CASE-BASED APPROXIMATE REASONING

THEORY AND DECISION LIBRARY

General Editors: W. Leinfellner (Vienna) and G. Eberlein (Munich)

Series A: Philosophy and Methodology of the Social Sciences

Series B: Mathematical and Statistical Methods

Series C: Game Theory, Mathematical Programming and Operations Research

SERIES B: MATHEMATICAL AND STATISTICAL METHODS

VOLUME 44

Scope: The series focuses on the application of methods and ideas of logic, mathematics and statistics to the social sciences. In particular, formal treatment of social phenomena, the analysis of decision making, information theory and problems of inference will be central themes of this part of the library. Besides theoretical results, empirical investigations and the testing of theoretical models of real world problems will be subjects of interest. In addition to emphasizing interdisciplinary communication, the series will seek to support the rapid dissemination of recent results.

The titles published in this series are listed at the end of this volume.

CASE-BASED APPROXIMATE REASONING

by

EYKE HÜLLERMEIER
University of Magdeburg, Germany

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-5694-X (HB)
ISBN-13 978-1-4020-5694-9 (HB)
ISBN-10 1-4020-5695-8 (e-book)
ISBN-13 978-1-4020-5695-6 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

To Christiane, Jana, and Jaël

Table of Contents

Foreword

In the last three decades, important trends of research in artificial intelligence have been devoted to the design and the study of inference systems that exhibit human-like reasoning capabilities, with special emphasis on tolerance toward incomplete information and uncertainty. This program, which can be also related to cognitive psychology concerns, has led to the development of computational models, in particular for default reasoning that accommodates exceptions and inconsistency, for approximate reasoning with interpolative abilities, as well as for case-based reasoning.

Case-based reasoning (CBR) relies on the concept of similarity, and more particularly on the idea that situations recognized as similar in important aspects may be also similar in other respects. CBR thus appears as a yet simple instance of analogical reasoning, but as powerful as a general problem solving method. This explains the success encountered by case-based reasoning, and beyond that, the interest for similarity-based reasoning that has been increasing in the last ten years. The very idea of CBR is thus to solve new problems on the basis of experience that is represented by already solved problems of the same type, referred to as cases. Thus, a new problem is solved by adapting the solution of a similar case, hoping that the adaptation can be done with much less effort than solving the problem from scratch.

A CBR system requires efficient techniques for several important subtasks, such as organizing and maintaining the case base, retrieving cases (which are maximally similar to the problem) from the case base, and adapting stored cases to the problem at hand. The basic inference mechanism underlying CBR, where the concept of similarity plays a major role, is built upon the principle of instance-based learning and nearest neighbor classification.

CBR has always been motivated by real-world problems, and research in this area has largely focused on building efficient computer systems. Less work has been done, however, on the theoretical foundations of case-based and similarity-based inference. Note that this strongly contrasts with the situation in default reasoning, and to a less extent in approximate reasoning. The present book seeks to remedy this flaw. Its objective is to contribute to a sound foundation of CBR and related fields, such as instance-based learning and analogical reasoning, by providing formal models of similarity-based inference.

To accomplish this objective, Eyke Hüllermeier embeds case-based inference into different frameworks of knowledge representation and reasoning, namely constraint-based reasoning, probability theory, and fuzzy sets and possibility theory. His basic idea is to express the heuristic "similar problem-similar solution" principle in the form of an explicit model, using the formal language of the underlying framework. Proceeding this way, one can take advantage of the reasoning mechanisms offered by that framework. Thus, the author develops various alternative methods, realizing case-based inference as constraint propagation, as probabilistic inference, and as fuzzy set-based approximate reasoning. As convincingly shown in the book, this approach has several advantages. Particularly, case-based inference can benefit from the features offered by the different frameworks. As a noticeable example consider the issue of uncertainty representation. How reliable is a solution proposed by a CBR system? Questions of such kind, which have not received much attention in CBR so far, can adequately be approached by means of probabilistic methods. Likewise, a fuzzy set-based approach to case-based inference can adequately cope with imprecisely or vaguely described cases. Needless to say, the alternative formalizations of similarity-based inference developed by the author are complementary rather than competitive, and different applications will usually call for different methods.

