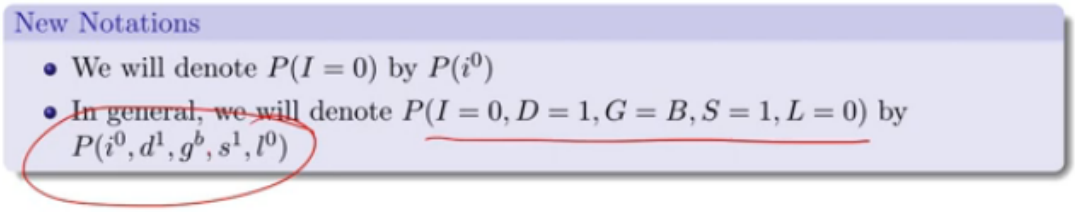
Lecture 16.5

Different types of reasoning

Encoded in a Bayesian Network

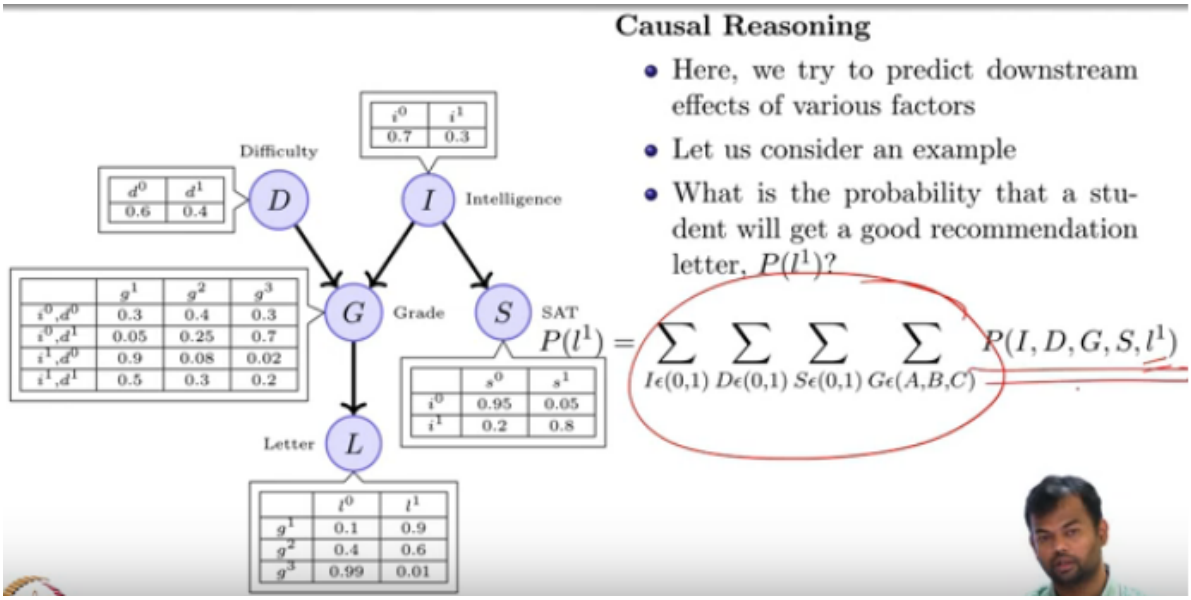
So, now this the next module .Right? This is again kind of two, motivate why do you need these joint distributions and, what kind of different reasoning that you could do and, we'll do some interesting reasoning on them and I hope to make a point, that these are important things which you could think of in various applications and that's why you need a joint distribution as opposed to just learning the conditional distribution, of probability of oil given all the other factors .Right? Because, if you have the joint distribution, you could do much more reasoning on top of the joint distribution .Okay?

Refer Slide Time :(0:50)



So, let's look at the different types of reasoning that you could do in a Bayesian network, so from now on we will use this notation, that if I want to say that intelligence equal to high or low, I'll just call it as a zero, so zero means low, high one means .Okay? In particular this entire thing I will compactly represent it as the following. Just make sure you're comfortable with this it's nothing great .Okay? I've been fine with this .Okay?

