Systems Analysis And Design

Prof. V. Rajaraman
Dept of Super Computer Education & Research
Indian Institution of Science, Bangalore

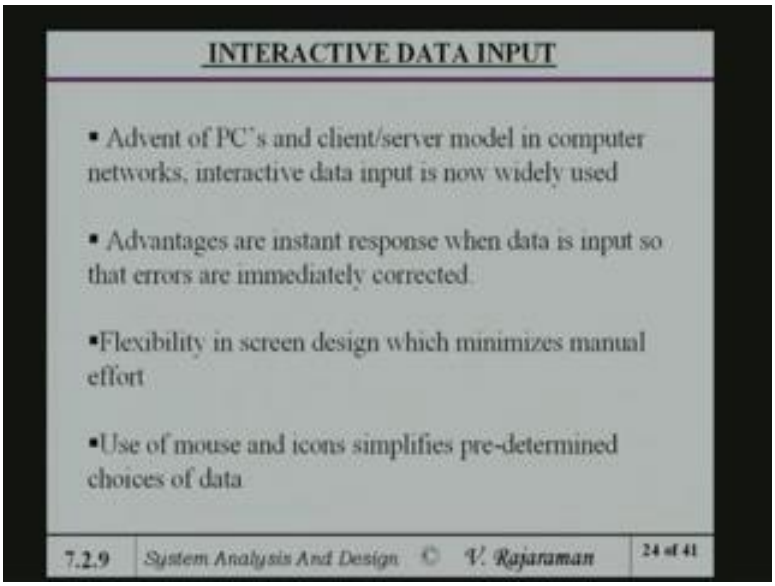
Lecture - 21

Last time, we were talking about data input methods, I will continue from there. The last point, I was trying to make was that if there are form in which data is filled up data is filled up, that is, the manually entered by an operator and manual errors are common. And because, one has to take abundant caution to plan the input form, such a way that the probability of error is reduced. But it will be ideal, if you can have a machine which can automatically read the form; interpret, whatever is in the form. And then, convert it into a computer readable file.

Unfortunately that technology is not yet there. Other words the technology to be able to read handwriting and interpret it absolutely correctly without any errors, which is of essential. When you input particularly forms with critical information is not matured at this time. Hopefully in the few year's time, it will mature. But at this time, it does not does not mature, the result that one has no alternative, but to enter that manually.

Except in certain cases, where you are asked to darken certain boxes like in multiple choice questions. In a examination, where last large number of candidates take part. There automatic computer readability is possible, because there is no real handwriting recognition involved. So, it is only a mark sensing, which is needed. Mark sensing is quite mature and automatic grading of papers, which are really blackened. Now, some box are blackened by the candidate is, now commonly used in many organizations. So, now, we will move on to another method of inputing data called interactive data input. Interactive data input has become currently very popular with the advent of PC's and also so called client server model computing. Other wards, there is a large computer in the background which is called a server, which is got a huge database methods of archiving data and so on.

(Refer Slide Time: 04:38)



And connected to a local area network to a number of PC's, the PC's are roudimentary pcs. In other wards they are inexpensive PC's without too much of memory and does not need lot of lot of desk and so on. So, these are used for data input low cost PC's might say. And so that the input is interactive other wards the instead of the form being filled up. There are number of methods have interact to data inputs, that I have been discussed them, where the operator inputs data on line that is PC connected to the server.

And while, it is being connected to the server is there any error in the data input. Instantaneously a program inside the PC will detects such an error, if it is an obvious error. Like entering a numeric in a alphabetic field or vise versa in 7 alphabetic field putting a number. And in a currency field in 7 entering dollar or rupees or euro, if we enter some other like he enters something like percentage and so on. Then, it will automatically point out the error. And before any damage is done he has to correct it right away.

So, the amount of time required to later on check them out is reduced. So, interactive data input has the advantage have instant response even error is made. And you can design the screen in a flexible manner. In other words you did not have to cramp the entire form into one screen. The screens can be number of screens one after the other and the forms can be very well laid down and one can enter.

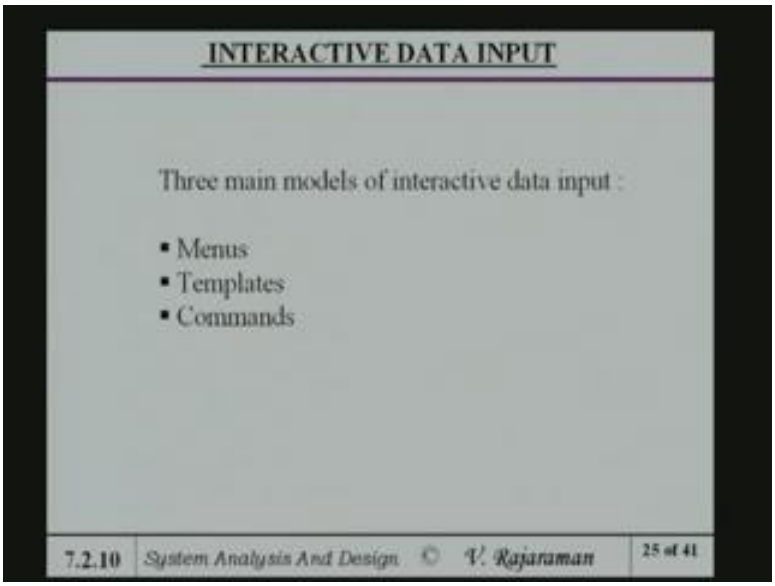
The forms in a much easier fashion, then in by inputting manually from a form on to that, so because, is not always true in every case, but there are many situations, where this is possible. Use of icons and mice mouse simplifies the predetermined size of data in other words the picking of a mouse and moving an arrow and so on. Is so simple, that it is very convenient to input the data and interact to the input as gained an importance particularly because a lot of input is now entered by users like if we have an ecommerce site.

Like you want to book a ticket by train in the on the computer, then you have to have the user you have to enter all the data, we taken by the reservation program. So, the data entry particularly by the user has to be checked, while it is being entered to flow out incorrect data. Otherwise, there will be a problem downstream. In other words, when it goes to a reservation program reservation program will not know what to do.

So, you have to be able to correct or any errors the user makes in data entry. Because, alter all users are not expert typist. They are now is typist. They use a single finger to input data and probability of making error is quite large. So, you should really be able to make sure that for particular users the forms are designed for interactive data input or fool proofs. Otherwise, it is a can be done by any non expert, who does not has to know much about the computer except that the keyboard is there.

And we have to enter it something on the keyboard. And that is a reason, why again inter actuary data input has gained an importance. There is there are two situation, one is a user enters data by himself. And the other is the operator also may be entering data by an interactive fashion on the machine.

