Fifth IFSA World Congress (1993), 1262-1265

APPROXIMATIVE INFERENCE IN HIERARCHICAL STRUCTURED RULE BASES

László T. Kóczy
Dept. of Telecommunication
and Telematics
Technical University of
Budapest
Sztoczek u. 2
H-1111 Budapest, Hungary

Kaoru Hirota
Dept. of Instrumentation
and Control Engineering
Hosei University
College of Engineering
Kajino-cho, Koganei-shi
Tokyo 184, Japan

Abstract

The paper discusses the problem of controlling systems with a very high number of input variables effectively by fuzzy If...then rules. The basic idea is the partition of the state space into domains, which step can be done even iteratively several times,and every domain has its own sub rule base referring to a considerably lower number of variables than the original space. In this manner the number of necessary rules is drastically reduced and time complexity of the control algorithm remains acceptable.

1. Introduction

In the control of very complex systems, often the number of input variables is very high. Time complexity of even the traditional (Zadeh or Mamdani style) fuzzy control algorithms is too high for real time control. However, often a few variables dominate the system in a certain area of the state space. This subset of variables changes when the working point in the state space changes. In this paper, a solution applying hierarchical structured rule systems is proposed.

The principle of fuzzy rule based control was invented by Zadeh in his seminal paper [1] where he described the compositional rule of inference. Disadvantage of the original algorithm is its exponential complexity in terms of the number of input variables [2]). Another algorithm invented by Mamdani et al. [3] has polynomial complexity (cf. [2]). A comparison of the two can be found in [4].

2. Rule approximation

A special problem occurs when the rule base does not cover the whole observation space: it is sparse. Sparseness reduces complexity. Also rules received by tuning the rule base [5] might be located with "gaps" among the antecedents - resulting in a partially sparse base. The need of data compression e.g. for fast transmission [6] is another motivation for discussing such systems. If the source of evidence is heterogenous, the rule base is often conflicting (e.g. an expert system comprising the knowledge of several specialists).

In sparse and/or conflicting rule bases conclusion can be calculated by approximating the fuzzy mapping $\mathcal{R}$: $X \rightarrow Y$. A theoretical foundation for the approximation of $\mathcal{R}$ combines the "graph view" interpretationin [7] and the Resolution Principle, and is proposed in [8].

Applying the fuzzy mapping view combined with the resolution principle, traditional function approximation techniques can be applied for every level set. An overview of such techniques including extrapolation, regression and other methods can be found in [9].

3. State spaces with a very high number of variables

In a state space of k input variables and maximum resolution of N, having a base with r rules, the rule base itself has

$$\mathcal{I} = O(rkN)$$

uniform complexity. With the CRI, calculation of a conclusion takes

$$\mathcal{C}_1 = O(rN^{k+1}),$$

steps, while Mamdani's algorithm only

$$\mathcal{C}_2 = O(rkN).$$

(For these results, see [2], for an introduction to complexity etc. see e.g. [10].) The second algorithm is very well tractable, while the first one is applicable only in the case of low k.

In some industrial systems, a serious difficulty arises. If there are several thousand input variables, even the complexity of the very rule base ($\mathcal{I}$) becomes rather high for practical purposes: the difficulty is in the value of r. In an acceptably complete rule base, the number of rules is very high: even if the number of linguistic terms for every input variable is bounded by L, L^k input combinations exist; with arbitrary membership functions, even more. A solution is offered by the fact that certain areas of the state space are dominated by small subsets of the variables. By identifying the state of the system, attention can be reduced to a small group of variables. If a crisp partitioning of the state space is possible, the original problem is reduced to several subproblems:

$$X = \bigcup_{i=1}^{p} D_i, \text{ where } D_i \cap D_j = \emptyset \text{ if } i \neq j.$$

For every D_i another subset of $\{X_i\}$ is dominant ($X = \times_{i=1}^{k} X_i$).The subset for D_i is S_i = $\{X_{i1},\ldots,X_{in_i}\}$ ($n_i \ll k$), and the S_i-s are usually not disjoint. For every S_i, another rule base exists, where the number of rules is considerably lower than in the space without partition. The total number of all rules in is approximately

$$r' = \sum_{i=1}^{p} L^{n_i}, \text{ where } n_i \ll k, \text{ as stated}$$

previously.

The complete hierarchical rule system includes also the partitioningrules in the form "If X is D_i then R_i", their number can be approximated by cp, where c is constant. The total number of rules in the system is in the order

$$r = cp + \sum_{i=1}^{p} L^{n_i} \ll L^k.$$

Succesful applications of a hierarchical, structured rule system have been published by Sugeno et al. (see e.g. [11]).

4. Fuzzy partition of the state space

In many industrial processes no partition of X exists in the crisp sense. There are disjoint areas of X where a certain subset of X_i-s dominates the process, however, these do not cover the whole X:

$$\bigcup_{i=1}^{p} D_i \subset X \text{ ($\subset$ denotes proper containment).}$$

For observations lying in

$$T = X \setminus \bigcup_{i=1}^{p} D_i$$

the effect of several subgroups of variables dominates jointly: none of the rule bases is applicable alone for computing the conclusion. Because of this fact, crisp definition of $\{D_i\}$ is a rough estimation. The influence domain of a certain subset of variables is in reality a fuzzy subset of X. A fuzzy partition or cover is $\bigcup_{i=1}^{p} supp(\tilde{D}_i) = X$,

where $core(\tilde{D}_i) \cap core(\tilde{D}_j) = \emptyset$. (Otherwise the base is contradictory). Usually, $supp(\tilde{D}_i) \cap supp(\tilde{D}_j) \neq \emptyset$, even if $i \neq j$. Observations in such a system might be of two types: their support is completely contained in a single domain or they lie completely or partially in T, i.e. their support is contained in at least two domain supports.

