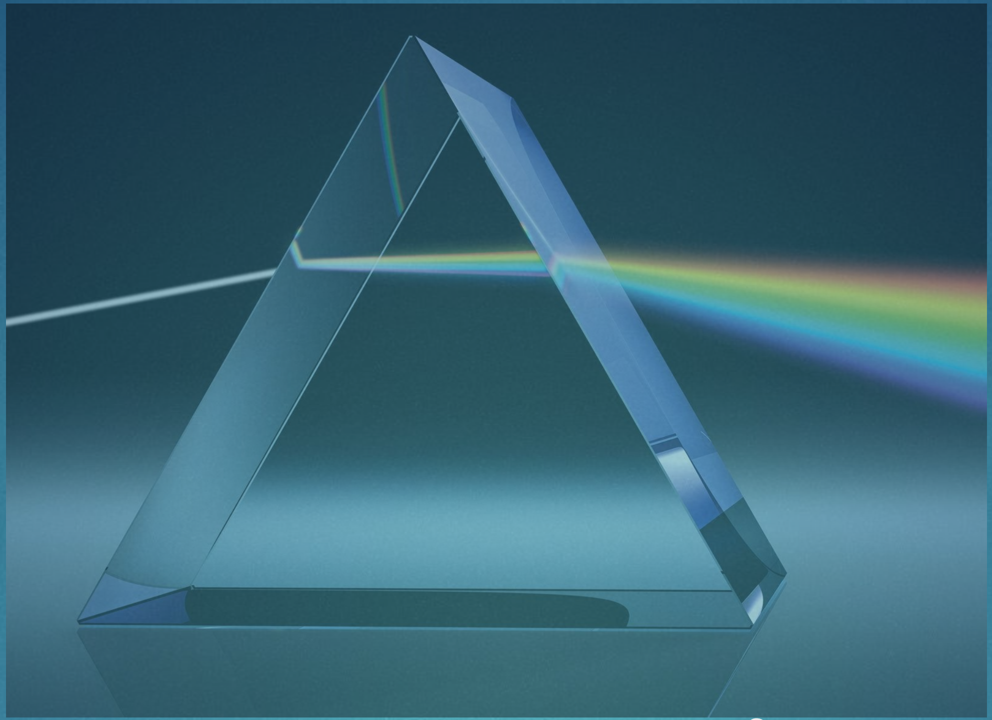


# Cendrowska's PRISM

-ANKIT CHOUDHARY, CST, UG Batch of '17
-MAYANK KUMAR, CS, UG Batch of '17

# GENERAL INFORMATION





- A simple covering algorithm developed by Cendrowksa in 1987.
- Uses Rule Sets instead of decision tree
- General strategy: for each class find rule set that covers all instances in it (excluding instances not in the class). This approach is called a *covering* approach because at each stage a rule is identified that covers some of the instances.
- Uses *separate-and-conquer* algorithm.

# ALGORITHM



For each class C
    Initialize E to the training set
        While E contains instances in class C
            Create a rule R with an empty left-hand side that predicts class C
            Until R is perfect (or there are no more attributes to use) do
                For each attribute A not mentioned in R, and each value v,
                    Consider adding the condition A = v to the left-hand side of R
                    Select A and v to maximize the accuracy p/t
                        (break ties by choosing the condition with the largest p)
                Add A = v to R
            Remove the instances covered by R from E

Learn One Rule

# The Algorithm Simplified




8/14/2015

If the training set contains instances of more than one classification, then for each classification, $\delta_n$, in turn:

- Step 1: calculate the probability of occurrence, $p(\delta_n | \alpha_x)$, of the classification 6n

for each attribute-value pair $\alpha_x$,

- Step 2: select the $\alpha_x$ for which $p(\delta_n | \alpha_x)$ is a maximum and create a subset of the

training set comprising all the instances which contain the selected $\alpha_x$,

- Step 3: repeat Steps 1 and 2 for this subset until it contains only instances of class

$\delta_n$. The induced rule is a conjunction of all the attribute-value pairs used

in creating the homogeneous subset.

- Step 4: remove all instances covered by this rule from the training set,
- Step 5: repeat Steps 1-4 until all instances of class $\delta_n$ have been removed.

When the rules for one classification have been induced, the training set is restored

to its initial state and the algorithm is applied again to induce a set of rules covering

the next classification.

