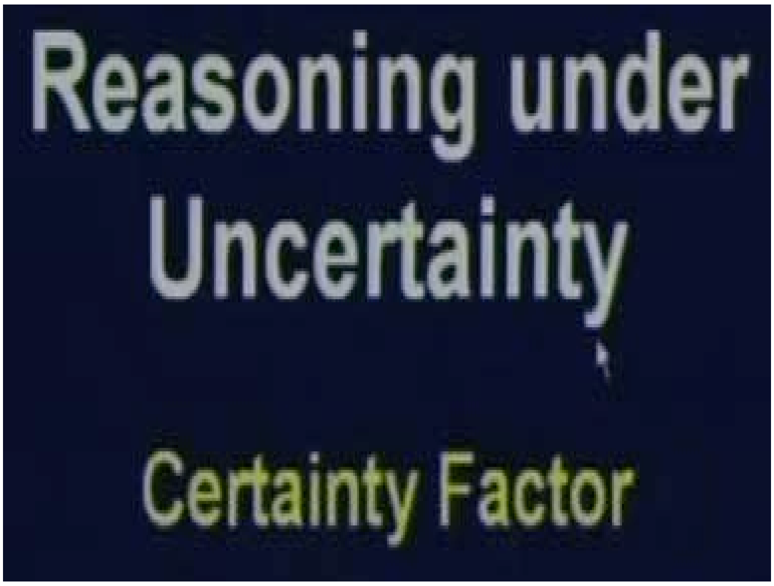
Artificial Intelligence
Prof. Anupam Basu
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur
Lecture # 26
Reasoning with Uncertainty - I

In the last lecture we discussed about rule based expert systems and we also gave the example of MYCIN. It was shown that MYCIN not only provides expert advice but also interacts with the user in a very natural way. And often it is not possible to really distinguish between whether you are communicating with a real doctor or you are communicating with an expert system that is a computer program for that matter. Today in our lecture we will start discussing about uncertainty management. We have seen also in the case of MYCIN that MYCIN rules often put in some suggestive evidence.

For example, if the symptoms are so and so then there is suggestive evidence .7 that this is this and so on. Now this suggestive evidence .7 or .8 whatever the number might be is actually talking of some uncertainty or some weakness in the strength of belief of the expert in the conclusion if we were absolutely certain. For example, if the two sides of a triangle are equal then the triangle is isosceles so in that case there is no uncertainty in this conclusion.

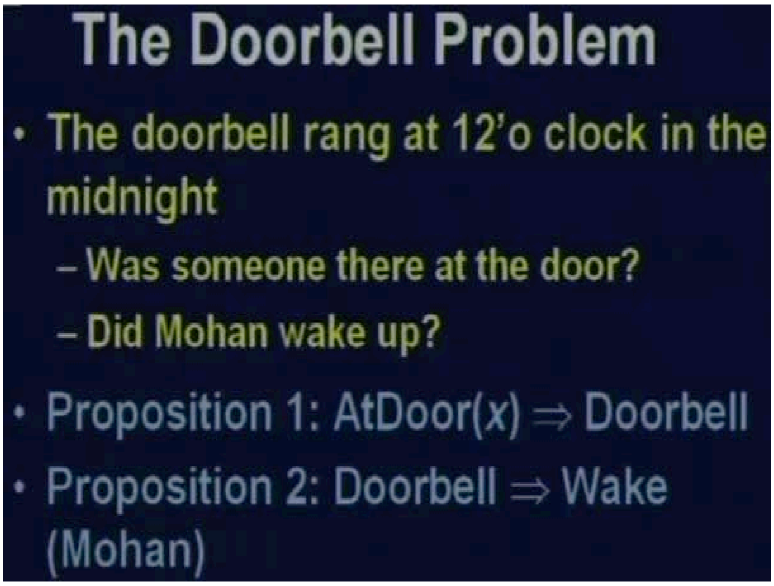
However, because of several reasons in real life scenarios like medical reasoning, Automobile diagnosis and many other things we cannot be very sure about either the evidence or if the evidences are known and its relationship with the conclusion there is some uncertainty lying somewhere. In the next few lectures we will see how such uncertainty is can be handled in the case of Artificial Intelligence. To start with we will look into certainty factors because we have seen a glimpse of these. So our lecture today will deal with reasoning under uncertainty.

(Refer Slide Time: 03:29)



There are different types of uncertainties. What are the different ways in which you can deal with that? To start with we will be talking about certainty factors. Let us look at this example.

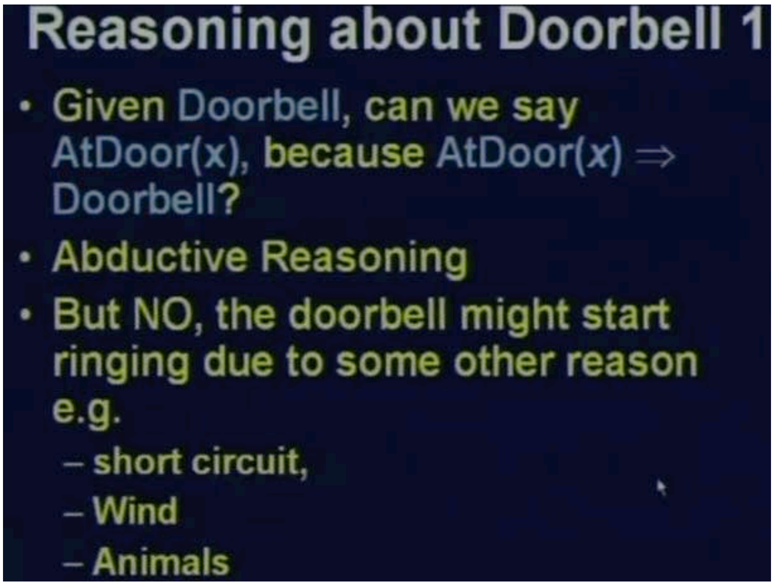
(Refer Slide Time: 04:44)



We call it the doorbell problem. The doorbell rang at 12 O'clock at midnight. Now the question we want to answer is, was someone there at the door? Was there anybody at the door?
And Mohan was sleeping in the room. Did Mohan wake up when the doorbell rang? Suppose I want to answer these two questions. My fact is that the doorbell rang at 12 O'clock in the midnight. Therefore if we place the propositions in the logic form as we have already seen at door x means if someone is at the door then he rings doorbell. And there is another proposition, if there is doorbell then that wakes up Mohan. This should have gone here so wake Mohan. So these are the two propositions. If someone is at the door then he rings the doorbell. If the doorbell rings that wakes up Mohan. Now the problem is suppose the doorbell rang, we know that by the first proposition at door x implies doorbell. If someone is at the door that comes from the previous proposition at door x implies doorbell so the fact is that doorbell has rang.

Can we say that there is someone at the door?
If you recall the classical proposition and logic p implies q means if p is true then q is necessarily true but if p is false then q maybe true or q maybe false. Therefore that is the implication and the implication operator we are showing here. So, according to deductive reasoning if we know doorbell as rung then we cannot say for sure that there was someone at the door using normal implication because p implies q means if p is true then q is true. But q can be true even if p is false so if q is true then p may or may not be true.

