GENERAL COMMENTS AFTER READING THE THREE EDITIONS:

What strikes me about this book after exploring it in recent weeks? First, I had forgotten how much emphasis Simon puts on *design*. This intrigues me, because I would argue that design is not a central part of classical cognitive science, but it certainly is important to embodied cognitive science. Second, even when discussing an approach that is designed to fit with the classical notion of a physical symbol system, Simon repeatedly points out the importance of the environment (and the general simplicity of internal processes). This is not surprising, given that versions of the parable of the ant appear throughout the three editions, but again this is an emphasis that has waned in modern variants of the classical approach. Third, at points in the later editions Simon blurs the division between the inner and the outer by noting that memory can serve as the environment for internal processes, but the division between the world and the environment may be arbitrary. He also talks about the need to act on the world; at points he seems to be predicting the leaky mind, although I suspect that he would not accept this radical notion. Fourth, the classical sandwich, and the disembodiment of mind, arise in explicit quotations, particularly in the latter two editions of the book.

THE SCIENCES OF THE ARTIFICIAL, FIRST EDITION, 1969

Below are margin notes, and thoughts, on this book. These notes were written by Prof. Michael R.W. Dawson in March, 2010. At the end of these notes, new notes appear that examine changes in this book of Simons in the second and third editions.

PREFACE

"Certain phenomena are ‘artificial’ in a very specific sense: They are as they are only because of a system's being molded, by goals or purposes, to the environment in which it lives." (Page ix).

The theme of the book is how to categorize artificial phenomena. Do artificial phenomena fall within the domain of science?

"The genuine problem is to show how empirical propositions can be made at all about systems that, given different circumstances, might be quite other than they are." (Page x).

Artificial systems can change because they are contingent upon their environment. This is a problem of artificiality that is wide spread.

"Engineering, medicine, business, architecture, and painting are concerned not with the necessary but with the contingent -- not with how things are but with how they might be -- in short, with design." (Page xi).

The science of the artificial is the science of design. Simon will illustrate this with two examples, the psychology of cognition and engineering design.

"Artificiality is interesting principally when it concerns complex systems that live in complex environments." (Page xi)

CHAPTER 1: UNDERSTANDING THE NATURAL AND THE ARTIFICIAL WORLDS

Natural science is about a class of things. Natural science tries to show that the wonderful is not incomprehensible. However our world is more artificial than natural. For instance a significant part of our environment consists of symbols. Biological and natural are not synonyms. Artifacts are not apart from nature, but are shaped by man's aims. Can there be a science of artificial objects and phenomena? First task is to get rid of the pejorative notion of the artificial.

"Our language seems to reflect man's deep distrust of his own products." (Page 4)

Simon draws a distinction between the artificial and the synthetic, but notes that both synthesis and artifice belong to the realm of engineering.

"The engineer is concerned with how things *ought* to be -- ought to be, that is, in order to *attain goals*, and to *function*. Hence, a science of the artificial will be closely akin to a science of engineering." (Page 5)

Simon notes for differences between the artificial and the natural: artificial things are synthesized; artificial things imitate appearances but lack the reality; artificial things are characterized in terms of functions, goals, adaptation; artificial things are often discussed in terms of imperatives. If an artificial object is to fulfill a purpose, then we must look at relations between three things: goals, the nature of the artifact, and the environment. The latter two fall in the domain of natural science.

An artifact is an interface between inner and outer environments; this notion of interface applies nicely to many natural systems too.

"If the inner environment is appropriate to the outer environment, or vice versa, the artifact will serve its intended purpose." (Page 7)

"Given an airplane, or given a bird, we can analyze them by the methods of natural science without any particular attention to purpose or adaptation, without reference to the interface between what I have called the inner and outer environments." (Page 7)

The inner/outer distinction is convenient, even if it is not necessary for Simon, functional means purposeful. Explaining by reference to purpose demands an understanding of the outer environment.

"Thus the first advantage of dividing outer from inner environment in studying an adaptive or artificial system is that we can often predict behavior from knowledge of the system's goals and its outer environment, with only minimal assumptions about the inner environment." (Page 8)

Designers insulate the inner system from the environment to attain an invariant relationship between the inner environment and the goal. Note: this seems to be the opposite of embodiment.

