

Glossary

Abduction (*abductive reasoning*). It is a thinking procedure establishing some relation between different knowledge units, in such a way that the established relationship should not necessarily be based on the deep nature of the associated units, but on an external criterion established by the scientist or by merely a practical reason.

Activity Area. This is the label received by any distinctive part of an *archaeological site* where it has been proved that some specific activity was performed.

Activity Theory. It is a social theory focusing on social actions as *practiced* by human actors in reference to other human actors and emphasizing human motivation and purposefulness. It was originally suggested by Leont'ev.

Adaptive algorithms. Instead of being defined a priori from specification, this kind of computer programs uses external data to set automatically their parameters. This means that they are made “aware” of their output through a performance feedback loop, so that the program output improves with respect to the desired *goal*.

Agent-Based Modeling. It is a computational modeling paradigm, in which phenomena are modeled as dynamical systems of interacting “agents”. Each agent is just a computer program, usually implemented as an *Expert System* or a *Neural Network*. It is then a computer program made of interacting computer programs.

Analogy. We refer with this name to the cognitive process of transferring information from a particular subject (the analogue or source) to another particular subject (the target). In a narrower sense, analogy is an *inference* from a particular to another particular. The word *analogy* can also refer to the relation between the source and the target themselves, which is often, though not necessarily, a *similarity* relationship.

A Priori. It refers to the state of some internal model prior to the current learning experience. It contrasts a classical usage of the term, which refers to “God-given” unmodifiable contents that transcend all experience.

Archaeological Data. Data are composed by the subset of *recognized* elements having been *observed* at an archaeological situation. It is the

result of a rational thinking operation on what can be seen at the archaeological site.

Archaeological Record. It is the generic set of all potentially sensed elements perceived at an archaeological situation (*archaeological site excavation, museum collection, laboratory experiment*). Once those *observables* are recognized as particular elements, they become *archaeological data*.

Archaeological Site. The place in physical space where social action was once performed, and some of their material consequences are still preserved, although indirectly.

Archaeometry. Archaeometric data are sets of measurements possible on *archaeological data*. In general, they are referred as chemical and physical determinations of any archaeological artifact.

Archaeozoology. The scientific study of animal bones found at archaeological sites. It includes taxonomic determinations, ecological inferences, and social practices of hunting, herding, butchery, and meat consumption.

Artifact. It is any material consequence of human intentional *action*. Pots and knives are artifacts, as huts, settlements, political territories, and socially modified landscapes. The unconscious material consequences of social action constitute another important category of *archaeological data*, but they are not *artifacts*.

Associationism. This is the name given to the theory that thinking and reasoning are performed in accordance with the law of association, only in terms of simple and ultimate elements derived from sense experiences.

Associative Memory. Content-addressed or associative memory refers to a memory organization in which the memory is accessed by its content (as opposed to an explicit address). Thus, reference clues are “associated” with actual

memory contents until a desirable match (or set of matches) is found.

Automata. They are information-processing machines transforming *inputs* into *outputs*. Simply stated, an automaton is a discrete processing mechanism, characterized by internal states.

Automated Archaeologist. It is a machine able to act as any of us, human archaeologists, learning through experience to associate archaeological observations to explanations, and using those associations to solve archaeological problems. It should have its own “cognitive core” and should interact with some explicitly simulated world to make changes or to sense what is happening. In this book, it has been described as a cognitive robot.

Bayes Theorem (*also known as Bayes’ rule or Bayes’ law*). It is a result in probability theory, which relates the conditional and marginal probability distributions of random variables. The probability of an event A conditional on another event B is generally different from the probability of B conditional on A . However, there is a definite relationship between the two, and Bayes’ theorem is the statement of that relationship.

Bayesian Network. Bayesian networks are directed acyclic graphs whose nodes represent variables, and whose arcs encode the conditional dependencies between the variables. The arcs specify the independence assumptions that must hold between the random variables. These independence assumptions determine what probability information is required to specify the probability distribution among the random variables in the network. According to authors like Pearl and Glymour, they have a causal interpretation. In this case, we are assuming that there is an arrow from X to Y in a causal graph involving a set of variables V just in case X is a direct cause of Y relative to V . The model consists in the causal graph together with the probability distribution of each variable conditional on its direct causes.

