Causal Cognitive Architecture 2 (CCA2): A Solution to the Binding Problem

DRAFT 1 – WILL REDUCE IN SIZE TO TIME ALLOCATION

Howard Schneider
Sheppard Clinic North, Richmond Hill, Canada

BICA Workshop at IVA-2021

2021 Annual International Conference on Biologically Inspired Cognitive Architectures 2021
September 14, 2021 University of Fukuchiyama, Kyoto, Japan (online)

Before moving to a solution to the binding problem, let's think about thinking….

$$\Upsilon(\pi) := \Sigma_{\mu \epsilon E} 2^{-K(\mu)} V^{\pi}_{\mu}$$

"universal intelligence"

(expected performance Υ of agent π)

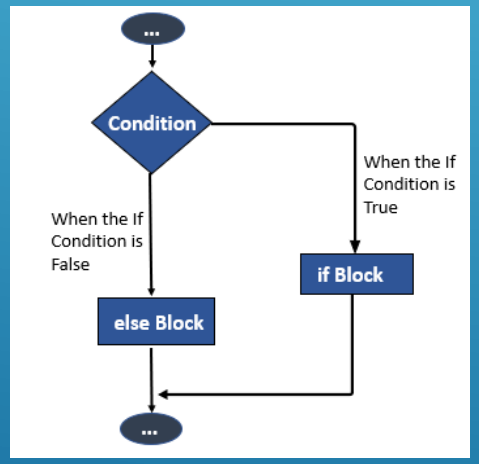
LEGG & HUTTER (2007) "UNIVERSAL INTELLIGENCE":

not computable

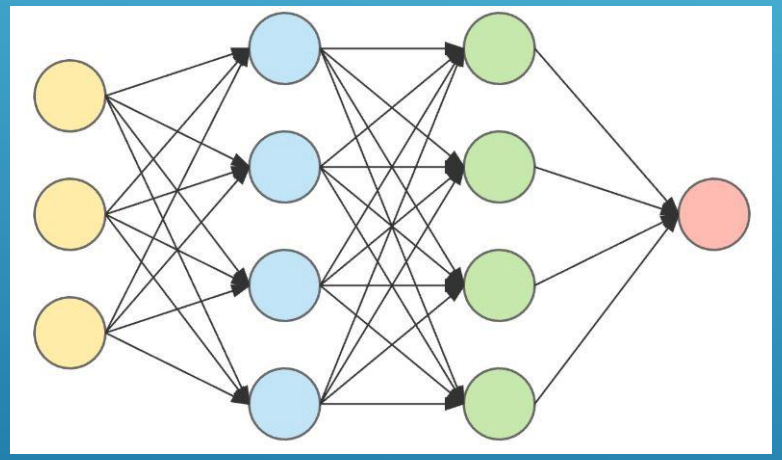
$$\Upsilon(\pi) := \Sigma_{\mu \epsilon E} 2^{-K(\mu)} V^{\pi}_{\mu}$$

algorithmic probability distribution of the space of environments $2^{-K(\mu)}$ times the value function V of agent π operating in environment μ

More useful to think of problem in terms of mechanisms we can use to make decisions….



Symbolic Logic GOFAI



Neural Networks

Navigation Maps

Navigation Maps with Causality

Navigation Maps

- Different way of making decisions
- Most animals – invertebrates and vertebrates use some sort of navigation system

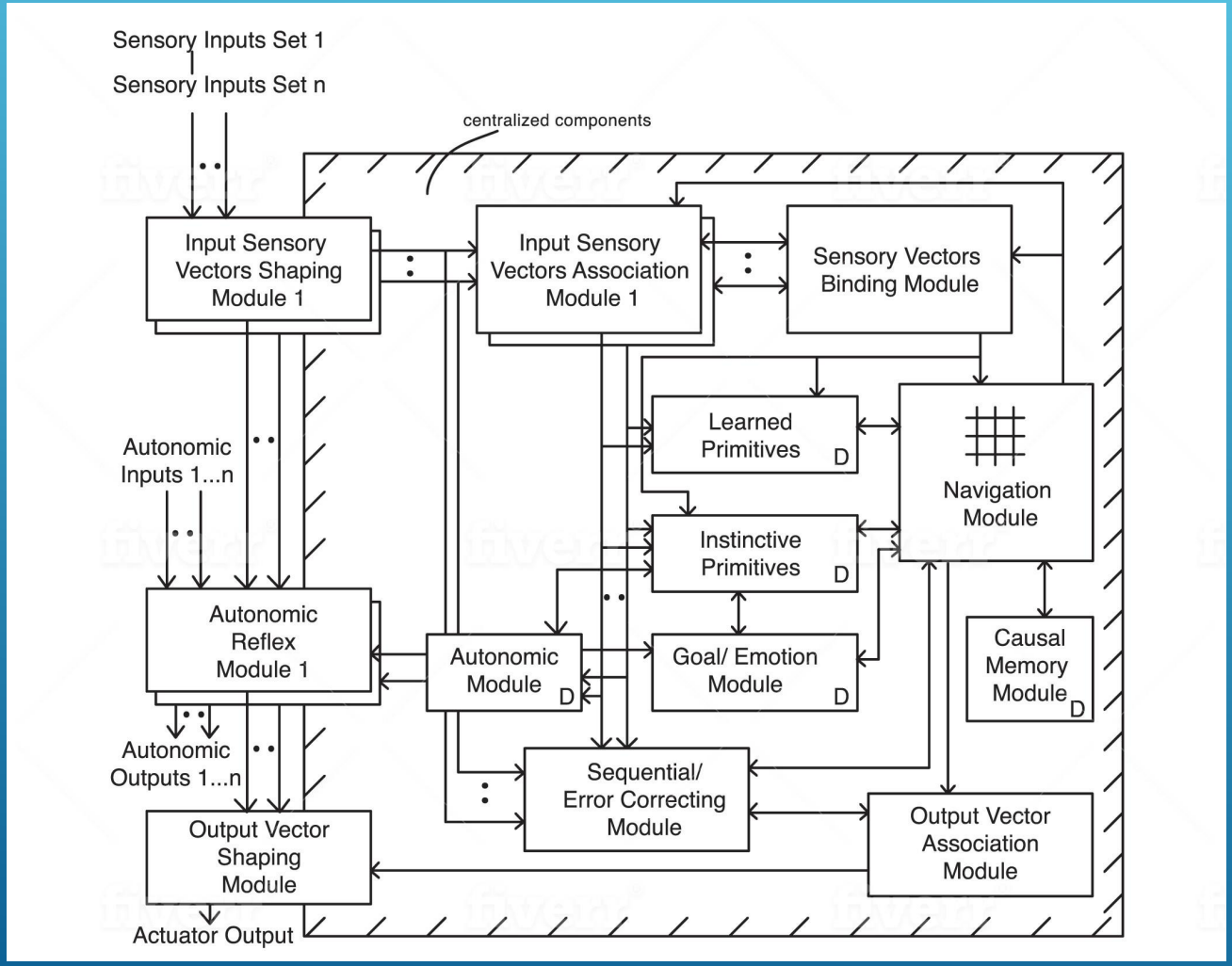
Navigation Maps

- Vertebrates – all have formal navigation systems similar to mammalian hippocampus
- Mammalian hippocampus place and grid cells

Navigation Maps

- We can use in an artificial cognitive architecture not just for navigation but all decisions
- →Causal Cognitive Architecture

Causal Cognitive Architecture 1 (CCA1)
BICA 2018, 2019, 2020



Hiker lost in the woods…..

Robot goes to the forest to save the hiker….

← Robot

← *Controlled by an CCA1*

As convenience, I will say: ***"CCA1"***

"CCA1" *= Robot + CCA1*

Choose pre-causal functioning of CCA1

Command Prompt - cca1_2020

```
Please choose type of "hippocampus"/"brain" which, of course,
 only loosely approximates the biological equivalent:
1. Lamprey hippocampal/brain analogue
2. Fish hippocampal/telencephalon analogue
3. Reptile hippocampal/pallium analogue
4. Mammalian hippocampus - note: meaningfulness, precausal
5. Human hippocampus - note: meaningfulness plus full causal features
6. Augmented Human level 1 - simultaneous multiple navigational threads
7. Augmented Human level 2 - algorithm center in each navigational module
Please make a selection:
```

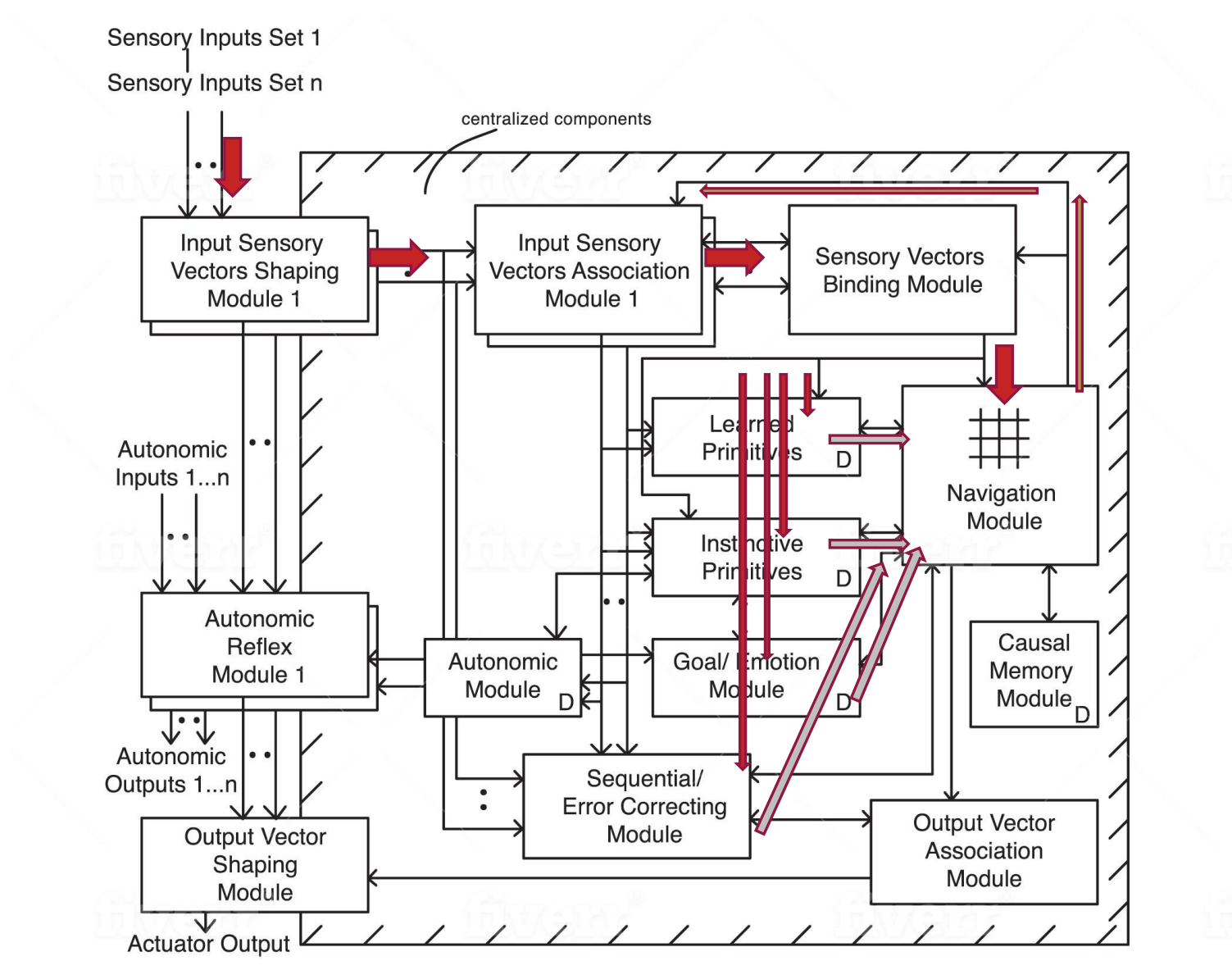
CCA1 must navigate to the lost hiker's square

Command Prompt - cca1_2020

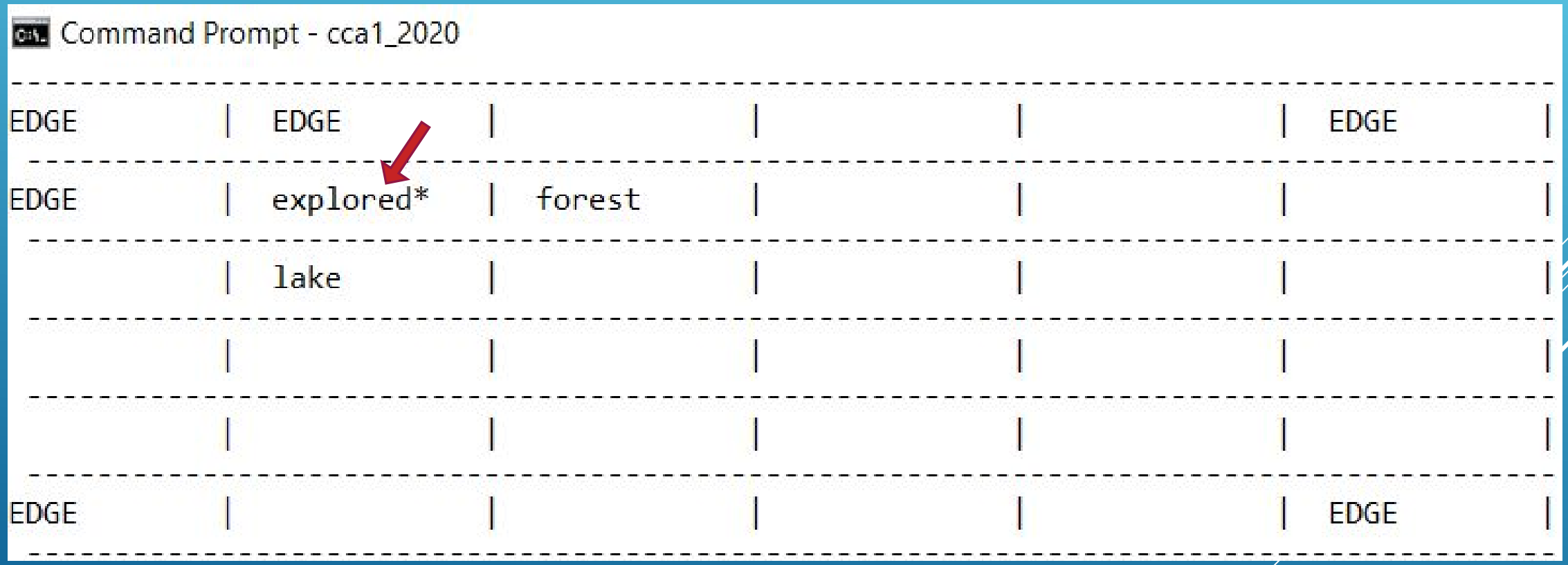
```
hiker position set to:  4 2

Bird's-Eye View of Forest (CCA1 does not have this view)
```

EDGE	EDGE	EDGE	EDGE	EDGE	EDGE
EDGE	CCA1 *	forest	sh_rvr	forest	EDGE
EDGE	lake	forest	forest	forest	EDGE
EDGE	forest	wtrfall	forest	forest	EDGE
EDGE	forest	hiker	forest	forest	EDGE
EDGE	EDGE	EDGE	EDGE	EDGE	EDGE

