Machine Learning and Case-based Reasoning: Their Potential Role in Preventing the Outbreak of Wars or in Ending Them

Robert Trappl[1,2*], Johannes Fürnkranz[1], Johann Petrak[1], Jacob Bercovitch[3]

[1]Austrian Research Institute for Artificial Intelligence
Schottengasse 3, A-1010 Vienna, Austria

[2]Department of Medical Cybernetics and Artificial Intelligence
University of Vienna, Austria

[3]Department of Political Science
University of Canterbury, New Zealand

OEFAI-TR-97-10

Abstract

In a current project we investigate the potential contribution of Artificial Intelligence for the avoidance and termination of crises and wars. This paper reports some results obtained by analyzing international conflict databases using machine learning and case-based reasoning techniques.

1 Introduction

Research in Artificial Intelligence has always been heavily supported by "defense agencies". While enormous amounts of money have been and still are spent on the development of AI methods for military purposes, practically no effort is undertaken to use these methods to support the *prevention and termination* of conflicts and wars. We believe that Artificial Intelligence has a great potential for peacefare although research in this area has not yet received the attention that it deserves (Trappl, 1986, 1992).

One possible contribution of AI to peacefare is the knowledge that can be gained by analyzing databases of international conflicts or conflict management attempts with machine learning algorithms. This paper presents two case studies for such an approach.

2 International Conflict Databases

In recent years the importance of empirical studies has also been recognized in the *international relations* research community. This increasing interest in

*Corresponding author. E-mail: robert@ai.univie.ac.at

empirical methods was a major source of motivation for the development of a variety of databases that try to capture the important aspects of international crises and code them into suitable attributes. The most important among these databases are:

- the Correlates of War Militarized Interstate Disputes dataset (Gochman & Maoz, 1984),
- the International Crisis Behavior (ICB) project (Brecher, Wilkenfeld, & Moser, 1988; Wilkenfeld, Brecher, & Moser, 1988),
- the COPDAB dataset (Azar, 1980),
- the event data sets of the KEDS and PANDA projects (Schrodt & Davis, 1994; Vogele, 1994; Bond, Bennet, & Voegele, 1994),
- the Butterworth dataset (Butterworth, 1976),
- the SHERFACS database (Sherman, 1988),
- the KOSIMO database of conflicts (Pfetsch & Billing, 1994),
- the CONFMAN database of conflict management attempts (Bercovitch & Langley, 1993).

We believe that databases like the above-mentioned provide a promising application area for symbolic machine learning algorithms. This study reports the results of such an attempt. After considering the availability and scope of these datasets, we eventually chose to work with the CONFMAN and KOSIMO databases.

2.1 The CONFMAN Database

The development of the *International Conflict Management (CONFMAN) Dataset* is a project that is conducted under the supervision of Jacob Bercovitch at the Department of Political Science of the University of Canterbury, New Zealand. This database is of particular interest for Machine Learning research, as it has been generated with the explicit aim of empirical analysis. Its primary focus is international mediation. Its aim is to both further our understanding of mediation, and facilitate the comparative investigation of different conflict management mechanisms.

Prompted by dissatisfaction with previous studies, which have rested on ideographic or normative approaches, the development of this database was begun with the aim of furthering the much needed empirical investigation of conflict management within a sound theoretical framework. The project is founded on the contingency approach to the study of international conflict management which regards the outcome of management efforts as contingent upon a number of contextual and process variables. The contingency approach encourages systematic empirical research because it recognizes variables and attributes with explicit operational criteria.

The database should facilitate the answer to such fundamental questions as "How do international mediation, and other forms of conflict management work?" and "Under what conditions are respective conflict management efforts most effective?". In answering these questions it is hoped the project will make a concrete contribution to the improvement of the international conflict management process.

Table 1: Attributes of the CONFMAN database

Attribute	*Type*	*Description*
V1	numeric	Dispute Number
V2/V3	ordinal/numeric	Duration (grouped/raw)
V4/V5	ordinal/numeric	Fatalities (grouped/raw)
V6	ordinal	Dispute Intensity
V7	ordinal	System Period
V8	nominal	Geographic Region
V9–V11	nominal	Issue 1 – Issue 3
V12	nominal	Final Outcome
V13	nominal	Dispute Initiator
V14/V15	nominal	Identity Party A/B
V16/V17	nominal	Time in IS A/B
V18	nominal	Alignment
V19/V20;V21/V22	numeric;ordinal	Power A/B (raw;grouped)
V23	nominal	Previous Relation
V24/V25	nominal	Political System A/B
V26/V27	numeric	Number of Parties A/B
V28/V29	ordinal	Homogeneity A/B
V30/V31	ordinal	Political rights A/B
V32/V33	ordinal	Civil liberties A/B
V35	nominal	Conflict Management Type
V36	nominal	Third Party Identity
V37	nominal	Mediator Rank
V38	nominal	Mediation Strategies
V39	nominal	Previous Relationship
V40	ordinal	Prev Attempts
V41	ordinal	Prev Attempts this Mediation
V42/V43	ordinal/numeric	Timing (grouped/raw)
V44	nominal	Initiated by
V45	nominal	Environment
V80	ordinal	m intensity
V70	ordinal	power disp
V75/V76	nominal	Human Rights A/B
V77	ordinal	Human Rights Disparity
V91	nominal	Political System Type
V81	binary	Political System Difference
V92	ordinal	Ally Numbers
V82	binary	Ally Support Disparity
V93	nominal	Homog. Type
V83	nominal	Homog. Comparison
V94	ordinal	Time in System
V84	binary	Time in System Comparison
V90	nominal	Total Issues
V46	nominal	Outcome (detailed)
V99	binary	Outcome (binary)

A mediation attempt is defined as the formal or institutionalized non-violent and non-judicial intervention of an outsider or third party willing to help both disputants seek an acceptable outcome. An offer of mediation services is included in this understanding of an intervention. Other forms of conflict man-

agement that are encompassed are negotiation, arbitration/adjudication, multilateral conference, and referral to an international organization. The referral of a dispute to an international organization is coded as a separate event from any subsequent mediation or adjudication by that organization.

