Final Report

submitted to

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
GEORGE C. MARSHALL SPACE FLIGHT CENTER, ALABAMA 35812

December 31, 1990

for NAS8-36955 Delivery Order # 64

entitled

Expert Systems for Superalloy Studies

by

Gary L. Workman, PhD.
and
William F. Kaukler, Ph.D.
Co-Investigators

H.E.E.E.S
COMPONENTS OF AN EXPERT SYSTEM

THE USER

ES SHELL USER INTERFACE

INFERENCE ENGINE

DECLARATIVE KNOWLEDGE
DATABASE OR FACTS

KNOWLEDGE BASE
OF HEURISTICS

Johnson Research Center
The University of Alabama in Huntsville
Huntsville, Alabama 35899

(NASA-CR-184236) EXPERT SYSTEMS FOR
SUPERALLOY STUDIES Final Report (Alabama
Univ.) 41 p

Unclas
G3/26 0321272
TO: Center for AeroSpace Information  
Attn: Mr. Phil French  
FROM: CN22/Joyce E. Turner  
SUBJECT: Final Reports, UAH Contract NAS8-36955

As discussed in our phone conversation of October 7, 1991, I am enclosing 27 FF427 Forms on NAS8-36955. These reports were sent to us by CASI for determining availability.

The COTR has reviewed each report and determined that each can be made publicly available. However, as we discussed, he was unable to determine the proper Headquarters office for approval. He has written the enclosed memo to serve as approval for entering each report.

I have assigned NASA CR numbers to each report. If any other information is needed, please call me at FTS 824-4528.

Thanks for your assistance in this matter.

Joyce E. Turner  
Technical Information Specialist

Enclosures  
COTR Memo  
FF 427 Forms
TO: CN22/Joyce Turner
FROM: EM11/Lawrence J. Smith
SUBJECT: Final Reports, UAH Contract NAS8-36955

Enclosed are forms FF427, NASA Scientific and Technical Document Availability Authorization, for the following Delivery Orders: 12, 14, 21, 23, 31, 33, 34, 35, 36, 37, 41, 43, 45, 46, 47, 50, 56, 59, 64, 67, and 69.

After considerable research, it has been determined that it is virtually impossible to name a NASA Headquarters' technical person responsible for the funding on each of these Delivery Orders. This determination is based on the fact that funding for these Delivery Orders involves suballotted, reimbursable, and as far back as eight-year money.

The final reports have been reviewed and accepted by the individual alternate COTR on each Delivery Order. All of these reports are deemed publicly available documents.

Lawrence J. Smith
COTR
NAS8-36955

Enclosures
TABLE OF CONTENTS

1.0 INTRODUCTION ................................................. 1

1.1 USE OF EXPERT SYSTEMS IN MATERIALS SCIENCE APPLICATIONS ................................................. 4

1.2 COMPARISON OF EXPERT SYSTEMS AND CONVENTIONAL SOFTWARE ................................................. 5

2.0 RESULTS ............................................................. 7

2.1 NEXPERT OBJECT DEVELOPMENT ........................................ 8

2.2 PHACOMP COMPUTATIONS ........................................ 11

2.3 STATISTICAL ANALYSES ......................................... 16

2.4 HEEES PROCEDURES ............................................ 21

2.5 USING IXL SOFTWARE ........................................... 25

3.0 SUMMARY OF HEEES PROJECT ..................................... 27

3.1 CURRENT STATUS OF THE PROJECT ............................... 28

3.2 HEEES EXPERT SYSTEM MODULES ................................ 29

3.3 LIST OF FILES PERTINENT TO HEEES ............................ 30

4.0 ACKNOWLEDGEMENTS ............................................. 31

5.0 REFERENCES ..................................................... 31

APPENDIX A. SAMPLE HEEES REPORT .................................. A-1
1.0 Introduction

There are many areas in science and engineering which require knowledge of an extremely complex foundation of experimental results in order to design methodologies for developing new materials or products. Materials science and engineering represent a very fertile discipline for such activity. Many of the research activities in materials science and engineering require knowledge contained in voluminous publications. Some of this knowledge resides in the open literature, some in proprietary documents, and some in the heads of researchers; as yet undisclosed to the rest of the world. Putting this knowledge into one database to assist in the development of new products is obviously a challenge. For that reason, it can be noted that historically some of the methodologies used for developing particular products can appear to either emerge from an alchemist's laboratory or as well thought 'scientific logic', obvious to everyone.

Superalloys are an area which fit quite well into this discussion in the sense that they are complex combinations of elements which exhibit certain characteristics; most desirable, but some undesirable. Highly desirable characteristics primarily include the high strength at high temperature capabilities which provide for most of the appeal for superalloy materials in the aerospace and the nuclear industries. Highly undesirable characteristics can impede the use of superalloys for particular applications and include items such as formation of deleterious phases with particular compositions, corrosion properties, and in some cases, the cost required to produce a given superalloy. Relating the contribution of each entity of the superalloy to the mechanical properties represents a huge undertaking. The trade-offs between each entity and the ultimate properties achieved in the superalloy do not sum as a linear combination of entities.

Obviously the use of superalloys in high performance, high temperature systems such as the Space Shuttle Main Engine (SSME) is of interest to NASA. Figure 1. shows an illustration of the SSME with some superalloy components identified.
Figure 1. Drawing of SSME showing some superalloy implementations

SSME Wrought Alloys
Waspaloy (Disks and Shafts)
INCO 903
A-286
Rene 41

SSME Cast Alloys
DS Mar-M-246+Hf (Turbine Blades)
CC Mar-M-246+Hf
INCONEL 718 (Housing)

Figure 2. Flowchart of manufacturing processes important to superalloys.
The superalloy manufacturing process is complex and the implementation of an expert system within the design process requires some thought as to how and where it should be implemented. Figure 2 shows a pictorial representation of the manufacturing flow from starting materials to final products.

