**Course: CS 836**

**Name: Chao Zhang**

**Student#: 200383834**

1. **Reducts**

*• Give a definition of a reduct of an information table.*

In a information table T, if a subset of attributes RAT satisfies the following conditions:

Existence: DEFAT(T)=DEFAT(T)

Sufficiency: DEFR(T)=DEFAT(T)

Minimization: aR (ER-{a}≠EAT)

We call R an attribute reduct of AT

*• Give a definition of a reduct of a consistent decision table.*

Existence: IND(C)IND(D)

Sufficiency: IND(R)IND(D)

Minimization: aR (﹁(IND(R-{a})IND(D)))

*• Give a definition of a reduct of a decision table. Show that your definition is applicable to both consistent and inconsistent decision table.*

In a decision table T with AT=CD, if a subset of condition attributes RC satisfies the following conditions:

Existence: POS(πD|πC) = POS(πD|πC)

Sufficiency: POS(πD|πR) = POS(πD|πC)

Minimization: aR (﹁(POS(πD|πR-{a}) = POS(πD|πC)))

We call R a relative attribute reduct of C.

For a consistent table which is IND(C)IND(D), according to the reduct R, we can get IND(R)IND(D) andaR (﹁(IND(R-{a})IND(D))). It is the same as POS(πD|πR) = POS(πD|πC) and for every single attribute a in R is not superfluous. For an inconsistent table, the definition illustrates the partition induced by R with respect to D in positive region is equal to the partition induced by C with respect to D and for every single attribute a in R is not superfluous to give the same partition in positive region.

*• Give a general definition of a reduct to cover reducts in an information table, a consistent decision table, and a decision table.*

Generally in a table, we can consider two subsets of attributes X,YAT. For these two subsets are arbitrary and may have a non-empty overlap. Then the notion we have will be “a reduct of X with respect to Y”. Then we have following conditions:

Existence: POS(πY|πX) = POS(πY|πX)

Sufficiency: POS(πY|πX) = POS(πY|πR)

Minimization: aR (﹁(POS(πY|πR-{a}) = POS(πY|πC)))

For a information table, these two subsets can be X=Y=AT.

For a decision table, these two subsets can be X is the set of condition attributes and Y is the set of decision attributes.

**2. Probabilistic rough sets**

*• Give a definition of probabilistic rough set approximations. Show that Pawlak rough set approximations are a special case.*

For a subset X of OB, its rough membership function is given by the conditional probability Pr(X|[x])=. For general probabilistic approximations, we use a pair of parameters α,β[0,1] and αβ to make sure that the upper approximation is greater than the lower approximation. So the general probabilistic approximations can be defined:

(α)={XOB | Pr(X|[x])α}, (β)={XOB | Pr(X|[x])>β}

For Pawlak rough set approximations:

(α)={XOB | Pr(X|[x])1}, (β)={XOB | Pr(X|[x])>0} where α=1 and β=0 in the general definition, so it is a special case.

*• Explain* *the motivations for introducing probabilistic rough sets.*

In the standard rough set model proposed by Pawlak, the lower and upper approximations are defined based on the two extreme cases according to the relationships between an equivalence class and a set. The lower approximation requires that the equivalence class is a subset of the set and the upper approximation requires the equivalence class must have a non-empty overlap with the set. A lack of consideration for the degree of their overlap unnecessarily limits the applications of rough sets and it is the main motivation for introducing probabilistic rough sets.

**3. Decision-theoretic rough sets**

*• Describe the Bayesian decision procedure.*

Let Ω={w1,w2…ws} be a finite set of s states. Let A={a1,a2…am} be a finite set of m possible actions. Let Pr(wj|x) be the conditional probability of an object x being in state wj given that the object is described by x. Let λ(ai|wj) denote the loss, or cost, for taking action ai when the state is wj . For an object with description x, suppose action ai is taken. Since Pr(wj|x) is the probability that the true state is wj given x, the expected loss associated with taking action ai is given by: R(ai|x)=. The quantity R(ai|x) is also called the conditional risk. Given a description x, a decision rule is a function τ(x) that specifies which action to take. That is, for every x, τ(x) takes one of the actions, a1,a2…am. The overall risk R is the expected loss associated with a given decision rule. Since R(τ(x)|x) is the conditional risk associated with action τ(x), the overall risk is defined by: R=, where the summation is over the set of all possible descriptions of objects. If τ(x) is chosen so that R(τ(x)|x) is as small as possible for every x, the overall risk R is minimized. Thus, the Bayesian decision procedure can be formally stated as follows: for every x, compute the conditional risk R(ai|x) for i=1,…,m defined by R(ai|x)= and select the action for which the conditional risk is minimum. If more than one action minimizes R(ai|x), a tie-breaking criterion can be used.

