AGI: STRUCTURES DISCOVERING

Mykola Rabchevskiy

In this chapter, we will continue the topic of ***generating new concepts by the AGI system*** - in this case, by detecting ***structures that represent previously unknown objects*** in the environment. Before discussing this, let's recall the details mentioned earlier.

A specific AGI-driven system is equipped with sensors and actuators according to the mission for which the system is being created. For the intelligent operation of sensors and actuators and the interpretation of the data received from them, the set of congenital knowledge includes entities corresponding to these commands and data. For the intelligent operation of sensors and actuators and interpretation of the data received from them, the set of ***congenital knowledge*** includes entities corresponding to these commands and data. For example, a message from a sensor might be classified as dangerous, beneficial, or neutral. Of course, there are congenital concepts "***sensor***", "***actuator***", "***command***", "***response***", and "***event***", which make it possible to distinguish these primary atomic entities from others.

According to the adopted terminology (Semantic storage), commands transmitted to sensors and messages received are primary non-removable atomic entities (i.e., not used as signs of other entities). ***Neither sending a command nor receiving a response creates new logical entities***, but ***references*** to such entities can be added to Chronicle.

Along with the information received in response to a request to a sensor or a command to an actuator, there may be messages about ***events*** in the external environment or the controlled system. They are also primary atomic entities, with the difference that they are not received as a response to a request.

Earlier, we discussed two ways of generating new entities: detecting repeating patterns in a sequence of events (DIY pattern mining) and detecting simple causal relationships (Causality).

The third "branch" in the concept generation system is the detection of ***unknown objects*** and, accordingly, the generation on this basis of ***concepts*** that describe the structure of these objects.

In the first of the listed tasks, the desired object of the search is initially clear - it is a sequence of logical entities - and the criterion for finding is obvious: an exact coincidence of sequences of essences. The task essentially boils down to finding identical objects (sequences of entities) among the set of known ones.

The second task is more complex: an unknown cause of a known consequence-event is found. To find the desired dependence, it is required to ***generate hypothetical dependencies*** of the effect on various possible causes and then test the hypotheses by comparing the predictions arising from the hypotheses with reality. The task can be characterized as finding unknown relationships between known entities.

The third task is even more difficult. When looking for a cause, we know the result, the cause of which should be found. When discovering unknown objects, we don't know anything specific; the ***combination of elements forming a stable structure is the desired new concept***. An approach based on the formulation of hypotheses with subsequent testing is acceptable for solving the problem, but what should a hypothesis look like in practice?

The most common aspects of the structure as a concept are:

- Presence of several components.
- Constancy in time of a specific set of parameters depends on the components' attributes, that is, the presence of invariants.

It is understood that component syndromes, describing semantic relationships between the components of an object, are unchanged, and quantitative attributes are functions of time. So ***initial data for structure detecting is a set of dynamic models of potential components***.

For example, suppose a dynamic model of a situation includes (along with others) points **A, B,** and **C**. These three points can be viewed as a triangle **ABC** with sides whose lengths depend on time: **L**ab(**t**), **L**bc(**t**), **L**ca(t). If

Lab(**t**) / **L**bc(**t**) = **1** and **L**bc(**t**) / Lca(**t**) = **1**

at the time interval (**t**1, **t**2), this means that these three points formed an "***equilateral triangle***" structure/object in the indicated interval. The hypothesis, in this case, is reduced to the choice of three ***points*** as the components of the structure and the assumption of the ***constancy of the values of the two functions*** that determine the ***quantitative relations*** between the components.

The hypothesis test checks that the values of the functions-relations are calculated for a series of moments in time; if they are constant in time, then the situation does not contradict the hypothesis. Note that if in the above example the values of both functions ***are constant but not equal to 1***, this is a structure representing a triangle; the size, position, and orientation may change while the angles at the vertices remain constant. Thus, the constant values of the functions-relations (***invariants of the structure***) are, in fact, the ***attributes*** (quantitative parameters) ***of the new "triangle" concept***.

Obviously, the type of sensor (measuring coordinates in this case) and the choice of possible quantitative relations-functions are closely related: if sensors measure something in space, it is reasonable to construct functions using geometric relations in space. The set of functions used as "building blocks" combined into invariant functions depends on the environment in which the sensors and actuators operate and reflect the fundamental features of the environment ***without being directly related to the system's mission***. It is natural to include both sensors and functions-invariants and fundamental concepts that characterize the environment in congenital knowledge. For example, in the case of the natural environment, they may include universal concepts for space:

- ***position*** (coordinates)
- ***distance***
- ***size***
- ***orientation***

The universal concept of ***size*** and ***position*** for an arbitrary composite object can be the ***diameter of the minimum circumscribed circle and its center*** (in the case of two-dimensional space; a sphere in 3D space). Whatever the mission of a system operating in a natural environment, the above concepts will be used in one way or another. ***Invariants*** for a pair of objects can be:

- ***Relative size*** as a ratio of component sizes.
- ***Relative distance*** as the ratio of the distance to the sum of the component sizes.
- ***Orientation*** relatively to the segment connecting the centers.