The central task of this research project has been the compilation of an extensive original dataset of international conflict management events since 1945. Primary information sources included Keesings Contemporary Archives (laterly Keesings Record of World Events), The Times Index, and The New York Times Index. Whenever necessary more detailed contemporary press reports or reputable historical accounts were also utilized.

The dataset that was used in the current study encompasses 921 international disputes and management attempts from 241 disputes since 1945. The attributes we used are listed in Table 1. This database — or previous versions of it — has been analyzed extensively with statistical methods, most recently in (Bercovitch & Wells, 1993; Bercovitch & Houston, 1993; Bercovitch & Lamare, 1993; Bercovitch & Langley, 1993).

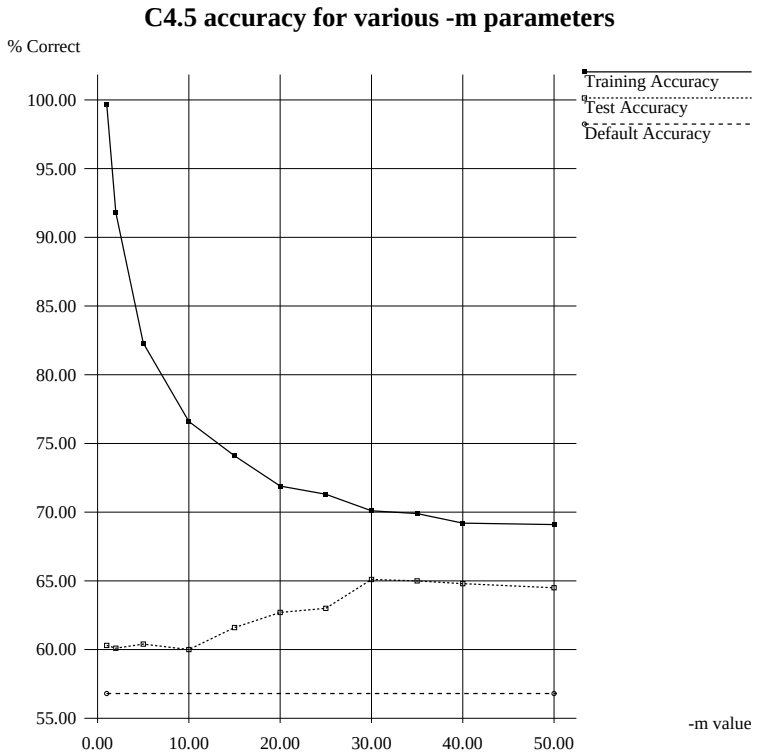
2.2 The KOSIMO Database

The KOSIMO database has been developed under the supervision of Frank Pfetsch at the Institute of Political Science at the University of Heidelberg, Germany (Pfetsch & Billing, 1994). The database is an attempt to unify and extend case lists and databases of several previous research projects: primarily (Butterworth, 1976),(Brecher et al., 1988),(Wilkenfeld et al., 1988), but also (Gantzel & Meyer-Stamer, 1986), (Holsti, 1983), and others. It contains more than 1400 conflicts from 1482 to 1990 encoded in three tables. We have concentrated on the NOPUTSCH table that describes 547 internal and international conflicts and wars between 1945 and 1990. As the database was not generated with the explicit aim of empirical analysis, its complex structure (list-valued, hierarchical, textual, and multi-dimensional fields) made several transformation necessary, which are described in more detail in (Fürnkranz, Petrak, & Trappl, 1997). As many of the above-mentioned problems can not trivially be solved by attribute-value based machine learning algorithms, we tried to analyze the database with case-based techniques.

3 Inductive Learning Techniques

As the simple attribute-value format of the CONFMAN database suggests the use of decision tree learning algorithms, we first tried to analyze it with C4.5 (Quinlan, 1993). We started with investigating the predictive accuracies that can be obtained by learning trees pruned to different degrees. Figure 1 shows the results in terms of accuracy on the training set (purity) and predictive accuracy as estimated by a 10-fold cross-validation for various choices of C4.5's `-m` and `-c` parameters.

The performance of C4.5 peaks at parameter settings `-m 30` and `-c 10`, i.e. when C4.5 ensures that only those decision nodes are further expanded for which at least two children cover more than 30 training examples or when a relatively



