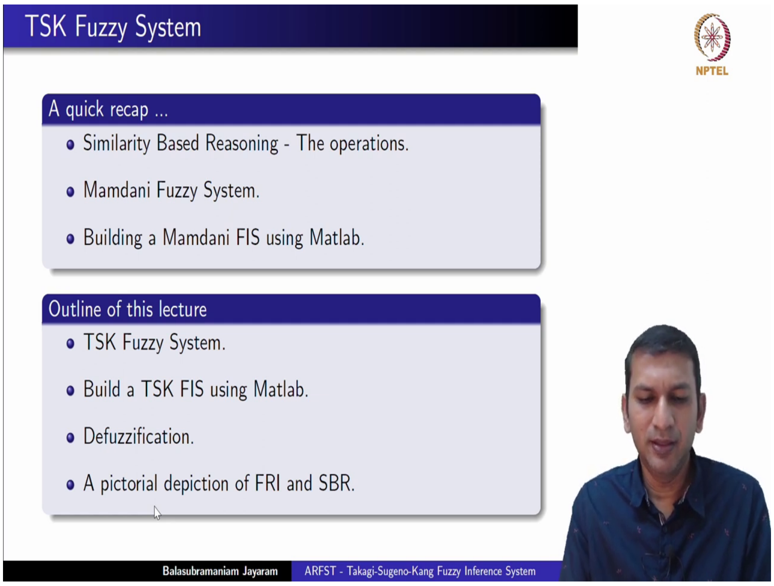
Approximate Reasoning using Fuzzy Set Theory
Prof. Balasubramaniam Jayaram
Department of Mathematics
Indian Institute of Technology, Hyderabad

Lecture - 38
Takagi-Sugeno-Kang Fuzzy Systems

Hello and welcome to the last of the lectures in this week 7 of this course titled Approximate Reasoning using Fuzzy Set Theory. A course offered over the NPTEL platform.

(Refer Slide Time: 00:33)

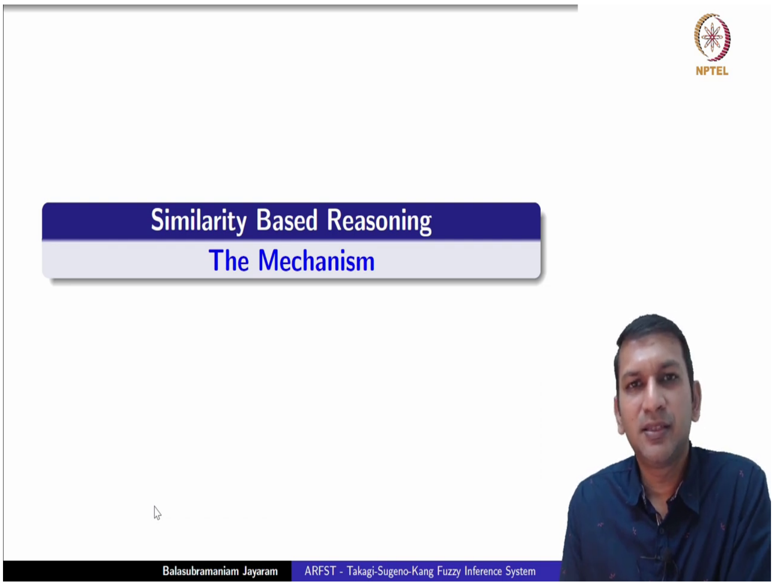


During this week of lectures, we have concentrated on similarity based reasoning. We have seen there are two major types of similarity based reasoning that of the fuzzy inference system proposed by Mamdani and Assilian which is typically called Mamdani fuzzy system. We have seen how to build a Mamdani fuzzy inference system using the fuzzy logic toolbox in Matlab.

We have approximated some specific given mathematical functions and we have also seen how you could build a Mamdani fuzzy inference system in a given practical application. In this lecture we will look at the second of the major similarity based reasoning inference schemes that of the Takagi Sugeno Kang fuzzy system. Well once again build a simple TSK fuzzy inference system using Matlab fuzzy logic toolbox in Matlab.

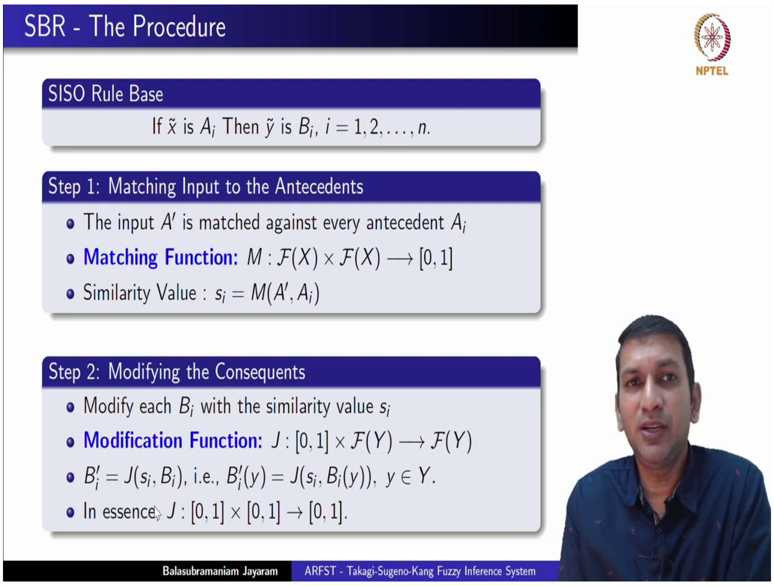
We will also visit the defuzzifier the defuzzification process briefly and finally, we will give a pictorial depiction of both the fuzzy relational inference mechanisms and the similarity based reasoning schemes perhaps you could then easily see some of the similarities that exist.

(Refer Slide Time: 01:53)



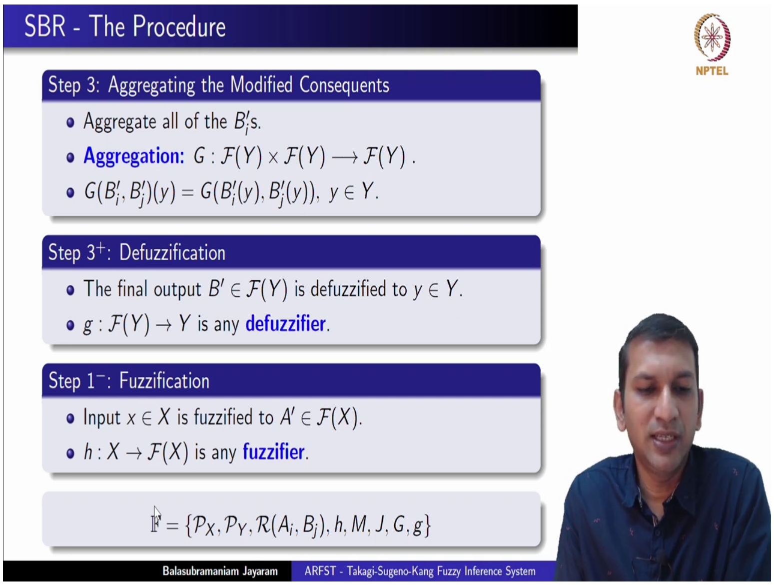
Let us revisit the mechanism underlying the similarity based reasoning scheme.

(Refer Slide Time: 01:59)



So, we are given a single input single output rule base of this form if x tilde is A i then y tilde is B i. In the first step given an A dash which is a fuzzy set on x we match this A dash the given input to each of the antecedent A i using a matching function and obtain the similarity value s i which is typically in the interval [0,1] in the range of [0,1]. The second step we use the similarity value to modify the consequence of each of the rules. So, given s i and B i we modify the consequent to a B i dash which is again a fuzzy set on y.

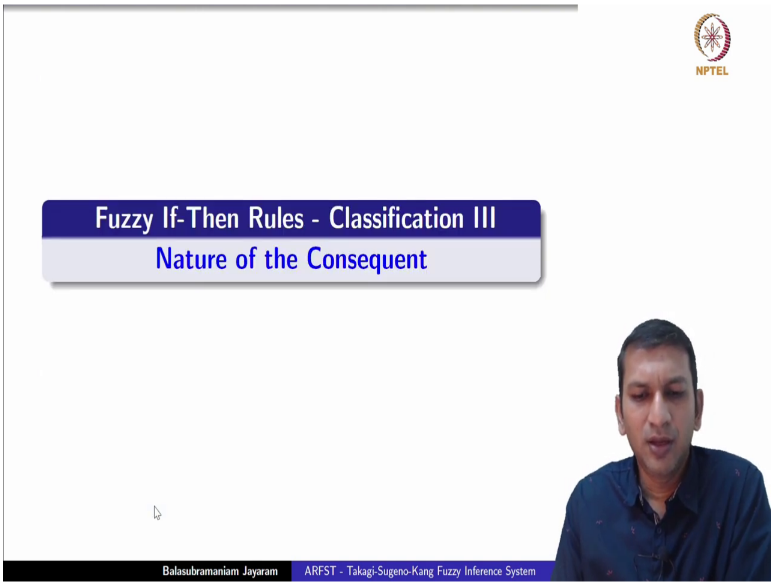
(Refer Slide Time: 02:40)



Now, we obtain such modified consequence B i for each of these is and the third step we aggregate all of them to a single B dash. In the final step if required we defuzzified; that means, map this B dash which is a fuzzy set on y to some value in the domain of the set y. There is also a pre-processing step that is often we may be given an input which is a real value not a fuzzy set, then in that case we apply the fuzzification operation to obtain the A dash the given the input to be given to the system.

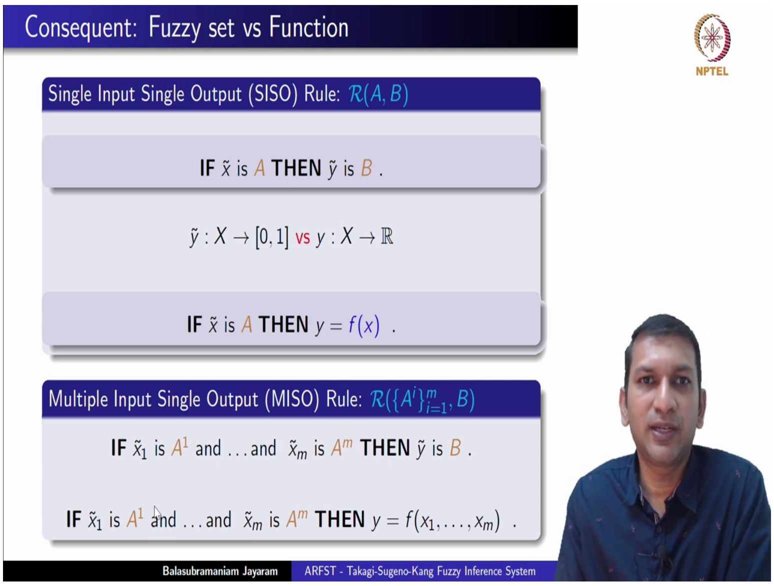
We have seen these five steps in depth. Note that when we need to specify a similarity based reasoning scheme these are the parameters we need to specify P x and P y are the fuzzy coverings on x and y R of A i B j are the rules A is are the antecedents which come from P x B js are the consequence which are picked from t y and these are associated to form the rule base h is the fuzzifier, M is the matching function, J is the modification function, G is the aggregation function and small g is the defuzzifier.

(Refer Slide Time: 04:09)



Also recall when we discussed the different classification that you could have on fuzzy if then rules one type of classification we came up with was based on the nature of the consequence.

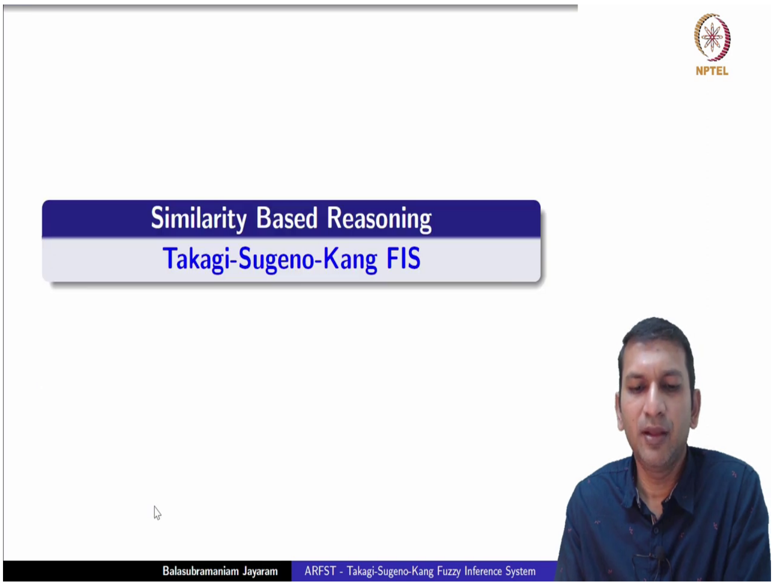
(Refer Slide Time: 04:22)



We have seen that typically fuzzy thumb rules have fuzzy sets on both antecedent and consequence; however, we could also have just a function on x real valued function on x or y real valued function on x as the y. So, in that case the rule would read like this if x tilde as A then y is equal to f(x). So, essentially this y is obtained directly from the given input x.

