PANIMALAR INSTITUTE OF TECHNOLOGY

DEPARTMENT OF CSE

IV YEAR - VII SEMESTER

CS8079 – HUMAN COMPUTER INTERACTION (R 2017)

UNIT III MODELS AND THEORIES

HCI Models: Cognitive models: Socio-Organizational issues and stakeholder requirements –Communication and collaboration models-Hypertext, Multimedia and WWW.

3.1 COGNITIVE MODELS

- One way to classify the models is in respect to how well they describe features of the competence and performance of the user. Competence models tend to be ones that can predict legal behavior sequences but generally do this without reference to whether they could actually be executed by users.
- Performance models not only describe what the necessary behavior sequences are but usually describe both what the user needs to know and how this is employed in actual task execution. The presentation of the cognitive models in this chapter follows this classification scheme, divided into the following categories:

a) hierarchical representation of the user's task and goal structure
b) linguistic and grammatical models
c) physical and device-level models

- The first category deals directly with the issue of formulation of goals and tasks. The second deals with the grammar of the articulation translation and how it is understood by the user. The third category again deals with articulation, but at the human motor level instead of at a higher level of human understanding.

3.1.1 Goal and Task Hierarchies

- Many models make use of a model of mental processing in which the user achieves goals by solving sub goals in a divide-and-conquer fashion. We will consider two models, GOMS and CCT, where this is a central feature. Imaginewe want to produce a report on sales of introductory HCI textbooks.
- To achieve this goal we divide it into several sub goals, say gathering the data together, producing the tables and histograms, and writing the descriptive material.
- Concentrating on the data gathering, we decide to split this into further sub goals: find the names of all introductory HCI textbooks and then search the book sales database for these books. Similarly, each of the other sub goals is divided up into further sub goals, until some level of detail is found at which we decide to stop. We thus end up with a hierarchy of goals and sub goals. The example can be laid out to expose this structure:

```
produce reportgather data
. find book names
. . do keywords search of names database
        <<further subgoals>>
```

. . sift through names and abstracts by hand

<<further subgoals>>

. search sales database

<<further subgoals>> layout tables and histograms

<<further subgoals>>write description

<<further subgoals>>

- Different design issues demand different levels of analysis. This most abstract task is referred to as the unit task. The unit task does not require any problem- solving skills on the part of the user, though it frequently demands quite sophisticated problem-solving skills on the part of the designer to determine them.

1. GOMS

- The GOMS model of Card, Moran and Newell is an acronym for Goals, Operators, Methods and Selection [56]. A GOMS description consists of these four elements:

a. **Goals**: These are the user's goals, describing what the user wants to achieve. Further, in GOMS the goals are taken to represent a 'memory point' for the user, from which he can evaluate what should be done and to which he may return should any errors occur.

b. **Operators**: These are the lowest level of analysis. They are the basic actions that the user must perform in order to use the system. They may affect the system (for example, press the 'X' key) or only the user's mental state (for example, read the dialog box).

c. **Methods:** There are typically several ways in which a goal can be split into sub goals. For instance, in a certain window manager a currently selected window can be closed to an icon either by selecting the 'CLOSE' option from a pop-up menu, or by hitting the 'L7' function key. In GOMS these two goal decompositions are referred to as methods, so we have the **CLOSE-METHOD** and the **L7-METHOD**:

GOAL: ICONIZE-WINDOW
. [select GOAL: USE-CLOSE-METHOD
. . MOVE-MOUSE-TO-WINDOW-HEADER
. . POP-UP-MENU
. . CLICK-OVER-CLOSE-OPTIONGOAL: USE-L7-METHOD
. . PRESS-L7-KEY]

The dots are used to indicate the hierarchical level of goals.

d. **Selection** From the above snippet we see the use of the word select where the choice of methods arises. GOMS does not leave this as a random choice, but attempts to predict which methods will be used. This typically depends both on the particular user and on the state of the system and details about the goals.

- For instance, a user, Sam, never uses the L7-METHOD, except for one game, 'blocks', where the mouse needs to be used in the game until the very moment the key is pressed. GOMS captures this in a selection rule for Sam:

  User Sam:

  - ✓ Rule 1: Use the CLOSE-METHOD unless another rule applies.
  - ✓ Rule 2: If the application is 'blocks' use the L7-METHOD.

The goal hierarchies described in a GOMS analysis are almost wholly below the level of the unit task defined earlier. A typical GOMS analysis would therefore consist of a single high- level goal, which is then decomposed into a sequence of unit tasks, all of which can be further decomposed down to the level of basic operators. The goal of decomposition between the overall task and the unit tasks would involve detailed understanding of the user's problem-solving strategies and of the application domain.

2. **Cognitive complexity theory**

- Cognitive complexity refer to the number of mental structures an individual uses, how abstract they are and how they interact to shape his discernment or an individual difference variable linked with a wide range of communication skills and associated abilities.
- Individuals with high cognitive complexity have the capacity to analyze a situation to discern various constituent elements and explore connections and possible relationships among the elements. These individuals think in a multidimensional way. The assumption of the complexity theory is that the more an event can be differentiated and parts considered in novel relationships, the more sophisticated the response and successful the solution. Whereas less complex individuals can be trained to understand a complicated set of detailed differentiations for a specific context, highly complex individuals are highly flexible in creating distinctions in new situations.
- Individuals with high cognitive complexity are open to new information, attracted to other individuals of high complexity, highly flexibility, socially influential, problem solvers, strategic planners, highly creative, effective communicators and generally good leaders.
- CCT has two parallel descriptions: one of the user's goals and the other of the computer system (called the device in CCT). The description of the user's goals is based on a GOMS-like goal hierarchy, but is expressed primarily using production rules. CCT uses generalized transition networks, a form of state transition network. The production rules are a sequence of rules:

  if condition then action

  where condition is a statement about the contents of working memory. If the condition is true then the production rule is said to fire.
- An action may consist of one or more elementary actions, which may be either changes to the working memory, or external actions such as keystrokes. The production rule 'program' is written in a LISP-like language.
- Rules in CCT need not represent error-free performance. They can be used to explain error phenomena, though they cannot predict them. For instance, the rules above for inserting a space are 'buggy' – they do not check the editor's mode.
- The CCT rules are closely related to GOMS-like goal hierarchies; the rules may be generated from such a hierarchy, or alternatively, we may analyze the production rules to obtain the goal tree:

```
  GOAL: insert space
. GOAL: move cursor – if not at right position
. PRESS-KEY-I
. PRESS-SPACE
. PRESS-ESCAPE
```

3.1.2 Linguistic Models

The user's interaction with a computer is often viewed in terms of a language, so it is not surprising that several modeling formalisms have developed centered around this concept. BNF grammars are frequently used to specify dialogs.

The models here, although similar in form to dialog design notations, have been proposed with the intention of understanding the user's behavior and analyzing the cognitive difficulty of the interface.

a. BNF

- Representative of the linguistic approach is Reisner's use of Backus–Naur Form (BNF) rules to describe the dialog grammar. This views the dialog at a purely syntactic level, ignoring the semantics of the language. BNF has been used widely to specify the syntax of computer programming languages, and many system dialogs can be described easily using BNF rules.
- For example, imagine a graphics system that has a line-drawing function. To select the function the user must select the 'line' menu option. The line-drawing function allows the user to draw a polyline, that is a sequence of linearcs between points. The user selects the points by clicking the mouse button in the drawing area. The user double clicks to indicate the last point of the polyline.

    draw-line::= select-line + choose-points+ last-point
    select-line ::= position-mouse + CLICK-MOUSE
    choose-points ::= choose-one| choose-one + choose-points
    choose-one ::= position-mouse + CLICK-MOUSE
    last-point ::= position-mouse + DOUBLE-CLICK-MOUSE
    position-mouse ::= empty | MOVE-MOUSE + position-mouse

- The names in the description are of two types: non-terminals, shown in lower case, and terminals, shown in upper case. Terminals represent the lowest level of user behavior, such as pressing a key, clicking a mouse button or movingthe mouse.
- Non-terminals are higher-level abstractions. The non-terminals are defined in terms of other non-terminals and terminals by a definition of the form name ::= expression
- The '::=' symbol is read as 'is defined as'. Only non-terminals may appear onthe left of a definition. The right-hand side is built up using two operators '+' (sequence) and '|' (choice).
- For example, the first rule says that the non-terminal draw-line is defined to be select-line followed by choose-points followed by last point.All of these are non-terminals, that is they do not tell us what the basic user actions are.
- The second rule says that select-line is defined to be position mouse (intendedto be over the 'line' menu entry) followed by CLICK-MOUSE. This is our first terminal and represents the actual clicking of a mouse. To see what position- mouse is, we look at the last rule.
- This tells us that there are two possibilities for position-mouse (separated by the '|' symbol). One option is that position-mouse is empty – a special symbol representing no action. That is, one option is not to move the mouse at all.
- The other option is to do a MOVE-MOUSE action followed by position-mouse. This rule is recursive, and this second position-mouse may itself either beempty or be a MOVE-MOUSE action followed by position-mouse, and so on. That is, position-mouse may be any number of MOVE-MOUSE actions whatsoever.
- Choose-points is defined recursively, but this time it does not have the option of being empty. It may be one or more of the non-terminal choose one which is itself defined to be (like select-line) position-mouse followed by CLICK-MOUSE.
- The BNF description of an interface can be analyzed in various ways. One measure is to count the number of rules. The more rules an interface requiresto use it, the more complicated it is. This measure is rather sensitive to the exact way the interface is described. For example, we could have replaced therules for choose points and choose-one with the single definition,
  choose-points ::= position-mouse + CLICK-MOUSE| position-mouse + CLICK-MOUSE + choose-points

More robust measure also counts the number of ‘+’ and ‘|’ operators. This would, in effect, penalize the more complex single rule. Another problem arises with the rule for select-line. This is identical to the choose-one rule.

b. Task–action grammar

- Measures based upon BNF have been criticized as not ‘cognitive’ enough. They ignore the advantages of consistency both in the language‘s structure and in its use of command names and letters. Task–action grammar (TAG) attempts to deal with some of these problems by including elements such as parameterized grammar rules to emphasize consistency and encoding the user’s world knowledge (for example, up is theopposite of down).
- To illustrate consistency, we consider the three UNIX commands: cp (for copying files), mv (for moving files) and ln (for linking files). Each of these has two possible forms. They either have two arguments, a source and destination filename, or have any number of source filenames followed by a destination directory:

    copy ::= ‘cp’ + filename + filename
    | ‘cp’ + filenames + directory
    move ::= ‘mv’ + filename + filename
    | ‘mv’ + filenames + directory link ::= ‘ln’ + filename + filename
    | ‘ln’ + filenames + directory

- Measures based upon BNF could not distinguish between these consistent commands and an inconsistent alternative – say if ln took its directory argument first. Task–action grammar was designed to reveal just this sort of consistency. Its description of the UNIX commands would be

    file-op[Op] := command[Op] + filename + filename
    | command[Op] + filenames + directory command[Op=copy] := ‘cp’ command[Op=move] := ‘mv’ command[Op=link] := ‘ln’

- Hierarchical and grammar-based techniques were initially developed when most interactive systems were command line, or at most, keyboard and cursor based. There are significant worries, therefore, about how well these approaches can generalize to deal with more modern windowed and mouse-driven interfaces. Pressing a cursor key is a reasonable lexeme, but moving a mouse one pixel is less sensible. In addition, pointer-based dialogs are more display oriented. Clicking a cursor at a particular point on the screen has a meaning dependent on the current screen contents. This problem can be partially resolved by regarding operations such as ‘select region of text’ or ‘click on quit button’ as the terminals of the grammar. If this approach is taken, the detailed mouse movements and parsing of mouse events in the context of display information (menus, etc.) are abstracted away. Goal hierarchy methods have different problems, as more display-oriented systems encourage less structured methods for goal achievement. Instead of having well-defined plans, the user is seen as performing a more exploratory task, recognizing fruitful directions and backing out of others. Typically, even when this exploratory style is used at one level,

    WRITE_LETTER
    .FIND_SIMILAR_LETTER
    .COPY_IT
    .EDIT_COPY

3.1.3 PHYSICAL AND DEVICE MODELS

a. Keystroke-level model

- The human motor system is well understood. KLM (Keystroke-Level Model) uses this understanding as a basis for detailed predictions about user performance. It is aimed at unit tasks within interaction – the execution of simple command sequences, typically taking no more than 20 seconds. Examples of this would be using a search and replace feature, or changing the font of a word. It does not extend to complex actions such as producing a diagram. The assumption is that these more complex tasks would be split into subtasks (as in GOMS) before the user attempts to map them into physical actions.
- It does not extend to complex actions such as producing a diagram. The assumption is that these more complex tasks would be split into subtasks (asin GOMS) before the user attempts to map them into physical actions. Thetask is split into two phases:
  - acquisition of the task, when the user builds a mental representation of the task;
  - execution of the task using the system's facilities.
- KLM only gives predictions for the latter stage of activity. During the acquisition phase, the user will have decided how to accomplish the task using the primitives of the system, and thus, during the execution phase, there is no high-level mental activity the user is effectively expert.
- KLM is related to the GOMS model, and can be thought of as a very low level GOMS model where the method is given. The model decomposes the execution phase into five different physical motor operators, a mental operator and a system response operator:
  a. K Keystroking, actually striking keys, including shifts and other modifier keys.
  b. B Pressing a mouse button.
  c. P Pointing, moving the mouse (or similar device) at a target.
  d. H Homing, switching the hand between mouse and keyboard.
  e. D Drawing lines using the mouse.
  f. M Mentally preparing for a physical action.
  g. R System response which may be ignored if the user does not have towait for it, as in copy typing.
- The execution of a task will involve interleaved occurrences of the various operators. For instance, imagine we are using a mouse-based editor. If we notice a single character error we will point at the error, delete the character and retype it, and then return to our previous typing point. This is decomposed as follows:
  1. Move hand to mouse H[mouse]
  2. Position mouse after bad character PB[LEFT]
  3. Return to keyboard H[keyboard]
  4. Delete character MK[DELETE]
  5. Type correction K[char]
  6. Reposition insertion point H[mouse]MPB[LEFT]
- 5 Notice that some operators have descriptions added to them, representing which device the hand homes to (for example, [mouse]) and what keys are hit (for example, LEFT – the left mouse button). The model predicts the total time taken during the execution phase by adding the component times for each of the above activities.

