Quantitative Discovery and Qualitative Reasoning about Failure Mechanisms in Pavement.

by
Nayel S. el-Shafei
B.S., Cairo University
(1981)

Submitted to the Department of Civil Engineering
in partial fulfillment of the
requirements for the degree of
Master of Science
at the
Massachusetts Institute of Technology
June 1986

Copyright © 1986 Nayel S. el-Shafei

The author hereby grants to M.I.T. permission to reproduce and to distribute publicly copies of this thesis document in whole or in part.

Signature of Author ______________________________
Department of Civil Engineering
June 1, 1986

Certified by ______________________________
Professor Fred Moavenzadeh, Thesis Supervisor

Accepted by ______________________________
Chairman, Department Committee

Abstract

Highway maintenance is a major national problem aggravated by the fact that our knowledge about the mechanical behavior of pavement is incomplete and/or unreliable. The complexity of pavement mechanics may stem from the heterogeneity and non-uniformity of its ingredients. Several theories have been proposed for explaining the failure mechanisms in pavement. None of them can stand alone to explain a considerable part of this behavior. The problem we are tackling in this thesis is understanding the nature of knowledge in the domain of failure mechanisms in pavement and also in the other evolving domains in engineering and science. By understanding the nature of knowledge we mean recognizing the significant parameters, how they interact, and finally finding the relationships that describe such interactions. Such understanding enables us to model the actual mechanical behavior of pavement. The approach we have developed involves hybridizing the symbolic and numeric techniques of quantitative learning from observations, explanation-based generalization, qualitative reasoning, and adaptive control. The hybridization of these techniques yields a generic tool, the *HOTEP* system for quantitative discovery and qualitative reasoning, under development in this research.

Thesis Supervisor: Fred Moavenzadeh
Title: Professor of Civil Engineering.

بسم الله الرحمن الرحيم

To my Parents

Acknowledgement

I dedicate this thesis to my parents who have given to me their love, support, advice, and encouragement. The effort put into this thesis is but a small symbol of my gratitude and appreciation for them.

I sincerely thank Prof. Fred Moavenzadeh, my thesis advisor, for putting my feet onto the right track of academic research, and for his patient urging, pushing me towards the best I can do. Prof. Moavenzadeh's continuous support, encouragement, open-mindedness towards unorthodox research methodologies are highly appreciated.

Exposure to the insights of Prof. Marvin Minsky has influenced me greatly, and will continue to do so. I'm very much indebted to Prof. Ryszard Michalski, always eager to enter into intellectual debate which greatly helped in shaping my ideas.

I would like also to thank Prof.Jerome Connor for his insightful visions, and Dr. Ken Maser for several constructive conversations.

A lot of people helped in making my research pursuit, and my stay at MIT, much easier. They are too numerous to be counted, but my office mate Rick Lathrop is one of them.

I would like to thank the MIT Artificial Intelligence Laboratory for providing superb computing resources and an environment which is intellectually stimulating.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1. Highway Maintenance Dilemma

> Maintaining the nation's 4 million miles of state and local road network currently requires more than a third of the total highway budget, and the share of highway resources going to maintenance is growing.
> In spite of this spending, continued deterioration of the nation's road system shows the need for more effective maintenance. Methods, equipments and materials have not changed significantly in recent years.
>
> TRB Report[July'84]

The nation's highway network is once again a focal point of public interest. While two decades ago the issue was one of new construction and development, the current topic of concern is that of maintenance . The primary question is how to maintain this network wisely, that is to stop the rapidly spreading deterioration. Without some action being taken, the highway network may no longer continue to play its crucial role in maintaining the nation's interconnection.

The relatively recent concerns about the highway network shed some light upon the level of knowledge we have about pavement itself. When we want to predict the behavior of pavement, it becomes obvious how unsatisfactory our knowledge of pavement really is. This lack of knowledge is not due so much to an insufficient research effort as it is to the inherent complexity of the heterogeneity and non-uniformity of pavement.

The problem in the domain of failure mechanisms in pavement is understanding the nature of knowledge in this evolving domain. By understanding the nature of knowledge we mean recognizing the significant parameters, how they interact, and finally identifying the relationships that describe such interactions. With our knowledge about the failure mechanisms of pavement incomplete and/or unreliable, there have been several competing theories proposed over the last 35 years. However, each of these theories that attempted to decipher the mechanical behavior of pavement cannot explain the nature of knowledge in the whole domain of pavement mechanics by itself. Therefore, either a hybridization of these theories, or the development of a completely different theory is required. But the effort of scrutinizing the available theories, hybridizing them, or forming a new theory if needed, is far beyond human capability, especially when we are overwhelmed by a large number of theories and variables. Due to this immense effort required in this evolving domain from both the quantitative and the qualitative points of view, we thought that the field of pavement mechnics presents a challenging arena in which to apply machine learning techniques of artificial intelligence.

At the same time, we do believe in the potential capabilities of machine learning techniques in building the knowledge of evolving domains. The quantitative aspects of discovery are essential for establishing new domains of science (or engineering). Therefore, this research concentrates on devising a generic tool for quantitative discovery, and applying it to the domain of failure mechanisms in pavement. The work on the epistemological, and then the qualitative aspects of discovery has advances substantially in the last decade (Simon, 1977). Quantitative discovery did not receive similar attention because it is on the border between artificial intelligence and applied mathematics.

The approach followed in this research is hybridizing the symbolic and numeric techniques of learning from observations, explanation-based generalization, adaptive control, and qualitative reasoning into a generic quantitative discovery tool. *Learning from observations* is suitable for building knowledge in evolving domains, where nature (or the studied system) supplies too many instances, positive and negative, without any more clues other than a success/failure criterion. The learning program tries to induce relationships that justify the behavior under consideration, in both its positive and negative instances. *Explanation-based generalization* is very useful in domains where we have several competing theories about the behavior under consideration. It tries to explain any detail of the actual behavior using one ,or more, of the competing theories. Explanation-based generalization enables us to use the learning systems as knowledge-based (expert) systems simultaneously while it is learning. The diagnosis, or prediction, will be given according to the current level of knowledge in the system. Coupling either, or both, of the above-mentioned techniques with a feedback loop of *adaptive control*, we get a means for evaluating the induced relationships and/or the deduced explanations. The feedback comes from the actual system to evaluate the previously made diagnosis, or prediction, through comparing it with the real behavior. For the purpose of this study, we propose a feedback on diagnoses after a year from implementing the remedial action required to cure the diagnosed distress, and six months after predicting the future behavior if no remedial action was prescribed. *Qualitative reasoning* ensures both of the induced relationships, and the deduced explanations, are logically consistent with the general, solid facts of the domain, and with each other.

All the above-mentioned machine learning capabilities makes the discovery system very flexible in adapting to the local environment where it is implemented. That is to say, the knowledge base of the system is updated (modified) to conform to the peculiarities of the environment. Thus, the versions of the same knowledge base distributed to various places may end up differing from each other after awhile, because of the immaturity of the domain, where not all the parameters are discovered yet. So, one may worry about the emergence of different knowledge bases at the various places, and therefore, the concept of *harvest* is introduced to counteract this phenomenon. At regular intervals (every year or so), all the different versions of the knowledge base are harvested to a super-generalization (learning) process. After this process, a new, more global, robust version is broadcasted to the different users.

The task of hybridizing all these features may seem difficult, but it is already under way. The implementation of these concepts is the *HOTEP* system for quantitative discovery and qualitative reasoning. For more information about it refer to chapter 5.

1.2. Conceptual Overview- The Behavioral Modeling of Pavement

In response to recent reports concerning the spreading deterioration of the highway network (Balta, 1984), a great deal of research effort has been aimed at maintenance, rehabilitation and reconstruction over the past 7 years (Moavenzadeh, 1977). Research into pavement maintenance is divided into two major areas: behavioral modeling of pavement, and pavement management systems, where the maintenance policies are optimized. In the rest of this section we will deal with the first of these topics, and in the next section the second topic is dealt with.

The motivation for the development of HOTEP system stems from the numerous evolving fields in engineering. There we are overwhelmed by the large number of variables which we should study. Understanding their possible interactions would contribute to better understanding of the field under consideration.

A good example of such a developing field is failure mechanisms in pavement, where our understanding of the mechanical behavior is far from satisfactory. (Although much valuable work has been done in this area since 1950, the problems demand that a new directions now be taken). Several theories and models had been proposed for representing the mechanical behavior of pavement.[1] The theories proposed range from plastic behavior to visco-elastic behavior (Moavenzadeh, 1977). Several studies tried to follow the fracture analysis approach, but for some reason or another, they stopped [(Kakel, 1968) and (Majidzadeh, 1977)]. Each theory (or model) was too fragile to expand over more than a narrow field of application. The complexity of the pavement mechanics may be imputed to the heterogeneity of pavement and non-uniformity of the behavior of its ingredients.

However, strenuous efforts were -and are still- being exerted to establish better understanding

---

[1] A good survey can be found in (Yoder, 1975) and (Yang, 1972).

for the failure mechanisms in pavement (Markow, 1984). Faced by about 70-100 variables, the researchers found themselves apt to develop several empirical formulae, each of which deals with a few of the many variables. The grouping of variables was a divide-and-conquer strategy rather than being based on some sound conceptual clustering. Thus the empirical formulae were developed including those variables that the researchers *thought* of as related/relevant ones. Such empirical formulae may give us the right answer for specific circumstances, but often they do not apply under other circumstances.

Thus we may rephrase the problems facing the evolving knowledge of developing fields as follows:

Empirical Formulae:

Most of the relationships in the developing fields are merely empirical formulae which are unable to offer an integral view and understanding of the domain.

Lack of Physical Meaning:

Those empirical formulae lack any physical meaning mostly because of their violation of dimenensional homogeneity between their terms. To maintain dimensional homogeneity, a formula must not add a term in "feet" to another term in "psi", for example, leaving the task of dimensional homogenization to some ambiguous dimensionful constants. Therefore, the empirical formulae do not allow us to gain insight into the nature of the domain, and hence they cannot become a law or a theory.

Inability of humans to handle many variables at the same time:

Both of the above-mentioned problems are because of our inability to handle many variables at the same time. On the other hand, we cannot leave these variables to super-statistical computer packages to detect all the possible tendencies between the different variables. Insight requires more than number-crunching powers in conducting the different processes intelligibly. Thus we find ourselves looking at an perfect field for applying the machine learning techniques.

In the following paragraphs we will try to classify these behavioral models into major categories according to the algorithmic approaches used in them.

1.2.1. The Mechanistic Simulation Models

These models try to simulate the mechanical behavior of pavement. Therefore they are consisted of sets of equations that express the different distress symptoms of pavement in terms of independent variables (structural, traffic or environnmental). These relations belong to one of the following categories:

I. Theoritical relations :

which are based upon a sound theoritical basis. Unfortunately, the known facts about the mechanical behavior of pavement do not cover but a small portion of the basics of the domain due to our low level of knowledge about pavement. Therefore they are augmented by different kinds of hypotheses to make full modeling for the domain. These augmentations exposes the models to the same criticism made against the other categories.

II. Hypothetical relations :

which are constructed according to (or considering) a specific hypothesis (or postulate) that lacks proof that similar behavior is exhibited by pavement. Examples for such hypotheses are the assumption of either plastic or elastic behaviour of pavement, and the assumption of an equivalent layer. Examples of system that implement the postulate of viscoelasticity of pavement are VESYS (Kenis, 1977), and VISTRA (Battiato, 1978).

III. Empirical formulae :

which are based upon a regression analysis for the different variables that are thought of as influential with respect to the distress symptom under consideration. The resulting relationship may hold true only under a restricted combination of environments. An example for a system that implements empirical formulae is EAROMAR-II (Markow, 1984).

Hybrid models are frequently built implementing more than one of the above-mentioned approaches. From the above-mentioned categories, we can determine the main limitations of this approach as:

- *Lack of sound theoritical basis* in most of the cases.
- *Locality* due to constructing the regressional (or hypothetical) relationship within specific environments. So, the validity of these relations are bound to their original environments.

1.2.2. Statistically-based Models

These models are based upon the statistical manipulation of a large record-database (case library) for all the maintenance works that have been done in the last two decades or so, in the district under consideration. The mechanistic behavior of pavement is represented as a state space, where the mechanistic behavior is expressed by various combinations of physical properties, each distinctive combination is called a state. The mechanistic behavior is represented as a graph, where the nodes represent the different states, and the edges (links) represent either the deterioration, or the upgrading of pavement from one state to another. Each edge may have a probabilistic weight attached to it, representing the possibility of such a state switch. The first proposal of adapting state space in representing the mechanistic behavior of pavement is found in (Findakly, 1971).

Both theory of probability and Bayesian theory are used in studying the statistical patterns of changing the status of pavement from a physical state to another. Thus an inter-state matrix of switching probabilities covering all the possible switchings between all the available states is built. Out of the physical state matrix, an "event tree" and its corresponding "probability tree" are developed. These event trees are representing the statistical relation between any distress manifestation and its possible basic causes.

The major limitations for this approach are the following:

- *The necessity of having a well-documented database* for all the deterioration manifestations, their diagnosis, and the maintenance works in the concerned area for the last 10-20 years is almost an impossible requirement.
- The obvious *Locality* of such a model. This locality stems from the dependence of such a model upon the peculiarities of a specific area, the maintenance policies followed in this area, and the style of maintenance' documentation, such as, the numerical versus alphabetical grading for different values, and the index references used by the local agency[2].

---

[2] Different values maybe given to the the same deterioration severity in different places due to the various index systems, which are explained in chapter PAVEMENTMAINTENANCESYSTEMS.

- The application of the State space concept to our current level of qualitative knowledge about pavement perpetuates our dysinformation about pavement and dampens the endeavor to enhance our understanding of pavement, both on the quantitative, and qualitative levels.

A good example for this probabilistic approach is the model used by Arizona Department Of Transportation, ADOT, (Golabi, 1983)

1.2.3. Stochastic Models

These models, as their name indicates, implement the stochastic processes such as theory of control and Markov processes in the field of pavement management systems. The closed-form equations are used for representing the cyclical nature of the various deterioration/maintenance series of processes. These stochastic models are based upon statistical manipulation of maintenance record databases. A considerable advantage of this representational form is its simplicity compared to the finite element methods.

This approach is yet under development, therefore aside from the performance point of view, we can detect the following limitations in its underlying conceptual basis:

- The efforts are aimed at coping with a fixed *"life Cycle"* , rather than learning what are the parameters that affect this life cycle and then trying to prolong it.
- As far as we do not know the real mechanical behavior of pavement, we cannot be sure about validity of the maintenance action as remedy of a specific deterioration manifestation, therefore, the developed closed-form equations are based on sheer statistical coincidence.
- Presuming the compatibility of the deterioration/maintenance pairs, the *fatigue* factor of the repetition of this pair of processes is yet ignored.

An example for an early implementation of this technique is a stochastic model is implemented on a hand-held programmable calculator at MIT (Balta, 1984).

1.3. Optimizing The Maintenance Policies

Beside the behavioral model, every pavement maintenance system should have a module for optimization of the maintenance policies. In this part, every deterioration symptom has, at least, one maintenance policy (or countermeasure) to prevent it or remedy it. Each of these optimization systems consists of the following elements:

- *Policy Database*: where the different policies are stored. They are categorized by deterioration manifestation which they can tackle.
- *Objective Function*:which represents the priorities and strategy of the user highway agency.
- *The Constraints*: which may be financial, planning-related, or for continuity and feasibility.

The way used in forming the objective function, and the superimposed constraints for such optimization problem are the points of difference between the various maintenance optimization systems. An example of such a pavement management systems is EAROMAR-II which was developed at M.I.T. (Markow, 1984).

1.4. Problem Configuration

For all the above-mentioned approaches, the same basic problem can be found: *lack of sufficient, and reliable knowledge about pavement*. All the different techniques mentioned in the previous section have attempted to compensate for this defect either by posting simplifying hypotheses or by restricting their models to limited applications or regions. Unfortunately, these systems have not met much success. The reason might be that none of these systems (models) has attempted to raise the level of knowledge about pavement mechanics that would improve our capability in maintaining our highways optimally.

Thus, the problem we are tackling is understanding the nature of knowledge in an evolving domain. By understanding the nature of knowledge we mean recognizing the significant parameters, how they interact, and then finally finding the relationships that describe such interactions. Such understanding enables us to model the actual mechanical behavior of pavement. Our domain is failure mechanisms in Pavement where knowledge is incomplete and/or incorrect. The approach we follow is a hybrid of qualitative reasoning, explanation-based generalization, and mathematical tools.

Despite their differences, all the pavement maintenance techniques share a common hypothetical approach which limits the advancement of knowledge concerning pavement behaviour. In the following paragraphs, major characteristics of this approach will be listed, for more details about these drawbacks refer to section [3.1]. For the pragmatic obstacles met in the application of these systems refer to section 3.2.5.1.

1. They do not map the actual behaviour of pavement. In the same time, they do not recognize the need to establish a methodology for the gradual building of knowledge that would assure continuous improvement in our understanding of the actual behavior.

2. The hypothetical/empirical formulae which form the foundation of most of the techniques are unreliable. While in the proposed system, we use the currently-reached level of knowledge to give diagnosis to the deterioration case in hand. Later on, the system learns from the follow up of its diagnosis in site, and consequently, it updates the knowledge base.

3. These techniques require a human expert, to read, analyze, and interpret the deterioration symptoms. All of the techniques until now also require an expert to separate sensible solutions from unrealistic ones, when reading the output of these programs. However, these experts are few, and their teaching process is very expensive and time-consuming. So, the cost of providing an expert for each jurisdiction to be responsible for maintaining the local highway network makes these programs impractical.

4. All of the previous systems are not *self-correcting* as they do not take into account any feedback of the actual results of applying the system's decisions in the field. Therefore, the reliability and credibility of theses systems are seriously in question. The flaws, detected by feedback evaluation, may be either in the diagnosis process, or it may be deeply-residing in the behavioral model.

1.5. Knowledge-based Systems - Background for a new Approach

In the seventies, not only did the need for systematic maintenance of highways emerge, but it was also a period during which a new approach for tackling non-algorithmic problems was developed. This approach is the knowledge-based (expert) systems - a new wave of programming using artificial intelligence techniques that attracted surmounting attention not only to itself, but also to AI in general.

1.5.1. What is Artificial Intelligence

"Artificial intelligence is that field of computer science that studies ideas that enable computers to do the things that make people seem intelligent."

(Winston, 1984).

An attempt to define some of those "things" mentioned by Winston, is given by (Nilsson, 1980) as follows:

- The ability to acquire and apply knowledge,
- The ability to manipulate and communicate ideas, and
- The ability to reason.

Computer scientists in the mid-fifties were motivated by an overwhelming dream of the infinite capabilities of the computer - the ability to solve any kind of problem. This was one of the high expectations that emerged after the development of the computer. An embodiment for these high expectations is "the general problem solver" (Davis & Lenat, 1982) which is a robust universal program that was thought to be applicable to any problem in any domain.

It was not until the late fifties that computer scientists began to get convinced of the vast distance that separates them from their goals/dreams. At that point they began thinking about what are the features that enable humans to solve problems, and how those features are acquired or constructed. This was the first step in the field of artificial intelligence and it was taken at MIT (Minsky, 1969).

Artificial intelligence is now a well-established field that includes many successful sciences, such as, robotics, natural language, pattern recognition, logic programming, machine learning, automatic reasoning, theorem proving, and vision.

Most of the recent interest in AI from the various disciplines is due to the emergence of a wave of programs called *expert system*. These programs implement different AI techniques, especially logic programming, in solving the non-algorithmic (qualitative) problems. We feel that this new trend deserves a separate section to describe it, and to unveil some of the misconceptions shrouding it. Also a trial is made to differentiate between expert system and artificial intelligence.

1.5.2. Why They Are called Expert Systems

The rapid advancement in AI sciences and the promising implementation in areas of practical concern encouraged computer scientists to prove how much progress has been achieved in AI. Therefore they started applying the different AI techniques to some of the long-waiting problems in several practical domains. The common feature of these problems is their unsolvability by any algorithmic approach because of their need for several unwritten rules of thumb and commonsense. These problems needed either a completely logical approach or a hybrid of logic and algorithmic approaches. The resulting computer programs have been called "knowledge-based systems" or "expert systems" (Hayes-Roth, et al, 1983).

The first expert system, DENDRAL, was developed in Stanford university to recognize the composition and structure of any chemical compound (organic or inorganic), a non-trivial problem with an 8-digits number of candidate solutions (Feigenbaum, et al, 1971).

Dendral was followed by a long series of expert systems in several areas. One of these is MACSYMA; an expert system for solving almost any mathematical problem at the level of a graduate student in mathematics department. This system was developed at MIT (Moses, 1967). Another system is PROSPECTOR, a system which performs mineral prospecting by reading and interpreting the geological maps. It was developed by SRI in 1979 (Duda, et al, 1979).

Medicine was the field that attracted most of the interest of expert systems builders, because of being the field of the highest percentage of unwritten rules. The ill-structuredness resulting from those unwritten rules made the medical domain a perfect arena for applying the knowledge-based techniques (Szolovits, 1978). In response to this need, several systems were developed for medical diagnosis in various areas, such as: "MYCIN" (for microbial infection blood diseases) (Shortliffe, 1976); "PUFF" (for the diseases of the respiratory system) (Freiherr, 1980); "CASNET" (for glaucoma diseases); "INTERNIST/CADUCEUS" (for internal diseases) (Pople, 1979); etc.

In 1978, a major advance was made in the field of knowledge-based systems. This came about with the introduction of the "domain-independent systems". The first of these being "EMYCIN", a similar to the above-mentioned "MYCIN" but after evacuating it from the domain knowledge, what remains is called the "inference engine" which can operate on any set of rules of any domain. Several domain-independent systems followed EMYCIN, such as, "KAS" (derived from "PROSPECTOR"), (Hendrix, 1975); "EXPERT", (Weiss & Kulikowski, 1981); "ROSIE", (Fain, et al, 1982); "HEARSAY III", (Balzer, et al, 1980);etc. One limitation of these domain-independent systems is that each system is built according to *specific assumptions and scope*, and based on *specific logic* in problem-solving. Therefore, these systems would be of great value to someone with a problem that fits one of the molds of the domain-independent system.

Another major step in the advancement of the expert systems' techniques was the development of the "representation languages". These languages are composed of several macro-commands, each of which maps a task or function performed by the human expert. However, the final structure is left to be defined by the user by using these macros as building blocks, and the overall environment is to be designed around them. Another important feature of these languages is the ability to construct new macros from the available set of commands to meet specific requirements. These representation languages include: OPS-5 (Forgy & McDermott, 1977), AGE (Nii & Feigenbaum, 1978), RLL (Greiner & Lenat, 1980), LOOPS, etc.

The difference between expert systems and artificial intelligence should be declared, especially because of the blurring misconception of considering them synonyms. Indeed expert systems are logic programs that are used in coding the verbal rules[3]. Therefore, they are useful in applications where the judgement is mostly qualitative, and where the current level of knowledge of the experts is satisfactory. An essential feature of intelligence is learning. Therefore, a valid accusation of un-intelligence is directed to the expert systems because of their perpetuation of the current level of knowledge without improving it. No expert in real can live up to his claim of expertise without a continuous process of Learning. Therefore, we see the central role of learning in all the sciences of artificial intelligence.

1.5.3. How Can AI Address Our Problem

After this brief overview of the expert system technique, now, we go back to our basic problem, that of highway maintenance. Unfotunately, the expert systems approach described above would not be helpful in the case of failure mechanisms in pavement as it emphasizes the qualitative nature of the problem to be solved and the completeness of the knowledge about the domain, neither of the requirements is available in the case of pavement mechanics.

Due to the evolving nature of the domain of failure mechanisms in pavement, we think it would be an excellent domain to apply machine learning techniques.

Learning processes include the acquisition of new declarative knowledge, the development of

---

[3] The major tool in the coding of the qualitative knowledge is the concatenated and/or nested *if-then* rules. These rules are searched either in a forward chaining mode or a backward chaining one.

cognitive skills through instruction or practice, the organization of new knowledge into general, effective representations, and the discovery of new facts and theories through observation and experimentation. The study, and computer modeling, of learning processes in their multiple manifestations constitutes the subject matter of machine learning (Carbonell, 1984).

The drawbacks of the above-mentioned pavement maintenance systems can be overcome quite effectively by implementing the concepts of quantitative learning.

The area of quantitative learning has received relatively little recent attention because of being on the border between Logic and numerical analysis. To understand and mimic the cognitive process of Learning -whether by discovery or otherwise- we need symbolic techniques, while to deal with quantities we certainly need accurate numerical tools. Attempts to ignore either of these aspects have proved insufficiently powerful, e.g., the BACON series (Langley, 1984), and ABACUS (Falkenhainer, 1984).

Any stable relationship in nature should render itself susceptible to some logico-mathematical formulation. Our ability to disclose such relationship is a completely different issue, and here comes the role of discovery. Therefore,we may say that the process of discovery is an application of some cognitive processes to logical (or mathematical) probes (or tools) (Michalski, 1984). At least within the realm of engineering research we may redefine the quantitative discovery as adaptively-controlled inductive learning. Thus to develop a system that is capable of quantitative discovery we need to build the appropriate cognitive processor armed with the suitable logico-mathematical tools. A vital feature of such a system is parsimony, which set it aside from any exhaustive search by enumeration.

1.6. Outline of the Thesis

Chapter 2 describes the objectives of the research and defines a scope of the work. In chapter 3, a survey of the major pavement maintenance systems is given. Chapter 4 offers an epistemological study of the domain of failure mechanisms in pavement. Chapter 5 describes the system configuration of the proposed HOTEP system, and the current progress in implementing its various parts. The sixth chapter deals with polymorphism as a solution of the ill-structured domains, and how it helps in cutting down the data acquisition effort through the lookahead feature. The last chapter is the conclusion of the thesis, and it shows the limitations of the current research as guidelines for the future work. An annotated bibliography is covering the areas of pavement mechanics, pavement maintenance systems, and artificial intelligence, among other related areas. Then, an appendix is devoted for the description of the QUDS, an implemented part of HOTEP system, and some illustrative, real examples, drawn from various fields, are given to demonstrate the performance of the QUDS system. The other appendix describes the Π-theorem of dimensional analysis.

Chapter 2

Objectives and Scope of work

2.1. Objectives

We have two major objectives for developing the HOTEP system described in chapter 5. The objectives are:

1. Enhancing our knowledge about pavement and fracture mechanics, in general, and failure mechanisms in pavement in particular. In order to reach this goal, we apply some of the artificial intelligence concepts, crystalized in the HOTEP system that is under development in this research.
2. Integrating symbolic discovery concepts with quantitative inductive techniques into a general-purpose tool for quantitative discovery. The quantitative discovery system relies on a qualitative reasoning unit in ensuring the plausibilty of the discovered hypotheses. The crystalyzation of these concepts is acchieved by building HOTEP, a quantitative discovery and reasoning tool.

