From: Report to CEC, Subprogram FAST, 1982
MISMATCH BETWEEN MACHINE REPRESENTATION AND HUMAN CONCEPTS: dangers and remedies
By D. Kopec and D. Michie

CHAPTER 1

FAILURE AT THE HUMAN INTERFACE: FOUR CASE STUDIES

Gen 7

1.1 Three Mile Island

1.2 Air Traffic Control

1.3 NORAD Military Computer

1.4 Royal Dutch Steel

1.5 Viewpoint to be tested in the Main Body of this Study.

1. Failure at the human interface: four case studies

The year 1980 may be labelled by historians and scientists as the "Year of Mishaps" for American technology. Instances abounded in the domain of control and operation of civil and military aerospace alone. Some conspicuous examples are:

(1) The numerous near-misses at Kennedy and other busy national and international airports.

(2) The technical problems with the operation of the Hercules Sea Stallion helicopters which resulted in aborting the rescue mission to free the American hostages in Iran.

(3) The NORAD military computer which falsely signalled a Soviet nuclear attack on three separate occasions.

(4) Numerous crashes of fighter planes on training missions.

These incidents (cases (1) and (3) will be among those investigated more specifically) dramatise the idea that as technology rapidly advances and computers are given more responsibilities, the role of the man-machine interface becomes more critical.

We shall address the issue

Problems of the "Human Window" -- the need for satisfactory conceptual interfaces between man and machine, as related to

(1) Three Mile Island

(2) Air Traffic Control

(3) NORAD Military Computer

(4) Royal Dutch Steel

Investigations have been carried out into the above four particular cases as examples of "user-inscrutability" of complex and sophisticated systems. In each case automation, the use of machinery to perform tasks that previously had been accomplished exclusively by humans, plays a key role. In the past two decades automation has been inherently related to computerisation as has been the case in the areas of interest in this report. Since at least for the forseeable future man intends to be "in charge" of communications with machines, and not vice-versa, it is necessary

that the relationship between man and machine be humanised rather than further computerised.

The mismatch between a technological system and the humans who operate it can be either at a "syntactic" or "semantic" level. A knowledge representation can be syntactically correctable. But if its semantic structure is wrong, then no amount of human technology can correct it. The nature of the underlying causes of mismatch between man and machine, whether essentially syntactic or semantic is the issue to which our investigations will be oriented.

It is no longer just laymen who are unable to comprehend how computers and the advanced information technology under which they operate, can co-ordinate within large sophisticated systems; control rooms in nuclear power stations, in air traffic, and on oil platforms are instances where large systems are now generally beyond the technical and scientific sophistication of those who operate them.

1.1 Three Mile Island

The nuclear power station at Three Mile Island, Pennsylvania (known as T.M.I.-2 since it is one of two plants there) is one particular case in point. In its conclusions in "The Accident At Three Mile Island" the Report of the President's Commission summarizes the "Causes of the Accident" (p11, bottom):

Trans

> "In conclusion, while the major factor that turned this incident into a serious accident was inappropriate operator action, many factors contributed to the action of the operators, such as deficiencies in their training, lack of clarity in their operating procedures, failure of organizations to learn the proper lessons from previous incidents, and deficiencies in the design of the control room."

In this overview of the "Causes of the Accident" the Commission stresses that it was the "lack of attention to the human factor in nuclear safety" which was to blame for the seriousness of the accident. While the training of operators and operation of the control room may have been adequate under normal circumstances, they were seriously deficient under accident conditions. Operators, even senior ones, did not have a sufficiently deep understanding of nuclear power under the complex prevailing circumstances. The specific procedures which operators had to follow as a result of the accident were at the least very confusing, and could reasonably have been interpreted to lead to the actions

which the operators had mistakenly taken. Furthermore, the lessons from previous accidents (which might have prevented the T.M.I.-2 accident altogether) did not result in new clear instructions being passed on to the operators.

The control room also proved lacking in many ways, particularly with regard to the "human interface". The commission report continues:

> "The control panel is huge, with hundreds of alarms, and there are some key indicators placed in locations where the operators cannot see them. There is little evidence of the impact of modern information technology within the control room."