Operator	Remarks	Time (s)
K	Press key	
	good typist (90 wpm)	0.12
	poor typist (40 wpm)	0.28
	non-typist	1.20
B	Mouse button press	
	down or up	0.10
	click	0.20
P	Point with mouse	
	Fitts' law	$0.1 \log_2(D/S + 0.5)$
	average movement	1.10
H	Home hands to and from keyboard	0.40
D	Drawing – domain dependent	–
M	Mentally prepare	1.35
R	Response from system – measure	–

wpm = words per minute

Table 3.1: Times for various operators in the keystroke level model

b. Three-state model

- We saw that a range of pointing devices exists in addition to the mouse. Often these devices are considered logically equivalent, if the same inputs are available to the application. That is, so long as you can select a point on the screen, they are all the same. These different devices – mouse, trackball, light pen – feel very different.
- Buxton has developed a simple model of input devices the three-state model, which captures some of these crucial distinctions. He begins by looking at a mouse. If you move it with no buttons pushed, it normally moves the mouse cursor about. This tracking behavior is termed state 1. Depressing a button over an icon and then moving the mouse will often result in an object being dragged about. This he calls state 2.

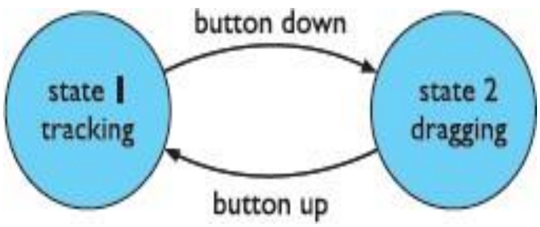


Fig. 3.1: Mouse Transitions: States 1 and 2

- Instead we consider a light pen with a button, it behaves just like a mouse when it is touching the screen. When its button is not depressed, it is in state 1, and when its button is down, state 2. However, the light pen has a third state, when the light pen is not touching the screen. In this state the system cannot track the light pen's position. This is called state 0.

  A touchscreen is like the light pen with no button. While the user is not touching the screen, the system cannot track the finger that is, state 0 again. When the user touches the screen, the system can begin to track state 1. So a touchscreen is a state 0–1 device whereas a mouse is a state 1–2 device. As there is no difference between a state 0–2 and a state 0–1 device, there are only the three possibilities we have seen.

- The only additional complexity is if the device has several buttons, in which case we would have one state for each button: 2left, 2middle, 2right.
- At first, the model appears to characterize the states of the device by the inputs available to the system. So, from this perspective, state 0 is clearly different from states 1 and 2. However, if we look at the state 1–2 transaction, we see that it is symmetric with respect to the two states.

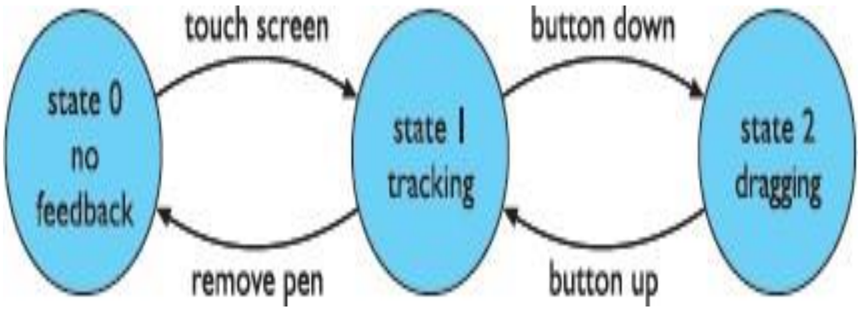


Fig.3.2: Light pen transitions: three states

- In principle, there is no reason why a program should not decide to do simple mouse tracking whilst in state 2 and drag things about in state 1 State 2 requires a button to be pressed, whereas state 1 is one of relative relaxation (whilst still requiring hand–eye coordination for mouse movement). There is a similar difference in tension between state 0 and state 1.
- It is well known that Fitts' law has different timing constants for different devices. Recall that Fitts' law says that the time taken to move to a target of size S at a distance D is:

$$a+b\log_2((D/S)+1)$$

- The constants a and b depend on the particular pointing device used and the skill of the user with that device.

	Device	*a* (ms)	*b* (ms/bit)
Pointing (state 1)			
	Mouse	–107	223
	Trackball	75	300
Dragging (state 2)			
	Mouse	135	249
	Trackball	–349	688

Table 3.2: Fitt's law coefficients

3.1.4 Cognitive Architectures

- The concept of taking a problem and solving it by divide and conquer using subgoals is central to GOMS. CCT assumes the distinction between long- and short-term memory, with production rules being stored in long-term memory and 'matched' against the contents of short-term (or working) memory to determine which 'fire'. The values for various motor and mental operators in KLM were based on the Model Human Processor (MHP) architecture of Card, Moran and Newell. Another common assumption, which we have not discussed in this chapter, is the distinction between linguistic levels – semantic, syntactic and lexical – as an architectural model of the user's understanding.

- The problem space model rational behavior is characterized as behavior that is intended to achieve a specific goal. This element of rationality is often used to distinguish between intelligent and machine-like behavior. In the field of artificial intelligence (AI), a system exhibiting rational behavior is referred to as a knowledge-level system. A knowledge- level system contains an agent behaving in an environment. The agent has knowledge about itself and its environment, including its own goals. It can perform certain actions and sense information about its changing environment. As the agent behaves in its environment, it changes the environment and its own knowledge. We can view the overall behavior of the knowledge-level system as a sequence of environment and agent states as they progress in time. The goal of the agent is characterized as a preference over all possible sequences of agent/environment states. The search proceeds by moving from one state to another possible state by means of operations or actions, the ultimate goal of which is to arrive at one of the desired states. This very general model of computation is used in the ordinary task of the programmer. Once she has identified a problem and a means of arriving at the solution to the problem (the algorithm), the programmer then represents the problem and algorithm in a programming language, which can be executed on a machine to reach the desired state. The architecture of the machine only allows the definition of the search or problem space and the actions that can occur to traverse that space. Termination is also assumed to happen once the desired state is reached.

- The new computational model is the problem space model, based on the problem-solving work of Newell and Simon at Carnegie–Mellon University. A problem space consists of a set of states and a set of operations that can be performed on the states. Behavior in a problem space is a two-step process. First, the current operator is chosen based on the current state and then it is applied to the current state to achieve the new state. The problem space must represent rational behavior, and so it must characterize the goal of the agent. A problem space represents a goal by defining the desired states as a subset of all possible states. Once the initial state is set, the task within the problem space is to find a sequence of operations that form a path within the state space from the initial state to one of the desired states, whereupon successful termination occurs.

- We can highlight four different activities that occur within a problem space: goal formulation, operation selection, operation application and goal completion. The relationship between these problem space processes and knowledge-level activity is key. Perception that occurs at the knowledge level is performed by the goal formulation process, which creates the initial state based on observations of the external environment. Actions at the knowledge level are operations in the problem space which are selected and applied. The real knowledge about the agent and its environment and goals is derived from the state/operator information in the problem space. Because of the goal formulation process, the set of desired states indicates the knowledge-level goal within the problem space. The operation selection process selects the appropriate operation at a given point in time because it is deemed the most likely to transform the state in the problem space to one of the desired states; hence rational behavior is implied.

- Interacting cognitive subsystems (ICS) provides a model of perception, cognition and action, but unlike other cognitive architectures, it is not intended to produce a description of the user in terms of sequences of actions that he performs. ICS provides a more holistic view of the user as an information-processing machine. The emphasis is on determining how easy particular procedures of action sequences become as they are made more automatic within the user.

- ICS attempts to incorporate two separate psychological traditions within one cognitive architecture. On the one hand is the architectural and general-purpose information-processing approach of short-term memory research. On the other hand is the computational and representational approach characteristic of psycholinguistic research and AI problem-solving literature.

- The architecture of ICS is built up by the coordinated activity of nine smaller subsystems: five peripheral subsystems are in contact with the physical world and four are central, dealing with mental processes. Each subsystem has the same generic structure. A subsystem is described in terms of its typed inputs and outputs along with a memory store for holding typed information. It has transformation functions for processing the input and producing the output and permanently stored information. Each of the nine subsystems is specialized for handling some aspect of external or internal processing. For example, one peripheral subsystem is the visual system for describing what is seen in the world.

3.2 SOCIO-ORGANIZATIONAL ISSUES AND STAKE HOLDER REQUIREMENTS

There are several organizational issues that affect the acceptance of technology by users and that must therefore be considered in system design:

- systems may not take into account conflict and power relationships
- those who benefit may not do the work
- not everyone may use systems.

In addition to generic issues, designers must identify specific stakeholder requirements within their organizational context.

• Socio-technical models capture both human and technical requirements.
• Soft systems methodology takes a broader view of human and organizational issues.
• Participatory design includes the user directly in the design process.
• Ethnographic methods study users in context, attempting to take an unbiased perspective.

3.2.1 Organizational Issues

- We shall look at some of the organizational issues that affect the acceptance and relevance of information and communication systems. These factors often sit ‘outside’ the system as such, and may involve individuals who never use it.

➢ **Cooperation or conflict?**

The term ‘computer-supported cooperative work’ (CSCW) seems to assume that groups will be acting in a cooperative manner. This is obviously true to some extent; even opposing football teams cooperate to the extent that they keep (largely) within the rules of the game, but their cooperation only goes so far. People in organizations and groups have conflicting goals, and systems that ignore this are likely to fail spectacularly.

Imagine that an organization is already highly computerized, the different departments all have their own systems and the board decides that an integrated information system is needed. The production manager can now look directly at stocks when planning the week‘s work, and the marketing department can consult the sales department‘s contact list to send out marketing questionnaires.

The storekeeper always used to understate stock levels slightly in order to keep an emergency supply, or sometimes inflate the quoted levels when a delivery was due from a reliable supplier. Also, requests for stock information allowed the storekeeper to keep track of future demands and hence plan future orders. The storekeeper has now lost a sense of control and important sources of information. Members of the sales department are also unhappy: their contacts are their livelihood. The last thing they want is someone from marketing blundering in and spoiling a relationship with a customer built up over many years. Some of these people may resort to subverting the system, keeping ‘sanitized’ information online, but the real information in personal files.

- **Changing power structures**

The identification of stakeholders will uncover information transfer and power relationships that cut across the organizational structure. Indeed, all organizations have these informal networks that support both social and functional contacts. The official lines of authority and information tend to flow up and down through line management.

The physical layout of an organization often reflects the formal hierarchy: each department is on a different floor, with sections working in the same area of an office. If someone from sales wants to talk to someone from marketing then one of them must walk to the other‘s office. Their respective supervisors can monitor the contact.

In face-to-face conversation, the manager can easily exert influence over a subordinate: both know their relative positions and this is reflected in the patterns of conversation and in other non-verbal cues. Email messages lose much of this sense of presence and it is more difficult for a manager to exercise authority. The 'levelling' effect even makes it possible for subordinates to direct messages 'diagonally' across the hierarchy, to their manager‘s peers, or, even worse, to their manager‘s manager!

- **The invisible worker**

The ability to work and collaborate at a distance can allow functional groups to be distributed over different sites. This can take the form of cross-functional neighbourhood centers, where workers from different departments do their jobs in electronic contact with their functional colleagues. If the approach in an organization is 'management by presence', that is you know someone is working because they are in the office, then there is no way a remote worker is going to be trusted. If, on the other hand, the style is 'management by objectives', that is you know your subordinates are working because they are doing their jobs and producing results, then remote working is not so problematical.

- **Who benefits?**

In these systems the sender has to do work in putting information into fields appropriately, but it is the recipient who benefits. Another example is shared calendars. The beneficiary of the system is a manager who uses the system to arrange meeting times, but whose personal secretary does the work of keeping the calendar up to date. Subordinates are less likely to have secretarial support, yet must keep up the calendar with little perceived benefit. Of course, chaos results when a meeting is automatically arranged and the subordinates may have to rearrange commitments that have not been recorded on the system. The manager may force use by edict or the system may simply fall into disuse. Many such groupware systems are introduced on a 'see if it works' basis.

- **Free rider problem**

A system may still not function symmetrically, which may be a problem, particularly with shared communication systems. One issue is the free rider problem. Take an electronic conferencing system. If there is plenty of discussion of relevant topics then there are obvious advantages to subscribing and reading the contributions. When considering writing a contribution, the effort of doing so may outweigh any benefits.

The total benefit of the system for each user outweighs the costs, but for any particular decision the balance is overturned. A few free riders in a conference system are often not a problem, as the danger is more likely from too much activity. In addition, in electronic conferences the patterns of activity and silence may reflect other factors such as expertise. It is easy for the number of free riders gradually to increase and the system slide into disuse. It is hard to enforce equal use, except by restrictive schemes such as round-robin contributions (everyone contributes something however short). In the real world, such problems are often solved by social pressure, and the free rider reacts to the collective censure of the group. Increasing the visibility of participants' contributions might also help these social mechanisms.

- **Critical mass**

When telephones were only in public places, their use as a form of pervasive interpersonal communication was limited. However, once a large number of people have telephones in their homes it becomes worthwhile paying to have a telephone installed. In cost/benefit terms, the early subscribers probably have a smaller benefit than the cost. Only when the number of subscribers increases beyond the critical mass does the benefit for all dominate the cost.

