# Predictive General Intelligence

**Jacob F. Valdez**
Limboid AI
`jacob.valdez@limboid.ai`

## Abstract

The abstract paragraph should be indented ½ inch (3 picas) on both the left- and right-hand margins. Use 10 point type, with a vertical spacing (leading) of 11 points. The word **Abstract** must be centered, bold, and in point size 12. Two line spaces precede the abstract. The abstract must be limited to one paragraph.

## 1   Introduction

REMOVE THIS FIRST PARAGRAPH Our objective is not "can we make artificial general intelligence¿' but "how general can we make artificial intelligence?" Intelligence has been defined as "an agent's ability to achieve goals in a wide range of environments." (Legg and Hutter, 2007 CITE) and "skill-acquisition efficiency" with respect to available information (Chollet 2019 CITE). Although these definitions imply there is no upper bound on generality and that artificial general intelligence in the superlative sense is an illusion (Human ne AGI CITE), they do provide a metric we can put to our informal understanding of the concept and a reward signal to follow.

The life sciences and AI disciplines have identified numerous techniques underlying principles of intelligence which lead to performance increase over a diverse array of tasks. This work aims to consolidate some of those principles into a working implementation of task-agnostic, open-ended artificial 'general' intelligence together with experiments and ablation studies. Our paper is organized as follows: section 2 reviews key principles of intelligence; section 3 composes those critical ideas into a theoretical AI system and describes its software implementation; section 4 presents experiments with observations over a diverse set of open-ended learning environments and numerous close-ended tasks and benchmarks and also describes ablation studies; section 5 presents descriptive and comparative analyses of individual experimental results; and finally, section 6 gives a general discussion of this work along with its broader impact and a conclusion.

## 2   Principles of Intelligence

We take lessons from physical and life sciences as well as artificial intelligence. Our aim is not to exhaust every thought and theory but only to extract those critical for our AI system. See CITE for a more elaborate discussion. Energetic grounding Intelligence begins with information (CITE) which, in turn, depends on energy. Information theory even equates information with energy by $E(x) = -\log p(x)$. (CITE) With energy or information is equated to the negative log likelihood of an event occurring, events that are more probable demand less energy while those that are less likely demand more energy. It is important to note: this is an allocentric measure. The information carried by a signal $x_i$ may be very useful to another human, yet $-\log p(x_i)$ only measures the information content of a signal with respect to a particular probability distribution $p$. This probability distribution may represent various information systems. For instance, in discrete signal communication, $p$ represents a prior expectation of signal occurrence. Optimized codebooks use short strings to identify frequently occurring events while longer stings discriminate between infrequently occurring ones. The effect is, transmitting codes for likely occurrences consumes less physical energy while the

opposite is true for more informative, less expected signals. In quantum physics, a particle position in space may be represented by a probability wavefunction. Natural living systems represent a prior expectation over their environmental state. This is especially apparent under electroencephalogram observation of the brain where increased neuronal glucose metabolism is correlated with less expected signals from outside or inside the brain.

In the context of machine learning, the model represents a probability distribution over input-output pairs. Training the model on a dataset of high-energy, novel input-output pairs is analogous to allowing a cold object to absorb some of the entropic energy from a warmer mass. On the other hand, the more likely an input-output pair is to be produced by a model, the more 'neutral' its temperature becomes and less informative it is to the model. This explains the traditional log improvement of training with respect to the amount of information processed. (SHOW FIGURE) Continuing the thermodynamics parallel, machine learning systems even share properties of states and transitions during training. (figure XX. CITE FIGURE SHOWING SOLID-LIQUID-GAS AND NEURAL NETWORK TRAINING)

When its expectation is taken over the event space, probabilistic energy becomes entropy $H(p) = \mathbb{E}\left[-\log p(x)\right]$. The entropy of a probability distribution $q$ with respect to a prior expectation $p$ gives the cross entropy $H(p, q) = \mathbb{E}_p\left[\log q(x)\right]$ and asymmetric energy loss from one distribution to another may be defined by the Kullback-Leibler Divergence metric $D_{KL}(p \parallel q) = H(p, q) - H(p) = \mathbb{E}_p\left[\log \frac{p(x)}{q(x)}\right]$. Kullback-Leibler Divergence gives an objective measure to the minimum information required to transform a prior to a posterior, and if we let the uniform prior $\mathcal{U}$ represent zero information, the Kullback-Leibler divergence of any other distribution $q$ from it $D_{KL}(q \parallel \mathcal{U})$ then gives an absolute measure to the information contained in $q$.