Later we shall return in further detail to the role of computers and other human-interface related factors which were involved in the accident at T.M.I.-2. Let us now consider another area of growing concern for information technology.

1.2 Air Traffic Control (A.T.C.)

An article in The New Scientist (17 July, 1980) entitled "Near Misses in the Sky", focuses on the problems of computers in air traffic control. We cite some examples and summarize its main points.

On 31 October, 1979 at 6.52 P.M. a faulty connection in one memory element in the IBM-9020 computer at the air traffic control (ATC) centre in Leesburg, Virginia nearly resulted in a mid-air disaster. The computer went down for 6 minutes during which two planes, a northbound Boeing 737 and a southbound Delta Airlines Lockheed L-1011 some 320 km. away, at altitudes of 8700 and 9300 m. respectively, were heading for Wilmington, North Carolina. Then at 7:06, when the pilot of the L-1011 asked permission to descend to 8400 m. no-one at the control centre noticed the potential hazard. It was only because the 737 pilot saw the descending Delta and managed to swerve sharply that a mid-air collision was averted.

A routine fault analysis seconds after this incident indicated that a malfunctioning memory component had attempted to contact each controller's display console. However, the memory component was part of a newly reconfigured system and had been misprogrammed to contact one more display console than existed. This mistake and the subsequent non-response from the "missing" termi-

nal, resulted in the shutdown of the entire IBM-9020 system.

When this and other such incidents were reported to the U.S. Federal Aviation Administration (F.A.A.) which oversees A.T.C. centres as well as airport control towers, the response was: "no accident has ever been directly or indirectly associated with a computer failure."

Read A.T.C

Similar computer malfunctions occurred more than 850 times in 1979 at the 20 ATC centres across the United States. After the particular incident above, ATC controllers complained, and one Leesburg controller stated: "This type of accident will happen again. We can't keep defending this machinery. If the equipment is obsolete, admit it and take corrective action."

Perhaps this air traffic controller has missed the main cause of his problems. It is not the failures of machine hardware which are entirely to blame -- it is the methodology of fault diagnosis and how these faults are conveyed to the humans in charge that is responsible. In this sense the correct functioning of the hardware is the <u>syntax</u> component of the system. However, once a fault occurs somewhere along the control system, the issue of <u>semantics</u> just "what is going on" in humanly comprehensible terms, becomes the key. When aircraft identity and altitude data (which normally appear on the controller's display screen along with the aircraft's radar images) suddenly disappear, then the air traffic controller is plunged into a vacuum, the cognitive equivalent in an automobile of the discovery that the windscreen has fogged and the foot brake is not functioning.

The New Scientist article goes on:

> "A few weeks later, on 25 November, the A.T.C. centre at Fort Worth, Texas, experienced a power fluctuation which blew 32 fuses. The computer system failed, leaving display screens blank, and depriving controllers of even raw radar data for 4 minutes. No near-misses were reported, but the mishap caused another failure in the I.B.M. computer. This led the system to shut down again on 28 November.
> During this second failure, controllers at Fort Worth were following the progress of 19 aircraft, including two American Airlines Boeing 727's. Just before the breakdown, the controllers gave permission to one 727 to descend to 5400 m., the exact altitude at which the other 727 was flying only 130 km. away. When the displays flickered back to life 4 minutes later, an alert controller saw the conflict and advised the descending airliner to slow down. The two aircraft passed with only 180 m. between them."

In 1979 the Leesburg centre alone had 200 failures, while Fort Worth had 74. New York's A.T.C. centre at Kennedy Airport had the longest breakdown on 14 November, 1979 lasting 13 hours. Fortunately there were no close calls, but the costs in tension, discomfort and dollars can be surmised when we consider the fact that 1.2 million gallons of excess fuel were burned.

The events reported above and many more incidents which were not recorded specifically as "near-misses" led the Professional Air Traffic Controllers Organisation (PATCO) to conduct a survey of computer failures in the F.A.A. air traffic control centres. The survey led to allegations at a Congressional hearing (April 9,1980) that although the F.A.A. has $3000 million to spend on computers, it persists in operating a computer system that is unreliable and "possibly senile". Just weeks before this hearing Congress received a report from the Government Accounting Office which concluded that the F.A.A.'s "past efforts to deal with mid-air collisions have been hampered by a lack of internal coordination, and disagreements over policy, approach, timing and direction" (ibid., p189). These problems sound very similar to those of the Nuclear Resulatory Commission and their role in directly or indirectly causing the accident at Three Mile Island.

