An Attempt to Generalize AI Part 19: Wish-Fulfillment in Action Selection

By Paul Almond

24 August 2010

Website: http://www.paul-almond.com
E-mail: info@paul-almond.com

This is the nineteenth in a series of articles attempting an overview of how minds may work and how similar systems could be implemented in computers. The model is supposed to do most of the planning, projecting the system's behavioral history into the future, but there is no history of competence to start with and noise would also cause problems. To establish an appropriate behavioral history, an *action selection process* is needed to act externally on the system and push behavior slightly in the preferred direction. The action selection process is now being changed and the new process is described. The new action selection process involves artificially and modestly increasing a predicted evaluation function score so that the model will produce a narrative which tends to result in that higher score. Some of this narrative will involve good luck, but some of it will involve improved behavior. Providing that the score increase was sufficiently modest, the improved behavior will still be an improvement on its own terms. A prediction for an imminent output can then be obtained from the model and the output made, thereby contributing to an improvement in the behavioral history.

Table of Contents

Table of Figures

List of Abbreviations

AI	artificial intelligence
AGI	artificial general intelligence
BERP	basic, exploratory relevance process
EFS	evaluation function score
ERP	exploratory relevance process
RMP	relevance measurement process

1 Introduction

This is the nineteenth in a series of articles attempting an overview of how minds may work and how similar systems could be implemented in computers. A complete description of the cognitive model developed so far is provided in *An Attempt to Generalize AI – Part 15: A Complete Description*, available at http://www.paul-almond.com/AI15.pdf.[1]

(Readers already familiar with the existing action selection process, and who have an understanding of its potential problem of accessibility, particularly readers who have read the above article, may wish to go straight to Section 3: The New Action Selection Process: Wish-Fulfillment, on page 8.)

In this view of cognition, the hierarchical model itself is supposed to do most of the planning, projecting and abstracting the system's behavioral history into the future, but there is no history of competence to start with and imperfections in modeling and noise would also cause degradation of behavior. To establish an appropriate behavioral history, an *action selection process* is needed to act externally on the system and push behavior slightly in the preferred direction. Such an action selection process was previously described.

The action selection process relied on different output values being tried for an imminent output, to try to find a value that would give an improvement in behavior. Once found and used, the improved behavior would form part of the behavioral history, and predictions in future use of the action selection process would be made within the context of this behavioral history. This presented a possible issue of "accessibility" similar to that in Darwinian evolution. Darwinian evolution can only "access" improvements available through sequences of beneficial, random variations and this kind of system can only "access" improvements in the action selection process available through sequences of beneficial output value changes.[2] It is not certain that such improvements would always be available without diminishing returns.

The action selection process is now being changed and a new action selection process will be described in this article. The action selection process described here does not have to be used just with this particular cognitive model: It could be usable in other systems provided that they meet certain criteria.

[1] Almond, P., 2010. *An Attempt to Generalize AI - Part 15: A Complete Description*. [Online] paul-almond.com. http://www.paul-almond.com/AI15.pdf or http://www.paul-almond.com/AI15.doc.

[2] That could be misunderstood. It specifically means single output value changes *in the action selection process*. In reality, changing an output in the action selection process will tend to change the outputs predicted by the modeling system to follow after that.

2 The Previous Action Selection Process

The action selection process described previously works as follows.[3]

The modeling system is informed of inputs and outputs that occur and is capable of probabilistically predicting future inputs and outputs: In the proposed system this is achieved by means of a probabilistic hierarchy.

An evaluation function score (EFS) is continually computed and provided as inputs, so that the modeling system "knows" about past inputs corresponding to the EFS and can be required to predict future inputs corresponding to the EFS probabilistically, meaning that the modeling system is effectively predicting a future value of the EFS.

When an output is imminent with values of 0/1, each value is tried, and the modeling system is updated as if the output had occurred with that value.[4] The modeling system is made to update the model's probabilities and the predictions for a future EFS are obtained from the model. The output value which resulted in the higher predicted EFS is actually used as an output. The output is made and the model updated accordingly.