Processing at each phase adds some features to the overall properties of the final product. The details of these processing steps and the resulting characteristics in the superalloys provide knowledge which an expert system can use to help in the design process. Obviously feedback from each of these intermediate, as well as the final stages, should be incorporated into the overall process. A major motivation for this research is to develop a methodology to assist metallurgists in the design of superalloy materials using current expert systems technology.

Figure 3. Overall view of expert system to assist in the design of superalloys.
Although the ultimate final product in this work will eventually be an expert system used to design superalloys for any environment or application, this research project has focused on a very important requirement for SSME uses. Hydrogen embrittlement is disasterous to rocket engines and the heuristics (or rules of thumb) can be very complex. Attacking this problem as one module in the overall design process represents a significant step forward. In order to best describe the objectives of the first phase implementation, we have designated the expert system **Hydrogen Environment Embrittlement Expert System** or **HEEES**.

1.1 Use of Expert Systems in Materials Science Applications

The use of expert systems in various materials science applications has started to show an increase in the last six years. Typically expert systems are beneficial in solving complex problems in which "expert" capability needs to be applied either to assist in focussing on the significant details of the problem or to retain "expertise" of humans who have many years experience in building up a knowledge base that is irreplaceable. In reality, the problem which is to be undertaken in this work, satisfies both of these criteria.

As mentioned above, only in the last six years has significant work been performed in using expert systems for materials science applications. The most relevant to the proposed study are References 83-87, which were published in the proceedings of two conferences hosted by The Metallurgical Society and which were dedicated to materials processing applications of expert systems. The first was held in Orlando in 1986 and the second in Anaheim in 1990. Needless to say, the time difference required for the third conference will be much shorter than the time it took between the first and second conferences.

A systematic approach in the attempt to predict hydrogen embrittlement in Space Shuttle Main Engine Alloys has been undertaken. This "expert systems" approach is
unique in that it tries to take advantage of the data available from many sources about superalloy behavior. Such an approach has not been attempted before.

1.2 Comparison of Expert Systems and Conventional Software

Expert systems are not new. They have been around since the 1960's. The increasing number of applications in which expert systems have been implemented provides the major innovations in expert systems today. For instance expert systems have been developed to assist in the diagnosis of mass spectral data (DENDRAL, 1968), bacteriological blood infections (MYCIN, 1973), geological formations for mineral deposits (Prospector, 1978), and to configure VAX computers (R1, 1983). Recent trends have included many design functions such as integrated circuits and other large complex systems in which the multitude of rules and contraints are just too cumbersome for humans to effectively weed through for a proper design process.

Conventional software normally consists of code which can interrogate a database and perform calculations or massage data from the database. If new data needs to be massaged, frequently new coding needs to be performed to account for the different database contents; or even more frequently, the concepts for the analysis change and the old program needs to be modified to generate a new algorithm for proper determination of the desired goals.

Expert systems offer a different approach to that problem. Figure 4. shows the basic elements contained in a typical expert system.
Another attribute of expert systems is the environment in which expert systems are developed. A number of shells are available in which domain experts can prepare useful elements of knowledge in which the expertise required to solve particular problems can be assembled into one or more knowledge bases. Implementing a strategy for utilization of the knowledge still requires considerable effort.

After reviewing many different shells, NEXPERT OBJECT was chosen to be the development platform for this task. Originally, it was felt that both forward and backward chaining would be necessary in the superalloy design process and that feature was not available in all the expert systems shells reviewed prior to receiving the task order. Another feature that was offered in NEXPERT OBJECT was the graphical environment based on Microsoft Windows. Inherently this feature was a plus in the development phase in that the screens could be transferred into other software products. The delivery package uses the run-time system, which does not provide a graphical
interface and does not share those attributes. However, the finished system does not require all the graphical support used by the development system.

2.0 Results

The end result of this project is the first phase of a software environment using the run-time version of NEXPERT OBJECT which can make some logical inferences from the superalloy database compiled during this contract about the hydrogen embrittlement characteristics of superalloy materials. We have called this module HEEES, as explained above. The HEEES module can be represented graphically by Figure 5 below.

Figure 5. The HEEES architecture.
2.1 NEXPERT OBJECT DEVELOPMENT

NEXPERT OBJECT uses rules and objects to represent knowledge. The objects describe and define the environment in which reasoning will take place, and the rules perform this reasoning through a series of tree-like gates. To apply the expert system to the hydrogen embrittlement problem, the environment must first be defined. After this environment has been established, rules can be implemented to simulate human reasoning and evaluate an alloy in terms of its hydrogen embrittlement.

The expert environment is explicitly defined by objects, their classes, and their properties. The object names a thing in the environment, for example, composition. This object is uniquely defined by its properties: cobalt, chromium, nickel, etc. These properties refer to the weight percent of each element needed to compose the alloy in question. The composition also has a property, name, which allows information about the alloy to be retrieved from the LOTUS spreadsheets.

Objects with similar properties can be grouped together accordingly by classes. One feature of the expert system is that it will search for a set of alloys similar to the one in question. This helps the user associate other alloys which might be useful or detrimental based on the results of the original alloys' evaluation. The similar alloys are objects of a class called similar_alloys. These objects have common properties and the values of these properties determine whether each object is in or out of the class.

