T213 - Modelling and simulation: digital experimentation, demographic and environmental simulation by Nitish

Definitions for Modeling:

Modeling: a preliminary sculpture in wax or clay from which a finished work can be copied **Model**: the act of representing something (usually on a smaller scale)

Definitions for Simulation:
Simulating a design through software programs that use models to replicate how a device will perform in terms of timing and results.

Simulation – creating computer versions of real-life
Simulation is mimicry. Never as good or true as reality but as good as the basic model governing it (our synthesized perception of reality) can be. Notice that modeling and simulation are necessarily close-tied together. Simulation allows us to experiment with a "reality" that is, for example, too dangerous to let happen, not yet realized or hypothetical, in the course of development (where we want to know more about e.g., feasibility, performance, reliability, etc.), and so on. (WWW)

Simulation is very important for network performance evaluation. In fact studies show simulation is a very important tool for decision making too. Simulation can be used in various fields - contention, collisions, momentary network overload, crash of a node, etc.

Impact of Simulation:
We hear simulation studies regarding nuclear waste disposal, water management in developing countries, and so on. We all use simulators for our training, studies and most importantly, decision making.

http://www.acm.org/crossroads/xrds9-2/elzas.html

However, simulation is still considered by many as the "method of last resort." When all other methods fail, we use simulation. To go one step further, in some cases (flawed) simulation results are used as justification, or should I say alibi, for decisions that otherwise would not stand any criticism. Is it time to coin the expression "lies, damn lies, and simulation?"

The size of the lie depends on the knowledge that we can put in the model. In this way, we can cover the whole field from realistic performance assessment to pure conjecturing. The first applies to "hard" systems (technical like cars, airplanes or computers without taking human interaction into account); the second to "soft" systems (with living, e.g., human, components like we find in sociological models, economic models, interaction models, etc.) It all depends on how far we can get with validation, i.e., proving that the model or the simulation is "right" by comparing the results with real-life experiments. Unfortunately, in many cases, the required experiments are impossible to carry out, at least with the prerequisite degree of detail and accuracy that is needed for creating sufficient trust in the model. It is especially in these cases where the modeler/simulationist has to be candid about the liberties he has taken and the relative merits of his/her results.

How will a code of conduct change common practices and how much time will be needed to achieve that?

A code of conduct, of course, will not automatically change common practice. Only if there is an understandable penalty on misbehaving (however small the penalty is, e.g., membership of professional society being rescinded after warning), will such a code have any influence. The other side of this is that as soon as the public starts up malpractice cases for professionals acting against the Code(s) of Conduct, the professionals will, in general, hurry to conform. It must be noted, however, that even in professions where this occurs already (medical doctors, accountants, etc.) there still are (groups of) individuals that do not toe the line. Human nature and "money is the root of all evil," I presume. In some cases, you need a federal (or global?) authority to force people to behave (case in point: the SEC in the US).

Topic 214
Visualization: mapping by Simon
Scientific visualization, sometimes referred to in shorthand as SciVis, is the representation of data graphically as a means of gaining understanding and insight into the data. It is sometimes referred to as visual data analysis. This allows the researcher to gain insight into the system that is studied in ways previously impossible.

3.5 Science and the Environment

It is important to differentiate between scientific visualization and presentation graphics. Presentation graphics is primarily concerned with the communication of information and results in ways that are easily understood. In scientific visualization, we seek to understand the data. However, often the two methods are intertwined.

From a computing perspective, SciVis is part of a greater field called visualization. This involves research in computer graphics, image processing, high performance computing, and other areas. The same tools that are used for SciVis may be applied to animation, or multimedia presentation, for example.

As a science, scientific visualization is the study concerned with the interactive display and analysis of data. Often one would like the ability to do real-time visualization of data from any source. Thus our purview is information, scientific, or engineering visualization and closely related problems such as computational steering or multivariate analysis. The approaches developed are general, and the goal is to make them applicable to datasets of any size whatever while still retaining high interactivity. As an emerging science, its strategy is to develop fundamental ideas leading to general tools for real applications. This pursuit is multidisciplinary (concerning morals of other users) in that it uses the same techniques across many areas of study.

Visualization, in the presentation sense, is not a new phenomenon. It has been used in maps, scientific drawings, and data plots for over a thousand years. Mapping is a form of visualization.

The main reasons for scientific visualization(mapping) are the following: it will compress a lot of data into one picture (data browsing), it can reveal correlations between different quantities both in space and time, it can furnish new space-like structures beside the ones which are already known from previous calculations, and it opens up the possibility to view the data selectively and interactively in `real time'.

By following the formation and the deformation as well as the motions of these structures in time, one will gain insight into the complicated dynamics. As was mentioned before, you also want to integrate our simulation codes into a visualization environment in order to analyze the data 'real time' and to by-pass the need to store every intermediate result for later analysis.

This is possible by means of processing in which the simulation is distributed over a set of high-performance computers and the actual visualization is done on a graphical distributive workstation. It is also very useful to have the possibility to interactively change the simulation parameters and immediately see the effect of this change through the new data. This process is called computational steering and it will increase the effective use of CPU time.