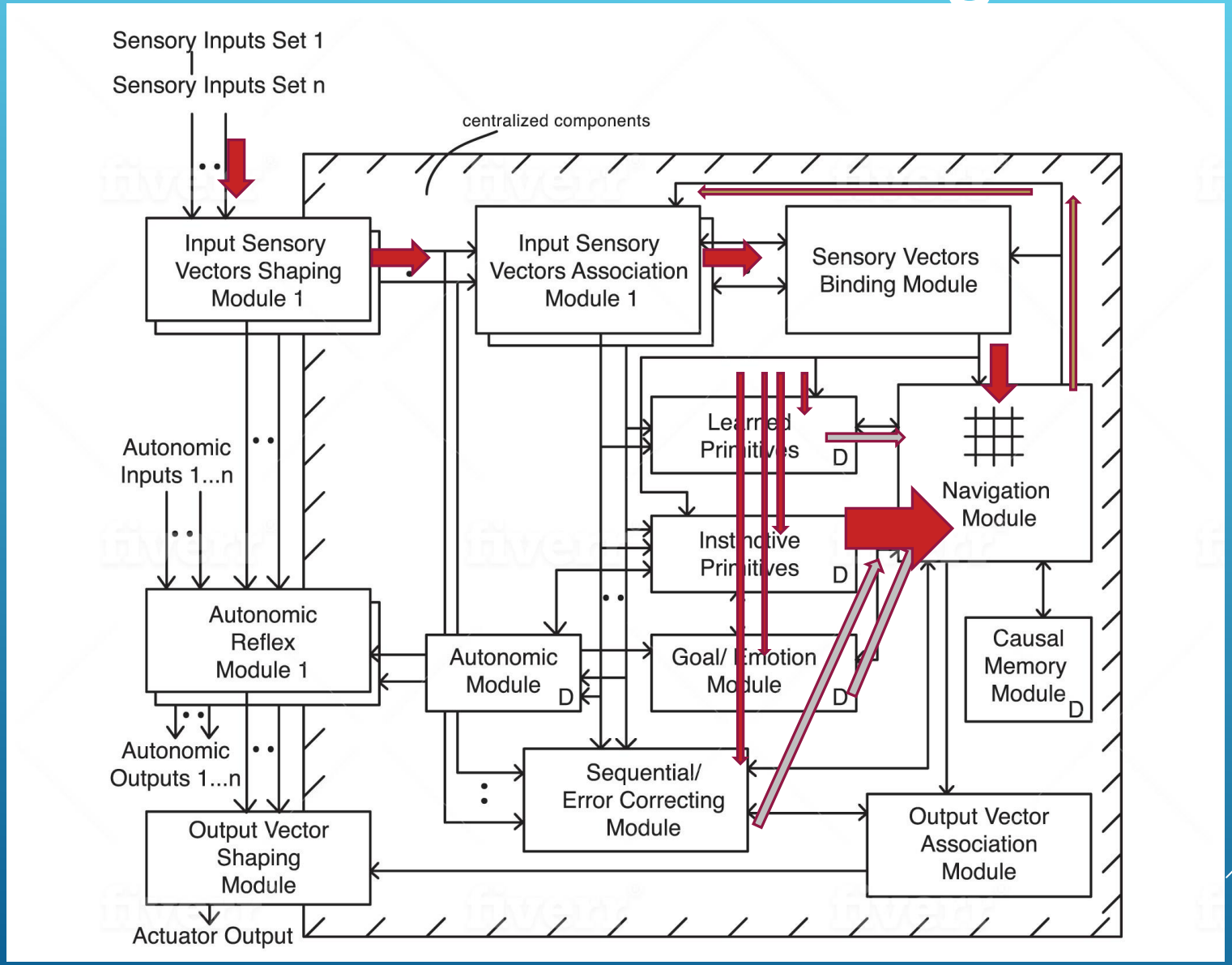
CCA1 – perception….



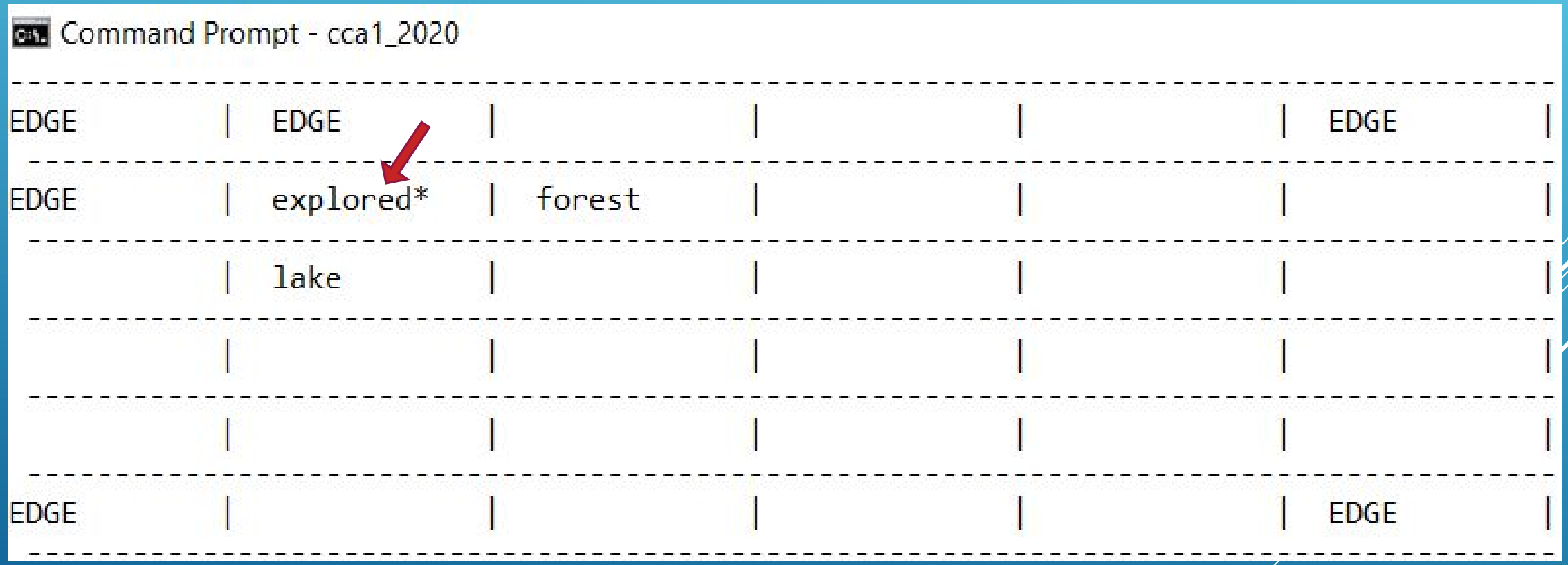
CCA1 builds up internal map from perceptions (and processing) in N, E, S, W directions



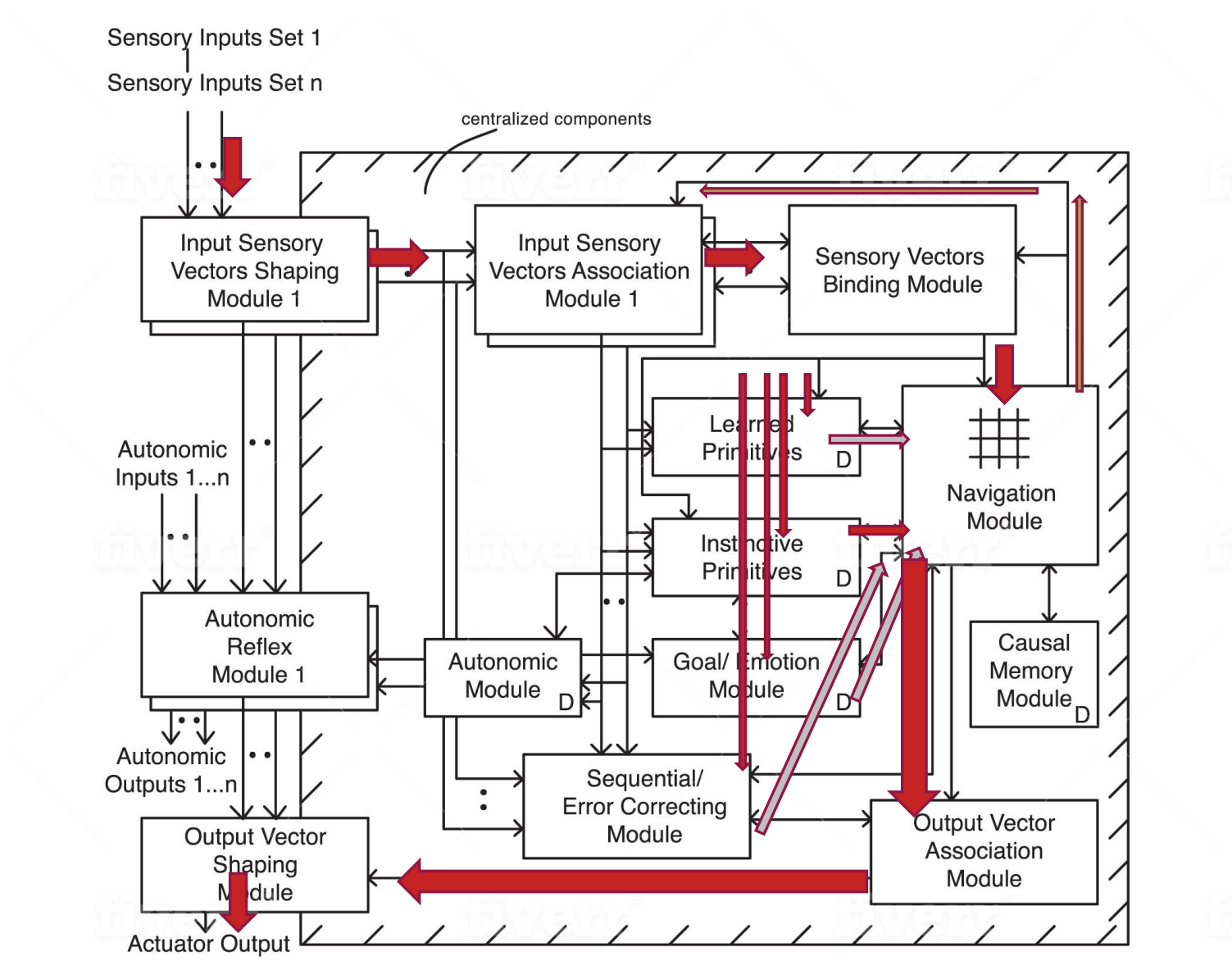
Lake (deep water) – Instinctive Primitive – do not go
Forest – Instinctive Primitive – no signal



CCA1 builds up internal map from perceptions (and processing) in N, E, S, W directions



Navigation to the East (to the forest square)

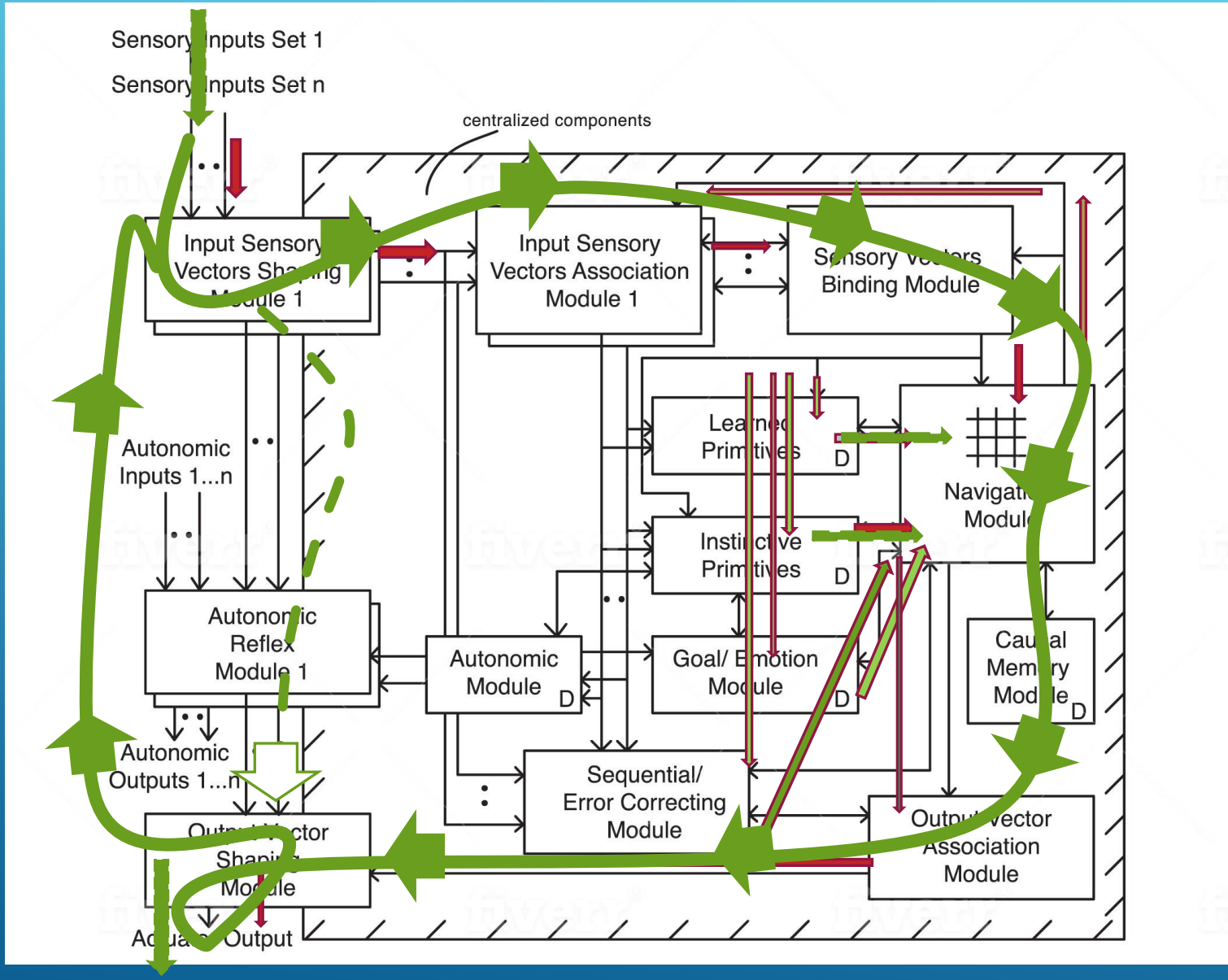


CCA1 moves East into 'forest' square

Command Prompt - cca1_2020

```
CCA1 moved from (1, 1)  1,2

Bird's-Eye View of Forest (CCA1 does not have this view)
-------------------------------------------------------------------------------
EDGE       |  EDGE       |  EDGE       |  EDGE       |  EDGE       |  EDGE       |
 ------------------------------------------------------------------------------
EDGE       |  forest     |  CCA1  *    |  sh_rvr     |  forest     |  EDGE       |
 ------------------------------------------------------------------------------
EDGE       |  lake       |  forest     |  forest     |  forest     |  EDGE       |
 ------------------------------------------------------------------------------
EDGE       |  forest     |  wtrfall    |  forest     |  forest     |  EDGE       |
 ------------------------------------------------------------------------------
EDGE       |  forest     |  hiker      |  forest     |  forest     |  EDGE       |
 ------------------------------------------------------------------------------
EDGE       |  EDGE       |  EDGE       |  EDGE       |  EDGE       |  EDGE       |
```

