Causal Cognitive Architecture 1: Integration of Connectionist Elements into a Navigation-Based Framework

Howard Schneider
Causal Cognitive Architecture 1: Integration of Connectionist Elements into a Navigation-Based Framework

Howard Schneider
Sheppard Clinic North, Toronto, ON, Canada
howard.schneider@gmail.com

Abstract. The brain-inspired Causal Cognitive Architecture 1 (CCA1) tightly integrates the sensory processing capabilities found in neural networks with many of the causal abilities found in human cognition. Sensory input vectors are processed by robust association circuitry and then propagated to a navigational temporary map. Instinctive and learned objects and procedures are applied to the same temporary map, with a resultant navigation signal obtained. Navigation can similarly be for the physical world as well as for a landscape of higher cognitive concepts. Causality emerges from the architecture, with good explainability for causal decisions. An emotional system reduces the class imbalance problem. A simulation of the CCA1 controlling a search and rescue robot is presented with the goal of finding and rescuing a lost hiker within a grid world.

Keywords: Cognitive Architecture; Causality; Spatial Navigation; Artificial General Intelligence

1. Introduction

Artificial neural networks (ANNs) can recognize patterns and perform reinforcement learning at a human-like proficiency [1,2]. However, compared to a four-year-old child, in terms of logically and causally making sense of their environment or a problem at hand, especially if training examples are limited, they perform poorly [3,4].

A number of cognitive architectures integrate subsymbolic and symbolic processing to varying degrees [5,6,7]. A review of the field by Langley [8] notes that while early models were mainly symbolic, many of the modern architectures are more hybrid. Lake and colleagues [9] propose that thinking machines should build causal models of the world, and discuss intuitive physics and psychology present in infants. Taatgen [10] discusses the need to incorporate learning at multiple levels of abstraction, while Laird and Mohan [11] discuss combining more innate architectural learning mechanisms.


Despite many of the above designs and implementations combining ANNs and symbolic elements, they do not approach the causal abilities seen in human children.

The Meaningful-Based Cognitive Architecture (MBCA) [18,19,20] combines connectionist and symbolic elements in a biologically-plausible manner in which sensory input vectors are processed causally. A collection of intuitive and learned logic, physics, psychology and goal planning procedural vectors (essentially acting as small algorithms) are applied against inputs, and intermediate causal results can be fed back to the sensory input stages and processed over and over again. However, it must clumsily move data around to make a decision about choosing its next action.

In this paper, the Causal Cognitive Architecture 1 (CCA1) is introduced. The CCA1 improves upon the MBCA with a more straightforward navigation-based logic system that works well with streams of vectors rather than needing more classical-like symbolic representations. Causal properties easily emerge from its architecture. As well, the CCA1 is more realistic with regard to modeling the evolution of the mammalian and human brain.
2. Architecture and Operation of the CCA1

2.1 Sensory and Autonomic Inputs and Processing

An overview of the architecture of the CCA1 is shown in Figure 1. Sensory Inputs 1..n from different sensory systems 1..n, propagate to the Input Sensory Vectors Association Modules 1..n, with a module dedicated for each sensory system. Each such module contains a conventional neural network [1] or a hierarchy of Hopfield-like Network units (HLNs) [18], or other roughly similar technology that can robustly associate an input sensory vector with other vectors within the CCA1. Further binding of the processed input sensory vectors, via straightforward temporal mechanisms, or more complex global feedback mechanisms, occurs within the Sensory Vectors Binding Module.

Autonomic inputs, as in the physiological sense coming from the CCA1 and its embodiment, e.g., low energy reserves or an abnormal temperature rise in a component, and so on, are processed in the Autonomic Reflex Modules as well as the Autonomic Module (within the centralized components of the CCA1) which bidirectionally communicates with the former. The autonomic modules can offload a variety of routine, lower-level internal signal processing from the more complex CCA1 centralized components.