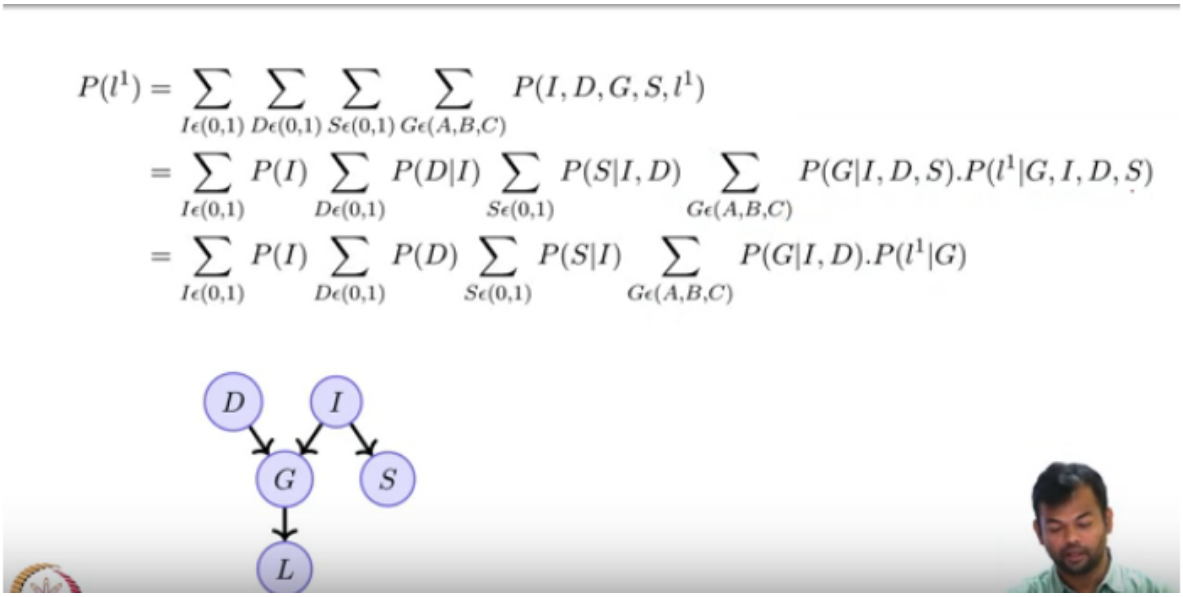
Refer Slide Time :(1:13)



Now, the first type of reasoning that we can do is causal reasoning and, in causal reasoning as the name suggests we try to do these downstream effects of various factors .Right? So, I would want to know, what is the probability that a student will get a good recommendation letter Okay? How, would you compute this probability how do you compute this is, this straightaway given to you anyway? No,

how will you compute this probability, I mean either you don't know? Or it's too obvious that you don't want to say, which one is it, if it's the latter please say it .Right? I mean otherwise how will I know, how will you compute P of, l1 what's that called you will marginalize over all the variables which are of not of interest to you .Right? So, this is how you will actually compute P of, l1 you'll just marginalize over all the other variables and just keep the value of L as, high is it fine but, do we have this actually what do we have.

Refer Slide Time :(2:07)



We have the factors .Right? so let's see I'll do it once, I thought I wouldn't have to do it but, you're encouraging response actually forces me to do it, so this actually factorizes as, various factors and now what I've done is just push the summations, I just adjusted the summation, so that the variables which are not important .Okay? So, now what I'm going to do is just use the factorized, form of this and I'll just adjust the summation so, that I'm doing minimum work .Right? Because, when I'm inside I'll just look at the appropriate variables and so on and now, this actually means the sum of these two terms, for all possible values of the grade .Okay? And so on.

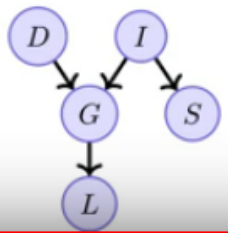
Refer Slide Time :(2:50)

$$P(l^1) = \sum_{I \epsilon (0,1)} P(I) \sum_{D \epsilon (0,1)} P(D) \sum_{S \epsilon (0,1)} P(S|I) \sum_{G \epsilon (A,B,C)} P(G|I,D) P(l^1|G)$$

$$= \sum_{I \epsilon (0,1)} P(I) \sum_{D \epsilon (0,1)} P(D) \sum_{S \epsilon (0,1)} P(S|I) 0.9(P(g^1|I,D)) + 0.6(P(g^2|I,D)) + 0.01(P(g^3|I,D))$$