"Processing Cycles" repeat over and over again

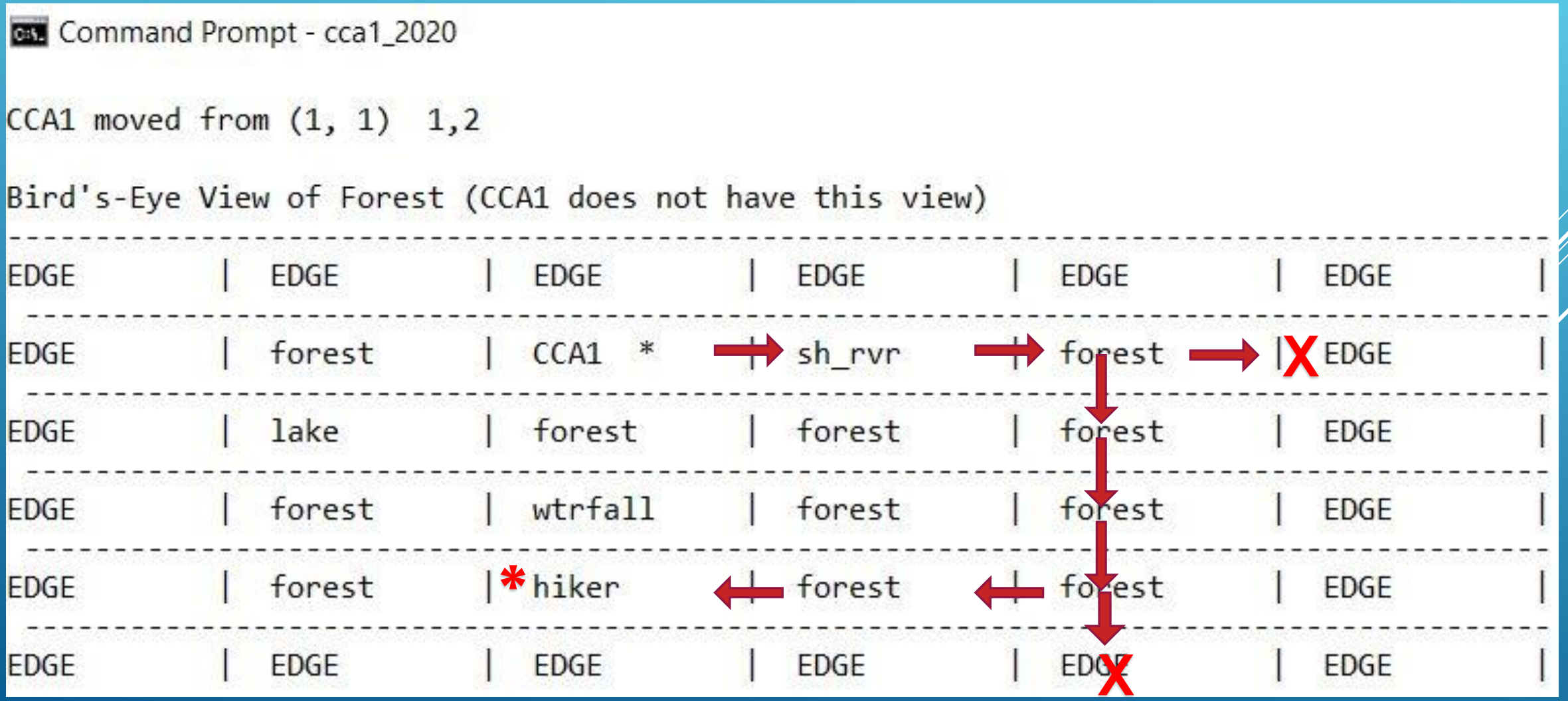


No Special Central Controlling Stored Program

No computer-like clock circuitry centrally controlling CCA1

Vectors propagated from circuit to circuit, and then the cycle is repeated

CCA1 eventually navigates to the hiker square, and rescues the lost hiker



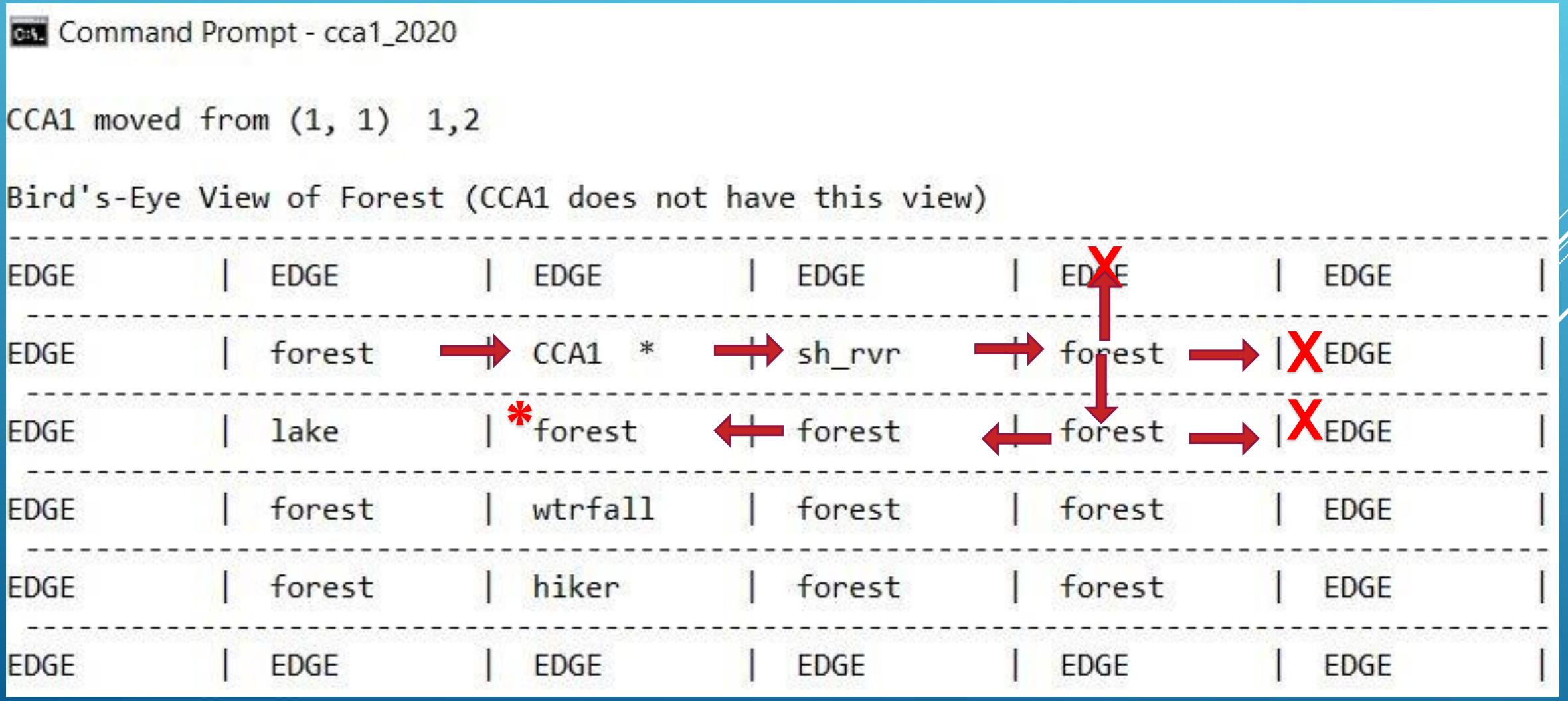
Start new CCA1 simulation….

Command Prompt - cca1_2020

```
hiker position set to:  4 2

Bird's-Eye View of Forest (CCA1 does not have this view)
```

EDGE	EDGE	EDGE	EDGE	EDGE	EDGE
EDGE	CCA1 *	forest	sh_rvr	forest	EDGE
EDGE	lake	forest	forest	forest	EDGE
EDGE	forest	wtrfall	forest	forest	EDGE
EDGE	forest	hiker	forest	forest	EDGE
EDGE	EDGE	EDGE	EDGE	EDGE	EDGE

CCA1 moves to north of the waterfall square….



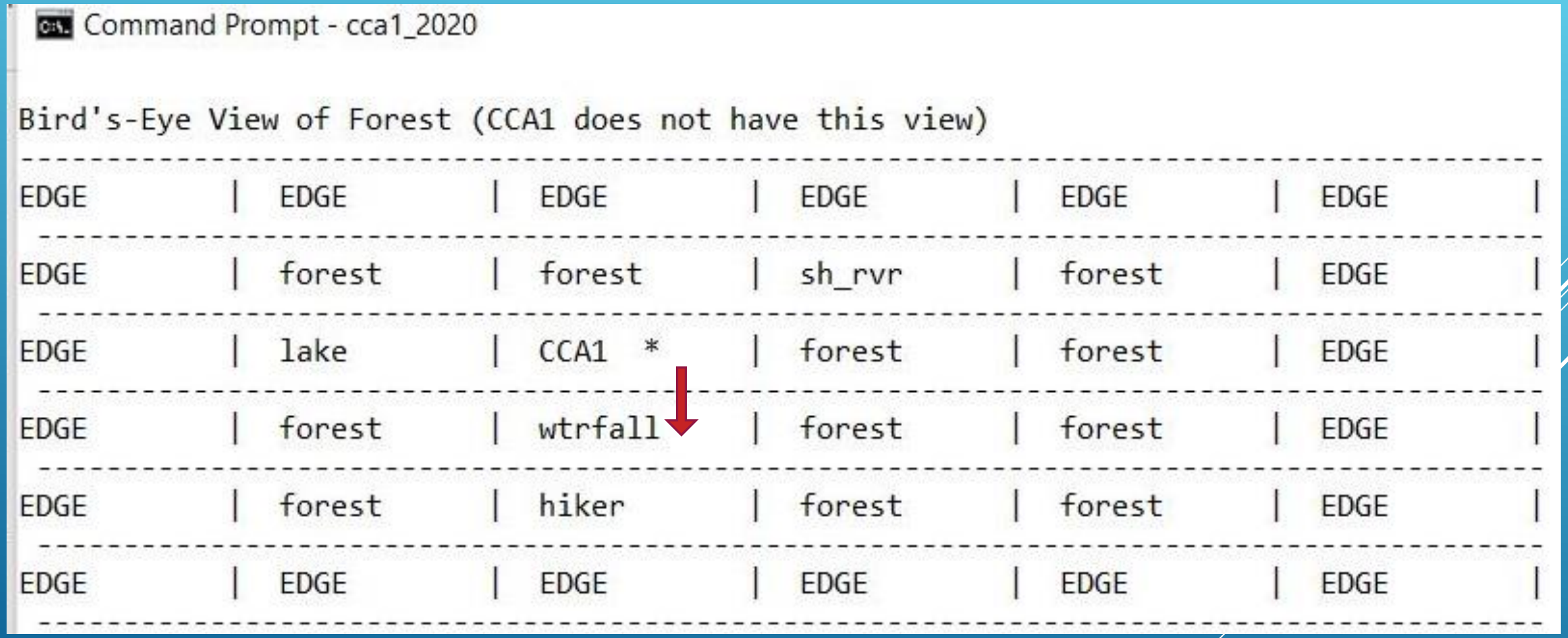
CCA1 has moved north of the waterfall square…

Command Prompt - cca1_2020

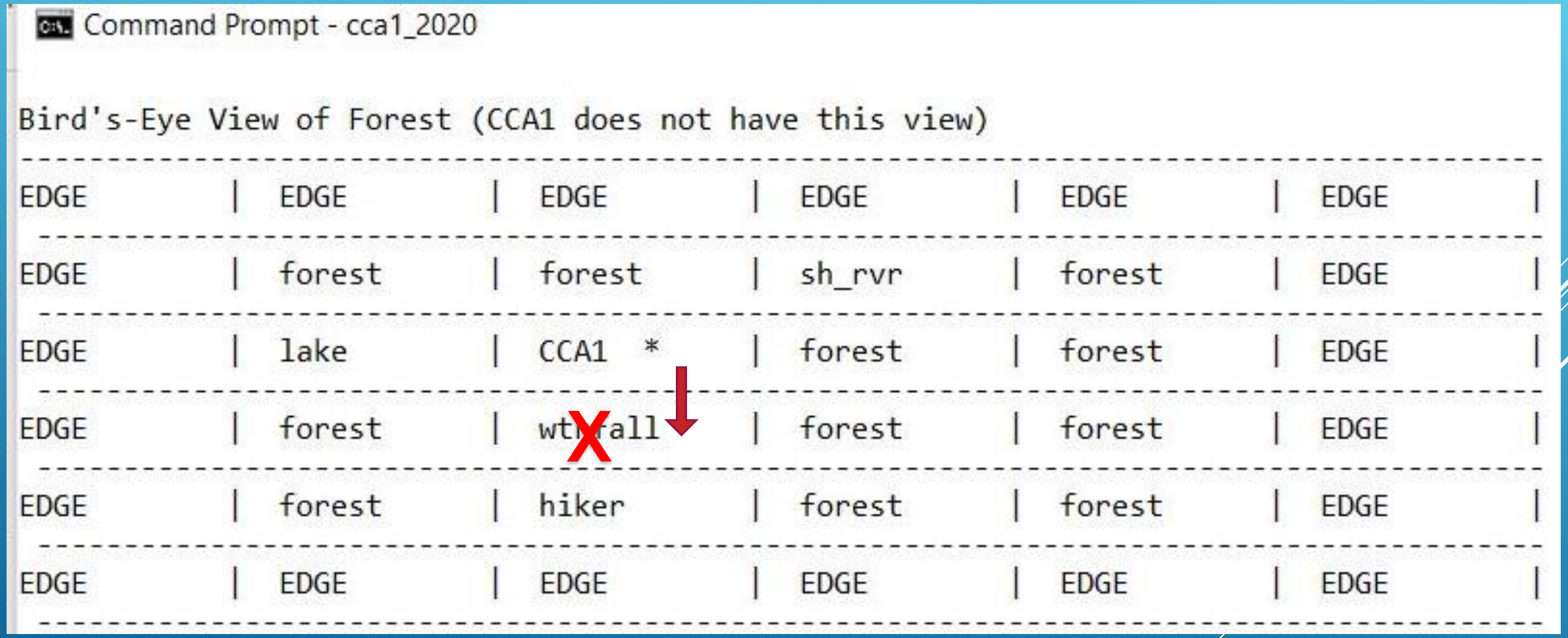
```
Bird's-Eye View of Forest (CCA1 does not have this view)
-------------------------------------------------------------------------------
EDGE       | EDGE       | EDGE       | EDGE       | EDGE       | EDGE       |
-------------------------------------------------------------------------------
EDGE       | forest     | forest     | sh_rvr     | forest     | EDGE       |
-------------------------------------------------------------------------------
EDGE       | lake       | CCA1  *    | forest     | forest     | EDGE       |
-------------------------------------------------------------------------------
EDGE       | forest     | wtrfall    | forest     | forest     | EDGE       |
-------------------------------------------------------------------------------
EDGE       | forest     | hiker      | forest     | forest     | EDGE       |
-------------------------------------------------------------------------------
EDGE       | EDGE       | EDGE       | EDGE       | EDGE       | EDGE       |
-------------------------------------------------------------------------------
```

CCA1 has never seen a waterfall before…. just sees a river (noisy….fast flowing)…. and is generally able to cross rivers

S – sees fast noisy river (does not see cliff part) Able to cross shallow rivers, so moves South



CCA1 moves S (south) and is swept off cliff of waterfall and is damaged – mission ends

