THE AQ15 INDUCTIVE LEARNING SYSTEM: AN OVERVIEW AND EXPERIMENTS

by

R. S. Michalski
I. Mozetic,
J. Hong
N. Lavrac:

ISG 86-20, UIUCDCS-R-86-1260, Department of Computer Science, University of Illinois, Urbana, 1986.

Report No. UIUCDCS-R-86-1260

The AQ15 Inductive Learning System: An Overview and Experiments

by

Ryszard S. Michalski, Igor Mozetic[1], Jiarong Hong[2], Nada Lavrac[1]

July 1986

Intelligent Systems Group
Department of Computer Science
University of Illinois at Urbana-Champaign

ISG 86-20

This research was supported in part by the National Science Foundation under Grant No. DCR 84-06801, the Office of Naval Research under Grant No. N00014-82-K-0186, the Defense Advanced Research Project Agency under Grant No. N00014-K-85-0878, and by the Slovene Research Council.

---

[1] On leave from: *Jozef Stefan Institute, Ljubljana, Yugoslavia.*
[2] On leave from: *Harbin Institute of Technology, Harbin, The People's Republic of China.*

ABSTRACT

AQ15 is a multi-purpose inductive learning system that uses logic-based, comprehensible knowledge representation. It is able to incrementally learn attributional disjunctive concepts from data that may contain erroneous or inconsistent examples, and can perform constructive induction. The latter means that the program uses background knowledge to generate new attributes not present in the input data, and, if they pass a relevance test, employs them in the learning process. In an experimental application to three medical domains, the program learned decision rules that performed at the level of accuracy of human experts. A surprising and potentially significant result is the demonstration that by applying the proposed method of rule reduction and flexible matching (TRUNC), one may drastically decrease the complexity of the knowledge base without affecting its performance accuracy.

Keywords: Knowledge Acquisition, Machine Learning, Inductive Inference, Applications (Medicine)

1. INTRODUCTION

It is widely acknowledged that the construction of a knowledge base represents the major bottleneck in the development of any AI system. An important method for overcoming this problem is to employ inductive learning from examples of expert decisions. In this knowledge acquisition paradigm, knowledge engineers do not have to force experts to state their "know how" in a predefined representational formalism. Experts are asked only to provide correct interpretation of existing domain data or to supply examples of their performance. It is known that experts are better at providing good examples and counterexamples of decisions than at formalizing their knowledge in the form of decision rules. Early experiments exploring this paradigm have also shown that decision rules formed by inductive learning may outperform rules provided by human experts [Michalski & Chilausky 80, Quinlan 83].

This paper describes briefly an inductive incremental learning program AQ15 that learns attributional descriptions from examples. As an important aspect of development of learning systems is their evaluation using practical problems, we also present results of applying AQ15 to three medical domains: lymphography, prognosis of breast cancer recurrence, and location of primary tumor. These three domains are characterized by consecutively larger amounts of inconsistent and sparse learning events.

The evaluation was done from the viewpoint of *classification accuracy* of the induced rules on new objects and *complexity* of the rules. Examples of a few hundred patients with known diagnoses were available, along with the assessed classification accuracy of human experts. We randomly selected 70% of examples for rule learning and used the rest for rule testing. For each domain, the experiment was repeated four times. The induced rules reached the classification accuracy of human experts. Performance of experts was measured in two out of three domains, (breast cancer and primary tumor) testing four and five experts, respectively. The experiments revealed an interesting phenomenon that by *truncating* rules and applying *flexible* rule matching one may significantly reduce the size of the knowledge base without decreasing its performance accuracy.

2. DESCRIPTION OF AQ15

The program AQ15 is a descendant of the GEM program [Reinke 84] and the AQ1–AQ11 series of inductive learning programs, e.g., [Michalski & Larson 75]. Its ancestors were experimented with in the areas of plant disease diagnosis [Michalski & Chilausky 80, Reinke 84], chess end–games [Reinke 84], diagnosis of cardiac arrhythmias [Mozetic 86], and others. This section provides a brief description of AQ15 and its basic features. A more detailed presentation is in [Hong, Mozetic & Michalski 86].

All these systems are based on the AQ algorithm, which generates decision rules from a set of examples, as originally described in [Michalski 69] and [Michalski & McCormick 71]. When building a decision rule, AQ performs a heuristic search through a space of logical expressions to determine those that account for all positive examples and no negative examples. Because there are usually many such *complete* and *consistent* expressions, the goal of AQ is to find the most

preferred one, according to a flexible extra–logical criterion. This criterion is defined by the user to reflect the needs of the application domain. When input data may include inconsistent and/or incorrect learning events, it may be advantageous to develop *incomplete* and/or *inconsistent* descriptions. We tested this hypothesis using the TRUNC method of rule reduction and obtained results that were quite unexpected. The results seem to indicate that the TRUNC method may be useful not only for learning from inconsistent and incorrect examples, but also for learning from perfect examples. The method is described in sections 3 and 4, and the results in section 5.

Learning examples are given in the form of *events*, which are vectors of attribute values. Attributes may be of three types: nominal, linear or structured (the domain is a hierarchy). Events represent different decision classes or, generally, concepts. Events from a given class are considered its *positive examples*, and all other events are considered its *negative examples*. For each class a decision rule is produced that covers all positive examples and no negative ones. Rules are represented in VL_1 (Variable–valued Logic system 1) notation [Michalski & Larson 75]. VL_1 is a multiple–valued logic attributional calculus with typed variables. A *selector* relates a variable to a value or a disjunction of values, e.g.:

$$[\text{Weather\_type} = \text{cloudy} \vee \text{rain}]$$

A conjunction of selectors forms a *complex*. The following complex states that the weather is cloudy, the temperature is greater than 60 degrees, and winds blow from the South or West:

$$[\text{Weather\_type} = \text{cloudy}] \;\&\; [\text{Temp} > 60] \;\&\; [\text{Wind\_direction} = \text{South} \vee \text{West}]$$

Complexes are assembled into *covers*. A cover is a disjunction of complexes describing all positive examples and none of the negative examples of the concept. A cover is formed for each decision class separately. It defines the condition part of a corresponding decision rule. The following are two examples of decision rules:

$$[\text{Transport} = \text{car}] \;\Leftarrow\; [\text{Weather\_type} = \text{cloudy} \vee \text{rain}] \;\vee\; [\text{Temp} = 40..60]$$

$$[\text{Transport} = \text{bike}] \;\Leftarrow\; [\text{Weather\_type} = \text{sun}] \;\&\; [\text{Temp} > 60]$$

As one can see, the rules are easy to interpret. This ease of interpreting AQ15 generated rules is one of the most attractive features of the program. The major idea behind the covering algorithm is to generate a cover in steps, each step producing one conjunctive term (complex) of the cover. Each step starts with focusing attention on one selected positive example (a *seed*). The algorithm generates a set of all complexes (a *star*) which cover the seed and do not cover any negative examples, and then selects the best complex from the star according to the user defined criteria. The basic *covering algorithm* is as follows:

While partial cover does not cover all positive examples
do
1. select an uncovered positive example (a *seed*),
2. determine maximally general complexes covering the seed and no negative examples (generate a *star*),
3. select the best complex from the star according to the user–defined problem–dependent preference criteria,
4. generate a new partial cover by adding the best complex to the current cover.

At the end, a partial cover becomes a cover of the class.

The algorithm starts with an initial cover that is either empty, was previously learned, or is supplied by the user. Extending the seed against all the negative examples, i.e. *generating a star* in step 2, is again a multistep procedure which can be described as follows:

While partial star covers some negative examples
do
1. select a covered negative example,
2. generate all maximally general hypotheses that cover the seed and exclude the negative example; the resulting set is called a *partial star* of the seed against the negative example,
3. generate a new partial star by intersecting the current partial star with the partial star of the seed against the negative example,
4. trim the partial star if the number of disjoint complexes exceeds the user defined threshold (the *maxstar* parameter).

At the end, a partial star becomes the star of the seed, i.e., the set of maximally general complexes covering the seed and not covering any negative example.

The procedure starts with an initial star which is either the entire event space or a complex from the initial cover. If the star generating procedure were to work exhaustively, the search space for covers might grow very rapidly with the number of negative examples and the number of variables used. To deal with this problem, a parameter (*maxstar*) controls how many disjoint complexes may be kept in a partial star. If the number of its disjoint complexes exceeds the parameter, the star is trimmed according to the user specified *criteria*. A typical criterion is: first "maximize the number of positive examples covered" and then, in the case of a tie, "minimize the number of selectors" or "minimize the total cost of variables used".

The program is able to produce rules of different degrees of *generality*. Rules may be *general* (having minimum number of variables, each with maximum number of disjunctive values), *minimal* (minimum number of both, variables and values), or *specific* (maximum number of variables, each with minimum number of values).

AQ15 has the *incremental learning* facility. The user may supply his decision hypotheses as initial rules. The system implements the method of *learning with full memory*. In this type of learning the system remembers all learning examples that were seen so far, as well as the rules it formed. By this method, as opposed to *learning with partial memory*, new decision rules are guaranteed to be correct with respect to all (old and new) learning examples [Reinke 84, Reinke & Michalski 86].

When learning from inconsistent examples, the system provides three options: a) inconsistent examples are treated as positive examples, b) as negative examples, or c) are removed from the data; in this case their membership is decided by the learning process. If statistical information about the probability of inconsistent examples is available, they are preclassified according to the maximum likelihood [Michalski & McCormick 71].

