Autonomous Machine Learning

An article on a path to true artificial intelligence

Sharad Nataraj

CMPE 297 - Advanced topics in deep learning

Agenda

- Introduction to Autonomous Machine learning
- Why I choose this topic?
- What is lacking with existing technology
- The proposed architecture
- JEPA
- Closing remarks

Introduction

- Our current ML is more system 1 than system 2
- System 1 vs System 2
  - System 1 is spontaneous, fast and superficial
  - System 2 is deliberate, logical and perceptive
- Humans do not always act based on policies.
  - We are capable of using knowledge during previously unseen scenarios.
  - Humans are capable of learning new tasks faster
  - Humans follow a hierarchical approach to problem solving.
  - Humans learn continuously
- True artificial intelligence is when a machine can learn and act like humans
- Currently machines are mostly capable of performing well only with explicit training.

Why I chose this topic

- Era of advanced machine learning.
- Possibility of achieving artificial general intelligence.
- Lacking in the existing technology of automated systems
- Relevance to the current industry
- Currently the state-of-the-art

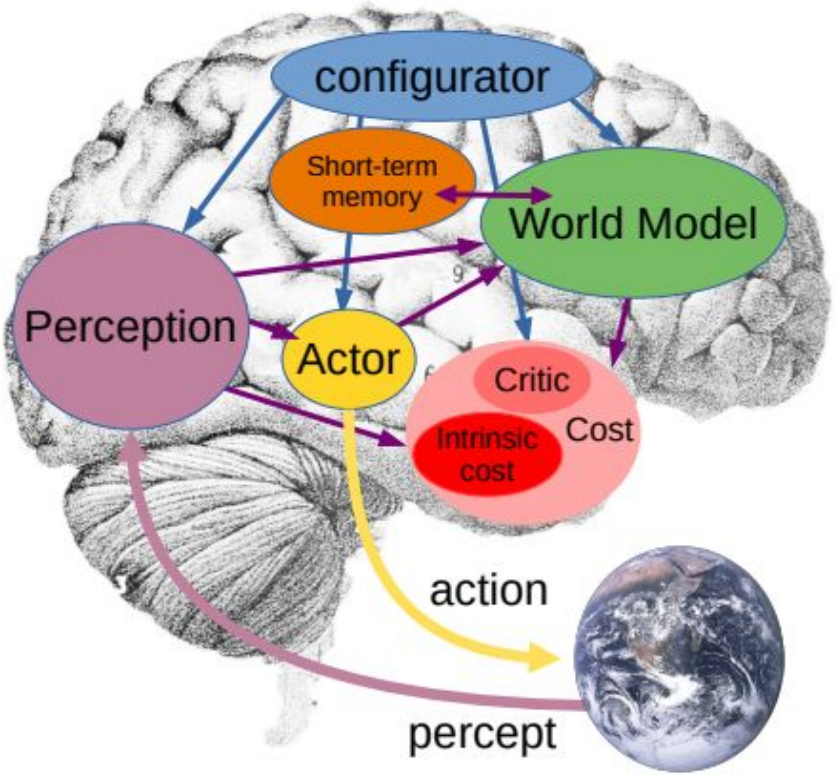
The story so far..

- **Multitask models** - we train one model to perform multiple tasks. By doing this we allow the model to learn the representations rather than the data itself so each task supports the learning of other tasks.
- **Multimodal models** - we build models to train on different kinds of data. This produces a rich representation of the data allowing for a better understanding of the task.
- **Simulation based training** - Used most commonly in autonomous systems where getting examples of every possible scenario is impractical, we teach the machine what to do by doing it.
- **Meta learning** - we need the model to learn progressively and continuously. So we teach the model how to learn.

The story so far..

- With such advances we have not yet achieved true artificial intelligence.
- We need a model that is capable of
  - Hierarchical learning
  - Learning mostly by observing
  - Learning based on rewards
  - Learns new tasks faster
  - Learns common sense rather than examples.
- So why don’t we use all the current models together..?
  - Turns out that's the whole idea of the paper!