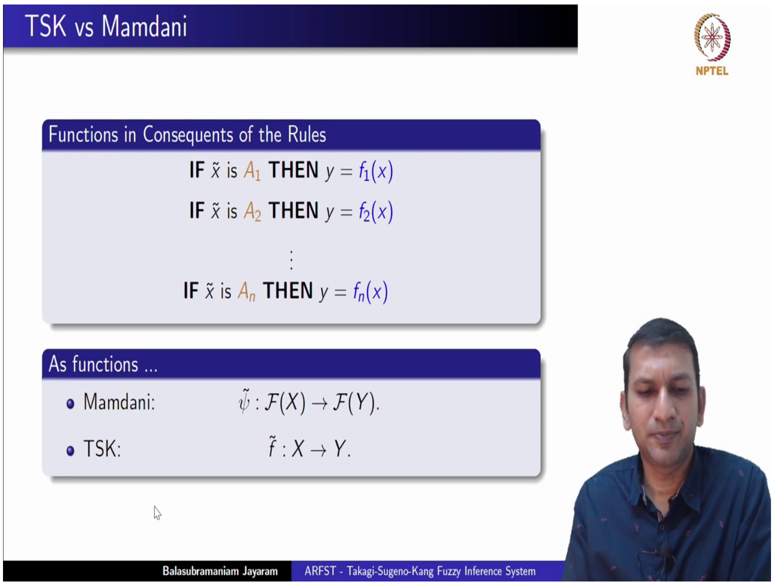
So, this is a function of x, y is given as a function of x typically a real valued function of x and we have seen that the same can happen in the case of multi input single output rule. So, the y is dependent on x 1 to x m the m dimensional input vector.

(Refer Slide Time: 05:23)



Now, let us look at the Takagi Sugeno Kang fuzzy inference system.

(Refer Slide Time: 05:28)

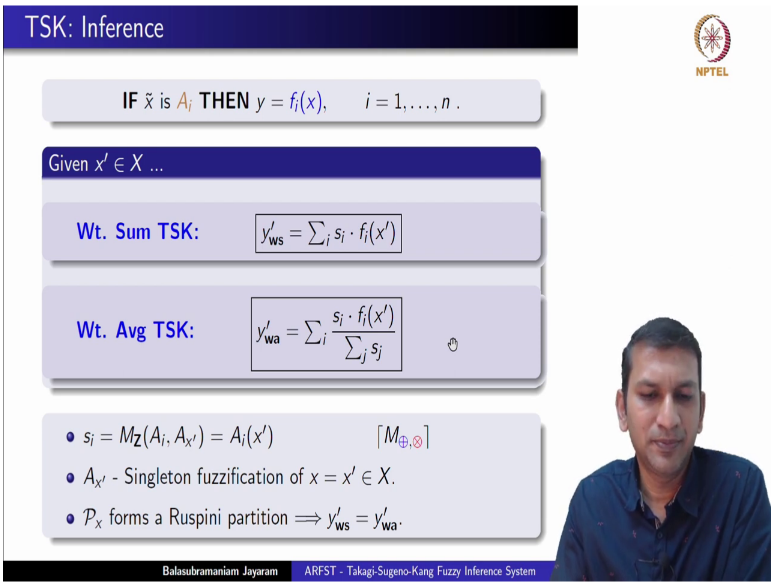


Let us compare it with the known Mamdani inference scheme the first difference between a TSK and Mamdani fuzzy inference system note that both of them are similarity based reasoning schemes. So, the first difference is in the consequence of the rule. Well, in the case of Mamdani fuzzy sets form the consequence in the case of Tagaki Sugeno Kang fuzzy systems what we have are actually functions of x as the consequence.

So, you would have rules of the form if x tilde is A 1 then y is equal to f 1 of x if x tilde is A n then y is equal to f n of x so, on and so, forth. Now, these functions f 1, f 2 so, on till f n they are real valued functions they could be polynomials, they could be linear function, they could be non-linear functions. There are no restrictions on the class of functions to which f is can be low this is the first difference.

Secondly, if you look at the overall inference scheme itself as functions then in the case of Mamdani typically we input a fuzzy set on x and we obtain a fuzzy set on y of course, you could always defuzzified and get an output over the domain y, but typically Mamdani fuzzy systems are looked at like this. So, in that sense they are mapping from F(X) to F(Y); however, TSK fuzzy systems typically are considered as mappings from x to y of course, we will suitably fuzzify it to process it, but you could look at them as just functions from x to y.

(Refer Slide Time: 07:04)



Let us look at the inference itself there lies another differential. So, this is the rule base that we are considering note that the output or the consequent functions y are functions of x. So, now, given an x dash in x, the weighted sum TSK is given like this the output y dash is obtained like this sigma over s i dot f i of x dash. Essentially it is a weighted sum weighted by s i and f i of x dash is the value that f i takes at the given input x dash.

Note that we are actually giving just a value x dash from x. In the case of weighted average TSK all we are going to do is use the similarity values as weights and do a weighted average of these weights with respect to the function values at x dash taken by the consequence of all the rules. So, essentially it is either a weighted sum of the similarity values into the function values at x dash or the weighted average of the similarity values into the function values at x dash.

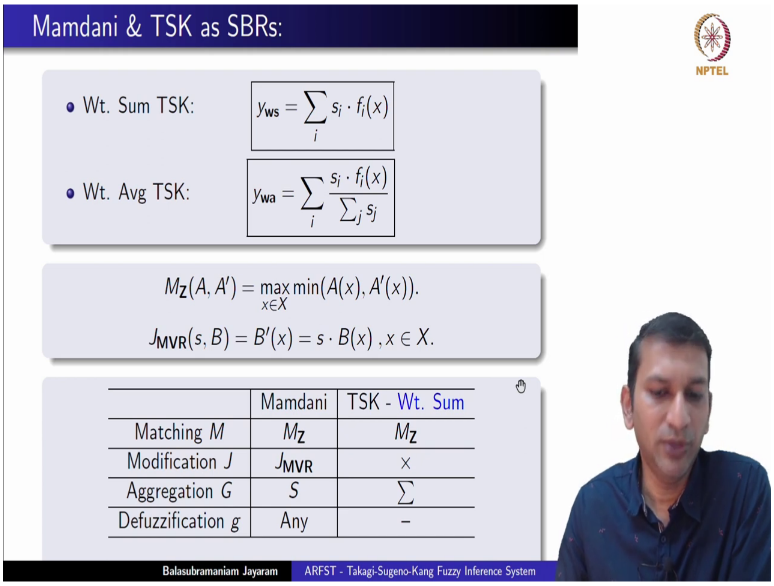
So, note that s is are the similarity values. So, typically you could use the M z function in which case we know that this similarity s i is nothing but the membership value of x dash at A i. Of course, you could use anyone of those family of matching functions where you suitably change these two operations plus and cross. Note that A x dash here symbolizes the singleton fuzzified x dash.

That means, that x is equal to x dash you are fuzzifying its in a singleton way, it takes the of the corresponding fuzzy set takes the value one at x is equal to x dash and 0 everywhere else. And if you are using such a matching function where s i is in fact, A i of x dash. If P x the fuzzy covering on x if it does form the Ruspini partition and if each one of the fuzzy sets features in the antecedent of the room.

So, essentially we have a complete rule base then what is easy to see is that the output from both the weighted sum TSK and the weighted average TSK they are going to be equal. Because in that case what we would have is sigma j s j s actually going to turn out to be 1 because of Ruspini partition these are essentially a j of x dash. So, they will turn out to be 1 in which case these 2 are in fact, identical.

So, TSK inference is pretty simple you first it is classified under SBR because given an input we are finding the similarity to each of the consequence, we are modifying the similarity to the antecedence we are modifying the consequence using these similarity values and aggregating them. But can we really see it as an SBR?

(Refer Slide Time: 10:15)



Well in some sense yes, look at it these are the formulae that we have if it these are M z and J MVR essentially this is a z s matching function and this could be looked at as just the product of the similarity value into the consequent fuzzy setting rule. So, essentially it is a product operation you could look at it as a product team.

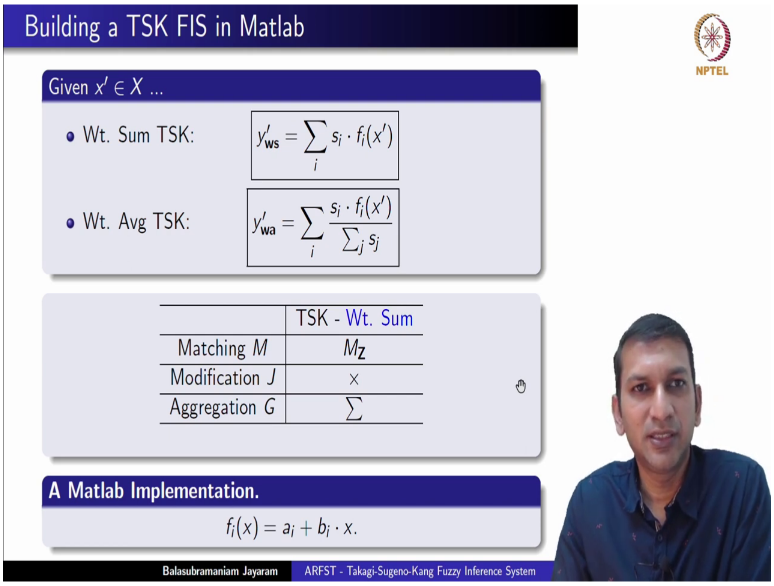
Now how do Mamdani and TSK compare as SBRs? Let us concentrate on the TSK weighted sum of course, we could also in some convoluted way look at weighted average TSK also as an SBR in the matching it or mapping it to the different operation that we are given of h M J G and g, but for the moment let us concentrate only on the weighted sum TSK fuzzy system.

So, we know that the matching function that you could use is typically the Zadeh's matching function of course, you could use any one of the major class of matching functions wherein you change those two operations plus and times cross. The modification function in the case of Mamdani fuzzy system is essentially this, you could also just look at it as product.

But in the case since we are dealing with only real numbers here, we could just look at it as product here as the corresponding modification function for the aggregation all we are doing is using the summation and since in a Mamdani fuzzy inference system you get an output which is a fuzzy set on why we need a defuzzification whereas, here you could just look at this as itself a real value.

So; that means, we are falling into the domain of y. Even though Matlab looks at this entire operation as a defuzzification operation we could look at this as modification followed by aggregation. So, this in a nutshell is TSK inference system this is how we infer, and this is how we could map it to the different stages of steps in a similarity based reasoning scheme. And we see some correlations or relationships the way Mamdani and TSK fuzzy systems either differ or also compare favorably with respect to each other.

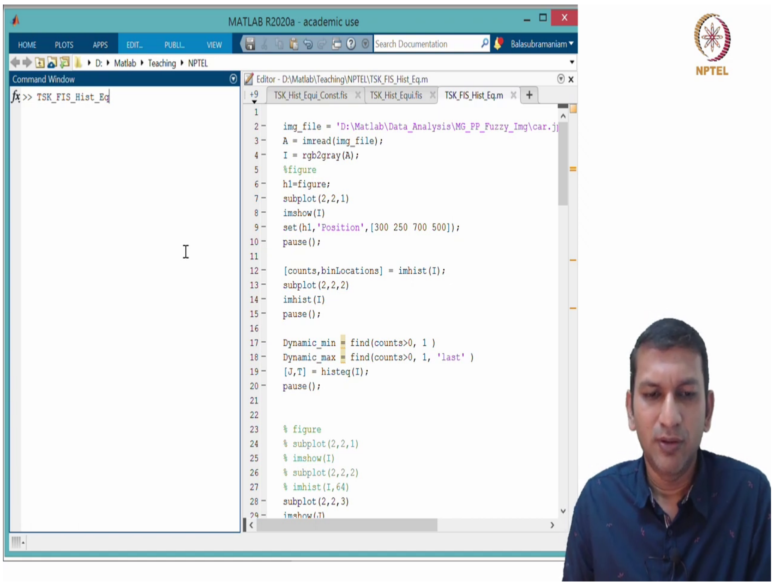
(Refer Slide Time: 12:32)



Now, we will take a look at how to build a particular TSK for fuzzy inference system using the fuzzy logic toolbox in Matlab. Note that these are the two things that we will be looking at and with respect to the TSK weighted sum fuzzy inference system this is how the different steps correspond to use M z as a matching function or one of those matching functions in those classes those class of operations where you only change the plus and the times, modification is product typically and aggregation is the summation.