# Accuracy





$$\text{Accuracy} = p/t$$

*t* : Number of instances covered by rule

*p* : Number of instances covered by rule that belong to the positive class

# Heuristics Used

Used when $p/t$ value is same for some attributes



*Opting for generality I:-*

- *If Accuracy of two attributes are same, then choose the one with higher p value*
- *Why? Since it is a possibility that the chosen attribute maybe irrelevant, but if an attribute with higher frequency is chosen then such possibilities decrease.*

# Heuristics Used





*Opting for Generality II*

- When both the Accuracy offered by two or more attribute-value pairs is the same and the numbers of instances referencing them is the same, PRISM selects the first.
- This is the only time that the order of input of the attributes affects the induction process, but in these cases it is still possible for an irrelevant attribute value pair to be selected.

# Proof of Generality II



Ankit, Mayank, IIESTS
8/14/2015

To illustrate how PRISM copes with this situation, suppose there are four attributes, a, b, c and d, each having three possible values, 1, 2 and 3, and the rules to be induced for class $\delta_1$ are:

Rule 1: c1 ^ dt → $\delta_1$,

Rule 2: *c2* ^ *d2*→ $\delta_1$,

Rule 3: c3 ^ d3 → $\delta_1$.

Thus, attributes a and b are irrelevant to $\delta_1$, whereas all values of attributes c and d are equally relevant. If the training set is complete, then $p(\delta_1 | \alpha_x)$ is the same for all $\alpha_x$ and PRISM selects $\alpha_1$. The subset containing only instances which have value 1 for attribute b also presents the same problem--$p(\delta_1 | \alpha_x)$ is equal for all $b_x$, so $b_1$ is selected, and so on.

# Result

The result is the following set of rules:

Rule 1: $a1 \wedge b1 \wedge c1 \wedge d1 \rightarrow \delta_1$,

Rule 2: $a2 \wedge b1 \wedge c1 \wedge d1 \rightarrow \delta_1$,

Rule 3: $a3 \wedge b1 \wedge c1 \wedge d1 \rightarrow \delta_1$,

Rule 4: $b2 \wedge a1 \wedge c1 \wedge d1 \rightarrow \delta_1$,

Rule 5: $b3 \wedge a1 \wedge c1 \wedge d1 \rightarrow \delta_1$.

Rule 6: $c2 \wedge d2 \rightarrow \delta_1$,

Rule 7: $c3 \wedge d3 \rightarrow \delta_1$.

Rule 8: $c1 \wedge d1 \rightarrow \delta_1$.

Since Rule 8 is the generalisation of Rules 1-5, Rule 8, 6 and 7 are chosen. Hence the problem is solved.

# Example : Contact lens problem





| Age | SpectaclePrescription | Astigmatism | TearProductionRate | RecommendedLenses |
|---|---|---|---|---|
| Young, | Myope, | No, | Reduced, | None |
| Young, | Myope, | No, | Normal, | Soft |
| Young, | Myope, | Yes, | Reduced, | None |
| Young, | Myope, | Yes, | Normal, | Hard |
| Young, | Hypermetrope, | No, | Reduced, | None |
| Young, | Hypermetrope, | No, | Normal, | Soft |
| Young, | Hypermetrope, | Yes, | Reduced, | None |
| Young, | Hypermetrope, | Yes, | Normal, | hard |
| Pre-presbyopic, | Myope, | No, | Reduced, | None |
| Pre-presbyopic, | Myope, | No, | Normal, | Soft |
| Pre-presbyopic, | Myope, | Yes, | Reduced, | None |
| Pre-presbyopic, | Myope, | Yes, | Normal, | Hard |
| Pre-presbyopic, | Hypermetrope, | No, | Reduced, | None |
| Pre-presbyopic, | Hypermetrope, | No, | Normal, | Soft |
| Pre-presbyopic, | Hypermetrope, | Yes, | Reduced, | None |
| Pre-presbyopic, | Hypermetrope, | Yes, | Normal, | None |
| Presbyopic, | Myope, | No, | Reduced, | None |
| Presbyopic, | Myope, | No, | Normal, | None |
| Presbyopic, | Myope, | Yes, | Reduced, | None |
| Presbyopic, | Myope, | Yes, | Normal, | Hard |
| Presbyopic, | Hypermetrope, | No, | Reduced, | None |
| Presbyopic, | Hypermetrope, | No, | Normal, | Soft |
| Presbyopic, | Hypermetrope, | Yes, | Reduced, | None |
| Presbyopic, | Hypermetrope, | Yes, | Normal, | None |