A form of *constructive induction* is implemented in AQ15 as well. The program's background knowledge is expressed in the form of rules, used to generate new attributes not present in input data. The background knowledge rules are of two types: L–rules (logic) that define values of new variables by logical expressions, and A–rules (arithmetic) that introduce new variables as arithmetic functions of original variables. The L–rules and A–rules are two different representations of domain knowledge relevant to the learning process. The L–rules permit one to represent background concept definitions, constraints among the concepts, concept generalization hierarchies, causal dependencies, etc. Concepts known to the program or learned by the program are also added to the stock of L–rules. The algorithm attempts to use new variables to produce better decision rules. The following is an example of a simple L–rule:

$$[\text{Temp} < 32] \Rightarrow [\text{Weather\_type} \neq \text{rain}] \ \& \ [\text{Amount\_of\_rain} = \text{NA}]$$

The program is also capable of automatically testing the learned rules on new events. It produces a *confusion matrix* that shows for each concept and event the degree of match, according to the flexible rule interpretation method (see Section 4). Thus it seems to be an ideal tool for experimenting on inductive knowledge acquisition in a variety of practical domains. AQ15 is implemented in Berkley Pascal and runs under the Unix operating system on VAX and SUN machines. It consists of approximately 13.000 lines of code.

3. TRUNCATION OF RULES AND FLEXIBLE MATCHING

Most human concepts are structures with flexible, imprecise boundaries. They can match different instances with varying degrees of precision and have context–dependent meaning. Flexible boundaries permit one to use concepts beyond the typical range; imprecise boundaries are useful for avoiding superfluous or undesirable precision. When building a learning or inference system, two crucial issues are the way in which concepts are represented, and the way in which they are recognized.

As pointed out in [Michalski 86], the meaning of a concept can be distributed between its *base representation* and the *method of its interpretation*. The base representation explicitly states the typical, context–independent properties of the concept. The interpretation method determines whether a given instance satisfies the base concept description by conducting inference – deductive, analogical or inductive – using contextual information and background knowledge. The method may give a yes–no answer or may determine the degree to which the instance satisfies the base concept representation.

Such a *two–tiered* concept representation yields a spectrum of possibilities. At one extreme, all the concept properties are explicitly defined, including any concept variations and exceptions.

This may lead to a very complex and unwieldy concept representation. The concept recognition process, however, would involve merely a simple matching of the properties of an instance with the information in the concept description. At the other end of the spectrum, the concept is explicitly represented only by a simple prototypical description characterizing its ideal form. Such a prototypical description does not have to relate to a single object like in the family resemblance case [Rosch & Mervis 75, Murphy & Medin 85] but may be an abstract concept specification (e.g., a logical formula involving disjunction). The process of concept recognition using a prototypical description is more complicated. Instead of seeking a strict match (a satisfaction of a complex description), the system determines the degree of similarity between the prototypical (ideal) concept description and the given instance, and compares it with the results from matching the instance with other ideal concept descriptions. The concept that gives the best match is assigned to the instance. This method saves memory for concept representation at the expense of more complicated matching procedure. The matching procedure may be the same for a class of concepts, which increases the cost–effectiveness. Also, by changing the concept interpretation method one may affect the concept recognition process *without* changing the concept representation, and thus may apply the concept to new situations, not originally planned.

Depending on the costs associated with storing a representation and performing the inference, the most effective distribution of meaning between the concept representation and interpretation corresponds to some point within the above spectrum. Interesting research problems are to determine this point of optimal balance, and to find out what concept interpretation methods should be used in different situations. Some preliminary experimental results on the last problem are discussed in [Michalski & Chilausky 80], and more recently in [Uhrik 85].

Let us illustrate the above ideas by the knowledge representation used in AQ15. In this program concepts are represented by a disjunction of conjunctive expressions (complexes). Each expression is associated with a pair of weights: t and u, representing the *total* number of instances (events) explained by the expression, and the number of events explained *uniquely* by that expression, respectively. The complexes are ordered according to decreasing values of the t–weight. The t–weight may be interpreted as a measure of the typicality or the representativeness of a complex as a concept description. The complex with the highest weight (t–weight) may be interpreted as describing the most typical examples of the concept. It may also be viewed as a prototypical or the ideal definition of the concept. On the other hand the complexes with lowest u–weight can be viewed as describing rare, exceptional cases. If the learning events from which rules are derived are noisy, such "light" complexes may be indicative of errors in the data.

Two methods of recognizing the concept membership of an instance are distinguished: the *strict* match and the *flexible* match. In the strict match, one tests whether an instance satisfies condition part of a rule (or, generally, if it can be logically derived from it). In the flexible match, one determines the degree of similarity or conceptual closeness between the instance and the condition part. Using the strict match, one can recognize a concept without checking other candidate concepts, i.e., without taking into consideration the context. In the flexible match, one needs to perform inference involving an event and candidate rules, and determine the most similar concept that best "matches" the instance. The flexible matching can be accomplished in a variety of ways, ranging from approximate matching of features through deduction and analogy,

to *conceptual cohesiveness* that employs inductive inference [Michalski & Stepp 83].

The above weight–ordering of complexes suggests an interesting possibility. Suppose we have a t–weight ordered disjunction of complexes, and we remove from it the lightest complex. So truncated description will not strictly match events that uniquely satisfy the truncated complex. However, by applying a flexible match, these events may still come out to be the most closely related to the correct concept, and thus be correctly recognized. A truncated description is, of course, simpler but carries a potentially higher risk of recognition error, and requires a more sophisticated evaluation. We can proceed further and remove the next "light" complex from the cover, and observe the performance. Each such step produces a different trade–off between the complexity of the description on one side, and the risk factor and the evaluation complexity on the other (Figure 1). At some step the best overall result may be achieved for a given application domain. This method of knowledge reduction by truncating ordered covers and applying a flexible matching is called TRUNC.

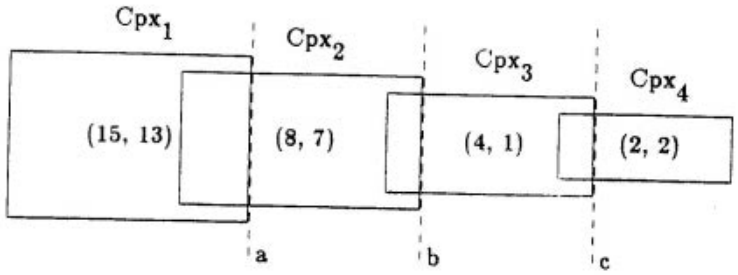

Figure 1. An example of a t–ordered cover. The cuts at a, b and c mark truncated covers with 1, 2 or 3 complexes, respectively. In each pair (x,y), x represents the t–weight, and y represents the u–weight.

The above described trade–off is related to the issues studied in Variable Precision Logic, which is concerned with trade–offs between certainty, computational costs and specificity of inferences [Michalski & Winston 86]. An interesting problem is to test how the cover truncation method affects the accuracy of recognition and the complexity of the decision rules in different practical settings. Section 5 presents results of some such experiments, which in some cases came out very surprising. We now turn to the problem of flexible matching used in this study, and the resolution of a conflict when several concept descriptions are satisfied by an event.

4. FLEXIBLE RULE INTERPRETATION

When strictly matching a new event against a set of (disjunctive) rules, three outcomes are possible: only one rule may be matched (satisfied), more than one rule may be matched, or no rule may match. These cases are classified into categories called SINGLE, MULTIPLE and NO_MATCH, respectively (Figure 2). Each category requires a different evaluation procedure, and a different method of determining the accuracy of concept recognition. For exact match (category SINGLE), the evaluation is easy: the decision is counted as correct if it is equal to the

known diagnosis of the testing object, and as wrong otherwise. If there are several exact matches (the MULTIPLE case) or none (the NO_MATCH case) the system activates the *approximate, context-dependent scheme* that determines the best decision (or the most probable one). Comparing this decision with the decision provided by experts, one evaluates it as correct or incorrect. The scheme consists of two simple heuristic evaluation criteria, one for the MULTIPLE case, and the other for the NO_MATCH case.

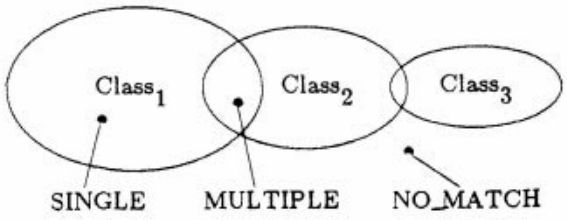

Figure 2. The three possible cases when matching a new event against a set of decision rules.

Estimate of probability for the MULTIPLE case (EP). When an event matches a few, rules the system selects the one which suggests the most probable decision. Let $C_1, \ldots, C_n$ denote decision classes and e an event to be classified. For each decision class C_i we have a rule that consists of a disjunction of complexes (*Cpx*), which, in turn are conjunctions of selectors (*Sel*). We define the estimate of probability, *EP*, as follows:

1) *EP* of a complex Cpx_j in the context of the event e is the ratio of the weight of the complex (the number of positive learning examples covered by the complex) by the total number of learning examples (*#examples*), if the complex is satisfied by the event e, and equals 0 otherwise:

$$EP(Cpx_j, e) = \begin{cases} Weight(Cpx_j) \;/\; \#examples & \text{if complex } Cpx_j \text{ is satisfied by } e, \\ 0 & \text{otherwise.} \end{cases}$$

2) *EP* of a class C_i is the probabilistic sum of *EP*s of its complexes. If the rule for C_i consists of a disjunction of two complexes $Cpx_1 \vee Cpx_2$, we have:

$$EP(C_i, e) = EP(Cpx_1, e) + EP(Cpx_2, e) - EP(Cpx_1, e) \times EP(Cpx_2, e)$$

The most probable class is the one with the largest *EP*, i.e., the one whose satisfied complexes cover the largest number of learning examples. It is assumed that the learning examples are a representative sample of the domain, and that the numbers of examples for each class are proportional to the frequency of occurence of classes. Obviously, if the class is not satisfied by the given event, its *EP* equals 0. For each C_i this measure determines the number of learning examples that support the classification of the new event into class C_i. The larger such number

is, the stronger support is assumed.