This way of selecting actions is crude, but it is important to note that it is not supposed to be the main process by which the system plans. The modeling system is basing its predictions on inputs *and outputs*. It is running a model of the world, which *includes the behavior of the system itself*. If the system has a history of competent or improving behavior, which conforms to a particular description, the model will extend and abstract that description into the future, just as it would extend and abstract other descriptions. In fact, if the system has a history of competent or improving behavior, it may seem that we do not need an external action selection process: Whenever an output was needed, we could just obtain the system's own prediction for it and use that to make the output. This will not work, however: The system does not start with a history of competent or improving behavior and even if it did, with nothing outside the model to impose any standard of preferred behavior, random noise in the model's computations would ultimately cause it to drift "off course". The action selection process described above was intended to deal with both of these issues by giving the system's behavior a small "nudge" in the preferred direction.

The action selection process is not the main planning process: That occurs in the model itself, in a way consistent with Metzinger's views of the "self" as part of a model.[5]

---

[3] Almond, P., 2010. *An Attempt to Generalize AI - Part 15: A Complete Description*. [Online] paul-almond.com. http://www.paul-almond.com/AI15.pdf or http://www.paul-almond.com/AI15.doc. pp.32-41.

[4] For simplicity, it will be assumed in this article that the system uses 0/1 values for input/output.

[5] Metzinger, T., 2003. *Being No One: The Self-Model Theory of Subjectivity*. Cambridge (MA): MIT Press. Metzinger, T., 2009. *The EGO Tunnel: The Science of the Mind and the Myth of the Self*. New York: Basic Books.

Because the action selection process is not the main planning process, nothing *profound* is expected of it, but even given that, it may still be rather crude. It can only find an improvement in behavior that is available by beneficially modifying a sequence of single output values. This limits the scope for improvement, and such improvements may not always be available: There is a possible issue of accessibility here. This may be analogous with Darwinian evolution, in which fitness can only be increased through a sequence of beneficial, random variations. A better action selection process is required which “nudges” the system in a preferred direction without the same potential issue of accessibility. This is the subject of this article.

3 The New Action Selection Process: Wish-Fulfillment

3.1 How the Action Selection Process Works

The modeling system is informed of inputs and outputs that occur and is capable of probabilistically predicting future inputs and outputs: In the proposed system this is achieved by means of a probabilistic hierarchy; however, it could be achieved by other means. Further, the modeling system is required to update the probabilities that it produces without regard to direction in time, so that informing the modeling system of the probability for an input/output in the future will affect predictions for other inputs/outputs, including earlier ones. (The hierarchical system of patterns and pattern instances described in this series of articles already meets this requirement.)

An evaluation function score (EFS) is continually computed and provided as inputs, so that the modeling system "knows" about past inputs corresponding to the EFS and can be required to predict future inputs corresponding to the EFS probabilistically, meaning that the modeling system is effectively predicting a future value of the EFS.

When an output is imminent, the following process is performed.

1. The modeling system is directly manipulated, with its probabilistic predictions for future inputs corresponding to future input of one or more EFS values being directly altered so that the expected EFS value(s) is/are slightly increased. (For example, if the modeling system's prediction is that, in one hour, the EFS value will be 26.0, we might manipulate the model directly to change this prediction to 26.3.)
2. The probabilistic prediction for the imminent output is obtained from the modeling system and stored.
3. The changes that have just been made to the model in Step (1) are reversed. (We have just been lying to the model about future predictions, so we do not want our lie, or its effects, to remain there.)
4. The imminent output is actually made with the most likely value, as indicated by the probabilistic prediction obtained from the model in Step (2). The modeling system's probability for this output is then updated to reflect the fact that the output is known, with certainty, to have occurred with that value. The modeling system then updates its other predictions, as usual.

and that is all there is to it. The process is actually simpler than the previous one: Because more of the process is absorbed into the model, there is less that has to be done explicitly outside it. (See Figure 1: The Action Selection Process, on page 9.)

