Armchair Missions to Mars: Using Case-Based Reasoning and Fuzzy Logic to Simulate a Time Series Model of Astronaut Crews

GERRY STAHL

Center for LifeLong Learning & Design

Department of Computer Science and Institute of Cognitive Science

( )) University of Colorado, Boulder, Colorado, USA 80309-0430

Gerry.Stahl@Colorado.edu
(303) 492-3912
(303) 492-2844

Abstract

Computer simulation of long missions in space can provide experience and predictions without the expense and risk of actual flights. Simulations are most helpful if they can model the behavior of key psychological factors of the crew over time, rather than simply predicting overall mission success. Because of the lack of experience with interplanetary trips and the problems of generalizing and adapting data on analog missions, it is not possible to formulate a set of formal rules adequate for building an expert system. Rather, a case-based reasoning approach to constructing a time series model is pursued. Even for this approach, however, the case base must be supplemented by adaptation rules. These rules of thumb are gleaned from the social science literature on small group interactions under extreme conditions of isolation and confinement. The non-quantitative nature of these rules lends itself to formulation and computation using fuzzy logic. The application domain presents several technical issues for traditional case-based reasoning: there is no natural hierarchy of parameters to use in optimizing installation and retrieval of cases, and there are large variations in behavior among similar missions. These problems are addressed by custom algorithms to keep the computations tractable and plausible. Thus, the harnessing of case-based reasoning for this practical application requires the crafting of a custom, hybrid system.

Keywords: *simulation*   *case-based reasoning*   *statistical modeling*

Preface

Background of the Research

During the period of space exploration around 1993, planners at NASA (the US space agency) were concerned about interpersonal issues in astronaut crew composition. The nature of astronaut crews was undergoing significant change. In the past, astronauts were primarily young American males with rigorous military training; missions were short; crews were small. Prior to a mission, a crew trained together for about a year, so that any interpersonal conflicts could be worked out in advance. The future, however, promised crews that would be far less homogeneous and regimented: international crews speaking different languages, mixed gender, inter-generational, larger crews, longer missions. This was the start of Soviet-American cooperation and planning for Space Station. While there was talk of a manned expedition to Mars, the more likely scenario of an international Space Station with six-month crew rotations was a realistic concern.

There was not much experience with the psychological effects on crews confined in isolated and extreme conditions for months at a time. The data from submarines and Antarctic winter-overs provided some indications, but it was limited, inappropriately documented and inconsistent. NASA was beginning to conduct some experiments where they could collect the kinds of data they needed. But they required a way of analyzing such data, generalizing it and applying it to projected scenarios.

The Soft Computing Algorithms

NASA wanted a way of predicting how a given crew – with a certain mix of astronauts – might respond to mission stress under different scenarios. This would require a complex model with many parameters. There would never be enough relevant data to derive the parameter values statistically. Given a modest set of past cases, the method of case-based reasoning suggested itself. A case-based system requires (1) a mechanism for retrieving past cases similar to a proposed new case and (2) a mechanism for adapting the data of a retrieved case to the new case based on the differences between the two.

For the retrieval mechanism, we defined a number of characteristics of astronauts and missions. The nature of our data and these characteristics raised a number of issues for retrieval and we had to develop innovative modifications of the standard case-based reasoning algorithms, as described in detail below.

For the adaptation mechanism, we developed a model of the mission based on a statistical approach known as interrupted time series analysis. In addition, we derived a set of adaptation rules based on the social science literature about confined and isolated crews. We formulated these rules in English and represented them in the software using fuzzy logic.
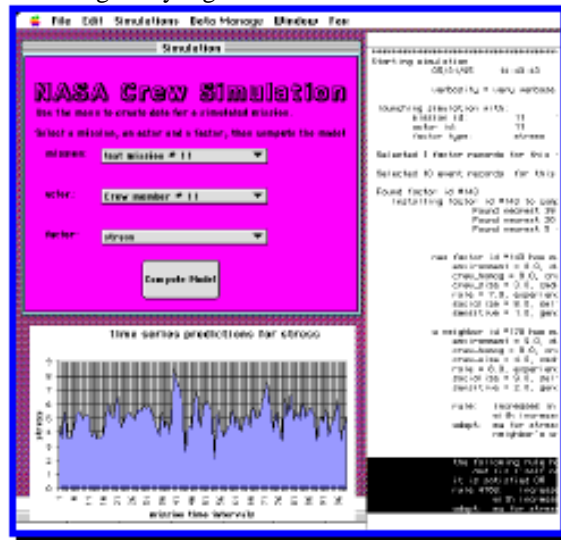

Figure 1. A view of the CREW interface. Upper left allows selection of mission characteristics. Menu allows input of data. Lower left shows magnitude of a psychological factor during 100 points in the simulated mission. To the right is a listing of some of the rules taken into account.

