Plasticity as the Mirror of Empowerment

David Abel	Michael Bowling	André Barreto	Will Dabney
Google DeepMind	Amii, University of Albe	rta Google DeepMind	Google DeepMind
Shi Dong	Steven Hansen	Anna Harutyunyan	Khimya Khetarpal
Google DeepMind	Google DeepMind	Google DeepMind	Google DeepMind
Clare Lyle	Razvan Pascanu	Georgios Piliouras	Doina Precup
Google DeepMind	Google DeepMind, Mila	Google DeepMind	Google DeepMind
Jonathan Richens	Mark Rowland	Tom Schaul	Satinder Singh
Google DeepMind	Google DeepMind	Google DeepMind	Google DeepMind

Abstract

Agents are minimally entities that are influenced by their past observations and act to influence future observations. This latter capacity is captured by empowerment, which has served as a vital framing concept across artificial intelligence and cognitive science. This former capacity, however, is equally foundational: In what ways, and to what extent, can an agent be influenced by what it observes? In this paper, we ground this concept in a universal agent-centric measure that we refer to as plasticity, and reveal a fundamental connection to empowerment. Following a set of desiderata on a suitable definition, we define plasticity using a new informationtheoretic quantity we call the *generalized directed information*. We show that this new quantity strictly generalizes the directed information introduced by Massey (1990) while preserving all of its desirable properties. Our first finding is that plasticity is the mirror of empowerment: The agent's plasticity is identical to the empowerment of the environment, and vice versa. Our second finding establishes a tension between the plasticity and empowerment of an agent, suggesting that agent design needs to be mindful of both characteristics. We explore the implications of these findings, and suggest that plasticity, empowerment, and their relationship are essential to understanding agency.

1 Introduction

Two capacities represent fundamental aspects of any agent, regardless of its design, physical substrate, or goals: In what ways can the agent be shaped by what it observes? And, what things can an agent do to shape aspects of its observable environment? A system that entirely lacks either capacity—for instance, a machine whose outputs cannot be influenced by its inputs—is likely lacking agency altogether (Barandiaran et al., 2009). In this paper, we highlight a fundamental connection between these two capacities, which we refer to as empowerment and plasticity.

Empowerment. Empowerment refers to an agent's ability to influence its observable future. For example, a car with a full tank of gas has higher empowerment than a car with an empty tank, precisely because the former can realize more outcomes than the latter. The concept of empowerment was first introduced by Klyubin et al. (2005b) as "an agent-centric quantification of the amount of control or influence the agent has and perceives" (p. 2). Since its inception, the concept has been expanded



Figure 1: The mirror of plasticity and empowerment: For any agent and environment, the plasticity of the agent $\mathfrak{P}(\lambda)$ is equivalent to the empowerment of the environment $\mathfrak{E}(e)$, and vice-versa.

to act as a plausible explanation of intrinsic drive in the absence of reward (Klyubin et al., 2005a; Mohamed and Jimenez Rezende, 2015; Ringstrom, 2023), framed key aspects of safety and alignment (Salge and Polani, 2017; Turner et al., 2021), skill discovery (Gregor et al., 2016; Campos et al., 2020; Baumli et al., 2021; Levy et al., 2023), social influence (Jaques et al., 2019), self-preservation (Ringstrom, 2023), and has played a meaningful role in explanatory theories of how people explore (Brändle et al., 2023).

Plasticity. Conversely, plasticity reflects an agent's capacity to "remain...adaptive... in response to significant events" (p. 2, Carpenter and Grossberg 1988), and has been studied extensively across neuroscience and machine learning. Within neuroscience, "plasticity" often refers to "synaptic plasticity" (Citri and Malenka, 2008), and reflects "the capacity of the neural activity generated by an experience to modify neural circuit function" (p. 1). It is a central building block for understanding memory and forgetting (Fusi et al., 2005), often studied through the lens of the stability-plasticity dilemma (Carpenter and Grossberg, 1988, 1987a,b). Within machine learning, plasticity is often studied under the lens of the *loss* of plasticity, (Dohare et al., 2021, 2024), which is especially prevalent within neural networks (Lyle et al., 2023; Abbas et al., 2023). Each of Chen et al. (2023), Raghavan and Balaprakash (2021), and Kumar et al. (2023) further study plasticity under the lens of the stability-plasticity dilemma in different learning settings. Across all of this work, it is clear plasticity is an important capacity for any agent—its connection to empowerment, however, beyond sometimes sharing an information-theoretic toolkit (Kumar et al., 2023), has remained distant. To this end, we develop a simple, universal definition of agent plasticity (Definition 4.1) that brings the concept on equal footing with empowerment.

The Mirror of Plasticity and Empowerment. Our first finding is that plasticity is the mirror of empowerment (Proposition 4.6), pictured in Figure 1. To reveal this mirror, we exploit the inherent symmetry in the exchange of symbols between agent and environment, as in the bidirectional communication model introduced by Marko (1973). This allows us to consider both the agent and environment's point of view: How much power does the environment have to shape the signals that the agent emits, and how malleable is the agent to be shaped by the environment? These two quantities constitute the agent's plasticity and the environment's empowerment, and Proposition 4.6 shows they are identical. To build toward this finding, we develop the generalized directed information (GDI, Definition 3.1) as a suitable formalism for the concept of plasticity under minimal assumptions. The GDI is a strict generalisation of the directed information (Proposition 3.2) introduced by Massey (1990). Through a new extension of the conservation law of information (Theorem 3.5), we show that all of the information exchanged between any agent and environment can be broken into either plasticity or empowerment. This conservation unlocks our second main finding, which proves that empowerment and plasticity can be in tension (Theorem 4.8)—in other words, the design of an agent implicitly faces a dilemma with respect to plasticity and empowerment. We take this result to have significant implications for the design and analysis of agents, and suggest that further examining this dilemma is an important direction for future research. We further highlight the necessary and sufficient conditions for an agent to admit non-zero plasticity (Lemma 4.2), and exploit these conditions to prove several intuitive properties hold of plasticity (Theorem 4.3). Collectively, we take these findings to provide a rich new toolkit with which to measure and study agents under minimal assumptions. We further speculate that plasticity, empowerment, and their relationship offer a simple, powerful conception of agency: An input-output system is an agent if and only if it has non-zero plasticity and non-zero empowerment.

2 Preliminaries: Directed Information, Agents, and Empowerment

Notation. Throughout, we let caligraphic capital letters $(\mathcal{X}, \mathcal{Y})$ denote sets, and capital italics X, Y denote discrete random variables over their respective domains \mathcal{X} and \mathcal{Y} . We let $X_{1:n} = (X_1, \ldots, X_n)$ express a sequence of n discrete random variables, and let $[a:b]_n$ denote an inclusive interval of integers such that $1 \le a \le b \le n$. Lastly, we let $\Delta(\mathcal{X})$ denote the probability simplex over the set \mathcal{X} . That is, the function $p: \mathcal{X} \to \Delta(\mathcal{Y})$ expresses a probability mass function $p(\cdot \mid x)$, over \mathcal{Y} , for each $x \in \mathcal{X}$. All notation and core definitions are summarized in Table 1 in Appendix A.

2.1 Directed Information

We assume the basic definitions of information theory (Shannon, 1948) are known to the reader, such as the Shannon entropy $\mathbb{H}(X)$ of a discrete random variable X supported on the set \mathcal{X} , and the mutual information $\mathbb{I}(X;Y)$ between two discrete random variables X and Y supported on \mathcal{X} and \mathcal{Y} respectively. We will, however, make heavy use of the *directed information*, developed by Massey (1990) and expanded on by Kramer (1998), and Massey and Massey (2005), defined as follows.

Definition 2.1. For any two sequences of discrete random variables, $X_{1:n} = (X_1, ..., X_n)$ and $Y_{1:n} = (Y_1, ..., Y_n)$ the **directed information** is

$$\mathbb{I}(X_{1:n} \to Y_{1:n}) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \mathbb{I}(X_{1:i}; Y_i \mid Y_{1:i-1}). \tag{1}$$

The directed information is designed to capture the influence that the past elements X_i have on future elements Y_i . This notation assumes each X_i precedes and is a possible cause of Y_i , though if we want to consider indexing where Y_i precedes X_i we make use of the \hookrightarrow in our notation,

$$\mathbb{I}(Y_{1:n} \hookrightarrow X_{1:n}) \stackrel{\text{def}}{=} \sum_{i=2}^{n} \mathbb{I}(Y_{1:i-1}; X_i \mid X_{1:i-1}). \tag{2}$$

Directed information is particularly useful for analyzing systems with *feedback*—as is clearly present in the case of the information exchange between agent and environment in reinforcement learning. In fact, directed information was originally proposed in the context of the bidirectional communication theory developed by Marko (1973). We recall two facts about the directed information that we will later exploit and generalize. First, directed information is non-negative and upper-bounded.

Proposition 2.2 (Adapted from Theorem 2 of Massey, 1990 and Property 3.2 of Kramer, 1998). For any two sequences of discrete random variables $X_{1:n} = (X_1, ..., X_n)$, $Y_{1:n} = (Y_1, ..., Y_n)$,

$$0 \le \mathbb{I}(X_{1:n} \to Y_{1:n}) \le \mathbb{I}(X_{1:n}; Y_{1:n}),\tag{3}$$

where equality holds in the left inequality if and only if $\mathbb{I}(X_{1:i}; Y_i \mid Y_{1:i-1}) = 0$ and in the right inequality if and only if $\mathbb{H}(Y_i \mid Y_{1:i-1}, X_{1:i}) = \mathbb{H}(Y_i \mid Y_{1:i-1}, X_{1:n})$, for all i = 1, 2, ..., n.

The conditions for equality on the upper bound indicate that the directed information is equal to the mutual information when there is only communication flowing in one direction. We next confirm this intuition by recalling one of the central facts about directed information proven by Massey and Massey (2005)—that directed information is *conserved* in a particular sense.

Proposition 2.3 (Conservation Law of Directed Information; adapted from Proposition 2 of Massey and Massey, 2005). *For any pair* $(X_{1:n}, Y_{1:n})$,

$$\mathbb{I}(X_{1:n}; Y_{1:n}) = \mathbb{I}(X_{1:n} \to Y_{1:n}) + \mathbb{I}(Y_{1:n} \hookrightarrow X_{1:n}). \tag{4}$$

where
$$\mathbb{I}(Y_{1:n} \hookrightarrow X_{1:n}) = \sum_{i=2}^{n} \mathbb{I}(Y_{1:i-1}; X_i \mid X_{1:i-1}).$$

That is, the amount that $X_{1:n}$ influences $Y_{1:n}$ together with the amount that $Y_{1:n}$ influences $X_{1:n}$ is exactly the total mutual information between $X_{1:n}$ and $Y_{1:n}$. This highlights why the presence of bidirectional feedback forces $\mathbb{I}(X_{1:n} \to Y_{1:n})$ to be strictly less than $\mathbb{I}(X_{1:n}; Y_{1:n})$ —notice that if $\mathbb{I}(Y_{1:n} \hookrightarrow X_{1:n}) = 0$, then we arrive at the identity $\mathbb{I}(X_{1:n}; Y_{1:n}) = \mathbb{I}(X_{1:n} \to Y_{1:n})$.

2.2 Agents

Alongside the tools of information theory, the primary objects of our study are *agents* and the *environments* they interact with. Inspired by the work on general RL (Hutter, 2004; Hutter et al., 2024) and following the recent trend to move beyond the Markov property (Dong et al., 2022; Lu et al., 2023; Bowling et al., 2023; Abel et al., 2023a, 2024; Bowling and Elelimy, 2025), we embrace a perspective on the agent-environment interaction that is intentionally light on assumptions. Our starting point is to define a pair of finite sets that characterize the possible signals exchanged between agent and environment (Dong et al., 2022).

Definition 2.4. The interface is a pair of finite sets, (A, \mathcal{O}) such that $|A| \geq 2$, $|\mathcal{O}| \geq 2$.

We refer to each element $a \in \mathcal{A}$ as an action, and each element $o \in \mathcal{O}$ as an observation. Together, these two sets form the space of possible interactions between any agent and environment that share an interface. Following the conventions of Hutter (2000); Dong et al. (2022); Abel et al. (2023a) and Bowling et al. (2023), we refer to this set of possible interactions of any length as *histories*. To accommodate the case where the environment emits the first symbol *or* the agent emits the first symbol, we distinguish between sequences that end with an action and end with an observation.

Definition 2.5. The histories, histories ending in observation, and histories ending in action, that arise from interface (A, \mathcal{O}) are defined as the sets

$$\mathcal{H} \stackrel{\text{def}}{=} \mathcal{H}_{...A} \cup \mathcal{H}_{...O}, \tag{5}$$

$$\mathcal{H}_{...A} \stackrel{\text{def}}{=} (\mathcal{O} \times \mathcal{A})^* \cup (\mathcal{A} \times \mathcal{O})^* \times \mathcal{A}, \tag{6}$$

$$\mathcal{H}_{..\mathcal{O}} \stackrel{\text{def}}{=} (\mathcal{A} \times \mathcal{O})^* \cup (\mathcal{O} \times \mathcal{A})^* \times \mathcal{O}. \tag{7}$$

We refer to any element $h \in \mathcal{H}$, $h \in \mathcal{H}_{...A}$, or $h \in \mathcal{H}_{...O}$ as a history.

An agent, then, is a stochastic mapping from a history to an action, as in the agent functions introduced by Russell and Subramanian (1994).

Definition 2.6. An agent relative to interface (A, \mathcal{O}) is a function, $\lambda : \mathcal{H}_{...\mathcal{O}} \to \Delta(A)$.