An object may also have subobjects. These subobjects have their own unique entities because they represent unique things but are related to the parent object because they are components of that parent object. Properties of the object and the subobject are usually quite different because they identify different things. Examples of object-class-subobject-property association is demonstrated below.

object: A_286
classes: similar_alloys
         wrought
subobjects: mechanical_properties
Within the expert system, data is stored in slots. The type of data stored depends upon the type designated by the property. Properties can be one of several types: boolean, string, integer, floating point, date, or time. The values assigned to the properties are directly tied to the appropriate object with the notation object.prop. Meta-slots are associated with the slots to customize the acquisition of data. Meta-slots can designate where to look for the value and in what order, or they can establish a default value.

Once the environment was established, rules were added to allow the expert system to evaluate alloys. Rules follow the basic format of IF... THEN... and DO.... The IF portion of the rule is a series of conditions which must be evaluated when the rule is initiated by the expert system. If all the conditions are met, the hypothesis, which is in the THEN portion of the rule, is found to be true and the list of actions following the hypothesis are then carried out.
The knowledge base is composed of many rules which can perform many operations. Spreadsheets can be retrieved, updated, and created to transfer data to and from the knowledge base. Diagrams and text can be displayed to present results or explanations. Comparisons and pattern matching operations can be used for various purposes. External routines can be executed to provide more information to the analysis within the knowledge base.

Rules are triggered by association with known information. For example, if the slot microstructure.eutectic_G_Gprime is known, eight rules are immediately placed on the agenda to be evaluated. Most rules in the knowledge base are intertwined so a complete evaluation can be triggered with only a selected amount of initial information. Of course, if a group of rules, called a frame, was not triggered, it very likely was not necessary for the evaluation.
2.2.0 PHACOMP Computations

A popular computational tool has been applied to predicting microstructures of superalloy materials. Several different algorithms of the Phase Computation program (or PHACOMP) have been devised over the years. An implementation of the PHACOMP algorithm has also been prepared to supplement the functionality of the HEEES. This program specifically determines from the composition of the superalloy in question whether or not any deleterious sigma phase could form. There are many versions of the basic routine which attempt to correct the accuracy of prediction for certain alloys which are not properly treated. The prediction is based on the electron energy state density caused by the 3d orbitals in the individual elements. The elements which make up superalloys are, in general, transition metals and electron compound (sigma phase) formation can, to a degree be predicted from these energy densities. The Groups (columns) of the transition elements were assigned the electron valence number, $N_v$, based on Pauling's density of states predictions. Since sigma phase forms from the

Figure 5. Useful elements for superalloy systems.

PERIODIC TABLE DISPLAY OF ELEMENTS USEFUL IN PREPARING SUPERALLOYS

![Periodic Table Display of Elements Useful in Preparing Superalloys]
austenitic matrix in a superalloy, when the weighted sum of the elements' $N_v$'s exceeds a certain value (about 3.6), sigma phase precipitates from the austenite solution.

Concurrent phase diagram studies revealed that sigma in austenitic regions is two-phased and since no sigma is desired, the critical value, $N_v$, for any sigma to form was empirically established at 2.5. PHACOMP permits a prediction of two-phase boundaries in quaternary phase diagrams. Because actual phase diagrams are not employed; yet predictions of phase formations are still made, PHACOMP has been called an early application of Artificial Intelligence\textsuperscript{43}.

### 2.2.1 Sigma Phase Formation

Sigma phase precipitates at the high temperatures of heat treatment. For sigma phase to form, the time and temperature must be sufficient for its precipitation. Some alloys are more difficult than others to form a sigma phase. Also the phase may precipitate in service at high temperatures.

**Figure 6.** Historical trends in understanding phase formations in superalloys.
The PHACOMP program yields a projection of possible sigma phase formation. As is typical of such compounds, with high brittleness and its highly acicular morphology, sigma phase readily embrittles an alloy. For any application, the sigma phase is highly undesirable. As a result, any hypothetical or developmental alloy should be screened for the sigma phase formation or the propensity for it.

2.2.2 PHACOMP Algorithm

In preparing the Super PHACOMP program, many versions of PHACOMP were researched. Variations dealt with temperature of phase formation, alterations of the individual $N_v's$ for certain elements to align the predictions for certain special alloys, use of other (5d) orbitals for similar predictions, and many variants on how to treat non-austenitic phases. Presently, three versions are implemented in the Super PHACOMP program. Since certain calculations are repeated in each, the versions were woven together for efficiency. It is the simple blending of the versions that make Super PHACOMP valuable to the HEEES project. During the PHACOMP calculations, a variety of parameters need to be calculated which have significance on several aspects of Hydrogen behavior in the superalloys.

Whether or not the alloy forms sigma phase is not the only function of the Super PHACOMP program. In this implementation, the routines for PHACOMP that were used came from: Decker et al.\textsuperscript{41}, Sims et al.\textsuperscript{10, 43}, and Morinaga et al.\textsuperscript{52}. The early work published on PHACOMP utilizes the electron vacancy number $N_v$ of the alloy to characterize the tendency for an austenitic alloy to precipitate the detrimental topologically close packed or TCP phases. The last is the New-PHACOMP (from Japan) which uses $M_d$ values instead of the $N_v$, where $M_d$ represents the average energy levels of d orbitals of the alloying transition metals. As mentioned above the use of $N_v$ as a characteristic parameter for metals was first presented by Pauling\textsuperscript{82}, whereas the new parameter $M_d$ correlates with the electronegativity and atomic size in the classical
approach of the Hume-Rothery theory of metals. The standard form adopted was from Sims et al given in the Superalloys II book. Two versions were given by Decker et al, and it is the version of Woodyatt/Sims/Beattie that was selected. In addition to the PHACOMP numbers which are characteristic of each alloy, the program determines if the alloy will form sigma phase or not based on the \( N_v \) or \( M_d \) and the alloy type.