The Software System

We developed a case-based reasoning software system named CREW. To make the retrieval of cases tractable, scalable and efficient, we developed the system in a database programming environment. We selected FoxPro because it was highly

optimized, included a general purpose programming language and was compatible with both Windows and the Macintosh.

Most of the software code consisted of the algorithms described in this chapter. Because CREW was intended to be a proof-of-concept system, its data entry routines and user interface were minimal. The user interface consisted of a set of pull-down menus for selecting a variety of testing options and a display of the results in a graph format (see Figure 1). Some of the steps in the reasoning were printed out so that one could study the reasoning process.

By the end of the project, we successfully demonstrated that the time series model, the case-based reasoning and the fuzzy logic could all work together to perform as designed. The system could be set up for specific crews and projected missions and it would produce sensible predictions quickly. The next step was to enter real data that NASA was just beginning to collect. Because of confidentiality concerns, this had to be done within NASA, so we turned over the software to them for further use and development.

The People Involved

The research was sponsored by the Behavioral and Performance Laboratory at Johnson Space Center in Houston, Texas, part of NASA's Astronaut Support Division. We worked closely with NASA researchers Dr. Joanna Wood and Dr. Albert Holland on the design of the software and the data. At the end of the project, we delivered the software to them to continue the work.

The research was conducted at Owen Research, Inc. (ORI) in Boulder, Colorado. ORI is a small research lab founded and run by Dr. Robert Owen. Owen is a physicist specializing in laser optics. He also has a Ph.D. in anthropology, and his dissertation in that field led to this research in modeling small group behavior using AI (artificial intelligence) techniques. I developed the technical approach and programmed the system. Dr. Brent Reeves assisted with the fuzzy logic algorithms. To help collect and analyze social science literature related to small groups in isolated conditions, we worked with Professor Russell McGoodwin of the Anthropology Department at the University of Colorado (CU) and his student, Nick Colmenares. In addition, I conducted several interviews of an experienced astronaut, Mike Lounge, and discussed our project with him.

I began this project immediately after completing my Ph.D. dissertation in computer science at CU, where I had specialized in AI. Since my undergraduate years at MIT in the mid-sixties and a Ph.D. in philosophy in the mid-70's, I have worked as a systems programmer, software developer and computer consultant. Following this project, I continued to work with ORI on software for an optical bench to stabilize interferometry equipment during space flight and on an Internet-based system for teachers to share curriculum ideas. After working at ORI, I returned to CU, where I am now a Research Professor in cognitive science and computer science. My current research involves Web-based environments for collaborative learning and knowledge-building – for further information or to contact me, see http://www.cs.colorado.edu/~gerry.

"Armchair missions to Mars" describes research conducted during a two year SBIR (Small Business Innovative Research) grant (Project 91-1-II-1201-9027) from NASA in 1993-1995. It was originally published (except for this Preface) in the

journal *Knowledge-Based Systems* in 1996, volume 9, pages 409-415, and is reprinted below with the permission of the publisher, Elsevier.

Introduction

The prospect of a manned mission to Mars has been debated for 25 years since the first manned landing on the moon. (Amer. Astro. Soc., 1966) It is routinely argued that this obvious next step in human exploration is too costly and risky to undertake, particularly given our lack of experience with lengthy missions in space. (McKay, 1985)

Social science research to explore issues of the effects of such a mission on crew members has focused on experience in *analog* missions under extreme conditions of isolation and confinement, such as Antarctic winter-overs, submarine missions, orbital space missions and deep sea experiments. (Harrison, et al., 1991) This research has produced few generalizable guidelines for planning a mission to Mars. (Collins, 1985)

We have undertaken to simulate the effects of interplanetary missions in a computer program named CREW. This program is for use by NASA to assist in astronaut crew selection and mission planning. (Owen, et al., 1993) Given descriptions of tentatively selected crew members and of scheduled activities, CREW simulates the mission and reports on the probable course of particular factors during the duration of the mission.

