *C* incorporates the principles of x-rays and decay relevant to how x-rays work? Not quite. After all, if it were a constraint that every time an x-ray had such a discoloration *your* tooth had an x-ray, you would be pretty miserable. The constraint is that every time an x-ray has a spot like that, *the tooth to which it was exposed* has a cavity. All we have so far is

The spot on the x-ray shows that the tooth to which *the x-ray itself* was  
exposed has a cavity

I call this *reflexive* information, because the informational content of the x-ray is about the x-ray *itself*. That is, the proposition *that the tooth to which the x-ray was exposed has a cavity*, ascribes a property not to you, or your tooth, but to the x-ray, the property of having been exposed to a tooth which has a cavity. With reflexive information what a signal shows is something about itself; one aspect of the x-ray, the pattern of discoloration, shows something else about the x-ray, something about its history.

To get from reflexive information to information about your tooth, we need to bring in circumstances, namely, the circumstance that your tooth is the very one to which the x-ray was exposed:

*Given* that the x-ray was exposed to your tooth, the spot on the x-ray  
shows that your tooth has a cavity.

This is a proposition about your tooth, not one about the x-ray. Our informational content is no longer reflexive. On this concept, the x-ray doesn't show something about itself, but something about the rest of the world. I call this *incremental information*: We've got the x-ray. What *else* does the world have to be like, *given* these circumstances and the state of the x-ray?

We now have three concepts of information to work with:

A signal *S* has the reflexive informational content that *P* relative to constraint *C* iff according to *C* the signal occurs only if it is the case that *P*

A signal *S* has the reflexive informational content that *P* relative to constraint *C* and given circumstance *F* iff according to *C*, given *F* the signal occurs only if it is the case that *P*.

A signal *S* carries the information that *P* iff *S* has the informational content that *P* relative to some factual constraint *C* and factual circumstance *F*.

It would be pleasant to find a direct link from informational to belief, but things are not so easy. We'll actually need a fourth concept of information, but before getting to that we need to look at the structure of action.

##### 5. *THE REFLEXIVE/CIRCUMSTANTIAL STRUCTURE OF ACTION*

I'll use "actions" for types of acts; acts are unrepeated events involving an agent executing some movement at some time. As I mentioned, actions come in two basic varieties, executing movements and accomplishing results; we do the latter by doing the former.

Suppose for example that I am typing. I move my right forefinger in a certain way; that's executing a movement. By doing that I depress the j-key; that's an accomplishment. By depressing the j-key, I bring it about that a "j" appears on the screen; that's another accomplishment.

Accomplishments we think of as bringing about results, so the canonical form of an action report is:

X brought it about that *P*.

This regimentation makes it much easier to see the connection between accomplishments and the propositional attitudes that motivate them.

What I bring about when I execute a movement depends on constraints and circumstances. Moving my finger as I did was a *way of* depressing the j-key

in the circumstance that my finger was poised over the j-key, according to mechanics and principles about the way fingers and computer keyboards work.

An action  $A$  by  $a$  at  $t$  is a way of bringing it about that  $P$  in circumstance  $F$  relative to constraint  $C$  iff according to  $C, F$  &  $a$  performs  $A$  at  $t$  implies that  $P$ .

$a$  performs  $A$  at  $t$  iff  $A$  is movement  $M$  and  $a$  executes  $M$  at  $t$  or  $A$  is an accomplishment *bring it about that*  $P$  and  $a$  brings it about that  $P$

The term "accomplishment" usually suggests that one has done what one wanted to do. It doesn't have that sense within this theory. One might say it includes the somewhat ironical sense: By tripping over the books on the floor of my office, I spill the coffee I was holding on a report on my desk that I just printed out. A colleague says, "My, look what you've accomplished, and it's only 9 a.m." On our usage, ruining the report is an accomplishment of mine: a result brought about by something I did, however unintended.

An act is *successful* or *unsuccessful* relative to a goal, or, if one wants to sound more scientific, a chosen end-state. The act is successful relative to the goal if the goal is one of the things the act brings about. An action  $A$  can be assigned success conditions relative to a goal and a constraint. The success conditions will be those circumstances in which the action will be a way of bringing about the goal, according to the constraint. For example, suppose I want to mail a letter. According to the way the post-office works, walking up to a metal box on a street corner and dropping a stamped letter in the box is a way of mailing a letter, *if* the metal box is an official mail depository box. So that is the success condition of my action, relative to the goal of mailing a letter and the constraints arising from the way the post office does things.