In all cases, preliminary calculations need to be made for phases that do not contribute to the austenite-to-sigma transformation. The appropriate carbides, borides and gamma-prime phases need to be accounted for and the elements used to form these phases removed from the overall constitution to leave the gamma phase composition from which the PHACOMP parameter or number, \( N_v \), can be determined. This is the result of having determined sigma phase could form from simple austenitic matrices if the \( N_v \) was of the appropriate value. The commercially interesting alloys have all the other phases present as well. This is where the limiting assumptions for PHACOMP begin. It was determined early in the development of PHACOMP that the precipitated phases did not react with the matrix. The secondary precipitate composition and the amount of elements consumed in their formation needed to be determined. This is the major area of interest for PHACOMP. These very calculations of phases formed are valuable to the determination of other attributes in the question of HEE. Microstructural data was lacking in the amounts of the carbides, borides and gamma-prime found in the alloy in question. Also, for a hypothetical alloy, these values could not be known a priori.

Several references identified stacking fault energy (SFE) as a significant mechanism in controlling the strength and resistance to hydrogen embrittlement of austenitic alloys. The SFE determines whether cross slip or coplanar motion of dislocations predominates. Superalloys have a wide range of SFE's and thus different deformation characteristics. Hydrogen is known to move, or be tied to,
dislocations\textsuperscript{40}. When the austenite phase has a low SFE, coplanar dislocation motion occurs and leads to dislocation pileups at obstacles to their movement. Higher SFE's permit cross slip and the pileups are avoided. Hydrogen concentrations are expected to be higher in these pileups since the dislocations carried atoms of hydrogen there. Molecular hydrogen could collect in these areas and embrittle the material once the stresses are relaxed. From our calculated gamma phase compositions and the equations given by Schramm and Reed\textsuperscript{50}, SFE's for the alloys are evaluated. A series of calculations are made according to the constitutional criteria set out for the austenitic composition. Comparison between these gives a degree of confidence in the calculation. At least one SFE is calculated for all alloys. The value of SFE is employed in HEEES for HEE evaluation. Since some HEE effects depend on knowing whether the alloy is Ni or Co or Fe based, the Super PHACOMP program identifies the alloy appropriately. This information is sent back to HEEES.

Based on a paper by Kusunoki et al \textsuperscript{73}, PHACOMP calculations are performed and supplemented with a Solution Index (SI) parameter. The SI represents the degree of solid solution strengthening available to the gamma-prime phase. It is with the amount of gamma-prime and it's SI that they claim (within limits outlined in the paper) to be able to predict the strength of the alloy. They and others\textsuperscript{46, 10, 41} note the superalloy strength is greatly controlled by the amount of gamma-prime. So too are the high temperature creep and corrosion resistance. The creep rupture life and the calculated tensile strength for the alloy are evaluated in Super-PHACOMP according the Kusunoki et al paper. If the alloy falls within the bounds they set aside, dealing with the PHACOMP number and SI number, then the calculated strength and creep rupture life will have a high confidence level. These calculations are fed into HEEES.
2.3 STATISTICAL ANALYSES

By definition, the Expert System (ES) is not capable of creating knowledge from the data made available to it; however, it can be used to extract information or infer results or relationships not intuitively obvious to everyone. It is instead more like a database where special information is kept and organized. This special information that will give value and apparent intelligence to the ES can be derived from a statistical analysis of the data collected or otherwise obtained. It is the knowledge-base made up of expert-derived facts in the form of rules that distinguishes the ES from a common database. The rules that make up the ES represent knowledge about a certain subject obtained from an expert on that subject. These rules are factual statements that can be made about some aspect of that subject. A rule may represent a feeling or educated guess that an expert expresses without justification. It may be that experience has shown that a certain idea is true but without direct causality.

Statistical analysis of groups or sets of data can provide rules similar to those of an expert. A correlation between two sets of data may be plausibly explained by some underlying common feature. Often, the correlation compels one to think of the cause for the correlation and thereby discover the communality. Experience and expertise play a significant role in determining the causes for the correlations. However, knowing the cause for the correlation is not a requirement for finding the correlation nor does it prevent the use of the correlation as a form of experience without causality. So, whether the reason for a correlation is known or not, one can use this statistical relationship in an ES rule. In addition, one doesn't have to be an expert in the subject to carefully apply statistical methods to data in any field of interest.

Two things result from the use of statistics, knowledge of which data sets are related to one another (and how strongly) and which data sets are not related to one another. An example of the latter case is the disproval of some theory that one event leads to another. A lack of statistical correlation either disproves the theory or renders it
useless. The parameters that have the highest correlation may never be discovered or known but statistics are needed to verify candidates. Expertise and experience play a role in finding the candidate parameters that can be screened. HEE is a problem that has many different causes and a multitude of factors which influence it. Research over the years has pointed to a variety of mechanisms and to this date, predictability of an alloy response to hydrogen is still an elusive goal.

During the research for this project, alloy parameters that represent the alloy response to hydrogen and other alloy properties were collected. The most significant data is the alloy composition and some mechanical tests in hydrogen. *None of the alloys surveyed were created for hydrogen service.* Instead, alloys with desirable properties were screened with respect to their hydrogen response. The best alloys have been improved by modification of heat treatment and composition. Evolution doesn't greatly improve them over their base capabilities. The list of 'best' alloys has been known for decades yet, as a group, no major communality has been found among them. There is, however, a high communality among them due to the service requirements. For example, for the SSME, the alloys are mostly Ni-based and have a high strength from the high volume fraction of gamma-prime. Coincidentally, these alloys are very oxidation and creep resistant as well. Unfortunately, attempts to strengthen the alloys reduce their resistance to hydrogen.

