CONTEXTUAL LEARNING FOR MULTITASKING AND CAUSATION LABELING

Thomas E. Portegys, portegys@gmail.com , ORCID 0000-0003-0087-6363

Dialectek, DeKalb, Illinois USA

ABSTRACT

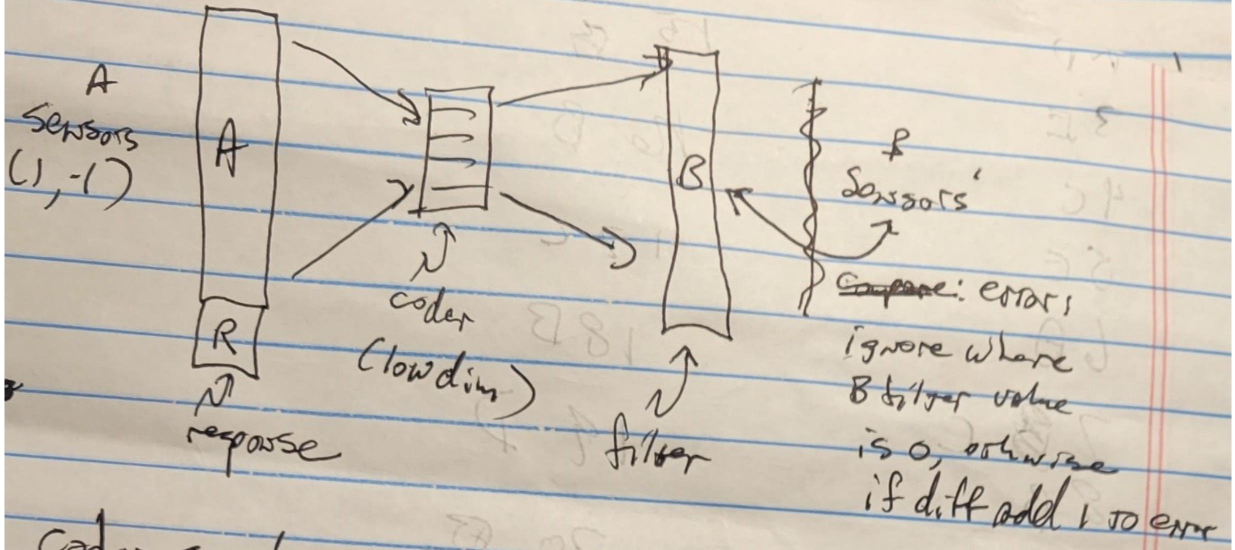
This project is an effort to combine the Morphognosis and Mona neural network models into a comprehensive model for learning and behavior called Mandala. Mona features a contextual causation learning with goal-directed motivation. Morphognosis features contextual multilayer perceptron (MLP) learning. Mandala achieves this by externally accumulating tiers of temporal information that are fed into an MLP at each time step. Natural environments abound in event streams that require multitasking. Mandala affords multitasking as it is robust in the presence of intervening events representing overlaid causation streams, a capability that conventional recurrent artificial neural networks (RNNs) struggle with. In addition, externally accumulating temporal information discretely labels hierarchical cause-and-effect relationships that can be used for augmented processing. In the case of Mona, channeling motivation through the network for the purpose of goal-seeking requires this feature.

Keywords: Mona, Morphognosis, causation learning, multitasking, artificial neural network, machine learning.

Notes:

Note: sparse example: A: (3,5) with two dimensions after dimension reduction from input, B: (1,4). Plot points into a 10x10 space. Repeat process going up each level. So a dimension value pivots to a range of dimension in the higher space (similar to DB pivot).

close



Coder stored as A→B in Mona network, which uses establment + motivation

eval of A→B is f(proximity of coder to prototype, error of B)

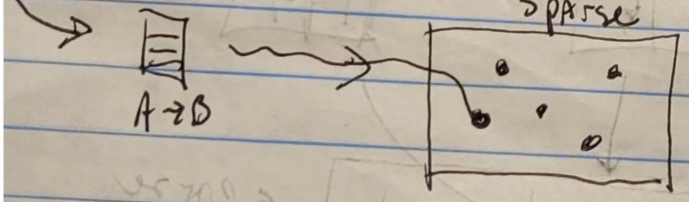
using clustering w/relative

closer weighs error more)

↗ is worse eval

Coder at lower level up sampled using quantized vector va

Input to entire NN is entire stack of ~~B~~ levels



Train from env + let natural causations arise

opefully delays in changes at higher levels results in larger time spanning causations

How to achieve variable causation latencies at levels. Causations in higher levels will have greater time spans than lower levels. One method for this would be to vary the proximity parameter for the code centroid spaces. Centroids that can be close together will capture small changes; centroid that are farther apart will each be mapped to by a wider variation of codes and thus require more changes to map to a different centroid.

