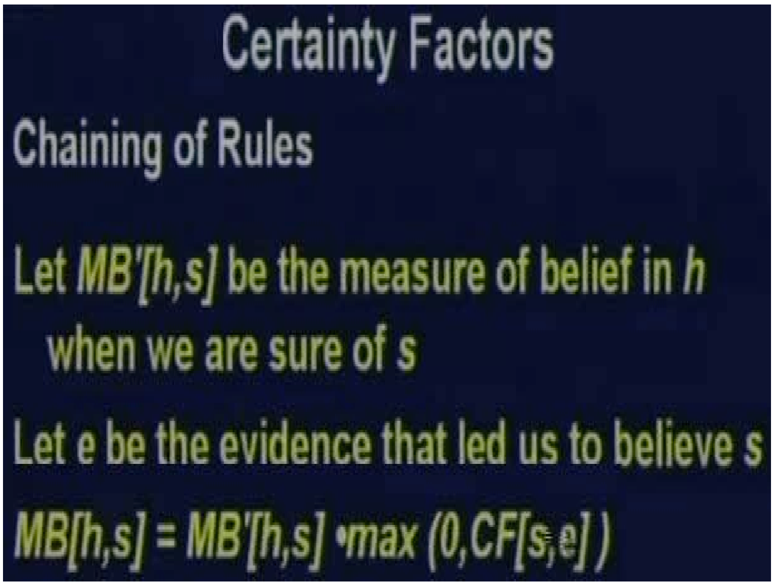
Artificial Intelligence
Prof. Anupam Basu
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur
Lecture # 27
Reasoning with Uncertainty - II

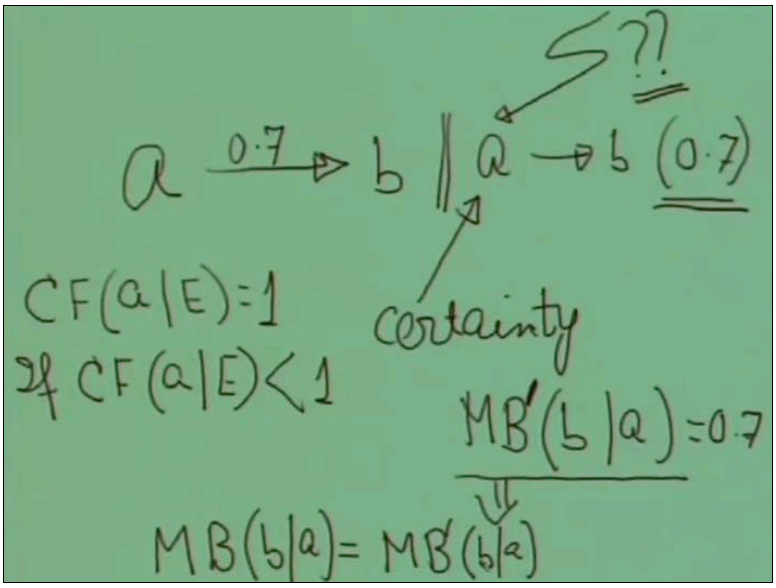
In the last lecture we discussed about certainty factors and briefly introduced certainty factors. In this lecture we will further deal with that. We have seen the different scenarios.

(Refer Slide Time: 01:18)



We will first start with the last point where we left. Certainty factors in the case of rule chaining. Now what was really attempted to be explained here is something like this; I have got some antecedent a pointing to some other antecedent b with some certainty say 0.7. This is equivalent to writing it as a implies b 0.7, what does it mean? It means that if a is known with certainty then I will infer b with confidence or belief of 0.7. So I would say that MB(b by a) is equal to 0.7 in the normal case but we know that this a may not be known with certainty there may be some uncertainty in this a itself. If the certainty factor of a given whatever evidences E be 1 then in that case it is a certainty but suppose if the certainty factor of a given some evidence E is less than 1 then obviously this 0.7 I get here should be modified. Therefore we say it is MB prime. MB prime is the measure of belief in b if a is known with certainty. But a is not known with certainty. Therefore this needs to be modified and I must have some MB(b by a) where this MB prime b by a must be modified.

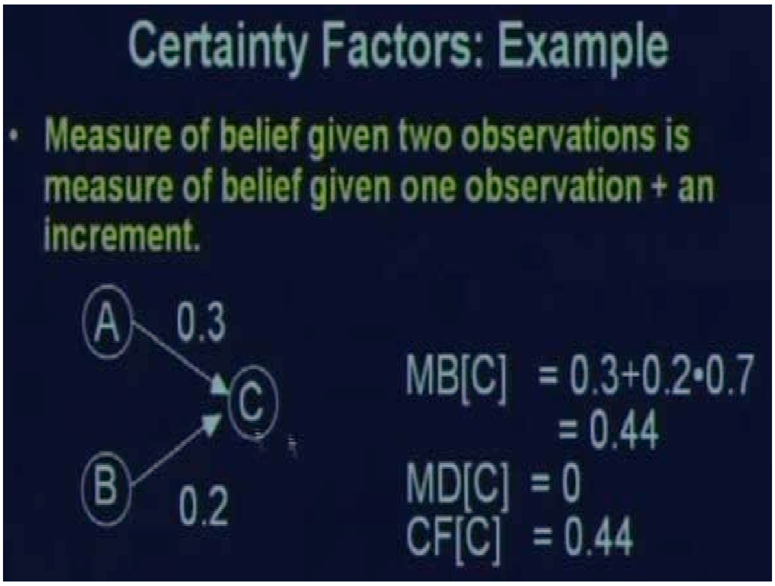
(Refer Slide Time: 03:33)



MB prime b by a must be modified is the measure of belief in b given a is known with certainty and that would have been 0.7. But a is not known with certainty so this has to be modified and so we multiply this with max of 0 and certainty factor of a given some evidence E. Suppose a was known this part was 0.8. a is not known with certainty but a is known with belief of 0.8. In that case this would have been modified to this is 0.7 times max of 0 and 0.8 so it will be 0.8 so it will be .56.