*• Derive the probabilistic rough set approximations by using the Bayesian decision procedure*

A decision-theoretic model formulates the construction of rough set approximations as a Bayesian decision problem with a set of two states and a set of three actions. The set of states is given by Ω={X, Xc} indicating that an element is in X and not in X, respectively. Corresponding to the three regions, the set of actions is given by A={aP, aB, aN}, denoting the actions in classifying an object x, namely, deciding xPOS(X), deciding xBND(X), and deciding xNEG(X), respectively.

The expected losses associated with taking different actions for objects in [x] can be expressed as:

R(aP|[x])=Pr(X|[x])λPP+Pr(Xc|[x])λPN,

R(aN|[x])=Pr(X|[x])λNP+Pr(Xc|[x])λNN,

R(aB|[x])=Pr(X|[x])λBP+Pr(Xc|[x])λBN.

The Bayesian decision procedure leads to the following minimum-risk decision rules as follows:

1. If R(aP|[x])R(aB|[x]) and R(aP|[x])R(aN|[x]), decide xPOS(X);
2. If R(aB|[x])R(aP|[x]) and R(aB|[x])R(aN|[x]), decide xBND(X);
3. If R(aN|[x])R(aP|[x]) and R(aN|[x])R(aB|[x]), decide xNEG(X).

Considering a special class of loss functions with λPPλBPλNP and λNN λBNλPN, with this condition and the equation Pr(X|[x])+Pr(Xc|[x])=1, we can get:

1. If Pr(X|[x])≥γ and Pr (X|[x])≥α, decide xPOS(X);

(B) If β≤Pr(X|[x])≤α, decide xBND(X).

(N) If Pr(X|[x])≤β and Pr(X|[x])≤γ, decide xNEG(X);

where

α=(λPN−λBN) / ((λPN−λBN)+(λBP−λPP))

γ=(λPN−λNN) / ((λNP−λ PP)+(λPN−λNN))

β=(λBN−λNN) / ((λBN−λNN)+(λNP−λBP))

Then we obtain the following condition on the loss function:

((λNP−λBP)/(λBN−λNN))((λBP−λPP)/(λBN−λNN))

It implies that 1αγβ0 and then we can get the simplified rules:

(P) If Pr(X|[x])≥α, decide xPOS(X);

(N) If Pr(X|[x])≤β, decide xNEG(X);

(B) If β<Pr(X|[x])<α, decide xBND(X).

So we can get the probabilistic approximations from the relationship between the three regions and approximations:

(α)={XOB | Pr(X|[x])α}, (β)={XOB | Pr(X|[x])>β}.

**4. Describe the main results of naive Bayesian rough sets.**

Naive Bayesian rough set model is a practical method for estimating the conditional probability. First, we perform the logit transformation of the conditional probability:

logit(Pr(X|[x]))=log=log, which is a monotonically increasing transformation of Pr(X|[x]). Then, we apply the Bayes’ theorem: Pr(X|[x])=. Similarly, for Xc we also have:

Pr(Xc|[x])=. Then using above three formulas, we can get logit(Pr(X|[x]))= log O(X|[x])= log= log=

log= log·= log+logO(X).

where O(X|[x]) and O(X) are the a posterior and the a prior odds, and  is the likelihood ratio. A threshold value on the probability can be expressed as another threshold value on logarithm of the likelihood ratio. For the positive region, we have:

Pr(X|[x])αlogloglog·log

loglog+=α’

Similar expressions can be obtained for the negative and boundary regions. The three regions can now be written as:

POS(α,β)(X)={xOB|logα’},

NEG(α,β)(X)={xOB|logβ’},

BND(α,β)(X)={xOB|β’<log<α’}.

where

α’=log+ log

β’=log+ log