A Comparison of Modern Synthetic Character Design and Cognitive Robotics Architecture with the Human Nervous System

Charles Jordan Gilliland
MIT Department of Brain and Cognitive Sciences
dinoriki@mit.edu

Abstract

We begin by discussing reasons for choosing artificial human life as a research goal, and then investigate the adaptiveness or appropriateness of three artificial life architectures to the design of artificial humans. Brooks’ Subsumption Architecture is examined, followed by Blumberg’s Synthetic Character Design, then by Yoon’s Affective Synthetic Character Design. Comparisons with the human Central Nervous System are made where appropriate as it is assumed that similarity to the human CNS is a good metric for evaluation of an artificial life architecture.

1 Introduction

In comparing the design of Synthetic Characters and Robots, henceforth collectively know as Artificial Life Forms (ALFs) with that of the human nervous system, there are several motivations in mind. The point is mainly to assume that if a human-like ALF is to be created, then the design will have to closely parallel that of an actual human. This is desirable because it is easier than building something that seems just like a human but has very little in common, and because if the design is based upon the actual inner workings of human beings, then a great deal of what is learned in designing the ALF will be applicable to understanding actual humans’ functioning, and vice versa.

So, we will assume that the ultimate goal of ALF design is to create an ALF that can interact with humans within the physical world and interact on a human level, i.e. as another human. This goal is by no means agreed upon or even mentioned by many of the people responsible for the designs to be discussed, but it nonetheless seems to be a well-thought-of but not-so-frequently mentioned goal at the heart most AI researchers and even laypersons that perhaps someday in the distant future, a fellow such as Lt. Commander Data (an “Android” or ALF from the popular TV show Star Trek) might possibly exist. Why is this goal so desirable? Because of the human need for companionship, a comfortable environment, and self-understanding. If ALFs are to be designed, then the implication is that hopefully someday many of them will be used, and if they are to be abundant, then they should have as positive an effect on the lives of the humans around them as possible. This is best accomplished if ALFs take human form, and adopt human practices.

A potential argument against assuming that an artificial human is the target of ALF research is that among the list of uses for ALFs there are many things that would not require a humanoid ALF, in fact there are many things that would require an ALF that is specifically not humanoid. To cast away further doubt, I will say that any ALF design must be flexible and robust. If any engineering lesson has been learned it is that both of
those properties are essential if countless hours of research are to be invested in a design. Thus by saying that the ultimate goal is a humanoid ALF does not preclude the possibility of various intermediate (or even more advanced) types, it simply states that whatever design is developed, it must be extensible to produce a humanoid ALF. This also means that we can compare the ALF design to what is known about the human design and not worry about steering off course, unless we actually craft the ALF design too tightly around the details of the human blueprints, thus losing generality.

Therefore, we will evaluate several existing ALF designs, basing our judgment on how easy it would be to adapt the system to humanoid design, while realizing that the systems may not have been designed with construction of humanoid ALFs in mind. Therefore, statements made will focus on the ALF architecture itself as opposed to specific implementations of ALFs. First we consider Brooks' Subsumption Architecture [Bro86] followed by Blumberg’s Synthetic Character Design [Blu96] then Yoon’s Affective Synthetic Character Design [Yoo00].

2 Subsumption Architecture

Brooks’ Subsumption Architecture [Bro86] approaches ALF design by creating a number of layers of ALF control, starting with the lowest layer, layer 0, and placing each new layer immediately above the most recently added layer. Each layer can subsume, or suppress the outputs of, any layer below it, but may not alter layers above it in any way. Furthermore, all layers have access to all sensor data, and all layers can produce motor output.

An unenforced rule (meaning not an a priori requirement, but a practical result) of the architecture is that the lowest layers handle the most primitive competency of the ALF and higher layers perform more advanced operations. The idea is that failure in a higher layer will allow lower layers to continue operation, and that lower layers are more fundamentally important because they tend to handle things like survival whereas higher layers allow the ALF to accomplish desirable things given that the lower layers are functional.

