Signal-to-Computation Ratio: A Measure of System Complexity through Information Extraction

J.L

September 11, 2024

Abstract

In this paper, we introduce the concept of Signal-to-Computation Ratio (SCR), a new metric for evaluating the computational efficiency required to extract information from a system. SCR is an intrinsic property of the system and is independent of the algorithm used. We define SCR mathematically and extend the concept to time-evolving dynamical systems.

1 Introduction

In many complex systems, the efficiency of extracting information from a system is of paramount importance. While the signal-to-noise ratio (SNR) is a well-known metric in communication theory, we propose an analogous concept, the Signal-to-Computation Ratio (SCR), to measure the computational cost relative to the amount of information extracted from a system, if at-all possible. The SCR is a system-specific property, reflecting how "difficult" it is to extract meaningful information from that system, irrespective of the specific algorithms used.

The goal of this paper is to formalize SCR, explore its properties, and examine its applications. Furthermore, we investigate it in the context of time-evolving dynamical systems.

2 Definitions

2.1 System

The family of systems is defined as follows:

$$X := \{x_t\}_{t \in T}$$

where T is the index set. This family represents the various system states, indexed by $t \in T$, t might represent time or any other indexing parameter.

2.2 Binary Query

We define the binary set $\mathbb{B}$ and the set of binary queries Q as follows:

$$\mathbb{B} := \{0, 1\}$$
$$Q := \{q : X \to \mathbb{B}\}$$

Every q essentially creates a macrostate of x_t, whose microstate may or may not obey equipartition.

Two binary queries q and q' are considered identical if they return the same output for all systems $x_t \in X$:

$$\forall q, q' \in Q, \quad q = q' \iff \forall x_t \in X, q(x) = q'(x)$$

Furthermore, two queries q and q' are considered equivalent upto a transformation f from the set $\amalg$ of inver-table functions:

$$\forall q, q' \in Q, \quad q \sim q' \iff \exists f \in \amalg, \forall x_t \in X, f \circ q(x_t) = q'(x_t)$$

2.2.1 Closure

2.3 Horizon

The horizon H is defined as the set of functions that map a system $x_t \in X$ to its obtainable data:

$$H := \{h : X \to D\}$$

Each $h \in H$ reflects the data obtainable about $x_t \in X$.

2.4 Sandbox

The sandbox S is defined as a system paired with its horizon:

$$S := X \times H$$

Notationally, instead of writing $(x_t, h) \in S$, we use the notation $x_t|h \in S$.

2.5 Binary Algorithm

A binary algorithm processes some data $d \in D$ and outputs a single bit:

$$A := \{a : D \to \mathbb{B}\}$$

For $a \in A$, $q \in Q$, and $h \in H$, an algorithm implementing a query under h is defined as:

$$a \underset{h}{\rightsquigarrow} q \iff \forall x_t \in X, q(x_t) = a(h(x_t))$$

Two binary algorithms a and a' are considered equivalent under h if they implement a same query:

$$a \underset{h}{\sim} a' \iff \exists q \in Q, a \underset{h}{\rightsquigarrow} q \wedge a' \underset{h}{\rightsquigarrow} q$$

The set of algorithms implementing a query q under horizon h is denoted as:

$$A_{q|h} := \{a \in A : a \underset{h}{\rightsquigarrow} q\}$$

2.6 Meaningful Query Set

The set of meaningful query under horizon h is defined as:

$$Q_h^* := \left\{ q \in \frac{Q}{\sim} : \exists a \in A, a \underset{h}{\rightsquigarrow} q \right\}$$

The set essentially eliminates unimplementable queries and redundant (defined by $\sim$) queries.

2.7 Computation Resource

The computation resource cost $\mathfrak{c}(a, d)$ is the cost of computing $a(d)$:

$$\mathfrak{c} : A \times D \to \mathbb{R}^+$$

The computation resource cost associated with a query q, given sandbox $x_t|h$, is the minimum cost of computing the query:

$$c : S \to (Q \to \mathbb{R}^+)$$
$$c_{x_t|h}(q) := \min_{a \in A_{q|h}} \mathfrak{c}(a, h(x_t))$$

In practice, the computation resource could be measured in terms of the number of operations (e.g., additions, multiplications) or energy consumption (e.g., joules).

2.8 Extraction Cost

The extraction cost $e_{x_t|h}(n)$ is the total computation resource cost to extract n bits of information from the sandbox $x_t|h$:

$$e : S \to (\mathbb{N} \to \mathbb{R}^+)$$
$$e_{x_t|h}(n) := \min_{\{\Omega \subseteq Q_h^* : |\Omega| = n\}} \sum_{q \in \Omega} c_{x_t|h}(q)$$

2.9 Signal-to-Computation Ratio (SCR)

The signal-to-computation ratio (SCR) of a sandbox, is a curve over n bits of information extracted, trending the increasing difficulty to extract more information,

$$SCR : S \to (\mathbb{N} \to \mathbb{R}^+)$$

$$SCR_{x_t|h}(n) := \frac{n}{e_{x_t|h}(n)}$$

3 Application to Time-Evolving Dynamical Systems

In many real-world applications, the systems of interest are time-evolving; $T = \mathbb{R}$ represents time. Let x_t represent the state of a dynamical system at time $t \in T$. In this context, the horizon often considers the entire trajectory of the system over a time interval, rather than a single snapshot.

$$h_{[t_0,t_1]}(x_t) := \{x_t\}_{t \in [t_0,t_1]}$$

The SCR to forecast a dynamic system Δt in the future is then:

$$SCR_{x_{t+\Delta t}|h_{[0,t]}}(n)$$

Especially in chaotic systems, as Δt increases, SCR decreases drastically, indicating the exponential difficulty to predict the far future.

4 Conclusion

In this paper, we introduced the concept of Signal-to-Computation Ratio (SCR) as a measure of the computational difficulty required to extract information from a system. We defined SCR mathematically and explored its properties, especially in the context of time-evolving dynamical systems.