### Algorithm\_Tutorial\_before\_midterm\_assignmentAndInclassQuestion\_20241031\_14:05

说话人1 00:00  
The elements like the, for example, it's usually close to 123, and the s one contain one, and s two contain two. And s three contain one and three. And thats water containing three. For example, if this cost is really, this cost is the one. This cost is two. For example, if you select s one and s two, total cost is a normal budget. It's a feasible solution. But we want to maximize the two elements that we cover this feasible solution. Only cover two elements in this two. For example, if we select s two and s three is also a physical solution, do not exceed the budget, but they can cover 3 and 123. This is a better solution than this solution. We need to clarify our objective. And then, let's see, we use in the problem statement. We use the function as for coverage function. As is if we use the index to denote for you to be invest in, not like what we select.

In this example, then the f one and two means we select as one as two. Then this function value equals because what? The success is selected, cover two elements, and then like 1 and 23, this equals to three.

We want to select several substance and maximize this as well. This is the, this is our problem. And let's see, this question gives us a greedy algorithm. This is really algorithm run like this. We start from an empty set, like ig initially equals to a an empty set. And each time we select a subset, how to say it has the it can covers the maximum uncovered element, but this number also divided by the cost. Can be said that each time without a set, which can for many uncovered elements, but also has low cost. We can see the line three I start. Initially, ig is

说话人2 03:28  
afc

说话人1 03:29  
set, and this I star inputs to this argument means we arguments I belongs to. M is the total index set.

This I belongs to m and this fig plus minus

说话人2 03:57  
nig

说话人1 04:00  
divided by cn we want to run this m select one I once a set maximize this ratio. And intuitively, this ratio means we want to select the success cover more uncovered elements, and at the same time is lower cost.

The line two just let us go through all success. We checked, we check the line four, check the weather. We can and a new success in our bundle. So if if add a new success exceed our budget, then we do not add this element. This success. If it is within our budget. So we add this subset is feasible, then we editing our ig and both see the next element. Go to the next subject ok so now let's see these questions. Do you have any problem about the problem statement and real reason? But do you know our objective is what and how they actually address?

These elements are very intuitive. Make sense? All right. Ok let's see the first question show that reading algorithms, approximation ratio cannot be better than when we're in by giving an example. So cannot be better than one word means means we give an example, and we show that the optimal solution is way better than the really come with them.

This solution is like, for example, if if we have an element, just like the issues, we have a set, for example, in this s one. Now we construct an example, give this s one just equals to the universal sense. We can see if we select as one, then we only need this set to cover all the elements. We can set this cost be very like the very large. For example, as to the only cover one element, we can say that is called the, for example, if this k two cost is the, then this ratio, the algorithm, when the algorithm see the first iteration of this algorithm, it will check

which set is the maximize this ratio.

Right now, maybe this is constant t and this a little bit within smaller is

说话人2 07:44  
ay

说话人1 07:46  
divided by two n then the ratio for we consider the first information on this reality.

For example, we only have these two success. And the the ratio of about this is n divided by b divided by two n two. In the first equation, ig is the empty set. It is the

说话人2 08:16  
mf

说话人1 08:17  
one

说话人2 08:17  
minus lmh

说话人1 08:23  
then for each of each is the, it is at this time and then divided tonight.

说话人2 08:30  
So then that is 4 points.

说话人1 08:36  
So this should be

说话人2 08:52  
how we call. Is it? How we create it? Go down can be. Yeah.

说话人1 09:00  
And so this is one, okay?

说话人2 09:05  
That's all right, ok

说话人1 09:08  
so that's right. Over. And this is f 2 %.

说话人2 09:15  
Good question. This is me. So everyone

说话人1 09:30  
has a short question is, this equation is not estimation. The university

说话人2 09:44  
equation, the

说话人1 09:46  
reality from this analysis. One set,

说话人2 09:49  
just the right

说话人1 09:50  
only cover one and it. But the optimal solution is this one as one cover an element. The radialism cannot achieve better ratio than 100. It just construction for such that the greedy algorithm will choose a very bad solution. The optimal solution can cover audience.