2.2. Scope of Work and Limitations :

Based upon the above-mentioned status of the highway network, the current pavement maintenance systems, and the present developments in machine learning, we are drawing hereafter the major features of the proposed system in conformance with a multi-phase plan. The plan, as discussed in the next section, was formed, such that, the fundamental features are implemented in the first two phases of this plan. These first two phases of the plan are the subject of this thesis. As mentioned before, the objective of this project is to develop HOTEP, a general tool for quantitative discovery and reasoning about any domain. HOTEP provides diagnosis (or reasoning) for the cases submitted to it, according to its current level of knowledge. Because of the reasons discussed in the previous chapter, failure mechanisms in pavement was chosen as an application domain. This system will have the following features :

1. *Giving diagnosis* for any pavement deterioration case according to the latest level of the knowledge attained by the system.

2. Predicting the future behavior of the pavement under study.

3. Prescribing, at request, the appropriate remedial action for the specific diagnosis.

4. Prescribing the least amount of remedy required to prevent a specific state of distress predicted in the future[4].

5. Providing, at request, the qualitative and quantitative reasoning for attaining a specific decision.

6. Acquiring data parsimoniously, using polymorphism and "synthesized links of generalization" (Pople, 1979) to keep the user's participation at a minimum.

7. In case of failing to diagnose the causes of a case due to lack of relevant knowledge, the system asks the user about it and adds the new piece of information permanently to its knowledge base after checking its compatibility with the rest of the knowledge base.

8. The system follows up every diagnosis (or prediction) it makes (through feedback from the field) to revise both the credit assigned to every bit of information in the knowledge base, and to revise the knowledge base itself (Holland, 1986) . The pavement will be field-checked after an appropriate period of time (probably a year) to check the decision effectiveness by comparing the expected versus actual results. Then, a re-evaluation of the diagnostic approach (credits and facts) is made. Every new case combined with its revised diagnosis, constitutes a complete case. All the cases, encountered by this system version, are saved into a *Case Library*. The impact of every complete case upon the current status of the knowledge base is assessed, and accordingly the knowledge base is revised. Thus the system is in a permanent process of learning.

9. As the human researchers do, the system will continuously observe the cases, both successful and unsuccessful, in its *Case Library*, and the approaches it used in tackling each of them. The goal of this scrutiny is finding out if there could have been a more comprehensive and better expressive relations to describe the pavement behaviour. Such discovery procedure will facilitate better understanding of the real mechanical behaviour of pavement, and, hopefully, it will lead to new relationships, or to the modification of existing ones. A quantitative discovery system, QUDS, has been developed as a part of the HOTEP system, to carry on this task.

10. The system keeps track of every case it encounters in a "Case Library". This Case Library is used for immediate diagnosis of similar cases, so, HOTEP saves considerable computational resources when it tackles a case similar enough to another case the system had encountered previously. The similarity between cases is judged through a *matching threshold system* .

---

[4] refer to preventive planning in (Doyle, 1984)

11. The learning capability of the system adapts its knowledge base to the peculiarities of the site in which it is applied. Because of the immaturity of pavement mechanics, the adaptation process will take every version of the system in a different direction[5], so, by the end of the year, we may find the copies of the knowledge base, that we broadcasted to the different sites, ended up diferent from each other. At regular periods, say annually, the would-be-different versions of the knowledge base should be *harvested* for a super learning session, after which a new release of the knowledge base is broadcasted for the different users. The learning capability of the system assures the adaptation of the system to the local environment, while, the harvesting process assures the robustness and globality of the acquired knowledge.

12. The system helps in investigating the new theories and hypotheses by extending them to their possible limits. Such automatic testing will allow the scrutiny of a larger number of theories and their mutations. This exploring feature is in the core of the discovery process.

13. Another feature of the system is robustness (or universality), which means that it would be applicable to any region with almost no initial adjustment required (compared to the network zoning and classification required by Arizona NOS system (Golabi, 1983)). But like the human expert, when still novice, it starts with the formal knowledge that a novice would get at school. By this formal knowledge, he can tackle most of the problems (but in a slower and unexperienced way), for the rest of the problems, the system will have a knowledge acquisition facility that would work for covering any knowledge gap that is found. As the system passes through the cases, it becomes more experienced, more accurate and faster. The system grows automatically with no programming effort at all, as it builds up its own expertise from past cases.

14. The system will check the compatibility, and the correctness of the input data through a semi-logic procedure before starting processing them. Also the system will attempt to extract the maximum benefit from the input data before aquiring any further data.

15. The system, in its final version, should be user-friendly and allows the operator's responses and directives that use simple English (natural language) such that no technical background is required from the operator. In the same time the user is enabled to go deep into the system to examine any intermediate result, or to trace, explicitly, the process of decision making.

16. One of the major uses of this system is as a brain for a proposed *road machine*, which is to be developed in the last phase of the long-term plan detailed in section 2.3 page 31. A closer look at the final integrated system reveals the following: The road machine is a multi-apparatus unit, hooked to an automobile that carries a personal computer(e.g. HP96) in which HOTEP is loaded as a brain for the road machine. A possible scenario would be as follows:

---

[5] This is because, in evolving domains, some of the important parameters are not discovered yet, and they are embedded inside coefficients, or inside other dependent parameters in the modeling relationships.

- The car begins its routine (or emergency) surveillance patrol of a certain portion of the highway.
- As the car goes on, the road machine takes readings (or pictures) and translates them into symptoms (through a pattern recognition program).
- These symptoms are transmitted spontaneously to the HOTEP, which immediately analyzes them, and gives either a diagnosis, or an order to the driver to stop, so that the road machine may collect further data (either readings or tests) while they are still in the same area. Thus the maintenance decisions are made very quickly, and yet, at low operating costs. The motivations for this plan has been discussed in chapter 1.

2.3. Work Breakdown Plan

As we mentioned previously, the proposed system is a part of a long-term plan. The optimal goal of this plan is the development of an *intelligent mobile maintenance station* to compensate for the severe shortage of pavement experts, which are demanded at every jurisdiction in the country to take care of its local highway network. The process of preparing an expert is very expensive and time-consuming, and that is the cause of the shortage.

This plan [2-0] is expected to take about seven years for complete development. It is designed such that its three phases are distributed over the study period in a series of projects. Each of the four phases is completely independent, from the points of view of system building, and implementation. In the following paragraphs, a brief exposition to these phases is made.

1. Phase(I): Literature Review for the following topics:
    a. Artificial intelligence tools, with special emphasis upon the different kinds of machine learning , theorem proving, qualitative reasoning, and knowledge-based systems.
    b. Pavement maintenance systems, with special emphasis upon the conceptual foundations for the different approaches in this domain.
    c. The epistemology of failure mechanisms and fracture mechanics in pavement.
2. Phase(II): Building HOTEP, a system for quantitative discovery and reasoning, and then applying it to the failure mechanisms in pavement. which would cover the area of diagnosis of flexible pavement. This phase includes the following steps:
    a. System Design.

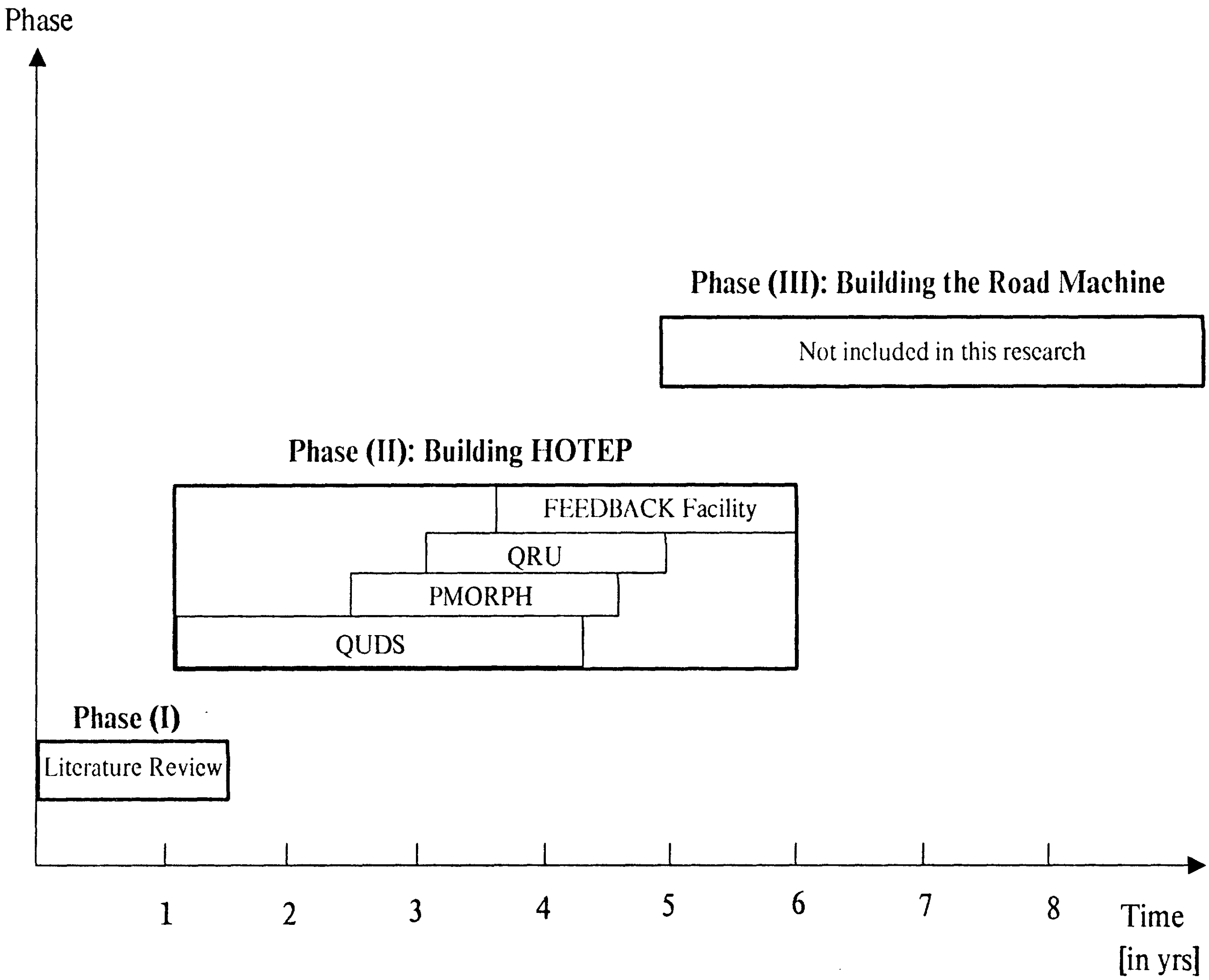


Figure 2-1: Work Plan

   b. System Construction.

   c. Conceptual Calibration.

   d. Testing and commissioning of the system by real data.

3. Phase(III): Development of the road machine, which would be the third stage of the project. This road machine will include several equipments for the detection of any kind of distress. One of the major equipments to be used in the road machine is the pulsed electromagnetic wave radar, PEWR, which will provide a real-time surveillance of the pavement condition without hindering the traffic, as it will be carried by the mobile road machine. The implementation of the PEWR is the point of research that, currently, takes place at MIT. For further information refer to (Maser, 1986). The major elements of this radar system will be the following:

   a. Antenna which will emit (pulse) the electromagnetic waves towards the pavement.

   b. The receiver of the radar will receive the reflected waves.

   c. The sampler will pick sample of the returned waves every specific period, and the transmit it to the analog device after filtering it from the accompanying noise.

   d. The analog device (oscilloscope) will rectify (orthogonize) the waves and then produce a full wave form. This full wave form is transmitted to an analog-to-digital converter

   e. The Analog-to-Digital Converter will digitize the wave form into a digital file processible by any ordinary computer algorithm.

   f. The digitized wave form will be transmitted to the signal preprocessor which is the interface between the radar and HOTEP system.

Chapter 3

Insights in Pavement Maintenance Systems

The purpose of this chapter is to give an insight into the epistemological[6] structure of the pavement maintenance process. After setting the taxonomical framework for classification of systems, the we review several pavement management systems and methodologies. The review discusses the main elements of the process and how do they interact, or cross-affect each other in forming the whole pavement maintenance (or management) system. Then, we address the guidelines of the epistemological study of the characteristics of the domain of pavement maintenance with an emphasis on the impact of the mechanistic hypotheses upon the overall performance of the system. Because of the profoundly ill-structured nature of the domain of pavement mechanics, a polymorphic[7] framework of representation is proposed as a solution that would shift the pavement mechanics knowledge from its current, shallow, unstructured nature towards deeper, betterly-structured form, that would improve our capability to comprehensively understand pavement. Such epistemological study should be pursued extensively to lay a required foundation for any thoughtful implementation of artificial intelligence tools in the domain of pavement maintenance. One of the major advantages of using a polymorphic representation is achieving a parsimonious data acquisition through a Look-ahead search implementation, as detailed in (pp. 82).

---

[6] Webster defines "epistemology" as "The study of the nature and grounds of knowledge esp. with reference to its limits and validity"

[7] Webster's defines "Polymorphism" as "The quality or state of being able to assume different forms; The property of crystallizing in two or more forms with distinct structures."

3.1. Methodological Taxonomy

the taxonomy proposed here in this chapter is constructed in compliance with the general classification used in the other researches conducted in the field of pavement maintenance systems at MIT (Balta, 1984), and (Elliot, 1969). Due to misnomination, the major *categories* were called *levels*. Therefore a caution should be paid in defferentiating between the major levels (that mean categories), and their real composing levels. In the following sections, we will explain the classification's concept used in each subsequent level of the taxonomy.

3.1.1. Level(I): The scope of system

At this basic level,we classify systems under one of two categories, according to the scope of interest of the system:

3.1.1.1. network-level Category

In this category, the highway network, under consideration, should be looked at as one entity, all over the course of analysis. The final objective of the system is determined according to a general strategy that takes into consideration several high-level, *general* aspects of financial, managerial, prioritizational constraints. A good advantage of this approach, in some implementations, is that it takes into account the interactions (cross-effects) between the different links (highways) of the network due to any change in the road characteristics (traffic, structural, or environmental). From the point of view of transportation systems analysis, the objective function of these systems is the maximization of the utility function all over the network.

3.1.1.2. Project-level Category

In this category, the links (portions of Highways) are taken into consideration individually. In most of the cases, this down-scoping results in more effort to be directed towards the technical, economical, and scheduling details. this means more *demand-responsive* decisions. From the point of view of transportation systems analysis, the objective function of these systems is the maximization of the utility function of the individual link under consideration.

3.1.2. Level(II): Types of Options

The types of options depend upon the underlying scope of system (level I),therefore, each scope has its relevant options as shown in the following paragraphs.

3.1.2.1. network-level systems

1. <u>Statistical, non-discreminant models :</u>

Another nomenclature for this category is *non-equilibration* models. A destinctive feature of these systems is the holistic, generalized trend used in dealing with the maintenance of highways. They deal with the frequency (probability of occurance) of each deterioration manifestation, all over the network regardless of which individual Highway suffers from which deterioration manifestation. Accordingly, the resources are allocated for the maintenance of specific (predicted) number of occurances of each manifestation all over the network.

```
              |-----|
     Decision |     | Level
           |-----------|
           |Processing |
|---------------------------------|
|     Entry Level                 |
|---------------------------------|
```

This probabilistic approach squeezes the wide, raw input-data level into a brief, statistical-information level. Consequently, the investment and maintenance decisions could be formed in general terms without allocating resources to specific highways. Such generality serves, primarily, in providing elasticity in spending the limited resources, and secondarily, in meeting the restrictive time limits of maintenance. Thus the resource allocation process depends upon the strength of the advocate of each region, at the time of resource allocation which is done manually outside the scope of the system. A good example of this approach is *The Network Optimization System, NOS* developed by Arizona Department Of Transportation, ADOT (Golabi, 1983).

The major weakness point of these systems is the neglect of the resource allocation to the specific projects and leaving it to the human decision, which may lead to short-term savings on the expense of the long-term economy due to the natural tendency of deferring the weakly-defended projects till their Highways turn into serious condition that needs extensive remedy. A partial solution for this point is made through regular survey of highway'stata. Two-year frequency makes a reasonable countermeasure.

2. <u>Equilibration systems :</u>

This is the most perfect, and consequently, the most expensive option to be used. the network here is studied at the lowest level of interaction between the various links of it. the network equilibration enables us to know the change in the traffic distribution all over the network due to any deterioration (or maintenance action) at any of the links. By the word *equilibration* we mean the eqilibrium between the demand of the users and the supply of the operators all over the network, So, we can determine all the traffic parameters of the deterioration problem, such as, type of traffic, volume, and modal distribution of traffic over each link. After getting these vital data, a more precise prediction could be made for the deterioration of pavement.

If we want to discuss the size of information at the various levels of the system, we can see that the original wide entry level is inflating considerably at the processing level due to the study of enormous number of cross relations between the different links of the network. Then, the decision level is reasonably limited in comparison with the previous two levels. Therefore, the space complexity of equilibration systems is relatively high (Aho, Hopcroft, & Ullman, 1974).

```
                   |-------------|
                   |Decision Lev.|
|-------------------------------------------|
|              Processing Level             |
!-------------------------------------------!
           |       Entry Level       |
           !-------------------------!
```

No complete example is available for the network eqilibration model because of its expensiveness. As a partial example, we can look at *the road investment analysis model, RIAM* (Moavenzadeh, 1977), which is limited to low-volume, rural Highways and it deals with the network equilibration in an integer programming mode which may imply an *all-or-nothing* strategy. Another partial example is the *intercity multimodal transportation, IMT* (Safwat, 1982). This model performs a sophisticated equilibration of the highway network, but stops before tackling the problem of Highway maintenance, but a maintenance module could be augmented very easily due to its extensively-detailed output about the traffic and economical impacts of any investment policy.

3.1.2.2. Project-level systems

1. Alternative Project :

This is one of the promising approaches due to its incorporation of a mechanistic simulation model of pavement, in addition to, several tools to meet the financial, engineering, mangerial, and traffic constraints of maintenance decision process with regarding a specific highway. It uses the concept of simulation models, in terms of regressive equations, to simulate the various deteriorational aspects of mechanistic behavior of pavement, based on different mechanistic assumptions and hypotheses.
Instead of predicting the distress behavior of pavement, some other systems use several empirical formulae for predicting the serviceability of the highway. The serviceability formulae are closed-form equation, derived stochastically using the control theory (Balta, 1984).

In order to meet the economical and constraints, this approach uses various techniques, such as integer programming for resource allocation, net present value and other methods for project evalaution. The space complexity (memory space) is by no means an obstacle in any of the project-level'options (Aho, Hopcroft, & Ullman, 1974). A good example of this approach is EAROMAR-II (Markow, 1984), which uses a simulation model to represent both the damage relationships, and the serviceability of the highway pavement. EAROMAR-II uses after that, economical analysis to study the cost and impact of any strategy used to rehabilitate the Highway under consideration, over a specified span of time. After economically comparing several alternative strategies, the user can chose the most appropriate strategy. A discussion of the methodology used in the assessment of pavement damage will be found in the next chapter.

2. Single project, or mechanistic exact solution:

The second title explains both the advantages and disadvantages of such approach, as the exact mechanistic behavior of pavement is yet unknown, to large extent, and all the relationships, used till now, do not exceed being hypotheses backed by different regression analysis results. These hypotheses have different strengthes at the different aspects of pavement behavior, such that, you cannot expect high performance from a hypothesis in predicting *all* the kinds of deterioration everywhere. However, several systems utilize several relationships, each of which is based on different hypothesis, such that for each distress manifestation, we use the most expressive relationship. This approach may lead to a pitfall if the different hypotheses are based on contradicting assumptions.

The advantage of this approach is having quick, reasonable simulation of the mechanistic behavior of pavement, so that we can get a fairly precise prediction for the impact of any maintenance action or strategy, using our limited knowledge of pavement mechanics. Due to the vagueness of the area of pavement mechanics, no simulation model is widely used. But that does not depreciate the value of some systems, such as the viscoelastic model, VISYS, ENI's model for pavement design (giambolini), and Viscoelastic layered system (Moavenzadeh, 1977).

3.2. Pavement Maintenance Systems : Topology of the field

This section discusses the main elements of the field, and how they interact or cross-affect each other. First, we will discuss the major elements (concepts) that should be included in any system (Balta, 1984).

1. The *decision criteria*, or priority ranking for the competitive policies.
2. *Time span* of the analysis.
3. The required *frequency of pavement inspection*, which parameters to be inspected, and how to gather the data (data acquisition).
4. The *mathematical procedures of the decision process* (e.g.,optimization).
5. Measures of *pavement conditions* (e.g., PSI, PSR, etc).
6. The type, detail, and intended use of the *results produced* by the system.
7. The predictive (simulating) capabilities, used in the *assessment of mechanistic damage relationships*.
8. The *success (impact) of each policy* in improving a specific criterion in pavement.

After brief description of each of the above-mentioned points, we concentrate on the last two points, elaborating on the potential improvements in these two aspects by using the different artificial intelligence techniques as detailed in chapter 5.

3.2.1. Decision Criteria

As we mentioned before, the decision criteria in highway maintenance depend upon several parameters including: financial, engineering, political, managerial, and of course, they depend upon the scope of the maintenance program (network level, or project level).

3.2.2. Data Acquisition

In the following paragraphs, we are discussing what are the data to be acquired, and how to acquire them. The manipulation of the acquired data is dealt with in section 3.2.5 [pp. 50].

3.2.2.1. The Acquired Data:

The acquired data represent two destinctive points of view in looking at the deterioration of highways, the first of which deals with deterioration as a mechanical problem, while the other point of view tackles the problem from a utilitarian aspect. Therefore, the acquired data are classified under one of the following categories:

1. Physical deterioration of pavement :
   This category includes all the parameters that represent the actual *pavement distress* in two dimensions:
   The extent of Distress, **and** *the severity* **of it.** The set of pavement distress manifestations includes the different kinds of cracks, rutting, potholes, etc. The parameters of this category are a subject of our research as we may find better, and more expressive parameters to represent the deterioration manifestations, or we may find new relationships, of better correlation between these (expected) parameters and their actual values.

2. Disutility of highway due to deterioration:
   This category is more controversial as it includes the parameters thought to represent the ill-defined concept of disutility of the highway by the users. Anyhow, this category generally addresses most of the following elements:
   - Ride Quality.
   - Roughness.
   - Skid Resistance.

There are two major areas of research in the field of data acquisition which need extensive efforts, which are parsimony of data acquisition, and the automation of it.

1. Parsimony of Data Acquisition :

This means the minimization of information acquired for the decision making process. This parsimony would never be achieved but through either of the following ways:

- Forming more precise relationships for deterioration assessment. Machine learning is a possible vehicle for attaining such a goal, and therefore, we implement different techniques of it, as described in chapter 5.
- Discovering new parameters (or relationships) in the area of pavement mechanics. A candidate field through which we may discover new parameters is remote sensing. For a rough description of such potentialities, refer to (Maser, 1986). The discovery of new relationships is one of the major goals of this thesis.

2. Automation of Data Acquisition :
Data Acquisition is one of the most constraining bottlenecks that hinders any pavement maintenance system from being widespread, because data acquisition process is, till now, very labor-intensive. The automation of data acquisition has two consequent stages:
    a. Mechanization of most of the processes of data acquisition.
    b. Full automation of the whole procedure of data acquisition including the decision process. A detailed discussion of this stage was made in the description of the road machine, which is the last phase of the proposed work plan discussed in section 2.3.

3.2.2.2. Methods of Data Acquisition :

Data Acquisition for the shear distress parameters (e.g. cracks, and potholes) is, generally, visual, despite some state agencies use measurement for some of the parameters such as Alaska, Arizona, California, and Florida. On the other hand, the parameters for roughness, deflection, and skid resistance are acquired by specially-devised equipments such as:

- **For Roughness:** Mays Ride Meter, PCA Car Roadmeter, DC Differential Transducer, etc.
- **For Deflection:** Benkelman Beam, Dynaflect, LaCroix Deflectograph, etc.
- **For Skid Resistance :** Mu Meter, ASTM Skid Trailer, etc.

A good summarization for the different equipments used for measuring pavement condition is in table 3-0, excerpted from (Balta, 1984).

	ROUGHNESS	DEFLECTION	SKID RESISTANCE	DISTRESS
ALASKA	Mays Ride Meter	Falling Weight Deflectometer		Rutting: Measured Cracking, Patching: Visual
ALBERTA	PCA Car Roadmeter	Benkelman Beam and Dynaflect		Visual
ARIZONA	Mays Ride Meter	Dynaflect	Mu Meter	Visual/ Measured
CALIFORNIA	PCA Car Ridemeter	Dynaflect or Deflectometer	K.J, Law, Inc., Skid Tester	Visual/ Measured
CERL				Visual
DENMARK	Servo-accelerometer Mounted in Car	Falling Weight Deflectometer		
FLORIDA	Mays Ride Meter	Dynaflect	ASTM Skid Trailer	Visual/ Measured

Table 3-1: Equipments for Measurement of Pavement Condition

	ROUGHNESS	DEFLECTION	SKID RESISTANCE	DISTRESS
LOUISIANA	Mays Ride Meter			Visual
NEW SOUTH WALES	NAASRA Roughness Meter	Lacroix Deflectograph	SCRIM*	Visual
NEW YORK	DC Differential Transducer			Visual
OHIO	Mays Roadmeter	Dynaflect	K.J. Law, Inc., Skid Tester	Visual
ONTARIO	Car Ride Meter	Benkelman Beam and Dynaflect	Brake-Force Trailer	Visual
RHODE ISLAND+	Mays Ride Meter	Dynaflect	K.J. Law, Inc. Skid Tester	Visual
UTAH	Cox Roadmeter	Dynaflect	Mu Meter	Visual
WASHINGTON	PCA Roadmeter	Benkelman Beam	ASTM Skid Trailer	Visual

---

*Sideways force co-efficient routine investigation machine.

+Recommendations made to Rhode Island

References: 6, 7, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21.