2.2 Pre-Causal Cognitive Processing and Output

Cognition in the CCA1 is at its core movement-based. At the simplest level, the CCA1’s embodiment navigates through physical space, although at higher cognitive levels navigation occurs through a space of concepts and analogies. This is touched upon briefly further below.

The Navigation Module holds a temporary map of a small part of the physical world, and on this small map are objects from the Sensory Vectors Binding Module, an object representing the CCA1 embodiment itself, and possible objects from the Instinctive Primitives Module and the Learned Primitives Module.

Fig 1. Causal Cognitive Architecture 1 (CCA1)
(Not all connections shown. D – Internal Developmental Timer)
The Instinctive Primitives Module and the Learned Primitives Module are triggered by processed vectors from the Input Sensory Vectors Association Module, the Sensory Vectors Binding Module, as well as by the Goal/Emotion Module and the Autonomic Module. The Instinctive Primitives Module and the Learned Primitives Module can link possible objects onto the temporary map in the Navigation Module, as well as manipulate these objects in the Navigation Module.

The Navigation Module propagates an output that effectively tells the Output Vector Association Module to produce the desired movement of the embodiment of the CCA1.

The Sequential/Error Correcting Module recognizes and learns temporal sequences in input sensory vectors as occur in the physical world, and provides a signal to the Navigation Module.

The Output Vector Association Module does not just look up in a table what actuators to actuate for a desired movement by the Navigation Module but in a robust way learns a variety of different patterns of Actuator Outputs. As well, proprioception (the position of the CCA1’s embodiment’s actuators), along with other processed sensory inputs, are propagated to the Sequential/Error Correcting Module which produces an error signal between the desired and actual outputs, further improving output accuracy.

2.3 Internal Developmental Timer and Learning

Note that in Figure 1 some modules have in the corner the letter D, which stands for an Internal Developmental Timer. The properties of these modules are strongly influenced by the developmental stage of the CCA1.

A newly created CCA1 is not a tabula rasa. For example, when the CCA1 starts operating for the first time there are certain patterns the Input Sensory Vectors Association Modules will recognize. As well the Instinctive Primitives Module, as noted above, contains a variety of objects it can map onto the temporary map of the Navigation Module and a variety of operations it can influence the Navigation Module to perform on these objects. It is advantageous for the CCA1 to learn and produce relevant actions at different stages of development. Thus, different Instinctive Primitives will be triggered at different stages, and different Learned Primitives will be learned or triggered at different stages.

Learning occurs throughout the CCA1. Feedback pathways are omnipresent in the architecture. At present, if ANNs are being used for association modules, then a separate learning period is required. If HLNs are being used instead, then continuous learning becomes more feasible [18].

2.4 Causal Cognitive Processing and Output

In pre-causal operation of the CCA1, as described above, associations occur in the Input Sensory Vectors Association Modules and other modules described above. The Navigation Module effectively allows a sort of pre-causal processing with objects and rules being applied by the Instinctive and Learned Primitives Modules, and the Navigation Module making a navigation decision.

In causal operation, there is a more substantial feedback pathway from the Navigation Module to the Input Sensory Vectors Association Modules and other modules. As such, intermediate results of a problem or a causal situation can be fed back to the sensory input stages, and processed again in the next processing cycle.

In a processing cycle sensory inputs are processed through the CCA1, and the Navigation Module makes a decision about an output vector, and then the next processing cycle starts again. However, in causal operation if an intermediate result from the Navigation Module is fed back to the sensory input stages, then in the next processing cycle, the CCA1 will take as the input the intermediate result which resides now in (or in the pathway of) the Input Sensory Vectors Association Modules. As such, intermediate results can be fed back to the sensory input stages and then processed by the CCA1 and Navigation Module, over and over again. As shown in the simulation example below, in this manner, causal processing of the inputs often results.