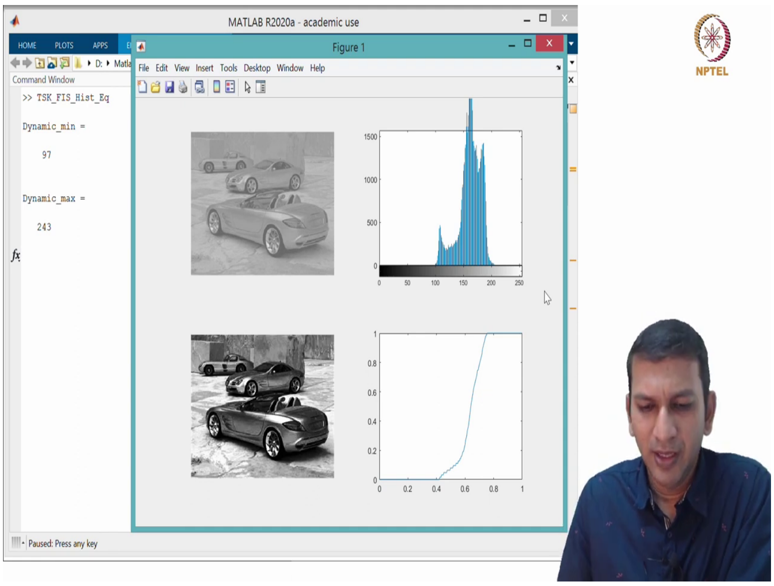
Let us look at a particular Matlab implementation and fuzzy logic toolbox allows you to implement the function on the consequent side either using a constant function or a linear function. That means, either you can have just a i where b i is 0 or it is a i plus b i x for each of those consequence and the n rules that you have.

(Refer Slide Time: 13:40)



Let us look at implementing a TSK FIS in Matlab. Well, what we will do is implement build a TSK fuzzy inference system to actually capture the behavior of histogram equalization as we saw in the previous lecture.

(Refer Slide Time: 13:57)

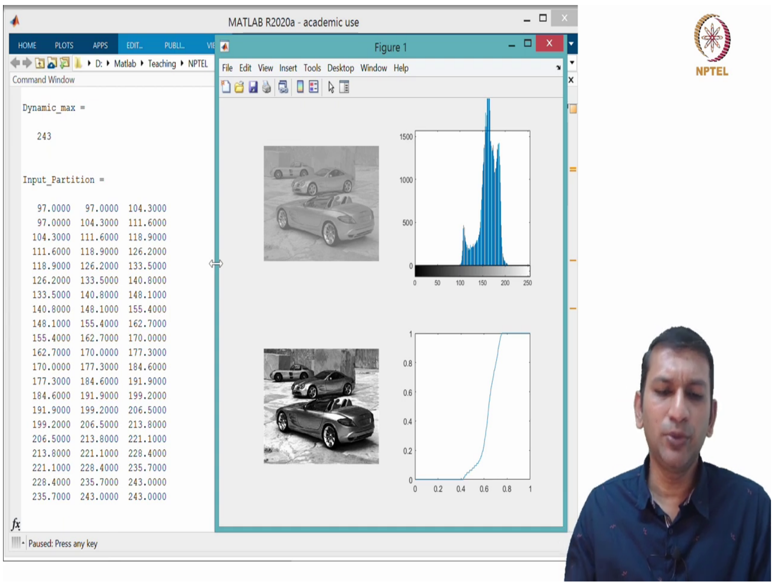


So, once again this is the image we are going to consider, clearly our goal is to enhance the contrast in this image. If you look at the corresponding histogram of pixel intensities we have seen the dynamic range of this is limited it is not the entire 0 to 55 interval and this is what we want to extend or expand. We have seen that the dynamic range is actually given by this interval 97 243 so; that means, the pixel intensities lie between these values.

If you apply the histogram equalization we have seen in the previous lecture this is the kind of contrast enhancement that you obtain and in the sense histogram is equalization can be seen as making such a monotonic transformation to this histogram and that is how it is expanding the dynamic range of the pixel intensities. We want to implement or capture this monotonic behavior using a fuzzy inference system. Remember it is the monotonic behavior of this expansion of this dynamic ring that we want to capture histogram equalization is one mechanical way of doing it.

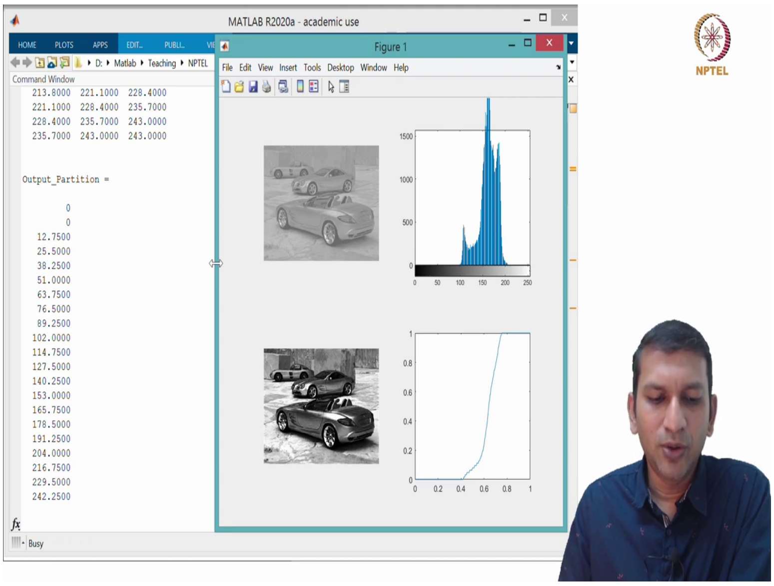
Now, we would like to implement a TSK fuzzy inference system to come up with a function which shows similar behavior that when applied on this image will allow us to obtain a contrast enhanced image. So, once again; that means, we need to actually find the fuzzy coverings on x and y. It is clear x is the dynamic range which is essentially 97 to 243 and y is the entire interval into which we want to extend or stretch this dynamic range into which is the 0 to 55 interval.

(Refer Slide Time: 15:53)



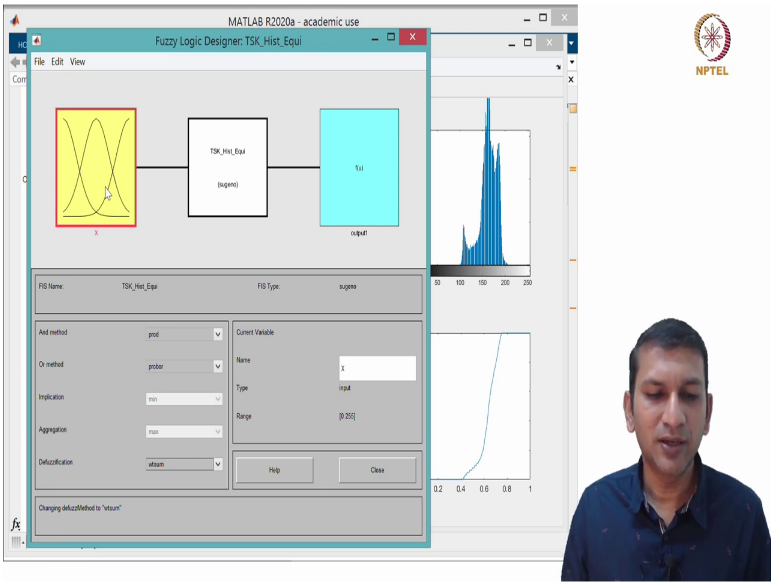
Let us once again fit split the dynamic range the actual dynamic range of the image into 20 equal intervals and fix triangular membership functions on them. So; that means, we are going to have 21 such membership functions this will form the fuzzy covering of the input space each one of them will also act as an antecedent. Now the question is to whom should we relate this to? For the moment let us only consider a TSK fuzzy inference system where the consequent function is a constant function.

(Refer Slide Time: 16:36)



So, essentially, we are only going to pick these points to be mapped into. These are 21 equal spaced points so; that means, you have 20 equal width intervals between 0 and 255.

(Refer Slide Time: 16:52)

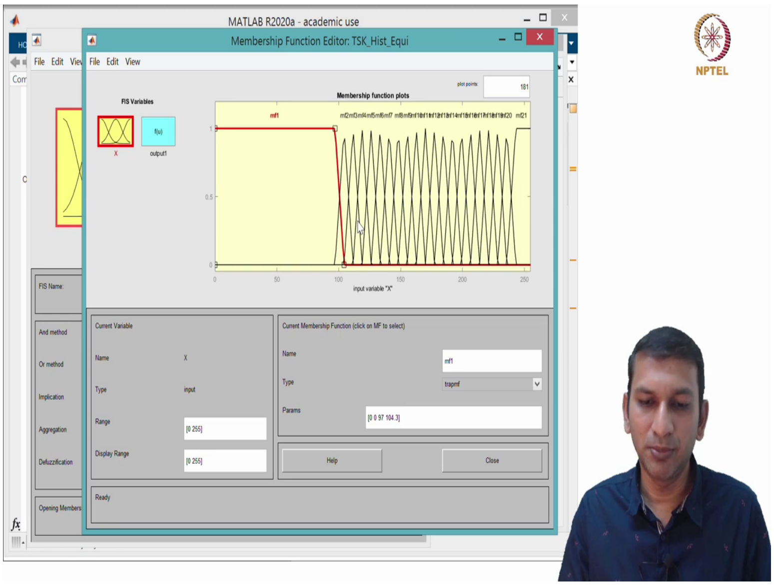


Let us look at how the TSK fuzzy infinite system itself would look like. As was mentioned you see here this is a matching function from the family of matching functions to which

Zadeh's and the (Refer Time: 17:04) matching functions below what we called as modification is implication here and aggregation is here you can see that Matlab calls weighted average or weighted sum as the defuzzification which of course, is a defuzzifying operation.

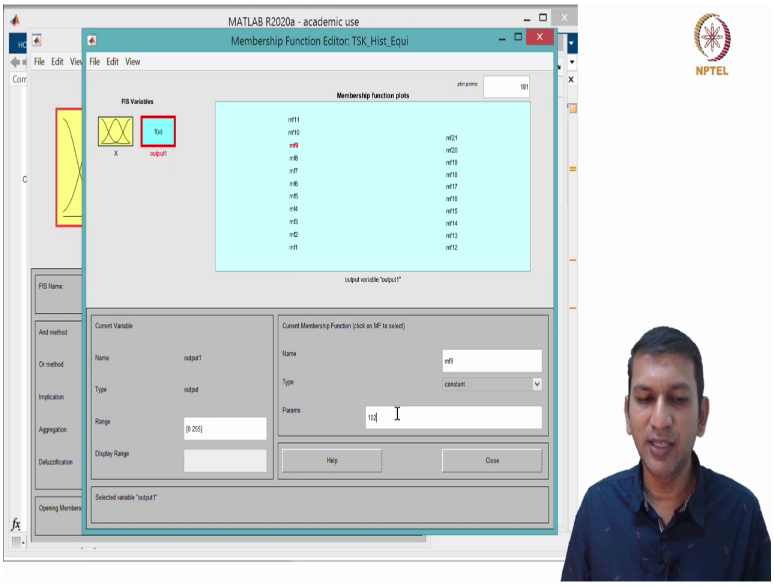
But instead of this defuzzification operation we can also look at the TSK file system being implemented as using the product for the modification or implication and the sum for the aggregation well.