Bronze Age. In Europe, it is a chronological period from 2000 B.C. until 700 B.C. It is the time where bronze metallurgy spreads to all this geographical areas. In the Near East, Bronze Age chronology is a bit earlier, from 4000 B.C. until 1200 B.C..

Boolean. This adjective, coined in honor of George Boole, is used along the book to a measurement that results in one of the truth-values ‘true’ or ‘false’, often coded 1 and 0, respectively.

Burial Analysis. The name usually refers to the archaeological study of ancient graves and cemeteries, including the study of human bones, the architecture of the graves, and the nature, frequency and spatial position of grave-goods. The goal is to infer the social organization from the differences in observed funerary practices.

Case-Based Learning. It is a kind of instance-based learning, which represents knowledge in terms of specific cases or experiences and relies on flexible matching methods to retrieve previously memorized cases and apply them to new situations. Decisions are made based on the accumulated experience of successfully solved cases.

Categorization. It is the process in which ideas and objects are *recognized*, differentiated, and *understood*. *Categorization* implies that objects are grouped usually for some specific purpose. Ideally, each group or category illuminates a relationship between the subjects and objects of knowledge. *Classification* and *Clustering* are kinds of categorization.

Cause (causality). It has been defined as “the way an entity becomes what it is”.

Chiefdom. In social evolution, this is a level of social organization, which is just before the formation of state and complex societies. Although such societies have some characteristics of complex structures, like inequality, they are still far from the characteristics of full-scale complex

societies, like class-struggle, coercitive power, capital accumulation, etc.

Classification. It is a form of categorization where the task is to take the descriptive attributes of an observation (or set of observations) and from this to label or identify the observation within a different phenomenological domain. The task of a classifier is to partition this feature space into disjoint regions that each represents a particular class, cluster, or pattern.

Clustering. It is the process of grouping input samples in similarity classes. Clustering algorithms partition the input space so that diversity may be explicitly recognized and encoded.

Computer Vision. It has been defined as a process of *recognizing* elements of interest in an image, and it can be described as the automatic logical deduction of structures or properties of the three-dimensional objects from either a single image or multiple images and the recognition of objects with the help of these properties.

Computational Intelligence. Computational Intelligence is a discipline domain within computer science and cognitive studies, which is based on the hypothesis that reasoning can be realized using computation.

Conceptual Space. This is an analogy that allows understanding a concept as represented in terms of its relationships to the most relevant concepts that are located in the neighborhood in the ordered “representation space”. Consequently, it is defined by a set of quality dimensions, which form the framework used to assign properties to concepts and to specify relationships among them.

Connectionism. Connectionism is a movement in cognitive science, which hopes to explain human intellectual abilities using artificial *neural networks*.

Constraint. In Artificial Intelligence, they are the expressly allowed values for variables. In some cases, problem solving methods are implemented as an evaluation or search that satisfies a restricted set of expressly allowed values.

Composition. It is said of the elements an entity is made of. Not any enumeration of constitutive elements is a *composition*. It is necessary that each component be expressed as a proportion of the total sum of components. Compositions should be expressed as vectors of data, which sum up to a constant, usually proportions or percentages.

Curve Fitting. This is the procedure of finding a curve, which matches a series of data points and possibly other *constraints*, in which the function must go exactly through the data points.

Deduction. In logic, it is a rigorous proof, or derivation, of one statement (the conclusion) from one or more statements (the premises)—*i.e.*, a chain of statements, each of which is either a premise or a consequence of a statement occurring earlier in the proof.

Distributed Representation. This is the most characteristic representation format in a *neural network*, where concepts appear as ephemeral patterns of activation across an entire set of units rather than as individuated elements or symbols. Different patterns capture different aspects of the content of the concepts in a partially overlapping fashion. Alternative concepts are simply alternative patterns of activation.

Emergence. It refers to the way complex systems and patterns arise out of a multiplicity of relatively simple interactions “Emergent” entities (properties or substances) ‘arise’ out of more fundamental entities and yet are ‘novel’ or ‘irreducible’ with respect to them.

Entropy. In information theory, it is a measure of the uncertainty associated with a random variable.