Addressing the A->B, X->Y co-occurrence event, in this case the code may lie between the A->B and the X->Y centroids. So why not process both of them? Neither will show up as strongly as if they were the sole causation event, but they should prove significant also.

2/13/2022

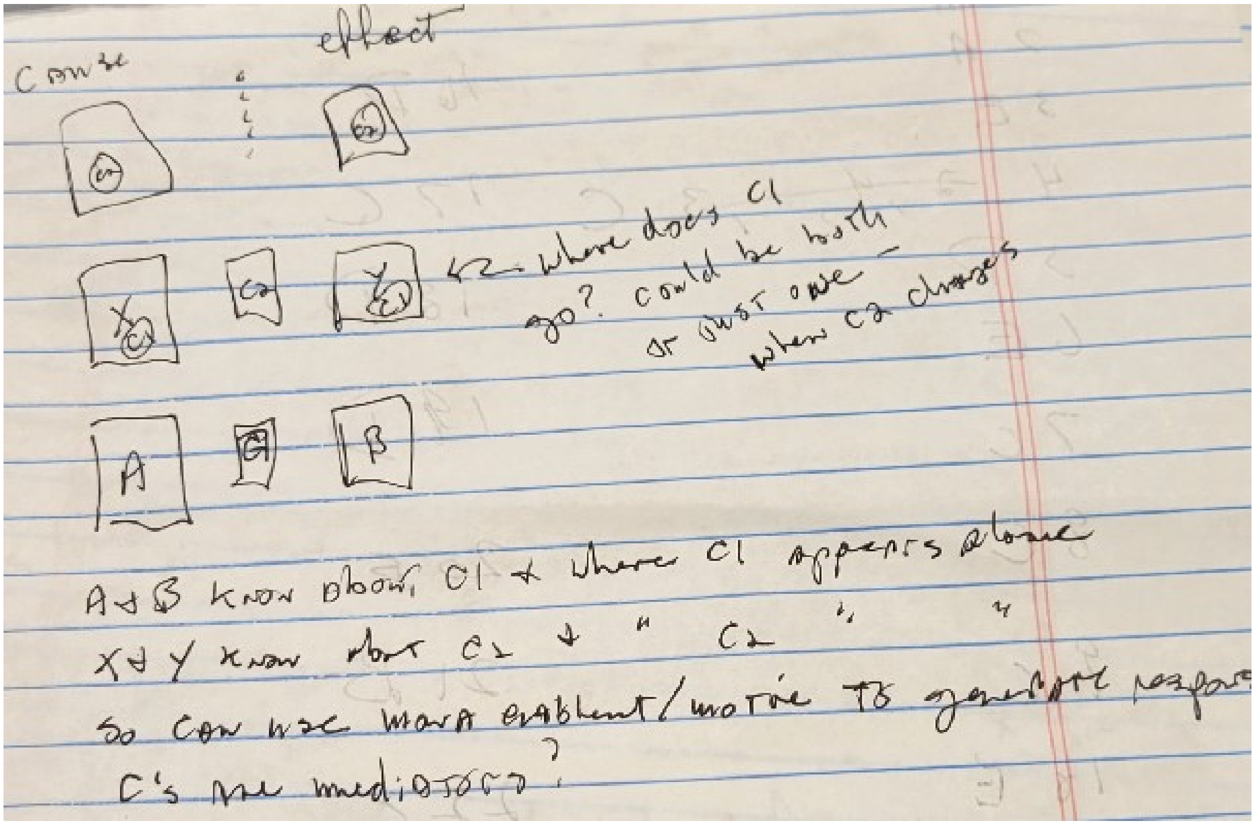
Let there be two computations:

Sensors at all levels -> response (not including need, as this is a model-based system as opposed to a model-free system aka Q-learning).

