Newell A (1980) Reasoning, problem solving and decision processes: the problem space as a fundamental category. In: Nickerson RS (ed.) *Attention and Performance VIII*, pp. 693–718. Hillsdale, NJ: Lawrence Erlbaum.
Rescher N (1980) *Induction: an Essay on the Justification of Inductive Reasoning*. Pittsburgh, CA: University of Pittsburgh Press.
Rips LJ (1994) *The Psychology of Proof: Deductive Reasoning in Human Thinking*. Cambridge, MA: MIT Press.
Sloman SA (1996) The empirical case for two systems of reasoning. *Psychological Bulletin* **119**: 3–22.

Reasoning under Uncertainty

Intermediate article

Francisco J Díez, UNED, Madrid, Spain
Marek J Druzdzel, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

CONTENTS

Most artificial intelligence applications, especially expert systems, have to reason and make decisions based on uncertain data and uncertain models. For this reason, several methods have been proposed for reasoning with different kinds of uncertainty.

INTRODUCTION

We often have to make decisions based on uncertain knowledge, not only in our private lives (which job to take, which house to buy, where to invest our money) but also in professional activities, such as medicine, economics, politics, engineering, and education. Therefore, any reasoning method that tries to replicate human reasoning must be able to draw conclusions from uncertain models and uncertain data. Models may be uncertain because of indeterminism in the real world or because of our lack of knowledge. Furthermore, data may be incomplete (pieces of information may be not available in a diagnostic case), ambiguous (a pronoun in a sentence may refer to different subjects), erroneous (patients may lie to their doctors, or sensors may be faulty), or imprecise (because of the limited precision of measuring devices, subjective estimations, or natural language).

This article reviews some of the uncertain reasoning methods that have been proposed in the field of artificial intelligence.

NAIVE BAYES

The oldest method applied in uncertain reasoning is probability theory. Probabilistic reasoning concentrates basically on computing the *posterior probability* of the variables of interest given the available *evidence*. In medicine, for example, the evidence consists of symptoms, signs, clinical history, and laboratory tests. A diagnostician may be interested in the probability that a patient suffers from a certain disease. In mineral prospecting, we may wish to know the posterior probability of the presence of a certain deposit given a set of geological findings. In computer vision, we might be interested in the probability that a certain object is present in an image given observation of certain shapes or shadows.

The probability of the diagnoses given the available evidence can be computed by the generalization of Bayes' theorem to several variables. However, the direct application of this method would need a prohibitive number of parameters (probabilities), which grows exponentially with the number of variables involved. Two assumptions were introduced to simplify the model. The first assumption is that the diagnoses are mutually exclusive; i.e. each patient can suffer from at most one disease and each device can have at most one failure at a time. It is then possible to consider a variable D taking n values, as many as the number of possible diagnoses. The second assumption is that the findings are conditionally independent given each diagnosis d_i, so:

$$P(f_1, \ldots, f_m \mid d_i) = P(f_1 \mid d_i) \cdot \ldots \cdot P(f_m \mid d_i), \quad 1 \le i \le n \tag{1}$$

where f_k represents one of the possible values of a finding F_k. Under these assumptions, the posterior probability of d_i can be computed as follows:

$$P(d_i \mid f_1, \ldots, f_m) = \frac{P(d_i) \cdot P(f_1|d_i) \cdot \ldots \cdot P(f_m|d_i)}{\sum_j P(d_j) \cdot P(f_1|d_j) \cdot \ldots \cdot P(f_m|d_j)} \tag{2}$$

In this simplified model, the number of parameters is proportional to the number of variables: it requires n prior probabilities $P(d_i)$ plus, in the case of m dichotomous findings, $n \times m$ conditional probabilities $P(f_k \mid d_i)$.

