International Journal of Modern Nonlinear Theory and Application, 2014, 3, 182-190
Published Online September 2014 in SciRes. http://www.scirp.org/journal/ijmnta
http://dx.doi.org/10.4236/ijmnta.2014.34020

The Computational Theory of Intelligence: Information Entropy

Daniel Kovach

Kovach Technologies, San Jose, CA, USA
Email: kovachtechnologies@gmail.com

Received 12 August 2014; revised 10 September 2014; accepted 18 September 2014

Abstract

This paper presents an information theoretic approach to the concept of intelligence in the computational sense. We introduce a probabilistic framework from which computation alintelligence is shown to be an entropy minimizing process at the local level. Using this new scheme, we develop a simple data driven clustering example and discuss its applications.

Keywords

Machine Learning, Artificial Intelligence, Entropy, Computer Science, Intelligence

1. Introduction

This paper attempts to introduce a computational approach to the study of intelligence that the researcher has accumulated for many years. This approach takes into account data from Psychology, Neurology, Artificial Intelligence, Machine Learning, and Mathematics.

Central to this framework is the fact that the goal of any intelligent agent is to reduce the randomness in its environment in some meaningful ways. Of course, formal definitions in the context of this paper for terms like "intelligence", "environment", and "agent" will follow.

The approach draws from multidisciplinary research and has many applications. We will utilize the construct in discussions at the end of the paper. Other applications will follow in future works. Implementations of this framework can apply to many fields of study including General Artificial Intelligence (GAI), Machine Learning, Optimization, Information Gathering, Clustering, and Big Data, and extend outside of the applied mathematics and computer science realm to even more areas including Sociology, Psychology, and Neurology, and even Philosophy.

1.1. Definitions

One cannot begin a discussion about the philosophy of artificial intelligence without a definition of the word

How to cite this paper: Kovach, D. (2014) The Computational Theory of Intelligence: Information Entropy. *International Journal of Modern Nonlinear Theory and Application*, **3**, 182-190. http://dx.doi.org/10.4236/ijmnta.2014.34020

"intelligence" in the first place. With the panoply of definitions available, it is understandable that there may be some disagreement, but typically each school of thought generally shares a common thread. The following are three different definitions of intelligence from respectable sources:

1) "The aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment."[1].

2) "A process that entails a set of skills of problem solving enabling the individual to resolve genuine problems or difficulties that he or she encounters and, when appropriate, to create an effective product and must also entail the potential for finding or creating problems and thereby providing the foundation for the acquisition of new knowledge."[2].

3) "Goal-directed adaptive behavior." [3].

Vernon's hierarchical model of intelligence from the 1950's [1], and Hawkins' *On Intelligence* from g 2004 [4] are some other great resources on this topic. Consider the following working definition of this paper, with regard to information theory and computation: Computational Intelligence (CI) is an information processing algorithm that

1) Records data or events into some type of store, or memory.

2) Draws from the events recorded in memory, to make assertions, or predictions about future events.

3) Using the disparity between the predicted and events and the new incoming events, the memory structure in step 1 can be updated such that the predictions of step 2 are optimized.

The mapping in 3 is called learning, and is endemic to CI. Any entity that is facilitating the CI process we will refer to as an agent, in particular when the connotation is that the entity is autonomous. The surrounding infrastructure that encapsulates the agent together with the ensuing events is called the environment.

1.2. Structure

The paper is organized as follows. In Section 2 we provide a brief summary of the concept of information entropy as it is used for our purposes. In Section 3, we provide a mathematical framework for intelligence and show discuss its relation to entropy. Section 4 discusses the global ramifications of local entropy minimization. In Section 5 we present a simple application of the framework to data analytics, which is available for free download. Sections 6 and 7 discuss relevant related research, and future work.

2. Entropy

A key concept of information theory is that of entropy, which amounts to the uncertainty in a given random variable, [5]. It is essentially, a measure of unpredictability (among other interpretations). The concept of entropy is a much deeper principal of nature that penetrates to the deepest core of physical reality and is central to physics and cosmological models [6]-[8].

