Artificial Intelligence

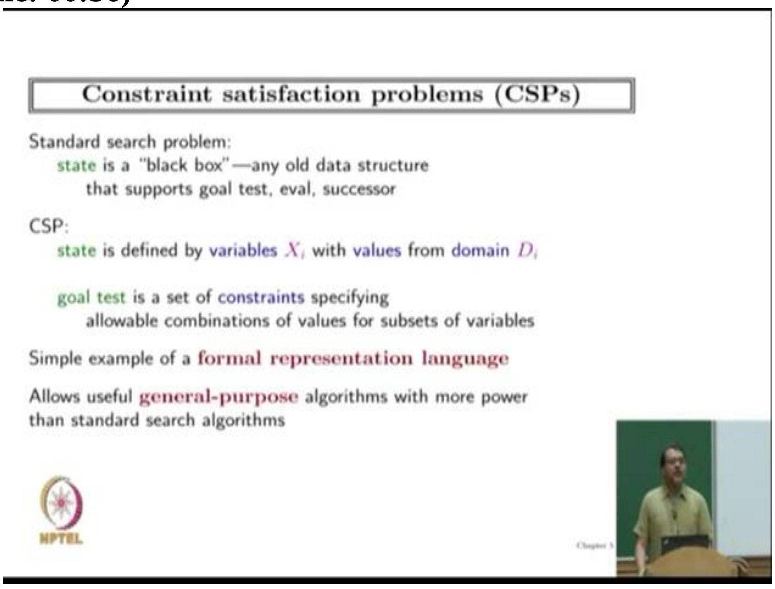
Prof.Mausam
Department of computer Science and Engineering
Indian Institute of Technology – Delhi

Lecture - 6
Constraint Stratification Problems: Variable and Value Ordering in Backtracking Search, Part – 4

So, we have been discussing constraint satisfaction problems and I give you a one minute summary of you know what we have been saying so far. So, we started by saying that we need to go inside the state and we need to define a representation, any representation that allows us to have a nuanced understanding of what is happening inside the state and one such representation is the constraint satisfaction problem representation.

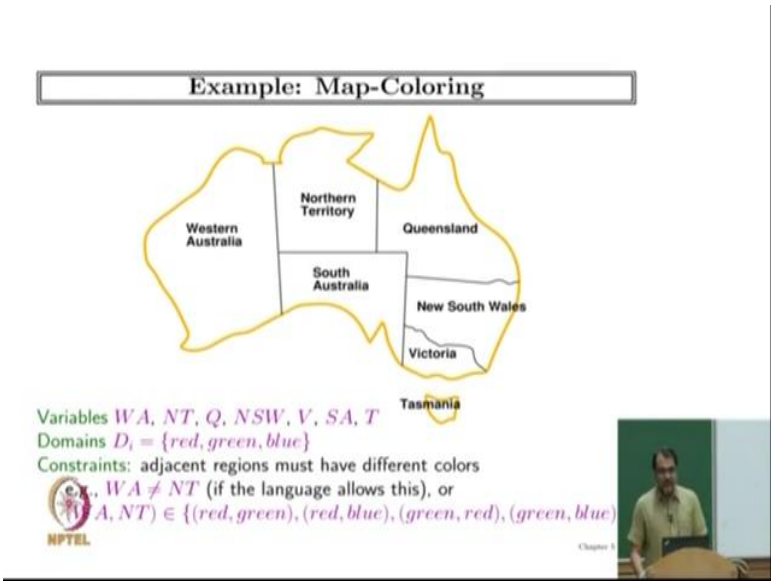
It is a very simple representation it says that each state has a set of variables and variable state values and for now, we will focus on discrete variable so that means it takes discrete values and values are from the domain Di.