Epistemic Action. According to most dictionaries, “epistemic” is an adjective meaning “knowledge”, “cognitive”; therefore, an epistemic action is some operation involving the creation or transformation of knowledge. *Problem solving* is an example of a series of epistemic actions.

Ethnoarchaeology. This is the study of material consequences of social action as perceived at an observed controlled situation, like in an ethnographic context. It usually implies the use of archaeological tools and methods for studying social *evidence* whose function and origin are known because they have been seen how a living population produced and used them.

Evidence. In its broadest sense, it refers to anything that is used to determine or demonstrate the truth of an assertion. In scientific research, evidence is accumulated through observations of phenomena that occur in the natural world, or which are created as experiments in a laboratory. Archaeological evidence usually goes towards supporting or rejecting a hypothesis. In some cases, it can be used as a synonym for *archaeological record*, or archaeological observables.

Experimentation. This is a scientific method, which tests through repeated controlled experiences the likelihood of some hypothesis. It is usually distinguished from mere observation, because experiences are not merely “observed”, but performed by the observer (the experimentalist).

Expert System. The name refers to a computer program implementing a series of Production Rules (If...Then pairs), which is used to solve diagnostic problems.

Explanandum (*Latin*). In a *problem*, it is the statement that needs to be explained

Explanans (*Latin*). In a *problem*, it is the statement that explains the problem.

Explanation. This is a statement pointing to causes, context, and consequences of some object, process, state of affairs, etc. An explanation can only be given once *understanding* has been reached.

Function. In this book, this term refers to two different concepts. In logics, mathematics, and computer science, it is an abstract entity that associates an input to a corresponding output according to some rule. In Archaeology and biology, it has been argued that to ascribe a *function* to something means to relate it to the “intended” use of something else, or the role it should play in bringing something about. A functional predicate is a logical symbol that may be applied to an object term to produce another object term.

Fuzzy logic. It is derived from fuzzy set theory dealing with reasoning that is approximate rather than precisely deduced from classical predicate logic. Fuzzy truth values represent membership in vaguely defined sets.

Generalization. What is true for a set of elements should be true for all elements that are similar to the prior set, or are related in some way. It is used as a synonym for *induction*.

Genetic Algorithm. This is a computing search technique used to find exact or approximate solutions to optimization problems. It uses analogies inspired by evolutionary biology such as inheritance, mutation, selection, and crossover. Genetic algorithms are implemented as a computer simulation in which a population of abstract representations of candidate solutions evolves toward better solutions.

Geostatistics. It is a branch of statistics that deal specifically with spatial relationships. It involves the analysis and prediction of spatial or temporal phenomena, and it implies a class of techniques used to analyze and predict values of a variable distributed in space or time.

GIS (Geographic Information System).

This is a computer program integrating a spatial database and a cartographic representation, in such a way that spatial data are automatically visualized cartographically, and database queries can be formulated by selecting geographic areas in the map representation.

Goal. This is usually a synonym of an objective or desired outcome. We refer to goals as desired state of affairs of a person or of a system, that is to say, a state of the domain of activity of an intelligent entity which she/he/it tries to achieve.

Gradient. It is the direction of maximum rate of some quantitative values, and a scalar measurement of this rate.

Grave-Goods. The objects and materials placed at a grave and supposed to accompany the dead person.

Heuristic. This is a reasoning procedure based on simple, efficient rules, instead of formal proofs, which have been proposed to explain how people make decisions, come to judgments, and solve problems. Although such rules hardly generate “true” results, they allow obtaining good enough results under most circumstances.

Holocene. This epoch is a geological period, which began approximately 11,550 calendar years BP (about 9600 B.C.) and continues to the present.

Household. More than a synonym of “house”, we use this term in archaeology and anthropology to indicate all activities and work operations usually performed at a domestic level.

Hunter-Gatherer. A kind of human society whose subsistence is satisfied without the cultivation of plants or animal husbandry, but just in terms of hunting wild animals and gathering wild plants.

Iconography. This word literally means “image writing”, and is used to indicate the

identification, description and the interpretation of the content of images. Discussing imagery as iconography implies a critical “reading” of imagery that often attempts to explore social and cultural values.