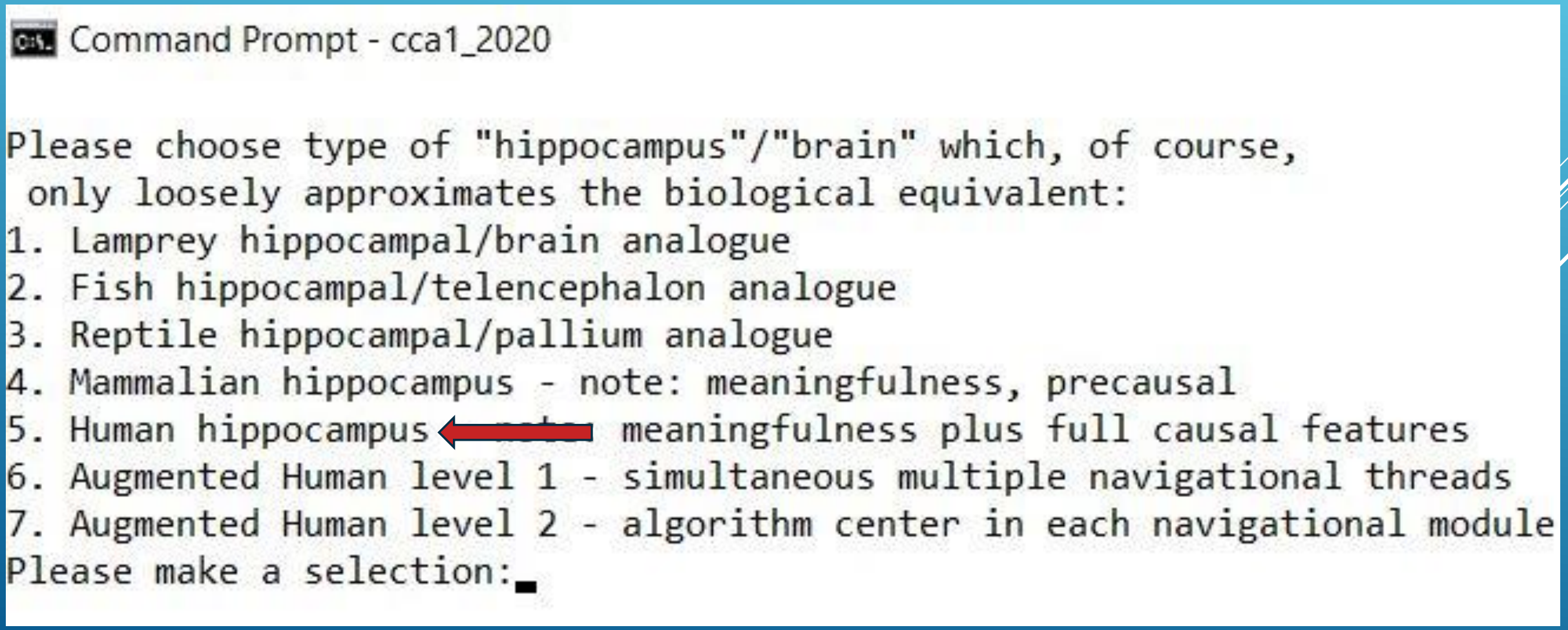


Failure of mission

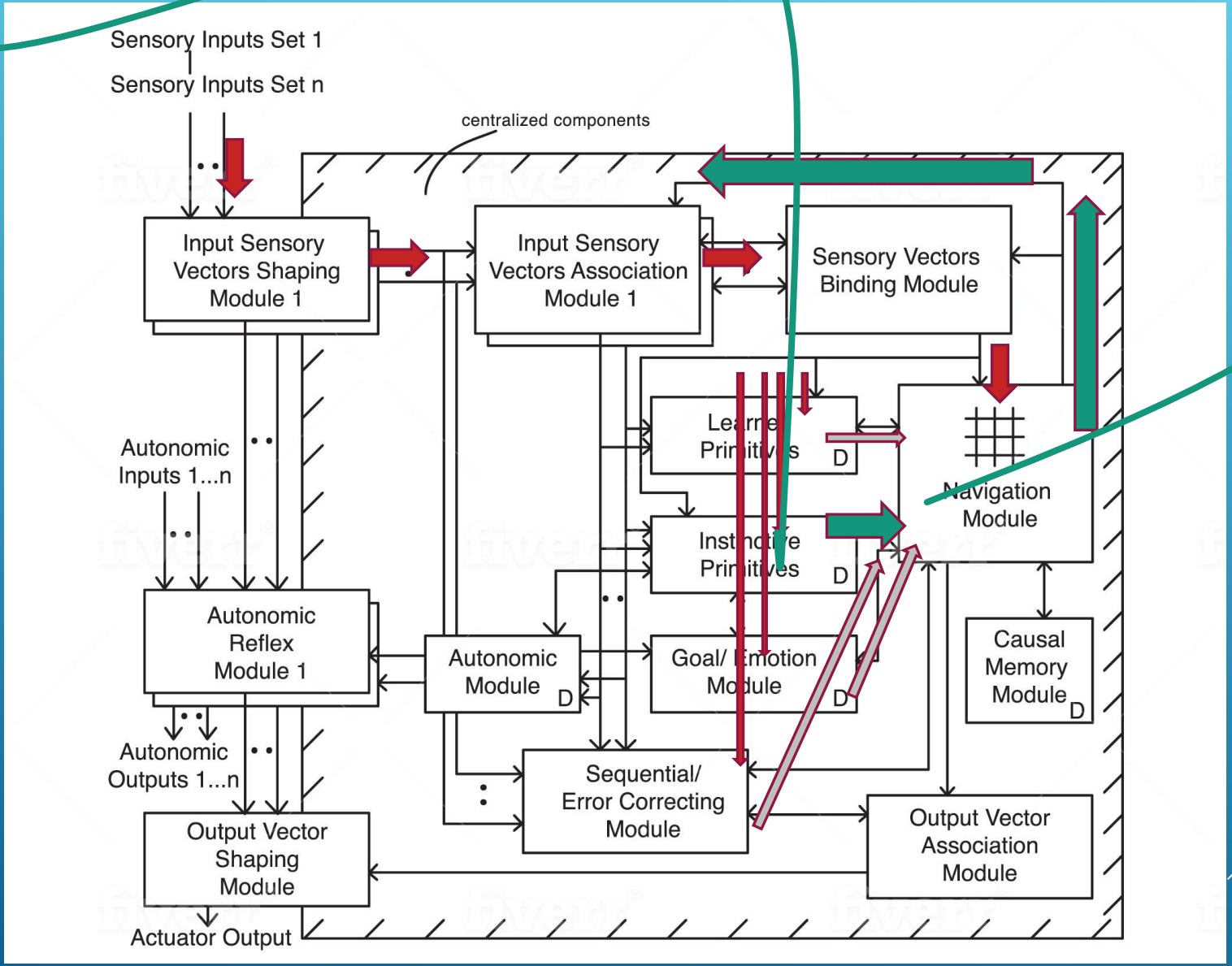
Associative Learning Does Occur

-If repaired and it goes out into the forest on another mission
-Sees fast flowing river with much noise
-Triggers in Goal/Emotion Module and Learned Primitives Module not to go there
-Makes another choice for direction of move

-New Simulation
-Use full causal features of architecture



{“water”} + {“fast flow” + “noise”} → {“water” + “push”}



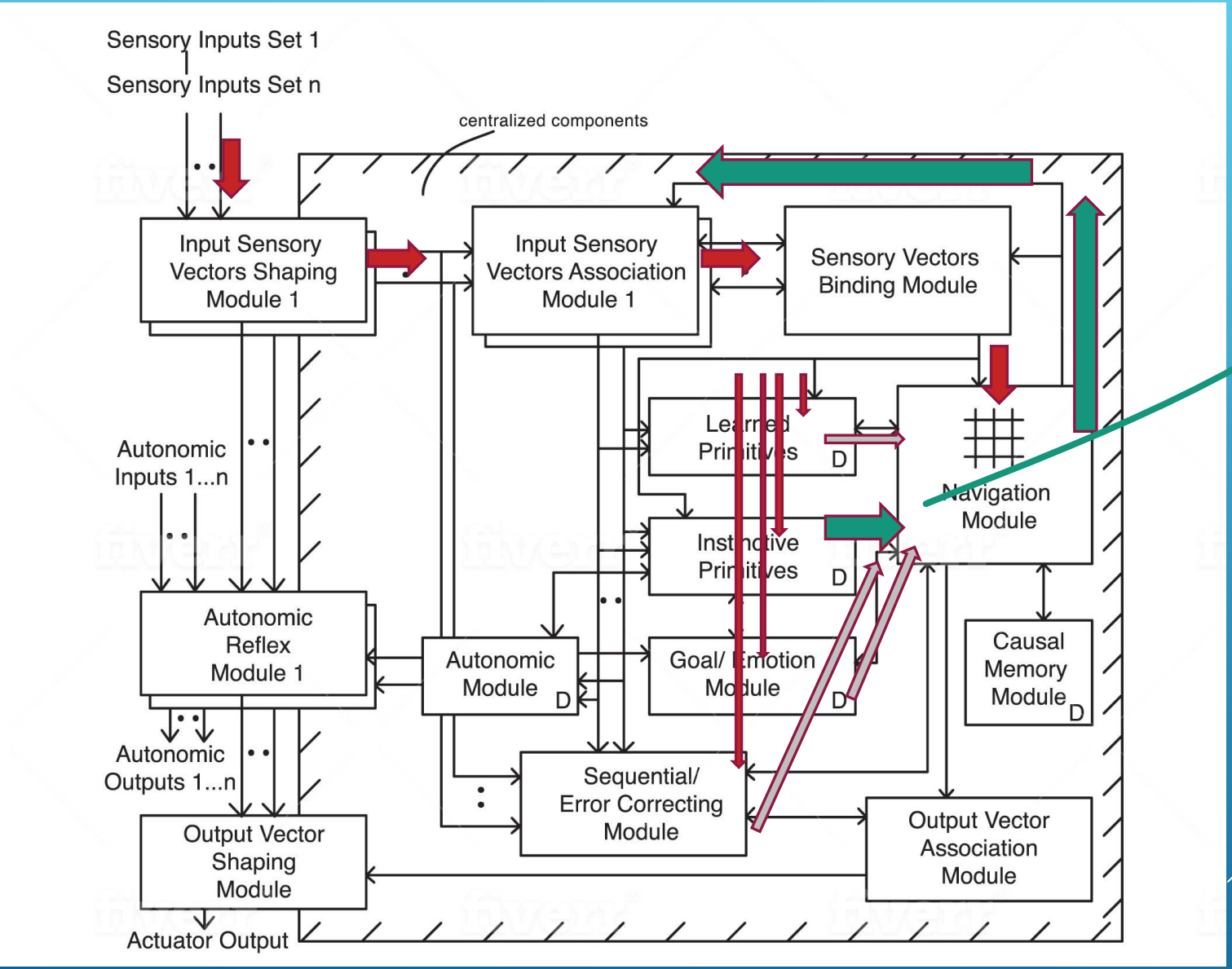
Temporary map → {“CCA1 under water”}

Command Prompt - cca1_2020

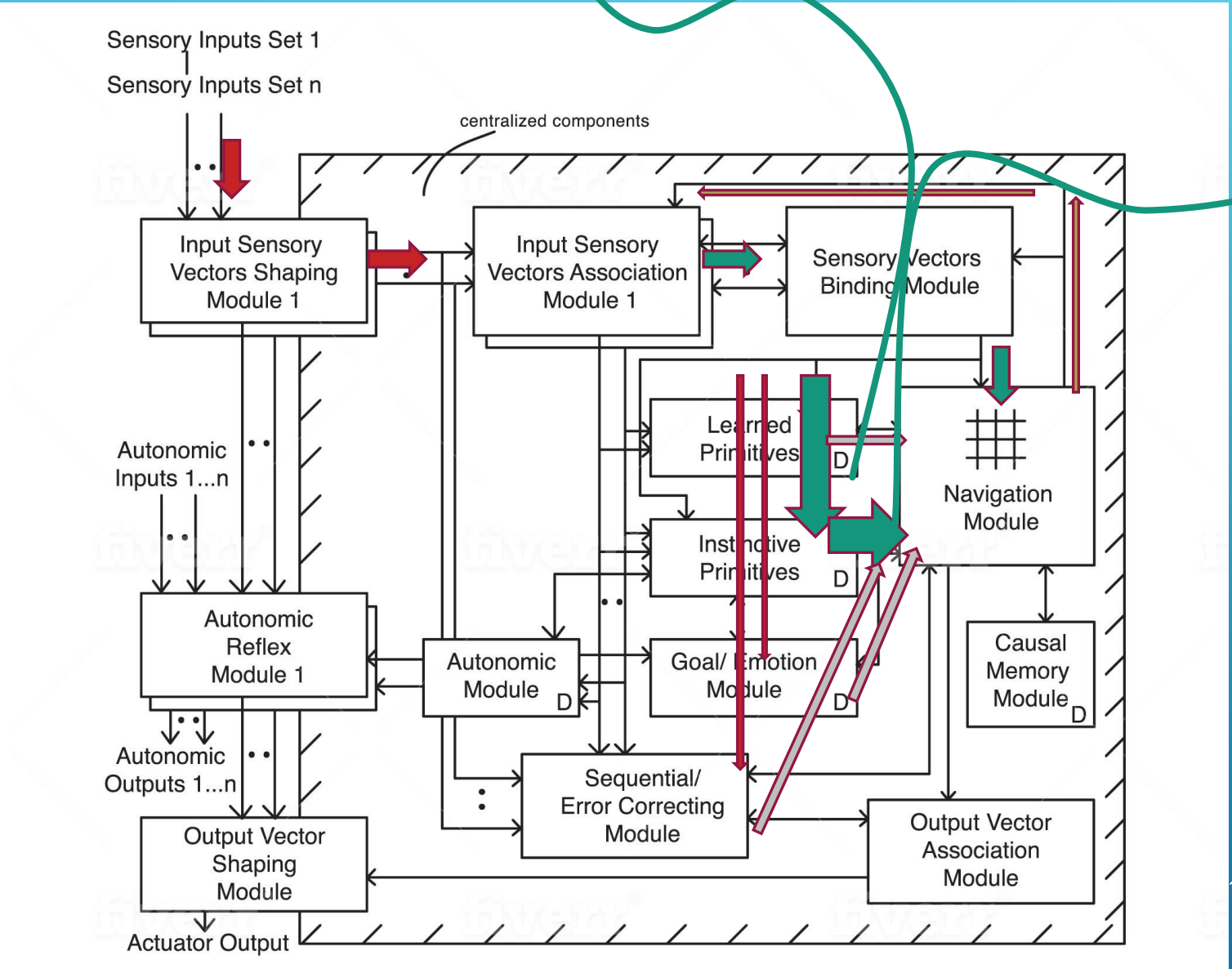
Internal Map From Stack

air	air	air	air	air	air
water	water	water	water	water	water
water	water	water	water	water	water
water	water	water	CCA1 *	water	water
water	water	water	water	water	water
water	water	water	water	water	water
water	water	water	water	water	water

{“CCA1 under water”} is fed back to sensory input module



{“CCA1 under water”}

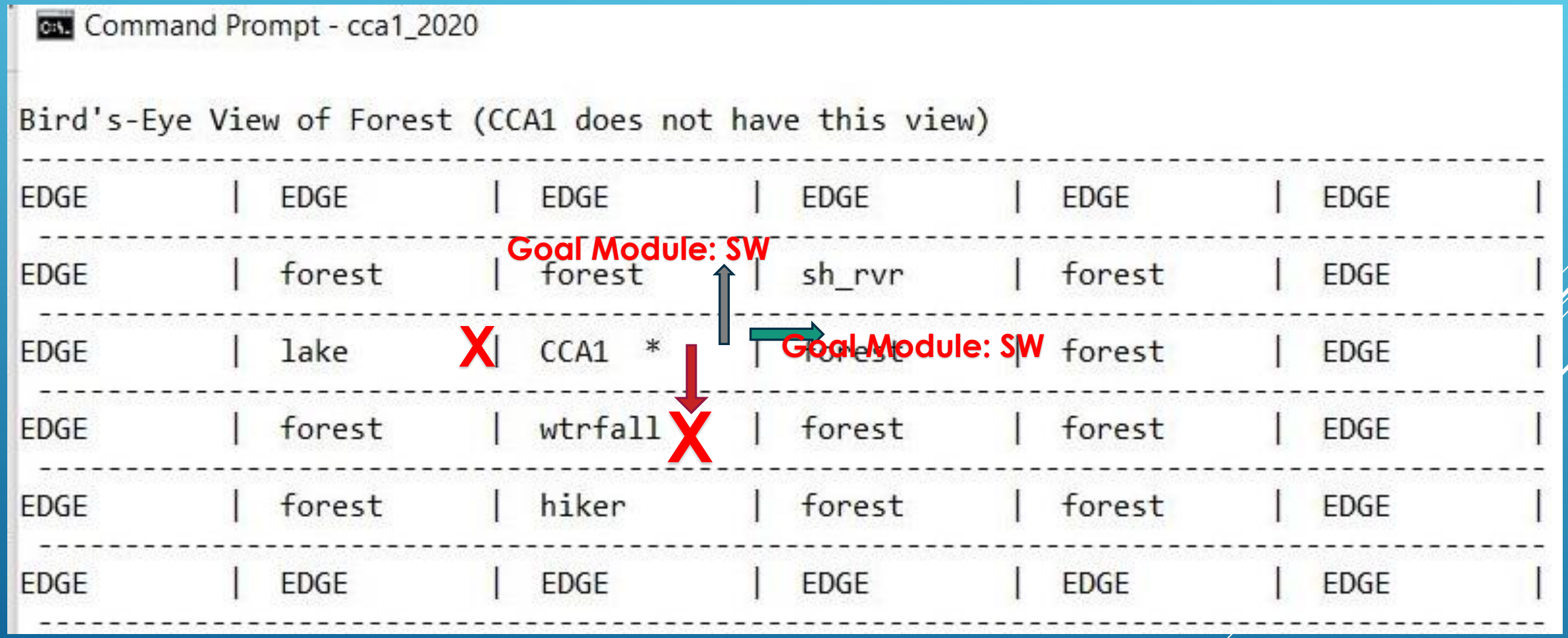


“do not go”

