Glossary

This glossary contains terms important to the study of computational intelligence, neural networks, fuzzy logic, and evolutionary computation. In many cases, these terms have been absorbed from the related disciplines of engineering, mathematics, biology, physics, statistics, and other scientific or business fields.

Activation function (activation rule) An algorithm for computing the activation value of a processing element as a function of its inputs. The activation function is sometimes the weighted sum of the inputs (inner product of the weight vector and the input vector), followed by a threshold or squashing function. More often, it comprises only the threshold or squashing function.

Activation state The collective activation values of a group of processing elements in a neural network.

Activation value The output value of a processing element. Values may be continuous or discrete. Continuous values may be bounded to some interval. Discrete values may be restricted to some small set of values.

Adaline Adaptive linear element, a processing element implementing a weighted sum of inputs. It was developed by Bernard Widrow.

Adaptation Any process whereby a structure is progressively modified to give better performance in its environment (Holland 1992). *Dynamic adaptation* is the ability of a system to adapt in a changing (dynamic) environment.

Adaptive resonance theory Adaptive resonance theory was developed by Grossberg as a theory of human cognitive information processing. It is implemented by binary and analog versions of neural networks.

Algorithm A step-by-step procedure for finding the solution to a problem (Chapman 2000). It is usually readable and understandable by a human.

Allele An individual value a gene may assume.

Ambiguity See *Fuzziness*.

Antecedent The *if* part of an if-then rule. Also called the *precondition* or *assumption*.

Approximate reasoning See *fuzzy inference*.

Architecture The specification of the layers, number of processing elements per layer, and interconnections between the layers in a neural network.

Arity The number of arguments required by a function.

Artificial intelligence (AI) The study of how to make computers do things at which, at the moment, people are better (IEEE Neural Networks Council 1996). AI is concerned with intelligent behavior in artifacts. Intelligent behavior, in turn, involves perception, reasoning, learning, communicating, and acting in complex environments (Nilsson 1998).

Artificial neural network (ANN) An analysis paradigm that is roughly modeled after the massively parallel structure of the brain. It simulates a highly interconnected, parallel computational structure with many relatively simple individual processing elements (PEs).

Associative memory Memory with the property to recall an entire pattern given a partial or noisy pattern.

Autoassociative network A neural network that uses single patterns rather than pattern pairs.

Axon The sending fiber of a neuron. The action potential of the neuron cell body is transmitted along the axon, from which it is received via synapses by the dendrites of other neurons.

Back-propagation A learning rule for multilayer feedforward networks in which weights are adjusted by backward propagation of the error signal from the outputs to the inputs. It uses gradient descent learning.

Batch training (epoch training) A procedure for training neural networks in which weights are adjusted after each epoch. See *Interactive training*.

Bias An additional node in a neural network with a constant activation value. The bias can be connected to any number of layers in a network, via fixed or variable weighted connections, which act as offsets to the processing elements.

Building blocks In a genetic algorithm, this term generally refers to compact (short) schemata that are part of highly fit individuals, and which, with high probability, appear in ever-increasing numbers in future generations.

Cell See *Neuron*.

Cellular automaton (CA) A very simple virtual machine that can result in complex, even lifelike, behavior (Kennedy, Eberhart, and Shi 2001).

Chromosome In evolutionary computation, a chromosome is a string of symbols, bits, and/or real numbers that represents one possible solution in a population of solutions.

Classifier system A system that takes (a set of) inputs and produces (a set of) outputs that indicate some classification of the inputs. An example might take inputs from sensors in a chemical plant and classify them in terms of: "running OK," "needs more water," "needs less water," and "emergency." (IEEE Neural Networks Council 1996)

Closure property In genetic programming, for the closure property to be satisfied, each function must be able to successfully operate on any function in the function set and on any value of any data type assumable by a member of the terminal set.

Codebook vector An input pattern that generates a maximum or nearly maximum activation value for a given output processing element (classification) of a neural network.

Competitive learning An unsupervised learning scheme in which processing elements compete for the right to respond to a given subset of inputs. The response of a processing element to an input pattern tends to inhibit other units. After learning, processing elements become feature detectors. See *Lateral inhibition*.