# Class→ Recommendation=Hard





Step1: - Calculate Accuracy

| | p/t |
|---|---|
| Age = Young | 2/8 |
| Age = Pre-presbyopic | 1/8 |
| Age = Presbyopic | 1/8 |
| Spectacle prescription = Myope | 3/12 |
| Spectacle prescription = Hypermetrope | 1/12 |
| Astigmatism = no | 0/12 |
| Astigmatism = yes | 4/12 <== tie |
| Tear production rate = Reduced | 0/12 |
| Tear production rate = Normal | 4/12 <== tie |

Use Heuristic II, Select the first One

# Step 2: Instances covered by Astigmatism=Yes



Rule till now, if astigmatism=yes, then recommendation=hard

| Age | Spectacle Prescription | Astigmatism | Tear production Rate | Recommended lenses |
|---|---|---|---|---|
| Young | Myope | Yes | Reduced | None |
| Young | Myope | Yes | Normal | Hard |
| Young | Hypermetrope | Yes | Reduced | None |
| Young | Hypermetrope | Yes | Normal | hard |
| Pre-presbyopic | Myope | Yes | Reduced | None |
| Pre-presbyopic | Myope | Yes | Normal | Hard |
| Pre-presbyopic | Hypermetrope | Yes | Reduced | None |
| Pre-presbyopic | Hypermetrope | Yes | Normal | None |
| Presbyopic | Myope | Yes | Reduced | None |
| Presbyopic | Myope | Yes | Normal | Hard |
| Presbyopic | Hypermetrope | Yes | Reduced | None |
| Presbyopic | Hypermetrope | Yes | Normal | None |

# Check for attribute of next best accuracy

Rule till now, if astigmatism=yes, then recommendation=hard

```
                                                p/t
                                                ---
Age = Young                                     2/4
Age = Pre-presbyopic                            1/4
Age = Presbyopic                                1/4
Spectacle prescription = Myope                  3/6
Spectacle prescription = Hypermetrope           1/6
Tear production rate = Reduced                  0/6
Tear production rate = Normal                   4/6 <== winner
```

# Again check for remaining instances





Rule till now, If astigmatism = yes and tear production rate = normal
then recommendation = hard

```
                                         p/t
                                         ---
Age = Young                              2/2 <== tie
Age = Pre-presbyopic                     1/2
Age = Presbyopic                         1/2
Spectacle prescription = Myope           3/3 <== tie
Spectacle prescription = Hypermetrope    1/3
```

Use Heuristic I, Select the one with higher 'p' value

# Finally Rule 1



Ankit, Mayank, IIESTS


On going further, we find that all the values for recommendation is hard, so the rule 1 is:-

If astigmatism = yes and tear production rate = normal and Spectacle Specification= myopia,

then recommendation = hard

Other rules on next page

# All rules



```
If spectacle prescription = Myope and astigmatic = yes and
   tear production rate = normal
   then recommendation = hard
If tear production rate = reduced
   then recommendation = none
If age = young and astigmatic = no and
   tear production rate = normal
   then recommendation = soft
If age = pre-presbyopic and astigmatic = no and
   tear production rate = normal
   then recommendation = soft
If age = presbyopic and spectacle prescription = Myope  and
   astigmatic = no
   then recommendation = none
If spectacle prescription = Hypermetrope and astigmatic = no  and
   tear production rate = normal
   then recommendation = soft
If age young and astigmatic = yes and tear production rate = normal
   then recommendation = hard
If age = pre-presbyopic and spectacle prescription = Hypermetrope and
   astigmatic = yes
   then recommendation = none
If age = presbyopic and spectacle prescription = Hypermetrope and
   astigmatic = yes
   then recommendation = none
```

# Limitations

- Does not work properly for highly noisy datasets
- PRISM algorithm silent on
  - Order with which classes are explored
  - Order with which attributes are explored
- Standard PRISM also demands that all attributes are added
- Standard PRISM has no support-based pruning

# Induction from Incomplete training sets- Highly noisy





Suppose there are four attributes, a, b, c and d. Attribute a has five possible values (1,2, 3, 4, 5), attributes b and c each have four possible values (1, 2, 3, 4) and attribute d has three possible values (1,2, 3). Thus a complete training set would consist of 5 x 4 x 4 x 3 = 240 instances. Suppose that the rule set governing class $\delta_1$ is

Rule 1: $a4 \wedge d2 \rightarrow \delta_1$,

Rule 2: $c1 \wedge d1 \rightarrow \delta_1$,

Rule 3: $a2 \wedge c4 \wedge d2 \rightarrow \delta_1$,

Rule 4: $a5 \wedge c4 \wedge d2 \rightarrow \delta_1$,

and that the 40 instances are listed in table in next page.