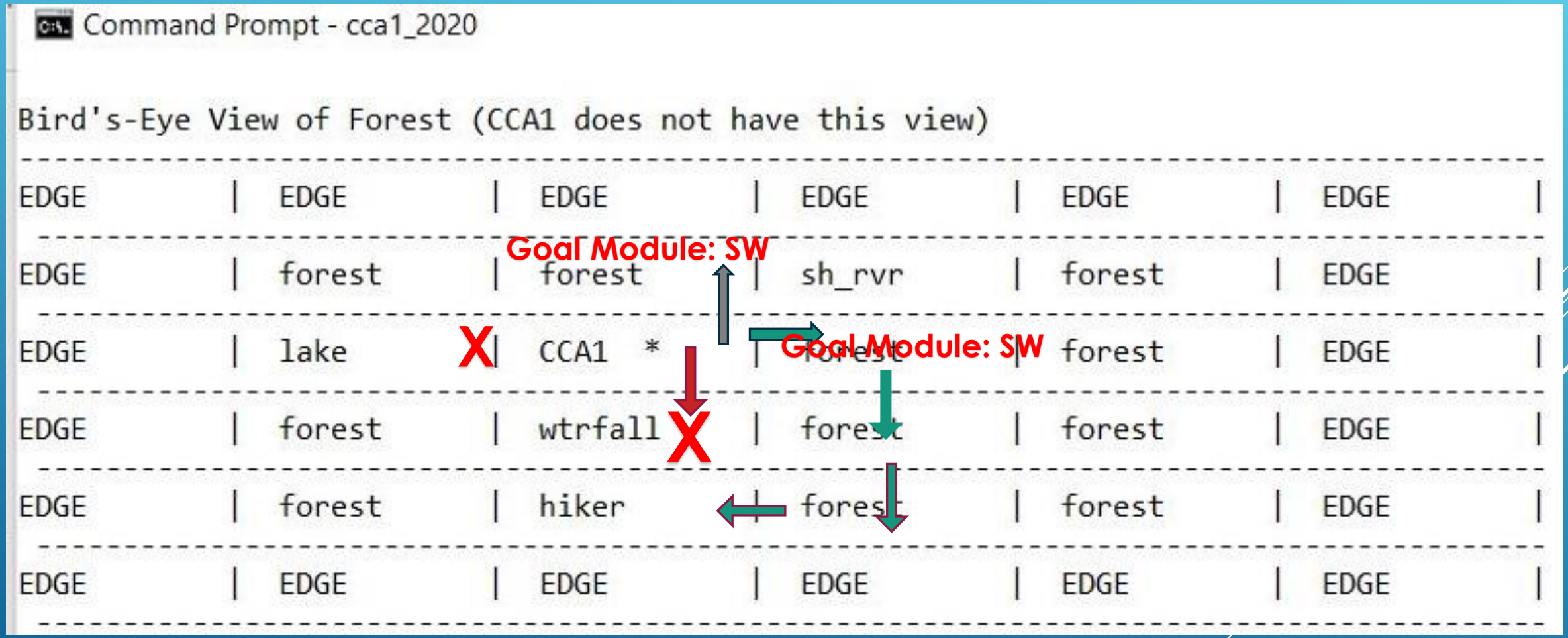
->retrieve previous temporary map

->do not go south

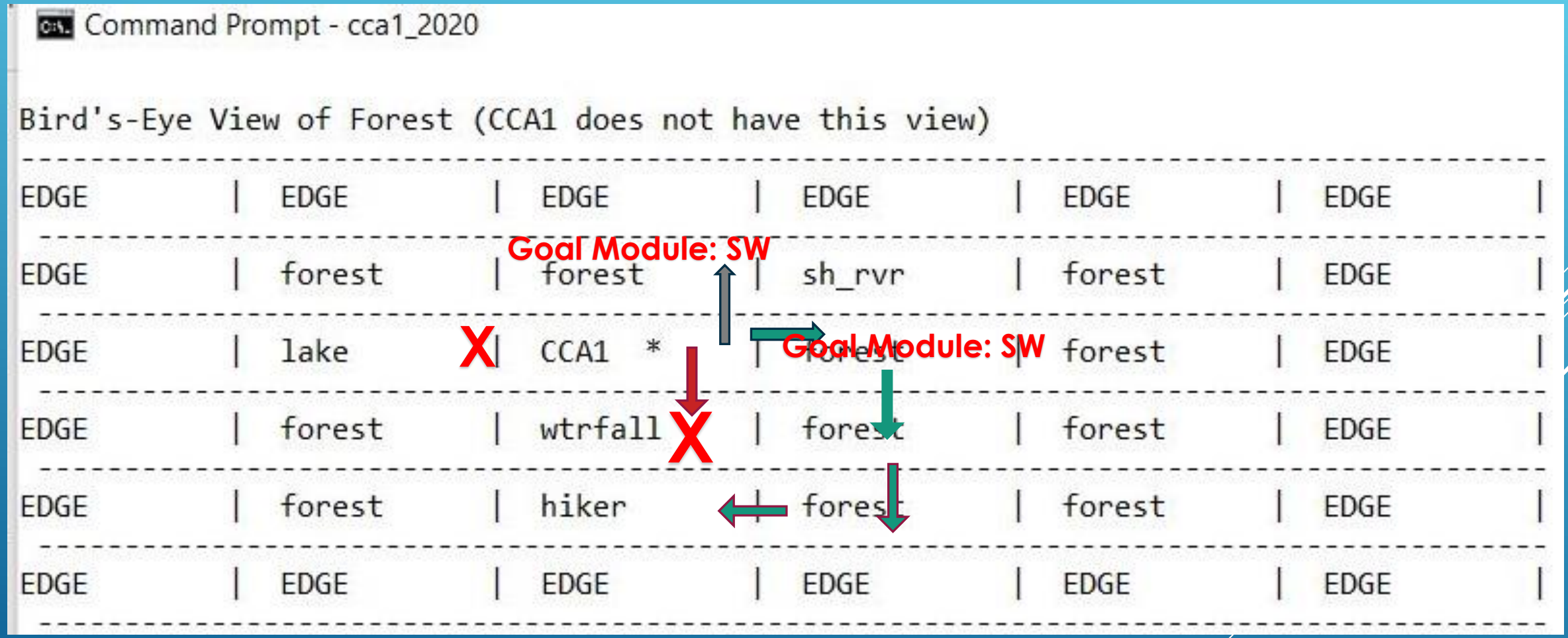
Do not go south – goes east even though bias from Goal Module to go south or west.



Continues south and then west…. and…. Rescues the lost hiker



Even though CCA1 had never seen a waterfall before, **it causally avoided this danger**

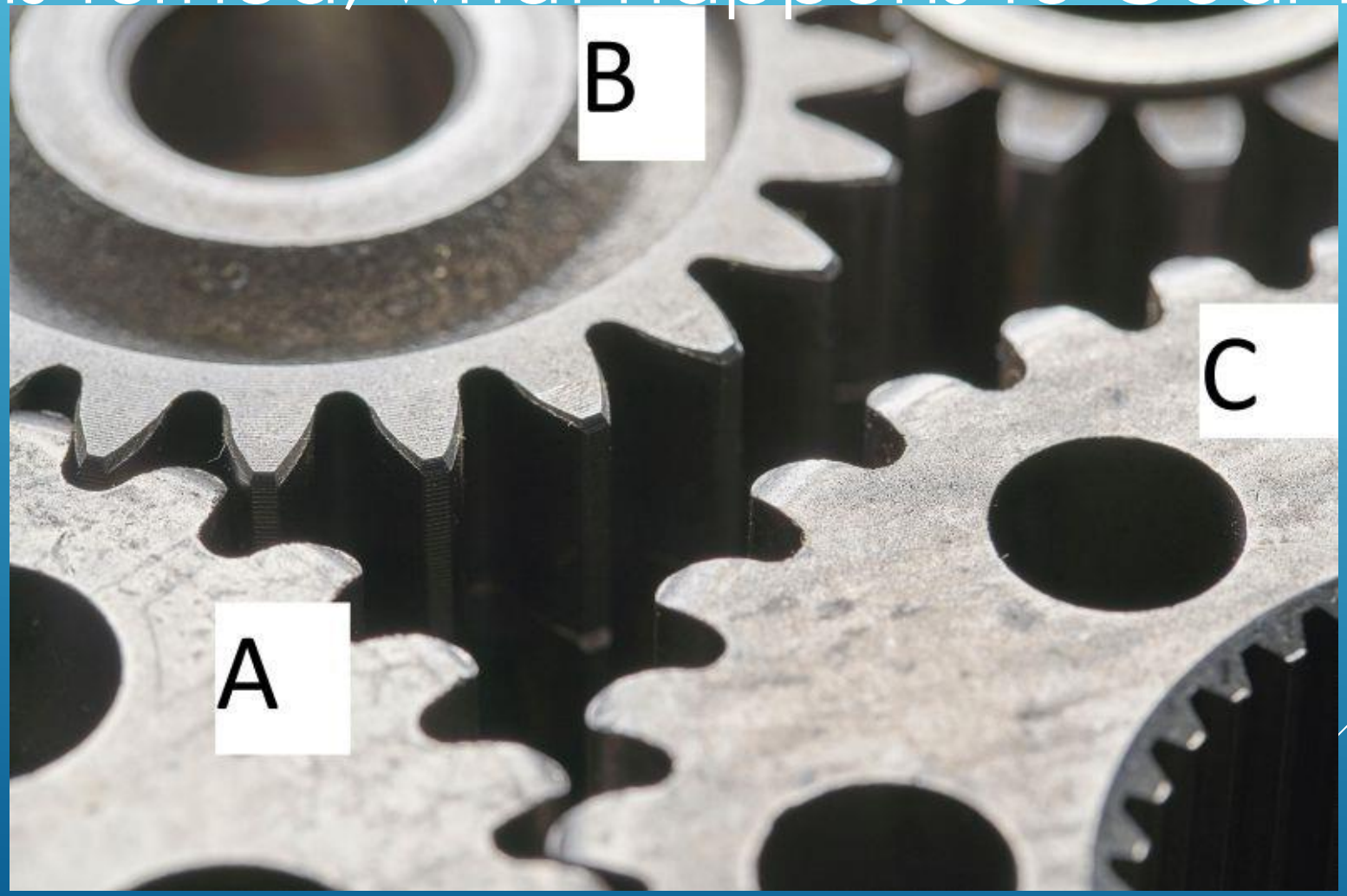


Causality emerges from the architecture

No central controlling stored program other than the repeating processing cycles of the CCA1

New simulation – CCA1 is inspecting a broken machine it has never seen before.

If Gear C is turned, what happens to Gear B?



Gear C is recognized and added to create a new temporary map

Command Prompt - cca1_2020

```
Internal Map From Stack
-------------------------------------------------------------------------------
air*        |  air        |  air        |  air        |  air        |  air        |
 ------------------------------------------------------------------------------
air         |  *push      |  air        |  air        |  air        |  air        |
 ------------------------------------------------------------------------------
C           |  A;moves    |  B;moves    |  air        |  air        |  air        |
 ------------------------------------------------------------------------------
air         |  air        |  air        |  air        |  air        |  air        |
 ------------------------------------------------------------------------------
air         |  air        |  air        |  air        |  air        |  air        |
 ------------------------------------------------------------------------------
air         |  air        |  air        |  air        |  air        |  air        |
 ------------------------------------------------------------------------------
air         |  air        |  air        |  air        |  air        |  air        |
```

-Cannot fully repair a machine with 100's of parts by associations only (unless very common reasons for the breakdowns)

– even if only move a few parts there are millions and millions of combinations that need to be tried and learned by association ->simply not possible/practical

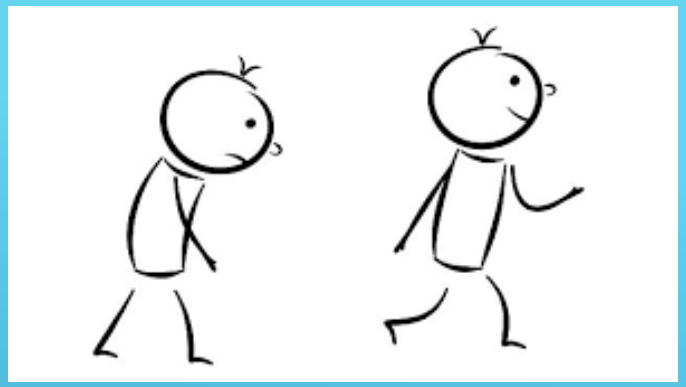
- Causality allows repairing a machine the CCA1 has never seen before.

⟹ **Causality emerges from the architecture**

Plausible evolutionary transition from Associative Behavior to a Causal Behavior

Small enhancements in circuitry allow this as shown in pre-causal to fully causal operation of the CCA1

Analogies



Should rescue CCA1 spend more time with person A or person B?
-Person B smiles a lot but is noisy, compared to Person A

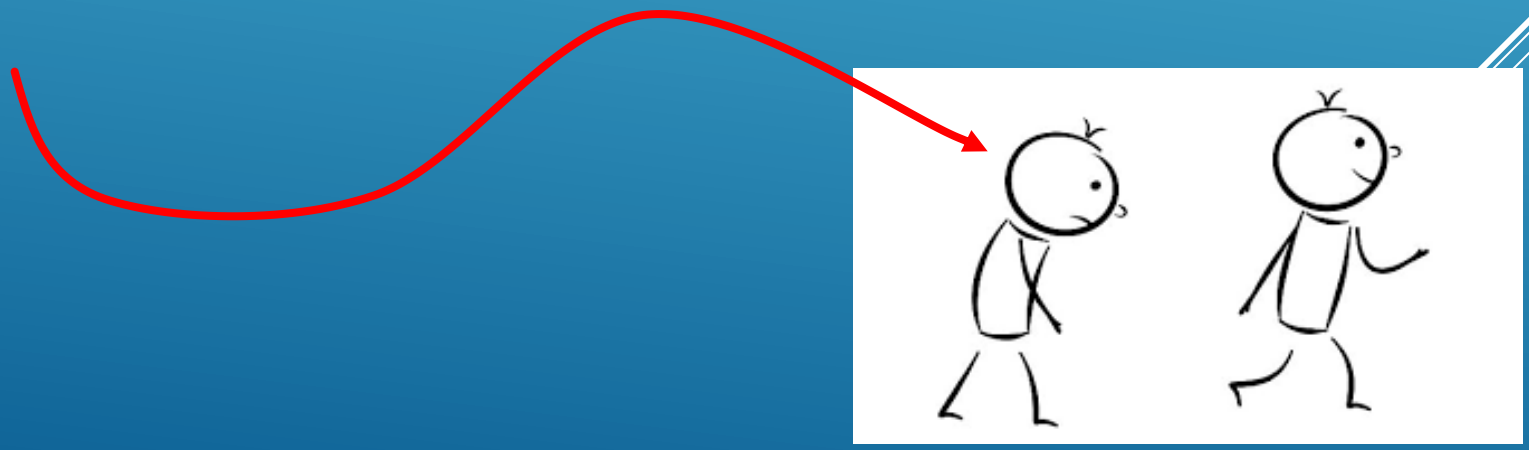
-Who to chose?

→ Question for a philosopher!!

.

