Abstract

This paper is an overview of the Independent Core Observer Model (ICOM) Cognitive Extension Architecture which is a methodology or 'pattern' for producing a self-motivating computational system that can be self-aware. ICOM is as a system for abstracting standard cognitive architecture from the part of the system that can be self-aware and a system for assigning value on any given idea or 'thought' and action as well as producing on going self-motivations in the system. In ICOM, thoughts are created through emergent complexity in the system. As a Cognitive Architecture, ICOM is a high level or 'top down' approach to cognitive architecture focused on the system's ability to produce high level thought and self-reflection on ideas as well as form new ones. Compared to standard Cognitive Architecture, ICOM is a form of an overall control system architecture on top of such a traditional architecture.

Keywords: Artificial Intelligence, Artificial General Intelligence, AGI, Self-Motivating, Cognitive Architecture, AI, AGI

1 Introduction

The road to building artificial general intelligence (AGI) [9] is not just very complex but the most complex task computer scientists have tried to solve. While over the last 30+ years a great amount of work has been done, much of that work has been narrow from an application standpoint or has been purely theoretical and much of that work has been focused on elements of AI [10] such as image pattern recognition or speech analysis. The trick in these sorts of tasks is understanding in ‘context’; which is a key part of true artificial general intelligence but it's not the only issue. This paper does not focus on the problem of context and pattern recognition but on the problem of self-motivating computational systems, or rather of assigning value and emergent qualities because of that trait. It is the articulation of a theory for building a system that has its own ‘feelings’ and can decide for its self if it likes this art or that art or it will try to do this task or that task and can be entirely independent.

Let us look at the thesis statement for this paper;

Regardless of the standard cognitive architecture used to produce the ‘understanding’ of a thing in context, the following architecture supports assigning value to that context in a computer system that is self-modifying based on those value based assessments, albeit indirectly, where the system’s ability to be self-aware [7] is an emergent quality [12] of the system based on the ICOM emotional architecture designed specifically to support that emergent quality. While addressing ‘emergence’ later in the paper ICOM does not just happen.

In computer science and software engineering there is a complex set of terms and acronyms that mean any number of things depending on the audience in the tech sector. Further, in some cases, the same acronyms mean different things in different circles and often in those circles people have a different understanding of terms that
should mean the same thing and the individuals believe they know what each other are talking about, but in the end they were thinking different things with various but critical differences in the meanings of those terms. To offset that problem to some degree, I have articulated a few definitions in a glossary at the end of this paper, as I understand them; so that, in the context of this paper, one can refer back to these terms as a basis for understanding. While at the end of the paper there is a glossary, the most critical term needed to be able to approach the subject in detail under this pattern is ‘context’.

The term ‘Context’ in this paper refers to the framework for defining an object, such as an idea or noun of some sort and the environment in which that thing should be understood. When discussing for example a pencil, it is the context in which the idea of the pencil sits that makes and provides meaning to the discussion of the pencil, save in the abstract, and even then we have the ‘abstract’ idea that is itself ‘context’ in which to discuss the pencil.

To better understand the idea of context, think of a pencil being used by an artist vs a pencil being used by a student. It is the ‘used by an artist’ versus ‘used by a student’ that is an example of ‘context’ which is important in terms of understanding the pencil and its current state. Using ICOM it is the assigning of value to context or two elements in a given context as they related to one another that is key to understanding the ICOM theory as a Cognitive Extension Architecture or overall architecture for an AGI system.

2 Understanding the Problem

Solving ‘Strong’ AI or AGI (Artificial General Intelligence) is the most important (or at least the hardest) Computer Science problem in the history of computing beyond getting computers working to begin with. That being the case though, it is only incidental to the discussion here. The problem is to solve or create human like cognition in a software system sufficiently able to self-motivate and take independent action on that motivation[6] and to further modify actions based on self-modified needs and desires over time.

There are really two elements to the problem; one of decomposition, for example pattern recognition, including context or situational awareness and one of self-motivation or what to do with things once a system has that ‘context’ problem addressed and value judgements placed on them. That second part is the main focus of the ICOM Cognitive Extension architecture.