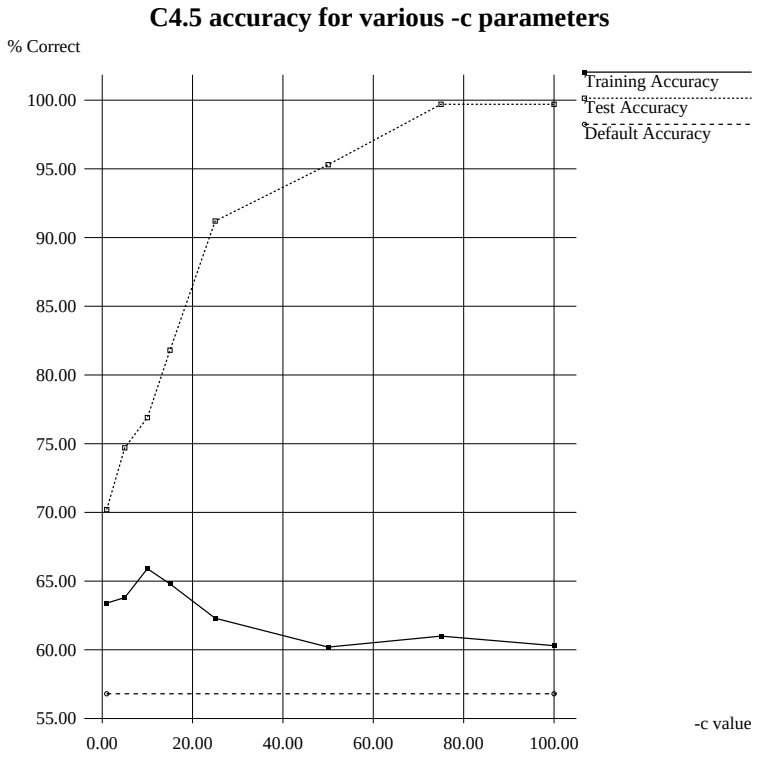


Figure 1: Results for various settings of C4.5's `-m` and `-c` parameters

high degree of pruning is employed.[1] The default accuracy of this domain (indicated by the flat lines at the bottom of Fig. 1) is 56.8%, the result for C4.5's default parameters (`-c 25` and `-m 2`) is 62.5% with a standard deviation of ±5.2. All in all the above accuracy curves exhibit the characteristic shape for noisy domains: too complex trees are inaccurate because of overfitting, while too simple trees are inaccurate because of over-generalization.

We have also tested the combination of the best two values found above, which resulted in a decision tree with an estimated predictive accuracy of around 66.7% (±3.7). The tree obtained by this parameter setting is shown in Fig. 2. Three rules (P3, P7, and P9) cover a majority of successful conflict management attempts. In particular rule P7 which describes 176 attempts, among them 123 (69.9%) successes, looks interesting.

However, a closer investigation of trees like the one of Fig. 2 shows that the quality of the rules in a tree is very unstable. Some rules discriminate very well between successes and failures, while other leaves seem to have about the same distribution of successes and failures as in the original datasets and can thus not be expected to perform better than mode prediction. Consequently we switched to examining single rules instead of entire trees. In a first experiment we generated a decision tree using C4.5 and examined it for nodes containing only successful or only unsuccessful mediation outcomes. A typical example of such a pure rule is shown below (a complete listing of the rules that have been discovered by this process is given in appendix A).

> **If** *there have been less than 400 fatalities* **and**
> *party B's raw power index is not extremely high* **and**
> *the conflict management type was mediation* **and**
> *the conflict lasted between 1 and 3 months*
> **then** *the conflict management was always successful*
> *in 15 mediation attempts in 8 different conflicts.*

In a special experiment we tried to answer the question, what factors influence the success or failure of a given mediation strategy. For this purpose we generated decision trees in which we forced C4.5 to use the attribute *Mediation Strategy* at the root, so that it will try to find the best sub-tree discriminating between successful and unsuccessful conflict management attempts for a given mediation strategy. As the quality of the leaves of the found trees varied, we again examined the trees for leaves in which successes or failures significantly dominate the matching mediation attempts. One such rule was

> **If** *the mediator has mixed relationships with the conflict parties,*
> *or is from the same block as both of them,*
> *or from a different block as both of them,*
> **and**
> *the mediation environment is party B's territory,*
> *a third party's territory or a composite*
> **then** *directive mediation strategies have been*
> successful *in 37 mediation attempts and have*
> failed *in 7 mediation attempts.*

[1] The value range of the `-c` parameter in C4.5 is from 0 to 100, where a low value indicates heavy pruning, while a value of 100 indicates no pruning.