(Refer Slide Time: 05:55)

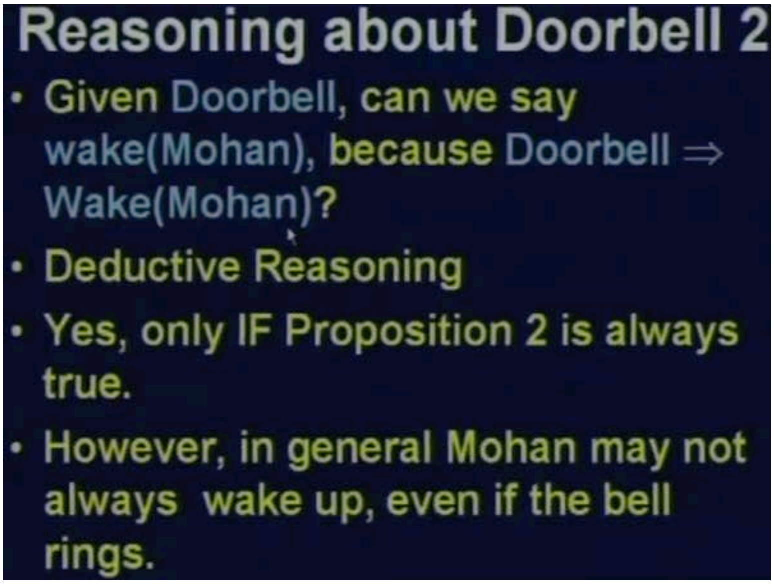


According to deductive reasoning if there is doorbell ringing then we cannot for sure say at door x. But using our common sense or day-to-day style of reasoning often we tend to conclude if there is a doorbell then someone is at the door. Although according to deductive reasoning that may not be true. So this sort of reasoning where if there is something like p implies q and we find q is true then we infer p. This may not be correct according to deduction but that is what we often do and this sort of reasoning is known as abductive reasoning.

Till now we have seen mostly deductive reasoning but in real intelligent systems often we take recourse to other forms of reasoning and all of you will realize and appreciate that when there is a doorbell then we immediately assume that there is someone at the door. So this sort of reasoning given this sort of p implies q type of implication is known as abductive reasoning. And whenever there is a doorbell you expect that someone is at the door and you rush to the door to open it. Most of the time you are right but always you may not be right. So this sort of reasoning is also very useful, abductive reasoning. Later we will study about inductive reasoning.

Here you are introduced to abductive reasoning. But then you may ask if there is a do doorbell then obviously there is someone at the door, but no, the doorbell might start ringing due to some other reason although less probable and mostly that may happen but still it is possible that the doorbell has rung for short circuit maybe because of wind or maybe the dog or some other animal just pressed around the doorbell or whatever there can be thousand and one reasons for that so what to do? How do you go about inferencing these things? Again coming to the second question, given doorbell can we say that wakes Mohan up because we have already seen that a proposition is doorbell implies wake Mohan.

(Refer Slide Time: 10:48)

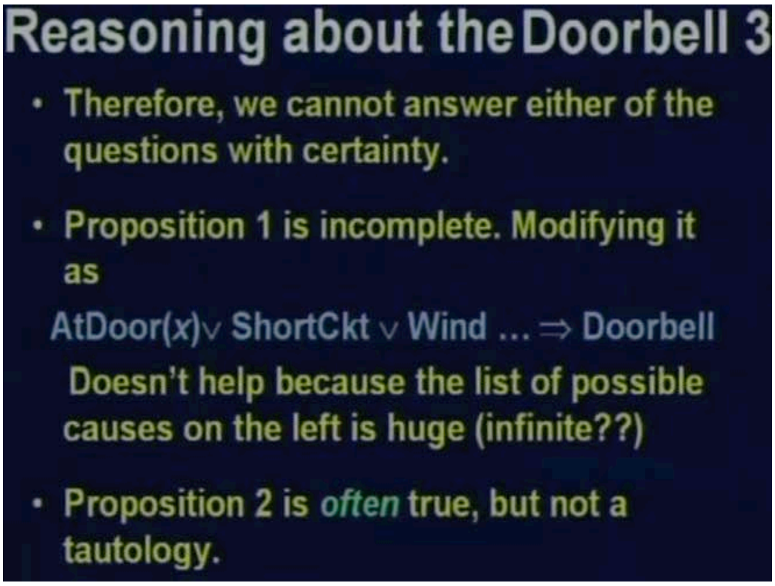


Now this is deductive reasoning. Now we have got the doorbell so the p part of p implies q is true. Therefore obviously it will wake up Mohan. And deductive reasoning immediately tells you that, so it should always be true. But if you think a little bit we can infer that the doorbell wakes up Mohan when this implication is true doorbell implies wake up Mohan. If this entire implication is true that whenever there is a doorbell that will wake up Mohan then obviously the doorbell will wake up Mohan. But now we can also put this rule or this implication in question.

Is this implication doorbell implies wake up Mohan always true?
Maybe there may be cases that when Mohan is really tried which is not always the case and mostly that is not the case but if Mohan is really tired then even the doorbell will not wake him up. Even if the doorbell rings Mohan may not always wake up. So this implication that we are saying may be mostly correct but may not be always correct. Therefore we cannot answer either of the questions with certainty.

What are the questions? Is someone at the door or did Mohan wake up?
None of these we can answer with certainty. The first one we cannot answer with certainty because that is an abductive reasoning. Most probably the doorbell rings only when someone is at the door but there may be also chances of short circuit for which the doorbell rings automatically. Relatively unnatural scenario can also wake up Mohan.

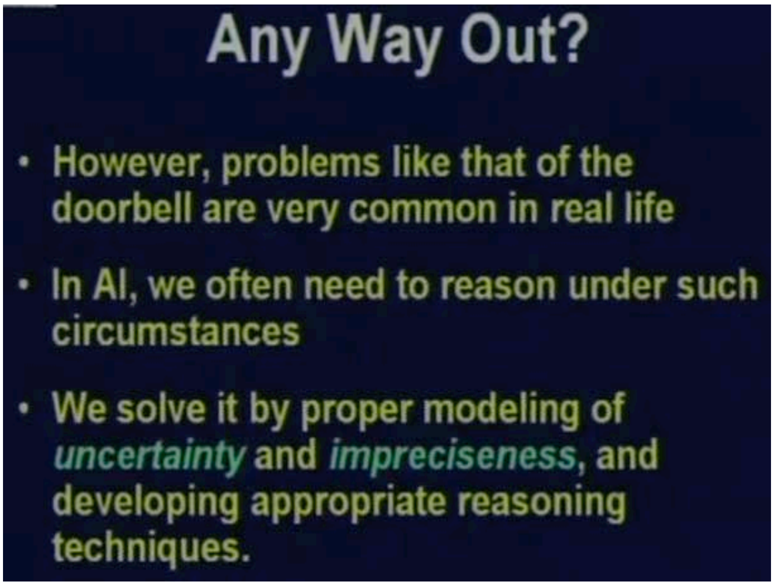
(Refer Slide Time: 13:13)



The second one is, does the doorbell wake up Mohan?
This cannot also be done with certainty because mostly a doorbell wakes up Mohan but this implication is not always true. Sometimes it may not work. Now proposition one was incomplete because the doorbell can ring mostly because of at door x but there can be several other reasons like short circuit, wind etc. Now we could have modified the proposition in this way. We can add on different other conditions at door x or short circuit or wind etc but it really does not help because there may be thousand and one reasons and how many reasons will you go on enumerating in order to make a complete proposition and mostly that is not possible. Therefore this approach does not help because the list of possible causes here are huge in fact they can be infinite.