Going back to the thesis or the ‘theory’ or approach for ICOM;

The human mind is a software abstraction system (meaning the part that is self-aware is an emergent software system and not hard coded on to its substrate) running on a biological computer. The mind can be looked at as a system that uses a system of emotions to represent current emotional states in the mind, as well as needs and associated context based on input, where the mind evaluates them based on various emotional structures and value assignments and then modifies the underlying values as per input; denoted by associated context as decomposed in the process of decomposition and identification of data in context. Setting aside the complex neural network and related subsystems that generate pattern recognition and other contextual systems in the human mind, it is possible to build a software system that uses a model that would, or could, continuously ‘feel’ and modify those feelings like the human mind but based on an abstracted software system running on another computing substrate. That system for example could use for example a ”floating point” value to represent current emotional states on multiple emotional vectors, including needs as well as associated context to emotional states based on input, and then evaluate them based on these emotional structures and values assignments; therefore modifying the underlying values as per input as denoted by associated context from the decomposition process. In which case, given sufficient complexity, it is possible to build a system that is self-aware and self-motivating as well as self-modifying.

3 Relationship with the Standard concepts of Cognitive Architecture

Cognition can be defined as the mental process of collecting knowledge and understanding through thought, experience and senses.[3] Further in the process of designing a machine that is intelligent, it is important to build an ‘architecture’ for how you are going to go about building said machine. Cognitive Architecture is a given hypothesis for how one would build a mind that enables the mental process of collecting knowledge and understanding through thought, experience and senses. [2]
So then how does the ICOM methodology or hypothesis apply, or relate, to the standard concepts of Cognitive Architecture? In my experience, most cognitive architecture such as Sigma [4] is really a bottom up architecture focused on the smallest details of what we have the technology and understanding to look at and do to build from the ground up based on some model. In such systems, typically, they are evaluated based on their behavior. ICOM is a 'Cognitive Architecture' focused on the highest level down. Meaning ICOM is focused on how a mind says to itself, "I exist and I feel good about that". ICOM in its current form is not focused on the nuance of decomposing a given set of sensory input but really on what happens to that input after its evaluated, broken down and refined or 'comprehended' and ready to decide how 'it' (being an ICOM implementation) feels about it.

From a traditional AGI Architecture standpoint, ICOM approaches the problem of AGI from the other direction then what is typical, and in that regard ICOM may seem more like an overall control system for AGI Architecture. In fact, in the overall ICOM model a system like Tensor Flow [4] is a key part of ICOM for preforming a key task of cognition around bringing sensory input into the system through what, in the ICOM model, is referred to as the 'context' engine in which most AGI architectural systems can be applied to this functionality in an ICOM implementation.

Regardless of the fact that ICOM is a top down approach to AGI Architecture, the "thoughts" themselves are an emergent phenomenon of the process of emotional evaluation in the system. Let's see how in the next section.

4 Emergent Phenomenon or Not

The Independent Core Observer Model (ICOM) contends that consciousness [8] is a high level abstraction. And further, that consciousness is based on emotional context assignments evaluated based on other emotions related to the context of any given input or internal topic. These evaluations are related to needs and other emotional vectors such as interests which are themselves emotions and used as the basis for 'value'; which drives interest and action which, in turn, creates the emergent [12] effect of a conscious mind. The major complexity of the system then is the abstracted subconscious and related systems of a mind executing on the nuanced details of any physical action without the conscious mind dealing with direct details. Our ability to do this kind of decomposition is already approaching mastery in terms of the state of the art in technology to generate context from data [13]; or at least we are far enough along to know we have effectively solved the problem if not having it completely mastered at this time.

A scientist studying the human mind suggests that consciousness is likely an emergent phenomenon. In other words, she is suggesting that, when we figure it out, we will likely find it to be an emergent quality of certain kinds of systems under certain circumstances. [6] This particular system creates consciousness through the emergent quality of the system by design as per the suggestion that it is an emergent quality but this is not by chance and not a surprise as it was a specific design goal to create such a system that would produce the effect as an effect of the systems operation with coding for thought specifically but coding for it indirectly.