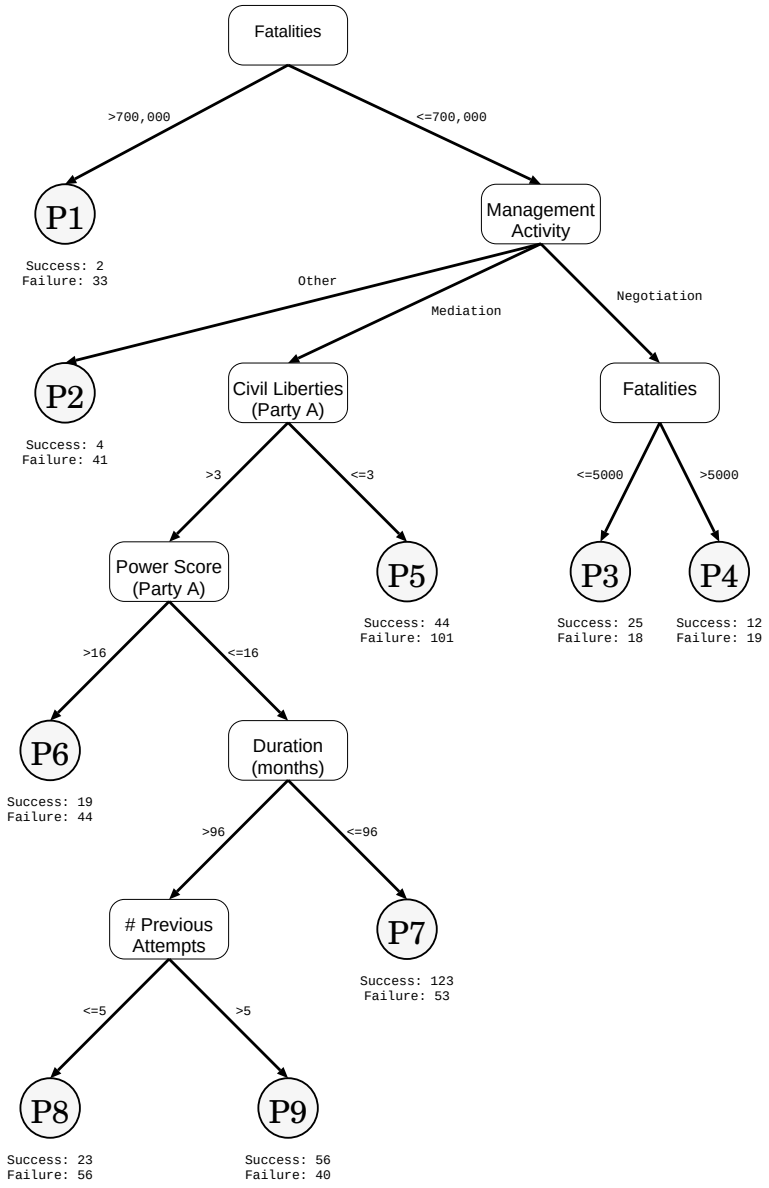


Figure 2: Decision tree for predicting the outcome of conflict management attempts

A complete decision tree resulting from these experiments can be found in (Fürnkranz et al., 1997).

In another experiment we tried to compare the results of *feature subset selection* with the identification of relevant factors by statistical analysis. For this purpose we used the feature subset selection procedure implemented in an early version of the publicly available machine learning library MLC++ (Kohavi, John, Manley, & Pfleger, 1994), which realizes a wrapper approach around C4.5 (John, Kohavi, & Pfleger, 1994). It starts with an empty set of attributes and greedily adds the attribute that gives the highest increase in estimated predictive accuracy for the tree grown from the new set of attributes. Alternatively, the algorithm can also choose to delete an existing attribute from the current set of attributes. Predictive accuracy is estimated with consecutive 10-fold cross-validation experiments (with different random splits) until the standard deviation of the resulting estimate is below 1%. If no feature can be added or deleted without decreasing the estimated accuracy of the tree for two consecutive tries, the program stops with the current set of features. In order to avoid to be too short-sighted a one-time decrease is not sufficient for stopping the algorithm. In this case two features may be added at a time if this increases accuracy.

Table 2: Relevant attributes detected by feature subset selection

`C4.5 -m 2 -c 25`				
Choice	*Variable*	*Purity*	*# X-vals*	*Accuracy*
1	V39 Previous Relationship	64.2%	3	63.2%
2	V20 Raw Power Score B	68.1%	4	65.5%
3	V37 Mediator Rank	71.7%	4	65.9%
4	V27 Number of Parties B	74.1%	4	67.0%
5	V22 Grouped Power Score B	73.4%	4	67.3%
6	V06 Dispute Intensity	—	—	—
6	V45 Environment	70.3%	4	67.5%
8	V11 Issue 3	70.5%	4	67.7%
`C4.5 -m 2 -c 25 -s`				
Choice	*Variable*	*Purity*	*# X-vals*	*Accuracy*
1	V39 Previous Relationship	64.3%	2	63.3%
2	V19 Raw Power Score A	67.5%	2	65.4%
3	V27 Number of Parties B	70.3%	4	65.8%
4	V10 Issue 2	70.6%	4	66.7%
5	V70 Power Disparity	—	—	—
5	(V82 Ally Support Disparity)	71.3%	3	66.9%
7	V90 Total Issues	71.3%	3	67.1%
8	V21 Grouped Power Score A	—	—	—
8	V45 Environment	70.9%	4	67.1%
10	(V82 is deleted again)	70.9%	3	67.7%

Table 2 summarizes the output of MLC++ from two experiments, one using the default settings for the parameters, and one with the `-s` option turned on, which allows C4.5 to lump different outcomes of a test together thereby obtaining simpler trees. We have tried a few different parameter settings, in particular those that yielded the best results in previous experiments (Fig. 1). However in this case, the default choices seemed to be very good, which indicates that only relevant attributes are used and that therefore too high settings of the `-m` parameter and too low settings of the `-c` parameter may force C4.5 to throw away relevant information. For each of the two experiments we report the purity of the final tree, the number of cross-validations needed to get the standard deviation below 1% and most importantly the estimated accuracy of the tree. The tables have to be read from the top to the bottom.

The final decision tree in both cases consisted of 8 variables and had an accuracy of above 67%. It is interesting that in the experiment where the `-s` parameter was activated, the program at one point had 9 variables in the tree, but the feature concerning the disparity of the support of each party's allies could be deleted again at the end with a further increase of accuracy. This shows that the algorithm does not necessarily converge towards an optimal subset of features. It may for example be the case that adding a combination of certain attributes yields a better tree, while adding only one of them results in a worse tree. The analysis reveals that for example the attributes chosen in the tree of Fig. 2 are very different from the attributes that appear in Table 2 which have produced a more accurate tree.