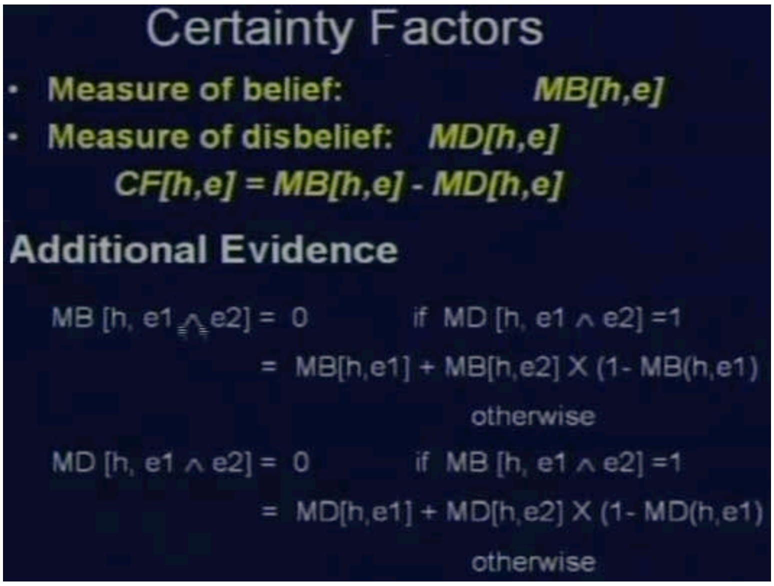
Suppose e is the evidence that led us to believe s and the certainty factor of s given the evidence e was CF(s, e) so we take the max of that and multiply MB prime (h, s) with this. That leads us to the modified measure of belief when the rules are chained. Next let us look at some examples.

(Refer Slide Time: 06:03)



This is the traditionally old example. A is implying C with a certainty of 0.3, B is implying C with a certainty of 00.2 now what would be the total certainty of C? Suppose first it was only known that A is true, B is true is not yet known so this rule was enabled so A implies C has been inferred with 0.3, now forget about this part. Now we get another evidence B and there is a rule which says if B is true then also C is true but that is even a weaker rule where this relation is strengthened with the certainty factor of 0.2. So we have seen in our formulae we can see that MB(h by e_1 and e_2) that means in our case given B and A we will have the measure of belief of h by A plus measure of belief of h by B times 1 minus measure of belief of h by A because it is coming twice.

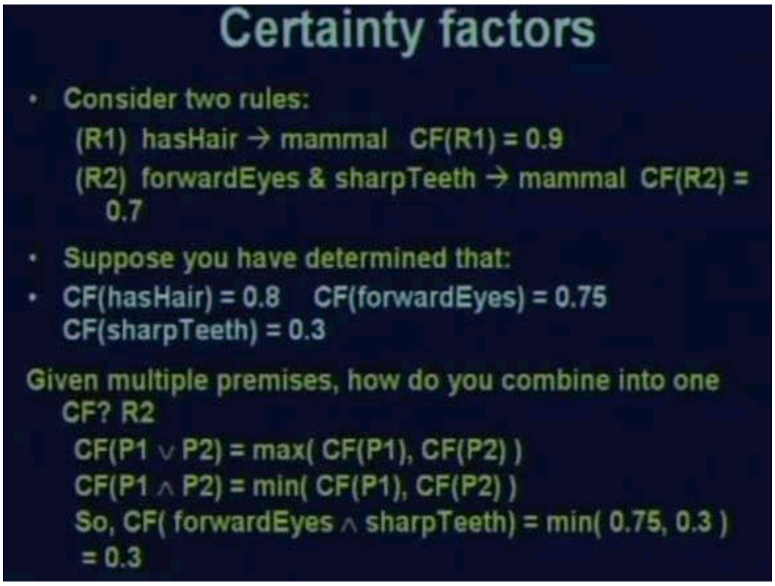
(Refer Slide Time: 07:54)



Then we are applying this formula and we get 0.3 plus 0.2 times 0.7 that is 0.44. And none of these rules spoke against C. Therefore measure of disbelief in C is 0. Therefore certainty factor is MB minus MD so that will be 0.44 minus 0 is equal to 0.44. Let us take a little more complicated example.

Let us consider two rules where the first rule says if something some animal has hair then it is a mammal. And this implication has got strength of 0.9. The second rule is saying if the animal has forward bulging eyes and sharp teeth then it is also a mammal and the strength of this rule that means the strength of this implication is 0.7. So we have got two rules and we have got some facts as well.

(Refer Slide Time: 08:44)



What are these facts?
Suppose from prior observations or information we have found that certainty factor of this antecedent has hair is 0.8. So we know this antecedent not for sure because it might be you were driving in a car and the animal just moved in front of your car and you could have only a glimpse of that and you have got some information and some degree of belief in what you saw. With eighty percent confidence you are saying that the animal that ran across had hair and had forward eyes and that was 0.7.

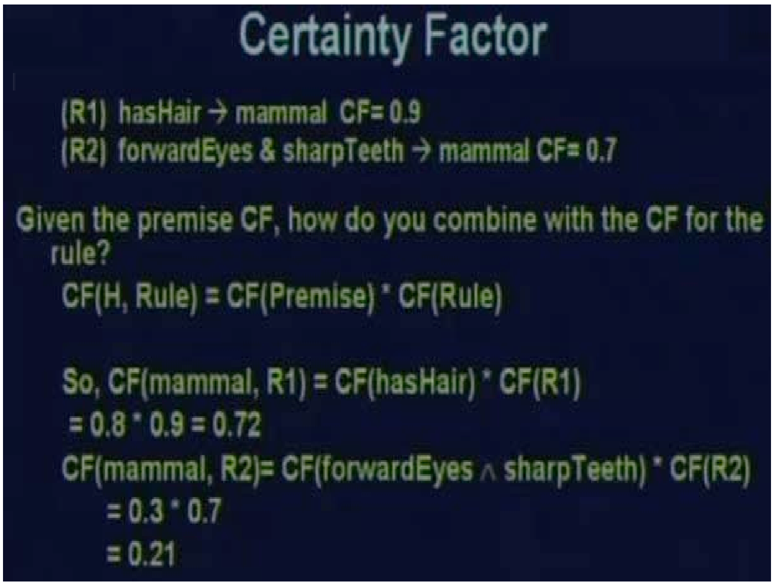
Now note that might be you have looked at the hair part and said you are eighty percent confident that the animal had hair and your friend who was in the car took a glimpse of the eyes and said I am seventy five percent confident that it had forward eyes and another friend said I am thirty percent confident that it had sharp teeth. So might be that from different sources you have got these information So the database is like this now; C F in has hair is 0.8, C F in forward eyes is 0.75, CF in sharp teeth is 0.3 and I have got these rules.