(Refer Slide Time: 00:56)



So we have a variable Xi which can take a value from domain Di and then we have constraints and these constraints may be written in some of the presentation also, let us say I allow inequality so for example, in the case of map colouring

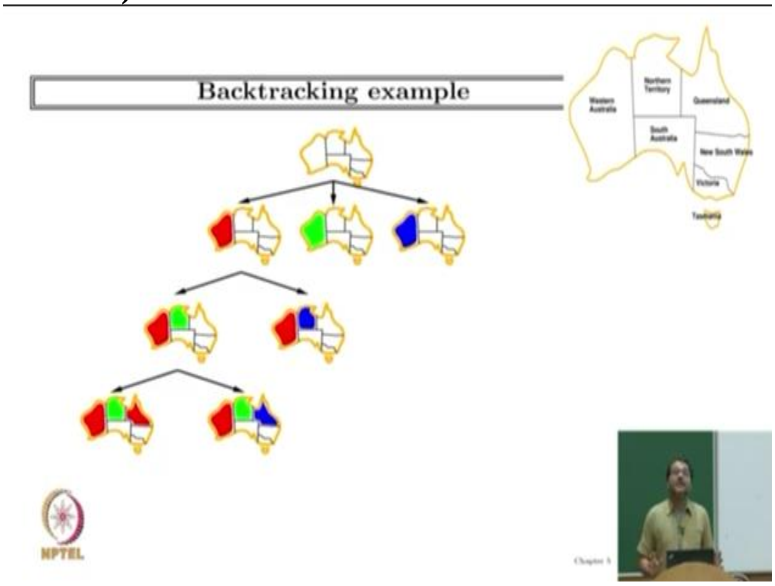
(Refer Slide Time: 01:10)



Let us say we are trying to cover the map of Australia, the constraints would be the 2 adjoining States cannot be equal to each other. For example, Western Australia is not equal to Northern Territory and we have been taking this as an example. So in this case, we have 7 states Western Australia on the west, Northern Territory on the north, South Australia on the south and on the east we have Queensland, New South Wales.

And Victoria and in the as an Island state, we have this one. Let us say you want to colour them using red, blue and green and so those are our values so we discussed that and we then define a backtracking search algorithm.

(Refer Slide Time: 01:51)



And the backtracking search algorithm essentially said that I will start from an empty state, I will start from an empty assignment. No variable has been assigned any value and then I were start to search you must have discussed any problem solving from an AI point involves some kind of search and so we are going to use the search where we will decide on one variable and we discussed last time that the order of the variable sort of does not matter.

And so we can decide variables in any order, it can be a dynamic order, but at any given node, we will only choose one variable and branch on that and not choose different variables and branch on different variables. Because if we do that it will be unnecessarily exploring permutations of variables which is not important.

So, we will choose a variable let us say we pick Western Australia and then we will branch on all values of it so in this case, we branch on 3 values. Western Australia being red, green, and blue and say we go to red node and then we again branch and let us say we branch on Northern Territory and so this is how we branches notice that there is no point on branching on Northern Territory being red.

So even if you have branch on Northern Territory being red the state is going to say, constraint violated backtrack and whenever a state says constraint violated backtrack, your going to backtrack and this keeps going. For example, at some point you may branch on Western Australia red, Northern Territory green and you know some value for Queensland now just intuitively look at this and tell me if there is any state you think cannot lead to a solution which one cannot lead to any solution?

The right bottom which is Western Australia red, Northern Territory green and Queensland blue does everybody see this? Why it not lead to a solution would because South Australia has no value the meaning all its values are gone. However, at this present moment of time, my algorithm does not know that there is no solution there and that is actually an important bit for you to think about.