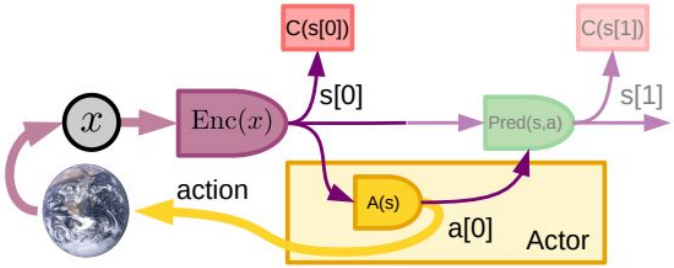
Proposed Architecture

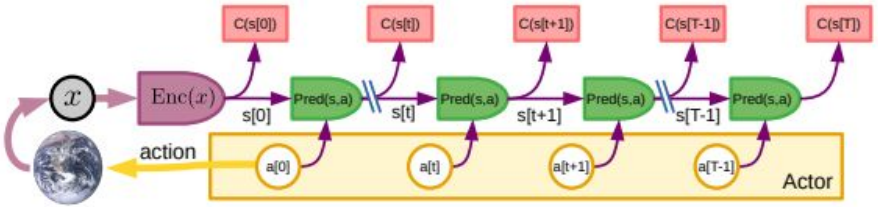


Perception is the key to intelligence

- The model is capable of perceiving the world in two modes. System 1 and system 2 or (mode 1 and mode 2)
- Mode 1 is policy based, fast and spontaneous.
  - Actor receives the current state of the world.
  - Produces a list of actions based on existing policy
  - Agent acts using these actions.
- Mode 1 may run into issues during unseen scenarios.
- Mode 2 is perceptive, cost based and optimized.
  - Actor receives the current state of the world. Cost module produces immediate energy.
  - Produces a list of actions.
  - Evaluated by critic and cost module using simulations from world model.
  - Actor optimizes actions based on the cost.
  - Executes and stores experience in buffer.

Mode 1 vs Mode 2 perception





World Models

- The core aspect of this architecture is the world model.
- It is a complex system on its own.
- The goal is the learn the world.
  - It predicts the future state of the world based on the current state and actions.
- The world is not always predictable
  - The world model generates of list of plausible scenarios.
  - It also pieces together missing information(generative model) to fill gaps in information.

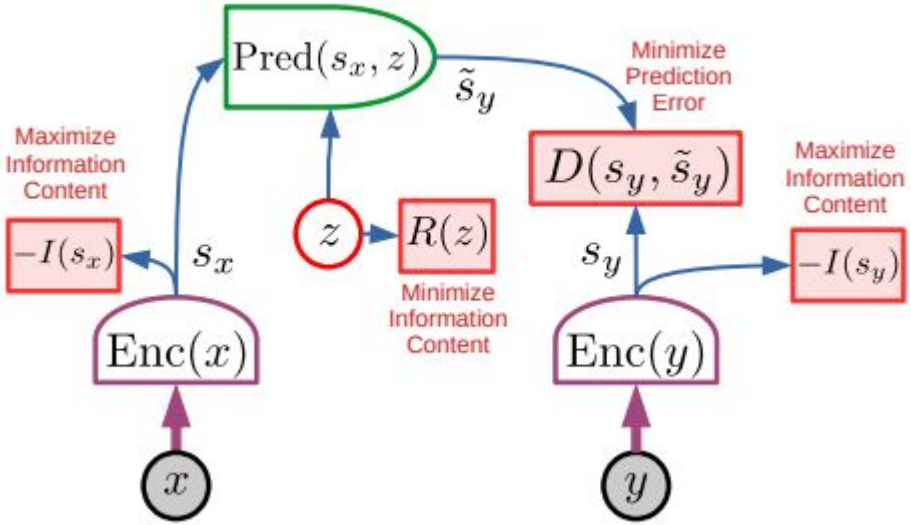
World models

- In the context of our architecture
  - The world model fills in gaps in information that were not acquired by perception.
  - The perception is may be an intrinsic module in the model or like in our case a 3rd party component.
- The role of the world model is configured by the configurator
- The role of the world model is comparable to that of a trainable world simulator
- Requirements to build a world model
  - Should learn the world at multiple abstraction levels
  - Should be able to handle uncertainty
  - Should learn a diverse set of data with less examples
  - Should generate scenarios without explicit training

Where does JEPA contribute to this architecture

- Turns out JEPA is just a part of the whole system and not the system itself.
- The key contributor to the entire system is the world model.
- An ideal world model should
  - Have diverse set of example it is trained on. Quality rather than quantity.
  - Handle uncertainty as the real world in chaotic.
  - Be able to plan and make predictions on various time scales and abstraction levels
- Currently EBM(Energy based models) are used.
  - EBM do not work well with high dimensional features like encodings of real world observations.
  - Prone to information collapse.

