

# Examples of Awareness of Intentions<sup>1</sup>

Brian Epstein

Tufts University, Medford

Michael D. Ryall

University of Toronto

June 12, 2021

<sup>1</sup>Rough examples for the paper

We consider the problem of Brian's Toddler. The toddler, Individual 1, is faced with a decision as to which of two toys, labeled A and B, to obtain. The issue is whether A or B is best.

## Parsimonious game theoretic treatment

The parsimonious game theoretic treatment is shown in Fig. 1. The uncertainty of individual 1 with respect to which toy is actually best is represented by including an initial move by Nature at time  $t = 1$  – i.e., the two possible states are A-Best and B-Best, one of which is true and the other of which is counterfactual. The bold lines indicate the choices of the players. Individual 1 is then faced with a choice as to whether to get A or get B. As illustrated by the dashed line connecting 1's decision nodes, 1 does not know with certainty which is the true state but, as indicated, believes it is most likely that A is best. Therefore, 1 chooses to get A in  $t = 2$ . As a result, 1 obtains A in  $t = 3$ .

The features to note are: 1) the world simply presents 1 with a decision; 2) although 1 is uncertain about which toy is best, he is aware of the counterfactual possibilities – indeed, 1 knows everything about the game, including what will happen as a consequence of his actions; 3) all of 1's cognitive processes associated with the decision are compressed into the act of making a decision. The decision could be elaborated as one involving probabilistic beliefs on the part of 1, but this is not necessary. For whatever reason, at the time of his decision, 1 believes (with some measure of uncertainty) that A is best. Given these beliefs, and a desire for possessing the best toy, 1 chooses to get A.

## A four-phase decision process

Our goal is to expand the parsimonious treatment of individual 1's cognitive process to include some of the features debated in the philosophy literature. With an eye toward adopting the unawareness formalism of game theory, we begin by elaborating what 1 is aware of about the world, how he thinks about his decision at a given moment in time, and how this evolves dynamically.

First, what is the decision process? One useful disaggregation of the decision process for our context is: **Phase 1** Individual finds himself in a state of the world in which a decision is called for (here, we count deciding not to pursue the decision further as, itself, a decision); **Phase 2** Individual selects focal elements for the decision analysis, analyses them, and forms an intention;

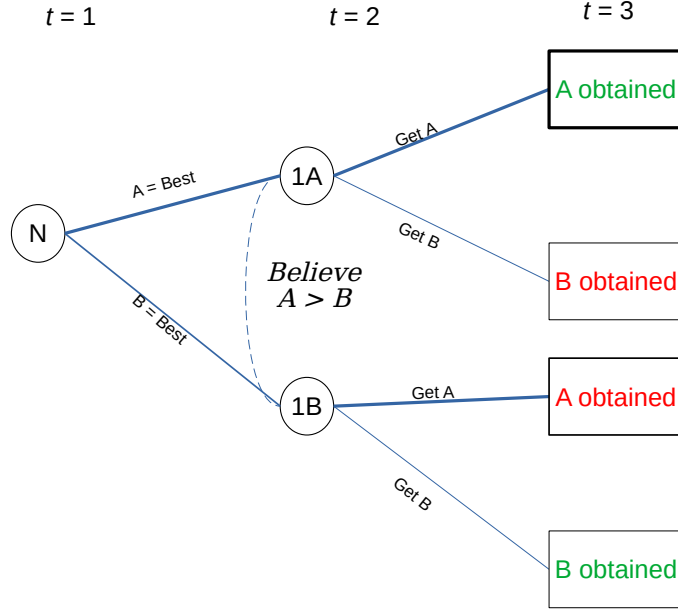


Figure 1: Brian’s Toddler, parsimonious game-theoretic treatment **Phase 3**

**Phase 3** Individual forms a plan to effectuate the intention; **Phase 4** Individual acts in accordance with the plan.