We are working with staff at the psychology labs of NASA's astronaut support division, so we have focused on psychological factors of the crew members, such as stress, morale and teamwork. NASA has begun to collect time series psychological data on these factors by having crew members in space and analog missions fill out a survey on an almost daily basis. As of the conclusion of our project (June 1995), NASA had analyzed data from an underwater mission designed to test their data collection instrument, the IFRS (Individualized Field Recording System) survey, and was collecting data from several Antarctic traverses. The IFRS survey was scheduled to be employed on a joint Soviet-American shuttle mission. Its most likely initial use would be as a tool for helping to select crews for the international Space Station.

Our task was to design a system for incorporating eventual IFRS survey results in a model of participant behavior on long-term missions. Our goal was to implement a proof-of-concept software system to demonstrate algorithms for combining AI techniques like case-based reasoning and fuzzy logic with a statistical model of IFRS survey results and a rule-base derived from the existing literature on extreme missions.

This paper reports on our system design and its rationale. The CREW system predicts how crew members in a simulated mission would fill out their IFRS survey forms on each day of the mission, that is, how they would self-report indicators of stress, motivation, etc. As NASA collects and analyzes survey data, the CREW program can serve as a vehicle for assembling and building upon the data—entering empirical cases and tuning the rule-base. Clearly, the predictive power of CREW will depend upon the eventual quantity and quality of the survey data.

Modeling the Mission Process

NASA is interested in how psychological factors such as those tracked in the IFRS surveys evolve over time during a projected mission's duration. For instance, it is not enough to know what the average stress level will be of crew members at the end of a nine-month mission; we need to know if any crew member will be likely to be particularly stressed at a critical point in the middle of the mission when certain actions must be taken. To obtain this level of detail of prediction, we created a *time series* model of the mission.

The model is based on standard statistical time series analysis. McDowall, et al. (1970) argue for a stochastic ARIMA (Auto Regressive Integrated Moving Average) model of interrupted time series for a broad range of phenomena in the social sciences. The most general model takes into account three types of considerations: (1) trends, (2) seasonality effects and (3) interventions. An observed time series is treated as a realization of a stochastic process; the ideal model of such a process is statistically adequate (its residuals are white noise) and parsimonious (it has the fewest parameters and the greatest number of degrees of freedom among all statistically equivalent models).

(1) *Trends.* The basic model takes into account a stochastic component and three structural components. The stochastic component conveniently summarizes the multitude of factors producing the variation observed in a series which cannot be accounted for by the model. At each time t there is a stochastic component a_t which cannot be accounted for any more specifically. McDowall, et al. claim that most social science phenomena are properly modeled by first-order ARIMA models. That is, the value, Y_t of the time series at time t may be dependent on the value of the time series or of its stochastic component at time $t-1$, but not (directly) on the values at any earlier times. The first-order expressions for the three structural components are:

autoregressive: $Y_t = \alpha_t + \phi\, Y_{t-1}$

differenced : $Y_t = \alpha_t + Y_{t-1}$

moving average : $Y_t = \alpha_t + \theta\, a_{t-1}$

We have combined these formulae to produce a general expression for all first-order ARIMA models: $Y_t = a_t + f\, Y_{t-1} + q\, a_{t-1}$

This general expression makes clear that the model can take into account trends and random walks caused by the inertia (or momentum) of the previous moment's stochastic component or by the inertia of the previous moment's actual value.

(2) *Seasonality.* Many phenomena (e.g., in economics or nature) have a cyclical character, often based on the 12 month year. It seems unlikely that such seasonality effects would be significant for NASA missions; the relevant cycles (daily and annual) would be too small or too large to be measured by IFRS time series data.

(3) *Interventions.* External events are likely to impact upon modeled time series. Their duration can be modeled as exponential decay, where the nth time period after an event at time e will have a continuing impact of $Y_{e+n} = d^n\, w$ where $0 <= d <= 1$. Note that if $d = 0$ then there is no impact and if $d = 1$ then there is a

permanent impact. Thus, d is a measure of the rate of decay and w is a measure of the intensity of the impact.