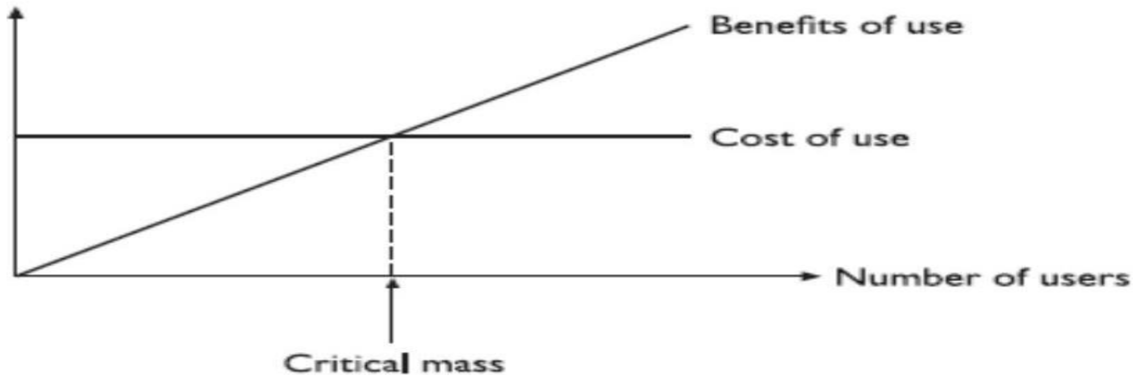


Fig. 3.3: Cost benefit of system use

The telephone was useful for subgroups before it became beneficial for all. Even when only a small proportion of the population had personal telephones, they still formed a significant proportion of their social group, so these cliques of use could grow gradually over time.

- **Automating processes – workflow and BPR**

Organizations have many such processes, and workflow systems aim to automate much of the process using electronic forms, which are forwarded to the relevant person based on pre-coded rules. Some workflow systems are built using special purpose groupware, often based on a notation for describing the desired workflow.

The rigid form of a typical workflow system is an example of global structuring. The danger with any form of global structuring is that it may conflict with or inhibit more informal and less structured patterns of activity which also contribute to the organization‘s free running.

A more radical approach to organizational processes is found in business process re-engineering (BPR). Traditionally, organizations have been structured around functions: sales, accounts, stores, manufacturing. However, the purpose of an organization can be seen in terms of key business processes. The ordering/delivery process described above is a typical and important example. In BPR these processes are recorded and analyzed. Problems in the current process are noted and the whole process may be redesigned in order to make the path of the process more efficient. For example, instead of sending an order to the accounts department to approve, a list of customer credit limits could be given to the sales executives. They could then check the credit rating of the customer whilst on the phone and only forward the order to accounts if there are any unusual problems.

- **Evaluating the benefits**

The benefits from cooperative systems, especially organization-wide systems such as email or electronic conferencing, are in terms of job satisfaction or more fluid information flow. Some, such as the video call, are expected primarily to help social contact within the organization. It may be possible to measure contentment and job satisfaction using attitude questionnaires.

3.2.3 Capturing Requirements

- We begin by capturing and analyzing requirements, but we need to do this within the work context, taking account of the complex mix of concerns felt by different stakeholders and the structures and processes operating in the workgroups.
- need to take account of
  - stakeholders
  - work groups and practices
  - organisational context

- We consider several approaches:
  - socio-technical modeling
  - soft systems methodology
  - participatory design
  - ethnographic methods and contextual inquiry.

- Who are the stakeholders? can be defined as anyone who is affected by the success or failure of the system. It can be useful to distinguish different categories of stakeholder, and the following categorization from the CUSTOM approach is helpful for this:

1. Primary stakeholders are people who actually use the system – the end-users.
2. Secondary stakeholders are people who do not directly use the system, but receive output from it or provide input to it (for example, someone who receives a report produced by the system).
3. Tertiary stakeholders are people who do not fall into either of the first two categories but who are directly affected by the success or failure of the system (for example, a director whose profits increase or decrease depending on the success of the system).
4. Facilitating stakeholders are people who are involved with the design, development and maintenance of the system.

Example: Classifying stakeholders – an airline booking system

An international airline is considering introducing a new booking system for use by associated travel agents to sell flights directly to the public.

Primary stakeholders: travel agency staff, airline booking staff

Secondary stakeholders: customers, airline management

Tertiary stakeholders: competitors, civil aviation authorities, customers'travelling companions, airline shareholders

Facilitating stakeholders: design team, IT department staff

- All of the approaches we are considering here are concerned with understanding stakeholders within their organizational context.

- **Socio-technical models**
  - Technological determinism, the view that social change is primarily dictated by technology, with human and social factors being secondary concerns, was prevalent. The socio-technical systems view came about
  - to counter this technology-centric position, by stressing that work systems were composed of both human and machine elements and that it was the interrelationship between these that should be central.

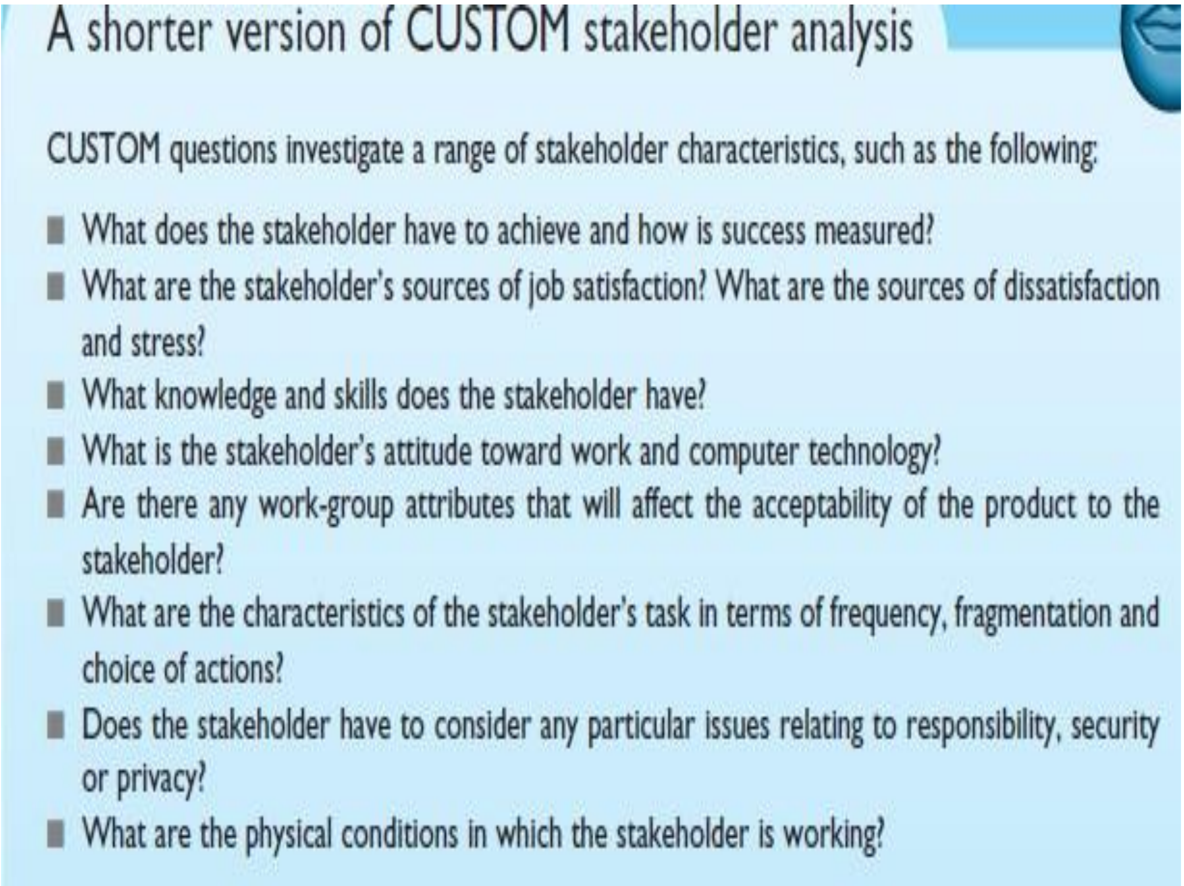
- Socio-technical models for interactive systems are therefore concerned with technical, social, organizational and human aspects of design. They recognize the fact that technology is not developed in isolation but as part of a wider organizational environment. It is important to consider social and technical issues side by side so that human issues are not overruled by technical considerations.
- The key focus of the socio-technical approach is to describe and document the impact of the introduction of a specific technology into an organization. Methods vary but most attempt to capture certain common elements:
  - The problem being addressed: there is a need to understand why the technology is being proposed and what problem it is intended to solve.
  - The stakeholders affected, including primary, secondary, tertiary and facilitating, together with their objectives, goals and tasks.
  - The workgroups within the organization, both formal and informal.
  - The changes or transformations that will be supported.
  - The proposed technology and how it will work within the organization.
  - External constraints and influences and performance measures.
- Information is gathered using methods such as interviews, observation, focus groups and document analysis. The methods guide this information-gathering process and help the analyst to make sense of what is discovered.
- By attempting to understand these issues, socio-technical approaches aim to provide a detailed view of the role technology will play and the requirements of successful deployment. We will compare two approaches to illustrate how this may work in practice.

- **CUSTOM methodology**
  - CUSTOM is a socio-technical methodology designed to be practical to use in small organizations. It is based on the User Skills and Task Match (USTM) approach, developed to allow design teams to understand and fully document user requirements.
  - CUSTOM focuses on establishing stakeholder requirements: all stakeholders are considered, not just the end-users. It is applied at the initial stage of design when a product opportunity has been identified, so the emphasis is on capturing requirements.
  - It is a forms-based methodology, providing a set of questions to apply at each of its stages. There are six key stages to carry out in a CUSTOM analysis:

  a. Describe the organizational context, including its primary goals, physical characteristics, political and economic background.
  b. Identify and describe stakeholders. All stakeholders are named, categorized (as primary, secondary, tertiary or facilitating) and described with regard to personal issues, their role in the organization and their job.
  c. Identify and describe work-groups. A work-group is any group of people who work together on a task, whether formally constituted or not. Again, work-groups are described in terms of their role within the organization and their characteristics.
  d. Identify and describe task–object pairs. These are the tasks that must be performed, coupled with the objects that are used to perform them or to which they are applied.
  e. Identify stakeholder needs. Stages 2–4 are described in terms of both the current system and the proposed system. Stakeholder needs are identified by considering the differences between the two. For example, if a stakeholder is identified as currently lacking a particular skill that is required in the proposed system then a need for training is identified.

f. Consolidate and check stakeholder requirements. Here the stakeholder needs list is checked against the criteria determined at earlier stages.

- Stages 2 to 4 are described in terms of the current situation (before the new technology is introduced) and the proposed situation (after deployment). Stakeholders are asked to express their views not only of their current role and position but of their expectations in the light of the changes that will be made.
- The stakeholder concerns and goals are elaborated. In addition, the impact of the technology on working practices is considered (Stage 3) and the transformations that will be supported by the system specified (Stage 4). The changes from the current position to the proposed position represent the issues that need to be addressed to ensure successful deployment, and these are made explicit during Stages 5 and 6.
- CUSTOM provides a useful framework for considering stakeholder requirements and the use of forms and questions (a ‘manual’ for its use is available makes it relatively straightforward; if somewhat time consuming, to apply. For less complex situations, a shortened version of CUSTOM stakeholder analysis is available .This also provides a checklist for investigations for stages 2– 4.



- **Open System Task Analysis (OSTA)**
  - OSTA is an alternative socio-technical approach, which attempts to describe what happens when a technical system is introduced into an organizational work environment.

- Like CUSTOM, OSTA specifies both social and technical aspects of the system. However, whereas in CUSTOM these aspects are framed in terms of stakeholder perspectives, in OSTA they are captured through a focus on tasks. OSTA has eight main stages:

1. The primary task which the technology must support is identified in terms of users' goals.
2. Task inputs to the system are identified. These may have different sources and forms that may constrain the design.
3. The external environment into which the system will be introduced is described, including physical, economic and political aspects.
4. The transformation processes within the system are described in terms of actions performed on or with objects.
5. The social system is analyzed, considering existing work-groups and relationships within and external to the organization.
6. The technical system is described in terms of its configuration and integration with other systems.
7. Performance satisfaction criteria are established, indicating the social and technical requirements of the system.
8. The new technical system is specified.

OSTA uses notations familiar to designers, such as data flow diagrams and textual descriptions.

- **Soft systems methodology**
  - The socio-technical models we have looked at focus on identifying requirements from both human and technical perspectives, but they assume a technological solution is being proposed. Soft systems methodology (SSM) arises from the same tradition but takes a view of the organization as a system of which technology and people are components.
  - There is no assumption of a particular solution: the emphasis is rather on understanding the situation fully. SSM was developed by Checkland to help designers reach an understanding of the context of technological developments and the influences and concerns that exist within the system under consideration.
  - SSM has seven stages. A distinction is made between the 'real-world' stages (1–2, 5–7) and the systems stages (3–4). The first stage of SSM is the recognition of the problem and initiation of analysis. This is followed by a detailed description of the problem situation: developing a rich picture. This will include all the stakeholders, the tasks they carry out and the groups they work in, the organizational structure and its processes and the issues raised by each stakeholder.
  - Any knowledge elicitation techniques can be used to gather the information to build the rich picture, including observation (and video and audio recording), structured and unstructured interviews and questionnaires, and workshops incorporating such activities as role play, simulations and critical incident analysis.
  - Less structured approaches are used initially to avoid artificially constraining the description. The rich picture can be in any style – there are no right or wrong answers – but it should be clear and informative to the designer.