Table 3-0, continued

3.2.3. Time Span of the Analysis :

By "time span" it is meant the period covered by the forecast of the *project-level*pavement maintenance system, while in *network-level* systems, another parameter is used, which is the *frequency of data collection* that shows us how frequently the system is updating its knowledge about the status of the network, in order to get its forecast as close as possible to the real behavior of pavement. Anyhow, it does not mean that the "frequency of data collection" parameter is inapplicable in the "project-level"systems, but up until now, it is not used. So, as one of the potentials of extending the capabilities of the network-level systems, there should be an augmented database (inventory) of the history of each highway in the user's jurisdiction. Such inventory (or case library) has several uses, as described in the implementation of HOTEP's case library, section 5.2.1. The decision process of the system should also be modified, such that, it would utilize this case database in reaching more adaptive, precise decision in a shorter time.

The frequency of data collection ranges generally between one year to five years. Arizona Department Of Transportation requires annual inspection of pavement conditions. While California Department Of Transportation (CALTRANS), and Washington State Department of Transportation require it every second year. Ontario Department Of Transportation collects data every three years.

For the current "project-level" pavement maintenance systems, the available tools of both mechanistic simulation and economical analysis extend the range of forecast effectively, according to the reliability of the mechanistic relationships. EAROMAR-II, as example, has the ability of long-term forecast, upon which the economical analysis is performed. From the point of view of time span, we find the same approach in Washington State Department of Transportation, California Department Of Transportation (CALTRANS), and Texas Rehabilitation and Maintenance Optimization System.

The long-term strategies, despite its appeal, is too theoritical and unrealistic. The precision of the prediction deteriorates rapidly with the time, therefore, it is better to base the economical analysis upon the following:

1. Reasonable frequency of pavement inventory survey, to ensure the representation of the real problems. As the pavement history (inventory) upon which we base our forecast includes implicitly the impact of the old strategies techniques which is mostly different from the present ones.
2. Focusing on short-term project strategies in the decision process, while keeping an eye upon the long-term project as a general guideline. This approach is followed by California Department Of Transportation (CALTRANS), and Texas Rehabilitation and Maintenance Optimization System, RAMOS.

3.2.4. Mathematical Procedures of Decision Process :

There are two different approaches in dealing with the evaluation of pavement status, these approaches are summarized hereafter.

3.2.4.1. Pavement Condition Index or Rating Number

In this approach, a scoring system is set to represent the status of pavement such that :

- Each deterioration symptom, with specific severity and extent, corresponds to some discount of the score.
- Whenever the pavement condition reaches certain score, a remedial action should be taken.

These indexes are either set by experiencial judgement in some systems, or by some measurements in the other systems. Look at table 3-1 (Balta, 1984).

There are two crucial disadvantages in the use of such indexes :

1. You may reach the same index number by two completely different cases of deterioration.
2. The absolute number does not give you any insights about the history of pavement or even the kind of deterioration.

Table 3-2: Pavement Condition Indices

NAME OF INDEX	INDEX EQUATION	INDEX TYPE*	CONDITION PARAMETERS ACCOUNTED FOR
ALASKA			
Surface Condition Index	$SCI = 1.38\ R^2 + 0.01\ (A+P)$	I	RUTTING (R), ALLIGATOR CRACKING (A), FULL WIDTH PATCHING (P)
Condition Value	$CV = \frac{\text{MAYS RIDEMETER SCORE} + SCI}{2}$	II	ROUGHNESS (Mays Meter), DISTRESS (SCI)
Composite Value Score	$CVS = (CV \times V\text{-}C\ \text{RATIO} \times \text{ACCIDENT RATING})^{1/3}$	II	ROUGHNESS AND DISTRESS (CV), TRAFFIC (Volume: Capacity Ratio), SAFETY (Accident Rating)
CERL			
Pavement Condition Index	$PCI = C - \sum_{i=1}^{P} \sum_{j=1}^{M_i} a(T_i, S_j, D_{ij}) \times F(t,q)$	I	DISTRESS TYPE (Ti) which includes LEVEL OF SEVERITY (Sj) and DENSITY OF DISTRESS (Dij) c = constant (usually 100); a = weighting value; a = f(T,S,D); F = adjustment factor for multiple distress types
DENMARK			
Present Serviceability Rating	$PSR = 5.0 - 1.9\ \lg t\ \left(1+K\left(\frac{RMS}{V}\right)^2\right)$	I	ROUGHNESS: SUSPENSION PROPERTIES OF MEASURING VEHICLE (K), VERTICAL ACCELERATION OF A PASSENGER AUTO (RMS), VELOCITY (V)
FLORIDA			
Ride Rating	$RR = a + bx$	I	ROUGHNESS FROM MAYS METER (X)
Defect Rating	$DR = 100 - \Sigma$ (defect points)	I	DISTRESS (DEFECT POINTS) WHICH INCLUDES CRACKING, RUTTING, AND PATCHING
Basic Rating	$BR = \sqrt{RR \times DR}$	II	ROUGHNESS (RR), DISTRESS (DR)

NAME OF INDEX	INDEX EQUATION	INDEX TYPE*	CONDITION PARAMETERS ACCOUNTED FOR
Adjusted Basic Rating	$SR = f(ADT,BR)$	II	ROUGHNESS AND DISTRESS (BR), TRAFFIC (ADT)
Engineering Rating	$ER = \sqrt{ORxSR}$	II	ROUGHNESS, DISTRESS, TRAFFIC (BR,ADT); ABILITY OF ROAD TO HANDLE TRAFFIC (OR)
LOUISIANA			
Relative Priority	$R.P. = 25\frac{\Sigma(S^2xD)}{875} + 15\frac{\Sigma(S^2xD)}{250} + 20[1-0.476(PSI)] + 20[0.0008(ADT)] + 5[1-0.02857(SN)]$	II	CRACK/RAVEL/PATCH (1st term), RUT/DISTORTION (2nd term), MEASURED DISTORTION (3rd term), ADT (4th term), SKID NUMBER (5th term)
NEW SOUTH WALES, AUSTRALIA			
Maintenance Index	MI = 0.25 (ROUGHNESS RATING) + 0.25 (VISUAL ASSESSMENT RATING) + 0.50 (DEFLECTION RATING	II	Self-explanatory
Safety Index	SI = 0.75 (SCRIM RATING) + 0.25 (ROUGHNESS RATING)	II	SCRIM is a measurement of skid resistance
NEW YORK STATE			
Present Rideability Index	PRI = f (ROUGHNESS METER DATA, SUBJECTIVE RATING)	I	ROUGHNESS
Pavement Surface Rating	PSR - Subjective, 0-10	I	DISTRESS
Base Rating	BR - Subjective, 0-10	I	STRUCTURAL
Maintenance Index	MI - Subjective, 0-10	I	INDICATION OF MAINTENANCE PERFORMED IN THE PAST YEAR

NAME OF INDEX	INDEX EQUATION	INDEX TYPE*	CONDITION PARAMETERS ACCOUNTED FOR
Structural Rating	$\text{STR.R.} = 3(\text{PSR}) + 4(\text{BR}) + 3(\text{MI})$	II	DISTRESS, STRUCTURAL ADEQUACY, MAINTENANCE PERFORMED DURING THE PAST YEAR
Sufficiency Rating	$\text{SUFF.R.} = \frac{(\text{V/C Ratio}) + (\text{STR.R.})}{2}$	II	DISTRESS, STRUCTURAL ADEQUACY, MAINTENANCE DONE LAST YEAR, TRAFFIC CAPABILITY
OHIO			
Present Serviceability Index	$\text{PSI} = 4.18 - 0.007(\text{RC})^{0.658} - 0.01\sqrt{C+P} - 1.34(\text{RD})^2$	II	ROUGHNESS (RC), DISTRESS (C,P,RD) - AASHO Equation
Skid Number	SN = from K.J. Law Skid Tester	I	SKID RESISTANCE
Pavement Condition Rating	$\text{PCR} = 100 - \sum_{i=1}^{n} \text{DEDUCT}_i$	I	DISTRESS, where DEDUCT_i = (Weight for Distress) x (Weight for Severity) x (Weight for Extent)
ONTARIO			
Pavement Condition Rating	PCR = Subjective	II	Subjective Assessment of ROUGHNESS and DISTRESS
Distress Manifestations	$\text{DM} = \sum_{i=1}^{27} C_i(S_i+d_i)$	I	DISTRESS, where C, S, and d are weighting values for type, severity, and density of distress, respectively
Distress Index	$\text{DI} = (0.1 \times \text{RCR})^{1/2} \times \frac{320-\text{DM}}{320}$	II	DISTRESS (as above), ROUGHNESS (RCR [measured])

NAME OF INDEX	INDEX EQUATION	INDEX TYPE*	CONDITION PARAMETERS ACCOUNTED FOR
UTAH			
Present Serviceability Index	$PSI = 4.18-0.007(RC)^{0.658} -0.01\sqrt{C+P} -1.34(RD)^2$	II	ROUGHNESS (RC), DISTRESS (C,P,RD): AASHO Equation
Distress Index	$DI = \frac{(2A + 2M + L + T)}{36}$	I	DISTRESS (ALLIGATOR (A), MAP (M), LONGITUDINAL (L), AND TRANSVERSE (T) CRACKING)
Structural Index	SI = Scale of 0-10 correlated with predicted years to failure from Dynaflect data	I	DEFLECTION
Final Index	$FI = 0.47 [F_1(PSI)^{1.5} + F_2(SI)^{1.5} + F_3(DI)^{1.5}]$	II	ROUGHNESS (PSI); DEFLECTION (SI); DISTRESS (PSI,DI); ADT, 18-Kip LOADS, FUNCTIONAL CLASS, SPEED LIMIT (Weights F_1, F_2, F_3)
WASHINGTON STATE			
Pavement Rating	$PAV'T\ RT. = 100 - \Sigma D$	I	DISTRESS (D = Weighted deduct points)
Ride Rating	$RIDE\ RT. = [1.0-0.3 (\frac{CPM}{5000})^2]$	I	ROUGHNESS (CPM)
Pavement Condition Rating	PCR = (PAV'T RT. x RIDE RT.)	II	DISTRESS, ROUGHNESS

---

*INDEX TYPES: Type I -- Index represents raw data for only one pavement condition parameter (e.g., DISTRESS, ROUGHNESS, DEFLECTION, SKID, etc.)

Type II -- Index represents a combination of more than one pavement condition parameter.

References: 6, 7, 9, 12, 13, 14, 15, 16, 18, 19.

3.2.4.2. Decision Trees, Condition States, and Condition Matrices

These are some techniques that form an alternative to the use of indexes. California Department Of Transportation (CALTRANS) uses decision trees, and Texas Rehabilitation and Maintenance Optimization System(RAMOS) uses the condition matrices, while Arizona Department Of Transportation(ADOT) uses an intermediate stage between the two techniques, which is the condition states.

The discussion of theses techniques is well beyond the scope of this research. Good documentation could be found in the bibliography of (Balta, 1984).

An interesting, study-worthy notice is the strong similarity between the Markovian Simulation of the damage process, proposed in (Findakly, 1971), and the Network Optimization System (NOS), developed for Arizona Department Of Transportation (ADOT) eleven years later (Golabi, 1983). Both of the systems depend on concept of *condition states* supported by a *transition probability matrix* and a *damage assessment matrix*. CERL also uses the probabilistic approach in the forecast of damage evolution, assuming a normal distribution of the occurance of each pattern of damage (or deterioration).

The pitfall of the use of these stochastic techniques is the assumption of some statistical distributions of the evolutionary life-times for the different distress manifestation. Such assumption restricts (localizes) the applicability of the system to a specific environment, as every enviro-structural combination has its own probabilistic distribution and intensity.

3.2.5. Pavement Mechanics- An epistemological exploration

In this part, an epistemological exploration is made for the kind of relationships used for the simulation of damage evolution in the pavement maintenance systems. The purpose of this part of research is gaining some insights about our current knowledge of Pavement Mechanics, esp. with respect to the following aspects:

1. The nature of the underlying assumptions.
2. The universality of the relationships, i.e., whether it is including all the possible parameters that may be thought of as relevant, and correlating them in the proper way.
3. The limitations of the relationships due to either the underlying assumptions, or pragmatic considerations of application.

This kind of study is called epistemological research. The benefits from such study are innumerable, as the study of the epistemological structure of any field should be a prerequisite foundation before conducting any research in this field, to ensure good understanding of the backgrounds of the problems met in this field.

Another benefit of studying the epistemological structure of the pavement mechanics, is the high potentiality of constructing better knowledge representation that may hasten the progress of research in this stagnant field. When the research in a field approaches a stagnation point, as it is the case in mavement mechanics, a breakthrough is required. Therefore a scrutinizing revision should be made for the basic foundations of the domains, that have been taken for granted. We do not need to prove that such breakthroughs are attained by unorthodox profound tackling of the deepest levels of knowledge of the field. This scrutiny is one of the motivations for this research.

To avoid being too theoritical, or impractical, we would take one of the currently-working pavement maintenance systems as an example in which we would examine the above-mentioned points. This system is EAROMAR-II.

3.2.5.1. Epistemological Remarks: Pragmatic Obstacles

The following is a preliminary epistemological study, and by no way could be considered a complete one. As a first-degree limitation, we will concentrate on the flexible pavement relationships. Indeed this down-scoping would not restrict the soundness of the drawn remarks, and conclusions, as they are related to some general research trends which are not restricted to flexible pavement. The major remarks are listed hereafter.

1. Universalization of localized relationships :
   Some of the relationships are seriously deficient, and *localized* as they are derived from data belonging to one specific area. This derivation means that the relationship is bound to the following:

   - The environmental zone of the research locale.
   - The kind of maintenance policies followed in this area.
   - The kind of materials and technologies used in this area.

   Some examples for those attempts to generalize some very localized formulae are mentioned hereafter:

   a. The relationships of *linear cracking* and *design temperature* are developed from data for Ontario and South Manitoba (pp.79). Of course, we can expect its performance in, say, Louisiana or Texas.

   b. The correlation between "roughness and serviceability"(pp.84) is based upon comparisons between ratings from Ontario and Minnesota which lie in the same environmental region.

   c. The *joint faulting equation*(pp.112) is derived from regression of data from Minnesota, Wisconsin, and North Dakota which also lie in the same climatic region.

2. The Statistical Soundness of Regression Analyses :
   The sample size from which the relationship is developed, is sometimes unreliably small. An example for that is the *linear cracking* equation (pp.79) which is based upon the regression analysis of only 32 pavement sections in Canada. Nothing is mentioned in the technical report of EAROMAR-II about the Standard Deviation of any regression analysis despite the big number of poorly-fitted curves all over the report.

3. Defective Simplification:
   Several equations are simplified, easily by giving a constant value to some parameters. The examples are too much but include :

   a. Fixing the type of subgrade (d) to *loam* in equation (24) (pp.80).

b. The simulation of all *linear cracking* occurs during the *coldest* design temperature, which is derived from the *freezing index*(pp.80). The question here is about the effectivness of the relationship in non-freezing regions such as California and New Mexico ?!

4. Homogeneity of the different underlying Assumptions :
The developers of EAROMAR-II apparently did not pay attention to this point, and therefore the technical report came out, as most of the literature about pavement mechanics, without any discussion about the underlying assumptions of the research.

5. Impracticality of more powerful tools:
A good example for this problem is the *viscoelastic model* (VESYS) and the *finite element approach* in pavement mechanics where their lack of closed-form type of equations entraps them from the widespread use.

3.2.5.2. The Role of AI

After the above-mentioned brief discussion of the current obstacles to implement EAROMAR-II as a universal pavement maintenance system, we will discuss what artificial intelligence can offer towards the solution of these problems. It is very obvious, from the above-mentioned discussion, that the reliability of the mechanistic part of EAROMAR-II is the major impedence to its success. EAROMAR-II has very powerful economical and traffic tools but the malfunction lies in the most crucial part, the mechanistic model. To get rid of this malfunction, we have to revise the damage equations in the mechanistic part. The area of *learning from observations* (or learning by discovery) offers big promises for tackling this problem, as it enables us to learn in domains, where the experiments are conducted continuously and the learning program should analyze the raw results of these experiments, and then draw its conclusions and rules. The program discovers the correlated parameters, and what kind of relationship they may have, with no ,or minimal, assistance from the outside. So, we may gain some unprecedented insights, and also we may discover some relationships we did not think about before. Several interesting applications of different methods of learning are found in the Machine Learning book (Michalski, 1984). By applying these Learning techniques, we can improve the relationships of damage evolution in the following way:

1. Instead of starting from the vacuum, as all other Learning from observations systems do, we would take the current relationships as a preliminary basis. After running the

system through some cases, we may find some new parameters added to the relationship, some removed, some numbers changed, or the whole form changed.

2. This learning from observations (or quantitative discovery) could be controlled to the desired levels of reliability by setting all our specifications, such as, the minimum sample size, maximum standard deviation, and the prefered forms of relationships.
3. By including these learning capabilities in the Pavement Maintenance System, the system could be implemented in any new region without any adaptational effort, as it would start modifying its knowledge base automatically as it encounters new cases.

Chapter 4

Failure Mechanisms

In this chapter we are focusing upon three points:

1. The underlying assumptions (hypotheses) of the mechanistic behavior of flexible pavement, and how they are affecting the performance of the pavement maintenance systems.
2. The failure mechanisms of flexible pavements, their causes, multiplication, and interrelationships.
3. Finally, describing the main features of the proposed *representation network*, which will include all the above-mentioned failure mechanisms. This network is a major element of the HOTEP system, refer to chapter 5, as it represents the expandable knowledge base of the system.

4.1. introduction

The pavement design process is one of the most interesting fields to watch the impact of lacking knowledge upon the design methodology. This lack of knowledge concentrates in the area of pavement mechanics, where we do not know the exact behavior of pavement in the various mechanical aspects. There are *three axioms for the rational understanding of pavement*, these axioms have to be fulfilled before claiming the success of any design (or analysis) technique:

1. Establishing (or adopting) a theory to predict the failure or distress parameters.
2. Evaluation of pertinent material properties necessary for the theory selected.
3. Determination of the relationship between the magnitude of the parameter in question to the failure or performance level desired.

Despite of the incomplete fulfillment of the above-mentioned axioms, the only *two approaches* in the pavement design are:

1. Theoritical design of pavement, in which a hypothesized behavior of pavement is utilized in the design and proportioning of pavement.

2. Performance-based design, where also hypothesized relationships of distress development are used in the design of pavement with a specific level of performance.

The best comment on both of the approaches comes from the renowned, authoritative expert in pavement mechanics, E. Yoder as he says (Yoder, 1975):

> Neither of [performance-based design nor theoritical design of pavement] is satisfactory within itself. Complete reliance upon pavement performance represents a static condition wherein *one must wait a relatively long period before new concepts can be proven out*. On the other hand, theoritical equations are generally based upon *simplified assumptions* and many times do not apply to conditions as they exist in the field.

Then, in another place, he concludes:

> In the strictest sense, there is at present no truly fundamental or rational design procedure that is widely accepted in the pavement design industry.

4.2. The basic mechanics of multilayered systems

The common feature of all the design techniques, developed up until now, is the consideration of the highway as an *elastic multilayered system*, consequently, the following assumptions are the foundation for almost all the design techniques:

1. **Homogeneous material properties:** This assumption could not be considered realistic but only for the mechanistic behavior of subgrades with modular ratio (of pavement virsus subgrade) close to unity. This case is exemplified by flexible granular base/subbase pavement structure having thin asphalt concrete surface course.

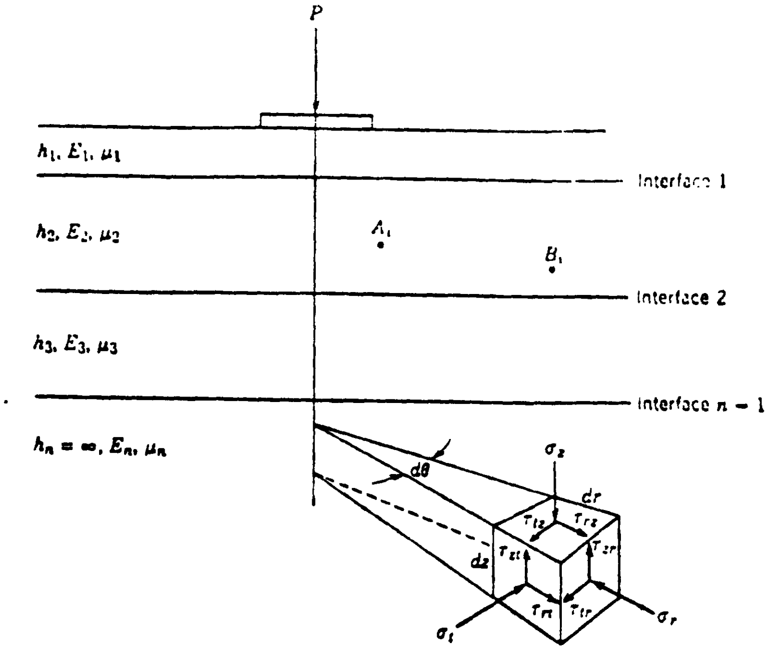
2. Each layer has a **finite thickness** except the lower one, and all are infinite in lateral directions.

3. Each layer is **isotropic.**

4. **Full friction is developed between layers** at each interface.

5. **Surface shearing forces are not present at the surface:** This assumption and the previous one are overriden by some of the modern computerized design procedures such as the multilayered shell, and BISAR system developed by Shell.

The type of theory used is generally destinguished by reference to three properties of the material behavior response:



Generalized multilayered elastic system.

Figure 4-1: The generalized multilayered elastic system.

1. The relationship between stress and strain [Linear or Nonlinear] (figure 4-2.a). The relationship could be expressed in terms of the following three equations:

$$\varepsilon_z = \frac{1}{E}[\sigma_z - \mu(\sigma_r + \sigma_t)]$$

$$\varepsilon_r = \frac{1}{E}[\sigma_r - \mu(\sigma_t + \sigma_z)]$$

$$\varepsilon_t = \frac{1}{E}[\sigma_t - \mu(\sigma_r + \sigma_z)]$$

Where E = Young's Modulus,
μ = Poisson's ratio,
σ's = Stresses,
{refer to figure 4-1},
ε's = Strains,
{refer to figure 4-1}.

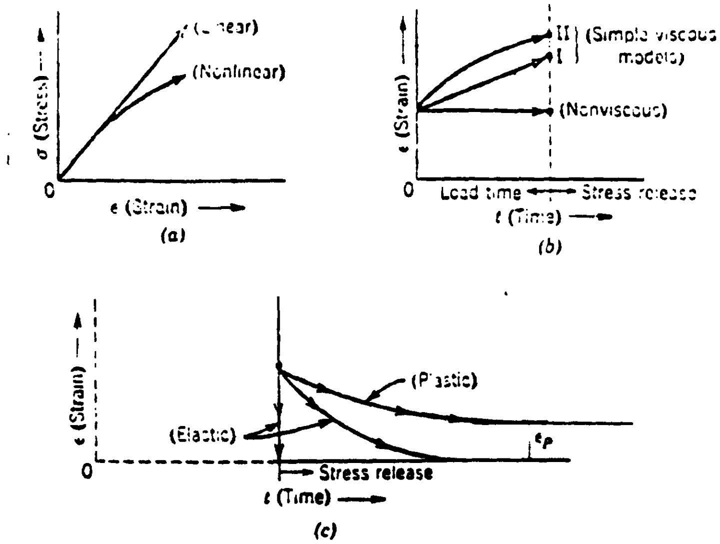
2. The time dependency of strain under a constant stress' (load') level [Viscous or Non-viscous] (figure 4-2.b).

- **Viscous :**

a. Maxwell-Type :

$$\varepsilon(t) = \frac{\sigma}{E} - \frac{\sigma t}{\eta}$$

A spring & a dashpot in series.



Material characteristics. (a) linearity; (b) viscous effects; (c) recoverable effects.

Figure 4-2: Material Characteristics.

b. Kelvin-Type :

$$\varepsilon(t)=\frac{\sigma}{E}(1-e^{-Et/\eta})$$

A spring & a dashpot in parallel.

Different combinations of the above-mentioned simple types are used sometimes.

- **Non-viscous :**

t = Constant

where,
t = Time,

η = Coefficient of viscosity of the dashpot system.

3. The degree to which the material can rebound or recover strain after stress removal (plastic or elastic) {figure 4-2-c }.

From the practical point of view, most of the pavement maintenance systems use *linear viscoelastic theory* to simulate the pavement mechanics. To override the dichotemy of "plastic virsus elastic" recovery of strain, these systems assume an elastic behavior but with a very long span of recovery time, which turns the system into a quasi-plastic one.

4.3. Multilayered solutions

4.3.1. One-layer Systems :

It is based upon Boussinesq's equations originally developed for an homogeneous, isotropic, and elastic medium due to a point load at the surface, refer to table 4-1. The central (core) formula in this technique is the one used to determine the vertical stress at any depth below the surface due to a point load at the surface.

$$\sigma_z = k\frac{P}{z^2}$$

$$k = \frac{3}{2\pi}\frac{1}{[1 + (r/z)^2]^{5/2}}$$

where:

r = distance radially from point load, and z = depth.

An important assumption in this technique is that the pavement portion (above the subgrade) does not contribute any partial deflection to the total surface deflection. Thus the deflection occurs in the subgrade from its topmost layer to infinity.

<u>Disadvantages:</u>

1. The calculated stresses are much higher than the measured ones.
2. For the same stress, the calculated deflection is higher than the measured one.
3. The impact of the strength of the reinforcing layers upon the dispersion of stresses is completely neglected.