"We might hope to be able to characterize the main properties of the system and its behavior without elaborating the detail of *either* the outer or inner environments." (Page 9)

Can we depend on the simplicity of the interface? "Description of an artifice in terms of its organization and function -- it's interface between inner and outer environments -- is a major objective of the invention and design activity." (Page 10) Key aspect of describing an artifact is using the goals that link the inner environment to the outer environment.

"If the inner system is properly designed, it will be adapted to the outer environment, so that its behavior will be determined in large part by the behavior of the latter, exactly as in the case of 'economic man'." (Page 11)

Behavior takes on the shape of the task environment. Designing a perfect inner system is not easy. In imperfect design, the properties of the inner system will show through. Note: this theme of the limiting properties of the inner system showing through when the design is not perfect is central to this book. For instance, Simon often describes cognitive psychology as though its purpose is to discover such limiting properties.

The artificial imitates the real via the interface. Simulation predates the existence of the digital computer. Simon says that there are two ways that a simulation can tell us things that we did not already know. First, simulation derives difficult implications from premises. Note: the homeostat is an example that I use to illustrate the same. Simulation also helps when we don't know much about a system of natural laws. Abstraction is an aid to simulation in this case; here Simon is thinking about simulation in a top-down way.

Artificial systems are particularly susceptible to simulation via abstract or simplified models. Here are Simon adopts a more traditional notion of functionalism:

"Resemblance in behavior of systems without identity of the inner systems is particularly feasible if the aspects in which we are interested or rise out of the *organization* of the parts, independently of all but a few properties of the individual components." (Page 17)

The digital computer is so useful because it only makes organizational properties available: the hardware is invisible. Any computer can be assembled out of a small set of basic elements; a micro theory of internal structure is not required. Note: here Simon is appealing to a traditional notion of multiple realisation.

"The important point, for our present discussion, is that the parts could as well be neurons as relays, as well relays as transistors." (Page 19)

Simon proceeds to list the basic nature of computers, including central processing unit, memory, program, serial processing, and basic operations. An empirical theory of computers is possible. Simon discusses the development of time-sharing computers, and in so doing describes a synthetic approach -- the empirical use of simulation. He ends the chapter by foreshadowing the parable of the ant as applied to computers. That is, the complex behavior of computers is due to their environment and not due to their complex internal nature. He then foreshadows a synthetic approach in terms of using computers to study thought:

"For if it is the organization of components, and not their physical properties, that largely determines behavior, and if computers are organized somewhat in the image of man, then the computer becomes an obvious device for exploring the consequences of alternative organizational assumptions for human behavior. Psychology can move forward without awaiting the solutions by neurology of the problems of component design -- however interesting and significant these components turn out to be." (Page 22)

CHAPTER 2: THE PSYCHOLOGY OF THINKING

Simon begins this chapter with the parable of the ant, emphasizing the complexity of the ant's path.

"Viewed as a geometric figure, the ant's path is irregular, complex, hard to describe. But laxity is really a complexity in the surface of the beach, not a complexity in the ant." (Page 24)

Simon immediately sites the robot tortoise and goal-seeking automata, and then repeats his claim about the ant. The ant's simplicity is at the level of organization which is an abstract level. The parable of the ant might also apply to man. Why? Because man is an adaptive system, and proper adaptation reflects the outer environment and hides the inner environment as was noted in Chapter 1.

"*A man, viewed as a behaving system, is quite simple. The apparent complexity of his behavior over time is largely a reflection of the complexity of the environment in which he finds himself.*" (Page 25)

Most of man's thinking and problem-solving is artificial. Problem solving can be described as search. Man is too slow to do exhaustive systematic search. So how are problems solved? One approach is to apply serial systematicity: look for inconsistencies to reduce the search space. One can also search for contradictions. In short, you reduce the search space by combining search with reason. How can we go further to remove trial-and-error search? The answer: look for more promising avenues to explore. If we discovered this approach, by studying human problem-solving, then we would be learning about strategy, but not learning about biology.