Problems with computer breakdowns in A.T.C. were highlighted by two further incidents.

The first event occurred on 18 January, 1980. Gerald O'Brien, became the first controller ever to be charged by the F.A.A. with tampering with A.T.C. equipment. Allegedly he willfully removed data from computerised radar scopes and thereby contributed to the potential endangerment of a Soviet Aeroflot jetliner approaching Kennedy International Airport. This particular Aeroflot flight included the Soviet Ambassador, Anatoly Dobrynin on it. The incident occurred at a time when the local chapter of PATCO was publicly opposing the handling of Soviet or Iranian planes, because of the Soviet intervention in Afghanistan and the holding of American hostages in Iran. After a shift change at Kennedy's Instrument Flight Rules (IFR) room, an alert supervising controller became aware that the Aeroflot jet's radar blip did not carry any computer-generated alphanumeric identifying data block, normally indicating airline flight number, altitude and ground speed. The supervisor then called the Air Route Traffic Control Centre (ARTCC) to clarify the situation. The fact that there was another plane in the same air sector as the Aeroflot led to a confused instruction from the supervisor telling the Aeroflot to descend 10 miles too early in the normally very busy air space over Long Island. Fortunately there was no other traffic in the area so that no serious danger was present.

The second incident was just one of a series of close calls during the summer months at New York City's airports. On 9 July, 1980, about 10 miles east of Kennedy, a British Airways Boeing 707 and a small private twin-engine Cessna passed within 150 m. of each other. This was not directly caused by any computer or human error, despite the fact that numerous mechanical breakdowns and human operational errors had been occurring at just this time. The computer in the IFR room did not automatically provide the British plane's data block, next to the plane's "blip" on a radar scope, but this was obtained manually.

The real problem lies in the crude "see and avoid" system intended to assure safe separation of controlled and uncontrolled aircraft mingling in the same airspace. Theoretically, under this system, since the weather was clear, the pilots of the planes involved should have been able to see each other and avoid danger. But the small craft was only detectable as a radar blip and was only tracked due to the initiative of a controller who gave specific instructions for the computer to obtain further data.

We can see that present methods of A.T.C. need to be carefully scrutinized, for in a number of cases disasters have been avoided only by astute human efforts beyond the call of duty. Not enough attention has been given to how operators must prevail in situations where computers malfunction. This is perhaps more confusing and hazardous than when humans can rely on radar and radio transmissions alone.

There is a clear split between long-term U.S. and European views of what the roles of people should be within A.T.C. In Hedley Voysey's article, "Problems of mingling men and machines" (New Scientist, 18 August 1977) the U.S. view is presented by a quote from Andres Zellweger of the Advanced Concepts Staff of the F.A.A.:

From / Read

> Today's A.T.C. system, which employs over 25,000 air traffic controllers, is overly labour intensive and, with the current traffic control procedures, will become even more so in the future. The F.A.A. plans call for increased automation of controller function with a human role change from controller of every aircraft to A.T.C. manager who handles exceptions while the computer takes care of routine A.T.C. commands.

Read

The attitude and direction of American A.T.C. is diametrically opposed to the European outlook. Peter Sturgeon, of the Applied Psychology Department of Aston University, has for some years been examining various aspects of the A.T.C. man-to-machine rela-

tionship. He sees the European approach as aimed at a partnership (symbiosis) between man and machine, which is more far-sighted and superior to either working alone. Due to real doubts over the reliability of U.S. methods, European and U.K. controllers have been reluctant to adopt the U.S. approach. The main concerns of the Aston University group are:

Trans

(1) Will the controller who has to intervene in an exceptional case be properly placed to do so?

(2) Over long periods of time the sheer lack of verbal communication between A.T.C. and individual aircraft may lead to mistakes in instructions or

(3) a generally increasing reluctance of humans to intervene in the system at all.

The F.A.A. plans extend almost to the end of the century and thus it is unlikely that there will be a shift in the U.S. approach for some time. Later we shall investigate what efforts are being made employing artificial intelligence methods to improve A.T.C.

