Available online at www.sciencedirect.com
ScienceDirect
Engineering Applications of Artificial Intelligence 21 (2008) 14–25
Engineering Applications of ARTIFICIAL INTELLIGENCE
www.elsevier.com/locate/engappai

A model for parameter setting based on Bayesian networks

Reyes Pavón[a,*], Fernando Díaz[b], Victoria Luzón[a]

[a]*Depto. de Informática, Universidad de Vigo, Campus Universitario As Lagoas s/n, 32004 Ourense, Spain*
[b]*Depto. de Informática, Universidad de Valladolid, Plaza Santa Eulalia 9-11, 40005 Segovia, Spain*

Received 13 July 2005; received in revised form 17 November 2006; accepted 2 February 2007
Available online 8 May 2007

Abstract

One of the difficulties that the user faces when using a model to solve a problem is that, before running the model, a set of parameter values have to be specified. Deciding on an appropriate set of parameter values is not an easy task. Over the years, several standard optimization methods, as well as various alternative approaches according to the problem at hand, have been proposed for parameter setting. These techniques have their merits and demerits, but usually they have a fairly restricted application range, including a lack of generality or the need of user supervision. This paper proposes a meta-model that generates the recommendations about the best parameter values for the model of interest. Its main characteristic is that it is an automatic meta-model that can be applied to any model. For evaluation purposes and in order to be able to compare our results with results obtained by others, a real geometric problem was selected. The experiments show the validity of the proposed adjustment model.

Keywords: Setting parameters; Bayesian networks; Learning; Genetic algorithms; Constructive geometric constraint solving

1. Introduction

There are a variety of situations in which a researcher is faced with a modeling problem. Modeling is a process through which a *model M* (function or algorithm) is constructed to explain the behavior of the system and to predict unknown answers. Once the model of the studied system is established based on a set of *parameters* $\mathbf{\Theta}$, the parameter values which make the model generate the best results should settle down.

In this work, we are interested in the parameter setting of models used in the artificial intelligence (AI) field, such as artificial neural networks, genetic algorithms (GAS), cluster algorithms and so on (Duda et al., 2001). Parameter setting of a model $M(\mathbf{\Theta})$ can be carried out either by applying specific mechanisms to the model (Friedrichs and Igel, 2004) or by applying general optimization techniques used in system modeling (Fletcher, 2000; Rao, 1996). In many other cases, the choice of the parameters involves consulting the specialized technical literature or simply resorting to a trial and error technique.

This work has utility in the engineering field, where a variety of AI models and techniques are used to solve a whole range of problems, but such techniques are not known in depth. Our work could help in these situations. As an example to illustrate its usefulness, a problem of graphic design was selected (*the root identification problem*).

In a deeper analysis, it can be distinguished three approaches to the problem of parameter setting in AI: the *evolutionary* approach (used by evolutionary algorithms: GAS (Goldberg, 1989); evolution strategies (Schwefel, 1995); and evolutionary programming (Fogel, 1999)), the *model selection* approach and the *statistical* approach.

The *evolutionary* approach is based on adapting the parameters during the evolutionary algorithm run. Parameter control techniques can be sub-divided into three types: deterministic, adaptive, and self-adaptive (Eiben et al., 1999). In deterministic control, the parameters are changed according to deterministic rules without using any feedback from the search. The adaptive control takes place when there is some form of feedback that influences the

*Corresponding author.
E-mail addresses: pavon@uvigo.es (R. Pavón), fdiaz@infor.uva.es (F. Díaz), luzon@uvigo.es (V. Luzón).