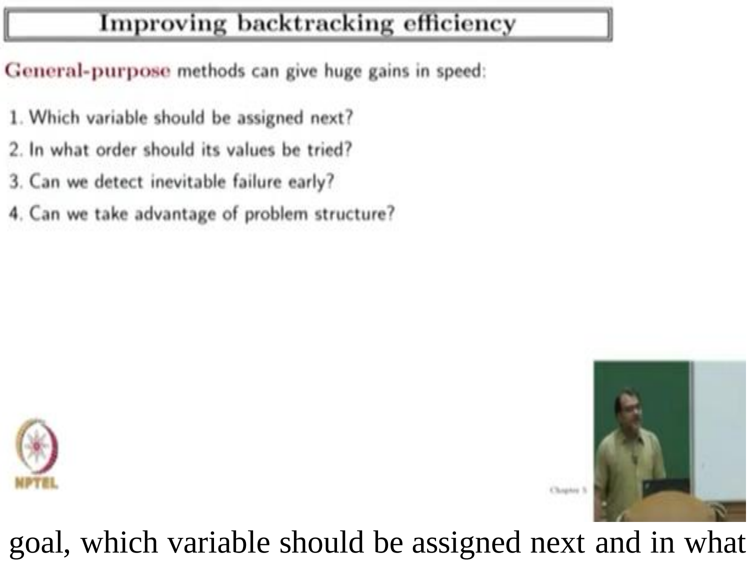
This is called the idea of detecting failures early if you look at the state with Western Australia red, Northern Territory green and Queensland blue, we do not have any value remaining for South Australia, even though we do not have any value remaining for South Australia and there is no way that the search can proceed and lead to a solution at this point in time, we do not know this or at least the algorithm does not notice.

The algorithm thing that will keep searching and at some point it might find a solution and of course you and I notice we can quickly see and say up to just no point in searching. So, what we did is what is called inference? We made an inference internally in our brain and said you know, there is no point going forward because they will be a failure and that was we made up we actually proved it, we have a proof.

If you think about what is search? Search is let us try, let us try this option and let us try that option. Let us try that option, at some point we will succeed what is inference, we are saying we have a proof that we should not go forward because it will not lead to a solution or we have a proof alternatively, that we should go forward because it will absolutely lead to a solution and if you if I can do this, I do not need to do search.

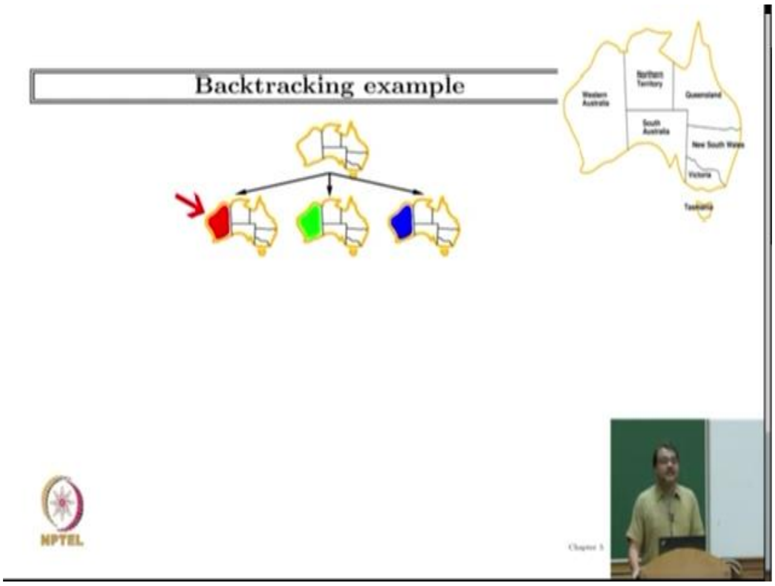
Because I know exactly what value to give after inference and so in todays lecture, we will try to understand the dichotomy between inference and search. But before we do this before they get too interesting inference we have 2 more things to discuss what are the 2 more things to discuss first let us say we are trying to run the backtracking search algorithm. Then how do we pick which variable to set first and how do we pick which value of that variable to set first

(Refer Slide Time: 06:20)



So, this is our first goal, which variable should be assigned next and in what order should the values be tried? Let us first answer these 2 questions these are easy enough questions and then we will talk about detecting failures. So, I need suggestions from you which variable do you want to assign next for example let us say that I am at

(Refer Slide Time: 06:45)



Let us say this variable, this node, what node is this? Let us say Western Australia has been assigned red nothing else has been assigned so far, which variable do you want to assign next? I need a suggestion. What is your name? misord, yes misord immediate adjacent states no notion of adjacency is a notion of map colouring. But let us think about variables and values here they are general variables.