# Table

**TABLE 3**

*Example of incomplete training set*

| *a* | *b* | *c* | *d* | δ | *a* | *b* | *c* | *d* | δ | *a* | *b* | *c* | *d* | δ | *a* | *b* | *c* | *d* | δ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 3 | 3 | 2 | 2 | 1 | 2 | 2 | 2 | 3 | 2 | 1 | 1 | 1 | 4 | 3 | 2 | 2 | 1 |
| 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 3 | 2 | 4 | 1 | 2 | 4 | 4 | 1 | 3 | 2 |
| 1 | 2 | 3 | 1 | 2 | 2 | 2 | 4 | 2 | 1 | 3 | 2 | 4 | 2 | 2 | 4 | 4 | 3 | 1 | 2 |
| 1 | 3 | 1 | 3 | 2 | 2 | 3 | 2 | 1 | 2 | 3 | 3 | 1 | 1 | 1 | 5 | 1 | 1 | 2 | 2 |
| 1 | 3 | 3 | 2 | 2 | 2 | 3 | 3 | 1 | 2 | 3 | 3 | 1 | 2 | 2 | 5 | 1 | 3 | 2 | 2 |
| 1 | 4 | 1 | 3 | 2 | 2 | 3 | 3 | 3 | 2 | 3 | 3 | 2 | 2 | 2 | 5 | 2 | 2 | 2 | 2 |
| 1 | 4 | 4 | 1 | 2 | 2 | 4 | 1 | 3 | 2 | 3 | 4 | 2 | 1 | 2 | 5 | 3 | 1 | 2 | 2 |
| 2 | 1 | 1 | 1 | 1 | 2 | 4 | 2 | 1 | 2 | 4 | 1 | 3 | 2 | 1 | 5 | 3 | 2 | 3 | 2 |
| 2 | 1 | 1 | 3 | 2 | 3 | 1 | 1 | 1 | 1 | 4 | 1 | 4 | 2 | 1 | 5 | 4 | 1 | 3 | 2 |
| 2 | 1 | 2 | 1 | 2 | 3 | 1 | 4 | 3 | 2 | 4 | 2 | 1 | 3 | 2 | 5 | 4 | 4 | 3 | 2 |

# Results by using Prism Algorithm





The set of rules induced by PRISM for the class $\delta_1$ is

Rule A: $a4 \wedge d2 \rightarrow \delta_1$,

Rule B: $a3 \wedge c1 \wedge d1 \rightarrow \delta_1$,

Rule C: $a2 \wedge c4 \rightarrow \delta_1$ ,

Rule D: $b1 \wedge d1 \wedge c1 \rightarrow \delta_1$.

# Problems in such cases



These are known as overfitting

1. Failure to induce a rule

   A rule will not be induced if there are no examples of it in the training set. This applies to all induction programs. Even human beings cannot be expected to induce rules from non-existent information.

   e.g. Rule 4 above:-

   Rule 4: a5 ^ *c4* ^ *d2*→ δ1

# Problems in such cases





2. OVER-GENERALIZATION

An induced rule may be too general if there are no counter-examples to it in the training set. Any attempts to specialize automatically would have unwanted side-effects on rules which were not too general.

e.g. Rule C and Rule 3 above:-

Rule C: $a2 \wedge c4 \rightarrow \delta 1$

Rule 3: $a2 \wedge c4 \wedge d2 \rightarrow \delta 1$,

# Problems in such cases

3. OVER-SPECIALIZATION

- Theoretically, the induction algorithm is based on finding the $\alpha_x$ for which $p(\delta_1 | \alpha_x)$ is a maximum. In practice, for an incomplete training set, the true probability of occurrence p is unknown, and is approximated by the relative frequency, $f(\delta_1 | \alpha_x)$.This approximation of p introduces errors in the estimation of accuracyof each $\alpha_x$, which becomes significant for small training sets, resulting in the selection of an irrelevant attribute-value pair as the best representative of $\delta_1$

e.g. Rule B and Rule D and Rule 2 above:-

Rule B: $a3 \wedge c1 \wedge d1 \rightarrow \delta_1$,

Rule D: $b1 \wedge d1 \wedge c1 \rightarrow \delta_1$,

Rule 2: $c1 \wedge d1 \rightarrow \delta_1$

# Scope Of Improvement





- Pruning using Information gain criterion.
- Stopping at the attributes where support is low
- Modify the evaluation criteria for each rule, e.g. replace p/t with entropy, lift, etc.