These mathematically pure yet rich ideas from information theory, give basis to formalize connections across numerous domains of science and identify the overarching motif of free energy minimization. With respect to any particular system, free energy is the complement of bound and entropic energy. For instance, note the similarity between Gibbs free energy $G$, Helmholtz free energy $F$, grand free energy $\phi$, and Kullback-Leibler divergence $D_{KL}(p \parallel q)$:

$$
\begin{aligned}
G &\equiv E + PV &&- TS \\
F &\equiv E &&- TS \\
\phi &\equiv E &&- TS - \mu N \\
\underbrace{D_{KL}(p \parallel q)}_{\text{free energy}} &\equiv \underbrace{H(p, q)}_{\text{bound energy}} &&\underbrace{-H(p)}_{\text{entropy}}
\end{aligned}
$$

In all domains of definition, the universe fights free energy. This involves entropy maximization. For instance, the Casmir effect provides an attractive or repulsive force without invoking any of the fundamental forces. Instead it influences particles in a way that directly maximizes the uniformity and entropy of space. The principle of least action states that XXXX.

Life represents a defiance of this overall trend, yet even in its struggle, organisms do not expend unnecessary amounts of energy. Blood sugar is minimized until necessary. YYY. Allostatic control theory shows that life cannot tolerate excessive amounts of energy and such unusual physiological states as those fostered by modern society give rise to diabetes, chronic hypertension, and other lifestyle problems.

In the mind, Friston's free energy principle posits that - define - discuss examples of action in humans - discuss neuronal principle and experiment of neural controlled car in "Free Energy and the Brain" - explain that the EEG is observing areas of extra glucose consumption - aka free energy

INTRODUCE INFORMATION THEORY HERE. THEN DISCUSS HOW LIFE IS A DEFIANCE OF THE TREND. THEN HOW LIFE EATS, RESTS, AND ACTS. FINALLY MAKE PARALLELS WITH ARTIFICIAL INTELLIGENCE.

Natural intelligence does not need to be spoon fed data to acquire Diverse, adapted skills and behaviors. This should not be the case for artificial intelligence either.

The theoretical world of machine learning meets a hard reality when training time comes. State of the art AI systems consume thousands of RANT ABOUT FINANCIAL AND ENVIRONMENTAL ASPECTS OF LEARNING. In contrast, human intelligence finds relative ease in acquiring new

skills, and the carbon footprint of physiologically comparable species does not approach computing systems. (CITE)

How does natural intelligence cope with the reality of bioenergetics? By keeping it foremost in mind. Energetic demands provide strong guidance to the

Allostatic control theory IS THAT THE RIGHT NAME FOR THIS THEORY? carries this thought further with "predictive regulation" which "proposes that efficient regulation requires anticipating needs and preparing to satisfy them before they arise." (Allostasis: A model of predictive regulation CITE) DISCUSS PRINCIPLES OF ALLOSTATICS PAPER.

After feeding, the animal typically enters a 'rest and digest' period where a subset of its peripheral nervous system – the parasympathetic system – dominates decreasing respiration, heart rate, and blood pressure. Through these are natural processes promoting adaptive, diverse intelligent animal life, importantly, they are not appropriate at all times. In excess, they can reduce the amount of mobile energy available for survival functions. Homeostatic control therefore must balance energy collection against mobilization.

data = increase budget mobile energy = budget information energy = decrease budget

I might be confusing information, data, and energy We extend these bioenergetic themes to intelligence with information theory's definition of MOVE THIS TO BIOLOGICAL PART BEFORE INTORDUCING

This brief yet rich consideration of information

We emphasize that data only loses expected energy with consumption. That is, for any machine learning model $f \colon X \mapsto Y$ where $X$ and $Y$ are the domains of inputs and outputs respectively, the probabilistic energy of the output is upper-bounded by the input $E(X) \geq E(Y)$. Using deep learning, we might imagine each layer of the neural network as 'absorbing' some of the signal's energy as it maps the signal onward $E(X) \geq E(H^1) \geq E(H^2) \geq E(H^l)$. In the specific framework of agent-environment interaction, this means following the route from perception to action.