Lastly, environments are stochastic functions that produce observations for an agent.

Definition 2.7. An environment relative to interface (A, \mathcal{O}) is a function, $e: \mathcal{H}_{...A} \to \Delta(\mathcal{O})$.

We let $\Lambda_{\rm all}$ and $\mathcal{E}_{\rm all}$ denote the set of all such functions λ and e defined over interface $(\mathcal{A},\mathcal{O})$ respectively, and let $\Lambda \subseteq \Lambda_{\rm all}$ and $\mathcal{E} \subseteq \mathcal{E}_{\rm all}$ denote non-empty subsets of $\Lambda_{\rm all}$ or $\mathcal{E}_{\rm all}$.

Together, an agent λ and environment e that share an interface interact by exchanging observation and action signals, inducing the stochastic process $A_1O_1, A_2O_2 \dots A_nO_n \dots$ that can carry on indefinitely. We assume that actions are emitted first, but this is purely for convenience and the decision is ultimately benign. Under this framing, it is useful to call attention to a symmetry between agent and environment.

Remark 2.8. Agents and environments, as defined, are symmetric in the sense that swapping sets \mathcal{A}, \mathcal{O} converts every agent into an environment, and every environment into an agent. To see this, notice that an environment, $e:\mathcal{H}_{...A}\to\Delta(\mathcal{O})$ becomes a function from $\mathcal{H}_{...O}$ to $\Delta(\mathcal{A})$, which is precisely the definition of an agent.

2.3 Empowerment

Empowerment was first introduced by Klyubin et al. (2005b) as "how much control or influence an [agent] has" (p. 2). More precisely, Klyubin et al. focus on the *open-loop agents* represented by a probability mass function $p(a^n) \in \Lambda_{a^n}$ over the sequence of n discrete random variables, $A_{1:n} = (A_1, A_2, \ldots, A_n)$. Then, the empowerment of an agent is the maximum mutual information between this sequence of actions, $A_{1:n}$, and the observation produced directly following the sequence, O_n , as follows.

$$\mathfrak{E}_{K}^{n}(\Lambda_{a^{n}}, e) = \max_{p(a^{n}) \in \Lambda_{a^{n}}} \mathbb{I}(A_{1:n}; O_{n}).$$

$$(8)$$

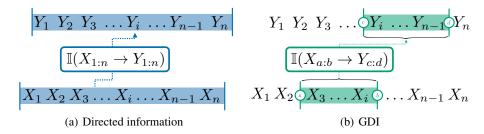


Figure 2: A visual of the directed information (left) and generalized directed information (right). In the case of GDI, the intervals can fully or partially overlap, or not overlap at all. When the intervals are identical, the two quantities are identical (Proposition 3.2).

Empowerment as Directed Information. Capdepuy (2011) build on Klyubin's definition of empowerment by allowing the agent and the environment to *both* communicate with one another. Thus, under Capdepuy's definition, the agents in question can be closed-loop, which modifies the definition of empowerment in two ways. First, the max is instead taken with respect to a richer class of functions such as $\Lambda_{\rm all}$, compared to Λ_{a^n} . Second, the mutual information is replaced by the direction information, as follows,

$$\mathfrak{E}_{\mathcal{C}}^{n}(\Lambda, e) = \max_{\lambda \in \Lambda} \mathbb{I}(A_{1:n} \to O_{1:n}). \tag{9}$$

In light of its generality, we take the directed information view of empowerment to be the appropriate starting point.

3 Generalized Directed Information

A significant limitation of directed information is that it requires both sequences $X_{1:n}$ and $Y_{1:n}$ to be of the same length, and start from the beginning of time. Thus, the existing measures do not allow us to measure the empowerment from one sequence of arbitrary length to another sequence of arbitrary length. To remedy this, we here develop a new measure that extends the directed information to arbitrary sequences $X_{a:b}$ and $Y_{c:d}$ that captures directed information as a special case (Proposition 3.2). We define this new measure as follows.

Definition 3.1. Let $X_{1:n}, Y_{1:n}$ be sequences of discrete random variables and let $[a:b]_n$ and $[c:d]_n$ be valid intervals. The **generalized directed information** (GDI) is defined as:

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) \stackrel{\text{def}}{=} \sum_{i=\max(a,c)}^{d} \mathbb{I}(X_{a:\min(b,i)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}). \tag{10}$$

As with directed information, this notation assumes X_i precedes and is a possible cause of Y_i . If we want to consider the indexing such that Y_i precedes X_i we again use \hookrightarrow in our notation (Equation 2), $\mathbb{I}(X_{a:b} \hookrightarrow Y_{c:d}) \stackrel{\text{def}}{=} \sum_{i=\max(a+1,c)}^d \mathbb{I}(X_{a:\min(b,i-1)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}).$

To ensure we capture the spirit of the original directed information, we verify that several properties hold of our generalized definition. Namely, that (1) when the sequences $X_{a:b}$ and $Y_{c:d}$ are the same length and both begin at the start of time, we recover directed information, and (2) the basic properties of directed information such as the conservation law are maintained. We first note that when a=c=1 and b=d=n, we recover the directed information, as desired.

Proposition 3.2 (GDI Strictly Generalizes DI). Let a = c = 1 and b = d = n. Then,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{1:n} \to Y_{1:n}). \tag{11}$$

We defer all proofs to Appendix B. Next, we show that if the first sequence comes *after* the second sequence, then the GDI is zero.

Proposition 3.3 (Temporal Consistency of GDI). If a > d then $\mathbb{I}(X_{a:b} \to Y_{c:d}) = 0$.

Thus, the quantity is only non-zero when at least some of the left-hand sequence of random variables precedes at least some of the right-hand sequence of random variables, as we would expect.

We next show that we can naturally subdivide the intervals along which we measure the GDI.

Proposition 3.4 (Summation of Intervals). Let $a \le k < b$ and $c \le \ell < d$.

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{a:b} \to Y_{c:\ell}) + \mathbb{I}(X_{a:b} \to Y_{\ell+1:d})$$
(12)

$$= \mathbb{I}(X_{a:k} \to Y_{c:d}) + \mathbb{I}(X_{k+1:b} \to Y_{c:d}). \tag{13}$$

Since each term on the right is also just a GDI term, we can thus decompose any pair of sequences of random variables into arbitrarily many valid sub-sequences, and preserve GDI.

We now prove the most significant property of the new measure: The GDI exhibits a conservation law similar to that of the directed information (Proposition 2.3). This conservation is especially important for our discussion of plasticity and empowerment, as it is critical for understanding their tension.

Theorem 3.5 (Conservation Law of GDI).

$$\mathbb{I}(X_{a:b}; Y_{c:d} \mid X_{1:a-1}, Y_{1:c-1}) = \mathbb{I}(X_{a:b} \to Y_{c:d}) + \mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b}). \tag{14}$$

Note that the sum of the two GDI terms is a *conditional* mutual information term. There are two perspectives that help to understand this conditioning. First, notice that when a=1 and c=1, the terms $X_{1:0}$ and $Y_{1:0}$ disappear, and we recover the same term as in the original conservation law (Proposition 2.3). Second, conditioning on past information allows us to remove possible confounders. For example, imagine $X_1=0$, and this emission ensures that both $X_3=0$ and $Y_4=0$. Now, if we measure $\mathbb{I}(X_{2:4} \to Y_{3:5})$, the total information flowing from X_3 to Y_4 needs to remove the potential information from X_1 —this is the reason for the conditioning.

We prove several other properties of the GDI including an extension of the data-processing inequality (Derpich and Østergaard, 2021), but defer these results to Appendix C.

4 Plasticity and Empowerment as Generalized Directed Information

We now motivate the GDI as a definition for plasticity. We start by enumerating a set of desiderata that characterize an account of what a suitable definition of plasticity might look like.

Our desiderata come in two forms. The first set offer a *structural* picture of the definition, while the second set focus on the *semantics* of the definition. While we hope that these desiderata are illuminating, we also note that they fall short of a complete axiomatic characterization of plasticity. In this sense, we anticipate that subtle aspects of the definition may need slight adjustment as we hone in on the right axioms for the concept (Lakatos, 1963). We suggest that identifying such a bedrock is an important direction for further research, but is here out of scope.

Structural Desiderata. A definition of agent plasticity should adhere to the following four structural criteria. First, the definition should be mathematically precise. That is, we should be able to codify all relevant commitments about our definition into a simple set of mathematical terms. Second, the definition should be simple to state, and only require minimal conceptual overhead in terms of new formalism. Third, the definition should not impose additional assumptions on our agent and environment setup beyond those already enumerated: Agent and environment share an interface consisting of two finite sets and interact in discrete time by exchanging signals from their interface. Fourth, the definition should provide flexibility to ground discussions of plasticity under a variety of settings. That is, the plasticity measure should be consistent with respect to arbitrary windows of time, or with or without a specific environment present. It is critical that we can measure the plasticity of an agent throughout its learning process, and not just the interval [1:n]. This fourth structural desiderata is what motivates the development of the GDI over regular directed information, as without it we cannot speak about plasticity over arbitrary intervals of experience.

Semantic Desiderata. Next, we identify desiderata on our definition that indicate what the definition should *mean*. These are inspired by a mixture of reflections from the broader communities that have studied plasticity, and a handful of thought experiments that stretch intuitions about plasticity. First, plasticity is always a non-negative scalar value. Second, agents that are unaffected by what they

observe have zero plasticity. For instance, a constant agent that always outputs the same action should have zero plasticity. Lastly, agents whose behavior is influenced by past observations have non-zero plasticity that is proportional to this influence.

4.1 Plasticity

These desiderata motivate the GDI as our definition of plasticity. Intuitively, the definition captures how much a sequence of observations influences a sequence of the agent's actions. The GDI allows us to select which observation sequence $O_{a:b}$ and action sequence $A_{c:d}$ are of interest, as follows.

Definition 4.1. The plasticity of agent λ relative to \mathcal{E} from $[a:b]_n$ to $[c:d]_n$ is defined as

$$\mathfrak{P}_{a:b}^{a:b}(\lambda,\mathcal{E}) \stackrel{\text{def}}{=} \max_{e \in \mathcal{E}} \mathbb{I}(O_{a:b} \to A_{c:d}). \tag{15}$$

For brevity, we will most often refer to an agent's plasticity relative to a set of environments as $\mathfrak{P}(\lambda,\mathcal{E})$ and obscure dependence on the interval when otherwise clear from context. Additionally, while we introduce the definition in its general form, we most often consider the plasticity of an agent interacting with a specific environment, e, which is recovered as a special case involving a singleton set, $\mathfrak{P}(\lambda,\{e\})$. We refer to such cases as $\mathfrak{P}(\lambda)$ or $\mathfrak{P}(\lambda,e)$ for brevity. This is similar to the distinction pointed out by Capdepuy (2011) on actual versus potential information flow: "It is important to understand that empowerment, because it is defined as a capacity, is a potential quantity. The actual controller of the agent is not considered." (p. 44). We embrace this distinction for plasticity as well, and modulate the set \mathcal{E} to refer to either instantiated or potential plasticity. Lastly, we note that the above definition uses the convention where, for every $i=1,\ldots,n$, each A_i precedes each O_i , but where needed, we can also make use of the \hookrightarrow notation.

GDI and the Desiderata. We next confirm that the GDI satisfies our presented structural and semantic desiderata. For the first two structural desiderata, we call attention to the fact that the quantity $\mathfrak{P}(\lambda,\mathcal{E})$ is mathematically precise, and requires minimal overhead (albeit subjectively): While the GDI is new, it provides a necessary enrichment of the directed information to account for (1) sequences of arbitrary length, (2) that may not start at the beginning of time, and (3) involve systems with bidirectional communication. These properties are critical to satisfy our remaining structural desiderata, and thus we believe this slight added complexity is earned. Third, our definition imposes no further assumptions on agent or environment. We inherit the general setting of an agent interacting with an environment in discrete time, and add no further restrictions. Lastly, as noted, our fourth structural desiderata is accounted for in virtue of the GDI: We are now able to range over arbitrary sequences, and can accommodate the cases when a specific environment e, a set e, or the set of all environments e and can accommodate the cases when a specific environment e and that isolates the necessary and sufficient conditions for plasticity relative to a specific environment to be non-zero.

Lemma 4.2 (Necessary and Sufficient Conditions for Positive Plasticity). *An agent has positive plasticity relative to e if and only if there is a time-step where actions are influenced by observation.*

More formally, for any pair (λ, e) and indices $[a:b]_n$, $[c:d]_n$, there exists an $i \in [\max(a,c):d]$ such that

$$\mathbb{H}(A_i \mid O_{1:a-1}, A_{1:i-1}) > \mathbb{H}(A_i \mid O_{a:\min(b,i)}, O_{1:a-1}, A_{1:i-1})$$
(16)

if and only if $\mathfrak{P}_{a:b\atop a:d}(\lambda,e)>0$.

Notice that the only difference between the left and right side of the inequality is that the right entropy term also conditions on $O_{a:\min(b,i)}$ —consequently, the inequality holds when learning about the observations from $[a:\min(b,i)]$ lowers uncertainty about A_i . As a result, a pair (λ,e) will have non-zero plasticity from interval $[a:b]_n$ to $[c:d]_n$ if and only if the observations in $[a:b]_n$ provides information about the actions in $[c:d]_n$. We make heavy use of this lemma to elucidate the following general properties of plasticity.

Theorem 4.3. The following properties hold of the function \mathfrak{P} , for all intervals $[a:b]_n$, $[c:d]_n$:

- (i) For all (λ, \mathcal{E}) , $\mathfrak{P}_{a:b\atop a:d}(\lambda, \mathcal{E}) \geq 0$.
- (ii) For any a < d, there exists a pair (λ, \mathcal{E}) where λ is deterministic and $\mathfrak{P}_{a:b}(\lambda, \mathcal{E}) > 0$.