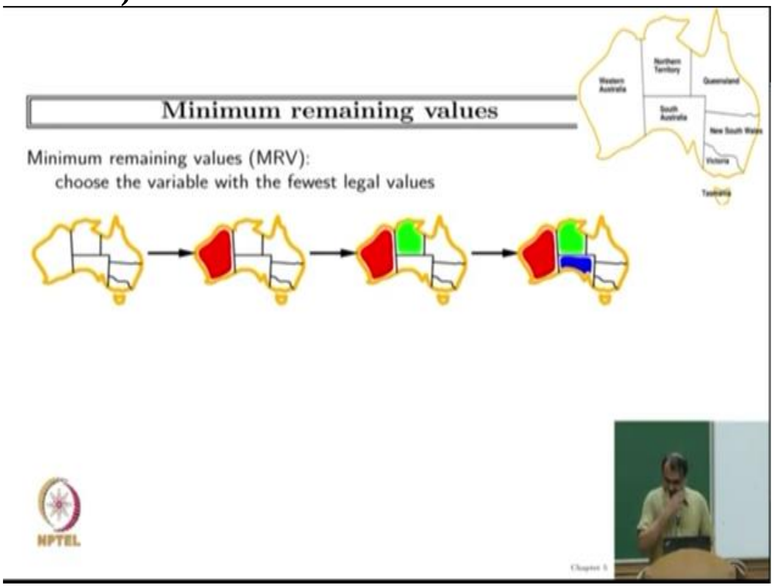
The general values and you have some constraints. So, which variable do you want to assign next? Yes, what is your name? You want to assign South Australia next why? It is it has maximum constraints. What is your name? Yash, Yash says I want to assign South Australia next because it has maximum constraints, maximum constraints. So, intuitively the idea is that any variable which is most constrained should we assigned first it is like this.

Suppose you know, the administrative person needs to schedule a meeting between me and the director. Who does this administrative person asked first? Me other director, you guys are too nice. But seriously you will not get a job as an administrative person if you did that. So who do you assign first? The director and let us say there was a meeting between me and shahrukh Khan who do you ask first?

Yes, I mean, they do not know that I am actually busy shahrukh khan. So, who would you assign for as the busy person or well known person? Now you get the right answer. If you had said the well known person and I would know what you think of me as, so yes, so you always want to check the busy person first. Why? Because they have too many more constraints, it is possible that you know I say let us meet wednesday afternoon.

And then the director said I am busy then I say let us meet wednesday morning, the director says, I am busy. So in that case you know, I do not the director say when he wants to meet and then I will see if I can say, I am busy or not. So, that is the point now, what are the indicators? What are the various indicators for how busy a variable is or how constrained a variable is? Very good one is what he has said. Let the degree of the variable and there is one more that is very important. Yes, what is your name? Baadh set of possible values set is not enough size, size of possible values.

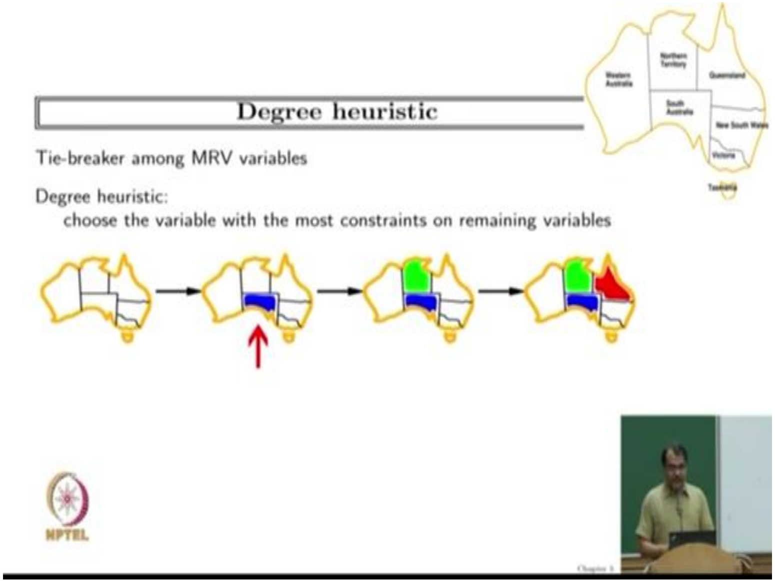
(Refer Slide Time: 10:00)



