Aus dem Institut für Multimediale und Interaktive Systeme der
Medizinischen Universität zu Lübeck (Direktor: Prof. Dr. rer. nat. M. Herczeg0

Sicherheitskritische Mensch-Maschine-Systeme

M. Herczeg

Unser tägliches Leben wird zunehmend von der Verfügbarkeit und Funktionsfähigkeit technischer Systeme abhängig. Wir erleben diese Systeme beispielsweise in Form von Fahrzeugen, Produktionssystemen, Kraftwerken und medizintechnischen Geräten. Die meisten dieser komplexen Systeme müssen von Menschen überwacht und gesteuert werden. Aufgrund des hohen Sicherheitsbedarfs bei diesen Anwendungen sprechen wir auch von *sicherheitskritischen Mensch-Maschine-Systemen.*

Dieser Bedeutung werden die Grundlagen und Methoden zu Analyse, Entwicklung und Betrieb dieser Systeme nur bedingt gerecht. Die besondere Schwierigkeit liegt in der Verknüpfung der in ihren Eigenschaften unterschiedlichsten Teilsysteme *Mensch* und *Maschine*. Diese Problematik zeigt sich in gelegentlichen Störfällen und Unfällen dieser Systeme deren Ursachen im allgemeinen ähnlich unangemessen beschrieben werden, wie es schon die Konzeption der Systeme widerspiegelt.

Mensch-Maschine-Systeme

Mensch-Maschine-Systeme sind Konstrukte, bei denen Menschen mit Hilfe von Maschinen, heutzutage vor allem Computern, Dinge konzipieren, produzieren, speichern, wiederfinden, ändern oder auch überwachen und steuern.

So muss beispielsweise ein Maschinenbau-Ingenieur an einem CAD-Arbeitsplatz als ein Mensch-Maschine-System verstanden werden. Dasselbe gilt für einen Autofahrer hinter dem Lenkrad, einen Piloten im Cockpit oder auch für einen Mediziner an einem Diagnosesystem.

Es gibt ein grundsätzliches, gewissermaßen naturgegebenes Problem mit Mensch-Maschine-Systemen, nämlich dass Mensch und Maschine nicht besonders gut zusammenpassen. Dieses Problem deutet sich bereits an, wenn man versucht, eine Busfahrkarte an einem Automaten zu lösen, einen Videorecorder zu programmieren oder eine komplexe Zeichnung mit einem Graphiksystem zu erstellen.

Zur Lösung des Problems gibt es unterschiedliche Vorgehensweisen:

1. die Maschine wird passend gemacht
2. der Mensch passt sich an
3. beide passen sich an

Man setzt bislang üblicherweise auf Variante 2, die Anpassung des Menschen an die Maschine.

Sicherheitskritische und risikobehaftete Systeme

Wenn wir von sicherheitskritischen Mensch-Maschine-Systemen sprechen, meinen wir damit *stark risikobehaftete Mensch-Maschine-Systeme*. Damit wurde ein Begriff durch einen anderen ersetzt. Allerdings ist der Begriff *Risiko* definiert, nämlich als das Produkt aus der *Wahrscheinlichkeit des Eintretens eines ungewünschten Ereignisses*, gewissermaßen eines Systemfehlers, und dessen *Tragweite oder Kosten*, die durch das eingetretene Ereignis entstehen.

Risiko = Wahrscheinlichkeit des Ereignisses * Tragweite des Ereignisses

Beispiele für sicherheitskritische Mensch-Maschine-Systeme:

- Transportsysteme (z. B. Autos, Flugzeuge, Schiffe, Bahnen)
- Verkehrsüberwachung (z. B. Flugsicherung, Schiffsleitzentralen, Bahnstellwerke)
- Produktions- und Versorgungssysteme (z. B. Chemiewerke, Wasserwerke, Gaswerke, Kraftwerke)
- Medizintechnische Systeme (z. B. Anästhesiesysteme, Diagnostiksysteme, Bestrahlungssysteme)

Die Mensch-Maschine-Schnittstelle dieser Systeme nennen wir oft auch *Prozessführungssysteme*, *Supervisory-Control-Systeme* oder *Leitwarten*. Ihnen liegt ein dynamischer Prozess zugrunde, der mit ihrer Hilfe überwacht und gesteuert werden muss.