Note two factors; the rules are also not hundred percent certain, they are certainty factors with these implications and neither are these facts hundred percent certain. So our problem is that, given multiple premises the different premises 1 2 3 how do I combine them into one certainty factor when I take this rule?

Now certainty factor of P1 or P2 is max of certainty factor of P1 and certainty factor of P2 and certainty factor of P1 and P2 is min of certainty factor of P1 and certainty factor of P2. So if I consider rule R2 the certainty factor of forward eyes and sharp teeth if I apply this formula then I should take the min of sharp teeth and forward eyes confidence. So min of 0.75 and 0.3 and min of 0.75 and 0.3 is 0.3. So this antecedent part of the rule R2 is 0.3. Now these are the rules has hair implies mammal with certainty factor 0.9, forward eyes and sharp teeth implies mammal with certainty factor 0.7. Now, has hair is a single antecedent so its certainty factor is already known it is 0.8. Therefore now I know the certainty factor of all these antecedent parts it is 0.8

and these 2 together is 0.3 it was 0.75 and 0.3 so since they are handed I have taken the min of those so it is 0.3.

Refer Slide Time: 13:55)

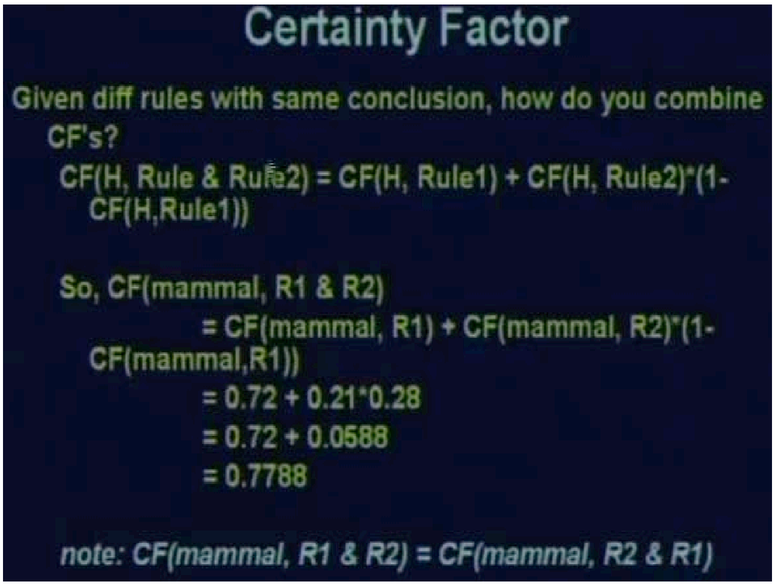


Now how do you combine with the certainty factor for the rule?
The formula is that certainty factor of a hypothesis given a rule. Now this rule is an evidence. We know that certainty factor of H by e is a certainty factor of the premise that is this part and this has to be multiplied with a certainty factor of the implication. So, for the first one it is certainty factor of has hair times certainty factor of R1. So the certainty factor of has hair is 0.8 and the certainty factor of the rule is 0.9 so that is 0.7 2. The first rule tells me that I infer that the animal that ran across the front of my car is a mammal with certainty 0.72.

Now if we look at the second rule that given R2 what is the certainty that it is a mammal?
Then we know that the certainty factor of the antecedent part the forward eyes and sharp teeth is 0.3 and the certainty factor of the rule itself is 0.7 so if we multiply them it is 0.21. Now we come to a point when two rules are stating that it is a mammal, one with a confidence of 0.72 and another with a confidence of 0.21. Then we have got different rules with the same conclusion.

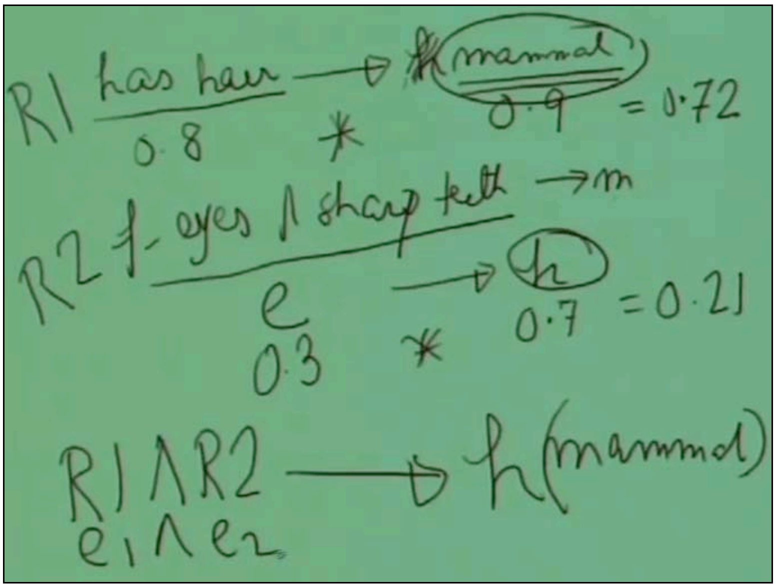
Note that it is the same as here we can consider the rules to be evidences, my evidences are changing, when I had this has hair the CF in that was 0.8 then this was an evidence that was talking of the hypothesis that is mammal so I combine them by multiplying. This was 0.9 so these two got multiplied and I got 0.72.