"The more sophisticated, in a certain sense, the strategy became, the less search was required. But it is important to notice that, once the strategy was selected, the course of the search depended only on the structure of the problem, not on any characteristics of the problem solver." (Pages 30-31)

A variety of cognitive tasks points to only a small set of limits on internal mechanisms. For instance, consider a standard concept attainment paradigm. If you know how many binary features are used to characterize a concept then you can estimate the time required to attain the concept. The time should very with the logarithm of the number of stimulus dimensions. But people may not be able to discover the most efficient strategy because they face memory limitations. Indeed two key internal constraints are the capacity of short-term memory and the time required to move a chunk from short-term memory into long-term memory. Indeed, strategies may be many, and may be learned, and will outnumber the internal limits.

"Our knowledge of behavior must be regarded as sociological in nature rather than psychological -- that is, as revealing what human beings in fact learn when they grow up in a particular social environment." (Page 35)

Experimental psychology needs to estimate parameters of a small number of internal limits. But psychology has been more interested in hypothesis testing Banyan parameter estimation. For instance they report trials to learn instead of time to learn. The few studies that have looked at that time required to learn a chunk of information have revealed a constant in the order of 10 seconds. The rate of learning can be affected by meaningfulness and similarity which has been modeled in simulations like EPAM. A second limiting property is the capacity of short-term memory which has been identified as being 7+ or -2 chunks by Miller. Interfering tasks appear to drop short-term memory capacity to two chunks. Simon goes on to say that there are many other limiting properties to be found.

"What it does imply is that we should not look for great complexity in the laws governing human behavior, in situations where the behavior is truly simple and only its environment is complex." (Page 41)

The inner system organizes stimulate into chunks. Simon now talks about what is made explicit by different representational formats, comparing images to list structures. Organize chunks implies storing information about relations. Again, this is taken as evidence in favor of adopting a functional approach according to the traditional notion of functionalism. Simon moves on to discussing natural language and faces a conflict between the adaptive nature of thought and nativism. He argues that nativism applies to the general internal characteristics or limitations of the inner world that he has been discussing it earlier in the chapter. He believes that the kind of limitations that he has been discussing are exactly the sorts of limitations that you would expect to find in a system designed to deal with language. So, language is both natural and artificial. Only the thinkable is expressible.

Simon concludes the chapter by returning to the parable of the ant. He points out that human behavior is the result of a fairly small and simple set of constraints that apply to the inner mechanisms. The mechanisms only reveal themselves to the extent that adaptation is not perfect which leads Simon to discuss the disembodiment of mind.

"The main reason for this disembodiment of mind is, of course, the thesis that I have just been discussing. The difference between the hardware of a computer and the 'hardware' of the brain has not prevented computers from simulating a wide spectrum of kinds of human thinking -- just because both computer and brain, when engaged in flop, our adaptive systems, seeking to mold themselves to the shape of the task environment." (Page 54)

CHAPTER 3: THE SCIENCE OF DESIGN

One task of engineering is to teach about artificial things. A central concern is the process of design. But natural science has driven the science of the artificial from professional schools in an attempt to seek academic respectability.

"A science of artificial phenomena is always in imminent danger of dissolving and vanishing. The peculiar properties of the artifact lie on the scene interface between the natural laws within it and the natural laws without." (Page 57)

"The artificial world is centered precisely on this interface between the inner and outer environment; it is concerned with attaining goals by adapting the former to the latter. The proper study of those who are concerned with the artificial is the way in which that adaptation of means to environments is brought about -- and central to that is the process of design itself." (Page 58)

But, a science of design is emerging. Natural science is concerned with how things are while design is concerned with how things ought to be. One can adapt ordinary declarative logic to describe the requirements of design. In particular Simon discusses the logic of optimize a nation methods, such as constrained optimization. The logic of optimization makes use of possible worlds; it moves from would to should. The issue is to develop efficient and practical methods for constrained optimization. This leads Simon to introduce satisficing, illustrating it with the traveling salesman problem.