- (iii) For all $\mathcal{E}_{\text{small}} \subseteq \mathcal{E}_{\text{big}}$, $\mathfrak{P}_{\text{add}}^{a:b}(\lambda, \mathcal{E}_{\text{small}}) \leq \mathfrak{P}_{\text{add}}^{a:b}(\lambda, \mathcal{E}_{\text{big}})$.
- (iv) (Informal) The following agents all have zero plasticity relative to every environment set: Open-loop agents, agents that always output the same action distribution, agents whose actions only depend on history length, and agents whose actions only depend on past actions.

Together, this theorem and lemma indicate that the semantic desiderata are satisfied from a number of perspectives. By the necessary and sufficient conditions of Lemma 4.2, we find an intuitive test for whether an agent has plasticity in a given window: Are its actions impacted by its observations? The *magnitude* of this impact directly determines the magnitude of the agent's plasticity, aligned with the fourth desiderata. The lemma also indicates that many popular learning algorithms such as UCB1 (Auer et al., 2002) or Q-learning (Watkins and Dayan, 1992) have non-zero plasticity in most environments, as expected. Theorem 4.3 illustrates basic properties of plasticity: Point (i) confirms that plasticity is a non-negative scalar, while point (ii) shows that a deterministic agent can be plastic. Point (iii) highlights a form of monotonicity to plasticity—as the environment set under consideration is enriched, an agent's plasticity cannot go down. Lastly, property (iv) shows that many trivial agents have zero plasticity, such as agents whose actions are determined solely by factors *other* than what they observe such as open-loop or constant agents.

4.2 Revisiting Empowerment through GDI

In addition to capturing plasticity, the GDI can enrich our definition of empowerment as well. As discussed in our exposition of empowerment, the definition of Capdepuy (Equation 9) moves from the mutual information to the directed information. Due to the generality of the GDI, empowerment can further benefit from the flexibility of the GDI as follows.

Definition 4.4. The **empowerment** of Λ in environment e from $[a:b]_n$ to $[c:d]_n$ is defined as

$$\mathfrak{E}_{a:b\atop c:d}(\Lambda, e) \stackrel{\text{def}}{=} \max_{\lambda \in \Lambda} \mathbb{I}(A_{a:b} \to O_{c:d}). \tag{17}$$

As we might hope, when 1 = a = c and b = d = n, we recover the definition from Capdepuy, and when we restrict to the set of open loop agents Λ_{a^n} , we recover the definition from Klyubin et al.

Proposition 4.5.
$$\mathfrak{E}_{1:n}^{1:n}(\Lambda, e) = \mathfrak{E}_{C}^{n}(\Lambda, e)$$
 and $\mathfrak{E}_{1:n}^{1:n}(\Lambda_{a^{n}}, e) = \mathfrak{E}_{K}^{n}(\Lambda_{a^{n}}, e)$.

Thus, we can specialize to either past definition as desired. However, empowerment now benefits from the flexibility of the GDI to refer to an agent's empowerment over arbitrary windows of experience that need to begin at the start of time. Furthermore, we can recover a meaningful notion of the empowerment of a *single agent* by considering a singleton set, $\mathfrak{E}(\{\lambda\}, e)$, as with plasticity. We will again refer to this simply by the notation $\mathfrak{E}(\lambda)$ or $\mathfrak{E}(\lambda, e)$ for brevity when clear from context.

4.3 Analysis: Plasticity and Empowerment

We next analyze the relationship between plasticity and empowerment. First, we highlight the fact that plasticity and empowerment are really two sides of the same coin. The key is to observe that the environment and the agent, as defined, are in fact symmetric objects: They are each a stochastic function from a history to a symbol. The names of the symbols and the functions are so far arbitrary. As a result, we can also ask about the *empowerment of the environment*, or the *plasticity of the environment* simply by taking the environment-centric view point suggested by Remark 2.8.

Proposition 4.6 (Plasticity is the Mirror of Empowerment). Given any pair (λ, e) that share an interface, and valid intervals $[a:b]_n$, $[c:d]_n$, the agent's empowerment is equal to the plasticity of the environment, and the agent's plasticity is equal to the empowerment of the environment:

$$\mathfrak{E}_{a:b}(\lambda) = \mathfrak{P}_{a:b}(e), \qquad \mathfrak{P}_{a:b}(\lambda) = \mathfrak{E}_{a:b}(e).$$
 (18)

As a corollary, we find that despite our best efforts to design an agent with high plasticity or empowerment, some environments force *all* agents they interact with to have zero of either quantity.

Corollary 4.7. For all
$$(\lambda, e)$$
, $\mathfrak{P}(e) = 0 \iff \mathfrak{E}(\lambda) = 0$ and $\mathfrak{E}(e) = 0 \iff \mathfrak{P}(\lambda) = 0$.

In other words, even the most pliable or powerful agents cannot do anything meaningful in specific environments; the environment can always just choose to ignore the agent's actions, which places a firm limit on how empowered an agent can be. Similarly, even if we design a highly plastic agent, the agent can only make use of that plasticity in an environment with information.

Example: Two-Player Game of Poker. To build intuition, consider the case where the environment is itself another agent. That is, two agents—Alice and Bob—interact by exchanging actions. The interface is defined by two sets, \mathcal{A}_{alice} , \mathcal{A}_{bob} , where the actions of Alice are Bob's observations, and vice-versa. In this case, over any choice of intervals, *Alice's plasticity is identical to Bob's empowerment* and vice-versa. Imagine further that Alice and Bob are playing poker against each other, with some initial pool of chips. If either agent were to run out of chips, they are removed from play. Notice that the set of available actions to each agent is a strictly increasing function of their available pool of chips, as the larger pile of chips yields more bets that can be placed. Therefore, the more chips Alice or Bob has, the more empowered they are. And, by contrast, if Alice has more chips, this raises how plastic Bob can be, as he has a higher potential to be influenced by Alice's actions.

Our next result highlights a fundamental tension between plasticity and empowerment. That is, if we consider a specific agent λ interacting with an environment e, an important question arises: Is there any connection between the agent's level of plasticity, and the agent's level of empowerment? We answer this question by establishing a tight upper bound on the sum of the two quantities, which establishes a tension between an agent's plasticity and its empowerment over the same interval.

Theorem 4.8 (Plasticity-Empowerment Tension). For all intervals $[a:b]_n$, $[c:d]_n$ and any pair (λ, e) , let $m = \min\{(b-a)\log |\mathcal{O}|, (d-c)\log |\mathcal{A}|\}$. Then,

$$\mathfrak{E}_{a:b}(\lambda, e) + \mathfrak{P}_{c:d}(\lambda, e) \le m, \tag{19}$$

is a tight upper bound on empowerment and plasticity. Furthermore, under mild assumptions on the interface and interval, there exists a pair $(\lambda^{\diamond}, e^{\diamond})$ such that $\mathfrak{P}_{\stackrel{\text{cid}}{a:b}}(\lambda^{\diamond}, e^{\diamond}) = m$, and there exists a pair (λ', e') such that $\mathfrak{E}_{\stackrel{\text{cid}}{a:b}}(\lambda', e') = m$, thereby forcing the other quantity to be zero.

Therefore, an agent cannot simultaneously maximize its plasticity and its empowerment. In fact, the achievable levels of empowerment are strictly determined by the agent's plasticity (and vice versa). The latter part of the result indicates that in extreme cases, a high plasticity can ensure that the agent has zero empowerment, and vice versa. Notably, the tension only comes about when we consider $\mathfrak{P}_{a;d}^{c:d}$ and $\mathfrak{E}_{a;d}^{a:b}$ over the same pair of intervals, $[a:b]_n, [c:d]_n$. Therefore, this tension may dissipate as the intervals change; in this way, agents may oscillate between phases of high plasticity and high empowerment, but cannot simultaneously maximize both. A deeper examination of this trade-off and its implications for the design of agents is a significant opportunity for future research.

5 Discussion

In this paper, we have presented a careful study of the concepts of plasticity and empowerment, and established a connection between them. Our first technical contribution is the development of the generalized directed information (Definition 3.1), which we take to be of independent interest across the many fields that make use of information theory. We identify several fundamental properties of the GDI, including a new conservation law (Theorem 3.5). We then motivate the GDI as an appropriate definition for the concept of plasticity, and provide the first universal definition of plasticity. We then identify the necessary and sufficient conditions for an agent to have non-zero plasticity (Lemma 4.2). Our analysis further reveals that plasticity admits a number of intuitive properties (Theorem 4.3), and that the GDI generalizes existing notions of empowerment (Proposition 4.5). We then reveal two general facts about plasticity and empowerment: First, an agent's plasticity is identical to the environment's empowerment and vice-versa (Proposition 4.6); Second, that plasticity and empowerment pose a new trade-off (Theorem 4.8), suggesting that we need to thoughtfully engage with both concepts in the study and design of agents. We take these results to serve as a meaningful bedrock from which we can study agency under minimal assumptions. We speculate that an input-output system is best thought of as an agent if and only if it has non-zero plasticity and non-zero empowerment. We defer a full conceptual analysis of this sort for future work, but refer the reader to the works by Wooldridge and Jennings (1995), Barandiaran et al. (2009), Kasirzadeh and Gabriel (2025) for similar discussions.

5.1 Plasticity via Actions, Policies, or Agent States?

As discussed, the original conception of plasticity from the neuroscience literature is terms of *neural* or *synaptic* plasticity—that is, the degree of flexibility inherent to some neural circuitry internal to a biological organism. By contrast, our conception of plasticity focuses purely on *behavior*: We

measure the influence of observation on action. A simple variation of our definition can instead consider the degree that observations influence other aspects of an agent, such as the agent's internal parameters, memory policy or agent state more generally. We can choose from a wide variety of these options, and the formalism (and the corresponding theoretical results) will largely look the same. In the case of policies, we can the view that every agent actively maintains a policy that belongs to a finite class of representable policies, $\Lambda_{\rm B} \subset \Lambda_{\rm all}$ (the "agent basis" from Abel et al., 2023a). Then, letting $L_1, \ldots L_n$ denote a sequence of discrete random variables over Λ_B , we can express a policy-focused notion of plasticity in terms of the quantity $\mathbb{I}(O_{a:b} \to L_{c:d})$. This measures how much observations impact an agent's choice of policy, rather than a choice of action. Or, we can shift the influence to agent state. In this case, we suppose the agent maintains an internal summary of its history known as agent state (Dong et al., 2022; Abel et al., 2023b; Kumar et al., 2023). Letting S_1, \ldots, S_n denote a sequence of random variables over this agent state space S, we can further expand our conception of plasticity to focus on agent states by building around $\mathbb{I}(O_{a:b} \to S_{c:d})$. We believe this view is especially promising to examine agency under resource constraints (Simon, 1955; Russell and Subramanian, 1994), the extended mind (Clark and Chalmers, 1998; Harutyunyan, 2020), and the stability-plasticity dilemma (Carpenter and Grossberg, 1988), but leave exploration of these directions for future work. We believe our conception of plasticity accommodates a clean account of the stability-plasticity dilemma (and, as such, a natural definition of stability), though properly telling this full story is out of scope for the present paper.

5.2 Optionality, Information, and Causality

Both empowerment and plasticity center around how the agent or environment could have acted or unfolded differently. In the case of empowerment, we ask: Had the agent acted differently, how much could have changed in the environment? For plasticity: If the environment had unfolded differently, how much would that have impacted what the agent did? In this way, both concepts implicitly deal with considerations regarding free will, counterfactuals, causality, and determinism. For example, empowerment is defined by maximizing over the set Λ , reflecting the different ways an agent could respond to stimuli. Constraining Λ —for instance, by restricting actions, compute, or memory—will in turn reduce the agent's ability to influence future observations and, consequently, its empowerment. In the most general case we can take the max over all agents, $\Lambda_{\rm all}$, reflecting every possibly way experience could give rise to action over a given fixed interface. However, often times there are further restrictions in place, such as the morphology of a robot or a limited available computational budget such restrictions can be reflected by considering different subsets $\Lambda \subset \Lambda_{\rm all}$, thereby encoding the different possible ways the agent could act. Capdepuy (2011) discusses this distinction in Section 4.1.3, noting that *potential* information flow involves taking the max over available policies, while actual information flow is relative to a single policy. Similarly, plasticity (Definition 4.1) involves maximization over a set of environments \mathcal{E} . This choice raises the question: What constitutes environment optionality? How could an environment unfold differently? Interventions, as described by (Pearl, 2009), offer a formal framework for grounding the notion of environment optionality from a causal perspective. Here, an external experimenter modifies the environment and observes the responses of its sub-systems (such as the agent). This aligns with the continual learning literature, where plasticity is treated as the agent's capacity to adapt to interventions on the data-generating process. For example, in (Dohare et al., 2021) the input distribution is repeatedly and randomly intervened on, and in (Dohare et al., 2024) the labels are randomly permuted, whereas in (Abbas et al., 2023) the agent is switched between different Atari games. These distributional shifts can be formalized as interventions t on a baseline environment e, where each intervention creates a new, distributionally shifted environment e' = t(e). In each study there is an assumed set of interventions \mathcal{T} (such as label permutations), which generates the environment set $\mathcal{E} = \{t_i(e) \mid t_i \in \mathcal{T}\}$. In essence, an agent's plasticity is its capacity to adapt to environmental changes (interventions), defined relative to a set of environments \mathcal{E} generated by these interventions. And, analogous to constraining the policy set for empowerment, plasticity is non-increasing if the set of possible environmental changes is constrained (as in point iv of Theorem 4.3). We suggest that drawing out a full causal picture of both plasticity and empowerment is a compelling direction for further work, building on

As a consequence of the definition that focuses on observation and action, we find that every deterministic environment forces zero plasticity. That is, if e is deterministic in the sense that it can be written as a map, $e: \mathcal{H}_{...A} \to \mathcal{O}$, then for every agent and intervals, $\mathfrak{P}_{\frac{a:b}{c:d}}(\lambda, e) = 0$. This may appear puzzling at first, but on closer inspection, there are several intuitions to elucidate this fact. First,

from the agent's point of view, a deterministic environment presents no optionality. This is similar to a causal conception of plasticity relative to the empty set of interventions. Since the environment *could not have been otherwise*, there is no way that the agent can react to the different unfoldings of the environment. However, a variant of plasticity that builds around the internal components of the agent (such as the agent's beliefs or memories) will overcome this in a simple way. Consider an agent with limited memory. That is, at most the agent can remember five observations. Then, if the agent is at time-step seven, it cannot possibly remember all prior observations due to its memory constraints. Therefore, from the agent's memory-constrained point of view, the environment may appear stochastic. While these threads play an important conceptual role in plasticity and empowerment, we defer a full exploration of their significance to future work.