(Refer Slide Time: 16:02)

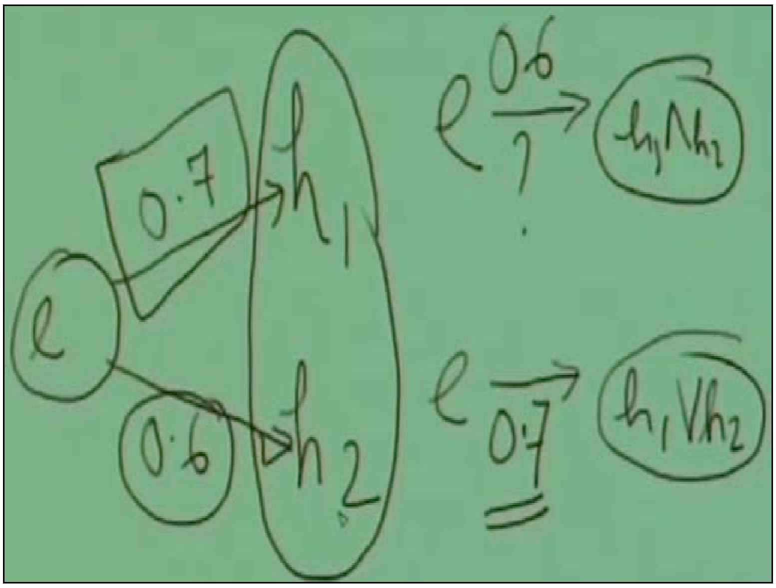


Similarly, I had another rule forward eyes and sharp teeth so this was my e and that was giving to my hypothesis that is mammal and this came to 0.3 and this was something like 0.7 this implication was 0.7 so that got multiplied and I got 0.21. Now this is again mammal so I have got R1 and R2. Both these rules are talking in favor of mammal. So the common hypothesis h that is mammal is being supported by two rules R1 and R2. So this is evidence one and this is evidence two.

(Refer Slide Time: 18:02)



(Refer Slide Time: 20:02)

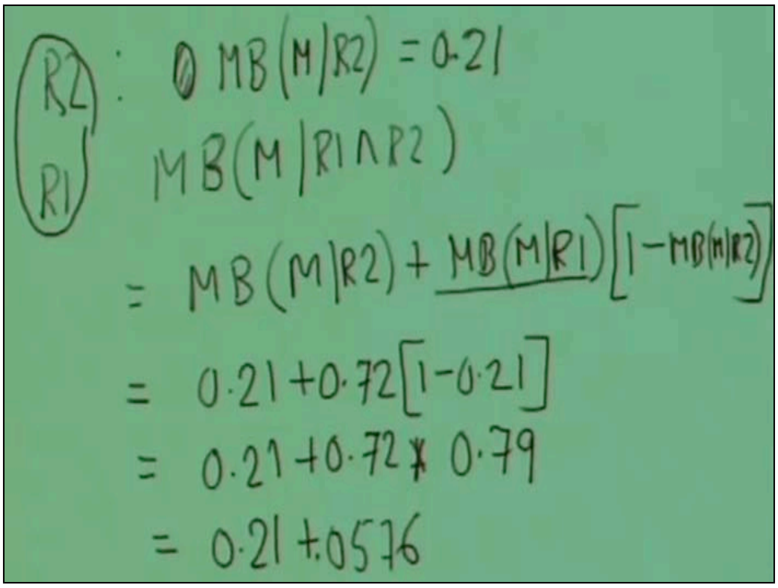


So the same thing can be applied. Combining these rules is the same as looking at it as e_1 and e_2 is leading to h or both R1 and R2 are leading to mammal and this came up with 0.72 and this came up with 0.21. Now how do I combine these two rules?
If we take the rule R1 then I will infer mammal with a confidence of 0.72.But the second rule which has got in itself 0.21 will be added to this but before adding it should be multiplied with 1 minus MB(h, e_1) the same old formula we did here. Now by applying this formula we get 0.72 plus .588 so it is 0.788. Now a very relevant question can be asked.

If you recall the architecture of or the mode of working of an expert system rule based system there are different rules R1 R2 R3 and the inference machine will ultimately decide which rule to fire. So this computation we have just now shown is, we first apply R1 and infer that the animal is a mammal with a certainty of 0.7 2 by applying rule R1 and we apply rule R2 after that then we get this figure 0.788. But there is no assurance that the inference machine and the conflict resolution strategy will always follow the same order of firing the rules.

What would happen if the rules are fired in the alternative way?
R2 is fired first and we infer that the animal is a mammal with certainty factor 0.21 and after that we fire rule R1. Then what would be the computation?
First if rule R2 is fired then we infer with a measure of belief that mammal given R2 to be 0.21. Then we fire rule R1. Then we have to combine these two and we will apply the formula that mammal given R1 and R2 to be MB(M by R2) plus MB(M by R1)(1 minus MB(M by R2)). It is just the complementary, the opposite.

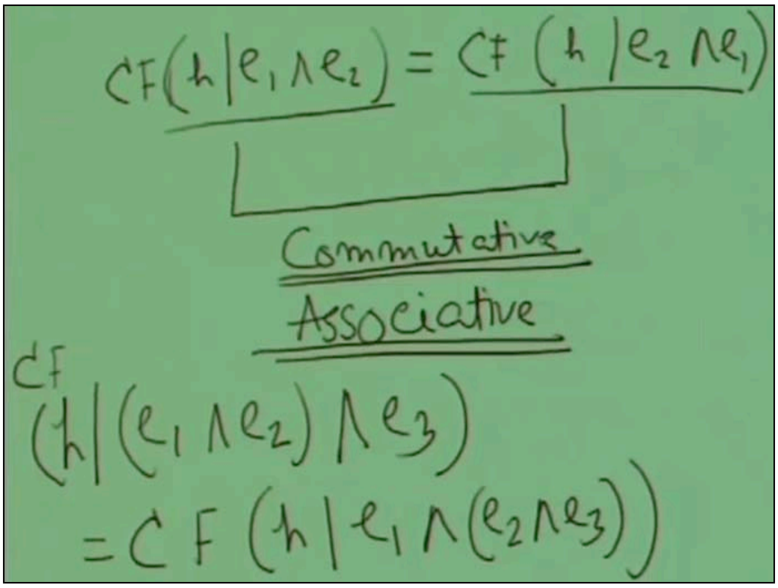
(Refer Slide Time: 23:40)



Here you see this is 0.21 plus 0.72(1 minus 0.21) is equal to 0.2 1 plus 0.72 times 0.79. And if you multiply that it will be 0.21 plus 0.576 if you do it with 0.8 for example it will little less than this as shown. So if we combine this using the same formula we will get the same value if we do it in this way. Therefore what would this factor get changed to? It will be 0.79 so this should be 0.72 times 0.79 so this part up will add up to the same value so we will find that it will be approximately 0.788. So what does it mean?