Induction. It can be defined as the way of concluding that facts similar to those observed are true in cases not examined. Inductive learning tools are trained to recognize patterns or to predict outcomes by *generalizing* from a group of measurements for which the desired outcome is known (training data) to a larger set of circumstances.

Inequality. Applied to a social context, it refers to the degree of social differences in a human community, where not all individuals have the same access to resources or to social life means.

Inference. It is the act or process of deriving a consequence based solely on what one already knows.

Influence Diagram. This is a directed graph, where an arrow denotes an influence expressing available knowledge about the “relevance” of one variable to explain another.

Input. It refers to external information entering into a system.

Intelligence. It refers to any goal-directed functioning.

Intention. An agent’s *intention* in performing an action is their specific purpose in doing so, the end or *goal* they aim at, or intend to accomplish.

Interaction. It is a kind of action that occurs as two or more entities have an effect upon one another. Social interaction can be broadly defined as social relationships generated through the movement of people, commodities, capital, and/or information over geographic space.

Interface. In this book, the term is used to indicate a boundary between two entities. It has also been used to indicate the ways to link entities that may be related, but are not properly connected, for instance, a computer and its human user.

Interpolation. This is a method of constructing new data points from a discrete set of known data points.

Inverse Reasoning. Inverse problems refer to problems in which one has observations on the response, or part of the response, of a system and wishes to use this information to ascertain properties that are more detailed. Inverse reasoning entails determining unknown *causes* based on observation of their effects.

Kinematics. It refers to the study of movements that can be made using a tool or an object.

Labor. In this book, it refers to work of any kind.

Learning. It refers to the acquisition and development of memories and behaviors, including skills, knowledge, understanding, values, and wisdom. It is the product of experience.

Lithics. It refers to those tools made of stone (flint, quartz, obsidian, etc.), which were the only cutting and scrapping tools before the invention of metallurgy.

Localized Representation. This is the most usual representation format, where concepts appear as individuated elements or symbols: one word for each concept.

Luminance. It is a photometric measure of the density of luminous intensity in a given direction. It describes the amount of light passing through or emitted from a particular area.

Machine. It is simply a device, which given a particular input, generates a corresponding output. In other words, it transforms an input signal into an output response.

Machine Learning. This is a discipline concerned with programming computers to optimize a performance criterion using example data or past experience.

Mapping. In mathematics and related technical fields, the term *map* is often a synonym for *function*. In many branches of mathematics, the term denotes a function with a property specific to that branch. In formal logic, the term is sometimes used for a *functional predicate*, whereas a function is a model of such a predicate in set theory.

Mechanism. It is a device designed to perform a particular function. When used out of its proper meaning, this term can be used as an analogy to elements or processes connected in such a way that they produce outputs. Throughout the book, the word it is used to refer to a set of computations designed to bring about a certain outcome.

Megalithism. We refer with this word to different prehistoric phenomena produced since the *Neolithic*: (a) monumentality (great monuments, of stone, earth, or wood); (b) multiple sepultures; (c) megaliths themselves (from Greek: *mega* great, *lith* stone). These phenomena (if they appeared) could have been simultaneous in specific areas or just one or two of them could have taken place. They have enough in common to be referred to with the same word, although in many cases they are different historical events. In Western Europe, megaliths are usually dated from 4500 to 2100 B.C. Some of these monuments could have remained at certain areas until 1000 B.C.

Mesolithic. It is the historical period that preceded the *Neolithic*, and the origin of agriculture and husbandry. In Western Europe, this period is usually dated from 10000 B.C to 5000 B.C.

Modeling. A *model* is a pattern, plan, representation, or description designed to show the structure or workings of an object, system, or concept. Scientific modeling is the process of generating explanatory models of perceived phenomena.

Morphometry. It refers to the measurement and quantitative analysis of *shape* features.

Monotonic. In mathematics, a monotonic function (or monotone function) is a *function*, which tend to move in only one direction as *x* increases or decreases, and therefore is either entirely nonincreasing or nondecreasing.

Neolithic. This is the historical period in which agriculture and animal domestication were discovered. Its chronology is very variable in different parts of the world, ranging from 9000 B.C. in the Near East, to 5000 B.C. or even 3000 B.C. in other regions.