Although this method was used in the construction of diagnostic medical systems in the 1960s, it was severely criticized because its assumptions are usually unrealistic (Szolovits and Pauker, 1978). In fact, the assumption of exclusive diagnoses is a reasonable approximation only when the probability of the simultaneous presence of two diseases or two failures is very low. In medicine, however, it is common for a patient to suffer from multiple disorders. Also, the assumption of conditional independence is unrealistic when there are causal associations among findings other than those due to the diagnoses included in the model. Because of the crudeness of such assumptions, this method is often called 'naive Bayes' or 'idiot Bayes'. Even under these assumptions, the model still requires a large number of parameters, which may be difficult to obtain.

MYCIN'S CERTAINTY FACTORS

MYCIN was a rule-based expert system developed in the 1970s as a decision support tool for antibacterial therapy (Buchanan and Shortliffe, 1984). To accommodate uncertainty, MYCIN associated a certainty factor with each rule and, consequently, with each proposition.

The certainty factor of each rule 'if E then H', $CF(H,E)$, is a measure of the degree to which evidence E confirms hypothesis H. When E increases the probability of H, so that $P(H|E) > P(H)$, then $0 < CF(H,E) \le 1$. The higher the increase in probability, the higher the certainty factor. When E contributes evidence against H, so that $P(H|E) < P(H)$, then $-1 \le CF(H,E) < 0$. The value of each certainty factor in MYCIN's rules was obtained from human experts when formulating the rules.

Analogously, MYCIN assigned a certainty factor to each assertion. The result was a set of quadruplets of the form illustrated in Table 1. Thus we know with absolute certainty that the patient's name is John Smith; there is strong evidence indicating that the form of organism-1 is rod, weak evidence that it is staphylococcus, weak evidence that it is not a streptococcus, and absolute certainty that the form of organism-2 is not rod.

When the user introduces a piece of evidence, such as A = 'the form of organism-1 is rod', MYCIN states that $CF(A) = 1$. Given the rule 'if A then B' with certainty factor $CF(B,A)$, the certainty factor for B can be computed as $CF(B) = CF(A) \cdot CF(B,A)$. If there is a second rule 'if B then C', then $CF(C) = CF(B) \cdot CF(C,B)$. This value $CF(C)$ might, in turn, be used in the application of a third rule 'if C then D', and so on.

There was also an equation for combining convergent rules, such as 'if E_1 then H' and 'if E_2 then H', which support the same hypothesis H. The certainty factor of a composed antecedent, such as 'if A and not B', was computed by applying these equations:

$$CF(\text{not } E) = 1 - CF(E) \tag{3}$$

$$CF(E_1 \text{ and } E_2) = \min(CF(E_1), CF(E_2)) \tag{4}$$

$$CF(E_1 \text{ or } E_2) = \max(CF(E_1), CF(E_2)) \tag{5}$$

Although the performance of MYCIN was comparable to that of human experts in the field of infectious diseases, the certainty factor model was

Table 1. Certainty factors of assertions as represented in MYCIN

Object	*Attribute*	*Value*	*CF*
patient	name	John Smith	1.0
organism-1	morphology	rod	0.8
organism-1	identity	staphylococcus	0.2
organism-1	identity	streptococcus	−0.3
organism-2	morphology	rod	−1.0

soon criticized for its mathematical inconsistencies. One of them is that it does not consider correlation between propositions. For instance, if there are two hypotheses, H_1 = 'organism-1 is a streptococcus', with $CF(H_1) = 0.6$, and H_2 = 'organism-1 is a staphylococcus', with $CF(H_2) = 0.3$, then $CF(H_1$ and $H_2) = \min(0.6, 0.3) = 0.3$, whereas it should be $CF(H_1$ and $H_2) = 0$, because the hypotheses are mutually exclusive.

Another problem is the lack of sensitivity in eqns 4 and 5; for example $\min(0.2, 0.9) = \min(0.2, 0.2)$ and $\max(0.9, 0.9) = \max(0.9, 0.2)$.