5.3 The Plasticity-Empowerment Tension

Lastly, we present a short example expanding on the discussion of the tension between plasticity and empowerment highlighted by Theorem 4.8. Consider the example environment depicted in Figure 3. The environment contains n+1 rooms, each containing a switch and a light. The environment is conditioned on a latent Bernoulli random variable, X, corresponding to the probability to turn the light on or off each time step. The mouse only observes whether the light is on or off in the room they currently occupy. The action space of the mouse enables it to move to a neighboring room, or pull the lever. In each room ℓ_i a pull of the lever has a probability of i/n to change the state of the light for the next time step, with the remaining probability n - i/n determined by X—hence, in room ℓ_0 , the mouse has no control over the light, and in room ℓ_n , the mouse fully controls whether the light is on or off. Otherwise, the random variable X dictates the state of the light. Suppose X is a fair coin, that is, $X \sim \text{Bernoulli}(\theta = 0.5)$. So, without intervention, the lights in all rooms switch state with probability 0.5. In this environment, the far left room maximizes plasticity, whereas the far right room maximizes empowerment, with all rooms in the middle smoothly interpolating between these two extremes. On the left, the mouse gets to receive the most new information from the environment, while on the right, the mouse fully controls the state of the light. Theorem 4.8 suggests that this trade-off is always present to some extent.

To emphasize the tension even further, we can imagine that the latent random variable X is non-stationary. That is, consider the sequence of smoothly evolving Bernoulli random variables, X_1, X_2, \ldots Here, again, if the agent always inhabits the far right room, it fully closes off the possibility of learning about the evolution of $\{X_i\}_{i=1}^{\infty}$. However, if the agent is in the far left room, then the light is fully determined by the evolution of X_i , and thus the agent has maximized its potential to learn about the sequence from the environment. Each room in between these two interpolates between these two extremes, roughly corresponding to a walk along the Pareto frontier of the trade-off expressed by the theorem.

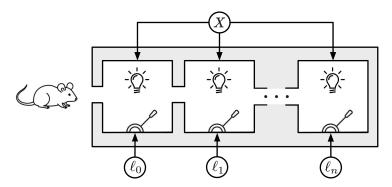


Figure 3: An example environment in which an agent (the mouse) is presented with a gradual change in the balance between their realized plasticity and empowerment. The mouse inhabits a corridor consisting of n+1 rooms, each equipped with a light that can be (stochastically) controlled by the lever in that room.

5.4 Future Work

There are many directions for future work. First, as a consequence of the conditions in Lemma 4.2 and the tension highlighted by Theorem 4.8, there are new opportunities to inform the design of agents. In particular, we speculate that it is prudent to *design agents to balance plasticity and empowerment* tactfully, since over-optimizing one can impact the other. We foresee an opportunity to recover new fundamental results connecting the limits of learning, the design of safe agents, and this balance—if an agent has insufficient plasticity, it cannot adapt, and if it has insufficient empowerment, it cannot steer. Additionally, it may be valuable to better understand how people handle the tension revealed by Theorem 4.8, similar to understanding how people address the explore-exploit dilemma. Second, we foresee a direct connection between plasticity, empowerment, and many standard concepts within game theory and multi-agent systems such as the emergence of social dilemmas, cooperation, and the dynamics of multi-agent systems more generally (Jaques et al., 2019). Lastly, a natural next step is to design efficient estimators for the GDI, and by consequence, plasticity and empowerment. Following the work by Jiao et al. (2013), we suspect that a similar algorithmic template can underlie an estimator for GDI. This estimator can unlock experimental study examining how plasticity and empowerment fair in contemporary agents as we vary different aspects of their design or the environment.

Acknowledgments

The authors are grateful to Pablo Samuel Castro, Claudia Clopath, Iason Gabriel, and Joel Z Leibo for their thoughtful comments on a draft of the paper. We would further like to thank Barel Skuratovsky and the Agency team for inspirational conversations.

References

- Z. Abbas, R. Zhao, J. Modayil, A. White, and M. C. Machado. Loss of plasticity in continual deep reinforcement learning. In *Proceedings of the Conference on Lifelong Learning Agents*, 2023.
- D. Abel, A. Barreto, B. V. Roy, D. Precup, H. van Hasselt, and S. Singh. A definition of continual reinforcement learning. In *Advances in Neural Information Processing Systems*, 2023a.
- D. Abel, A. Barreto, H. van Hasselt, D. Precup, B. Van Roy, and S. Singh. On the convergence of bounded agents. *arXiv preprint arXiv:2307.11044*, 2023b.
- D. Abel, M. K. Ho, and A. Harutyunyan. Three dogmas of reinforcement learning. *Reinforcement Learning Journal*, 1, 2024.
- P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47:235–256, 2002.
- X. E. Barandiaran, E. Di Paolo, and M. Rohde. Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, 17(5):367–386, 2009.
- K. Baumli, D. Warde-Farley, S. Hansen, and V. Mnih. Relative variational intrinsic control. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.
- M. Bowling and E. Elelimy. Rethinking the foundations for continual reinforcement learning. *arXiv* preprint arXiv:2504.08161, 2025.
- M. Bowling, J. D. Martin, D. Abel, and W. Dabney. Settling the reward hypothesis. In *Proceedings of the International Conference on Machine Learning*, 2023.
- F. Brändle, L. J. Stocks, J. B. Tenenbaum, S. J. Gershman, and E. Schulz. Empowerment contributes to exploration behaviour in a creative video game. *Nature Human Behaviour*, 7(9):1481–1489, 2023.
- V. Campos, A. Trott, C. Xiong, R. Socher, X. Giró-i Nieto, and J. Torres. Explore, discover and learn: Unsupervised discovery of state-covering skills. In *Proceedings of the International Conference on Machine Learning*, 2020.
- P. Capdepuy. *Informational principles of perception-action loops and collective behaviours*. PhD thesis, University of Hertfordshire, 2011.

- G. A. Carpenter and S. Grossberg. ART 2: Self-organization of stable category recognition codes for analog input patterns. *Applied optics*, 26(23):4919–4930, 1987a.
- G. A. Carpenter and S. Grossberg. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer vision, graphics, and image processing*, 37(1):54–115, 1987b.
- G. A. Carpenter and S. Grossberg. The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*, 21(3):77–88, 1988.
- Q. Chen, C. Shui, L. Han, and M. Marchand. On the stability-plasticity dilemma in continual meta-learning: Theory and algorithm. Advances in Neural Information Processing Systems, 2023.
- A. Citri and R. C. Malenka. Synaptic plasticity: multiple forms, functions, and mechanisms. *Neuropsychopharmacology*, 33(1):18–41, 2008.
- A. Clark and D. Chalmers. The extended mind. Analysis, 58(1):7–19, 1998.
- M. S. Derpich and J. Østergaard. Directed data-processing inequalities for systems with feedback. *Entropy*, 23(5):533, 2021.
- S. Dohare, R. S. Sutton, and A. R. Mahmood. Continual backprop: Stochastic gradient descent with persistent randomness. *arXiv preprint arXiv:2108.06325*, 2021.
- S. Dohare, J. F. Hernandez-Garcia, Q. Lan, P. Rahman, A. R. Mahmood, and R. S. Sutton. Loss of plasticity in deep continual learning. *Nature*, 632(8026):768–774, 2024.
- S. Dong, B. Van Roy, and Z. Zhou. Simple agent, complex environment: Efficient reinforcement learning with agent states. *Journal of Machine Learning Research*, 23(255):1–54, 2022.
- S. Fusi, P. J. Drew, and L. F. Abbott. Cascade models of synaptically stored memories. *Neuron*, 45 (4):599–611, 2005.
- K. Gregor, D. J. Rezende, and D. Wierstra. Variational intrinsic control. *arXiv preprint arXiv:1611.07507*, 2016.
- A. Harutyunyan. What is an agent? http://anna.harutyunyan.net/wp-content/uploads/2020/09/What_is_an_agent.pdf, 2020.
- M. Hutter. A theory of universal artificial intelligence based on algorithmic complexity. *arXiv* preprint cs/0004001, 2000.
- M. Hutter. *Universal artificial intelligence: Sequential decisions based on algorithmic probability*. Springer Science & Business Media, 2004.
- M. Hutter, D. Quarel, and E. Catt. *An introduction to universal artificial intelligence*. CRC Press, 2024.
- N. Jaques, A. Lazaridou, E. Hughes, C. Gulcehre, P. Ortega, D. Strouse, J. Z. Leibo, and N. De Freitas. Social influence as intrinsic motivation for multi-agent deep reinforcement learning. In *Proceedings of the International Conference on Machine Learning*, 2019.
- J. Jiao, H. H. Permuter, L. Zhao, Y.-H. Kim, and T. Weissman. Universal estimation of directed information. *IEEE Transactions on Information Theory*, 59(10):6220–6242, 2013.
- A. Kasirzadeh and I. Gabriel. Characterizing AI agents for alignment and governance. *arXiv* preprint arXiv:2504.21848, 2025.
- A. S. Klyubin, D. Polani, and C. L. Nehaniv. All else being equal be empowered. In *European Conference on Artificial Life*, pages 744–753. Springer, 2005a.
- A. S. Klyubin, D. Polani, and C. L. Nehaniv. Empowerment: A universal agent-centric measure of control. In *IEEE Congress on Evolutionary Computation*, volume 1, pages 128–135. IEEE, 2005b.
- G. Kramer. Directed information for channels with feedback, volume 11. Citeseer, 1998.

- S. Kumar, H. Marklund, A. Rao, Y. Zhu, H. J. Jeon, Y. Liu, and B. Van Roy. Continual learning as computationally constrained reinforcement learning. *arXiv* preprint arXiv:2307.04345, 2023.
- I. Lakatos. Proofs and refutations. Nelson London, 1963.
- A. Levy, S. Rammohan, A. Allievi, S. Niekum, and G. Konidaris. Hierarchical empowerment: Towards tractable empowerment-based skill learning. *arXiv preprint arXiv:2307.02728*, 2023.
- X. Lu, B. V. Roy, V. Dwaracherla, M. Ibrahimi, I. Osband, and Z. Wen. Reinforcement learning, bit by bit. *Foundations and Trends in Machine Learning*, 16(6):733–865, 2023. ISSN 1935-8237.
- C. Lyle, Z. Zheng, E. Nikishin, B. A. Pires, R. Pascanu, and W. Dabney. Understanding plasticity in neural networks. *arXiv preprint arXiv:2303.01486*, 2023.
- H. Marko. The bidirectional communication theory-a generalization of information theory. *IEEE Transactions on communications*, 21(12):1345–1351, 1973.
- J. Massey. Causality, feedback and directed information. In *Proceedings of the International Symposium on Information Theory and its Applications*, volume 2, 1990.
- J. L. Massey and P. C. Massey. Conservation of mutual and directed information. In *Proceedings*. *International Symposium on Information Theory*, 2005. ISIT 2005., pages 157–158. IEEE, 2005.
- S. Mohamed and D. Jimenez Rezende. Variational information maximisation for intrinsically motivated reinforcement learning. In *Advances in Neural Information Processing Systems*, 2015.
- J. Pearl. Causality. Cambridge university press, 2009.
- K. Raghavan and P. Balaprakash. Formalizing the generalization-forgetting trade-off in continual learning. *Advances in Neural Information Processing Systems*, 2021.
- T. J. Ringstrom. Reward Is Not Necessary: Foundations for Compositional Non-Stationary Non-Markovian Hierarchical Planning and Intrinsically Motivated Autonomous Agents. PhD thesis, University of Minnesota, 2023.
- S. J. Russell and D. Subramanian. Provably bounded-optimal agents. *Journal of Artificial Intelligence Research*, 2:575–609, 1994.
- C. Salge and D. Polani. Empowerment as replacement for the three laws of robotics. *Frontiers in Robotics and AI*, 4:260425, 2017.
- C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3): 379–423, 1948.
- H. A. Simon. A behavioral model of rational choice. *The quarterly journal of economics*, 69(1): 99–118, 1955.
- A. M. Turner, L. Smith, R. Shah, A. Critch, and P. Tadepalli. Optimal policies tend to seek power. *Advances in Neural Information Processing Systems*, 2021.
- C. J. Watkins and P. Dayan. Q-learning. Machine learning, 8(3-4):279–292, 1992.
- M. Wooldridge and N. R. Jennings. Intelligent agents: Theory and practice. *The knowledge engineering review*, 10(2):115–152, 1995.