Complexity The interaction of many parts of a system, giving rise to behaviors and/or properties that are not found in the individual elements of the system (Kennedy, Eberhart, and Shi 2001).

Computational intelligence Comprises practical adaptation concepts, paradigms, algorithms, and implementations that enable or facilitate appropriate actions (intelligent behavior) in complex and changing environments.

Confusion matrix A matrix in which the man-made (rows) and the machine-made (columns) classifications per class are entered. A perfect classifier has only the diagonal populated. Errors appear in nondiagonal positions. The confusion matrix is an efficient way to observe the separability between classes (Principe, Euliano, and Lefebvre 2000).

Connection (link) A unidirectional information pathway between processing elements in a neural network.

Connection strength (weight) The strength of a connection between two processing elements, which determines the net effect of one processing element on the other. Connections can have a positive (excitatory), zero, or negative (inhibitory) effect.

Connectionism Generally related to neural networks. Refers to highly distributed systems in which computations are based on the inter-PE weights (connections).

Consequent The *then* part of an if–then rule. Also called the *conclusion*.

Constriction factor A coefficient used in particle swarms to ensure convergence of the particle swarm optimization algorithm.

Content addressable memory See *Associative memory.*

Converged (as used in evolutionary computation) A gene is said to have converged when a high percentage of the chromosomes in the population all contain the same allele for that gene. In some circumstances, a population can be said to have converged when all genes have converged. The term is sometimes used (or misused) informally to indicate that an evolutionary algorithm population has not found an improved solution in a large number of generations.

Core In a fuzzy set, the crisp set of objects that have a membership of 1 in the fuzzy set. It can also be defined as the alpha-cut when alpha = 1 (IEEE Neural Networks Council 1995).

Crisp logic The 0 and 1 logic used by digital computers. Sometimes called Aristotelian logic. Something is either *true* or *false*, *on* or *off*. Truth is all or nothing, absolutely true or absolutely false, with no middle ground possible.

Crossover In genetic algorithms, formation of a new chromosome when a reproduction operator combines parts of each of two "parent" chromosomes. The simplest form is single-point crossover, in which an arbitrary point in the chromosome is picked. All the information from parent A is copied from the startup to the crossover point; then all the information from parent B is copied from the crossover point to the end of the chromosome. The new chromosome thus gets the head of one parent's chromosome combined with the tail of the other. Variations exist that use more than one crossover point or combine information from parents in other ways. (IEEE Neural Networks Council 1996)

Crossover point In fuzzy logic, the lowest membership point between two adjacent membership functions (linguistic terms) over a domain.

Crossover rate In a genetic algorithm, the probability that given two parent chromosomes, the crossover process will occur.

Deception The condition where the combination of good building blocks leads to reduced fitness rather than increased fitness. Proposed by Goldberg (1989) as a reason for the failure of GAs on many tasks. (IEEE Neural Networks Council 1996)

Defuzzification A fuzzy set maps an object into a membership value. The inverse of this process takes a fuzzy set, a set of fuzzy sets, or some system that has fuzzy uncertainty, and produces a single crisp object that is most representative in some well-defined

mathematical sense of the set, sets, or system. Several methods are used for defuzzification; most involve computing some type of a mean or mode of a fuzzy set, or sets, or system (such as center-of-area, mean-of-maximum, or height-based methods). (after IEEE Neural Networks Council 1995)

Degree of membership See *membership function*.

Delta rule See *Widrow–Hoff rule*.

Dendrite The receiving fibers of a neuron. Dendrites receive action potentials via synapses from other neurons or from the environment.

Discrete recombination In evolution strategies, an operation that comprises selection of parameter values from either of two parents.

Distributed representation An information representation scheme in which entities are represented by patterns of activity, distributed over many processing elements. Long-term memory is represented by connection weights. Short-term memory is represented in recurrent networks by the activation states of processing elements. Stored knowledge cannot be isolated to a single location as in the Von Neumann computer. The robustness of neural networks is due to this property. See *parallel distributed processing*.

Dot-product See *inner product*.

Elitist strategy In a genetic algorithm, ensuring that the individual chromosome with the highest fitness is always copied into the next generation.

Entropy The tendency of systems to lose energy and order and to settle to more homogeneous states (Kennedy, Eberhart, and Shi 2001).