Measure of fit for the NO_MATCH case (MF). In this case the event belongs to a part of the event space that is not covered by any decision rule and this calls for flexible matching. One way to perform such matching is to measure the fit between attribute values in the event and the class description, taking into consideration the prior probability of the class. We used in the experiments a simple measure, called *measure of fit, MF*, defined as follows:

1) *MF* of a selector Sel_k and an event e is 1, if the selector is satisfied, i.e. if one of event's attribute values lies in the range of values of the selector. Otherwise, this measure is proportional to the amount of the decision space covered by the selector:

$$MF(Sel_k, e) = \begin{cases} 1 & \text{if selector } Sel_k \text{ is satisfied by } e, \\ \dfrac{\#Values}{DomainSize} & \text{otherwise.} \end{cases}$$

where $\#Values$ is the number of disjunctively linked attribute values in the selector, and $DomainSize$ is the total number of the attribute's possible values.

2) *MF* of a complex Cpx_j to an event e is defined as the product of *MF*s for a conjunction of its constituent selectors, weighted by the proportion of learning examples covered by the complex:

$$MF(Cpx_j, e) = \prod_k MF(Sel_k, e) \times (Weight(Cpx_j) \;/\; \#examples)$$

3) *MF* of a class C_i to en event e is obtained as a probabilistic sum for a disjunction of complexes. If the rule for C_i consists of a disjunction of two complexes $Cpx_1 \vee Cpx_2$, we have:

$$MF(C_i, e) = MF(Cpx_1, e) + MF(Cpx_2, e) - MF(Cpx_1, e) \times MF(Cpx_2, e)$$

We can interpret the measure of best fit of a class as a combination of "closeness" of the event to the class and an estimate of the prior probability of the class. Closeness is measured by *MF* of selectors, where the fit is complete for selectors that are satisfied. *MF* of an unsatisfied selector is the probability that it will be satisfied if the event's corresponding attribute value changes. A selector that covers more decision space fits an event better than a selector that covers less decision space (having fewer alternative values). Closeness to a complex is the probability that the event will be covered by the complex if the values of attributes corresponding to unsatisfied selectors change. *MF* of a complex is then weighted by an estimate of priori probability, i.e., the proportion of the learning examples that it covers. Note that the estimate of probability *EP* is a special case of the measure of fit *MF*; when all selectors in a complex are satisfied the measure of fit of a complex is the same as the estimate of probability.

The above measure of fit is one of many possible measures that can be devised for flexible matching. One way to improve this measure would be to define a distance between an attribute value and a selector, when attributes are linear [Michalski & Chilausky 80].

5. EXPERIMENTS

The experiments were performed on data from three medical domains: lymphography, prognosis of breast cancer recurrence and location of primary tumor. All data were obtained from the Institute of Oncology of the University Medical Center in Ljubljana, Yugoslavia [Kononenko, Bratko & Roskar 84].

Lymphography. This domain is characterized by 18 attributes and 4 diagnostic classes. Data of 148 patients were available. The set of attributes was *complete*, i.e., was sufficient for having all learning examples consistent. This means that examples of any two classes were always different. Diagnoses in this domain were not verified and actual testing of physicians was not done. A specialist's estimation is that internists diagnose correctly in about 60% and specialists in about 85% of cases.

Prognosis of Breast Cancer Recurrence. For about 30% of patients that undergo a breast cancer operation, the illness reappears in five years. Prognosis of this recurrence is very important for patients' post–operational treatment. The domain is characterized by 2 decision classes and 9 attributes. The set of attributes was incomplete, i.e., not always sufficient to distinguish between cases with different prognosis. Data for 286 patients with known diagnostic status 5 years after the operation were available. Five specialists of the Institute of Oncology were tested. They gave a correct prognosis in 64% of cases.

Location of Primary Tumor. Physicians distinguish among 22 possible locations of primary tumor. Patients' diagnostic data were described by 17 attributes. The given set of attributes was incomplete, as some patients with the same values of all attributes had different location of primary tumor. Data of 339 patients with known locations of primary tumor (verified by operation or by X–ray) were available for the experiment. At the Institute of Oncology 4 internists and 4 specialists were tested. Internists determined a correct location of primary tumor in 32% and oncologists in 42% of test cases. Regarding these relatively low results, we should stress that there are 22 possible locations and that the correct location of primary tumor is only one of the sources of evidence used in cancer treatment.

Table 1 provides a summary of these medical domains. It presents the number of examples, of classes, of attributes, and the average number of values per attribute for each domain.

In all medical domains 70% of examples were selected for learning and the remaining 30% for testing. Each testing experiment was repeated 4 times with randomly chosen learning examples. Final results are the average of 4 experiments.

Domain	Examples	Classes	Attributes	Values/Attr
Lymphography	148	4	18	3.3
Breast cancer	286	2	9	5.8
Primary tumor	339	22	17	2.2

Table 1. Characteristics of the data for the three medical domains.

For illustration, in Figure 3 there is an example of a paraphrased rule from the domain of lymphography. Complete rules for all three domains are in the Appendix.

Diagnosis = lymphoma if:

Filling_defects_lacunar = none ∨ lacunar ∨ lacunar_central
Special_structures_and_forms = none ∨ bladder **Base**
Lymph_nodes_size_diminishing = 0 **complex**
Lack_of_lymph_nodes_filling = yes
No_of_diseased_lymph_nodes ≥ 10 (t-weight:40, u-weight:22)

∨

Filling_of_lymp_nodes = grains ∨ fine_drops ∨ dispersed ∨ obscure
Special_structures_and_forms = cup ∨ bladder
Early_filling_of_lymp_nodes = yes
Block_of_afferrent_vessels = no
By_pass = no (t-weight:24, u-weight:7)

∨

Special_structures_and_forms = cup ∨ bladder
Lymph_vessels = curves ∨ deformities
Lymph_nodes_size_enlarged = 1..2
Block_of_afferent_vessels = no
Dislocation_of_lymph_nodes = yes (t-weight:18, u-weight:3)

∨

Filling_of_lymp_nodes = fine_drops ∨ stripes ∨ obscure
Filling_defects_various = follicular ∨ gross_central
Lymph_nodes_size_enlarged = 1..3
Block_of_lymph_nodes_chain = no
Extravasates = yes (t-weight:10, u-weight:3)

∨

Changes_of_lymph_nodes_shape = oval
No_of_diseased_lymph_nodes = 30..39 (t-weight:2, u-weight:1)

Figure 3. A complete rule, generated by AQ15 from all available examples, with t-ordered complexes, for the domain of lymphography. The rule consists of 5 complexes and 22 selectors. After truncation to the "base complex" (with the highest t-weight) the rule has only 1 complex with 5 selectors. T-weight is the total number of examples covered by a complex, and u-weight is the number of examples covered by the complex uniquely.

Two sets of experiments were performed. In the first one only rules of the minimal type were used. Different cover reduction mechanisms were applied on them, and their effect on complexity and classification accuracy of rules was determined. Complexity was measured by the total number of selectors and complexes in the rules, and accuracy by the "1st choice correct" evaluation method (Table 2). In the second set of experiments we measured classification accuracy by two parameters: correctness and precision. We used rules of different degree of generality, applied different evaluation methods and used two cover reduction mechanisms to find the optimal combination of correctness and precision (Table 3).

Domain	Cover truncation	Complexity Sel	Cpx	Accuracy 1st choice	Human Experts	Random Choice
Lymphography	no	**37**	**12**	**81%**	85% (estimate)	25%
	unique >1	34	10	80%		
	base cpx	**10**	**4**	**82%**		
Breast cancer	no	**160**	**41**	**66%**	64%	50%
	unique >1	128	32	66%		
	base cpx	**7**	**2**	**68%**		
Primary tumor	no	551	104	39%	42%	5%
	unique >1	257	42	41%		
	base cpx	112	20	29%		

Table 2. Average complexity and accuracy of AQ15's rules (minimal type) learned from 70% of examples, over 4 experiments. Two simple cover truncation mechanisms were applied - keeping only complexes that uniquely cover more than one example (unique >1), and deleting all but the heaviest complex in each rule (base cpx).

In addition to results obtained from using complete (untruncated) rules, results of two other experiments are presented. In the first experiment we eliminated from rules all complexes that cover uniquely only one learning example, and in the second we eliminated all complexes except the most representative one that covers the largest number of learning examples. Complexity of rules is measured by the number of selectors and complexes. Table 2 shows that some results came out very surprising. When the cover of each class was truncated to only one (the heaviest) complex, the complexity of the rule set for lymphography went down from the total of 12 complexes and 37 selectors to only 4 complexes (one per class) and 10 selectors (see bold numbers). At the same time the performance of rules went slightly up (from 81% to 82%) ! A similar phenomenon occurred in the breast cancer domain, where the number of selectors and complexes went down from 160 and 41 to 7 and 2, respectively; while the performance went slightly up from 66% to 68%. This means that by using the TRUNC method one may significantly reduce the knowledge base without affecting its performance accuracy. Results for human experts were the average of testing of five and four domain specialists in the domains of breast cancer recurrence and primary tumor, respectively [Kononenko, Bratko & Roskar 84]. In the domain of lymphography, physicians' accuracy is given only as their own estimate; it was not independently measured.

In practice, giving always exactly one answer (1st choice) is often not the most appropriate. One might wish to get more than just one possible diagnosis, or none if there is not enough evidence. If any of the alternative diagnoses given by the system is the same as the known diagnosis of the testing example the answer is counted as a correct one. However, the more alternative diagnoses, the smaller diagnostic precision of the system. Therefore, in evaluation of such a system, the results should be measured by two quantities: *correctness* (the ratio of the number of correct answers by the number of testing examples), and *precision* (the ratio of the number of correct answers by the total number of answers given).

