AGI: KNOWLEDGE REPRESENTATION

Mykola Rabchevskiy

The functioning of an intelligent system consists of the manipulation of knowledge. The width of possibilities for manipulating knowledge, in turn, is determined by how knowledge is represented and is characterized by a ***set of operations on the knowledge*** that is available for the selected representation variant.

Knowledge includes ***logical entities*** and ***relationships between them*** (which are also logical entities) and ***attributes of entities***, which represent parameters of entities that are not relationships: numerical values, words, texts.

The maximum possible set of operations on knowledge includes the ability to ***add and remove logical entities*** from the available knowledge set, ***add and remove relationships*** between entities, and corresponding operations on entity attributes.

We will compare the possibilities for the most famous and used implementations:

- a natural way realized in ***humans and animals***
- knowledge representation systems based on a ***fixed set of rules***
- ***neural networks with an invariable graph of the structure*** of connections
- ***semantic variable structure graphs***
- storing knowledge in the form of ***natural language texts***
- ***hybrid systems*** including components of the types listed above

Potential intelligence capabilities include:

- ***adaptation*** to changing environmental conditions
- detection and identification of ***known*** objects/situations
- detection and memorization of ***unknown*** objects/situations
- detection of ***causation***
- ***exchange*** of information with other systems

The ***detection of unknown*** objects/situations and ***cause-and-effec***t relationships is *de facto* the ***generation of new logical entities*** and together constitute the essence of ***self-learning***. The ability to exchange is, in turn, a necessary element of learning with an "external teacher".

Human and animals

A person's ability to manipulate personal knowledge includes everything described ***except for deliberate forgetting***. The impossibility of voluntary forgetting results in the ***asymmetry*** (well known to psychologists) between the ***replenishment*** of knowledge by adding new concepts and connections between them and the ***correction*** of knowledge, which requires not only the addition of new concepts/relations but also the ***forgetting of mutually exclusive statements***: after correcting previously acquired knowledge, logically ***contradictory statements coexist for some time***. It, however, does not prevent the implementation of all the listed intellectual capabilities. The effectiveness of the realization of intellectual abilities, naturally, is the highest in humans, but to one degree or another, these capabilities are present in animals.

Systems based on a fixed set of rules

The impossibility of changing the rules eliminates the memorization of new entities and relationships while allowing the ability to recognize known objects/situations and adapt.

Fixed structure neural networks

The structure of a neural network is described by a directed graph, as well as semantic graph structure, but the ***technique of using the semantic graph is radically different for these two ways of representing knowledge*** and entails a dramatic difference in capabilities.

In a semantic graph, its ***vertices represent logical entities*** - exactly one vertex corresponds to each logical entity. Accordingly, ***adding/removing*** an entity means adding/deleting a graph's vertex and ***changing its structure***. Adding/removing logical links is performed by adding/deleting the graph's edges and changing its structure. The data attached at the vertices and edges are used only for naming/indicating entities and setting the values of their attributes; that is, the ***graph's structure sets the body of knowledge***. Thus, logical entities in the representation of knowledge by a semantic graph are ***addressable, enumerable, and individually changeable***.

In a neural network, the structure is assigned by the developer and ***remains unchanged*** in the future; knowledge manipulation is reduced to ***varying the numerical parameters*** associated with the vertices of the graph (neurons) and the connections between them. Such a representation can be interpreted as a point of a high-dimensional multidimensional space, and the learning process can be interpreted as moving this point. It is ***impossible to select data related to a specific logical entity or relation*** in such a ***distributed representation*** - the entire set of parameter numbers determines the whole set of "spread" entities and their relationships. Distributed representation has dramatic consequences: ***any change in one of the numerical parameters potentially affects all concepts and relations*** "spread" over the network. This is the reason for the laboriousness of forming acceptable parameters by "training" the network, which provides the required response to possible input data: ***correction of the response to a specific input data set distorts the rest to one degree or another***. Distributed knowledge representation, ***excluding the possibility of selective correction*** of knowledge by adding/ excluding a concept or a link between concepts, ***excludes all the listed intellectual capabilities except for detecting known objects***.