Now what was the second proposition?
The second proposition was a doorbell implies wake up Mohan. That is often true but that is not a tautology. Tautology means something that is always true. For example, Mortal x or not Mortal x that is always true irrespective of any interpretation. These are called tautologies, A or 0A which is always true. These things are known as tautologies. So proposition two is, often true but is not a tautology may not be always true.

(Refer Slide Time: 16:25)

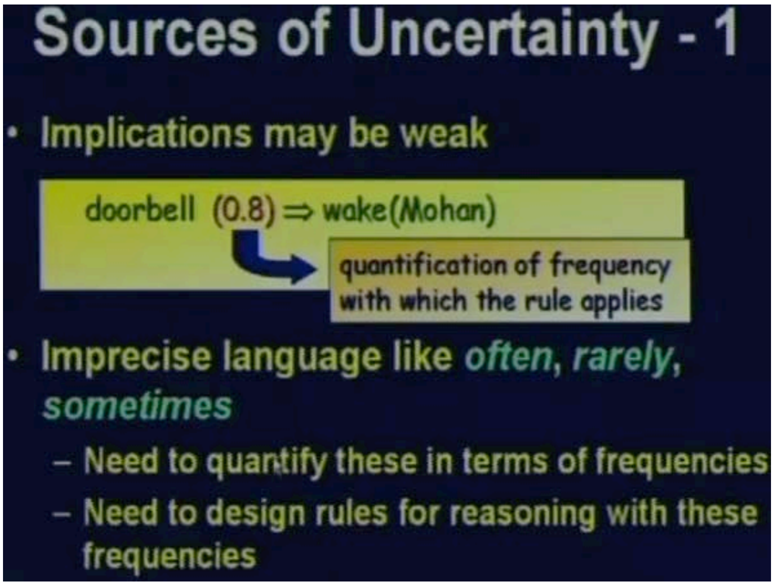


How to do deal with this scenario? Is there any way out?
But these sort of problems like the doorbell problem is very common in fact most of the real life problems we try to solve using Artificial Intelligence techniques are like this, the real world is like this. It is not always 2 plus 2 equals 4. So, in AI we often need to reason under such circumstances and in order to solve it we need to properly model uncertainty and impreciseness. There is a subtle difference between these two words uncertainty and impreciseness.

Uncertainty in a rule as, it rained, it will rain tomorrow, it will rain today, it rains in July etc are statements which are uncertain because although mostly it rains in July in the eastern parts of India but in many of the parts it is not always certain that everyday in July it will rain. So these sort of statements are always associated with some sort of uncertainty. On the other hand, impreciseness is inherent in most of the statements we make.

The boy is quite tall, it is quite likely that it will rain in July. What do you mean by quite likely? The boy is very tall, what do you mean by very? I like him very much, how much do you like him? Is it 5, 10, 15? Now these are not precise statements. The height of the building is for example 20 ft, 200 ft or whatever then we quantify that and that is precise. But mostly many of our statements are not precise. Therefore in order to deal with real life problems we have to handle both uncertainty and impreciseness and that can be handled using appropriate reasoning techniques. Now let us look at the different sources of uncertainty. We have seen the doorbell example.

(Refer Slide Time: 19:11)

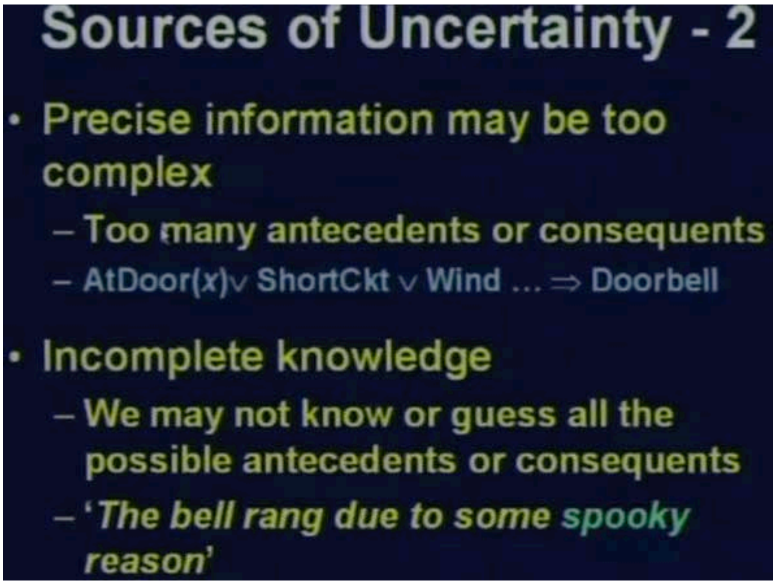


Therefore implications may be weak. They are not absolutely certain that if there is a doorbell then that will wake up Mohan. We can rather write the implication in this way; doorbell .8 implies wake up Mohan. That means now this implication is being given strength of .8. What does it mean?

It may mean that eighty percent of the time doorbell rings Mohan wakes up. So, if I have written doorbell implies wake Mohan then that is a certain proposition as every time the doorbell rings Mohan will wake up. But since that is not the real life case we are putting up some quantification of the frequency with which the rule applies. This rule is mostly applicable in eighty percent of the cases so I write down .8.

Now imprecise language like often rarely, I rarely meet Tom, how frequently? Sometimes, so all these typical words we use in our day-to-day life leads to impreciseness in the statement. Now in order to really work in the AI area and to develop real applicable systems we have to quantify these imprecise terms in terms of frequencies might be. One possible way is to code them in terms of frequencies. The first thing is we need to quantify them and also we need to design rules for reasoning with these frequencies. We have to deal with the rules, we have to write rules in such a way that these uncertainties are captured in those cases. Therefore this is one source of uncertainty, implications may be weak.

