ATTRIBUTE	POSSIBLE VALUES
age	old, midlife, new
competition	no, yes
type	software, hardware

AGE	COMPETITION	TYPE	PROFIT
old	yes	swr	down
old	no	swr	down
old	no	hwr	down
mid	yes	swr	down
mid	yes	hwr	down
mid	no	hwr	up
mid	no	swr	up
new	yes	swr	up
new	no	hwr	up
new	no	swr	up

Gain(X,T) = Information(X) - Information(X, T)

Information(X) = $p(1)*\log_2(p(1)) + \ldots. p(n)*\log_2(p(n))$
where pn the the fraction of the data set for which the value of the decision variable is the nth value. For instance p(down) = 5/10

Information(X,T) = sum over values of T((|Ti| / |T|)*Information(T))
So for Age: sum(old,mid,new) (3/10*info(the 3 old) + 4/10*info(the 4 mid) …

Algorithm

Compute Information among current set of examples
Compute Information Gain from splitting on all available features
Add node to tree for splitting on the max gain.
Split examples into subsets according to this new node in tree
Return to step 1 with each subset