Some parameters that relate to alloy properties were calculated from Super-PHACOMP. Complex, multi-element relationships are used to determine properties such as the amount of gamma-prime or the susceptibility to sigma phase formation. Such information was found to have a high statistical significance to the mechanical hydrogen test results.

A variety of statistical techniques were applied to the data collected in order to efficiently prepare it for the use in the ES. Simple techniques such as sorting the compositions by Ni, Cr or other elements were used as well as the more sophisticated
methods of covariance, partial correlations, canonical correlations, principle component analysis, discriminant analysis and casement plots. Some of these methods have similar effects on the data and reveal the same correlations.

At first, the compositions of the alloys alone were analyzed. Relationships among alloys selected for SSME use verses non-SSME candidates were sought. No such relationship was found. Only when the data of notched and smooth tensile ratios were included did meaningful results begin to develop. The best results came from covariance and partial correlation analysis of the large matrix of composition and mechanical test data. Statistical analysis is only valid for complete data sets. Different sets of data had to be studied since complete information on all alloys was never available.

Covariance and partial correlations are very powerful statistical methods. They are also relatively standard and popular as a result. The theory behind these methods will not be discussed here since any statistics textbook will cover the subjects. Principle components is a technique which uses normalized covariances among multiple parameters to determine which among them contribute the most to the overall pattern in the data. What that pattern actually is is not known a-priori. A dozen or more parameters can be assembled to test for principle components. As many components can contribute to the overall pattern. However, if the parameters are properly selected, only a few components may be the result, indicating that only a few parameters are actually needed to describe the pattern. Parameters which have a high negative correlation are just as useful as those that correlate positively. The significant feature of plotting the second against the first principle component is that the interrelationship between the selected parameters for the analysis is displayed. This way, one can find the parameters which are strongly correlated to the one of interest and determine the degree of correlation graphically. The alternative is to evaluate the large covariance (or partial correlations) matrix of numbers manually.
One major item of information that was found is that the patterns obtained always required at least 4 to 5 components to describe them. This is representative of the hydrogen embrittlement problem. This has also been the underlying difficulty in this work. The problem cannot be broken down into merely a few critical parameters. For example, the alloys that show the best notch ratios do not also have the best smooth tensile ratios. This also underlines another inherent problem in that there is no best alloy for comparison. The best alloy depends on the criteria. Collectively, all criteria do not identify any single alloy as the best but develops a conflict. Such a conflict cannot be resolved by any method, statistical or otherwise.

Once Super-PHACOMP was written, a new variety of parameters were added to the statistical analysis. The various PHACOMP numbers, each derived by a different algorithm, a calculated stacking fault energy and the weight percent gamma prime in the alloy were used in principle component analysis, discriminant analysis and of course, covariance analysis. Canonical relationships were also found but only substantiated the other analyses. A high correlation was found between the notched ratio data and the New-PHACOMP number while the smooth ratio data and other PHACOMP numbers were not highly correlated to the notch ratio. Instead, the PHACOMP-A number was highly positively correlated to the smooth ratio data and the New-PHACOMP with the weight-percent gamma prime were very negatively correlated. These observations have no reasonable cause for their outcome. However, the correlations are so high that one can use these parameters to predict the degree of embrittlement. This was done through the use of discriminant analysis. After the alloys with notch and smooth ratio data were classified into low, medium, and severe embrittlement categories, discriminant analysis was applied. The purpose was to use the parameters with a high correlation to the hydrogen embrittlement factor (in the form of the three classes for each type of test) to define a discriminating set of equations. Given the parameter values, it is then possible to determine which of the three classes the parameter set falls within.
After much trial and error, a suitable set of classes were established for the discriminant analysis. (This is not a task that can be automated.) Different classes were needed for each of the two tests, notch and smooth. These classes are listed below. There is no relationship between these classes and those established elsewhere. Four classes were selected originally but the results of the statistical analysis showed that one should not break down the data into this many classes. The discriminant analysis was repeatedly applied on different classification sets until the best quality of discrimination could be made. A set of equations were calculated and now can be used to classify an alloy based on PHACOMP numbers, stacking fault energy and wt. percent gamma prime. With the data available at that time, the quality of discrimination was about 70-80 percent. This refers to the confidence in predictability one would expect. It is certain that the equations and quality of predictability would change if more data were added. The equations as determined from data available up to June 1990 in this project are the ones used in the ES.

**Classes Used in Discriminant Analysis**

<table>
<thead>
<tr>
<th>Notched Tensile Ratio</th>
<th>Smooth Tensile Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe &lt;0.699</td>
<td>Severe &lt;0.899</td>
</tr>
<tr>
<td>Moderate 0.7-0.969</td>
<td>Moderate 0.9-0.989</td>
</tr>
<tr>
<td>Low(best) &gt;=0.97</td>
<td>Low(best) &gt;=0.99</td>
</tr>
</tbody>
</table>

These class values were also used for the IXL Knowledge Discovery process.

One could use only two classes such as embrittled and not embrittled, but the quality of the alloys permitted in the not embrittled category may not have suitable properties. In addition, since one goal of HEEES is to tailor an alloy, one needs more than yes/no answers. It is better to sense a direction so that one can follow the path to a
better alloy. (For example, going from severe to moderate after a change in the alloy points to an improvement. However, under the alternate class system, both alloys would still be embrittled.)