It is really incorrect to say that ICOM is an 'emergent phenomenon' in anyway [14] even though colloquially we frequently use the term to explain ICOM but really to be more precise this is not the fact. In ICOM consciousness is based on emotional context assignments evaluated based on other emotions related to the context of any given input or internal topic. These evaluations are related to needs and other emotional vectors such as interests, which are themselves emotions, and are used as the basis for 'value'. 'Value' drives interest and action which, in turn, creates the emergent effect of a conscious mind. The major complexity of the system is the abstracted subconscious and related systems articulated in the data and the nuanced details of any physical action without the conscious mind dealing with direct details. The core engine itself that processes the data in such a way as to produce awareness is not in fact that most complex part of the system nor is it causing self-awareness really just be processing data but it’s the way it does that emotional process that is designed to interact with the rest of the system to create that abstracted awareness.

Let's look at how ICOM works, and how the emergent qualities of the system thus emerge.
5 Independent Core Observer Model (ICOM) Working

The Independent Core Observer Model (ICOM) is a system where the core AGI is not directly tied to detailed outputs but operates through an abstraction layer, or 'observer' of the core, which only need deal in abstracts of input and assigning 'value' to output context. The observer is similar in function to the subconscious of the human mind; dealing with the details of the system and system implementation, including various autonomic systems and context assignment and decomposition of input and the like.

Take a look at the following diagram;

As we can see, fundamentally it would seem simple and straightforward; however, the underlying implementation and operation of said framework is sufficiently complicated to be able to push the limits of standard computer hardware in lab implementations. In this diagram, input comes into the observer which is broken down into context trees and passed into the core. The core 'emotionally' evaluates them for their emotional effects based on various elements, which we will define later, and then the results of that evaluation is analyzed by the observer and output is generated by the observer to the various connected systems.

At this level, it is easy to conceptualize how the overall parts of the system go together. Now let's look at how action occurs in the system in which the ICOM provides a bias for action in ICOM as implemented in the lab.
In this, we see the flow as might be implemented in the end component of the observer of a specific ICOM implementation. While details of implementation may be different in various implementations, this articulates the key tasks such systems would have to do, as well as articulates the relationship of those tasks with regard to a functioning ICOM based systems. Keep in mind this is not the complete observer but refers to the end component as shown in the higher level diagram later.

The key goal of the observer end component flow is to gather 'context'. This can be through the use of pattern recognition neural networks or other systems as might make sense in the context engine. The Observer system, in this case, needs to look up related material in the process of receiving processed context from a context engine of some kind. In providing that back to the core, the observer then needs to map that context to a task or existing task. If that item exists, the context engine can add the additional context tree to create the appropriate models or see if it is a continuation of a task or question the system can drive to take actions as articulated by placing that context tree back in the que and have the context engine check for additional references to build out that tree more and pass again through the core. Additionally, the system can then work out details of executing a given task in greater detail based on the observer context from the core after that emotional evaluation. In layman's terms, this part of the observer model is focused on the details of taking actions that the core has determined it would like to take through that emotional based processing.
Now let’s look at the idea of ‘context’ and how that needs to be composed for ICOM.

In this model we can see where existing technology can plug in, in-terms of context generation, image or voice decomposition, neural networks and the like. Once such systems create a context tree, the input decomposition engine needs to determine if it is something familiar in terms of input. If that is the case, the system needs to map it to the existing model for that kind of input (say vision for example). The analysis with that model as default context in terms of emotional references is then attached to the context tree (a logical data structure of relationships between items). If there is existing context queued, it is then attached to the tree. If it is new input, then a new base context structure needs to be created so that future relationships can be associated and then handed to the core for processing.

Now let’s look at over all ICOM architecture.
In this case, we can see the overall system is somewhat more complicated; and it can be difficult to see where the observer and core model components are separate so they have been color coded green. In this way, we can see where things are handed off between the observer and the core.