→CCA1's architecture and temporary maps, readily form and use analogies

There is a navigation output to navigate to object A (i.e., person A)



Explainability

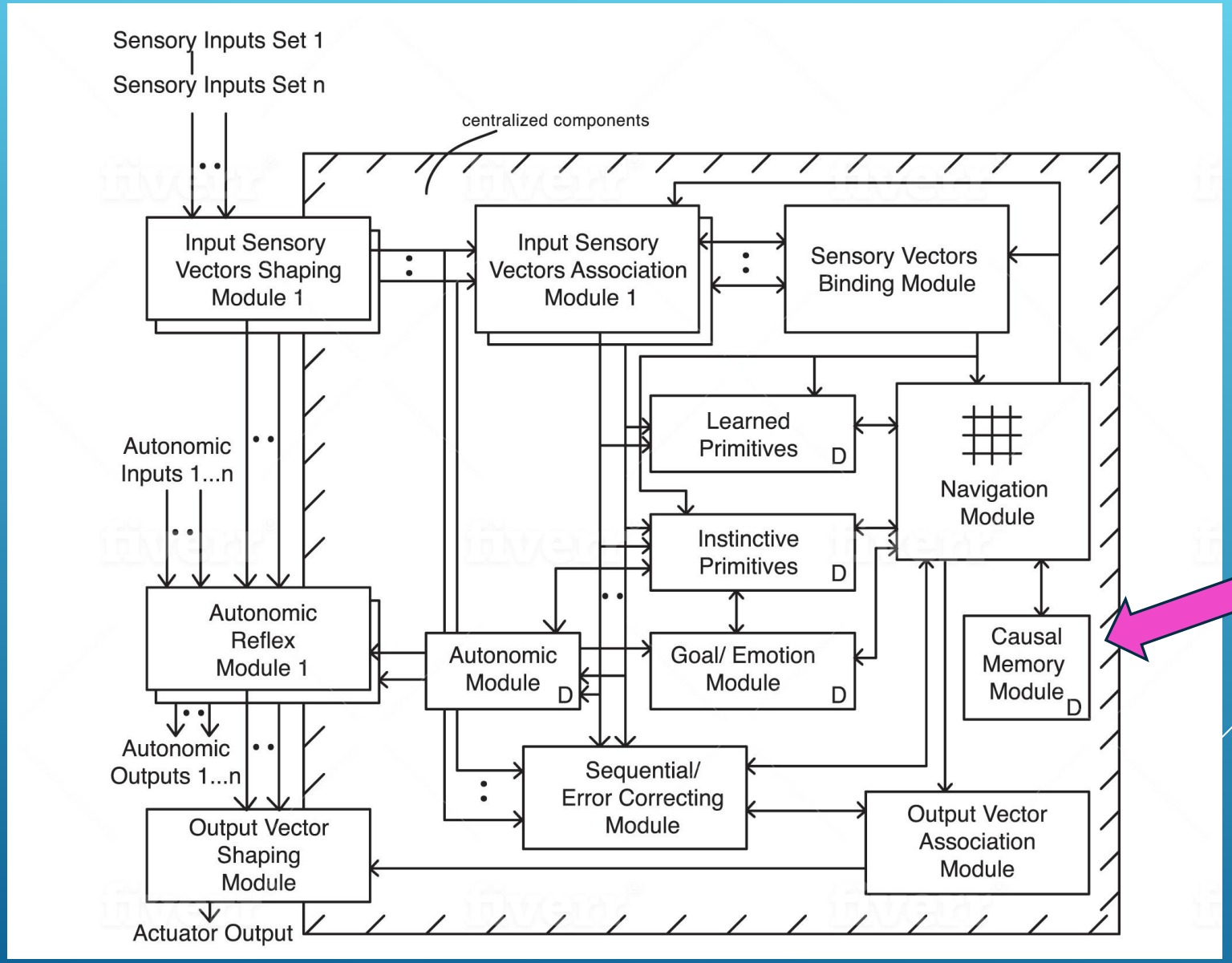
After being used, ‘temporary maps’ are actually stored permanently in the Causal Memory portion of the Navigation Module

CCA1 Supports Schneider Psychosis Hypothesis

Schneider –BICA 2019:
-Imperfect functioning in going from precausal to full causal behavior (more complexity, feeding partial results back to sensory modules) can result in psychotic behavior (hallucinations, delusions and reduced cognition)

Causal Cognitive Architecture 1 (CCA1)

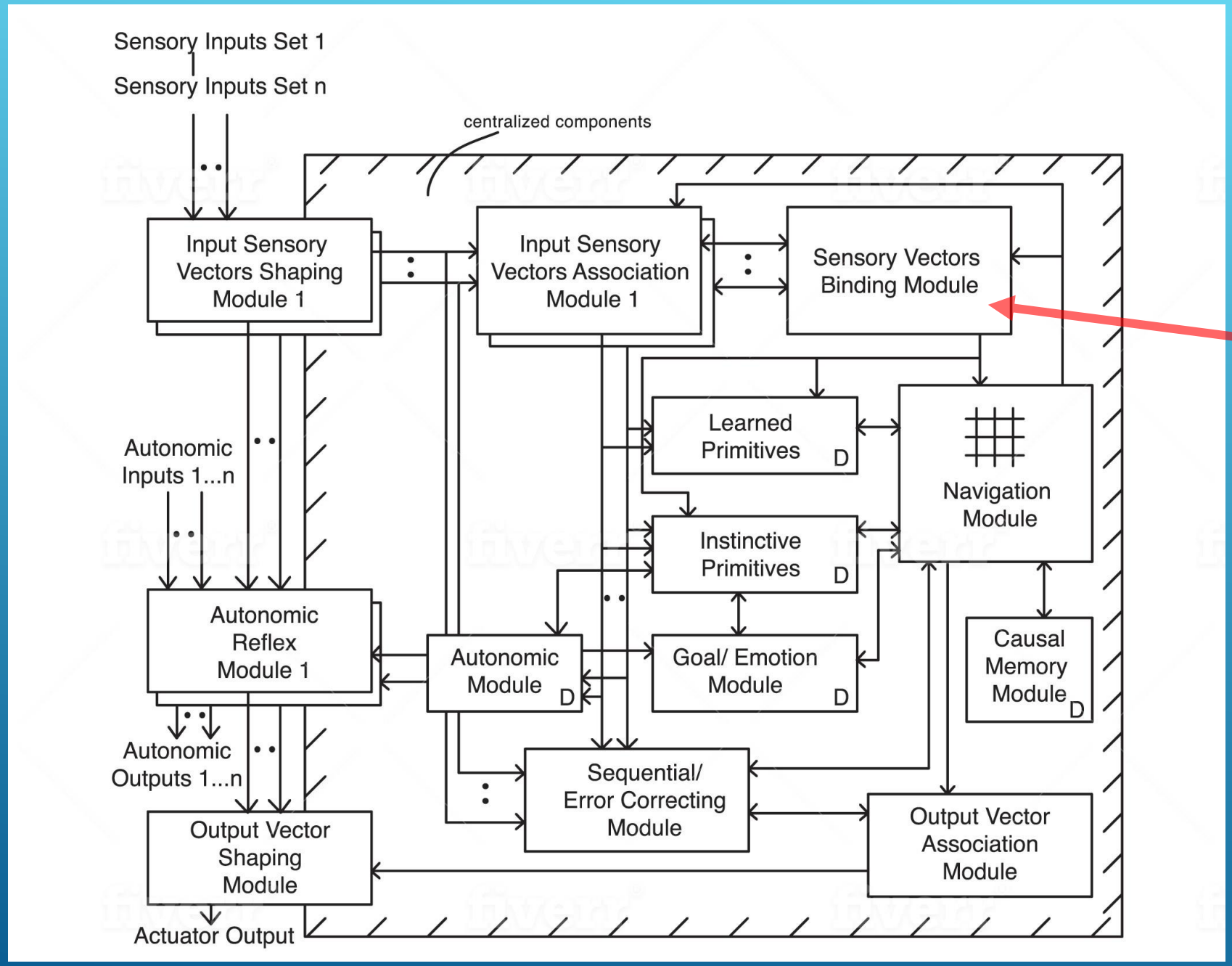
BICA 2020



- CCA1 handles toy problems
- Want a more robust version of CCA1

but….problems Arising in Attempts to Enhance the CCA1

- Problem is that Sensory Vectors Binding Module must output some vector which represents object/environment it has detected by fusing sensory features together
- Then use neural network-like pattern recognition to identify the objects and sensory scene

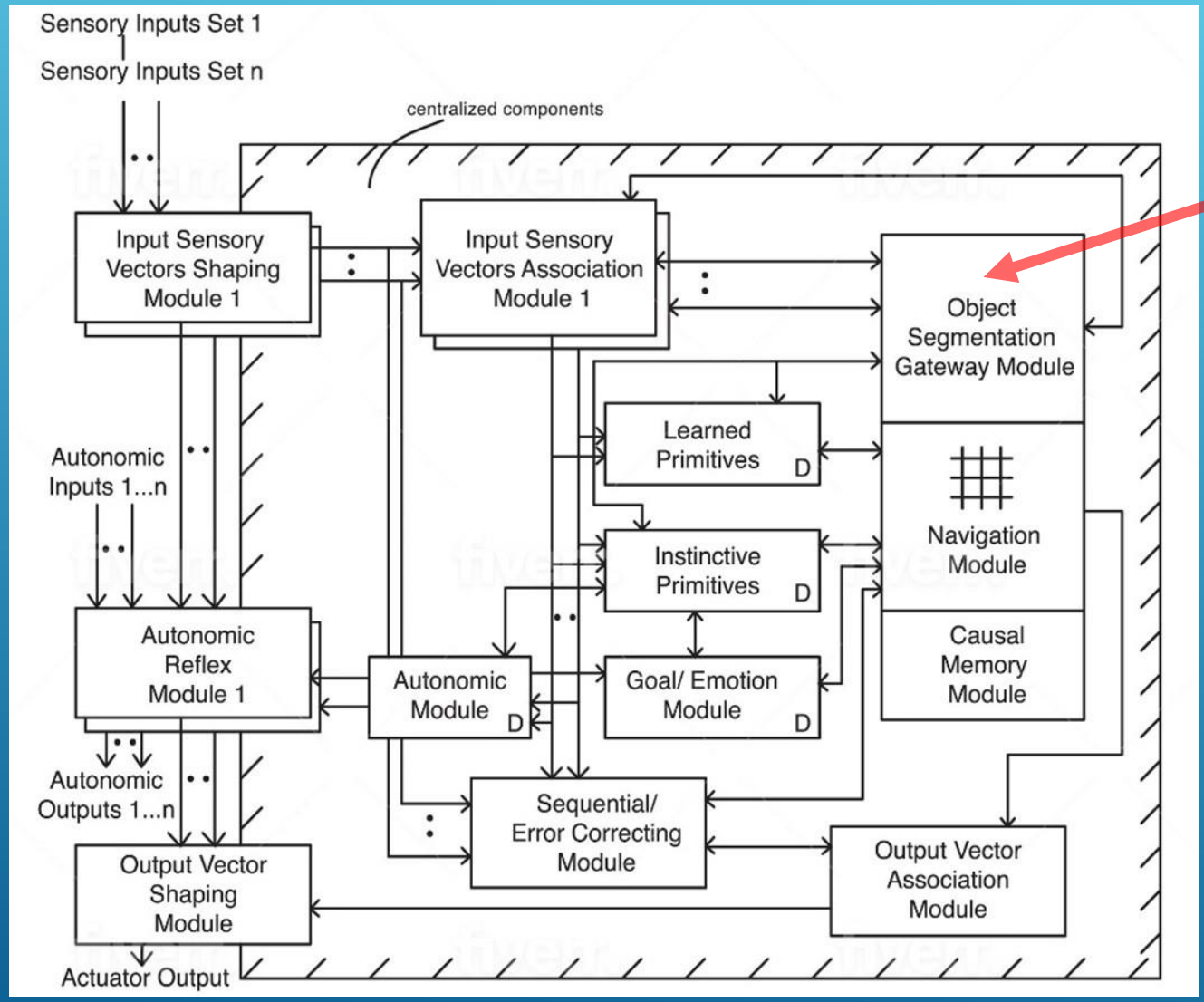

Sensory Inputs Set 1
Sensory Inputs Set n
centralized components
Input Sensory Vectors Shaping Module 1
Input Sensory Vectors Association Module 1
Sensory Vectors Binding Module
Learned Primitives D
Navigation Module
Instinctive Primitives D
Autonomic Inputs 1...n
Autonomic Reflex Module 1
Autonomic Module D
Goal/ Emotion Module D
Causal Memory Module D
Autonomic Outputs 1...n
Sequential/ Error Correcting Module
Output Vector Association Module
Output Vector Shaping Module
Actuator Output

- Problem is that Sensory Vectors Binding Module must output some vector which represents object/environment it has detected by fusing sensory features together
- **How to label different combinations?**
- **Need a binding language of sorts**

Solution → Problems Arising in Attempts to Enhance the CCA1

1. Eliminate the Sensory Vectors Binding Module
2. Bind sensory inputs directly in the Navigation Module

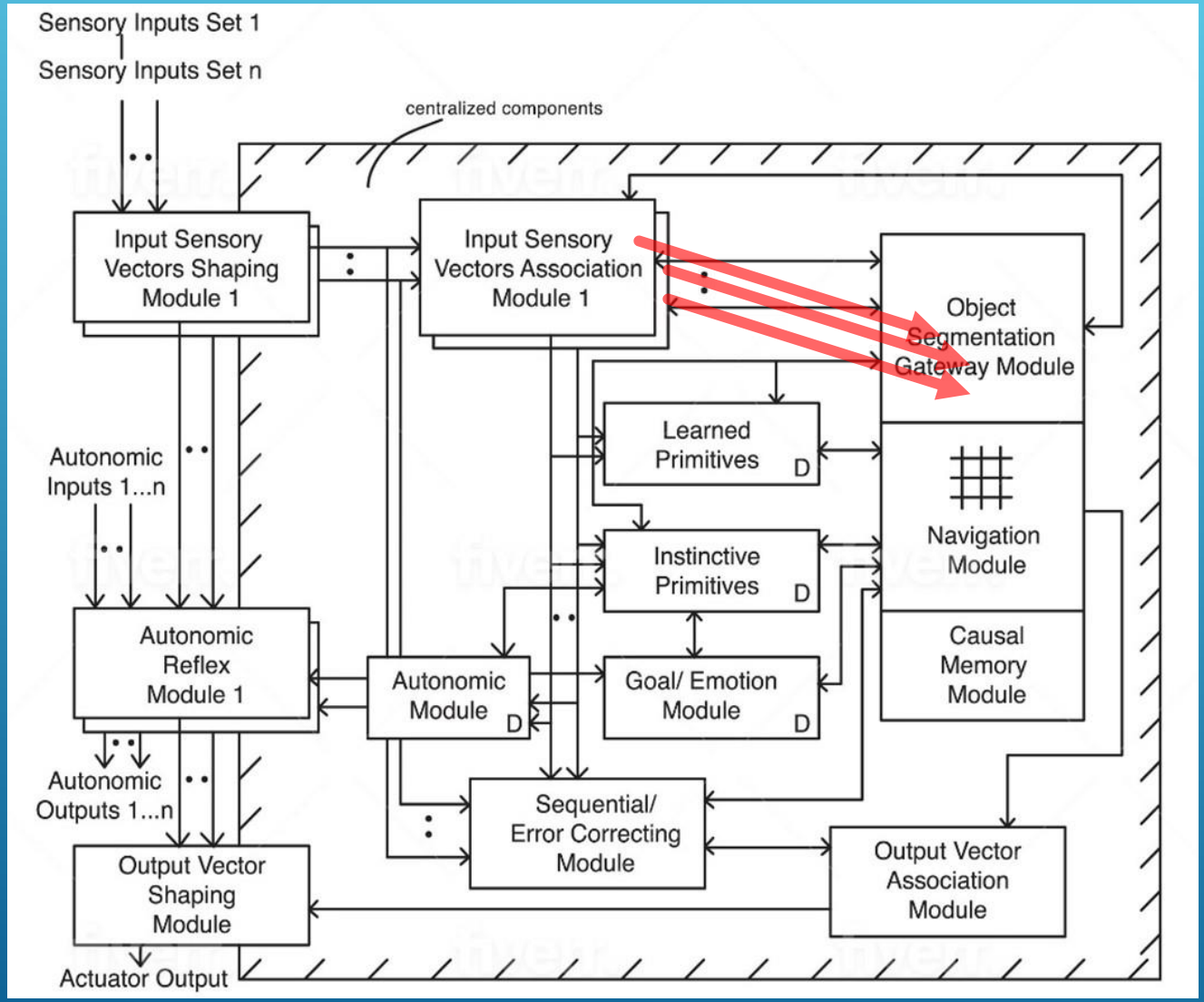
Solution → Problems Arising in Attempts to Enhance the CCA1