## 6. *HARNESSING INFORMATION*

Intelligent design often means designing something so that information guides action in a way that produces the result one wants. A mousetrap is a relatively

simple device that illustrates this. Old-fashioned mouse traps really weren't traps, but mouse-killers. Now one can purchase a more humane device. There is a little tunnel, with an opening at one end. The tunnel is mounted off-center on a little fulcrum, so the whole thing acts like a teeter-totter. If nothing is in the tunnel except a little peanut butter (the recommended bait), the open end will be on the floor and the closed end, where you put the peanut butter, will be about  $3/8$  of an inch above the floor. The open end has a door that swings shut from above. When the open end is on the floor, the door stays open, perched above the opening, its center of gravity just slightly to one side of directly above the hinge. When the open end rises even a little bit, the center of gravity of door passes over the hinge and it swings shut. When a mouse enters and goes down the tunnel to get the peanut butter, as it passes above the fulcrum the weight shifts, the closed end goes to the floor, the open end goes up in the air, and the door shuts. The mouse is trapped but unharmed, and can be released humanely in the neighbor's yard.

Here is how one can look at this as an information-using device. The change of weight as the mouse passes the fulcrum is a signal. It contains the information that there is a mouse in the tunnel, relative to the constraint that only mice will enter the tunnel. It is the job of the user to put the mousetrap in some place where this constraint will hold. The closing of the door is an action, which will be successful if there is a mouse in the tunnel, relative to the goal of trapping a mouse, and the constraints about mice and plastic, e.g., that mice cannot walk through plastic doors or open them, at least not those of this kind. The device is constructed so that the very event that carries the information that certain circumstances obtain (the change of weight) is also a cause of an action that will be successful in just those circumstances. This gives us a new concept of content, which I'll call "pragmatic content." The pragmatic content of a state is the success conditions of the actions it causes. Pragmatic content is relative to architecture, circumstance, and goal. One wants the informational carried by a state to guarantee the pragmatic content of the state. Then the actions the state

causes will be successful. It is pragmatic content, rather than informational content, that is our model for the content of beliefs.

The information that the door will close is what we call *architectural* information. It is incremental information rather than reflexive, in that it is not information about the carrier itself. But it is information about something that is part of the same device, as opposed to something that is outside the device. The architecture is the key to harnessing information; the design creates the information by creating causal connections.

The sort of scales that they have or at least used to have in doctor's offices provides another example of architectural information. You step on the scale, and the doctor moves some little weights on a little bar; the position of the weights when the bar balances shows your weight. He lowers another bar to the top of your head. That bar is attached to a vertical bar that slides in a holder which is fastened to the scale just in front of the platform on which you stand to get weighed. An arrow on the vertical bar points to calibrations on the holder, and this shows how tall you are. Let's suppose that you are  $z$  and that the pointer points to 6' 3." Then we have:

The position of the pointer has the reflexive informational content *that the person's whose head is stopping its downward movement is 6' 3" tall*, according to constraints about metal and human bodies.

The position of the pointer has the architectural informational content *that the person who is standing on the weight platform is 6' 3" tall*, according to constraints about metal and additional constraints about human bodies and the architectural constraints, and given the architectural circumstance that it is attached to the scale.

(The additional constraint is that the way human bodies are built; this is required to establish that the person who stands on a weight platform of the scale will be the person whose head stops the downward movement of the height bar. If

humans were built like the leaning tower of Pisa, this might not be so; the height bar might stop at the head of the person standing on the next scale.)

The position of the pointer has the incremental informational content that *z is 6' 3" tall*, according to constraints about metal and human bodies, given the circumstance that *z* is the person who is standing on the scale.