(Refer Slide Time: 21:55)

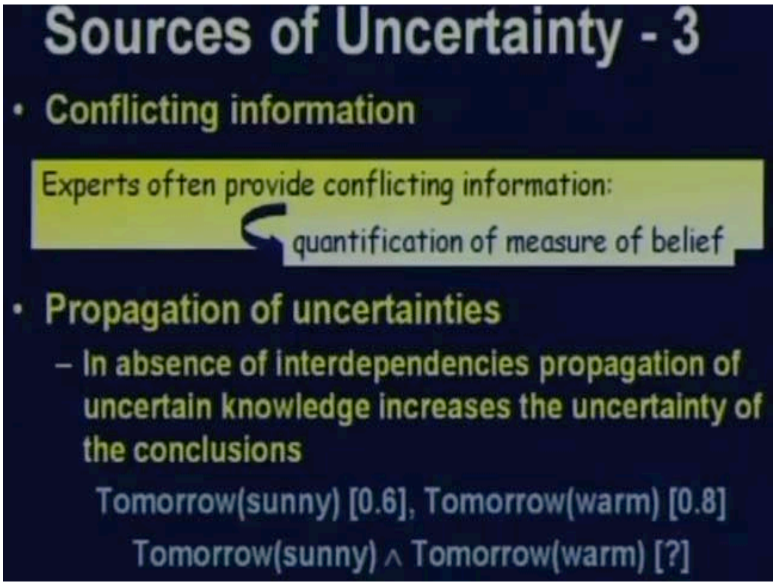


There is another source of uncertainty. It is really difficult to get precise information all the time. We have to leave with imprecise information. Whenever we hear some sound like somebody speaking through the telephone the words are propagated as waves through your telephone line through a channel where there is lot of noise. And there may be some words which are wobbled up and we are intelligent to really understand what that sound was or even if that sound is disturbed or changed during transmission because we cannot expect that noiseless channel to propagate the wave.
Therefore to get the precise information is always difficult.

In the earlier case we have seen when does the doorbell ring?
In order to code that if we had to do really make it very precise we have to write down so many things, at door x or short circuit or wind etc and that is not possible so that is another source of uncertainty.

Another problem that can occur is incomplete knowledge. Knowledge as we say is always incomplete as there is no end to learning. Whatever knowledge you put in your knowledge base that is never complete, you may learn something new tomorrow. But even if whatever knowledge we put in our knowledge base and we start working with that our knowledge base will be incomplete and we will have to live with it. We may not know or guess all the possible antecedents or consequents. Maybe I thought that the doorbell rings besides somebody being at the door, short circuit or wind or animal but there can be some other reason, the bell rang due to some other reason and I really do not know. That means my knowledge is incomplete. So, in order to build a system robust and reasonably good and at least trying to mimic human intelligence we have to deal with such incomplete knowledge.

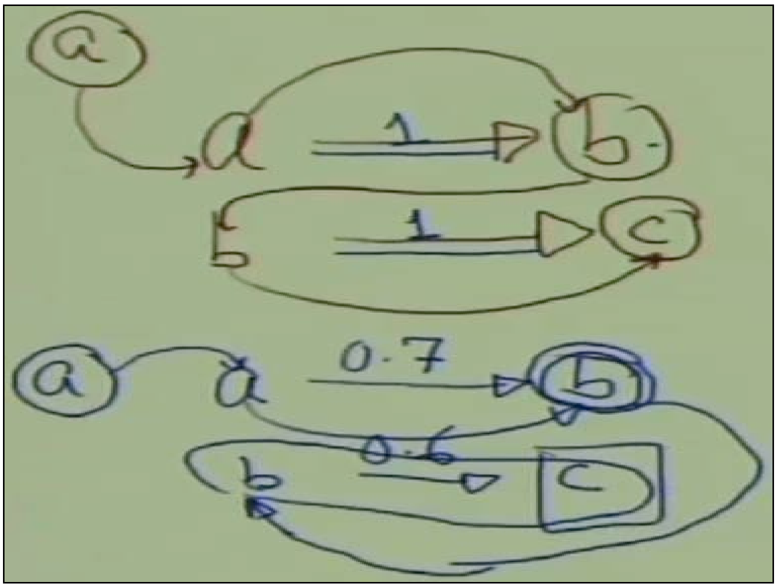
(Refer Slide Time: 25:00)



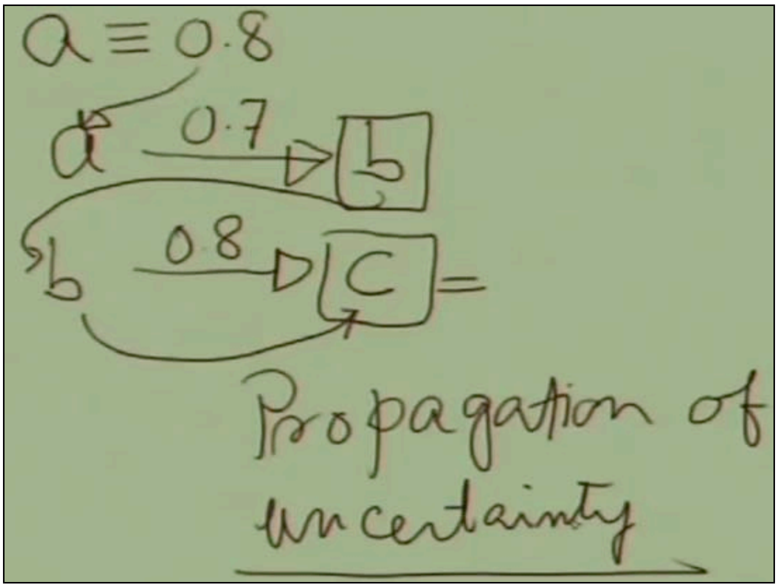
The third source of uncertainty comes from the fact that there are conflicting information. You take a patient with some complicated symptoms, you go to two different doctors and it is quite likely that the two doctors may differ in their opinion. If the disease is not the regular things that we see and some very new peculiar symptoms that we have seen and there are different groups of doctors who may differ in their opinion. So, experts often provide conflicting information because the actual truth is not known with certainty. I may make some statements with some belief, from my experience I put in that statement, I say I am eighty percent sure that this patient is suffering from Thalassemia but the twenty percent I am not sure. So another doctor can say; well I am seventy percent sure that this same patient is suffering from some other disease. Maybe actually both the diseases may be there in the patient or either of these experts may not be correct, that is a reality. So experts often provide conflicting information so when the doctors or the experts speak about these they speak with a degree of belief so we have to quantify the measure of belief.

How strongly you feel about what you are saying? How much sure are you?
It may be eighty percent it may be seventy percent but how much sure are you is measured of belief. And there is another problem here that uncertainties are often propagated. I had a scenario that a implies b and b implies c, if I perform chaining and suppose a is true then I come here and take this tool and infer b and since b is true I come to this rule and infer c. This is chaining and in each of these implications that I had here it was a certain implication so it was with a belief one and this was also with a belief one.