(Refer Slide Time: 08:57)



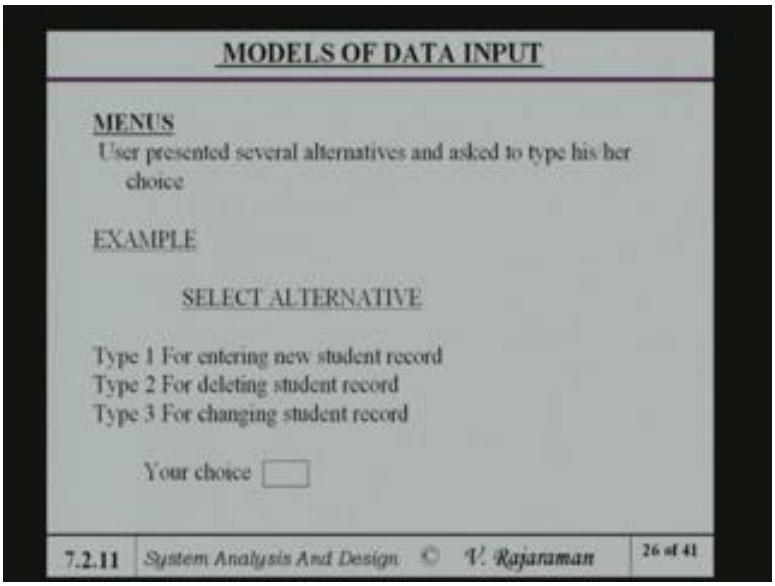
There are many methods of interactive data input. There are three principle methods which I talk about one is called menu driven and the other is called template driven and third is called command driven.

Menu driven as the name implies is that there are certain number of choices given to the user. And based on the choice, there is a kind of you might say tree which is their depending on one choice. It will go to one branch along the tree and provide more branches, which are relevant for the tree. And that way, the interactive data input is effectively adaptive type of a data input by the user depending upon the input of the particular data. It goes to one of number of different alternative parts along the tree.

Templates are essentially equivalent to forms, the forms are in the screen. And we enter on the form in the screen. And the advantage of template is that gain is very clearly demarkets, what is to be entered, where is to be entered and so on. And while entering there will be a background program it will check the data entry. And right away throw out data which is incorrect.

And third one is command based, where the computer ask you certain questions. And based on those questions if we answers those questions. So, illustrate all these three methods of entering data by some example some simple examples.

(Refer Slide Time: 14:55)



Here for instance, you want to enter a new student record do you want to enter a new student record. There is another set of database application, database of students. And when database is updated by the by the administrator or by an operator who has told to update it. One has to kind up ask the update program laws, whether you have to enter a new data record for a student or it is for deleting a existing student record or changing what is existing as a student record.

So, you type 1 2 or 3. So, it is an inter active, you like 1for instance enter new student record. Then, it will go to a branch, where you say give a name of a student. And give the roll number that will start with asking roll number, because roll number is a key. So, enter the roll number of the student, enter the name of the student, enter the year he is end enter the department he is joining and so on. It may be a longer information including things like date of birth, permanent address, hostel address and so on.

So, it will go along the direction, when a new student record is to be entered. If it is a deleting a student, record really ask do you really want to delete. Because, it is about the mistake you type 2 and if you are deleted one some paralogise gone. So normally, a program when you type 2 as a choice it will ask a question you really want to delete. And then, you have to say yes, if you say yes, then it will say give the roll number of the student.

And you enter the roll number and it will bring back the students name and roll number. And again ask is it a right name of the student, because you must have may entered roll number incorrectly. So if then, you have to again say yes, then only it will go ahead. And start the delete operation. And for changing student's record similarly they do want to change the student record. You answer the tree; it will again ask do you want to change the student record. If you say yes it will say give the roll number and if you give the roll number.

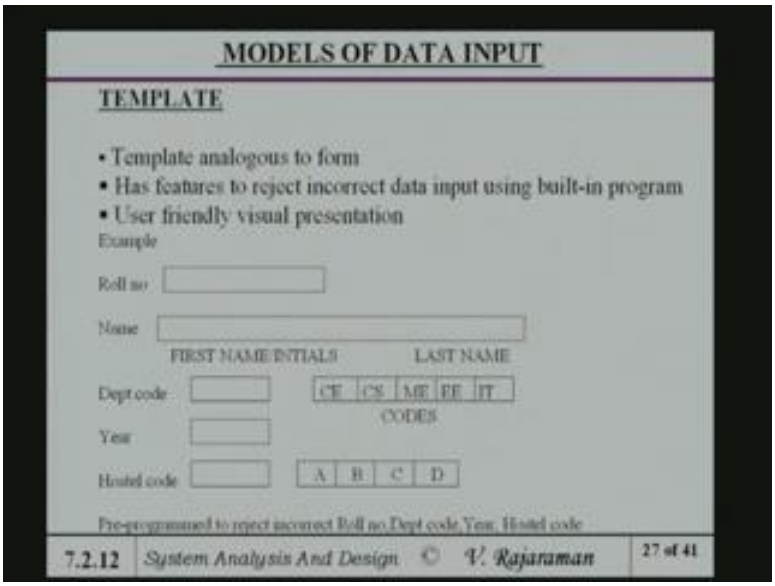
There it will bring out the student record and display it in front of you. And will ask you which field you want to change, then you give the say you want to change the address field. And so, it will then hold the address so and so. And then, it say type the new address it will say and you type the new address and as soon as type new address, you say enter. And it will overwrite the old address will be gone and the new address will be give put it in that place.

So in other words, in the menu driven thing, whenever an alternative is picked. There is a progressive conversation is go on. And the conversation is design in such a way there is appropriate for any other given for usual choices. So, the root matches root of the tree is the first choice. And then, the branches go along depending upon what you choose for the root.

And so, the design of a menu driven thing, menu driven data input is very effective. Because, it is kind of old proof it gives reasonable amount of control for the administrator to design the menu and kind of guides the user. Or in fact, the skill operator to all the steps, because it does not assume and also assumes more or less might say that the any others or a new person who enters is still will he will not make mistakes. And that for incorrect entry in the database can be really fatal. So, that the menu based system is very very useful.

Particularly because, there is a turn over people in any organization and you do not have to go through any large training because almost common sense you have to follow certain path which is predetermine for you if.

(Refer Slide Time: 15:49)



Now, the other is that the second method is template, template analogous to the form. Features reject incorrect data input with a built in program, while you are entering. And it is a user friendly visual presentation. In fact example for a sense entering a student name in the previous case for instance.

We are saying enter new student that is in (Refer Slide Time: 14:55) the in menu alternative was enter a new student. Then, the type 1 it may go to the form. So, the final entry find make that the tree are not exclusive. They can be combined have appropriate. So, this operators of the user enters choice 1, then it will go to base ask for the student roll number.