2.1. Mathematical Representation

Although terms like Shannon entropy are pervasive in the field of information theory, it will be insightful to review the formulation in our context. To arrive at the definition of entropy, we must first recall what is meant by information content. The information content of a random variable, $X = \{x_1, \cdots, x_N\}$ denoted $I[X]$, is given by

$$I[X] = \log\left[\frac{1}{\mathbb{P}[X]}\right] = -\log\left[\mathbb{P}[X]\right]. \tag{1}$$

where $\mathbb{P}[X]$ is the probability of X. The entropy of X, denoted $\mathbb{E}[X]$, is then defined as the expectation value of the information content.

$$\mathbb{E}[X] = E\left[I[X]\right] = -E\left[\log\left[\mathbb{P}[X]\right]\right]. \tag{2}$$

Expanding using the definition of the expectation value, we have

$$\mathbb{E}[X] = -\sum_i^N \mathbb{P}[x_i]\log\left[\mathbb{P}[x_i]\right]. \tag{3}$$

2.1.1. Relationship of Shannon Entropy to Thermodynamics

The concept of entropy is deeply rooted at the heart of physical reality. It is a central concept in thermodynamics, governing everything from chemical reactions to engines and refrigerators. The relationship of entropy as it is known in information theory, however, is not mapped so straightforwardly to its use in thermodynamics.

In statistical thermodynamics, the entropy $\boldsymbol{S}$, of a system is given by

$$\boldsymbol{S} = -k_b \sum\nolimits_i^N p_i \ln\left[p_i\right], \tag{4}$$

where p_i denote the probability of each microstate, or configuration of the system, and k_b is the Boltzmann constant which serves to map the value of the summation to physical reality in terms of quantity and units.

The connection between the thermodynamic and information theoretic versions of entropy relate to the information needed to detail the exact state of the system, specifically, the amount of further Shannon information needed to define the microscopic state of the system that remains ambiguous when given its macroscopic definition in terms of the variables of Classical Thermodynamics. The Boltzmann constant serves as a constant of proportionality.

2.1.2. Renyi Entropy

We can extend the logic of the beginning of this section to a more general formulation called the Renyi entropy of order α, where $\alpha \geq 0$ and $\alpha \neq 1$ defined as

$$\mathbb{H}_\alpha\left[X\right] = \frac{1}{1-\alpha} \sum\nolimits_i^N \log\left[\mathbb{P}\left[x_i\right]^\alpha\right] \neq 1. \tag{5}$$

Under this convention we can apply the concept of entropy more generally to extend the utility of the concept to a variety of applications. It is important to note that this formulation approaches 1 in the limit as $\alpha \to 1$. Although the discussions of this paper were inspired by Shannon entropy, we wish to present a much more general definition and a bolder proposition.

3. Intelligence: Definition and Assumptions

$\mathbb{I}: S \to O$. The function $\mathbb{I}$ represents the intelligence process, a member of $\mathcal{I}$, the set of all such functions. It maps input from set S to output in O. First, let

$$\mathbb{I}_t\left[s^i\right] = o_t^i \tag{6}$$

reflect the fact that $\mathbb{I}$ is mapping one element from S to one element in O, each tagged by the identifier $i \in \mathbb{N}$, which is bounded by the cardinality of the input set. The cardinality of these two sets need not match, nor does the mapping between $\mathbb{I}$ need to be bijective, or even surjective. This is an iterative process, as denoted by the index, t. Finally, let O_t represent the collection of o_t^i.

Over time, the mapping should converge to the intended element, $o^i \in O$, as is reflected in notation by

$$\mathbb{I}_t\left[s^i\right] = o^i. \tag{7}$$

Introduce the function

$$\mathbb{L}_t = f\left(O, O_t\right). \tag{8}$$

which in practice is usually some type of error or fitness measuring function. Assuming that $\mathbb{L}_t$ is continuous and differentiable, let the updating mechanism at some particular t for $\mathbb{I}$ be

$$\mathbb{I}_t = \mathbb{I}_{t-1} + \nabla \mathbb{L}_{t-1}. \tag{9}$$

In other words, $\mathbb{I}$ iteratively updates itself with respect to the gradient of some function, $\mathbb{L}$. Additionally, $\mathbb{L}$ must satisfy the following partial differential equation

$$\frac{\partial}{\partial t}\mathbb{L} = \rho(t) d\left(O - O_t\right), \quad \rho \mapsto \mathbb{R}, \tag{10}$$

where the function d is some measure of the distance between O and O_t, assuming such a function exists, and ρ is called the *learning rate*. A generalization of this process to abstract topological spaces where such a distance function is a commodity is an open question.

Finally, for this to qualify as an intelligence process, we must have

$$\lim_{t\to\infty} d\left(O-O_t\right)\to 0. \tag{11}$$

3.1. Unsupervised and Supervised Learning

Some consideration should be given to the sets S and O. If $O=P(S)$ where $P(S)$ is the power set of S, then we will say that the mapping $\mathbb{I}$ is an *unsupervised* mapping. Otherwise, the mapping is *supervised*. The ramifications of this distinction are as follows. In supervised learning, the agent is given two distinct sets and trained to form a mapping between them explicitly. With unsupervised learning, the agent is tasked with learning subtle relationships in a single data set or, put more succinctly, to develop the mapping between S and its power set discussed above [9] [10].

