

Object-Oriented Analysis and Design
Prof. Partha Pratim Das
Department of Computer Science and Engineering
Indian Institute of Technology – Kharagpur

Lecture – 29
Identification of Classes, Objects and Relationship in LMS

Welcome to module 17 of object-oriented analysis and design. In the last module, we have discussed about techniques of classification and identification of key abstraction for actually bringing a system specification into a concrete design. We have talked about some examples and outlined an approach. From this module actually we will continue this current exercise over a number of modules, we will start then this module and continue over the next couple of modules.

(Refer Slide Time: 01:09)

Module Objectives

- Perform an exploratory exercise to identify potential classes, objects, and relationships for the Leave Management System (LMS)
- The following will be used as primary source of information for LMS (as received from the client)
 - Written English description of LMS (PDF document as shared)
 - Conversation between client and vendor (Tutorial video as uploaded)
- This will be a multi-module exercise spanning Module 17 to 20

Please keep these handy while you study these modules. We shall frequently refer to those

Here our objective is to take the leave management system exercise and actually try to apply all the different techniques, principles, measures that we have talked about so far. Try to apply them may be in a small scale and see how they help us, get to an object oriented design of the leave management system.

So from this point onwards, it will more an exploratory exercise, so you will also have to as you listen to this lecture, it will be good to have the leave management system exercise specification handy, we had uploaded that earlier and there was a video where two of our TAs Tanwi and Srijoni interacted as perceived customer and vendor to talk about the leave management system.

So before you go through this modules, please revise on that specification and the video and keep those handy with you and also keep some worksheet alongside so that you could at times pause the lecture and try to work out things by yourself. What since even though leave management system is a small problem, you will find that it is quite big to be covered in full in terms of this lecture. So I will just try to work out the representative parts and leave the rest of the details of you to go through.

(Refer Slide Time: 02:51)

Module Outline

- Input and Output Specification
- Informal English Description of LMS
- Identification of classes
 - Extraction of Nouns
 - Quality Check
 - Identification of Attributes
 - Structural Clustering
 - Quality Check
 - Behavioral Clustering
 - Quality Check
 - Identified Classes

So with this we will start off with the module, the outline for this particular one would be to start with the English description of LMS and try to identify classes. That is the first thing that needs to be done. So this module, we will focus on identification of classes.

(Refer Slide Time: 03:10)

Input and Output Specification

Input

Informal English Description ⇒

Output

- Classes
- Attributes
- Responsibilities

Class Name	
Attributes	Responsibilities
•	•
•	•
•	•

Naturally, broadly I specify the input output that we are trying to achieve. The input therefore of this exercise is informal English description, the leave management system description and the output would be classes associated with that if in the process we can identify some attributes, some responsibilities, we will try to do that.

(Refer Slide Time: 03:31)

Informal English Description of LMS

A Company wants to manage the attendance and leave of its employees through LMS. The requirement specifications are:

- ① The company has three categories of employees:
 - *Executive*: Employees who work as individual contributors and report to a Lead.
 - *Lead*: Every Executive reports to a Lead who approves / regrets her / his leave. A Lead reports to the Manager.
 - *Manager*: Every Lead reports to the Manager who approves / regrets her / his leave. There is *only one* Manager.



So this is, what is our input, now this is if you refer, please refer to the PDF LMS document at this stage and you will find this is directly lifted from that document. So it talks about the basic system, there is a company which wants to manage attendants, leave of its employees and it starts saying that the company has three categories of employees and so on.

So certainly, in this space I cannot the copy, I do not need to also, copy the whole specification as given in those two pages, you can always refer to that, that is what we are talking about as an informal description to get started with.

(Refer Slide Time: 04:09)

Identification of Noun

- Linguistic analysis of the problem description - extraction of Nouns
- Rules of thumb¹ :
 - Elimination of irrelevant terms ✓
 - Elimination of names of values ✓
 - Elimination of vague terms ✓
 - Identification of attributes ||
 - Identification of operations ||
 - Elimination of terms which are in fact relationships ✓



And please always supplement that with the discussions that Tanwi and Srijoni had been doing for this. So the first task that we will do first attempt that will take is taking this English description, we will try to identify, we will try to extract nouns. So this is an approach specifically used for textual description of interactive notes is called a linguistic analysis.