Let's see the second question. Don't really algorithms. Approximation ratio can be very small when the number of elements n is very large, if we place an instant n the total elements, the number of them, elements very large than the ratio of very small, and which means the performance of this really are each other is very bad. We will see this area can serve a building block for a modified area, which is constant information.

Now, suppose 94 is always true during the first l information about. Now we have a an il

说话人2 11:23  
it's not too much.

说话人1 11:27  
Suppose now we use the this term might help to, like trace the progress of our algorithm. So you can see our video with them will go through

说话人2 11:41  
all

说话人1 11:41  
success, right? I am not single to all the tests.

说话人2 11:51  
That's

说话人1 11:57  
it will see every such as right. We use this il indicate the apparently ig becomes

说话人2 12:07  
watch.

说话人1 12:09  
So let's use this il

说话人2 12:11  
two

说话人1 12:11  
I one I I

说话人2 12:15  
I am,

说话人1 12:20  
which means which our algorithm first select the si one.

So know that these are index of the set.

But is that not so bad? So this is not impossible. We can say that universal situation,

because we support the language always

说话人2 12:48  
during

说话人1 12:49  
the personnel integration of our practice. It means our video is the first select I want to set. And then I two

说话人2 12:59  
subject,

说话人1 12:59  
and then I have to set. Then I start with the index sets, set of sets selected by output. For example, I started the optimal solution campaign, several concepts that is our problem for the convenience of analysis about the problem two things to see. So let's let's consider such a situation. So this il equals to this first first assumption is the algorithm, the line, four of these algorithm always evaluate. It would be

说话人2 13:43  
true

说话人1 13:44  
about this sets.

And the second situation is the il is the first one that maybe the line four of these algebra evaluation

说话人2 13:57  
force,

说话人1 13:59  
which means, which means the algorithm and I one, I two dot il minus one. When this algorithm select il when at the iteration that il is the success that maximize the ratio. But this area cannot be added into the bubble, because the language

说话人2 14:31  
false, which means

说话人1 14:32  
as the im and the budget and the two possible exceed the budget.

Let's

说话人2 14:39  
consider these

说话人1 14:40  
two situations, and we will see that the two a will hold the inequality of this employee report in these two situations.

A let us show that fi

说话人2 15:00  
star

说话人1 15:00  
minus

说话人2 15:02  
fi this is

说话人1 15:08  
ok first, we intuitively interpret this inequality, which means fi star is the number of elements covered by optimal solution. This il is in the else in the l's iteration before or the totally l equations that the radiology as a reality element select this process, this is aa difference between optimal solution and pretty solution. We can and the race in the health step of race team. This one is the difference in optimal solution and the previous step of gradient. It means our reading algorithm is

not make progress. So it does cover more and more elements approaching the. Let's show this

说话人2 16:24  
is how do you

说话人1 16:25  
so we have a hint.

说话人2 16:27  
Is it more? So

说话人1 16:28  
this

说话人2 16:28  
is how you want

说话人1 16:34  
it. So this means we can use this pin to bound the number of fi right? Now, we actually want to show that the outcome solution is not is not so large. It can be bound by the algorithm.

I do it. Today. We will like the substitute the I with sr if we substitute I with sr and j with our really algorithm solution, then it will be like I star, ok this is a hint. And I just substitute. I started is I and just I with I star and sub to the j with il minus one, chinese. And then this will is equal to I divide this by ci and be still one ci out. This should be equal because this ci and this ci can cancel. This actually equivalent to the last session.

So why we want to like to change it like this? Because we want to make this ratio show up, because our reading algorithm behavior, the behavior is it. Each time it selects a new success, a new item to maximize this ratio, we can see the sum over this. I belongs to those elements, which

contained in optimal solution, but doesn't contain that isn't contained in the io minus one. Right? So note that these elements, where is the elements? What are those elements? This is a very important to transform. So we can see these elements are in outcome solution, but not really solution. But now really, I will, each time that's my situation.

If we change this I into l in il then those il will maximize the solution compared with this r we can take it out in this form, this. I it's different, right? Because design belongs to this. If we change all these things, all I in this form into the il which is the which is the one to maximize this ratio.

Then we can take it out, leave this in here. This is smaller or equal. Just because the reading is about that. Do you agree with

说话人2 21:25  
this estimation?