3.2. Overtraining

Further, we should note that just because we have some function $\mathbb{I}:S\to O$ satisfying the definitions and assumptions of this section does not mean that this mapping be necessarily meaningful. After all, we could make a completely arbitrary but consistent mapping via the prescription above, and although this would satisfy all the definitions and assumptions, it would be complete memorization on the part of the agent. But this, in fact is exactly the definition of overtraining a common pitfall in the training stage of machine learning and about which one must be very diligent to avoid.

3.3. Entropy Minimization

One final part of the framework remains, and that is to show that entropy is minimized, as was stated at the beginning of this section. To show that, we consider $\mathbb{I}$ as a *probabilistic* mapping, with

$$\mathbb{P}_t\left[s_j^i\right]=\mathbb{P}\left[\mathbb{I}\left[s_j^i\right]=o^j\right], \tag{12}$$

indicating the probability that $\mathbb{I}$ maps $s^i\in S$ to some particular $o^j\in O$. From this, we can calculate the entropy in the mapping from S to O, at each iteration t. If the projection $\mathbb{I}\left[s^i\right]$ has N possible outcomes, then the Shannon entropy of each $s^i\in S$ is given by

$$\mathbb{E}_t\left[s^i\right]=-\sum_j^N \mathbb{P}_t\left[s_j^i\right]\log\left[\mathbb{P}_t\left[s_j^i\right]\right]. \tag{13}$$

The total entropy is simply the sum of $\mathbb{E}_t\left[s^i\right]$ over i. Let $|S|=M$, then for the purposes of standardization across varying cardinalities, it may be insightful to speak of the normalized entropy $\mathbb{E}_t\left[S\right]$,

$$\mathbb{E}_t\left[S\right]=\frac{1}{M}\sum_i^M \mathbb{E}_t\left[s^i\right]. \tag{14}$$

As $t\to\infty$, the mapping from each s_j^i to its corresponding o^j converges, andwe have

$$\lim_{t\to\infty}\mathbb{P}_t\left[s_j^i\right]=1,\quad i=1,2,\cdots,M\ . \tag{15}$$

Therefore

$$\lim_{t\to\infty}\mathbb{E}_t\left[S\right]=0. \tag{16}$$

Further, using the definition for Renyi entropy in 5 for each t and i

$$\mathbb{H}_{t,\alpha}\left[s^i\right]=\frac{1}{1-\alpha}\log\sum_j^N \mathbb{P}_t\left[s_j^i\right]^\alpha . \tag{17}$$

To show that the Renyi entropy is also minimized, we can use an identity involving the p-norm

$$\mathbb{H}_{t,\alpha}\left[s^i\right]=\frac{\alpha}{1-\alpha}\log\left\|\mathbb{P}_t\left[s_j^i\right]\right\|_\alpha, \tag{18}$$

and show that the log function is maximized $t\to\infty$ for $\alpha>1$, and minimized for $\alpha\in[0,1)$. The case $\alpha\to 0$ was shown above with the definition of Shannon entropy. To continue, note that

$$\sum_j^M \mathbb{P}_t\left[s_j^i\right]=1, \tag{19}$$

since the summation is taken over all possible states $o^j\in O$. But from 15, we have

$$\left\|\mathbb{P}_t\left[s_j^i\right]\right\|_\alpha<1,\ \alpha>1, \tag{20}$$

forfinite t and thus the log function is minimized only as $t\to\infty$. To show that the Renyi entropy is also minimized for $\alpha\in[0,1)$, we repeat the abovelogic but note that the with the sign reversal of $\frac{\alpha}{1-\alpha}$, the quantity $\left\|\mathbb{P}_t\left[s^i\right]\right\|_\alpha$ is *minimized* as $t\to\infty$.

Finally, we can take a normalized sum over all i to obtain the total Renyi entropy of S, $\mathbb{H}_{t,\alpha}[S]$. By this definition, then the total entropy is minimized along with its components.