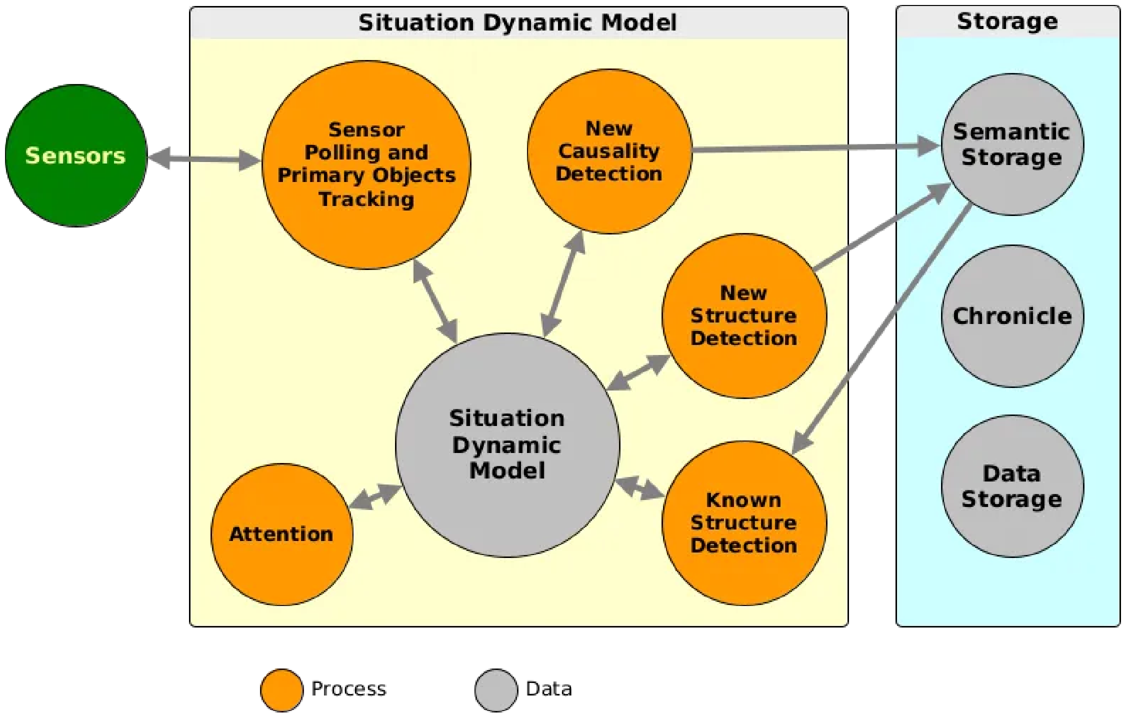
When a previously unknown structure is discovered, its definition must be memorized as a new concept. Two aspects of the created concept are essential: first, the definition is actually a sequence of operations for testing the corresponding hypothesis, that is, the ***definition of structure is a routine***. Secondly, the new concept obviously has no name.

The structures discovered during the analysis of the current dynamic model replace the components that formed up the detected object; recursive application of structure search allows building multilevel structures. The simplest version of the composite concept combines two elements with specific values of the above quantities as invariants.

Discovering and memorizing new structures as new concepts are combined with attempts to detect already known (previously learned and memorized) objects by their descriptions/definitions.

Technically, the search for new structures and the discovery of already known composite objects are performed in parallel. The balance in the distribution of computing resources may vary depending on the situation.

The tasks of detecting known objects and discovering unknown structures are closely related to ***object tracking*** and the ***priorities assigning*** to recognized objects by the ***attention*** module; since the situation model is updated after each significant event, the presence of the most important objects is checked (and their parameters are updated) in the first place:



SUMMATION

- ***New concepts*** *are formed by the AGI system when detecting* ***repeating sequences*** *of events, when detecting* ***cause-effect relationships*** *and when* ***detecting stable structures*** *in a dynamic model of a situation.*
- *The detection of structures, like cause-and-effect relationships, is* ***based on generating hypotheses*** *with their subsequent confirmation or refutation.*
- *The hypothesis of structure` presence is reduced to the assumption of the* ***existence of quantitative invariants*** *(the arguments of which are the attributes of the objects of the dynamic model of the situation).*
- *Definition of structure memorized as a* ***routine*** *that tests the presence of invariants.*
- *The functions used as potential invariants are constructed from the component functions included in the AGI system as part of congenital knowledge (along with the*

description of sensors and actuators) that reflect the ***fundamental properties of the environment and do not depend on the mission****.*

- *Unknown structures discovering and detecting objects with the known structure is* ***efficiently parallelizable****.*
- *The attention module controls the* ***distribution of computing resources*** *between tracking detected objects, identifying objects as the environment changes, and searching for new structures, depending on the assessment of the situation.*

Subscribe to AGI engineering

By Mykola Rabchevskiy · Launched 2 years ago

AGI: fundamentals, architecture, implementation, source code