- Similarly using the other tables, we can evaluate this equation

$$P(l^1) = 0.502$$

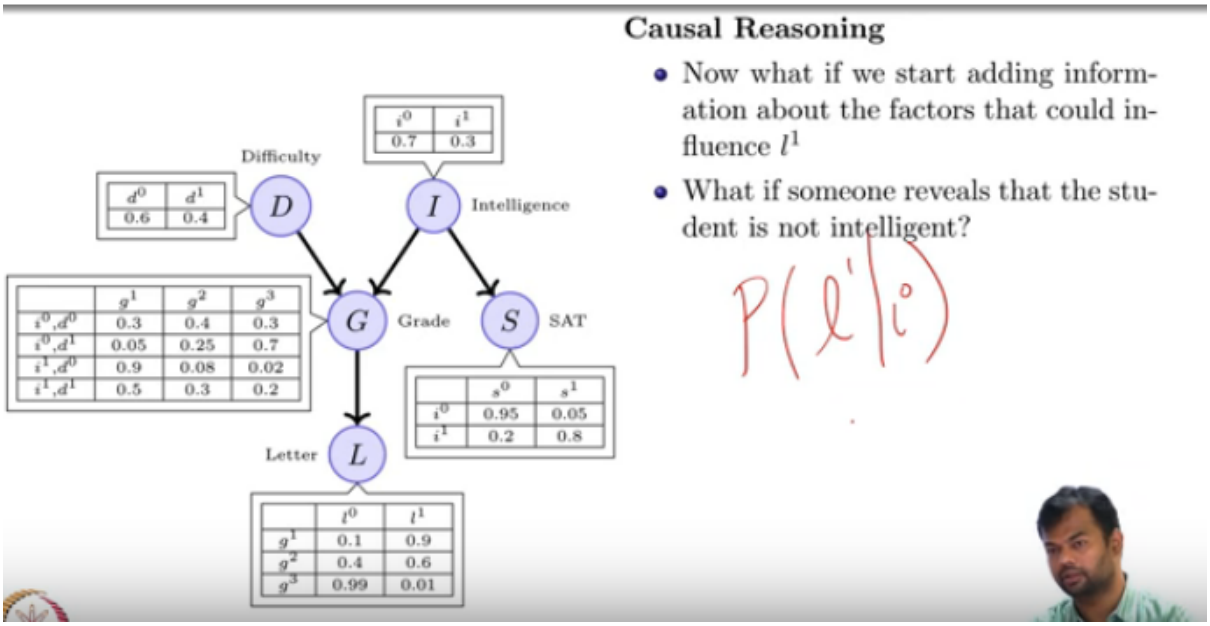


	l^0	l^1
g^1	0.1	0.9
g^2	0.4	0.6
g^3	0.99	0.01

	g^1	g^2	g^3
i^0,d^0	0.3	0.4	0.3
i^0,d^1	0.05	0.25	0.7
i^1,d^0	0.9	0.08	0.02
i^1,d^1	0.5	0.3	0.2

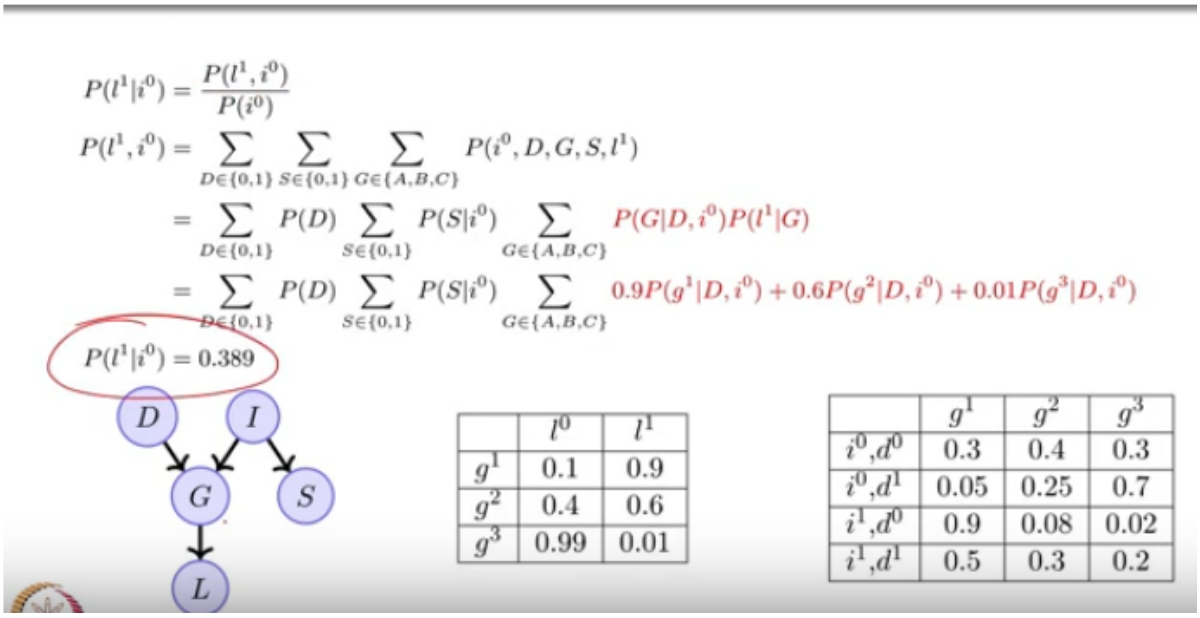
It so, I will just keep doing that so let's, so this is a simplified form that I get because, **the last Y only** depends on the grade, case I have just I have written out the full chain rule and, then I just simplified the chain rule based on the independence assumptions that we have so, far clear .Okay? Now, let's look at the last form that we have and we'll focus on this guy, the red guy .Okay? Now how do you compute this, you consider all possible values of the grade so, can you compute this last quantity what will you have to do? How many of you find this plain obvious I mean please raise your hands, please is them high up .Okay? Not many why is it so, I mean this was just meant for the sake of completeness I did not, really intend to go over this, can you compute this or not, everyone can compute this how many if you can compute this. probability from the given tables please .Okay? good so, then we'll not really go over this because that's not the main point of the discussion so, it turns out that quality of getting a good recommendation letter through all this compute and, this compute depends, on these five tables that you had turns out to be 0.5 0 2, what you can do is, you can go back and do this computation and see whether you get this value .Okay? just I mean for those of you I see some doubtful faces, I'm not sure why those is not very difficult but, you can just go back and check for now, it's suffices to say that after this computation you left with the following value this is the probability of getting a recommendation letter is 0.5 0 2, what I want you to.