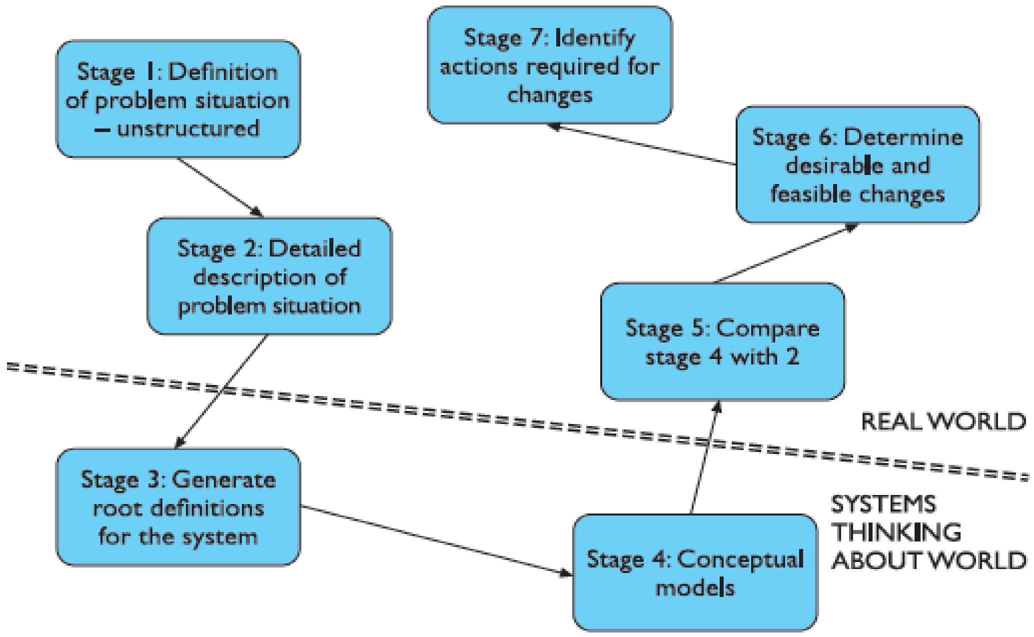


Fig3.4: Seven stages of soft systems methodology

- At the next stage in SSM we move from the real world to the systems world and attempt to generate root definitions for the system, which define the essence of what the system is about. There may be several root definitions of a system, representing different stakeholder perspectives, for example. Root definitions are described in terms of specific elements, summarized using the acronym, CATWOE:

1. Clients – those who receive output or benefit from the system.
2. Actors – those who perform activities within the system.
3. Transformations – the changes that are effected by the system. This is a critical part of the root definition as it leads to the activities that need to be included in the next stage. These 'transform' the inputs of the system into the required outputs.
4. Weltanschauung – (from the German) meaning world view. This is how the system is perceived in a particular root definition.
5. Owner – those to whom the system belongs, to whom it is answerable and who can authorize changes to it.
6. Environment – the world in which the system operates and by which it is influenced.

Root definition for airline management: an airline booking system
Client: customer **Actor:** travel agency staff **Transformation:** customer's intention and request to travel transformed into sale of seat on flight and profit for organization **Weltanschauung:** profits can be optimized by more efficient sales **Owner:** airline management **Environment:** Regulations of international civil aviation authorities and national contract legislation. Local agency policies worldwide

- **Participatory design**

In participatory design:
workers enter into design context
In ethnography (as used for design):
designer enters into work context

- Participatory design is a philosophy that encompasses the whole design cycle. It is design in the workplace, where the user is involved not only as an experimental subject or as someone to be consulted when necessary but as a member of the design team.

- Users are therefore active collaborators in the design process, rather than passive participants whose involvement is entirely governed by the designer.

- The argument is that users are experts in the work context and a design can only be effective within that context if these experts are allowed to contribute actively to the design process.

- It therefore aims to refine system requirements iteratively through a design process in which the user is actively involved.

- Participatory design has three specific characteristics. It aims to improve the work environment and task by the introduction of the design. This makes design and evaluation context or work oriented rather than system oriented. Secondly, it is characterized by collaboration: the user is included in the design team and can contribute to every stage of the design.

- Finally, the approach is iterative: the design is subject to evaluation and revision at each stage. The participatory design process utilizes a range of methods to help convey information between the user and designer. They include

a. **Brainstorming** This involves all participants in the design pooling ideas. This is informal and relatively unstructured although the process tends to involve 'onthefly' structuring of the ideas as they materialize. All information is recorded without judgment. The session provides a range of ideas from which to work. These can be filtered using other techniques.
b. **Storyboarding** Storyboards can be used as a means of describing the user's day-to-dayactivities as well as the potential designs and the impact they will have.
c. **Workshops** These can be used to fill in the missing knowledge of both user and designer and provide a more focussed view of the design. They may involve mutual enquiry in which both parties attempt to understand the context of the design from each other's point of view.
d. **Pencil and paper exercises** These allow designs to be talked through and evaluated with very little commitment in terms of resources. Users can 'walk through' typical tasks using paper mock-ups of the system design. This is intended to show up discrepancies between the user's requirements and the actual design as proposed.

Example:

DESIGN FOCUS

Tomorrow's hospital – using participatory design

The nurse walks around the ward to a patient's bedside. She takes her PDA from her pocket and enters the patient's name. Her PDA is connected via a wireless network to the central patient treatment database and the patient's information comes on screen including reminders of treatments needed. First is a blood pressure and heart rate check. She does these checks and enters the results. A mild painkiller is also prescribed. She gets the tablet, which is individually wrapped in a bar coded packet. Her PDA has a built-in bar code reader and she scans this across the tablet. The system registers the drug's use and removes it from the to-do list as she gives the patient the tablet with a glass of water.

A picture of tomorrow's hospital? In fact, this is exactly the day-to-day activity in Hospital da Trofa, just outside Porto, Portugal. There have been numerous attempts to use PDA-based systems in hospitals. But most have failed. So why is this one being used? The hospital is part of a European Commission funded project Team-HOS and the system was designed using a methodology that has a strong *participatory* focus [316]. From the beginning, nurses, doctors, pharmacologists and dieticians were involved in and in control of the design. It has benefited from their knowledge and experience, which is why it does the right things for the context. Furthermore, the pride with which the hospital staff describe the system shows that they really feel it is their system, not one imposed from above.

See the book website for an extended case study: /e3/casestudy/trofa/

Effective Technical and Human Implementation of Computer-based systems (ETHICS)

- ETHICS [243] is a method developed by Enid Mumford within the socio-technical tradition, but it is distinct in its view of the role of stakeholders in the process. In the ETHICS methodology, stakeholders are included as participants in the decision making process.

- ETHICS considers the process of system development as one of managing change: conflicts will occur and must be negotiated to ensure acceptance and satisfaction with the system. If any party is excluded from the decision-making process then their knowledge and contribution is not utilized and they are more likely to be dissatisfied. However, participation is not always complete. Mumford recognizes three levels of participation:
  - Consultative – the weakest form of participation where participants are asked for their opinions but are not decision makers.
  - Representative – a representative of the participant group is involved in the decision making process.
  - Consensus – all stakeholders are included in the decision-making process.

- The usual practice is that design groups are set up to include representatives from each stakeholder group and these groups make the design decisions, overseen by a steering committee of management and employee representatives.

The design groups then address the following issues and activities:

1. **Make the case for change**. Change for its own sake is inappropriate. If a case cannot be made for changing the current situation then the process ends and the system remains as it is.
2. **Identify system boundaries**. This focuses on the context of the current system and its interactions with other systems, in terms of business, existing technology, and internal and external organizational elements. How will the change impact upon each of these?
3. **Describe the existing system**, including a full analysis of inputs and outputs and the various other activities supported, such as operations, control and coordination.
4. **Define key objectives**, identifying the purpose and function of each area of the organization.
5. **Define key tasks**: what tasks need to be performed to meet these objectives?
6. **Define key information needs**, including those identified by analysis of the existing system and those highlighted by definition of key tasks.
7. **Diagnose efficiency needs**, those elements in the system that cause it to underperform or perform incorrectly. If these are internal they can be redesigned out of the new system; if they are external then the new system must be designed to cope with them.
8. **Diagnose job satisfaction needs**, with a view to increasing job satisfaction where it is low.
9. **Analyze likely future changes**, whether in technology, external constraints (such as legal requirements), economic climate or stakeholder attitudes. This is necessary to ensure that the system is flexible enough to cope with change.
10. **Specify and prioritize objectives based on efficiency**, job satisfaction and future needs. All stakeholders should be able to contribute here as it is a critical stage and conflicting priorities need to be negotiated. Objectives are grouped as either primary (must be met) or secondary (desirable to meet).

- The final stages of the ETHICS approach focus on the actual design and evaluation of the system. Necessary organizational changes are designed alongside the technical system. These are then specified in detail, implemented and evaluated.
- The ETHICS approach attempts to reach a solution that meets both user and task requirements by having specialist teams negotiate objectives and rank potential solutions.

➢ **Ethnographic methods**

- Real action is situated action; it occurs in interaction with the materials and people of the workplace. In extremis, it is claimed that an action can only be understood in the place, in the social situation, and at the time at which it occurred.

- Such a level of contextualization is obviously useless for design, and its advocates will in practice generalize from their observations, even if they ostensibly eschew such generalization.

- Many branches of sociology and anthropology have long recognized that one cannot study people divorced from their social and cultural context. Ethnography has become very influential, particularly in the study of group systems.

- Ethnography is based on very detailed recording of the interactions between people and between people and their environment. It has a special focus on social relationships and how they affect the nature of work.

- The ethnographer does not enter actively into the situation, and does not see things from a particular person's viewpoint. However, an aim is to be encultured, to understand the situation from within its own cultural framework. Culture here means that of the particular workgroup or organization, rather than that of society as a whole. Ethnographers try to take an unbiased and open-ended view of the situation.

Contextual inquiry

- Contextual inquiry has much in common with the ethnographic tradition: it studies the user in context, trying to capture the reality of his work culture and practice. However, it is also an approach rooted in practice and it differs in a number of significant ways from pure ethnographic study: the intention is to understand and to interpret the data gathered, and rather than attempting to take an open-ended view, the investigator acknowledges and challenges her particular focus. In addition, the explicit aim is to design a new system, whereas in a pure ethnographic study, it would be open ended.
- The model of contextual inquiry is of the investigator being apprenticed to the user to learn about his work. Interviews take place in the workplace so that the objects, artifacts and relationships of the work can be better understood. Examples of work are collected and both verbal and non-verbal communication is studied. The idea is to be as comprehensive in the data gathering as possible and to be concrete. Another central notion of contextual inquiry is that of partnership: the user is the expert in the workplace and is therefore encouraged to lead the investigation. The investigator is not a passive observer. Her objective is to gain a shared understanding of how the work happens and, to do so; she questions meaning and offers interpretations of what she observes. The aim is to draw out the implications of comments and actions and understand (rather than assume) what they really mean. In order to do this honestly and effectively the investigator must know her focus – her pre-existing beliefs and assumptions about the situation – and be prepared to challenge and adjust them in the face of new information.
- A number of models of the work are developed to capture what is important in the user's work situation:
- The sequence model elaborates the steps required to complete a specific task, as well as the triggers that initiate that sequence of steps.
  - The physical model maps the physical work environment and how it impacts upon work practice, for example, an office plan showing where different work activities happen.
  - The flow model shows the lines of coordination and communication between the user and other participants within and outside the workplace.
  - The cultural model reflects the influences of work culture and policy and shows the scope of these influences. This may include official or unofficial codes of behavior, common expectations (which may or may not be explicit) and value systems.
  - The artifact model describes the structure and use of a particular artifact within the work process.

3.3 COMMUNICATION AND COLLABORATION MODELS

- single-user or multi-user
- We need to understand normal human–human communication:
- face-to-face communication involves eyes, face and body
- conversation can be analyzed to establish its detailed structure. Look at several levels – minutiae to large scale context:

- face-to-face communication
- conversation
- text based communication
- group working

3.3.1 FACE-TO-FACE COMMUNICATION

- Face-to-face contact is the most primitive form of communication primitive, that is, in terms of technology. The first thing to note is that face-to-face communication involves not just speech and hearing, but also the subtle use of body language and eyegaze. It has a range of these phenomena, and how they influence our use of computer-mediated communications.

a. Transfer effects and personal space

When we come to use computer-mediated forms of communication, we carry forward all our expectations and social norms from face-to-face communication. People are very adaptable and can learn new norms to go with new media. the rules of face-to-face conversation are not conscious, so, when they are broken, we do not always recognize the true problem.

- Personal space also differs across cultures: North Americans get closer than Britons, and southern Europeans and Arabs closer still. This can cause considerable problems during cross-cultural meetings

b. Eye contact and gaze

- Our eyes tell us whether our colleague is listening or not; they can convey interest, confusion or boredom. Sporadic direct eye contact (both looking at one another's eyes) is important in establishing a sense of engagement and social presence. People who look away when you look at them may seem shifty and appear to be hiding something.

- The relative frequency of eye contact and who 'gives way' from direct eye contact is closely linked to authority and power. Naturally, all these clues are lost if we have no visual contact. The problems with direct eye contact, many signals can be easily read through a video channel. This involves not just the eyes, but the whole facial expression, and this is apparent even on poor- quality video or very small (pocket- TV-sized) monitors

c. Gestures and body language

- We use our hands to indicate items of interest. This may be conscious and deliberate as we point to the item, or may be a slight wave of the hand or alignment of the body to allow our colleagues to read our movements. This can be a serious problem since our conversation is full of expressions such as 'let's move this one there', where the 'this' and 'there' are indicated by gestures (or eyegaze). This is called <u>*deictic reference.*</u>

4 Back channels, confirmation and interruption

- It is easy to think of conversation as a sequence of utterances: A says something, then B says something, then back to A. This process is called *turn- taking* and is one of the fundamental structures of conversation. However, each utterance is itself the result of intricate negotiation and interaction. Consider the following transcript:

Alison: Do you fancy that film . . . *er* . . . 'The Green' *um* . . .
it starts at eight.
Brian: Great!