We have made some refinements to the standard time series equations, to tune them to our domain and to make them more general. First, the stochastic component, $a_i(t)$, consists of a mean value, $m_i(t)$, and a normal distribution component governed by a standard deviation, $s_i(t)$. Second, mission events often have significant effects of anticipation. In general, an event j of intensity w_{ij} at time t_j will have a gradual onset at a rate e_{ij} during times $t < t_j$ as well as a gradual decay at a rate d_{ij} during times $t > t_j$. The following equation incorporates these considerations:

$$Y_{i(t)} = a_{i(t)} + f_i Y_{i(t-1)} + q_i a_{i(t-1)}$$

$$+ \sum_{j=1}^{n} \left[e_{ij}^{(t_j - t)} w_{ij} \right] + \sum_{j=1}^{n} \left[d_{ij}^{(t - t_j)} w_{ij} \right]$$
$$\left(for\ t < t_j \right) \qquad\qquad \left(for\ t \geq t_j \right)$$

where:

$Y_i(t)$ = value of factor i for a given actor in a given mission at mission time t

t_j = time of occurrence of the jth of n intervening events in the mission

a = noise: a value is generated randomly with mean m and standard deviation s

m = mean of noise value	$0 <= m <= 10$
s = standard deviation of noise	$0 <= s <= 10$
f = momentum of value	$-1 <= f <= 1$
q = momentum of noise	$-1 <= q <= 1$
e = rise rate of interruption	$0 <= e <= 1$
d = decay rate of interruption	$0 <= d <= 1$
w = intensity of interruption	$-10 <= w <= 10$

The model works as follows. Using IFRS survey data for a given question answered by a given crew member throughout a given mission and knowing when significant events occurred, one can use standard statistical procedures to derive the parameters of the preceding equation: m, s, f and q as well as e, d and w for each event in the mission. Then, conversely, one can use these parameters to predict the results of a new proposed mission. Once one has obtained the parameters for a particular psychological factor, a crew member and each event, one can predict the values that crew member would enter for that survey question i at each time period t of the mission by calculating the equation with those parameter values.

This model allows us to enter empirical cases into a case-base by storing the parameters for each *factor* (i.e., a psychological factor for a given crew member during a given mission) or *event* (i.e., an intervention event in the given factor time series) with a description of that factor or event. To make a time series prediction of a proposed factor with its events, we retrieve a similar case, adapt it for differences from the proposed case and compute its time series values from the model equation.

Using Case-Based Reasoning

The time series model is quite complex in terms of the number of variables and factors. It must produce different results for each time period, each kind of mission, each crew member personality, each question on the IFRS survey and each type of intervention event. To build a rule-based expert system, we would need to acquire thousands of formal rules capable of computing predictive results for all these combinations. But there are no experts on interplanetary missions who could provide such a set of rules. Nor is there data that could be analyzed to produce these rules. So we took a case-based reasoning approach. We take actual missions—including analog missions—and compute the parameters for their time series.

Each survey variable requires its own model (values for parameters m, s, f and q), as does each kind of event (values for parameters e, d and w). Presumably, the 107 IFRS survey questions can be grouped into several factors—although this is itself an empirical question. We chose six psychological factors that we thought underlay the IFRS questionnaire: crew teamwork, physical health, mental alertness, psychological stress, psychological morale and mission effectiveness. In addition, we selected a particular question from the survey that represented each of these factors. The CREW system currently models these twelve factors.

There is no natural taxonomy of events. Our approach assumes that there are categories of events that can be modeled consistently as interventions with exponential onsets and decays at certain impact levels and decay rates. Based on the available data, we decided to model eight event types: start of mission, end of mission, emergency, conflict, contact, illness, discovery, failure.

The case-base consists of instances of the 12 factors and the 8 event types. Each instance is characterized by its associated mission and crew member, and is annotated with its parameter values. Missions are described by 10 characteristics (variables), each rated from 0 to 10. The mission characteristics are: harshness of environment, duration of mission, risk level, complexity of activities, homogeneity of crew, time of crew together, volume of habitat, crew size, commander leadership and commander competence. Crew member characteristics are: role in crew, experience, professional status, commitment, social skills, self reliance, intensity, organization, sensitivity, gender, culture and voluntary status. In addition, events have characteristics: event type, intensity and point in mission.