Furthermore, MYCIN might assign a higher certainty factor to a hypothesis H_2 even if it is less probable than another hypothesis H_1. This anomaly could occur because certainty factors of rules were defined as measures of confirmation (the relative increase in belief), while certainty factors of propositions were interpreted as measures of absolute belief.

It was also pointed out that the combination of convergent rules is valid only under some conditions. These conditions resemble those of conditional independence in the naive Bayes model. However, while the assumption of conditional independence might in particular cases be justified by means of causal arguments, no argument can be made for the assumptions that are implicit in the combination of rules. Therefore, MYCIN's model, rather than solving the main problem of the naive Bayes, ran into more serious troubles (see Buchanan and Shortliffe, 1984, ch. 10–12; Pearl, 1988, sec. 1.2). (*See* **Learning Rules and Productions**; **Expert Systems**; **Resolution Theorem Proving**; **Deductive Reasoning**; **Rule-based Thought**)

PROSPECTOR'S BAYESIAN MODEL

PROSPECTOR was an expert system for geological prospecting developed in the 1970s (Duda *et al.*, 1976). Like MYCIN, PROSPECTOR used if–then rules, but instead of using ad hoc certainty factors, it was based on the theory of probability. Each rule had two parameters, namely the likelihood ratios, which were obtained from human experts' estimations.

The propagation of evidence in PROSPECTOR consisted in computing the probability of the consequence of a rule given the probability of the antecedent. In order to simplify the computation, PROSPECTOR made several assumptions of conditional independence, in addition to approximations and interpolations aimed at smoothing the inconsistencies in the probabilities elicited from human experts.

PROSPECTOR became the first commercial success of artificial intelligence when it assisted in the discovery of a deposit of molybdenum worth about one million dollars. However, this achievement did not lessen criticism of the use of probabilistic methods in expert systems. The assumptions and approximations required by PROSPECTOR were still largely unjustified.

DEMPSTER–SHAFER THEORY

In 1968, Dempster proposed a probabilistic framework based on lower and upper bounds on probabilities. In 1976, Shafer developed a formalism for reasoning under uncertainty that used some of Dempster's mathematical expressions, but gave them a different interpretation: each piece of evidence (finding) may support a subset containing several hypotheses. This is a generalization of the 'pure' probabilistic framework in which every finding corresponds to a value of a variable (a single hypothesis).

The main criticism of this theory from a semantic point of view is the lack of robustness of the combination of evidence. Given three single hypotheses and two findings, it may happen that a hypothesis receiving almost no support from any individual finding is confirmed by the combination of them, while the other two hypotheses are discarded (Zadeh, 1986). Also, a small modification of the evidence assignments may lead to a completely different conclusion. This paradox poses no problem for Dempster's (1968) interpretation (lower and upper probabilities) or to Pearl's (1988, sec. 9.1) interpretation (probability of provability), but seems counterintuitive in Shafer's (1976) interpretation, and for this reason some researchers have proposed alternative formalisms based on different combination rules.

The main problem of the Dempster–Shafer theory in its original formulation is that its computational complexity grows exponentially with the number of hypotheses. One of the solutions proposed consists in building a network of frames of discernment (in fact, a network of random variables), whose axiomatic definition is reminiscent of the properties of conditional independence of Bayesian networks.

BAYESIAN NETWORKS

A Bayesian network (Pearl, 1988) is a probabilistic model that consists of a finite set of random variables $\{V_i\}$ and an acyclic directed graph whose nodes represent those variables and whose arcs represent, roughly speaking, probabilistic dependences between variables.

Dependences are quantified by means of a set of conditional probability distributions (CPDs). Each CPD involves a node and its parents in the graph: $P(v_i \mid pa(v_i))$. (Node V_i is a parent of V_j if there is a link $V_i \rightarrow V_j$ in the graph.) In the case of discrete variables, each CPD is given by atable of probabilities. The product of all the CPDs is a joint probability over all variables, from which it is possible to obtain any other marginal or conditional probability, such as the *a posteriori* probability of any variable given a set of evidence.