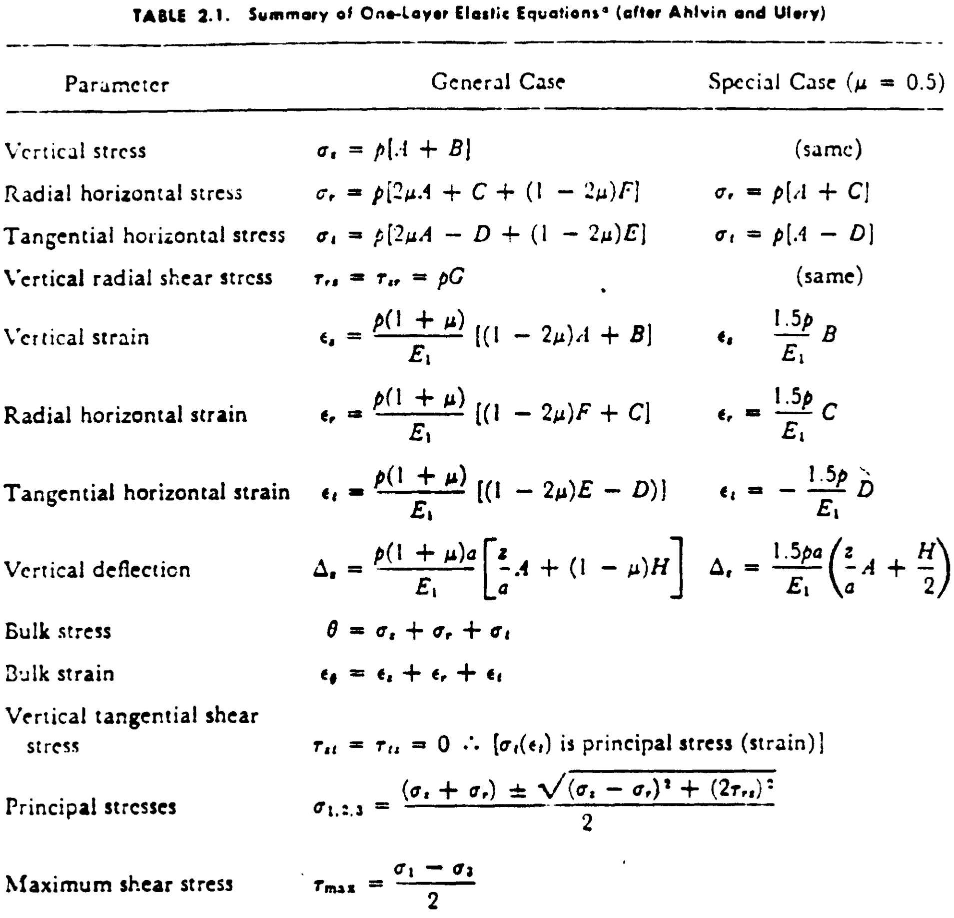


TABLE 2.1. Summary of One-Layer Elastic Equations[a] (after Ahlvin and Ulery)

Parameter	General Case	Special Case ($\mu = 0.5$)
Vertical stress	$\sigma_z = p[A + B]$	(same)
Radial horizontal stress	$\sigma_r = p[2\mu A + C + (1 - 2\mu)F]$	$\sigma_r = p[A + C]$
Tangential horizontal stress	$\sigma_t = p[2\mu A - D + (1 - 2\mu)E]$	$\sigma_t = p[A - D]$
Vertical radial shear stress	$\tau_{rz} = \tau_{zr} = pG$	(same)
Vertical strain	$\epsilon_z = \frac{p(1 + \mu)}{E_1}[(1 - 2\mu)A + B]$	$\epsilon_z \quad \frac{1.5p}{E_1} B$
Radial horizontal strain	$\epsilon_r = \frac{p(1 + \mu)}{E_1}[(1 - 2\mu)F + C]$	$\epsilon_r = \frac{1.5p}{E_1} C$
Tangential horizontal strain	$\epsilon_t = \frac{p(1 + \mu)}{E_1}[(1 - 2\mu)E - D))]$	$\epsilon_t = - \frac{1.5p}{E_1} D$
Vertical deflection	$\Delta_z = \frac{p(1 + \mu)a}{E_1}\left[\frac{z}{a}A + (1 - \mu)H\right]$	$\Delta_z = \frac{1.5pa}{E_1}\left(\frac{z}{a}A + \frac{H}{2}\right)$
Bulk stress	$\theta = \sigma_z + \sigma_r + \sigma_t$	
Bulk strain	$\epsilon_\theta = \epsilon_z + \epsilon_r + \epsilon_t$	
Vertical tangential shear stress	$\tau_{zt} = \tau_{tz} = 0 \therefore [\sigma_t(\epsilon_t)$ is principal stress (strain)$]$	
Principal stresses	$\sigma_{1,2,3} = \frac{(\sigma_z + \sigma_r) \pm \sqrt{(\sigma_z - \sigma_r)^2 + (2\tau_{rz})^2}}{2}$	
Maximum shear stress	$\tau_{max} = \frac{\sigma_1 - \sigma_3}{2}$	

Table 4-1: Summary of One-Layer Elastic Equations.

4.3.2. Two-layers Systems

Typical flexible pavements are composed of layers so that the modulus of elasticity decreases with the depth, to reduce the stresses and deflections in the subgrade from those obtained in the ideal homogeneous case. This technique which was developed by Burmister is based on the same assumptions mentioned in the beginning of this chapter. The technique depends upon the strength ratio of the two layers (E_1/E_2) as shown in figure 4-3 in which we can compare Boussinesq's technique ($E_1/E_2 = 1$) to more realistic ratios.

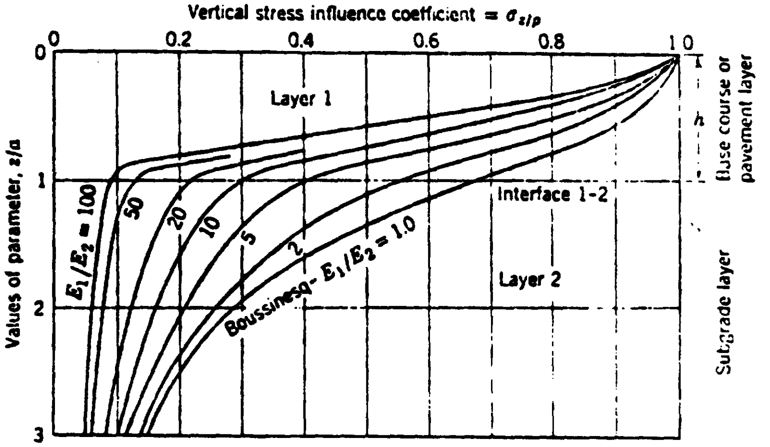


Figure 2.6. Basic pattern of Burmister two-layer stress influence curves. (From Burmister, Highway Research Board Bulletin 177.)

Figure 4-3: Basic Pattern of Burmister two-layer stress influence curves.

4.3.3. Three-layers Systems

Huang expanded Burmister's equations to deal with three-layers systems to gain more reality and expressiveness by using the *interface deflection factor*, which is illustrated in several charts in any pavement design handbook.

Another technique for three-layers systems was developed by Acum and Fox. This technique produced the first extensive tabular summary of normal and radial stresses in three-layers systems at the intersection of the plate axis with the layer interfaces.

4.4. Fundamental Design Concepts

4.4.1. Subgrade Stress

Figure 4-4 reveals the following facts :

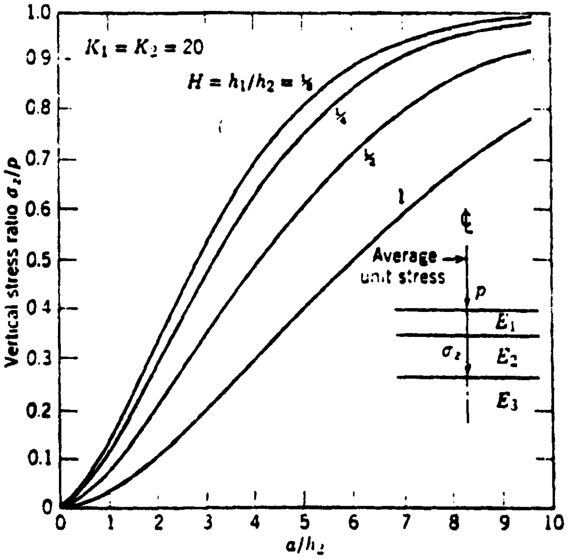


Figure 2.11. Vertical stress ratio at the second interface of a three-layer system as a function of a/h_2 and h_1/h_2. (From Nielson.)

Figure 4-4: Subgrade Stresses

1. The vertical compressive stress ratio,VCSR, (σ_z/p)for three-layer pavement system is proportional to A ($=a/h_2$).

   $\sigma_z/p \;\alpha\; a/h_2$

   i.e.,for a given load (contact radius a), a decrease in vertical compressive stress ratio could be attained by an increase in the base-course thickness.

2. For constant (h_2):

   $$VCSR \;\alpha\; \frac{1}{h_1}$$

   i.e.,for a given load and a constant base-course thickness, the stress level can be reduced by increasing the surface course thickness.

3. $VCSR \;\alpha\; \frac{1}{k_1}$

   $\cdot k_1 = E_1/E_2$

   $$VCSR \;\alpha\; \frac{1}{k_2}$$

   $\cdot k_2 = E_2/E_3$

   i.e., in order to reduce the vertical compressive subgrade stress, we have to design the pavement such that the higher we go from the subgrade, the higher is the strength of these layers.

4.4.2. Deflection

Most of the deflection of pavement is contributed by the subgrade (70-95%). As far as the deflection is the integration of vertical srains over depth, a good way to reduce the subgrade deflection is by reducing the subgrade compressivestress. This goal could be attained by either :

- Increasing the thickness of layers above subgrade.
- Increasing the strength of layersabove subgrade, which is more effective than the previous one.

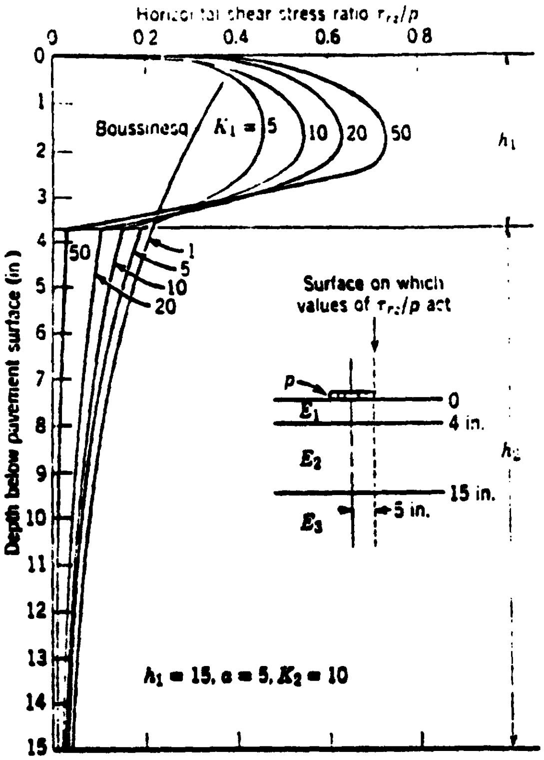
4.4.3. Shear Stress

The increase of rigidity or thickness of the layers above subgrade is good for both the subgrade stress and the deflection. But this goodness is on the expense of the *shear stress.*

The maximum horizontal shear stress occurs at the mid depth of the surface course, and it shifts to the point of the upper third as (h_1) increases. From figure 4-5, we can see also how effective is the increase of the surface thickness in decreasing the horizontal shear stress.

4.4.4. Tensile Stress

If the relative expensiveness of asphalt persuades the designers to increase the thickness or rigidity of the base-course instead of the surface-course, the horizontal tensile stresses would stop them from reaching very thin membranes of asphalt as we can see the proportional relationship between the horizontal tensile stress and (h_1/h_2) in figure 4-6.



Typical distribution of shearing stresses in a three-layered system. (From Nielson.)

Figure 4-5: Impact of Shear Stress

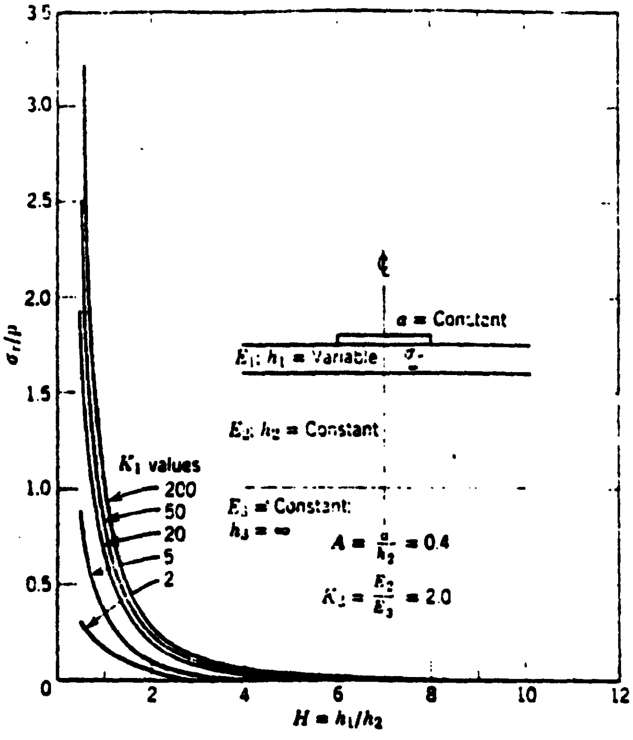
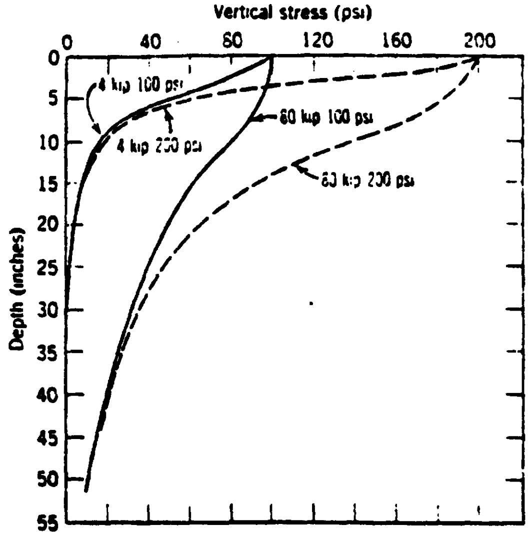


Figure 2.14. Horizontal tensile stress ratio at the bottom of the first layer of a three-layered system as a function of K_1 and H.

Figure 4-6: Impact of thickness proportioning upon Horizontal Tensile Stresses

4.4.5. Effect of Tire Pressure and Total Load

Both of tire pressure and total load have vital impact upon the pavement as shown in figure 4-7. The impact of tire pressure vanishes almost at a depth of 35 inches. The impact of dual wheels is very slight and negligible.



Variation of vertical stress with depth, Boussinesq problem.

Figure 4-7: Effect of Tire Pressure & Total Load

4.5. Pavement Rehabilitation

The term *Pavement Rehabilitation* covers three major areas:

1. Failure Mechanisms: In which a trial is done to define what constitutes a failure, how it is developed, and whether the distress is progressive or non-progressive. This area is very central to the topic of this research.

2. Methods of Evaluation : This area is out of the scope of this research. Good theoritical basis is found in [Witczak,74], and good applications'survey is in (Balta, 1984).

3. Methods of Overlay design and reconstruction : To have an effective method of overlay design, we should establish a good understanding of failure mechanisms. Therefore I would not tackle this area but only after gaining some substantial insights in the failure mechanisms through this research. The only part to be considered at

this stage of research is the tentative remedy of each distress category as is believed in the current level of knowledge. These tentative remedies will be annexed after the discussion of each category of failure mechanisms.

4.5.1. Failure Mechanisms

There are two kinds of failure:

1. Functional Failure The functional failure denotes the status of pavement at which it can no longer carry out its intended function and it depends primarily upon the degree of surface roughness.

2. Structural Failure This kind of failure indicates the breakdown of one or more of the pavement components. The structural failure is the failure category that needs a considerable research effort, and that is why it is the focal point of this research. One specialization will be considered, that is the emphasis upon the flexible pavement. The different causes of structural failure in flexible pavement are :

   - Surface Fatigue.

   - Consolidation.

   - Shear in either of the following: subgrade, subbase, base, or surface.

4.5.2. Description of Various Failure Mechanisms

The next facts would be expressed in the form of brief points as a preparatory stage for translating them into production rules in a knowledge-based program.

4.5.2.1. General Notes :

1. If we have rutting, combined with upheaval , and the distance between them is bigger than $0.75\,[H_1 + H_2]$, then it is subgrade shear failure.

2. If we have rutting, upheaval, and the distance between them is smaller than H_1, then it is surface shear.

3. If we have rutting, and upheaval other than the above-mentioned two cases, then it is base shear.

4. Consolidation (or settlement) can *not* coexist with upheaval.

4.5.2.2. Bleeding :

1. It may be caused by :
    - Excess bitumen in the mix.
    - Excessively-high (bitumen/blotter) ratio in the mix.
    - In joints, excessive amounts of asphalt cement (or tar).
    - Too soft asphalt for the climate.
    - Consolidation of the surface (Yoder, 1975), refer to [4.5.2.22].
2. It is traffic-related and occurs in the wheel tracks.

4.5.2.3. Bump :

1. It is a kind of :
   Upheaval, refer to [4.5.2.1].
2. It may be caused by :
    - Frost heave,refer to [4.5.2.28].
    - Upward thrust of ice forming under the pavement.
    - Evolution of shoving, refer to [4.5.2.31].
    - Swelling (vertical movement) of an underlayer.

4.5.2.4. Cracking, Alligator

1. It may be caused by :
    - Evolution of block cracking, refer to [4.5.2.5].
    - Excessive vertical movement of one or more of the underlayers (Yoder, 1975),and/or
    - Surface fatigue (Yoder, 1975), refer to [4.5.2.25].

2. It is mostly progressive.
3. It may cause :
   Potholes, refer to [4.5.2.19].
4. If it is observed with potholes, the pavement is structurally inadequate.

4.5.2.5. Cracking, Block

1. It may be caused by :
   - Shrinkage of surface, refer to [4.5.2.29].
   - Excessive vertical movement of underlayers ,especially the subgrade.
2. It may cause :
   Alligator cracking in the presence of a moisture source.

4.5.2.6. Cracking, Contraction (Shrinkage)

It may be caused by :

- Shrinkage of surface.
- Weak prime or tack coat.
- Loss of flexibility of the surface-course.

4.5.2.7. Cracking, Edge

1. It may be caused by :
    - Insufficient thickness of surface.
    - In case of embankment, either of the following:
        - Heavy traffic with either settlement or lateral displacement of the embankment.
        - In case of no traffic, Both of settlement and lateral diplacement of embankment.

4.5.2.8. Cracking, Longitudinal

1. It may be caused by :
    - Horizontal movements.
    - Sher failure in the subgrade IF the cracks are outside the tire cracks.
    - Cold or improperly-constructed joints between pavement sections.
    - Reflection Cracking ,refer to [4.5.2.9].
    - Subgrade settlement under heavy traffic, refer to [4.5.2.22].
    - Surface shrinkage.
    - Insufficient surface thickness.
2. It may exceed 1/4 inch, permitting surface infilteration.
3. It includes :
    - Edge Cracking, refer to [4.5.2.7].

4.5.2.9. Cracking, Reflection

1. It may be caused by :
    - Horizontal movement of underlayers, refer to [4.5.2.18].
    - In case of overlayed highways, lack of bridging over underlying cracks or joints due to the following:
        - insufficient surface thickness, <u>and/or</u>
        - Evolution of a crack in the underlying surface-course.
2. It may cause :
   Surface infilteration.

4.5.2.10. Cracking, Slippage

It may be caused by :

- Braking/ Starting wheel'thrusts, associated with either of the following:
    - Evolution of corrugation, refer to [4.5.2.12].
    - Evolution of shoving, refer to [4.5.2.31].
- Unstable sand mix surface-course [quality mistake by either designer or contractor] (HRB, 1970).

4.5.2.11. Cracking, Transversal

1. It may be caused by :
    - Shrinkage of asphalt, refer to [4.5.2.6].
    - Shrinkage of an underlayer, which is a type of horizontal movement, refer to [4.5.2.18].
    - Insufficient surface-course thickness.
    - Reflection cracking, refer to [4.5.2.9] if it is in an overlayed highway.

4.5.2.12. Corrugation

1. It may be caused by :
    - High traffic, with either of the following:
        - o Soft Surface.
        - o unstable underlayer.
    - Shoving, refer to [4.5.2.31].
    - Weak prime coat.
    - Very soft surface-course.
2. It may cause :
   Slippage cracking, refer to [4.5.2.10].

4.5.2.13. Depression

It may be caused by :

- Subsidence(Settlement) of the subgrade, refer to [4.5.2.22].
- In case of fill sections, improperly-compacted fill.

4.5.2.14. Imprint

It may be caused by :

- Very soft asphalt.

4.5.2.15. Indentation

It may be caused by :

- Sharp objects dragged over the surface.

4.5.2.16. Movement of Underlayers :

1. It includes :
    - Vertical movement, refer to [4.5.2.18].
    - Horizontal movement, refer to [4.5.2.17].

4.5.2.17. Movement of underlayers, Vertical

1. It may be caused by :
    - Volume change in an underlayer, in which the case should be conjectured with horizontal movement, refer to [4.5.2.18] (Yoder, 1975).
2. It includes :
    - Settlement.
    - Swelling.

4.5.2.18. Movement of Underlayers, Horizontal

1. It is a kind of :
   Movements of Underlayers [4.5.2.16].

2. It may be caused by :

   - Lack of internal friction in an underlayer, e.g., The clayey subgrades. Refer to Coulomb's Law: $s = c + \sigma \tan\varphi$.

   - Frost Heave, refer to topic, refer to [4.5.2.28].

   - Volume change in the subgrade soil, in which the case should be conjectured with vertical movement [4.5.2.17] (Yoder, 1975).

   - Fill settlement.

   - Sliding of side slopes, refer to [4.5.2.30].

4.5.2.19. Potholes :

1. It may be caused by :

   - Poor bituminous concrete, i.e., low(Bitumen/Blotter) ratio.

   - A development of *alligator cracking*, in this case there should be some alligator cracks around the potholes.

   - Traffic over localized weakened spots on the surface.

4.5.2.20. Ravelling

It may be caused by :

- Uneven distribution of bitumen from the spray bar.

- Low (Bitumen/Blotter) ratio, in which case it can not coexist with Bleeding, refer to [4.5.2.2].

- Traffic on a weak surface.

4.5.2.21. Rutting :

1. It may be caused by :
    - Consolidation of several layers of the highway (Yoder, 1975), refer to [4.5.2.22].
    - Localized and channeled wheel traffic over unstable pavement-course or foundation (HRB, 1970).
2. Rutting width equals, Approximately, the depth of the failed layer (Yoder, 1975).

4.5.2.22. Settlement or Consolidation

1. It may be caused by : Weak base-course or resilient subgrade.
2. Improperly-compacted underlayer.
3. The combined action of the following:
    - Heavy axle loads, and
    - Insufficient thicknesses (Design defect)

   i.e.,either ($ESAL/h_1$) or ($ESAL/h_2$) is higher than the permissible limits. In this case, consolidation takes place in the subgrade mainly, to lesser extent in the base-course, and seldomly, in the surface course (despite it is mentioned in (Yoder, 1975) as a possible cause of bleeding,4.5.2.2!).

4.5.2.23. Shear Failures :

1. It may be caused by :

    - Lack of cohesion and internal friction in the pavement'base structure.

   It includes *edge failures*, refer to [4.5.2.7] when the shear failure develops in the subgrade.

4.5.2.24. Streaking

It may be caused by :

- Construction mistake: uneven distribution of bitumen from the spray bar.

4.5.2.25. Surface Fatigue :

1. It may be caused by :
   Excessive vertical movements of one or more of the underlayers (Yoder, 1975), refer to (Movement of Underlayers), refer to [4.5.2.16].

2. It includes :

    - Embrittlement of surface over resilient subgrade, which is a very common cause, 85% .

4.5.2.26. Swelling :

1. It may be caused by :

    - High content of clay in the underlayer, <u>and</u>

    - moisture source.

2. It may cause :
   Bump, refer to [4.5.2.3].

4.5.2.27. Weathering

It may be caused by :

- Drying out (or loss) of bitumen due to climatic conditions.

4.5.2.28. Frost Heave :

1. If it is localized, the cause is a local one, such as silt pocket.
2. Refer to the section of the *environmental impact*.

4.5.2.29. Shrinkage of Surface

It may be caused by :

- Badly constructed suface, due to either:
  - Low (Bitumen/aggregate) ratio.
  - The construction took place at a very hot weather.

4.5.2.30. Sliding of Side Slopes :

It may be caused by :

- Shear failure.
- Unstabilized slopes close to wetness source.

4.5.2.31. Shoving :

It may be caused by :

- Too soft surface course to resist the horizontal pressure, and
- Weak prime coat.

4.5.2.32. Waves

1. It is due to bad subgrade soils.
2. It develops *gentle, stretched swells*.
3. It develops some Shear Failure especially in semicircular or longitudinal cracks which is the first stage of the development.

4.5.2.33. Adverse Ground-water Conditions

1. It causes subgrade weakening.
2. The remedy is the removal of the water source.
3. Minor resurfacing would never help.
4. A drainage system is a must.

4.5.3. Example of utilizing the network

For a diagramatic configuration of part of the above-mentioned relationships, refer to 4-7.

An illustrative example to show us how the implementation of <u>Looking-ahead</u> feature in the network search slashes down the required data acquisition, is given in 5-1. This feature is a direct result of the polymorphic construction the representation network, refer to chapter 5.2.2.

style(LineWidth 6.2 inches, LeftMargin 1 in, Spacing 2)

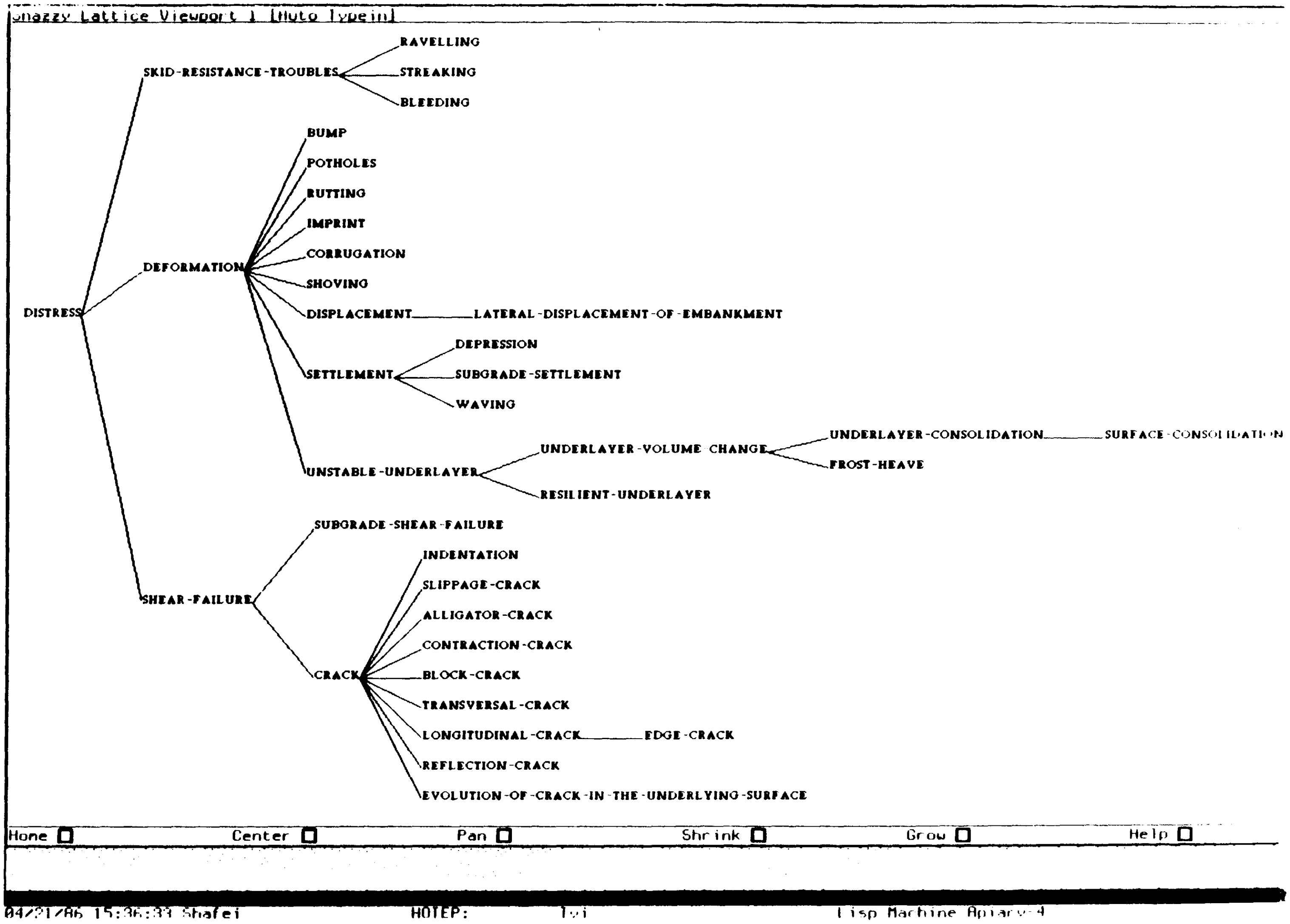


Figure 4-8: Part of the Representation Network

Chapter 5

System Configuration

In this chapter we will configure the basic features of the proposed system and explain how they could interact conceptually with each other. Such interaction is explained in the system architecture 5-0. The terminology used in the system architecture is thoroughly explained in the subsequent parts of this chapter.