Cue the JEPA


$\text{Pred}(s_x, z)$
$\tilde{s}_y$
Minimize Prediction Error
$D(s_y, \tilde{s}_y)$
Maximize Information Content
$-I(s_x)$
s_x
z
$R(z)$
Minimize Information Content
s_y
Maximize Information Content
$-I(s_y)$
$\text{Enc}(x)$
$\text{Enc}(y)$
x
y

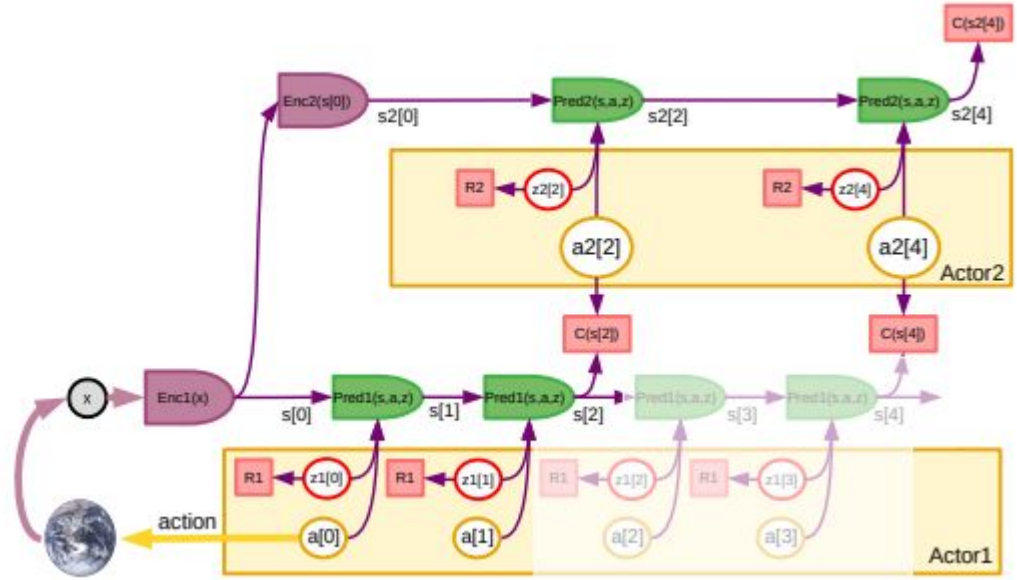
JEPA

- Embedding based architecture - not a model
- The goal is to create a rich representation of x and y using embeddings
- Use a latent variable z to estimate y from x.
- We need not predict y from x exactly.
- We need to understand the relationship between y and x through z.
- This helps us create rich representations of the data without learning the data itself.
- We use this representation to make predictions.

JEPA

- JEPA is key to building world models.
- To obtain the best results the following principles are followed when designing a JEPA based model.
    - Maximize the information of the encoding (sx) obtained from x
    - Maximize the information of sy obtained from y
    - Make sy easily predictable from sx
    - Minimize the information contained in the latent variable z

H-JEPA


C(s2[4])
Enc2(s[0])
s2[0]
Pred2(s,a,z)
s2[2]
Pred2(s,a,z)
s2[4]
R2
z2[2]
R2
z2[4]
a2[2]
a2[4]
Actor2
C(s[2])
C(s[4])
x
Enc1(x)
s[0]
Pred1(s,a,z)
s[1]
Pred1(s,a,z)
s[2]
Pred1(s,a,z)
s[3]
Pred1(s,a,z)
s[4]
R1
z1[0]
R1
z1[1]
R1
z1[2]
R1
z1[3]
action
a[0]
a[1]
a[2]
a[3]
Actor1

H-JEPA

- Provides the ability to use different levels of abstraction.
- Is able to breakdown the complex tasks into smaller subtasks each associated with a JEPA
- Higher level JEPAs learn from the embeddings produces by lower level JEPAs
- Has the potential to make short term and long term predictions.
- Perfect for a world model!

Closing Statements

- We have seen the proposed architecture to achieve system 2.
- Identified the key components for such a system.
- Identified the importance of world models.
- Identified JEPA as the best architecture for world models
- This is a paradigm shift in machine learning.
    - Systems capable of true artificial intelligence.
    - Leads to fully autonomous systems which are practical and efficient.

Thank You!!

Please feel free to contact me for any questions ,discussions or feedback.

sharad.nataraj@sjsu.edu

Medium article