

TIME, SPACE, CAUSALITY, NEURAL NETWORKS, AND TEXT

Mykola Rabchevskiy

Reader reactions to chapters “[AGI IN A WORLD OF TRAJECTORIES](#)”, “[CAUSALITY](#)”, “[STRUCTURES DISCOVERING](#)” suggests the need to discuss the aspects touched upon there in more detail due to their fundamental nature.

Take a look at a short video illustrating the topic of the chapter:

Stopping the video at any position, we see nothing more than randomly scattered points in space. However, when observing the process in dynamics, that is, in a situation where space is combined with time, we immediately discover some rotating object, clearly distinguishable from the chaotic movement of other points. What allows us to do this? The presence of constancy in a dynamic process. What prevents us from finding an object in a static frame? Lack of dynamics.

The **dynamism of the observed environment** combined with the **ability to find that constant (invariant)** that connects a group of elements into a single whole, observed over a **finite time interval**, allow us to detect an **unknown object**.

If there are some known objects, it is possible to detect them by comparing them with their stored descriptions; it is obviously impossible to detect unknown objects by comparing them with known ones. The only way to find them is to **search for the structure** in a dynamic environment, as described in “[Structure discovering](#)”. The choice of points as elementary objects of the scene in the video guarantees the absence of known objects for the "purity of the experiment".

The observed effects explain the essential features of the animal's and human behavior: since the dynamics of a situation are determined by the **relative movement** of the observer and the environment, an active observer can turn a static situation into a dynamic one by moving relative to the environment. This behavior is an integral part of **active sensing**.

Movement relative to objects in the environment allows you to detect unknown (and, accordingly, unidentifiable) objects and the analysis of their apparent "displacement" relative to each other (that is, the relative displacement of the projections of objects on a two-dimensional subspace) allows you to **rank objects in terms of distance** to them (and even estimate the distance quantitatively if there are recognized objects of known size).

Structural detection described in “[Structure discovering](#)” is based on the search for **invariants** constructed from functions that reflect the specifics of space: distance between objects in space, mutual orientation in space, and all kinds of their **combinations as functions of time**. Thus, the use of Spatio-temporal relationships of the natural environment is the basis for constructing structured models of environmental objects.

Invariants that allow detecting unknown objects can use other types of attributes besides spatial ones, for example, **color** and **brightness** (in the case of computer vision as a source of information). If spatial attributes are not used and the attributes of the potential object are constant over time, dynamism is not required. It is the basis of the simplest methods of visual detection of objects, based on the **segmentation of static images** by color and brightness. However, as is clear from the video, these approaches significantly limit the ability to detect unknown objects; **mimicry and camouflage successfully suppress detection capabilities** in the absence of dynamics.

It is easy to see common elements in the spatial structure search method and the cause-effect search method described in “[Causality](#)”: in both cases, what is to be detected **cannot be calculated** like calculating velocity, correlation or volume; a **search/selection/construction of a specific function** required that describes the structure or cause-and-effect relationship. It should be especially noted that the sought invariant function, which implicitly describes the structure, has as arguments the spatial attributes of all objects-components of the structure, that is, the **type** of the function and the **number of arguments is unknown in advance**.

An essential feature of **constructing an invariant function** corresponding to the desired object is that it is per se a **combinatorial problem**, the complexity of which grows rapidly both with an increase in the number of potential elements of an unknown structure and with an increase in the number of elementary relation functions used as "building blocks".

An essential consequence of the **impossibility of "calculating"** the structure and the need to construct an invariant function is that this operation, in principle, **cannot be performed by a neural network**, which, regardless of size and complexity, is nothing more than a function that calculates the answer using a **fixed number of arguments**. This circumstance fundamentally interferes with the ability of a neural network to detect objects that it has not been "trained" before it used for its intended purpose and thus limits the scope of applicability of neural networks as a fundamental element of an AGI system: **performing the most "intelligent" operations of detecting structures and cause-and-effect relationships by a neural network are impossible**.

Important conclusions can also be revealed by comparing the search for unknown objects in the natural environment and the search for cause-and-effect relationships with the processes of **text analysis** by artificial intelligence systems.

The apparent difference between the natural environment and the text is the primitiveness of the relations between the elementary objects of the text symbols: for two symbols A and B, the relations are reduced to the order (A precedes B or vice versa). The differences from the natural environment are radical: first, the elementary relation for the text environment is **one and only**; secondly, there is **no time**. A remote analog of time can be considered the **index of the current symbol** in the whole sequence - but then the **analog of the current situation** in a natural environment that contains many objects **is a single current symbol**. An instantaneous situation in a natural environment is characterized by the **relative position of many objects, their relative sizes**, and there is nothing similar for a single symbol. There also is **no division of situations into static and dynamic**. As a result, an analog of the search for unknown spatial structures in a natural environment is the search for unknown chains of symbols, which is much simpler both due to discreteness and one-dimensionality.

At the same time, there is a commonality: **unknown textual patterns of arbitrary length cannot be "calculated" as a result of a function with a fixed number of parameters** either; using neural networks for text analysis faces the same problems as when analyzing situations in a natural environment.

SUMMATION

- **Detection of previously unknown objects in a natural environment requires a dynamic situation.**
- **Dynamism can be created by the movement of the object controlled by the AGI system.**
- **Detection of unknown objects and causal relationships cannot be reduced to calculating the value of a particular previously known function, as is the case when calculating velocity, volume, correlation, and so on; a search/construction of an invariant function is required, the structure and complexity of which are not known in advance.**
- **Constructing an invariant function is a combinatorial task.**
- **A neural network cannot perform the construction of an invariant function.**
- **The text analysis is much simpler due to the one-dimensionality and the absence of time as a property of the "text environment".**

Subscribe to AGI engineering

By Mykola Rabchevskiy · Launched 2 years ago

AGI: fundamentals, architecture, implementation, source code