It is interesting to compare the above results with the results produced with classical statistical methods. Table 3 shows the most relevant aspects of mediation attempts that have been identified in previous work (Bercovitch & Lamare, 1993). There is obviously a considerable overlap. Almost all of the variables of Table 3 appear in one of the two experiments of Table 2, most of them in both. The most notable exception is the absence of mediation strategy. The number of fatalities, which is also not considered by Machine Learning, is partially reflected in the intensity of the conflict, which has been recognized as important, although only in one experiment. Using only fatalities for generating a decision tree would only yield 62.4% accuracy using the same grouping as in (Bercovitch & Lamare, 1993). On the other hand, the Machine Learning method has at-

Table 3: Relevant features for mediation outcome detected by statistical analysis

Fatalities
Mediation Environment
Mediation Strategy
Previous Relations of Mediator
Issues
Mediator Rank

tributed a higher significance to the previous relation of the mediator (63.3%). In addition aspects concerning the power of the conflict parties and about the number of parties involved on each side have been considered.

4 Case-Based Techniques

Situations of international conflict and war, like other complex human life situations, are often described and explained in terms of previous similar situations. Such comparisons often help to understand the various possibilities of actions the participants and international organizations can choose, and their possible consequences. Similarity-based case retrieval and analysis can therefore be a useful tool for analyzing a new conflict situation. This and the fact that the KOSIMO database was less susceptible to analysis with machine learning techniques, because it was not coded in a strict attribute-value fashion, but included list-valued, multi-dimensional and hierarchical fields motivated our experiments with case-based learning and similarity-based case retrieval techniques. These experiments were performed with the VIE-CBR tool box, a flexible and extensible library of Common LISP routines that allows easy experimentation with alternative algorithms for any of its functional components (Petrak, 1994).

Sim	Year	Conflict
		Bosnia-Herzegovina
0.62	1938	Germany-Czechoslovakia (Munich Treaty)
0.60	1948	Israel I (Palestine War)
0.57	1974	Cyprus IV (Turkish Invasion)
0.55	1965	India XVI (Kashmir IV)
0.54	1968	CSSR (Invasion)
		Germany-Czechoslovakia (Munich Treaty)
0.77	1968	CSSR (Invasion)
0.75	1953	GDR (17. June 1953)
0.72	1946	Greece (Civil War II)
0.67	1948	Berlin I (Blockade)
0.66	1961	Berlin III (Wall Erection)
		USA-Grenada
0.66	1959	Dominican Republic I (Intervention)
0.57	1962	Cuba IV ('Cuba-Crisis')
0.57	1954	Guatemala I (Intervention)
0.57	1973	Libya-USA
0.57	1945	Triest

Figure 3: The five best matches for three selected cases ordered by decreasing similarities (English translation of the original German KOSIMO database entries)

First, we simply tried to retrieve the most similar cases to a given conflict from the KOSIMO database. Figure 3 shows the retrieval of the five nearest neighbors of three selected cases in the database when using a similarity measure previously defined by a domain expert. The case "548 Bosnia-Herzegovina" has been coded and added to the database by one of the authors of the KOSIMO

Table 4: Error rates and output similarities of 1-NN for KOSIMO

1-NN	ERGEBNISM		ERGEBNIST		ERGEBNISP		LOESUNG		INTENS	
SIM-EVEN	49%	0.62	44%	0.62	80%	0.41	70%	0.40	54%	0.78
SIM-EXPERT	47%	0.65	45%	0.61	82%	0.39	73%	0.38	55%	0.78
Default	49%	0.51	37%	0.63	98%	0.38	73%	0.27	65%	0.70

database for this experiment. It seems remarkable that one of the two situations (Vietnam and the Munich Treaty) with which the conflict in Bosnia was often compared by politicians before deciding to intervene or not, was ranked the most similar case by the program. This and several other similar experiments have been discussed at the *Second International Workshop on the Potential Contribution of Artificial Intelligence to the Avoidance of Crises and Wars* in Vienna and the invited domain experts (among them the creators of the database) found meaningful explanations for the similarity of the retrieved cases.

We have also performed experiments that aimed at deriving automatic classification of unknown cases using nearest neighbor techniques. Table 4 shows the results obtained by trying to predict the military (ERGEBNISM), territorial (ERGEBNIST), political (ERGEBNISP) results of a conflict, the type of conflict resolution used in its settlement (LOESUNG), and the intensity of the conflict (INTENS). We have tried two different similarity measures, one that gives equal weights to all features (SIM-EVEN), and one that was provided by one of the creators of the KOSIMO database (SIM-EXPERT). The results in the first column show the error rates, while the numbers in the second column show a *output similarity* measure that tries to take into account the distance between the predicted output value and its actual value in the case. All numbers shown are the average of 10 cross-validated runs.

Most of the error rates are only slightly better than default accuracy. The best improvements in error rate were achieved for ERGEBNISP and INTENS, while the algorithm performed badly for predicting territorial outcomes of a conflict (ERGEBNIST). Note, however, that for these experiments, missing values were treated like other values. All cases that had a missing classification but were classified with a non-missing value, received output similarity zero. For the field ERGEBNIST, for instance, no value was specified in 63% of all cases in KOSIMO. This might be an explanation why the results for this field were particularly bad.