Because there is only a small handful of cases of actual IFRS data available at present, additional cases are needed to test and to demonstrate the system. Approximate models of time series and interventions can be estimated based on space and analog missions reported in the literature, even if raw time series data is not available to derive the model statistically. Using these, we generate and install supplemental demo cases by perturbating the variables in these cases and adjusting the model parameters in accordance with rules of thumb gleaned from the literature on analog missions. This data base is not rigorously empirical, but it should produce plausible results during testing and demos. Of course, the database can be recreated at a later time when sufficient real data is available. At that point, NASA might change the list of factor and event types to track in the database or the set of variables to describe them. Then the actual case data would be analyzed using

interrupted time series analysis to derive empirical values for m, s, f and q for the factors.

Users of CREW enter a scenario of a proposed mission, including crew composition and mission characteristics. They also enter a series of n anticipated events at specific points in the mission period. From the scenario, the system computes values for m, s, f and q for each behavioral factor. For events $j = 1$ through n, it computes values for d_j, e_j and w_j. The computation of parameters is accomplished with case-based reasoning, rather than statistically. The missions or events in the case-base that most closely match the hypothesized scenario are retrieved. The parameters associated with the retrieved cases are then adjusted for differences between the proposed and retrieved cases, using rules of thumb formulated in a rule-base for this purpose. Then, using the model equation, CREW computes values of Y_t for each behavioral factor at each time slice t in the mission. These values can be graphed to present a visual image of the model's expectations for the proposed mission. Users can then modify their descriptions of the mission scenario and/or the sequence of events and re-run the analysis to test alternative mission scenarios.

CREW is basically a database system, with a system of relational files storing variable values and parameter values for historical cases and rules for case adaptation. For this reason it was developed in the FoxPro database management system, rather than in Lisp, as originally planned. FoxPro is extremely efficient at retrieving items from indexed database files, so that CREW can be scaled up to arbitrarily large case-bases with virtually no degradation in processing speed. CREW runs on Macintosh and Windows computers.

The Case Retrieval Mechanism

A key aspect of case-based reasoning (CBR) is its case retrieval mechanism. The first step in computing predictions for a proposed new case is to retrieve one or more similar cases from the case base. According to Schank (1982), CBR adopts the dynamic memory approach of human recall.

As demonstrated in exemplary CBR systems (Riesbeck & Schank, 1989), this involves a hierarchical storage and retrieval arrangement. Thus, to retrieve the case most similar to a new case, one might, for instance, follow a tree of links that begins with the mission characteristic, harshness of environment. Once one followed the link corresponding to the new case's environment, one would select the link for the next characteristic and so on until one arrived at a leaf of the tree with a particular case. The problem with this method is that not all domains can be organized in such a hierarchy meaningfully. Kolodner (1993) notes that some CBR systems need to define non-hierarchical retrieval systems. In the domain of space missions, there is no clear priority of characteristics for establishing similarity of cases.

A standard non-hierarchical measure of similarity is the n-dimensional Euclidean distance, which compares two cases by adding the squares of the differences between each of the n corresponding variable values. The problem with

this method is that it is intractable for large case-bases because you must compare a new case with every case in the database.

CREW adopts an approach that avoids the need to define a strict hierarchy of variables as well as the ultimately intractable inefficiency of comparing a new case to each historic case. It prioritizes which variables to compare initially in order to narrow down to the most likely neighbors using highly efficient indices on the database files. But it avoids strict requirements even at this stage.

The retrieval algorithm also responds to another problem of the space mission domain that is discussed in the section on adaptation below, the fact that there are large random variations among similar cases. This problem suggests finding several similar cases instead of just one to adapt to a new case. The case retrieval algorithm in CREW returns **n** nearest neighbors, where **n** is a small number specified by the user. Thus, parameters for new cases can be computed using adjusted values from several near neighbors, rather than just from the one nearest neighbor as is traditional in CBR. This introduces a statistical flavor to the computation in order to soften the variability likely to be present in the empirical case data.