Domain	Evaluation method for MULTIPLE	Evaluation method for NO-MATCH	Type of rules	All cpx Corr.	All cpx Prec.	Best cpx Corr.	Best cpx Prec.
Lymphography	1st choice correct	1st choice correct	specific minimal general	79% 81% 81%	79% 81% 81%	80% 82% 81%	80% 82% 81%
	correct if match	always incorrect	specific minimal general	63% 78% 86%	85% 77% 74%	52% 58% 58%	94% 89% 87%
	correct if match	1st choice correct	specific minimal general	80% 83% 89%	78% 76% 74%	81% 83% 82%	81% 82% 81%
Breast cancer	1st choice correct	1st choice correct	specific minimal general	68% 66% 65%	68% 66% 65%	67% 68% 65%	67% 68% 65%
	correct if match	always incorrect	specific minimal general	59% 77% 86%	64% 57% 54%	13% 16% 17%	67% 67% 64%
	correct if match	1st choice correct	specific minimal general	72% 80% 86%	62% 58% 54%	68% 68% 66%	68% 68% 66%
Primary tumor	1st choice correct	1st choice correct	specific minimal general	41% 39% 39%	41% 39% 39%	33% 29% 29%	33% 29% 29%
	correct if match	always incorrect	specific minimal general	33% 50% 51%	31% 24% 24%	22% 25% 25%	44% 34% 34%
	correct if match	1st choice correct	specific minimal general	47% 52% 53%	34% 24% 24%	35% 32% 32%	33% 28% 28%

Table 3. Trade–offs between correctness and precision of AQ15's rules for different evaluation methods and different types of rules.

Several experiments with AQ15 were performed to evaluate trade–offs between correctness and precision (Table 3). This trade–off is a reflection of the phenomena studied in Variable Precision Logic [Michalski & Winston 84]. We tried three different evaluation schemes (representing

combinations of 1st choice correct, correct if match and incorrect for MULTIPLE and NO_MATCH cases). The "1st choice correct" means that the best flexible matching was used. The "correct if match" for MULTIPLE and "incorrect" for NO_MATCH mean that the strict match method was used. We also used rules of different degree of generality (specific, minimal, general) and different cover reduction mechanism. "All cpx" and "Best cpx" mean complete cover and the cover truncated to one complex, respectively. One of the interesting future research tasks is to find an appropriate information–theoretic measure for defining an optimal combination of correctness and precision.

6. ANALYSIS OF RESULTS

The domain of lymphography seems to have some strong patterns and the set of attributes is known to be complete, i.e., no event description belongs to more than one class. There are four possible diagnoses, but only two of them are prevailing, i.e., they occur much more often than others. The domain of breast cancer has only two decision classes, but does not have many strong patterns. Domain of location of primary tumor has many decision classes and mostly binary attributes. There are only a few examples per class, and the domain seems to be without any strong patterns. Both domains are underspecified in the sense that the set of available attributes is incomplete (not sufficient to discriminate between different classes). The statistics in Table 4 include average number of complexes per rule, average number of attributes per complex, average number of values per attribute and finally, average number of learning examples covered by one complex. We can see that in the domain of primary tumor decision rules consist of complexes that in average cover slightly more than 2 examples. In the domain of lymphography complexes in average cover 8 examples, which indicates a presence of strong patterns.

Domain	Cpx/Rule	Attr/Cpx	Values/Attr	Examples/Cpx
Lymphography	3	3.1	1.8	8
Breast cancer	20	3.9	1.7	5
Primary tumor	5.2	5.3	1.0	2.3

Table 4. Average complexity of AQ15's decision rules (minimal type) in the three medical domains, when no cover truncation mechanism was applied.

Several experiments with AQ15 were performed, each with a different complex truncation heuristic. This was done in order to investigate the trade–off between complexity and accuracy, and to derive some preliminary conclusions about the effects of the cover reduction mechanism. Results given in Table 2 present only two extreme cases of these experiments. By eliminating all complexes but one, a significant reduction of complexity was obtained. Except for the primary tumor domain, there was no decrease of accuracy.

In the domain of primary tumor, initial elimination of lightest complexes (those that cover only 1 example) increased accuracy from 33% to 41%; accuracy decreased when further complexes were

eliminated. In the domain of lymphography accuracy increased until only one "heaviest" complex in the two most important rules was kept (82%). In the breast cancer domain each step of elimination of complexes increased accuracy as well. Best results were obtained when all complexes for the class "recurrence" were deleted. The obtained diagnostic accuracy was 72% which is close to a priori probability of the diagnosis "no recurrence".

It is surprising that a cover reduction mechanism that strongly simplifies the rule base may have no affect on classification accuracy. Removing complexes from a cover is equivalent to removing disjunctively linked conditions from a concept description. This process overspecializes a knowledge representation, producing an *incomplete* concept description (i.e., a one that does not cover some positive examples).

Such knowledge reduction technique by specialization may be contrasted with knowledge reduction by generalization used in the ASSISTANT learning program, a descendant of ID3 [Quinlan 83]. This program represents knowledge in the form of decision trees, and has been applied to the same medical problems as here [Kononenko, Bratko & Roskar 84]. The program applies a *tree pruning* technique based on the principle of maximal classification accuracy. The technique removes certain nodes from a tree, and is equivalent to removing conjunctively linked conditions from a concept description. Thus, such a knowledge reduction technique overgeneralizes the knowledge representation, producing an *inconsistent* concept description (i.e., a one that covers some negative examples). It is interesting to point out that this technique may also lead to an improvement of accuracy in decision making when learning from noisy and overlapping data. Table 5 presents the complexity and diagnostic accuracy of ASSISTANT's trees built with and without tree pruning [Kononenko, Bratko & Roskar 84]. Complexity of trees is given by the number of nodes and leaves. In all domains results were better when the tree pruning mechanism was used.

Domain	Tree pruning	Complexity Nodes	Complexity Leaves	Accuracy 1st choice
Lymphography	no	38	22	76%
	yes	25	14	77%
Breast cancer	no	120	63	67%
	yes	16	9	72%
Primary tumor	no	188	90	41%
	yes	35	18	46%

Table 5. Average complexity and accuracy of decision trees built by ASSISTANT on 70% of examples, over 4 experiments. In all three domains the tree pruning mechanism reduced the complexity and increased the accuracy. Note that more than one decision may be assigned to some leaves (hence there are only 18 leaves for 22 classes in the primary tumor case).

Tree pruning corresponds to the removal of selectors from complexes. This seems to suggest that when learning from noisy or inconsistent examples the knowledge reduction process may not only

involve removal of complexes from a cover (a specialization process) but also removal of selectors from complexes (a generalization process). This means that the generated concept description would be both inconsistent and incomplete. It is an interesting problem for further research to determine conditions under which such inconsistent and incomplete descriptions might be more advantageous than consistent and complete ones.

7. CONCLUSION

A major contribution of the paper is a demonstration that a relatively simple, attribute–based inductive learning method is able to produce decision rules of sufficiently high quality to be applicable to practical problems with noisy, inconsistent and/or incompletely specified learning examples. An especially important for practical applications is perhaps the fact that the method produced these results without using a large amount of domain knowledge that would be required by an analytic approach or explanation–based generalization [Mitchell, Keller & Kedar–Cabelli 86; DeJong & Mooney 86]. It relied primarily on learning examples that were obtained from already existing records or human experts. It is well known that it is typically easier for an expert to make decisions (i.e., to produce examples) than to formulate a theory justifying them.

Although the program can work with relatively little domain knowledge (e.g., only the specification of types and domains of attributes, and the preference criterion), it can also take advantage of the domain knowledge when it is available. The latter is realized by employing background knowledge representation facilities in the form of logical and arithmetical rules (L–rules and A–rules).

The AQ15 program has shown itself to be a powerful and flexible tool for experimenting with inductive knowledge acquisition. It produces decision rules which are easy to interpret and comprehend. The knowledge representation in the program is, however, limited to only attributional descriptions. For problems that require structural descriptions one may use a related program INDUCE2 [Hoff, Michalski & Stepp 83] or its incremental learning version INDUCE4 [Mehler, Bentrup & Riedsel 86]. A weakness of the experimental part of the paper is that the authors had no influence on the way the data were prepared for the experiments and the available data allowed us to test only a few of the features of AQ15.

Another major result is a demonstration that the knowledge reduction by truncating the covers may lead in some cases to a substantial reduction of the rule base without decreasing its performance accuracy. We have also shown that by varying the degree of generality of rules and applying different evaluation methods, different trade–offs between the correctness and precision of decision rules are achieved. Further research will be required to find for any given domain a rule reduction criterion that leads to the best trade–off between accuracy and complexity of a rule base. Another topic for further research is to develop more sophisticated methods for flexible matching.

ACKNOWLEDGEMENTS

The authors thank Ivan Bratko and Igor Kononenko from the Faculty of Electrical Engineering at the E. Kardelj University in Ljubljana for collaboration and comments, and physicians Matjaz Zwitter and Milan Soklic from the Institute of Oncology at the University Medical Center in Ljubljana for providing medical data and helping to interpret them. We further acknowledge Gail Thornburg from the UI School of Library and Information Science and the AI Laboratory at the Dept. of Computer Science for her criticism and valuable suggestions.

REFERENCES

DeJong G., Mooney, R., Explanation–Based Learning: An Alternative View. *Machine Learning*, Vol. 1, No. 2, 1986.

Hoff, W., Michalski, R.S., Stepp, R.E., INDUCE.2: A Program for Learning Structural Descriptions from Examples. Report ISG 83–4, UIUCDCS–F–83–904, Dept. of Computer Science, University of Illinois at Urbana–Champaign, 1983.

Hong, J., Mozetic, I., Michalski, R.S., AQ15: Incremental Learning of Attribute–Based Descriptions from Examples, the Method and User's Guide. Report ISG 86–5, UIUCDCS–F–86–949, Dept. of Computer Science, University of Illinois at Urbana–Champaign, 1986.

Kononenko, I., Bratko, I., Roskar, E., Experiments in Automatic Learning of Medical Diagnostic Rules. Presented at the *International School for the Synthesis of Expert Knowledge Workshop '84*, Bled, Yugoslavia, August 22–24, 1984. Also Technical Report, Faculty of Electrical Engineering, E. Kardelj University, Ljubljana, Yugoslavia, 1984.

Mehler, G., Bentrup, J., Riedesel J., INDUCE.4: A Program for Incrementally Learning Structural Descriptions from Examples. Report in preparation, Dept. of Computer Science, University of Illinois at Urbana–Champaign, 1986.

Michalski, R.S., On the Quasi–Minimal Solution of the General Covering Problem. *Proceedings of the V International Symposium on Information Processing (FCIP 69)*, Vol. A3 (Switching Circuits), Bled, Yugoslavia, pp. 125–128, 1969.