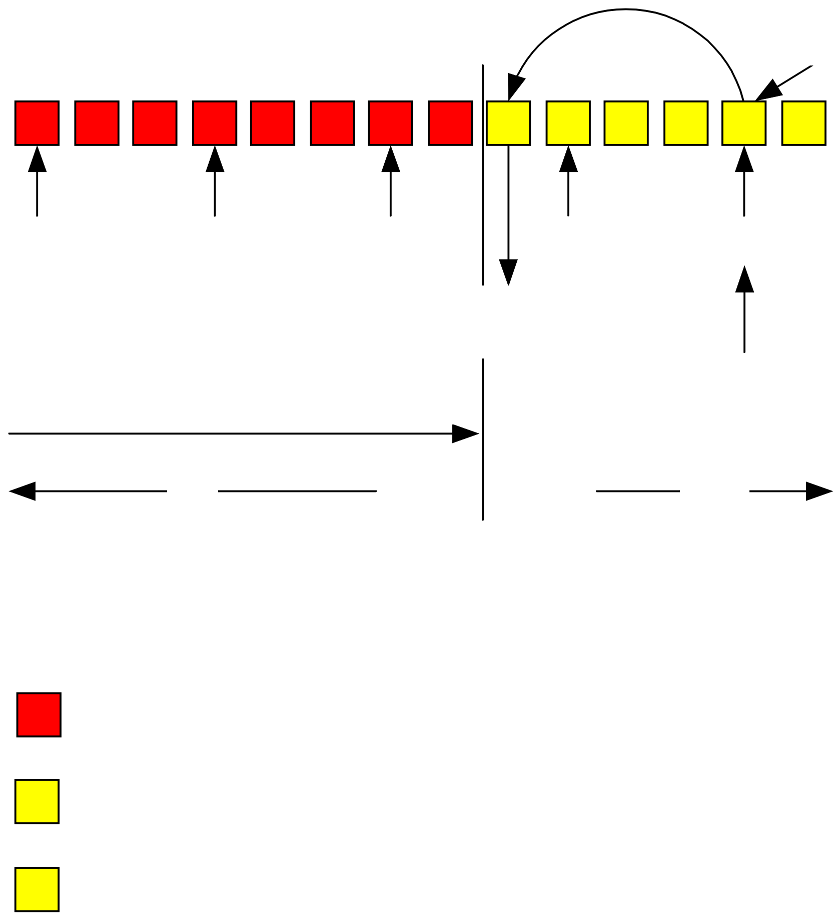

Figure 1: The Action Selection Process

3.2 *Why* the Action Selection Process Works

Suppose that the system already had a history of behaving competently and/or of improving behavior. A sufficiently powerful modeling system would have an automatic tendency to continue predicting this kind of behavior as part of its general modeling of the world. The model's predictions of the system's own behavior would be predictions of competent and/or improving behavior. When we needed to make an output, we would merely need to obtain a probabilistic prediction for that output and make the output accordingly.

This is not the case, of course. Initially, there is no history of competent and/or improving behavior and in any case, the issue of random noise would soon cause problems. This is dealt with by artificially increasing the expected EFS value(s).

The modeling system is effectively producing a *narrative* to fit what is known.[6] Given information about past inputs/outputs, it creates a probabilistically expressed narrative about the future, which fits in with the history of inputs/outputs. This also applies when the narrative is required to fit *future* events. Given high confidence information about past inputs/outputs *and* a future input, the modeling system will create a probabilistically expressed narrative of what happens from the present until this input is supposed to occur. If we manipulate future inputs to increase expected future values for the EFS artificially, then we are making the future look "better" in the model, and we are requiring the model to fill in the rest of the narrative to be consistent with this. The modeling system is being required to predict that, while retaining as much consistency with previous events as possible, *better things will happen* to match up with the better future.

You may think that "better things will happen" is suspect, and to some extent you would be right. "Better things" can be predicted to happen in two ways.

- The system's *own behavior* is better.
- The *rest of the world's behavior* is better: The system has "good luck".

For example, suppose the system were controlling a vehicle and it were required to avoid bumping into other vehicles. Artificially increasing predicted EFS values might create a narrative in the model in which the system's driving is improved, but it might also create a narrative in which other vehicles are less likely to get in the way.