And ask you to give the name and while entering name, it also tell you, what has you has enter first what you have to enter last. And first name as wishes as many people particularly in south India. Do not really has a final out have the first name last name concept. So, you can assume that you have an initial like my name is vvb is my initial my name is Rajaraman. So, I effectively kept v as the first initial and enter r a j r a m n as the last name.

For this name or it is a first name, last name business, then you can put a first name some people have first name initial last name. And there are some people have multiple initials like vkrv Rao. And put all those initials and put the last name. And in case of surname

like Rao, the initial become very important. Because there are even if, you take the institute, there are enormous institute may be about at least 30 or 40 different Raos.

And only way, we distinguish between them is by looking at the initials. And so, initials may be more than one, one initials that, why I say first name as initial. Then, you put a department code; there is a given choice CE CS ME EE IT, whatever. It may be a two letter code or it could be a two three letter code took. Normally, we use a code with somewhat meaning full like CS could normally mean Computer Science, CE could normally mean Civil Engineering and ME Mechanical EE Electrical and so on.

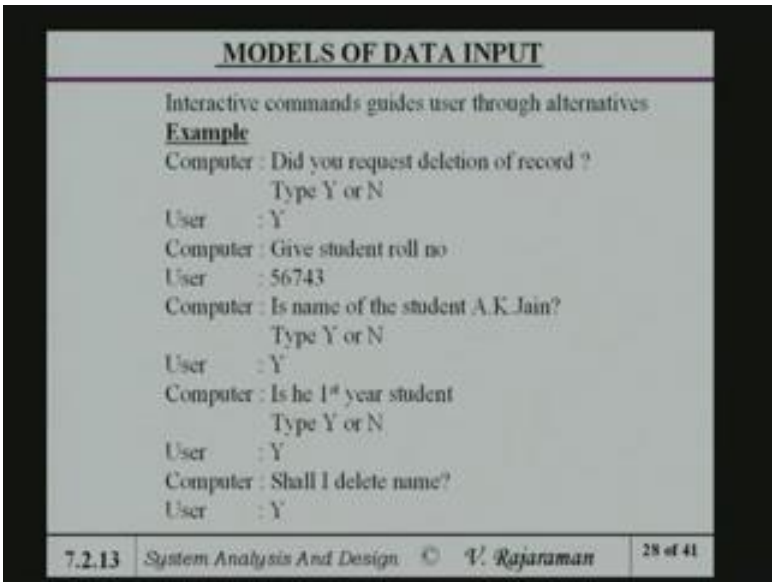
So, you normally try to put a meaningful name meaningful abbreviation might say or code. And if suppose large number of departments which are similar like civil engineering and chemical engineering both could be CE. So, you try to put a one more letter like CHE for chemical engineering and CVE or just CE for civil engineering. So, that depends upon the situation is of in a normally two or three letters would suffice.

Mostly three letters will be more than sufficient for representing a huge number of different disciplines. By enlarge people are satisfied with three letters. Because, the point of course is the machine is going to process. But, the putting more letter is again going to meet to error possibility. Lesser the number of letters the better of you are the lesser the data entry the better of you are. So, the principle point data entry is minimize, the number of keys close. So, one would like to minimize as are of possible.

And then the year, now again in this case the year 1 1 1 box., suppose a person enters 0 6 it might not accept it. So in that, it will say enter the full year 2006. So, you are asked to enter alpha digits. Or if you want to be lot work careful on the right hand side, you might say enter all four digits of the year. Or in the year box itself, you might put four boxes and say enter all boxes and similarly enter hostel code and so on.

So, the pointer is it will guide you. So, the template will guide you to each of them. And ask you to enter, if there is an error made or if there is a misunderstanding at that time itself, the machine is going to pointer out.

(Refer Slide Time: 23:56)



Interactive command based menu example is gain. The menu is somewhat like the menu thing. Did you commands is their like the menu thing you had gives your choice 1 or 2 and so on, whereas in this case the machine will ask you a question did you request deletion of record. The user says yes computer give student roll number is that. You want to abundantly cautious; you can even ask the computer can really ask you really wanted to delete the record type yes or no. And you again type yes or no and it will go ahead the computer give student roll number.

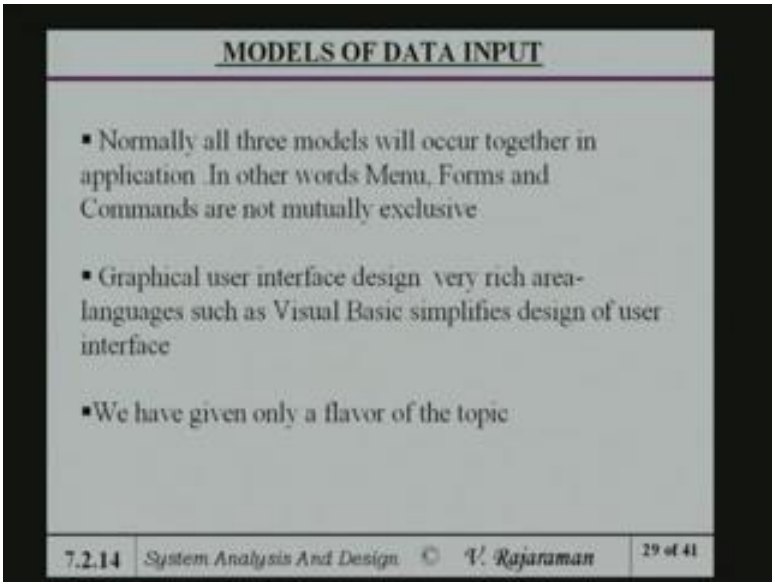
And you give the roll number under roll number is a key frame; really an error is made at the entry of the roll number. Then, that could be a later on difficulty, which arise while processing. In fact, one of the most common error which is made in data entry what is called transcription error instead of putting 56743. One might in or the roll number may be written in by hand. And 7 may look like a 1. And so, he may enter 56143 and 56143 may be a legal roll number it may belong to some other student name.

So to avoid that, you normally would like to have some automatic error detection in key database. In fact later in this lecture, I am going to talk about methods of automatic detection of errors particularly in key database like roll number, account number, pin codes of like that. There is any identification number or identification how is a process may have. So that will talk about later, but I am assuming at this time that. There is this any error is made the machine going to point out. There is not a legal roll number.

And then, it will go and retrieve the file and ask the question is the never the student Akj type yes or no. And you because, you wanted a delay you have entered data about that student. And this is a first year student yes or no you initialized delayed may again it ask and you say yes. Then because, after the name is given and after the year is given may be you made a mistake. In terms of the original roll number and say said if you enter 56143 is of the 56743 142 belong to some other student.

So, it will ask us question is the name of the student Akj and what you would have is Aakj Rao. Then, you say may you wake up and say it not really Ak Rao. So, you could know then it will say retype the roll number. And that is the way it will go in terms of command. So, the advantage is of any entry error which is made is automatically detected while entry is being made. That is the greatest advantage how interactive data input.