(Refer Slide Time: 17:41)



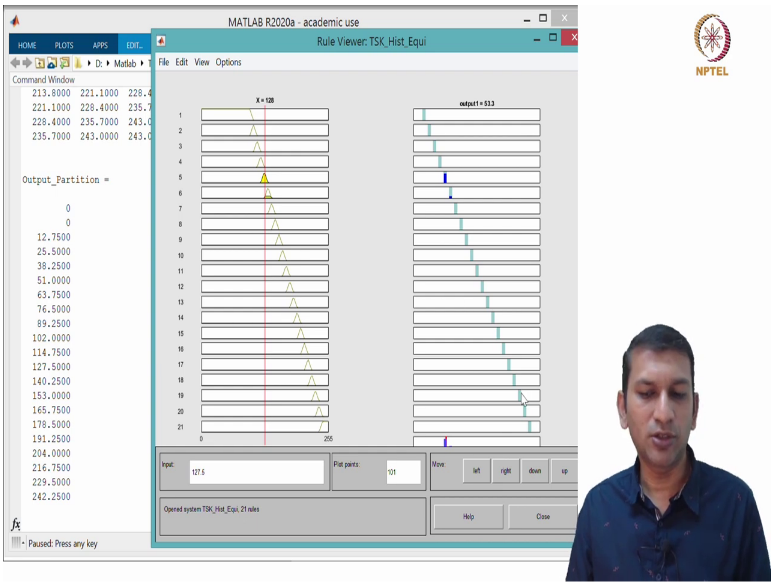
Now, let us look at how the rules are. This essentially what we had yesterday or the previous lecture. So, these are the input membership functions the where we have done it in when we are trying to capture the behavior of histogram equalization through Mamdani fuzzy inference system.

(Refer Slide Time: 17:58)



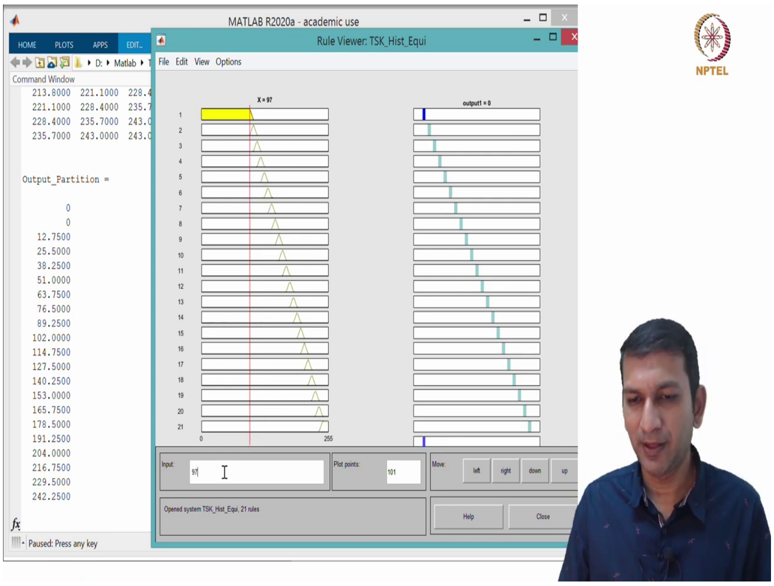
Now, you will see that in the case of output membership functions while it is not shown here you will see that it is in fact, so, the ninth membership function is mapped to a constant. So, f 9 of x is the constant function number 2 that is what we want and f 10 is the constant function 1 14.8 and f 11 is the constant function 127.5.

(Refer Slide Time: 18:31)

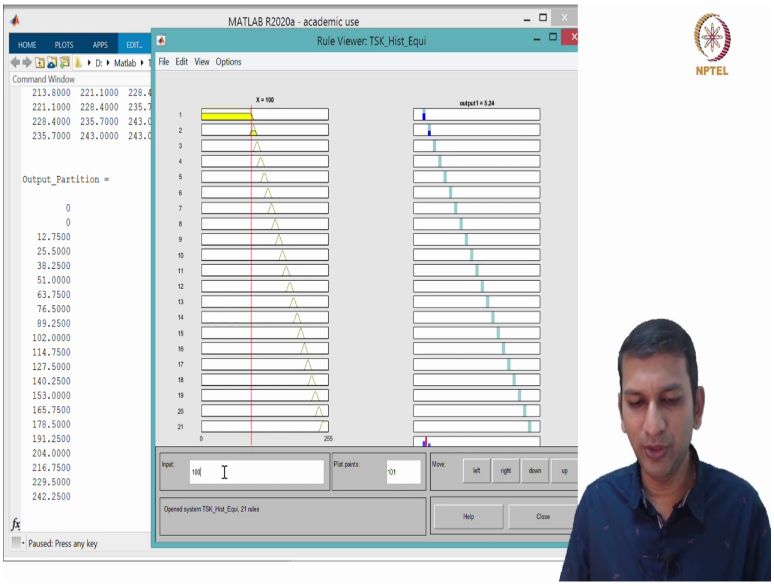


So, this is what we have here let us look at the rules themselves. It is easy to see now that these are the input fuzzy sets which become antecedents these are the output constant functions which are the consequence. Now just as it happened in the previous case the dynamic range starts from 97.

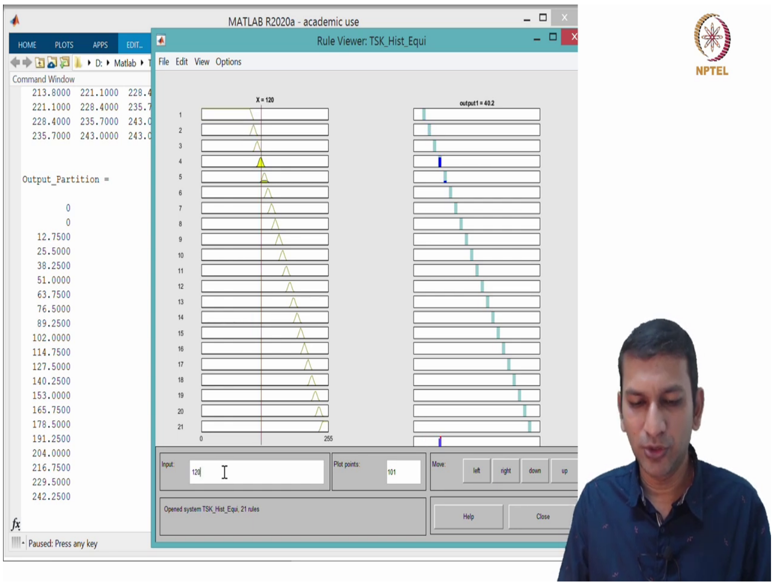
(Refer Slide Time: 18:50)



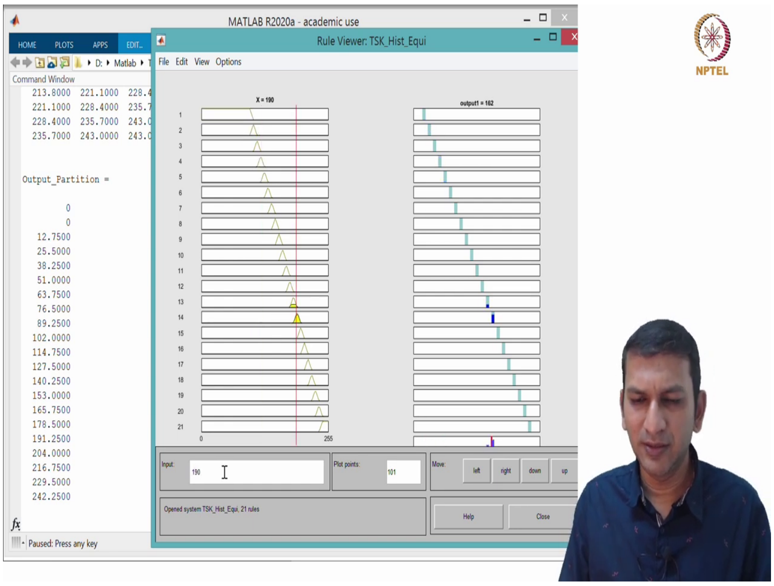
(Refer Slide Time: 18:54)



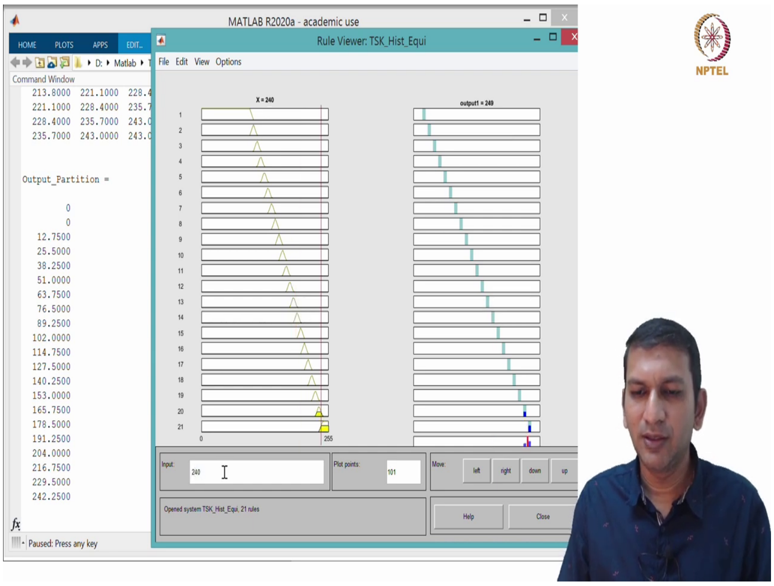
(Refer Slide Time: 18:57)



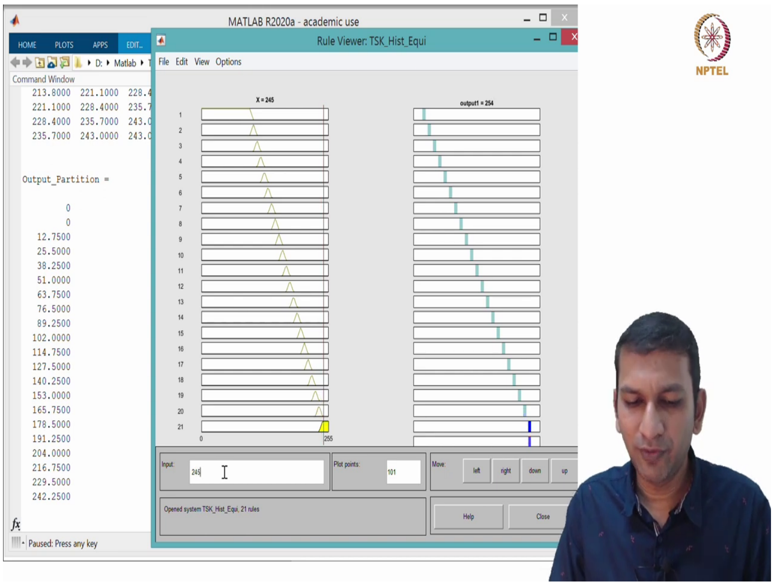
(Refer Slide Time: 19:00)



(Refer Slide Time: 19:03)

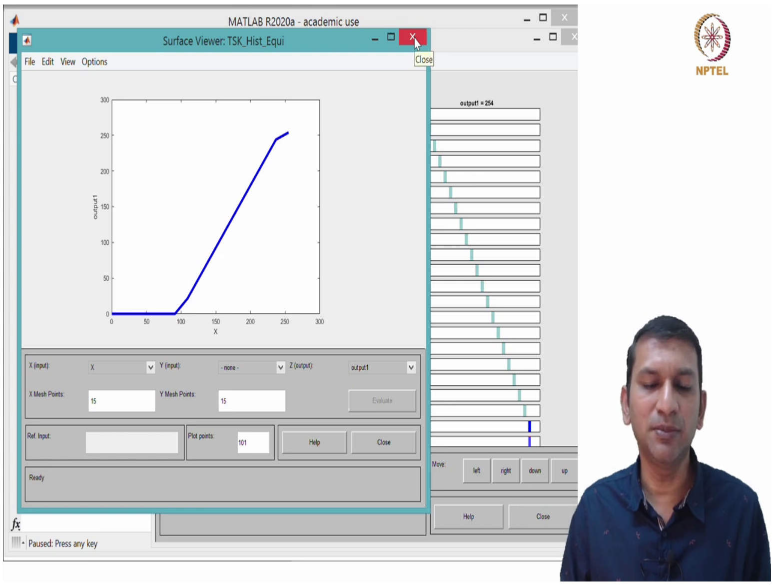


(Refer Slide Time: 19:08)



So, the input is 97 the output is 0, if it is 100 it starts to increase it goes to 5.2 if it is 120 it goes to 40 190 it goes to 162, if it is 240 it is almost 249, it is beginning to saturate at 245 it is 254. So, you see that we are actually getting a monotonic transformation of this dynamic range into the entire interval 0 to 55.