Neural Network. An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems process information. The key element of this paradigm is that information is processed by a large number of *neurons* working in unison to solve specific problems. ANNs, like people, learn by example, adjusting the synaptic connections that exist between the *neurons*. In more practical terms neural networks are nonlinear statistical data modeling or decision making tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data.

Neurocomputing. It is the field of research that deals with behavior of artificial neurons and artificial neural networks.

Neuron. An artificial *neuron* is a mere input-output computing mechanism, where the output is a weighted transformation of incoming input. Throughout the book, the terms *neuron*, *unit*, and *node* are used indistinctly to indicate the same.

Non-Monotonic. This term covers a family of formal frameworks devised to capture and represent *defeasible inference*, i.e., that kind of *inference* of everyday life in which reasoners draw conclusions tentatively, reserving the right to retract them in the light of further information. Such inferences are called “non-monotonic”

because the set of conclusions warranted on the basis of a given knowledge base does not increase (in fact, it can shrink) with the size of the knowledge base itself.

Observation. It is the activity of *sensing*, which assimilates the knowledge of a phenomenon in its framework of previous knowledge and ideas. Observation is more than the bare act of seeing: To perform observation, an agent must seek to add to its knowledge.

Output. The outcome of any function, process, or mechanism.

Parallel Distributed Processing. The prevailing *connectionist* approach today was originally known as Parallel Distributed Processing (PDP). PDP was a neural network approach that stressed the parallel nature of neural processing, and the distributed nature of neural representations.

Perception. In psychology and the cognitive sciences, *perception* is the process of acquiring, interpreting, selecting, and organizing sensory information.

Pixel. A *pixel* (short for picture element) is a single point in a graphic image. Each such information element is not really a dot, nor a square, but an abstract sample. A pixel is generally thought of as the smallest complete sample of an image.

Planning. It refers to the process of thinking about the activities required to create a desired future on some scale. This thought process is essential to the creation and refinement of a plan, or integration of it with other plans. The term is also used to describe the formal procedures used in such an endeavor, the objectives to be met, and the strategy to be followed.

Post Depositional. All kind of disturbance processes that altered the original location and characteristics of archaeological materials after they were originally placed as a consequence of some *social action*.

Problem. We have a problem, when we are in a situation at which an *intention* or *goal* cannot be achieved directly.

Problem Solving. It is any directed sequence of rational cognitive operations intended to achieve some objective.

Processualism. Processual archaeology is a form of archaeological theory advocating the study of processes, that is to say, the way humans did things, and the way things decayed, in terms of an explicitly scientific methodology. Although the processualists are often criticized as omitting the social aspects of human behavior, they were the first archaeologists who re-created archaeology as the study of human behaviors and social processes. The critical reaction to processualism borne in the 1990s is called post-processualism, which is largely based on a critique of the scientific method for studying human and social phenomena. The general critique is that archaeology is not an experimental discipline. Since theories on social behavior cannot be independently verified experimentally then what is considered “true” is simply what seems the most reasonable to archaeologists as a whole. Since archaeologists are not perfectly objective then the conclusions they reach will always be influenced by personal (and social, political) biases.

Prototype. It is an original type, form, or instance of some entity serving as a typical example, for other entities of the same category. When the regularities extracted for a given archaeological data share a common set of attributes, this set can be said to define a prototype.

Recognition. It is a process that occurs in thinking when some event, process, pattern, or object recurs. Thus, in order for something to be recognized, it must be familiar. When the recognizer has correctly responded, this is a measure of *understanding*.

Recursive. A data structure that is partially composed of other instances of the data structure.

Recursive functions are characterized by the process in virtue of which the value of a function for some argument is defined in terms of the value of that function for some other arguments, as well as the values of certain other functions. In order to get the whole process started a certain class of functions needs to be singled out, whose values do not in turn depend of their values for smaller arguments. These are called the *initial* functions.

Remote sensing. In the broadest sense, remote sensing is the short or large-scale acquisition of information of an object or phenomenon, by the use of either recording or real-time sensing device(s) that is not in physical or intimate contact with the object (satellite imaging, magnetic resonance, Laser ranger, etc.)

Robot. It is a mechanical or virtual, artificial agent, which, by its appearance or behavior, conveys a sense that it has *intentions* or agency of its own. The word *robot* can refer to both physical and virtual software agents.