(Refer Slide Time: 24:39)



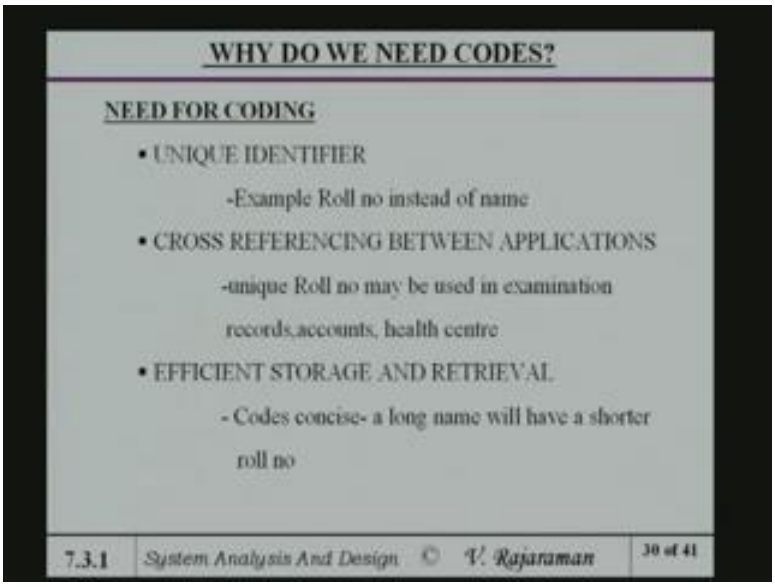
As I said all three models will occur together they are exclusive. Menu forms and commands are not exclusive mutually exclusive. Graphical user interface design is nowadays very simple simplified. Languages like visual basic and others allow you to work on the screen very effectively. And so, the screen you can do the work. And screen based system are easy to design.

And because this particular examples, I given only a flavor of the subject; one has to actually sit down and do a lot of his or her own design for conversations as well as

templates and menus and so on. And then, you will learn a particularly they if you are given a project, it is normally a case in all undergraduate courses. In the project, if you have doing something like hostel management system or hotel management system or reservation system or what have.

You will be required to design all the necessary input form and you go back. And use their ideas as applicable for the particular and exercise and that you will learn more.

(Refer Slide Time: 26:10)



Now, the next topic in data input is coding of data, what is coding? Coding is to give unique identifier is of something which is a non unique like for instance, why do we use roll number of a name? The reason why you use roll numbers of a name is because the name is not a unique identifier. But by take Bangalore telephone directory and look at say r Rajenna or something like that. There are five of Rajennas such a common name in Karnataka that r rajenna is so common and so, r rajenna cannot be a unique identifier.

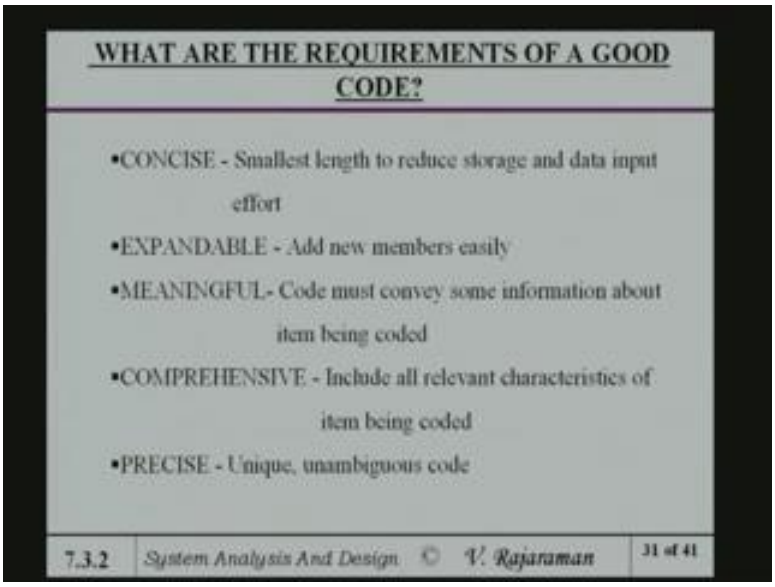
So, we want to kind of give different identifiers you really have to have some method of coding that name into a number and because, you are distinguished between the really different Rajennas. In the case of course, telephone directory telephone numbers are different. And the addresses of course, the address are normally different. And the addresses will effectively you allow you to kind of tracked out a particular r rajenna required.

But the point, I want to make is that names are not unique, whereas some codification is unique. So, you need a unique identifier. And cross referencing between applications, what I mean by that? Is that in a in a student roll number for instance is a very important kind of a data for a student. Because, the roll number is used in the examinations it is used in your result card. If it is a residential campus, it is also used in the hostel accommodation.

And it is used in many many situations in residential campuses the roll number is also used for if you want to use the medical facility and so on. So it is a key identifier, the key identifier is something which is important. And it will be used in accounts department for your fees. So, you need a number key number 1 number 1 unique number to cross the reference between different applications. Put in account applications like fee collection an examination results application and so on.

And third important reason, why we code, is that storage and retrieval is efficient. Because, the length of a code is much shorter than say length of a name. The name, that the particular person's name is long name. And then, as said apart from the point of result is not unique. It also becomes a lot of data entry, whereas these work more concise. In terms of the amount of storage interface and also retrieval based on a key is much faster. Because, it has to only go and search in a given subsets of may be a numbers that is a roll number.

(Refer Slide Time: 29:39)



What are requirements of a good? Good code because there are many many coding methods. I we will look at various coding methods. Before, we said we know now unique identification cross referencing and efficient storage, why and how do we delete and what are the requirements of a good code. And then, how do we do it that is way we go, why what and how.

And what are the requirements of a good coding scheme to be concise. A small well, Smaller the length, the smaller the amount of storage required and less is the data input effort. And as I said the more the data input will more is the error probability. So, it should be expandable. Other words, when you when no system is static. When a student new student joins in a college, he has assign a new roll number for that person.

And so, the code should be such that when somebody new joins he should be expandable easily. You can add the roll number without any great difficulty. Like in a bank, you know savings account, current accounts, many accounts are there. At everyday somebody starts new accounts. And somebody starts new account you have to give a unique identification number for their account to be able to retrieve and so on. So, it is an expandable.

Any system, unless there is a capability of expend it, it is useless. To if it is meaningful it is very nice, what a meaningful is that? By just looking at the code you must have some idea about what that code represents. So meaningfulness like, we have a roll number is just a substring of digits, it is not meaningful. Similarly if, we have a string of digits for an account number in a bank, it is not meaningful.

