Optimizing Parameters of Expert Rules

September 14, 2010

1 Introduction

Orbotech provided an example of expert rules used for identifying a particular defect, along with a set of 390 samples for experimenting. The expert rules are parameterized by 10 parameters. While reasonable values of the parameters can be guessed by the expert, the performance of the expert rules can be improved by finding the best values of the parameters for the given training set.

The search space is rather large: assuming that the parameters have discrete domains of size N, the total number of possible combinations is N^{10} (10^{10} for 10 possible values for each parameters), and exhaustive search is infeasible. Informed search, such as implemented in the UNCERTIMA Toolkit, helps, as shown later; however, direct application of the algorithm is not efficient, and the problem requires adaptation.

2 Problem: Parameter Tuning

Rules, in the form of a decision tree, are provided. The rules are parameterized by 10 parameters—thresholds in the nodes of the decision tree; reasonable value ranges for each of the parameters are specified by the expert.

The rules are tuned using a training set (390 samples). Parameters which maximize the classification accuracy over two classes (the hypothesis that there is a defect of particular kind is either confirmed or rejected) must be selected.

3 Solution Approaches

There quite a few multi-parameter optimization algorithms, such as

- iterative sequential optimization (ISO);
- Latin Hypercube sampling (LHS) followed by local search (LS);
- Simulated annealing;

and others. Depending on the structure of the decision tree, they may or may not find good combinations of parameter values. A significant drawback of most known algorithms is that their termination conditions are either not well-defined, or cannot change flexibly with varying computational resources (the algorithms are not 'any-time').

A family of informed search algorithms based on the notion of Value of Information has been developed during work on IMG4 projects and implemented in the UNCERTIMA Toolkit for Optimization under Uncertainty. These algorithms cope well with problems in which just a few of many available parameter combinations may be observed before a good combination is selected. Measurement selection and algorithm parameter tuning are examples of such problems.

4 Adaptation of the Problem to VOI-based Search Algorithms

The VOI-based search algorithms work best when the measurements (or trials) are expensive, and the search space is relatively small. These two conditions allow to estimate Value of Information of a trial for each combination of parameters, and to perform the most promising trial. However, in the case of tuning expert rules parameters on a relatively small training test, the situation is just opposite:

- The search space is huge: there may be billions of possible parameter combinations under reasonable settings.
- A single trial is very cheap, cheaper than estimating its value of information.

Thus, an exhaustive search would be more efficient than a direct naive application of the informed search algorithm, and for a successful application of the informed search the problem must be adapted: the size of the search space must be reduced, and an atomic computation must be made sufficiently large to be worth the meta-computation.

It turns out that due to a particular shape of the expert rules, some parameters are responsible for breaking the samples into big groups, and the rest are used for fine adjustments of the classification within the group. The role and the influence of each of the parameters can be estimated without opening the black box of the expert rules by evaluating the effect of changing the parameter value with the domain. The stronger the change in the accuracy over the domain, the more influential is the parameter.

Splitting the parameters into two groups makes possible application of the informed search algorithm.

- Denote the two subsets of parameters as S (strategical) and T (tactical).
- Define the search space for the informed search on the parameter combinations from S only.

- For each combination of parameters from S, use a relatively simple search algorithm (ISO, LHS-LS) to find the best combination of parameters from T (given an assignment to S). Consider an invocation of the simple search as a single computation.

By varying the sizes of subsets S and T, it is possible to conform to the two requirements of the VOI-based informed search quoted earlier:

- small search space (while still infeasible for exhaustive search);
- expensive atomic computation.

5 Application to Orbotech Data

It can be easily discovered that among ten parameters

```
NmissingContour
GrayDiff
StdLimit
HuesNumber
MinContour
ColorBal
Yellow
LumVar
minDI
Border
```

the parameters MinContour, Yellow, minDI affect the classification accuracy much stronger than the rest. Therefore, the set of parameters is split as $S = \{MinContour, Yellow, minDI\}$, $T = \{NmissingContour, GrayDiff, StdLimit, HuesNumber, ColorBal, LumVar, Border\}$. Each measurement runs iterative sequentail optimization on the parameters from T, and a grid dependency model is defined on S to guide the search. The problem in UNCERTIMA is defined as follows:

```
problem rules {


  space (MinContour, Yellow, minDI) {
    MinContour = [5, 10 .. 50];
    Yellow = [ 20, 40 .. 200 ];
    minDI = [1 .. 10];
```

```
  }

  observe accuracy {
    1=> accuracy = 0.01 / 0.01;
  }

  find max util(MinContour, Yellow, minDI) = accuracy;

  model lattice(MinContour, Yellow, minDI) {
    accuracy = (0.5, 1.0) / (0.25, 0.25, 0.25);
  }
}
```

This approach helped improve the accuracy on the provided sample set from 0.948 to 0.969, and achieving the solution takes only a few minutes on a modest laptop (for the myopic search with rational VOI recomputation). The found values of the parameters are:

```
NmissingContour  50
GrayDiff  50
StdLimit  50
HuesNumber  0.5
MinContour  15 # was 5
ColorBal  0.5
Yellow  140 # was 100
LumVar  500
minDI  3 # was 5
Border  0.5
```

The chosen values of parameters from the subset T came out slightly different from the original ones, but leaving them unchanged does not affect the accuracy.

6 Conclusion

VOI-based informed search can be an efficient tool for optimizing algorithm parameters. A problem should and can be adapted for application of the search algorithm. Application of the approach to tuning the parameters of the expert rules provided by Orbotech improved the accuracy by about 2%, and the computation was cheap.