Architectural information is very important in understanding how human technology works, for technology often depends on vast systems of signals that contain information about each other according to a myriad of different constraints. Particularly important are technologies that allow for the *flow of incremental information*. Consider, for example, the pattern of pixels on my television on the day of the 1980 NFC Playoff game, which carried the information that Dwight Clark had caught a pass from Joe Montana, defeating Dallas, and sending San Francisco to the Super Bowl for the first time. The process started at Candlestick Park, where Clark caught the ball. The camera that was focussed on him---call it *c*---went into an internal state  $\phi$  that carried the reflexive information that the person on whom *it* was focused had caught a football, relative to constraints about people, footballs, light, cameras and so forth. It carried the incremental information that Dwight Clark had done so, relative to those constraints, and given the fact the camera was focussed on him. Slightly simplifying, the camera caused a satellite *s* to go into state  $\gamma$ , which had the reflexive information that the camera sending signals to *it* was in state  $\phi$ , relative to constraints about cameras, wireless transmission, and satellites. The satellite's being in  $\gamma$  carried the architectural information that *c* was in state  $\phi$ , relative to the (architectural) circumstance that *c* was the camera sending signals to it. It also carried the architectural information that the person on whom *c* was focussed had caught a football, relative to all the constraints listed so far. Relative to all those constraints, and given that *c* was focussed on Dwight Clark, it carried the information that he had caught the ball. Finally, in our very simple

system, the satellite  $s$  caused a pattern of pixels on my television screen, which carried the reflexive information that the satellite that transmitted the pattern to  $it$  was in state  $\gamma$ , relative to constraints about satellites, wireless transmission, cables, televisions sets and the like. The pattern of pixels carried the architectural information  $s$  was in state  $\gamma$ , given the architectural fact that  $s$  sends signals to my screen. The pattern of pixels inherits the architectural information that the fact that  $s$  was in state  $\gamma$  carried, that the person on whom  $c$  was focussed caught a football, relative to the constraints so far plus those relevant to the step from  $c$  to  $s$ . And finally the pattern of pixels on my screen carries the incremental information that Dwight Clark caught the ball, relative to all the constraints listed so far, given the architecture of the system, and given the fact that  $c$  was focussed on Clark.

A flow of information system like this is designed to preserve incremental information. The form in which it is preserved is important in the last step; the display of pixels carries the information in a way that I, the person who pays for the satellite hook-up and buys the advertised beer, can recognize. I don't care much about the system that lies between my set and the field. A cable rather than a satellite, or a transmitter and antenna system, would be equally satisfactory from my point of view, if the picture were as good. The connections and constraints contain all sorts of information relative to all sorts of constraints, but it is the flow of incremental information at which the whole system is aimed. I'm not interested in the state of the satellite, or of the camera; I'm interested in Dwight Clark and the 49-ers.

## 7. *INDIRECT CLASSIFICATION AND ATTUNEMENT*

The statement

$s$  shows that Dwight Clark caught the ball

can be given a relational analysis; one thing, the signal, has a certain relation, showing (carrying the information) to another, the proposition that Dwight Clark

caught the ball. There is another way of looking at it, however, which I call "indirect classification." The picture is that we are really not classifying a pair of objects by a relationship that holds between them, but classifying one object, the signal, by the state that it is in, identified in a very roundabout way. It is a bit like

My car is the color of that ripe tomato.

We can look at this as classifying a pair of things, my car and the tomato, as having the same color. Or we can look on "the color of that ripe tomato" as a somewhat roundabout way of identifying a color, which is being predicated of my car, the tomato really just serving at an aid to identifying the color. This possibility of looking at the same statement in two ways will prove crucial in seeing how propositional attitudes are able to do the work that they do.

Suppose, having seen the pattern of pixels on my television, I say, "Dwight Clark caught the ball." What is my evidence? I saw it on TV. That is, the picture, i.e., pattern of pixels, on my television screen shows that Dwight Clark caught the ball. The pattern of pixels is a local event, one that could be described without any mention of Dwight Clark. It seems like a rather roundabout way of describing the state of my screen, to bring in the activities of some individual a forty miles away connected with my television by a hunk of metal in space.<sup>i</sup> I am indirectly classifying the screen in terms of what it shows, relative to the constraints and connections that I pay my monthly Dish TV bill to be able to exploit. The reason I do this is as part of an explanation of why I believe that Dwight Clark caught the ball. The incremental information, in terms of which I describe the picture on my television, is the very same proposition that I come to believe, by watching it. It is a bit of a miracle, if one thinks about it like that.