Using statistics properly resulted in the worthwhile results above. This requires time and knowledge and luck. Beneficial correlations can be found by using the Knowledge Discovery Program called IXL from Intelligenceware. While they also use statistics in their evaluations, they also incorporate several of their own techniques to find knowledge in data. Once the data is made available to the program and set up, the program systematically combs through the data looking for significant patterns. Since the process is time consuming, the program is arranged so that it can function unattended. It also is not biased in its search and so often finds patterns that were never thought could exist. (It also finds useless patterns like 'people under 12 years old earn less than $1000 a year' if a salary survey database was used.)

This program was obtained to address the question of finding a pattern or correlation in the HEEES database that may not have any theoretical basis and therefore could not be known even by experts in the field. The program stands alone and is not part of HEEES. However, the IXL program was run with the data in HEEES and a few useful rules or pieces of knowledge were added to HEEES. These mostly relate to patterns IXL found within the classes given above. The most useful ones found that were used related compositions. Only rules that had very high confidence factors and low margins of error were used. Since the rules are untested, less value should be placed on them. Also, it is likely that these rules may conflict with others in HEEES. However, if all the rules or even most of the rules in HEEES tell the same story, then a high degree of corroboration is established. (Luck comes in when the high corroboration is found for a low embrittlement condition.)

2.4. HEEES Procedures
The HEEES is accessed through the Nexpert Forms program. A HEEES.BAT file in the root directory of the 386 computer with all the required files in it will start the process. Forms can be started from the NXPPROT subdirectory as well. For automatic HEEES access through Forms, the HEEES.RTD (run time definition) file is needed with the Forms. To run HEEES from Forms manually, and not the easy way with the prepared BAT file, type NXPFORMS /fHEEES.RTD while in the NXPPROT subdirectory. The HEEES.BAT file will bring these programs together properly to begin a session.

Forms is a shell which allows the user to answer questions, make choices, and provides input to the HEEES. HEEES is a compiled knowledge base with the built-in capability to find information on Lotus-123 files that may be needed in a session. There are menu options provided by Nxpforms that are accessible to the user. A series of small 'forms' files are sequenced by the Forms shell to receive information in an organized fashion. After answering some questions, and waiting occasionally for responses within the system, a report is presented on the screen which is several pages (screens) long. This report, once seen on the screen, can be printed on the HP-Laser printer connected to the computer by pressing Control-P (the CTRL and p simultaneously). The Forms shell then formats the report and sends it in its entirety to the printer. While it is possible to have the report automatically print at the end of the session, this generally creates excess paper. However, since the report contains the critical input information used, it is wise to keep a record of important session results. By adding the command

#Print#

at the end of the ASCII file called HEEESRPT.TXT (in the C:\NXPPROT\NXPFORMS directory), automatic report printing will occur at the end of each session. A special key that was provided with the Runtime System Software is required to execute the Runtime Routines of which Forms is one. Without the key installed properly in the parallel printer port, HEEES cannot be run. In addition, the
computer must have 1 megabyte of extended memory minimum to run HEEES. The computer therefore must have at least 2 megabyte of memory total.

HEEES can be used to solve several different user problems. Since one main application is alloy development, and alloy development involves composition changes, the report should be printed so that alloys not in the database can be properly input. It is rare to find a person who is able to mentally perform a mass balance for fifteen elements to two decimal places. HEEES will try to provide answers even when the input composition is not realistic. The report informs the user if the composition selected is out of bounds.

Several times during a single session, the computer performs searches and does calculations. Often, several minutes of computer time are needed. One must wait patiently until the next new screen is presented. The time taken for these periods will not deviate much from run to run. HEEES contains a considerable amount of knowledge and calculation ability. These aspects of HEEES are fixed by assembling them in the NEXPERT-OBJECT Development System. The final version of all the HEEES functions are then compiled and the HEEES.CKB results. This file cannot be altered and is the most essential single file in HEEES. Forms has several features which allow the user access to some parts of HEEES. The Data and the Objects and the Properties are some of these. The user may restart a session without the help of HEEES.RTD (which can be recalled by a sequence.. ALT-S cursor down to load RTD hit return, find HEEES.RTD in the NXPPROT directory, cursor to it and finally hit Return). By restarting, the user can volunteer data and start the Knowledge Access (Knowcess) process by entering CTRL-K. The data menu is accessed by CTRL-D. Escape, ESC, will return you to the main menu. Alternate forms and reports can be prepared and used from the Forms main menu or appropriately installed in the RTD file which defines the flow of events in Forms. The user has no control over the flow of the process within the compiled knowledgebase. Any modifications to the forms, their flow, their appearance,
their logic etc. can be made but the user must read the NEXPERT FORMS manual that describes the specialized structures of the forms and the other support files used.

A 'Case Status' and 'Full Report' can be called up (ALT-R and select) after Knowcess to tell what the final values or case of the hypotheses are and what the conditions of the rules that were used in the run were set to. Once on the screen, these too can be printed by CTRL-P. This is the only way to find out how the rules in HEEES work collectively. Note that only those rules needed for the run are shown. Usually, this represents only a third or less of all the rules in the knowledgebase.