Self-Organizing. The internal organization of a system increases in complexity without being guided or managed by an outside .

Seriation. Formally speaking, it is a way of situating an object within a series. In archaeology, *seriation* is a relative dating method in which artifacts from numerous sites, in the same culture, are placed on chronological order.

Shape. This term refers to the external configuration of some thing — in contrast to the matter, content, or substance of which it is composed. In geometry, two sets have the same shape if one can be transformed to another by a combination of translations, rotations, and uniform scaling. In other words, the *shape* of a set is all the geometrical information that is invariant to location, scale, and rotation. Shape can also be more loosely defined as the “outline”, silhouette, contour, or surface. This definition is consistent with the above, in that the shape of a set does not depend on its position, size

or orientation. However, it does not always imply an exact mathematical transformation.

Similarity. This is the degree of resemblance between two objects or entities. Given that resemblance can be defined as the correspondence in appearance or superficial qualities, similarity can be equated with a measure of correspondences between two entities.

Small-Scale Society. Generally, it is a society of a few dozen to several thousand people who live by foraging wild foods, herding domesticated animals, or non-intensive horticulture on the village level. Such societies lack cities as well as complex economies and governments.

Social Action. It can be defined in terms of purposeful changing of natural and social reality. In fact, it is the pattern of interactions between social agents with the world. Social actions are goal-directed processes that must be undertaken to fulfill some need or motivation. They are conscious (because one holds a goal in mind), and different actions may be undertaken to meet the same goal. However, an action can be an intentional action without the actor having to be aware of the intention from moment to moment.

Soft Computing. It differs from conventional (hard) computing in that, unlike hard computing, it is tolerant of imprecision, uncertainty, partial truth, and approximation. In effect, the role model for soft computing is the human mind. The guiding principle of soft computing is: “exploit the tolerance for imprecision, uncertainty, partial truth, and approximation to achieve tractability, robustness and low solution cost”.

Spatial Analysis. This is the study of spatial location of archeological observations. It tries to discover the existence of regularities and dependencies between places in physical space where archaeological data have been recognized.

Supervised Learning. It is a *machine learning* technique for creating a *function* from training

data, consisting of pairs of *input* objects, and desired *outputs*. The aim is to predict a class label of the input object (called *classification*) after having seen a number of training examples (i.e. pairs of input and target output). To achieve this, the learner has to *induce* from the presented data to unseen situations.

Taphonomy. It is the study of a decaying organism over time. Although the term was introduced to paleontology, it can be used to describe the study of the transition of remains, parts, or products of social action, from its original context, to the archaeological record *i.e.* the creation of fossil assemblages. The primary motivation behind the study of taphonomy is to better understand biases present in the fossil or archaeological record.

Taxonomy. In biology, it is the practice and science of species determination. Taxonomies, or taxonomic schemes, are composed of taxonomic units known as taxa (singular taxon), arranged frequently in a hierarchical structure.

Texture. Visual properties of a surface. It usually corresponds to observed irregularities in color, shape, roughness, etc.

Truth. There is no single definition of truth about which the majority of philosophers agree. A practical definition would be “the way the world really is, and not what I suppose it is”.

Typology. It literally means the study of types. In archaeology, it refers to the *taxonomy* of artifacts according to their characteristics.

Understanding. It is a cognitive process related to an abstract or physical object, such as, evidence, person, situation, or message whereby one is able to think about it and use concepts to deal adequately with that object. Understanding is a set of concepts in the systems cognitive core along with interrelationships between them. The degree of understanding is related to the complexity of a system and to the richness of connections of a given concept to the entire body of knowledge available to the system.

Unsupervised Learning. It is a *machine learning* technique where a model is fit to observations. It is distinguished from *supervised learning* by the fact that there is no *a priori output*. In unsupervised learning, *input* samples are grouped in similarity classes (*clustering*), and in so doing, *prototypes* are built as surrogates for learned concepts.

Vector. Informally speaking, any object that may be scaled and added. It is an array of numbers specifying the dimensionality of some entity.

Vector Quantization. This is a general term used to describe the process of dividing space into several connected regions, using spatial neighborhood as an analogue of similarity.