Epistasis In evolutionary computation, any kind of strong interaction among genes, not just masking effects. A possible definition is the interaction between different genes in a chromosome. It is the extent to which the contribution to fitness of one gene depends on the values of other genes. Problems with little or no epistasis are trivial to solve (hill-climbing is sufficient). But highly epistatic problems are difficult to solve, even for GAs. High epistasis means that building blocks cannot form, and there will be deception. (IEEE Neural Networks Council 1996)

Epoch The presentation of a complete set of patterns to a neural network.

Epoch training See *batch training*.

Error term (**error signal**) A measure of the difference between the observed state and the desired state of a processing element.

Euclidean distance The geometric distance between two points, given by the square root of the sum of the squared differences between vector components.

Euclidean normalization The normalization of a vector to unit length, obtained by dividing each component of the vector by the vector length.

Evolution The process of change that is assured given a reproductive population in which there are varieties of individuals, with some varieties being heritable, of which some varieties differ in fitness (reproductive success). (IEEE Neural Networks Council 1996)

Evolution strategies A type of evolutionary algorithm developed in the early 1960s in Germany. It employs real-coded parameters and, in its original form, relied on mutation as the search operator and a population size of one. Since then it has evolved to share many features with genetic algorithms. (IEEE Neural Networks Council 1996)

Evolution window In evolution strategies, the step-band size, or window, within which mutation operations result in fitness improvement.

Evolutionary computation (EC) Machine learning optimization and classification paradigms roughly based on mechanisms of evolution such as biological genetics and natural selection. The evolutionary computation field includes genetic algorithms, evolutionary programming, genetic programming, evolution strategies, and particle swarm optimization.

Evolutionary programming An evolutionary algorithm developed in the mid-1960s. It is a stochastic optimization strategy, which is similar to genetic algorithms but dispenses with both "genomic" representations and crossover as a reproduction operator. (IEEE Neural Networks Council 1996)

Exploitation When traversing a search space, the process of using information gathered from previously visited points in the search space to determine which places might be profitable to visit next. An example is hill-climbing, which investigates adjacent points in the search space and moves in the direction giving the greatest increase in fitness. Exploitation techniques are good at finding local maxima. (IEEE Neural Networks Council 1996)

Exploration The process of visiting entirely new regions of a search space to see if anything promising may be found there. Unlike exploitation, exploration involves leaps into the unknown. Problems that have many local maxima can sometimes only be solved by this sort of random search. (IEEE Neural Networks Council 1996)

Fan-in The number of connections to the input of a processing element or unit.

Fan-out The number of connections from the output of a processing element or unit.

Feedback network (recurrent network) A network with feedback paths. Feedback can occur within layers or between layers. The current activation state of a layer is a function of the previous activation state and the current inputs.

Feedforward network A network ordered into layers with no feedback paths. The ordering is from input (lowest), through one or more hidden layers, to the output layer (highest).The outputs of a given layer are connected to higher layers only, and its inputs originate from lower layers only.

Finite state machine A transducer that can be stimulated by a finite alphabet of input symbols, can respond in a finite alphabet of output signals, and possesses some finite number of different internal states (Fogel 1991).

Fitness A value assigned to an individual (candidate solution) that reflects how well the individual solves the problem being addressed, usually by an evolutionary algorithm.

Full approach In genetic programming, each limb of the program tree extends for the maximum allowed depth.

Function An expression that accepts one or more input values and calculates a single result from them.

Function space Contains the ranges of results of functions (Kennedy, Eberhart, and Shi 2001).

Fuzzification Combining the antecedent fuzzy sets, that is, the sets on the "if-" side of a fuzzy rule, using fuzzy operators.

Fuzziness Refers to nonstatistical imprecision and vagueness in information and data.

Fuzzy inference The main tool for applying fuzzy logic, comprising evaluation of if–then rules consisting of fuzzy antecedents and consequents (fuzzy propositions). It features parallel firing of all rules.

Fuzzy logic (FL) One of many logics of "approximate reasoning." Fuzzy logic comprises operations on fuzzy sets, including equality, containment, complementation, intersection, and union; it is a generalization of conventional (two-valued, or "crisp") logic.