# ENTROPY





The entropy of a set of events has been defined as a measure of the ’freedom of choice’ involved in the selection of the event, or the ’uncertainty’ associated with this selection (Edwards, 1964, Goldman, 1968, Shannon and Weaver, 1949). Given a training set, S, if the above assumptions hold, then each instance is classified correctly and uniquely, i.e. there is no uncertainty about the classification. The entropy of S is 0. The entropy of a decision tree or rule set, which fully describes S is also 0, but in most cases the decision tree  is a generalization of S, which implies that some information offered by the training set is redundant.

# Evaluating Entropy



If all that is known about the classifications is their probabilities of occurrence, p(δi ; i = 1, 2, 3), then the entropy of the set of classifications,

$$H = -\sum_i p(\delta_i) \log_2 p(\delta_i)$$

# For the contact lenses problem





$$H = -\, p(\delta_1)\log_2 p(\delta_1) - p(\delta_2)\log_2 p(\delta_2) - p(\delta_3)\log_2 p(\delta_3)$$

Here

$\delta_1$ = Probability of prescribing a hard contact lens

$\delta_2$ = Probability of prescribing a soft contact lens

$\delta_3$ = Probability of prescribing no contact lens

Hence $p(\delta_1) = 4/24$

$p(\delta_2) = 5/24$

$p(\delta_3) = 15/24$

# Continued …





$H = -4/24\log_2 p(4/24) - 5/24\log_2 p(5/24) - 15/24\log_2 p(15/24)$

$H = 0.4308 + 0.4715 + 0.4238$

$H = 1.3261$ bits.

The induction algorithm partitions the training sets into subsets such that entropy reduces maximally and continues doing so ntil entropy achieves 0 .

# Reducing Entropy





If the training set, S, is divided according to the values of some attribute, α, then unless the classification, δ, is completely independent of or, the values will contain some information about δ. The total entropy of the subsets is known as the conditional entropy of S with known α, $H( S \mid \alpha )$. Let $p ( \alpha_x )$ be the probability that attribute α has value x, and let $p(\delta_n \cap \alpha_x)$ be the probability that classification is $\delta_n$ and value of α is x ,

$$H(S|\alpha) = H(\delta_n \cap \alpha_x) - H(\alpha)$$

# Reducing Entropy



where,

$$H(\delta_n \cap \alpha_x) = -\sum_x \sum_n p(\delta_n \cap \alpha_x) \log_2 p(\delta_n \cap \alpha_x)$$

$$H(\alpha) = -\sum_x p(\alpha_x) \log_2 p(\alpha_x)$$

Possible to minimize the entropy of S by dividing it into subsets according to the value of that attribute for which $H( S \mid \alpha )$ is minimum.

Here comes the use of Frequency Table :

# Frequency Table



| No. of instances referencing | $a_1$ | $a_2$ | $a_3$ | Total |
|---|---|---|---|---|
| $\delta_1$ | 2 | 1 | 1 | 4 |
| $\delta_2$ | 2 | 2 | 1 | 5 |
| $\delta_3$ | 4 | 5 | 6 | 15 |
| Total | 8 | 8 | 8 | 24 |

$$H(S \mid a) = H(S \cap a) - H(a)$$

$$= -\sum_x \sum_n p(\delta_n \cap a_x) \log_2 p(\delta_n \cap a_x) + \sum_x p(a_x) \log_2 p(a_x)$$

$$= -3 \times \frac{2}{24} \log_2 \left(\frac{2}{24}\right) - 3 \times \frac{1}{24} \log_2 \left(\frac{1}{24}\right) - \frac{4}{24} \log_2 \left(\frac{4}{24}\right)$$

$$- \frac{5}{24} \log_2 \left(\frac{5}{24}\right) - \frac{6}{24} \log_2 \left(\frac{6}{24}\right) + 3 \times \frac{8}{24} \log_2 \left(\frac{8}{24}\right)$$

$$= \frac{1}{24} (3 \times 8 \log_2 8 - 3 \times 2 \log_2 2 - 2 \times \log_2 1 - 4 \log_2 4$$

$$- 5 \log_2 5 - 6 \log_2 6)$$

$$= 1{\cdot}2867 \text{ bits.}$$

$a \longrightarrow$ AGE

$a_1 \longrightarrow$ Young

$a_2 \longrightarrow$ Pre-Presbyopic

$a_2 \longrightarrow$ Pre-Presbyopic

# Continued …





Similarly :

H( S | b ) : 1.2867

H( S | c ) : 0.9491

H( S | d ) : 0.7773 ***MINIMUM***

b = SPECTACLE PRESCRIPTION

c = ASTIGMATISM

d = TEAR PRODUCTION RATE

Hence entropy S can be decreased maximally if the subsets are based on division over the attribute ‘d’ which is Tear Production Rate .