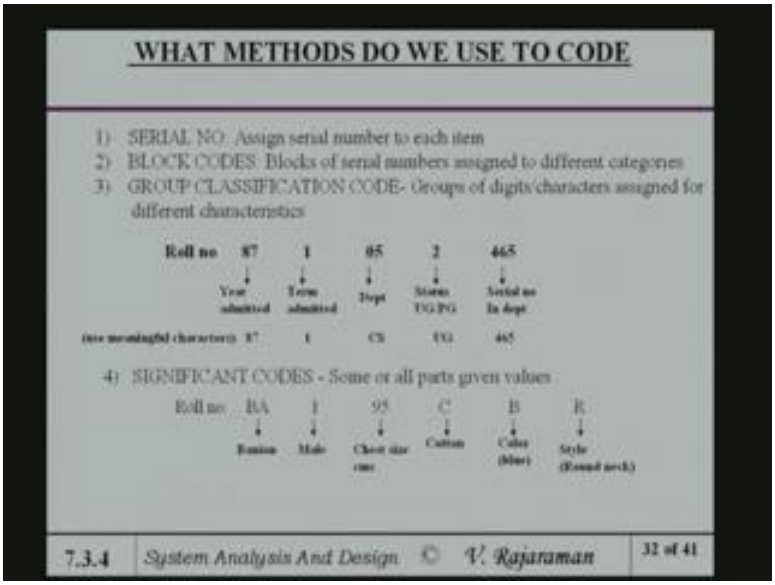
Whereas, if you say SA for savings cu for current and then say and then give some number. Then at least, it is savings account and if it is CU, it is current account and it is bits more meaningful than just a number. So meaningfulness, I explained how you can make meaning meaningful code? Some extent if I want to make code little bit meaningful it become less concise.

In other words is that the requirements are not, what about say not conflicting. If you want to meet one requirement, you may have to kind of let go or kind of loser some other requirement. But, you have ultimately seen, what is the best coding method? Must comprehensive all relevant characteristics of the item should be included say. And again, you want to make it comprehensive and meaningful it may not be concise.

And precise unique unambiguous concises and precises are different. Concises is smallest length precise this for instance. Meaningful code may be precise, because is unique unambiguous. In fact precision is a necessary requirement for all codes, say if a code is not precise it is not really a code, because a code is a unique identifier. So, precision is essential or necessary condition for a code. And the conciseness is not a necessary condition for a code.

It is a desirable condition for a code; you might say that there are certain things are essential. And there are certain things which are desirable. Preciseness is essential, whereas others are desirable. So, depending upon whatever code you pick. The essential qualification must fulfill. And then, you look at the desirable qualifications.

(Refer Slide Time: 34:04)



Now, let us look at methods of coding set serial number code is the simplest, what I mean by serial number code is? As manner student joins, I give next serial number to that student, there were 358 as the last serial number, last whole number. When a new student joins I put 359. Similarly in a bank, when a new account number, new stud, new sub the bank subscriber joins wants to become account holder.

You look at the last account number which has assigned and you just add one and give next account number to that person and that is the serial number. Serial numbering is by nature not meaningful. Because, it is just a number assumed as a procedure just as giving serially one after the other, but it is not meaningful. Block codes are where, you give

blocks of number. Within the block, you give serial number like some instance a block may decide that any number between 1 and 1000.

In a current account number may be which is between 2000 and 10000 is a series account number. So, the two different blocks and I whether it is 1000 and 2000. If it is say the of course, zero you will not normally assign may be comes from 1 to 1000, whatever 1 to 1000 or it is between 2000 or higher whether it is a current account or saving account.

The block codes are expandable, because it is also it is little bit more meaningful than serial number. Because, from the blocks you can know this particular account type it is. Similarly, roll number also they may give blocks. Like, you may give first block of 100 numbers to the civil engineering departments, next 100 numbers to mechanical engineering and so on. So, you may give such blocks and call a block code. And do classification code, let it not be more meaningful, where groups are give the characteristics of the item which is being coded.

Now in this case, the roll number is being coded. The roll number to make that more meaningful, you might assign a roll number which is not just a serial number. But, they it might say the year admitted, but first two years may be the year admitted. Term admitted was admitted in the first term or second term. And department code may be there. 05 may be computer science. In fact, you want to make it a little bit more meaningful instead of 05. You have put CS, because character codes bit for meaningful.

Then, the numeric code status UG or PG, you can either put 1 or 2 or you put UG or PG. And serial number in the department, so in that department UG students, they are given certain serial number only required. So, it is entire 213569 9 digit code 9 character code. You might say now becomes roll number, but it is not more meaningful. As soon as the root locate the roll number which year the student join which term he join which department he is in and whether he UG or a PG and so on.

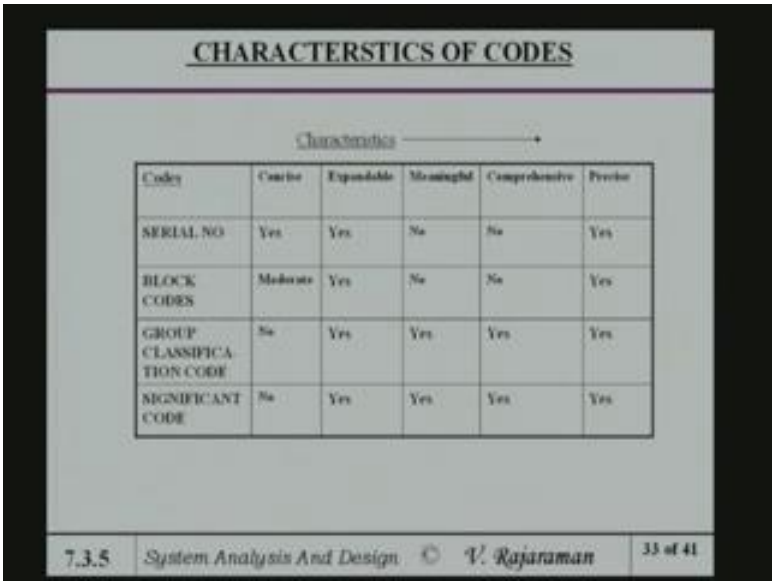
And this will be useful in certain other situations well like know fee calculation or any scholarship to be given. Because, PGs may be given scholarship, UGs may not be given scholarships. So that kind of a thing, you can be used later on in certain programs is fact. So, it is called a more meaningful. Meaningful characters can be used or numbers can be used certain ideas. Other wards these are called group classification. This something

called as significant code, where code itself has some parts which represent the value of the code like for instance manufacturer may code.

The actual wrest or banyans are make as the derivating baniyans and under wear is making socks and so on. So depending upon, what the item is, you may give BA for baniyan. And for itself meant for ladies or male or female, if it is a 1 or 2 depending on that. So, look at the structure size see that is now a quantitative description of that item. Because, this is actually a number, which represent some value corresponding to their baniyan which is mainly chest size 35 is chest 95 is the significant chest size. So that number is given.