In this example, I am *attuned* to two things. First, I know how to interpret pictures on a television screen; I know what they depict. If the pattern of pixels had been like that, but I had been watching a movie, I wouldn't have taken it to show anything about what happened at Candlestick Park, but I would have

taken it to depict a man catching a football. Secondly, I am attuned to the way television broadcasting systems work. If I know that a game is being televised live, I take the event depicted on the screen to have happened; I take the pixels on the screen to carry information about the events that a camera, or something like it, is trained on at the site of the game.

To say that I am attuned to the way the television broadcasting systems work is not to say that I could list or state or understand the nature of the constraints and connections involved. I might not even know whether the television I'm watching uses cable, satellite, or antenna. To say I am attuned is to say that my beliefs track the information carried by the events on my screen, given the connections and according to the constraints that are part of the system. It is this attunement that is necessary to *use* the system, although of course it would not suffice to repair it or to build it in the first place. The common sense concept that handles this is *know-how*. When one knows how to drink a glass of water, or ride a bike, or get the news on TV, one is able to do things that achieve goals that depend on all sorts of connections and constraints that one doesn't have the concepts or the need to explicitly believe. One has a positive doxastic attitude, included under our wide use of "belief" in section 2, but not the sort of thing one ordinarily calls beliefs.

My being attuned to the system is a bit like my becoming part of the system. Each link in the system, the camera, the satellite, and my television, are set up to carry the same incremental information, depending on different constraints and connections. So am I, just like another link. The pattern of pixels causes me to believe that Dwight Clark caught the ball and to say so. What I believe and what I say is just the very incremental information that was carried by the camera, the satellite and the TV screen. Here the relational point of view helps; we think of the incremental information as a proposition to which each of these events is related in different ways.

Let's return for a moment to the mousetrap. A mouse crawls in the trap and the door shuts. Why did the door close? We might say, "Because the trap knew a mouse had crawled inside." This is clearly metaphorical. Can we more carefully say, "Because the trap was in a state that showed that a mouse had crawled inside?" What sort of explanation would that be? It seems to me that we have the same sort of dual aspect explanation that we saw is typical of attitude explanations. We have given both a reason and a cause. To see the reason, we look at our statement as relational. Assuming the goal of catching mice, the proposition shown is the success-conditions for the action taken, the door closing. On the other hand, taken as classifying a state indirectly, it the statement gives cause of the door closing, the movement of the trap. We are giving an explanation in the context of an understanding of the goal of the trap, and the assumption that it embodies an information-harnessing design. We give both a reason and a cause, and they mesh.

#### 8. *INFORMATION, ACTION AND INTENTIONALITY*

Our concept of "carrying the information that" is factive: if  $s$  carries the information that  $P$ , then it's true that  $P$ . It doesn't provide a very promising candidate for a helpful analysis of belief, our paradigm propositional attitude, which is quite non-factive.

Our concept of "having the informational content that" is a bit more promising. The constraints relative to which a signal has informational content need not be true, so even if an actual signal  $s$  has the informational content that  $P$ , it might be false that  $P$ . For example, it might be that every so often a glitch in the network makes pixels appear depicting a football player catching a ball when he didn't. The ball actually hit something out of bounds, ending the play, and then bounced back into the player's outstretched hands while he still hung in the air. Cosmic rays caused by sunspot activity erase the part of the signal the camera sends to the satellite registering the bounce. It's very rare. It would cost

a lot to fix. So the system gets by with informational content, rather than information.

Birds of various species became attuned, over centuries or millennia, to the constraint that the path to any object that is clearly visible is unobstructed. This hasn't been true since humans started making transparent windows; a certain number of birds fly into the windows and die, because all their visual states have is informational content rather than information. Still, most birds do well in spite of being attuned to mere informational content.

I became attuned to the information contained in broadcasts of football games over many decades before the problem with cosmic rays showed up. Even though no cosmic rays were involved on the NFC playoff day, all I really had available from my TV was the informational content that Dwight Clark caught the ball. I was attuned to a false constraint, because the broadcasting system had been designed in terms of one.