This means that certainty factor should be independent. The law of combination should be independent of the order in which the rules have been fired. So certainty factor of some hypothesis given e_1 and e_2 should be the same as certainty factor of h by e_2 and e_1. This means the combination rule should be such that it should be commutative. Also, another property this combination rules should support is that it should be associative. That means I first do e_1 e_2 and then take e_3 that means I am combining e_1 and e_2 first and then combining with e_3 and given that the certainty factor of h given this should be the same as if I had combined it in a different way that is h by e_1 I have just done and later on I do say e_2 and e_3 should be same. So that is the associative property. And the other one we discussed about is the commutative property.

(Refer Slide Time: 25:55)



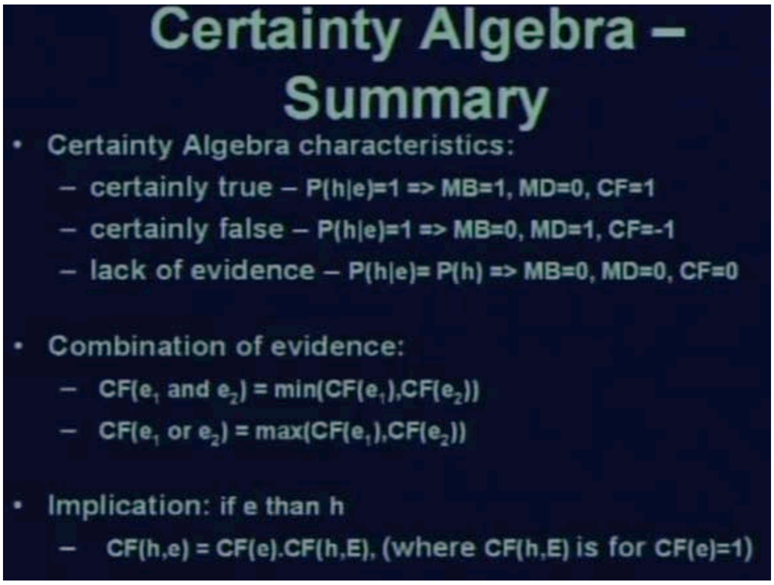
Now these two properties must be satisfied by the certainty factor combination rules. And unfortunately the certainty factor combination rules were designed in a way that these properties are satisfied.

(Refer Slide Time: 26:26)

So, if we are given multiple premises how do you combine them into one certainty factor?

The summary of certainty factor algebra we have discussed till now is that if it is certainly true.

(Refer Slide Time: 26:50)

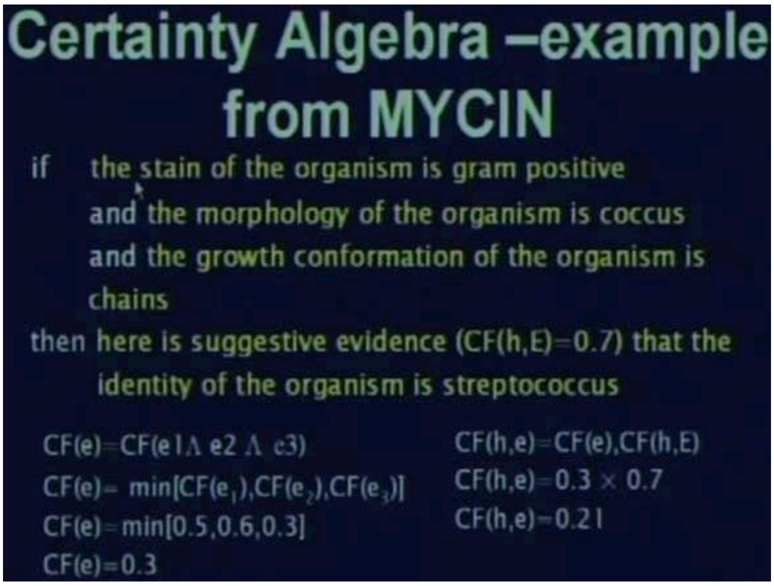


Now look at this;
Let us discuss the difference between probability and certainty factor later. We are probably familiar with the term conditional probability that given an evidence e, given a particular symptom headache what is the probability that the patient has got migraine?
So if we know for sure in that case we can say that the probability of h by e is 1. Now this is equivalent to in our certainty factor algebra that MB(h by e) is 1 and MD(h by e) is 0 and therefore the certainty factor of h by e is 1. If it be certainly false then probability of h by e is 0 then MB is 0, MD is 1 and certainty factor is minus 1. Therefore the range of certainty factor can be from 1 to minus 1. Whereas the range of measure of belief is from 0 to 1 and MD is also 0 to 1 but certainty factor is from minus 1 to plus 1.

When I do not know the relationship between the evidence e and h then whatever probability h has is the thing and there is no question of any conditional probability. In that case I have got no measure of belief, measure of disbelief and the certainty factor is 0, therefore I cannot do anything. Combination of evidences is what we already saw, certainty factor of e_1 and e_2 is min of certainty factor of e_1 and certainty factor of e_2 or disjunction is max of e_1 and e_2. Then implication that is if e than h this CF(h, e) is CF(e) (CF(h by E) where E is the certainty with which we know h if this e was known with certainty. Note that there is a case where we have done for rule chaining.

MB prime times the certainty factor the same thing. This is the way the certainty factors are computed. Next let us have a revisit with a MYCIN rule because MYCIN first introduced certainty factors. So let us see how the MYCIN rules do the same thing with certainty factor.