← CCA2

Binding Sensory Inputs in the Navigation Module

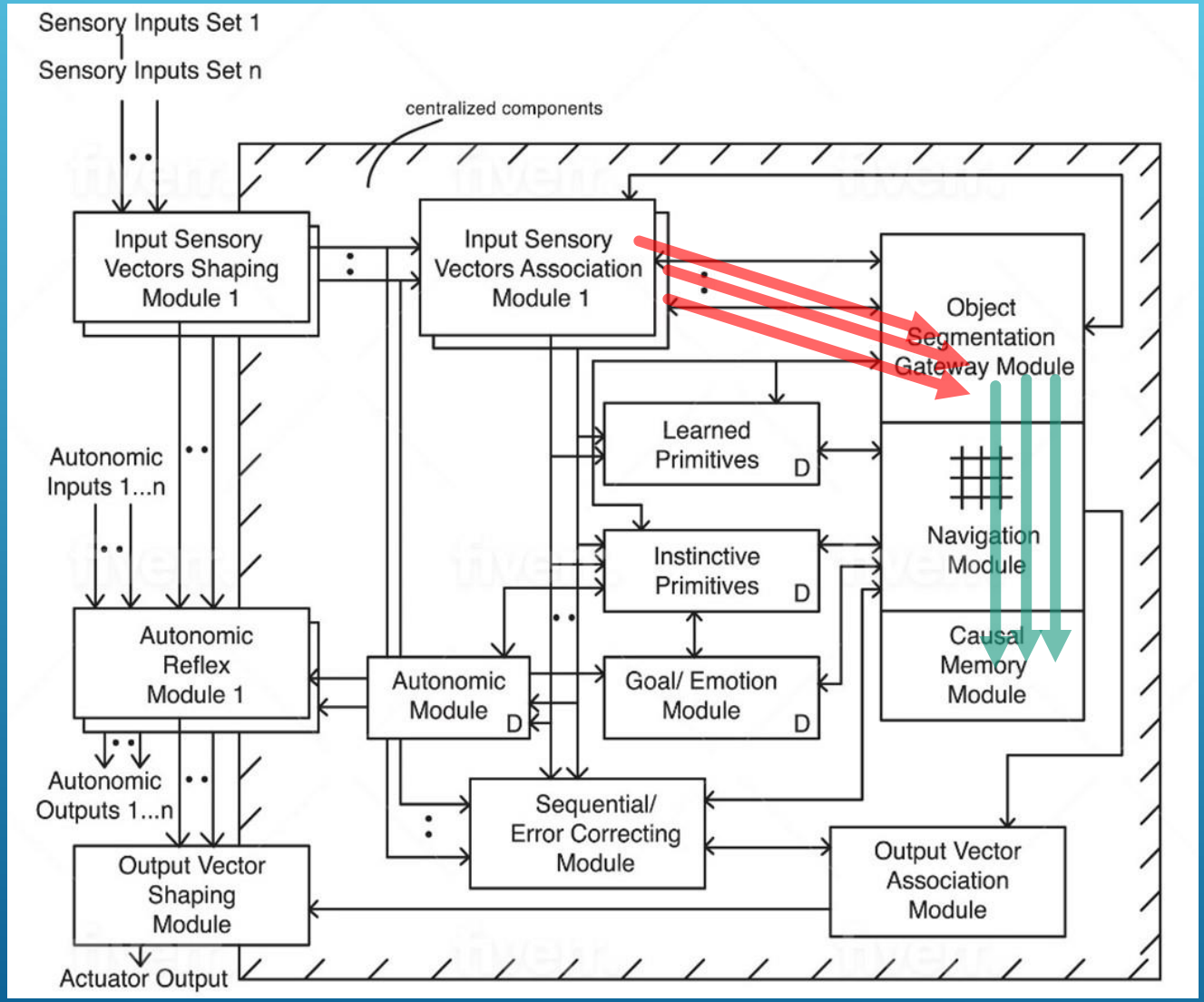
1. Each Input Sensory Vectors Association Module (visual, auditory, etc) creates a local sensory (ie, visual, auditory, etc) navigation map in the Navigation Module
2. Objects are segmented in these maps into multiple maps and composite maps


Sensory Inputs Set 1
Sensory Inputs Set n
centralized components
Input Sensory Vectors Shaping Module 1
Input Sensory Vectors Association Module 1
Object Segmentation Gateway Module
Learned Primitives
D
Navigation Module
Instinctive Primitives
D
Causal Memory Module
Autonomic Inputs 1...n
Autonomic Reflex Module 1
Autonomic Module
D
Goal/ Emotion Module
D
Autonomic Outputs 1...n
Sequential/ Error Correcting Module
Output Vector Shaping Module
Output Vector Association Module
Actuator Output

← CCA2

3. Each Input Sensory Vectors Association Module (visual, auditory, etc) creates a local sensory (ie, visual, auditory, etc) navigation map in the Navigation Module

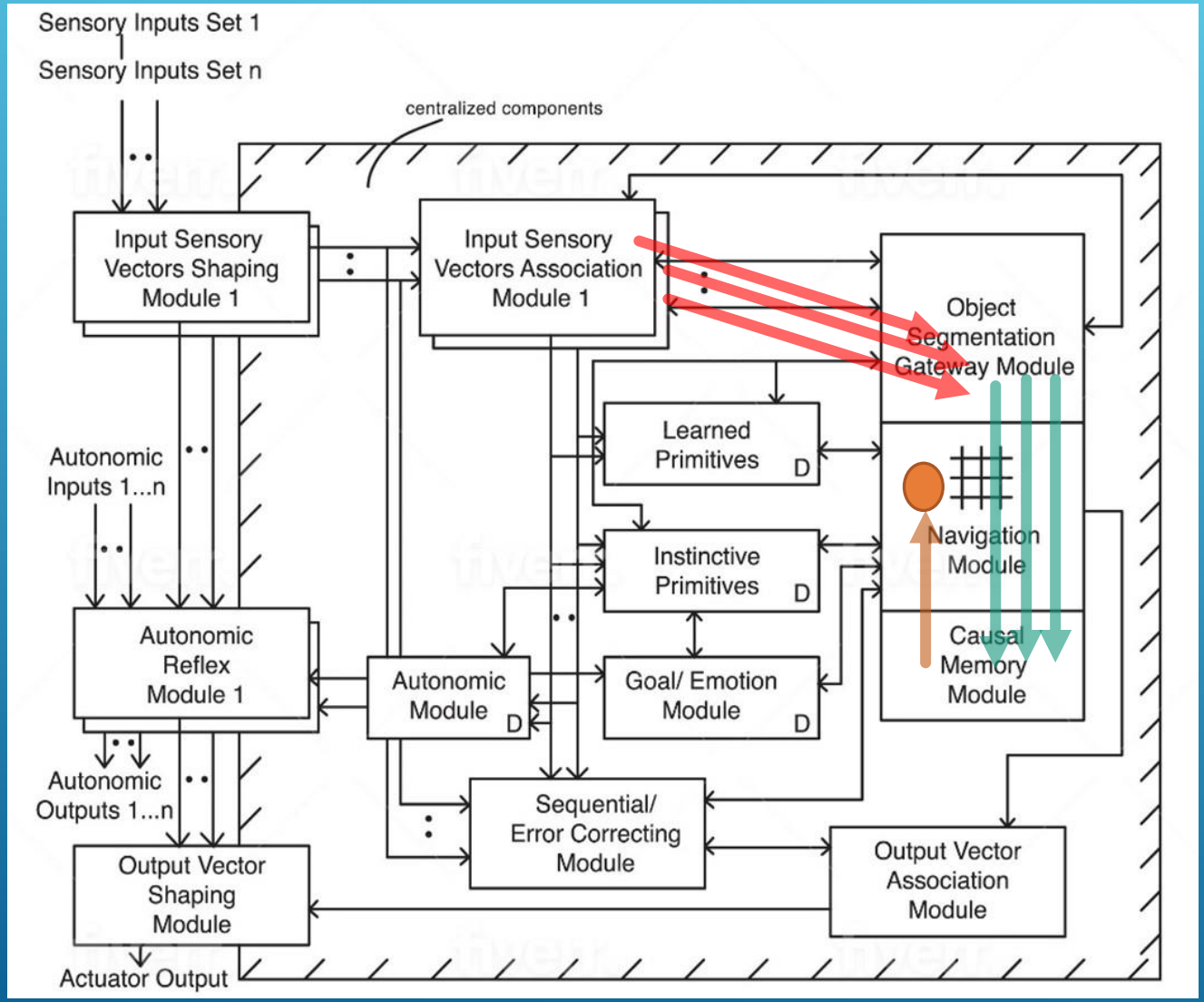
4. Match all the local navigation maps against previous navigation maps stored in the Causal Memory Module


Sensory Inputs Set 1
Sensory Inputs Set n
centralized components
Input Sensory Vectors Shaping Module 1
Input Sensory Vectors Association Module 1
Object Segmentation Gateway Module
Learned Primitives
D
Autonomic Inputs 1...n
Navigation Module
Instinctive Primitives
D
Causal Memory Module
Autonomic Reflex Module 1
Autonomic Module
D
Goal/ Emotion Module
D
Autonomic Outputs 1...n
Sequential/ Error Correcting Module
Output Vector Association Module
Output Vector Shaping Module
Actuator Output

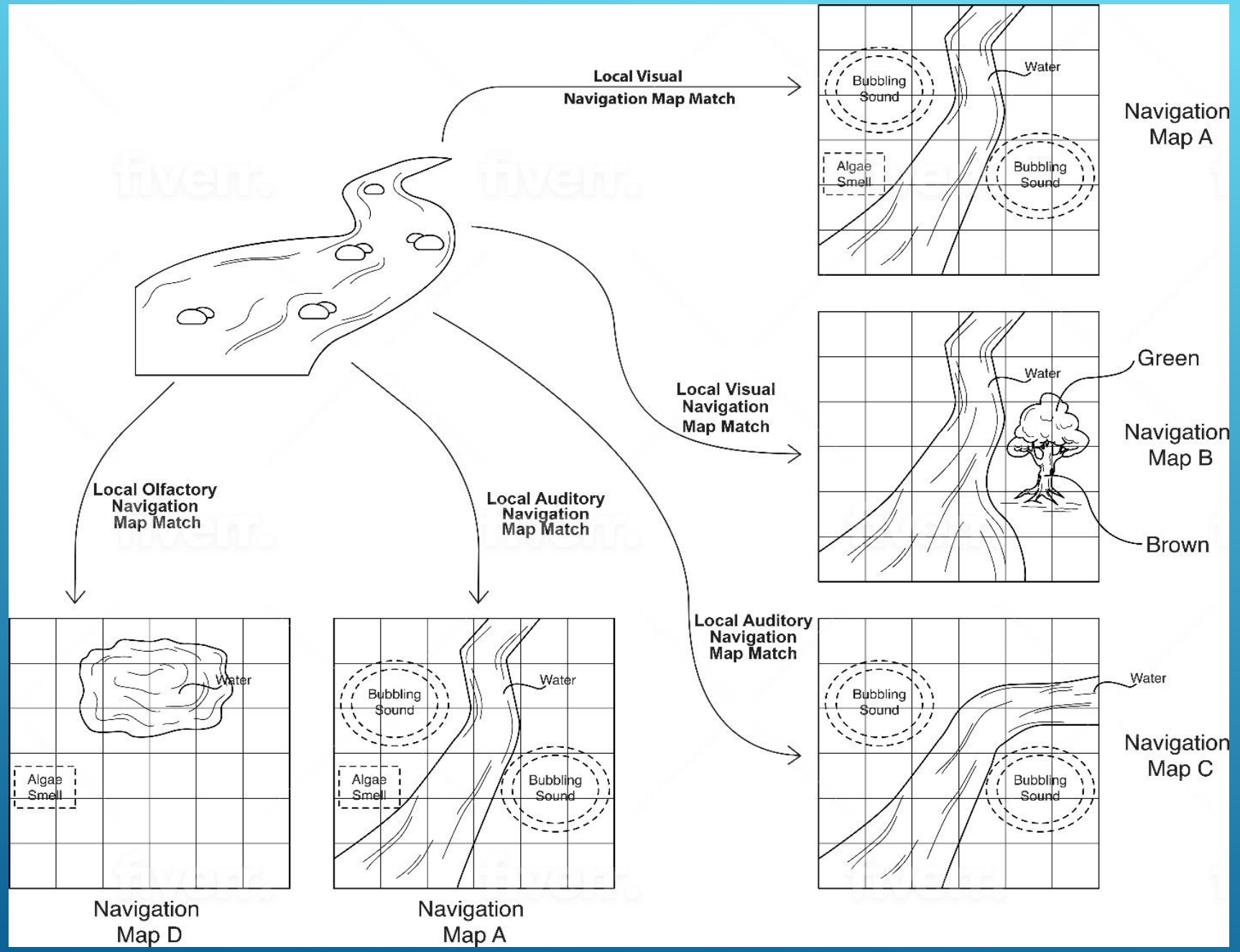
← CCA2

5. Retrieve the best matching navigation map(s)

6. Actually retrieve best maps and then settle on one best multi-sensory navigation map – this is CCA2's perception of the moment


Sensory Inputs Set 1
Sensory Inputs Set n
centralized components
Input Sensory Vectors Shaping Module 1
Input Sensory Vectors Association Module 1
Object Segmentation Gateway Module
Learned Primitives
D
Autonomic Inputs 1...n
Navigation Module
Instinctive Primitives
D
Autonomic Reflex Module 1
Autonomic Module
D
Goal/ Emotion Module
D
Causal Memory Module
Autonomic Outputs 1...n
Sequential/ Error Correcting Module
Output Vector Shaping Module
Output Vector Association Module
Actuator Output

← CCA2

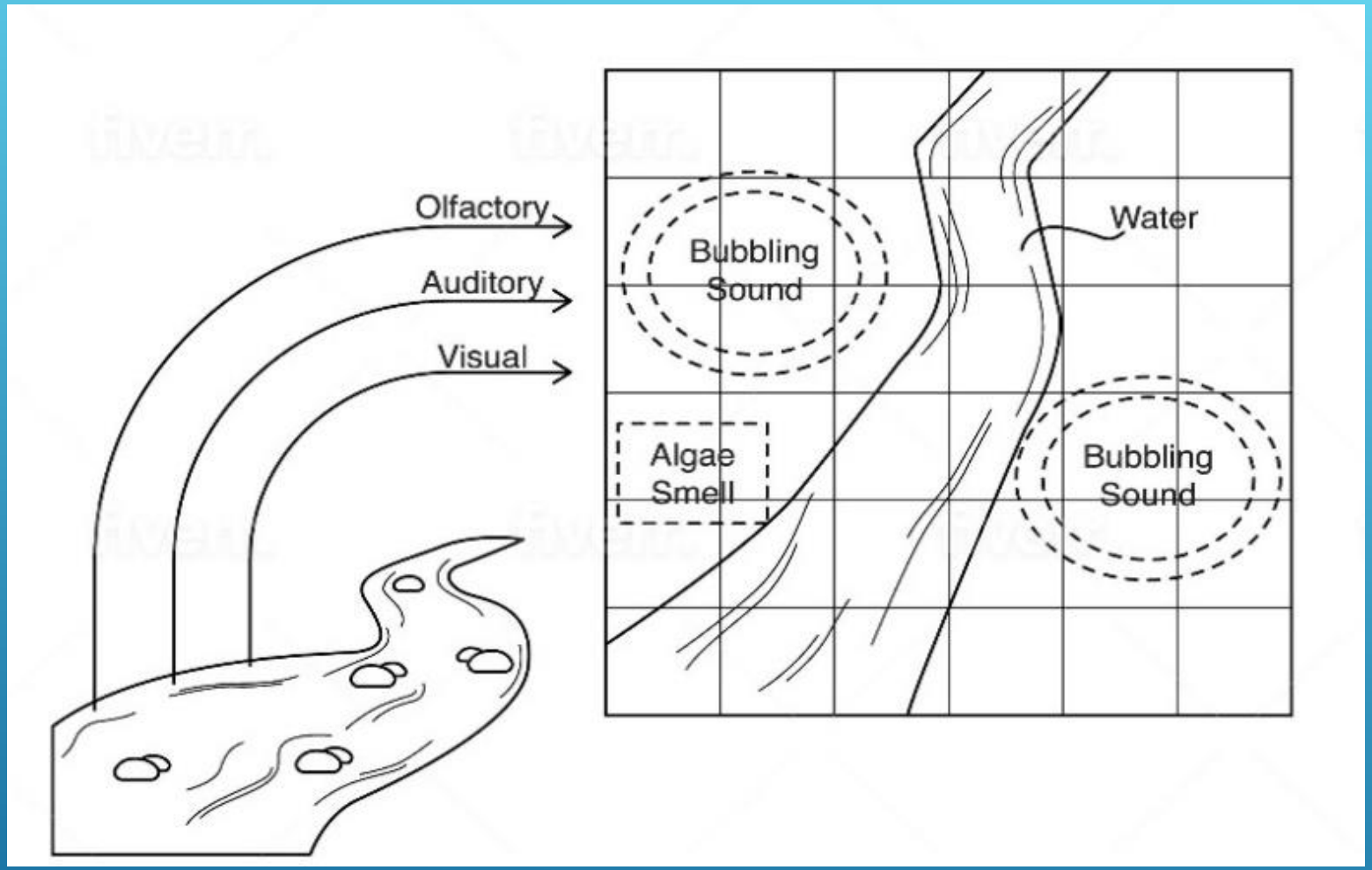

Local Visual
Navigation Map Match
Bubbling
Sound
Water
Algae
Smell
Bubbling
Sound
Navigation
Map A
Local Visual
Navigation
Map Match
Water
Green
Navigation
Map B
Brown
Local Olfactory
Navigation
Map Match
Local Auditory
Navigation
Map Match
Local Auditory
Navigation
Map Match
Water
Algae
Smell
Navigation
Map D
Bubbling
Sound
Water
Algae
Smell
Bubbling
Sound
Navigation
Map A
Bubbling
Sound
Water
Bubbling
Sound
Navigation
Map C