5.1. System Architecture

The system architecture of HOTEP 5-0 is the consolidation of the various objectives mentioned in chapter 2.

Each task (stage, or sub-system) represents a macro function corresponding to a similar one performed by the researchers/experts when they attempt to solve a similar problem. The System is written in both LISP (for the logical part) and Pascal (for the numerical part). The Quantitative Discovery System, QUDS, and a initial knowledge base, PMORPH, have been already developed in a prototype.For more details about HOTEP refer to (El-Shafei, 1986a). For description of the various macros in the system refer to the next section.

5.2. Task Description

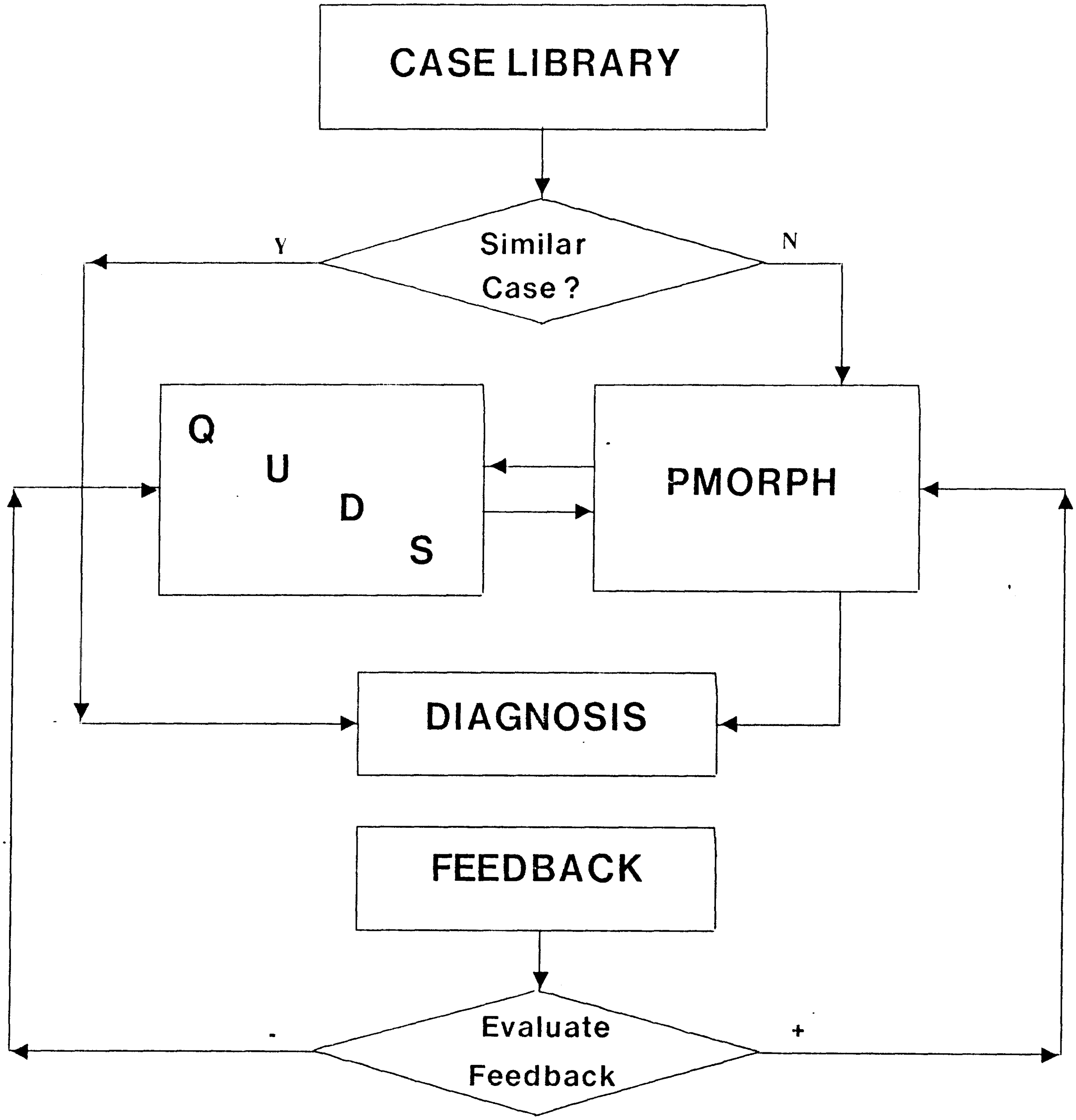


Figure 5-1: The System Architecture

5.2.1. Input Cases and the Case Library

The input case should include a descripton of the various structural, environmental, and traffic properties of the pavement under consideration, it should also include a detailed description of the distress manifestations and their current level of deterioration.

The system searches the "Case Library" for a similar case through a "Threshold Matching System". If a matching case is found the same diagnosis will be given. Thus considerable computational resources are saved.

The "Case Library" is built from all the cases that the system has ever encountered. Each stored case is augmented by its prescribed diagnosis, and a credit assigned for the reasoning path that led to that diagnosis.

5.2.2. The Polymorphic Representation Network, PMORPH

PMORPH is the knowledge base which contains all that HOTEP knows about the domain. The knowledge of the failure mechanisms in pavement is envisaged in terms of distinct physical states. Each state has its own characteristics and its own relations with the other states (El-Shafei, 1986a). The discrete nature of the physical states makes the Frames (Minsky, 1975) the most appropriate representation (Pauker, 1976). The first proposal to represent the mechanistic behavior of pavement as a State Space was made by (Findakly, 1971) to assign stochastic weights for any state transition. A statistical implementation was carried out by (Golabi, 1983) for Arizona Dept.of Transportation. Our approach is completely different. We concentrate on the causal and taxonomical semantics of this State Space. Such approach enhances our understanding even within the same level of knowledge, not to mention freeing our knowledge from any repitition (or redundancy). Following is a typical frame used in PMORPH :

A simplified Example of a file in the database .

case:	Route 128, Milepost 123.00->128.00		
Type of pavement:	Flexible		
AADT:	15,000 veh.per lane		
% Trucks :	24 %	Ave.Speed :	49 mph
Lane width:	12'	Subgrade Texture :	Gravel(WG)
# of lanes:	3	Dry Density :	135 pcf
Shoulder width:	8'	Corrected CBR (%) :	86,37,5 %
Pavement thickness:	8"	Surface Stiffness :	120 ksi
Base thickness:	15"	Base Elas.Modulus :	20 ksi
CBR:	3.0	Subgrade Modulus :	1500 psi
Temprature variation:	Summer 85, Fall 60, Winter 30,Spring 50		

Symptoms	Severity
Shoving	1-3"
Transv.Cracks	0.5-2."
Settlement	1-2"
Shldr.Embank.Washaway	Med.

Diagnosis:

A- Surface is too soft under the given traffic to resist the horizontal pressure and it has a poor bond with base, and this may evolve into "Shoving".

B- Surface thickness is inadequate, and this may evolve into "Transversal-Cracking".

C- There are Cracks, and this makes surface pervious. It is a rainy region. Surface layer is pervious and penetratable by rains, and this may evolve into "Base-Washaway" and "Shldr-Embank-Washaway" at some spots.

Remedial Action:

A- Construct a 3" overlay after applying a good amount of tack coat.

B- Reconstruct both the base and the surface after 2 years. Until that time apply the usual maintenance policies.

Follow-up Result: Positive.

Reliability**: B

A typical frame in PMORPH:

STATE: longitudinal-Cracking

STATE-TYPE: Symptom

CLASS: Structural

IS-A-KIND-OF: Cracks

REPRESENTATIVE-VAR: LC

VARIABLE-DIMENSIONS: 1(1)

SEVERITY-THRESHOLD: 2 ft

ASSOCIATED-SYMPTOMS: nil

NECESSARY-SYMPTOMS: nil

CONTRADICTS-WITH: nil

MAYBE-CAUSED-BY: and(Block-Cracking, or(Water-percolation,
Resilient-Subgrade))
(Excessive-vl-mvmnt-of-underlayer)
Surface-fatigue

MAY-CAUSE: Potholes

MAYBE-RELATED-TO: Surface-Stiffness(S), Base-Elas-Modulus(E),
Subgrade-Modulus(SM), Freezing-Index(FI),
EQUIV.THICKNESS(H), Asphalt-Age(A),
Ave.Annual Daily traffic(AADT), SURFACE-HEAT(T),
California Bearing Ratios(CBR's),
Heat-Capacity(C), ACCUM-LOAD/YR(X)

5.2.2.1. The state Frames

To understand the nature of a state we may pay a closer look at the components of a frame. The TYPES of states represented by frames are either *intermediate* states that are used within the diagnostic process after checking their plausibility (through associated symptoms), *computational* states where some synthesized parameters are calculated[8], or *symptoms* that are observed and need no proof other than their existence. Frames are of different CLASSes. Most

---

[8] The computational states are embodyment of some of the assumed concepts and hypotheses. They could be synthesized by the system automatically if a constellation of variables is observed frequently in the same configuration. For examples of computational states, refer to A.6.7. By the use of these computational states [daemons] we save the computational resources required for the computation till we need to evaluate this daemon. Once the daemonic state is evaluated, its value is associated with it not to be evaluated twice (Winston, 1984).

of them are *Structural* (physical) states of pavement and some of them represent *Environmental* or *Traffic* condtions that influence other states. Each state (such as *Alligatoring*) could be A-KIND-OF a more general category (*Cracks* in this case). Most of the states have their own characteristically REPRESENTATIVE-VARIABLEs which, in turn, has their own dimensions. The global dimensions are: mass, length, time, temperature, and charge. Every state, as well, has its ASSOCIATED-SYMPTOMS which could be reduced to a smaller set of NECESSARY-SYMPTOMS. For the last two fields, we do not mean symptoms literally. They may rather be intermidiate states, equation, or unequality. A state may also CONTRADICTS-WITH some other states. Finally, each state MAYBE-CASED-BY some other states and they also MAY-CAUSE some other states. The causality relationships are coupled with a precedence score. As a support of genuine discovery, the causally-related states are not the only ones to be kept under scrutiny. The otherwise suspected states collectively with the causally-related ones -all represented by their variables- are included in a MAYBE-RELATED-TO field. If the state has been fired once before, the variables should have been clustered into dimensionally-homogeneous groups in a previous session. Thus we save the computational resources needed for Dimensional Analysis. This feature is an implementation of the concept of K-lines of (Minsky, 1986).

The K-line is a wire-like agent which connects whatever mental agents are active when you solve a problem or have a ggod idea. In our case, we replace the mental agents by physical states and the wire-like agent by a dimensionless and/or related group of variables.

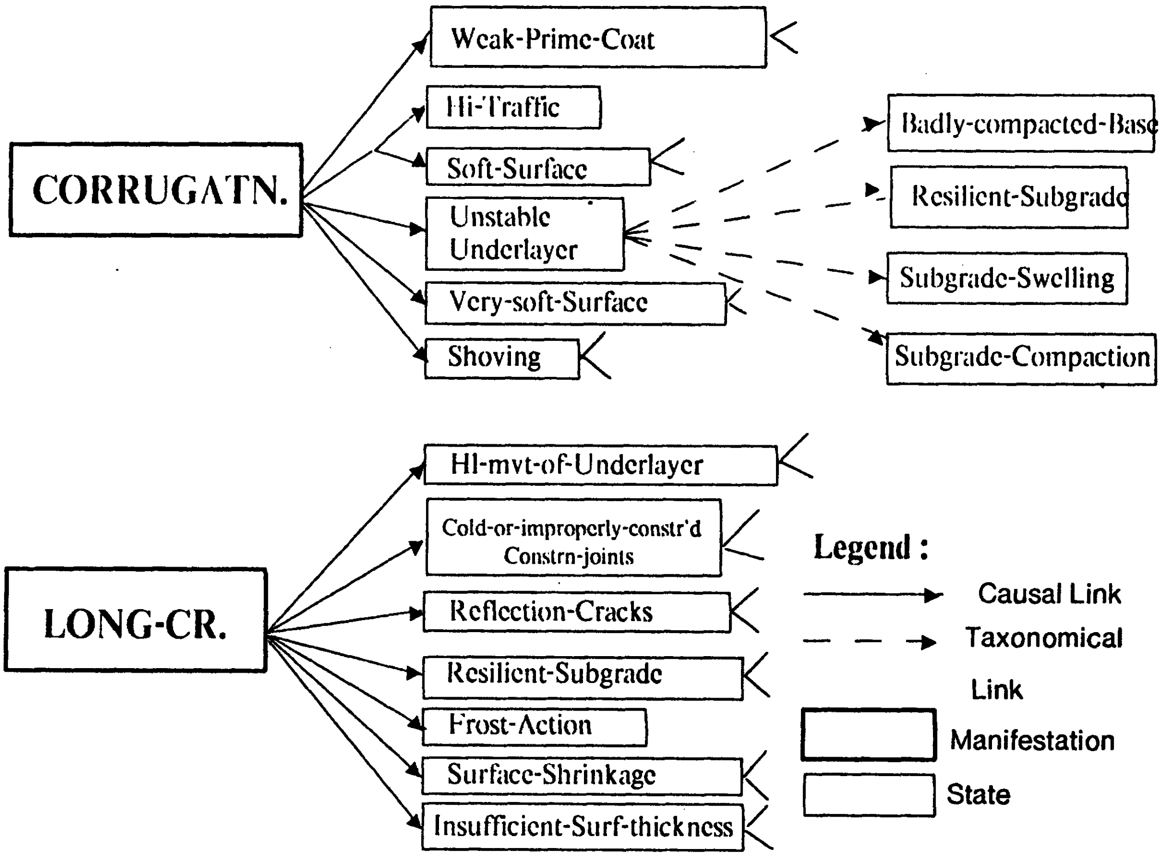
5.2.2.2. Polymorphism and Look-ahead

We can observe at least two different kinds of representation used in the problem solving, namely, the *Causal Representation* and the *Taxonomical Representation*. This representational redundancy, or *Polymorphism* is utilized in two aspects:

Look-ahead Facility:

> We assume -until proven otherwise- that multiple manifestations of a deterioration problem stem from the same underlying failure mechanism. Polymorphic look-ahead, achieved by the implementation of the *Synthesized links of Generalization* as suggested by (Pople, 1979) over several levels of

the tangled hirarchies, cuts the number of candidate diagnosis tasks to a minimum. Subsequently, the differential diagnostic efforts are reduced drastically. Thus, the look-ahead facility keeps the participation required from the user at a minimum level. An illustrative example is shown in figure 5-1.



If we are using any of the traditional search techniques, we've to conduct at least[(5+4)+7=16] tests to confirm (or refute) each possible cause. Meanwhile, The use of the Lookahead facility would require only one test to confirm the Subgrade Resiliency. Notice that this (1:16) parsimony was attained looking ahead one level up the hierarchy, while the intended depth of the facility is 3 levels.

Figure 5-2: An illustrative Example for the Look-ahead Facility

Parsimonious Data Acquisition:

When climbing the network in one plane (or representation) is hindered by lack of a link, the program tries to override this gap through other (representational) planes. This bridging process takes place if there is a (grand) parent node and a (grand) child node that include the gap in between and also they, combinedly, belong to more than one representational plane.

In addition to the polymorphic representation of pavement, PMORPH includes individual rules that describe some theories and facts about the failure mechanisms of pavement.

5.2.3. The Quantitative Discovery System, QUDS

QUDS is consisted of several interacting modules (figure 5-2), their descriptions and the way they interact is as follows. For space limitatation, it was impossible to show a complete example. Several examples are given in (El-Shafei, 1986a).

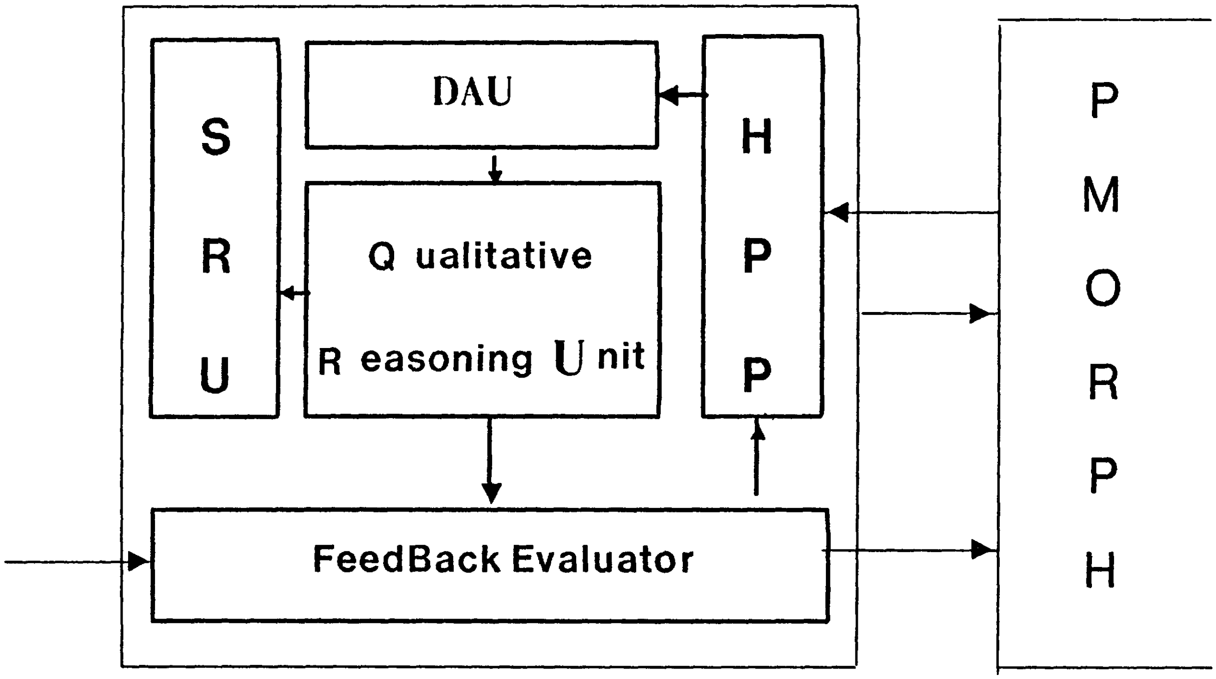


Figure 5-3: The Quantitative Discovery System, QUDS

5.2.3.1. The Heuristic Pre-Processor, HPP

The task of Heuristic Pre-Processor is to collect all the variables that may relate to the problem/variable under consideration. Then it piles them up into heaps according to their classes. It also transfer some control commands about the nature of the problem, e.g., the need for dimensional homogeneity, and the sophistication of the expected relation (linear, polynomial, logarithmic,..etc.). HPP works according to the following:

1. The current causal network, i.e., the <*CAUSED-BY*> fields in PMORPH.
2. Respecting the <*CLASS*> categories as much as possible in forming the dimensionally-homogeneous groups.
3. The <*MAYBE-RELATED-TO*> fields.
4. User specifications.

5.2.3.2. The Dimensional Analysis Unit, DAU

Both this unit and the Stepwise Regression Unit represent a mathematical facet for inductive learning which is integrated by the Qualitative Reasoning Unit to form a reasonable Discovery System.

This mathematical inductive learning tool is completely different from both the BACON series and ABACUS, as they are tackling the formation of a relationship on a heuristically-numerical basis. That means that they exhaustively search all the possible combinations of variables (raised to different integer powers) till they reach satisfactory correlation. Such kind of search, which ignore any algebraic capability, does not ensure finding the right relation, and resulting empirical formula has no physical meaning at all. Needless to mention its severe limitations discussed in the section on The Related Works. On the other hand, we have developed the quantitative facet of our system based on the formal Π-theorem of dimensional analysis and stepwise regression analysis. Therefore, we deal with the definite mathematical problems with the appropriate tools rather than wasting our resources in exhaustive search.

It implements the Π-Theorem of dimensional Analysis (Taylor, 1974) in clustering the variables into dimensionally-homogeneous groups. The variables in each group belong to the same class of variables. DAU isolates the variable(s) under consideration in one group with a unary power of exponentiation. Through Dimensional Analysis, we make sure that the resulting relationship is physically-meaningful, and if there is a law or theory that *may* hold between these variables, it is the induced relationship. DAU is invoked by a request from the HPP which decides on the invocation depending on the nature of the problem to be solved. The DAU utilizes a sort of K-lines (Minsky, 1986) to keep in memory the possible combinations of variables that form dimensionless groups. The dimensionally-homogeneous groups are passed from the DAU to the Qualitative Reasoning Unit (QRU) to check viability and semantics of the grouping as described later. After that the approved groups are passed from the QRU to the Stepwise Regression Unit (SRU).

5.2.3.3. The Qualitative Reasoning Unit, QRU

This unit plays a central role in the qualitative learning through conducting two major tasks:

1. No quantitative knowledge could be accepted (or induced) until after checking its qualitative viability. For example, we can not say that (X = 2.5 Y) till we know that X *IS-PROPORTIONAL-TO* Y. After this qualitative check we can go forward in discovering the right quantitative expression. Therefore the QRU reason about (or checks the correctness of) both the intra-group relationships between the different variables and the inter-group relationships between the various dimensionally-homogeneous groups from the qualitative point of view. Both causal, temporal, and probably nosological (connectedness) reasoning are germane to this unit (Forbus, 1984) (Williams, 1984) (Doyle, 1984). The approved groups are passed to the Stepwise Regression Unit (SRU).

2. Working as an *Explanation-Based Generalizer, EBG* . Based on the domain knowledge stored in PMORPH, the EBG will recognize any new mechanism (or plan). Once recognized, these observed mechanisms are generalized as far as possible while preserving the underlying explanation of their success.

Unfortunately, almost all of the work that has been done in reasoning, concentrated on the qualitative reasoning. Therefore we try to develop a system that is quantitatively capable of reasoning because we do believe that such a capability is a prerequisite for almost all the problems in the engineering domains.

5.2.3.4. The Stepwise Regression Unit, SRU

This unit implements the formal technique of stepwise regression analysis (Smillie, 1966) which enables it to develop either linear or non-linear equations to correlate any number of terms (or groups). The SRU can also expand any argument as a polynomial in terms of another argument. The characteristic feature of this unit is its ability to construct the relationship gradually (Smillie, 1966) by adding one new term at a time to the previous stage of relation construction. At every stage the unit picks out of the remaining candidate groups the most influential one and then adds it to the relation if its contribution to the total correlation exceeds a minimum threshold. This feature enables us to weigh the tradeoffs between the complexity of the relation and its correlation.

The SRU works in parallel with the Explanation-based generalizer (EBG) embedded in the Qualitative Reasoning Unit. Any conflict between them should be solved by the QRU.

5.2.3.5. Feedback Evaluator, FBE

This is the unit responsible for the evaluation of the results of the feedback and accordingly modifying the contents or the reliability of either the QRU rules, the HPP rules, or the PMORPH network.

5.2.4. The Feedback Unit

As we mentioned before, each case is registered in a file with its prescribed diagnosis. After a specific period of time elapses since the diagnosis, feedback is required to help in evaluating the reliability of the reasoning chain used for reaching that diagnosis. The feedback should include the current values of all the variables submitted the first time.

Chapter 6

Polymorphism for the ill-structured domains

Differential Diagnosis of pavement deterioration is a vast, ill-structured domain where the singular representational structures face great failure because of the unfitting of the domain into one of those Structures. This failure is due to the complex multiplicity of the aspects of the problems in the domain. More specifically, The major three criteria of the representation dilemma in the domain of Pavement Mechanics are:

1. Simultaneity : The task of pavement diagnosis is severely aggrevated by the existence of several distress causes working upon the pavement simultaneously. Till now no work has been done in the area of seggregating the symptoms according to their possible causes. In other words, no approach has been developed up until now for the assessment of the combined action of multiple causes upon one kind of deterioration manifestation.

2. Interaction : Most of the instances of simultaneity of causes are accompanied by interaction between these causes. I do believe that this interaction is the major reason for our misunderstanding of pavement. The reason for this misunderstanding is the cross-effects between the different deterioration causes in an unstudied way.

3. Incompleteness or Incorrectness of our knowledge : This fact is due to the huge amount of assumptions that we postulate in order to fit the pavement into one of the already-available patterns of mechanistic behavior neglecting the semantics (expressiveness of reality) of such fitted pavement.

The paper will discuss the Polymorphism[9] as a way of representation which promises solutions for most of the problems of representation met by the other ways, especially, the first two of the

---

[9] Webster's defines "polymorphism" as "The quality or state of being able to assume different forms; the property of crystallizing in two or more forms with distinct structures"

above-mentioned aspects of the problem(Simultaneity and Interaction). Surprisingly enough, the Polymorphism was proposed in three different areas of AI at the same time, namely learning, Theory of heuristics, and Medical Diagnosis. A comparison will be done, and an automatic crystallization of such Polymorphism is proposed. Some examples from other domains such as the Medical Diagnosis would be helpful because of the extensive research done over there on the cognitive criteria of the ill-structured domains.