That is you basically try to parts the description in terms of parts of speech and then based on the parts of speech you try to identify concepts and then you start relating them. So if you are talking about the noun being the major part of speech, then our approach would be to identify nouns which is a easiest to identify and then in the process, we would like to eliminate irrelevant terms and you will find lot of nouns which basically does not have a sense in this whole thing.

We will try to identify, eliminate the names of values, they may not give as classes. We will eliminate vague terms, but will certainly try to focus on identifying which nouns are representing attributes operations, so that they can be grouped into in terms of classes and so on and we may also eliminate the certain nouns which represent relationship and take them up when we are actually talking about relationships.

(Refer Slide Time: 05:43)

Identification of Noun

A Company wants to manage the attendance and leave of its employees through LMS. The requirement specifications are:

- The company has three categories of employees:
 - Executive: Employees who work as individual contributors and report to a Lead.
 - Lead: Every Executive reports to a Lead who approves / regrets her / his leave. A Lead reports to the Manager.
 - Manager: Every Lead reports to the Manager who approves / regrets her / his leave. There is only one Manager.



So let us get ahead with this exercise, I am again reproducing the beginning of the document and now you can see that in this I have highlighted certain words in red. Now these words as you can see the company, let me make the pen tip as blue now, so the company is certainly a noun. If you keep reading once to manage the attendance is a noun and leave is a noun, of its employees is a noun.

Through LMS, LMS also is a noun but we have not highlighted here, I have not highlighted here because it is obviously an irrelevant terms so far as the nouns or classes of the domain are concerned. Requirement specification is a noun but again we have not picking it up because that is a irrelevant term. But company we have already identified, employee we have identified employee.

Then we find a next noun in terms of individual contributor, we found a noun in lead, we found a noun in manager and so on. So that is a relatively straight forward process that you can do but of course you will have to keep on applying your intelligence, common sense and the domain sensitivity to see what could be nouns and what may not be nouns.

(Refer Slide Time: 07:11)

Identification of Noun

- The **company** has provisions for the following categories of **leave** associated with the respective **leave rules**:
 - **Casual Leave (CL)**:
 - 10 **CL**'s are available in a **calendar year**. All **CL**'s are credited to an **employee** on 01-Jan. For **employees** joining in the middle of the **year**, the number of **CL**'s are prorated. **CL**'s cannot be carried over to the next **calendar year**.
 - More than 2 **CL**'s cannot be availed at a time. **CL**'s cannot be clubbed with other **types of leave**. Total **period** of **absence** including **holidays** cannot be more than 4 **days**. **Holidays** intervening the **absence** are not counted as **leave**.
 - **CL**'s do not need **pre-approval**; but must be approved within 2 days of its availing.
 - ...



So if we move forward, this is as it continues the company has provisions for the following categories of leave. These have already been identified as noun. So we already know that. But then it comes with a noun which is leave rules. So we still do not know whether leave rules are relevant term or less relevant or irrelevant but we need to mark that. The causal leave which is short named as CL, so I am just using the CL as a noun.

Then you find several nouns coming here, calendar year, employee had come earlier, year, then we get absence, we get holidays, we get types of leaves, we get pre approval. These are according to part of these are all nouns. So the basic exercise is to actually go through this and keep on identifying the nouns. So I would not proceed further. I would suggested that at this point, you pause the lecture and take your time of to go back to the document and identify all the nouns in the document.

(Refer Slide Time: 08:22)

List of all Identified Nouns *

Company	Attendance	Leave	Employees
Contributors	Lead	Executive	Manager
Leave Rules	Days	Year	Name
Type of Leave	Period	Absence	Holiday
PL	CL	EL	DL
SL	ML	LWP	UL
Pre-approval	Month	Service	Quarter
Medical	Parenthood	Disciplinary	Administration
Certificate	Certificate	Action	Function
Daily	Personal	Calender	Batch Task
Attendance	Details	Year	
Account	Balance	Designation	SysAdmin
Parent	Salary	Week	List
Privilege	Right	Login ID	Leave Status
Employee Code			



Presuming that you have already taken the break, you have identified the nouns then more or less you should get this set of nouns. I am saying more or less because there is always some possibilities of interpretations, some possibilities of missing out items but these are the typical nouns you will get.