说话人1 21:32  
We can see if this il is the first element that mean that make our current model exceed the budget, then this also hold right. Here we can see this cost is all the total cost of these elements. It must be smaller or equal to the elements of optimal solution, the power to the cost of the optimal system, because this is only a part of optimum. And our optimal solution must be a feasible solution. The feasible means it does not exceed the budget.

So it is, how are you to be? So if you just change it to be, and you rearrange the items, rearrange these terms, and you can get the inequality, which means. So basically, three key steps, why is that you correctly use the hint? And the second, you observe how you observe the property of readiness. Reading algorithm is maintaining, and

说话人2 23:04  
you

说话人1 23:05  
use the probably use the reading property and then you observe that is six. So combine these three comments, you can get the a unit part. Okay? About the b show that actually, this hint tells you everything. So this will be very easy for me, because you get to solve the aid, equality, recursive, which means

since now, we have mi star minus seven, rl something minus one.

Then we can see that it is smaller or equal to this thing squared. Is il minus two. Right? And that about, finally, we reach mi zero, which is empty set and empty set cover the number of elements, which is zero. That here will come back here. Ii shouldn't say this is so here is one minus the ilb and times one minus cil minus one, rmb right?

And then, finally, you will get this new product. X for example, all these times mi star minus m like I zero, which is the enemy set. Good, it will be able to einstein. And then you can you can use the inequality in the hint, which is one minus x smaller input would be minus x you change it to into this. Take us down over the over these powers, and then you can get to the east quality, then foresee. We'll see. So there we see that if the il is the first element, making the lab for evaluate to be false, then that those inequality in a and b also hope so. Then we can simply substitute the ig with ig plus I star. Here is this. I star is the first element. That means the line for evaluate which we evaluate reports, which means after we select the success ig if we add one more success I star, then the budget constraints will be violated.

This will we just use the invited, indeed. It will be larger than one minus as very many since so. Here, we need to use that in the line board is false. Ok which means this cost already be larger than b if it is larger than b then this will be larger than one, right? And this gender, this will be smaller than - 1. Right? And together with this will be smaller than one over e and take the minus here, it will be larger than 1 - 1 already recommend. It will be 1 - 1 will be larger than this gfi I see several students like said the they do not use this I star, they just it at the larger than. So they use the inequality in b with il into ig but then you cannot use this budgets, violation property, right?

So no seven, whether this is a larger or smaller, this is c and let's see the third one, although although the greedy algorithm in the first question, we see that greedy, this greedy algorithm actually performs bad, right? Not so good. But since we have the c in quality, then we can use it as a subroutine of the modified radio with them.

Now, our one factory algorithm just combine two solution, just like the we learn in the lecture in the lecture, also a and there is a combined solution.

Combine. The first one is the brady algorithm solution, and the other is the

it's just a single success that is visible and also cover the maximum number of elements. Then we can say, so I just follow the step in the answer I upload. Here you can on the message. This I first, our solutions, our final solution. If I is the set returned by our new modified video with them, this is aa number of elements covered by current solution. This is my because it takes the maximum one, then it will be at least half the average of these two solutions.

This is the first property we need to use, which is the maximum one. If we take maximum of the two, then it will be larger than the average of these two.

说话人2 31:18  
Right?

说话人1 31:21  
And then here we need to use the, since I this I star actually is not.

Now we want to use the two c in inequality. This I star is

说话人2 31:39  
a

说话人1 31:40  
is the success with maximum cover the number. It is larger than it plus e star. Here we need e this e star is the element is a subset, makes the line form of the video algorithm. First evaluated to be false. Fy star must be larger than an e star, because the definition of s the s star just is the maximum value.

Also, we have an assumption that each single element, each cost of the success is smaller than b fe star is an arbitrary kind of arbitrary success with that economics in the budget. This ipr is a single success that also does not exceed the budget and cover the maximum element. This fx must be larger than fp stuff. And then is the coverage property, which we can put the ig and e star both into the bracket. Why is this? Because e star and ig they may have intersection if they have some difference intersection, then those elements in the intersection will be added twice, right? When we come to twice here and here. But if we first union, this last means union, right? If we reverse the union and then count the, then calculate the cardinality of this set. So it will be smaller than this, because, yeah, because the part of your system, and here just use the two c's inequality, which is 1 - 1 origin, fi stuff here is then we get to this half of 1 minus 1 point.