So, this leads us to the choice of variable ordering it is that we should first assign the variable which has the minimum remaining values. For example, if I have assigned red to Western Australia and green to northern Italy, then I must assign South Australia next because it has only one value when value the remaining. If I have assigned Western Australia initially I can assign any one of us to Northern Territory and South Australia.

Because both have 2 variables 2 values remaining so, this is called the minimum remaining values heuristic and then sometimes this minimum remaining values heuristic gives the same number for multiple variables.

(Refer Slide Time: 10:45)



In such a case, we use Yash's idea and we have a tie breaking heuristic called that degree heuristic. The degree heuristic says choose the variable with the most constraints on the remaining variables. Notice that this has to do at every step of the search so, we have to talk about the remaining variables remaining values and so on so forth. So for example, if I actually use minimum remaining values and tie breaker degree heuristic.

I will start with South Australia because it has the maximum degree and therefore, assigning it would make our life simpler and then after we assign South Australia now, that degree the values of all the other variables there are only 2 so again degree heuristic comes with a picture. So the degree heuristic says assign the one which has the maximum degree.

Now there are many which have degree 2 like Northern Territory, Queensland, New South Wales they all have degree 2. So, now you can do the random diversity so you assign Northern Territory first. Once you assign Northern Territory now you have to talk about minimum remaining values of Western Australia has been the winning value 1, Queensland has minimum remaining value 1.

So you assign one which has the higher degree between the 2 so between the 2 Queensland has a higher degree because it has a degree of 1 with New South Wales whereas Western Australia has degrees 0 with the unassigned variables and so you assign Queensland next and this process goes on. So, this is my variable ordering heuristic it is called a heuristic and not an algorithm.

The algorithm is a backtracking search algorithm. The heuristic and this is not realistic function of admissibility, this is different, this is a heuristic this is a way to the order you could have heard in any which way and different thing will lead to valid algorithm but by choosing an order perfectly or not perfectly but at least intelligently you will be saving on the search significantly.

So, now the next question is which value do assign first? Now, you may not get this from this example. But intuitively which value should I assign first? Value, which is shared by least number of people is one suggestion? Again, it is a very good suggestion for map colouring, but we have to do something general we have to talk about general variables, general values, general constraints and so on and so for that.

So what is the generalization of that particular idea? Value, which is shared by least number of variables? What is the generalization? What do you mean the value which can be satisfied? Again, you are assuming that you have only inequality constraints, we may not only have inequality constraints, so even your ideas that are not you are not able to generalize from map colouring to a general constraint satisfaction problem.

So the intuition is right if I have not used read so far, let me use red or if red has been assigned the least number of times so far then let me use red, that sort of intuition. What is your name? Hass, that sort of the intuition Hass but at this point, we need to go beyond that, because we may not only have colouring problem, we not only have inequality constraints and so what would you how would you generalize this, which value?