Per level + response -> effect (this will create the code to be incorporated at higher level)

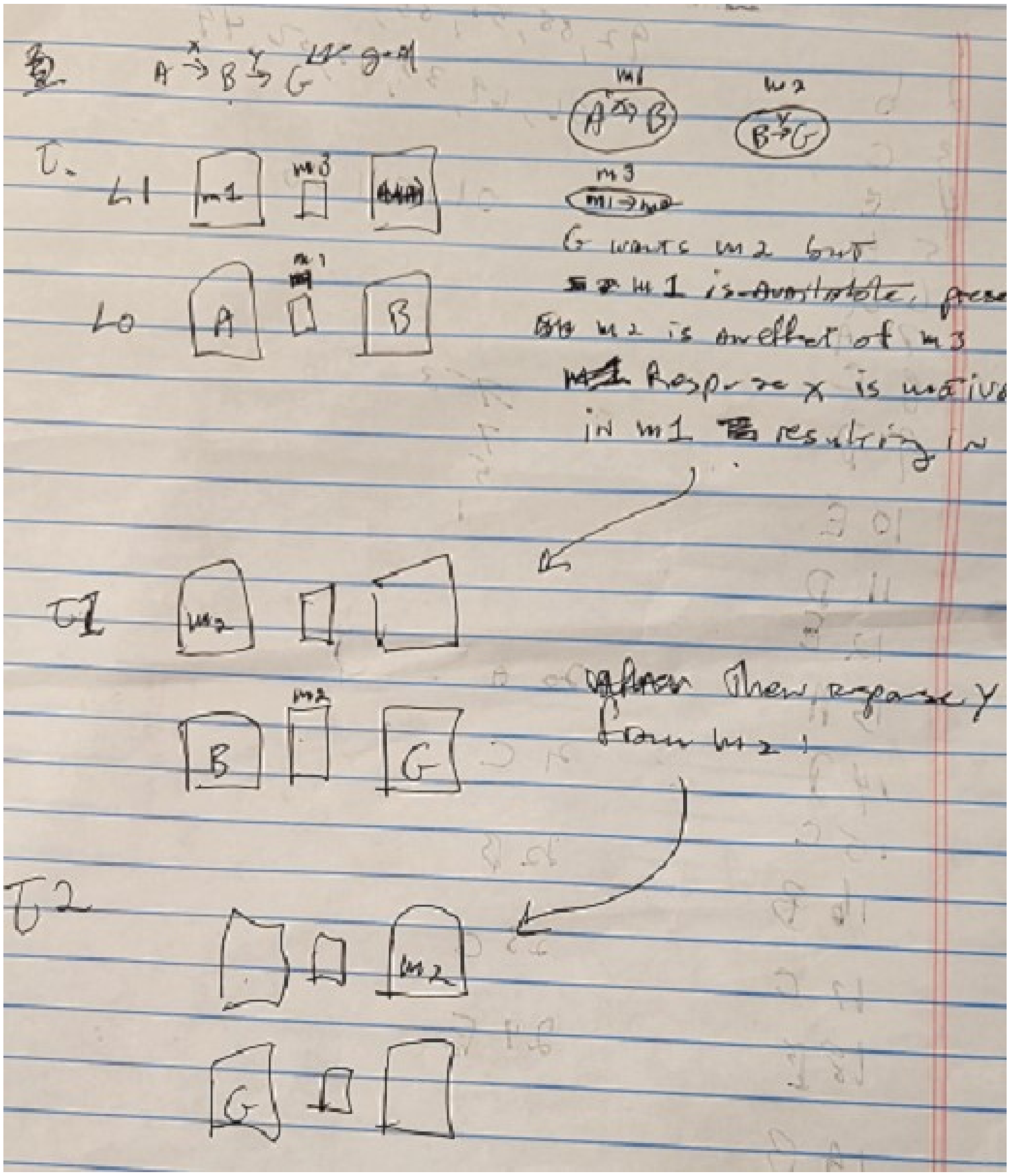
Can this scheme dispense with the Mona network and centroids? (no)

2/14/2022



← where does C1 go? could be both or just one – when C2 changes

A & B know about C1 & where C1 appears alone
X & Y know about C2 & " C2 " "
So can use Mona enablement/motive to generate respons
C's are mediators?