# Continued …

Ankit, Mayank, IIESTS
8/14/2015

The induction algorithm is continued and hence rep[eated over the two sets of value for $d_1$ and $d_2$ until the all the leafs at a particular level has entropy 0;

For example :

for, $d_1$ i.e. under the condition of normal Tear Production Rate :

$\delta_3$ is always applicable. Mathematically :

Since , $H(S \mid d) = H(\delta_n \cap d_x) - H(d)\,(x = 1,2)$

$$H(\delta_n \cap d_x) = -\sum_x \sum_n p(\delta_n \cap d_x) \log_2 p(\delta_n \cap d_x)$$

$$H(\alpha) = -\sum_x p(d_x) \log_2 p(d_x)$$

# One More Step...



| No. | d1 | d2 | TOTAL |
|---|---|---|---|
| δ1 | 0 | 4 | 4 |
| δ2 | 0 | 5 | 5 |
| δ3 | 12 | 3 | 15 |
| TOTAL | 12 | 12 | 24 |

The Entropy for d1 -> 0

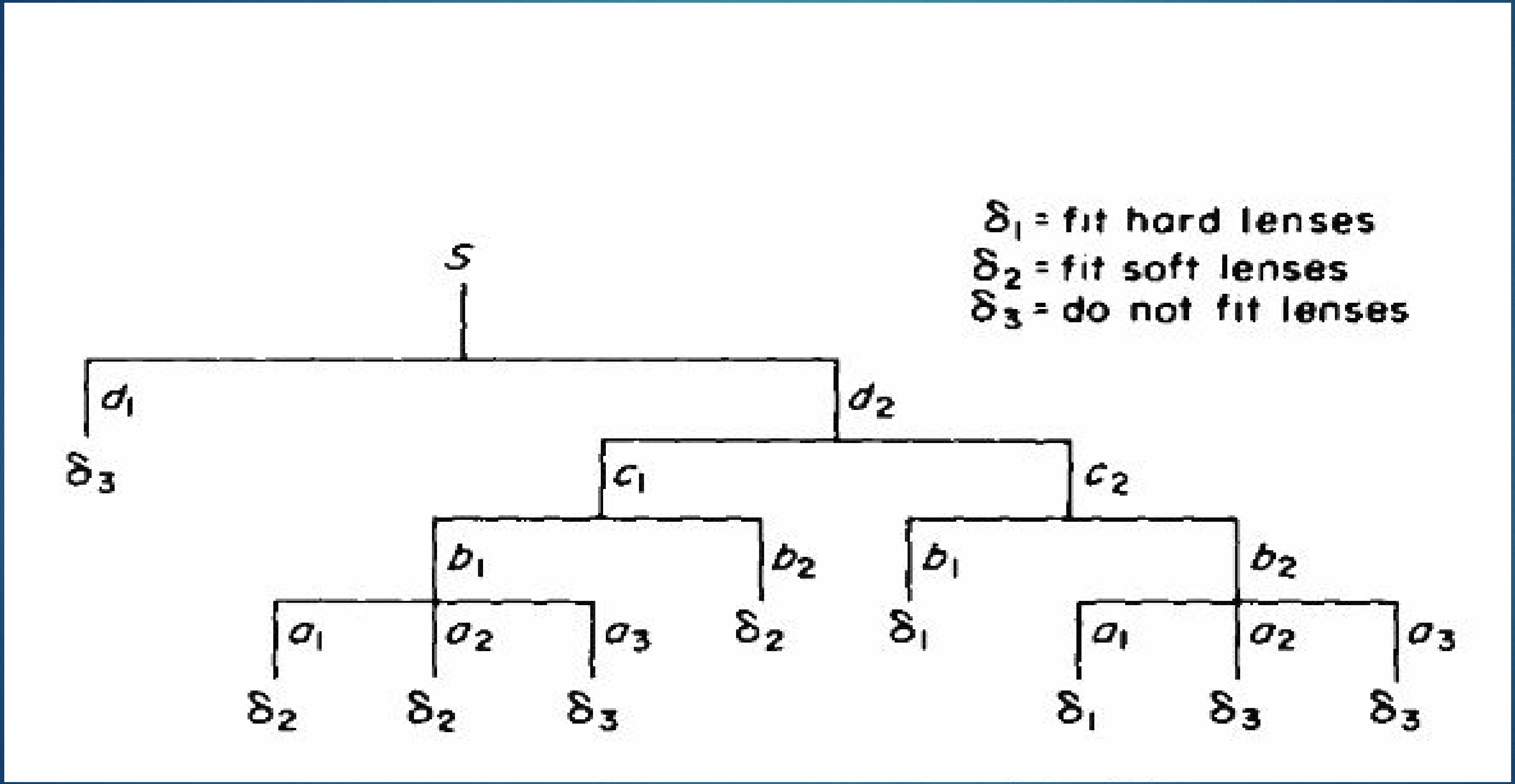
Hence the Branching stops for the branch d1