Part of the motivation for the Subsumption Architecture lies in its practicality for the actual construction of robots, in that the layered architecture makes certain engineering issues simpler, such as by providing computational parallelism by allowing the layers to run asynchronously. Another engineering strength of the architecture is that it is relatively simple in terms of the connections between layers and thus the engineering problem is simplified. This explicit desire for a simple architecture is also rooted in Brooks’ notion that an ALF need not be complex to exhibit complex behavior, that if the environment the ALF is situated in is complex, that the interactions of the ALF with the environment can result in the desired behavioral complexity.

In comparing the Subsumption Architecture to the human CNS, little similarity is to be found. Although there are some aspects of the CNS related to inhibition that resemble subsumption, the structure of the CNS as a whole is far from being purely subsumptive. The large amount of recip-
local and feedback connectivity in the CNS between various systems of various levels of sophistication, e.g. communication from lower layers such as the cerebellum to higher layers such as the neocortex, make placement of the CNS within the framework of the Subsumption Architecture senseless. The thalamus, for example, projects to and is projected to by various structures throughout the CNS. Placement of the thalamus as a single layer within the Subsumption Architecture would be impossible because it would have to be both above and below a given other layer simultaneously.

Another major shortcoming of the Subsumption Architecture is its assumption that the ALF does not contain any internal representation of the world, that it uses the world itself as it is the “best” existing model of the world. The problem with this is that a huge number of the problems an ALF would potentially hope to solve require faster-than-real-time search through a model of the world. For example, if I wish to decide upon a course of action, I use my internal representation of the world as a simulation of the real world in an attempt to find the best behavior. The whole point is to avoid performing the operations on the real world because we want the operation we do ultimately perform on the real world to be the best operation among many, and we only want that operation to occur once—in fact there may only be one chance to act. Hence, despite Brooks’ statement regarding complex behavior stemming from a simple ALF and a complex environment, I believe that the simplicity of the ALF can go only so far, and internal representations of the world are required even for many simple operations.

When comparing the architecture to the CNS in the light of its lack of internal representation, it becomes even more apparent how disparate the two systems are, for a great deal of the architecture of the CNS depends on the existence of internal representations, such as the role of prefrontal areas in making judgments based upon the world. Various systems integrate with the internal representation if not via direct connections, then by functional connection corresponding to an indirect physical connection.

Therefore, although the Subsumption Architecture may be useful for designing simple and practical ALFs, the complexity of the creature is far too limited by the architecture for the architecture to be of much long term use.

3 Ethologically Inspired Synthetic Characters

The Synthetic Character Design used in Blumberg’s “ethologically inspired tool kit for building autonomous animated creatures (i.e. Hamsterdam)” [Blu96] differs considerably from the Subsumption Architecture. The design is actually less restrictive, consisting of three basic layers, the behavior, motor, and geometry layers, largely due to the fact that the system was based in part of ethological studies. The latter two layers can more or less be combined for purposes of this discussion because the motor layer basically just decides how to intermix multiple motor commands to result in a set of geometric movements of the ALF, in a 3D graphics environment in this case.

The behavior layer consists of various behavior objects that all com-
pete for control of the motor system at a given time and this is functionally similar to the decision making process present in an actual organism, as Blumberg notes. Even if a behavior does not become selected, it can give suggestions to the motor layer as to how it might still help the losing behavior and the motor system then attempts to service losing behaviors while performing the winning behavior if it does not interfere with performing the winning behavior.

The ALF contains a primitive form of short term memory, which is used to store the recently performed behaviors and the recent objects of interest in the world. Hence we already have the beginnings of a world model in the form of a set recent objects of interest. The set of recently performed behaviors is basic, but nonetheless provides enough information for primitive planning to occur, in the form of action-selection, a major feature of Blumberg’s ALF design. The combination of this short term memory and internal state variables regarding level of interest allows life-like action-selection to take place. It is life-like in that just as the human CNS contains attention mechanisms that make more novel stimuli of greater importance, Blumberg’s ALF will pay more attention to the idea of a novel action than to a frequently-performed one because the level-of-interest for a frequently-performed activity declines in order to allow other actions to take place. Blumberg notes that this issue of persistence, i.e. how long a creature attends to a particular task before doing something else, is an optimality problem already well-solved by real organisms to ensure the proper balance of persistence and opportunism.