Fuzzy-neural system This form combines fuzziness and neural algorithms, but the focus is on a traditional neural network whose operations have been "fuzzified" rather than having neural nodes which implement the operations required for fuzzy inference. Since a conventional neuron performs a weighted sum of its inputs and uses a (possibly soft) threshold for outputting a value, a fuzzy neuron does the same operation except that all inputs, weights, and operations may be defined on fuzzy numbers. Neural algorithms such as back-propagation can be extended to cover the fuzzified network as well. (IEEE Neural Networks Council 1995)

Fuzzy reasoning See *fuzzy inference*.

Fuzzy set A real-valued function mapping a collection of objects, or the points within the dynamic range of a variable, to values between 0 and 1. The value associated with each object or point is the degree to which that object or point belongs to the fuzzy set. Unlike a *crisp set*, an object or point may belong partially to the fuzzy set. Examples of fuzzy sets are membership functions that represent linguistic predicates such as small and tall, or a notation such as $x >> 1$ (whereas $x > 1$ would define a crisp set). (after IEEE Neural Networks Council 1995)

Gene A specific segment of a chromosome that is associated with a specific function. In evolutionary computation, a subsection of a chromosome that (usually) encodes the value of a single parameter. (IEEE Neural Networks Council 1996)

Generalization The property of a system that abstracts features from the input space to define a particular class.

Generalized delta rule The delta rule for semilinear activation functions in layered feedforward networks.

Generation In an evolutionary algorithm, an iteration of the measurement of fitness and the creation of a new population by means of reproduction operators. (after IEEE Neural Networks Council 1996)

Generation gap In an evolutionary algorithm, the percentage of the population replaced in each generation.

Genetic algorithm (GA) A search algorithm that imitates natural evolution mechanisms, including crossover, mutation, and survival of the fittest. Genetic algorithm paradigms work on populations of individuals rather than on single data points or vectors. They are more often used for optimization but also for classification. Implementations typically use fixed-length character strings to represent their genetic information, together with a population of individuals that undergo crossover and mutation in order to find promising regions of the search space. (after IEEE Neural Networks Council 1996)

Genetic programming A version of an evolutionary algorithm used to directly evolve computer programs. The original version of genetic programming utilized LISP expressions to evolve hierarchical tree structure programs.

Genetic operator A search operator acting on a coding structure that is analogous to a genotype of an organism (e.g., a chromosome) (IEEE Neural Networks Council 1996).

Genotype The collection of chromosomes required to completely specify an organism. See also *structure*.

Genome The entire collection of genes (and hence chromosomes) possessed by an organism. (IEEE Neural Networks Council 1996)

Global method In evolution strategies, the use of the entire population of individuals as potential sources from which individual components for the new individual can be obtained.

Gradient descent An algorithm for minimizing some error function by making small incremental adjustments toward a local or global optimum. Also called the steepest descent.

Gray coding A kind of binary coding that removes the Hamming cliffs (Kennedy, Eberhart, and Shi 2001).

Hamming cliff In order for a binary number to increment by one numeric unit, it is sometimes necessary to flip all its bits, for example, 01111 (15) to 10000 (16). A numeric algorithm may encounter difficulty moving over this "cliff" (Kennedy, Eberhart, and Shi 2001).

Heteroassociative network A neural network that uses pattern pairs.

Hebbian learning rule A fundamental law of learning, formulated by Hebb. Essentially, the connection strength between two processing elements is adjusted in proportion to the product of their activation values. Many other learning rules are founded on Hebb's original rule.

Hidden layer A layer of processing elements in a neural network with no direct connections to the environment. All connections from a hidden layer are to other layers within the system.

Hierarchical network A network consisting of several layers with connections between layers chosen to achieve some kind of processing in stages. Such networks can have feedforward and/or feedback connections and can combine different learning paradigms at each stage.

Hill climbing An algorithm for maximizing some function using gradient descent. See *gradient descent.*

Hinton diagram A compact-graphical display of activation values or weights in a neural network developed by Hinton. For example, values can be represented by rectangles whose sizes are proportional to the magnitudes and whose signs are represented by colors.

Hyperplane A surface defined in hyperspace.

Hyperspace A Cartesian coordinate space of high dimension, typically higher than three dimensions.

