consists of a number of conjunctive clauses for each class of the concept. An example of the concept is classified by any rule that it matches. DNF rules are considered to be an intuitive form for expressing many concepts, and are used widely in knowledge-based systems.

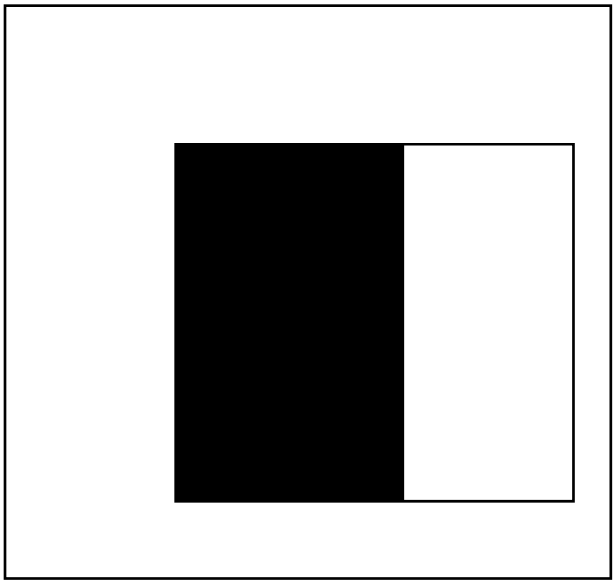
1.1

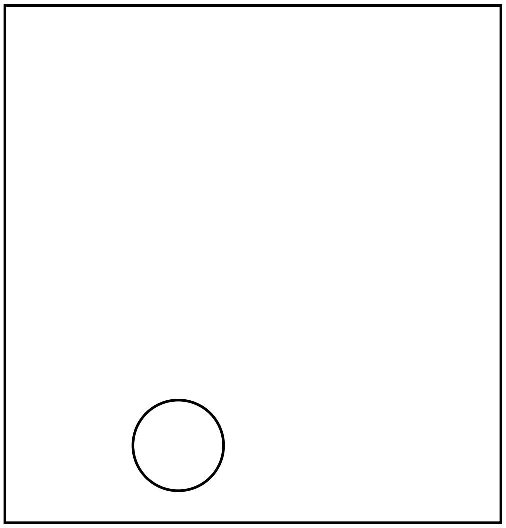
ID3 trees are different again. Like CART trees they specify only a single attribute at each node. However, if it is symbolic, the node has a branch for each possible value of the attribute, so the tree need not be binary. Ripple-down rules have the advantage that the rule at each node may be as specific (many terms) or as general (few terms) as is desired. Thus, a ripple-down rule tree that distinguishes two classes of a concept may have a single node where an ID3 or CART tree requires several.

The use of a tree-like structure has another advantage over the use of DNF rules. Often two classes will differ in only a few respects. A learning algorithm may have to use several of the classes similarities to distinguish them from other classes. In a tree-like structure this may result in the two classes being distinguished by a node some way down the tree. A rule-based concept description, however, would duplicate large parts of rules in order to first differentiate examples from the other classes. The final few terms of these rules would distinguish the examples of the two classes in question.

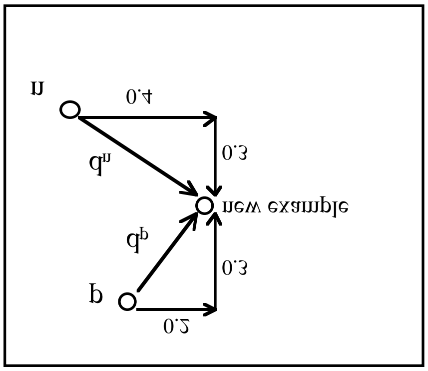
1.3

Utgoff (1989) applied a similar, though more complicated, technique to the ID3 decision tree learner (Quinlan, 1986), allowing it to update the existing decision tree as new training examples were presented. Utgoff s algorithm, ID5R, restructures parts of the decision tree as new examples filter down and are misclassified. ID5R keeps track of unused attributes and values of the examples that generated a leaf node, and uses these attributes to split the leaf node when a new example with a different class arrives. The subtree is also reordered, if necessary, to keep the attribute with the greatest information gain at the root. This is a complicated operation, and may have to be performed recursively back up to the root of the tree. However, the algorithm guarantees to produce the same tree as ID3, given the same examples in any order. The ripple-down rule approach is much simpler because exceptions are only activated in the context of their parent rule. The addition of exceptions to existing rules does not affect those higher up in the concept description.