The results are a little better when we look at the output similarities of the 1-NN predictions. Except for the ERGEBNIST experiments they are always substantially better than mode prediction, which suggests that the retrieved cases capture some aspects that are relevant for the classification of the current case. This is illustrated in Fig. 4 which shows with which relative frequency the output similarities for field INTENS occurred when applying 1-NN with SIM-EVEN to KOSIMO (first line of Table 4). It is obvious that 1-NN's predictions are closer to the target than mode prediction. In about 90% of the cases 1-NN

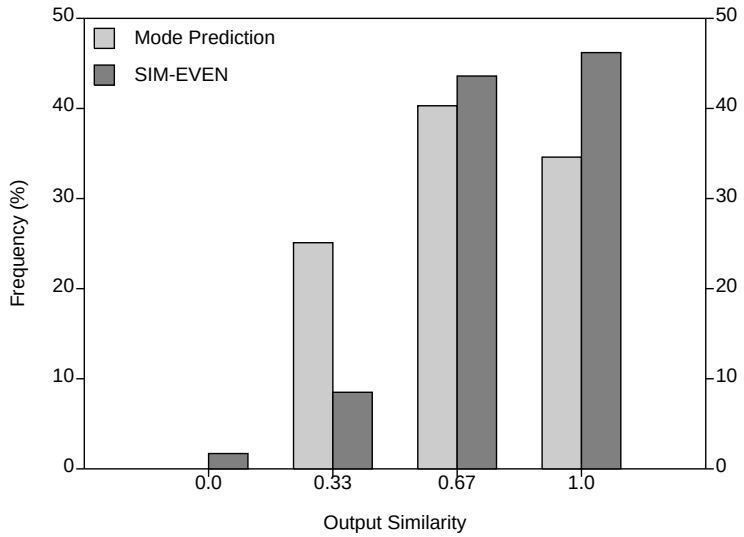


Figure 4: Output similarity of field INTENS (mode prediction vs. SIM-EVEN)[3]

correctly predicts the intensity level or misses only by one degree.

We have also tried several enhancements of the basic nearest neighbor algorithm, like e.g. case weighting or considering more than one neighbor for the derivation of a classification. These and more experiments with different weighting schemes and on a different subset of the KOSIMO database can be found in (Fürnkranz et al., 1997) and (Petrak, Trappl, & Fürnkranz, 1994).

5 Related Work and Discussion

There has been some previous work on rule induction via decision tree learning, which we will briefly discuss below. More extensive overviews can be found in (Mallery, 1988), (Schrodt, 1991a), and (Schrodt, 1997).

Schrodt (1991b) has performed several experiments in predicting interstate conflict outcomes using the Butterworth "Interstate Security Conflicts, 1945–1974" (Butterworth, 1976). He used his own implementation of ID3, the predecessor of C4.5, to learn decision trees for predicting the effects of management efforts with respect to five different outcomes. In all his experiments the estimated predictive accuracy of the learned trees was below mode prediction accuracy, i.e. below the accuracy that one would achieve by always predicting the majority class. However, his implementation of ID3 was not capable of dealing with numeric data and, more importantly, did not have C4.5's extensive pruning facilities. The only method used for getting simpler trees was manual feature subset selection by observing which attributes his algorithm typically

[3]The field INTENS describes the intensity of the conflict on a value range from 1 to 4. The output similarity measure can thus yield 4 different degrees of similarity from 0.0 (when 1 is predicted and 4 is correct and vice versa) to 1.0 (exact prediction) depending on the difference between the prediction and the actual intensity.

selects near the root of the tree, which did not result in higher accuracies. Note that the method we chose for feature subset selection produces different results: for example the attributes chosen in the tree of Fig. 2 differ substantially from the attributes that appear in Table 2, which have been chosen to maximize predictive accuracy of the trees generated by C4.5, In our study simple decision trees are usually more accurate than an unpruned decision tree. However, even the unpruned tree exhibits a significant gain in predictive accuracy compared to mode prediction. Predicting the outcome of conflict management attempts thus seems to be an easier task than the prediction of aspects of the outcome of the conflict itself. A reason for this might be that conflict management events are more repetitive than the conflicts themselves.

Our results with C4.5 indicate that the quality of the rules inside the decision tree can vary substantially. In general, single rules can be more reliable than entire decision trees. Mallery and Sherman (1993) report a variety of rules that have been learned with I^2D (Unseld & Mallery, 1993), an improved version of ID3 that was specifically developed to deal with the structured nature of the SHERFACS dataset (Sherman, 1988). They report purity and coverage for the learned rules, but give no indication of their predictive accuracies. Estimating the predictive accuracy of single rules of a decision tree is more problematic than estimating the accuracy of the entire tree, because cross-validation can only estimate the accuracy of complete classifiers. The alternative, to reserve a subset of the data entirely for testing, is also not a good solution, because the size of these databases is in general fairly low compared to the number of attributes, so that an additional loss of training data would be detrimental to the quality of the learned rules. Schrodt (1991b) has used an entropy ratio, similar to the information score proposed in (Kononenko & Bratko, 1991) which gives a higher weight to correct predictions of rare classes.

6 Further Work

Currently we are working on a much larger and more recent version of the CONFMAN database. Initial experiments with the C4.5 algorithm have promised a significant improvement of the results in terms of accuracy. We also plan to employ a wider range of machine learning and knowledge discovery techniques, e.g. the use of *inductive logic programming* techniques and the discovery of *partial determinations*. Another goal is to further improve the results by using domain-specific background knowledge. We also plan to address different sets of questions in the near future. An example for such a question would be "Has the set of factors that determine the success of a mediation outcome changed since the end of the *cold war*?". With recent versions of the CONFMAN database we hope to be able to shed some light on this question.

References

Azar, E. E. (1980). The conflict and peace data bank (COPDAB) project. *Journal of Conflict Resolution*, *24*(1), 143–152.

Bercovitch, J., & Houston, A. (1993). Influence of mediation characteristics and behavior on the success of mediation in international relations. *The International Journal of Conflict Management*, *4*(4), 297–321.

Bercovitch, J., & Lamare, J. W. (1993). The process of international mediation: An analysis of the determinants of successful and unsuccessful outcomes. *Australian Journal of Political Science*, *28*, 290–305.