The Causal Memory Module shown in Figure 1 stores memories of operations of the Navigation Module. If similar events occur again, the Causal Memory Module can feed back into the Navigation Module which actions were taken in the past. The Causal Memory Module also gives the CCA1 a sense of time – what state and operation of the Navigation Module occurred in the past before or after another operation of the Navigation Module. Note that this also allows the CCA1 reasonable interpretability and explainability of causal decisions it makes, even with the handicap of associative processing of sensory data. The Causal Memory Module is functionally part of the Navigation Module, and access to the Causal Memory Module occurs via the Navigation Module.
3. CCA1 Simulation and Examples

3.1 CCA1 Simulation Overview

In the simulation of the CCA1, the embodiment of the CCA1 plus the CCA1 itself (together which are simply called here “the CCA1”) must enter a grid world forest, and find and rescue a lost hiker. This is intentionally similar to the environment used by the MBCA simulation [20] although the environment used by the CCA1 includes a waterfall, allowing a more realistic demonstration of pre-causal versus causal behavior.

The simulation is written in the Python language, and via a menu driven user interface allows various modules to be used, to be partially used or not to be used, enabling a spectrum of pre-causal to fully causal models to be simulated. The small size of the grid world allows a simulation of an ANN or a simulation of an equivalent network of HLN to be selected by menu for use in the various association modules.

Due to the proof of concept stage of the work, a measure of a CCA1 model’s future expected value is simply taken as the reciprocal of the number of moves required to find a lost hiker over a number of simulation runs (which involve some randomness in perception of the environment).

Figure 2 shows the starting position of the CCA1 in the grid world. The lost hiker is in another square, and must be found (i.e., the CCA1 moves to this square) to be considered rescued by the CCA1. Please note that this is a bird’s-eye view for the reader—the CCA1 does not see this information but must construct its own internal map of the world around it.

The CCA1 is able to safely navigate through the “forest” portions of the grid world. Edges of the grid world can be detected and simply do not allow movement. If the CCA1 enters a “cliff” square, then it will be considered to have fallen off the cliff and to be damaged, and thus the mission (i.e., the simulation) will end. The CCA1 is able to cross shallow rivers. However, if it walks into what may seem like a river but is part of a waterfall, it will be considered to have fallen off the cliff portion of the waterfall and to be damaged, and thus the mission (i.e., the simulation) ends.

Each processing cycle the CCA1 propagates the sensory inputs through its architecture, and over four cycles, attempts to recognize the squares south, west, north, and east adjacent to its current position, storing this information in the temporary map within the Navigation Module.

<table>
<thead>
<tr>
<th>CCA1</th>
<th>forest</th>
<th>shallow river</th>
<th>forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>cliff</td>
<td>forest</td>
<td>forest</td>
<td>forest</td>
</tr>
<tr>
<td>forest</td>
<td>waterfall</td>
<td>forest</td>
<td>forest</td>
</tr>
<tr>
<td>forest</td>
<td>lost hiker</td>
<td>forest</td>
<td>forest</td>
</tr>
</tbody>
</table>

Fig. 2. Birds-Eye View of the Starting Positions of the CCA1 and the Lost Hiker
(Note: The CCA1 does not have this information but must build up its own internal temporary map within the Navigation Module)

3.2 Pre-Causal CCA1 Simulation Examples

A new simulation starts with the CCA1 and the lost hiker in the positions as shown in Figure 2. The CCA1 recognizes to the south features which the Instinctive Primitives Module considers a cliff (name is arbitrary) object with a signal to the Navigation Module to not navigate there. Thus there is no output from the Navigation Module in this processing cycle. In the next two processing cycles the CCA1 recognizes to the west and north features which the Instinctive Primitives Module considers edge objects with a signal to the Navigation Module to not navigate there. All these recognized objects are stored in the temporary map within the Navigation Module, i.e., the map is built up as the CCA1 navigates through the world.