In our example: **Phase 1** Individual 1 is presented with a choice of A or B; **Phase 2** Individual 1 consults beliefs and forms an intention to obtain A; **Phase 3** Individual 1 plans to effectuate the intention by crawling to the location of Toy A and picking it up; **Phase 4** Individual 1 crawls to Toy A and picks it up.

Note that Phase 2 may involve limiting attention to a subset of what we might call Individual 1’s “field-of-awareness” (FOA for short, which could also abbreviate field-of-attention if we wish). That is, given all the elements of the world of which 1 is passively aware at the start of the decision process, he may decide to limit his attention to a strict subset of elements thought to be decision-relevant. Similarly, the individual may call to mind (make himself aware) elements that expand his FOA. The plan formed in Phase 3 may be simple, but it may also be state-contingent (and, typically, will be for the satisfaction of complex real-world intentions). A key assumption is that the plan, once formulated, will be enacted as long as the world unfolds in a way that is “sufficiently consistent” with it (the precise meaning of this will need to be worked out).

Therefore, we begin by defining the states of the world at  $t = 1$  as  $S_1^0 \equiv \{(A, n), (B, n), (A, y), (B, y)\}$ .

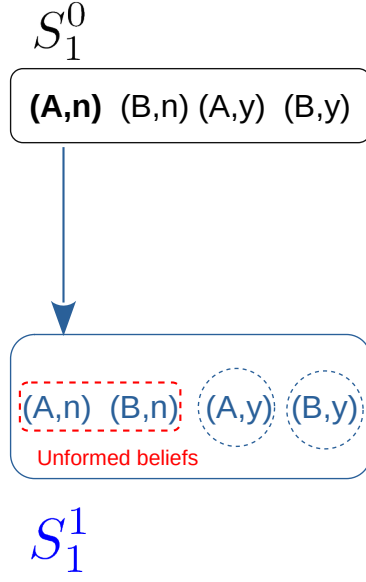


Figure 2: Toddler's Field-of-Awareness (blue)

This follows our earlier notation where  $S$  indicates a state space, 0 indicates Nature, and 1 indicates the time period. The states are  $(x_1, x_2)$  where  $x_1$  is which toy is truly best and  $x_2$  is whether a toy is obtained or not (note: for completeness, the state should elaborate *which* toy is obtained, but the example will work without the extra clutter).

Begin with the case of a fully aware decision maker. Assume the state is  $(A, n)$ . In this state, Individual 1's FOA is  $S_1^1 = S_1^0$  as illustrated in Fig. 2 by the blue part of the diagram. The red dashed line indicates which states 1 believes *could be* true. Beyond being aware of the states about which 1 is unsure is the case, 1 also knows the counterfactual states  $(A, y)$  and  $(B, y)$  (where the dashed circles around them indicate that 1 is also aware that were, e.g.,  $(A, y)$  to be true he would know it). We also indicate that, in this state, 1's beliefs have not yet been formed (in a Bayesian setting, are uninformative).

In addition to what the individual is aware of with respect to the present situation, he also has thoughts about the future. Tracking all four phases, our individual considers the future unfolding as shown in Fig. 3. Moving to Phase 2, which is an analysis and commitment step, 1 may discover that either A or B is preferred. Here, analysis takes one time period. Once the analysis is complete, and the intention is formed to obtain the preferred toy, 1 formulates a plan. The planning takes one time period.