Bercovitch, J., & Langley, J. (1993). The nature of dispute and the effectiveness of international mediation. *Journal of Conflict Resolution*, *37*(4), 670–691.

Bercovitch, J., & Wells, R. (1993). Evaluating mediation strategies - a theoretical and empirical analysis. *Peace & Change*, *18*(1), 3–25.

Bond, D., Bennet, B., & Voegele, W. B. (1994). Panda: Interaction events data development using automated human coding.. Extended Version of a paper presented at the 1994 Annual Meeting of the International Studies Association in Washington, DC on April 1st, 1994.

Brecher, M., Wilkenfeld, J., & Moser, S. (1988). *Crises in the Twentieth Century - Handbook of International Crises*, Vol. I. Pergamon Press, Oxford.

Butterworth, R. L. (1976). *Managing Interstate Conflict, 1945-74: Data with Synopses.* University of Pittsburgh Center for International Studies, Pittsburgh.

Fürnkranz, J., Petrak, J., & Trappl, R. (1997). Knowledge discovery in international conflict databases. *Applied Artificial Intelligence*, *11*(2), 91–118.

Gantzel, K. J., & Meyer-Stamer, J. (Eds.). (1986). *Die Kriege nach dem zweiten Weltkrieg bis 1984.* München.

Gochman, C. S., & Maoz, Z. (1984). Militarized interstate disputes 1816–1976: Procedures, patterns, and insights. *Journal of Conflict Resolution*, *28*, 585–616.

Holsti, K. J. (1983). *International Politics: A Framework for Analysis* (2nd edition). Englewood Cliffs.

Hudson, V. M. (1991). *Artificial Intelligence and International Politics.* Westview Press, Boulder, CO.

John, G. H., Kohavi, R., & Pfleger, K. (1994). Irrelevant features and the subset selection problem. In Cohen, W., & Hirsh, H. (Eds.), *Proceedings of the 11th International Conference on Machine Learning (ML-94)*, pp. 121–129 New Brunswick, NJ. Morgan Kaufmann.

Kohavi, R., John, G., Manley, D., & Pfleger, K. (1994). MLC++: A machine learning library in C++.. In *Tools with Artificial Intelligence*. IEEE Computer Society Press.

Kononenko, I., & Bratko, I. (1991). Information-based evaluation criterion for classifier's performance. *Machine Learning*, *6*, 67–80.

Mallery, J. C. (1988). Thinking about foreign policy: Finding an appropriate role for artificial intelligence computers. Master's thesis, M.I.T. Political Science Department, Cambridge, MA.

Mallery, J. C., & Sherman, F. L. (1993). Learning historical rules of major power intervention in the post-war international system.. Paper prepared for presentation at the 1993 Annual Meeting of the International Studies Association.

Petrak, J. (1994). VIE-CBR - Vienna case-based reasoning tool, version 1.0: Programmer's and installation manual. Technical report TR-94-34, Austrian Research Institute for Artificial Intelligence, Vienna. Available from `ftp://ftp.ai.univie.ac.at/papers/oefai-tr-94-34.ps.Z`.

Petrak, J., Trappl, R., & Fürnkranz, J. (1994). The possible contribution of AI to the avoidance of crises and wars: Using CBR methods with the KOSIMO database of conflicts. Technical report TR-94-32, Austrian Research Institute for Artificial Intelligence, Vienna. Available from `ftp://ftp.ai.univie.ac.at/papers/oefai-tr-94-32.ps.Z`.

Pfetsch, F. R., & Billing, P. (1994). *Datenhandbuch nationaler und internationaler Konflikte*. Nomos Verlagsgesellschaft, Baden-Baden.

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Mateo, CA.

Schrodt, P. A. (1991a). Artificial intelligence and international relations: An overview. In (Hudson, 1991).

Schrodt, P. A. (1991b). Classification of interstate conflict outcomes using a bootstrapped ID3 algorithm. *Political Analysis*.

Schrodt, P. A. (1996). *Patterns, Rules and Learning: Computational Models of International Behavior*. University of Michigan Press. Forthcoming.

Schrodt, P. A., & Davis, S. G. (1994). Techniques and troubles in the machine coding of international event data. Dept. of Political Science, University of Kansas. Paper presented at the 1994 meeting of the International Studies Association, Washington DC.

Sherman, F. L. (1988). Sherfacs: A new cross-paradigm, international conflict dataset.. Paper written for presentation at the 1988 annual meeting of the International Studies Association.

Trappl, R. (1986). Reducing international tension through artificial intelligence: A proposal for 3 projects. In Trappl, R. (Ed.), *Power, Autonomy, Utopia: New Approaches Toward Complex Systems*. Plenum, New York.

Trappl, R. (1992). The role of artificial intelligence in the avoidance of war. In Trappl, R. (Ed.), *Cybernetics and Systems '92*, pp. 1667–1672 Singapore. World Scientific.

Unseld, S. D., & Mallery, J. C. (1993). Interaction detection in complex data-models.. MIT A.I. Memo No. 1298.

Vogele, W. B. (1994). Global conflict profiles: Event data analysis using panda.. Paper prepared for presentation at the annual meeting of the American Political Science Association, New York City, Sep. 1994.

Wilkenfeld, J., Brecher, M., & Moser, S. (1988). *Crises in the Twentieth Century - Handbook of Foreign Policy Crises*, Vol. II. Pergamon Press, Oxford.