Individual In evolutionary computation, a single member of a population. Each individual represents a possible solution to the problem being solved; it is a single point in the problem space (Kennedy, Eberhart, and Shi 2001).

Inertia weight A coefficient used in particle swarms to enhance global exploration at the beginning of a run and local exploitation near the end of the run.

Inner product (dot product, scalar product) The scalar sum of the products of the components of two vectors. The activation value of a processing element is typically a function of the inner product of the weight and input vectors.

Input layer A layer of processing elements receiving inputs from the external environment. These inputs may be sensory inputs or signals from other systems outside the one being modeled.

Input vector The set of inputs to a processing element of a neural network or to the network as a whole.

Intensification A fuzzy logic linguistic hedge concept which is a kind of combination of concentration and dilation. For original membership values between 0.5 and 1.0, membership values are increased, while original membership values between 0.0 and 0.5 are decreased.

Interactive training A procedure for training neural networks in which the weights are adjusted after each pattern is presented. See also *batch training*.

Interconnections per second A performance measure of a neural network. It is a function of the number of processing element interconnections calculated per second.

Iteration In a neural network, the process of setting activation states of layers and applying their activation rules according to the update procedure.

Intermediate recombination In evolution strategies, it involves setting each parameter value for a child at a point between (often midway between) the values for the two parents.

Lateral inhibition The inhibitory action of processing elements within a neural network layer, whereby strong positive activation of one processing element causes negative influence on the activations of neighboring processing elements. Typically used in competitive learning.

Law of the excluded middle A law from traditional logic which states that the union of a set with its complement results in a universal set of the underlying domain. The law does not hold and is irrelevant in fuzzy logic.

Law of noncontradiction A law from traditional logic which states that the intersection of a set with its complement results in an empty or null set. It does not hold and is irrelevant in fuzzy logic.

Law of sufficiency If a solution is good enough, fast enough, and cheap enough, it is sufficient. (Being good enough simply means it meets specifications.)

Layer A slab in a network with strict hierarchical ordering between groups of elements. See *Slab*.

Learning A procedure whereby a neural network adapts its weights by successive applications of the learning rules. The term *adaptation* is preferred to describe this procedure.

Learning rate A parameter used to regulate the magnitude of weight changes during learning in a neural network.

Learning rule An algorithm for adjusting the weights and connections of a neural network based on inputs from the environment.

Least mean square rule See *Widrow–Hoff rule*.

Linguistic hedge A linguistic variable that changes the shape or position of a membership function. (Zadeh 1972)

Linguistic variable A variable whose values are words or sentences in a natural or artificial language. (Zadeh 1975)

Link See *Connection*.

Local method of recombination In evolution strategies, this involves forming one new individual using components from two randomly selected parents.

Local storage Storage local to a given processing element that is not accessible to other units. Local storage includes the weight vector, holding the weights for the input connections, and other variables required in computing outputs.

Membership function The fuzzy set is completely defined by the real-valued mapping from a set of objects to [0,1]. This mapping is also known as the membership function of the fuzzy set. It is frequently represented as $\mu_A(x)$: *X maps to* [0,1], where *X* is the *universe of discourse* and *A* is the fuzzy set with membership function $\mu_A(x)$. These functions can have discrete or infinite domains, but if they map to only 0 or 1, they represent crisp sets. Example shapes of the graphs used for membership functions are triangular, trapezoidal, Gaussian, and sigmoid. Each is characterized by a few numerical parameters, including *center, support, core*, and *curvature*. These may be set manually or may be tuned automatically, depending on the availability of expertise and data. (after IEEE Neural Networks Council 1995)

Mexican hat A function used in self-organizing networks to adjust the weights of processing elements in the neighborhood of a winning processing element. Weights at the center are most strongly excited; weights farther away are slightly inhibited; and weights a long way out are not changed at all. The distribution of these weight changes resembles a Mexican hat.

Momentum factor A parameter used to promote stability of learning in a neural network. To dampen possible oscillations, weight changes are moderated by a term proportional to the previous weight change and the momentum factor.

Multimodality Exists when a problem has multiple solutions, global optima, or good local optima (Kennedy, Eberhart, and Shi 2001).

Mutation An evolutionary algorithm operator that forms a new chromosome by making (usually small) random alterations to the values of genes in a parent chromosome (after IEEE Neural Networks Council 1996).