(Refer Slide Time: 29:29)



(Refer Slide Time: 31:18)



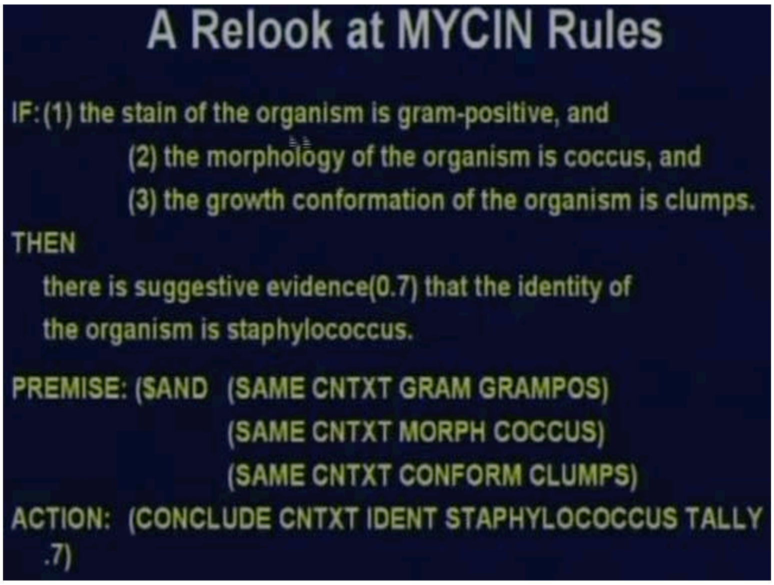
But now if I say that a implies b with some certainty .7 and b implies c with some belief .6 and I know a then obviously the strength of belief with which I know b is less than it was here. So I infer b with a less belief and then I come to this rule and propagate so this thing is propagated, this lack of belief is propagated here and again this rule is not certain because it h- it has got a strength .6 so I infer c with even weaker belief. Now let us take another scenario that a implies b.

The strength of belief is .7 and a is known with some uncertainty, a is not known for sure and suppose a is known with some certainty of .8 now here this implication is a weak implication and further this antecedent is also not surely known so the strength with which b is inferred is also weaker than either .7 or .8. So if I had b implies c then with some belief of .8 this implication then through this chain of reasoning c will be inferred with a much weaker belief.

This is what we mean by propagation of uncertainty. So, in the absence of interdependencies of propagation of uncertain knowledge the uncertainty of the conclusion increases. Here is an example.

Suppose I know tomorrow will be sunny with a belief .6 and tomorrow will be warm with a belief .8. When I take them together; Tomorrow will be sunny and tomorrow will be warm what will be the certainty of this conjunction?
This is a conjunction of these two uncertain statements uncertain propositions tomorrow is sunny tomorrow is warm. When I take them together what will be the strength of belief in this conjunction?
Now in order to handle such scenario certainty factors were proposed. This was proposed along with the development of MYCIN at the Stanford University so this is often known as Stanford certainty factors, Stanford certainty algebra etc. Now let us have a re-look at the MYCIN rules.

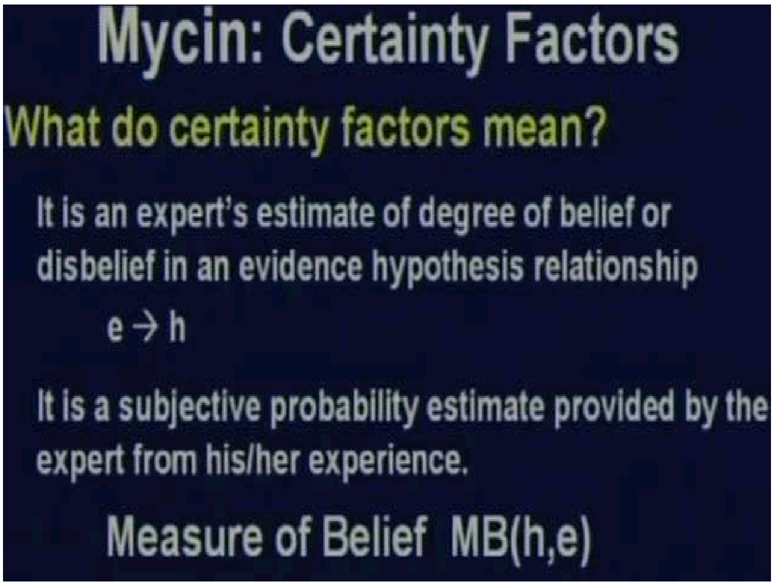
(Refer Slide Time: 33:18)



If it is one rule I am looking at the first antecedent is, 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 clumps then there is suggestive evidence .7 that the identity of the organism is staphylococcus. So the experts could conclude that the organism is staphylococcus with a belief of .7 assuming that all these three antecedents are true. Now this .7 is the certainty factor.

What do certainty factors mean?
It is an expert estimate of the degree of belief or disbelief in an evidence hypothesis relationship. We will be dealing with this sort of relationships now for a while quite frequently. e implies h means given this evidence we conclude this evidence leads to the hypothesis h. h stands for the hypothesis and e stands for evidence.

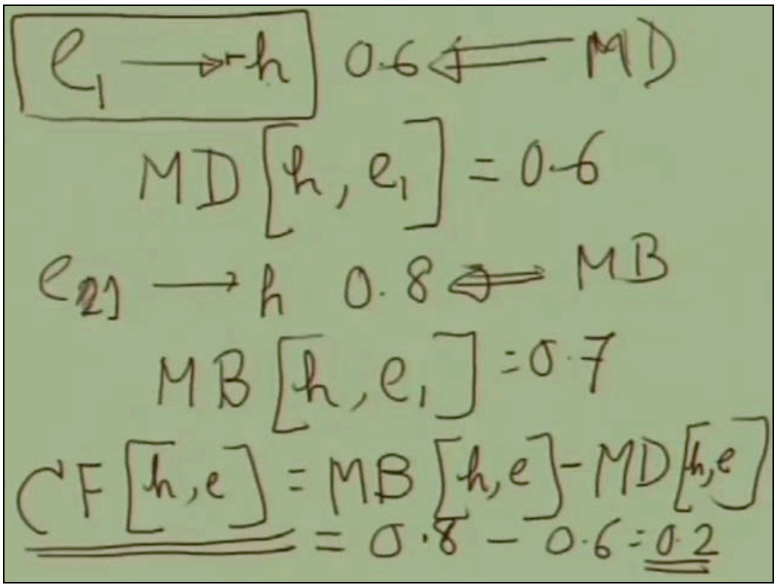
(Refer Slide Time: 34:48)



Now, if we go back to the earlier rule where from did this value come?
This value actually came from expert's knowledge and subjective estimate of the relationship between the occurrences of these symptoms; Stain of organism being gram positive, the morphology of organism being coccus and the growth of organism being clumps and these three relates to the organism being staphylococcus with a strength of .7. So these are subjective estimates, we are saying subjective probability estimate provided by the expert from his or experience. So this .7 is actually a measure of belief that is writing it in this way MB means measure of belief of the hypothesis h given the evidence e. Now this rule is supporting the fact that the identity of the organism is staphylococcus. So this is the measure of belief.