Although brilliant is a is bad, but it doesn't provide us something that you can use. Do you still have some question about this? Which step do you think not natural or

说话人2 34:40  
not? Right?

说话人1 34:52  
Please.

Now I will simply talk about the in class exercise. First about the of the scheme, which is if you have some days, if there is something wrong with your problem statement, because I believe professor lee always says that in class exercise is used to like the attendance. And so you need to listen carefully about what he said. So the problem statement, it is like 20, 20 marks for each each question. So if there is something wrong with your problem statement, you will be deducted by 20 marks. If there is something wrong with your answer, only five marks will be deducted. If your answer is probably rather than ten marks will be that. So actually, I I didn't very carefully check your answer. Oki just carefully check your problem statement. Let me see.

The first in class question is about the final problems that is not any hard, but can be represented, can be formulated by linear programming. So I see some of you said the short past problem. That's a wonderful choice. And assignment problem also, but some people says it will be a very example with the specific number. So it's about the difference between problem and instance. The problem is a very general with parameters. And the instance is a problem

with those parameters specify. So if you just give me an instance of problem with all the premises, be the specific numbers, then it will be definitely not hard to solve, right? Because its size is very small, but but the problem became grammar. If those parameters are very large, then the problem could be very hard. You need to give a problem, instead, of instance, which is with parameters, with general parameters.

And also at the same time, your problem is not very hard to solve. Someone also gives a like a special instance of that set things, so which is definitely not right, and that is the part. But if you give very a simple example, next step, then it can be solved. The second one is what? It's about the zero one aspect problems, really, the randomized reading problem. The professor actually asks that whether that randomized algorithm can find optimal solution always, but not that whether it can sometimes find the optimal solution. It definitely can sometimes buy powerful solution, but not always is the key score points.

And the answer is not always. You can find some counter example, which is, for example, let's first review the recognize the reality. The red one, 01, that's a problem. Every item has a size and profit, you and you have a bag with certain size and how to give items within size budget to maximize profit. Then he said we start with a really solution. So note that we have two possible reading principles. First is each time we choose an item is the largest profit. This is first, and the second is really reasonable, is each time you take an item with the largest

profit per cost, the profit per size.

This one mind now is a start with a great solution and remove randomly chosen items from the nasa and replace it with other items, and then gradually says possibly using the same reality to choose so that if the result is better, you can also go back to previous to show that this recognized algorithm cannot always find out the solution. You need to give a counter example where the algorithm can never reach the optimal solution.

For example, we have three angles. For example, let's try the the breeding reasonable, which is each time you mean, item is the largest value precise. For example, ababc is reactant standing here. Science is 114, and the value is 237. It is very nice, but it suppose it is four. Then this value presented to three. And this is like a one . 75. If we each time select the item with largest value design, then we first select b right? And then second select a and then cc cannot be added into our work bundle, right? Ok now our right hand, everyone start with this solution, and then try to throw away rather than they throw away one item. For example, is throw, either is throw a or throw b then that c cannot be added, right? Because each time you can only throw one hand, but either you believe me, but he cannot be added, or are you delete a remain me? But in most situations, we cannot go into, right? But the optimal solution is actually still right, because it has the largest values and within the budget.

So in this situation, you can never reach the solutions, the outputs of the agency. The other countries are always like the if we start with the greedy solution, we should keep the item with large value. Then we can construct like abc before item, and we have signs like new ones. This value is the 6637. For example, if we start with a really solution, it's the largest value, then the really solution will select 90. Now our side is three, so we only have 1 %. Then we will ask you, this is at least. And so either we throw c or d if we throw c then nothing can be added, right? Because we cannot add a or b if it's relevant choose d this role, then we can add a or b so like we can reach this da or db I don't know it's wrong on cdo ca or cb then it will get value is larger, better.

Number nine DC a it is just DC and the instruction it will be. And it changes smaller. So its maximum is ten, right? Ok and that's the optimal solution you want. A and b is 12 ok so it just gets stuck. Stop here, stop the CD okay, so some of you says it will get stuck in some local optimum. But you need a company now is the end of this class. I will be the first in the basic 21, 3, and 56123, and motion.