6.1. Introduction

The researchers in the area of Highway maintenance for the last decade were busy in building different programs that deal with the problem of pavement deterioration from different aspects, but no one really paid any attention to the nature of the knowledge of the domain. Meanwhile, the last two decades were the arena for a lot of research aimed towards the cognitive processes of clinical diagnosis. The major part of these processes is the differential diagnosis that is used by the physician to manipulate the available manifestations and data about the patient out of which he decides parsimoniously whether he needs any further data before making his diagnosis. Some important observations were made about the process of differential diagnosis. In the following paragraphs, I will expose those observations which are relevant to the problem of Highway maintenance:

1. The sequence of the differential diagnosis is very vital as it is at least the way to keep the "Principle of Parsimony" intact.
2. the differential diagnosis process for the same case may differ completely from a specialist to another, still both of them reach the same diagnosis.
3. Physicians, as well as pavement experts, generate tasks early in the patient encounter, long before definitive data are available to ensure that the hypothesized task is indeed an appropriate one to pursue in the case at hand (Pople, 1979).
4. Consultation with other specialists when evaluating difficult clinical problems helps to restore a breadth of perspective to ensure that reasonable alternatives to the primary conceptualization of the problem by the physician will not be overlooked. Such a width of perspective is sacrificed as a price for the early formation of possible diagnoses.

The first two points may be summed up into one point which is:

> Paucity of information about the cognitive process of differential diagnosis which is accounted for in large part by the lack of suitable analytic tools that simulate the physician's thought process (Pauker, 1976).

This missimulation is very obvious in the third observation as the early conceptualization of the problem implies the determination of the strategy of information acquisition and the appropriate representation structure According to this conceptualization. No work has been done to reveal the cognitive process that leads to this early conceptualization of the problem.

Looking at the fourth observation, we can tell why most of the diagnostic consultation programs cannot provide that breadth of perspective. That is because they deal with the process of decision making in well-structured situations where the differential diagnosis is given a priori,which is not the case of the real situations where the differential diagnosis is to be formulated.

This is What made some AI scientists, such as Pople and Simon, consider the domain of differential diagnosis an ill-structured one (Pople, 1979), while all the programs developed for Highway maintenance as well as those of Artificial Intelligence in Medicine, AIM, were assuming a well-structured problem before hand. That was the reason of the failure of those programs to attract the attention of either the Highway maintenance media or the Medical media. The next chapter will discuss the shortcomings of this misconsideration (or misrepresentation) as depicted in PIP, INTERNIST/CADUCEUS ,and CASNET,which are programs of Artificial Intelligence in medicine.

6.2. Limitations of Monomorphic Representations

Theoritically speaking, there is no domain where we cannot apply a single representation structure which can cover it completely for the purpose of search, this is, at least, from the semantic point of view. But from the pragmatic point of view, if the domain under consideration is very vast and (or) we cannot acquire all of its real (hidden) structure even if we can acquire the whole real structure, it would be too detailed to allow us to look ahead through it or to have focus on few hypotheses within limited time. In this case the best way to deal with such a domain is by having several limited capabilities as in the example of "the elephant and the blind people", as it is obvious that one blind man cannot recognize the elephant. Similarly, applying a single structure to an ill-structured domain would end up with similar results. The most crucial shortcoming of considering an ill-structured problem as a well-structured one is believing in one conceptualization of the problem and hence constructing a representation that is compatible with that conceptualization. Most of the efforts after that are wasted in fitting the problem into that representation mold, and consequently, any solution that may come from this representation would be the farthest from reality. The above-described singular representation is called monomorphic[10] representation.

In PIP, an expert system for the diagnosis of renal failure, we observe that due to the imposition of a specific representation, namely the causal network of state frames with a triggering control mechanism applied in all the stages of differential diagnosis (Pauker, 1976), the following problems rise:

1. An invocational explosion of hypotheses is the major problem that may keep the program busy on increasingly idle hypotheses for very long time. This is due to the lack of the feature of early conceptualization.

2. The program is unable to detect multiple coexisting diseases.

In the other side,INTRNIST-1, an Expert System for internal diseases, uses another monomorph, the causal hierarchy, which has the virtue of controlling the proliferation of active hypotheses during the diagnostic process. But it still has the following problems:

1. The probabilistic scoring mechanism has no means for representing "must-not-have" relation.

2. If we have two completely different diseases sharing some manifestations, the INTERNIST picks only one of them. For example, Both Constrictive Pericarditis (a cardiological disease) and A.Glomerunephritis (a nephrological one) share the major manifestation of Nephrotic Syndrome.

---

[10] Webster's defines monomorphic as "having but a single form, structural pattern; .."

"CASNET", which is a diagnostic program for the Glaucoma diseases (Szolovits, 1978), makes a strong point about our entire discussion in this chapter. It works in a domain where normal and diseased states are well understood in physiological detail. Thus, we are in well-structured domain. Therefore, monomorphism is a sole and successful option as no aspects are feared to be lost if the singular representation structure is well built.

6.3. Polymorphism as a Solution

As a prelude for preparing the stage for the proposal of Polymorphism, [(Pople, 1979),pp.131] launches an exagerated attack upon the usage of "Binary Choice (True/False)Tasks saying:

> The main difficulty with the binary-choice approach is that it fails to aggregate diagnostic possibilities into decision sets

He should not raise this point as a justification for his proposal because :

1. This point could be very easily countermeasured by a set of good descriptive,interconnective relations, which is already a part of the heuristics used by the physicians.
2. The failure of monomorphism in dealing with ill-structured problems is very obvious and needs no further justifications.

Anyhow, we saw how the monomorphism in the ill-structured problems leads to several shortcomings which no single structure can solve collectively. Therefore Pople also thought that as far as the same collection of data evokes different conceptualizations in the minds of the different specialists as we showed in the fourth observation of consultation, in the introduction section. This variety of conceptualizations saves a lot of questions that were to be asked to the maintenance staff (or patient) had we had a single conceptualization. Thus, a diagnosis for the difficult cases is attained effectively and parsimoniously. This is how did Pople conclude the importance of having at least two structures of representation to provide the breadth of perspective required for the effective differential diagnosis.

Fortunately,in the Highway maintenance area we can find easily three axioms upon which we can scaffold three different representational structures, namely Structural (Topological), Environmental, and traffic properties of the highway. These three axes are just preliminary proposals, hoping that further discussions with experts of Highway Maintenance would lay a sound cognitive foundation for future developments in this domain. Also in the Medical Domain we can find two types of structures very easily, they are the causal structure and the nosological (locational) structure. Both of the structures are constructed upon the same knowledge base (of states,diseases, and manifestations). Therefore, if in one structure we need to acquire some information externally for a specific node, instead of that we override this node by detouring around it via the perpendicular representation structure.

6.4. The same Concept: Different Nomenclatures for Different Domains

Surprisingly enough,we find the "Principle of Polymorphism " proposed in three completely different areas in the field of Artificial Intelligence almost at the same time with different nominations.

We have discussed this principle in the area of clinical diagnosis in the preceding chapter as proposed by (Pople, 1979). In the area of "Theory of Heuristics" and its applications in both the theory of Numbers and the Biological Evolution, we can trace almost the same theme under the name of "Heuristic Discovery" in Lenat's paper in Machine Learning (Lenat, 1983). And in the area of "Learning by Analogy" we can detect the same theme in (Winston, 1984), where every new example, near-miss, or counter-example could be considered as a new facet (or structure) for the same polymorph.

6.4.1. Comparison

1. In the learning process of Specialize/Generalize proposed by Winston, if we considered each example as a representation, the same concept could be depicted in the use of the *Planning Link* in Pople's paper to override any instantiation node by utilizing the *Generalized Synthesis Operators* proposed by him in pp.167, which have the goal of generating the Most Specific Conjectures, MSC (Michalski, 1984) in its process of generaliation. The Specialize/Generalize process exists under the same nomination in Lenat's paper for the asme uses Winston made them for.

2. Pople's *Rapid Focussing* using the Generalized Synthesis Operators for developing Planning and Spanning Links, which was the solution for the problems emerged due to the detailed causal network, is very similar, in some degree, to the *Shallow Tree Pattern* proposed by Lenat's paper (pp.273) to avoid getting lost in a huge number of intermediate states especially when we need to *look-ahead*.

6.5. Automatic Crystallization of Polymorphs: Recommendations

A study-worthy topic is the relationship and similarity between differential diagnosis and discovery. For the first moment they may sound completely different but if we scrutinized them from the points of view of Upward Search mechanisms and upward Most-Specific-Conjectures generalization, we would find that differential diagnosis is a discovery process with predefined set of goals (or generally, predefined upper layer in the search hierarchy) and guided by some combinatorial heuristics (which is available also for the discovery process most of the time). By partial relaxing the condition of the predefinition of the set of goals, we reach the stage of Inductive Learning which is the solution of the last problem mentioned in the beginning of this' chapter, i.e.,the incmpleteness or incorrectness of our knowledge about pavement. Therefore, I suggest, for the case of vast, ill-structured domains such as pavement deterioration with all its different implications and all the ambiguity enclaving the nature of its basic components (pavement and concrete), that we should develop a Learning Program which will be able to acquire all the information it needs by different means -say to have full-automatic experimentation'researcher guided by a set heuristics for directioning the experimentations in the field of pavement which would be a good framework for representation of the knowledge if nothing else - and then formulate it in the appropriate form according to a set of heuristic rules, which I would not think they may exceed 200 rules. If we managed to build such a program, the domain of differential diagnosis of pavement I be shifted from its ill-structured nature towards a well-structured one.

The next step, by the use of another set of heuristics, we can automatically construct multiple facets of the representational structure (rather than staying limited to the structural,environmental and traffic facets only). Those facets may be beyond our temporary capabilities to be developed or even beyond our imagination to be grasped (say, Thermological, Electromagnetic, or classification of the different elements of the roadway in the case of pavement Diagnosis in comparison with the rigidly-fixed, current pathological representation).

It may sound a bit imaginative but never to be impossible, besides, may an imaginative proposal be the breakthrough required in the stagnant pavement deterioration domain.

Chapter 7

Conclusion & Future Work

The conducted research, described in this thesis, has a double-folded goal. The goal is enhancing our knowledge about an evolving domain through the development of a generic tool for quantitative discovery. The evolving domain here is the failure mechanisms in pavement. The approach followed in the development of this tool is hybridizing the symbolic and numeric techniques of learning from observations, explanation-based generalization, adaptive control, and qualitative reasoning into a generic quantitative discovery tool. The crystalization of the above-mentioned concepts is HOTEP system, which is under development within this research. HOTEP enables us to build our knowledge in a gradual form, such that at any point of time we can use the current level of knowledge in making decisions about a problem at hand. So, we can use the system for diagnosis and prediction while it is still in its permanent learning process. The more cases the system encounters, the more experienced it will be. This experiential performance is a principal feature of any human expert.

The current status of development of HOTEP is as follows. The quantitative discovery system, QUDS, is almost completely developed. Pilot projects have been developed for both of the case library and the polymorphic representation network. The feedback facility is not implemented yet.

The system tested in the areas of failure mechanisms of pavement, soil mechanics, and theoritical physics. Current efforts are exerted towards the application of the system in the areas of econometric modeling and shear failure in steel.

7.1. Limitations & Future Work

1. The conceptual design of the feedback facility needs to be implemented.
2. Implementing the concept of explanation-based generalization in handling the competing theories.
3. The quantitative discovery system, QUDS, should be able to handle the partially (or totally) nominal set of variables.
4. Determining the optimal number of representational levels to be spanned by both the lookahead feature and the synthesized links of generalization.
5. Prescribing the least amount of remedy required to prevent a specific state of distress predicted in the future (preventive planning).
6. Providing, at request, the qualitative and quantitative reasoning for attaining a specific decision.
7. Developing the proper mechanism for harvest of knowledge. Every year or so the would-be-different versions of the knowledge base should be *harvested* for a super learning session after which a new release of the knowledge base is broadcasted for the different users. The learning capability of the system assures the customization of the system to the local environment. Meanwhile, the harvesting process assures the robustness and globality of the acquired knowledge.

Bibliography

Abdallah, Said. *Dimensional Analysis for Experimental Fluid Mechanics.* Cairo University Press, 1979.

Abelson,Hal & Jerry Sussman. *Structure and Interpretation of Computer Programs.* Cambridge, Mass.:MIT Press, 1985. It is the text book of the course 6.001 that inaugurate the MIT freshmen for a strivingly competitive curriculum.

Aho, Alfred, John Hopcroft, and Jeffrey Ullman. *The Design and Analysis of Computer Algorithms.* Reading, MA:Addison-Wesley Publishing Co, 1974.

Akmajian, Adrian, Richard Demers, & Robert Harnish. *Linguistics - An Introduction to Language and Communication.* Cambridge:MIT Press, 1984. Second Edition.

Alkire, Bernard D. Effect of variable-drainage freeze-thaw tests on post-thaw Shear Strength. *Transportation Research Record TRR*, 1981, *1(Rec.809)*, 115-139.

Alkire, B., and J.Morrison. Change in soil structure due to freeze-thaw and repeated loading. *Transportation Research Record TRR*, 1983, *2(Rec.918)*, 15-22.

Applegate, James, Michael Douglas, Yekta Gursel, Gerald Sussman, & Jack Wisdom. *The Outer Solar System for 210 million years.* Technical Report AI Memo 877, Massachusetts Institute of Technology, February 1986. The paper discusses some results of a special-purpose computer developed for Astrodynamics. One of the results is assuring that no collision will happen between Pluto and Neptune within the next 210 Myr, despite their intersecting orbits. Previous calculations could not go beyond 5 Myr. Also more insights are given into the Three-body problem.

Committee on Fracture in Compressive Stress Fields. *Fracture in Compression of Brittle Solids.* Technical Report NMAB-404, National Materials Advisory Board, Commission on Engineering and Technical Systems, National Research Council, 1983.

ASCE. Pavement Management Systems - Can they save pavements, cut rehab costs ? *Civil Engineering*, November 1979, *1*, 62-65. A good history of Pavement Maintenance Systems.

Backus, J. Can Programming be librated from the Von Neumann Style? A Dunctional Style and its Algebra of Programs. *Communications ACM*, August 1978, *21(8)*, 613- 641.

Balta, Wayne S. Optimization of Maintenance Policies . Master's thesis, Massachusetts Institute Of Technology, 1984. A good literature survey for the available Pavement Management Systems.

Balzer,R., L.Erman, P.London, & C.Williams. *HEARSAY-III: A domain-independent framework for Expert Systems*, pages 108-110. AAAI 1, 1980.

Bartell, C & K. Kampe. *Development of the California Pavement Management System - System Description*. Technical Report FHWA/CA/HM/7139/78/03, FHWA, 1978.

Bathe, Klaus Jurgen. *Finite Element Procedures for Engineering Analysis.* Reading, MA:Addison-Wesley Publishing Co., 1980.

Battiato, G. et al. *Design of Flexible Pavements by the Viscoelastic VISTRA method*. Technical Report, Agip, a subsidiary of Eni, 1978.

Bavel, Zamir. *Math Companion for Computer Science.* Reston, VA:Reston Publishing Co, 1982.

Belongia, David L. Highway Applications of Geotextiles. *Civil Engineering-ASCE*, May 1981, *1*, 60-63.

Benjamin, Jack, and Allin Cornell. *Probability, Statistics, and Decision for Civil Engineers.* New York, N.Y.:McGraw-Hill Book Co, 1970.

Brachman,R.J., Fikes,R.E.& Levesque,H.J. *KRYPTON: A Functional Approach to Knowledge Representation*, pages 67-73. IEEE, October, 1983.

Bradley, Stephan, Arnold Hax, & Thomas Magnanti. *Applied Mathematical Programming.* Reading, MA:Addison-Wesley Publishing Co., 1977.

Bridgman, P.W. *Dimensional Analysis.* New Haven, CT :Yale University Press, 1931.

Campbell, J.A. *Ellis Horwood Series Artificial Intelligence. Implementations of PROLOG.* Chichester, U.K.:Ellis Horwood, 1984.

Carbonell,J.; R.Michalski; T.Mitchell. An Overview of Machine Learning. In Michalski,R. et al. (Ed.), *Machine Learning,* Palo Alto,CA: Tioga Pub., 1984.

Charniak, Eugene, Drew McDermott. *Introduction to Artificial Intelligence.* Reading, MA:Addison Wesley Co., 1985.

Charniak,Eugene, Christopher Riesbeck, & Drew McDermott. *Artificial Intelligence Programming.* Hillsdale, N.J.:Lawrence Erlbaum Associates, 1980.

Chen, W.F., and E.C.Ting. *Fracture in Concrete.* New York, N.Y:ASCE, 1980. Proceedings of a Session sponsored by the Committee on Properties of Materials of the ASCE Engineering Mechanics Division.

Chisholm,R.A., and W.A.Phang. Measurement and Prediction of Frost Penetration in highways. *Transportation Research Record, TRR*, 1983, *1(Record 918)*, 1-9.

Chomski, Noam. *Aspects of the Theory of Syntax.* Cambridge, MA:MIT Press, 1965.

Chomski, Noam. *Language and Mind.* London, UK:Harcourt, Bruce,and Jovanovich, 1968. Good introduction to the concepts of Performance and Competence.

Clocksin, W.F., & C.S. Mellish. *Programming in Prolog.* Berlin:Springer-Verlag, 1981.

Cox, Brian. Pavement Maintenance Management. In Peter Pell (Ed.), *Developments in Highway Pavement Engineering,* London, UK: Applied Science Publishers, 1978.

Davis,R. & Lenat,D. *Knowledge-Based Systems in Artificial Intelligence.* New York:McGraw Hill, 1982. A presentation for the Ph.D. dessertations of the authors that describe in detail the AM system for Learning by Discovery in the domain of Mathematics, and Tereisias the reasoning system of Mycin expert system.

de Kleer, Johan. *Multiple Representations of Knowledge in a Mechanics Problem-Solver*, pages 299-304. IJCAI, Cambridge, MA, 1977. A good proposal for integrating real quantities in qualitative reasoning.

Doyle,Richard. *Hypothesizing and Refining Causal Models.* Technical Report AIM 811, Massachusetts Institute Technology, December 1984.

Dreyfus, Hubert. *Harper Colophon Books. What Computers Can't Do - The Limits of Artificial Intelligence.* New York, N.Y.:Harper & Row, 1979.

Duda, R.O., J.G.Gaschnig, & P.E.Hart. Model Design in the PROSPECTOR consultant system for mineral exploration. In Michie, D. (Ed.), *Expert Systems in the micro-electronic age*, Edinburgh University Press, 1979.

el-Shafei, Nayel S. Utilzing Quantitative Discovery in Qualitative Reasoning about Failure Mechanisms in Pavement. Master's thesis, Massachusetts Institute of Technology, 1986.

el-Shafei, Nayel S. *Quantitative Discovery and Reasoning about Failure Mechanisms in Pavement*. Society of Computational Mechanics, Ashurst, Southampton, UK, 1986.

Elliot, John F. *Distress and Failure of Pavement Systems.* PhD thesis, Massachusetts Institute of Technology, 1969.

Esch,D. & McHattie,R. Prediction of roadway strength from soil properties. *Transportation Research Record TRR*, 1983, *1(Rec.918)*, 30-34.

Ewalds,H.L., & R.J.H.Wanhill. *Fracture Mechanics.* Delft, The Netherlands:Delftse Uitgevers Maatschappij, DUM, 1984.

Fain,J., Hayes-Roth,F., Sowizral,H., Waterman,D. *Programming in ROSIE: An introduction by means of examples*. Technical Report And-1646-ARPA, Rand Corp., 1982.

Falkenhainer,Brian. *ABACUS - Adding Domain Constraints to quantitative Scientific Discovery.* Technical Report UIUC-DCS-F-84-1195 (ISG 84-7), University of Illinois, Urbana-Champaign, 1984.

Feigenbaum, E.A., Buchanan, B.G., Lederberg, J. On Generality and problem solving: A Case Study using the DENDRAL program. In Meltzer, Be., Michie, D. (Ed.), *Machine Intelligence*, Edinburgh University Press, 1971.

Feynman, Richard. Simulating Physics with Computers. *International Journal of Theoritical Physics*, 1982, *21(6/7)*, 467-488.

Findakly, Hani K. Stochastic Approach to the analysis of highway pavements. Master's thesis, Massachusetts Institute Of Technology, 1971.

Flanagan, Owen J., Jr. *A Bradford Book. The Science of the Mind.* Cambridge, MA:MIT Press, 1984.

Fodor, Jerry. *A Bradford Book. The Modularity of Mind.* Cambridge, MA:MIT Press, 1983.

Forbus,Kenneth. *Qualitative Process Theory.* Technical Report AI-TR-789, Massachusetts Institute of Technology, AI Lab, July 1984.

Forgy,C. & J.McDermott. *A Domain-independent production system language*, pages 933-939. IJCAI 5, 1977.

Freiherr, G. *The seeds of Artificial Intelligence*. Technical Report NIH no. 80-2071, SUMEX-AIM, 1980.

Gabel, David. The Project Management Act. *PC Week*, February 1986, *3(5)*, 57-81. A good review for the PC-based project control systems available in the market.

Gifford,G. et al. Advice for evaluation of Thaw weakening of frost-susciptible soil. *Transportation Research Record TRR*, 1983, *1(Rec.918)*, 10-15.

Golabi,Kamal, & N. Kulkarni. Statistical State Transition in Pavement Management Systems. *ASCE Interfaces*, 1983, *1983*, 30-55.

Gorry, G.A., H.Silverman, & S.G.Pauker. Capturing Clinical Expertise: A Computer Program that Considers Clinical Responses to Digitalis. *American Journal of Medicine*, March 1978, *64*, 452-460.

Greiner,R. & Douglas Lenat. *A Representation Language Language, RLL*, pages 165-169. AAAI 1, 1980.

Gunaratne,M., J.Chameau, and A.Altschaeffl. Introduction to Fuzzy Sets in Pavement Evaluation. *Transportation Research Record, TRR*, 1984, (985)}, 16-25. The concept of Type-2 Fuzzy Sets is introduced for evaluation of pavement performance.

Haugeland, John. *A Bradford Book. Mind Design - Philosophy. Psychology. Artificial Intelligence.* Cambridge, MA:MIT Press, 1981.

Hayes-Roth, Fredrick. *Teknowledge Series of Knowledge Engineering.* An investigation of Tools for building Expert Systems. In Hayes-Roth, Fredrick, Donald Waterman, Douglas Lenat (Ed.), *Building Expert Systems,* Reading, MA: Addison-Wesley, Inc, 1983.

Heisey,J.S.& K.H.Stokoe. Moduli of pavement systems from spectral analysis of surface waves. *Transportation Research Record, TRR*, 1982, *1(Record 852)*, 22-32.

Hendrix, G.G. *Expanding the utility of semantic networks through partitioning*, pages 115-121. IJCAI 4, 1975.

Hillis, William D. *The Connection Machine*. PhD thesis, Massachusetts Institute of Technology, 1986.

Hofstadter, Douglas R. *Godel, Escher, Bach - An Eternal Golden Braid.* New York, N.Y.:Vintage Books, 1980.

Holland, John H. Escaping Brittleness: The possibilities of General Purpose Learning Algorithms applied to parallel rule-based systems. In Michalski, Ryszard, Jaime Carbonell, & Tom Mitchell (Ed.), *Machine Learning - An Artificial Intelligence Approach - Part 2,* Los Altos, CA: Morgan Kaufmann Publishers, Inc., 1986. A discussion of the problem of Credit Assignment.

Hozayyen, Aly. *Numerical Methods for Engineering Systems.* Cairo, Egypt :Cairo University Press, 1978.

Highway Research Board. *Standard Nomenclature and Definitions for Pavement Components and Deficiencies.* Special Report 113, TRB, 1970.

Jones, R.H. Frost Heave Damage and its Prevention. In Peter Pell (Ed.), *Developments in Highway Pavement Engineering,* London, UK: Applied Science Publishers, 1978.

Kakel, William W. Utilzing Quantitative DFracture in Asphaltic Mixtures. Master's thesis, Massachusetts Institute of Technology, 1968.

Kenis, William J. *Predictive Design Procedure - A design method for flexible pavements using the VESYS structural subsystem.* Univ.of Michigan, Ann Arbor, MI, 1977.

Kernigan, Brian, & Dennis Ritchie. *The Software Series. The C Programming Language.* Prentice-Hall Co., 1978.

Kerr,A.D., and P.J.Shade. Analytic Approach to Concrete Pavement Blowups. *Transportation Research Record, TRR*, 1983, *1(Rec.930)*, 78-83.

Kuhn, Thomas. *International Encyclopedia of Unified Science, Vol.II, no.2. The Structure of Scientific Revolutions.* University of Chicago Press, 1970.

Kuipers, Benjamin. *Computational Models of Cognition and Perception - A Bradford book.* Commonsense Reasoning about Causality: Deriving Behavior from Structure. In Daniel Bobrow (Ed.), *Qualitative Reasoning about Physical Systems,* Cambridge, MA: MIT Press, 1985. The qualitative differential equations are used as a language to obtain a behavioral description out of the structural description.

Lahue S. & Botelho, F. Sulfur-extended Asphalt. *Civil Engineering-ASCE*, May 1981, *1*, 57-59.

Langer, Susanne. *An Introduction to Symbolic Logic.* New York, N.Y.:Dover Publication Co., 1967.

Langley,P. & Simon,H. *The Search for Regularity : Four Aspects of Scientific Discovery.* Technical Report CMU-RI-TR-84-20, Carnegie-Mellon University, 1984.

Lavrac, Nada, Igor Mozetic, & Ivan Bratko. Analysis of Two Learning Algorithms: A Comparison Study in Three Medical Domains. To be published as a Technical Report in both UIUC-ISG, and Josef Stefan Institute, Ljubljana, Yugoslavia.

Lenat,D. The role of heuristics in learning by discovery. In Michalski, Carbonell & Mitchell (Eds.), *Machine Learning,* Palo Alto,CA: Tioga Pub. Co., 1983.

Lenat, Douglas, & John Seely Brown. Why AM and EURISKO Appear to Work. *Artificial Intelligence Journal*, 1984, *23*, 269-294. A defense against a criticizing article by Ritchie in the same issue.

Lewis, Harry, and Christos Papadimitriou. *Prentice-Hall Software Series. Elements of the Theory of Computation.* Englewood Cliffs, N.J.:Prentice-Hall, Inc, 1981.

Li, Deyi. *Computer Engineering Series. A PROLOG Database System.* Letchworth, Hertfordshire, England:Research Studies Press, 1984.