0952-1976/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2007.02.013

Nomenclature

D	database of runs
$M(\mathbf{\Theta})$	model whose parameters have to be fixed
M^*	meta-model or adjustment model
P	problem to solve
AI	artificial intelligence
BAM	Bayesian adjustment model
BN	Bayesian network
CAD	computer-aided design
CBR	case-based reasoning
DB	database analysis
EM	expectation maximization algorithm
GA	genetic algorithm
HMM	hidden markov model
MAP	maximum A posteriori technique
ML	maximum likelihood technique
MSBN	microsoft Bayes networks
STAT	Barreiro's statistical analysis

parameter specification. Examples of adaptive control are the works of Davis (1989), Julstrom (1995) and Smith and Smuda (1995). Finally, self-adaptive control is based on the idea that evolution can be also applied in the search for good parameter values. In this type of control, the operator probabilities are encoded together with the corresponding solution, and undergo recombination and mutation. Self-adaptive evolution strategies (Beyer and Schwefel, 2002) are an example of the application of this type of parameter control.

In the *model selection* approach, a model selection criterion can be used to select parameters, more generally to compare and choose among models which may have a different capacity or competence. When there is only a single parameter one can easily explore how its value affects the model selection criterion: typically one tries a finite number of values of the parameter and picks the one which gives the lowest value of the model selection criterion. Most model selection criteria have been proposed for selecting a single parameter that controls the "complexity" of the class of functions in which the learning algorithms find a solution, e.g., the structural risk minimization (Vapnik, 1982), the Akaike (1974) Information Criterion, or the generalized cross-validation criterion (Craven and Wahba, 1979). Another type of criteria are those based on held-out data, such as the cross-validation estimates of generalization error (Kohavi and John, 1995). These are almost unbiased estimates of generalization error (Vapnik, 1982) obtained by testing $M(\mathbf{\Theta})$ on data not used to choose parameter $\mathbf{\Theta}$. This approach is less applicable when there are several parameters to simultaneously optimize. In this case, the conditions which must be satisfied are more restrictive (Bengio, 2000).

Finally, the general goal of *statistical* approach (or statistical estimation theory) is to estimate some unknown parameter from the observation of a set of random variables, referred to as "the data". The maximum likelihood (ML), maximum a posteriori (MAP), expectation maximization (EM) or hidden Markov models (HMM) are examples of estimation techniques. Briefly, the ML is a classical technique which estimates a parameter directly from the data, assuming that data are distributed according to some parameterized probability distribution functions (Jeffreys, 1983). The MAP technique is a Bayesian approach, i.e., a priori available statistical information on the unknown parameters is also exploited for their estimation (DeGroot, 1970). When it is assumed that the data are drawn from a mixture of parameterized probability distribution functions, where not just the parameters, but also the mixture components have to be estimated, the EM algorithm is useful (Dempster et al., 1977; Redner and Walker, 1984; Jordan and Jacobs, 1994). As a last step, the HMM assume that the mixtures evolve in time according to a Markov process (Rabiner, 1989).

In this paper we suggest the use of a *meta-model* M^* that generates recommendations about the best parameter values $\mathbf{\Theta}$ for the model of interest M and that differs from other systems in that it is:

- *general*, in the sense of being a system that can be applied to any model $M(\mathbf{\Theta})$;
- *automatic*, in the sense of being a system built from a set of data, without any model user supervision.

Most general, meta-modeling is the analysis, construction and development of the frames, rules, constraint, models and theories applicable and useful for the modeling in a predefined class of problems. In the context of this work, a model can be viewed as an abstraction of phenomena in the real world, and a meta-model is yet another abstraction, highlighting properties of the model itself.

The rest of the work is focused on defining such a system for parameter setting. As will be shown in the followings sections, Bayesian networks (BNs) are identified as the suitable formalism to define the meta-model M^*. Moreover, the proposed model has the added benefit of removing the need for, or at least reducing the effect of, user decisions about parameter values.

The structure of the paper is as follows. Section 2 is devoted to describing the proposed model for parameter setting. In Section 3, notions about BNs are briefly presented, including learning from databases, and we outline their suitability for building a framework for parameter setting. After this, Section 4 describes the problem used throughout the paper to illustrate the application of the proposed adjustment model, the experiments carried out for evaluating it and the results attained.