3.4. Entropic Self Organization

In section 3 we talked about the definitions of intelligence via the mapping $\mathbb{I}:S\to O$. Here, we seek to apply the entropy minimization concept to $P(S)$ itself, rather than a mapping. Explicitly, $\sigma\subset P(S)$, where

$$\sigma=\left\{\boldsymbol{s}\in P(S)\right\}, \tag{21}$$

and for every $s\in S$, there is aunique $\boldsymbol{s}\in\sigma$ such that $s\in\boldsymbol{s}$. That is, every element of S has one and only one element of σ containing it. The term *entropic self-organization* refers to finding the $\Sigma\subset P(S)$ such that $\mathbb{H}_\alpha[\sigma]$ is minimized over all σ satisfying 21

$$\Sigma=\min\mathbb{H}_\alpha[\sigma]. \tag{22}$$

4. Global Effects

In nature, whenever a system is taken from a state of higher entropy to a state of lower entropy, there is always some amount of energy involved in this transition, and an increase in the entropy of the rest of the environment greater than or equal to that of the entropy loss [6]. In other words, consider a system S composed of two subsystems, s_1 and s_2., then

$$S=s_1+s_2. \tag{23}$$

Now, consider that system in equilibrium at times t_1, and $t_2 t_1>t_2$ and denote the state at each S^1 and S^2, respectively. Then due to the second law of thermodynamics,

$$S^2\geq S^1. \tag{24}$$

Therefore

$$s_1^2+s_2^2\geq s_1^1+s_2^1. \tag{25}$$

Now, suppose one of the subsystems, say, s_1 decreases in entropy by some amount, Δs between t_1, and t_2, *i.e.* $s_1^2=s_1^1-\Delta s$. Then to preserve the inequality, the entropy of the rest of the system must be such that

$$s_2^2\geq s_2^1+\Delta s. \tag{26}$$

So the entropy of the rest of the system has to increase by an amount greater than or equal to the loss of entropy in s_1. This will require some amount of energy, ΔE.

Observe that all we have done thus far is follow the natural consequences of the Second Law of Thermodynamics with respect to our considerations with intelligence. While the second law of thermodynamics has been verified time and again in virtually all areas of physics, few have extended it as a more general principal in the context of information theory. In fact, we will conclude this section with a postulate about intelligence:

Computational intelligence is a process that locally minimizes and globally maximizes Renyi entropy.

It should be stressed that although the above is necessary of intelligence, it is not sufficient in the justification of an algorithm or process as being intelligent.

5. Applications

Here, we implement the discussions of this paper to practical examples. First, we consider a simple example of unsupervised learning; a clustering algorithm based on Shannon entropy minimization. Next we look at some simple behavior of an intelligent agent as it acts to maximize global entropy in its environment.

5.1. Clustering by Entropy Minimization

Consider a data set consisting of a number of elements organized into rows. The experiment that follows, we consider 300 samples, each a vector from $\mathbb{R}^3$. In this simple proof of concept we will group the data into like neighborhoods by minimizing the entropy across all elements at each respective index in the data set. This is a data driven example, so essentially we use a genetic algorithm to perturb the juxtaposition of members of each neighborhood until the global entropy reaches a minimum (entropic self organization), while at the same time avoiding trivial cases such as a neighborhood with only one element.

We leverage the Python framework heavily for this example, which is freely available for many operating systems at [11].

Please note that this is a simple prototype, a proof of concept used to exemplify the material in this paper. It is not optimized for latency, memory utilization, and it has not been optimized or performance tested against other algorithms in its comparative class, although dramatic improvements could be easily achieved by integrating the information content of the elements into the algorithm. Specifically, we would move elements with high information content to clusters where that element would otherwise have low information content. Furthermore, observe that for further efficacy, a preprocessing layer may be beneficial, especially with topological data set. Nevertheless, applications of this concept applied to clustering on small and large scales will be discussed in a future work.

We can visualize the progression of the algorithm and the final results, respectively, in **Figure 1**. For simplicity, only the first two (non-noise) dimensions are plotted. The accuracy of the clustering algorithm was 8.3% error rate in 10,000 iterations, with an average simulation time: 480.1 seconds.

Observe that although there are a few "blemishes" in the final clustering results, with a proper choice of parameters including the maximum computational epochs the clustering algorithm will eventually succeed with 100% accuracy.

Also pictured in **Figure 2** are the results of the clustering algorithm applied to a data set containing four additional fields of pseudo-randomly generated noise, each in the interval $[-1,1]$. The performance of this trial was worse than the last in terms of speed, but was had about the same classification accuracy. The accuracy of the clustering algorithm was 6.0% error rate in 10,000 iterations, with an average simulation time: 1013.1 seconds.