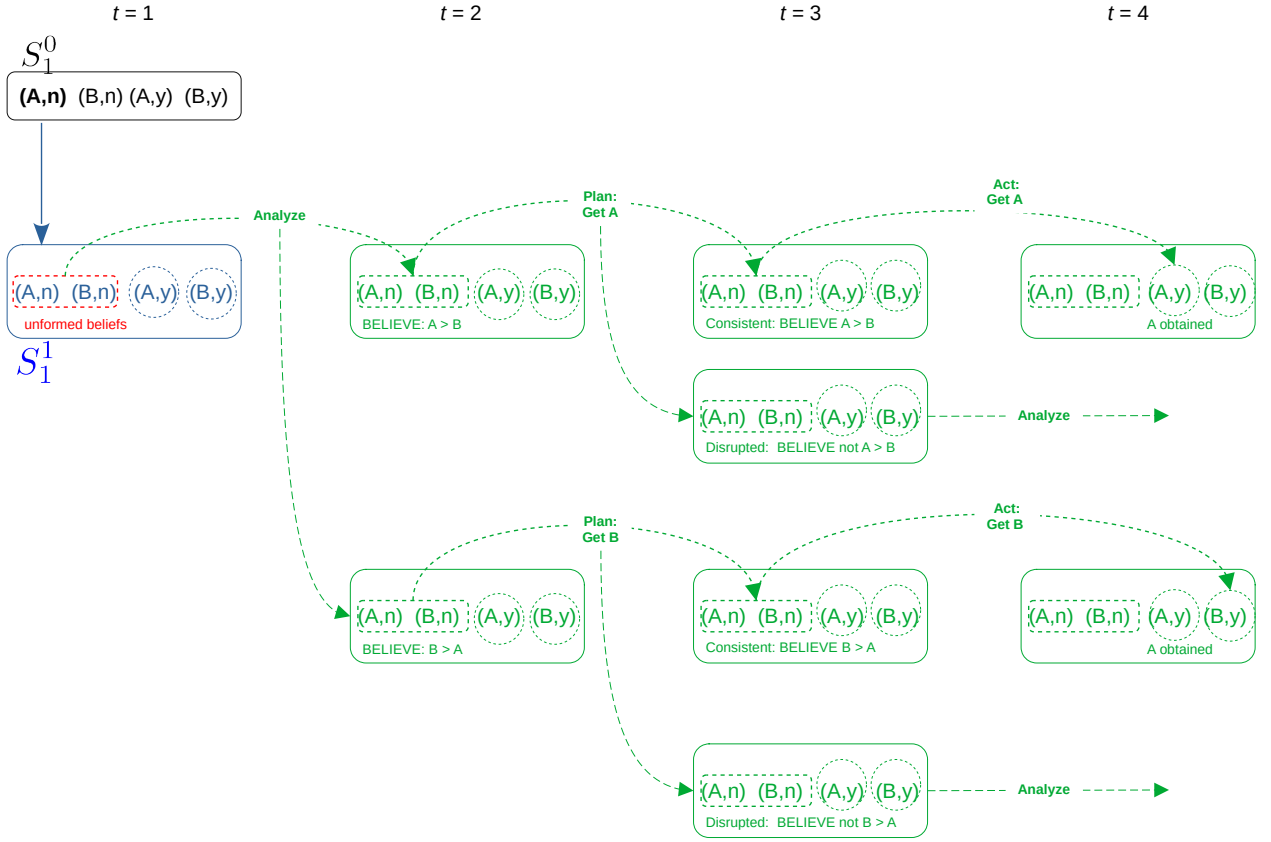


Figure 3: At  $t = 1$ , Toddler thinks about the future (green)

Here, we make an important assumption: while the individual is thinking, analyzing, planning, and acting, the world evolves. Now, from  $t = 1$  to  $t = 2$ , this evolution is implicit as occurring and possibly affecting the analysis. The individual considers things from the perspective of his FOA, which may evolve during the interval between periods. Still, the outcome of that process is that 1 comes to believe either that A or B is best. What happens from  $t = 2$  to  $t = 3$  is distinct from the analysis phase. During this interval, a plan to get A or to get B is being shaped. However, real-world events may intrude upon the process in ways that disrupt the plan. This is illustrated in the figure. For example, 1 may be planning to get A yet experience something that happens to indicate that B is really best (e.g., Toy B starts making a ringing sound). When the plan is

disrupted, we assume that the decision maker must backtrack to an earlier stage – either a new analysis or a new planning cycle. In the figure, we show that the process reverts to a new analysis stage.