Majidzadeh, K et al. *Application of Fracture Mechanics for improved design of bituminous concrete.* Technical Report FHWA/RD-76/91, FHWA, 1976.

Majidzadeh, K. et al. *Development and Field Verification of a mechanistic structural design system in Ohio.* Univ.of Michigan, Ann Arbor, MI, 1977.

Margolus, Norman. Physics-Like Models of Computation. *Physica*, 1984, *10D*, 81-95.

Margolus, Norman, Tommaso Toffoli, & Gerard Vichniac. Cellular Automata Supercomputers for Fluid Dynamics Modeling. Cellular Automata Group at MIT Laboratory for Computer Science.

Markow, Michael & Brian Brademeyer. *EAROMAR-II , Technical Report.* Technical Report FHWA/RD-82/086, FHWA, 1984.

Marks, Charles. *A Bradford Book. Commissurotomy Consciousness & Unity of Mind.* Cambridge, MA:MIT Press, 1981.

Marr, David. *Vision - A Computational Investigation into Human Representation and Processing of Visual Information.* San Francisco, CA:W.H. Freeman & Co., 1982.

Maser, Kenneth. *Radar.* Society of Computational Mechanics, Ashurst, Southampton, UK, 1986.

McCullough,B.& A.Taute. Use of deflection measurements for determining pavement material properties. *Transportation Research Record, TRR*, 1982, *1(Record 852)*, 8-15.

McGilton, Henry, and Rachel Morgan. *Byte Book: McGraw-Hill Software Series for Computer Professionals. Introducing the Unix System.* New York, N.Y.:McGraw-Hill Book Co, 1983.

Michalski,Ryszard. A Theory and Methodology for Inductive Learning. In Michalski,R. et al. (Ed.), *Machine Learning*, Palo Alto,CA: Tioga Pub., [illegible]

Minsky, Marvin. *Semantic Information Processing.* Cambridge, MA:MIT Press, 1969.

Minsky,Marvin. A Framework for representing knowledge. In Winston,Patrick (Ed.), *The Psychology of Computer Vision,* New York: McGraw Hill Book Co., 1975.

Minsky, Marvin Lee. *NATO Conference Series: Systems Science.* Volume 16: Adaptive Control: From Feedback to debugging. In Selfridge, Oliver, Edwina Rissland, & Michael Arbib (Ed.), *Adaptive Control of Ill-Defined Systems,* New York: Plenum Press, 1984.

Minsky,Marvin. *Society of Mind.* Cambridge,MA:Simon & Schuster, 1986.

Mitchell,T., Keller,R., & Kedar-Cabelli,S. *Explanation-Based Generalization, A Unifying View.* Technical Report ML-TR-2, Lab for C.S.Research,The State Univ.of N.J.,Rutgers, August 1985.

Moavenzadeh, F. & Brademeyer, B. *A Stochastic Model for Pavement Performance and Management.* Univ.of Michigan, Ann Arbor, MI, 1977.

Mooney,Raymond & Gerald DeJong. *Learning Schemata for Natural Language Processing,* pages 681-687. IJCAI, Los Angeles,CA, 1985.

Moses, Joel. *Symbolic Integration.* PhD thesis, Massachusetts Institute of Technology, 1967.

Nazarian,S. et al. Use of Spectral Analysis of Surface wave Method for determination of moduli and thicknesses of pavement systems. *Transportation Research Record, TRR*, 1983, *1(Record 930)*, 38-45.

Nii,H.P., and E. Feigenbaum. Rule-base understanding of signals. In Waterman, D., and F.Hayes-Roth (Ed.), *Pattern-directed Inference Systems,* New York: Academic Press, 1978.

Nilsson,Nils. *Principles of Artificial Intelligence.* Palo Alto, CA:Tioga, 1980.

Osman, Abdel-Monem. *Highway Engineering.* Cairo, Egypt:Cairo University Press, 1981.

Pan, Victor, & John Reif. Efficient Parallel Solution of Linear Systems. *ACM Linear Systems,* 1985, }, 143-152.

Patil, Ramesh, P.Szolovits, & W.Schwartz. *Causal Understanding of Patient illness in Medical Diagnosis*, pages 893-899. IJCAI, 1981.

Pauker,S.et al. Towards the simulation of clinical cognition. *American Journal of Medicine*, 1976, *60*, 981 -996.

Polya, G. *How to Solve it - A new aspect of mathematical method.* Princeton, N.J.:Princeton University Press, 1973.

Pople,H.Jr. Heuristic methods for imposing structure on ill-structured problems:The structuring of Medical Diagnosis. In P.Szolovits (Ed.), *Artificial Intelligence in Medicine,* AAAS, 1979. The idea of Synthesized Links of Generalization as a vehicle for Polymorphism is proposed as a remedy.

Potter, J.L. *MIT Press Series in Scientific Computation. The Massively Parallel Processor.* Cambridge, MA:MIT Press, 1985.

Rauhut,Brent, Robert Lytton. Serviceability Loss due to Roughness caused by Volume Change in Expansive Clay Subgrades. *Transportation Research Record, TRR*, 1984, (993)}, 12-18. The impact of volume change in subgrade upon roughness is emphasized.

Rich,Charles. *The Layered Architecture of a System for Reasoning about Programs*. IJCAI, Los Angeles,CA, 1985.

Rieger, Chuck. Spontaneous Computation and its Roles in AI Modeling. In Waterman,D.A., & Frederic Hayes-Roth (Ed.), *Pattern-Directed Inference Systems,* New York: Academic Press, 1978.

Rieger, Chuck; and Grinberg, Milt. *The Causal Representation and Simulation of Physical Mechanisms*. Technical Report TR-495, University of Maryland, November 1976.

Rieger, Chuck, and Grinberg, Milt. *The Declarative Representation and Procedural Simulation of Causality in Physical Mechanisms*. IJCAI, 1977.

Roos, Daniel. *ICES System Design*. Cambridge, MA:MIT Press, 1966. An attempt to develop a unified approach for the different domains in Civil Engineering.

Safwat, K.Nabil. *The Simultaneous Prediction of Equilibrium on large-scale networks: A unified consistent methodology for transportation planning*. PhD thesis, Massachusetts Institute Of Technology, 1982.

Scott, John III. Viscoelastic Characterization of rubber-additive Asphalt. Master's thesis, Massachusetts Institute Of Technology, 1969.

Seltz-Petrash Ann. What's happened to the quality of asphalt? *Civil Engineering-ASCE*, December 1980, *1*, 43-46.

Shahin, M.Y., and J.M.Becker. Development of Performance Prediction Models for Airfield Pavements. *Transportation Research Record, TRR*, 1984, (985)}, 20-34. The concept of Local Models for the individual sites is introduced.

Shortliffe, E.H. *Computer-based medical consultation: MYCIN.* New York:American Elsevier, 1976.

Simmons,Reid. Temporal Representation for Planning. This MIT-AI Area Exam Paper is a good Survey for the topic.

Simon,Herbert. *Pallas Paperbacks. Models of Discovery.* Boston, MA:D.Reidel Co., 1977.

Simon, Herbert. *The Sciences of the Artificial.* Cambridge, MA:MIT Press, 1982.

Smillie, K.W. *An Introduction to Regression and Correlation.* London, U.K.:Academic Press, 1966.

Soo, V.W., C.A. Kulikowski, & D. Garfinkel. *A Framework for Representation of Expertise in Experimental Design for Enzyme Kinetices.* Technical Report CBM-TR-146, Lab for Computer Science Research, Rutgers University, 1985.

Stefik, Mark, & Daniel Bobrow. Object-Oriented Programming: Themes and Variations. *AI Magazine*, Winter 1986, *6(4)*, 40-62. A good Survey for the major concepts, and languages of this trend of programming.

Steinway,J.D. et al. *Locating Voids beneath Pavement using Pulsed Electromagnetic Waves.* Technical Report 237, NCHRP, 1981.

Stoy, Joseph E. *MIT Press Series in Computer Science. Denotational Semantics - The Scott-Strachey Approach to Programming Language Theory.* Cambridge, MA:MIT Press, 1977.

Svobodova, Liba. File Servers for Network-based Distributed Systems. *Computing Surveys*, December 1984, *16(4)*, 353-398.

Szolovits,P.& Pauker,S. Categorical and Probabilistic reasoning in Medical Diagnosis. *Artificial Intelligence*, 1978, *11*, 115-144.

Taylor, E.S. *Dimensional Analysis for Engineers.* Oxford, U.K.:Clarendon Press, 1974.

Thompson,M.R., and M.S.Hoffman. Concepts for developing an NDT-based design procedure for determining Asphalt Concrete Overlay thickness. *Transportation Research Record, TRR*, 1983, *1(Record 930)*, 12-18.

Toffoli, Tommaso. CAM: A High-Performance Cellular-Automation Machine. *Physica*, 1984, *10D*, 117-127.

Toffoli, Tommaso. Cellular Automata as an Alternative to (rather than an Approximation of) Differential Equations in Modeling Physics. *Physica*, 1984, *10D*, 117-127.

Towle, S. *Report of Consulting Engineer on Pavements.* Technical Report, Department of Public Works, 1892. The oldest document that deals with Pavement Technology.

Vorhees, Harry. *Vision Utilities.* Working Paper 281, MIT-AI Lab, December 1985.

Weiss, S., Kulikowski, C. *Expert Consultation Systems: The EXPERT and CASNET projects.* Technical Report no. 3, Machine Intelligence, 1981.

Welborn,York. Physical Properties as related to asphalt durability: State of the art. *Transportation Research Record, TRR*, 1984, (999)}, 31-36. A good record of the genesis of the testing methodology in pavement.

Weld,Daniel. The Use of Aggregation in Causal Simulation. *Artificial Intelligence Journal*, August 1986, *29(3)*, 10-25. The concept of Aggregation,i.e.,Cycles recognition is used for the switching between the two models.

Wiener, Norbert. *God, Golem, Inc. - Where Cybernetics impinges on religion.* Cambridge, MA:MIT Press, 1973.

Wilensky, Robert. *LISPcraft.* New York, N.Y.:W.W. Norton & Co, 1984.

Williams,Brian C. *Qualitative Analysis of MOS Circuits.* Technical Report AI-TR-767, Massachusetts Institute of Technology, AI Lab, July 1984.

Winston,Patrick H. *Learning and Reasoning by Analogy: The Details.* Technical Report AIM 520, Massachusetts Institute of Technology, May 1980. An influential paper in the area of Learning by Analogy.

Winston,P.H. *Artificial Intelligence.* Reading,MA:Addison Wesley Pub., 1984. Chapters 11,12 convey Winston's perspective of Learning by Analogy.

Yandell,W.O. Possible Effect of Relative Plastic Behavior of Pavement layers on Pavement life. *Transportation Research Record, TRR*, 1983, *1(Record 930)*, 86-90.

Yang, Nai. *Design of Functional Pavements.* New York:McGraw Hill, 1972.

Yoder, E. & Witczak M. *Principles of Pavement Design.* New York:John Wiley Inc., 1975.

Appendix A

QUDS pilot project- Description & Results

This appendix discusses the limitations of the preceeding quantitative discovery systems. These limitations were the motivation for building a new quantitative discovery system. The new system, QUDS, is described in section 5.2.3. The differences between QUDS and the precedent systems are shown. After that, the necessity of incorporating some numerical inductive tools, is highlighted. Then a brief discription of the current status of development of the different elements of QUDS, its current limitations, and scheme for future work are given. Finally, some demonstrative examples are given to show the analytical capabilities of QUDS.

A.1. Limitations of Precedent Works

The major two systems that dealt with quantitative discovery are ABACUS (Falkenhainer, 1984) and BACON series (Langley, 1984). Both of them suffer from severe limitations in form of numerical restrictions that keep these systems away from any pragmatic application. By restrictions, It is meant the following limitations:

1. Limited use of operators and powers,
2. Confining the proposed relationships to a mold of single-termed (or at most double-termed) relationships,
3. Unawareness of any prospective physical meaning by violating the dimensional homoGeneity of the proposed relations to keep it within the realm of empirical formulae,
4. Lacking the ability to build the relationship gradually in accordance with relative contribution of each argument to the overall expressiveness of the relationship, and
5. Several other features discussed within the paper.

In the rest of the chapter we will deal primarily with ABACUS system as it shares the same approach with BACON series of systems. Therefore the critique for ABACUS extends for BACON. AB algorithm, the part of ABACUS responsible for quantitative discovery, tries to approach the lower (numerical) level of the problem logically while it is a definite numerical one. The symbolic approach of learning should be used in the higher, abstractual levels of the problem where the numerical methods fall short, such as picking the variables to be included in the relationship, the form of the relationship, when to use it, and how to modify it.

A.1.1. The form of relationships

1. The Relation-finding routine is founded upon the assumption of "One-term relationship" in which all arguments are related by either 'multiplication' or 'division'. The goal of the routine is to find the integer powers for which the arguments should be raised to get the multiplication product equal to constant for a cosiderable percentage of the data.

   The Relation-finding routine will be rendered to obsolescence once the "addition" is incorporated. At that time, the only way for developing relationships will be Regression Analysis.

2. Even within the realm of the one-term relationships the program has an undisclosed condition. This condition is that when a percentage of the data fits on a relationship, these data are banded over a confined region,i.e., all the fitted data are adjacent and occupying a specific range of values for one of the arguments, at least, rather than being scattered all over the scattergraph, intermingled with the rest of data.

   Thus we have the problem of deciding whether :

   a. The found relationship is global, the case of which a threshold correlation should be exceeded; or

   b. The found relationship is local over a specific range of values for an argument , and the search should be resumed for the rest of data.

3. The Relation-Finding routine does not care about the dimensional homogeneity. Thus the relations deduced do not exceed, generally, being empirical formulae. The difference between the relationship and the empirical formula is that the first one is an actual, robust relationship that has a (global) physical meaning while the second one is just an equation given to describe the possible correlation between the values of different sets of arguments under specific circumstances. Such empirical formulae have no possible physical meaning, e.g., we may derive an equation by regression analysis between the airplanes passing over a certain area and the soil consolidation in the same area but it would be completely meaningless. The consent

with the development of just empirical formulae contradicts with the mere concept of real learning.

4. Despite the claim of indifference towards the dependent and independent variables, the sorting of the values of a specific variable implies it is the independent (or the most important) one.

5. ABACUS uses only multiplication(*) and division(/) in building the relationships ignoring the following operators:

   - Adding(+) and subtraction(-).
   - Real (fractional) powers.
   - Exponential and logarithmic relationships.
   - Expansion of series.
   - Non-linear regression (even if the proposed "adding" capability was used).
   - Binary variables (as in Integer programming) to represent the multi-phased relationships.

6. The AB algorithm intends to implement a naive, primitive version of Dimensional Analysis as a differentiative criterion between *multiplication* and *addition*. This method, according to an extensive discussion with B. Falkenhainer, has the following disadvantages :

   a. The technique uses "try-and-error" strategy which is dreadfully inefficient especially in wide search spaces, not to speak about the unlimited space of all possible permutations. The situation becomes more astonishing if we knew that this kind of problems has a definite solution using the same concept of dimensional homogeneity but more effectively, which is Dimensional Analysis.

   b. An immense number of combinations of the arguments and their powers is generated with no practical, restricting criteria to stop the combinatorial explosion, especially in the real-world problems where the number of the arguments is originally very big, not to speak about about raising each argument to all the possible integer powers (- -> +).

   c. An exhaustive check of dimensional homogeneity is applied to every permutation of the arguments and their integer powers to determine whether the two arguments under consideration are to be multiplied or added to each other. To be precise, the algorithm checks the units'homogeneity rather than the dimensional homogeneity, i.e., all arguments expressing time (such as age of pavement) or including time within its dimensions(such as speed) should express time in the same unit (say years), otherwise the program is unable to detect the homogeneity. Consequently, no possible relation could be thought of according to this technique.

A.1.2. Search

1. In the technical report (Falkenhainer, 1984), we find the following statement:

   > [In] the relation finding routine ...,new levels of description may be generated without placing too much emphasis on verifying if the chosen path is the best one to take.

   Such behavior is completely impractical and would waste a lot of CPU time due to the combinatorial explosion. The solution to this problem is using an *Optimization Technique*,i.e., A preference criterion.

2. The ABACUS performance is heavily damaged by the presence of *Irrelevant variables* as mentioned in the technical report. This is due to the following of the "Depth-first" technique in the absence of any preference criterion.

3. In (Falkenhainer, 1984) page 13, we find how the ABACUS is severely affected by the tautologies. Falkenhainer confesses the inability of the algorithm to stop the proliferation of such tautologies.

A.1.3. Statistical Handling of Data

As a limitation on ABACUS, Falkenhainer says:

> The problem of noise was never properly addressed in this implementation.

Indeed the problem is much more profound than being ignored as a noise. In the practical applications we are not knowing before hand the the final form of the relationship and in most of the cases there would be competitive, candidate forms and relations, therefore what is considered a noise in one candidate form may not be so in the other. The solution for this problem is the implementation of *Formal Statistics* in the form of *thresholds* of:

- Correlation Factors.
- Standard Deviation.

A.2. A Proposal for Remedy

We can observe that a major part of the limitation of ABACUS and BACON stems from the attempt to ignore the numerical nature of the quantitative discovery process, and hence ignoring the powerful tools that the algebraic processes can provide. We do believe that utilizing some inductive numerical tools is so vital in simulating the human process of quantitative discovery.

The limitations discussed in the previous section could be overriden by the use of the following numerical tools, coupled with a qualitative reasoning unit, explained in 5.2.3:

1. Dimensional Analysis, which would enable ABACUS build the general form of relationships, especially in having addition and subtraction, i.e., having multiple-termed relationship. ABACUS system tried lately to acquire this feature through an exhaustive comparison between all the different combinations of variables and powers. Despite the combinatorial explosion of the technique, ABACUS limits the maximum power of exponentiation by 3, and of course it cannot handle fractional powers. This implementation of dimensional analysis does not need knowledge intensive processing. To form the relationship between the various dimensionless groups coming out of this analysis, we need the application of the following three knowledge-based features.
    a. *Variable classification*, by which we group the variables according to some common features,e.g., variables of environmental, geometrical,or physical properties. This kind of classification is either to be supplied by the user or to be built into the program in case of predefined domain of application.

    b. *Variable Exclusion*, which we use in case we have more than one dimensionless group and we like to study the possible relation between two specific variables. In such case we formulate the dimensionless groups such that each of these variables is exclusively contained in one group (or term). Generally speaking, in case of multiple-termed relationships, we can isolate the parameter under consideration in one side and the rest of the parameters in the other side.

    c. *Physical Interpretation*, which would save us a lot of effort by telling us before hand what are the possible dimensionless groups without calculations, e.g.,if we have the parameters Velocity, Viscosity,and Density we can say that the dimensionless group is Reynold's number.

    The user can deactivate the dimensional analysis algorithm if he is not worrying about the dimensional homogeneity of the proposed relationships, in which case the program produces results similar in form to those produced by ABACUS but in a far more efficient way due to the formal statistical procedures included in it[11].

2. Stepwise Regression Analysis is a tool to form the additive relationship between the dimensional groups we got out of the dimensional analysis. By this tool we are able to construct linear regression, non-linear regression, and polynomial expansions, and the most important feature of the technique is gradual building of the relationship. The stepwise (gradual) building of the relationship means that the relationship is built in stages. At each stage, the system picks, out of the remaining terms, the most influential term with respect to the the parameter under

---

[11] In a trial to compare the performance of ABACUS to that of the stripped down QUDS, we ran the same problem of the law of velocity of particles (vele) on the same machine, vax 11/750. It took 62.4 seconds, CPU time, by ABACUS to get the same answer that QUDS gets in 1.7 seconds. We should not that the problem did not need polynomial expansion, fractional powers of exponentiation, big number of terms, or any other things that ABACUS cannot handle at all.

consideration, and adds it to the relationship. Every stage relationship has its own correlation. Thus we have a trade off between complexity of relationship (by adding new terms) and the expressiveness of the relationship (represented by the correlation). Decision criterion for this trade off should be set by the user.

A.3. The Present Status of QUDS

A.3.1. Program Description

The program has the following parts:

1. *The Heuristic PreProcessor* currently has the following capabilities:
    - Isolating the parameter under consideration in a separate term, in a separate side of the relationship, with a unary power of exponentiation.
    - Sorting the variables according to their categories before the starting the formation of the dimensionally-homogeneous groups.
    - Discarding the redundant variables that could be expressed in terns of the other variables. This task is partly heuristic (background knowledge), and partly analytical (check of singularity).
2. *Dimensional Analysis* Algorithm, DA, which has the following advantages :
    - Assuring the dimensional homogeneity of the various terms of any relationship.
    - Based on the preceeding advantage, we ensure the possible existance of a physical meaning for the relationship, i.e., we transform it from the area of empirical formulae to a the area of laws. Thus we are contributing to the level of our knowledge.
    - By the use of the DA algorithm coupled with Stepwise Regression Analysis, we can use all the basic operators (+, -, *, /) in an eventually, natural form of equations.
    - The DA lgorithm is invoked on request by the user, i.e., if the user isn't worrying about the dimensional homogeneity, he may not permit the activation of the DA algorithm. In such a case, the program will activate the Stepwise Regression Analysis directly which will give similar results to those of ABACUS but in far more <u>efficient</u> way due to the formal statistical measures followed in it.
3. *The Qualitative Reasoning Unit*, QRU, is central control unit of QUDS that checks the plausibility of the generated dimensionally-homogeneous groups according to the domain knowledge and logical deduction. It should also plays the role of generating explanation-based generalizations for the instance at hand in conformance with the

domain knowledge. The current status of development is just having some simple facts about the domain that should not be violated.

4. *Stepwise Regression Analysis*, SRA, which has the following options :

   - Developing Linear and Non-Linear Regression (LR & NLR) equations to correlate any number of terms linearly. Linear regression could be extended to non-linear regression by performing the regression on the logarithms of the terms instead of the terms themselves. By this logarithmic approach, we can develop any kind of curvilinear relationships including exponential relationships.

   - Polynomial Expansion (PE) of one argument in terms of another one. The definition of the series could be stored previously and it could be called later on.

   - The Stepwise Regression constructs the relationship gradually by adding one new term at a time to the previous stage of relation construction. At every stage the algorithm picks out of the remaining terms the most influential one and then adds it to the relation if its contribution to the total correlation exceeds a minimum threshold. At the end of every stage, the relationship is written down with its degree of correlation. This feature helps the evaluation of the impact of each individual term on the enhancement of the correlation. Thus, some of the terms could be discarded if their contribution does not overweigh the complexity added by the term.

5. *The Feedback Unit* is not implemented yet, as it cannot work but after building a fairly integrated system.

N.B.:Each of the above-mentioned numerically inductive algorithms (or their options) could be used separately.

A.3.2. Input Procedure <inphase>

A.3.3. Input Files

The nominal input file <case1> represents the physical input file <cases>, which should have the following features :

1. The case title should be in the second line and should not exceed 60 characters .

2. The names of the fundamental units (or dimensions)should be in the fourth line. Each name should not exceed one character for the time being.

3. Each of the following lines corresponds to one variable. The variable name comes in

the first column (1 character), then come the dimensions of that variable in the same order that was used for writing the units'names.

4. The title <Data (or event) Set> preceded by <* > denotes the beginning of the event readings.

5. The following line should have the number of events (readings) <setsize> unpreceeded by any other character.

6. The next line includes the variables' names as column headers for the the data that will follow.

7. The next line is not to be read. It is for improving the readability of the input file for the humans.

8. Each of the following lines corresponds to one event, with all the values of the variables to be listed under the corresponding column headers.

A.3.4. The Order of Input

The variables should be listed in the following order :

1. In case of using the Dimensional Analysis Algorithm, The variables should be ordered in accordance with its importance, such that the most important variable comes last.

2. In case of not using DA algorithm or using it with the option of polynomial expansion, the order of variables is the reverse of the above-mentioned one.

A.4. The Structure of the Program

The program is mainly composed of the following procedures :

- <inphase>
  : to read the input data from both the screen and the physical input file <cases>.

- <nonvanish>
  : to check the non-singularity of the matrix of the exponents or any of its submatrices, as the singularity means that one of the dimensions is not independent and could be derived from the other dimensions.

- <evaldet>
  : to get the value of the determinant.

- <submatrix>
  : to get a submatrix out of a bigger matrix by eliminating a specific row and a specific column.

- <formgroups>
  : to solve the matrix of exponents in accordance with the formation criteria set by <preference>, getting in the end some dimensionless [or dimensionally-homogeneous] groups composed out of the initial arguments.

- <preference>
  : to set the criteria according to which the dimensional groups will be formed in <formgroups>. This procedure will be the field for most of the knowledge-based techniques that will be utilized in the ultimate version of Automatic Calibrative Updating System.

- <regression>
  : to manipulate the data'(or event') set statistically, as a preprocessor for the final stepwise regression process <gauss>.

- <correlation>
  : to perform the correlation calculations between the various variables. This process is a preprocessor for the <regression> procedure.

- <gauss>
  : to perform the stepwise regression analysis upon the correlation matrix produced by <correlation>.

A.5. Limitations for Future Work

1. Incorporation of Nominal Arguments

   We do believe it is a quiet easy task to make the program perform in the following sequence:

   a. If the value of the nominal argument is constant all over the data set, the program will ignore this argument completely and goes ahead in its usual procedures.

   b. If the nominal argument has several values, the program will divide the data set into several subsets each of which has a constant value of the nominal argument. Then the program deals with each subset as an individual subproblem; otherwise the program ignores the nominal argument and proceeds into its regular procedures.

2. The Dimensional Analysis Algorithm should be more flexible in the formation of the dimensionally-homogeneous groups. The following are the specific prospects of this proposed flexibility features :

   a. The procedure <preference> should be more user friendly and interactive for accepting proposals from the user and checking their plausibility.

   b. The DA Algorithm should have a knowledge base that will enable it suggest different advices for the user either in a general way or in a domain-specific way.

3. The program should be able to decide the best form of the relationship whether it is with or without D.A. and deciding which option in each of the two techniques to be used.
4. The program should use the Aq algorithm as a postprocessor for discovering the common characteristics of each group of the relationships if such a classification is not available in the domain knowledge.
5. The program should be able to develop multi-phased broken relationships. Indeed, all we need is getting the first partial derivatives of the relationship and study the continuity (topology) of the function.
6. The program should be integrated into bigger programs that use Data Acquisition techniques heavily, so ACUS gets a continuous (and automatic) supply of data sets upon which the program can work.
7. The program should study the all the possible combinations of the different terms. The caveat of this feature is the combinatorial explosion as what happens in ABACUS. The solution of such a problem is achieved through the development of a knowledge base that will control the introduction of new combinations.
8. In case of discovering the dependency of the supposedly-independent dimensions, the program should be able to eliminate one of the dimensions -according to some rules- and then continues its tasks.
9. The program should have an automatic procedure to decide the powers to which the dimensional terms will be raised, to enhance the expressiveness. A half-baked idea is the trial of each term (which is a dimensionally-homogeneous group) raised to the powers of 1,2,and 3 with the rest of terms.