Classification of visualization techniques is often based on the dimension of the domain of the quantity that is visualized, i.e. the number of independent variables of the domain on which the quantity acts, and on the type of the quantity, i.e. scalar, vector, or tensor .

The rise of the "Information Age" and the ascendancy of Computer Graphics come together in the area of information visualization, where interactive graphical interfaces are used for revealing structure, extracting meaning, and navigating large and complex information worlds.

Increasing amounts of data and information and the availability of fast digital network access (e.g., in the information highway environment) have created a demand for querying, accessing, and retrieving information and data. However, information technology will not transform business, science, medicine, engineering, and education if the users cannot use it easily and efficiently. Technology must come to the users, taking their needs into account. If we do not involve the users, we will develop useless systems. One of the concerns of this field is the human-information interface, and how advances in interactive computer graphics hardware, mass storage, and data visualization could be used to visualize information.

The success of visualization not only depends on the results which it produces, but also depends on the environment in which it has to be done. This environment is determined by the available hardware, like graphical workstations, disk space, color printers, video editing hardware, and network bandwidth, and by the visualization software. For example, the graphical hardware imposes constraints on interactive speed of visualization and on the size of the data sets which can be handled.

Many different problems encountered with visualization software must be taken into account. The user interface, programming model, data input, data output, data manipulation facilities, and other related items are all important. The way in which these items are implemented determines the convenience and effectiveness of the use of the software package as seen by the scientist. Furthermore, whether software supports distributive processing and computational steering must be taken into account.

T215 - Data search, collection, processing and analysis by Joseph

When you have a question, whether it is for science or just any question in life, you need to do a series of steps to answer it. These include the scientific method of data search, collection, processing, and analysis. First you got to find data by searching for them, and then you collect some data to experiment or use on, process and record information, and finally analyze the answers you receive.

How can you search for data? This is relatively simple ever since the Internet came out. There are tons of information and ideas over the Internet, in which you can receive quick and easy information. You can also search for data yourself, by going outside and looking for stuff. By getting the data you would then go advance to the next step, collection.

Collecting data, like searching is very simple, it actually is basically the same. Like when you go outside to find the data, you would bring in something useful to your topic. On the Internet, you can find certain items you need and buy them from a shop, or look for it. You can even make it, which leads to the actual doing, data processing.

To data process is to use the data you collected and process it. There are many ways to do this. You can make an experiment like in science or so. But in computer terms, this is a series of operations performed on data to receive information. You can input the data into the computer, which would process it and give you information.

But the information is not answers yet; it is just a bundle of facts. To get them to be answers, you must analyze the information received from processing and work them out in a way that it can answer your question. For example you can receive the information one plus one is two when your question is how many apples would you have if you pick one from one farmer than one from another. The information one plus one is two tells you a broad answer that is not clear. By analyzing, you can realize that there is one apple from one farmer plus one apple from another farmer which makes it two apples.

or

Data search, collection, processing and analysis by Chaan

Recently I had a great experience of working at this hotel. The job had to do something me giving out brochures while my friends were compiling data about people. The data had information like why the people were there (for the fair), their age, what universities they attended etc. The people wrote all this info on application forms and since I wanted more money, I took the responsibility of entering all that data onto an excel spreadsheet. Now, one may wonder why I got paid so much for a job that doesn't look that challenging.

This is where the topic of the paper comes in. Data search and collection are two very important things in today's world. Data collection had already been done at the hotel by my friends and by the people who had filled up the forms. Now data search was for the people who were paying me to do this work. The fair wanted to know what kind of people, from what age group and other kinds of information like that. In order to find this, they made me enter all the data in an excel spreadsheet because it becomes easy to search for information through computers. They also thanked me for organizing all the data in the way I did because it made it easier for people to search through the document!

Data processing refers to the process of converting data into information. It's usually automated and ran on a computer. Data is only useful when it's presented clearly and organized in a clear way. So the act of entering data into excel was "data processing." I processed all the data in a nice way so that the people can easily view it and access it, which brings us to data analysis. It can also refer to the converting of data back into information. You may want to know why one would do that. One of the reasons could be that the way data was processed is not acceptable by the company and thus they would want to do it their own way or simply because it could be done in a better way.

According to wikipedia, data analysis refers to data analysis is the act of transforming data with the aim of extracting useful information and facilitating conclusions. Unlike data mining, data analysis is usually more narrowly intended as not aiming to the discovery of unforeseen patterns hidden in the data, but to the verification or disproval of an existing model, or to the extraction of parameters necessary to adapt a theoretical model to (experimental) reality.

Topic 216 - Forecasting by Ronald Chu

The basic definition of forecasting in technological terms is “forecasting the future characteristics of useful technological machines, procedures or techniques.” Forecasting is the process of estimation in unknown situations. It is similar to the word “prediction” but is more general and used in the discussion of time-series data. There are two main aspects of technological forecasting. Primarily, a technological forecast deals with the characteristics of technology, such as levels of technical performance, like speed of a military aircraft, the power in watts of a particular future engine, the accuracy or precision of a measuring instrument, the number of transistors in a chip in the year 2015, etc. Secondly, technological forecasting usually deals with only useful machines, procedures or techniques. This is to exclude commodities, services or techniques intended for luxury or amusement.