Incidents und Accidents – Störfälle und Unfälle

Gelegentlich erfahren wir von der Problematik des Versagens dieser Systeme in Form von *Störfällen* oder

Unfällen. Meist wird in den Medien von *menschlichem* oder von *technischem Versagen* berichtet. Die wirkliche Problematik liegt bei genauerer Betrachtung in den meisten Fällen dazwischen. Im Folgenden soll dies an einigen bekannten dramatischen Beispielen dargestellt werden.

Atom-Unfall in Tschernobyl am 26. April 1986

Der Unfall ereignete sich am 26. April 1986 während eines Tests. Durch diverse vorbereitende Arbeiten und betriebliche Randbedingungen war der Atomreaktor vom unter Zeitdruck stehenden Bedienpersonal fälschlicherweise auf unter 1 % anschließend wieder auf 7 % der Nennleistung gebracht worden, wobei ein Leistungsbetrieb unter 20 % nicht zulässig ist. Der Reaktor befand sich dadurch in einem gefährlichen instabilen Zustand.

Im Folgenden die letzte Minute vor der Katastrophe:

Der Test beginnt mit dem Schließen der Turbinenschnellschlussventile.

Der Druck und die Leistung im Reaktor steigen an.

30 Sekunden nach Testbeginn steigt die Leistung so stark an, dass das automatische Regelsystem die Leistungssteigerung nicht verhindern kann.

6 Sekunden später gibt der Schichtleiter den Auftrag, den Reaktor notabzuschalten. Der Notschalter wird betätigt.

Wenige Sekunden später erfolgen Alarmmeldungen über hohe Reaktorleistung und ein jäher Leistungsanstieg.

Innerhalb von 4 Sekunden schaukelt sich die Energieabgabe auf nahezu das 100fache der Nennleis-tung des Reaktors auf. Das Schnellabschaltesys-tem der Steuerstäbe benötigt für das Wirksamwerden (Bremsen) 18-20 Sekunden. Diese Zeit stand nicht mehr zur Verfügung.

Der eigentliche Druckaufbau im Reaktorkern dauerte etwa eine Zehntel Sekunde. Die obere, etwa 1000 Tonnen schwere Reaktorabdeckplatte, wurde angehoben und der obere Teil des 64 Meter hohen Reaktors zerstört. Es gab im Abstand von wenigen Sekunden mehrere große Explosionen.

Der Grund des Unfalles war, soweit wir es heute zurückverfolgen können, eine komplexe Kette von Fehlhandlungen und Fehlfunktionen, die im Wesentlichen auf mangelndes Verständnis des aktuellen Prozessgeschehens sowie auf mangelnde technische Einflussnahmemöglichkeiten auf den Prozess zu begründen sind. Ein klassischer und tragischer Fall von grober Fehlkonzeption eines besonders sicherheitskritischen Mensch-Maschine-Sys-tems. In offiziellen und inoffiziellen Darstellungen zum Vorgang findet sich als Ursache je nach politischem oder wirtschaftlichem Interesse die Feststellung, es sei menschliches oder auch es sei technisches Versagen gewesen. Die Anlage galt bis zum Unfall als mustergültig konstruierte und betriebene Anlage.

Airbus-Unfall in Warschau am 14. September 1993

Am 14. September 1993 verunglückte ein Airbus A320 der Lufthansa bei der Landung in Warschau. Aufgrund einer schwierigen Landesituation (Nässe, Windscherungen, unzulängliche Wettermeldungen, kurze Landebahn) wollte der Pilot zur Sicherheit frühzeitig die Schubumkehr einschalten und die Radbremsen bedienen, nachdem er mit dem ersten Hauptfahrwerk aufgesetzt hatte, das andere aber noch nicht mit vollem Andruck am Boden war. Er war wegen des Rücken- und Seitenwindes in einem vorgeschriebenen Landeverfahren weiter vorne in der Landebahn und in leichter Schräglage aufgesetzt.

Die Maschine verhinderte 9 Sekunden lang den Bremsvorgang und fuhr entsprechend auch nicht automatisch die Störklappen (Spoiler) aus. Dass die dann einsetzende Schubumkehr mit nur 71% Triebwerksleistung einsetzt, war eine für Piloten nicht übersteuerbare Entscheidung der Konstrukteure, um die Triebwerke zu schonen.

Was war das Hauptproblem dieser missglückten Landung? Die Maschine stufte das Flugzeug noch als fliegend ein, da ein Hauptfahrwerk noch nicht ausreichend eingefedert war und die Räder sich noch nicht ausreichend schnell drehten. Eine solche Schutzfunktion nennt man *Interlock*, eine technische Schutzverriegelung um Fehlbedienungen oder technische Fehler zu kompensieren.