There are basically two ways of building a Bayesian network. The *automatic* process involves taking a database and applying one of the many algorithms that yield both the structure and the conditional probabilities. The *manual* process involves two stages: building the structure of the network by selecting the variables and drawing causal links among nodes, as in Figure 1; and then estimating the corresponding conditional probability distributions. (*See* **Machine Learning**)

Ideally, those probabilities should be obtained from objective data, such as databases or epidemiological studies, but in practice the lack of objective data often forces the knowledge engineer to obtain the probabilities from human experts' estimations.

Bayesian networks overcome the limitations of the naive Bayes method in two ways. Firstly, they can diagnose the simultaneous presence of several diseases or failures, because each disease can be represented by a different node. Secondly, the properties of conditional independence are justified either by the statistical independencies in the database, in the case of automatic construction, or by the use of a causal graph elicited from a human expert.

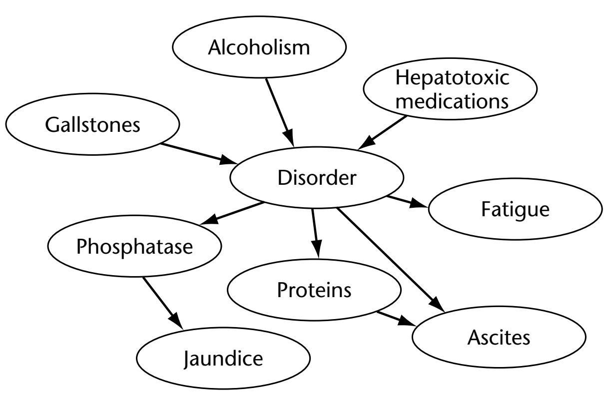

Figure 1. A simplified fragment of HEPAR-II, a medical Bayesian network.

Bayesian networks are also superior to PROSPECTOR in the justification of conditional independencies, but at the price of an increase in computational complexity (the time spent in computing the posterior probabilities).

Causal Bayesian networks have additional advantages, such as the ease with which the model can be extended or refined and its reasoning explained to users.

The main criticisms of Bayesian networks are the difficulty of building the network and the computational complexity of evidence propagation, which is NP-hard: the time required by exact probability updating algorithms depends mainly on the structure of the network, while the complexity of stochastic algorithms depends mainly on the numerical parameters (probabilities). (*See* **Computability and Computational Complexity**; **Bayesian Belief Networks**)

INFLUENCE DIAGRAMS

Influence diagrams are extensions of Bayesian networks which, in addition to random variables, capture available decision options and preferences (utilities). Random variables are represented by circles or ovals, decision nodes as squares or rectangles, and utility nodes as diamonds or parallelograms. Influence diagrams are decision support tools. They permit one to select the optimal decision, the decision that maximizes the expected utility. They overcome the limitations of Bayesian networks in their explicit representation of utilities and in the possibility of selecting the questions to ask or the tests to perform (goal-oriented reasoning). Influence diagrams are closely related to decision trees and Markov decision processes (see Pearl (1988) chapter 6). (*See* **Markov Decision Processes, Learning of**; **Decision-making**)

FUZZY LOGIC AND FUZZY SETS

Some of the sentences that we use in our daily life, such as 'it is cold today', are neither completely true nor completely false. These propositions are called *fuzzy*. In fact, most of the adjectives that we use daily could be interpreted as fuzzy predicates (e.g., 'young', 'rich', 'tall', 'happy', 'healthy', 'big', 'good', 'cheap', 'dark', 'crowded', 'heavy', 'fast', 'modern').

Since in classical logic it is usual to assign 0 to false propositions and 1 to true propositions, some logicians have built multivalued logics in which $v(p)$, the truth-value of proposition p, might also take values between 0 and 1. The truth-value of a composed proposition (negation, conjunction, disjunction, implication, etc.) is a function of the truth-values of the propositions that compose it: for instance, $v(p \wedge q) = f_\wedge(v(p), v(q))$.