In the following processing cycle the CCA1 recognizes to the east features which the Instinctive Primitives Module considers (i.e., links properties and functions onto the temporary map of the Navigation Module) a forest object (name is arbitrary). The other parts of the temporary map do not allow movement, and thus the Navigation Module (in conjunction with a procedural vector from the Goal/Emotion Module) makes a decision to move east.
In this example, the CCA1 continues to navigate through the landscape, and after a number of moves, makes it to the square “forest” east of the lost hiker. In the next few processing cycles its internal map will become further updated, and the Navigation Module will make the decision to move to the west to the square “lost hiker.” The CCA1 is considered to have rescued the lost hiker and the simulation ends.

Learning occurs throughout the CCA1. Temporary maps the Navigation Module builds up, such as the one for this grid world, are stored, and can be associated and recalled in the future. In a complex environment, even if a non-exact map is associated and recalled, it can be useful in guiding behavior to repeat successful moves and avoid unsuccessful ones.

Consider a brand new CCA1, starting off again as shown in Figure 2. In this example, the CCA1 eventually navigates to the square “forest” just north of the square “waterfall.” (Note: “waterfall” is labeled in the map in Figure 2 for the convenience of the reader. The CCA1 does not know this square is a waterfall—it must try to build up its own internal map.) The CCA1 has already spent a number of moves in the north and east areas of the grid world without success in finding the lost hiker. As a result, a vector from the Goal/Emotion Module now favors navigation moves west and south.

In the next few processing cycles the CCA1 recognizes a cliff to the west (which the Instinctive Primitive Module signals to avoid movement to) and to the south a shallow river with fast flowing noisy water (the cliff part of the waterfall is not visible). A shallow river does not trigger any prohibitions in the Instinctive Primitives Module since the CCA1 is able to cross shallow rivers. As such, the CCA1 moves to the square labeled “waterfall” in Figure 2, where it is swept over the waterfall’s cliff, and is damaged. The CCA1’s Goal/Emotion Module makes a strong memory of the damage which occurred with moving towards such a fast flowing, noisy river.

If the CCA1 is then repaired, and the next day goes on another rescue mission, and if it recognizes a fast flowing river with much noise, this will trigger in the Goal/Emotion Module and the Learned Primitives Module a signal to the Navigation Module not to navigate to this square.

### 3.3 Causal CCA1 Simulation Examples

Consider a new simulation run with a brand new CCA1, starting off again as shown in Figure 2. This CCA1 takes full advantage of the architecture.

This CCA1 so happens to navigate again to the square “forest” just north of the square which in Figure 2 (intended only for the reader) is labelled a “waterfall.” The CCA1 has already meandered in the north and east areas of the grid world forest without success, and navigation south and west is as result now favored.

In the next few processing cycles the CCA1 recognizes a cliff to the west (which the Instinctive Primitive Module signals to avoid movement to) and to the south a shallow river with fast flowing noisy water (the cliff part of the waterfall is not visible). This CCA1 has never seen a waterfall before. However, \{“water”\} + \{“fast flow” + “noise”\} triggers in the Instinctive Primitives Module \{“water”\} + \{“push”\}.

The Navigation Module cannot process a vector representing \{“water” + “push”\} and so feeds it back to the Input Sensory Vectors Association Module. In the next processing cycle, the Input Sensory Vectors Association Module ignores the external sensory inputs, but instead propagates the intermediate result \{“water” + “push”\}.

\{“water” + “push”\} triggers in the Instinctive Primitives a vector causing the Navigation Module to bring up a new temporary map. On this new map is an object representing the CCA1 and an object representing water on much of the map, and the Instinctive Primitives module procedural vector causes the object representing the CCA1 moved into the water in the temporary map, with water on top of it. The Navigation Module cannot process this any further and so it feeds \{“CCA1 under water”\} back to the Input Sensory Vectors Association Module. In the next processing cycle, the Input Sensory Vectors Association Module ignores the external sensory inputs, but instead propagates the intermediate result \{“CCA1 under water”\}. This triggers in the Instinctive Primitives module a signal “do not go” which is sent to the Navigation Module.