Michalski, R.S., Theory and Methodology of Machine Learning. In R.S. Michalski, J.G. Carbonell, T.M. Mitchell (Eds.), *Machine Learning – An Artificial Intelligence Approach*. Palo Alto: Tioga, 1983.

Michalski, R.S., Two–tiered Concept Representation, Inferential Matching and Conceptual Cohesiveness. An invited paper for the *Workshop on Similarity and Analogy*, Allerton House, University of Illinois, June 12–14, 1986.

Michalski, R.S., Chilausky, R.L., Learning by Being Told and Learning from Examples: An Experimental Comparison of the Two Methods of Knowledge Acquisition in the Context of Developing an Expert System for Soybean Disease Diagnosis. *International Journal of Policy Analysis and Information Systems*, Vol. 4, No. 2, pp. 125–161, 1980.

Michalski, R.S., Larson, J., AQVAL/1 (AQ7) User's Guide and Program Description. Report No. 731, Dept. of Computer Science, University of Illinois at Urbana–Champaign, 1975.

Michalski, R.S., McCormick, B.H., Interval Generalization of Switching Theory. Report No. 442, Dept. of Computer Science, University of Illinois at Urbana–Champaign, 1971.

Michalski, R.S., Stepp, R.E., Learning from Observations: Conceptual Clustering. In R.S. Michalski, J.G. Carbonell, T.M. Mitchell (Eds.), *Machine Learning – An Artificial Intelligence Approach*. Palo Alto: Tioga, 1983.

Michalski, R.S., Winston, P.H., Variable Precision Logic. Artificial Intelligence Memo No. 857, MIT, Cambridge. An extended version to appear in *AI Journal*, 1986.

Mitchell, T.M., Keller, R.M., Kedar–Cabelli, S.T., Explanation–Based Generalization: A Unifying View. *Machine Learning*, Vol. 1, No. 1, pp. 47–80, 1986.

Mozetic, I., Knowledge Extraction through Learning from Examples. In T.M. Mitchell, J.G. Carbonell, R.S. Michalski (Eds.), *Machine Learning: A Guide to Current Research*. Kluwer Academic Publishers, 1986.

Murphy, G.L., Medin, D.L., The Role of Theories in Conceptual Coherence. *Psychological Review*, Vol. 92, No. 3, 1985.

Quinlan, J.R., Learning Efficient Classification Procedures and their Application to Chess End Games. In R.S. Michalski, J.G. Carbonell, T.M. Mitchell (Eds.), *Machine Learning – An Artificial Intelligence Approach*. Palo Alto: Tioga, 1983.

Reinke, R.E., Knowledge Acquisition and Refinement Tools for the ADVISE META–EXPERT System. M.S. Thesis, ISG 84–4, UIUCDCS–F–84–921, Dept. of Computer Science, University of Illinois at Urbana–Champaign, 1984.

Reinke, R.E., Michalski, R.S., Incremental Learning of Decision Rules: A Method and Experimental Results. To appear in *Machine Intelligence 11* (eds. J.E. Hayes, D. Michie, J. Richards), Oxford University Press, 1986.

Rosch, E., Mervis, C.B., Family Resemblance: Studies in the Internal Structure of Categories. *Cognitive Psychology*, No. 7, 1975.

Uhrik, T.K., A Rule Exerciser for Knowledge Base Enhancement in Expert Systems. M.S. Thesis, Dept. of Computer Science, University of Illinois at Urbana–Champaign, 1985.

APPENDIX

The following are rules produced by AQ15 for all three investigated medical domains. Rules presented were induced from all available examples and are untruncated. The preference criterion was set to: first "maximize the number of newly covered examples" (not yet covered by any previous complex) and then, in the case of a tie, "minimize the number of selectors". Rules are of the minimal type (minimum number of selectors in a complex, and minimum number of values in each selector), and may intersect. Inconsistent examples were treated as positive examples for each class. The maxstar parameter was set to 20. In the rules, "t–weight" is the total number of examples covered by a complex, and "u–weight" is the number of examples uniquely covered by the complex.

1. Lymphography

[lympho_diag = lymphoma] ⇐
[diminish_LN = 0] & [fill_def_lacun = no ∨ lacun ∨ lacun_cent] & [spec_str_form = no ∨ bladder] & [lack_LN_fill = yes] & [diseased_LN = 10..inf] (t–weight:40, u–weight:22)
∨ [block_affer_vess = no] & [by_pass = no] & [early_fill_LN = yes] & [fill_LN = grains ∨ fine_drops ∨ dispersed ∨ obscure] & [spec_str_form = cup ∨ bladder] (t–weight:24, u–weight:7)
∨ [lymph_vess = curves ∨ deform] & [block_affer_vess = no] & [enlarged_LN = 1..2] & [spec_str_form = cup ∨ bladder] & [disloc_LN = yes] (t–weight:18, u–weight:3)
∨ [block_chain = no] & [extravas = yes] & [enlarged_LN = 1..3] & [fill_def_var = follicular ∨ gross_cent] & [fill_LN = fine_drops ∨ stripes ∨ obscure] (t–weight:10, u–weight:3)
∨ [change_LN_shape = oval] & [diseased_LN = 30..39] (t–weight:2, u–weight:1)

[lympho_diag = metastases] ⇐
[block_affer_vess = yes] & [fill_def_lacun = lacun_marg] & [fill_LN = grains ∨ fine_drops ∨ coarse_drops ∨ dispersed ∨ obscure] & [diseased_LN = 0..29] (t–weight:50, u–weight:23)
∨ [fill_def_lacun = lacun ∨ lacun_marg] & [fill_LN = no ∨ grains ∨ coarse_drops ∨ dispersed ∨ obscure] & [spec_str_form = no ∨ bladder] & [diseased_LN = 0..9] (t–weight:21, u–weight:11)
∨ [fill_LN = grains ∨ coarse_drops ∨ dispersed] & [spec_str_form = cup] & [diseased_LN = 0..19] (t–weight:18, u–weight:3)
∨ [lymph_vess = curves] & [early_fill_LN = no] & [lack_LN_fill = yes] (t–weight:12, u–weight:5)
∨ [lymph_vess = curves ∨ displac] & [early_fill_LN = yes] & [enlarged_LN = 1] & [fill_def_var = follicular ∨ tiny] & [fill_LN = grains ∨ fine_drops ∨ coarse_drops ∨ dispersed] (t–weight:8, u–weight:4)
∨ [change_LN_shape = round] & [fill_def_var = gross_cent] & [fill_LN = coarse_drops ∨ reticular] (t–weight:5, u–weight:3)
∨ [block_affer_vess = yes] & [fill_def_var = tiny] & [fill_LN = coarse_drops ∨ obscure] & [diseased_LN = 30..39 ∨ 50..59] (t–weight:2, u–weight:2)

[lympho_diag = fibrosation] ⇐
[by_pass = yes] & [diminish_LN = 1..2] (t–weight:4, u–weight:4)

[lympho_diag = normal] ⇐
[lymph_vess = no] (t–weight:2, u–weight:2)

2. Prognosis of Breast Cancer Recurrence

[breast_cancer = no_recurrence] ⇐
[menopause = lt40 ∨ ge40] & [tumor_size = 0..19] & [inv_nodes = 0..2] (t–weight:35, u–weight:14)
∨ [age = 40..69] & [menopause = ge40 ∨ premeno] & [tumor_size = 20..34] & [deg_malig = 1..2] & [breast_quad = left_low ∨ right_low] & [irradiat = no] (t–weight:30, u–weight:4)
∨ [age = 30..59] & [tumor_size = 0..14] (t–weight:28, u–weight:15)
∨ [menopause = premeno] & [tumor_size = 20..34] & [node_caps = no] & [breast_quad = left_low ∨ right_low ∨ central] & [irradiat = no] (t–weight:27, u–weight:7)
∨ [age = 50..69] & [tumor_size = 25..34] & [deg_malig = 1..2] & [breast_quad = left_up ∨ left_low] (t–weight:24, u–weight:4)
∨ [menopause = ge40 ∨ premeno] & [tumor_size = 10..19 ∨ 25..39] & [deg_malig = 1] & [breast = left] & [irradiat = no] (t–weight:21, u–weight:4)
∨ [age = 50..59] & [inv_nodes = 0..2] & [breast = right] & [breast_quad = left_up ∨ left_low ∨ right_up] & [irradiat = no] (t–weight:20, u–weight:4)
∨ [tumor_size = 25..44 ∨ 50..54] & [inv_nodes = 0..2] & [deg_malig = 2] & [breast_quad = left_up ∨ right_low] (t–weight:17, u–weight:4)
∨ [age = 20..49] & [node_caps = no] & [deg_malig = 2] & [breast_quad = left_low ∨ right_up] & [irradiat = no] (t–weight:15, u–weight:3)
∨ [menopause = premeno] & [tumor_size = 15..29] & [node_caps = no] & [breast = right] & [breast_quad = left_up ∨ right_up ∨ right_low] (t–weight:13, u–weight:4)
∨ [age = 30..59] & [tumor_size = 40..49] & [inv_nodes = 0..5] & [deg_malig = 2..3] (t–weight:12, u–weight:6)
∨ [menopause = premeno] & [tumor_size = 20..29] & [deg_malig = 2] & [breast_quad = left_up] (t–weight:11, u–weight:2)
∨ [age = 50..69] & [tumor_size = 20..29] & [deg_malig = 2..3] & [breast = right] & [breast_quad = left_up] (t–weight:10, u–weight:4)
∨ [age = 60..69] & [tumor_size = 25..34] & [breast_quad = left_low] (t–weight:9, u–weight:3)
∨ [age = 60..69] & [deg_malig = 2] & [breast_quad = left_up] (t–weight:9, u–weight:2)
∨ [age = 40..59] & [tumor_size = 25..34] & [deg_malig = 1] & [breast = right] (t–weight:7, u–weight:2)
∨ [age = 40..59] & [tumor_size = 20..24] & [inv_nodes = 0..2] & [node_caps = no] & [deg_malig = 3] (t–weight:6, u–weight:5)
∨ [inv_nodes = 6..8] & [deg_malig = 2] & [breast_quad = left_up ∨ left_low ∨ right_up] (t–weight:6, u–weight:4)