7. Updated best navigation map with current sensory inputs

8. OR if too many updates to make, then make a copy of it and make a new navigation map

9. Updated (or new) multisensory navigation map will be stored in Causal Memory Module

10. Do current operations on this updated multi-sensory navigation map


Olfactory
Auditory
Visual
Bubbling Sound
Water
Algae Smell
Bubbling Sound

→ This binding of sensory inputs to a navigation navigation map solves many of the problems of allowing the CCA1 (which is now the CCA2) to handle larger non-toy problems

Solution to the Binding Problem (Feldman, 2013):

1. General coordination of objects and activities
2. The subjective unity of perception
3. Visual Feature-Binding
4. Variable Binding such as the binding of words in a sentence that allow reasoning

1. Sub-problem: General coordination of objects and activities

- Use of navigation maps as a basic data element
- Instinctive Primitives and Learned Primitives are applied against objects on the current navigation map
- As such, a coordination of objects and activities occurs

2. Sub-problem: The subjective unity of perception

- Best match navigation map represents the CCA2's perception of reality of the sensory scene in front of it
- Current best match navigation map will be updated with current input sensory information, and represents CCA2's perception of the world
- There is a subjective unity perception

3. Sub-problem: Visual Feature-Binding

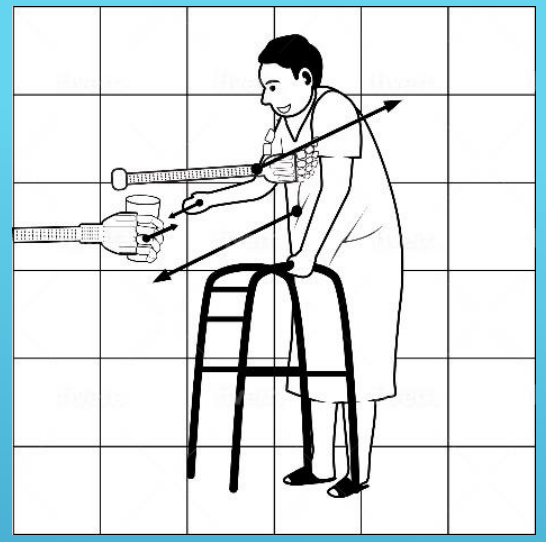
- Spatially mapping visual features onto a spatial navigation map solves this binding sub-problem
- No longer require a binding language; rather, binding occurs in the Vectors Association module and the Navigation Module

4. Sub-problem: Variable Binding such as the binding of words in a sentence that allow reasoning

- Not considered in detail in this paper, but such binding and actually language seems to emerge from the architecture.
- Verbs and nouns provide explanations to the user
- Explanations generated via saved navigation maps

arms because the map is from its point of view.)

Explainability yields language

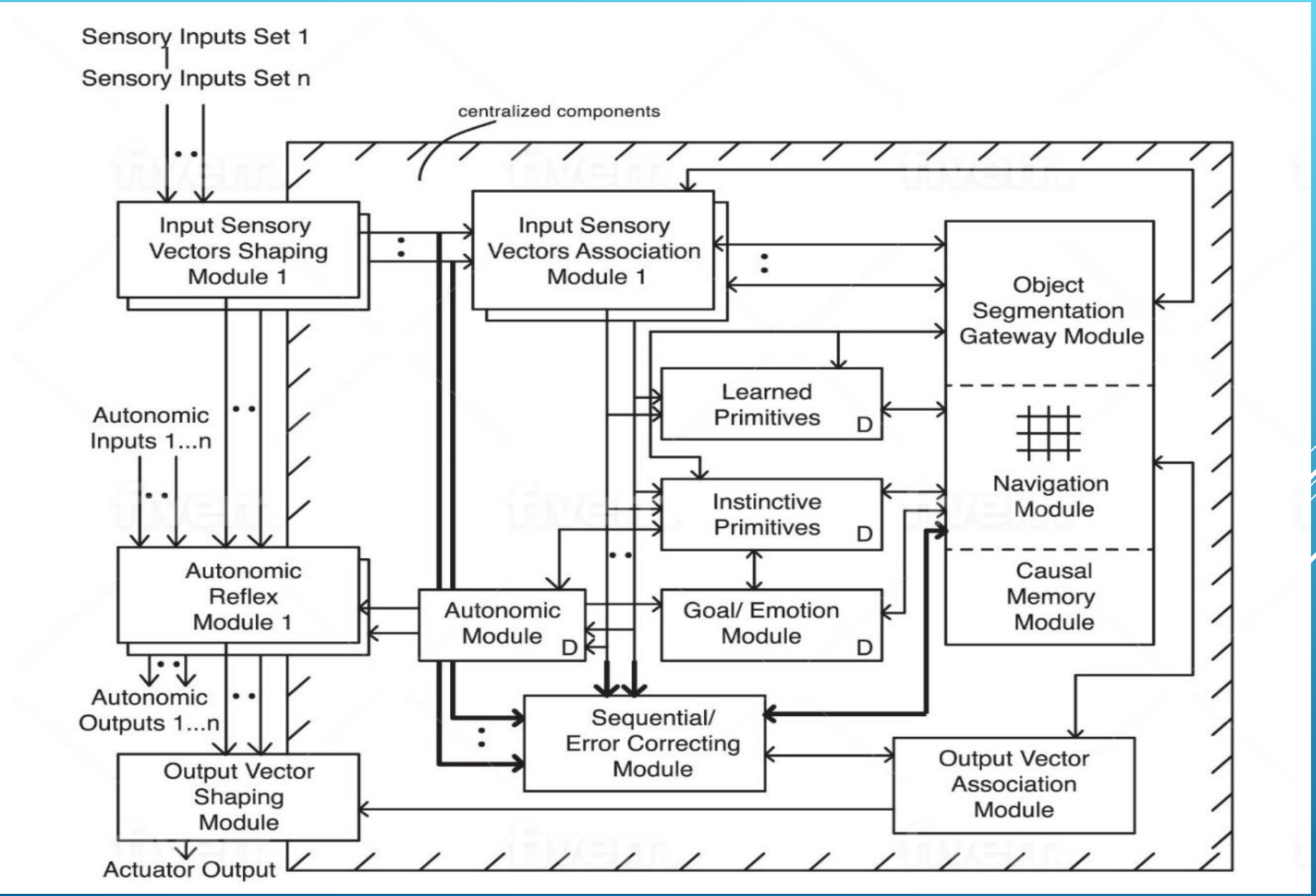


```
patient ask glass
cca3 hold right hand glass
cca3 move +45 degrees right hand glass
patient move -45 degrees right hand
patient move -45 degrees body
patient move body to ground
cca3 no move body to ground
cca3 move +45 degrees left hand
```

but need to bind both Space and Time….

- Most definition of the ‘Binding Problem’ do not take time into account, ie, binding changes
- However, CCA2 shows changes in sensory inputs with time, that *must* bind time also
- **CCA3** – bind space *and* time

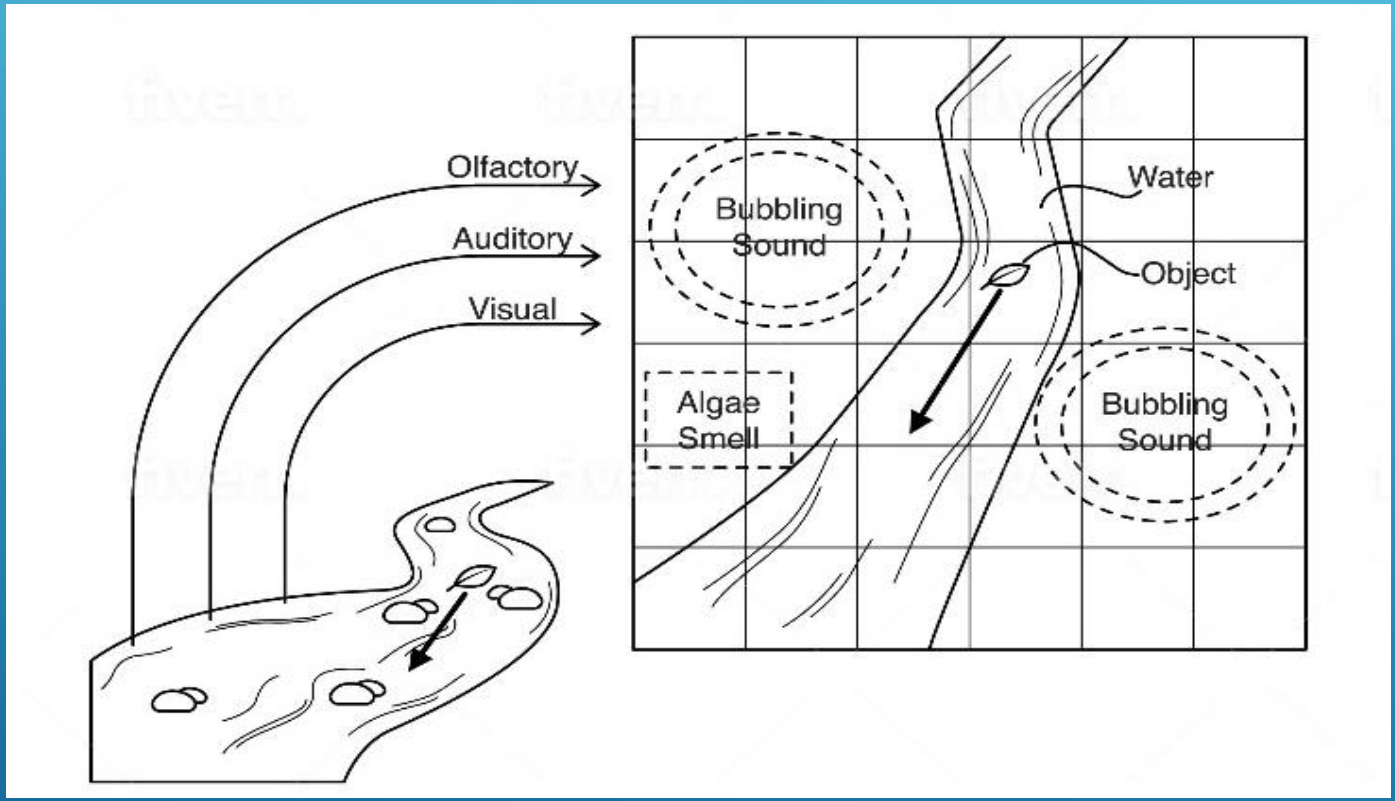
CCA3


Sensory Inputs Set 1
Sensory Inputs Set n
centralized components
Input Sensory Vectors Shaping Module 1
Input Sensory Vectors Association Module 1
Object Segmentation Gateway Module
Learned Primitives
D
Navigation Module
Instinctive Primitives
D
Causal Memory Module
Autonomic Inputs 1...n
Autonomic Reflex Module 1
Autonomic Module
D
Goal/ Emotion Module
D
Autonomic Outputs 1...n
Sequential/ Error Correcting Module
Output Vector Shaping Module
Output Vector Association Module
Actuator Output

CCA3

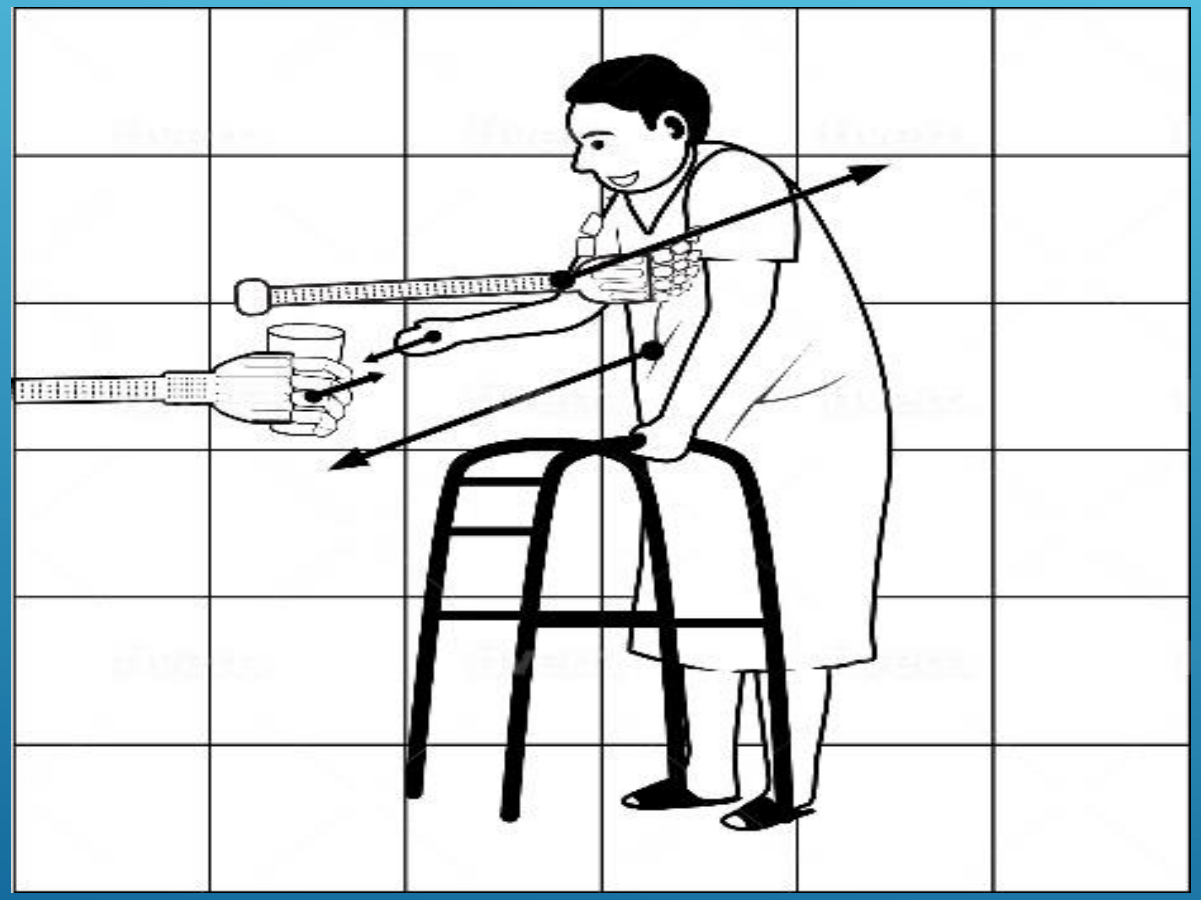
- Motion prediction vectors Most definition of the ‘Binding Problem’ do not take time into account, ie, binding changes
- However, CCA2 shows changes in sensory inputs with time, that *must* bind time also
- CCA3 – bind space *and* time

- Generate motion prediction vectors
- Incorporate much like the other objects onto a navigation map



There is now a moving object (a leaf) on the river. Its motion is represented by a motion prediction vector.

arms because the map is from its point of view.)



→Desirable Properties of CCA3:

- Seems more able to go beyond toy problems
- Pre-Causal Behavior
- Fully Causal Behavior – closing of neurosymbolic gap
- Supports Schneider's psychosis hypothesis
- Solution to the Schizophrenia Paradox
- Analogies emerge automatically
- Explainability emerges automatically
- Lifelong ('continual') learning
- Abilities readily generalize to new and novel environments

hschneidermd@alum.mit.edu