Neighborhood-neural networks The set of processing elements topologically adjacent to a particular processing element in a neural network, especially a self-organizing feature map. Particle swarm optimization: The set of particles topologically adjacent to a particle in a particle swarm. The neighborhood generally remains constant during a run.

Neural-fuzzy system Any system that uses a combination of fuzzy logic and neural networks can be referred to as a neural-fuzzy system. The two may be different subsystems that are merely loosely coupled, as when they are used in series or parallel or to control each other's parameters. The advantage of combining the two forms is that top-down knowledge-driven rules can be combined with bottom-up data-driven processing. (IEEE Neural Networks Council 1995)

Neural network An adaptive, generally nonlinear distributed system that is a hardware or software implementation of a neural network algorithm.

Neural network attributes The distinguishing characteristics of a neural network, describing its architecture, activation rule, learning rule, and update procedure.

Neural network model The abstract specification of a neural network paradigm. A model is independent of any implementation and can be simulated in software or implemented as a neurocomputer.

Neural network paradigm A particular set of neural network attributes. An example is *back-propagation.*

Neurocomputer A hardware implementation of a neural network, using electronic, optical, or other components.

Neuro-fuzzy system A *neural-fuzzy system* where the fuzzy logic inference and neural algorithms are integrated into a single form, rather than being maintained as two subsystems interacting with each other. Usually, this combination takes the form of a multilayer neural network that maps inputs to outputs. The individual nodes in the network perform the processing associated with the stages of *fuzzy inference*. This includes representing the membership functions in the nodes, performing partial match, *min* or *product* operations, *defuzzification*, and others. The forward pass of the neural net implements fuzzy inference, and the backward pass and learning change the parameters of the fuzzy membership functions and the other operations. Such a neuro-fuzzy combination may be used for supervised, reinforcement, or unsupervised learning. (IEEE Neural Networks Council 1995)

Neurode See *Processing element*.

Neuron A nerve cell composed of a body, an axon, and dendrites. The fundamental unit of neural activity in biological systems. In artificial neural networks, a neuron is modeled by a processing element with weighted connections.

Node See *Processing element*.

Normalized fitness In an evolutionary algorithm, an individual's normalized fitness is its raw fitness value divided by the sum of the raw fitness values for all population members.

Normalization of vectors The adjustment of vector components to a limited range, typically 0 to 1 or −1 to 1. See also *Euclidean normalization*.

NP-complete A problem sufficiently complex that any deterministic search technique that completely searches the problem domain will probably not find an acceptable answer in an acceptable time.

NP-hard See *NP-complete*.

Offspring In evolutionary computation, an individual generated by any process of reproduction (IEEE Neural Networks Council 1996).

One-point crossover In a genetic algorithm, one-point crossover involves selecting a single crossover point at random and exchanging the portions of the individual strings to the right of the crossover point.

Operator A rule for changing a proposed problem solution.

Ordinal A number that designates the place (1^{st}, 2^{nd}, 3^{rd}, etc.) occupied by a measurement in an ordered sequence.

Overselection In genetic programming, for populations of 1,000 or more, highly fit individuals are sometimes given an even greater probability of selection than their normalized fitness would indicate.

Output layer In a neural network, a layer of processing elements sending output signals to the external environment. These may be motor signals or signals sent to systems external to the system being modeled.

Output vector In a neural network, the set of activation values of a layer.

Parallel distributed processing A processing paradigm consisting of a network of many processing elements, operating in parallel, and information storage distributed across the network.

Parameter space Comprise of the legal values of all of the elements that can be entered into a function to be optimized (Kennedy, Eberhart, and Shi 2001).

Parent In evolutionary computation, an individual that takes part in reproduction to generate one or more other individuals, known as offspring, or children (IEEE Neural Networks Council 1996).

Particle An infinitesimal, moving, usually multidimensional representation of a vector (Kennedy, Eberhart, and Shi 2001).

Particle swarm optimization (PSO) A form of swarm intelligence in which each individual (called a particle) in the population flies stochastically toward its personal best position and its neighbor's best position found so far, so that the whole population has the tendency to fly toward better and better search areas.

Pattern See *Input vector*.