1.3 NORAD Military Computer

Within an 8-month period during 1979-1980 the U.S. experienced 3 false alerts indicating that it had been attacked by Soviet missiles. These were all due to computer error. The first reported false alert occurred on 9 November, 1979 and was the result of a mechanical error when a war game information tape was inadvertently fed into live channels, setting off early warnings of a nuclear missile attack from Soviet submarines probably located in the north Pacific. This meant that 10 jet interceptors from three bases in the U.S. and Canada were scrambled aloft and missile bases throughout the U.S. were put on low-level alert.

While the six-minute alert had been considered sceptically enough not to notify the President or Secretary of Defence, if it had lasted just one more minute the incident would have been brought to their immediate attention. There have been several such false alarms in the past, notably in the late 1950's and early 1960's caused by computer failures, natural phenomena, and test firings, but this was the first where the command went out from the NORAD (North American Air Defence Command) centre in Colorado Springs, Colorado, to the vast complex of defence centres chained across the United States.

A second false alert occurred on 3 June, 1980, again within NORAD due to a computer error. The warning indicated a missile attack from the Russian mainland and from submarines. Instantly the alert spread to the U.S. strategic air command in Nebraska and to the National Military Command Centre in the Pentagon.

It required only 90 seconds for the Command Staff buried 500 m. below Colorado's Cheyenne Mountain to check the alarm against radar and satellite information to confirm that there was actually no evidence of a Soviet attack. A minute later it was decided that the nuclear alert which had been transmitted to U.S. military command posts around the world, was cancelled. Thus the false alarm lasted only 3 minutes, but it required 20 more minutes for the U.S. strategic forces to stand down.

On 6 June there was yet another alert, but this was an intentional one, in an effort to duplicate the circumstances surrounding the first event on 3 June. As previously, the alert was not deemed serious enough to notify President Carter or Secretary of Defence Harold Brown, although the "situation room" in the White House and the President's command post were told.

A two-week investigation headed by Gerald P Dineen, an Assistant Secretary of Defence, revealed that the false alerts were caused by a single faulty integrated circuit.

Trans

The key precaution built into the NORAD alert system is that it is an 8-stage process whereby a key human decision must be made at each stage. This prevents machinery alone from ordering a nuclear strike. Defence officials at Cheyenne Mountain gave the further assurance: "If there are eight different stages, then we were only at step one ..."

However, the real concern of U.S. officials are two circumstances which could directly result from a false alarm:

Trans

1) the "shrinking time factor" and 2) the possibility of "escalating responses".

1) the "shrinking time factor" is the critically short period of time in which humans in either the U.S. or Soviet Union must be able to read computer signals and make decisions. This is directly related to the fear that since the Soviet warning systems and technology are less sophisticated than those in the U.S., they would have less time to consume for making a decision and testing the certainty of their information. Perhaps even more pertinent to this shrinking time factor is that approximate-

ly 75% of the Soviet nuclear strike force is on land-based launchers. This means that they are most vulnerable targets for U.S. attack and that they require more time to reach their targets as compared with the more balanced land-based, airborne and seaborne force in the United States.

2) The possibility of "escalating responses". Though these nuclear alerts were discovered to be false and the result of some computer error in 6 and 3 minutes respectively, it is quite significant that 20 minutes were required to bring U.S. forces down from their higher state of alert. An alert, whether or not false, means that there is sudden, great activity at military command posts. Bomber crews start their engines, some planes even take off, land-based missile silos are brought closer to firing, and ballistic missile submarines receive signals. There is little doubt that the Russians could spot at least some of these movements; they could then respond quickly and with more magnitude. These "escalating responses" could continue until a full scale nuclear confrontation could result from a simple misunderstanding of normal precautionary steps due to a false alarm.

1.4 Royal Dutch Steel

The study of automation and how it affects the operators of a large control system requires the cooperation of many different parties. At the highly automated Hoogovens hot strip mill of Royal Dutch Steel such a rare cooperation between management, psychologists, ergonomists and workers was achieved. The study was carried out by specialists of the British Steel Corporation, the Technical University of Delft and Hoogovens, together with production management for a period of 18 months from February, 1975. The productivity of the Hoogovens hot strip mill had dropped abruptly and the key question was, to what extent was this problem caused by a newly installed highly automated control system.