- Alison has asked Brian whether he wants to go to the cinema (or possibly to watch the television at home). She is a bit vague about the film, but Brian obviously does not mind! As Alison says 'that film*er* . . .', she looks at Brian. From the quizzical look on his face he obviously does not know which film she is talking about. She begins to expand 'The Green um . . .',and light dawns; she can see it in his eyes and he probably makes a small affirmative sound 'uh huh'.
- The nods, grimaces, shrugs of the shoulder and small noises are called *backchannels*. They feed information back from the listener to the speaker at a level below the turn-taking of the conversation.

- Back channels -media effects

Restricting media restricts back channels
video – loss of body language audio – loss of facial expression
half duplex – lose most voice back-channel responses
text based – nothing left!

d. Turn-taking

- As well as giving confirmation to the speaker that you understand, and indications when you do not, back channels can be used to interrupt politely. Starting to speak in the middle of someone's utterance can be rude, but one can say something like 'well uh' accompanied by a slight raising of the hands and a general tensing of the body and screwing of the eyes.
- This tells the speaker that you would like to interrupt, allowing a graceful transition. In this case, the listener *requested* the floor. *Turn-taking* is the process by which the roles of speaker and listener are exchanged. Back channels are often a crucial part of this process

4.3.2 Conversation

- It focuses on two-person conversations, but this can range from informal social chat over the telephone to formal courtroom cross-examination. As well as the discipline of *conversational analysis*, there are other sociological and psychological understandings of conversation.
- There are three uses for theories of conversation in CSCW.

1. First, they can be used to analyze transcripts, for example from an electronic conference. This can help us to understand how well the participants are coping with electronic communication.
2. Secondly, they can be used as a guide for design decisions – an understanding of normal human–human conversation can help avoid blunders in the design of electronic media.
3. Thirdly, and most controversially, they can be used to drive design structuring the system around the theory. We will concentrate mainly on the first goal, although this will have implications throughout for design

Basic conversational structure

- Imagine we have a transcript of a conversation production of such a transcriptis not a simple task. For example, a slightly different version of Alison and Brian's conversation may look like this:

> Alison: Do you fancy that film?
> Brian: The *uh* (*500 ms*) with the black cat – '
> The Green whatsit'?
> Alison: Yeah, go at *uh* . . . (*looks at watch - 1.2 s*) . . . 20 to?
> Brian: Sure.

- This transcript is quite heavily annotated with the lengths of pauses and even Alison's action of looking at her watch. The most basic conversational structure is *turntaking*. On the whole we have an alternating pattern: Alison says something, then Brian, then Alison again.
- The speech within each turn is called an *utterance*. There can be exceptions to this turn-taking structure even within two-party conversation. Often we can group the utterances of the conversation into pairs: a question and an answer, a statement and an agreement. The answer or response will normally follow directly after the question or statement and so these are called adjacency pairs . The requirement of adjacency can be broken if the pair is interposed with other pairs for clarification, etc.:

> Brian: Do you want some gateau?
> Alison: Is it very fattening?
> Brian: Yes, very.
> Alison: And lots of chocolate?
> Brian: Masses.
> Alison: I'll have a big slice then.

- This conversation can be denoted: B-x, A-y, B-y, A-z, B-z, A-x. Adjacency pair 'x' ('Do you want some gateau?'–'I'll have a big slice then') is split by two other pairs 'y' and 'z'. One would normally expect the interposed pairs to be relevantto the outer pair, seeking clarification or determining information needed for the response

Context

- Take a single utterance from a conversation, and it will usually be highly ambiguous if not meaningless: 'the *uh* with the black cat – "The Green whatsit"'. Each utterance and each fragment of conversation is heavily dependent on *context*, which must be used to *disambiguate* the utterance. We can identify two types of context within conversation:

  internal context – dependence on earlier utterances.
  external context – dependence on the environment.

Example

> **Brian:** (*points*) that post is leaning a bit
> **Alison:** that's the one you put in

Two types of context:

- external context – reference to the environment
  e.g., Brian's '*that*' – the thing pointed to ← *deictic reference*
- internal context – reference to previous conversation
  e.g., Alison's '*that*' – the last thing spoken of

Common Ground

Resolving context depends on meaning
⇒ participants must share meaning
so must have shared knowledge

Conversation constantly negotiates meaning
… a process called *grounding*:

Alison: So, you turn right beside the river.
Brian: past the pub.
Alison: yeah …

Each utterance is assumed to be:
relevant – furthers the current topic
helpful – comprehensible to listener

Topics, focus and forms of utterance

- Given that conversation is so dependent on context, it is important that the participants have a shared focus. We have addressed this in terms of the external focus the objects that are visible to the participants – but it is also true of the internal focus of the conversation.

Alison: Oh, look at your roses . . . Brian: Mmm, but I've had trouble with greenfly. Alison: They're the symbol of the English summer. Brian: Greenfly? Alison: No roses silly!

Tracing topics is one way to analyse conversation.

- Alison begins – *topic* is roses
- Brian shifts topic to greenfly
- Alison misses shift in focus … *breakdown in communication*

- Alison began the conversation with the *topic* of roses. Brian shifts to the related, but distinct, topic of greenfly. However, for some reason Alison has missed this shift in focus, so when she makes her second utterance, her focus and Brian's differ, leading to the *breakdown* in communication. The last two utterances are a recovery which re-establishes a shared *dialog focus*.
- Another way of classifying utterances is by their relation to the task in hand. At one extreme the utterance may have no direct relevance at all, either a digression or purely social. Looking at the task-related conversation,
- The utterances can be classified into three kinds:

1. **substantive** directly relevant to the development of the topic;
2. **annotative** points of clarification, elaborations, etc.;
3. **procedural** talking about the process of collaboration itself.

- Procedural utterances may be related to the structure of collaboration itself, or may be about the technology supporting the collaboration. The latter is usually in response to a breakdown where the technology has intruded into the communication.

Breakdown and repair

- At a lower level, we may see breakdown due to incorrectly read gestures or eyegaze, and through missed or inappropriate back channel responses. Despite the frequency of breakdowns in normal speech, our communication is not usually significantly affected because we are so efficient at repair.
- Redundancy, frequency of turn-taking and back channels, all contribute to the detection of breakdown and its rapid repair. Electronic communications often reduce redundancy (a single channel), reduce the frequency of turn-taking and reduce back channels. The problem is thus not so much breakdowns in communication, but a reduced ability to recover from them.

Constructing a shared understanding

- The major difference between a book and conversation is that the latter is interactive. The shared knowledge used in a book is static, whereas that used during a conversation is dynamic, as the participants increase their understanding of one another and as they shift their focus from topic to topic.
- When participants come to a conversation, they may come from different backgrounds and bring different knowledge. Even close colleagues will have different recent experiences, and as we have seen in previous examples, have different foci.
- The participants do not try to unify their knowledge and background indeed, they could not fully do so without living one another's lives. Instead, they seek to obtain a *common ground*, a shared understanding sufficient for the task in hand.
- Establishing this common ground will involve negotiating the meanings of words and constructing shared interpretations of the world. Clark and Schaefer refer to this process as *grounding*. The aim of grounding is to construct a meaning *in the conversation* which is sufficient for the task. Two guiding principles for our utterances are that they should be relevant *and helpful*.

Speech act theory

- A particular form of conversational analysis, *speech act theory*, has been both influential and controversial in CSCW. Not only is it an analytic technique, but it has been used as the guiding force behind the design of a commercial system, Coordinator.
- Speech act theory has origins going back over 25 years, but was popularized by Winograd and Flores in the design of Coordinator The basic premise of speech act theory is that utterances can be characterized by what they *do*. If you say 'I'm hungry', this has a certain *propositional meaning* that you are feeling hungry.
- Speech act theory concerns itself with the way utterances interact with the actions of the participants. Individual speech acts can contribute to a conversation. The basic structure of conversations can then be seen as instances of generic conversations. One example of such a generic structure is a *conversation for action* (*CfA*).
- It represents the stages two participants go through in initiating an external action that one of them should perform. There are two variants, the one shown representing a conversation where the first speaker (A) is requesting that the other participant (B) does something. The other, similar, variant is where the first speaker begins with an offer. where the first speaker begins with an offer.
- The numbered circles in Figure are 'states' of the conversation, and the labeled arcs represent the speech acts, which move the conversation from state to state. Note that the speech acts are named slightly differently in different sources (by the same author even!), but the structure of a CfA is the same. The simplest route through the diagram is through states 1–5.

- The network has some nodes marked with a double circle. These are the completion nodes, and at these points neither party expects any more acts by the other as part of this conversation. So the fragment above which left Alison and Brian in state 3 must continue. Of these completion nodes only state 5 represents conclusions where the request has been satisfied.

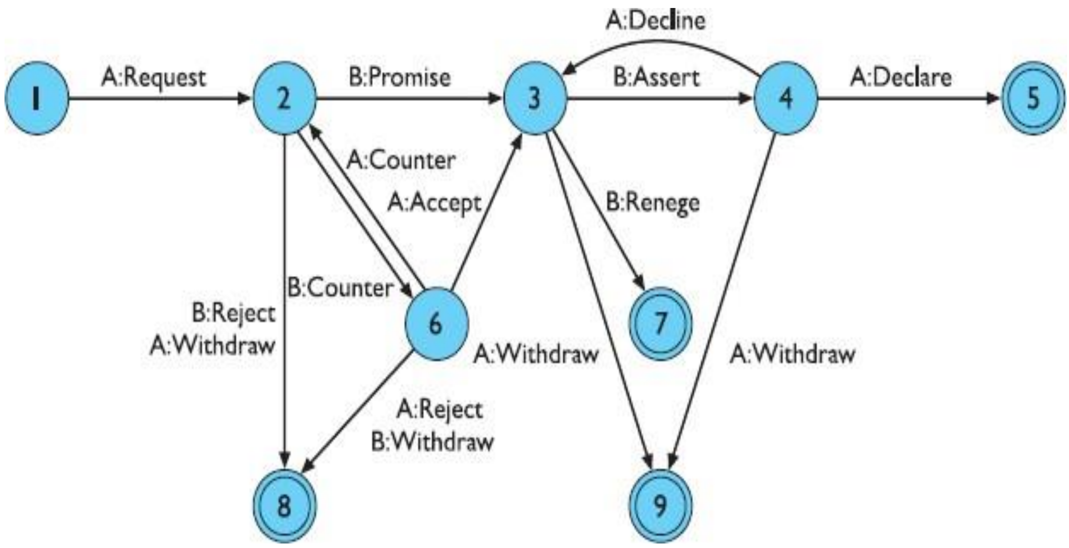


Fig.3.5: Conversation for action

Circles represent 'states' in the conversation

Arcs represent utterances (speech acts)

Stages between Conversations

• Request , • Promise , • Assert ,• Decline , • Reject , • Withdraw , • Counter , • Accept , • Renege , • Declare

There are other generic conversation forms as well as CfA. These include:

1. **conversation for clarification** usually embedded within a CfA to clarify the required action (different from countering a request);
2. **conversation for possibilities** looking toward future actions;
3. **conversation for orientation** building up a shared understanding.

CfA in action

- Simplest route 1–5:

Alison:	have you got the market survey on chocolate mousse?	*request*
Brian:	sure	*promise*
Brian:	there you are	*assert*
Alison:	thanks	*declare*

- More complex routes possible, e.g., 1–2–6–3 ...

Alison:	have you got ...	*request*
Brian:	I've only got the summary figures	*counter*
Alison:	that'll do	*accept*

4.3.3 TEXT-BASED COMMUNICATION

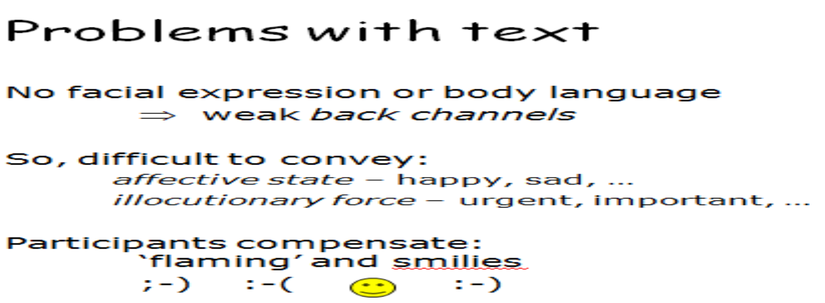
- For *asynchronous* groupware (and even some synchronous systems), the major form of direct communication is text based. There are exceptions to this, for instance voice messaging systems and answer phones, and other media may be used in addition to text such as graphics, voice annotation or even video clips. But despite these, text is still the dominant medium.
- Types of electronic text:
  - discrete directed messages, no structure (email)
  - linear messages added (in temporal order)
  - non-linear hypertext linkages
  - spatial two dimensional arrangement
- Text-based communication is familiar to most people, in that they will have written and received letters. However, the style of letter writing and that of face-to-face communication are very different. The text-based communication in groupware systems is acting as a speech substitute, and, thus, there are some problems adapting between the two media. There are four types of textual communication in current groupware:

a) **discrete** – directed message as in email. There is no explicit connection between different messages, except in so far as the text of the message refers to a previous one.
b) **linear** – participants' messages are added in (usually temporal) order to the end of a single transcript.
c) **non-linear** – when messages are linked to one another in a hypertext fashion.
d) **spatial** – where messages are arranged on a two-dimensional surface.