Refer Slide Time :(4:32)



Appreciate is that if you look at these tables, it was not directly obvious from anywhere that probability of getting a highlighter good liquor recommendation later 0.502 but, after doing all this computation involving these five tables you can arrive, at that value that's the only take away from what I have done so far .Okay? But, there was no causal reasoning there .Right? it was just that what's the probability of getting a high recommendation later but, typically we are interested in the questions of the following form, what if I told you that the student is not intelligent, what am I actually asking you, in terms of probability what is the priority that I'm asking you, probability of l1 given a zero .Okay? Now can you compute this probability, can you, what would it look like it would be something, over something, or something into something .Okay?

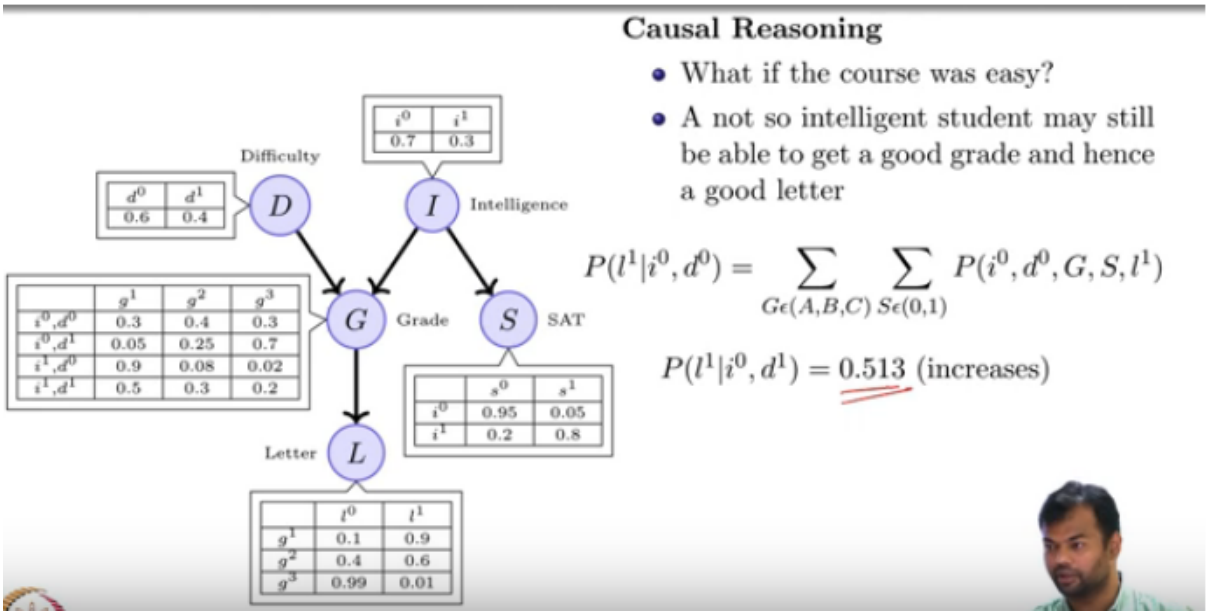
Refer Slide Time :(5:27)



So, this is how you will do it, I'd you'll take the Joint Distribution and, divided by the marginal distribution that's simple .Okay? And now, each of these we can compute, from the Joint Distribution