And cotton C is for cotton and color blue color or red color whatever and style if it is round neck or v neck sleeveless or whatever. And so, depending upon other than use different codes, so this is called significant code, significant code somewhat meaningful it also incorporates some values which are of importance in the item.

(Refer Slide Time: 40:07)



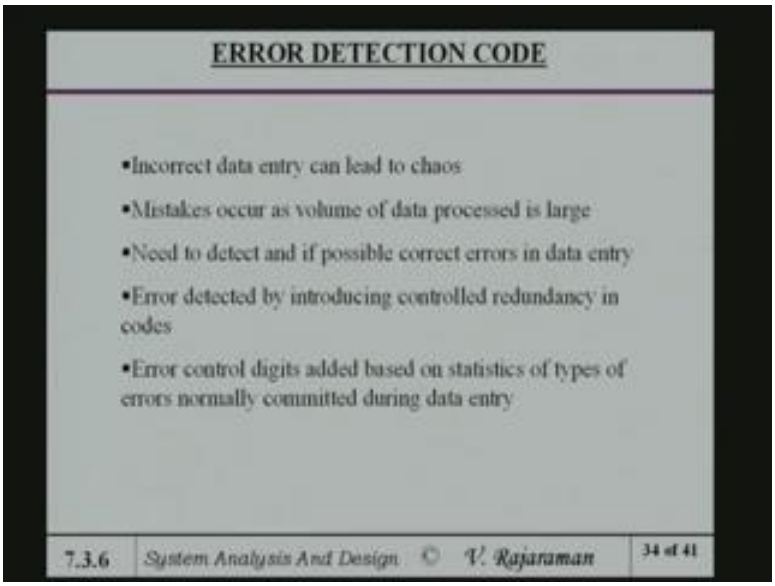
So, we compare serial number code, block code, group classification code, code and significant code. There are number of qualities, which I talked namely concise expandable meaningful comprehensive and precise. You see the last column all of them are precise, because I said pres the precision is a essential qualification. So, all of them fulfill the essential qualification.

Serial number is not concise, it is a expandable and it is not meaningful. It is not comprehensive, what we mean by comprehensive is does it really kind of cover all characteristics of a code. In other words, you might look at the some extent the roll number somewhat comprehensive, because it gives a lot of information about that. Similarly significant code is also somewhat comprehensive.

So, block codes are moderately concise, because there are blanks in between blocks. Expandable it is not meaningful, it is not comprehensive. Because, precise group classification code it is not concise. Because, added more the necessary. It is expandable, it is meaningful and it is comprehensive. Expandable is very simple, because I have the serial number at the end.

And if new students join in the computer science department, I give the next serial number to that person. And significant codes are also not concise, they are expandable, they are meaningful they are comprehensive and they are precise. So, if you really want to have a reasonably expandable meaningful comprehensive code. It is better to pick other group classification codes or significant codes not serial number or block, so even though they are not concise.

(Refer Slide Time: 42:15)



As I final out incorrect data entry can lead to chaos. Mistakes occur as volume of data process is large. Other wards, major difference would become problems. You do in projects in the college. And real field situation is of in the real situation the number of

data items might may where it will be lakhs like CT number of students occurring in CT will be a very large number.

So, the volume is very large. So, I need to that when you volume becomes there is definitely a need to detect and correct errors at data entry stage itself otherwise you get into problem later on. Actually error is detected by introducing control redundancy inputs. Other words detection of errors always requires some extra digits to put in to the code which help you to detect the error.

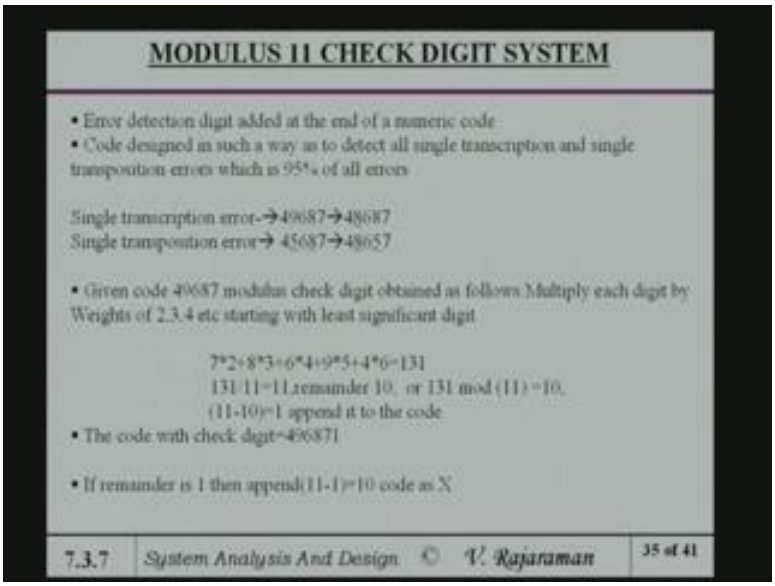
Like all of you in a digital computer design codes or logic codes would have talked about the parity bit. The parity bit is to set number of bits of database you can add a parity bit. And total together even a parity fields. Then, you can say there is a single error in one of the bits. And code say error detecting code.

Similarly, in of you use a code like decimal number then instead of parity bit I must have a extra digital number at the extra digit which is a error detecting digit. It is also possible to have a error correction also. But, as error correction as I said redundancy becomes too much. And so, normally people are satisfied with error detection. So, you have a control redundancy there is 1 extra digit. Say if you got a 8 digit number you add a 1 extra digit.

It does not mean that this kind of error detection can be done only in numeric codes. It can be done in any kind of codes in numeric as well as alpha numeric. But, I restrict myself in my discussion to numeric codes to illustrate the points. And it can be expanded to alpha numeric I tell you how to do that. And error control digit varied to the given code to actually detect the error. Not before detecting any error, it ordered to kind detect any error. First of all, you got to find out what type of errors do, people normally make in detect.

So, in case of parity bit we assume that the error occurs in transmission. And so, it is a random error any one of the bits can go wrong. And people assume that only single bit can go wrong. Single bit parity detection is what is done, in with odd parity or even parity. So in this case, it is it could not need, not be a random kind of a thing. But, there are because human beings, they are entering the data. So, the human beings have certain tendency to make certain type of errors the biased to certain errors. So, you must have some data about the type of mistakes do which people make.

(Refer Slide Time: 45:52)



So, people have actually done this statistical study. And they found out that the kinds of errors they make are of two types. One is called a single transcription error like for instance you look at the number 49687. If a person enters 48687, then it is called a single transcription error. Single digit has been changed from 9 to 8 is called a transcription error.