And you can see that well this exercise just by itself may not be helping us a lot because you can just count how many we have, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13. We have 13 rows, 12 full rows, so four in each, 48 and last one. We have in this exercise itself we have identified 49 nouns. So we will have to see that how far have we actually understood about the system by then this exercise.

(Refer Slide Time: 09:20)

Quality Check

- The Metrics at this Stage
 - ① **Coupling:** Very low
 - ② **Cohesion:** Very low
 - ③ **Sufficiency:** Unclear
 - ④ **Completeness:** Unclear
 - ⑤ **Primitiveness:** Unclear
- Actions required
 - Identify Attributes
 - Perform Structural Clustering

So what all continue to do is I introduced you to the different quality check parameter. So of

and on we will go back to that and just do a quick check. Basic metrics, coupling, cohesion, sufficiency, completeness and primitiveness, what do you know. In terms of coupling and cohesion, we have very low coupling, cohesion.

We just do not really know which is related to what and whether these nouns are coherent ones, these are, how they are really coupled and so on. So we have very low information about that and regarding sufficiency, completeness and primitiveness, we are absolutely unclear.

So the next step in the design process, we pick up in terms of the what we learnt in the earlier module, that we will try to identify the attributes out of this nouns because an attribute is expected to be a noun and we will perform some structural clustering to see if we can group these concepts better. So let me take you that.


(Refer Slide Time: 10:20)

Attributes

Company	Attendance	Leave	Employees
Contributors	Lead	Executive	Manager
Leave Rules	Days	Year	Name
Type of Leave	Period	Absence	Holiday
PL	CL	EL	DL
SL	ML	LWP	UL
Pre-approval	Month	Service	Quarter
Medical Certificate	Parenthood Certificate	Disciplinary Action	Administration Function
Daily Attendance	Personal Details	Calender	Batch Task
Account	Balance	Designation	SysAdmin
Parent	Salary	Week	List
Privilege	Right	Login ID	Leave Status
Employee Code			

Noun used only as value should be treated as attribute

- Login ID
- Name
- Employee Code
- Personal details
- Designation
- Salary



So here on top again, you have the set of nouns, as we have extracted and let us look at what could be an attribute. So what is the characteristic of an attribute, the characteristic of an attribute is, it should typically be a noun which is used only as a value. That is an attribute does not have a behaviour. Attribute does not have its other attributes or operations and so on. So but it is something concrete.

So you will expect it to appear as a noun but not have anything other than a value. So if we look at some say let us say, if we look at, say leave or if we look say contributor, if we look at a manager, we will typically know these cannot be attributes because we have seen, please go

back to the LMS document very frequently.

If you see through that document, you will find that manager does this, contributor work, leave is approved, leave is regretted. So all these kinds of, so there are several actions associated with this. But if we look at say the salary, or if we look at designation, if we look at personal details, login ID, so on so forth.

If we look into these and there are some more like employee code. If we look into this, these are nouns which adjust use as values. So we can go through that and say that these will be the attributes. So I just trying to do a basic process of classification okay, using this property. So this is also a kind of structural clustering only that we are doing, using the property that we expect nouns to be, nouns treated as values to be attributes, we are tried to get the attributes out. So we can say that these are the attributes that we can get to see right now.

(Refer Slide Time: 12:50)

Structural Clustering			
Company	Attendance	Leave	Employees
Contributors	Lead	Executive	Manager
Leave Rules	Days	Year	Name
Type of Leave	Period	Absence	Holiday
PL	CL	EL	DL
SL	ML	LWP	UL
Pre-approval	Month	Service	Quarter
Medical Certificate	Parenthood Certificate	Disciplinary Action	Administration Function
Daily Attendance	Personal Details	Calender Year	Batch Task
Account	Balance	Designation	SysAdmin
Parent	Salary	Week	List
Privilege	Right	Login ID	Leave Status
Employee Code			

- People
 - Manager, Lead, Executive
 - SysAdmin
- Organizations
 - Company
- Things (Documents)
 - Medical Certificate, Parenthood Certificate



Then you can do a more organised structured clustering by trying to kind of go in, kind of a discovery mode and say that okay, do we see people in this, like since this is a moderately generic domain, this is not a very technical domain like you know aircraft model or a very specific domain model. So we will try to throw in some concepts and see that do we have clusters here which shared the properties of that concept.