Perceptron A simple neural network consisting of an input layer connected to a single processing element. The activation function of this unit is a linear threshold function, applied to the inner product of the input and weight vectors.

Perceptron convergence procedure A learning rule for a perceptron.

Phenotype The expressed traits of an individual (IEEE Neural Networks Council 1996).

Population In evolutionary computation, a group of individuals that may interact, for example, by mating and producing offspring. Typical population sizes in EC range from 1 (for certain evolution strategies) to many thousands (for genetic programming). (IEEE Neural Networks Council 1996)

Position (as used in particle swarm optimization) The location representation of a particle in a population. See *Individual*.

Positive predictive value The likelihood that a signal of an event is associated with the event; give that a signal occurred.

Presentation of a pattern The application of a pattern to the input layer of a neural network.

Processing element (PE, unit, node, neurode, cell) An active computational unit of a neural network. A processing element consists of an activation function, a set of input connections, a single output (which can be distributed to any number of other processing elements) and local storage. The processing element computes an output value when activated by an update procedure.

Radial basis function A function $f(x)$ defined over a metric space is radially symmetric if there exists a center c such that $f(x)$ depends only on the distance between x and c and not on the direction of the vector from x to c. A function is said to be a basis function if it is a member of a family of functions that are linearly independent and span the space—any "reasonable" function on that space can be expressed as a linear combination of these basis functions. (IEEE Neural Networks Council 1995)

Ramped half-and-half method In genetic programming, a method of creating programs with evenly distributed depth parameters. Within each subpopulation of a given depth, one-half of the programs are created using the grow approach, and one-half using the full approach. (Kennedy, Eberhart, and Shi 2001)

Recombination In evolution strategies, the method used to exchange a portion of each individual between two population members. It is analogous to crossover in a genetic algorithm.

Recurrent network See *Feedback network*.

Region In neural networks, a group of processing elements with common attributes. See also *Slab*.

Reinforcement adaptation (reinforcement learning) A variation of supervised learning, in which a network is given a correct or incorrect signal for each input pattern.

Reproduction The creation of a new individual from two parents (sexual reproduction). *Asexual reproduction* is the creation of a new individual from a single parent. (IEEE Neural Networks Council 1996)

Reproduction operator A mechanism that influences the way in which genetic information is passed on from parent(s) to offspring during reproduction. Reproduction operators fall into three broad categories: crossover, mutation, and reordering operators. (IEEE Neural Networks Council 1996)

Response See *Output vector*.

Scaling To adjust a set of values in some orderly way (linear, logarithmic, etc.) so that they fall between two specified end points. For example, 100 data points with a minimum value of 3.1 and a maximum value of 125.4 could be scaled to be between 1 and 2. Note that this is not the same thing as normalization.

Schemata (singular: **schema**) In a genetic algorithm, each schema defines a subset of strings with identical values at specified string locations. Schemata provide a means by which similarities among the individual population members can be described and exploited.

Search space If the solution to a task can be represented by a set of N real-valued parameters, then the job of finding this solution can be thought of as a search in an N-dimensional space. This is referred to simply as the search space. More generally, if the solution to a task can be represented using a representation scheme, R, then the search space is the set of all possible configurations that may be represented in R. (IEEE Neural Networks Council 1996)

Selection A number of problem solutions (chromosomes or patterns of features) are proposed and tested; those that do well in the test are more likely to produce offspring in the next generation, while those that perform poorly are more likely to be eliminated. This is what is meant by "survival of the fittest."

Self-organization The ability of some systems to generate their form without external pressures, either wholly or in part. It can be viewed as a system's incessant attempts to organize itself into ever more complex structures, even in the face of the incessant forces of dissolution described by the second law of thermodynamics.

Self-organizing feature map An unsupervised learning neural network developed by Kohonen (1982a, 1982b). It adjusts its weights to reflect the probability distribution of features that are found in the input patterns and implements the concept of a

neighborhood that shrinks over time; weights are adjusted according to the location with respect to the neighborhood.

Sequential training A procedure for training neural networks in which weights are adjusted after each input of a training pattern.

Sigmoid function A nonlinear activation function whose output is a nondecreasing and differentiable function of the input, with maximum and minimum saturation values. See also *Squashing function*.