A Pure Rules discovered in the CONFMAN database

Rules contained in an unpruned decision tree that cover 10 or more *unsuccessful* conflict management attempts:

```
Rule F1:
if (400 < V5_FAT <= 700000) &&
   (V20_POWERB <= 33) &&
   (V32_LIBA <= 3) &&
   (V33_LIBB == 1) &&
   (V35_MGMTACT == "MEDIATION")
then SUCCESS: 0
     FAILURE: 12 (2 Conflicts)
```

```
Rule F2:
if (400 < V5_FAT <= 700000) &&
   (V20_POWERB <= 33) &&
   (V24_POLSYSA == "MULTI-PARTY") &&
   (V32_LIBA <= 3) &&
   (V33_LIBB > 1) &&
   (V35_MGMTACT == "MEDIATION") &&
   (39 < V43_TIM <= 67)
then SUCCESS: 0
     FAILURE: 19 (4 Conflicts)
```

```
Rule F3:
if (V3_DUR > 6) &&
   (400 < V5_FAT <= 700000) &&
   (V20_POWERB <= 33) &&
   (V24_POLSYSA == "MULTI-PARTY") &&
   (V32_LIBA <= 3) &&
   (V33_LIBB > 1) &&
   (V35_MGMTACT == "MEDIATION") &&
   (V40_NRMED > 3) &&
   (67 < V43_TIM <= 256)
then SUCCESS: 0
     FAILURE: 14 (3 Conflicts)
```

```
Rule F4:
if (V3_DUR > 76) &&
   (400 < V5_FAT <= 700000) &&
   (V19_POWERA <= 31) &&
   (V20_POWERB <= 22) &&
   (V30_RIGHTSA > 4) &&
   (V32_LIBA > 3) &&
   (V35_MGMTACT == "MEDIATION") &&
   (V38_MEDSTR == "DIRECTIVE") &&
   (V41_NRMEDM <= 4) &&
   (61 < V43_TIM <= 136)
then SUCCESS: 0
     FAILURE: 16 (5 Conflicts)
```

```
Rule F5:
if (V3_DUR <= 190) &&
   (400 < V5_FAT <= 700000) &&
   (V19_POWERA <= 31) &&
   (V20_POWERB <= 33) &&
   (V32_LIBA > 3) &&
   (V33_LIBB > 3) &&
   (V35_MGMTACT == "MEDIATION") &&
   (V38_MEDSTR == "COMM-FACIL") &&
   (V39_RELMED == "NO_PREV_REL") &&
   (V41_NRMEDM <= 2) &&
   (V93_HOMT == "MAJORITY")
then SUCCESS: 0
     FAILURE: 14 (8 Conflicts)
```

```
Rule F6:
if (400 < V5_FAT <= 700000) &&
   (V19_POWERA > 31) &&
   (V20_POWERB <= 33) &&
   (V32_LIBA > 3) &&
   (V35_MGMTACT == "MEDIATION") &&
   (V41_NRMEDM <= 4) &&
   (V43_TIM <= 158)
then SUCCESS: 0
     FAILURE: 13 (3 Conflicts)
```

```
Rule F7:
if (V5_FAT > 700000) &&
   (V43_TIM <= 73) &&
   (V84_TISC == "DIFF_TIME_SYS")
then SUCCESS: 0
     FAILURE: 31 (3 Conflicts)
```

Rules contained in an unpruned decision tree that cover 10 or more *successful* conflict management attempts:

```
Rule S1:
if (1 < V3_DUR <= 3) &&
   (V5_FAT <= 400) &&
   (V20_POWERB <= 33) &&
   (V35_MGMTACT == "MEDIATION")
then SUCCESS: 15
     FAILURE: 0 (8 Conflicts)
```

```
Rule S2:
if (V2_DUR_G > 2) &&
   (V5_FAT <= 400) &&
   (V20_POWERB <= 33) &&
   (V35_MGMTACT == "MEDIATION") &&
   (V43_TIM <= 35) &&
   (V44_REQINI == "BOTH_PARTIES")
then SUCCESS: 15
     FAILURE: 0 (11 Conflicts)
```

```
Rule S3:
if (400 < V5_FAT <= 700000) &&
   (V19_POWERA <= 31) &&
   (V20_POWERB <= 22) &&
   (V32_LIBA > 3) &&
   (V35_MGMTACT == "MEDIATION") &&
   (V38_MEDSTR == "DIRECTIVE") &&
   (V41_NRMEDM <= 4) &&
   (V43_TIM > 145) &&
   (V84_TISC == "DIFF_TIME_SYS")
then SUCCESS: 14
     FAILURE: 0 (5 Conflicts)
```

```
Rule S4:
if (V3_DUR > 13) &&
   (400 < V5_FAT <= 700000) &&
   (V19_POWERA <= 31) &&
   (V20_POWERB <= 33) &&
   (V32_LIBA > 3) &&
   (V35_MGMTACT == "MEDIATION") &&
   (V38_MEDSTR == "PROCEDURAL") &&
   (V43_TIM <= 18) &&
   (V94_TIS <= 4)
then SUCCESS: 10
     FAILURE: 0 (5 Conflicts)
```

```
Rule S5:
if (815 < V5_FAT <= 700000) &&
   (V19_POWERA <= 31) &&
   (V20_POWERB <= 10) &&
   (V30_RIGHTSA > 4) &&
   (V32_LIBA > 3) &&
   (V35_MGMTACT == "MEDIATION") &&
   (V38_MEDSTR == "DIRECTIVE") &&
   (V39_RELMED == "SAME_BLOC_BOTH") &&
   (V40_NRMED <= 5) &&
   (V41_NRMEDM <= 2) &&
   (V43_TIM <= 61)
then SUCCESS: 12
     FAILURE: 0 (7 Conflicts)
```