So people, let us look at what could be people and certainly we can find out managers, employees, executive, lead, these are all people because they will have the ability to act, take decisions and so on and we also get something like a SysAdmin, here I have kept these two in

little bit two different buckets because their specification says that SysAdmin is a batch.

So we are really not sure that SysAdmin in one sense can be a people that is there could be a physical person and therefore the person can act but it could also be an automated process which could act. So kind of we will put them in people but just try to remember that okay some however it said that SysAdmin is a match. If you look at organisation, we get company that is pretty straight forward, then one certainly main thing to cluster structurally are things.

So we get a whole lot of them, the different kinds of documents like parenthood certificate, medical certificate, these are certainly different things or documents that we can get so in the whole system.

(Refer Slide Time: 14:54)

Structural Clustering			
Company	Attendance	Leave	Employees
Contributors	Lead	Executive	Manager
Leave Rules	Days	Year	Name
Type of Leave	Period	Absence	Holiday
PL	CL	EL	DL
SL	ML	LWP	UL
Pre-approval	Month	Service	Quarter
Medical Certificate	Parenthood Certificate	Disciplinary Action	Administration Function
Daily Attendance	Personal Details	Calender Year	Batch Task
Account	Balance	Designation	SysAdmin
Parent	Salary	Week	List
Privilege	Right	Login ID	Leave Status
Employee Code			

- Events
 - Holiday, Disciplinary Actions
 - CL, ML, UL ...etc
- Temporal Nouns
 - Date, Year, Time, Days, Calender year, Week, Month
- Incidental Nouns
 - Leave Rule, Administrative Functions, Period



If we proceed in doing structural clustering, one kind of clusters that often happen in the system are known as event where something is going to take place. Usually events have a temporal sense, events have a beginning and end, events have a sense of activation and certain side effects and so on. So that is typically an event. So let us look for other events, the certain things happening.

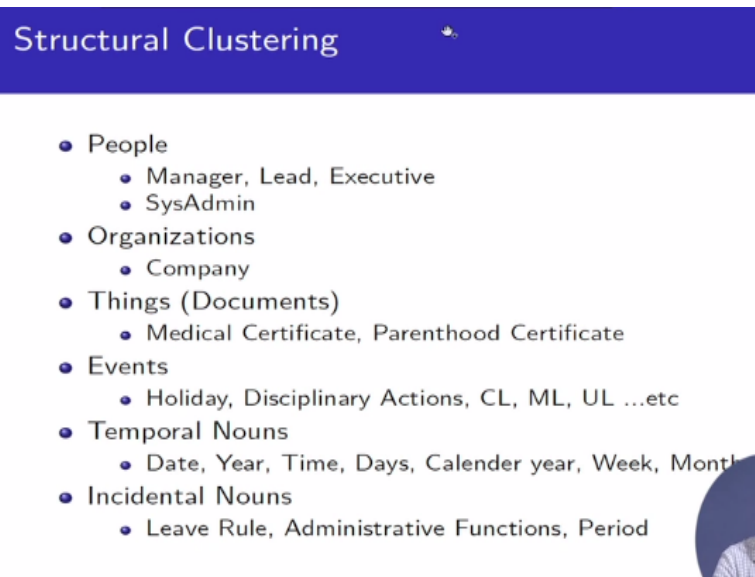
And we find that plenty of them are here. For example, leave is an event, is an event where an employee is going, will be absent from attendance in the company as it starts at a certain point, ends at a certain point and so on. Similarly, all of these different types of leave as discussed, the all these are kind of events. So is a holiday, holidays are not event because that is a designated day on which the company does not work as a whole.