And that is very common 1 digit, because you can instead of one putting a 7 or instead of 4 putting a sometimes 7 and things like that. So, it is kind of careless mistakes mighty. And the other is the transposition error, because you type pass you transpose the S the 445687 becomes 48657. So, it is called transposition error. That is two digits are transposed.

So, 99 percent of all errors are either transposition errors or transcription error. So, if you can come up with the code which will detect 95 percent of the errors, then you are reasonably alright. And in fact, the method I am going to talk about guarantee to detect all transcription error, single transcription errors and all single transposition errors. It is not guaranteed detect multiple transposition errors or multiple transcription errors.

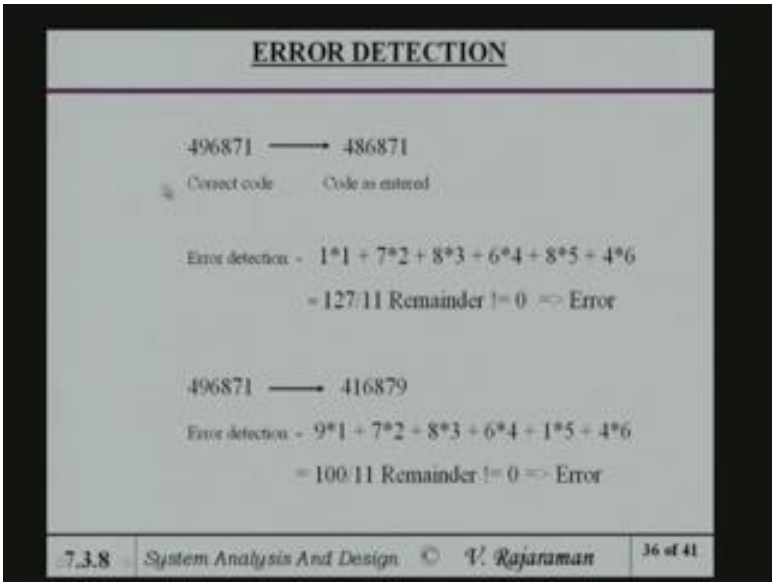
But, there is a certain probability. They also in the countdown probability with which they can be found out are reasonably a large. And people are found that over something like 10000 data items. They use an error detecting digit at least 9000 out of 10000 items, 9998 will get only two will be left undertaking. So, the single error detection code is a

very useful and a column code called method of coding is suppose you are given a code say 78678694, 78694 is the code.

In fact I have taken the same thing. So, I put you know the least single digit here 78694. So, I start with the least significant digit multiply by 2 next digit multiply by 3 next by 4, next by 5 next by 6, add all of these. And if I add I get 131. Divide 131 by 11, I will get a remainder of 10. So, there is 131 mod 11 is 10. So, append 11 minus 10, 1 that is whatever is the remainder. So, 11 is the number I took, while I picked 11 that is not a magic number. I will explain why 11 is picked 11 minus 1 is 10, I append 1.

So, new code becomes 49687 and then added error detecting digit 1. Suppose in this particular division, the remainder becomes 10, what do you want 1 digit represent one place. So, I use X if it is remainder is 10, I use the X if the last digit place. So, this a way the code is formed.

(Refer Slide Time: 49:54)



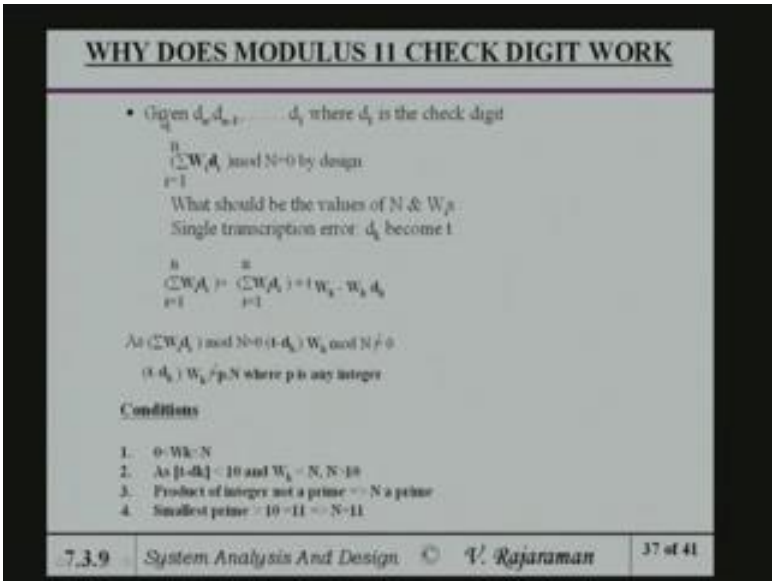
Now, the how do I detect the error. So, this correct code 496871. Somebody entered 486871. Instead of 9, it would have 8. Then, error detection is again you take this code 1 into 1 7 into 2 8 into 36 into 4 8 into 5 plus 4 into 6 become 127 by 11. Remainder is not equal to 0, the remainder, because the way in which designed is sum of all this weighted sum is divisible by 11. That is how I made that is, why I used 11 minus 1. The remainder has the appended number.

So that is of why I decide. So, if it is not zero; that means, it is an error. So, immediately I know that is an error. So, if I 496871 and just now doing a transposition error like without the extra digit I added can also be errorless. Other words somebody can make an error by entering this and this in place of 9 and 9 in last place. So now, it comes single transposition error.

This case also error detection is there, because for that any single transposition error or any single transcription error and be detected by this coding scheme which is called as modulus 11 coding scheme. Because, 11 must use as the number by which I divide and 11 is a prime number. And why I use 11 will be clearer when I explain theory. So now, if I take this number and again 9 multiplied by 1 and so on and add and divide 100 by 11 the remainder is not 0.

So, I know immediately it is an error. So, the machine what will it do is some code is entered it will have some programming in it, which will do the simple arithmetic calculation divide by 11 and find out the answer is 0 or non zero. The answer is non zero it will say it is an error. If it is a 0 of course, it will accept and go ahead.