Then, if the state of the world in  $t = 3$  is consistent with the plan plan, individual 1 proceeds to act. So, in the top row of individual 1’s projected future, following the plan to Get A, a state of the world occurs in which 1 continues to believe that A is best and, hence, 1 acts to obtain A. In  $t = 4$ , therefore, the final state is  $(A, y)$  (and 1 knows that this is the state).

## A two-phase reduction

Before continuing with the example let us compress Phases 1-3 into one – i.e., there is a thinking, deciding, and planning phase followed by an acting phase. This seems to give us a sufficient level of elaboration to investigate the kinds of issues in which we are interested. Refer to the compressed phase as “planning.” The revised version of Fig. 3 is shown in Fig. 4.

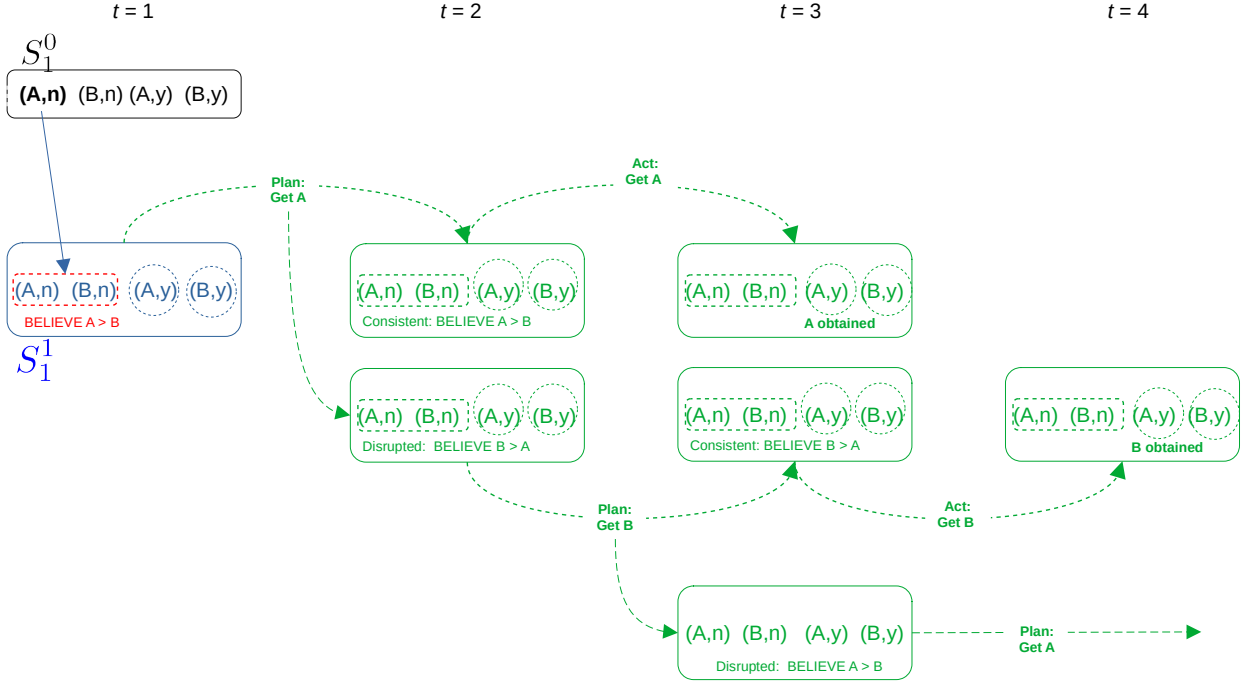


Figure 4: The two-phase version

All of this is by way of setting up the dynamic analysis. Based upon the, now, compressed analysis, decision, and planning phase, 1 develops a plan to get A. Fig 5 shows the situation at the start of  $t = 2$ . The world has evolved to  $S_2^0$  in which state  $(A, n)$  continues to hold. Reality is illustrated on the top row. Individual 1's act "Plan Get A" happens and leads to  $S_1^0$ .