So several these kinds of events including disciplinary action is an event because certain things specifically is done. So we can say that events tell us that these are the possible clusters that may be present here. Then certain things emerge from what you get to read if you go through, go back to the annotated document that you have created by marking the nouns, you will find that the word, the nouns like week, year, particularly days, month.

They occur not only the exist, but they occur very frequently. So there is a dominant presence for them and we can easily identify that they are structurally similar because all of them talk about time. So we can say that these are the temporal. So here I create a group which is kind of a temporal nouns.

Then there are lot of incidental nouns like leave rules, administrative functions, period of time so. You really have to go back to the document again and proceed further, work on the refinements to understand whether they care concepts nouns for concepts or they are just incidentally came in providing the description.

(Refer Slide Time: 17:20)



So given this, these are the concepts that abstractions that we have been able to identify through structural clustering and let us quickly do a quality check. Obviously, we have improved somewhat in terms of cohesion and coupling because, cohesion because we have been able to, we knew nothing almost about all these concepts that came up as nouns and now we have been able to group them into some seven eight broad classes.

So those classes I can say are coherent and in terms of, so if they are coherent they are similar by property. So there are certain amount of coupling that is coming up but certainly we have not made any progress in terms of the other metric parameters and we will go ahead and do the next which is the behavioural clustering.

(Refer Slide Time: 18:09)

Behavioral Clustering

- What takes place in the system
 - Leaves – CL, ML, UL ...etc
- Initiator and Participant
 - Initiator – Lead and Manager; SysAdmin
 - Participant – Executive, Lead and Manager

In terms of the behavioural clustering, the main approach is to look for dominant behaviours and one way to look for dominant behaviour is to look for what happens in the system, what takes place in the system. So we will see in subsequent modules that they will be more of discovery of behaviour through identification of verbs, adverbs and so on but having read this specification by now for three four times I presume, because if you have sincerely done the noun extraction and have trying to do the structural clustering.

Then you would have got a sense of certain thing taking place in the system which is leave, various kinds of leaves. This is a leave management system, they are certainly that is a leave behaviour is kind of a dominant concept that will exist and that is coming out and proliferating through a variety of different leave types.

The other that you look for typically in a behavioural clustering is to find initiators and participants that is in terms of a behaviour, we say this is an initiative which in terms of the final model will turn out to be clients and participants are kind of serves or just passive participants.

Do we find that those kind of behaviour and if you look and go through the specification

again through these expectations through this classes, we will find very well defined initiators like Lead, Manager, SysAdmin who are frequently talking about doing something in the system. Initiator is would does something in the system, who is responsible for starting some activity action in the system.

And you also find lot of participants like executive, lead and manager again because if the lead is approving or regretting the leave application of an executive, then the executive is a participant in the whole thing. Of course, I mean, these are not exhaustive, I would suggest strongly that you go through the document and try to identify more of the actions that are taking place in the system and more of the initiators participants in the systems and make the list more big.

(Refer Slide Time: 20:34)

Abstractions Summary from Linguistic Analysis

- **Key Abstractions:** Leave, Employee, CL, EL, ...
- **Non-Key (Supporting) Abstractions:** Medical Certificate, Parenthood Certificate, Disciplinary Actions, ...
- **Incidental Abstractions:** Leave Status, Administrative Functions, ...
- **Attribute (Property) Abstractions:** name, personal details, employee code, ...

Now having done this, then we can summarize the abstractions that we have found from the simple linguistic analysis which is been based only on nouns, will do more of that later and we can say that we have the key abstractions now, we can identify the key abstractions as something like leave, employee, casual leave, earned leave and so on.

And they do satisfy the requirements of the key abstraction we had specified earlier that is they come directly from the vocabulary of the domain, they are dominant concepts in the fact that they are repeatedly talked of in the specification. They are involved in a major part of the activities that are proposed in the system and so on. So looking into this symptoms, I propose that we accept leave, employee these as key abstractions.