Forms related files are found in the NXPPROT directory and in the NXPFORMS directory. The compiled knowledgebase is in the NEXPERT directory along with other files created by the development system and not usable with Forms. The Lotus-123 databases are collected in the HEEES directory. The databases (spreadsheets) are organized individually to function with NEXPERT. The overall structure of the databases cannot be altered without completely crashing HEEES. Data can be added to the end of the appropriate columns in each sheet but HEEES will not see these additions until the worksheet ranges are properly redefined to include these new rows of data. PHACOMPH.EXE and KILL.EXE are files found in HEEES directory and are critical to HEEES. These are compiled programs and cannot be altered. Object and source code for these are in the same directory. NEXPDATA.PRN and PHACNEXP.OUT are special files which are used for data transfer between HEEES and PHACOMPH. These files change on every run. BASRUN20.EXE and BRUN20.EXE are needed to support the PHACOMPH and KILL programs and therefore must not be removed or altered.

Some work can be performed through Windows386 or 3.0. This shell (and a mouse) is needed to run the NEXPERT OBJECT Development System which cannot function on the computer without the proper key. The newer development system will run under Windows 3.0. WIN386 is not available on the computer but PIF files that
support NEXPERT will be left for future applications when a development system is reinstalled with Windows 3.0. or 386.

2.5 Using IXL Software

The IXL program provided is a supplementary package to HEEES. This set of programs is designed to take a very large database of information and find patterns in the data that can be printed in an expert system rule format. This has been called knowledge discovery. Some HEEES database information has been converted to dBase format and run through IXL. The rules IXL created were examined and worthwhile candidates installed in HEEES. If these rules 'fire', the report will signify the rule was originally from IXL.

The instructions for IXL give the details of how to set it up. IXL resides in its own directory and must be path'd. A database (I found dBase files to be the most foolproof) or set of databases that will be used for a search are placed in a separate directory. It is within this directory that all the subsequent analyses and intermediate data sets will be stored. When within the directory of data, IXL is invoked and the preparation for knowledge discovery begins.

IXL has other features that make it unique in the Artificial Intelligence field. One such feature is the Data Dictionary. With this, a range of numeric data can be described by a more useful English term (an object). For example, given a database with people's ages, rather than have cumbersome numerical ranges, the Data Dictionary allows the user to define 'Child' as 1-12 years, 'Adolescent' as 13-19, 'Adult' as 20-50 and 'Elderly' as 66-110. An IXL pattern that may be discovered could then be:

If 'Child'= TRUE AND 'Elderly'= TRUE
THEN Income < $10,000
In a similar vein, patterns can be established from data that is not-numeric, such as countries, cities, names, model numbers on parts etc. An IXL pattern that could be discovered may be:

If Country = 'Mexico' AND 'Elderly' = TRUE

THEN $1 <= Pension <= $20000.

Using the Data Dictionary, embrittlement categories were established for pattern searching. Using both Notched and Smooth H$_2$/He ratios, three categories, Low, Moderate and Severe (relating to HEE) were created. Patterns or correlations were desired for each choice so that the conditions for each could be detected. While IXL can be allowed to find any patterns it finds by itself, the user can set desired goals to find patterns for first. For example, it was most desirable to find the correlations that led to a Low condition of HEE. Only a couple of correlations were found that had significant confidence that could be used. For each 'rule' found, IXL gives the confidence factor and estimate of error that the rule is true when applied to the database used to generate the rule. This is the catch, your rules are only as good as your data. In addition, IXL is intended for use with Extremely Large Databases. This was not the case for HEEES. A database of 1000 rows is considered small for IXL. IXL rejects rows that have blanks anywhere in them, as a result, editing the data before processing is needed. Such preparations prevent full automation of the process. Once the user becomes familiar with IXL and how it works, the power and the limitations will become apparent.

The most difficult part of IXL Knowledge Discovery is setting up the search conditions for effective use of the computer and the user's time. Meaningless rules by the hundreds could be the outcome if goals and confidence limits are not well defined. One function of IXL is that it partitions numeric data sets into pieces which then become individual 'objects' for pattern searching. The user can specify how many pieces all the
data sets can be cut up into. Too many pieces and IXL will run for hours; too high a confidence level, and perhaps no rules could be found. An example of a meaningless (to us) rule from HEEES would be:

If Alloy= 'Waspalloy' OR 'Astroloy'
Then 10 < Nickel Content < 80 AND 0 < Chrome Content < 50.

3.0 Summary of HEEES Project

The goal of this project is to design a knowledge base to assist in developing improved SSME alloys for use in turbine blades. Improved, in this case, refers to reduced hydrogen embrittlement with minimal reduction in strength. The steps which were taken to achieve the goals were as follows.

1. Collect data on SSME and non-SSME alloys.
2. Generate rules, based on the data, to guide the analysis of the alloy.
3. Create a knowledgebase with these rules to perform the alloy analysis.

The completed project so far includes the following achievements:

1. Collected data and researched causes of hydrogen environment embrittlement.
2. Created a database to accommodate important information pertaining to the alloys within the database.
3. Analyzed the collected data statistically and created a version of PHACOMP which integrates several methods of PHACOMP evaluation.
4. Developed an expert system with rules to determine the level of embrittlement, the tensile strength, and other miscellaneous information about an alloy in question. This alloy can be selected from the database or provided by the user.
5. Created an interactive system to allow user input by forms and output by a report which contains a summary of the available information on the alloy and possible ways to improve the alloy.

3.1 Current Status of the Project

The first stage of development for this expert system has been completed, leaving us with a functional shell. To improve this shell and create a truly useful expert system, several areas can be improved upon.