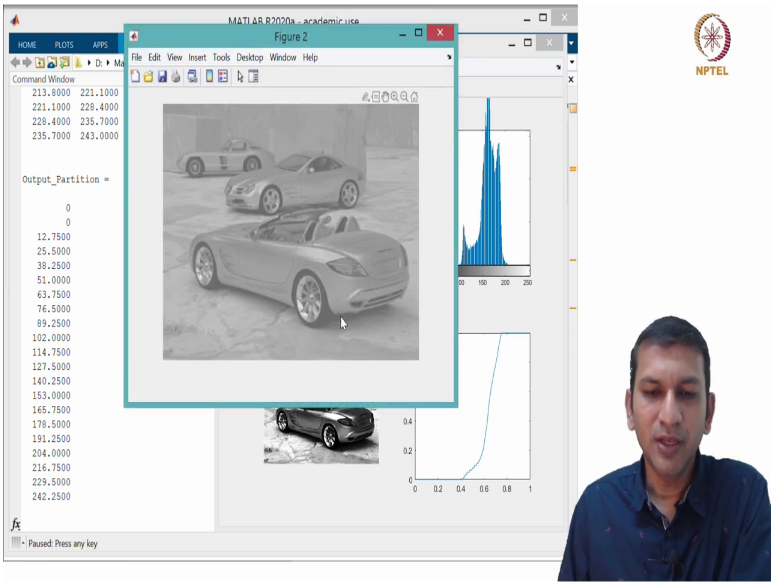
(Refer Slide Time: 19:21)



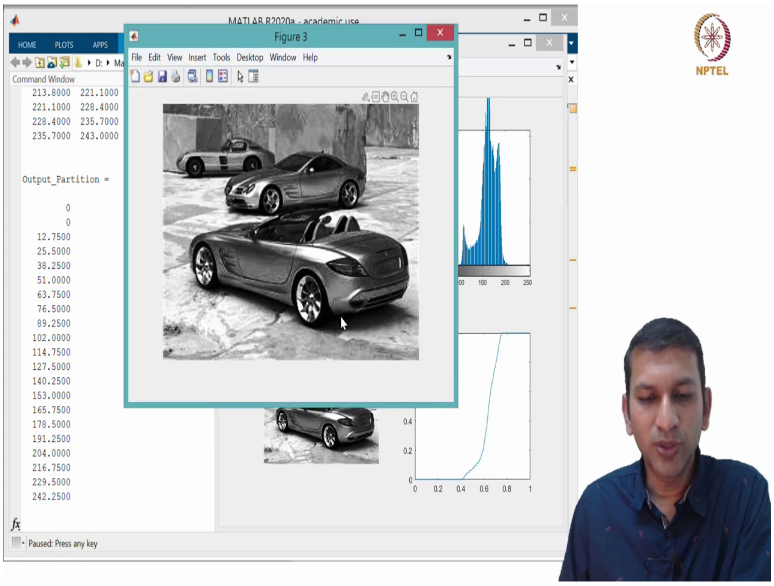
Let us look at the surface. So, this essentially the input output mapping that we had from the Mamdani inference system that we built. Of course, the inputs are same the outputs we have taken it to be just the centres of the output membership function that we had for the Mamdani fuzzy inference case.

So, that both of these are matching is not really surprising. So, this is what we have as the output function this is the mapping that this fuzzy inference system actually captures right. Now, let us look at applying this fuzzy inference system TSK fuzzy inference system to these images.

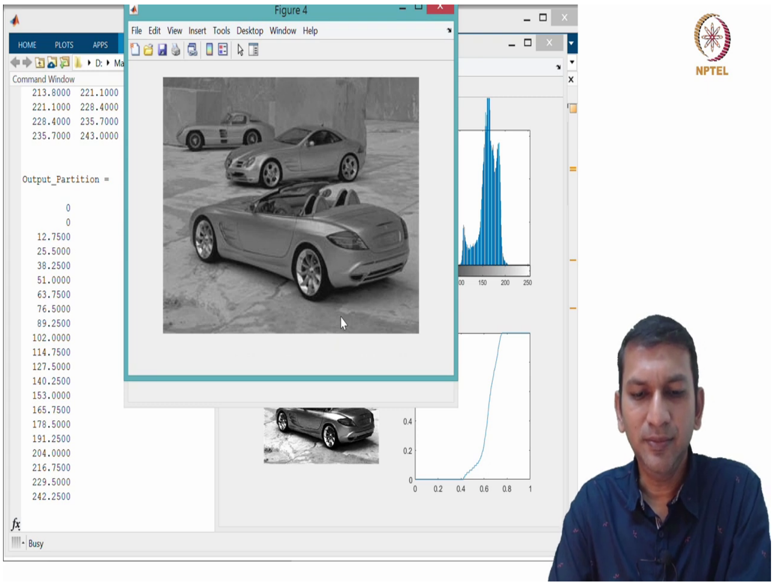
(Refer Slide Time: 20:10)



(Refer Slide Time: 20:15)

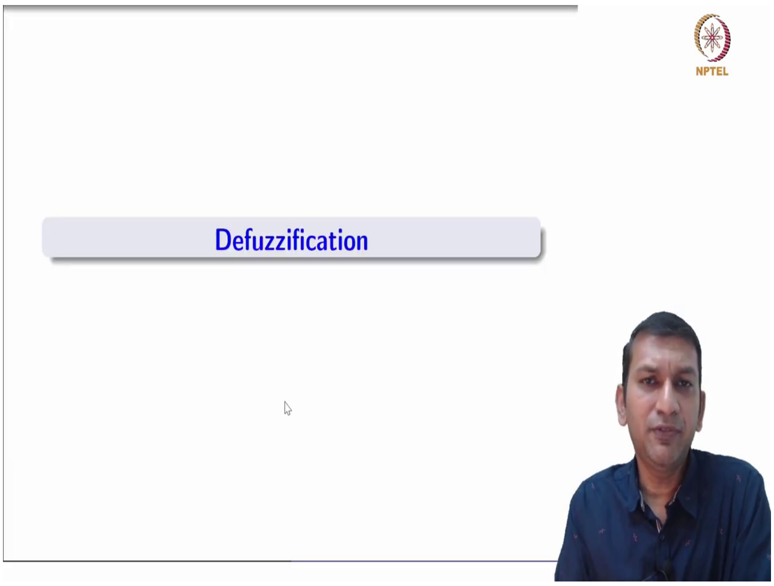


(Refer Slide Time: 20:19)

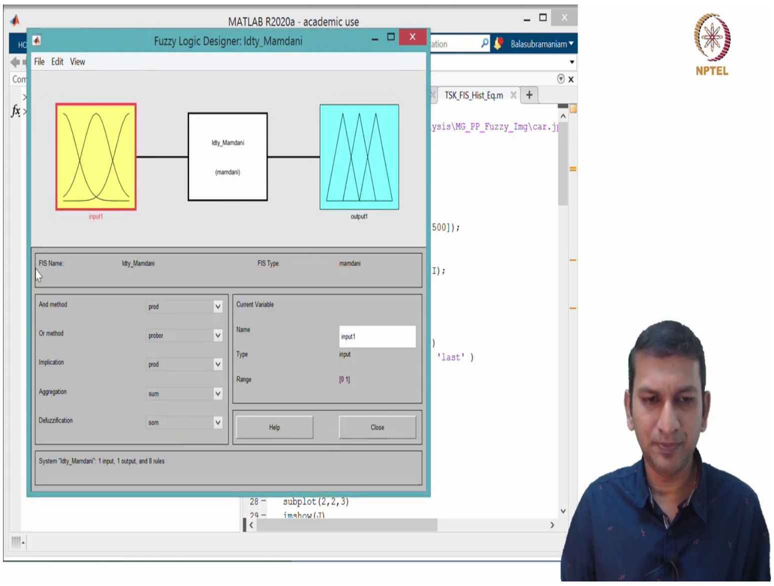


So, this is the original image we had, there is the histogram equalized image and this is the fuzzy inference system applied image. So, when you when we gave this input image to the TSK fuzzy inference system, this is the contrast enhanced image that we obtained.

(Refer Slide Time: 20:35)

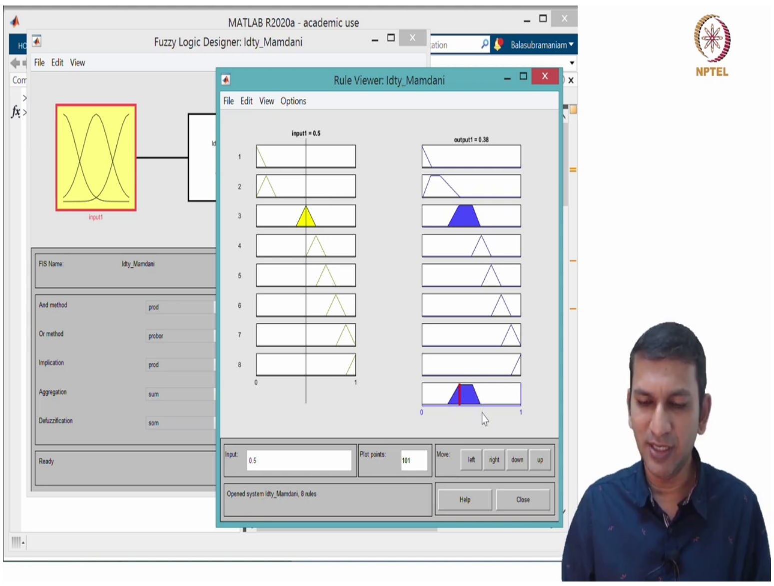


(Refer Slide Time: 20:42)



Let us take a brief detour to look at the defuzzification mechanism itself for that let us consider a simple Mamdani fuzzy inference system.

(Refer Slide Time: 20:52)



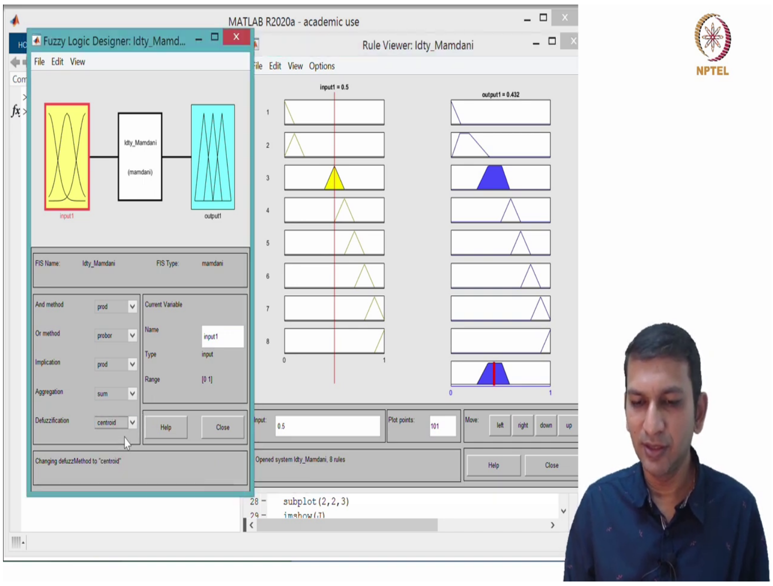
In this we are actually trying to approximate the identity function, but clearly this function the rule base that we have constructed is not going to approximated, but that is not the point here. We want to see the effect of different defuzzification procedures or mechanisms operations on the output. Now if you want to approximate the identity function then what we want is when we give the input to be 0.5 we are expecting 0.5 to be the output.

However, we see here this is the overall output fuzzy set B dash and this is the corresponding defuzzification mechanic operation that we have used. SOM stands for smallest of maxima. So, we see here in this fuzzy set the kernel belongs to this part and the smallest of maxima means, the first of the points to which the kernel belongs. So, we take the kernel which is part of the support over which the fuzzy set assumes the value 1 or it could also be that part of the support where it attains its maximum value the height the support of the height.