(Refer Slide Time: 52:20)



Now the simple theory the theory, explain in detail in this book on system analysis and design. And in fact, in chapter 11 the full theory is given very detail. But, I very quickly say what the idea is, what we did was we took the individual digits. And multiply any each one of them by a weight. And so, we are two quantities we used. One is the weights

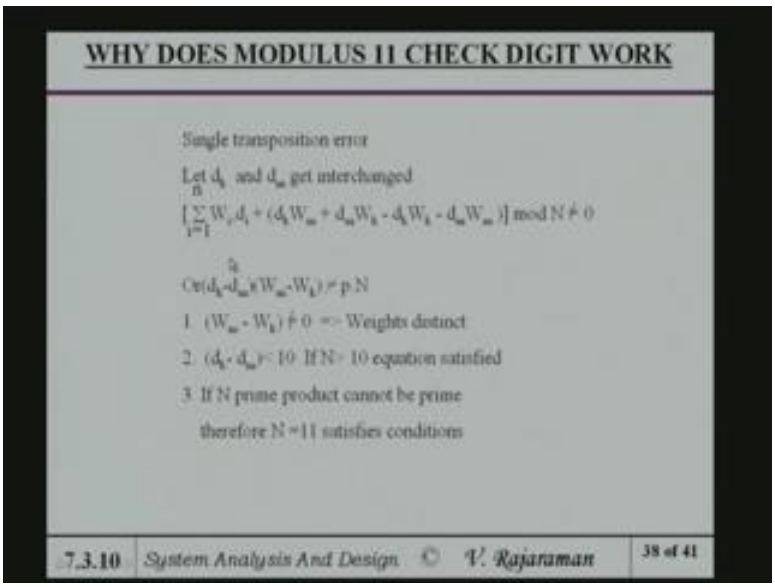
the one question is, what weight should I pick. And then I did model W sigma W I d I model should be 0 by design. So, what should be the value of N and what should be the value of W.

Now I had find out, what value is W? Wi's and what all the N will work. And that is done by just looking at suppose the single transcription error and decay the K digit becomes t. Now sigma W I d I becomes sigma W I t I d I plus t W K minus W K d K. because I 2 N W I d I already had got everything in it, because d K became t. This is a extra term it having because this two are equal.

Now sigma W I d I model is 0 by design. So, t minus d K t minus d K times W K model should not be equal to 0. If it is 0, there are coincidence, even if it an error it will say there is no error. So, conditions are W K should be less than 1. Because, if it is greater than one model can be non zero t minus d K must be less than 10, which of course,, is satisfied. Because, everyone is a digit there, the weight also should be less than N.

The size be greater than 10, the weight should be less than 10. Products of a integer cannot be a prime. So, 10 should be a prime, because the point is this is a product of two integers this product of two integers should not be divisible by N. So, this N should be a prime. So, smallest prime is greater than 10 is 11. Therefore, 11 is prime.

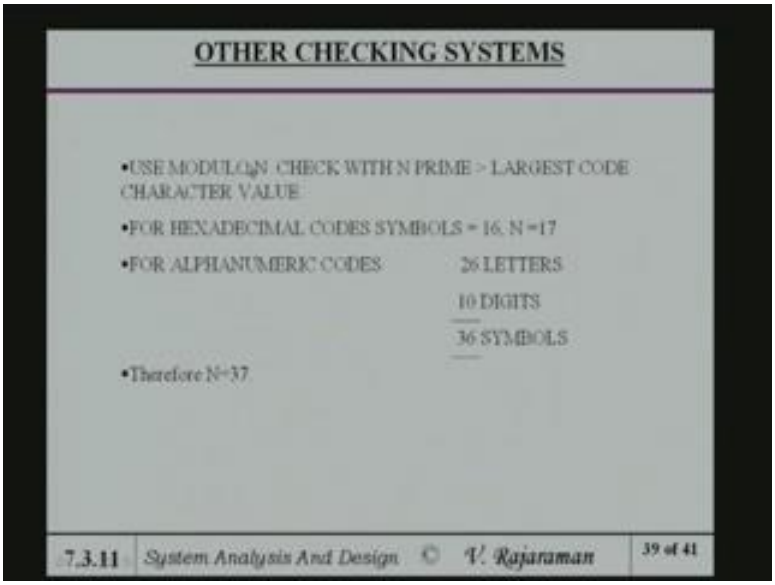
(Refer Slide Time: 55:11)



Now the weights for single transcription error, single transcription error means d k and d m get interchanged. So now, rewrite this formula again W I d I and d k W m is a digit which became k and m got interchanged. And I can write this how and I get d k minus d m times W m W k should not be equal to p times N, when p is an integer; that means, it is it should not divisible by mod N.

One thing is that W m minus W k should not be equal to 0; that means, weight should be distinct and d k minus d m is less than 10. If n is greater than 10 this is always satisfied if n is prime product cannot be a prime. Therefore, N is 11 is satisfied. So, what it means is that the weight should be distinct and unique. They N value should be 11 is the higher lowest prime greater than 10 and no weight can be greater than 10.

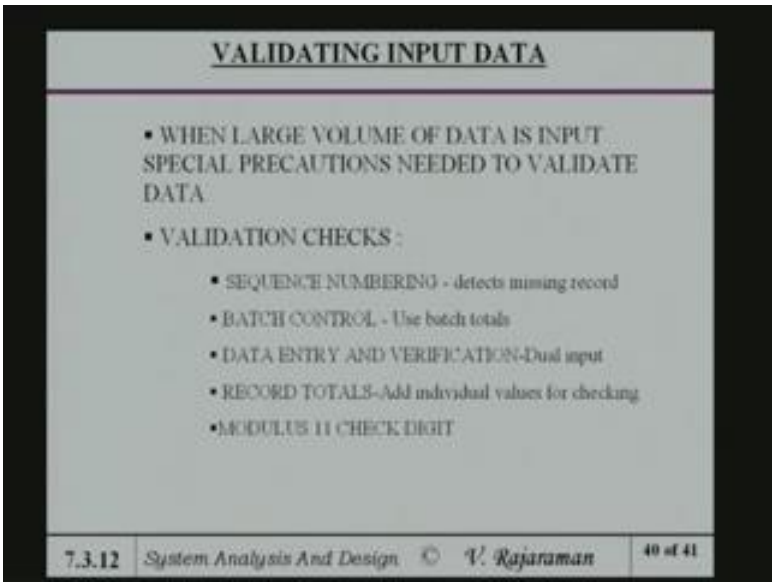
(Refer Slide Time: 56:26)



So, what it really means is that this will modulation check digit system will work, only if the number of digits, in a code is less than equal to or less than 9. Because, the 10h digit will be the error detecting digit, if you use there are more than 9 digits and if you say 11 or 12 digits. Then, I go to the next higher prime. There is suppose there are 12 digits.

Then I can go upto to 12 and next prime is 13. So, it may modulus 13 if it is hexadecimal it will modulus 17. Because, 16 is the largest number in hexadecimal for alpha numeric codes there are 26 36 symbols. And in alpha numeric code I code A B C D as after 10 11 12 13 14 and so on. And I can use a modulus 37, because 37 is the large next prime greater than 36.

(Refer Slide Time: 57:30)



So, there are other way validating checks which are there. And then, large volumes of data is input you have some certain checks called sequence numbering, batch control data entry verification record totals and modulus 11 check digit. And I think I had explained each one of them in much greater detail, what is mean by sequence numbering, what is batch control, why they are required and what is data entry verification, what is data total of course, modulus 11 check digit is, what we already talked about.

So, next time I start from here and talk about validation of data inputs apart from what is there in modulus 11check digit.