And Continues for rest…

# Few More Steps...



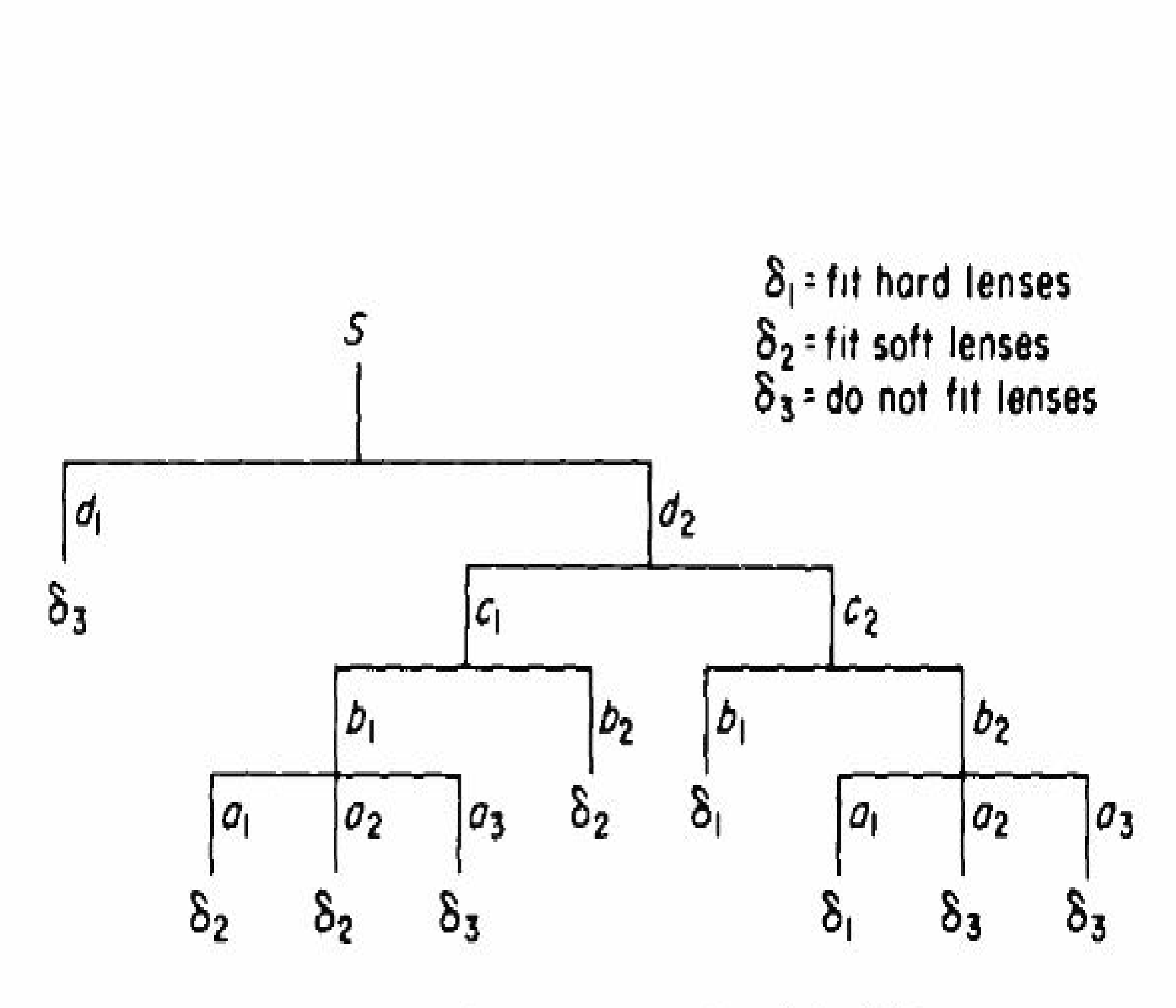




# Rules

The final set of Rules are :