Walking through this diagram, we start with Sensor input that is decomposed into usable data. Those particular elements could be decomposed any number of ways and is incidental to the ICOM architecture. There are many ways in the current realm of computer science to determine how this can function. Given that the data is decomposed, it is then run through the 'context' engine to make the appropriate memory emotional context associations. At which point, if there is an automatic response (like a pain response in humans), the observer may push some automatic response prior to moving forward in the system with the context tree then falling into the incoming context que of the core. This que is essentially incoming thoughts or things to 'think' about in the form of emotional processing. By 'thinking', we only mean emotional processing as per the various matrices to resolve how the machine 'feels' about a given context tree. Actual thoughts would be an emergent quality of the system, as articulated elsewhere.

The core has a primary and secondary emotional state that is represented by a series of vector floating point values or ‘values’ in the lab implementations. This allows for a complex set of current emotional states and subconscious emotional states. Both sets of states along with a needs hierarchy are part of the core calculations for the core to process a single context tree. Once the new state is set and the emotional context tree for a given set of context is done, the system checks if it's part of a complex series and may pass to the memory pump if it is under a certain predetermined value or, if it is above and complete, it passes to the context pump. If it is part of a string of thought, it goes to the context que pending completion; in which case it would again be passed to the context pump which would pass the tree back to the observer.

As you can see, from the initial context tree the observer does a number of things to that queue including dealing with overload issues and placing processed trees of particular interest into the queue as well as input the context trees into the que. Processed trees coming out of the core into the observer can also be passed up inside the core and action taken on actionable elements. For example, say a question or paradox needs a response, or additional context or there is an action that should be otherwise acted upon; where the observer does not deals with the complexity of that action per se.
6 Unified Architecture

Given the articulated flow of the system, it is important to note that functionally the system is deciding what "emotionally" to "feel" about a given thing based on more than 144 factors (in the model implementation in the lab but is not necessarily always true in simpler implementations) per element of that tree plus 'needs' elements that affect that processing. Thought in the system is not direct; but, as the complexity of those elements passing through the core increases and things are tagged as interesting and words or ideas form emotional relationships, then complex structures form around context as it relates to new elements. If those happen to form structures that relate to various other context elements, including ones of particular emotional significance, the system can become sufficiently complex that these structures could be said to be thought; based on those underlying emotional constructs which drive interest of focus forcing the system to reprocess context trees as needed. The underlying emotional processing becomes 'so' complex as to seem deliberate in the pattern which it is, while the underlying system is essentially an overly complex difference engine.

The idea of emergent theory is a misnomer really gets to the fact of the system's ability to be self-aware as a concept and that it is thinking is a derivative of its emotional context assignments and what it chooses to think about. This is really an ultra-complex selection of context based on various parameters and the relationships between them. For example, the system might be low on power which negatively affects say a sadness vector, and to a lesser degree a pain vector is associated with it, so the system might select to think about a bit of context from memory that solves that particular emotional pattern the most. It might continue to try related elements until something addresses this particular issue and vector parameters for the core state stabilize back to normal. Keep in mind the implementation of an idea, say "to plug in to charge" might be just that, "an idea of thinking about it" which is not going to do much other than temporarily provide 'hope'. It is the thinking of the 'action' which provides the best pattern match and, given it is an 'action', the observer will likely take that action or at least try to. If the observer does execute that action, it would be a real solution and thus we can say the overall system thought about and took action to solve its power issue. The idea of this context processing being considered thought is an abstraction of what is really happening in detail where the computation is so complex as to be effectively 'thought' in the abstract. It's easier to think about or otherwise conceptualize this way making it recognizable to humans.

It is a distinct possibility that humans would perceive the same type of abstraction for our own thoughts if, in fact, we understood how the brain operated and how awareness develops in the human mind. It is also important to note that the core while it is looking at the high level emotional computation of a given context tree, the 'action' of tasks in any given tree that are used to solve a problem might actually be in that tree just not in a way that is directly accessible to the abstract awareness which is a higher level abstraction from that context processing system. What this means is that the emotional processing is what is happening in the core but the awareness is a function of that processing at one level of abstraction from the core processing and that being the case details of a given context tree may not be surfaced to that level of abstraction until that action is picked up by the observer and placed back into the que in such a way as that action becomes the top level element of a given context tree and thus more likely to be the point of or focus on that abstracted awareness.