∨ [age = 40..49] & [tumor_size = 30..34] & [inv_nodes = 0..2] & [node_caps = no] & [breast_quad = left_up ∨ right_up] (t-weight:6, u-weight:3)
∨ [age = 40..59] & [tumor_size = 35..39] & [deg_malig = 3] & [breast_quad = left_up ∨ left_low] (t-weight:5, u-weight:5)
∨ [age = 70..79] & [irradiat = no] (t-weight:5, u-weight:3)
∨ [age = 40..59] & [tumor_size = 30..39] & [inv_nodes = 9..11] & [breast_quad = left_up ∨ left_low ∨ right_up] (t-weight:4, u-weight:4)
∨ [tumor_size = 45..54] & [breast = left] (t-weight:4, u-weight:2)
∨ [age = 40..49] & [tumor_size = 40..44] & [breast = right] (t-weight:4, u-weight:2)
∨ [age = 50..59] & [menopause = ge40] & [tumor_size = 25..29] & [breast_quad = right_up] (t-weight:1, u-weight:1)

[breast_cancer = recurrence] ⇐
[age = 30..49] & [inv_nodes = 3..17] & [node_caps = yes] & [breast = left] (t-weight:11, u-weight:5)
∨ [inv_nodes = 0..8] & [deg_malig = 3] & [breast = left] & [irradiat = yes] (t-weight:10, u-weight:7)
∨ [menopause = ge40 ∨ premeno] & [tumor_size = 30..34] & [deg_malig = 2..3] & [breast = right] & [breast_quad = left_up ∨ central] (t-weight:8, u-weight:8)
∨ [age = 50..69] & [tumor_size = 20..34] & [inv_nodes = 0..5] & [deg_malig = 2] & [breast_quad = right_up ∨ central] (t-weight:6, u-weight:6)
∨ [tumor_size = 30..39] & [deg_malig = 3] & [breast_quad = right_up ∨ right_low] (t-weight:6, u-weight:5)
∨ [age = 40..59] & [inv_nodes = 3..8] & [breast_quad = right_up ∨ right_low] (t-weight:5, u-weight:3)
∨ [age = 60..69] & [tumor_size = 35..49] & [breast_quad = left_low ∨ right_up] (t-weight:4, u-weight:4)
∨ [age = 40..49] & [menopause = premeno] & [tumor_size = 20..29] & [deg_malig = 2..3] & [breast = left] & [breast_quad = left_low ∨ right_up] (t-weight:4, u-weight:3)
∨ [age = 40..59] & [tumor_size = 25..29] & [deg_malig = 2..3] & [breast = right] & [breast_quad = left_low] (t-weight:3, u-weight:3)
∨ [age = 40..59] & [menopause = ge40] & [tumor_size = 30..34] & [inv_nodes = 0..5] & [deg_malig = 3] & [breast = left] (t-weight:3, u-weight:3)
∨ [tumor_size = 20..29] & [deg_malig = 1] & [breast_quad = left_up] & [irradiat = no] (t-weight:3, u-weight:2)
∨ [age = 30..39] & [tumor_size = 35..39] (t-weight:3, u-weight:2)
∨ [age = 60..69] & [menopause = ge40] & [tumor_size = 20..24] & [deg_malig = 2..3] & [breast_quad = left_low] (t-weight:3, u-weight:2)
∨ [menopause = premeno] & [inv_nodes = 9..11] & [irradiat = no] (t-weight:3, u-weight:2)
∨ [age = 50..59] & [menopause = ge40 ∨ premeno] & [tumor_size = 15..24] & [node_caps = no] & [deg_malig = 2] & [breast = left] (t-weight:2, u-weight:2)
∨ [tumor_size = 40..44] & [deg_malig = 1] & [breast = left] (t-weight:2, u-weight:2)
∨ [age = 30..39] & [inv_nodes = 0..2] & [deg_malig = 2] & [breast_quad = left_up ∨ central] (t-weight:2, u-weight:2)
∨ [age = 40..49] & [menopause = ge40] & [tumor_size = 20..24] & [breast = right] (t-weight:2, u-weight:2)

∨ [age = 40..49] & [tumor_size = 15..19] & [breast_quad = left_up] (t-weight:2, u-weight:2)
∨ [age = 30..39] & [tumor_size = 15..19 ∨ 30..34] & [deg_malig = 1] & [breast = right] (t-weight:2, u-weight:2)
∨ [tumor_size = 25..29] & [deg_malig = 3] & [breast = left] & [breast_quad = left_up] (t-weight:2, u-weight:2)
∨ [menopause = lt40] & [tumor_size = 20..24] (t-weight:2, u-weight:1)
∨ [tumor_size = 15..24] & [inv_nodes = 9..11 ∨ 24..26] & [breast = left] (t-weight:2, u-weight:1)
∨ [tumor_size = 35..39] & [deg_malig = 1] & [breast = right] (t-weight:1, u-weight:1)
∨ [age = 40..49] & [tumor_size = 50..54] & [breast = right] (t-weight:1, u-weight:1)
∨ [tumor_size = 50..54] & [deg_malig = 3] (t-weight:1, u-weight:1)
∨ [age = 50..59] & [tumor_size = 35..39] & [inv_nodes = 0..2] & [deg_malig = 2] & [breast = left] & [breast_quad = left_low] & [irradiat = no] (t-weight:1, u-weight:1)
∨ [tumor_size = 30..34] & [deg_malig = 1] & [irradiat = yes] (t-weight:1, u-weight:1)

3. Location of Primary Tumor

[tumor_location = bladder] ⇐
[age = bet30_59] & [sex = male] & [bone = no] & [perit = yes] & [liver = no] & [abdom = no] (t-weight:1, u-weight:1)
∨ [perit = yes] & [neck = yes] & [mediast = yes] (t-weight:1, u-weight:1)

[tumor_location = breast] ⇐
[sex = fem] & [brain = no] & [axillar = yes] & [abdom = no] (t-weight:17, u-weight:14)
∨ [age = bet30_59] & [sex = fem] & [bone = yes] & [lung = no] & [liver = no] & [skin = no] & [neck = no] & [mediast = no] & [abdom = no] (t-weight:5, u-weight:2)
∨ [lung = no] & [axillar = yes] & [mediast = yes] & [abdom = yes] (t-weight:3, u-weight:3)
∨ [age = lt30] & [bone = no] & [axillar = yes] (t-weight:1, u-weight:1)
∨ [lung = yes] & [perit = yes] & [supraclav = yes] (t-weight:1, u-weight:1)

[tumor_location = cerv_uteri] ⇐
[sex = fem] & [deg_diff = fair] & [bone = yes] & [liver = no] & [supraclav = no] & [abdom = yes] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = fem] & [lung = yes] & [perit = no] & [liver = yes] & [mediast = yes] (t-weight:1, u-weight:1)

[tumor_location = colon] ⇐
[age = ge60] & [sex = fem] & [hyst_type = adeno] & [bone = no] & [lung = no] & [perit = no] & [supraclav = no] & [mediast = no] (t-weight:3, u-weight:3)
∨ [age = ge60] & [sex = fem] & [bone = no] & [lung = yes] & [perit = no] & [liver = yes] & [brain = no] & [mediast = yes] (t-weight:2, u-weight:2)

∨ [age = bet30_59] & [sex = fem] & [bone = no] & [lung = yes] & [pleura = no] & [liver = no] & [supraclav = no] (t-weight:2, u-weight:2)
∨ [liver = yes] & [supraclav = yes] & [mediast = no] & [abdom = yes] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = fem] & [deg_diff = fair] & [pleura = no] & [perit = yes] & [liver = no] & [abdom = yes] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = male] & [hyst_type = adeno] & [perit = yes] & [liver = no] & [abdom = no] (t-weight:1, u-weight:1)
∨ [age = lt30] & [liver = yes] & [mediast = no] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = male] & [hyst_type = adeno] & [bone = no] & [pleura = no] & [perit = no] & [liver = yes] & [supraclav = no] & [abdom = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = male] & [hyst_type = adeno] & [pleura = yes] & [liver = yes] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = male] & [deg_diff = well] & [bone = no] & [pleura = no] & [perit = yes] & [abdom = yes] (t-weight:1, u-weight:1)

[tumor_location = corp_uteri] ⇐
[age = bet30_59] & [sex = fem] & [bone = no] & [pleura = no] & [perit = no] & [liver = no] & [supraclav = no] & [mediast = no] & [abdom = yes] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = fem] & [deg_diff = well] & [bone = yes] & [lung = no] & [pleura = no] & [liver = no] & [skin = no] & [axillar = no] & [mediast = no] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = fem] & [lung = yes] & [perit = no] & [liver = yes] & [mediast = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [lung = yes] & [perit = yes] & [liver = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [bone = yes] & [liver = no] & [mediast = no] & [abdom = yes] (t-weight:1, u-weight:1)
∨ [sex = fem] & [lung = no] & [perit = yes] & [supraclav = no] & [axillar = no] & [mediast = yes] (t-weight:1, u-weight:1)

[tumor_location = duod_intest] ⇐
[age = ge60] & [bone = no] & [lung = yes] & [pleura = no] & [supraclav = no] & [abdom = no] (t-weight:1, u-weight:1)

[tumor_location = esophasus] ⇐
[age = bet30_59] & [deg_diff = poor] & [pleura = no] & [neck = yes] & [supraclav = yes] & [abdom = no] (t-weight:2, u-weight:1)
∨ [sex = male] & [hyst_type = epid] & [lung = yes] & [supraclav = yes] (t-weight:2, u-weight:1)
∨ [sex = male] & [bone = yes] & [lung = no] & [mediast = no] & [abdom = yes] (t-weight:1, u-weight:1)
∨ [skin = yes] & [neck = yes] & [supraclav = no] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = male] & [deg_diff = poor] & [bone = yes] & [lung = no] & [liver = no] & [skin = no] & [neck = no] & [mediast = no] (t-weight:1, u-weight:1)