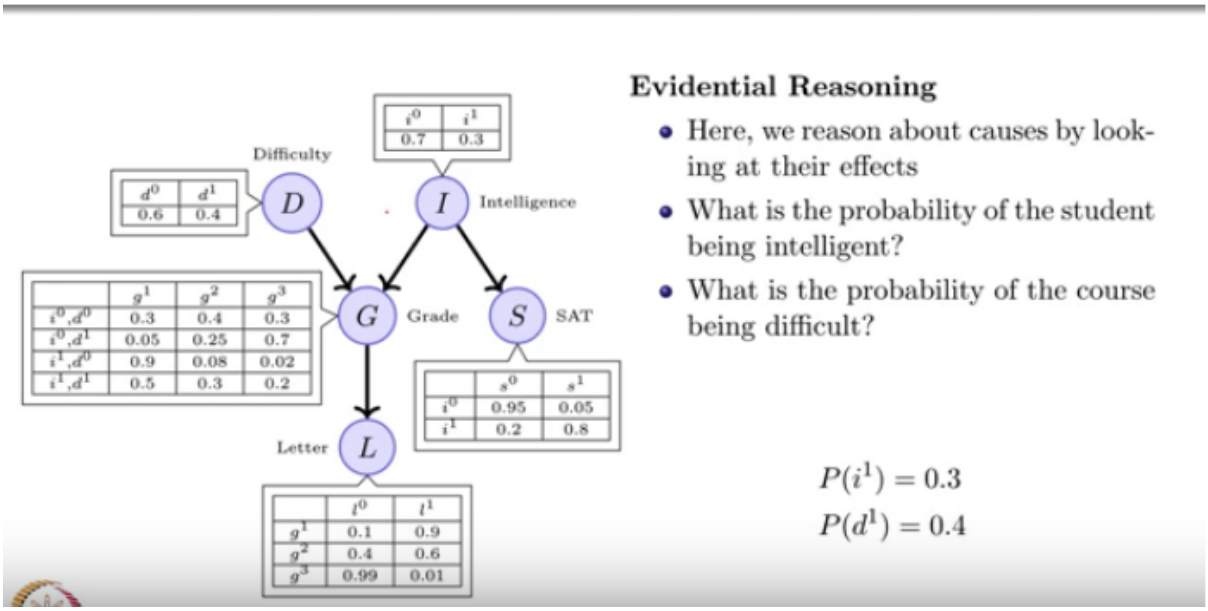
by, marginalizing over whatever variables are not of interest .Okay? And this is how you would get it and, now at the end of it, again this computation is not important what happened here, what was your initial probability of someone getting a high good recommendation later now, what has happened .Right? So, this kind of reasoning is important it you would want to know, that if I change a certain factor, now think of it in terms of sports or anything .Right? if you say that this particular player is, going to play today or not play, then how much does the chance of winning change, of course I don't know why anyone would be interested in knowing that chance, for any good reason but, why would you be interested in knowing that or you are into betting is it .Okay? Good to know that .Okay? So, I mean for various reasons like all for more practical things that if someone has this symptom, man if that symptom was not there .Right? Then what would have happened or if the symptom is added, what would have happened and so on .Right? what's the probability of a certain disease, increasing, or decreasing and so, this kind of causal reasoning where you change some factors in your, where you set some values for variables in your joint distribution and see what how, it affects variables which are downstream why do I say they are downstream, well these variables depend on the variables that you are setting in particular .Right? In this example intelligence actually determines grade, which in turn determines the letter of recommendation .Right? Hence when you change something in the intelligence you will see an effect in the letter of recommendation that's fine .Okay? And that this kind of reasoning is very important the graph gives you a nice way of finding these dependencies and, the factorization gives you a very efficient way of computing these probabilities .Right? .Okay? so that is that any kind of thing, of any other kind of reasoning here, this is causal reasoning what are the kind of reasoning that could be

Refer Slide Time :(7:29)