The case retrieval mechanism consists of a procedure for finding the **n** most similar factors and a procedure for finding the **n** most similar events, given a proposed factor or event, a number **n** and the case-base file. These procedures in turn call various subprocedures. Each of the procedures is of computational order **n**, where **n** is the number of neighbors sought, so it will scale up with no problem for case bases of arbitrary size. Here are outlines of typical procedures:

nearest_factor(new_factor, n, file)
1. find all factor records with the same factor type, using a database index
2. of these, find the 4n with the **nearest_mission**
3. of these, find the n with the **nearest_actor**

nearest_mission(new_mission, n, file)
1. find all mission records with environment = new mission's environment $\pm$ 1 using an index
2. if less than 20n results, then find all mission records with environment = new mission's environment $\pm$ 2 using an index
3. if less than 20n results, then find all mission records with environment = new mission's environment $\pm$ 3 using an index
4. of these, find the 3n records with minimal |mission's duration - new mission's duration| using an index
5. of these, find the n records with minimal $S\ dif_i{}^2$

nearest_actor(new_actor, n, file)
1. find up to n actor records with minimal $S\ dif_i{}^2$

Note that in these procedures there is a weak sense of hierarchical ordering. It is weak in that it includes only a couple of levels and usually allows values that are not exactly identical, depending on how many cases exist with identical matches.

Note, too, that the n-dimensional distance approach is used (indicated by "minimal ? $\text{dif}_i{}^2$"), but only with 3*$\mathbf{n}$ cases, where $\mathbf{n}$ is the number of similar cases sought. The only operations that perform searches on significant portions of the database are those that can be accomplished using file indexes. These operations are followed by procedures that progressively narrow down the number of cases. Thereby, a balance is maintained that avoids both rigid prioritizing and intractable computations.

Case based reasoning often imposes a hierarchical priority to processing that is hidden behind the scenes. It makes case retrieval efficient without exposing the priorities to scrutiny. The preceding algorithms employ a minimum of prioritizing. In each instance, priorities are selected that make sense in the domain of extreme missions based on our understanding of the relevant literature and discussions with domain experts at NASA. Of course, as understanding of the domain evolves with increased data and experience, these priorities will have to be reviewed and adjusted.

Rules and Fuzzy Logic

Once $\mathbf{n}$ similar cases have been found, they must be adapted to the new case. That is, we know the time series parameters for the similar old cases and we now need to adjust them to define parameters for the new case, taking into account the differences between the old and the new cases. Because the database is relatively sparse, it is unlikely that we will retrieve cases that closely match a proposed new case. Adaptation rules play a critical role in spanning the gap between the new and the retrieved cases.

The rules have been generated by our social science team, which has reviewed much of the literature on analog missions and small group interactions under extreme conditions of isolation and confinement, e.g., (Radloff, 1968). They have determined what variables have positive, negligible or negative correlations with which factors. They have rated these correlations as either *strong* or *weak*. The CREW system translates the ratings into percentage correlation values. For instance, the rule, "teamwork is strongly negatively correlated with commander competence" would be encoded as a -80% correlation between the variable *commander competence* and the factor *teamwork*.

The rules function roughly as follows in CREW: one rule, for instance, is used to adjust predicted *stress* for a hypothetical mission of length *new-duration* from the stress measured in a similar mission of length *old-duration*. Suppose that the rule states that the correlation of psychological *stress* to mission *duration* is +55%. All mission factors, such as stress, are coded on a scale of 0 to 10. Suppose that the historic mission had its duration variable coded as 5 and a stress factor rating of 6, and that the hypothetical mission has a duration of 8. We use the rule to adapt the historic mission's stress rating to the hypothetical mission given the difference in mission durations (assuming all other mission characteristics to be identical). Now, the maximum that stress could be increased and still be on the scale is 4 (from 6 to 10); the *new-duration* is greater than the old by 60% (8 - 5 = 3 of a possible 10 - 5 = 5); and the rule states that the correlation is 55%. So the predicted stress for the

new case is greater than the stress for the old case by: 4 x 60% x 55% = 1.32—for a predicted stress of 6 + 1.32 = 7.32. Using this method of adapting outcome values, the values are proportional to the correlation value, to the difference between the new and old variable values and to the old outcome value, without ever exceeding the 0 to 10 range.

There are many rules needed for the system. Rules for adapting the four parameters (**m**, **s**, **f** and **q**) of the 12 factors are needed for each of the 22 variables of the mission and actor descriptions, requiring 1056 rules. Rules for adapting the three parameters (**e**, **d** and **w**) of the 8 event types for each of the 12 factors are needed for each of the 24 variables of the mission, actor and intervention descriptions, requiring 6912 rules. Many of these 7968 required rules have correlations of 0, indicating that a difference in the given variable has no effect on the particular parameter.