If we regard beliefs as states that have informational content, rather than carry information, we can allow for false beliefs, in virtue of attunement to constraints that may be quite reliable, but are not exceptionless. This is a start, but it doesn't seem to provide much leverage. Lots of false and fallible beliefs seem to arise in ways that don't have much to do with attunement to false constraints.

Things get more promising if we look at how fallible informational content can be in the context of our concept of success conditions. Consider our mousetrap. There are really a lot of ways that the trap can be moved so that the door shuts, other than a mouse crawling in. It usually happens several times just setting the trap in place. A cat might rest its paw on top of the trap. There might be an earthquake. A marble might roll in the trap. A mouse might crawl across the top of the trap on his way to the floor from the pasta shelf. And so on. The constraint that whenever the center of gravity shifts across the fulcrum, there is a mouse in the tunnel of the trap isn't even close to being true. Still, it doesn't have

to be true for the mousetrap to be a pretty good one. No great harm is done if the door shuts for some other reason. Even the old-fashioned mouse-killing mouse "traps," with the blade that snaps when a cheese pedal wiggles, are still a successful product, even though the results of a "false positive" can lead to a crushed toe or finger or paw. We can imagine attunement to a myriad of overlapping, not all that reliable, constraints, as long as the benefits of success are high and the costs of failure low.

Designing mechanisms to act on the basis of constraints that are not true, invariable connections among types of events, but merely somewhat probable, even in the most favorable conditions, can be a good strategy. Another simple example is the automatic pencil sharpener. You can trigger these things with all sorts of lead-pencil-shaped objects, including automatic pencils, ballpoint pens, pencils stuck in with the eraser first, and so on. It's not an entirely unpleasant way to kill time. Even in a well-designed office, an absent-minded academic can be counted on to falsify the constraint a couple of times a month. Still, it's a good product, for somewhat lazy and somewhat responsible people.

Let's call devices that are attuned to constraints that are only sometimes right "information-content harnessing devices." My philosophical hypothesis is: (i) that human beings are naturally occurring information-content harnessing devices; (ii) our system of using propositional attitude reports as explanations of actions (including internal acts, such as theoretical and practical inferences) is a system of dual-purpose indirect classification, which involves attunement to the way humans work as information-content harnessing devices; (iii) our concept of persons and personal identity reflects this attunement.

(I also think this whole system has gone somewhat berserk, as a scheme for survival of the species or anyone's genes, but on the other hand much of what we value in human life is a product of this craziness. I won't defend expand on these deep thoughts in this essay.)

This hypothesis obviously needs to be worked out in considerably more detail; perhaps it even deserves to be. However, for the purposes of this essay I will simply consider one rather weighty objection to the whole.

#### 9. *PAINS, PLEASURES, AND ORIGINAL INTENTIONALITY*

One can object to this idea as follows. The broadcasting system, the mousetrap, and the pencil sharpener may well be describable in terms of informational content and success conditions. Dennett has show how we can take the "design stance" and the "intentional stance" towards lots of artifacts and naturally occurring processes (Dennett, 1987), and these applications of the "informational stance" are in the same boat; perhaps the informational stance is a species of the intentional stance. But, as Searle has emphasized, humans don't merely have this sort of "attributed" intentionality; our beliefs and desires and hopes and fears really have content (Searle, 1992). When we take the intentional stance the content ultimately comes from *us*. We are describing the operations of the network, the mousetrap, and the pencil sharpener in terms of what we use them for. If our goal were to ruin ballpoint pens, the success conditions of the automatic pencil sharpener's beginning to spin would not be that a pencil has been inserted, but that a ballpoint pen had. It's just up to us, and that isn't the case with our beliefs and desires. Their content is not up to us.

The problem is that pragmatic content, unlike information or informational content, is relative to goals. Our attribution of success conditions to the mousetrap door closing, or the pencil sharpener beginning to spin, are clearly based on our goals for creating and using the mousetraps and pencil sharpeners, not the goals of the mousetraps and pencil sharpeners themselves. What sense can we make of the attribution to systems of goals that are intrinsic to the system, rather than coming from the outside to a system? We cannot start with desires, in attempting to show that our system of propositional attitudes can be seen naturalistically. We need to build a bridge from goals that can be

attributed to an organism on the basis of its own situation, to our ordinary concept of desire.