The following summary of the main conclusion is from Hedley Voysey (New Scientist, 18 Aug. 1977, pp. 416-7):

> "The operators became so unsure of themselves that, on some occasions, they actually left the pulpits used for control unmanned ... The operators also failed fully to understand the control theory of the programs used in the controlling computer, and this reinforced their attitude of "standing well back" from the operation - except when things were very clearly going awry. By intervening late, the operators let the productivity drop below that of plants using traditional control methods. So automation had led to lower pro-

ductivity and operator alienation simultaneously."

An idea which had appealed to the plant's designers was to enclose the steel strips being rolled. However this seriously obstructed the ability of the plant's experienced operating staff to assess when there was a serious computer failure, either in the complex programs or electrical input sensors.

Read

The remedy for this unsatisfactory situation at the Hoogovens plant was not simple either. It required that operators be taught the intricacies of the system they were running. This involves the control methods, the operation of programs in process terms, and all visual cues which might be helpful in detecting problems. With the complex physics of the system, the theory of the control system is intricate as well. Again, as in the case of T.M.I.-2, we see the need that operators should not just be button pushers who can only deal with a limited number of situations, but highly skilled and trained individuals, with appropriate salaries for their responsibilities.

We summarize some of the further conclusions of the Hoogovens study, which was entitled: Human Factors Evaluation Hoogovens No. 2 Hot Strip Mill.

(a) There is no evidence that automation at the Hoogovens plant forced people into jobs which are socially unacceptable.

(b) Automatic systems should be designed so that it is possible for the operator to anticipate potential problems and take preventive action, rather than react to problems only after they have already arisen.

(c) Information displays should be designed to help the operator predict performance and to help him understand the decisions being taken by the automation, as opposed to the use of displays only to indicate the state of a process.

(d) The visual and auditory information which the operators use in the present mill is of paramount importance to their task, and special attention should be given in future designs to providing the best possible view of the process.

(e) A suitable form of back-up is required in the event that the process computer is out of action or inaccurately set up.

(f) The facilities for operator interaction with the computer system should be extended, particularly to enable him to

monitor and improve the performance of the automation; this would include the alteration of incorrect or suspect data to produce a "better set up"; by "game playing" with the computer to examine the desirability of alternative courses of action; and by using, where necessary, the operators' corrections to the automatic control as well as automation feedback loops, to update the computer set ups.

(g) Operator training should emphasize the function and performance of the automation. Operators should be trained so that they can define and articulate their technical needs clearly to production and automation engineers.

These conclusions are in line with the general European view of what the relationship between man and computer should be: a partnership which is superior to man or computer working alone.

1.5 Viewpoint to be tested in the main body of this study

For certain special tasks which are socially critical, stand-alone information systems should not be entrusted with operational supervision. Certification should only be granted to systems which demonstrably augment the user's understanding of his task environment. A clear distinction between "surface" (cosmetic) and "structural" (conceptual) causes of misunderstanding in an information system is needed. Some tasks may be too complex for elimination of the latter cause of user-inscrutability to be possible. This theoretical case can be demonstrated in model domains on a laboratory scale, and must be kept in mind as real-life task environments of ever-increasing complexity are penetrated by machine systems. An analogy exhibits this last point: If a patient were to enter a doctor's office complaining of a boil on the thigh, lancing could be the indicated treatment. If the true problem were dislocation of the hip, surface treatment of any kind would be ineffective. In software today, the issue of "deep" inscrutability has yet to be addressed by the profession at large.

Commitment must therefore be to interactive man-machine systems wherever possible. Within this context the only computer-based decision structures which can be regarded as safe are those amenable to conceptual debugging at run time. No commercially available software of the present state of the art has this property, with the exception of little-known packages of the "expert systems" type. These special A.I. systems can be extended to incorporate "high-level expertise" in fault diagnosis and correction both in other programs and also in themselves. We do not see any other path to ultimate safe solutions. We have therefore taken the step of including in Annex B.1 a survey and analysis of the

feasibility of automatic fault-diagnosis and debugging of programs.