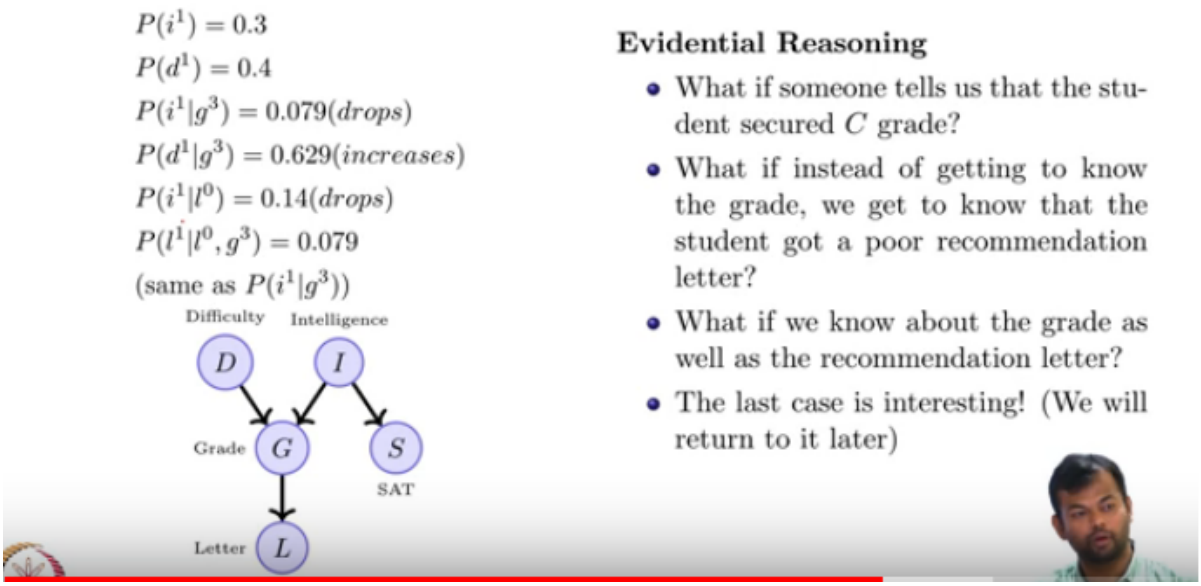
Now, there's several other types of causal reasoning that equal to .Right? Now what if the course was easy, what would happen your p of l1 was, 0.5 0 2, now what am I asking you to compute P of l1, given D 0 .Right? Easy is 0, what do you expect to happen? Increase or decrease? Increase, that was of course was easy probably the student got a better grade hence he or she got a better accommodation later .Right? So, all of these things are did not unfortunately increase by, a lot but still increase by something. Okay? So, all of these are important reasoning that you would want.

Refer Slide Time :(8:10)



To make in various real-life situations now, the other kind of reasoning is the reverse of this, this was causal reasoning and, the other kind of reasoning is evidential reasoning .Right? So, here we reason about the causes by, looking at their effects, can you give me an example now, probability of someone being, intelligent given that, that person got a good or bad recommendation .Right? .Right? That's also important .Right? So, know you know, you have seen some causes, you have seen some effects that this person has a certain set of symptoms now you're trying to reason about the different causes that could have caused that symptom so, it's again very important to do this kind of reasoning so, you know? That say the probability of someone being intelligent if you do the same meagre calculation you will come up to 0.3, oh no you don't need to do a meagre calculation this is directly sorry Okay? This is because, this just does not depend on anything else you can directly read it off from the table so, probably of someone being intelligent is 0.3 but, now if I ask and probably of the course being difficult is point 0.4.

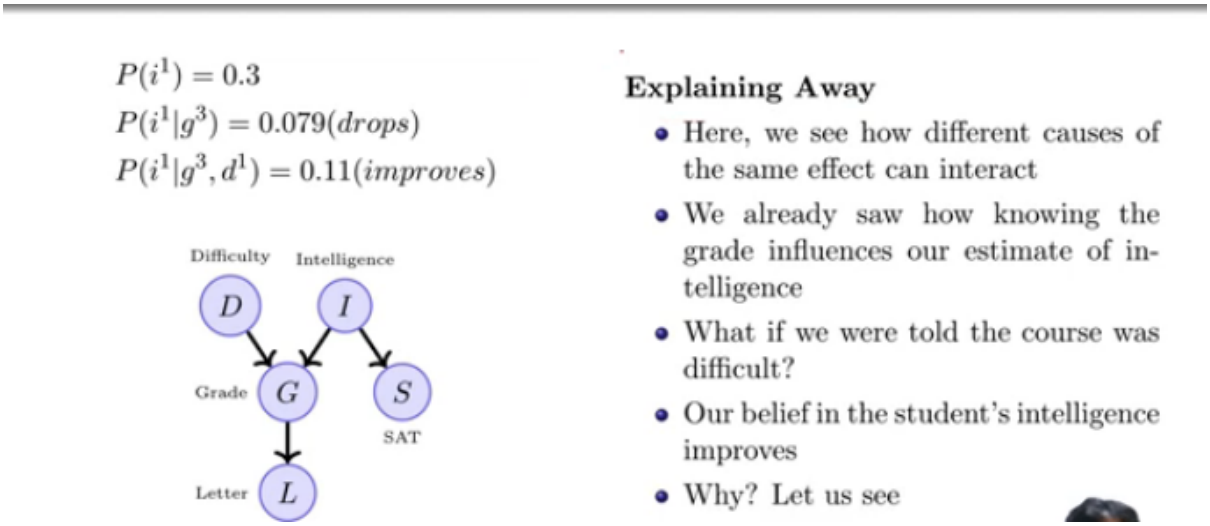
Refer Slide Time :(9:14)