The temporary map in the Navigation Module is that of the CCA1 object and water, and “do not go” is fed back to the Input Sensory Vectors Association Module. In the next processing cycle, the Input Sensory Vectors Association Module ignores the external sensory inputs, but instead propagates the intermediate result “do not go” which triggers in the Instinctive Primitives a trigger to the Goal/Emotion Module which triggers in the Navigation Module to retrieve the temporary map of the grid world forest, and “do not go” applies to the square south on that temporary map (and which in Figure 2 is labelled as a “waterfall”).
In the next processing cycle the CCA1 recognizes the square to the east as forest and the Navigation Module instructs the CCA1 to navigate to the east, as there are no possibilities to the west and south. Even though the CCA1 had never seen a waterfall before nor knew about its properties, it causally avoided this danger. This CCA1 continues to navigate, and a few moves later it finds and rescues the lost hiker.

Although a number of processing cycles took place in deciding to avoid the waterfall, note that all these cycles were essentially triggered from one to another, with no special controlling program, other than the architecture of the CCA1.

If the CCA1 was operating in a different environment where it was, for example, constructing a machine (and had experience related to this environment, of course), then the same mechanism allows it to easily predict that if part A pushes on and moves part B, then if part C pushes on and moves part A (with this intermediate result fed back and then the movement of part A propagated in the next processing cycle), then part B will move, for example. Causality readily emerges from the architecture of the CCA1 and its instinctual and learned contents.

4. Discussion

4.1 Hopfield-like Networks (HLNs)

For pragmatic reasons, the CCA1 can be built with ANNs for associative functions and modified logic gates for symbolic operations. However, the CCA1 can alternatively use hierarchies of HLN as well as the logic-like structures of HLN, similar to use in the MBCA [18,21]. HLN in the MBCA model give three putative advantages which also occur in the CCA1 so constructed:

Better biological and evolutionary plausibility. The HLN are inspired by mammalian cortical minicolumns [22, 23]. Evolutionary precursors to the mammalian cortex appear to go back to the earliest vertebrates [24]. Small changes in the arrangement of the HLN would seem to allow a transition in the mammalian brain from pre-causal behavior to full causal behavior [20].

A generative aspect to processing information. In the MBCA model [18] the HLN can be dynamically reconfigured a number of times each processing cycle to extract what is described as maximal “meaningfulness” from the input sensory vectors as well as in other aspects of the architecture, where the measure of meaningfulness is the reciprocal of the Shannon entropy, favoring activation of the maximal number of HLN further downstream. Empirically this often gives better recognition of input vectors [18], although a proof of its usefulness is lacking.

An ability to quickly increase/decrease weights between different HLN. The Goal/Emotion Module can reduce the class imbalance problem by large changes of weights between HLN associated with infrequent but significant events, e.g., damage to the CCA1’s embodiment.

4.2 Advantages of the CCA1 compared to the MBCA

More Efficient Movement of Data. The Meaningful-Based Cognitive Architecture (MBCA) [18] contains Logic/Working Memory units which must request data from different parts of the architecture at different times. The CCA1 instead contains a Navigation Module which receives feeds of vectors and may trigger by its outputs other vectors in other parts of the architecture, but there are no discrete requests for data, and its operation is less complex than the operation of the MBCA Logic/Working Memory unit.