"In the real world we usually do not have a choice between satisfactory or more solutions, for we only rarely have a method of finding the optimum." (Page 64)

"We satisfice by looking for alternatives in such a way that we can generally find an acceptable one after only moderate search." (Page 65)

Satisficing might be particularly appropriate for design. The problem is that design alternatives must be synthesized if one views design is searching through a problem space. The key question raised: how does one search for candidate solutions when alternatives must be synthesized? Simon argues in favor of means-end analysis. But in so doing, Simon makes explicit the classical sandwich:

"The condition of any goal-seeking system is that it is connected to the outside environment through two kinds of channels: the afferent, or sensory, channels, through which it receives information about the environment; and the efferent, or her motor, channels, through which it acts on the environment." (Page 66)

"Goal-directed action depends on building this kind of bridge between the afferent and the efferent worlds." (Page 67)

Simon illustrates this with his discussion of GPS, which works by representing the desired situation, differences between the current and the desired situation, and actions that change situations. The basic problem for GPS is what to do next. It is the choice of action under uncertainty. The reasoning that GPS performs is not valid in the sense of standard logic, because there is no certainty of accomplishing goals.

"Problem-solving systems and design procedures in the real world do not merely *assemble* problem solutions from components but must *search* for appropriate assemblies." (Page 69)

Satisfactory design must conserve scarce resources and also the design process must manage the resources of the designer. The cost of design must be considered in a cost-benefit analysis. This can be used to define a search through a problem space in terms of heuristic values of cost or gain.

"Hence it is more useful to think of the values as estimates of the game to be expected from further research along the path then to think of them as 'values' in any more direct sense." (Page 71)

All kinds of information can be used to assign a path's value. This leads to a new view of search processes:

"Processes for gathering information about problem structure that will ultimately be valuable in discovering a problem solution." (Page 72)

Simon now talks about hierarchy and design in a way that reminds me of Cummins and iterative functional analysis. He talks about decomposing a complex structure into a set of semi-independent components that correspond to functional parts. Simons also notes that multiple decompositions are possible. This leads to a notion design as testing that can be either top-down or bottom-up:

"Think of the design process as involving first the generation of alternatives and then the testing of these alternatives against a whole array of requirements and constraints." (Page 74)

The direction of design -- top-down or bottom-up --can affect the nature of the final product. Note: this strikes me as interesting in terms of comparing analytic versus synthetic theories. Problem representation can also affect design. Simon illustrates this with number Scrabble by showing how you can make it easy by representing it in a way that equates it to tick tack toe. Solving a problem means representing it in a way that the solution is transparent.

"All mathematical derivation can be viewed simply as change in representation, making evident what was previously true but obscure." (Page 77)

We don't yet have a taxonomy of representations. Simon and this the chapter by listing seven characteristics of a curriculum for teaching design where these characteristics are all topics that have been raised in the chapter. He points out that design and psychology are pretty much equivalent. He describes music as a science of the artificial, talking about composition and design and also noting that design is a unifying language.

"The ability to communicate across fields -- the common ground -- comes from the fact that all who use computers in complex ways are using computers to design, or to participate in the process of design." (Pages 82-83)

"If I have made my case, then we can conclude that, in large part, the proper study of mankind is the science of design, not only as the professional component of a technical education but as a core discipline for every liberally educated man." (Page 83)

CHAPTER 4: THE ARCHITECTURE OF COMPLEXITY

Can some principles be applied to systems, that is complex systems, in general? Cybernetics is one attempt to answer this question in the affirmative. Simon explores the questions by focusing on the complex systems of the behavioral sciences. He goes on to consider four main themes: hierarchy, evolution, dynamics and decomposability, and descriptions. As shown by the quote below, for Simon a complex system is a nonlinear system.

"Roughly, by a complex system I mean one made up of a large number of parts that interact in a non-simple way. In such systems, the whole is more than the sum of the parts." (Page 86)

Simon starts with a very general definition of a hierarchical system, noting that in a hierarchy and elementary subsystem might actually be complex. His notion of hierarchy is less restrictive than typical. He notes that social biological and physical systems are clearly hierarchical. He talks about the span of a hierarchy is being a term that might affect when hierarchical the composition is applied. Social hierarchies are not spatial. Books and music are hierarchical according to Simon's definition.