A.6. Examples

In this appendix, several examples are listed to exhibit the various capabilities of the program. Different combinations of the options of the major algorithms are tried separately as listed before each example.

A.6.1. Example for demonstrating Stepwise Regression Analysis alone

We can see that by firing SRA alone, we got results similar to those of ABACUS but with much higher performance, time-wise. By the observation of the total correlation at each construction stage, we can evaluate manually or automatically the contribution of each argument. The example is a real one for the readings of a slaughter house (Smillie, 1966).

```
Carcass cut-out experiment.

     $     f    l
Y    0     0    0
A    0     1    0
B    0     0    1
C    0     0    0

* Data Set :
20
Y       A     B     C

28.29  160  1.40  31.9
30.16  154  1.13  36.0
30.01  153  1.30  36.1
28.64  155  1.47  32.4
28.86  146  1.23  33.2
29.23  156  1.43  33.4
28.21  158  1.47  30.9
28.23  154  1.50  31.2
28.60  158  1.77  32.3
28.24  160  1.73  31.4
28.93  168  1.43  32.7
28.87  155  1.30  33.1
28.20  162  1.57  31.7
28.72  157  1.60  32.3
28.01  151  1.37  30.4
29.47  157  1.27  34.6
29.54  150  1.23  34.3
30.27  152  1.07  37.3
28.60  152  1.57  31.9
26.55  162  1.70  27.9


Case Title : Carcass cut-out experiment.

# of arguments          = 4
Dimensional Analysis    = 'no'

The data (or event) set is :

   2.8e+01    1.6e+02    1.4e+00    3.2e+01
   3.0e+01    1.5e+02    1.1e+00    3.6e+01
```

```
   3.0e+01   1.5e+02   1.3e+00   3.6e+01
   2.9e+01   1.6e+02   1.5e+00   3.2e+01
   2.9e+01   1.5e+02   1.2e+00   3.2e+01
   2.9e+01   1.6e+02   1.4e+00   3.3e+01
   2.8e+01   1.6e+02   1.5e+00   3.1e+01
   2.8e+01   1.5e+02   1.5e+00   3.1e+01
   2.9e+01   1.6e+02   1.8e+00   3.2e+01
   2.8e+01   1.6e+02   1.7e+00   3.1e+01
   2.9e+01   1.7e+02   1.4e+00   3.3e+01
   2.9e+01   1.6e+02   1.3e+00   3.3e+01
   2.8e+01   1.6e+02   1.6e+00   3.2e+01
   2.9e+01   1.6e+02   1.6e+00   3.2e+01
   2.8e+01   1.5e+02   1.4e+00   3.0e+01
   2.9e+01   1.6e+02   1.3e+00   3.5e+01
   3.0e+01   1.5e+02   1.2e+00   3.4e+01
   3.0e+01   1.5e+02   1.1e+00   3.7e+01
   2.9e+01   1.5e+02   1.6e+00   3.2e+01
   2.9e+01   1.6e+02   1.7e+00   2.8e+01
The possible solutions are :

(1)  Y =  -15.264 +    0.283 X1
                                 Correlation = 79.3

(2)  Y =  -10.252 +    0.255 X1 -    0.426 X2
                                 Correlation = 80.0

(3)  Y =  -10.132 +    0.254 X1 -    0.512 X2 +    0.006 X3
                                 Correlation = 80.1
```

A.6.2. Example for D.A. and Polynomial Expansion

```
Polynomial relation:  Y = X^3 - 2X^2 + 3X - 4

var\unit->
Y
X

* Data set
15
Y         X

86.00     5
4.487     2.3
-3.759    .085
-4.000    0
12740     24
-10.00    -1
-2056     -12
-4.321    -.1
1118      11
2966      15
2.000     2
-2.000    1
127600    51
-57880    -38
-3182     -14
```

```
Case Title : Polynomial relation:  Y = X^3 - 2X^2 + 3X - 4

The data (or event) set is :

   8.6e+01   5.0e+00
   4.5e+00   2.3e+00
  -3.8e+00   8.5e-02
  -4.0e+00   0.0e+00
   1.3e+04   2.4e+01
  -1.0e+01  -1.0e+00
  -2.1e+03  -1.2e+01
  -4.3e+00  -1.0e-01
   1.1e+03   1.1e+01
   3.0e+03   1.5e+01
   2.0e+00   2.0e+00
  -2.0e+00   1.0e+00
   1.3e+05   5.1e+01
  -5.8e+04  -3.8e+01
  -3.2e+03  -1.4e+01

The possible solutions are :

(1)  Y = -569.514 +     0.9775 X3
                                   Correlation = 0.9990
```

```
(2)  Y =    3.064 +    1.0016 X3 -    2.0216 X2
                                   Correlation = 1.0000

(3)  Y =   -3.969 +    1.0000 X3 -    2.0005 X2 +    2.9956 X1
                                   Correlation = 1.0000

        X2 = X1^2

        X3 = X1^3
```

A.6.3. Example for Dimensional Analysis & R.A.

This example illustrates the "Law of Vibration of Strings". It has one degree of freedom (# of arg. - # of fundamental units = 1).

```
Frequency of vibrating rope -- f = .5 * (u^-.5) * l^-1 * (T^.5)

var\unit-> m   l   t
u          1  -1   0
l          0   1   0
T          1   1  -2
f          0   0  -1

* Data set
10
u     l     T     f

.10   6     61    2.058
.30   3     72    2.582
.05   10    39    1.396
.15   12    56    0.805
.05   5     77    3.924
.25   9     89    1.048
.20   8     43    0.916
.10   12    96    1.291
.30   7     91    1.244
.15   10    80    1.155
```

```
Case Title : Frequency of vibrating string :
        f = .5 * (u^-.5) * l^-1 * (T^.5)

The data (or event) set is :

   1.0e-01   6.0e+00   6.1e+01   2.1e+00
   3.0e-01   3.0e+00   7.2e+01   2.6e+00
   5.0e-02   1.0e+01   3.9e+01   1.4e+00
   1.5e-01   1.2e+01   5.6e+01   8.1e-01
   5.0e-02   5.0e+00   7.7e+01   3.9e+00
   2.5e-01   9.0e+00   8.9e+01   1.0e+00
   2.0e-01   8.0e+00   4.3e+01   9.2e-01
   1.0e-01   1.2e+01   9.6e+01   1.3e+00
   3.0e-01   7.0e+00   9.1e+01   1.2e+00
   1.5e-01   1.0e+01   8.0e+01   1.2e+00

Number of Dimensional Groups :     1


Dimensional Group # 1 :

power of the argument u =   0.500
power of the argument l =   1.000
```

```
power of the argument T =  -0.500
power of the argument f =   1.000

THE REGRESSION EQUATION IS :

  0.50      1.00     -0.50      1.00
u       * 1       * T       * f        =     0.500
```

A.6.4. Example for D.A. & R.A.

This made-up example is made up to exhibit the capability of the program to manipulate any kind of data and to propose the most appropriate relationship that expresses it.

```
Dimensional Analysis, Three Degrees of Freedom

var\unit-> w   x   y   z
A          1   2   0  1.5
B          2   5   8   1
C          3   1   1   3
D          4   0   2  2.4
E          5   1  0.5  5
F          6   9   3   8
G          7   1   6   2

* Data set
10
A     B     C     D     E     F    G

1     0.3   0.5   0.3   1     1    240.37
0.02  0.1   0.08  0.02  0.05  0.03 0.365
0.6   0.8   0.1   0.1   0.02  0.5  115.066
0.08  0.09  1.3   0.2   0.4   0.5  0.0474
0.1   1.2   0.45  0.98  2     0.1  5072.38
1     0.2   0.3   0.35  0.25  1    722.79
0.5   0.3   0.2   1.1   0.4   0.5  142410
0.04  0.01  0.1   0.03  0.03  0.1  0.747766
0.02  0.05  0.04  0.06  0.07  0.15 2257.37
1     0.22  0.31  0.13  0.7   0.9  76.8


Case Title : Dimensional Analysis, Three Degrees of Freedom

The data (or event) set is :

1.0e+00  3.0e-01   5.0e-01   3.0e-01   1.0e+00   1.0e+00  2.4e+02
2.0e-02  1.0e-01   8.0e-02   2.0e-02   5.0e-02   3.0e-02  3.6e-01
6.0e-01  8.0e-01   1.0e-01   1.0e-01   2.0e-02   5.0e-01  1.2e+02
8.0e-02  9.0e-02   1.3e+00   2.0e-01   4.0e-01   5.0e-01  4.7e-02
1.0e-01  1.2e+00   4.5e-01   9.8e-01   2.0e+00   1.0e-01  5.1e+03
1.0e+00  2.0e-01   3.0e-01   3.5e-01   2.5e-01   1.0e+00  7.2e+02
5.0e-01  3.0e-01   2.0e-01   1.1e+00   4.0e-01   5.0e-01  1.4e+05
4.0e-02  1.0e-02   1.0e-01   3.0e-02   3.0e-02   1.0e-01  7.5e-01
2.0e-02  5.0e-02   4.0e-02   6.0e-02   7.0e-02   1.5e-01  2.3e+03
1.0e+00  2.2e-01   3.1e-01   1.3e-01   7.0e-01   9.0e-01  7.7e+01

Number of Dimensional Groups :    3


Dimensional Group # 1 :
```

```
power of the argument A =  -0.152
power of the argument B =   0.166
power of the argument C =  -1.525
power of the argument D =  -0.151
power of the argument E =   1.000
power of the argument F =   0.000
power of the argument G =   0.000

Dimensional Group # 2 :

power of the argument A =  -2.951
power of the argument B =  -0.318
power of the argument C =  -1.507
power of the argument D =   0.527
power of the argument E =   0.000
power of the argument F =   1.000
power of the argument G =   0.000

Dimensional Group # 3 :

power of the argument A =  -1.229
power of the argument B =  -0.251
power of the argument C =   2.713
power of the argument D =  -3.352
power of the argument E =   0.000
power of the argument F =   0.000
power of the argument G =   1.000
The possible solutions are :

(1)  Y = 3264.074 +    1.0086  X2
                                     Correlation = 0.9999

(2)  Y =   17.691 +    0.9980  X2 +  976.2998 X1
                                     Correlation = 1.0000


   where :
             -2.95     -0.32     -1.51      0.53     1.00
     X1 = A       * B       * C       * D       * F       *

             -0.15      0.17     -1.52     -0.15     1.00
     X2 = A       * B       * C       * D       * E       *

             -1.23     -0.25      2.71     -3.35     1.00
     Y  = A       * B       * C       * D       * G
```

A.6.5. Example for D.A.,R.A., and P.E.

This example illustrates the "Velocity Law" implementing Dimensional Analysis, Regression Analysis, and Polynomial Expansion.

```
Newton's 2nd Law of Motion:

var\unit-> l   t
V          1   -1
g          1   -2
d          1    0
v          1   -1

* Data Set
10
V           g     d     v

80.331      9.81  328  4.2
123.695     9.81  620  -56
67.824      9.81  63   58
4.193       9.81  .89  -.35
26.205      9.81  35    .12
44.433      9.81  100  3.5
78.251      9.81  2    -78
44.235      9.81  85   17
47.227      9.81  28   -41
18.885      9.81  -4.3 21

Case Title : Newton's 2nd Law of Motion:

The data (or event) set is :

   8.0e+01   9.8e+00   3.3e+02   4.2e+00
   1.2e+02   9.8e+00   6.2e+02  -5.6e+01
   6.8e+01   9.8e+00   6.3e+01   5.8e+01
   4.2e+00   9.8e+00   8.9e-01  -3.5e-01
   2.6e+01   9.8e+00   3.5e+01   1.2e-01
   4.4e+01   9.8e+00   1.0e+02   3.5e+00
   7.8e+01   9.8e+00   2.0e+00  -7.8e+01
   4.4e+01   9.8e+00   8.5e+01   1.7e+01
   4.7e+01   9.8e+00   2.8e+01  -4.1e+01
   1.9e+01   9.8e+00  -4.3e+00   2.1e+01

Number of Dimensional Groups :    2


Dimensional Group # 1 :

power of the argument V =  -2.000
power of the argument g =   1.000
power of the argument d =   1.000
power of the argument v =   0.000
```

```
Dimensional Group # 2 :

power of the argument V =  -1.000
power of the argument g =   0.000
power of the argument d =   0.000
power of the argument v =   1.000
The possible solutions are :

(1)  Y =    0.500 -    0.5000 X2
                                  Correlation = 1.0000


   where :
               -2.00     1.00     1.00
       Y  = V       * g      * d      *

               -1.00     1.00
       X1 = V       * v


       X2 = X1^2
```

A.6.6. Example: Longitudinal Cracks due to Thermal Impact

This example demonstrates how useful QUDS could be for discovering new relationships. The study of the different possible relevant variables was steered manually as the system is not augmented yet to a real source of data. However, the resulting relationship is physically meaningful, syntactically elegant, and the more important is that it is statistically sound.

```
Longitudinal Cracking due to Thermal impact:

variable\unit-> m   l   t   d
H               1   1  -2   0   (Equiv. thickness =Thick. * Strength)
A               0   0   1   1   (Age of pavement * Ave.Ann.Temp.)
S               1  -1  -2   0   (Strength of Surface course, ksi)
T               0   2   0   1   (Surface Heat, 1000deg.sq.ft)
L               0   1   0   0   (Accumulated Longitudinal Cracks, ft)
C               1   0  -2  -1   (Heat Capacity per unit volume / mile,
                                 lb/sec2/deg.)

* Data set
10
H       A       S       T       L       C       F

15      2       120     -2      7       6.1     1
14      12      100      10     15      6.7     1
18      5       130     -20     25      5.8     1
14      4       90       25     40      7.4     1
11      3       110      35     20      7.6     1
```

```
20      6       125     45      38      4.9     1
16      5       122     0       5       6.8     1
25      4       115     15      0       7.2     1
10      9       102     12      45      6.5     1
15      7       95      23      55      5.2     1

Number of arguments: 6
Dimensional Analysis: y
Number of fundamental units: 4
Minimum enhancement: 1 %

Case Title : Longitudinal Cracking:

The data (or event) set is :

   1.5e+01   2.0e+00   1.2e+02  -2.0e+00   7.0e+00   6.1e+00
   1.4e+01   1.2e+01   1.0e+02   1.0e+01   1.5e+01   6.7e+00
   1.8e+01   5.0e+00   1.3e+02  -2.0e+01   2.5e+01   5.8e+00
   1.4e+01   4.0e+00   9.0e+01   2.5e+01   4.0e+01   7.4e+00
   1.1e+01   3.0e+00   1.1e+02   3.5e+01   2.0e+01   7.6e+00
   2.0e+01   6.0e+00   1.3e+02   4.5e+01   3.8e+01   4.9e+00
   1.6e+01   5.0e+00   1.2e+02   0.0e+00   5.0e+00   6.8e+00
   2.5e+01   4.0e+00   1.2e+02   1.5e+01   0.0e+00   7.2e+00
   1.0e+01   9.0e+00   1.0e+02   1.2e+01   4.5e+01   6.5e+00
   1.5e+01   7.0e+00   9.5e+01   2.3e+01   5.5e+01   5.2e+00

Number of Dimensional Groups :    2


Dimensional Group # 1 :

power of the argument H =  -0.500
power of the argument A =   0.000
power of the argument S =   0.500
power of the argument T =   0.000
power of the argument L =   1.000
power of the argument C =   0.000

Dimensional Group # 2 :

power of the argument H =  -1.500
power of the argument A =   0.000
power of the argument S =   0.500
power of the argument T =   1.000
power of the argument L =   0.000
power of the argument C =   1.000
The possible solutions are :

(1)  Y =    53.248 + 0.77631302 X1
                                          Correlation = 0.1485

(2)  Y =    43.509 + 2.45275252 X1  - 0.00037467 X3
                                          Correlation = 0.4396
```

```
(3)  Y =    19.122 + 0.99840349 X1 - 0.00199116 X3 + 0.14637679 X2
                                    Correlation = 0.6848

(4)  Y =     5.864 + 4.27666254 X1 - 0.01146692 X3 + 0.23041859 X2
                                                   + 0.00010266 X4
                                    Correlation = 0.8368


   where :
            -0.50    0.50     1.00
       Y  = H      * S      * L

            -1.50    0.50     1.00     1.00
       X1 = H      * S      * T      * C


       X2 = X1^2

       X3 = X1^3

       X4 = X1^4
```

A.6.7. Demonstrative Example: Longitudinal Cracks due to Thermal and Traffic impacts

This example demonstrates how the different parts of QUDS work in tackling a new case. In order to demonstrate the mechanism of discovery, we assume that the system has neither a sufficiently similar case in its *Case Library*, nor a relationship that predicts the development of longitudinal cracks. In this example, QUDS picks the relevant variables from a big number of variables. The criterion for picking up the variables is their existence in the *CAUSED-BY* and *MAYBE-RELATED-TO* fields in the PMORPH frames of the symptoms reported in the case under consideration. The causality triggers the states recursively till it reaches a fact or input.

1. *Input:*
   We use an exemplary case which has only one manifestation of distress, longitudinal cracks. For more details about the case, refer to the following case file.

```
case:                    Route 128, Milepost 123.00->128.00
Type of pavement:        Flexible
AADT:                    15,000 veh.per lane
% Trucks  :              24 %      Ave.Speed          :  49   mph
Lane width:              12'       Subgrade Texture   :  Gravel(WG)
# of lanes:              3         Dry Density        :  135  pcf
Shoulder width:          8'        Corrected CBR (%) :  86,37,5 %
Pavement thickness:      8"        Surface Stiffness :  120  ksi
Base thickness:          15"       Base Elas.Modulus :  20   ksi
CBR:                     3.0       Subgrade Modulus   :  1500 psi
Temprature variation: Summer 85, Fall 60, Winter 30,Spring 50
```

<u>Symptoms</u>	<u>Severity</u>
Long.Cracks	6 ft

2. *PMORPH:*

The only symptom mentioned in the input case file is longitudinal cracking. Therefore the state frame *LONG-CRACK* is triggered, refer to 5.2.2. The backward chaining ,through recursive causality, triggers other state frames including the computational frames (daemons) shown hereafter.

STATE: EQUIV.THICKNESS
STATE-TYPE: COMPUTATIONAL
CLASS: STRUCTURAL
REPRESENTATIVE-VAR: H
VARIABLE-DIMENSIONS: m(1), l(0), t(-2)
VALUE: nil
EQUALS: Surface-Stiffness * Surf-Thick
+ Base-Elas-Modulus * Base-Thick
+ Subgrade-Modulus * Sg-Equiv-Thick

STATE: SURFACE-HEAT
STATE-TYPE: COMPUTATIONAL
CLASS: ENVIRONMENTAL
REPRESENTATIVE-VAR: SH
VARIABLE-DIMENSIONS: l(2), deg(1)
VALUE: nil
EQUALS: Ave-Ann-Temprature * #-Lanes * Lane-Width * 5280

STATE: ACCUM-LOAD/YR
STATE-TYPE: COMPUTATIONAL
CLASS: TRAFFIC
REPRESENTATIVE-VAR: X
VARIABLE-DIMENSIONS: m(1), l(1), t(-2)
VALUE: nil
EQUALS: AADT * 2 * Equiv-Axle-load

3. *QUDS:*

After preparing the suspectedly-relevant (required) variables either by direct retrieval or daemonic computattion, which was explained in 5.2.2.1, The candidate variables moves into QUDS to have the following processing:

- The *Heuristic PreProcessor* classify the input variables into two groups, a structural group and an environmental (thermal) one. Still the groups may contain some irrelevant variables.

- The *Dimensional Analysis Unit* forms, at least one, dimensionally-homogeneous group out of every group classified by the heuristic preprocessor. Thus we get one environmental (thermal) dimensionally-homogeneous group, and two dimensionally-homogeneous structural groups one of which has the *REPRESENTATIVE-VARIABLE* of longitudinal cracking isolated and raised to the unary power. Such that, it is possible to have this latter group in the left hand side of the proposed relationship.

- The *Qualitative Reasoning Unit* has no facts or rules that are relevant to this case yet. Therefore, the default response from it is acceptance.

- The *Stepwise Regression Analyser* takes the values of the current case as an instance to be added to the other instances in the *Case Library*. We Assume that that case library had 11 other cases before. Thus the stepwise regression analyzer starts working on 12 cases. It builds the relationship gradually, in two stages. It shows us that the thermal properties have the major impact upon the development of longitudinal cracks. In the second degree, come the structural properties.

4. The *Feedback* facility is not implemented yet.

```
Longitudinal Cracking due to Thermal & Traffic Impacts:

variable\unit-> m   l   t   d
H               1   0  -2   0   (Equiv.Thickness, ksi.in.)
A               0   0   1   0   (Age of Asphalt, yrs)
S               1  -1  -2   0   (Strength of asphalt, ksi)
T               0   2   0   1   (Surf.Heat, 1000deg.sq.ft)
C               1   0  -2  -1   (Sect.Heat Capacity, ksi.mi./deg.)
X               1   1  -3   0   (Accum.Ld/yr , 1000 ton/yr)
L               0   1   0   0   (Accum.Length of Cracks, ft)

* Data set
12
H       A       S       T       C       X       L

15      2       120     -2      5.1     250     7
14      12      100     6       6.7     90      15
18      5       130     20      5.8     200     25
14      4       90      55      7.4     60      40
11      3       110     35      5.6     200     20
16      6       125     45      6.9     150     38
20      5       122     0       5.8     70      5
25      4       115     -5      6.2     80      0
10      9       102     22      6.5     350     45
12      7       95      33      7.1     380     55
16      6       140     53      7.2     200     30
12      3       190     57      7.5     290     58

Number of arguments: 7
Dimensional Analysis: y
Number of fundamental units: 4
Polynomial Expansion: n
```

```
Minimum enhancement: 1 %

Case Title : Longitudinal Cracking due to Thermal & Traffic Impacts:

Number of Dimensional Groups :    3


Dimensional Group # 1:

power of the argument H =  -3.000
power of the argument A =   0.000
power of the argument S =   2.000
power of the argument T =   1.000
power of the argument C =   1.000
power of the argument X =   0.000
power of the argument L =   0.000

Dimensional Group # 2:

power of the argument H =  -2.000
power of the argument A =   1.000
power of the argument S =   1.000
power of the argument T =   0.000
power of the argument C =   0.000
power of the argument X =   1.000
power of the argument L =   0.000

Dimensional Group # 3:

power of the argument H =  -1.000
power of the argument A =   0.000
power of the argument S =   1.000
power of the argument T =   0.000
power of the argument C =   0.000
power of the argument X =   0.000
power of the argument L =   1.000
The possible solutions are :

(1)  Y =  122.166 +    0.0001 X1
                                     Correlation = 0.8200

(2)  Y =   61.740 +    0.0001 X1 +    0.0001 X2
                                     Correlation = 0.9397


   where :
           -3.00      2.00      1.00      1.00
       X1 = H      * S       * T       * C
          = The thermal impact term.

           -2.00      1.00      1.00      1.00
       X2 = H      * A       * S       * X
          = The structural term.
```

$$Y = H^{-1.00} * S^{1.00} * L^{1.00}$$

= *The LHS, including the parameter under consideration with a unary power of exponentiation.*

Diagnosis: For implementational purposes, we set the diagnosis mechanism as follows: If the deterioration manifestation, predicted from the resulting relationship, did not deviate more than 20% from the actual deterioration, then give the causal path used in picking up the relevant variables as the diagnosis of the case.

Appendix B
Pi-Theorem

In this appendix, we try to illustrate the theoritical basis upon which the dimensional analysis is built. Dimensional analysis plays an important role as a numerically inductive tool in QUDS. Therefore we should pay a closer look upon the way the system forms the dimensionally-homogeneous groups. The whole technique is based upon Π-theorem (Bridgman, 1931).

If we have a complete[12] equation $\varphi(\alpha,\beta,\gamma,.....)=0$,with (n)variables and (m)fundamental, Then its solution has the form:

$$F(\Pi_1,\Pi_2,........)=0$$

where the Π's are the n-m products of the arguments α,β,...etc., which are in the fundamental units (Bridgman, 1931).

The typical Π has the form:

$$\Pi=\alpha^a\beta^b\gamma^c.....$$

where a,b,c,..etc. are chosen such that Π is dimensionless. Substituting the fundamental dimensions[13] (units)(U_1,U_2,U_3,....)for the variables α,β,γ,... ,we get as example:

$$\alpha=U_1^{\alpha_1}U_2^{\alpha_2}U_3^{\alpha_3}.....U_m^{\alpha_m}$$

where $\alpha_1,...\alpha_m$ are the dimensions (or the exponents) of the variable α in the fundamental units.

Thus we get (m) equations, each with (n)terms(unknowns):

---

[12] A complete equation is an equation independent of the units used to measure its dimensions.

[13] A possible set of fundamental dimensions may be composed of mass(M),length(L),and time(T)

$$\alpha_1 a+\beta_1 b+\gamma_1 c+....=0$$
$$\alpha_2 a+\beta_2 b+\gamma_2 c+....=0$$
$$..$$
$$\alpha_m a+\beta_m b+\gamma_m c+....=0$$

In general n will be greater than m. Thus there will be n-m independent sets of solutions, i.e.,there will be n-m independent dimensionless products and the arbitrary function F will be a function of n-m variables. The above-mentioned equations could be solved simultaneously by the use of any of the algebraic techniques.

B.1. Conditions

1. The theorem, which is the core of the Dimensional Analysis, is based upon the principle of Dimensional Homogeneity. This principle restricts the possible candidate relationships between the variables under consideration very effectively. It may be considered a good way to limit the combinatorial explosion of the candidate relationships.

2. The fundamental units (dimensions) should be independent. This condition could be checked by getting a *non-vanishing determinant of exponents*[14]. Therefore, a check of the exponents'determinant is very essential at the beginning of the analysis. If we got a zero value for the determinant, we should revise our set of fundamental units before applying the theorem otherwise,we should analyze the problem manually.

B.2. Limitations of Pi-Theorem

$\varphi(\alpha,\beta,......)=0$ is assumed to be the only relation between the variables included in it, otherwise, the partial differentiation used in the derivation of the theorem does not hold. Such assumption may hinder the approach from expressing the multi-phased relationships. We can override this assumption by the use of an integer variable in which case we divide the problem into several sub-problems according to its phases.

---

[14] formed from the last set of equations

Index