

Intuitive Agency

Author: Dustin Fast, CSCE A470, Fall 2018.

Inspiration

To explain the integration of information, we need only exhibit mechanisms by which information is brought together and exploited by later processes.

-David Chalmers, *Facing Up to The Problem of Consciousness*

Objective

General intelligence of the type we possess exists exclusively, for now, in the domain of the conscious human. Therefore, an understanding of the mechanisms leading to our conscious experience may be required if an artificial general intelligence is to one day be realized.

One defining aspect of human intelligence is our ability to subconsciously form new connections between abstract concepts, which then seem to "bubble up" to the forefront of our attention. This phenomenon, commonly called intuition, is responsible not only for our most startling "Aha!" moments, but also for the seemingly arbitrary changes in our awareness of, say, the ticking of a clock on the wall.

That we experience these changes of awareness unwillingly provides powerful clues to the underlying mechanisms of intuition and consciousness. With that in mind, this project aims to develop an ensemble learning system capable of rudimentary intuition by modeling an agent who's "attention" switches contexts according to its "intuition". If successful, the agent may then be used in a larger network of such agents, thus bootstrapping an increasingly advanced intuition.

The agent will be developed in Python using readily available machine learning and evolutionary programming libraries.

The Agent

The agent is composed of three layers, labeled *Conceptual*, *Intuitive*, and *Attentive*. They are described below and given by [Diagram 1](#) (attached). Data is mostly feed-forward, with recurrent feedback signaling the agent's current state and contextual fitness.

The Conceptual, Intuitive, and Attentive Layers

Conceptual Layer

The conceptual layer represents our existing knowledge base of abstract concepts. It consists of a set \mathcal{A} of artificial neural networks (ANNs) with the following properties -

Each ANN at this level has a set X of input nodes consisting of k feed-forward sensory input nodes and m feedback input nodes, defined as:

X	Channel	Description
x_0 to x_{k-1}	Feed-forward	Environmental sensory input, identical for every a in A
x_k to x_{k+m-1}	Feedback	Each x_k to x_n is mapped from a corresponding attentive-layer input node. This represents feedback of the agent's current context.

Each ANN is pre-trained (offline) for a different class of objects (ex: a_0 classifies digits, a_1 classifies letters, etc). During this pre-training, each x_k to x_{m-1} input-value should be randomized to simulate environmental noise.

Each ANN's output nodes are provided to the intuitive layer as input.

Intuitive Layer

The intuitive layer is a set of data pipes, one for each conceptual-layer ANN, connecting the conceptual layer to the attentive layer. On each state change, each pipe is weighted according to some fitness function that evolves in an online manner according to some genetic algorithm, who's fitness is received as feedback from the attentive layer.

These weights are subsequently used as a bias (possibly binary, logarithmic, etc) by the attentive layer. In this way, the agent's "intuition" learns how to best allocate the agent's "attention" while allowing "mistakes" to enter its awareness. These mistakes represent possible new connections between the conceptual layer's existing abstract concepts.

Attentive Layer

The attentive layer represents the context of our attention at any given moment. Context may be defined artificially as the agent's level of awareness of the symbolic concepts present in its current environment. It is at this level where fitness of the current context is determined, and that determination is then signaled back to the intuitive-layer's genetic algorithm.

This layer will be implemented as a singular ANN having a set M of input nodes, where $size(M) = size(A) * size(X)$. Each input node m_x is biased according to the current state's corresponding pipe weights as given by the intuitive layer. Training and validation of this ANN is described immediately below.

Proof of Concept

Upon completion of development, proof of concept will be attempted in the following way: The attentive-layer ANN will be trained on a set of known conceptual links between some set of concepts known by the conceptual-layer. The entire agent will then be subjected to the validation set. In this way, a successful intuitive agent will be demonstrated by the attentive layer's discovery of previously unlearned connections in the validation set.

Conceptual Layer
(Represents pre-learned abstract concepts)

Intuitive Layer
(Represents intuitive abstract connections)

Attentive Layer
(Represents current awareness)

Environment