Simon moves on to evolution, because he has a story to tell about how evolution and hierarchy can be interrelated:

"The time required for the evolution of a complex form from simple elements depends critically on the numbers and distribution of potential intermediate stable forms." (Page 93)

Simon returns to problem solving as search, considering trial and error from the perspective of selection. This makes contact again with the notion of heuristics guiding promising avenues of search.

"In problem solving, a partial result that represents recognizable progress toward the goal plays the role of a stable subassembly." (Page 96)

"When we examine the sources from which the problem-solving system, or the evolving system, as the case may be, derives its selectivity, we discover that selectivity can always be equated with some kind of feedback of information from the environment." (Page 97)

"Complex systems will evolve from simple systems much more rapidly if there are stable intermediate forms than if there are not. The resulting complex forms in the former case will be hierarchic." (Pages 98-99)

By looking at the interactions between subsystems in comparison with the interactions within subsystems, Simon now goes on to discuss nearly decomposable systems. In such a system, the interactions between subsystems exist, but are weak. In the short term, the behavior of one subsystem is independent of the short-term behavior of the other subsystems. In the long term, the behavior of a component does depend on the behavior of the other components, but only in an aggregate way. Nearly decomposable systems are common in the natural world. They are also common in social systems. Of course, these sorts of interactions provide the link between nearly decomposable systems and hierarchy.

"Hierarchies have the property of near decomposability. Intracomponent linkages are generally stronger than intercomponent linkages." (Page 106)

Towards the end of the chapter Simon links the notion of nearly decomposable systems to functional analysis in the traditional sense. That is, if a system is nearly decomposable, then to understand it functional analysis makes sense.

Simon now talks about state descriptions and process descriptions, where he equates state with sense and process with action.

"The distinction between the world as sense in the world is acted upon it finds the basic condition for the survival of adaptive organisms." (Page 111)

The quote above, in the quote below, again place Simon directly into proposing the classical sandwich.

"Thus, problem solving requires continual translation between the state and process descriptions of the same complex reality." (Page 112)

"The correlation between state description and process description is basic to the functioning of any adaptive organism, to its capacity for acting purposefully upon its environment." (Page 117)

THE SCIENCES OF THE ARTIFICIAL, SECOND EDITION, 1981

Below are margin notes, and thoughts, on this book. These notes were written by Prof. Michael R.W. Dawson in March, 2010. These notes were written after reading this edition immediately after reading the first edition, which is notated above. Because there is a substantial intersection between the two editions, the notes below focus on the new material. The first edition notes above apply perfectly well to the second edition, with the exception that the page numbers for the quotes require changing.

CHAPTER 1: UNDERSTANDING THE NATURAL AND THE ARTIFICIAL WORLDS

This chapter is essentially identical to Chapter 1 of the first edition, with the exception that Simon adds a new main section to the end “Symbol Systems: Rational Artifacts”. Here are my notes on this new material, which begins on page 26 of the second edition.

The computer is a physical symbol system, as is the mind and the brain.

"Symbol systems are almost the quintessential artifacts, for adaptivity to an environment is their whole raison d'être. They are goal-seeking, information-processing systems, usually enlisted in the service of the larger systems in which they are incorporated." (Page 27)

A physical symbol system holds symbols in memory and uses processes to manipulate the symbols. This conception leads Simon again to state the classical sandwich:

"Symbol structures can, and commonly do, serve as internal representations (e.g., 'mental images') of the environments to which the symbol system is seeking to adapt. They allow it to model that environment with greater or less veridicality and in greater or less detail, and consequently to reason about it. Of course, for this capability to be of any use to the symbol system, it must have windows on the world and hands, too. It must have means for acquiring information from the external environment that can be encoded into internal symbols, as well as a means for producing symbols that initiate action upon the environment." (Page 27)

Simon stresses that physical symbol systems are indeed physical, that is they are real-world devices. In this sense he is making the point that they are more embodied than mathematics and logic.

"The hypothesis is that a physical symbol system of the sword I have just described as the necessary and sufficient means for general intelligent action." (Page 28)

CHAPTER 2: ECONOMIC RATIONALITY

Simon's new chapter 2 on economic rationality is an addition to did not appear in the first edition.