Different choices of these logical functions lead to different logics. For instance, Lukasiewicz logic, Kleen logic, and standard fuzzy logic take the 'minimum' function for $f_\wedge$. Other fuzzy logics may use different triangular norms for $f_\wedge$. (A triangular norm is any function $f_\wedge$ that is commutative, associative, and monotone, and satisfies $f_\wedge(1, a) = a$.)

Similarly, there are several implication functions $f_\rightarrow$, all of which satisfy certain conditions. In principle, it would be possible to do inference with fuzzy propositions and fuzzy predicates, but in practice fuzzy inference is usually based on fuzzy sets and fuzzy relations, as shown below.

Given a set A and an element x, the truth-value of the proposition '$x \in A$' is usually called *membership degree* and is represented by $\mu_A(x)$. Whereas in the case of crisp (classical) sets $\mu_A(x)$ is either 0 or 1, in the case of fuzzy sets $\mu_A(x)$ may be any number in the interval [0, 1].

Each operation on fuzzy sets corresponds to a fuzzy-logical operation. For instance, intersection corresponds to conjunction, since $\mu_{A \cap B}(x) = v(x \in A \wedge x \in B) = f_\wedge(\mu_A(x), \mu_B(x))$. Therefore, fuzzy logic can be viewed as the basis of fuzzy set theory. However, it is more usual to view fuzzy set theory as the basis for fuzzy logic.

The rule 'if $P_A(x)$ then $P_B(y)$', where A and B are fuzzy sets and P_A and P_B are their associated predicates, is translated into a fuzzy relation given by $\mu_{A \rightarrow B}(x, y) = f_\rightarrow(\mu_A(x), \mu_B(y))$. Modus ponens consists in combining this rule with an assertion $P_{A'}(x)$ in order to obtain a new set B'. This process is performed by composing the set A' with the relation $\mu_{A \rightarrow B}$. The resulting set B' depends on the logical functions involved in the composition. In many applications of fuzzy logic this choice is made more or less arbitrarily – a typical choice is min as a norm, max as a conorm and Lukasiewicz's implication – and leads to inconsistencies and counterintuitive results (Fukami *et al.*, 1980). The correct way to approach this problem consists in determining the desirable properties of fuzzy inference and then selecting $f_\vee$, $f_\wedge$, and $f_\rightarrow$ coherently in order to ensure such properties (see, for instance, Trillas and Valverde, 1985).

The main criticism of fuzzy logic is the lack of a clear semantics, which leads to an arbitrariness in the application of fuzzy techniques. In particular, there are several definitions of degrees of membership, but apparently none of them is used when building real-world applications. All fuzzy systems use numbers between 0 and 1, but the semantics of those numbers and the way of assigning them differ significantly from application to application. Accordingly, there is no clear criterion for determining which norm or conorm to use in each case.

Additionally, there are several techniques of fuzzy reasoning. We have already mentioned that the properties of the fuzzy inference depend on the choice of logical functions, and knowledge engineers are often unaware of the inconsistencies that may result from an arbitrary choice. There are also other patterns of inference with fuzzy rules that are not based on the composition of relations, and other reasoning techniques, such as those involving fuzzy numbers and fuzzy clustering, not presented in this article. As a result, fuzzy logic consists in practice of a toolbox of heterogeneous techniques without clear indications for deciding which tool to use in any particular case. Users of fuzzy logic often devise ad hoc solutions for the representation and combination of knowledge and data. (*See* **Fuzzy Logic**; **Vagueness**)

ROUGH SETS

Rough sets (Pawlak, 1991) have been developed since the 1980s as a tool for data analysis. In simple terms, the starting point is a data table which represents, for each object, the values of some attributes and a label c: $(a_1, \ldots, a_n, c)$. The final objective of the analysis is to infer some classification rules of the form: 'If the attributes of a new object take values $(a_1, \ldots, a_n)$ then this object is c.'