G wants m2 but
m1 ~~is available~~, prese
m2 is an effect of m3
Response x is activ
in m1 resulting in

then response y
from m2 !

Ref:

Exploratory State Representation Learning

Astrid Merckling*, Nicolas Perrin-Gilbert, Alex Coninx and Stéphane Doncieux

Model-based scheme to explore and map env followed by RL on learned space.

Note: look up Reber learning comparison between humans and machines. Notes how humans appear to learn it piecemeal like Mandala.

Instead of using centroids, might use an RBF network to classify against prototype vectors.

Instead of generating random grammars, generate them according to rules:

1. Breadth: e.g. A, AB, ABC
2. Depth: A=A1 A2 A3
3. Repetition: AA, ABAB, ABCABC

At the edge of the tree, terminals are produced.

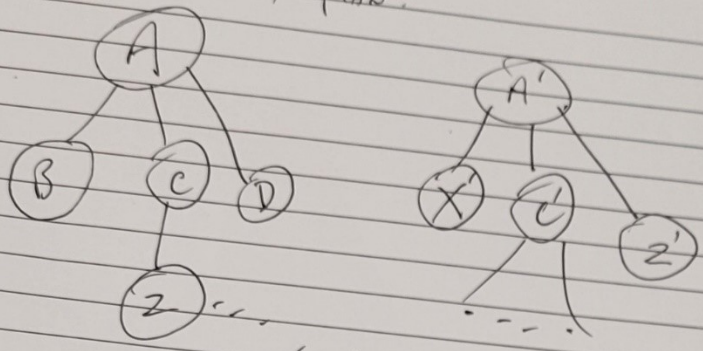
This will provide more reliable performance measurements.

Mandala prototyping with grammar. Instead of encoding causations, sum the occurrences, so if a and b occur in the lower level in the level's time duration, the one-hot representations of a and b are summed.

What sort of statistical analysis can be done on the mutual info/Baysian probs of the grammar that help explain performances?

Spruce up TDNN as a prototype test. Try testing entire length, accumulated values, copying values into multiple layers, and subsumption. These things helped for the maze project.

Modularity plan:



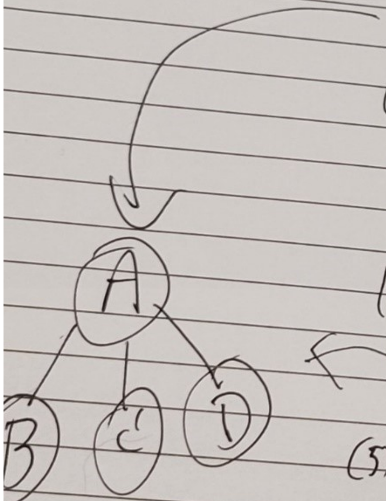
(1) learn both independently

(2) swap ~~all~~ corresponding nodes from A' to A e.g. C' → C

(3) ~~Main~~ challenge is ~~to~~ how well can navigate using new node even though modular has been learned