Rationality has the task of allocating resources. Simon describes the outer and inner worlds of economics. A company's goals define the inner environment, while cost and revenue curves define the outer environment. Simon distinguishes substantive rationality from procedural rationality; it is the former that applies to an artificial system. In principle, normative practices could be used for instance to determine the price at which a specific unit could be sold or to determine how many of the units should be constructed. In the real world though this is difficult to do. Real world economics is problematic, which leads Simon to discuss procedural rationality. This is because one of the inner constraints on the inner environment is uncertainty about the outer environment as far as real world economics is concerned. In short, real-world economic decision-making requires satisficing and heuristic search.

"Economic man is in fact a satisficer, a person who accepts 'good enough' alternatives, not because he prefers less to more but because he has no choice." (Page 36)

The market mechanism is a key process of social coordination, defining a set of interactive processes that lead to equilibrium. Central control is absent. However, when we think about systems that can be in equilibrium we don't want to see design without also seeing a designer. Note: this strikes me as an intuitive account by Simon in an attempt to explain the law of uphill and analysis and downhill synthesis.

"But somehow our intuitions about self-regulation without central direction do not carry over to the artificial systems of human society." (Page 40)

Estimating probability is computationally expensive, making worldly uncertainty difficult to deal with. Note: does this apply to perceptrons that match probabilities? Feedback is described as a simplifying mechanism for estimating probabilities. Feedback is described as correcting errors of the past, where feedforward is the more problematic formation of expectations about the future. The difficulty of feedforward processing is illustrated in classic problems like the prisoner's dilemma. These problems arise because goals, particularly goals in multiagent systems, may conflict partially or totally. One consequence is that equilibrium is not going to be achieved.

Simon considers the possibility of creating an artifact by letting it evolve in response to some kind of selective force. He notes that evolving fitness is not the same as seeking a global optimum.

CHAPTER 3: THE PSYCHOLOGY OF THINKING

This chapter is essentially the same as chapter 2 in the first edition. This is because he adds a complete new chapter on psychology that immediately follows this chapter in the second edition.

CHAPTER 4: REMEMBERING AND LEARNING

This new chapter is Simon's attempt to talk about what's happened in cognitive psychology in the decade between editions of the book.

In the 1950s and the 1960s the study of toy problems led to the hypothesis that human thought processes are actually quite simple. Does human thinking still looks simple when more semantically rich domains are studied? Simon is going to argue that more memory does not necessarily mean more complexity. He's going to use the parable of the end to do this, by pointing out that more memory can provide a more complex environment for the simple processes which in turn will lead to more complicated behavior. Interestingly, when he talks about memory is though there is an arbitrary boundary between the inner and the outer which seems to predict the leaky mind as hypothesized by modern embodied cognitive science. A theory of solving semantically rich problems requires a theory of memory. Intuitive leaps are acts of recognition. Simon illustrates this with his example of recognition and chess skill.

Very complex domains require external memories; that is, books and other external reference aids. Memory could be mainly a body of data, but Simon also talks about memory *for* skills. He begins to talk about productions in a production system not as data, but instead as skills. This leads him to claim that:

"The boundary between knowledge and skill is subtle." (Page 110).

Why do we need to represent? Simon talks about representing whenever it is costly to gain access to information in the real world. Note: this is an interesting issue to explore when comparing and contrasting classical cognitive science and embodied cognitive science. As the size of memory grows we might say that the system is more complex because it has a larger memory, or we might also say that it remains simple since its fundamental structure does not change.

"Human memory is best regarded as an extension (sometimes a large extension) of the environment in which human thought processes take place and not as an increment in the complexity of these processes." (Page 117)

"The external environments of thought, both the real world and long-term memory, undergo continual change. In memory the change is adaptive." (Page 118)

Learning is an invariant of the human cognitive system that deals with environmental change. Again Simon talks about data versus skill acquisition. He believes that learning of most kinds can be explained using the physical symbol system ideas that are at the core of the book. The information processing approach is currently vague with respect to how understanding aids learning.