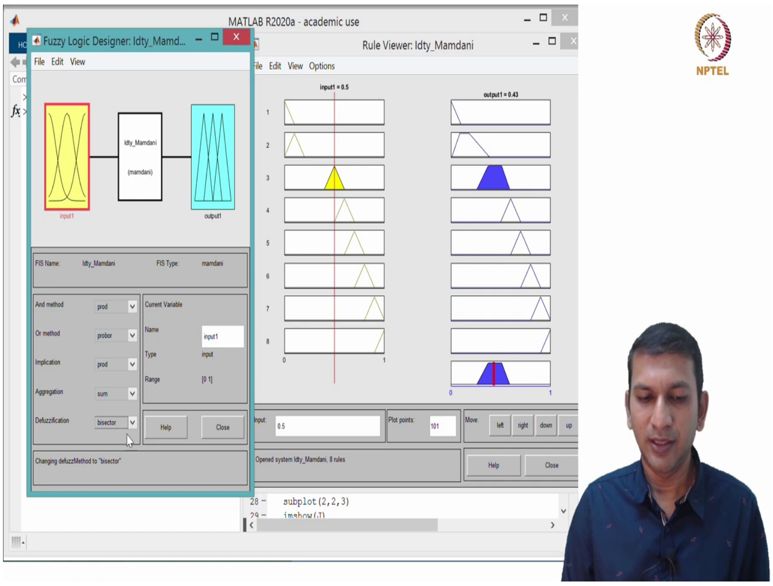
So, SOM stands for smallest of maxima which means essentially you are considering that interval over which attains its maximum value and taking the infimum of those points which in this case is 0.38. Instead of note that for 0.5 we are in fact, expecting 0.5 here, but that is not what we are getting we are getting 0.38. What if we instead take mean of maximum?

You see here it has moved from the first the infimum of the corresponding support to the middle point over which it takes the maximum membership values in this case it does look like it is 1 and you see already the output has moved to 0.445. So, from earlier it was 0.38 for the given input of 0.5 now it has moved to 0.45.

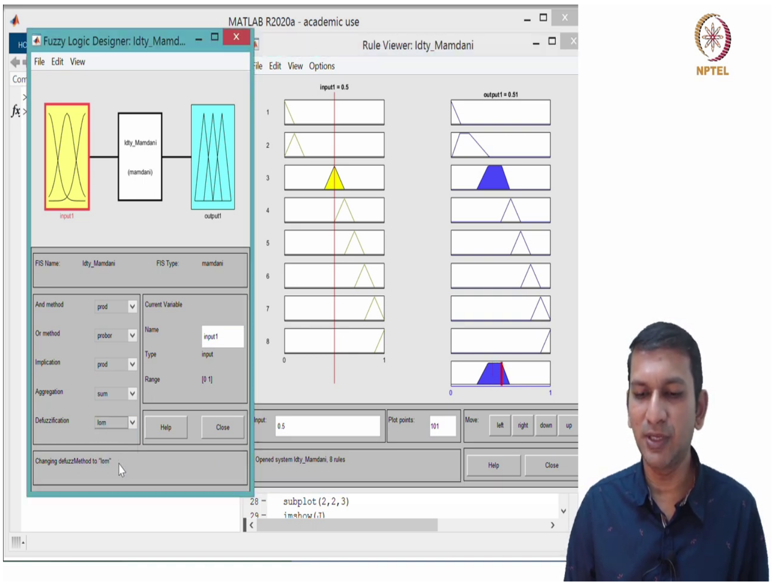
(Refer Slide Time: 23:11)



(Refer Slide Time: 23:18)

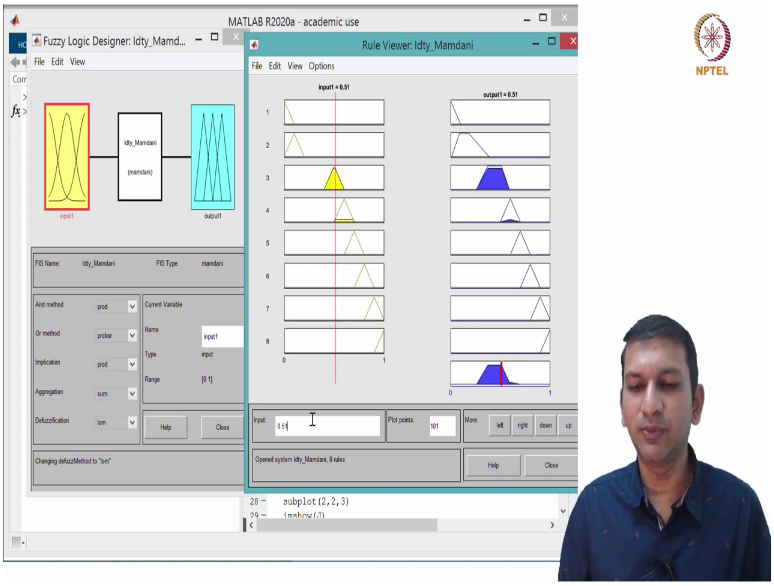


(Refer Slide Time: 23:27)



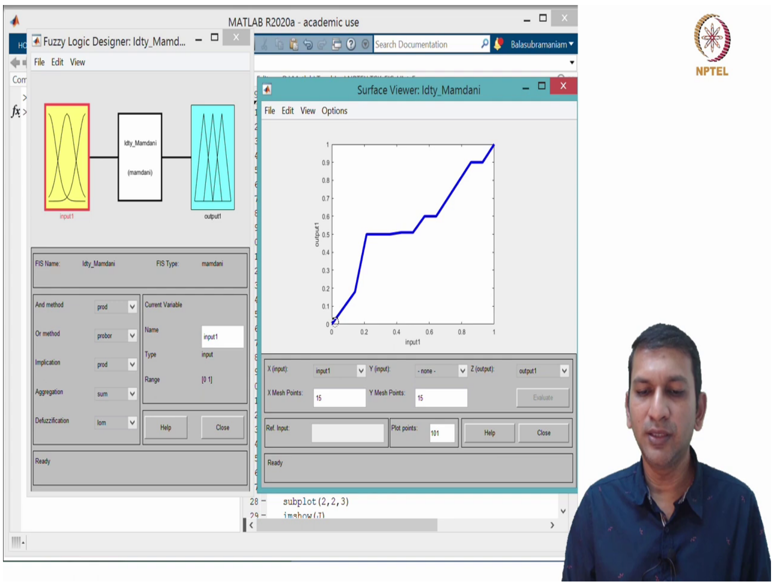
What if we take the centroid defuzzification? Let us move back a bit. So, it is 0.432. If you use the bisector defuzzification operation it is 0.43, but there is one more left LOM which stands for largest of maximum.

(Refer Slide Time: 23:35)



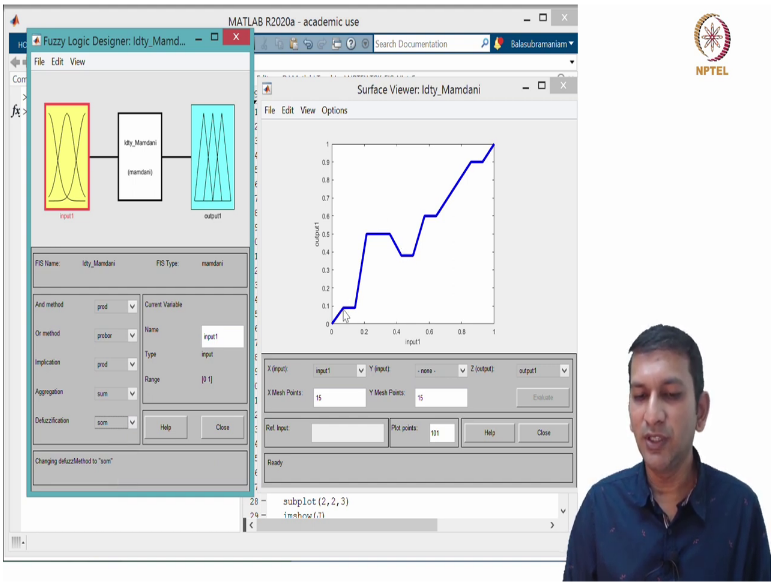
When we use that we see that its almost there. So, for 0.5 we are getting 0.5 and in fact, for 0.51 we would get 0.51. So, you see here by adjusting the defuzzification operation appropriately we can at times achieve the desired result.

(Refer Slide Time: 23:55)



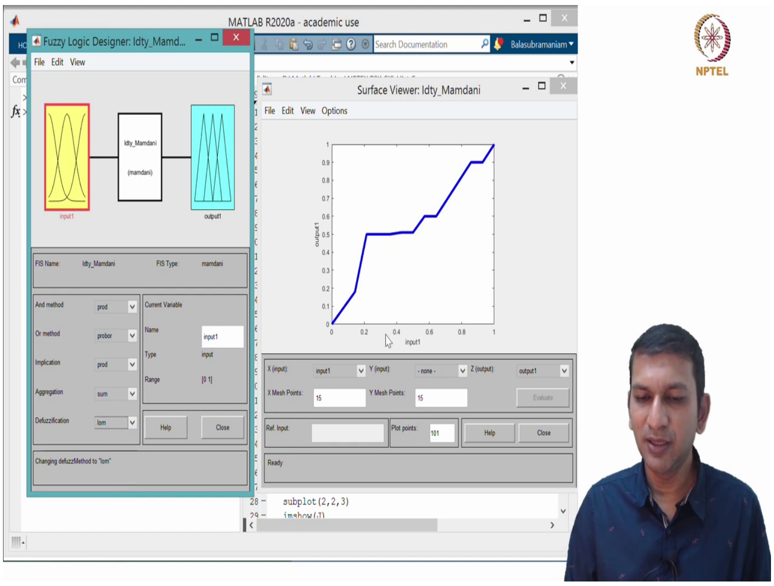
Let us stay with this for the moment let us look at the surface that we have caught here. Note that we actually want to and approximate the identity function; however, this is the mapping that we are obtaining.

(Refer Slide Time: 24:09)



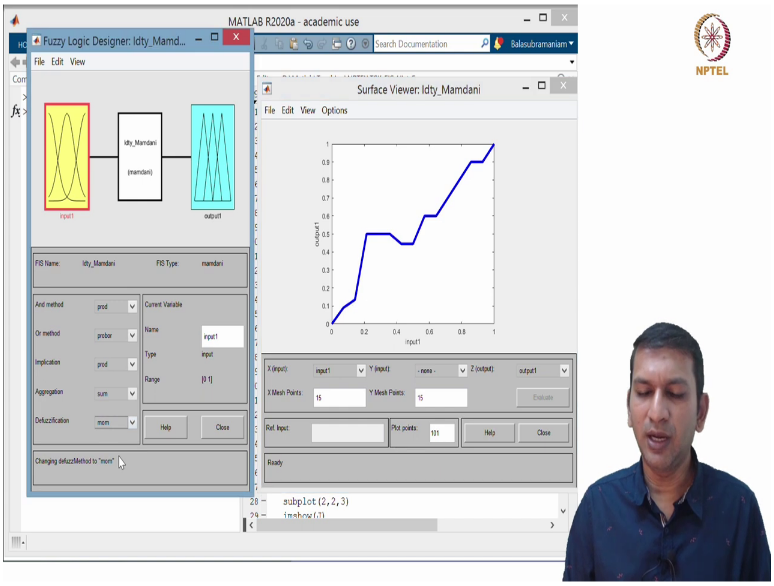
If you use smallest of maxima you see here this function is not only a not approximating the identity function, but is quite wiggly and it is also not monotonic even though identity function is monotonic on the unit interval.

(Refer Slide Time: 24:30)



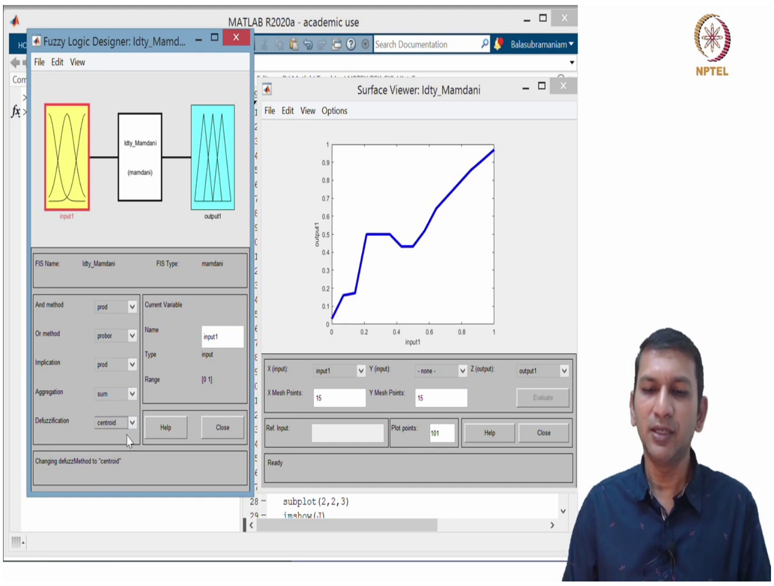
If you use the last of maxima well at around 0.5 it is doing a good job, it is in fact, hitting the diagonal of this function the identity function, but in other cases you see that it is not actually doing the job; that means, it is not the identity function. However, it is it seems to be monotonic; that means, once it goes up it does not come down.