A Notation

We first provide Table 1 summarizing all relevant notation.

Notation	Meaning	Definition
$\overline{\mathcal{X},\mathcal{Y}}$	Set	
X, Y, A, O	Discrete random variable	
$X_{1:n}$	Sequence of n random variables	(X_1,\ldots,X_n)
$[a:b]_n$	An interval of discrete time	$[a:b]$ s.t. $1 \le a \le b \le n$
$\mathcal A$	Set of actions	
\mathcal{O}	Set of observations	
$\mathcal{H}_{\mathcal{A}}$	Histories ending in action	$(\mathcal{O} \times \mathcal{A})^* \cup (\mathcal{A} \times \mathcal{O})^* \times \mathcal{A}$
$\mathcal{H}_{\mathcal{O}}$	Histories ending in observation	$(\mathcal{A} \times \mathcal{O})^* \cup (\mathcal{O} \times \mathcal{A})^* \times \mathcal{O}$
${\cal H}$	Set of all histories	$\mathcal{H} = \mathcal{H}_{A} \cup \mathcal{H}_{o}$
e	Environment	$e:\mathcal{H}_{\mathcal{A}} o \Delta(\mathcal{A})$
$\mathcal{E}_{ ext{all}}$	Set of all environments	$\Delta(\mathcal{O})^{\mathcal{H}_{\mathcal{A}}}$
${\cal E}$	Set of Environments	$\mathcal{E}\subseteq\mathcal{E}_{ ext{all}}$
λ	Agent	$\lambda:\mathcal{H}_{\mathcal{O}} ightarrow\Delta(\mathcal{A})$
$\Lambda_{ m all}$	Set of all agents	$\Delta(\mathcal{A})^{\mathcal{H}_{}\mathcal{O}}$
Λ	Set of agents	$\Lambda \subseteq \Lambda_{ m all}$
$\mathbb{I}(X_{1:n} \to Y_{1:n})$	Directed information	$\sum_{i=1:n} \mathbb{I}(X_{1:i}, Y_i Y_{1:i-1})$
$\mathbb{I}(X_{1:n} \hookrightarrow Y_{1:n})$	Directed information, X delay	$\sum_{i=2:n} \mathbb{I}(X_{1:i-1}; Y_i Y_{1:i-1})$
$\mathbb{I}(X_{a:d} \to Y_{c:d})$	GDI	$\sum_{i=\max(a,c):d} \mathbb{I}(X_{a:\min(b,i)}; Y_i X_{1:a-1}, Y_{1:i-1})$
$\mathfrak{P}_{\stackrel{a:b}{c:d}}(\lambda,\mathcal{E})$	Plasticity	$\max_{e \in \mathcal{E}} \mathbb{I}(O_{a:b} \to A_{c:d})$
$\mathfrak{P}_{rac{a:b}{c:d}}^{rac{a:b}{c:d}}(\lambda,e)$	Plasticity, single environment	$\mathfrak{P}_{a:b\atop a:d}\left(\lambda,\{e\} ight)$
$\mathfrak{P}(\lambda)$	Plasticity abbreviation	$\mathfrak{P}_{rac{a:b}{c:d}}^{rac{a:b}{c}}(\lambda,e)$
$\mathfrak{E}_{rac{a:b}{c:d}}(\Lambda,e)$	Empowerment	$\max_{\lambda \in \Lambda} \mathbb{I}(A_{a:b} \to O_{c:d})$
$\mathfrak{E}_{c:d}^{a:b}(\lambda,e)$	Empowerment, single agent	$\mathfrak{E}_{\stackrel{a:b}{c:d}}(\{\lambda\},e)$
$\mathfrak{E}(\lambda)$	Empowerment abbreviation	$\mathfrak{E}_{\stackrel{c:a}{a:d}}(\lambda,e)$

Table 1: A summary of notation.

B Proofs of Presented Results

We next provide proofs of each result from the paper, split into those results presented in Section 3 on GDI (Appendix B.1) and those results presented in Section 4 on plasticity and empowerment (Appendix B.2). Then, in Appendix C, we present additional results on the GDI, including a new data-processing inequality for the GDI.

B.1 Proofs of Results from Section 3 on GDI

Proposition 3.2 (GDI Stricly Generalizes DI). Let a = c = 1 and b = d = n. Then,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{1:n} \to Y_{1:n}). \tag{20}$$

Proof of Proposition 3.2.

The result is a straightforward consequence of the definition of GDI that can be illustrated by substitution. Let a=c and b=d be indices such that 1=a=c and b=d=n. Then, rewriting Equation 10, we find:

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \sum_{i=\max(a,c)}^{d} \mathbb{I}(X_{a:\min(b,i)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}), \tag{21}$$

$$= \sum_{i=1}^{n} \mathbb{I}(X_{1:i}; Y_i \mid X_{1:i-1}, Y_{1:i-1}), \tag{22}$$

$$= \underbrace{\sum_{i=1}^{n} \mathbb{I}(X_{1:i}; Y_i \mid Y_{1:i-1})}_{=\mathbb{I}(X_{1:n} \to Y_{1:n})}.$$

Proposition 3.3 (Temporal Consistency of GDI). If a > d then $\mathbb{I}(X_{a:b} \to Y_{c:d}) = 0$, and if $a \ge d$ then $\mathbb{I}(X_{a:b} \hookrightarrow Y_{c:d}) = 0$.

Proof of Proposition 3.3.

The property follows directly from the summation bounds of Equation 10, and its variant,

$$\mathbb{I}(X_{a:b} \hookrightarrow Y_{c:d}) \stackrel{\text{def}}{=} \sum_{i=\max(a+1,c)}^{d} \mathbb{I}(X_{a:\min(b,i-1)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}). \tag{23}$$

For the former, note that if a > d, then

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \sum_{i=\max(a,c)}^{d} \mathbb{I}(X_{a:\min(b,i)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}) = 0, \tag{24}$$

since $\max(a, c) > d$ when a > d.

For the latter, note that if $a \ge d$, then

$$\mathbb{I}(X_{a:b} \hookrightarrow Y_{c:d}) = \sum_{i=\max(a+1,c)}^{d} \mathbb{I}(X_{a:\min(b,i-1)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}) = 0,$$
 (25)

by similar reasoning that $\max(a+1,c) > d$.

Proposition 3.4 (Summation of Intervals). Let $a \le k < b$ and $c \le \ell < d$.

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{a:b} \to Y_{c:\ell}) + \mathbb{I}(X_{a:b} \to Y_{\ell+1:d})$$
(26)

$$= \mathbb{I}(X_{a:k} \to Y_{c:d}) + \mathbb{I}(X_{k+1:b} \to Y_{c:d}). \tag{27}$$

and similarly,

$$\mathbb{I}(X_{a:b} \hookrightarrow Y_{c:d}) = \mathbb{I}(X_{a:b} \hookrightarrow Y_{c:f}) + \mathbb{I}(X_{a:b} \hookrightarrow Y_{f+1:d}) \tag{28}$$

$$= \mathbb{I}(X_{a:e} \hookrightarrow Y_{c:d}) + \mathbb{I}(X_{e+1:b} \hookrightarrow Y_{c:d}), \tag{29}$$

Proof of Proposition 3.4.

The first equality of Proposition 3.4 consists of breaking the sum of Equation 10 into two sums with intervals $[\max(a, c), \ell]$ and $[\max(a, \ell + 1), d]$.

For the second equality observe that if b > d then we can change b = d as $i \not> d$ in the sum of Equation 10. Then, we can apply Equation 10, break up the first sum at k, and recombine terms, as follows:

$$\mathbb{I}(X_{a:k} \to Y_{c:d}) + \mathbb{I}(X_{k+1:b} \to Y_{c:d})$$

$$= \sum_{i=\max(a,c)}^{d} \mathbb{I}(X_{a:\min(k,i)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}) + \sum_{i=\max(k+1,c)}^{d} \mathbb{I}(X_{k+1:\min(b,i)}; Y_i \mid X_{1:k-1}, Y_{1:i-1}),$$
(31)

Then, we break the right sum into two intervals,

$$= \sum_{i=\max(a,c)}^{k} \mathbb{I}(X_{a:i}; Y_i \mid X_{1:a-1}, Y_{1:i-1}) + \sum_{i=\max(k+1,c)}^{d} \mathbb{I}(X_{a:k}; Y_i \mid X_{1:a-1}, Y_{1:i-1}) + \sum_{i=\max(k+1,c)}^{d} \mathbb{I}(X_{k+1:\min(b,i)}; Y_i \mid X_{1:k-1}, Y_{1:i-1}).$$
(32)

$$= \sum_{i=\max(a,c)}^{k} \mathbb{I}(X_{a:i}; Y_i \mid X_{1:a-1}, Y_{1:i-1}) +$$
(33)

$$\sum_{i=\max(k+1,c)}^{d} \left(\mathbb{I}(X_{a:k}; Y_i \mid X_{1:a-1}, Y_{1:i-1}) + \mathbb{I}(X_{k+1:\min(b,i)}; Y_i \mid X_{1:k-1}, Y_{1:i-1}) \right).$$

Applying the mutual information chain rule to the terms in the second sum and noticing if $i \le k$ then $i = \min(b, i)$ so we can change the bounds on the first sum to get

$$\mathbb{I}(X_{a:k} \to Y_{c:d}) + \mathbb{I}(X_{k+1:b} \to Y_{c:d}) = \sum_{i=\max(a,c)}^{k} \mathbb{I}(X_{a:\min(b,i)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}) + \sum_{i=\max(k+1,c)}^{d} \mathbb{I}(X_{a:\min(b,i)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}).$$
(34)

Now we combine the sums to get our result,

$$\mathbb{I}(X_{a:k} \to Y_{c:d}) + \mathbb{I}(X_{k+1:b} \to Y_{c:d})
= \sum_{i=\max(a,c)}^{d} \mathbb{I}(X_{a:\min(b,i)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}) = \mathbb{I}(X_{a:b} \to Y_{c:d}). \quad \Box \quad (35)$$

To prove the conservation law of GDI, we first introduce a short lemma that helps decompose the GDI into a difference of entropies.

Next, we present a further decomposition of any GDI term that will be a key step in our proof of the conservation law of GDI.

Lemma B.1. Let $[a:b]_n$, $[c:d]_n$ be valid intervals, and let $k = \max(a,c)$, $\ell = \min(b,d)$. Then,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{a:k-1}; Y_{k:\ell} \mid X_{1:a-1}, Y_{1:k-1}) + \mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) + \mathbb{I}(X_{a:\ell}; Y_{\ell+1:d} \mid X_{1:a-1}, Y_{1:\ell}),$$
(36)

and

$$\mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b}) = \mathbb{I}(Y_{c:k-1}; X_{k:\ell} \mid Y_{1:c-1}, X_{1:k-1}) + \mathbb{I}(Y_{k:\ell} \hookrightarrow X_{k:\ell}) + \mathbb{I}(Y_{c:\ell}; X_{\ell+1:b} \mid Y_{1:c-1}, X_{1:\ell}).$$
(37)

Proof of Lemma B.1.

Applying Proposition 3.4 multiple times,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{a:k-1} \to Y_{c:k-1}) + \mathbb{I}(X_{k:\ell} \to Y_{c:k-1}) + \mathbb{I}(X_{\ell+1:b} \to Y_{c:k-1}) + \\
\mathbb{I}(X_{a:k-1} \to Y_{k:\ell}) + \mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) + \mathbb{I}(X_{\ell+1:b} \to Y_{k:\ell}) + \\
\mathbb{I}(X_{a:k-1} \to Y_{\ell+1:d}) + \mathbb{I}(X_{k:\ell} \to Y_{\ell+1:d}) + \mathbb{I}(X_{\ell+1:b} \to Y_{\ell+1:d}).$$
(38)

Notice that either k=a or k=b, and similarly $\ell=c$ or $\ell=d$, so two terms are guaranteed to have empty ranges,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{a:k-1} \to Y_{c:k-1}) + \mathbb{I}(X_{k:\ell} \to Y_{c:k-1}) + \mathbb{I}(X_{\ell+1:b} \to Y_{c:k-1}) + \mathbb{I}(X_{a:k-1} \to Y_{k:\ell}) + \mathbb{I}(X_{\ell+1:d} \to Y_{k:\ell}) + \mathbb{I}(X_{\ell+1:d} \to Y_{k:\ell}) + \mathbb{I}(X_{\ell+1:d} \to Y_{\ell+1:d}).$$
(39)

We can now apply Proposition 3.3 to remove additional terms.