For such an observation, both (or, in general, all) variable subsets and sub rule bases must be combined. An example is A^* which overlaps with two domains in the fuzzy cover of $X = X_1 \times \ldots \times X_6$: D_2 and D_3, where the variable subsets $\{X_2, X_4, X_5\}$ and $\{X_2, X_3, X_6\}$ are occuring in the antecedent parts of the corresponding sub rule base. So, rules in R_2 have the form

"If X_2 is A_{2i} and X_4 is A_{4j} and X_5 is A_{5k} then Y is B_l"

where A_{pq} and B_l are fuzzy sets over X_p and Y, respectively. For A^*, all rules from R_2 and R_3 are considered and so the input variable set S = $\{X_2, X_3, X_4, X_5, X_6\}$ is used as the reduced state space. In this example, S contains only one less component than X itself, but if k is rather high, the subsets have much fewer elements, so their union S is also much less dimensional than X. All calculations can be done in S so, that all rules are transformed into S by the cylindric extension method: If the i-th sub rule base is described by $\{\mu_{ij}(x_{i1}, \ldots, x_{ik_i})\}$, and the j-th base adds the new variables $\{x_{jl_1}, \ldots, x_{jl}\}$, so that $\{x_{i1}, \ldots, x_{ik_i}\} \cup \{x_{jl_1}, \ldots, x_{jl_m}\} = \{x_{S1}, \ldots, x_{Sk_S}\}$ and $\{x_{i1}, \ldots, x_{ik_i}\} \cap \{x_{jl_1}, \ldots, x_{jl_m}\} = \emptyset$, the i-th base is extended to

$$\{\mu_{ij}(x_{S1}, \ldots, x_{Sk_S})\} =$$

$$\{\min\{\mu_{ij}(x_{i1}, \ldots, x_{ik_i}), 1(x_{jl_1}, \ldots, x_{jl_m})\}\} =$$

$$\{\mu_{ij}(x_{i1}, \ldots, x_{ik_i})\}.$$

Using this way of extension increases the complexity only minimally. For the observation, the new rule base is obtained by

$R_S = R_{ie} \cup R_{je}$ (subscript e stands for cylindric extension).

The conclusion for A^* can be obtained by any of the suitable algorithms (as the CRI, or Mamdani's technique) on only R_S.

4. 5. Approximations of $\mathcal{R}$: X → Y in the united subspaces

In the case of very large number of variables even the possibility of handling complete covers is questionable. A hierarchical rule base might be sparse in two different senses: either the sub rule bases do not cover the whole input state space (i.e. the union of the supports of all fuzzy domains is a proper subset of the state space) or the individual sub rule bases might have "gaps" inside the given domain. As a matter of course, the combination of both can also occur.

For the possible approximation techniques, especially the various types of interpolation, see e.g. [9]. Applying the method of cylindric extension to all sub rule bases flanking the observation (in the sense of the partial ordering introduced in [8]), including the case of partial overlapping, as well, the methods of rule interpolation are directly applicable for obtaining the conclusion, and this algorithm for the inference has a low computational complexity, maximally $O(kL^{k_{max}}) = O(kL^c)$, where k_{max} denotes $\max\{k_i\} << k$.

In this way, it is possible to construct hierarchical rule bases with sub bases only for the typical state areas. The size of the total information necessary to store is rather reduced. For most of the observations, a not very high number of sub bases must be united by cylindric extension, and in this combined base, rule approximation techniques are applied. The dimensionality of this subspace is still much lower than that of X, so complexity of these calculations is still acceptable.

References

[1] L. A. Zadeh: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. on Systems, Man and Cybernetics (1973). pp. 28-44.

[2] L. T. Kóczy: Computational complexity of various fuzzy inference algorithms. Annales Univ. Sci. Budapest 12 (1991). pp. 151-158.

[3] E. H. Mamdani and S. Assilian: An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. of Man-Machine Studies 7 (1978). pp. 1-13.

[4] L. T. Kóczy and Á. Juhász: Fuzzy rule interpolation and the RULEINT program. Proc. Joint Hungarian-Japanese Symposium on Fuzzy Systems and Applications, Budapest, 1991. pp. 91-94.

[5] D. G. Burkhardt and P. P. Bonissone: Automated fuzzy knowledge base generation and tuning. Proc. IEEE Conference on Fuzzy Systems, San Diego, 1992. pp. 179-188.

[6] L. A. Zadeh: Interpolative reasoning as a common basis for inference in fuzzy logic, neural network theory and the calculus of fuzzy If-then rules. Opening talk, 2nd Int. Conf. on fuzzy logic and neural networks, Iizuka, 1992.

[7] D. Dubois and H. Prade: Basic issues on fuzzy rules and their application to fuzzy control. Fuzzy Control Workshop, IJCAI-91, Sidney, 1991.

[8] L. T. Kóczy and K. Hirota: Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases, to appear in Information Sciences, 1992.

[9] L. T. Kóczy: Techniques of inference in insufficient and inconsistent fuzzy rule bases. Proc. Non-Classical Logics and their Applications, Linz, 1992. pp. 46-50.

[10] A. V. Aho, J. E. Hopcroft and J. D. Ullman: The design and analysis of computer algorithms. Addison-Wesley, Reading, etc., 1974.

[11] M. Sugeno, T. Murofushi, J. Nishio and H. Miwa: Helicopter control based on fuzzy logic. Presented at the Second Fuzzy Symposium on Fuzzy Systems and Their Applications to Human and Natural Systems, Tokyo - Ochanomizu, 1991.