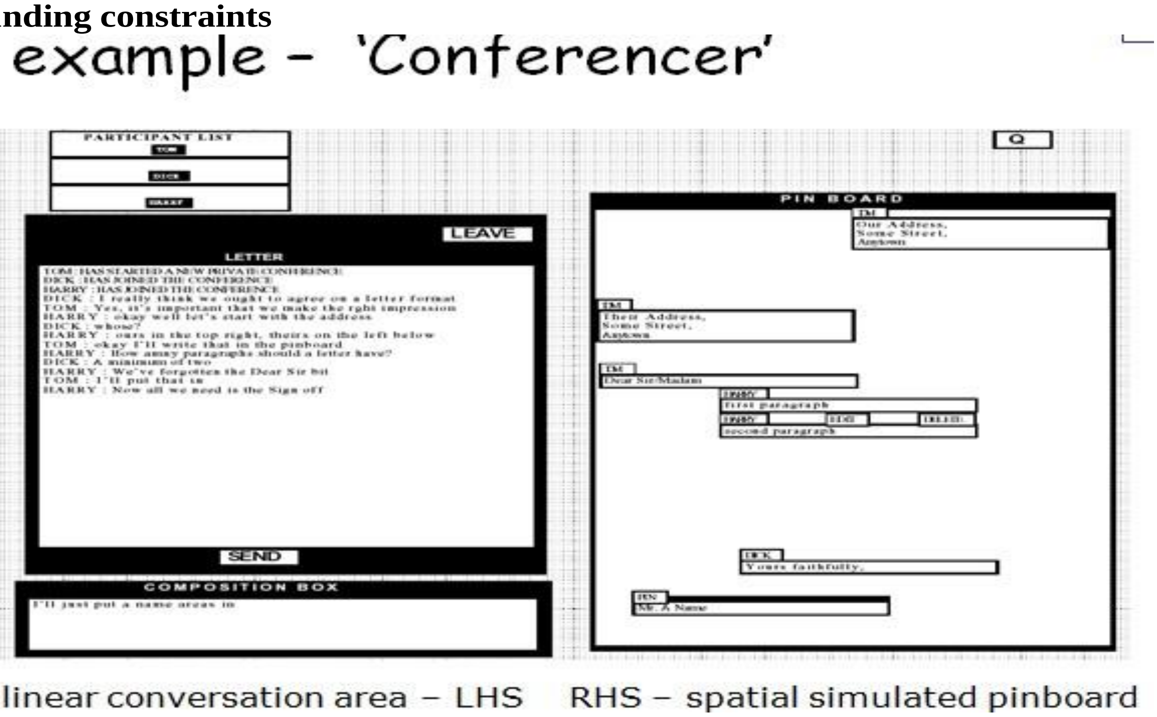
Back channels and affective state

- Much of the coordination of face-to-face conversation depends on back channels and interpretation of the listener's expressions. Text-based communication loses these back channels completely.
- Consider the effect of this on even a two-party conversation.The speaker would pause to seek back channel confirmation or to offer the floor, the text 'speaker' must either continue regardless, or finish the message, effectively passing the turn.
- One consequence of the lack of interruptions and more measured pace of interaction is that the utterances are more grammatical than speech. Inaddition to this loss of back channels, the speaker's tone of voice and body language are of course absent. These normally convey the *affective state* of the speaker (happy, sad, angry, humorous) and the *illocutionary force* of the message (an important and urgent demand or a deferential request).
- Email users have developed explicit tokens of their affective state by the use of flaming' and 'smilies', using punctuation and acronyms; for example:

  :-) – smiling face, happy
  :-(– sad face, upset or angry
  ;-) – winking face, humorousLOL – laughing out loud.



Grounding constraints



 This grounding process is linked strongly with the types of channels through which the conversants communicate. Clark and Brennan describe the properties of these channels in terms of *grounding constraints*. These include **cotemporality** – an utterance is heard as soon as it is said (or typed); **simultaneity** – the participants can send and receive at the same time; **sequence** – the utterances are ordered.

 These are all constraints which are weaker in text-based compared with face- to-face interaction. text-based system, different participants can compose simultaneously, but they lack cotemporality. Even if the messages appear as they are produced, they will not be read in real time. In addition, the messages may only be delivered when complete and even then may be delayed by slow communications networks.

 Linear transcripts obviously have some idea of sequence, but this is confused by the overlap and interleaving caused by the lack of cotemporality and imultaneity. Consider this typical interchange during the use of the York Conferencer system:

1. Bethan: How many should be in the group?
2. Rowena: Maybe this could be one of the four strongest reasons?
3. Rowena: Please clarify what you mean.
4. Bethan: I agree.
5. Rowena: Hang on.
6. Rowena: Bethan what did you mean?

 In a spoken conversation, Rowena and Bethan would have quickly corrected themselves if they began to speak at once, and the linearity would have reflected a *common* experience. The trouble is that the participants in the text- based conference each experienced the messages in a different order:

Rowena: 2 1 3 4 5 6
Bethan: 1 2 4 3 5 6

Turn-taking

- The fundamental structure of conversation was *turn-taking*. The last transcript was an example of a breakdown in turn-taking. Breakdowns are quite rare in two-party electronic conversations and are quickly corrected.

- In a pair of participants, turn-taking is simple; first one person says something, then the other. The only problem is deciding exactly *when* the exchange should happen. With three or more participants, turn-taking is more complex. They must decide *who* should have the next turn. This is resolved by face-to-face groups in a number of ways.

➢ First, the conversation may, for a period, be focused on two of the parties, in which case normal two-party turn-taking holds.
➢ Secondly, the speaker may specifically address another participant as the utterance is finished, either implicitly by body position, or explicitly: 'what do you think Alison?'
➢ Finally, the next speaker may be left open, but the cotemporality of the audio channel allows the other participants to negotiate the turn.

- Basically, whoever speaks first, or most strongly, gets in. These mechanisms are aided by back channels, as one of the listeners may make it clear that she wants to speak. In this case, either the speaker will explicitly pass the turn (the second option above), or at least the other listeners are expecting her to speak.
- The movement between effective two-party conversation (the first option) and open discussion will be mediated by back channel messages from the other participants.
- In an unstructured text-based conversation the third option is not available, nor, of course, are the back channels. Paired conversation is quite common and the second option, explicitly naming the next speaker, is possible.

Context and deixis

- Utterances are highly ambiguous and are only meaningful with respect to *external context*, the state o the world, and *internal context*, the state of the conversation. Both of these are problems in text-based communication.
- The very fact that the participants are not co-present makes it more difficult to use external context to disambiguate utterances. This is why many groupware systems strive so hard to make the participants' views the same; that is, to maintain WYSIWIS ('what you see is what I see').
- The means of direct communication, remote participants have difficulty in using deictic reference. They cannot simply say 'that one', but must usually describe the referrant: 'the big circle in the corner'. If their displays are not WYSIWIS then they must also ensure that their colleague's display includes the object referred to and that the description is unambiguous.
- Asynchronous participants have even more problems with deixis as there is no opportunity for their colleagues there are also problems with deictic reference to internal context

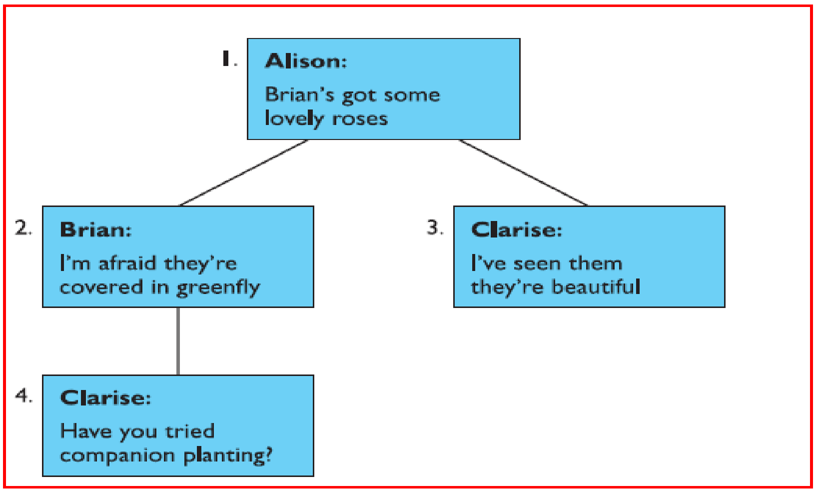


Fig3.6: Hypertext Conversation Structure

- In speech, the context is intimately connected to linear sequence and adjacency. As we have seen, even in linear text transcripts, overlap breaks the strict sequentiality of the conversation, and thus causes problems with indexicals and with context in general.

1. Alison: Brian's got some lovely roses. 2. Brian: I'm afraid they're covered in greenfly. 3. Clarise: I've seen them, they're beautiful.

- Hypertext-based systems avoid the implied sequentiality of a linear transcript. In the example, both Brian and Clarise replied to Alison's message at the same time. In a hypertext these would form parallel conversations. This is shown in above Figure , where in addition Clarise has sent a second message offering advice on Brian's greenfly. The use of 'they' in Clarise's message is now perfectly clear.

Pace and granularity

Pace of conversation – the rate of turn taking face-to-face – every few seconds telephone – half a minute email – hours or days **as the pace of a conversation reduces, there is a tendency for the *granularity* *to increase***

- The term *pace* is being used in a precise sense above. Imagine a message being composed and sent, the recipient reading (or hearing) the message and then composing and sending a reply. The pace of the conversation is the rate of such a sequence of connected messages and replies.
- Clearly, as the pace of a conversation reduces, there is a tendency for the *granularity* to increase. To get the same information across, you must send more per message. Even most monologs are interactive in the sense that the speaker is constantly looking for cues of comprehension in the listener.

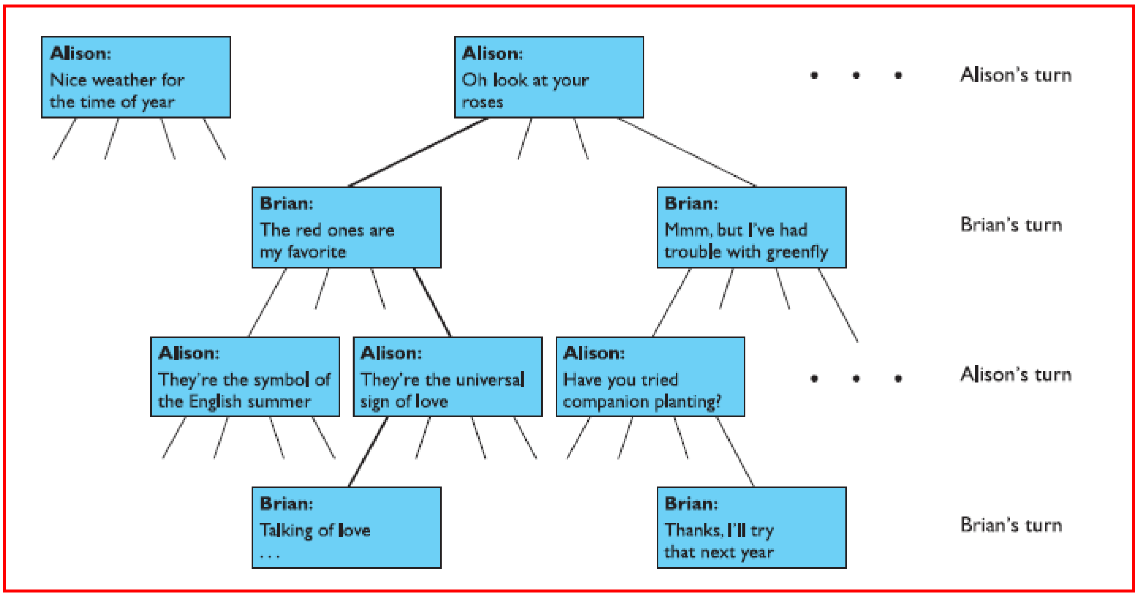


Fig3.7: Conversation Game

- Reducing the pace of a conversation reduces its *interactivity*. at the small scale of clarifying individual utterances, interactivity is important in determining the direction of a conversation. Imagine that the conversation is a little like a game, where the participants can make moves. In Figure , we can see some of the moves Alison and Brian can make whilst talking in the garden (Clarise has gone home).
- At each turn of the conversation, Alison or Brian can choose to say one thing which continues the discussion. That is, they gradually work out a path from the top of the tree downwards. A particularly promising conversation path is shown with bold lines. In a hypertext-based system one can expand several branches of a conversation tree, but in speech or in a linear text transcript the conversation follows one branch.
- Whatever medium is used, you cannot normally progress down the tree faster than the pace of the conversation. To overcome these limitations, people adopt several *coping strategies*. The simplest strategy is just to avoid conversation. This can be done by delegating parts of a task to the different participants. Each participant can then perform much of the task without communication. They must still communicate for large-scale strategic decisions, but have significantly reduced the normal communications.
- This approach reduces communication by reducing collaboration. More interesting in a cooperative work setting are two coping strategies which increase the chunk size of messages in order to reduce the number of interactions required to complete a task. These strategies are frequently seen in both text-based conferences and in letter writing.
- The first of these coping strategies is multiplexing. Basically, the conversants hold several conversations in parallel, each message referring to several topics. In terms of the conversation tree, this corresponds to going down several branches at once.

Linear text vs. hypertext

- Considerations of potential overlap suggest that hypertext-based communications may be better suited as a text-based communication medium. Similarly, the problems of pace may be partially solved in a hypertext.
- Multiplexed messages can be represented as updates to several parts of the hypertext, thus reducing the likelihood of breakdown and lost topics. In

addition, if the messages themselves can be mini-hypertexts, then eager messages listing several possible courses of action can be explicitly represented by the message.

- On the other hand, hypertext has its disadvantages. Even static hypertexts, which have been carefully crafted by their authors, can be difficult to navigate.
- A hypertext that is created 'on the fly' is unlikely to be comprehensible to any but those involved in its creation. Conklin and Begeman, themselves associated with the hypertext based argumentation tool gIBIS, conclude that 'traditional linear text provides a continuous, unwinding thread of context as ideas

4.3.4 GROUP WORKING

- We have been principally looking at the properties of direct communication and largely two-party conversations. Group behavior is more complex still as we have to take into account the dynamic social relationships during group working.
- We will begin by looking at several factors which affect group working, and then discuss the problems of studying group working.