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{a:k-1} \to Y_{c:k-1}) + \mathbb{I}(X_{k:\ell} \to Y_{c:k-1}) + \mathbb{I}(X_{\ell+1:b} \to Y_{c:k-1}) + \mathbb{I}(X_{a:k-1} \to Y_{k:\ell}) + \mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) + \mathbb{I}(X_{\ell+1:b} \to Y_{k:\ell}) + \mathbb{I}(X_{\ell+1:d} \to Y_{\ell+1:d}).$$

$$(40)$$

Recombining the last two terms with Proposition 3.4 gives us three remaining terms,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{a:k-1} \to Y_{k:\ell}) + \mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) + \mathbb{I}(X_{a:\ell} \to Y_{\ell+1:d}) \tag{41}$$

Applying Equation 10 to the first and last term,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \sum_{i=k}^{\ell} \mathbb{I}(X_{a:k-1}; Y_i \mid X_{1:a-1}, Y_{1:i-1}) + \mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) + \sum_{i=\ell+1}^{d} \mathbb{I}(X_{a:\ell}; Y_i \mid X_{1:a-1}, Y_{1:i-1}),$$

and applying the mutual information chain rule to these two summations gives,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{a:k-1}; Y_{k:\ell} \mid |X_{1:a-1}, Y_{1:k-1}) + \mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) + \mathbb{I}(X_{a:\ell}; Y_{\ell+1:d} \mid X_{1:a-1}, Y_{1:\ell}).$$
(43)

This completes the first case. The second case follows by similar reasoning,

$$\begin{split} &\mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b}) \\ &= \underbrace{\mathbb{I}(Y_{c:k-1} \hookrightarrow X_{a:k-1})}_{\mathbb{I}(Y_{c:k-1} \hookrightarrow X_{a:k-1})} + \underbrace{\mathbb{I}(Y_{k:\ell} \hookrightarrow X_{a:k-1})}_{\mathbb{I}(Y_{k:\ell} \hookrightarrow X_{k:\ell})} + \underbrace{\mathbb{I}(Y_{\ell+1:d} \hookrightarrow X_{a:k-1})}_{\mathbb{I}(Y_{\ell+1:d} \hookrightarrow X_{k:\ell})} + \\ &\mathbb{I}(Y_{c:k-1} \hookrightarrow X_{k:\ell}) + \mathbb{I}(Y_{k:\ell} \hookrightarrow X_{k:\ell}) + \underbrace{\mathbb{I}(Y_{\ell+1:d} \hookrightarrow X_{k:\ell})}_{\mathbb{I}(Y_{\ell+1:d} \hookrightarrow X_{\ell+1:b})} + \\ &= \mathbb{I}(Y_{c:k-1} \hookrightarrow X_{k:\ell}) + \mathbb{I}(Y_{k:\ell} \hookrightarrow X_{\ell+1:b}) + \mathbb{I}(Y_{c:\ell} \hookrightarrow X_{\ell+1:d}) \\ &= \mathbb{I}(Y_{c:k-1}; X_{k:\ell} \mid Y_{1:c-1}, X_{1:k-1}) + \mathbb{I}(Y_{k:\ell} \hookrightarrow X_{k:\ell}) + \mathbb{I}(Y_{c:\ell}; X_{\ell+1:b} \mid Y_{1:c-1}, X_{1:\ell}) \end{split}$$
 This completes both cases, and we conclude.

Theorem 3.5 (Conservation Law of GDI).

$$\mathbb{I}(X_{a:b}; Y_{c:d} \mid X_{1:a-1}, Y_{1:c-1}) = \mathbb{I}(X_{a:b} \to Y_{c:d}) + \mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b}). \tag{44}$$

Proof of Theorem 3.5.

The proof proceeds by decomposing the two GDI terms $\mathbb{I}(X_{a:b} \to Y_{c:d})$ and $\mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b})$ by application of Lemma B.1, then inducting on $\ell - k$ as in the original proof by Massey and Massey (2005).

First, we decompose both $\mathbb{I}(X_{a:b} \to Y_{c:d})$ and $\mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b})$ using Lemma B.1.

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{I}(X_{a:k-1}; Y_{k:\ell} \mid | X_{1:a-1}, Y_{1:k-1}) + \mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) + \mathbb{I}(X_{a:\ell}; Y_{\ell+1:d} \mid X_{1:a-1}, Y_{1:\ell}),$$
(45)

and

$$\mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b}) = \mathbb{I}(Y_{c:s-1}; X_{k\cdot\ell} \mid Y_{1\cdot c-1}, X_{1\cdot k-1}) + \mathbb{I}(Y_{k\cdot\ell} \hookrightarrow X_{k\cdot\ell}) + \mathbb{I}(Y_{c\cdot\ell}; X_{\ell+1\cdot b} \mid Y_{1\cdot c-1}, X_{1\cdot\ell}).$$
(46)

Let us now consider $\mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) + \mathbb{I}(Y_{k:\ell} \hookrightarrow X_{k:\ell})$, which is almost the term constrained by the law of conservation of directed information (when k=1). We use induction on $\ell-k$ as employed by Massey and Massey (2005) to show the sum is $\mathbb{I}(X_{k:\ell}; Y_{k:\ell} \mid X_{1:k-1}, Y_{1:k-1})$. First, if $\ell=k$ then $\mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) = \mathbb{I}(X_{k:\ell}; Y_{k:\ell} \mid X_{1:k-1}, Y_{1:k-1})$ and $\mathbb{I}(Y_{k:\ell} \hookrightarrow X_{k:\ell}) = 0$, thus establishing the base case. In the general case,

$$\mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) + \mathbb{I}(Y_{k:\ell} \hookrightarrow X_{k:\ell}) = \mathbb{I}(X_{k:\ell-1} \to Y_{k:\ell-1}) + \mathbb{I}(X_{k:\ell}; Y_e \mid X_{1:k-1}, Y_{1:\ell-1}) + \mathbb{I}(Y_{k:\ell-1} \hookrightarrow X_{k:\ell-1}) + \mathbb{I}(X_e; Y_{k:\ell-1} \mid X_{1:\ell-1}, Y_{1:k-1}).$$

By rearranging and applying induction,

$$= \mathbb{I}(X_{k:\ell-1}; Y_{k:\ell-1} \mid X_{1:k-1}, Y_{1:k-1}) + \mathbb{I}(X_{\ell}; Y_{k:\ell-1} \mid X_{1:\ell-1}, Y_{1:k-1}) + \mathbb{I}(X_{k:\ell}; Y_e \mid X_{1:k-1}, Y_{1:\ell-1}),$$

$$(47)$$

Then, applying the mutual information chain rule on the first two terms,

$$= \mathbb{I}(X_{k:\ell}; Y_{k:\ell-1} \mid X_{1:k-1}, Y_{1:k-1}) + \mathbb{I}(X_{k:\ell}; Y_{\ell} \mid X_{1:k-1}, Y_{1:\ell-1}), \tag{48}$$

and once more on the remaining terms,

$$= \mathbb{I}(X_{k:\ell}; Y_{k:\ell} \mid X_{1:k-1}, Y_{1:k-1}), \tag{49}$$

Thus by induction,

$$\mathbb{I}(X_{k:\ell} \to Y_{k:\ell}) + \mathbb{I}(Y_{k:\ell} \hookrightarrow X_{k:\ell}) = \mathbb{I}(X_{k:\ell}; Y_{k:\ell} \mid X_{1:k-1}, Y_{1:k-1}). \tag{50}$$

We now combine Equation 45, Equation 46, and Equation 50,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) + \mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b}) = \mathbb{I}(X_{a:k-1}; Y_{k:\ell} \mid X_{1:a-1}, Y_{1:k-1}) +$$
(51)

$$\mathbb{I}(X_{k:\ell}; Y_{c:s-1} \mid X_{1:k-1}, Y_{1:c-1}) + \mathbb{I}(X_{k:\ell}; Y_{k:\ell} \mid X_{1:k-1}, Y_{1:k-1}) + \tag{52}$$

$$\mathbb{I}(X_{a:\ell}; Y_{\ell+1:d} \mid X_{1:a-1}, Y_{1:\ell}) + \mathbb{I}(X_{\ell+1:b}; Y_{c:\ell} \mid X_{1:\ell}, Y_{1:c-1}). \tag{53}$$

Notice that at most one of the first two terms is non-zero (if a>c the top term is zero; if a< c the bottom term is zero), and similarly for the final two terms (if b>d the top term is zero; if b< d the bottom term is zero). We now apply the mutual information chain rule to the non-zero first term with the middle term. In either case we get,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) + \mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b}) \tag{54}$$

$$= \mathbb{I}(X_{a:\ell}; Y_{c:\ell} \mid X_{1:a-1}, Y_{1:b-1}) + \mathbb{I}(X_{a:\ell}; Y_{\ell+1:d} \mid X_{1:a-1}, Y_{1:\ell}) + \tag{55}$$

$$\mathbb{I}(X_{\ell+1:b}; Y_{c:\ell} \mid X_{1:\ell}, Y_{1:c-1}). \tag{56}$$

We can now apply the mutual chain rule with the non-zero last term. In either case, we get,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) + \mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b}) = \mathbb{I}(X_{a:b}; Y_{c:d} \mid X_{1:a-1}, Y_{1:b-1}), \tag{57}$$

which confirms the identity, and we conclude the proof.

B.2 Proofs of Results from Section 4 on Plasticity and Empowerment

Lemma 4.2 (Necessary and Sufficient Conditions for Positive Plasticity). Agents have non-zero plasticity relative to e if and only if there is a time-step where their actions are influenced by observation.

More formally, for any pair (λ, e) and indices $[a:b]_n$, $[c:d]_n$, there exists an $i \in [\max(a, c):d]$ such that

$$\mathbb{H}(A_i \mid O_{1:a-1}, A_{1:i-1}) > \mathbb{H}(A_i \mid O_{a:\min(b,i)}, O_{1:a-1}, A_{1:i-1})$$
(58)

if and only if $\mathfrak{P}_{a:b}(\lambda, e) > 0$.

Proof of Lemma 4.2.

We first establish necessity: If $\mathfrak{P}_{a:b\atop c:d}(\lambda,e)>0$, then there exists an $\exists_{i'\in[\max(a,c):d]}$ such that

$$\mathbb{H}(A_{i'} \mid O_{1:a-1}, A_{1:i'-1}) > \mathbb{H}(A_{i'} \mid O_{a:\min(b,i')}, O_{1:a-1}, A_{1:i'-1}). \tag{59}$$

Expanding the definition of $\mathfrak{P}_{a:b}(\lambda, e)$,

$$\mathfrak{P}_{a:b}^{a:b}(\lambda, e) = \mathbb{I}(O_{a:b} \to A_{c:d}) \tag{60}$$

$$= \sum_{i=\max(a,c)}^{d} \mathbb{I}(O_{a:\min(b,i)}; A_i \mid O_{1:a-1}, A_{1:i-1}).$$
(61)

Recall again that the conditional mutual information is non-negative. Hence, the sum

$$\sum_{i=\max(a,c)}^{d} \mathbb{I}(O_{a:\min(b,i)}; A_i \mid O_{1:a-1}, A_{1:i-1}), \tag{62}$$

is only greater than zero if at least one of its terms is non-zero. Since $\mathfrak{P}_{a:b\atop c:d}(\lambda,e)>0$, we conclude that there must be an $i'\in[\max(a,c):d]$ such that $\mathbb{I}(O_{a:\min(b,i')};A_{i'}\mid O_{1:a-1},A_{1:i'-1})>0$. We next apply the identity that conditional mutual information can be expressed as a difference of conditional entropy terms,

$$\underbrace{\mathbb{H}(A_{i'} \mid O_{1:a-1}, A_{1:i'-1}) - \mathbb{H}(A_{i'} \mid O_{a:\min(b,i')}, O_{1:a-1}, A_{1:i'-1})}_{\mathbb{I}(O_{a:\min(b,i')}; A_i \mid O_{1:a-1}, A_{1:i'-1})} > 0, \tag{63}$$

and therefore,

$$\mathbb{H}(A_{i'} \mid O_{1:a-1}, A_{1:i'-1}) > \mathbb{H}(A_{i'} \mid O_{a:\min(b,i')}, O_{1:a-1}, A_{1:i'-1}). \tag{64}$$

This completes the argument for necessity.

We next prove sufficiency: If there exists an $i' \in [\max(a, c) : d]$ such that

$$\mathbb{H}(A_{i'} \mid O_{1:a-1}, A_{1:i'-1}) > \mathbb{H}(A_{i'} \mid O_{a:\min(b,i')}, O_{1:a-1}, A_{1:i'-1}), \tag{65}$$

then $\mathfrak{P}_{a:b\atop a:d}(\lambda,e)>0$.

The same reasoning from the necessity argument may be applied in reverse. First, we can rewrite the above inequality and replace the difference with the conditional mutual information,

$$\mathbb{I}(O_{a:\min(b,i')}; A_{i'} \mid O_{1:a-1}, A_{1:i'-1}) > 0, \tag{66}$$

which in turn implies that

$$\sum_{i=\max(a,c)}^{d} \mathbb{I}(O_{a:\min(b,i)}; A_i \mid O_{1:a-1}, A_{1:i-1}) > 0, \tag{67}$$

and by the definition of plasticity, we conclude $\mathfrak{P}_{\frac{a:b}{c-d}}(\lambda,e)>0.$

Theorem 4.3. The following properties hold of the function \mathfrak{P} , for all intervals $[a:b]_n$, $[c:d]_n$:

- (i) For all (λ, \mathcal{E}) , $\mathfrak{P}_{a:b\atop a:d}(\lambda, \mathcal{E}) \geq 0$.
- (ii) For all a < d, there exists a pair (λ, \mathcal{E}) where λ is deterministic and $\mathfrak{P}_{a:b}^{a:b}(\lambda, \mathcal{E}) > 0$.
- (iii) For all $\mathcal{E}_{\mathrm{small}} \subseteq \mathcal{E}_{\mathrm{big}}$, $\mathfrak{P}_{cd}^{a:b}(\lambda, \mathcal{E}_{\mathrm{small}}) \leq \mathfrak{P}_{cd}^{a:b}(\lambda, \mathcal{E}_{\mathrm{big}})$.
- (iv) (Informal) The following agents all have zero plasticity relative to every environment set: Closed-loop agents, agents that always output the same action distribution, agents whose actions only depend on history length, and agents whose actions only depend on past actions.

Proof of Theorem 4.3.

We prove each property independently.

(i) For all (λ, \mathcal{E}) and any intervals $[a:b]_n$, $[c:d]_n$, $\mathfrak{P}_{a:b}(\lambda, \mathcal{E}) \geq 0$.