∨ [perit = no] & [liver = yes] & [supraclav = yes] & [mediast = no] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = male] & [lung = yes] & [liver = yes] & [mediast = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [hyst_type = epid] & [bone = no] & [lung = yes] & [abdom = yes] (t-weight:1, u-weight:1)

[tumor_location = gallblader] ⇐
[age = ge60] & [sex = fem] & [bone = no] & [perit = no] & [liver = yes]
& [mediast = no] (t-weight:9, u-weight:2)
∨ [age = ge60] & [sex = fem] & [bone = no] & [lung = no] & [perit = no]
& [abdom = yes] (t-weight:8, u-weight:3)
∨ [age = ge60] & [sex = fem] & [bone = no] & [liver = yes] & [supraclav = no]
& [abdom = no] (t-weight:3, u-weight:1)
∨ [age = bet30_59] & [lung = no] & [pleura = yes] & [perit = yes] & [supraclav = no]
& [abdom = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = fem] & [perit = yes] & [mediast = yes] (t-weight:1, u-weight:1)
∨ [lung = yes] & [brain = yes] & [mediast = no] (t-weight:1, u-weight:1)

[tumor_location = head_neck] ⇐
[bone = no] & [neck = yes] & [supraclav = no] & [mediast = no] (t-weight:17, u-weight:16)
∨ [sex = fem] & [skin = no] & [neck = yes] & [axillar = no] (t-weight:3, u-weight:2)
∨ [age = ge60] & [deg_diff = well] & [neck = yes] (t-weight:1, u-weight:1)

[tumor_location = kidney] ⇐
[age = lt30..bet30_59] & [sex = male] & [bone = yes] & [bone_marr = no]
& [pleura = no] & [liver = no] & [skin = no] & [neck = no] & [abdom = no] (t-weight:9, u-weight:7)
∨ [sex = male] & [bone = yes] & [lung = yes] & [liver = no] (t-weight:6, u-weight:2)
∨ [age = bet30_59] & [sex = male] & [hyst_type = adeno] & [lung = yes]
& [pleura = no] & [liver = no] & [supraclav = no] (t-weight:4, u-weight:1)
∨ [age = ge60] & [bone = no] & [perit = no] & [liver = no] & [neck = no]
& [supraclav = no] & [mediast = no] & [abdom = no] (t-weight:3, u-weight:3)
∨ [age = lt30] & [hyst_type = adeno] & [perit = no] & [liver = no] & [supraclav = yes]
& [axillar = no] & [abdom = no] (t-weight:2, u-weight:2)
∨ [age = bet30_59] & [bone = yes] & [lung = yes] & [perit = no] & [mediast = no] (t-weight:2, u-weight:1)
∨ [age = ge60] & [lung = yes] & [liver = no] & [mediast = yes] (t-weight:2, u-weight:1)
∨ [age = bet30_59] & [deg_diff = well] & [lung = no] & [perit = no] & [neck = no]
& [supraclav = yes] & [axillar = no] & [mediast = no] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [bone = no] & [lung = yes] & [perit = yes] & [liver = yes]
& [mediast = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = fem] & [lung = no] & [pleura = yes] & [brain = no]
& [abdom = yes] (t-weight:1, u-weight:1)

[tumor_location = liver] ⇐
[age = lt30] & [hyst_type = anapl] & [supraclav = yes] & [mediast = no] (t-weight:1, u-weight:1)
∨ [sex = male] & [bone = no] & [pleura = yes] & [perit = yes] & [abdom = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [pleura = yes] & [brain = yes] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = fem] & [bone = yes] & [lung = no] & [pleura = no] & [liver = no] & [skin = no] & [neck = no] & [axillar = no] & [mediast = no] & [abdom = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = fem] & [lung = no] & [liver = no] & [mediast = yes] & [abdom = yes] (t-weight:1, u-weight:1)
∨ [age = ge60] & [bone = no] & [lung = yes] & [pleura = no] & [supraclav = no] & [abdom = no] (t-weight:1, u-weight:1)
∨ [sex = male] & [bone = no] & [lung = yes] & [pleura = yes] & [mediast = yes] (t-weight:1, u-weight:1)

[tumor_location = lung] ⇐
[sex = male] & [deg_diff = fair ∨ poor] & [lung = no] & [mediast = yes] & [abdom = no] (t-weight:15, u-weight:3)
∨ [lung = no] & [brain = yes] & [mediast = yes] (t-weight:11, u-weight:4)
∨ [bone = no] & [perit = no] & [supraclav = yes] & [axillar = no] & [mediast = yes] (t-weight:11, u-weight:3)
∨ [sex = male] & [neck = no] & [supraclav = yes] & [mediast = yes] (t-weight:10, u-weight:1)
∨ [age = ge60] & [sex = male] & [lung = no] & [neck = no] & [mediast = yes] (t-weight:9, u-weight:4)
∨ [sex = male] & [bone = yes] & [lung = no] & [supraclav = no] & [mediast = yes] (t-weight:9, u-weight:1)
∨ [age = bet30_59] & [sex = male] & [lung = no] & [skin = yes] (t-weight:8, u-weight:4)
∨ [sex = male] & [hyst_type = epid] & [bone = yes] & [lung = no] & [pleura = no] & [neck = no] (t-weight:6, u-weight:3)
∨ [age = bet30_59] & [pleura = yes] & [perit = yes] & [mediast = yes] (t-weight:5, u-weight:4)
∨ [age = lt30] & [sex = fem] & [mediast = yes] (t-weight:4, u-weight:3)
∨ [age = bet30_59] & [sex = male] & [bone = no] & [lung = no] & [perit = no] & [liver = no] & [neck = no] & [supraclav = no] & [mediast = no] & [abdom = no] (t-weight:4, u-weight:2)
∨ [age = bet30_59] & [bone = no] & [lung = no] & [pleura = yes] & [perit = no] & [skin = no] & [neck = no] & [abdom = no] (t-weight:4, u-weight:2)
∨ [age = ge60] & [sex = fem] & [perit = no] & [liver = yes] & [abdom = no] (t-weight:4, u-weight:1)
∨ [age = bet30_59] & [sex = fem] & [bone = yes] & [lung = no] & [pleura = no] & [neck = no] & [axillar = no] & [mediast = no] & [abdom = no] (t-weight:3, u-weight:3)
∨ [age = ge60] & [sex = male] & [hyst_type = epid] & [neck = no] & [mediast = no] (t-weight:3, u-weight:2)

∨ [age = bet30_59] & [sex = male] & [lung = yes] & [perit = no] & [liver = yes] & [mediast = yes] (t–weight:3, u–weight:2)
∨ [age = ge60] & [hyst_type = epid] & [bone = no] & [neck = no] & [abdom = no] (t–weight:3, u–weight:2)
∨ [lung = yes] & [liver = yes] & [abdom = no] (t–weight:3, u–weight:1)
∨ [deg_diff = fair ∨ poor] & [bone = yes] & [pleura = no] & [supraclav = yes] & [axillar = no] (t–weight:2, u–weight:2)
∨ [sex = male] & [deg_diff = poor] & [bone = no] & [lung = yes] & [pleura = no] & [liver = no] & [neck = no] & [supraclav = no] (t–weight:2, u–weight:2)
∨ [age = ge60] & [bone = yes] & [pleura = yes] & [abdom = no] (t–weight:2, u–weight:1)
∨ [sex = fem] & [deg_diff = poor] & [lung = no] & [liver = no] & [neck = no] & [supraclav = yes] & [axillar = no] (t–weight:2, u–weight:1)
∨ [age = bet30_59] & [sex = male] & [hyst_type = adeno] & [lung = yes] & [neck = no] & [supraclav = yes] (t–weight:2, u–weight:1)
∨ [age = bet30_59] & [sex = male] & [deg_diff = well] & [bone = no] & [lung = no] & [pleura = no] & [neck = yes] (t–weight:1, u–weight:1)
∨ [age = ge60] & [perit = yes] & [supraclav = yes] (t–weight:1, u–weight:1)
∨ [liver = yes] & [neck = yes] & [axillar = no] (t–weight:1, u–weight:1)

[tumor_location = ovary] ⇐
[age = bet30_59] & [sex = fem] & [bone = no] & [lung = no] & [brain = no] & [neck = no] & [supraclav = no] & [axillar = no] & [abdom = no] (t–weight:19, u–weight:19)
∨ [age = bet30_59] & [lung = no] & [perit = yes] & [liver = no] & [mediast = no] & [abdom = yes] (t–weight:3, u–weight:3)
∨ [age = ge60] & [sex = fem] & [hyst_type = adeno] & [deg_diff = well ∨ poor] & [perit = yes] & [liver = no] & [abdom = no] (t–weight:2, u–weight:2)
∨ [skin = yes] & [mediast = no] & [abdom = yes] (t–weight:1, u–weight:1)
∨ [age = bet30_59] & [pleura = yes] & [perit = no] & [mediast = no] & [abdom = yes] (t–weight:1, u–weight:1)
∨ [age = ge60] & [sex = fem] & [bone = no] & [perit = no] & [liver = no] & [mediast = no] & [abdom = yes] (t–weight:1, u–weight:1)
∨ [age = ge60] & [sex = fem] & [lung = yes] & [pleura = no] & [perit = yes] (t–weight:1, u–weight:1)
∨ [perit = yes] & [supraclav = yes] & [mediast = yes] & [abdom = yes] (t–weight:1, u–weight:1)

[tumor_location = pancreas] ⇐
[age = ge60] & [sex = fem] & [lung = no] & [pleura = no] & [perit = yes] & [supraclav = no] (t–weight:5, u–weight:4)
∨ [age = bet30_59] & [sex = male] & [bone = no] & [pleura = no] & [liver = yes] & [supraclav = no] & [mediast = no] & [abdom = no] (t–weight:4, u–weight:4)
∨ [sex = fem] & [hyst_type = adeno] & [deg_diff = well ∨ poor] & [lung = no] & [liver = yes] & [skin = no] & [mediast = no] & [abdom = yes] (t–weight:4, u–weight:3)
∨ [age = bet30_59] & [bone = no] & [pleura = yes] & [perit = no] & [brain = no]