Thus, Blumberg’s ALF architecture is fairly unconstrained as it exposes internal variables to each unit in the behavior system, which is more similar to the way the CNS associates various different types of information in performing action-selection, with no single linear, top-to-bottom hierarchy involved. If anything, the CNS has more of a centralized control structure with the thalamus routing various types of information to different areas and with this information being combined in a decision-making process mainly in prefrontal areas.

Some hierarchy does exist within sensory systems, as primary sensory information is often combined within primary association areas before being used by other areas, but this is also true of Blumberg’s sensory system in that it allows for the hierarchical combination of sensory data. Thus the CNS at a very high level receives sensory input, performs processing on that information, routes the information along with internal state information to various processing centers, and then produces motor output. This is a crude description, but it does match the form of Blumberg’s system in which sensory information is received, processed, made available to all behavior objects which then combine that information with internal state variables to determine the next action which is then sent, possibly modulated by other losing actions to the motor/geometry system.

Blumberg’s system also allows a basic type of learning in which the short term memory is accessed in the context of a particular so-called releasing-mechanism or stimulus in an attempt to form the proper connection between the action performed and the stimulus received. Internal variables are used
to store the state of the ALF’s current “beliefs” regarding such connections, and thus we have the interaction of various systems—the short term memory system, long term memory system (in a very basic form, as internal state variables that can maintain their values indefinitely if they reach a point of equilibrium), behavior system, and sensory system which then determine future motor actions. The connection between learning and memory has already been established in the human CNS, so it’s only appropriate that Blumberg’s system have the two closely coupled.

Thus, the Synthetic Character Design benefits largely from its roots in ethology, making it both practical and robust, while leaving ample room for expansion into a more sophisticated system including more extensive internal state and world representations indicative of inferotemporal cortex, an affect system, a system governing the emotional state and emotional history of the ALF. Such a system is present in Yoon’s [Yoo00] Affective Synthetic Characters, a modification of Blumberg’s Synthetic Character Design. An ALF in this system can have a current mood, but it can also attach affective tags to objects and places, biasing its attitude towards that particular item. As in Blumberg’s work, Yoon’s affect system is inspired by actual organisms, and by design the system closely parallels that of the actual CNS.

As expected, the affect system is tightly coupled with the other systems in the ALF. The affect system is grouped together along with the drive system in a single motivation system which has direct connections to and from the other three systems, the perception system, motor system, and behavior system.

The motivation system sends output to the behavior system in order to request that a certain behavior be performed to satisfy a drive, which is also based upon the affect, or mood of the ALF. The motivation system sends output to the perception in order to allow drives and mood to affect the way the ALF perceives the world, i.e. to focus its attention on something its drives or mood regard as of special importance. The motivation system receives input from the perception system so that sensory input can affect mood or drives via releasing mechanisms. The motivation system receives
input from the behavior system because the result of a given behavior affects the mood and level of drives in the ALF. Finally, the motivational system receives input from the motor system, as feedback to the drive system, and sends output to the motor system to allow the mood and drive levels to affect the way a motor command is executed so that the ALF can convey its emotions and drives outwardly so others understand the ALF’s intentions and so that the ALF can communicate using its expressions much in the same way a human or animal conveys its emotional state for communicative purposes.