Apart from contributing to the formal foundations of case-based inference, the author's approach has further advantages. Integrating case-based reasoning with other computing paradigms such as probabilistic reasoning and fuzzy set theory can lead to efficient hybrid methods and flexible information processing systems. This can help to clarify differences between alternative methods, but also to show things they have in common. In fact, as one of the more recent trends in artificial intelligence and machine learning, the development of hybrid systems has already produced a number of successful applications. A very promising approach in this connection is the combination of case-based and rule-based reasoning: Since cases and rules can represent, respectively, individual facts and generalized knowledge, these approaches can complement each other in a reasonable way. Indeed, close connections at a formal level can be observed between case-based inference and rule-based inference techniques as realized in approximate reasoning. Worth mentioning is also the current interest, reflected by the organization of specialized workshops, in the combination of case-based reasoning and techniques such as fuzzy sets, neural networks and genetic algorithms, often called "soft computing" paradigms.

With these motivations in mind, Eyke Hüllermeier has done an excellent job in writing this book. His monograph, which is the first one of this type, presents state of the art information as well as novel ideas and new research results in a highly readable and intelligible form. From a theoretical point of view, it clearly constitutes an important contribution to the foundations of case-based and approximate reasoning. From a practical point of view, it should be of interest to everyone working in CBR, fuzzy sets, uncertain reasoning, and related fields.

Toulouse, July 2006 **Henri Prade**

Preface

Longstanding research in artificial intelligence and related fields has produced a number of paradigms for building intelligent and knowledge-based systems, such as rule-based reasoning, constraint processing, or probabilistic graphical models. Being one of these paradigms, *case-based reasoning* (CBR) has received a great deal of attention in recent years and has been used successfully in diverse application areas.

The CBR methodology is inspired by human problem solving and has roots in cognitive psychology. Its key idea is to tackle new problems by referring to similar problems that have already been solved in the past. More precisely, CBR proceeds from individual experiences in the form of *cases*. The generalization beyond these experiences is largely founded on principles of analogical reasoning in which the (cognitive) concept of *similarity* plays an essential role.

This book is an attempt to contribute to the theoretical foundations of CBR, which are not as fully developed as one might expect in light of the practical success of the methodology. To this end, we propose formal models of the fundamental though often implicitly used inference principle underlying CBR methods, namely the heuristic assumption that "similar problems have similar solutions". Proceeding from these models, concrete inference schemes are derived within different frameworks of approximate reasoning and reasoning under uncertainty.

The *case-based approximate reasoning* methods thus obtained especially emphasize the heuristic nature of case-based (similarity-based) inference. More specifically, we combine case-based reasoning with probabilistic methods as well as fuzzy set-based modeling and approximate reasoning techniques. This way, we hope to contribute to a solid foundation of case-based reasoning which is grounded on well-established concepts and techniques from the aforementioned fields, but also to inspire new approaches and to cast light on already existing ones.

Needless to say, the application of these reasoning methods is not restricted to CBR in a narrow sense. Instead, these methods suggest "case-based" approaches in other fields as well. In the final part of the book, we discuss models of *case-based decision making* which combine principles of both case-based reasoning and decision theory. Such models are motivated for reasons of cognitive plausibility as well as practical relevance, and can complement existing models, such as expected utility theory, in a reasonable way.

Much of the research underlying this book has been conducted during my stays at the *Institut de Recherche en Informatique de Toulouse* (IRIT) in the research group headed by Didier Dubois and Henri Prade. I would like to express my gratitude to both of them for providing this opportunity. The IRIT always offered an extremely stimulating research environment, and many of the ideas presented in this book emerged in discussions with Didier and Henri.

I am likewise indebted to my family who not only sustained the times of my absence and adopted the long hours that I have spent writing the manuscript, but also rendered every assistance and moral encouragement whenever needed. To them, Christiane, Jana, and Jaël, I dedicate this book.

Magdeburg, Germany Eyke Hüllermeier
March 2006

Notation

This list offers some of the basic notation that will be used throughout the book. Specific notation will be introduced in the main text as the need arises. Even though some symbols will be used for different purposes, the concrete meaning should always be clear from the context.