This follows as a direct consequence of the first inequality of Proposition C.2, together with the definition of plasticity. \checkmark

(ii) For all a < d, there exists a pair (λ, \mathcal{E}) where λ is deterministic and $\mathfrak{P}_{a:b\atop a:d}(\lambda, \mathcal{E}) > 0$.

It suffices to construct a single deterministic agent with non-zero plasticity for a fixed but arbitrary pair of intervals where a < d. Consider the singleton environment set containing e that outputs a uniform distribution on \mathcal{O} at each time-step.

To be precise, we define a deterministic agent as any non-stochastic mapping from histories to actions, $\lambda: \mathcal{H}_{..o} \to \mathcal{A}$, so the set of all such deterministic agents is the set of all such mappings,

$$\Lambda_{\text{det}} = \mathcal{A}^{\mathcal{H}_{..\mathcal{O}}}.$$
 (68)

Then, let $\lambda \in \Lambda_{\text{det}}$ be a deterministic function that outputs a_1 if the last observation was o_1 , and outputs a_2 otherwise. That is,

$$\lambda(ho) = \begin{cases} \delta\{a_1\} & o = o_1, \\ \delta\{a_2\} & \text{otherwise,} \end{cases}$$
 (69)

for all $h \in \mathcal{H}_{...A}$, $o \in \mathcal{O}$, where $\delta\{a_i\}$ is the Dirac delta distribution on a_i .

Then, for any choice of intervals where a < d, we see that for any $i \in [\max(a, c) : d]$,

$$\underbrace{\mathbb{H}(A_i \mid O_{1:a-1}, A_{1:i-1})}_{>0} > \underbrace{\mathbb{H}(A_i \mid O_{a:\min(b,i)}, O_{1:a-1}, A_{1:i-1})}_{=0}.$$
 (70)

Again by application of Lemma 4.2, this implies that $\mathfrak{P}_{a:b}(\lambda,e)>0$, and we conclude.

(iii) For all $\mathcal{E}_{\mathrm{small}} \subseteq \mathcal{E}_{\mathrm{big}}$ and any intervals $[a:b]_n, [c:d]_n, \mathfrak{P}_{\frac{a:b}{a:d}}(\lambda, \mathcal{E}_{\mathrm{small}}) \leq \mathfrak{P}_{\frac{a:b}{a:d}}(\lambda, \mathcal{E}_{\mathrm{big}}).$

The result follows directly from the use of the max operator in the definition of plasticity. That is, since

$$\max_{x \in \mathcal{X}_{\text{small}}} f(x) \le \max_{x \in \mathcal{X}_{\text{big}}} f(x), \tag{71}$$

for any sets $\mathcal{X}_{\text{small}} \subseteq \mathcal{X}_{\text{big}}$ and function f, then

$$\underbrace{\max_{e \in \mathcal{E}_{\text{small}}} \mathbb{I}(O_{a:b} \to A_{c:d})}_{= \mathfrak{P}_{a:b \atop c:d}(\lambda, \mathcal{E}_{\text{small}})} \le \underbrace{\max_{e \in \mathcal{E}_{\text{big}}} \mathbb{I}(O_{a:b} \to A_{c:d})}_{= \mathfrak{P}_{a:b \atop c:d}(\lambda, \mathcal{E}_{\text{big}})}, \tag{72}$$

as desired. ✓

(iv) (Informal) The following agents all have zero plasticity relative to every environment set: Agents that always output the same action distribution, agents whose actions only depend on history length, and agents whose actions only depend on past actions.

The argument follows directly for all three agent classes by application of Lemma 4.2: For each agent type, conditioning on observations tells us nothing about the actions the agent emits.

To be precise, we define a constant agent as an element of the set

$$\Lambda_{\text{constant}} = \{ \lambda \in \Lambda : \lambda(h) = \lambda(h'), \forall_{h,h' \in \mathcal{H}_{u,\mathcal{O}}} \}.$$
 (73)

That is, a constant agent outputs the same probability distribution over actions in every history.

For any such constant agent, we find that

$$\mathbb{H}(A_i \mid O_{1:a-1}, A_{1:i-1}) = \mathbb{H}(A_i \mid O_{a:\min(b,i)}, O_{1:a-1}, A_{1:i-1})), \tag{74}$$

since each $A_i \sim \lambda_c(H)$ is conditionally independent of $O_{a:\min(b,i)}$ given $A_{1:a-1}, O_{1:i-1}$, by definition of λ_c as a constant agent. By Lemma 4.2, this concludes the argument.

The same reasoning applies to agents whose actions only depend on the length of history, closed-loop agents, and agents whose actions only depend on past actions.

This concludes proof of all four properties.

Proposition 4.5.
$$\mathfrak{E}_{1:n}^{1:n}(\Lambda, e) = \mathfrak{E}_{C}^{n}(\Lambda, e)$$
 and $\mathfrak{E}_{1:n}^{1:n}(\Lambda_{a^{n}}, e) = \mathfrak{E}_{K}^{n}(\Lambda_{a^{n}}, e)$.

Proof of Proposition 4.5.

We show that Capdepuy's and Klyubin et al.'s definitions of empowerment are special cases in separate arguments.

(Capdepuy)

First, we can see that the GDI-based definition captures Capdepuy's definition when 1 = a = cand b = d = n, since

$$\mathfrak{E}_{\stackrel{a:b}{e:d}}(\Lambda, e) = \max_{\lambda \in \Lambda} \mathbb{I}(A_{1:n} \to O_{1:n}) = \mathfrak{E}_{\mathcal{C}}^n(\Lambda, e). \qquad \checkmark$$
 (75)

(Klyubin et al.)

Second, we show that when 1 = a = c and b = d = n and the set of agents in question are the open loop agents, Λ_{a^n} , then

$$\mathfrak{E}_{a:b}(\Lambda_{a^n}, e) = \mathfrak{E}_{\lim}(\Lambda_{a^n}, e) = \mathfrak{E}_{K}^n(\Lambda_{a^n}, e), \tag{76}$$

We start by expanding the GDI-based definition,

$$\mathfrak{E}_{\frac{1:n}{1:n}}(\Lambda_{a^n}, e) = \max_{\lambda \in \Lambda_{a^n}} \mathbb{I}(A_{a:b} \to O_{c:d}),\tag{77}$$

$$= \max_{\lambda \in \Lambda} \mathbb{I}(A_{1:n} \to O_{1:n}), \tag{78}$$

$$= \max_{p(a^n) \in \Lambda_{a^n}} \mathbb{I}(A_{1:n} \to O_{1:n}). \tag{79}$$

Now, since the agents are open loop, we know that $\mathbb{I}(O_{1:n} \hookrightarrow A_{1:n}) = 0$. By Proposition 2.3,

$$\underbrace{\mathbb{I}(O_{1:n} \hookrightarrow A_{1:n})}_{=0} + \mathbb{I}(A_{1:n} \to O_{1:n}) = \mathbb{I}(O_{1:n}; A_{1:n}), \tag{80}$$

and hence,

$$\mathbb{I}(A_{1:n} \to O_{1:n}) = \mathbb{I}(O_{1:n}; A_{1:n}). \tag{81}$$

Therefore,

$$\mathfrak{E}_{c:d}^{a:b}(\Lambda_{a^{n}}, e) = \max_{p(a^{n}) \in \Lambda_{a^{n}}} \mathbb{I}(A_{1:n} \to O_{1:n})$$

$$= \max_{p(a^{n}) \in \Lambda_{a^{n}}} \mathbb{I}(O_{1:n}; A_{1:n})$$
(82)

$$= \max_{p(a^n) \in \Lambda_{a^n}} \mathbb{I}(O_{1:n}; A_{1:n})$$
 (83)

$$=\mathfrak{E}_{K}^{n}(\Lambda_{a^{n}},e).\checkmark\tag{84}$$

This conclude both equalities, and we conclude the proof.

Proposition 4.6 (Plasticity is the Mirror of Empowerment). Given any pair (λ, e) that share an interface, and valid intervals $[a:b]_n$, $[c:d]_n$, the agent's empowerment is equal to the plasticity of the environment, and the agent's plasticity is equal to the empowerment of the environment:

$$\mathfrak{E}_{a:b}(\lambda) = \mathfrak{P}_{a:b}(e), \qquad \mathfrak{P}_{a:b}(\lambda) = \mathfrak{E}_{a:b}(e).$$
 (85)

Proof of Proposition 4.6.

The proof proceeds as a direct consequence of the definitions of empowerment and plasticity, and the symmetry highlighted by Remark 2.8.

First, we expand each definition for the agent,

$$\mathfrak{P}_{a:b}(\lambda) = \mathfrak{P}_{a:b}(\lambda, e) = \mathbb{I}(O_{a:b} \to A_{c:d}),\tag{86}$$

$$\mathfrak{E}_{a:b}(\lambda) = \mathfrak{E}_{a:b}(\lambda, e) = \mathbb{I}(A_{a:b} \to O_{c:d}). \tag{87}$$

Now, to arrive at analogous definitions for the environment, we exploit the symmetry highlighted by Remark 2.8 to recover notions of plasticity and empowerment when the environment is treated as an agent,

$$\mathfrak{P}_{a:b}(e) = \mathfrak{P}_{a:b}(e,\lambda) = \mathbb{I}(A_{a:b} \to O_{c:d}), \tag{88}$$

$$\mathfrak{E}_{a:b}^{a:b}(e) = \mathfrak{E}_{a:d}^{a:b}(e,\lambda) = \mathbb{I}(O_{a:b} \to A_{c:d}). \tag{89}$$

Now, notice that $\mathfrak{E}(\lambda) = \mathfrak{P}(e)$ by Equation 87 and Equation 88, and $\mathfrak{P}(\lambda) = \mathfrak{E}(e)$ by Equation 86 and Equation 89.

Theorem 4.8. For all intervals $[a:b]_n$, $[c:d]_n$ and any pair (λ, e) , let $m = \min\{(b-a)\log |\mathcal{O}|, (d-c)\log |\mathcal{A}|\}$. Then,

$$\mathfrak{E}_{a:b}(\lambda, e) + \mathfrak{P}_{c:d}(\lambda, e) \le m, \tag{90}$$

is a tight upper bound on empowerment and plasticity. Furthermore, under mild assumptions on the interface and interval, there exists a pair $(\lambda^{\diamond}, e^{\diamond})$ such that $\mathfrak{P}_{\stackrel{\text{cid}}{a:b}}(\lambda^{\diamond}, e^{\diamond}) = m$, and there exists a pair (λ', e') such that $\mathfrak{E}_{\stackrel{\text{cid}}{a:b}}(\lambda', e') = m$, thereby forcing the other quantity to be zero.

Proof of Theorem 4.8.

Let (λ, e) be a fixed but arbitrary agent-environment pair, $[a:b]_n, [c:d]_n$ be valid intervals, and let

$$\varepsilon = m - \mathbb{I}(O_{a:b}; A_{c:d} \mid O_{1:a-1}, A_{1:c-1}).$$

Then, by Theorem 3.5, we upper bound the sum,

$$\mathfrak{P}_{c:d}^{c:d}(\lambda, e) + \mathfrak{E}_{a:b}^{a:b}(\lambda, e) = \mathbb{I}(O_{a:b}; A_{c:d} \mid O_{1:a-1}, A_{1:c-1})$$
(91)

$$\leq_1 \min\left\{ \mathbb{H}(O_{a:b}), \mathbb{H}(A_{c:d}) \right\} \tag{92}$$

$$\leq_2 m.$$
 (93)

where \leq_1 follows by the standard upper bounds on conditional mutual information, and \leq_2 follows from the standard upper bound on joint entropy.

Thus, we can express m in terms of a sum,

$$m = \underbrace{\mathfrak{E}_{a:b}_{c:d}(\lambda, e) + \mathfrak{P}_{c:d}_{a:b}(\lambda, e)}_{=\mathbb{I}(...)} + \underbrace{\varepsilon}_{=m-\mathbb{I}(...)}.$$
(94)

Therefore, rearranging,

$$\mathfrak{E}_{a:b}^{a:b}(\lambda, e) = m - \mathfrak{P}_{a:b}^{c:d}(\lambda, e) - \varepsilon. \tag{95}$$

By Equation 93, $\varepsilon \geq 0$, and by Equation 91, $\mathfrak{P}_{c:d\atop a:b}(\lambda,e) \leq m$. Therefore,

$$\mathfrak{E}_{a:b}(\lambda, e) \le m - \mathfrak{P}_{c:d}(\lambda, e), \qquad \mathfrak{P}_{c:d}(\lambda, e) \le m - \mathfrak{E}_{a:b}(\lambda, e).$$
 (96)

Since each of (λ, e) , $[a:b]_n$, $[c:d]_n$ were arbitrary, we conclude these bounds hold for every agent-environment pair and all valid intervals.

To prove these are tight, we show there exists two pairs, $(\lambda^{\diamond}, e^{\diamond})$ and (λ', e') such that

$$m - \mathfrak{P}_{a:b}(\lambda^{\diamond}, e^{\diamond}) = 0, \qquad m - \mathfrak{E}_{a:b}(\lambda', e') = 0,$$
 (97)

We present the argument for plasticity, and recall that by symmetry (Remark 2.8), the same will hold of empowerment.

To do so, we require that $|\mathcal{A}| \leq |\mathcal{O}|$ and $(b-a) \geq (d-c)$. While the desired inequality will likely hold under softer conditions, the argument is cleanest under these conditions. For simplicitly, we also assume $a \leq c$.