Simon provides a general definition of a production system, and then makes a distinction that again makes me think of the leaky mind:

"A system whose behavior is governed by a perceptual productions is sometimes called *stimulus driven* or *data driven*; one governed by goals symbols in STM, goal driven." (Page 121)

Production systems can model learning by example.

"What constitutes novelty depends on what knowledge is already in the mind of the problem solver and what help is received from the environment in adding to this knowledge." (Page 123)

Simon talks about a couple of expert systems (AM and BACON) that can be viewed as solving problems without a goal because the way they work is to choose a direction of exploration that seems interesting. For this reason, Simon does not believe that discovery processes introduce new kinds of complexity into human cognition. He ends this chapter with a quote that revisits the parable of the ant, and again makes me think of the wiki mind:

"The inner environment, the hardware, is simple. Complexity emerges from the richness of the outer environment, both the world apprehended through the senses and the information about the world stored in long-term memory." (Page 127)

CHAPTER 5: THE SCIENCE OF DESIGN

This chapter is identical to chapter 3 of the first edition as far as I can tell, and so I have no new notes to add.

CHAPTER 6: SOCIAL PLANNING

This is a new chapter to does not appear in the first edition.

Simon notes that ambitious planners or designers are often interested in whole societies and their environments. He notes the blessings and curses of technology. Getting to the moon was a simple task because it was merely technological. Other achievements -- Simon's example is the Constitution -- were more complex. It succeeded because it was designed with limited objectives in mind.

"The success of planning on such a scale may call for modesty and restraint in setting the design objectives and drastic simplification of the real-world situation in representing it for purposes of the design process." (Page 163)

"The quality of design is likely to depend heavily on the quality of the data available." (Page 169)

Forecasting is fundamental to design, but data about the future is the weakest point in our armor of fact.

"The heart of the data problem for design is not forecasting but constructing alternative scenarios for the future and analyzing their sensitivity to errors in the theory and data." (Page 171)

Simon discusses homeostasis and feedback as stabilizers. Homeostasis is viewed as insulating the system from the environment, while feedback adapts the environment to outer variation.

"We are unable to think coherently about the remote future, and particularly about the distant consequences of our actions. Our myopia is not adaptive, but symptomatic of the limits of our adaptability. It is one of the constraints on adaptation belonging to the inner environment." (Page 180)

Simon returns to the notion of planning without goals, reminds the reader of expert systems for discovery that do not use goals, and suggests that this approach might be the most appropriate for the social design process. Note: here he makes me think a bit about the synthetic paradigm.

"The idea of final goals is inconsistent with our limited ability to foretell or determine the future. The real result of our actions is to establish initial conditions for the next succeeding stage of action." (Page 187)

Simon ends the chapter by proposing a curriculum for social design which is analogous to his approach in the original chapter on design which also appears earlier in the second edition.

CHAPTER 7: THE ARCHITECTURE OF COMPLEXITY

This chapter appears to be identical to the last chapter of the same name in the first edition.

THE SCIENCES OF THE ARTIFICIAL, THIRD EDITION, 1996

Below are margin notes, and thoughts, on this book. These notes were written by Prof. Michael R.W. Dawson in March, 2010. These notes were written after reading this edition immediately after reading the first and second editions, which are notated above. Because there is a substantial intersection between the three editions, the notes below focus on the new material.

CHAPTER1: UNDERSTANDING THE NATURAL AND THE ARTIFICIAL WORLDS

This chapter appears to be identical to the one in the previous edition.

CHAPTER 2: ECONOMIC RATIONALITY: ADAPTIVE ARTIFICE

Simon notes the in the preface to this edition that he has changed emphasis to reflect changes in his thinking about the roles of organizations and markets. He adds a subsection on “operations research and management science” beginning on page 27. There are a few paragraphs that have been rewritten, but none that affect my previous notes that are based on earlier editions of the book.

CHAPTER 3: THE PSYCHOLOGY OF THINKING: EMBEDDING ARTIFICE IN NATURE

In the preface, Simon claims that some material in this chapter has been updated. Early on, the only noticeable difference is in parameter estimations: he moves from 5 seconds per chunk of memorizing to 8 seconds. Just after his section on visual memory, he adds a few short paragraphs on “The Mind’s Eye”, indicating his awareness of the imagery debate (but not much more). I don’t see many other differences.