What happens if you observe some effects, what if someone tells us that the student secured a c-grade? Okay? Now, what is the question that I'm asking actually P of, i1, given G is C .Right? That's the priority that I'm asking so, what would happen? What's your guess, first of all can you compute this priority from the given set of data that I mean, from the given distributions so, that is something that you need to be very computer comfortable .Right? if you have the Joint Distribution either in its explicit form, or the factorized form, I can ask you all sorts of conditionals and modulus involving the random variables for which you have the Joint Distribution and you should be able to compute all of them, Hey it just involves using some conditionals and some marginal's .Right? That's all it boils down to .Okay? So, any question that I asked you, you should be able to do it, for example in a quiz .Right? so .Okay? So, what, what would happen? If the student got a C grade? The priority of that student being intelligent would probably drop, what about the difficulty of the course, it will increase .Right? so, these kind of this is now evidential reasoning because you're looking at some call and then trying to see what would have happened to the factors which could have influenced this cause what if instead of getting to know, that the student got a poor recommendation letter, suppose you know that the poor are storing at a poor recommendation letter, then what would happen? What would your assumption about the students intelligence be know, that the intelligence at least does not directly seem to influence the recommendation letter .Right? But, there's still a path through which this inference can flow .Right? So, you would your plain English reasoning would be oh he, or she, got a poor recommendation letter, which means the, grade was poor which means the, intelligence was not high .Right? So, again that is what would happen? The probability that the student is intelligent foot .Okay? So, all sorts of and these and other examples, we could do it and .Okay? Let's see now suppose we look at this case and this should have been I won no sorry, now suppose we know the grade as well as the recommendation later .Okay? When we knew the grade, and we knew the grade was bad, we saw that this drops from 0.3 to, 0.079, I'm saying that in addition to knowing the grade you also know, that the student got a bad recommendation letter, we still end up at the same value why is it so, how many forget that given the great, intelligence and the recommendation letter are independent, once you know the grade it completely decides I don't really need to go back and check whether the student was intelligent or not it doesn't matter because, irrespective of whether he, or she, was intelligent as far as the greatest a or, B or, C that completely determines what the recommendation letter is going to look like .Right? And this case is actually very interesting and we

will return back to it and when we talk about what are the different independencies encoded in a Bayesian network .Okay?

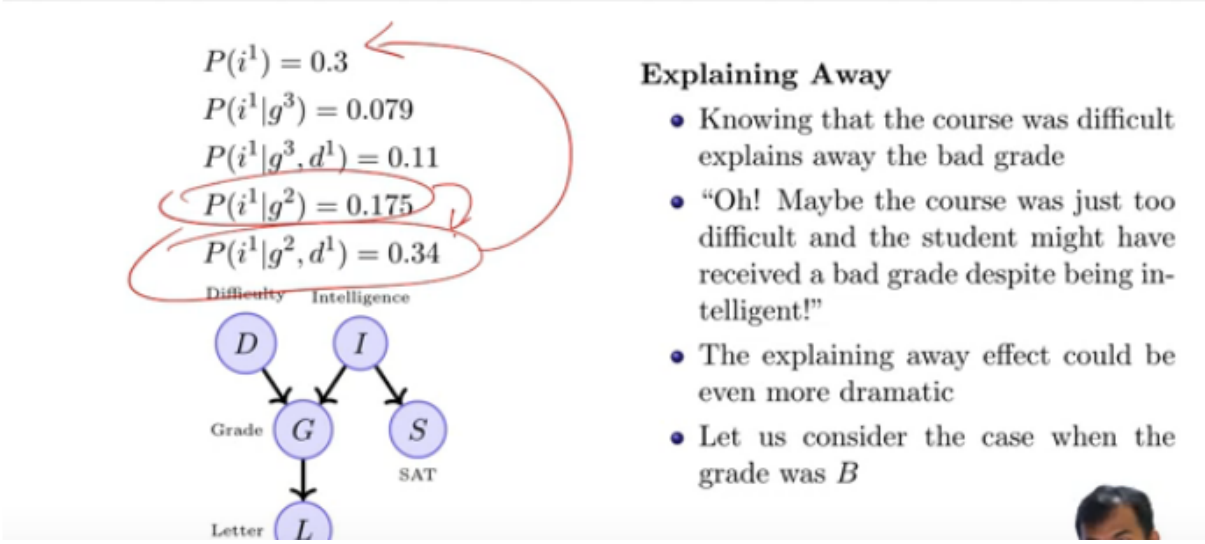
Refer Slide Time :(12:15)