Basic mathematical notation

$\forall$, $\exists$	universal and existential quantifier
$\wedge$, $\vee$, $\neg$	logical conjunction, disjunction, negation
$\Rightarrow$, $\Leftrightarrow$	logical implication, equivalence
$\overset{\text{df}}{=}$, $\equiv$	equality by definition, identity
$\mathfrak{N}$ ($\mathfrak{N}_0$)	set of natural numbers (including 0)
$\mathfrak{I}$	set of integers
$\mathfrak{Q}$	set (field) of rational numbers
$\mathfrak{R}$, $\overline{\mathfrak{R}}$	set (field) of real numbers (including $-\infty$ and ∞)
$\mathfrak{R}_{\geq\alpha}$ $\mathfrak{R}_{>\alpha}$	set of real numbers (equal to or) larger than α
$\aleph_0$	cardinality of the natural numbers
card(X), $\lvert X\rvert$	cardinality of the set X
$X \oplus Y$	addition of sets: $X \oplus Y = \{x + y \,\vert\, x \in X, y \in Y\}$
$\emptyset$	empty set
$\cup, \cap, \setminus$	set-theoretic union, intersection, difference
$X \times Y$	Cartesian product of sets X and Y
$\subset$ resp. $\subseteq$, $\subsetneq$	subset relation, proper subset relation
$\lvert\cdot\rvert$, $\lvert\cdot\rvert_p$	Euclidean distance (in $\mathfrak{R}^n$), $\mathcal{L}_p$-norm
$\lVert f\rVert_p$	$\mathcal{L}_p$-norm of the function f
$[\alpha, \beta]$, $(\alpha, \beta]$, ...	closed and (half) open intervals
diam(X)	diameter $\sup_{x,x'\in X} \Delta(x, x')$ of the set X
$\mathfrak{B}_\varepsilon(x)$, $\bar{\mathfrak{B}}_\varepsilon(x)$	open and closed ε-ball around x
$(x_n)_{n\geq1} \to x_0$	convergence of the sequence $(x_n)_{n\geq1}$ toward x_0
$(x_n)_{n\geq1} \nearrow x_0$	convergence from below
$(x_n)_{n\geq1} \searrow x_0$	convergence from above
$O(f), o(f)$	Landau symbols (with limit either 0 or ∞)
$\ll$	much smaller than
dom(f), rg(f)	domain and range of the function f
codom(f)	codomain of the function f

$f \wedge g$, $f \vee g$	lower and upper envelope of the functions f and g
$f\|A$	restriction of $f : X \longrightarrow Y$ to the set $A \subset X$
$f \circ g$	composition of functions f and g
$f \propto g$	function f is proportional to function g
$\mathbb{I}_X$	indicator function of the set X
id	identical function $x \mapsto x$
mod	modulo function (infix notation)
$\lfloor x \rfloor$	largest integer $\leq x$
max, min	maximum and minimum operator
inf, sup	infimum and supremum operator
exp	exponential function
ln, $\log_\alpha$	natural logarithm, logarithm with a base α
arg max	$y \in \arg\max_{x \in X} f(x) \overset{\text{df}}{\Leftrightarrow} f(y) = \max_{x \in X} f(x)$
$A = (a_{\jmath\jmath}), a^{\jmath}$	matrix A, $\jmath$-th column of A
A^t	transpose of matrix A
$A \times x$, $x \times y$	matrix and vector multiplication
e_k	k-th unit vector in $\mathfrak{R}^n$

Remark: We use the term "increasing" (decreasing) in the strict sense, i.e., in the sense of strictly increasing (strictly decreasing). The opposite of increasing (decreasing), i.e., decreasing (increasing) in the weak sense, is referred to as "non-increasing" (non-decreasing). Even though the arg max-operator actually returns a *set* of elements, we shall often write $y = \arg\max_{x \in X} f(x) \in X$; it is then assumed that the maximizing element is unique, or that y is simply an arbitrary choice among these elements.

Probability theory and statistics

$\mathcal{P}(\Omega, \mathcal{A})$, $\mathcal{P}(\Omega)$	class of probability measures over the measurable space $(\Omega, \mathcal{A})$, $\mathcal{P}(\Omega) \overset{\text{df}}{=} \mathcal{P}(\Omega, 2^\Omega)$
$\mathcal{F}(\Omega, \mathcal{A})$, $\mathcal{F}(\Omega)$	class of normalized uncertainty measures over $(\Omega, \mathcal{A})$ resp. $(\Omega, 2^\Omega)$
μ, $\mu_{Y\|(X=x)}$	probability measure, conditional measure
Bel, Pl, m	belief and plausibility function, mass distribution
$\mathbb{P}(X)$	probability of event X (informal notation)
$\otimes$	product of measures
$\lambda(\cdot)$	likelihood function
$\binom{n}{m}$	binomial coefficient
$\preceq$	stochastic dominance relation
$X \sim \mu$	X is distributed according to measure μ