(Refer Slide Time: 30:30)



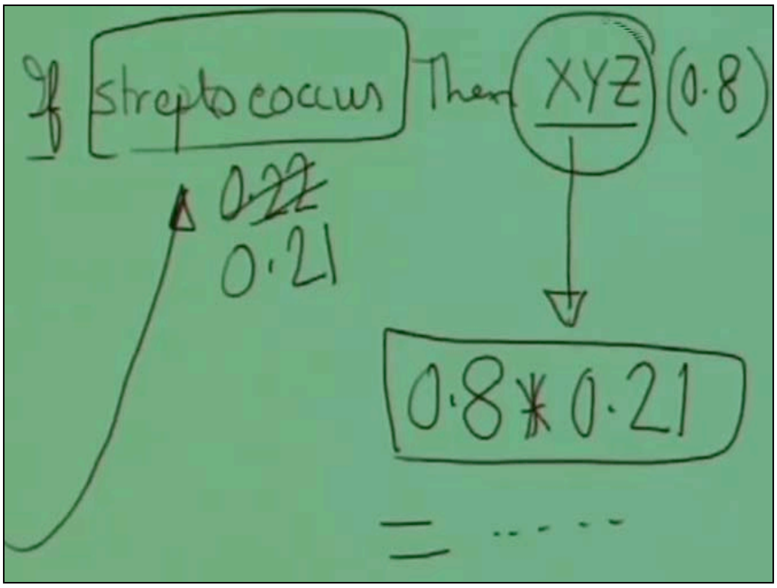
Here is a rule that we have seen earlier. If the stain of the organism is gram positive and the morphology of the organism is coccus and the growth conformation of the organism is chains there is suggestive evidence with certainty factor 0.7 that the identity of the organism is streptococcus. The earlier rule was staphylococcus and was clumps. But here it is chains and streptococcus. So this is altogether another rule. Probably this rule is a little different. So we have got the certainty factor.

What is my certainty factor in this complete evidence?
Suppose this is know with 0.5, the stain of the organism is gram positive that has been known from some other rule and that has been known with 0.5 and this has been known with 0.6 and this one has been known with 0.3. So the evidence of all these conjunction will be of dot mean of 0.5 0.6 0.3 so it will be 0.3 so it is the belief of the antecedent part. Now this entire rule is saying that this conclusion is true with 0.7 provided these are known for sure, these are certainly true.

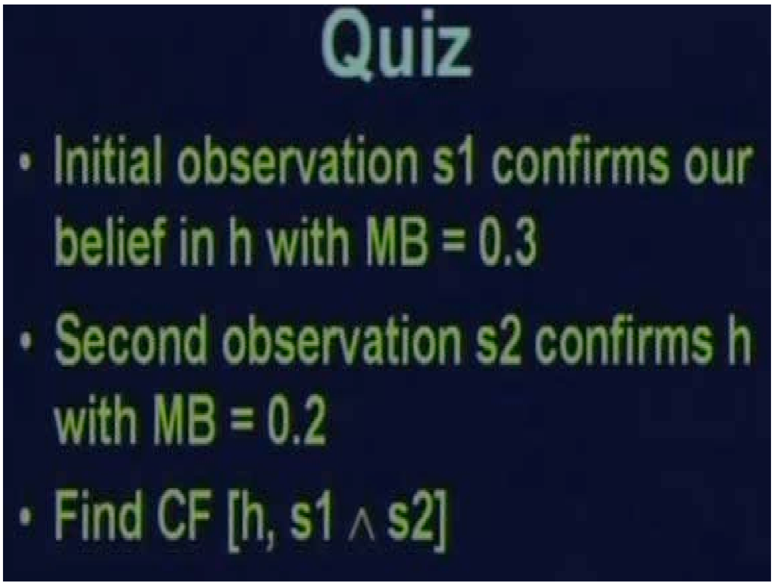
Unfortunately these have come from somewhere, the stain of the organism has been found through some test at some laboratory and you do not have hundred percent confidence in that so you had some .5 confidence into this. Therefore the strength of belief in this conclusion that the identity of the organism is streptococcus has to be modified. It has to be modified again as certainty factor of (h, e) certainty factor of e times this should be multiplication of certainty factor of (h, e). That means if these evidences where known for sure then it will be 0.7. But unfortunately these evidences are known with certainty of 0.3. So ultimately MYCIN inference machine will infer that identity of the organism is streptococcus with a confidence of 0.21.

And if there is some other rule, for example there is a rule which is something like this, if streptococcus then XYZ and the confidence the certainty factor of that is 0.8. Now from my earlier rule I have inferred streptococcus with a confidence of 0.21 and 0.8 will be 0.8 if this streptococcus was known with certainty but it should be multiplied by 0.21 and whatever comes through this multiplication that value will be the confidence the certainty factor that will be associated with XYZ in MYCIN.

(Refer Slide Time: 34:18)

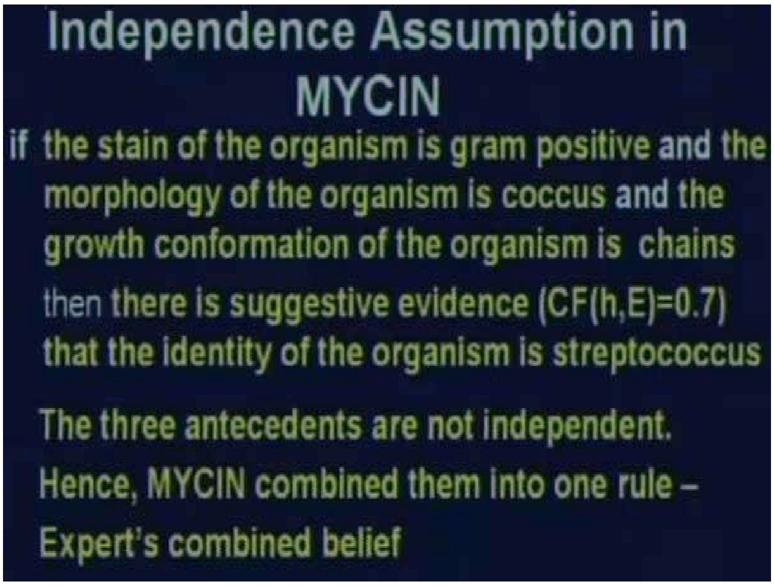


(Refer Slide Time: 34:29)



Quiz:
Compute the following:
Suppose we are trying to confirm some hypothesis h the initial observation is that s1 confirms our belief in h with a measure of belief 0.3 that means there is some rule like this; s1 implies h with certainty factor 0.3. Later on another test report comes from another lab and there is another rule s2 that is confirming the same hypothesis h with certainty factor 0.2. So s1 confirmed h with 0.3 and later on another report comes where s2 confirms h with 0.2 and now as we have got both these you are asked to find the certainty factor of h by s1 and s2. Now let us discuss a very important issue namely the MYCIN rules or in-certainty factors. This factor is behind the success of certainty factor. Or you can say that it is also a limiting factor of application of certainty factors known as independence assumption in MYCIN.