(Refer Slide Time: 24:55)



If you apply the mean of maxima, you see that not only is it not approximating it well. But it also loses its monotonicity property and while these are continuous you see the sharp edges and so, you know that these are not differentiable.

(Refer Slide Time: 25:13)



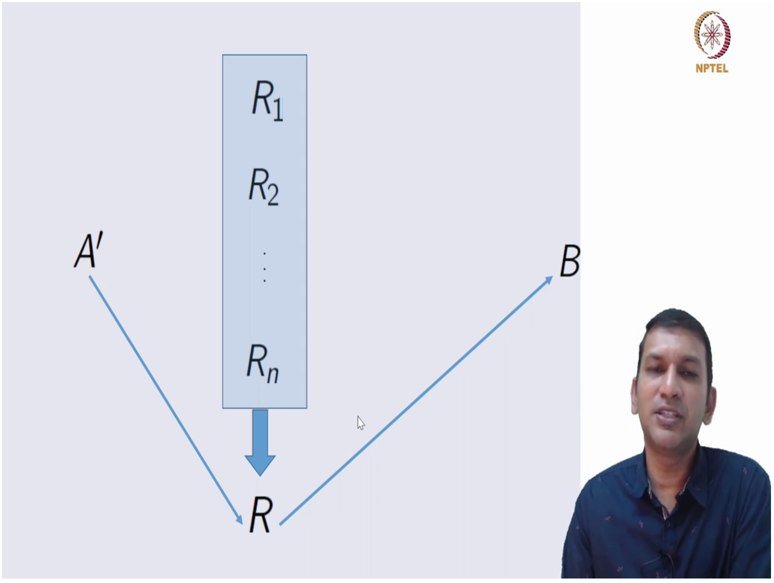
Instead, if you apply the centroid defuzzification scheme at least in some parts of the domain it seems more or less smooth. So, this should clearly illustrate demonstrate you that by appropriately changing the defuzzification operation we can in fact, expect to get to change the output function to a certain extent and typically when we want to tweak a function because we are going from some heuristics and building these fuzzy inference systems.

So, we have lots of degrees of freedom in the form of defuzzifier h fuzzifier h matching function m modification function g j aggregation function g and the defuzzifier function g. So, all of these degrees of freedom can be put to good use to actually come up with a function that we believe approximates the system function quite well.

(Refer Slide Time: 26:12)

Well in the last two weeks of lectures we have looked at fuzzy relational inference and the similarity base reasoning let us revisit them in a very pictorial way. So, please recall that in the case of fuzzy relation inference you had two inference strategies first aggregate then infer or first infer then aggregate.

(Refer Slide Time: 26:36)



Now, what we have to infer is, given an input A dash which is a fuzzy set on x we want the output to be a fuzzy set on y which we have indicated as B dash. To help us in this the knowledge base which is the ground truth consisting of fuzzy if then rules each of these rules is translated into a fuzzy relation. So, the first rule A 1 implies B 1 is captured in the form of R 1 fuzzy relation, similarly the other rules R 1, R 2 so, on till R n.

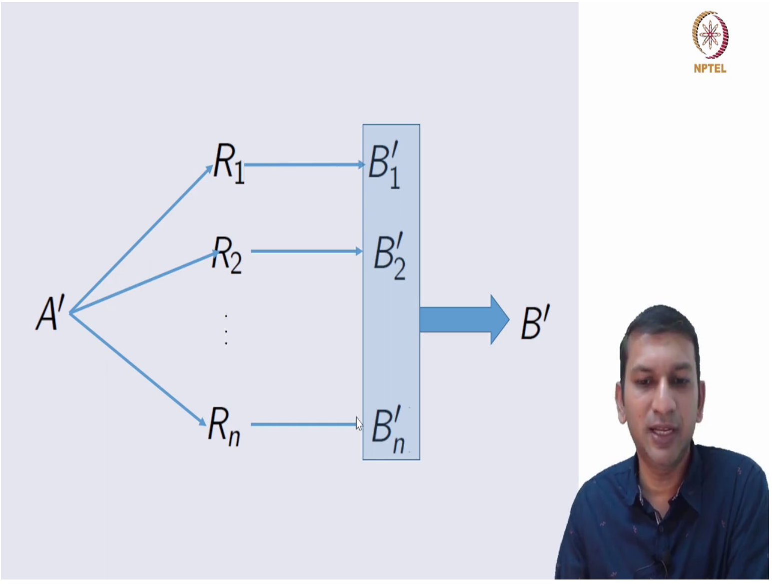
In the case of first aggregate then infer what we would do is combine all of these relations aggregate all these relations into a single relation R and use this to obtain the corresponding B dash. So, we compose A dash with R with any of the compositions either the sup t composition or the In-phi composition to obtain R B dash. So, if you use the sup t composition we call it the compositional rule of inference as proposed by Zadeh if you use the In-phi composition the Bandler cohort sub product composition we call it the BKS inference.

So, this is pictorially how we do fatty first aggregate then infer in fuzzy relation inference.

(Refer Slide Time: 27:58)

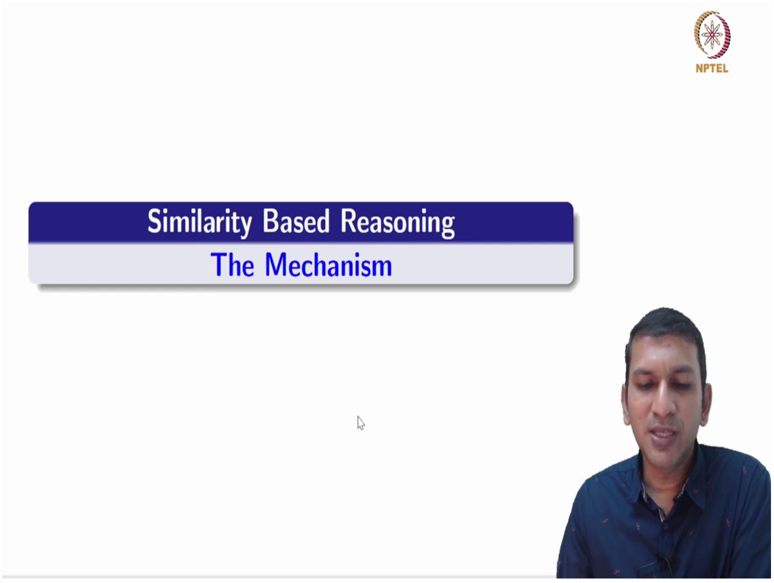
There is also an alternate inference strategy first infer then aggregate.

(Refer Slide Time: 28:04)

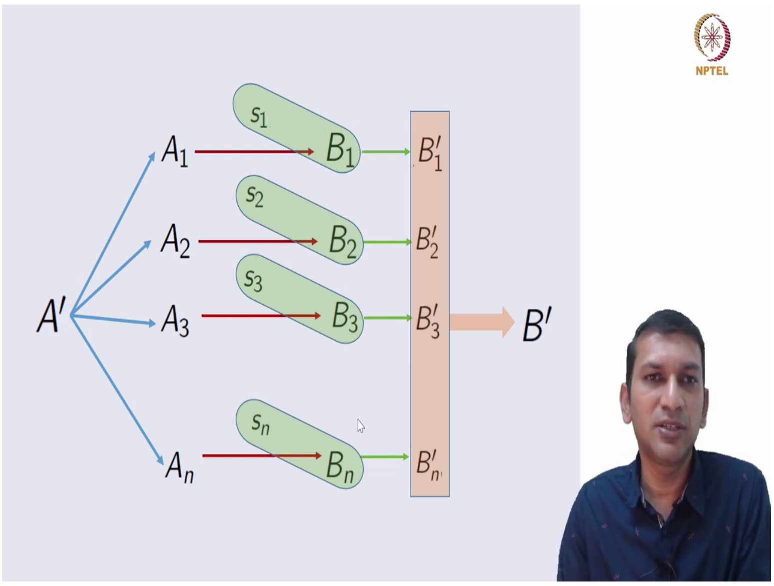


In this case given an A dash to obtain a B dash what we do is we take all these relations we do not aggregate them beforehand instead we infer B 1 dash from R 1 and A dash. Similarly, given this A dash we obtain a B dash through R 2 and so, on and so, forth till we obtain B 1 dash from A dash and R n. Now these locally inferred outputs B 1 dash B 2 dash on till B n dash all these B dashes we aggregate them. So, first we infer all these B i dash then we aggregate them into a B dash this is how we obtain the overall output fuzzy set B dash.

(Refer Slide Time: 28:51)



(Refer Slide Time: 28:55)

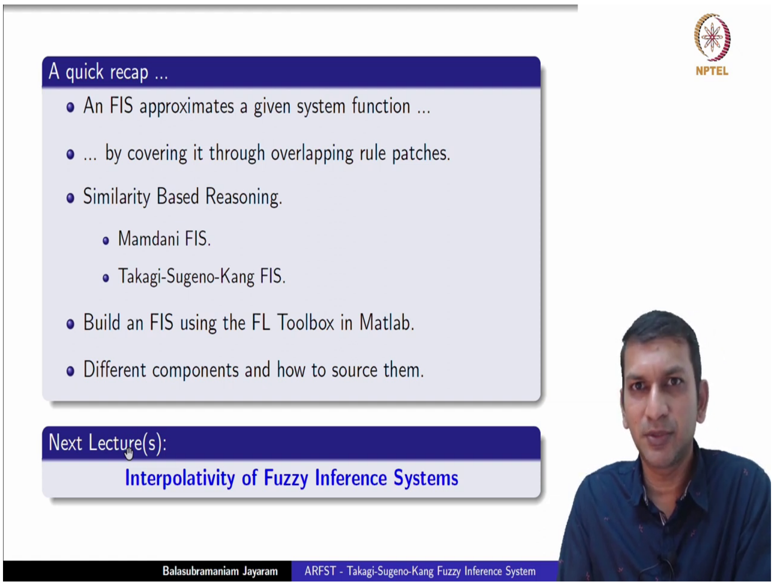


Now, if you look at similarity based reasoning once again we are given an A dash we would like a B dash but now we are keeping these rules as such. So, A 1 implies B 1, A 2 implies B 2, A 3 implies B 3 so, on till A 1 B implies B 1. Unlike in the case of fuzzy relation inference we do not change them into fuzzy relations instead we keep the rules as such given an A dash we match this A dash to each of the antecedents and find the corresponding similarity values.

So, A dash is matched with A 1 with a matching function the corresponding similarity value is s 1 the similarity of A dash and A 2 is given by s 2 similarity between A dash and A 3 is given by s 3 so, on and so, forth till s n. Now, using the similarity values we modified the consequence of the corresponding rules. So, s 1 modifies B 1 to give us B 1 dash s 2 modifies B 2 to give B 2 dash s 3 modifies B 3 to give B 3 dash and s 1 modifies B n to give us B n dash towards this end we use a modification function j.

Now the final step we aggregate all this modified consequence much like the way that we do in FITA and obtain the overall output fuzzy set B dash. Perhaps you may have seen some relations or relationship between or resemblances between FITA First Infer Then Aggregate and SBR scheme. In fact, later on in one of the lectures we will see how a fuzzy relational inference in with when the operators are specified in a certain way can also be looked at as a similarity based reasoning scheme.