The rules gleaned from the literature are rough descriptions of relationships, rather than precise functions. Because so many rules are applied in a typical simulation, it was essential to streamline the computations. We therefore made the simplifying assumption that all correlations were linear from zero difference between the old and new variable values to a difference of the full 10 range, with only the strength of the correlation varying from rule to rule.

However, it is sometimes the case that such rules apply more or less depending on values of other variables. For instance, the rule "teamwork is strongly negatively correlated with commander competence" might be valid only "if commander leadership is very low and the crew member's self reliance is low". This might capture the circumstance where a commander is weak at leading others to work on something, while the crew is reliant on him and where the commander can do everything himself. It might generally be good for a commander to be competent, but problematic under the special condition that he is a poor leader and that the crew lacks self reliance.

Note that the original rule has to do with the difference of a given variable (*commander competence*) in the old and the new cases, while the condition on the rule has to do with the absolute value of variables (*commander leadership*, *crew member's self-reliance*) in the new case. CREW uses fuzzy logic (Cox, 1994) to encode the conditions. This allows the conditions to be stated in English language terms, using values like *low*, *medium*, or *high*, modifiers like *very* or *not*, and the connectives *and* or *or*. The values like *low* are defined by fuzzy set membership functions, so that if the variable is 0 it is considered completely *low*, but if it is 2 it is only partially *low*. Arbitrarily complex conditions can be defined. They compute to a numeric value between 0 and 1. This value of the condition is then multiplied by the value of the rule so that the rule is only applied to the extent that the condition exists.

The combination of many simple linear rules and occasional arbitrarily complex conditions on the rules provides a flexible yet computationally efficient system for implementing the rules found in the social science literature. The English language statements by the researchers are translated reasonably into numeric computations by streamlined versions of the fuzzy logic formalism, preserving sufficient precision considering the small effect that any given rule or condition has on the overall simulation.

The Adaptation Algorithm

Space and analog missions exhibit large variations in survey results due to the complexity and subjectivity of the crew members' perceptions as recorded in survey forms. Even among surveys by different crew members on relatively simple missions with highly homogeneous crews, the recorded survey ratings varied remarkably. To average out these effects, CREW retrieves **n** nearest neighbors for any new case, rather than the unique nearest one as is traditional in CBR. The value of **n** is set by the user.

The parameters that model the new case are computed by taking a weighted average of the parameters of the **n** retrieved neighbors. The weight used in this computation is based on a similarity distance of each neighbor from the new case. The similarity distance is the sum of the squares of the differences between the new and the old values of each variable. So, if the new case and a neighbor differed only in that the new case had a mission complexity rating of 3 while the retrieved neighbor had a mission complexity rating of 6, then the neighbor's distance would be $(6-3)^2 = 9$.

The weighting actually uses a term called *importance* that is defined as (sum - distance)/(sum * (n-1)), where *distance* is the distance of the current neighbor as just defined and *sum* is the sum of the distances of the n neighbors. This weighting gives a strong preference to neighbors that are very near to the new case, while allowing all **n** neighbors to contribute to the adaptation process.

Conclusions and Future Work

The domain of space missions poses a number of difficulties for the creation of an expert system:

- Too little is known to generalize formal rules for a rule-based system.
- A model of the temporal mission process is needed more than just a prediction of final outcomes.
- The descriptive variables cannot be put into a rigid hierarchy to facilitate case-based retrieval.
- The case-base is too sparse and too variable for reliable adaptation from one nearest neighbor case.
- The rules that can be gleaned from available data or relevant literature are imprecise.

Therefore, we have constructed a hybrid system that departs in several ways from traditional rule-based as well as classic case-based systems. CREW creates a time series model of a mission, retrieving and adapting the parameters of the model from a case base. The retrieval uses a multi-stage algorithm to maintain both flexibility and computational tractability. An extensive set of adaptation rules overcomes the sparseness of the case base, with the results of several nearest neighbors averaged together to avoid the unreliability of individual cases.

Our proof-of-concept system demonstrates the tractability of our approach. For testing purposes, CREW was loaded with descriptions of 50 hypothetical missions involving 62 actors. This involved 198 intervention parameters, 425 factor

parameters and 4,047 event parameters. Based on our reading of the relevant literature, 7,968 case adaptation rule correlation figures were entered. A number of fuzzy logic conditions were also included for the test cases. Given a description of a crew member and a mission, the CREW system predicts a series of one hundred values of a selected psychological factor in a minute or two on a standard Macintosh or Windows desktop computer.