Keep in mind that the state labelled  $(A, n)$  in  $S_2^0$  is not identical to the one so labelled in  $S_1^0$ , even though the state spaces appear to be the same. First, the states in  $S_2^0$  includes the information about the sequence of preceding states (i.e.,  $s_1^0 = (A, n)$ ) as well as about the acts that caused them (i.e.,  $a_1^1 = \text{plan get A}$ ). Second, individual 1's state of mind has changed (and, remember, this is also summarized by the state). He recalls what he knew before,  $S_1^1$  as well as what action he took. This is indicated by the blue dashed line. Another difference is that he projects the decision process from the present into the future. This is indicated by the green objects.

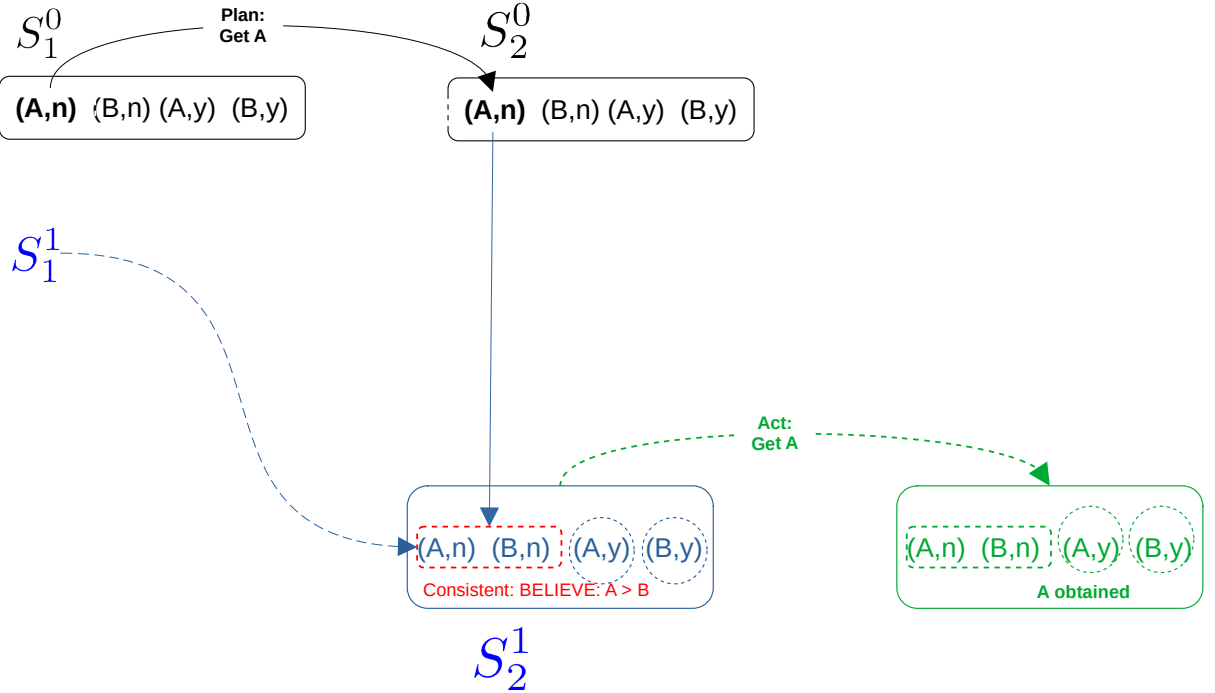


Figure 5: The plan proceeds without disruption

Having formed a plan to get A, nothing has happened to disrupt 1's decision to get A. His belief

remains that A is best. As we saw above, an alternative possibility was that an act of Nature, e.g., Toy B begins ringing, might have disrupted the plan. This emphasizes that what we are illustrating is one path of actual events analogous to the bold lines through the tree in Fig. 1. A diagram of all possible paths and FOAs would be quite complex to illustrate in a single diagram.