Both the theoretical foundations and the interpretation of rough sets are completely different from those of fuzzy sets. The lack of a precise boundary of a fuzzy set is a consequence of the vagueness of some concepts, such as 'big' or 'tall', and does not necessarily entail uncertainty. In contrast, the lack of a precise boundary of a rough set derives from the coarseness of the background knowledge implicit in the data table. However, the two theories are complementary and are usually categorized together as 'soft computing'.

NON-MONOTONIC LOGICS

Non-monotonic logics are an attempt to model a pattern of human reasoning that consists in making plausible, although fallible, assumptions about the world, and revising them in the light of new evidence. Upon hearing that 'Tweety is a bird', one might assume that 'Tweety can fly' just because 'most birds fly'; however, further evidence that

'Tweety is a penguin' will lead one to retract this assumption. The name 'non-monotonic' refers to the fact that the addition of new clauses may lead to either adding or retracting a previous conclusion. In classical logic new information never invalidates previous conclusions.

Several non-monotonic logics have been proposed, such as Reiter's logic, McCarthy's circumscription, Doyle's truth maintenance systems, and Cohen's theory of endorsements, each based on some implicit assumptions about uncertainty and preferences. Virtually all have been shown to lead to undesirable behavior under some circumstances.

All these logics are based on qualitative reasoning patterns, and thus differ from the other methods described above, which use and propagate numerical information. Qualitative models are easier to build, because they do not need to estimate quantitative parameters. On the other hand, they are unable to weight conflicting evidence. (*See* **Quantitative Reasoning**)

CONCLUSION

Probability is the oldest formalism for reasoning under uncertainty, and served as the first foundation for computer-based reasoning systems in the 1960s. The unrealistic assumptions of the naive Bayes model used in the first diagnostic systems, the desire to differentiate the evidence in favor of a hypothesis from the evidence against it, and the need for a rule-based reasoning method, led to the development of MYCIN's model of certainty factors. It was later shown that this model had serious inconsistencies and required even more unrealistic assumptions than those of naive Bayes. PROSPECTOR used a Bayesian framework for reasoning with rules, but it also relied on unjustified assumptions and approximations. Dempster–Shafer theory was an attempt to overcome some of the limitations of probabilistic reasoning, but its computational complexity prevented it from being used in practice. Fuzzy sets and fuzzy logic emerged as tools for representing the vagueness of natural language, and led to numerous applications in engineering, medicine, and many other fields.

These four methods – MYCIN, PROSPECTOR, Dempster–Shafer, and fuzzy sets – were developed in the 1970s. In those years, many computer scientists were convinced that probability was not an adequate framework for the problems addressed by artificial intelligence and expert systems. However, the emergence of Bayesian networks and influence diagrams in the 1980s proved that it was possible to build probabilistic models for real-world problems. Many expert systems and software packages based on these techniques became commercial products in the 1990s.

Some researchers nowadays take the position that probability is the only correct framework for uncertain reasoning. Others, while agreeing that probability theory is the best technique when there is enough statistical information, argue that it is hard to use in many practical cases, and for this reason artificial intelligence still needs alternative formalisms.

References

Buchanan BG and Shortliffe EH (eds) (1984) *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA: Addison-Wesley.

Dempster AP (1968) A generalization of Bayesian interference. *Journal of the Royal Statistical Society*, series B **30**: 205–247.

Duda RO, Hart PE and Nilsson NJ (1976) Subjective Bayesian methods for rule-based interference systems. *Proceedings of the AFIPS National Computer Conference*, New York **45**: 1075–1082.

Fukami S, Mizumoto M and Tanaka K (1980) Some considerations on fuzzy conditional interference. *Fuzzy Sets and Systems* **4**: 243–273.

Pawlak Z (1991) *Rough Sets. Theoretical Aspects of Reasoning about Data*. Dordrecht, Netherlands: Kluwer.

Pearl J (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann. [Revised second printing.]

Shafer G (1976) *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press.