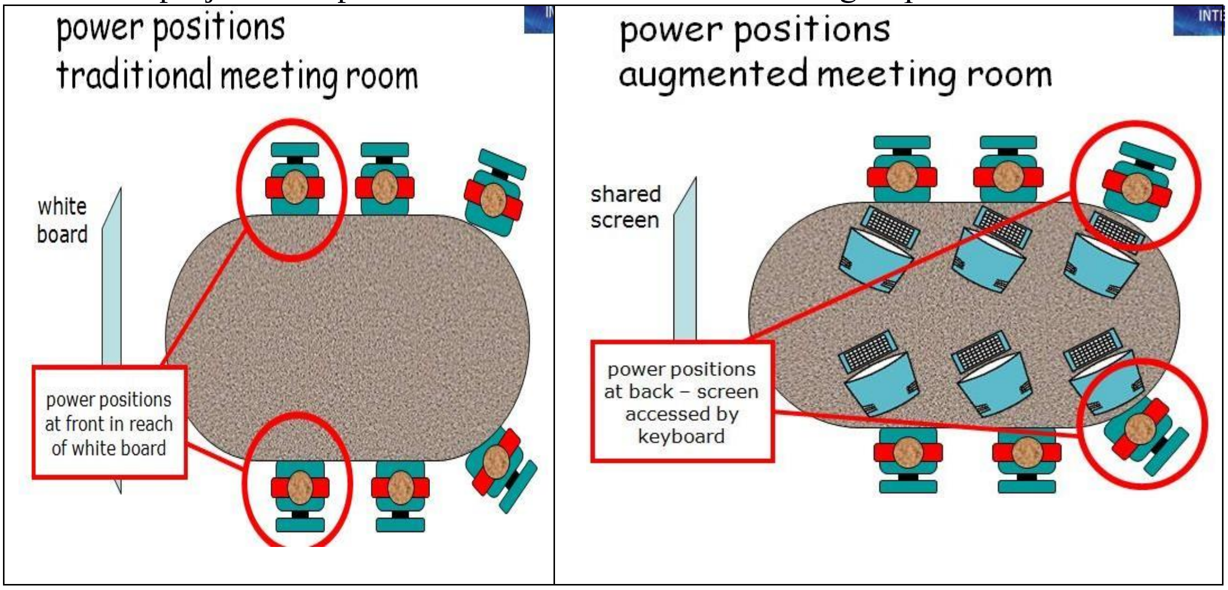
Group dynamics

- Whereas organizational relationships such as supervisor/supervisee are relatively stable, the roles and relationships within a group may change dramatically within the lifetime of a task and even within a single work session that systems, such as co-authoring systems, which use a formal concept of *role*, must allow these roles to change together with the socially defined roles.
- Even the naming of roles can cause problems. A person may be an author of a book or paper, but never write the words in it, acting instead as a source of ideas and comments. A particular case of this is the biographical story where the individual concerned and a professional writer co-author the book, but only the professional author writes.
- A co-authoring system such as Quilt would call the non-writing author a 'commentator' or a 'reviewer', but *not* an 'author'. Not only do the social relationships *within* the group change, but the group membership and structure can change in time.
- A member leaving or a new member joining can cause dramatic changes in the behavior of the group. Groupware systems, for example argumentation *tools*, can help in that they record the history of the group. Groupware designers should in general be aware that new members can and will enter the group and should design their software accordingly. The group may also divide into subgroups for detailed discussion and then reform. Tools must be able to support this.

Physical layout

- The physical layout of a room has a profound effect upon the working relationship of those in it. This is particularly obvious for meeting rooms, but should be considered in any group-working environment.
- If we wish to encourage conversation, as we do in a meeting room, the participants must be encouraged to look toward one another.
- Meeting rooms have a natural focus toward the screen at the front of the room, but inward-facing terminals can counteract this focus and thus encourage eye contact. the users still had some difficulty in adapting to the *power positions* in the electronic meeting room.
- At first sight, the electronic meeting room is not unlike a normal conference room. If the shared screen is a whiteboard or an overhead projector, then the most powerful position is toward the front of the room. Managers would

normally take this seat as they can then easily move to the whiteboard or overhead projector to point out some item and draw the group's attention.



Distributed cognition

- Human cognition, but the emphasis was, as in all traditional psychology, upon the activity *within* the person's head. A school of thinking has recently developed which regards thinking as happening not just within the head, but in the external relationships with things in the world and with other people. This viewpoint is called distributed cognition. This viewpoint is not as radical as it first appears.
- Traditional views talk about the movement of information between working memory and long-term memory: it is not so difficult then to regard bits of paper, books and computer systems as extensions to these internal memory systems.

3.4 HYPERTEXT, MULTIMEDIA AND WWW

3.4.1 Understanding Hypertext

- Hypertext attempts to get around these limitations of text by structuring it into a mesh rather than a line. This allows a number of different pages to be accessed from the current one, and, if the hypertext is well designed, the user should find it easier to follow his own particular idea through the mesh rather than being forced down one route.
- Typically, hypertext systems incorporate diagrams, photographs and other media as well as text. Such systems are often known as *multimedia* or *hypermedia* systems, although the three terms are often used interchangeably.
- A hypertext system comprises a number of pages and a set of *links* that are used to connect pages together. The links can join any page to any other page, and there can be more than one link per page. Thus a hypertext document does not simply start a linear progression and follow it to an end, but goes in lots of different directions, some of which terminate, while others link back into different parts of the document.
- There are many different ways of traversing the network, and so there are many different ways of reading a hypertext document – the intention is that the user is able to read it in the way that suits him best. Links can exist at the

end of pages, with the user choosing which one to follow, or can be embedded within the document itself.

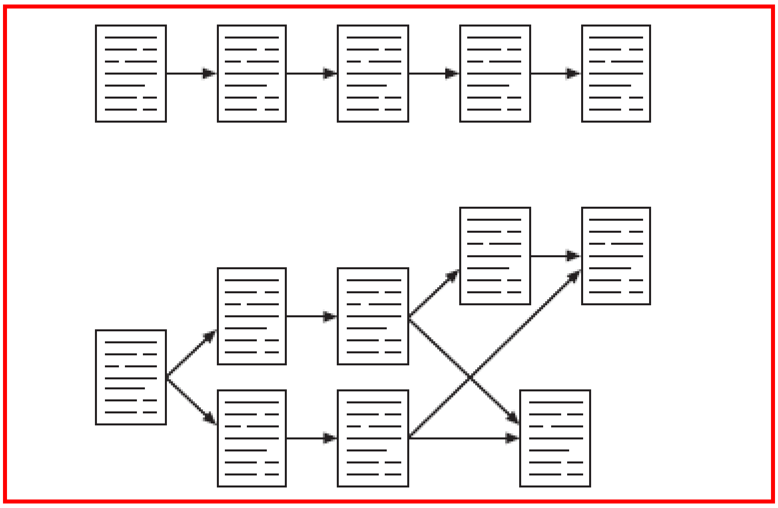


Fig3.8: Typical structures of linear text and hypertext

Rich content

As well as static material – text and static diagrams and photographs – hypertext systems may also include more dynamic material such as animation, video and audio clips, and even full computer applications**.**

Animation

- Animation is the term given to the addition of motion to images, making them move, alter and change in time. A simple example of animation in an interface is in the form of a clock. Digital clocks can flick by the seconds, whilst others imitate Salvador Dali and bend and warp one numeral into the next. Analog clocks have moving hour and minute hands, with an optional second hand sweeping round the clock face

Video and audio

- A media dominated world, there are strong arguments for using video or audio material as part of hypertext systems whether for education, entertainment or reference. Both audio and video material are expensive and time consuming to produce, but increasingly even home-PC systems include video and audio editing as standard.

Delivery technologyon the computer

- Many hypermedia systems are supplied on CD-ROM. This has the advantage of reasonably large capacity (650–700 Mbytes), but access is slower than with installed systems. For highly dynamic material, such as educational media, a special player is installed; alternatively, material such as software documentation may use a standard format such as web pages.

On the web

- World wide web is the best-known multimedia hypertext system of all. The World Wide Web offers a rich environment for the presentation of information.
- Documents can be constructed that are very different from paper versions; basic text can be augmented through the use of hypertext links to other documents, while graphics can easily be incorporated as pictures, photographs, icons, page dividing bars, or backgrounds. Pages can also have hypertext links embedded into different regions, which take the user to a different page or graphic if they are clicked on; these are known as activeimage maps.

On the move

- Mobile phones, PDAs (personal digital assistants), and notebook computers have all increased the demand to have hypermedia available on the move. Furthermore across many countries governments have sold franchises for high-bandwidth mobile services. After spending billions on these franchises the telecommunications giants really want people to use new mobile services!
- Notebook computers can use just the same mechanisms as desktop computers, using CD-ROM or DVD for standalone material, or connecting to the web through wireless access points or through modems linked to mobile phone networks. The fact that the computer is mobile means that location can be used as a key into context-aware hypermedia showing different content depending on location.
- The ‘stick e note’ system developed by the University of Kent uses a sticky note metaphor with notes stuck to particular locations. Only when you visit the location does the note become visible. This is a bit like an image map on a web page, but rather than clicking a mouse over an image to link to information, here you need to physically move to a location! Another example is the GUIDE system, which uses various means to detect location (closest network access point or GPS) and then delivers appropriate tourist information.

Application areas

The type of domains in which hypermedia systems have proved successful, looking briefly at some example systems.

Rapid prototyping

- Although now lacking the wealth of features expected of a hypermedia system, HyperCard on Macintosh computers has been very influential as a basis for experimental hypertext systems. HyperCard uses the metaphor of a card index, around which the user can navigate. Each card can hold text, diagrams, photographs, bitmaps and so on, and hot-spots on the cards allow movement between cards.
- Cards may also contain forward and backward buttons and a home icon, to allow the user to move sequentially and start from scratch respectively. HyperCard can be used for a range of applications including information management and teaching.
- HyperCard’s simple scripting language and easy to produce graphical interfaces meant it was also used extensively as a rapid prototyping tool for generating interactive systems. In fact, HyperCard stacks for both single-user and networked applications are available from the book website.

Help and documentation

- Hypermedia systems are ideally suited to online manuals and other help system applications. They allow user-oriented access to the information, and support browsing. In addition the information can be organized hierarchically, with successive selections providing more detailed information. This supports the varying needs that users have, such as quick reference, usage information, full details and so on. Many commercial help systems use hypermedia-style help.

Education and e-learning

- Hypertext and hypermedia are used extensively in educational settings, as they allow varied subjects to be related to each other in numerous ways so that the learner can investigate the links between different areas. In contrast to computer-aided learning (CAL) packages, hypermedia allows a student-controlled learning process. This is a hypermedia system built and used at

Brown University to support teaching in subjects as varied as English literature and biology.

- The system includes text, diagrams, photos and so on. Both learners and teachers can add information and links, giving students access to each other's opinions as well as those of their tutors. A map provides an overall view of the information for direct access and navigation, with links providing browsing facilities in the normal way. Intermedia has been successfully used for university-level teaching, and can be seen as a forerunner of the educational resources now facilitated by the web.

Collaboration and community

- Although strictly not hypertext, the web has become a central platform for collaborative applications and community. These use the hypertext structure of the web to structure and access shared resources and message areas.
- Establishing a sense of community can be very important on websites as it is one way to ensure loyalty and get visitors to return. This may involve explicit community features such as chat areas, or may simply be a matter of using a design, language and image that suggests a site which is open and listening to 'readers'.

Finding Things Lost in hyperspace

- Although the non-linear structure of hypertext is very powerful, it can also be confusing. It is easy to lose track of where you are, a problem that has been called 'lost in hyperspace'. There are two elements to this feeling of 'lostness'.
- The first is cognitive and related to content. In a linear text, when a topic is being described, the writer knows what the reader has already seen. In a hypertext, the reader can browse the text in any order.
- Each page or node has to be written virtually independently, but, of course, in reality it cannot be written entirely without any assumption of prior knowledge. As the reader encounters fragmentary information, it cannot be properly integrated, leading to confusion about the topic.
- The second is related to navigation and structure. Although the hypertext may have a hierarchical or other structure, the user may navigate by hyperlinks that move across this main structure. It is easy to lose track of where you are and where you have been.

Designing structure

- In some areas there may be preexisting understood structures to mirror; for example, the faculty and departmental structure of a university, or the main disciplines (circulatory, neurological, etc.) within medicine.

Making navigation easier

- No matter how well designed the site structure is, there will still be problems: because the user does not understand the structure; or because the user has individual needs that the designer has not foreseen; or because even a good structure is not perfect. Another type of hypertext takes the form of 'levels of access' to a document. Different levels of access privilege 'see' different amounts of information.
- A document structured in this way may provide one level of access that gives only a brief overview of the topic. The next level of access presents a fuller description of the system, while the next level may also include information regarding the precise meaning of technical terms used in the system. The final level of access may add historical information and such like

History, bookmarks and external links

- Hypertext viewers and web browsers usually have some sort of history mechanism to allow you to see where you have been, and a more stack-based system using the 'back' button that allows you to backtrack through previously visited pages. The back button may be used where a user has followed a hyperlink and then decided it was to the wrong place, or alternatively, when browsing back and forth from a central page that contains lots of links.
- The latter is called hub and spoke browsing. Although the back button is used extensively, it is used relatively little to go back more then one step. For error correction this makes sense, but for general revisiting one might think that moving back several steps would be common.

Indices, directories and search

- An index is not a complete list of all words in a document. If this were the case then the index for this book would be as big as the rest of the book! The words in an index are chosen because they are significant key phrases or words with a domain meaning, and not every occurrence of a word is indexed, only those deemed in some way important.
- The main difference between an electronic index and a paper one is that with the paper index you have to physically look up the page after finding the word in the index, whereas in an electronic index the links are 'live' so you can simply click to the content.