Simulated annealing An algorithm for the optimization of generic functions. The step length for each iteration is set by an arbitrarily set parameter that plays the role of temperature, as in the annealing of metals. The temperature is higher at first to facilitate exploration, then gradually reduced to facilitate exploitation. If a step, which may be deterministic or random, results in an increase of fitness, it is always accepted. If the step decreases fitness, it is accepted with a probability proportional to the temperature parameter.

Slab A group of processing elements that share the activation function and learning rule and have equivalent connection topologies. A network composed of slabs allows arbitrary connections between slabs, with no implied hierarchical ordering. A slab can be used to model any network topology.

Squashing function A function whose value is always between finite limits, even when the input is unbounded. See also *Sigmoid function, Threshold function*.

Structure In computational intelligence, the collection of patterns or strings needed to completely specify a system. More specifically, a candidate solution to a problem. See also *genotype*.

Sufficiency property In genetic programming, for the sufficiency property to be satisfied, the set of functions and set of terminals must be sufficiently extensive to allow a solution to be evolved.

Sum-squared error A measure of the error of a network or other computational intelligence implementation for a given set of output, target pairs. The error is the sum over all patterns of the square of the difference between the target (desired) response and the actual output response.

Supervised adaptation (supervised learning) An adaptation procedure in which a neural network is presented with a set of input pattern, target pairs. The network compares its outputs to the target values and adapts its weights by applying the adaptation rules. See also *Unsupervised adaptation*.

Support The support of a fuzzy set is the crisp set of objects that have a membership > 0 in the fuzzy set (IEEE Neural Networks Council 1995).

Swarm intelligence (SI) The adaptive behavior of a population of simple entities that interact either directly or indirectly. Swarm-intelligent systems often solve problems, converging or clustering in a problem space, and are usually stochastic. They often exhibit the property of self-organization.

Synapse The contact point between neurons in which the dendrites of several neurons are attached to an axon through the synaptic cleft.

Synaptic weight See *Weight*.

Takagi–Sugeno–Kang (TSK) rules These are fuzzy if–then rules where the consequent (or *then* part) is a set of functions of the inputs. Typically, each output function is a

polynomial, often a constant or linear function. The antecedent of the rule continues to be a conjunction of fuzzy sets. The TSK form can be naturally used for modeling a set of linear control laws, each having a different regime of applicability, as is the case for a hierarchical or mode-selection controller. Using TSK rules to blend the outputs leads to a smooth interpolation over the entire state space, producing mode melding instead of mode selection. (after IEEE Neural Networks Council 1995)

Target The desired response of a neural network corresponding to a given input pattern. See also *Supervised learning*.

Threshold function A function with two output levels, one for input below the threshold and one for input above the threshold.

Topology In a neural network, the pattern of PEs and interconnections, together with other attributes such as direction of data flow and PE activation functions.

Topology preserving map A neural network representation that preserves topological features of the environment such that the weight vectors are distributed to resemble the distribution of the input vectors.

Uniform crossover In a genetic algorithm, a random decision is made at each bit position in the string as to whether or not to exchange (cross over) bits between the parent strings (Kennedy, Eberhart, and Shi 2001).

Unit See *Processing element*.

Unsupervised adaptation (unsupervised learning) An adaptation procedure in which a neural network is presented with a set of input patterns, and the network adapts itself according to the statistical characteristics of the input patterns. See also *Supervised adaptation*.

Update procedure The timing of the application of the activation functions of elements in the network. The update procedure can be synchronous, in which case all activation values are updated simultaneously, or asynchronous, in which case activation values are updated in random sequence. Hierarchical networks use more elaborate update procedures, in which some layers are not activated at all until other layers have been sufficiently trained.

Vector length (magnitude) The Euclidean distance between the end points of the vector.

Velocity (as used in particle swarm optimization) The change rate of a particle's position, which is usually dynamically adjusted according to the particle's and its neighbor's flying experiences.

Weight See *Connection strength*.

Weight vector The set of weights of the input connections to a processing element. The dimension (number of components) of the weight vector is given by the fan-in of the processing element. A weight vector can also include an additional component for a bias input.

Widrow–Hoff rule A learning rule in which change of weight is proportional to the difference between the actual activation and the desired activation. The rule leads to minimization of mean-squared error.