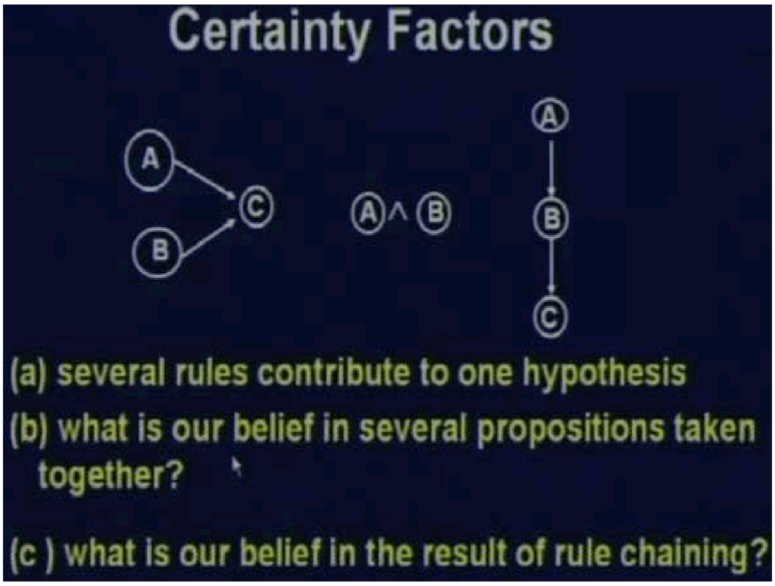
Similarly, there can be measure of disbelief which I would have written MD(h, e) measure of disbelief say an evidence e relates to the hypothesis h with a negation. That is, I do not believe that given this evidence this hypothesis will be supported so something like NOTh. If e tells NOTh with some strength .6 then we can say measure of disbelief in the hypothesis h given the evidence e is .6 and let this be evidence e1 and some other expert given this e1 says this h with .8. Then the measure of belief of h given e1 is .7. Now certainty factor CF(h, e) is measure of belief in (h, e) minus measure of disbelief given (h, e). So in this peculiar case certainty factor will be something like .8 minus .6 is equal to .2. So certainty factor is a measure of belief minus measure of disbelief.

(Refer Slide Time: 39:00)



So a rule can either say in favor of hypothesis or against a particular hypothesis. A measure of belief like this is in favor whereas this one is against so this is leading to MD whereas this is leading to MB. But the situation becomes a little more complicated when we have got more than one rule contributing to one hypothesis. For example, there is a rule here A implies C and there is another rule B implies C, Now both this rules are pointing to C but with different degrees of uncertainty, different degrees of belief and disbelief. So this is the case a; several rules contribute to one hypothesis.

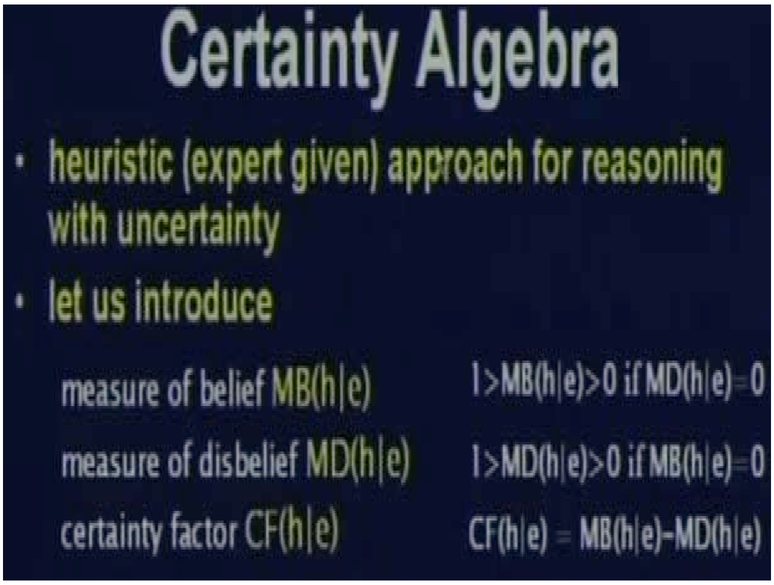
(Refer Slide Time: 40:18)



The second problem that can come up is what is our belief if several propositions are taken together?

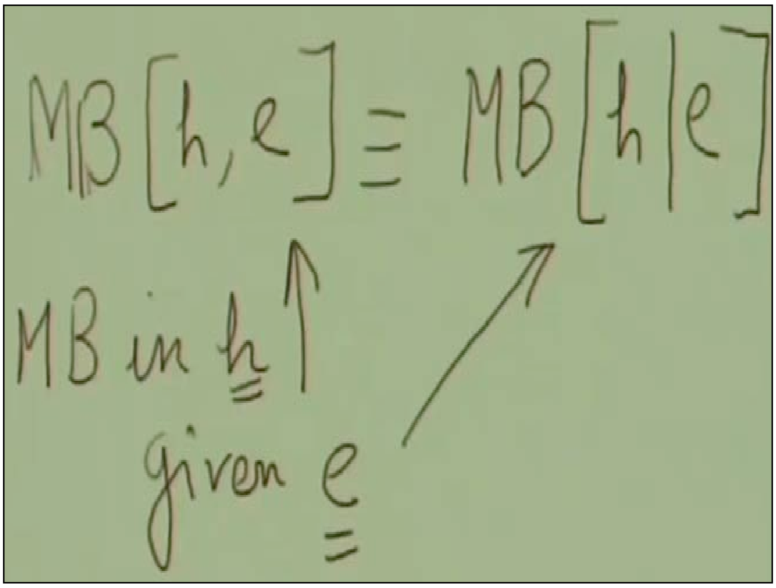
For example; here A is a proposition which is an uncertain proposition and B is another proposition which is also an uncertain proposition, If I take them together A and B what would be my combined certainty about this?
Obviously that will be less.
The third case that can occur is, what is our belief in the result of rule chaining?
There is a rule A implies B with some uncertainty maybe in A in this fact itself or in the implication itself so there is some uncertainty here and there is some uncertainty here. So when I chain them together what would be the certainty in this result C?
Obviously that will be even weaker but how do you compute that because ultimately we want to find some sort of a computable value. So these are some typical cases which we will have to deal with.

(Refer Slide Time: 41:48)



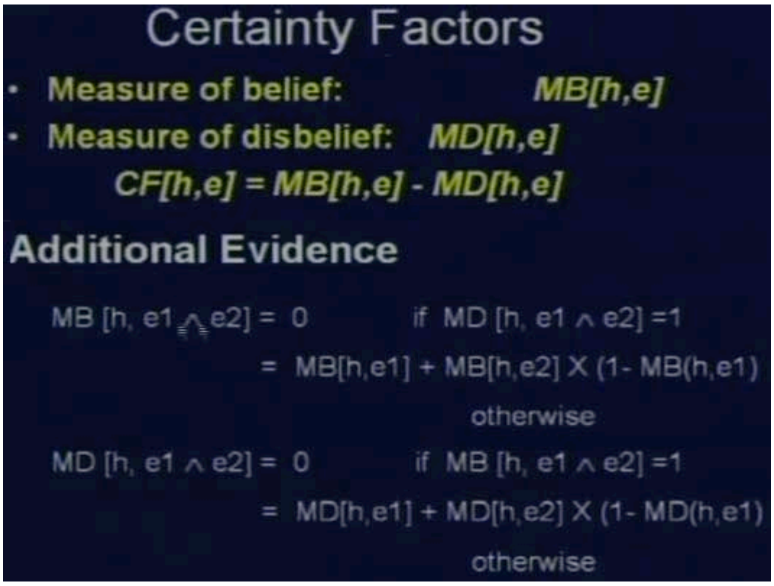
Now certainty algebra is a heuristic or rather expert given approach for reasoning with certainty. So measure of belief MB(h, e) now here you may face some rotational problems.