Szolovits P and Pauker SG (1978) Categorical and probabilistic reasoning in medicine. *Artificial Intelligence* **11**: 115–144.

Trillas E and Valverde L (1985) On mode and implication in approximate reasoning. In: Gupta MM, Kandel A, Bandler W and Kiszka JB (eds), *Approximate Reasoning and Experts Systems*, pp. 157–166. Amsterdam, Netherlands: North-Holland.

Zadeh LA (1986) A simple view of the Dempster–Shafer theory of evidence and its implication for the rule of combination. *AI Magazine* **7**: 85–90.

Further Reading

Castillo E, Gutiérrez JM and Hadi AS (1997) *Expert Systems and Probabilistic Network Models*. New York, NY: Springer-Verlag.

Cowell RG, Dawid AP, Lauritzen LS and Spiegelhalter DJ (1999) *Probabilistic Networks and Expert Systems*. New York, NY: Springer-Verlag.

Dubois D, Yager RR and Prade H (1993) *Readings in Fuzzy Sets for Intelligent Systems*. San Mateo, CA: Morgan Kaufmann.
Jensen FV (2001) *Bayesian Networks and Decision Graphs*. New York, NY: Springer-Verlag.
Klir GJ and Yuan B (1995) *Fuzzy Sets and Fuzzy Logic. Theory and Applications*. Upper Saddle River, NJ: Prentice Hall.
Krause P and Clark D (1993) *Representing Uncertain Knowledge. An Artificial Intelligence Approach*. Oxford, UK: Intellect Books.
Shafer G and Pearl J (1990) *Readings in Uncertain Reasoning*. San Mateo, CA: Morgan Kaufmann.
Shapiro SC (ed.) (1992) *Encyclopedia of Artificial Intelligence*, 2nd edn. New York, NY: John Wiley.

Receptive Fields

Introductory article

Aniruddha Das, Columbia University, New York, USA

CONTENTS

The receptive field of a neuron is the extended region of sensory space to which the neuron responds, including those stimuli that directly activate or suppress the neuron's activity, as well as those stimuli that can modulate the neuron's response to other stimuli, but have no direct effect by themselves.

INTRODUCTION

Neurons in the sensory systems of touch (somatosensory), hearing (auditory) and vision can be defined in terms of their receptive fields (RFs), namely the regions of sensory space that drive them most vigorously.

While neurons early in a sensory pathway have compact RFs that respond to simple stimuli, neurons at higher stages in the same pathway could have large RFs with complex properties, such as the 'face-selective' cells in the higher visual inferior temporal (IT) region. Determining the transformations in RF properties on passing from stage to stage in a sensory pathway and deducing the cortical circuitry that underlies such a transformation lies at the heart of understanding sensory processing.

The RFs of neurons, particularly at higher centers, can be highly plastic and modifiable. Their response properties may vary depending on the task in which the individual is engaged, may be modulated by attention, and can change over time with sensory training or changes in sensory stimulation. In particular, RFs can 'learn' to select features that are most important in the individual's environment and thus adapt constantly, although within limits, to changes in that environment.

This article will draw heavily on the visual system to illustrate the principles of RF structure, layout, response properties and plasticity, since vision is the best-studied sensory system. However, the principles described here apply to all sensory systems, and these parallels will be mentioned explicitly wherever possible.

HISTORICAL PERSPECTIVE

The concept of the receptive field (RF) was first suggested by the British physiologist Sir Charles Sherrington in the 1890s when he was studying the scratch reflex in dogs. He used this term in the book *The Integrative Action of the Nervous System* (published in 1906) to capture his observation that the reflex was spatially localized on the animal's body. Touching any spot on the dog's back elicited a scratch response directed to the same spot, which Sherrington named the receptive field for that reflex response.

In 1938, H. Keffer Hartline, then working at Johns Hopkins University, introduced the use of the term 'receptive field' to describe the response properties of single nerve fibers. Using a painstaking procedure he dissected out single fibers from the optic