$\phi_{\mu,\sigma}$	density function of the normal distribution with mean μ and standard deviation σ ($\phi \stackrel{\text{df}}{=} \phi_{0,1}$).
μ_{Ω}^{uni}	uniform measure over Ω
$\mathsf{med}(A)$	median of a set of numbers A
$\mathbb{E}(X)$	expected value of a random variable X
$\mathbb{V}(X)$	variance of a random variable X
$\mathsf{bias}(\theta^e)$	bias of an estimator θ^e
$\mathsf{MSE}(\theta^e)$	mean square error of an estimator θ^e

Remark: We shall often not distinguish between an element x and its singleton $\{x\}$. In particular, we use the same notation for a probability measure and the related probability distribution function, i.e., we write $\mu(x)$ instead of $\mu(\{x\})$.

Fuzzy sets, fuzzy measures, and possibility theory

$\mathfrak{F}(X)$	class of fuzzy subsets of a set X
$A, B, \ldots$	fuzzy sets (membership functions)
μ	membership function
$\mathrm{supp}(A)$	support of the fuzzy set A
A_α, A_0	α-cut of A ($0 < \alpha \leq 1$), closure of the support
$\top$, $\otimes$	triangular norm (t-norm)
$\oplus$	triangular co-norm (t-conorm)
$\rightsquigarrow$	generalized (multiple-valued) implication operator
δ, Δ	possibility distribution (measure)
π, Π	possibility distribution (measure)
$\mathcal{N}$	necessity measure
$\int^{ch}$	Choquet integral
$\int^{su}$	Sugeno integral

Remark: We do not distinguish between a fuzzy set and its membership function, i.e., we usually write $A(x)$ (rather than $\mu_A(x)$) for the degree of membership of an element x in the fuzzy set A.

Case-based inference

$\mathcal{S}$, $\mathcal{R}$, $\mathcal{C} = \mathcal{S} \times \mathcal{R}$	set of inputs, outputs, cases
$\mathcal{M}$, $\mathcal{M}^{\downarrow}$	memory of cases $\langle s, r\rangle$, projection of $\mathcal{M}$ to $\mathcal{S}$
φ	mapping $\mathcal{S} \to \mathcal{R}$ or relation between $\mathcal{S}$ and $\mathcal{R}$
$\widehat{\varphi}_{h,\mathcal{M}}$	approximation of φ based on similarity hypothesis h and memory of cases $\mathcal{M}$
σ, Δ	similarity measure, distance measure
$D_{\mathcal{S}}$, $D_{\mathcal{R}}$	ranges of the similarity measures $\sigma_{\mathcal{S}}$ (over $\mathcal{S}$) and $\sigma_{\mathcal{R}}$ (over $\mathcal{R}$)
Σ, $\langle \Sigma, s_0\rangle$	CBI setup, CBI problem with new input s_0
h_{Σ}, H_{Σ}	similarity profile, probabilistic similarity profile
h, H	similarity hypothesis, probabilistic hypothesis
$\mathcal{N}_{\alpha}(r)$	α-neighborhood of an outcome $r \in \mathcal{R}$
$\mathcal{N}_k(\mathcal{M}, s_0)$	k-selection from memory $\mathcal{M}$
$\mathcal{N}_k^{ex}(\mathcal{M}, s_0)$	extended k-selection
$\mathsf{SST}(\mathcal{M}, s_0)$	similarity structure
$\mathsf{pSST}(\mathcal{M}, s_0)$	partial similarity structure
$\mathsf{OST}(\mathcal{M}, s_0)$	outcome structure
$\mathsf{CST}(\mathcal{M}, s_0)$	case structure

Abbreviations

AI	artificial intelligence
CBA	case-based approximation
CBDM	case-based decision making
CBDT	case-based decision theory
CBI	case-based inference
CBL	case-based learning
CBLA, CBLP	case-based learning algorithm/process
CBR	case-based reasoning
EBDM	experience-based decision making
EUT	expected utility theory
ILP	integer linear programming
LS	least squares
ML, MLE	maximum likelihood (estimation)
PSP	probabilistic similarity profile
RCOP	repetitive combinatorial optimization
RSP	repeated search problem