[redacted]


p
n
d_p
d_n
0.2
0.3
0.3
0.4
new example

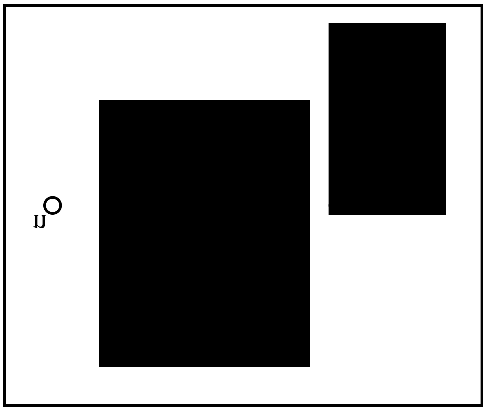
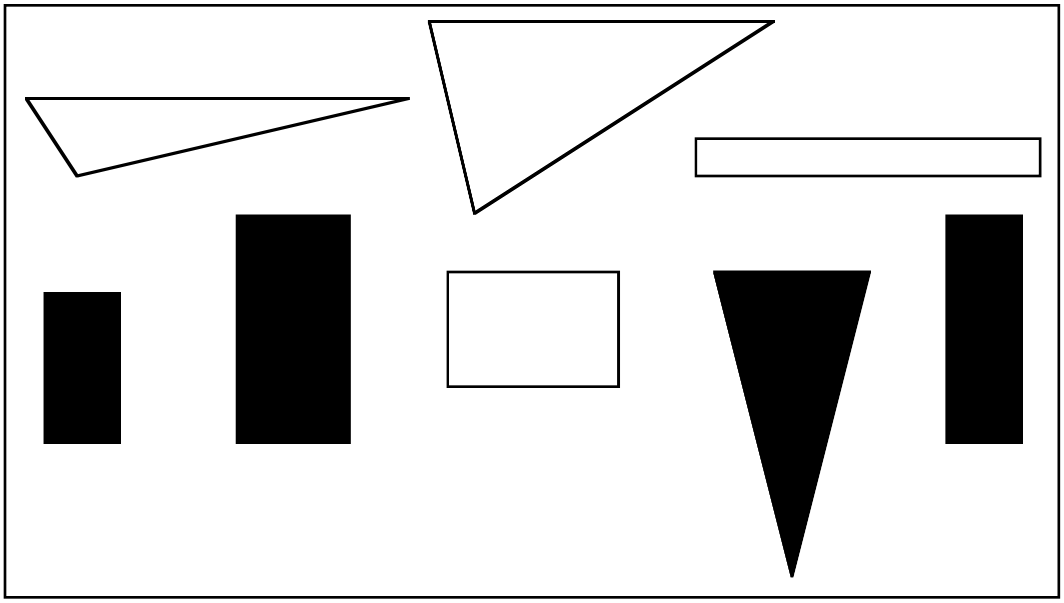


Figure 3.1 Eight blocks representing the standing up concept.



by such things as different scales. For example, attribute x is less than 10 times attribute y might discover the is bigger than relationship in the different units blocks data. However, the infinite number of possible relations, and the difficulty of selecting a small set that is meaningful in any particular context, make this impractical.

```
successor(word#0,word#1).          after(word#0,word#1).
successor(word#1,word#2).          after(word#0,word#2).
successor(word#2,word#3).          after(word#0,word#3).

before(word#1,word#0).                     near1(word#0,word#1).
before(word#2,word#0).                     near1(word#1,word#0).
before(word#3,word#0).                     near1(word#1,word#2).
```

Figure 3.9 Extensional definitions for relations used in the text categorisation example.

after(word#2,word#4)

after(word#2,word#4)

terms:	append(X,Y,Z) f(g(X),h,Y)
constants:	pi() jpl7j74()

Figure 3.16 Examples of valid terms and constants.

