ON THE EQUIVALENCE OF CAUSAL PATH ENTROPY AND EMPOWERMENT

A PREPRINT

Aidan Rocke
aidanrocke@gmail.com

December 6, 2018

ABSTRACT

The main contribution of this paper is to show that two intrinsic reward functions, the Causal Path Entropy and Empowerment, are equivalent only in deterministic environments. In non-deterministic environments, the Causal Path Entropy has several disadvantages over the Empowerment which are discussed.

1 Causal Path Entropy

We shall start by defining the notion of Causal Path Entropy, introduced in [1], show how it simplifies to a conditional Shannon entropy for digital organisms, and extend it so that it explicitly accounts for actions taken by an organism:

1. For any open thermodynamic system such as a biological organism we may treat phase-space paths taken by the system over a time interval $[0, \tau]$ as microstates and partition them into macrostates $\{X_i\}_{i \in I}$ using the equivalence relation:

$$x(t) \sim x'(t) \iff x(0) = x'(0) \tag{1}$$

As a result, we can identify each macrostate X_i with a present system state $x(0)$.

2. We may define the Causal Path Entropy S_c of a macrostate X_i associated with the present system state $x(0)$ as the path integral:

$$S_c(X_i, \tau) = -k_B \int_{x(t)} P(x(t)|x(0)) \ln P(x(t)|x(0)) Dx(t) \tag{2}$$

where k_B is the Boltzmann constant.

3. It must be noted that in order to calculate (2) we need the state-transition probability distribution $P(x(t)|x(0))$ which corresponds to an exact simulator of the agent's environment. Given that this is generally unknown to the agent at the instant $x(0)$, macrostates X_i are generally unknown to the agent as well, and therefore it's more epistemically sound to denote the Causal Path Entropy as:

$$S_c(x(0)) = -k_B \int_{x(t)} p(x(t)|x(0)) \ln p(x(t)|x(0)) Dx(t) \tag{3}$$

where $p(x(t)|x(0))$ denotes a subjective state-transition probability distribution.

4. If the organism is digital, i.e. simulated by a Turing machine, we may drop the Boltzmann constant and a discrete phase-space implies that $S_c(x(0))$ simplifies to the conditional Shannon entropy:

$$\begin{aligned} S_c(x_0) &= -\sum_{x_n} p(x_n|x_0) \ln p(x_0|x_0) \\ &= H(x_n|x_0) \end{aligned} \tag{4}$$

5. In order for the calculation of Causal Path Entropy to be useful for a digital organism we must explicitly account for its agency which is determined by its capacity for rational action in the world.

6. If $\mathcal{A}$ denotes a discrete action space and $\mathcal{X}$ denotes a discrete state space:

$$p(x_n|x_0) = \frac{p(x_n, a_{1:n}|x_0)}{p(a_{1:n}|x_n, x_0)} \tag{5}$$

where $a \in \mathcal{A}^n$ is an n-tuple of actions and $x_0, x_n \in \mathcal{X}$.

7. Now, we may note that $p(x_n, a_{1:n}|x_0)$ in the numerator of (5) may be expressed in terms of the agent's conditional distribution over n-step action sequences $w(a_{1:n}|x_0)$:

$$p(x_n, a_{1:n}|x_0) = w(a_{1:n}|x_0)p(x_n|a_{1:n}, x_0) \tag{6}$$

8. By combining (5) and (6) we have:

$$p(x_n|x_0) = \frac{w(a_{1:n}|x_0)p(x_n|a_{1:n}, x_0)}{p(a_{1:n}|x_n, x_0)} \tag{7}$$

9. Using (7) the Causal Path Entropy becomes:

$$S_c(x_0) = \max_w \mathbb{E}\big[\ln\big(\frac{p(a_{1:n}|x_n, x_0)}{w(a_{1:n}|x_0)p(x_n|a_{1:n}, x_0)}\big)\big] \tag{8}$$

Hence, we have:

$$S_c(x_0) = \max_w \mathbb{E}\big[H(a_{1:n}|x_0) - H(a_{1:n}|x_n, x_0) + H(x_n|a_{1:n}, x_0)\big] \tag{9}$$

(9) shall be useful in analysing the difference between the Causal Path Entropy and Empowerment, which we shall now introduce.

2 Empowerment

We shall introduce the n-step empowerment as was done in [3].

1. The n-step empowerment is defined by searching for the maximal mutual information $I(\cdot, \cdot)$ conditional on a starting state x_0 between a sequence of $n \in \mathbb{N}$ actions $a_{1:n}$ and the final state reached x_n:

$$\xi(x_0) = \max_w I(a_{1:n}, x_n|x_0) = \max_w \mathbb{E}\big[\ln\big(\frac{p(a_{1:n}, x_n|x_0)}{w(a_{1:n}|x_0)p(x_n|x_0)}\big)\big] \tag{10}$$

2. Hence, (10) may be expressed as the difference of two conditional Shannon entropies:

$$\xi(x_0) = \max_w \mathbb{E}\big[H(a_{1:n}|x_0) - H(a_{1:n}|x_n, x_0)\big] \tag{11}$$

where we used the fact that $p(a_{1:n}|x_n, x_0) = \frac{p(a_{1:n}, x_n|x_0)}{p(x_n|x_0)}$.

3 Equivalence of Empowerment and Causal Path Entropy

1. If we combine (9) and (11) we find that the Causal Path Entropy at x_0 may be expressed in terms of the Empowerment at x_0:

$$S_c(x_0) = \xi(x_0) + \max_w \mathbb{E}\big[H(x_n|a_{1:n}, x_0)\big] \tag{12}$$

2. Therefore, in order to have equivalence we must have:

$$H(x_n|a_{1:n}, x_0) = 0 \tag{13}$$

which is true if and only if $p(x_n|a_{1:n}, x_0) = 1 \oplus 0$ (where $\oplus$ denotes XOR) and this is the case only in deterministic environments.

3. It must be noted that in deterministic environments (12) simplifies to:

$$S_c(x_0) = \xi(x_0) = \ln N_{x_0} \tag{14}$$

where N_{x_0} represents the number of intrinsic options $a_{1:n} \in \mathcal{A}^n$ available at x_0.

4 Discussion

Whenever $S_c(x_0) \neq \xi(x_0)$ we must have $H(x_n|a_{1:n}, x_0) > 0$ so the Causal Path Entropy provides intrinsic compensation for:

1. Exploring unpredictable environments.
2. Exploring unknown environments.
3. Unreliable sensors.
4. Unreliable sensors.

Considering that there are more ways statistical uncertainty, rather than epistemic uncertainty, may contribute to $H(x_n|a_{1:n}, x_0) > 0$, on average the Causal Path Entropy does less than fully reward the agent for determining its number of intrinsic options. This is especially clear if we use (4) and (12) to re-formulate the Empowerment of the agent at $x(0)$:

$$\xi(x_0) = S_c(x_0) + \max_w \mathbb{E}\big[H(x_n|a_{1:n}, x_0)\big] = \max_w \mathbb{E}\big[H(x_n|x_0) - H(x_n|a_{1:n}, x_0)\big] \tag{15}$$

so $H(x_n|a_{1:n}, x_0) > 0$ corresponds to decreased Empowerment.

References

[1] Gross, A. Wissner. (2013) Causal Entropic Forces. Physical Review Letters.

[2] Salge, C., Glackin, C. & Polani, D. Empowerment-An Introduction. Arxiv.

[3] Mohamed, S., Rezende, D. Variational Information Maximisation for Intrinsically Motivated Reinforcement Learning. Arxiv.