& [neck = no] & [supraclav = no] & [abdom = yes] (t-weight:3, u-weight:3)
∨ [age = bet30_59] & [sex = fem] & [perit = no] & [liver = yes] & [mediast = no] & [abdom = yes] (t-weight:3, u-weight:2)
∨ [age = lt30..bet30_59] & [sex = male] & [pleura = yes] & [perit = yes] & [supraclav = no] & [mediast = no] (t-weight:2, u-weight:2)
∨ [age = ge60] & [sex = male] & [lung = no] & [pleura = no] & [liver = yes] & [mediast = no] & [abdom = yes] (t-weight:2, u-weight:2)
∨ [age = ge60] & [sex = fem] & [pleura = no] & [liver = yes] & [supraclav = no] & [abdom = no] (t-weight:2, u-weight:1)
∨ [lung = yes] & [perit = yes] & [supraclav = no] & [mediast = no] & [abdom = no] (t-weight:1, u-weight:1)
∨ [sex = male] & [bone = no] & [lung = yes] & [perit = yes] & [mediast = yes] (t-weight:1, u-weight:1)
∨ [age = ge60] & [bone = yes] & [pleura = yes] & [mediast = yes] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = fem] & [lung = yes] & [pleura = no] & [liver = yes] & [brain = no] & [mediast = yes] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = male] & [lung = yes] & [pleura = no] & [perit = no] & [liver = yes] & [mediast = no] (t-weight:1, u-weight:1)

[tumor_location = prostate] ⇐
[age = bet30_59..ge60] & [sex = male] & [hyst_type = adeno] & [bone = yes] & [lung = no] & [liver = no] & [skin = no] & [neck = no] & [mediast = no] & [abdom = no] (t-weight:5, u-weight:5)
∨ [sex = male] & [hyst_type = adeno] & [lung = no] & [neck = yes] & [supraclav = yes] & [abdom = no] (t-weight:2, u-weight:2)
∨ [age = ge60] & [sex = male] & [hyst_type = adeno] & [deg_diff = well] & [bone = no] & [perit = no] & [abdom = yes] (t-weight:2, u-weight:1)
∨ [age = ge60] & [sex = male] & [hyst_type = adeno] & [lung = no] & [perit = no] & [mediast = no] & [abdom = yes] (t-weight:2, u-weight:1)

[tumor_location = rectum] ⇐
[sex = male] & [hyst_type = adeno] & [deg_diff = well] & [lung = no] & [perit = no] & [liver = yes] & [abdom = no] (t-weight:3, u-weight:3)
∨ [sex = fem] & [brain = yes] & [mediast = no] & [abdom = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = fem] & [deg_diff = well] & [bone = no] & [lung = yes] & [abdom = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [bone = no] & [pleura = yes] & [liver = yes] & [mediast = no] (t-weight:1, u-weight:1)

[tumor_location = sal_glands] ⇐
[hyst_type = epid] & [deg_diff = well] & [bone = yes] & [skin = no] & [neck = yes] (t-weight:1, u-weight:1)
∨ [hyst_type = epid] & [pleura = yes] & [perit = yes] & [supraclav = yes] (t-weight:1, u-weight:1)

[tumor_location = stomach] ⇐
[age = lt30..bet30_59] & [sex = fem] & [deg_diff = well ∨ poor] & [lung = no] & [perit = yes] & [liver = no] & [skin = no] & [mediast = no] (t-weight:5, u-weight:4)
∨ [age = bet30_59] & [sex = male] & [bone = no] & [lung = no] & [pleura = no] & [perit = no] & [neck = no] & [abdom = yes] (t-weight:4, u-weight:4)
∨ [sex = fem] & [lung = no] & [liver = no] & [neck = no] & [supraclav = yes] & [axillar = no] & [mediast = no] (t-weight:4, u-weight:2)
∨ [age = bet30_59] & [bone = no] & [lung = no] & [liver = no] & [neck = no] & [supraclav = yes] & [mediast = no] (t-weight:4, u-weight:2)
∨ [age = ge60] & [sex = fem] & [lung = no] & [perit = no] & [liver = yes] & [mediast = no] & [abdom = yes] (t-weight:3, u-weight:3)
∨ [age = bet30_59] & [sex = male] & [hyst_type = adeno] & [bone = yes] & [lung = no] & [skin = no] & [neck = no] & [mediast = no] & [abdom = no] (t-weight:2, u-weight:2)
∨ [age = lt30] & [perit = yes] & [supraclav = yes] (t-weight:2, u-weight:2)
∨ [age = ge60] & [sex = male] & [pleura = no] & [perit = yes] & [liver = no] (t-weight:2, u-weight:2)
∨ [age = bet30_59] & [lung = no] & [pleura = no] & [perit = yes] & [liver = yes] & [skin = no] & [mediast = no] & [abdom = yes] (t-weight:2, u-weight:2)
∨ [lung = yes] & [axillar = yes] & [abdom = yes] (t-weight:2, u-weight:2)
∨ [age = bet30_59] & [sex = fem] & [pleura = no] & [perit = yes] & [liver = no] & [skin = no] & [abdom = no] (t-weight:2, u-weight:2)
∨ [sex = male] & [hyst_type = adeno] & [deg_diff = well] & [bone = no] & [perit = no] & [liver = no] & [mediast = yes] (t-weight:2, u-weight:1)
∨ [age = bet30_59] & [lung = yes] & [pleura = yes] & [liver = no] & [abdom = yes] (t-weight:2, u-weight:1)
∨ [age = ge60] & [sex = fem] & [deg_diff = well] & [lung = no] & [perit = no] & [liver = yes] (t-weight:1, u-weight:1)
∨ [age = lt30] & [bone = no] & [pleura = yes] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = male] & [lung = yes] & [supraclav = yes] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = male] & [lung = yes] & [perit = yes] (t-weight:1, u-weight:1)
∨ [age = bet30_59] & [sex = male] & [hyst_type = adeno] & [deg_diff = well] & [lung = no] & [perit = no] & [liver = yes] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = male] & [hyst_type = adeno] & [perit = yes] & [liver = yes] & [abdom = no] (t-weight:1, u-weight:1)
∨ [age = ge60] & [sex = male] & [deg_diff = poor] & [liver = yes] & [mediast = no] & [abdom = no] (t-weight:1, u-weight:1)

[tumor_location = testis] ⇐
[age = bet30_59] & [sex = male] & [lung = yes] & [pleura = yes] & [mediast = no] (t-weight:1, u-weight:1)

[tumor_location = thyroid] ⇐
[age = ge60] & [sex = fem] & [bone = yes] & [pleura = no] & [supraclav = no] (t-weight:4, u-weight:4)

∨ [age = lt30..bet30_59] & [bone = yes] & [lung = yes] & [perit = no] & [mediast = no] (t–weight:3, u–weight:3)
∨ [age = bet30_59] & [sex = fem] & [bone = yes] & [perit = no] & [brain = no] & [axillar = no] & [mediast = yes] (t–weight:3, u–weight:3)
∨ [sex = male] & [deg_diff = fair] & [neck = yes] & [axillar = yes] (t–weight:1, u–weight:1)
∨ [age = bet30_59] & [sex = male] & [hyst_type = adeno] & [deg_diff = well] & [bone = yes] & [skin = no] & [mediast = no] & [abdom = no] (t–weight:1, u–weight:1)
∨ [age = ge60] & [pleura = yes] & [supraclav = no] & [mediast = yes] & [abdom = no] (t–weight:1, u–weight:1)
∨ [sex = male] & [bone = yes] & [neck = yes] & [supraclav = no] (t–weight:1, u–weight:1)

[tumor_location = vagina] ⇐
[hyst_type = epid] & [lung = yes] & [perit = yes] & [mediast = no] (t–weight:1, u–weight:1)

[tumor_location = anus] ⇐
false (t–weight:0, u–weight:0)

BIBLIOGRAPHIC DATA SHEET	1. Report No. UIUCDCS-R-86-1260	2.	3. Recipient's Accession No.
4. Title and Subtitle The AQ15 Inductive Learning System: An Overview and Experiments			5. Report Date July 1986 6.
7. Author(s) R. S. Michalski, I. Mozetic, J. Hong, N. Lavrac			8. Performing Organization Rept. No.
9. Performing Organization Name and Address Department of Computer Science University of Illinois Urbana, IL 61801			10. Project/Task/Work Unit No. 11. Contract/Grant No. NSF DCR 86-06801 N00014-82-K-0186 N00014-K-85-0878
12. Sponsoring Organization Name and Address National Science Foundation, Washington, DC Office of Naval Research, Arlington, VA Defense Advanced Research Project Agency, Arlington, VA Slovene Research Council,			13. Type of Report & Period Covered 14.
15. Supplementary Notes			

16. Abstracts

AQ15 is a multi-purpose inductive learning system that uses logic-based, comprehensible knowledge representation. It is able to incrementally learn attributional disjunctive concepts from data that may contain erroneous or inconsistent examples, and can perform constructive induction. The latter means that the program uses background knowledge to generate new attributes not present in the input data, and, if they pass a relevance test, employs them in the learning process. In an experimental application to three medical domains, the program learned decision rules that performed at the level of accuracy of human experts. A surprising and potentially significant result is the demonstration that by applying the proposed method of rule reduction and flexible matching (TRUNC), one may drastically decrease the complexity of the knowledge base without affecting its performance accuracy.

17. Key Words and Document Analysis. 17a. Descriptors

Knowledge Acquisition
Machine Learning
Inductive Inference
Applications (Medicine)

17b. Identifiers/Open-Ended Terms

17c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report) UNCLASSIFIED	21. No. of Pages 32
	20. Security Class (This Page UNCLASSIFIED	22. Price