(4) next: how quickly can learn

← this



(5) point being how modular learning works

Adapt "horizon" w/ normalization

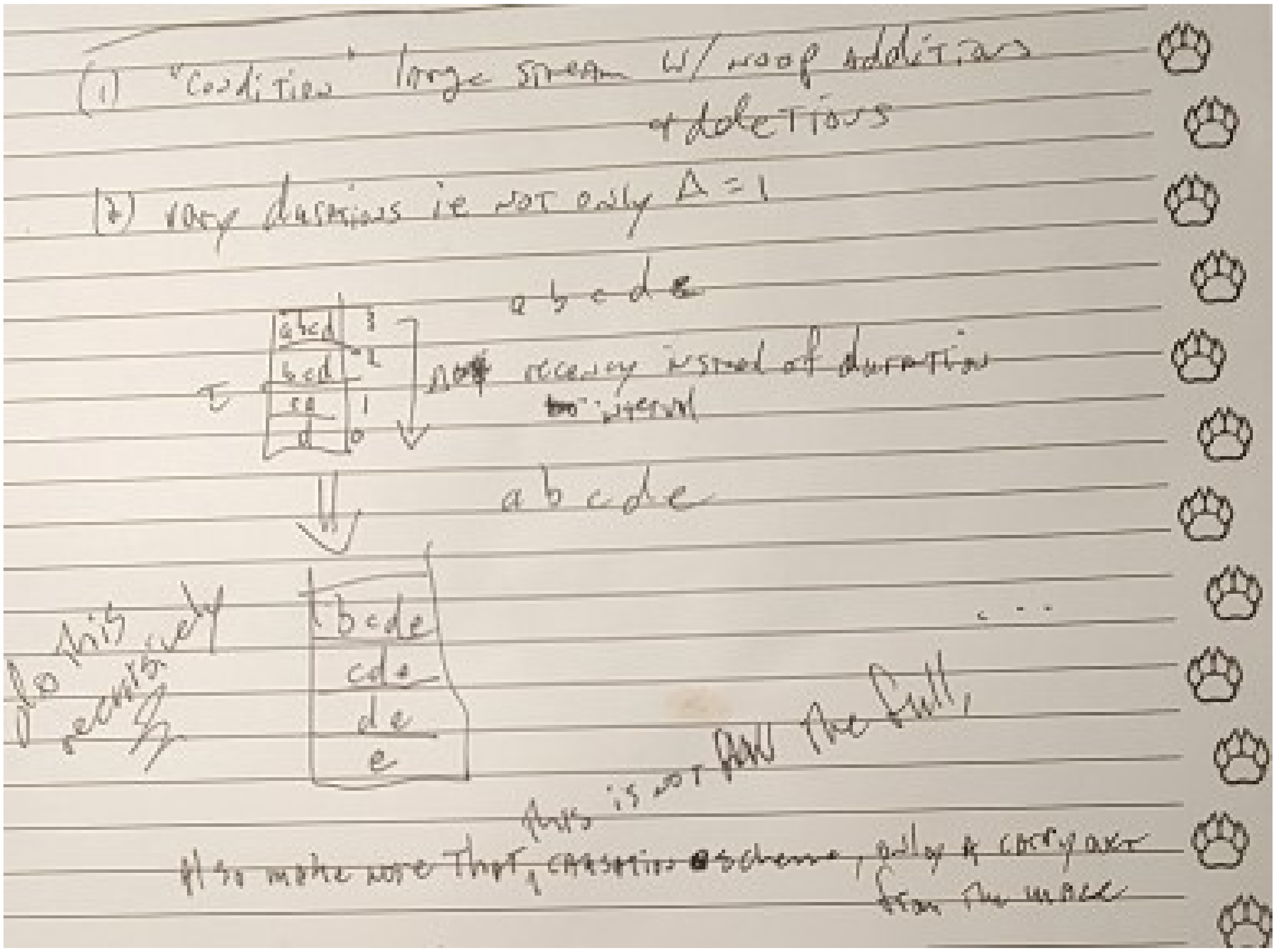
so if see a, b, c:

abc
~~a~~bc
c

A simplification:

Start -> Goal. In between are a number of subgoals in sequence. These can be learned separately/modularly. String some of them together and learn the whole stream. Then life intervenes and adds and subtracts subtasks from the stream. Test and measure this.

Let the Start -> Goal be one monolithic learning stream. Then the task is to complete it given deletion and insertions of other modules.



(1) "Condition" large stream w/ random additions + deletions

(2) vary durations ie not only Δ=1

a b c d e

abcd 3
bcd 2
cd 1
d 0

~~now~~ recency instead of duration
~~too~~ interval

a b c d e

bcde
cde
de
e

do this recursively

...

Also make sure that (this is not all the full,) a cascading schema, only a carry over from the [illegible]

Hierarchical Reinforcement Learning:
Assignment of Behaviours to Subpolicies by Self-Organization
Wilco Moerman
Cognitive Artificial Intelligence, Utrecht University
wilcom@phil.uu.nl, wilco.moerman@gmail.com

https://www.geeksforgeeks.org/deep-learning/sparse-autoencoders-in-deep-learning/

Remember to check EverNote and Melendey for refs.

Mona:

A pair of cooperating nest-building and foraging birds.

See: http://tom.portegys.com/research.html#nestingbirds

Morphognosis:

Honey bees forage for flower nectar cooperatively.

See: http://tom.portegys.com/research.html#honey_bees