By Proposition C.3, we express $\mathfrak{P}_{a:b}(\lambda^{\diamond}, e^{\diamond})$ as follows

$$\mathfrak{P}_{a:b}(\lambda^{\diamond}, e^{\diamond}) = \mathbb{I}(O_{a:b} \to A_{c:d}) \tag{98}$$

$$= \mathbb{H}(A_{\max(a,c):d} \mid O_{1:a-1}) - \mathbb{H}(A_{c:d} \mid\mid O_{a:b}). \tag{99}$$

Since we assumed $a \le c$, we rewrite

$$\mathfrak{P}_{c:d}^{a:b}(\lambda^{\diamond}, e^{\diamond}) = \mathbb{H}(A_{c:d} \mid O_{1:a-1}) - \mathbb{H}(A_{c:d} \mid\mid O_{a:b}). \tag{100}$$

Let $k = |\mathcal{A}|$, and let \mathcal{O}_k refer to the set containing the first k observations, $\mathcal{O}_k = \{o_1, o_2, \ldots, o_k\}$. Let e^{\diamond} be the environment that always emits o_1 along the initial (possibly empty) interval $[1:a-1]_n$, and emits a uniform random distribution over \mathcal{O}_k from time i=a onward. Now, let λ^{\diamond} be the agent that always plays a_1 until the start of the observation interval [a:b], then on seeing any observation o_j , it mirrors the same action back, a_j . More formally,

$$e(h_i) = \begin{cases} \delta\{o_1\} & i \in [1:a-1], \\ \text{Unif}(\mathcal{O}_k) & i \ge a, \end{cases} \qquad \lambda(h_i o_j) = \begin{cases} \delta\{a_1\} & i < a \lor j > k, \\ \delta\{a_j\} & \text{otherwise}, \end{cases}$$
(101)

for all $h_i \in \mathcal{H}_{...A}$, $o_j \in \mathcal{O}$, where $h_i o_j \in \mathcal{H}_{...\mathcal{O}}$ expresses the concatenation of h_i and o_j . Additionally, we let $\delta\{x_i\}$ express the Dirac delta distribution on x_i .

In this situation, observe that

$$\mathbb{H}(A_{c:d} \mid O_{1:a-1}) = (d-c) \cdot \log |\mathcal{A}|, \tag{102}$$

since the agent plays the uniform random distribution over A in this interval by mirroring the environment's uniform random distribution over O_k , and we assumed $(b-a) \ge (d-c)$. Furthermore,

$$\mathbb{H}(A_{c:d} \mid\mid O_{a:b}) = 0 \tag{103}$$

since learning about the observation at each time-step in the interval [a:b] fully determines the action in the interval [c:c+(b-a)]. Therefore, substituting into Equation 100

$$\mathfrak{P}_{c:d}^{a:b}(\lambda^{\diamond}, e^{\diamond}) = \underbrace{\mathbb{H}(A_{c:d} \mid O_{1:a-1})}_{\geq m} - \underbrace{\mathbb{H}(A_{c:d} \mid\mid O_{a:b})}_{=0}. \tag{104}$$

Hence, we conclude that

$$\mathfrak{P}_{a;b}^{a;b}(\lambda^{\diamond}, e^{\diamond}) \ge m,\tag{105}$$

and by the non-negativity of empowerment and plasticity,

$$\mathfrak{E}_{a:b}(\lambda^{\diamond}, e^{\diamond}) \le m - \mathfrak{P}_{a:b}(\lambda^{\diamond}, e^{\diamond}) = 0. \tag{106}$$

This completes the argument for the case of empowerment, and we call attention to the fact that by symmetry (Remark 2.8), an identical argument holds for plasticity.

C Other Results

Theorem C.1 (Data-processing inequality for GDI). Consider any tuple $(X_{1:n}, Y_{1:n}, Z_{1:n})$ where Z_i is conditionally independent of $X_{1:i}$ given Y_i , for each i = 1 ... n. Then,

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) \ge \mathbb{I}(X_{a:b} \to Z_{c:d}),\tag{107}$$

for any valid intervals $[a:b]_n$, $[c:d]_n$.

Proof of Theorem C.1.

The argument follows an identical structure to the standard argument for the data-processing inequality for regular mutual information, together with two inductive steps. The first inductive step applies to the interval $[a:b]_n$ and the second applies to the interval $[c:d]_n$. We assume across all arguments that each Z_i is conditionally independent of $X_{1:i}$ given Y_i for each i=1...n.

Base Case: a = b, c = d

First, we consider the base case when a=b and c=d, and assume $a \le d$ (if a>d, we know $\mathbb{I}(X_{a:b} \to Y_{c:d})=0$ by Proposition 3.3). Hence, we write $\mathbb{I}(X_{a:a} \to Y_{c:c})=\mathbb{I}(X_a \to Y_c)$. Expanding the definition for both relations X to Y and X to Z, we fine,

$$\mathbb{I}(X_a \to Y_c) = \sum_{i=\max(a,c)}^{d} \mathbb{I}(X_{a:\min(b,i)}; Y_i \mid X_{1:a-1}Y_{1:i-1}) = \mathbb{I}(X_a; Y_{\max(a,c)}),$$
(108)

$$\mathbb{I}(X_a \to Z_c) = \sum_{i=\max(a,c)}^{d} \mathbb{I}(X_{a:\min(b,i)}; Z_i \mid X_{1:a-1}Z_{1:i-1}) = \mathbb{I}(X_a; Z_{\max(a,c)}).$$
(109)

Then, we apply the standard proof technique for the data-processing inequality: We consider the mutual information $\mathbb{I}(X_a; Y_{\max(a,c)}, Z_{\max(a,c)})$, and use the identity that

$$\mathbb{I}(X_a; Y_{\max(a,c)}, Z_{\max(a,c)}) = \mathbb{I}(X_a; Y_{\max(a,c)}) + \mathbb{I}(X_a; Y_{\max(a,c)} \mid Z_{\max(a,c)}), \tag{110}$$

$$= \mathbb{I}(X_a; Z_{\max(a,c)}) + \underbrace{\mathbb{I}(X_a; Z_{\max(a,c)} \mid Y_{\max(a,c)})}_{-0}. \tag{111}$$

Where the final term is equal to zero by the assumption that each Z_i is conditionally independent of $X_{1:i}$ given Y_i for each $i = 1 \dots n$. Therefore, rearranging, we find that

$$\mathbb{I}(X_a; Y_{\max(a,c)}) + \mathbb{I}(X_a; Y_{\max(a,c)} \mid Z_{\max(a,c)}) = \mathbb{I}(X_a; Z_{\max(a,c)}), \tag{112}$$

and by the non-negativity of mutual information, we conclude

$$\underbrace{\mathbb{I}(X_a; Y_{\max(a,c)})}_{=\mathbb{I}(X_{a:a} \to Y_{c:c})} \ge \underbrace{\mathbb{I}(X_a; Z_{\max(a,c)})}_{=\mathbb{I}(X_{a:a} \to Z_{c:c})}, \tag{113}$$

as desired. This concludes the base case.

We next turn to the two inductive cases.

Inductive Case 1: $\mathbb{I}(X_{a:b} \to Y_{c:c}) \ge \mathbb{I}(X_{a:b} \to Z_{c:c})$

We assume as our inductive hypothesis that

$$\mathbb{I}(X_{a:b-1} \to Y_{c:c}) > \mathbb{I}(X_{a:b-1} \to Z_{c:c}),\tag{114}$$

and show that this implies

$$\mathbb{I}(X_{a:b} \to Y_{c:c}) \ge \mathbb{I}(X_{a:b} \to Z_{c:c}). \tag{115}$$

We apply Proposition 3.4 once on the terms $\mathbb{I}(X_{a:b} \to Y_{c:c})$ and $\mathbb{I}(X_{a:b} \to Z_{c:c})$ to decompose them into

$$\mathbb{I}(X_{a:b} \to Y_{c:c}) = \mathbb{I}(X_{a:b-1} \to Y_{c:c}) + \mathbb{I}(X_{b:b} \to Y_{c:c})$$
(116)

$$\mathbb{I}(X_{a:b} \to Z_{c:c}) = \mathbb{I}(X_{a:b-1} \to Z_{c:c}) + \mathbb{I}(X_{b:b} \to Z_{c:c}). \tag{117}$$

But then, by the inductive hypothesis and the base case, we know that

$$\mathbb{I}(X_{a:b-1} \to Y_{c:c}) \ge \mathbb{I}(X_{a:b-1} \to Z_{c:c}), \qquad \mathbb{I}(X_{b:b} \to Y_{c:c}) \ge \mathbb{I}(X_{b:b} \to Z_{c:c}). \tag{118}$$

Thus,

$$\underbrace{\mathbb{I}(X_{a:b} \to Y_{c:c})}_{=\mathbb{I}(X_{a:b-1} \to Y_{c:c}) + \mathbb{I}(X_{b:b} \to Y_{c:c})} \ge \underbrace{\mathbb{I}(X_{a:b} \to Z_{c:c})}_{=\mathbb{I}(X_{a:b-1} \to Z_{c:c}) + \mathbb{I}(X_{b:b} \to Z_{c:c})}, \tag{119}$$

as desired. ✓

Inductive Case 2: $\mathbb{I}(X_{a:b} \to Y_{c:d}) \ge \mathbb{I}(X_{a:b} \to Z_{c:d})$.

The argument for this inductive case is identical to the former, only we apply the inductive step to the interval $[c:d]_n$. \checkmark

This concludes the proof.

Next, as with most terms related to mutual information, the GDI is non-negative and upper bounded.

Proposition C.2 (GDI Upper Bound).

$$0 \le \mathbb{I}(X_{a:b} \to Y_{c:d}) \le \mathbb{I}(X_{a:b}; Y_{c:d} \mid X_{1:a-1}, Y_{1:c-1}). \tag{120}$$

Proof of Proposition C.2.

The two inequalities follow as a consequence of the definition of GDI, and its conservation law (Theorem 3.5).

First, note that $\mathbb{I}(X_{a:b} \to Y_{c:d})$ is non-negative by definition:

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \sum_{i=\max(a,c)}^{d} \underbrace{\mathbb{I}(X_{a:\min(b,i)}; Y_i \mid X_{1:a-1}, Y_{1:i-1})}_{>0}.$$
 (121)

since it is the sum of conditional mutual information terms that are each non-negative by definition. Second, by Theorem 3.5, we know that

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) + \mathbb{I}(Y_{c:d} \hookrightarrow X_{a:b}) = \mathbb{I}(X_{a:b}; Y_{c:d} \mid X_{1:a-1}, Y_{1:c-1}), \tag{122}$$

and since all three terms are non-negative, it follows that

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) \le \mathbb{I}(X_{a:b}; Y_{c:d} \mid X_{1:a-1}, Y_{1:c-1}).$$

Lastly, we can decompose the GDI into conditional entropy terms.

Proposition C.3 (GDI Kramer Decomposition).

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \mathbb{H}(Y_{\max(a,c):d} \mid X_{1:a-1}) - \mathbb{H}(Y_{c:d} \mid\mid X_{a:b}), \tag{123}$$

where

$$\mathbb{H}(Y_{c:d} \mid\mid X_{a:b}) = \sum_{i=\max(a,c)}^{d} \mathbb{H}(Y_i \mid Y_{1:i-1}, X_{a:\min(b,i)}), \tag{124}$$

is the generalization of Kramer's causal entropy to arbitrary indices.

Proof of Proposition C.3.

The result follows by expanding the definition of GDI and applying the conditional mutual information identity in terms of a difference of conditional entropy terms, as follows.

$$\mathbb{I}(X_{a:b} \to Y_{c:d}) = \sum_{i=\max(a,c)}^{d} \mathbb{I}(X_{a:\min(b,i)}; Y_i \mid X_{1:a-1}, Y_{1:i-1}), \tag{125}$$

$$= \sum_{i=\max(a,c)}^{d} \mathbb{H}(Y_i \mid X_{1:a-1}, Y_{1:i-1}) - \mathbb{H}(Y_i \mid X_{a:\min(b,i)}, X_{1:a-1}, Y_{1:i-1}), \tag{126}$$

$$= \sum_{i=\max(a,c)}^{d} \mathbb{H}(Y_i \mid X_{1:a-1}, Y_{1:i-1}) - \mathbb{H}(Y_i \mid Y_{1:i-1}, X_{1:\min(b,i)}), \tag{126}$$

$$= \sum_{i=\max(a,c)}^{d} \mathbb{H}(Y_i \mid X_{1:a-1}, Y_{1:i-1}) - \sum_{i=\max(a,c)}^{d} \mathbb{H}(Y_i \mid Y_{1:i-1}, X_{1:\min(b,i)}), \tag{127}$$

$$= \mathbb{H}(Y_{\max(a,c):d} \mid X_{1:a-1}) - \mathbb{H}(Y_{c:d} \mid X_{a:b}), \tag{127}$$

$$= \mathbb{H}(Y_{\max(a,c):d} \mid X_{1:a-1}) - \mathbb{H}(Y_{c:d} \mid X_{a:b}). \qquad \Box$$

As a direct corollary of Proposition C.3 together with Proposition 3.2, we recover Kramer's identity as a special case, as follows.

Corollary C.4. When 1 = a = c and b = d = n,

$$\mathbb{H}(Y_{\max(a,c):d} \mid X_{1:a-1}) - \mathbb{H}(Y_{c:d} \mid\mid X_{a:b}) = \mathbb{H}(Y_{1:n}) - \mathbb{H}(Y_{1:n} \mid\mid X_{1:n}), \tag{128}$$

and,

$$\mathbb{H}(Y_{c:d} \mid\mid X_{a:b}) = \mathbb{H}(Y_{1:n} \mid\mid X_{1:n}). \tag{129}$$