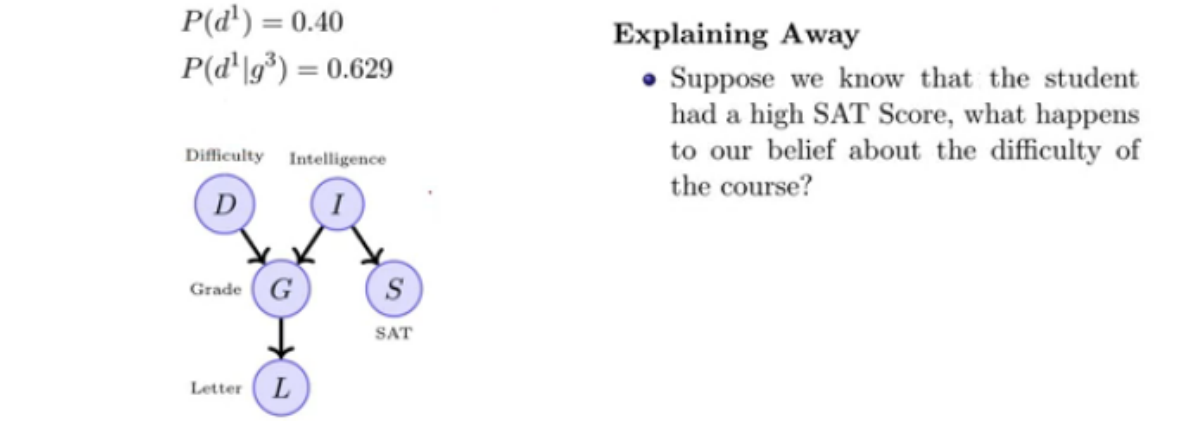
Now, the third kind of reasoning is known as explaining away and this is again very interesting so, here the idea is to see how, different causes of the same effect can interact with each other .Okay? So, we already saw that this happens, if I know the grade the probably of the student being intelligent drops .Okay? Now, suppose I was also told that the course was difficult what would happen now? All of what all of you get the question what am I asking now P of i1, given G 3 comma, D 1, what would happen? What will increase? The priority of the student being in why would it increase? grade and I mean the intelligence and difficulty have no relation with each other and irrespective of the Bayesian network or whatever if the intelligence of a student and the difficulty of a course has no relation with each other I mean I would independently think of setting them making the course as hard as possible irrespective of what's the level of the intelligence of the students and so, on .Right? So, why is it that knowing, the difficulty level of the course, should actually influence intelligence that doesn't make sense to me, what is happening here?

So, I believe actually improves .Right? It goes up as most of you said so, why is it happening? Actually can you give me reasoning for that .Right? So, that's what is known as explaining of you .Right? So, there are various factors in this case two factors, which could have caused the grade to be low .Right? Now I've given you one explanation that the grade is low because, the course was difficult .Right? So, that explains away why the grade was what it was .Right? So, now the intelligence does not have to be low to explain the poor grade, that's why they your estimate of what the intelligence is this case increases, then it increases by, some amount it's not, that it's going to become greater than 0.3, but it improves from where it was if you knew, only that the grade was poor now that, you know that, the grade of poor and, that course was difficulty it makes sense that the part, of the responsibility for the course for the grade being poor lies with the difficulty of the course and not solely with the intelligence of the person .Right? That's why your belief about the intelligence improves so, this is known as explaining away

Refer Slide Time :(14:35)



Refer Slide Time :(15:42)



So, this is what you reason .Right? "Oh! maybe the course was just too hard and the student may have received a bad grade despite being in tests and that's exactly what this number is telling you wait that's the English explanation for that and, the explaining away if it could even be more dramatic great so, let us consider the case when the grade was B actually so, now I give you i1, given G 2 and, I give you i1, given G 2 comma, d1 Okay? So, why I say this is dramatic is because 1, is almost a double increase, in your estimate of the probability once you know, that the course was difficult and now your estimate is actually better than the default probability that you would have been assigned to someone so, these are all different types of reasoning, that you could do with a Bayesian network, given the Joint Distribution and, given its factors, irrespective of whether you factorize or not all this reasoning can be done .Right? That's one thing that I want everyone to understand whether you have the factors or not, you can do this reasoning even if I give you the explicit Joint Distribution, having the factors just improves the computation and also make sure that the number of parameters that you are dealing with is much smaller. Okay? So, these are the two things that you achieve with factorization and I'll just not go into any more examples .Okay? Thank you.