1. $d_1 \rightarrow \boldsymbol{\delta_3}$
2. $d_2 \land c_1 \land b_1 \land a_1 \rightarrow \boldsymbol{\delta_2}$
3. $d_2 \land c_1 \land b_1 \land a_2 \rightarrow \boldsymbol{\delta_2}$
4. $d_2 \land c_1 \land b_1 \land a_3 \rightarrow \boldsymbol{\delta_3}$
5. $d_2 \land c_1 \land b_2 \rightarrow \boldsymbol{\delta_2}$
6. $d_2 \land c_2 \land b_1 \rightarrow \boldsymbol{\delta_1}$
7. $d_2 \land c_2 \land b_2 \land a_1 \rightarrow \boldsymbol{\delta_1}$
8. $d_2 \land c_2 \land b_2 \land a_2 \rightarrow \boldsymbol{\delta_3}$
9. $d_2 \land c_2 \land b_2 \land a_3 \rightarrow \boldsymbol{\delta_3}$



# Comparision

Approach 1 :

The final set of Rules are :

1. $d_1 \rightarrow \boldsymbol{\delta_3}$
2. $d_2 \wedge c_1 \wedge b_1 \wedge a_1 \rightarrow \boldsymbol{\delta_2}$
3. $d_2 \wedge c_1 \wedge b_1 \wedge a_2 \rightarrow \boldsymbol{\delta_2}$
4. $d_2 \wedge c_1 \wedge b_1 \wedge a_3 \rightarrow \boldsymbol{\delta_3}$
5. $d_2 \wedge c_1 \wedge b_2 \rightarrow \boldsymbol{\delta_2}$
6. $d_2 \wedge c_2 \wedge b_1 \rightarrow \boldsymbol{\delta_1}$
7. $d_2 \wedge c_2 \wedge b_2 \wedge a_1 \rightarrow \boldsymbol{\delta_1}$
8. $d_2 \wedge c_2 \wedge b_2 \wedge a_2 \rightarrow \boldsymbol{\delta_3}$
9. $d_2 \wedge c_2 \wedge b_2 \wedge a_3 \rightarrow \boldsymbol{\delta_3}$

Approach 2 :

The final set of Rules are :

1. $d_1 \rightarrow \boldsymbol{\delta_3}$
2. $c_1 \wedge d_2 \wedge a_1 \rightarrow \boldsymbol{\delta_2}$
3. $c_1 \wedge d_2 \wedge a_2 \rightarrow \boldsymbol{\delta_2}$
4. $a_3 \wedge b_1 \wedge c_1 \rightarrow \boldsymbol{\delta_3}$
5. $c_1 \wedge d_2 \wedge b_2 \rightarrow \boldsymbol{\delta_2}$
6. $c_2 \wedge d_2 \wedge b_1 \rightarrow \boldsymbol{\delta_1}$
7. $a_1 \wedge c_2 \wedge d_2 \rightarrow \boldsymbol{\delta_1}$
8. $b_2 \wedge c_2 \wedge a_2 \rightarrow \boldsymbol{\delta_3}$
9. $b_2 \wedge c_2 \wedge a_3 \rightarrow \boldsymbol{\delta_3}$

Although the number of rules in this set is the same in both the approaches, six of the rules have had redundant terms removed. The presbyopic patient with high hypermetropia and astigmatism no longer needs to undergo an examination to be told that she is not suitable for contact lens wear .

# Bibliography





- A,Cendrowksa ,JIMMS, 1987
- http://csee.wvu.edu/~timm/cs591o/old/Rules.html


рахмат
danke
謝謝
ngiyabonga
teşekkür ederim
Баярлалаа
spas
mersi
kia ora
barka
welalin
tack
спасибо
faafetai lava
vinaka
спасибі
blagodaram
dank je
misaotra
matondo
paldies
grazzi
mahalo
tapadh leat
kiitos
dankie
maururu
nanni
nandri
dhanyavad
thank you
gracias
хвала
asante
manana
obrigada
tenki
bedankt
bayarlalaa
grącie
hvala
köszönöm
dankon
ačiū
akun
enkosi
djiere dieuf
tau
mochchakkeram
murakoze
dziękuję
дякую
mamnun
chokrane
chnorakaloutioun
gratias ago
gràcies
sulpáy
go raibh maith agat
sobodi
děkuji
sagolun
sukriya
kop khun krap
taiku
obrigado
mèsi
ありがとう
tanemirt
grazie
arigatô
takk
dakujem
trugarez
didi madloba
najis tuke
terima kasih
rahmet
ευχαριστώ
diolch
dhanyavadagalu
shukriya
mercé
мерси
kam sah hamnida
rahmat
তোমাকে ধন্যবাদ
감사합니다
xièxie
merci