5.2. Entropy Maximization

In our next set of examples consider a virtual agent confined to move about a "terrain", represented by a three-dimensional surface, given by one of the two following equations, each of which are plotted visually in **Figure 3**, and defined by the following functions, respectively:

$$z = \exp\left[-\left(x^2 + y^2\right)\right], \tag{27}$$

and

$$z = \frac{1}{4}\exp\left[-\left(\frac{x}{10}^2 + \frac{y}{10}^2\right)\right]\left(\cos\frac{1}{2}\pi y + \sin\frac{1}{2}\pi x + 2\right). \tag{28}$$

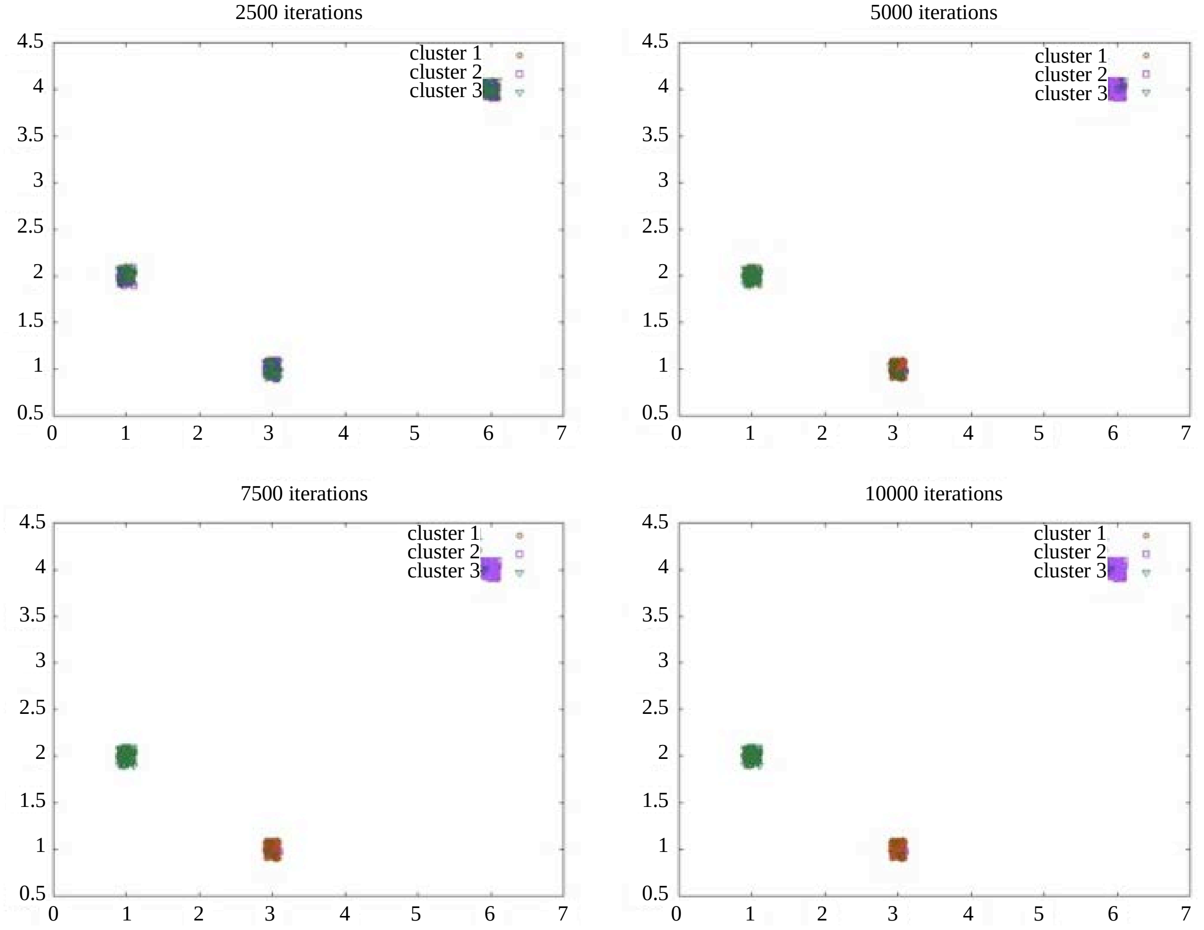


Figure 1. Entropic clustering algorithm results over time.

We will confine x and y such that $(x,y)\in([x_{\min},x_{\max}],[y_{\min},y_{\max}])$ and note that the range of each respective surface is $z\in[0,1]$. The algorithm proceeds as follows. First, the agent is initialized with a starting position, $p_0=(x_0,y_0)$. It updates the coordinates of the agent's position by incrementing or decrementing by some small value, $\varepsilon=(\varepsilon_x,\varepsilon_y)$. As the agent meanders about the surface, data is collected as to its position on the *z*-axis.

If we partition the range of each surface into equally spaced intervals, we can form a histogram H of the agent's positional information. From this H we can construct a discrete probability function, $\mathbb{P}_H$ and thus calculate the Renyi entropy. The agent can then use feedback from the entropy determined using H to calculate an appropriate ε from which it upates its position, and the cycle continues. The overall goal is to maximize its entropy, or timeout after a predetermined number of iterations.