As with physical energy, information theoretic energy – information – is always readily available. However, its relative availability may not be constant and the form it exists in may not be useful to the agent. Increasingly adapted and general intelligences should be able to anticipate demand and consume information in a variety of forms. This implies the agent should receive diverse modalities of information.

Just as stored information is analogous to energy, data consumption is to eating. Collecting training data gives the AI system information which may lower the negative log likelihood of future data. Regarding the 'nutrition' of consumed data, recent works have drawn attention to the danger of poisoning – injecting data into a training pipeline that decreases its performance or stability – as well as XXXX – which is comparable to malnutrition in our discussion. Most existing machine learning systems today are embryos in this respect. They are force fed data with little attention to feedback. In contrast, as animal life develops, it gains the capability to reject food that is presented to it, choose food from a range of options, and even independently search out food without the intervention of a trainer. AI systems should imitate natural intelligence accordingly.

The process of training may be likened to digestion. JUSTIFY. EXPLAIN.

Finally, using information theoretic energy allows us to define a reward system carrying rich implications from nature. NATURAL REWARD SIGNALS ARE HIGHLY CORRELATED WITH ENERGY ABUNDANCE. DISCUSS HOW NATURAL REWARD SYSTEMS - arise from a complex, chaotic "internal milieu" thus shaking out local minima - resist tampering. They break down in most cases to not forever favor 'local behavioral optima' - are intrinsic. So while the AI system is extrinsically reward-free, it generates its own intrinsic reward signals by information theoretic mechanisms

Triply grounding the AI system's energy in its compute platform, the interaction environment, and the AI's information theoretic energy helps align its objectives with our own. Open-ended learning (combine with previous paragraph) Few points deserve as much emphasis as reward-free (REMOVE), task-agnostic, open-ended learning in the domain of abstract principles of intelligence. Supervised learning plays a key role in superhuman narrow AI systems today, and supervised learning over large datasets - notably in language modeling - extracts information useful to tasks in unexpected

ways (CITE GPT3 & Switch-Transformer papers). Nonetheless, breaking present intelligence generality barriers demands zooming out from mean absolute errors, Inception scores, BLEU scores, IoU's - completely leaving the domain of classic supervised learning - to entropy, Kullback–Leibler divergence, and other 'pure' information-theoretic metrics.

Despite the advantages these principles of intelligence bring to an AI system, more inductive biases are necessary to approach human-like cognition and behavior (See Section 5 "Ablations"). Agent-Environment Interaction It should be clear that imitating natural intelligence demands adopting an agent-environment interaction framework. Unlike blind machine learning pipelines, the AI system must be able to perceive its actions and their consequences just as it would take in traditional observations.

However this does not necessarily limit the agent to a single body. Parallelization

Segmentation, Representation, and Composition TODO

Learning Styles Inverse language modeling

Learning styles Continual – or developmental – learning In context, few shot Environments

Measure intelligence on psychometric tests; Measure skills on classic SL, UL, and RL benchmarks: image classification image generation sentiment classification conditional text generation atari control 3D kinematics multiagent games language tasks

Language and demonstration guided Computer control Robot control

Each of these principles of intelligence provide unique advantages to AI systems, and proofs-of-concept such as RAINBOW reinforcement learning, X, and Y demonstrate that combining multiple inductive biases into one AI system synergizes the advantages of individual ideas. However, to our knowledge, no work has combined all of them in one system. Our work, Predictive 'General' Intelligence, does this.

# 3 Predictive 'General' Intelligence

Make brief statement on why prediction is a general principle of intelligence

# 4 Experiments

# 5 Analyses

# 6 Discussion

## Broader Impact

Authors are required to include a statement of the broader impact of their work, including its ethical aspects and future societal consequences. Authors should discuss both positive and negative outcomes, if any. For instance, authors should discuss a) who may benefit from this research, b) who may be put at disadvantage from this research, c) what are the consequences of failure of the system, and d) whether the task/method leverages biases in the data. If authors believe this is not applicable to them, authors can simply state this.

Use unnumbered first level headings for this section, which should go at the end of the paper. **Note that this section does not count towards the eight pages of content that are allowed.**

## Acknowledgments and Disclosure of Funding

# References

[] William Fulton. Introduction to intersection theory in algebraic geometry. In *Regional Conference Series in Mathematics*, number 54, 1983.