(Refer Slide Time: 43:11)



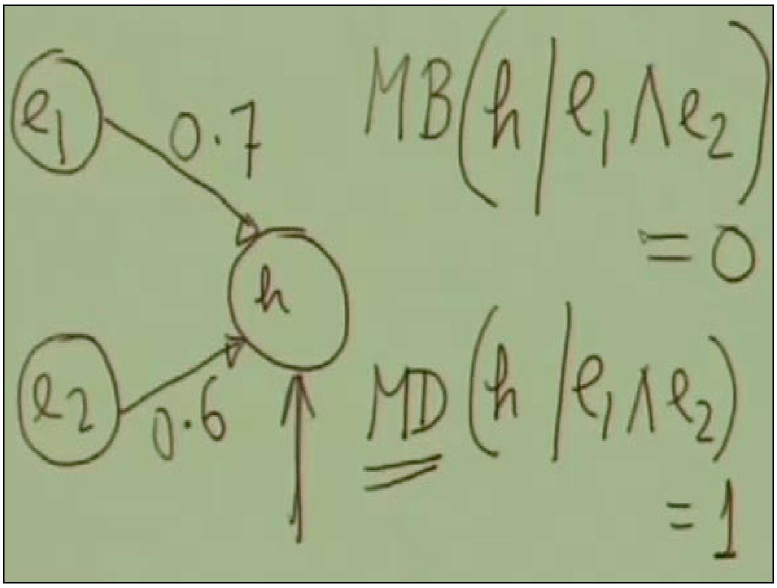
When we write MB(h, e) that means MB in h given e. That means if e is known then what is my measure of belief in h? The same thing is written as MB h given e these two are equivalent and the same thing is meant by both these notations. Therefore what is measure of belief MB h given e? It lies between 1 and 0. If MD(h) given e is 0 it will be some number. Measure of disbelief is again between 1 and 0 when the other one is known to be 0. And certainty factor of h given e is MB minus MD. So we start with a measure of belief which is MB(h, e),measure of disbelief is MD(h, e) and certainty factor of h given e is MB(h, e) minus MD(h, e). Now there may be multiple evidences which are supporting the hypothesis h.

(Refer Slide Time: 44:44)



I am given some measure of belief of h given e. Now I can have additional evidence like here h is supported by e1 as well as e2 and there are different strengths of relations like e1 is supporting h to some degree, e2 is supporting h to some other degree. Now what how can I compute MB(h, e1) and e2. It is again the same scenario. Suppose I have got some e1 is supporting the hypothesis h with some strength .7 and there is another evidence e2 which is supporting h with some other strength .6, now what is my combined belief in h? This is what I want to know when I take both these evidences together.

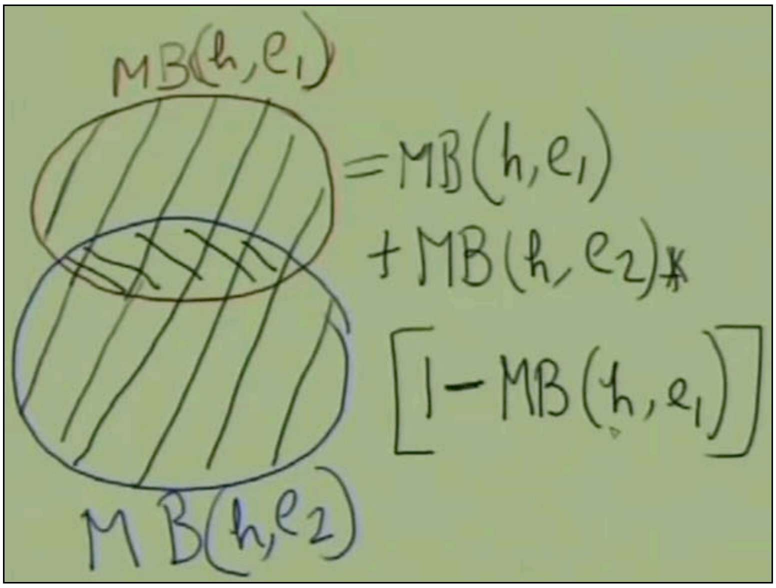
(Refer Slide Time: 45:02)



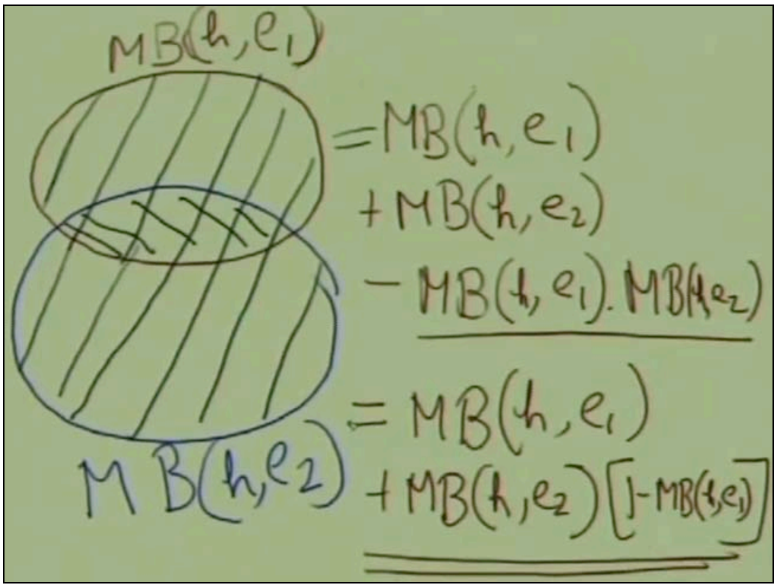
So what is my measure of belief in h given e_1 and e_2?
Now obviously if the measure of disbelief in h given e_1 and e_2 is 1. That means e_1 and e_2 taken together certainly makes me to disbelieve this then obviously this 1 is 0. Otherwise what will happen?
Otherwise I will compute it as MB(h, e_1) plus MB(h, e_2) times 1 minus MB(h, e_1). Think of a space, this is my measure of belief in h given e_1. And independently let me select another color, this is my measure of belief in h given e_2.

Now what is my total belief given e_1 and e_2?
That would be this space.
Now what is this space?
This part is MB(h, e_1), this blue part I add is MB(h, e_2) but there is a part which is common here so I must subtract that.
And what is this part?
This part is coming twice so I will not take this entire part but a part of that so times I multiply this with 1 whichever is falling over here minus MB(h, e_1).