In this particular simulation, the agent is initialized using a "random walk", in which is ε is chosen at random. Next, it is updated using feedback from the entropy function.

From the simple set of rules, we see emergent desire for parsimony with respect to position on the surface, even in the less probable partitions of z, as $z\to 1$. As our simulation continues to run, so tends $\mathbb{P}_H$ to a uniform distribution.

The **Figure 3** depict a random walk on surface 1 and surface 2 respectively, where the top and bottom right figures show surface traversal using an entropic search algorithm.

6. Related Works

Although there are many approaches to intelligence from the angle of cognitive science, few have been proposed from the computational side. However, as of late, some great work in this area is underway.

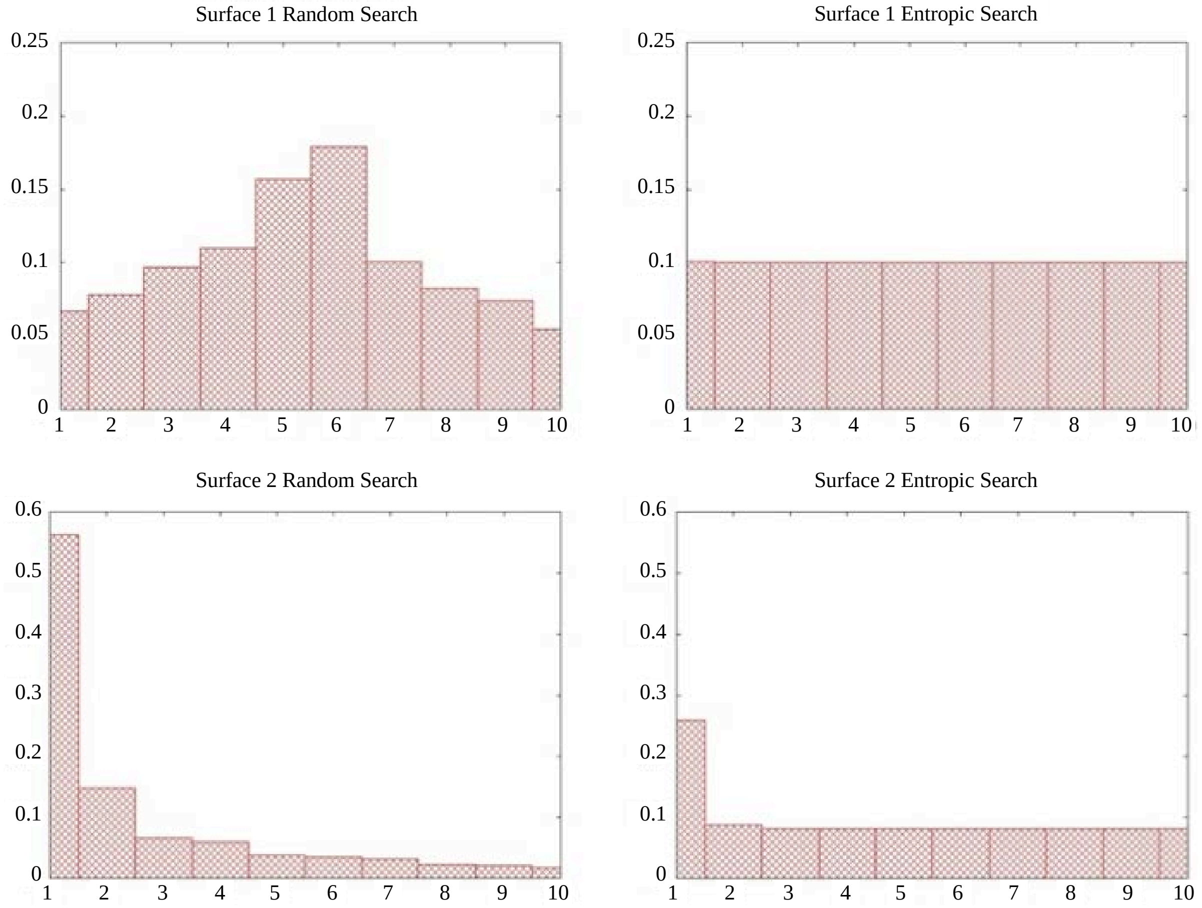


Figure 2. Entropic clustering algorithm results over time.