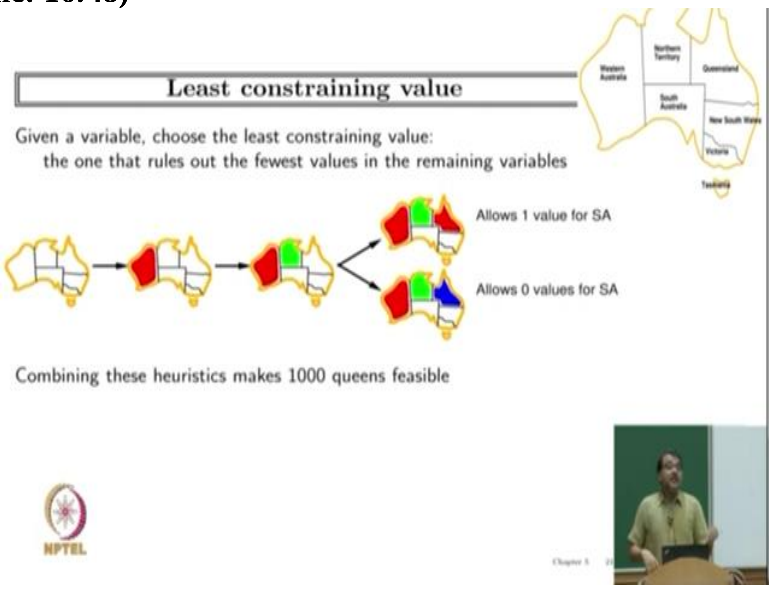
Do you try? What is your name? Chinmayi, yes Chinmayi, in the least recently used again, the same intuition you are saying least recent is better. Because I have let us say everybody has a different value. Every variable has a default value, and there is some inequality constraints like $x1=y1+1$ and so on so for that. I mean we are talking about general constraints or restriction not only map colour, but we are given constraints.

The value that can sort have fit in into large number of constraints that is appropriate. What is your name? Saigon so Saigon says values that can fit in into large number of constraints, let us think about what is the meaning of fitting in because if a value does not fit into some constraints, then it is a failure. If there is a value that cannot fit in into a constraint that does cannot satisfy a constraint, it is definitely a failure.

We need to move back we need to backtrack, but your intuition is form. So how do we operationalize this whole idea that the value has to sort of fit in into more constraints, but what does that mean? What does that mean? You are saying that after adding the value, we should narrow the search space that means that we less values for other variables and if you have less remaining values for the variables what is going to happen?

We might get into a failure, we may not have a solution because we have constrained the problem more, if we want a solution, what do we want? We want to have maximum values remaining for other variables. I understand your intuition path, but think about it if we narrow our search space, our search will move faster, but it is possible it will fail. We always know all constraines think about it. Let us discuss if there is a confusion. So the heuristic that comes out is called the least constraining value, the value of which constraints the other variables the least.

(Refer Slide Time: 16:48)



You have some variables that have been assigned, let us get away from them you have some variables that have been assigned. For all those variables I have picked one variable. I want to assign this I have different value, now, as soon as I assign a value, other variables may find some reduction in their value space and if I find that there is one value, which reduces everybody is space domain, possible values much more, and another space.

That reduces their domains less then we will pick the one that reduces their domain less in the hope that we will find a solution in this sub tree and backtrack what you said path would be better that we would take the one that is most constraint, but we do not want to succeed. So we will pick up every say it is like this you know that more faculty members are free on wednesday afternoon because they are no classes.

So now you talk to the director you said we want to do a meeting the director says do you know I can meeting on monday morning or wednesday afternoon, which value to try first. You try first the value which is wednesday afternoon, you try the value which most likely looks like least constraining to other people. Most likely the faculty members would not have a class at the time they will be free, so let us try this.

So even initially everything is equal so we put Western Australia red, then when we are assigning Northern Territory, we assign it green and then when we are going to Queensland, let us say now we figured out which one is least constraining. So by doing blue here, we have no possible value remaining for South Australia and by doing red here we have one possible value remaining for South Australia and this one is less constraining.

So let us try that first and if you try that first, then it will not fail you will actually find a solution. Now just combining these 2 heuristics remember that earlier we said that backtracking search can get to 25 queens, just by combining these 2 heuristics minimum remaining value with degree heuristic and least constraining value gets us 1000 queens problems. Notice by just small modification, we have done such a huge jump in our ability to solve constraint satisfaction problems.