(Refer Slide Time: 37:02)

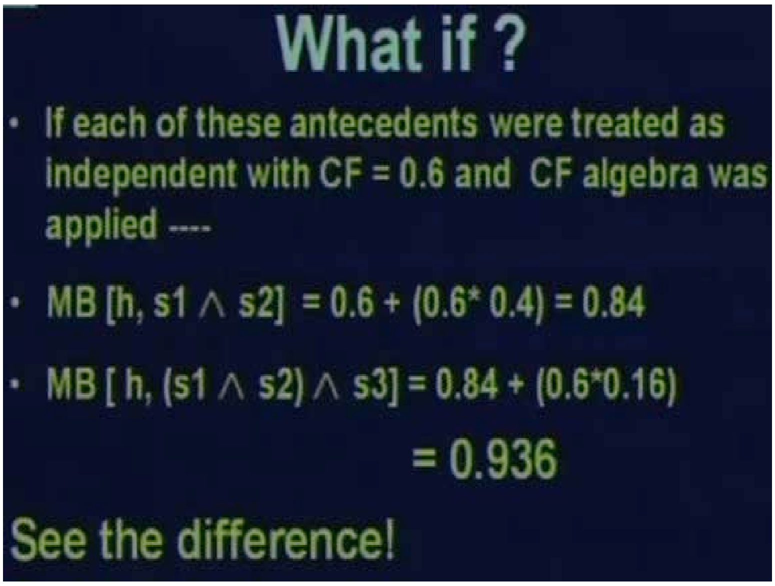


Let us look at this rule:
If the stain of organism is gram positive and the morphology of the organism is coccus and the growth conformation of the organism is chains then there is suggestive evidence with a certainty of 0.7 that the identity of the organism is streptococcus. Now all these three antecedents have been combined into a rule because these are not independent. Maybe that if it is gram positive then being gram positive also makes the morphology coccus so they are not independent. Since they are not independent they have been clubbed together and the expert has combined them together with a certainty factor of 0.7.

If we assumed that each of these antecedents were independent and all of them have got an independent certainty factor of .6 and certainty factor algebra was applied, now what do we mean by independent?
In that case that being gram positive does not affect causal, it is not something that the morphology being coccus. And the morphology being coccus is not casually related to the growth conformation in change. These are three independent observations.

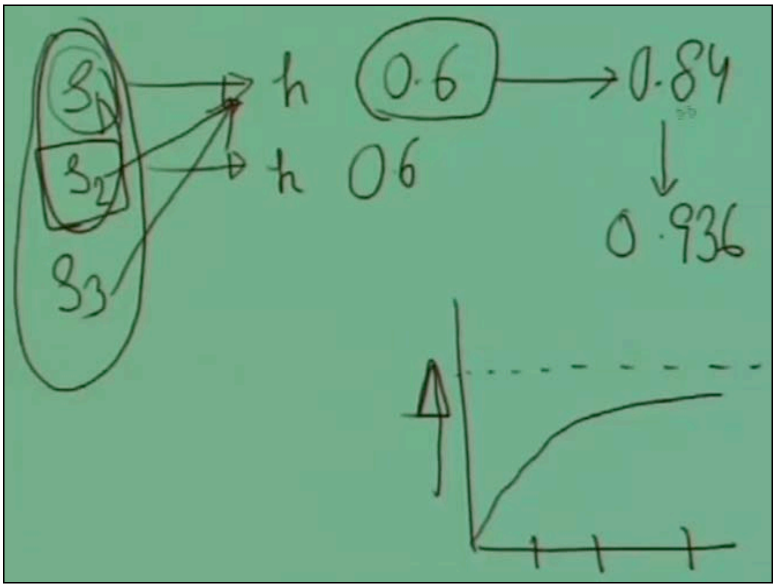
The boy is tall, or here is another example, if we say the father of the boy is rich and the boy has a car then in all priority these are causally related. But if there be something like that the father of the boy is tall and the boy has a car then these two are two independent statements which are not causally related and I independently find the truth in each of these propositions. So we say that two propositions are independent if they are not causally related. If they are not casually related in that case and if we assume that each of them have got certainty factor .6 then let us see what happens and we apply certain factor algebra. Then MB measure of belief in h by s1 and s2, s1 has got .6 and s2 is .6 so the combined belief is 0.84.

(Refer Slide Time: 40:33)



And if we take s3 so we are talking of this rule, this was 0.7, this is 0.6 and 0.6 so when we combine the first two then we get 0.84 and when we combine the third one with it then it becomes 0.936. The figures were .6 then 0.84 and then 0.936, let us see another interesting phenomena here.