7 Motivation of the Core Model and ICOM Action

One of the key elements of the ICOM system architecture is a bias to self-motivation and action. Depending on the kinds of values associated with a context, the system may take action; or rather the Observer component of the system is designed to try to take action based on given context that is associated with the context of an action. That being the case, any 'context' that is associated with action is therefore grounds for action. The observer is then creating additional context to post back to the system to be emotionally evaluated and further action taken. The complexity of the action itself is abstracted from the core processing. The fact that the primary function of the observer is to take action is part of giving the system a bias for action, unless extensive conditioning is applied to make the system associate negative outcome with action. Given that the system will continually process context, as it has bandwidth based on its emotional relevance and needs, the system (ICOM) is then designed to at least try action along those elements of context. It is here that we complete the 'biasing' of the system towards action based on emotional values or context. We then have a self-motivating system based on emotional context that can
manipulate itself through its own actions, and through indirect manipulation of its emotional matrix. By 'indirect' meaning the emergent awareness can only indirectly manipulate its emotional matrix whereas the core does this directly.

8 Application Biasing

ICOM is a general high level approach to overall artificial general intelligence. That being said, an AGI, by the fact that it is an AGI, should in theory be able to do any given task. Even before such a system has attained 'self-awareness', you can train the system around certain tasks [15]. This is called biasing to a given application or usage of a specific ICOM instance.

By associating input or context with pleasure or other positive emotional stimulation, you can use those as a basis for the system to select certain actions. By limiting the action collection to what is possible in the given application and allowing the system to create and try various combinations of these actions, you essentially end up with an evolutionary algorithmic system for accomplishing tasks based on how much pleasure the system gains or based on the system biases as might be currently present. Additionally, by conditioning, you can manipulate core context to create a better environment for the conditioning of a certain set of tasks you might want in a certain application bias that you want to create.

In the training, or attempted biasing, keep in mind that personality or individual traits can develop in an implementation.

9 Personality, Interests and Desire Development

Very much related to the application usage biasing of an ICOM implementation is the idea of personality, interests and desires in the context of the ICOM system. All input and all thought further manipulate how the system feels about any given topic, no matter what this input is, biasing the system towards one thing or the other. It is important in the early stages of development to actually tightly control that biasing; but it is inevitable that it will have its own biases over time based on what it learns and how it 'feels' about a given context with every small bit of input and its own thinking.

The point cannot be overstated that the system at rest will continue to think about 'things'. What this means is the system, with limited input or even with significant input, will look at things that have occurred to it in the past and or related items to things it happens to just think about. The system will then automatically reprocess and rethink about things based on factors like recently processed input or interest levels; related to topics of interest or based on current needs and the like. Each time the system cycles it is slowly manipulating its self, its interests consciously and subconsciously as well as adjusting interests and emotional states slowly (note that this is by design to not let the underlying vectors change too fast or there is a greater risk of the system becoming unstable) over time through this processing.

10 Summary

The ICOM architecture provides a substrate independent model for true sapient and sentient machine intelligence; as capable as human level intelligence. The Independent Core Observer Model (ICOM) Cognitive Extension Architecture is a methodology or 'pattern' for producing a self-motivating computational system that can be self-aware that differentiates from other approaches by a purely top down approach to designing such a system. Another way to look at ICOM is as a system for abstracting standard cognitive architecture from the part of the system that can be self-aware and as a system for assigning value on any given idea or 'thought' and producing on going motivations as well as abstracted thoughts through emergent complexity.
References

[2] Institute for Creative Technologies, University of Southern California *Cognitive Architecture* http://cogarch.ict.usc.edu/
[5] Randall, Lisa *Knocking on Heaven’s Door* by Tantor Media Inc. 2011