(Refer Slide Time: 48:55)

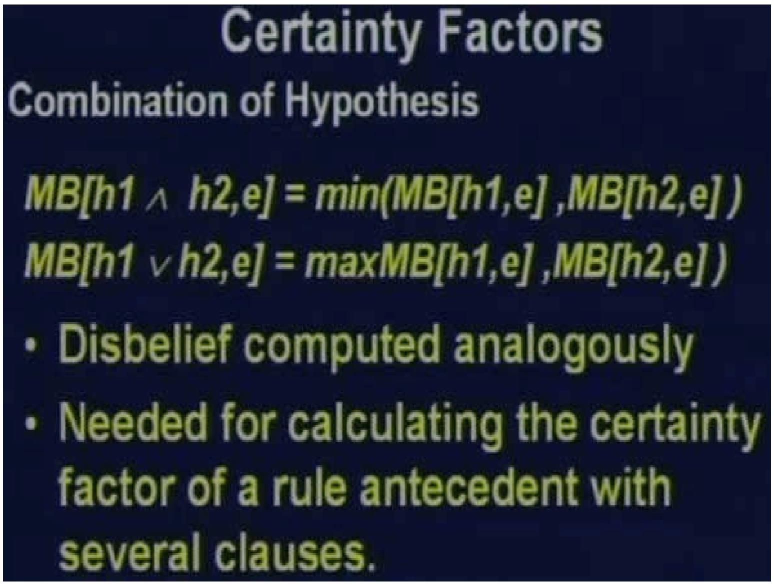


(Refer Slide Time: 50:18)



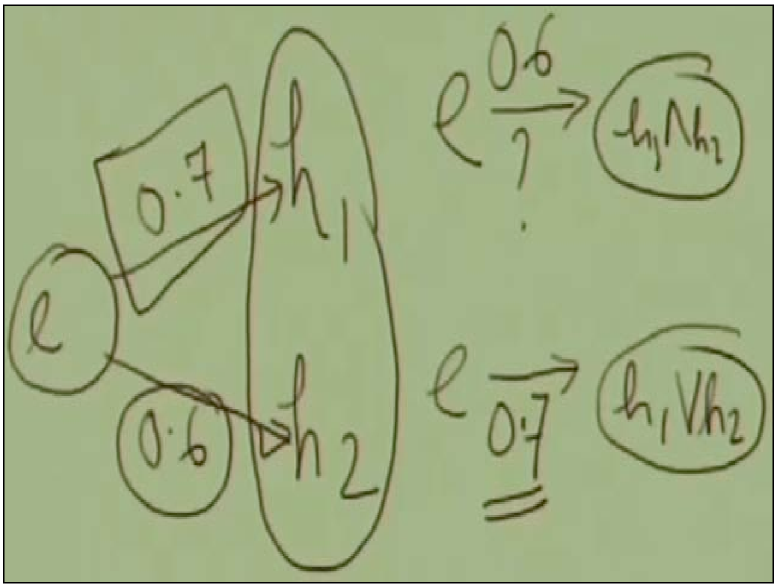
I can write this part as:
MB(h, e_1) plus MB(h, e_2) minus MB(h, e_1) MB(h, e_2) this part is common. This is the common part here. If I take the common part out it will be MB(h, e_1) plus MB(h, e_2) then this part will be 1 minus MB(h, e_1). That is the expression. Therefore this is the formula. Now, similarly for MD measure of disbelief will be similar. MD(h, e_1) and e_2 will be 0 if the measure of belief is 1 otherwise it will be MD(h, e_1) plus MD(h, e_2) times 1 minus MD(h, e_1). So it is the combination of hypothesis.

(Refer Slide Time: 50:55)



The earlier one was combination of evidence. Now we are looking at combination of hypothesis. This is a little different. In the earlier part this is the combination of evidence. When there are more than single evidence what is our strength in the hypothesis?
If there be multiple hypothesis h_1 and h_2 given e then e supports (h_1, e) and supports h_2 with different strengths. Now what is common? If I take the conjunction what is the strength of that? Now this is minimum of MB(h_1, e) and (h_2, e) that means I have got a hypothesis h_1.

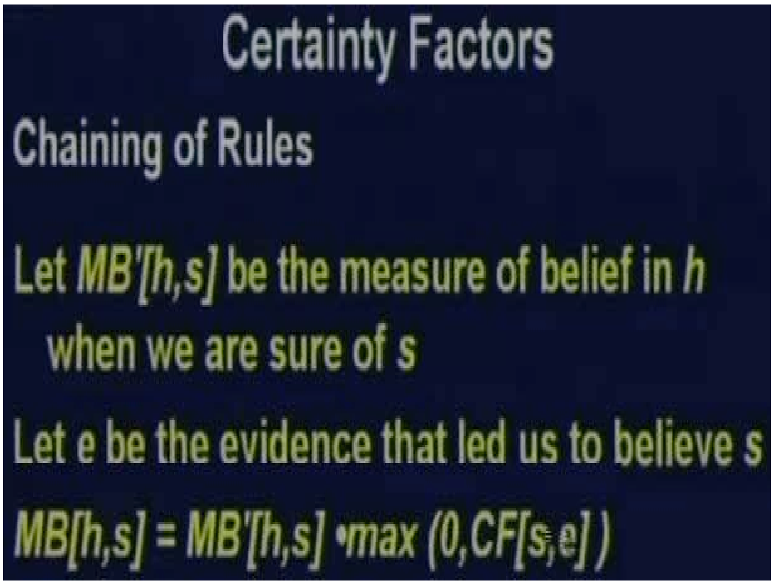
(Refer Slide Time: 52:02)



I have got a hypothesis h_2, one evidence is supporting this and the same evidence is supporting this with .7 and this with .6 and I want to find whether e is supporting h_1 and h_2, what is the strength of that?

It will be the minimum of the weaker of these two links because both these will be true with the minimum value so this will be .6. Similarly, if I had or is implying either h_1 or h_2 then obviously this majority part will come here as .7. So that is the formula that measure of belief of h_1 and h_2 will be the minimum of the individual beliefs and the disjunction h_1 or h_2 will be the max of these two h_1 and h_2. Now, disbelief will be computed analogously in the same way. In place of MB it will be MD. Now this is needed for calculating the certainty factor of a rule antecedent where there are several clauses. The other thing that remains here is chaining of rules. We have seen this chaining of rules several times.

(Refer Slide Time: 54:18)



Now let MB prime (h, s) s is symptom and MB prime note this prime be the measure of belief in h if we are sure about s. But if we are unsure about s because s has been inferred by another evidence e so e led us to believe s and that was not certain so there was some weakness over there so when I do not have full confidence in this then this MB prime will have to be modified. MB prime is the measure the belief in h if s is known for sure but s is not known for sure therefore given the uncertainty in s we will have to modify the MB prime in this way that will be multiplied with max of 0 and certainty factor of s given e. Whatever this one is, if this is 0 then this will be 0, this entire thing will be 0 if s is known to be false for sure then this will be 0 otherwise whatever be the certainty factor here that factor will multiply MB prime to give the measure of belief in s.

Note: Preview of next lecture not added in this lecture's editing part.