Enough Complexity for the CCA1 to Support the Psychosis Hypothesis. Schneider [20] hypothesizes that the complexity required to go from pre-causal to full causal operation of the MBCA results in imperfect functioning where the intermediate results sometimes are interpreted as actual external sensory inputs (i.e., hallucinations), which models the emergence of psychosis in Homo sapiens. This is supported by greater than a 10% prevalence of psychosis-like symptoms in humans [25] while psychosis is not naturally found in other animals, although versions of almost all other psychiatric disorders are [26]. However, while not possessing as much added complexity in going from pre-causal to causal operation, the CCA1 still has added enough features, and it can be argued the right amount from an evolutionary point of view (too many malfunctions and the species will not survive into the future despite the advantages of causal capabilities) particularly feeding back signals as intermediate results, to support the psychosis hypothesis.
More Biologically and Evolutionarily Plausible. Both the MBCA and CCA1 are biologically inspired cognitive architectures, although neither attempts to duplicate brain function down to the level of spiking neurons. A common feature of almost all animals with brains is the ability to perform navigation, and in mammals the hippocampal-entorhinal system encodes physical as well as more abstract relations [17]. Thus, it is biologically and evolutionarily more plausible for the CCA1 to have a dedicated Navigation Module, which is absent in the MBCA.

Simpler Transition to Higher Level Cognitive Processes. Consider the ability to make analogies. While in theory the Turing-complete nature of the Logic/Working Memory unit in the MBCA will allow just about any desired cognitive process, doing so requires complex routines and just the right data. On the other hand, the CCA1’s architecture and temporary maps, readily form and use analogies.

Consider an example where the same CCA1 in the above examples must consider if it should spend more of its free time with person A or person B who differs in being smiley but very noisy. Object A (i.e., person A) and object B (i.e., person B) are put onto its temporary map in the Navigation Module. The CCA1 decides whether it should navigate to object B—its Instinctive Primitives like smiling people. However, object B is noisy, and it results in pulling up the previous temporary map it had—the river seemed safe but made much noise also, and was considered a danger. Intermediate results are fed back to the sensory input stages and processed again, temporary maps are switched back again, and the noisy object B now may by analogy also be a danger. Thus in this example, there is a navigation output to navigate to object A (i.e., person A).

4.3 Conclusion and Future Work

The purpose of this paper is to introduce the Causal Cognitive Architecture 1 (CCA1). The Python-language simulation of the CCA1 demonstrates a proof of concept of the architecture.

An overview of the CCA1 was given. It is brain-inspired and tightly integrates the sensory processing capabilities found in neural networks with many of the causal abilities found in human cognition. Sensory input vectors are processed by robust association circuitry and then propagated to a temporary map. Instinctive and learned objects and procedures are also applied to this same map, resulting in a navigation signal output. Intermediate results can be fed back to the input stages and processed again in the next cycle, the whole process resulting in causal processing of the original sensory inputs. Navigation as such can be used not only for the physical world, but also for a landscape of higher cognitive concepts. Causal memory allows good explainability for causal decisions. The CCA1’s emotional system reduces many of the class imbalance problems seen in neural networks.

Future work to further develop the CCA1 includes:

1. Characterization and more formalization of the range of models with the basic properties of the CCA1;
2. Enlarging and enhancing the CCA1 simulation, including larger content in the Instinctive Primitives Module and fuller simulation of the HLN networks used in the association modules;
3. Quantitatively comparing the performance of different CCA1 models with each other, as well with non-CCA1 models. Given the compact grid world it is possible to create a corpus of relevant environments, somewhat similar to Chollet’s dataset [27], and Legg and Hutter’s [28] universal intelligence (1) can actually be numerically approximated for different CCA1 and non-CCA1 models to allow their comparison;

\[ \gamma(\pi) := \sum_{\mu} \text{for all environments} 2^{-K(\mu)} V^\pi_\mu \]

where \( \gamma \) is the expected performance of agent \( \pi \) over all \( \mu \) environments, \( K \) is the Kolmogorov complexity of \( \mu \), \( V \) is the future expected value;

4. Scaling up the architecture to perform a meaningful task.

Acknowledgments: This article builds upon work originally presented at BICA 2018 and 2019 (references 18, 19, 20).

References