(Refer Slide Time: 30:49)

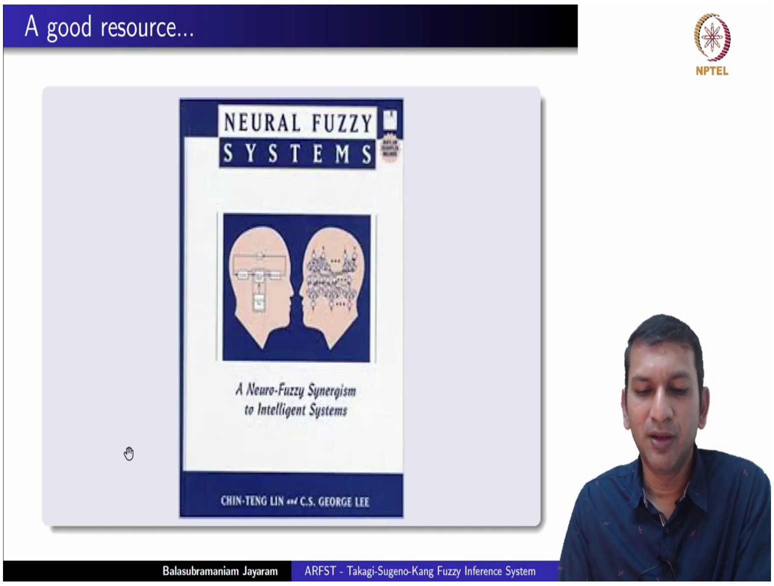


Well, a quick recap of the lectures that we have seen throughout this week we are now very aware that a fuzzy inference system approximates a given system function by covering it through overlapping rule patches. This is the concept that we have put to good use when we try to build fuzzy inference systems we have specifically discussed similarity based reasoning schemes in this week 2 of the major types are that of Mamdani and Takagi Sugeno Kang fuzzy inference systems.

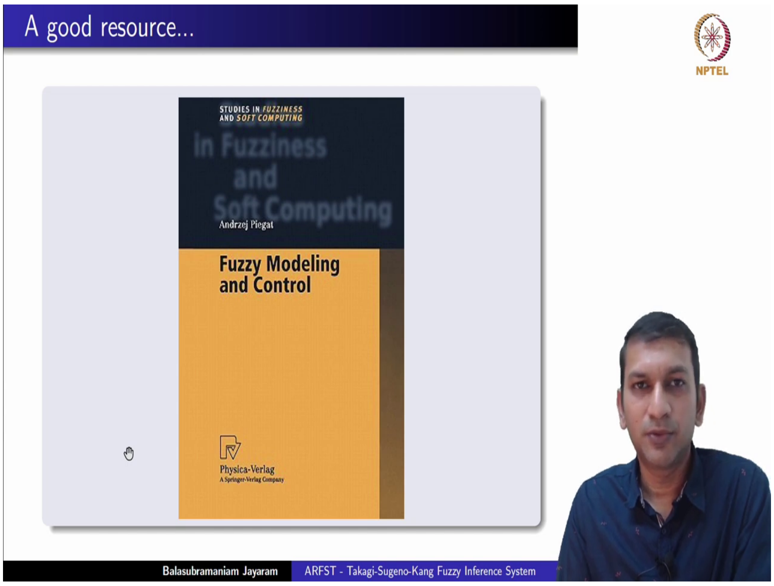
We have also seen how to build fuzzy inference systems both the Mamdani and the TSK type using the fuzzy logic toolbox available in Matlab to approximate a given specified function or even in a practical application, where we only had some qualitative features of the system function. Finally, we have also seen the different components of the fuzzy inference system and how to source them in a practical application. In the next few lectures we will discuss interpolativity of fuzzy infinite systems.

(Refer Slide Time: 32:07)

(Refer Slide Time: 32:16)

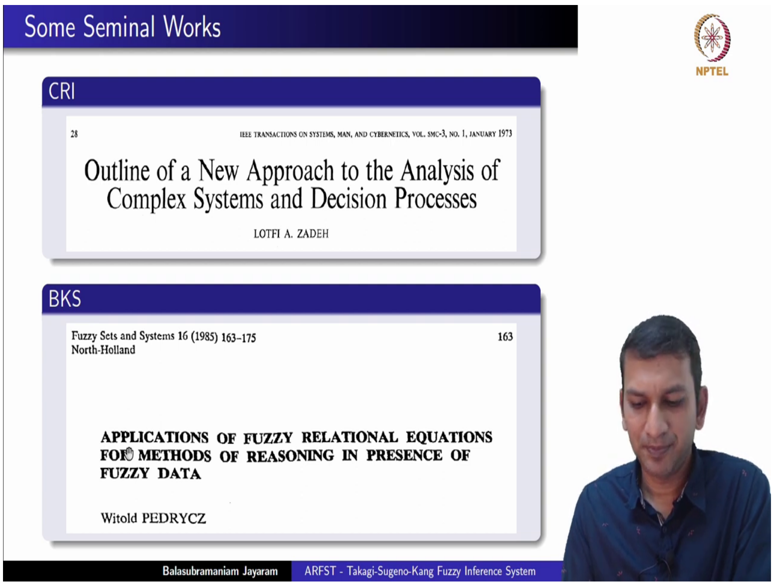


(Refer Slide Time: 32:21)



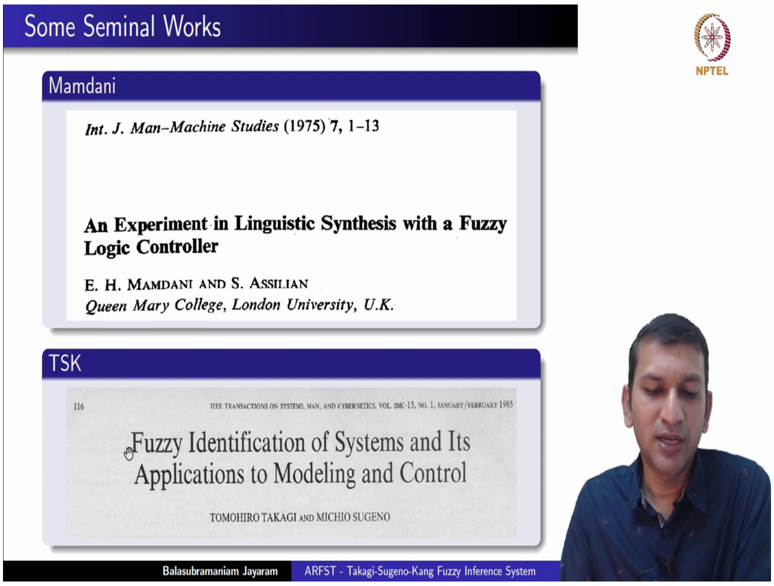
A good resource for the lectures that we have had through this week the topics covered in them is the book by Passino and Yurkovich, also the book of C T Lin and George lee and the book by Professor Piegat.

(Refer Slide Time: 32:26)



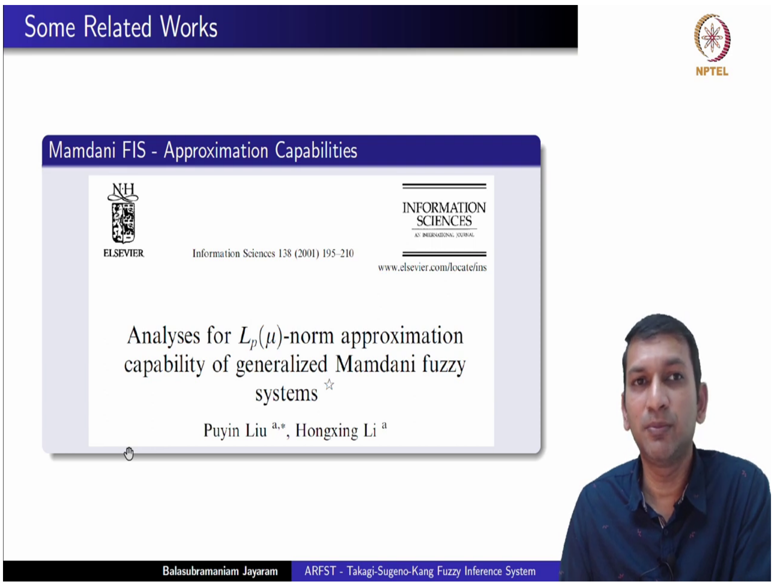
At this juncture I would also like to point out some seminal works that we have discussed throughout these lectures, it is this paper in which Professor Zadeh proposed the compositional rule of inference way back in 1973 in 1985 Pedrycz used the Bandler Cohort sub product composition to propose the BKS inference scheme.

(Refer Slide Time: 32:55)



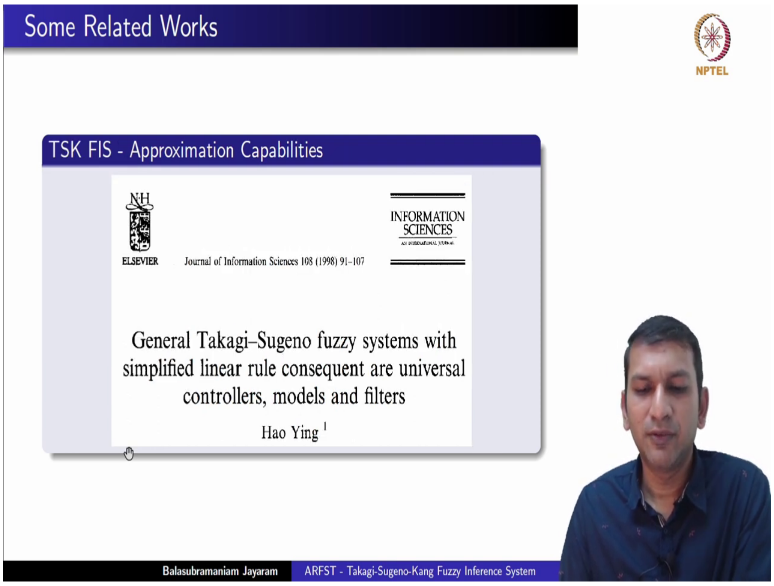
In this work it is in this work published in the journal of Man Machine Studies in 1975 Mamdani and S Assilian they proposed that Mamdani fazzy inference system in the year 1975 and Takagi and Sugeno they proposed the TSK fuzzy inference system in 1985 through this article published in the hydro play transactions on systems man and cybernetics.

(Refer Slide Time: 33:28)



Some related works it was mentioned during a few lectures that a Mamdani or a TSK fuzzy inference system is capable of approximating any continuous function to arbitrary accuracy. We have seen this with a few practical examples using the fuzzy logic toolbox in Matlab, but theoretically also these results are available. So, this is one paper that deals with Mamdani fuzzy systems showing their capability the universal approximation capability of Mamdani fuzzy systems.

(Refer Slide Time: 34:00)

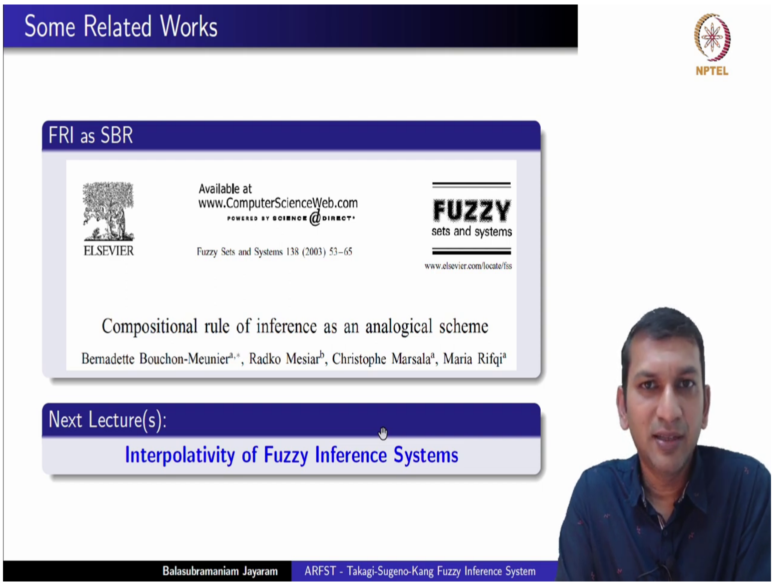


There is another paper way back in 1998 which discussed how even with linear rule consequence generalized Takagi Sugeno fuzzy systems are universal approximator. Please recall we have also used the consequent functions to be constant functions in the example that we have seen the case of contrast enhancement using TSK fuzzy system and we have seen that it has more or less captured what we had in mind.

(Refer Slide Time: 34:30)

This yet another paper discussing the uniform approximation or universal approximation capabilities of fuzzy relational inferences.

(Refer Slide Time: 34:40)



And finally, as was mentioned under some conditions on the operators and FRA can also be seen as an SBR. This is something that we will take up in one of the oncoming lectures, but this is one paper a work related to that.

So, in the next few lectures and from henceforth we will look at some of the properties that a fuzzy inference system should possess or we expect it to possess the desirable properties that of interpolativity, continuity, monotonicity and so on. In that quest we will begin by looking at the interpolativity of fuzzy inference mechanisms. Glad that you could join us today for this lecture I am hoping to see you in the next lecture.

Thank you again.