Now, your first thought, on reading that (or on realizing it if it already occurred to you) may be to reject the very idea that we could do planning like this. The system's method of action selection is based on wish-fulfillment: finding a way for things to be better without distinguishing between improved behavior (which is what we want to predict, because we can use the predictions to make it happen) and lucky coincidences (which are useless to us, because they are unlikely to happen and we cannot make them happen.)

Let us take an extreme example of this. Suppose the system is supposed to earn money, and the EFS represents how much money it has at any time. For next Wednesday, we have an expected EFS value of 27. We want the system to improve its behavior. We increase the EFS value in the model to 27 *million*. Surely, the system will predict some fantastic improvement in behavior to explain such a high EFS value? Unfortunately, this

[6] The idea that the model contains a *single* narrative is actually a simplification, but one that I will use throughout this article.

is not what happens. The “better narrative” predicted by the modeling system involves the system buying a lottery ticket and then winning the lottery. When we obtain our prediction for the imminent output from the modeling system, it is merely the first stage of the process of buying a lottery ticket. This would clearly be useless. The “better behavior” is not better behavior at all: It just seems that way within the context of a narrative in which *everything* is better.

We need to remember, however, that the action selection process described here is not supposed to be the basis for *all* of the system’s planning. The main basis for that is supposed to be the modeling system itself, with its ability to project and abstract patterns of competent and/or improving behavior into the future. The action selection process is only needed to improve the behavior very slightly, to establish a history of improving, and ultimately competent, behavior.

In practice, we would not do anything as extreme as artificially increasing the expected EFS from 27 to 27 million. We would use a very small increase, so that the system’s performance could gradually climb upwards. With a more modest EFS increase, the model will be predicting a narrative with better behavior and good luck, but the amount of good luck involved will be less outrageous, and it will be more likely that the better behavior is better on its own terms: that it would be better even *without* the good luck. We can still use this behavior, when we obtain the prediction for the imminent output and make it, even though it was found within the context of a very slightly optimistic outlook on reality, with the system engaging in general wish-fulfillment. It does not matter if the EFS that actually results is lower than the one on which the narrative was based, on account of the real world lacking happy coincidences. We do not want this improved behavior in itself: It is wanted to establish a behavioral history that can be built on later.

With this kind of action selection approach, the modeling system is not restricted to finding improved behavior by modifying a single action. Instead, the full power of the modeling system is available to find a full narrative of improved behavior which fits into the context of the behavioral history. The payoffs should more than compensate for contamination with predictions of “good luck”, particularly when only a modest increase is applied to the expected EFS.

3.3 The Exploratory Relevance Process (ERP)

The cognitive model and artificial intelligence (AI) approach, as described in previous articles, uses a basic exploratory relevance process (BERP) or other exploratory relevance process (ERP) to ensure that patterns, and pattern instances of patterns, are selected for use in the hierarchy according to their usefulness in reducing uncertainty in predictions that are of interest.

For now, with the new action selection process, the BERP or other ERP remains essentially the same. The only difference is that the bottom-level pattern instances that interest us have now changed. Previously, we just wanted to decrease uncertainty in the prediction for the future EFS value(s) that were used in the action selection process. We now also want the prediction for the imminent output – the one we are about to make – and which is almost at the present time.

Relevance is externally assigned to the pattern instances corresponding to input of the future EFS values that are taking part in the action selection process and to the pattern instance corresponding to the imminent output. (In practice it is likely to be assigned to outputs due to occur shortly afterwards as well.)

One issue is that of the degree of *linkage* between the future EFS values and the imminent output: What if we are predicting the EFS values and imminent output with low uncertainty, but the EFS predictions are not having a strong effect on the imminent output? I would point out that there should generally be some degree of connection because input/output events are what determine the EFS values; however, it may be that more needs to be explicitly done to establish this link and this may be discussed later.[7]

The imminent output will, of course, be continually changing: Each time an output is made, a new output becomes the imminent output.

3.4 Possible Objections

Objection 1: Isn't all this just an attempt to get something for nothing? You just tell your system that things are going to work out better, and then you wait for them to become better miraculously.