3.4.2 Web Technology And Issues Basics

- The web consists of a set of protocols built on top of the internet that, in theory, allow multimedia documents to be created and read from any connected computer in the world. The web supports hypertext, graphics, sound and movies, and, to structure and describe the information, uses a language called HTML (hypertext markup language) or in some cases, XML (extensible markup language). HTML is a markup language that allows hypertext links, images, sounds and movies to be embedded into text, and it provides some facilities for describing how these components are laid out.
- HTML documents are interpreted by a viewer, known as a browser; there are many browsers, and each can interpret HTML in subtly different ways, or support different levels of functionality, which means that a web page viewed through one browser can look very different from the same page viewed through another. web owes its success to many factors, including the robustness and (relative) ease of use offered by popular browsers from the very first graphical browser Mosaic, and continued in commercial browsers such as Netscape Navigator, Microsoft Internet Explorer and Opera. These offer a graphical interface to the document, controlled by the mouse. Hypertext links are shown by highlighting the text that acts as the link in an alternative color, and are activated by clicking on the link. A further color is used to indicate a link that has already been visited. Hypertext links can also be embedded into regions within an image.

Web servers and web clients

- Whereas a conventional PC program runs and is displayed on one computer, the web is distributed. Different parts of it run on different computers, often in different countries of the world. They are linked, of course, by the internet, an enormous global computer network.

- The pages are stored on web servers that may be on a company's own premises or in special data centers. Because they are networked, the webmaster for a site can upload pages to the server from wherever she is.

Network issues

- The fact that the web is networked raises a series of issues that can impact on usability. Network capacity is called bandwidth. This is a measure of the amount of information that can pass down the channel in a given time.
- For example, a typical modem speed is 56 kbs – that is 56 kilobits per second. This equates to about 6000 characters per second. This sounds fine until you realize that images may take many tens or hundreds of characters (bytes) to encode . . . this is why many have renamed the web the 'world wide wait'! bandwidth is not the only important measure.
- There is also the time it takes for a message to get across the network from your machine to the web server and back. This delay is called latency.
- Latency is caused by several factors – the finite speed of electrical or optical signals (no faster than the speed of light), and delays at routers along the way that take messages from one computer network and pass them on. This latency may not always be the same, varying with the exact route through the network traveled by a message, the current load on the different routers,etc. Variability in the latency is called jitter.

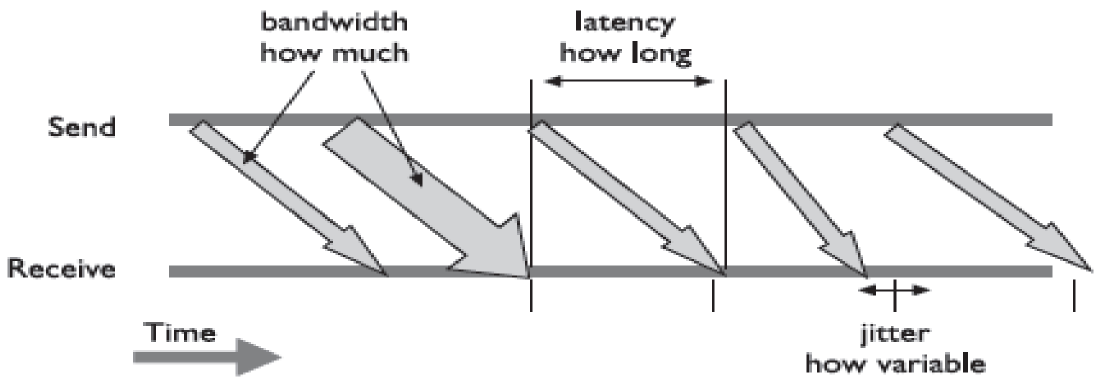


Fig.3.9: Bandwidth, Latency and Jitter

3.4.3 Static Web Content

The message and the medium

- Excellent page design can make useless material look attractive, but it still remains useless material. Poor design can mean that excellent material is never seen by potential readers, as they have become bored, or intolerant of the medium, or confused, or for a host of other reasons have aborted their attempts to download and view the information.
- Pages do have to look immediately interesting and attractive if people are to spend time, effort and, because of the communication costs, money, in viewing them; the user-centered nature of the medium makes this imperative.
- This is in marked contrast to television or cinema or other dynamic media, which are not under any direct user control, where information is presented to a passive audience. With web documents, people have actually to want to see the information, and make an effort to retrieve it, which clearly must have an influence on design.

Text

- Because web pages are displayed on many different machines, there are only a small set of fonts that can be guaranteed to be available: a standard font and a typewriter font (e.g. courier) with bold and italic versions in different sizes. It is possible to specify preferred fonts and many of these such as Arial, Verdana or

- Comic Sans are available on most web platforms. The various structured styles such as headings allow the web designer to create material that will lay out passably on all platforms. But these offer a fairly coarse level of control. The size and boldness of the heading should be chosen carefully; for example, huge dark fonts on a page can look loud and brash. There is an increasing desire to have fine control.
- Cascading style sheets (CSS) allow you to specify fonts, line spacing, size, etc., in a similar way to styles in a word processor or DTP package. The use of color is of great importance for web pages, but it is often abused.
- First, it should be remembered that a significant proportion of the potential viewers of the page will have problems with color, either because they are using older machines with a limited color palette, or because they have some form of color blindness. Color, when used, should not be the only cue available.

Graphics Obtaining graphics

- There are a number of sites on the web that contain archives of graphical images, icons, backgrounds and so on. There are also paint and image manipulation packages available on almost all computer systems, and scanners and digital cameras, where available, enable the input of photographs and diagrams.

Using graphics

- While graphics and icons tend to play a significant role in web page design, their use should be carefully thought out. Graphical images take longer to load than text, and this may become a problem. Text uses 8 bits to represent a character: some rough calculations show that approximately 2000 characters represent about a screenful of information, and so 16,000 bits (2 K) are required.

Icons

- Icons often appear on web pages, and while there are many available to choose from, they should be used with care. On web pages, icons are typically used in one of two ways. They are either visual cues, associating some small picture with different parts of the text (for example, some pages have icon-sized characters that appear next to instructions).
- Alternatively, they are used in much the same way as in a standard. WIMP interface to represent aspects of the functionality of the underlying pages. In this latter case, they must represent their associated functionality in either a concrete or an abstract form. This means that the design of the individual icon has to be carefully thought out, as a lot of information may have to be represented in a small area of screen estate.

Movies and sound

- Movies and sound are both available to users of the web, and hence to page designers. One problem associated with them is actually obtaining appropriate sound and video clips, as they usually require some sort of multimedia capability on behalf of the host machine in order to be able to digitize sound and capture and digitize video.
- Video suffers from the same problems as graphics, magnified by an order of magnitude or two; it can take extremely large amounts of time for a video segment to download.
- Video is also not well integrated into the web, requiring the creation of a process to run it that is separate from the page from whence it came. Not all receiving machines have the capability to play video, or sound, and so it is

unwise for a designer to rely on these dynamic media to convey information without replicating it elsewhere.

- The use of sound and video moves page design further away from the typesetter and toward the sound engineer and cinematographer; the integration of these cinematic media with the enhanced textual capabilities offered by the web is a new domain, in which the techniques that work and those that fail have not yet been fully explored, let alone understood.
- The need to download movies and sound puts sharp limits on the length of clip that can be shown. Streaming media over the internet, such as RealVideo, RealAudio and CuSeeMe, allow potentially unlimited sources.

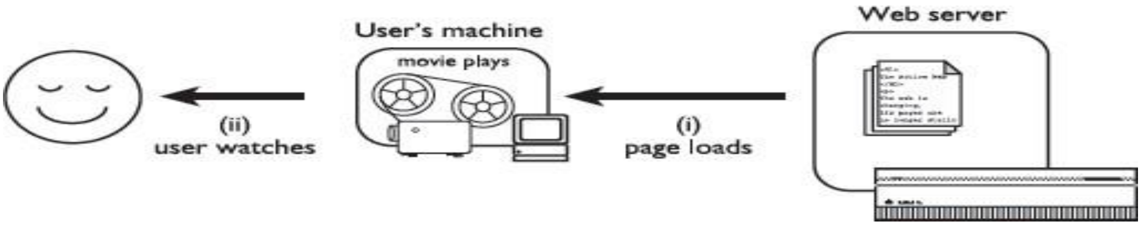


Fig.3.10: Animated GIF or movie needs to download completely

3.4.4 Dynamic Web ContentThe active web

- The web was simply a collection of (largely text) pages linked together. The material was static or slowly changing and much of it authored and updated by hand. Some pages were generated on the fly, in particular the gateways into ftp servers and to gophers, which were so important in adding 'free' content to the web.

What happens where

- When considering dynamic material on the web we need to take the external, user's viewpoint and ask what is changing: media, presentation or actual data; by whom: by the computer automatically, by the author, by the end-user or by another user; and how often, the pace of change: seconds, days or months? From a technical standpoint, we also need to know where 'computation' is happening: in the user's web-browsing client, in the server, in some other machine or in the human system surrounding it? The 'what happens where' question is the heart of architectural design.
- It has a major impact on the pace of interaction, both feedback, how fast users see the effects of their own actions, and feedthrough, how fast they see the effects of others' actions. The user view One set of issues is based on what the end-user sees, the end-user here being the web viewer. What changes? This may be a media stream (video, audio or animation) which is changing simply because it is the fundamental nature of the medium.

Technology and security

- The fundamental question here is where 'computation' is happening. If pages are changing, there must be some form of 'computation' of those changes. Where does it happen? Client One answer is in the user's web-browsing client enabled by Java applets, various plug-ins such as Flash, scripting using JavaScript or VBScript with dynamic HTML layers, CSS and DOM (Domain Object Model).
- Server A second possibility is at the web server using CGI scripts (written in Perl, C, UNIX shell, Java or whatever you like!), Java Servlets, Active Server Pages or one of the other server-specific scripting languages such as PHP.

- In addition, client-side Java applets are only allowed to connect to networked resources on the same machine as they came from. This means that databases accessed from clientside JDBC (Java database connectivity) must run on the web server (see below). Another machine Although the pages are delivered from the web server, they may be constructed elsewhere. For hand-produced pages, this will usually be on the page author's desktop PC. For generated pages, this may be a PC or a central database server. People Of course, as noted earlier, the process of production and update may even involve people!

Fixed content – local interaction and changing views

- Probably the most hyped aspect of the web in recent years has been Java. In fact, Java can be used to write server-end software and platform independent standalone programs (not to mention the embedded systems for which it was originally designed!),but the aspect that most people think of is Java applets.
- Applets are just one of the techniques that can be added to give client-end interaction (and about the least well integrated into the rest of the page). The most common alternatives are JavaScript, Flash and if you are prepared to limit yourself to Windows platforms, ActiveX plug-ins. These techniques share the characteristic that they are downloaded to the user's own machine and thereafter all interaction happens on the PC, not across the network

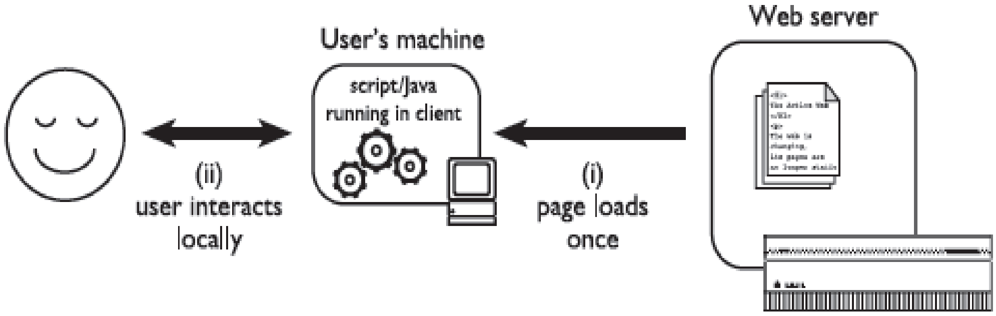


Fig.3.11: Java applet or Javascript running locally

Search

- Some user-driven interaction can be accommodated at the client end, but not all. Consider search engines. It would be foolish to download several megabytes of information so that a Java applet can search it online! Instead, all common websearch pages work by submitting forms to the server where CGI programs perform the searches and return results.
- An additional reason for this approach is that most browsers support forms, but some still do not support Java or scripting in a consistent manner. The web search engine for this book works in this way. The user's keywords are submitted to the server using an HTML form, they are compared against pre-prepared indexes at the server and all matching paragraphs in the book are returned

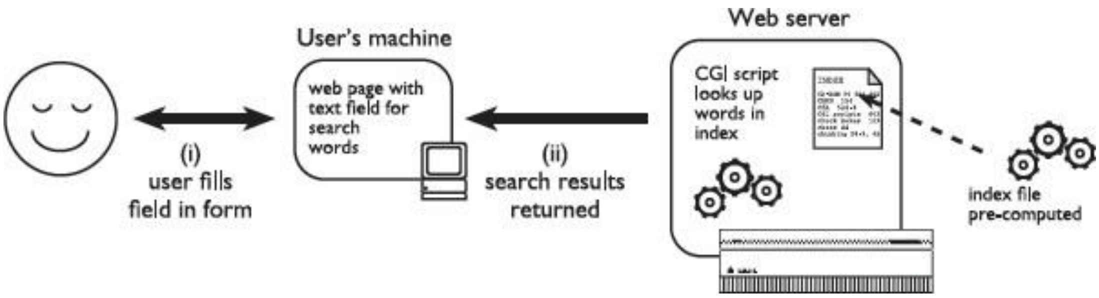


Fig.3.11: HCI Book Search

Dynamic content

- The mechanisms we have been discussing manage the feed through when the database is updated by some non-web means. Perhaps the most ‘active’ web pages are those where the content of the pages reacts to and is updateable by the web user. If pages are generated from database content using either the Java-applet/JDBC method or the CGI method, the same mechanisms can as easily be used to update as to access the database.
- The feedback of changes to the user is thus effectively instantaneous – you check for seat availability on the theatre web page, select a seat, enter your credit card details and not only is the seat booked, but you can see it change from free to booked on the web page.
- This sort of web application opens up many additional problems. You may need to add some form of login or authentication. If credit card numbers are supplied you need to ensure that the web server is secure.