Evolution teaches us that the fittest survive, and pass on their genes. So there is something naturalistic about the goal of surviving. Actions are successful if they promote the goal of survival. So the success conditions attributed to our actions ought to be relative to that natural goal.

This seems to work fairly well for some animals. The chicken sees a bit of grain and pecks. The success condition of the act was that an edible was at the ends of the arc of the peck, so that's the relevant informational content of the visual/brain state caused by the pattern of light reflected from the grain on the ground.

But we can't carry this too far. In the first place, modern biology seems to emphasize that the evolutionary importance of our fitness is for the survival of our genes, which seems one step further removed from our desires than our own survival. The goal of dispersing our genes doesn't seem to be a very good candidate for directly anchoring our propositional attitudes. My present motion towards the jar of chocolate chip cookies seems to be motivated by a desire to have the taste of a chocolate chip cookies flood the tastebuds in my mouth.

A more promising strategy is to ground our desires in the natural goals of avoiding pain and pursuing pleasure. Avoiding pains and pursuing pleasures are as natural as goals get. We might think of these as natural goals further grounded in promoting propagation of one's genes. For some animals, there are states it is like something to be in; some of these are pleasurable, some are painful. Animals try to avoid getting into, or try to remove themselves from painful situations, and try to get themselves into, or stay in, pleasurable ones. At one time, we may speculate, this natural tendency served the goal of gene dispersal well enough, there was a close correlation between actions that promoted the dispersal of our genes and pleasures, and actions that didn't promote them and pains. Eating is more fun than starving or being eaten and

procreating is fun while having the urge to do so without an outlet isn't. This is probably why states it is like something to be in were originally exploited by evolution. Pleasures and pains evolved as internal signs of what would help the genes and what might not.

The hypothesis, then, is that it is pains and pleasures that provide the intrinsic goals, that can naturalistically ground the concept of success for our actions, and hence the concept of pragmatic content, and hence the concept of doxastic states. That is, the basic architecture of doxastic states have informational content and, together with desires, cause actions, was originally in the service of the natural desires of avoiding pain and attaining pleasure. These desires are natural in two senses; pains are unpleasant, and we want to avoid them, and the opposite for pleasures. And during a long period of evolution, avoiding pains and pursuing pleasures was a good strategy for dispersing one's genes.

The link between painful and pleasurable activities, and activities that promote or don't the survival of one's genes, has largely been broken by civilization. Most of what we eat is bad for us. Successful acts of procreation mostly hasten the end of the world through overpopulation, so really don't serve anyone's genes.

If we are well brought up, our desires don't center on our pains and pleasures. Some ability to forgo pleasure or endure pain in certain circumstances evolved in the service of gene propagation, for example when one's offspring are at stake, or when a little pain now means a lot of pleasure in the future, or vice versa. Human invention, knowledge, culture, superstition and the rest have taken charge of this mechanism. I struggle to produce a philosophical essay when I could be eating a cookie. John Etchemendy chairs meetings trying to find a way to balance Stanford's Medical School's budget, when he could be writing a philosophical essay---or eating a cookie---or both, for that matter.

## 10. CONCLUSION

The speculations of the last section may be somewhat laughably naive to the biologist or ecologist, but I'm satisfied if naiveté is the only big problem. The point is to come up with a picture, even if it's a bit of a cartoon, that suggests how humans can have goals that can be naturalistically explained as intrinsic to them, and not just attributed to them by designers, theorists and the like. We need goals to have success; we need success to have success conditions; we need success conditions to have pragmatic content, and pragmatic content is my candidate for seeing doxastic states both have informational content and, in conjunction with desires, cause actions that, if the content is true, will succeed. And, I claim, seeing doxastic states and desires in this way is the key to solving the meshing problem in a naturalistic way.

---

<sup>i</sup> To tell the truth, neither dish satellite, nor even cable, were available in Palo Alto in 1980, where I saw a pattern of pixels on my neighbor's television that showed that Dwight Clark caught the ball in the NFC playoff game with Dallas, an event known, at least in the Bay Area, as "The Catch." The dish satellite story seems to me one of the more amazing cases of the flow of information, and the catch is one of the few historic events almost universally regarded as pleasant that I have watched on TV, so I decided to combine them ahistorically.