There are no real alternatives to technological forecasting. If a decision-maker has several alternatives open to him, he will choose the basis of which provides him the most desirable outcome. His only choice is whether the forecast is obtained by rational and explicit methods, or by intuitive means. There are four guidelines that must be followed when using rational methods. First of all, they can be taught and learned. Second of all, they can be described and explained. Thirdly, they provide a procedure possible to be followed by anyone who has gone through the necessary training. Finally, these methods are even guaranteed to produce the same forecast regardless of who uses them.

There are many methods of technological forecasting. They are the Delphi method, forecast by analogy, growth curves and extrapolation. Normative methods of technology forecasting — like the relevance trees, morphological models, and mission flow diagrams — are also commonly used. The Delphi technique is a method for obtaining forecasts from a panel of independent experts over two or more rounds. Here, experts are asked to predict quantities. In mathematics, extrapolation is the process of constructing new data points outside a discrete set of known data points. It is similar to the process of interpolation, which constructs new points between known points, but its results are often less meaningful, and are subject to greater uncertainty.

Although there are no “real” alternatives to technological forecasting, there are some exceptions. First of all, there can be no forecast at all. This means that we do not predict the future at all and just let things be. When something happens, it happens; we may or may not be prepared for it. Then, there is the “anything can happen” attitude. Similar to the first alternative, we do nothing to predict the future. This represents the attitude that the future is a complete gamble, that nothing can be done to influence it in a desired direction, and that there is no point therefore in attempting to anticipate it. Thirdly, there is the “glorious past” attitude. This is a bit different from the other two. This represents an attitude which looks to the past and ignores the future. Many organizations can point to significant achievements at some time or other in the past.

However, this attitude leads to the road of disaster. Fourthly, there is the window-blind forecasting alternative. This involves the attitude that technology moves on a fixed track, like an old-fashioned roller window blind and that the only direction is up. While this attitude does at least recognize that changes do take place and is therefore somewhat better than the other alternatives, it fails to recognize that there are other directions besides up. However, an organization that depends on window-blind forecasting will be taken by surprise, as some unexpected technological change will hit them hard. Finally, there is the genius forecasting alternative. This method consists in finding a genius, and asking him for his intuitive forecast. Many of these genius forecasts made in the past have been very successful. Unfortunately, there have also been many so wide of the mark as to be useless. In conclusion, it should be clear that where rational and explicit methods are available, they are much to be preferred.

T217 - Data logging by Dwarkesh

Data logging is the practice of recording sequential data, often chronologically. In computerized data logging, a computer program may automatically record events in a certain scope in order to provide an audit trail that can be used to diagnose problems. Examples of physical systems which have logging subsystems include process control systems, and the black box recorders installed in aircraft. Many operating systems and multitudinous computer programs include some form of logging subsystem. Some operating systems provide a syslog service (described in RFC 3164), which allows the filtering and recording of log messages to be performed by a separate dedicated subsystem, rather than placing the onus on each application to provide its own ad hoc logging system. In many cases, the logs are esoteric and hard to understand; they need to be subjected to log analysis in order to make sense of them. Other servers use a splunk to parse log files in order to facilitate troubleshooting; this approach may yield correlations between seemingly-unrelated events on different servers. Other enterprise class solutions such as those from Log Logic collect log data files in volume and make them available for reporting and real-time analysis.

The performance characteristics of a car can be monitored through its diagnostic port. There are several ways to do this; one way is by using a handheld device with software already installed and the appropriate cable to hook into the diagnostic port. Another method is to use a PDA loaded with specialized software and employ a special cable that plugs into the diagnostic port. Finally, a PC can be used, laptops generally work best, using specialized software and connecting it to the car with a specially-made cable.

Data logging a car allows one to observe how well the car is running. A data logger can be used to help tune the car, especially if modifications have been made to the vehicle. The reason it helps is that it logs various parameters in the engine such as ignition timing, injector pulse, RPM, knock/detonation, barometer pressure, temperature, oxygen sensors, and various other sensor data. Using the information that is provided (and either some sort of piggy-back or standalone system), a user can change the injector pulse width so that the engine is fed more or less fuel, increase or decrease ignition timing, and other changes. Such changes are made to make the car run more reliably or achieve higher performance.

When the computer is used for measurement, the tasks it performs can be summarized as follows:

Collect data from external sensors.
Store data in computer memory.
Display data in graphical form.
Process data; i.e. perform calculations.
The advantages of using a computer to collect data are:
Graphical display
Analyzing tools
Calculating aids
Continuous recording
Simultaneous recording
Period logging

Remote Logging. Data may be collected independently from the microcomputer using a data-logger. - Data-logger may be sited in locations which are not convenient for computers - Computer is made free for other activity - This is well-suited to longer term experiments. - The data may be viewed retrospectively e.g. Field work: climatic factors over few hours Laboratory experiments: germination of peas etc.