Answer

This objection would be based on a misunderstanding of what is happening. The improved narrative, with its improved behavior, is not expected to come from nowhere. Planning has been absorbed into the model, and it is the model that is expected to generate the narrative which matches what it is told about reality. This will only work with a powerful modeling system: something which I have been trying to develop in previous articles in this series. Such a powerful modeling system will be needed for an artificial general intelligence (AGI) anyway, and can be assumed to exist in any model of human cognition, so it makes sense to assume its availability: Without such a model, we will not be explaining human cognition and we will not be making any AGIs.

[7] This issue of linkage could also have done to receive more consideration with the previous action selection process as well – not that it matters now.

Objection 2: The new action selection process would have the same problems of accessibility as the previous one. At any time, the action selection process can only cause an improvement which can be achieved by an increase in the predicted EFS.

Answer

There *is* still an accessibility issue – and there will be some kind of accessibility issue in any system – but it is *much* less serious now. The requirement that an improvement result in an increase in the EFS is a much more general requirement than the requirement that an improvement result from a change to a particular output value. It does not indicate *how* the improvement has to occur, and the modeling system is free to use many different ways of achieving this, though the model that it has will impose other limitations.

3.5 Speculation: Wishful Thinking in Humans?

In 3.2, I used the example of a modeling system which predicts an improved narrative involving a lottery win, and I suggested applying only a modest increase to the EFS to deal with this issue. This raises the interesting issue, and I mention this only speculatively, of whether "wishful thinking" in humans could be related to this. Maybe human planning also looks for a "better narrative", without distinguishing much between improved behavior and improved *luck*. This might conceivably cause something like wishful thinking, particularly if the EFS increase, or its equivalent, is greater for some people or in some situations.

4 Conclusion

An action selection process has been described which replaces the previous action selection process in the cognitive model and artificial intelligence (AI) approach that is the subject of this series of articles.

The new action selection process involves artificially and modestly increasing a predicted evaluation function (EFS) score or scores so that the model will produce a narrative which tends to result in the higher score(s). Some of this narrative will involve good luck, but some of it will involve improved behavior. Given sufficiently modest score increases, the improved behavior will still be an improvement on its own terms. A prediction for an imminent output can then be obtained from the model and the output made, thereby contributing to an improvement in the behavioral history.

The advantage that this gives over the previous approach is one of *accessibility*. The problem of accessibility, here, is analogous to the situation in Darwinian evolution, in which only improvements which can be made by one or more random variations, each of which itself causes an improvement, are accessible. In the same way, in the previous action selection process, only improvements which could be made by one or more single, beneficial alterations to outputs in the action selection process were accessible. Even though the action selection process is not the main planning process, as that occurs in the model itself, this was still a limitation that would ideally be removed.

In a previous article I mentioned the idea of "coming in higher up" in the action selection process.[8] The new action selection process achieves this, but it also provides a much more comprehensive solution to the problem of accessibility than would have been provided by a modified version of the output selection process intended to "come in higher up" while keeping the basic process the same. Having the action selection process absorbed more into the model itself ensures that generation of the improved narrative can occur throughout the hierarchy and the modeling system's full powers are available to find improving behavior as part of a general narrative of improvement.

[8] Almond, P., 2010. *An Attempt to Generalize AI - Part 15: A Complete Description*. [Online] paul-almond.com. http://www.paul-almond.com/AI15.pdf or http://www.paul-almond.com/AI15.doc. p.40.

5 Bibliography

Almond, P., 2010. *An Attempt to Generalize AI - Part 15: A Complete Description*. [Online] paul-almond.com. Available at: http://www.paul-almond.com/AI15.pdf or http://www.paul-almond.com/AI15.doc [Accessed 24 August 2010].

Ibid. pp.32-41.

Metzinger, T., 2003. *Being No One: The Self-Model Theory of Subjectivity*. Cambridge (MA): MIT Press.

Metzinger, T., 2009. *The EGO Tunnel: The Science of the Mind and the Myth of the Self*. New York: Basic Books.