CHAPTER 4: REMEMBERING AND LEARNING: MEMORY AS ENVIRONMENT FOR THOUGHT

In the preface, Simon claims that some material in this chapter has been updated. The only notable difference that I see is that he has added a short subsection at the end of the chapter called “Finding New Problem Representations”. He notes that this is an important component of problem solving, but also notes that we don’t know much about how it happens!

CHAPTER 5: THE SCIENCE OF DESIGN: CREATING THE ARTIFICIAL

In the preface, Simon claims that some material in this chapter has been updated. There are a few paragraphs that have been rewritten, but none that affect my previous notes that are based on earlier editions of the book.

CHAPTER 6: SOCIAL PLANNING: DESIGNING THE EVOLVING ARTIFACT

In the preface, Simon claims that some material in this chapter has been updated. Midway through he adds a brief subsection on “Organizations in Social Design”. His main point is that organizational design needs to be added to the curriculum, because organizations can extend the abilities of individuals:

“From birth until death, our ability to reach our goals, even to survive, is tightly linked to our social interactions with others in our society.” (Page 154)

CHAPTER 7: ALTERNATIVE VIEWS OF COMPLEXITY

One goal of the chapter is to address complexity more generally. The 20th century has shown recurrent bursts of interest in complexity. Simon is interested in general in holism, and notes that there can be stronger or weaker interpretations of this concept. The strong notion of holism appeals to emergence, but at the same time rejects mechanistic explanations of emergence.

"In a weaker interpretation, emergent simply means that the parts of a complex system have mutual relations that do not exist for the parts in isolation." (Page 170)

The weak form of holism is compatible with reductionism. Weak emergence might be evident by appealing to the aggregate properties of a system. The short-run behavior of the component can be described independently of the behavior of the other components, because interactions between components make themselves evident only in the longer run. Simon believes in weaker notions of holism and emergence that our consistent with reductionism.

"In this pragmatic way, we can build nearly independent series for each successive level of complexity, but at the same time, build bridging series that show how each higher level can be accounted for in terms of the elements and relations of the next level below." (Page 172)

Simon begins with a discussion of cybernetics, and notes how feedback explains goal-directed adaptation and thus can avoid teleology.

Because of cybernetics "holism was brought into confrontation with reductionism in a way that had never been possible before, and that confrontation continues today in philosophical discussion of artificial systems." (Page 173)

The ideas that emerge from cybernetics have proven useful in a wide range of applications. General systems theory has succeeded to the extent that it has looked for common properties shared among broad classes of complex systems. Modern interest in complexity is revealed in a number of new fields including those that use the labels catastrophe, chaos, genetic algorithms, and cellular automata. Simon briefly discusses each of these topics in turn. He notes that catastrophe theory is concerned with systems that exhibit stable behavior followed by sudden shifts to disequilibrium. Simon does not believe that there are many practical situations to which this theory applies. He moves on to chaos theory, explaining it briefly and noting that because it is concerned with systems of nonlinear equations it has advanced via the synthetic approach; that is, through computers doing numerical simulation. Understanding of chaotic systems does not lead to the ability to predict them, though, as illustrated with the notion of the strange attractor. In the end, Simon doesn't seem to think that chaos theory as much of an impact on the topics that he has written about in the three editions of this book. He then turns to evolution and complexity, in particular genetic algorithms and cellular automata. He doesn't have much to say about either.

"Complexity is more and more acknowledged to be a key characteristic of the world we live in and of the systems that cohabit our world. It is not new for science to attempt to understand complex systems: astronomers have been added for millennia, and biologists, economists, psychologists, and others joined them some generations ago. What is new about the present activity is not the study of particular complex systems but the study of the phenomenon of complexity in its own right." (Page 181)

Note: it is interesting to me that one topic that doesn't get any treatment at all in this chapter is artificial neural networks.

CHAPTER 7: THE ARCHITECTURE OF COMPLEXITY: HIERARCHIC SYSTEMS

This chapter seems identical to previous editions.