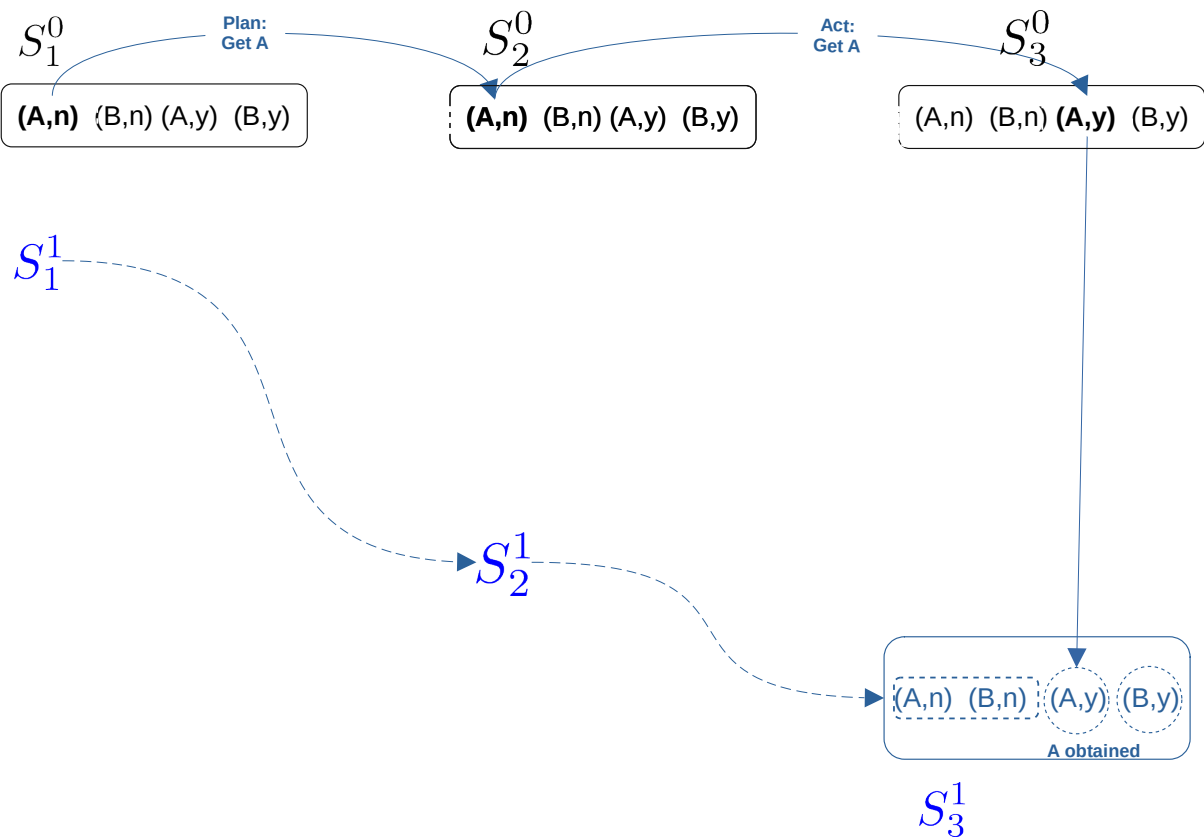


Figure 6: Decision and action resolve

Evolution to the final period,  $t = 3$  is shown in Fig. 6. At the conclusion, individual 1 obtains A and is certain that he obtained A.



## Can this be represented as an extensive-form game

So far, we have not written anything down that can't be shown as an extensive form game. This is not surprising given that, thus far in our example, individual 1 continues to have full awareness and, as well, his projection into the future is consistent with reality. One would need to be careful to ensure that feasible actions available to 1 would correspond to the logic of our assumptions. In particular, acts must follow plans and, presumably, some states could disrupt the plan which would leave 1 with one feasible act (restarting the planning process) or, if we include quitting the decision, two feasible acts. That game would be the one shown in Fig. 1 extended to give nature intervening moves (i.e., with the power to switch 1's assessment of which is the better toy) and adjusting the feasible acts as described.