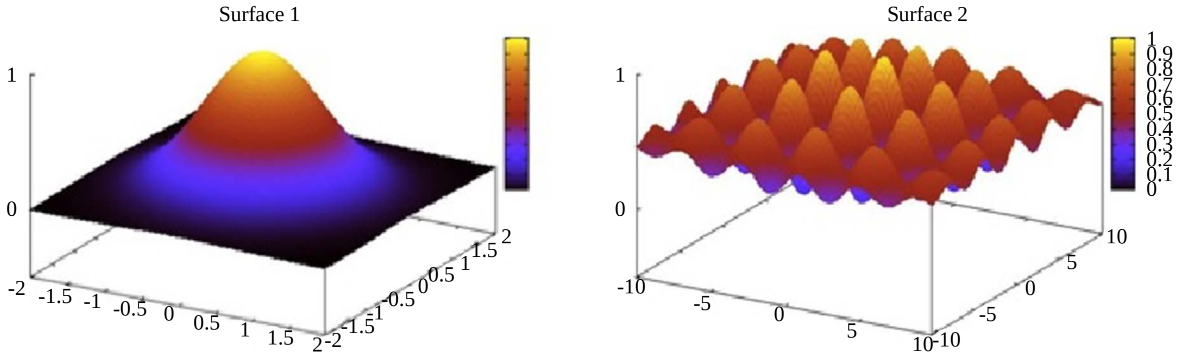


Figure 3. Surfaces for hill climbing agent simulation.

Many sources claim to have computational theories of intelligence, but for the most part these "theories" merely act to describe certain aspects of intelligence [12] [13]. For example, Meyer in [12] suggests that performance on multiple tasks is dependent on adaptive executive control, but makes no claim on the emergence of such characteristics. Others discuss how data is aggregated. This type of analysis is especially relevant in computer vision and image recognition [13].

The efforts in this paper seek to introduce a much broader theory of emergence of autonomous goal directed behavior. Similar efforts are currently under way.

Inspired by physics and cosmology, Wissner-Gross asserts autonomous agents act to maximize the entropy in their environment [14]. Specifically he proposes a path integral formulation from which he derives a gradient that can be analogized as a causal force propelling a system along a gradient of maximum entropy over time. Using this idea, he created a startup called *Entropica* [15] that applies this principal in ingenious ways in a variety of different applications, ranging from anything to teaching a robot to walk upright, to maximizing profit potential in the stock market.

Essentially, what Wissner-Gross did was start with a global principal and worked backwards. What we did in this paper was to arrive at a similar result from a different perspective, namely entropy minimization.

7. Conclusions

The purpose of this paper was to lay the groundwork for a generalization of the concept of intelligence in the computational sense. We discussed how entropy minimization can be utilized to facilitate the intelligence process, and how the disparities between the agent's prediction and the reality of the training set can be used to optimize the agent's performance. We also showed how such a concept could be used to produce a meaningful, albeit simplified, practical demonstration.

Some future work includes applying the principals of this paper to data analysis, specifically in the presence of noise or sparse data. We will discuss some of these applications in the next paper.

More future work includes discussing the underlying principals under which data can be collected hierarchically, discussing how computational processes can implement the discussions in this paper to evolve and work together to form processes of greater complexity, and discussing the relevance of these contributions to abstract concepts like consciousness and self awareness.

In the following paper we will examine how information can aggregate together to form more complicated structures, the roles these structures can play.

More concepts, examples, and applications will follow in future works.

References

[1] Wechsler, D. and Matarazzo, J.D. (1972) Wechsler's Measurement and Appraisal of Adult Intelligence. Oxford UP, New York.

[2] Gardner, H. (1993) Frames of the Mind: The Theory of Multiple Intelligences. Basic, New York.

[3] Sternberg, R.J. (1982) Handbook of Human Intelligence. Cambridge UP, Cambridge Cambridgeshire.

[4] Hawkins, J. and Sandra, B. (2004) On Intelligence. Times, New York.

[5] Ihara, S. (1993) Information Theory for Continuous Systems. World Scientific, Singapore.

[6] Schroeder, D.V. (2000) An Introduction to Thermal Physics. Addison Wesley, San Francisco.

[7] Penrose, R. (2011) Cycles of Time: An Extraordinary New View of the Universe. Alfred A. Knopf, New York.

[8] Hawking, S.W. (1998) A Brief History of Time. Bantam, New York.

[9] Jones, M.T. (2008) Artificial Intelligence: A Systems Approach. Infinity Science, Hingham.

[10] Russell, S.J. and Peter, No. (2003) Artificial Intelligence: A Modern Approach. Prentice Hall/Pearson Education, Upper Saddle River.