Future work includes expanding the fuzzy logic language syntax to handle more subtle rules. Our impression from conflicting conclusions within the literature is that it is unlikely that many correlation rules hold uniformly across entire ranges of their factors.

We would also like to enhance the explanatory narrative provided by CREW in order to increase its value as a research assistant. We envision our system serving as a tool to help domain experts select astronaut crews, rather than as an automated decision maker. People will want to be able to see and evaluate the program's rationale for its predictions. This would minimally involve displaying the original sources of cases and rules used by the algorithms. The most important factors should be highlighted. In situations strongly influenced by case adaptation rules or fuzzy logic conditions derived from the literature, it would be helpful to display references to the sources of the rules if not the relevant excerpted text itself.

Currently, each crew member is modeled independently; it is undoubtedly important to take into account interactions among them as well. While crew interactions indirectly affect survey results of individual members (especially to questions like: How well do you think the crew is working together today?), additional data would be needed to model interactions directly. Two possible approaches suggest themselves: treating crew interaction as a special category of event or subjecting data from crew members on a mission together to statistical analyses to see how their moods, etc. affect one another. Taking interactions into account would significantly complicate the system and would require data that is not currently systematically collected.

Use of the system by NASA personnel will suggest changes in the variables tracked and their relative priority in the processing algorithms; this will make end-user modifiability facilities desirable. In order to quickly develop a proof-of-concept system, we hard-coded many of the algorithms described in this paper. However, some of these algorithms make assumptions about, for instance, what are the most important factors to sort on first. As the eventual system users gain deeper understanding of mission dynamics, they will want to be able to modify these algorithms. Future system development should make that process easier and less fragile.

Data about individual astronauts, about group interactions and about mission progress at a detailed level is not public information. For a number of personal and institutional reasons, such information is closely guarded. Combined with the fact that NASA was just starting to collect the kind of time series data that CREW is based on, that made it impossible for us to use empirical data in our case base. Instead, we incorporated the format of the IFRS surveys and generated plausible data based on the statistical results of completed IFRS surveys and the public literature on space and analog missions. When NASA has collected enough empirical cases to substitute for our test data, they will have to enter the new

parameters, review the rule base, and reconsider some of the priorities embedded in our algorithms based on their new understanding of mission dynamics. However, they should be able to do this within the computational framework we have developed, confident that such a system is feasible. As NASA collects more time series data, the CREW database will grow and become increasingly plausible as a predictive tool that can assist in the planning of expensive and risky interplanetary missions.

References

American Astronomical Society (1966) *Stepping Stones to Mars Meeting*. Amer. Inst. of Aero. & Astro. New York.

Collins, D (1985) *Psychological Issues Relevant to Astronaut Selection for Long Duration Spaceflight: A Review of the Literature.* USAF Systems Command. Brooks Air Force Base. Texas.

Cox, E (1994) *The Fuzzy Systems Handbook*. Academic Press. Boston.

Harrison, A, Clearwater, Y, McKay C (1991) *From Antarctica to Outer Space: Life in Isolation and Confinement.* Behavioral Science. Vol. 34, pp. 253-271.

Kolodner, J (1993) *Case-Based Reasoning*. Morgan Kaufmann. San Mateo, CA.

McDowall, D, McLeary, R, Meidinger, E, Hay, R Jr. (1980) *Interrupted Time Series Analysis*. Sage. Beverly Hills, CA.

McKay, C (ed.) (1985) *The Case for Mars II.* Am. Astro. Society. San Diego, CA.

Owen, R, Holland, A, Wood, J (1993) A Prototype Case-Based Reasoning Human Assistant for Space Crew Assessment and Mission Management. In Fayyad, U, Uthurusamy, F (eds.) (1993) *Applications of Artificial Intelligence 1993: Knowledge-Based Systems in Aerospace and Industry.* SPIE vol. 1963, pp. 262-273. Bellingham, WA.

Radloff, R, Helmreich, R (1968) *Groups Under Stress: Psychological Research in Sealab II.* Appleton Century Crofts. New York.

Riesbeck, C & Schank, R (1989) *Inside Case-Based Reasoning*. Lawrence Earlbaum. Hillsdale, NJ.

Schank, R (1982) *Dynamic Memory*. Cambridge Univ. Press. Cambridge.