Our diagrams give a more elaborate description of what is going on in 1's mind (recalled history, FOA, and projected futures). However, this comes at the expense of providing a simultaneous illustration of the entire set of possible paths as we have in extensive form game trees. Later, if we introduce unawareness, our FOAs will provide an expressive capacity not available to standard extensive form games.

The next step is to illustrate what can go wrong – how an even fully aware individual (or, perhaps, a fully aware individual in particular) can get stuck in a “paralysis by analysis” situation. Then, we can show why intentional unawareness can help (or hurt) the situation.

## An interesting connection

While thinking about this example, I came across the idea of OODA Loops (Observe, Orient, Decide, Act), which gained popular use in the military. The wiki discussion is [here](#). I also downloaded a couple of short articles about this into the repo lit file. The diagram of the OODA Loop is shown in Fig.

## References

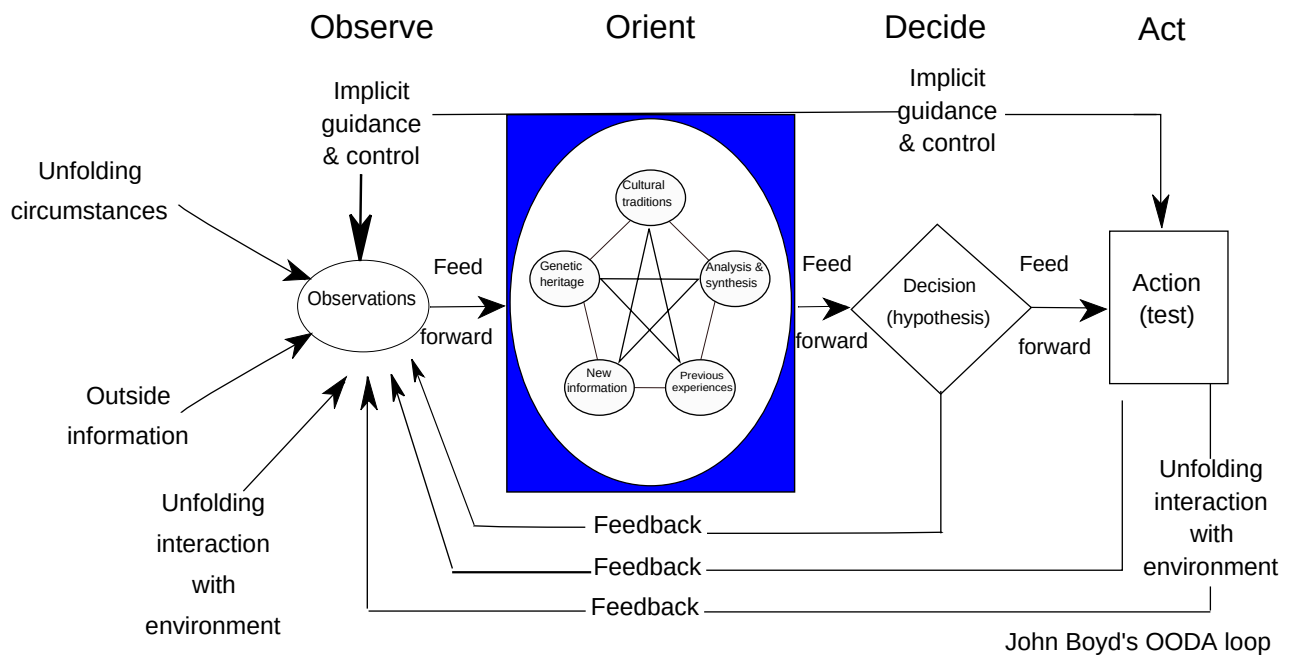


Figure 7: John Boyd's OODA Loop