[11] (2013) Download Python. N.p., n.d. Web. 17 August 2013. http://www.python.org/getit

[12] Marr, D. and Poggio, T. (1979) A Computational Theory of Human Stereo Vision. *Proceedings of the Royal Society B*: *Biological Sciences*, **204**, 301-328. http://dx.doi.org/10.1098/rspb.1979.0029

[13] Meyer, D.E. and Kieras, D.E. (1997) A Computational Theory of Executive Cognitive Processes and Multiple-Task Performance: Part I. Basic Mechanisms. *Psychological Review*, **104**, 3-65. http://dx.doi.org/10.1098/rspb.1979.0029

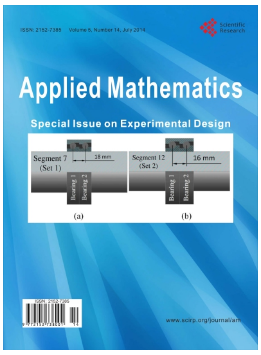
[14] Wissner-Gross, A. and Freer, C. (2013) Causal Entropic Forces. *Physical Review Letters*, **110**, Article ID: 168702. http://dx.doi.org/10.1103/PhysRevLett.110.168702

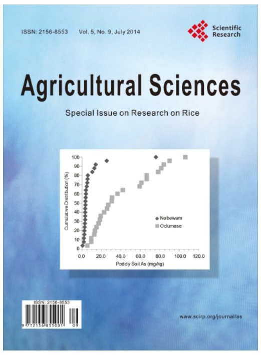
[15] (2013) Entropica. N.p., n.d. Web. 17 August 2013. http://www.entropica.com

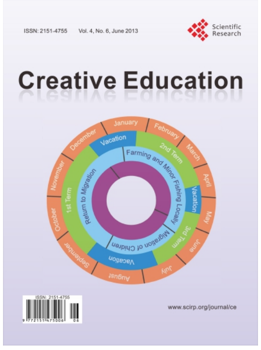
Advances in
Bioscience and Biotechnology

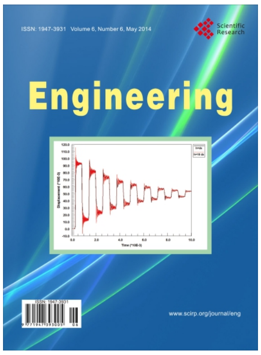
American Journal of
Analytical Chemistry

American Journal of
Plant Sciences
Special Issue on Photosynthesis

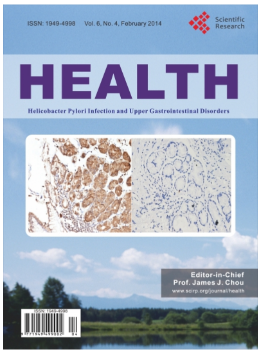

Applied Mathematics
Special Issue on Experimental Design


Agricultural Sciences
Special Issue on Research on Rice


Creative Education

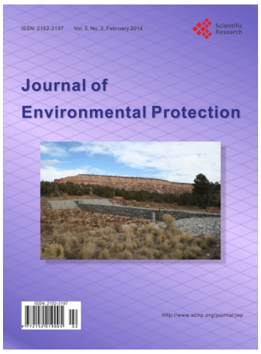

Engineering

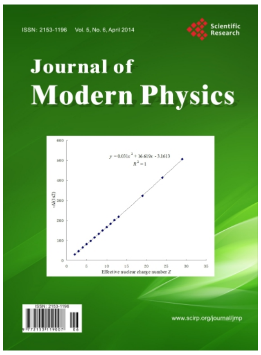
Food and Nutrition Sciences


HEALTH
Helicobacter Pylori Infection and Upper Gastrointestinal Disorders
Editor-in-Chief
Prof. James J. Chou

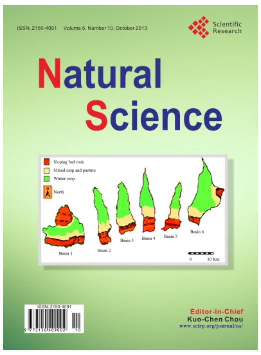
Journal of Computer
and Communications

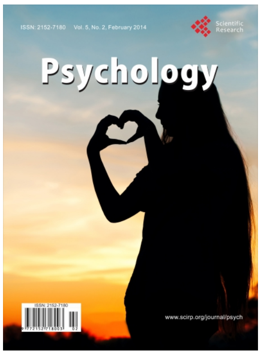
Journal of Cancer Therapy
Special Issue on Bone Tumor Research


Journal of
Environmental Protection


Journal of
Modern Physics

Modern Economy


Natural
Science
Editor-in-Chief
Kuo-Chen Chou


Psychology