(Refer Slide Time: 41:48)



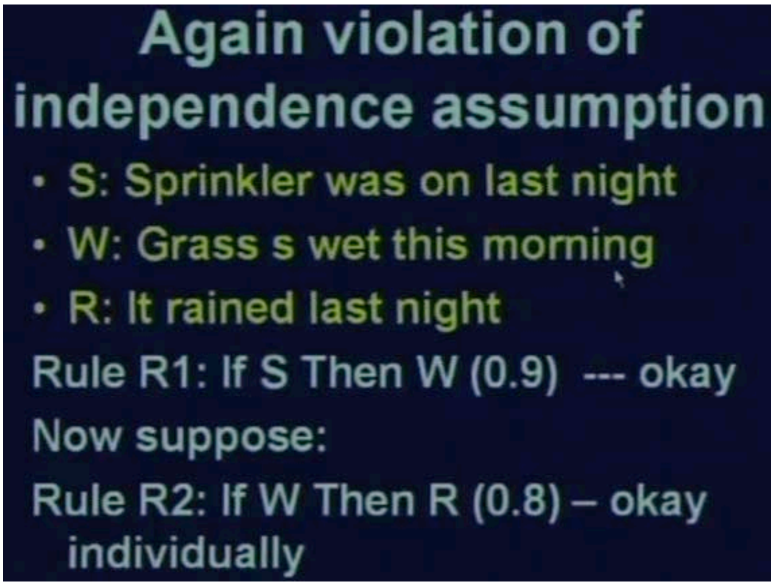
The first one here are independent, s1 supported h and it was with 0.6, and s1 was again with 0.6. Now I got another 0.6 so from this .6 my belief went to 0.84 because these two together now supported h. But when I put in s3 obviously since that is another supporting evidence which will further support this but this has gone up to 0.936. So as I am putting in more and more evidence my confidence in the overall thing is going up exponentially but is gradually going towards saturation where ultimately it can saturate to the value 1. This is one observation that one additional evidence s2 along with s1 really can support this in favor or if it be MD in disfavor.

But gradually as I put on more and more evidences then it tries to saturate and that is only natural. So it is coming to 0.936 which means the overall rule is coming to this.

So obviously it is not tallying with what the expert had said that it should be 0.7 the original rule has said that it is 0.7 but here if I take them independently then it is becoming different. But in real life many of the things will not be independent so here comes the real expertise that I am clubbing together all the dependent clauses and using my expertise I am putting in one particular confidence factor into it that is 0.7. When I do it then it becomes as if e implies h and these two antecedents are taken together.

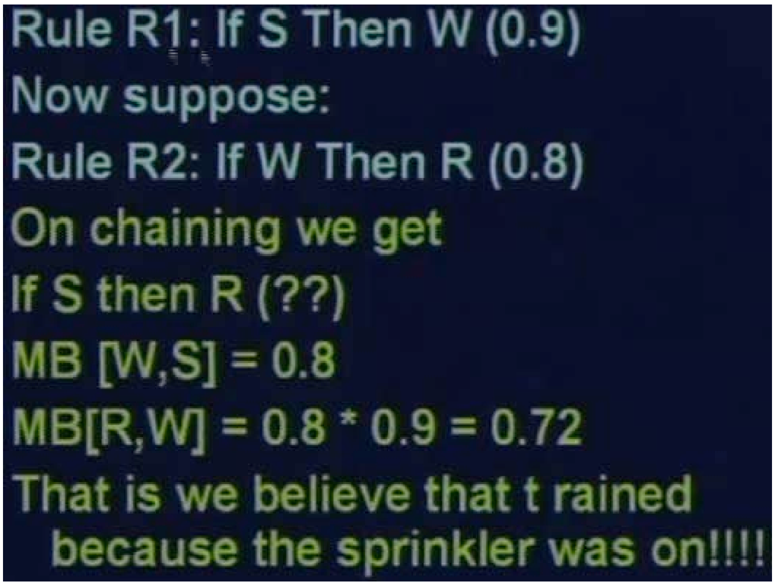
Here is another example. Let us consider these three. Here three notations are used S, W and R. These are three propositions. S is stating that sprinkler was used last night. For example, you are maintaining a lawn and you want to keep the grass green so you have put in a sprinkler which sprinkles water over there. Now S is sprinkler was on, it was put on last night and the statement W says grass is wet this morning today the grass is wet and R is another proposition it rained last night.

(Refer Slide Time: 45:54)



Now if I assume that all these are independent and i write a rule R1 if S that means is the sprinkler was on last night then W then today morning the grass will be wet. independently this rule is fine there is no problem with this and I have put in a confidence 0.9 because I was not too sure how long the sprinkler will be on but if this sprinkler is on in the morning the grass will be wet. Now suppose another rule has been written independently; if I find the grass wet today morning then R that means then I will assume that it rained last night. This rule is also right individually when these two are seen individually. I am making a statement that, if the sprinkler was on yesterday night then today the grass will be wet and I am putting 0.9 degree of belief or certainty factor to this rule. Another rule is saying that, if the grass is found wet today morning then I will assume that with 0.8 certainty that it rained yesterday night that is also very logical. Now the problem comes when we combine these rules.

(Refer Slide Time: 47:55)



If the rules are chained, if S then W and rule two is, if W then R then on chaining we will find a rule if S then R and what is the certainty of that?

If I go on combining MBWS it was 0.8 this rule and I combine this with this rule that is MBRW which will be 0.7 2 then because R was it rained last night I conclude that it rained last night because I found with 0.8 certainty that the grass was wet. So I have concluded that it has rained with a finite certainty while the actual reason was the sprinkler being on.

So when the sprinkler was on I just looked at the grass, and using my rule and assuming that these two rules are fine, independently I combined them. Actually the sprinkler was on, my intermediate effect was the grass being wet and I inferred that it rained. This is the problem and this is not correct. So we believe that it rained because the sprinkler was on. Again this is a violation of the independent assumption. So what is very important in order to do this is to understand the notion of causality and independence.

What is meant by causality here?

We can simply look at some event e_1 causes another event e_2. Maybe another event e_3 is causing e_4. So these are the causal links where one is causing the other e_1 is causing e_2 although e_3 and e_2 were not causally linked. So we must look at the scenario and if they are not causally related we should not combine them which is a very important consideration. So in this rule if the grass is wet, the action wet and rain should not be combined with this sprinkler business and if they have to be combined then this combined belief has to be given a separate certainty factor.

Now with this discussion of certainty factor next we will move to another very popular way of inferencing known as Probabilistic reasoning. Now certainty factor is not strictly following the formal rigor of probability. Later on we will once again see how it deals with uncertainty and then we will see how the theory of probability can be applied to deal with uncertainty.