1. The database has several voids in its composition, tensile and microstructural data. Even the list of alloys is limited. These voids need to be filled.
2. The spreadsheets can be better organized so the data entries within each spreadsheet parallel the other spreadsheets (column titles, etc.). But these changes cannot be made unless the knowledgebase is adjusted to accommodate these changes.
3. Within the knowledgebase, object naming needs to be unified and more structured. More comments would be useful for program modification.
4. Several items of data still need an order of sources to direct the system in discovering a value. In some instances the user is the second source if a value is unknown in the database when perhaps the user should not be approached at all.
5. The program, Nxpforms and the implemented expert system could be more user friendly. An alloy's composition cannot be directly modified if the composition is selected from the database. Also, if the user is asked to input a numeric value and the value is not known, the system provides no way to input 'notknown' as an answer.
6. When selectively retrieving data, the system is slow. The expert system only retrieves by query or sequentially when working with a relational database.
Implementing another database such as Oracle, and adjusting the knowledgebase appropriately would reduce the operating time considerably.

7. The system can be improved by increasing the versatility in the processing sequence. For example, the user should be able to provide select information and these results should override any calculated results or retrieved data generated by the system.

8. Augment the analysis with effects of various heat treatments and their effects on HEE and strength.

3.2 HEEES EXPERT SYSTEM MODULES

User Input: The user will provide the named and/or composition of the alloy in question. The user will also be able to provide any other information he may wish to specify.

Data Collection: The Super-PHACOMP program will be executed and results will be collected from it. Also, the available microstructural and mechanical data for the alloy will be retrieved.

Similar Alloy: The expert system will compare the provided composition with compositions in the database to select a similar alloy. This selection is based only on composition.

HEE Evaluation: The degree of HEE will be established based on the available information.

Results of HEE Evaluation: The degree of HEE suspected will be stated and explained. Improvements to be made on the alloy are suggested for other sessions with the expert system. These improvements will be specified by the user at the next session. The mechanical and microstructural data which was retrieved will be available for review by the user.

3.3 List of Files Pertinent To HEEES

Knowledge bases: HEEES.CKB HEEES.KB
4.0 ACKNOWLEDGEMENTS

We wish to thank Ms. Dianne Schmidt and Bryan MacPherson of the M&P Laboratory at Marshall Space Flight Center for their assistance in developing the expert knowledge required for this project. Also Ms. Beth Adams is to be commended for doing an outstanding job in setting up the databases and running NEXPERT OBJECT.

5.0 REFERENCES


APPENDIX A. SAMPLE HEEES REPORT

Aug 24 1990

HEEES - Hydrogen Environment Embrittlement Expert System

1. ALLOY COMPOSITION RELATED PROPERTIES

The alloy name given is: AF_56

Base type for alloy was determined to be Nickel based.

For this evaluation, the alloy process condition was given as CAST.

Composition of AF_56 with any modifications by the user:

Major elements of AF_56

<table>
<thead>
<tr>
<th>Element</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ni</td>
<td>61.40</td>
</tr>
<tr>
<td>Co</td>
<td>8.50</td>
</tr>
<tr>
<td>Ti</td>
<td>4.24</td>
</tr>
<tr>
<td>W</td>
<td>4.39</td>
</tr>
<tr>
<td>Nb</td>
<td>0.00</td>
</tr>
<tr>
<td>B</td>
<td>0.01</td>
</tr>
<tr>
<td>Zr</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Trace elements that are in this alloy

<table>
<thead>
<tr>
<th>Element</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hf</td>
<td>0.01</td>
</tr>
<tr>
<td>Mn</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Undesirable elements which were found in this alloy

<table>
<thead>
<tr>
<th>Element</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>0.00</td>
</tr>
<tr>
<td>Pb</td>
<td>0.00</td>
</tr>
<tr>
<td>O</td>
<td>0.00</td>
</tr>
<tr>
<td>Si</td>
<td>0.00</td>
</tr>
<tr>
<td>Bi</td>
<td>0.00</td>
</tr>
<tr>
<td>Cu</td>
<td>0.00</td>
</tr>
<tr>
<td>P</td>
<td>0.00</td>
</tr>
</tbody>
</table>

2. ALLOY SUSCEPTIBILITY TO HYDROGEN

Based upon the composition and other data given, the following was determined.

This alloy should be MODERATELY embrittled in hydrogen when performing a notched tensile test at room temperature. This conclusion is based on statistical analysis of the HEEES data. This alloy should be SEVERELY embrittled in hydrogen when
Aug 24 1990

HEEES - Hydrogen Environment Embrittlement Expert System

performing a smooth tensile test at room temperature. This conclusion is based on statistical analysis of the HEEES data.

3. SUGGESTED MODIFICATIONS TO THE ALLOY TO INCREASE HYDROGEN RESISTANCE

It was determined that the carbon content of this alloy should be reduced.
It was determined that the zirconium content of this alloy should be reduced.

It was determined that the stacking fault energy of the matrix in this alloy should be reduced.
It was determined that you should reduce the grain size of this alloy to improve HEE resistance.

4. NON-HYDROGEN SPECIFIC ALLOY PROPERTIES

SUPER-PHACOMP results:

New Phacomp, Japanese source: 0.78
Phacomp A, from Woodyatt/Sims/Beattie: 1.04
Phacomp B, from Rideout & Beck: 0.87

Sigma Phase will not form in this alloy. It is SIGMA-SAFE.
Sigma phase should not form in this alloy. The alloy is Sigma-SAFE. Sigma sum=0.

Calculated rupture time at 1093 deg C: 0.35 hours.
Calculated Tensile Strength at 760 deg C: 1249.51 psi.

The calculated stacking fault energy, SFE, for the matrix was found to be 374.79 ergs per sq. cm.
Weight percent Gamma-prime was calculated as: 41.29 %.

The grain size was: 23.00 micrometers.

END OF REPORT