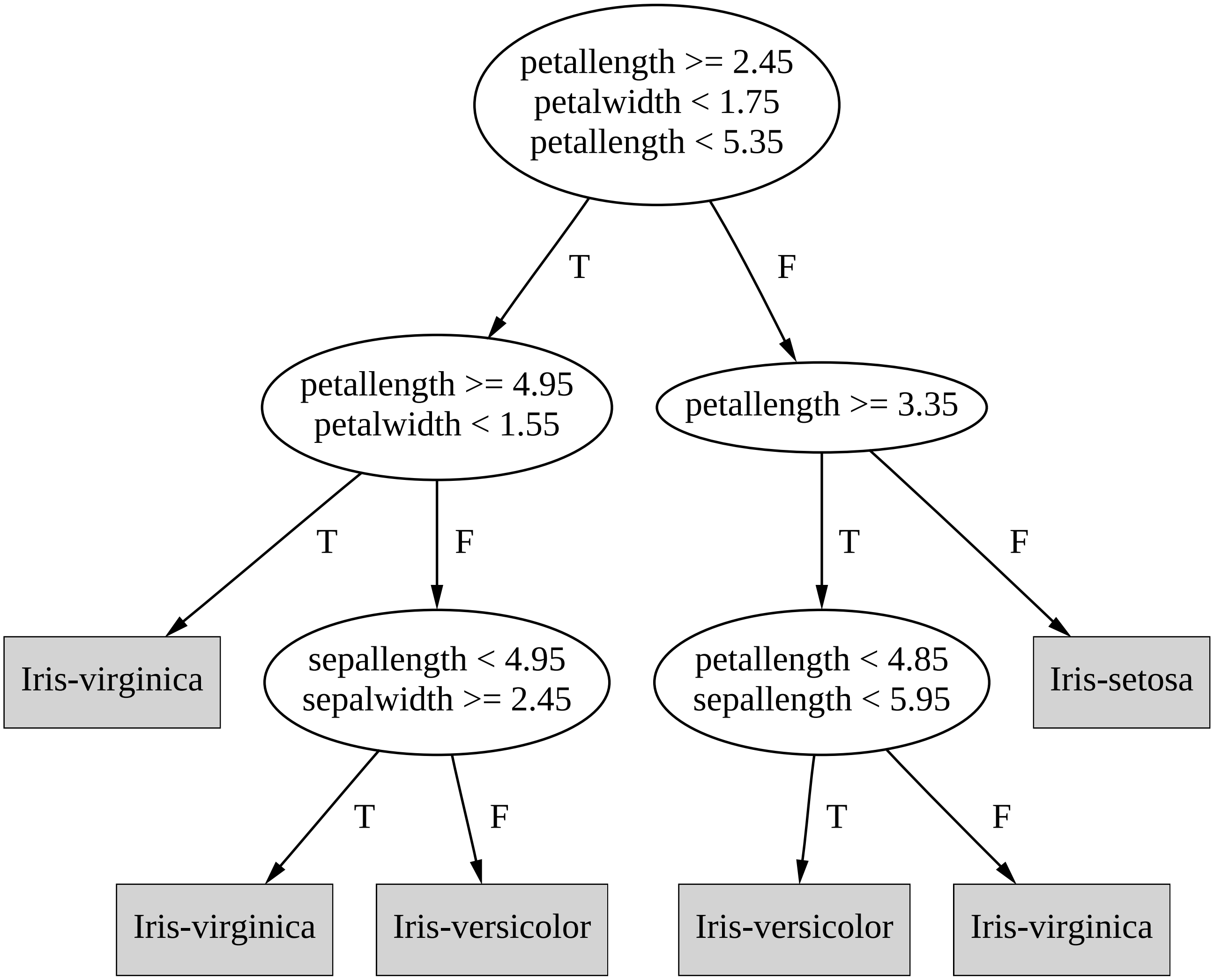
which is commonly abbreviated to

of the order of literals, computational efficiency often is not. For example, the two
clauses

Parameter	Description
—i arff file	Specifies the training data file (required).
—b	Use the binomial approximation instead of the hypergeometric function.
—c n	

combinations of variables and constants might prove useful. For example, a term such as

CreationDate:
Creator: dot version 95 (4-10-95)
Title: HenryDottyOutput


petallength >= 2.45
petalwidth < 1.75
petallength < 5.35
T
F
petallength >= 4.95
petalwidth < 1.55
petallength >= 3.35
T
F
T
F
Iris-virginica
sepallength < 4.95
sepalwidth >= 2.45
petallength < 4.85
sepallength < 5.95
Iris-setosa
T
F
T
F
Iris-virginica
Iris-versicolor
Iris-versicolor
Iris-virginica

example would match the term. A comparison of the two missing value methods is presented in Chapter˚5.

5.

in the manner described by Holte. ABE evaluated the ripple-down rules produced by each heuristic, using the missing values fail method.

identified in the implementation, HENRY should prove successful at combining the duties relational and propositional learning.