And there are few others which are supporting abstractions non-key kind of medical certificate, you know if there is a lady employee needs to go for maternity leave, she needs to produce a medical certificate. So that is an abstraction but that is kind of a non-key abstraction because we do not find many references of medical certificate. We do find another reference in terms of sick leave but we do not find too many of them.

So we probably would not like to put as a key abstractions and that is a judgement choice. I mean, may be some of you would like to make it key abstractions but probably I will not make it at this stage. Certain abstractions are also found in terms of structural and behavioural clustering is incidental abstractions kind of administrative function, leave status so on.

So we are kind of putting them on the side line because we really have to go through the document more, we really have to go through the video more to understand that if they are critical for the LMS system or they have just occurred in the specification because when you talk about the activities in certain context, the related subsidiary (()) (22:49) activities also get talked of.

For example, one that we will talk of in terms of the next module is the fact that does employees work because we know if their employees there, we will should work in work related work related abstractions are they in any way important here, we will talk about that. So these are kind of incidental abstractions which are of that similar type and of course we have identified a set of, this is not exhaustive, it is just indicating we have started identifying a set of property abstractions or attribute abstractions in the whole thing.

And mind you, this is this whole thing is kind of little bit structured than completely go, I mean, going wild by saying that you pick up the vocabulary of the problem domain and start making them, clustering them. We have taken at a little bit more organized approached in terms of doing a linguistic analysis which also is a very powerful tool.

(Refer Slide Time: 23:51)

Quality Check

- The Metrics at this Stage
 - ① **Coupling:** Low (was Low)
 - ② **Cohesion:** **Moderate** (was Low)
 - ③ **Sufficiency:** Unclear (was Unclear)
 - ④ **Completeness:** **Very low** (was Unclear)
 - ⑤ **Primitiveness:** **Very low** (was Unclear)
- Actions required
 - Decide on Classes (first level)

And just with a noun analysis, this is what we have been able to achieve and if you re-evaluate ourselves as to where we stand right now we find that coupling is still low but cohesion of course has been improving because we not only have been able to group in terms of structural and conceptual similarity but we have also applied more considerations to identify what are the key abstractions, what are the supporting abstractions, what are incidental and so on and so forth.

The sufficiency is absolutely unclear as yet. Completeness of course, I have no sense of completeness, but sounds very low sense of completeness is coming to forth, because now you have at least some confidence to say that well of the different concepts that you have identified through these nouns may not be 100 percent complete but they must follow from about 90 percent of the basic concept that exist.

You do not still know whether those concepts should be used in that form itself, whether the abstractions need refinement, how they will be organised on hierarchies and so on but you have some sense of the fact that it is not these 49 nouns have, may be, may be you have some more may be around 50 hard nouns is kind of the domain within which we will have to define our system.

And as we have done that then we have started seeing that certain things are keys, certain a less key, so you have also some early sense of primitiveness coming in as to what are the concepts that are completely on their own and what are the concepts which are defined in terms of the others. So given this, we can take the last step given this.

So you can see that, I have shown at three stages of acting, starting with the extractions of nouns as we go and apply some more considerations, revisit and refine, the quality, the whole purpose of this refinement is to keep on improving the quality metrics and which we are being able to do and at the objective of this particular module, we want to end with deciding on the classes of this key abstractions and the supporting ones.

(Refer Slide Time: 26:05)



Let us at least initially agree that all the key abstractions initially would be formed as classes, so if you put the key abstractions, so this I say is the resultant identified classes, so we have the all, we have the manager, lead, executive, employee, these kind of ones, we have leave and all different types of leave, we also have SysAdmin and we have the central company, the organisation on which.

So this is what from the particularly from behavioural clustering and from identification of key abstractions where we are tried to make an assessment on the density and importance of activities that different abstractions do, based on these we proposed that this could be at least our initial choice of classes that we make for the LMS system.

(Refer Slide Time: 27:01)

Module Summary

- We understand how can we identify classes and its attributes from an informal English description
- In the next module we will discuss the relation and hierarchy within the classes

So this brings us to the end of this current module but as I had specified at the beginning, this will be an ongoing process of doing this exercise of finding more classes, their relationships, their hierarchy and the objects and different operations and we will continue in the next module in terms of that.