Similarly in the CNS we see either direct or near-direct connections between the analogous areas, such as the connection from the amygdala to prefrontal cortex via the mediodorsal nucleus [Bro98], which is similar to a connection from the affect system to the behavior system. (We are assuming the behavior system is to include the role of higher order behavior. Otherwise, the behavior system is more similar to hindbrain in function.) Also, we know of connections from the amygdala to the visual areas of the inferotemporal cortex [Bro98] which is similar to a connection from the affect system to the perception system. The amygdala also receives all types of sensory information [Bro98], establishing the link from the perception system to the motivation system. The fact that the amygdala is crucial for the association between stimuli and their emotional value [Bro98], we see that the affect system is also responsible for the formation of affective tags used in Yoon’s system. It is also hypothesized that fibers running from the amygdala to the ventral striatum mediate the influence of emotions on movement [Bro98], corresponding to the connection from the motivational system to the motor system. We also know that stimulation of the basolateral nuclear group results in heightened attention [Bro98], an example of the affect system modifying the perception system.

This actually begs the question of what we mean by perception system, that is, do we mean those systems responsible for receiving and processing sensory information, or those that actually result in the final processed subjective experience of sensory information. We know, however, that the motivation system in Yoon’s design directly modifies the perception system for purposes of changing attention rather than requiring other systems such as the motor system to integrate information from the motivation and perception systems in order to derive the proper attention of the ALF when producing motor output. Thus the example regarding the basolateral nuclear group is valid.

Thus we see the numerous parallels between the connections between the affect system and other systems in Yoon’s ALFs and the connections between the limbic system and other systems within the human CNS. This in turn makes the structure well-suited to use in construction of artificial human life.

In fact, Yoon’s study includes the implementation of human-like creatures, albeit very simplified ones, that interact with each other and with actual human participants in a project called (void*): A cast of characters. Three ALFs were implemented using the Affective Synthetic Character design and it was shown that human participants almost universally perceived
emotions and personality traits in the ALFs identical to those intended by Yoon when the ALFs were created. And rather than just being hard-wired to pass this single test, the characters have a robust underlying system that parallels that of the human CNS.

5 Conclusion

In our survey of three recent systems for the development of ALFs we can see that there are various approaches to the problem of creating ALFs, and that these approaches vary in part due to the differing motivations of the researchers behind them.

Brooks’ Subsumption Architecture is effective for building simple systems that are basically stimulus response machines without any internal representation of the world whereas both Blumberg and Yoon have systems that allow for but do not currently contain such representations on more than an extremely basic level. In terms of information flow, Brooks’ system allows perception and motor output to occur at every level in the system unless specific subsumption occurs, and information flow among levels takes place strictly in the direction from lower to higher layers. Therefore, the Subsumption Architecture does not particularly resemble the human CNS and thus, as an assumption of this paper, would not be useful for construction of human-like ALFs.

The information flow in Blumberg’s and Yoon’s systems is significantly more like that of the human CNS, particularly in Yoon’s system which incorporates the affect system as well as the motor, behavioral, and perceptual systems of Blumberg’s systems. Yoon’s system takes things a step closer to the human CNS by making an explicit point of building a system that mirrors the CNS in structure and function, both for purposes of building a better system, and for further understanding human life itself, whereas Blumberg’s work was more focused on the production of believable, but not necessarily CNS-accurate ALFs, that would serve the purpose of giving the impression of being alive without necessarily having the underlying similarity. Thus Yoon’s system seems to be the most promising of the three, given that a functionally and architecturally accurate artificial human being is the desired goal.

Possible future extensions of Yoon’s system include addition of a substantial internal representation of the world, and of course implementations of solutions to all of the traditional AI problems like machine vision and object recognition, natural language processing, large-scale planning and goal seeking, and logical deduction and problem-solving. In terms of the architecture, the main missing component is that of a system handling long-term memory and the world model. As far as the connections between systems in Yoon’s existing model are concerned, a high level cognition and planning system such as the highly developed prefrontal cortex in humans could be added to either the behavior/action-selection system or to the motivational system but when considering the human CNS, it should be added to the behavior system assuming that the current behavior system is playing a prefrontal role. If the role of the behavior system is to remain fairly low-level, then it might more appropriately map to hindbrain functions, thus a new, separate planning system in the so-called creature kernel (referring to the
core set of the four systems of motivation, behavior, motor, and perception) might be needed.

References