Am 26. Mai 1991, eineinhalb Jahre zuvor, stürzte eine B767 der Lauda-Air über Thailand ab. In über 7000 m Höhe hatte sich die Schubumkehr fälschlicherweise eingeschaltet. Und genau das sollte ein solches Interlock verhindern.

Der Airbus „wusste" natürlich nichts von der Intention des Piloten, die Maschine so schnell wie möglich abzubremsen. Die Maschine konnte nicht mehr sicher vor Ende der Landebahn, die unzulässiger Weise auch noch mit einen Erdwall endete, zum Stehen gebracht werden.

Zum Glück verhinderte die Maschine nicht die letzte Strategie des Piloten, das Flugzeug mittels Einschlag des Seitenruders zum Schleudern zu bringen und quer zu stellen und dadurch die Fahrt bis zum Aufprall wesentlich zu verlangsamen. Diese unkonventionelle Methode, die wir auch als unkontrollierte Notreaktion bei Autofahrern kennen, rettete viele, leider nicht alle Menschenleben. Ein umfassendes Anti-Schleudersystem hätte auch diese, aus Sicht der Maschine gesehen, völlig unsinnige Aktion vereitelt, wenn es denn realisiert gewesen wäre.

Strahlentherapie-Unfallserie Therac-25

Bei der Realisierung von Gegenmaßnahmen gibt es große Unterschiede hinsichtlich der Zuverlässigkeit dieser Schutzmechanismen. So waren in den Jahren 1985 bis 1987 durch die Umstellung von Hardware- auf Software-Interlocks in einem medizinischen Bestrahlungsgerät für Elektronen- und Röntgenstrahltherapie namens Therac-25 mindes-tens sechs Menschen durch um mindestens den Faktor 100 zu hohe Strahlendosen ums Leben gekommen.

Das Bedienpersonal des Systems hatte mangels einer geeigneten Benutzungsschnittstelle an einem Bildschirm Fehlschaltungen herbeigeführt, ohne es zu merken. Eine Kopierfunktion erlaubte zur Bequemlichkeit die Übernahme von Ist-Parametern als Soll-Parameter mittels der Return-Taste. Dies war von den Benutzern gewünscht worden und wurde nach kurzer Zeit routiniert blind angewandt. Ständige Fehlermeldungen wie „Malfunction 64" führten zusätzlich zu einer erheblichen Desensibilisierung des Bedienpersonals. Eine Vielzahl weiterer Gestaltungsfehler könnten noch aufgezählt werden.

Beim Vorgängersystem Therac-20, das software-technisch praktisch genauso unzulänglich war, wurde dies durch ein hardware-basiertes Sicherheitskonzept abgefangen. Das Sicherheitskonzept des neuen Systems bestand aber fast nur noch aus Software-Interlocks. Die Wirkungen von Software und vor allem potentiellen Software-Fehlern war dabei praktisch nicht berücksichtigt worden. Software-technisch bestand der Hauptfehler darin, dass ein Zähler nur mit 8 bit ausgelegt war und beim unerwarteten Überlauf wieder bei 0 begann, was allerdings einen falschen Systemzustand symbolisierte.

In der Informatik wurde Therac-25 in der Vergangenheit praktisch nur in Bezug auf die Fragen sicherer Software und Verfahren zur Entwicklung sicherer Software diskutiert. Dabei wurde übersehen, dass in der Praxis jedes relevante Softwaresystem Fehler enthalten wird und dass das Beseitigen dieser Fehler in sicherheitskritischen Mensch-Maschine-Systemen durch geeignete Verfahren zwar so weit wie methodisch und wirtschaftlich möglich getrieben werden sollte, dass dies aber nie das grundsätzliche Problem ihres Vorhandenseins löst.

Ein Statement der FDA (US Food & Drug Administration) lautete: *„Ein großer Teil sicherheitskritischer Systeme kommt von kleinen Firmen, die resistent oder uninformiert bzgl. System-Sicherheit oder Software-Engineering sind."*

Dies mag prinzipiell richtig sein, ein ergänzender Kommentar dazu müsste jedoch lauten: *„... und selbst das würde vorerst nicht viel helfen, da der State-of-the-Art im Software-Engineering die Konzeption von Mensch-Maschine-Systemen noch nicht angemessen berücksichtigt."*

Modelle der Prozessführung

Wenn wir Prozessführungssysteme modellhaft betrachten, sehen wir drei Hauptkomponenten: den *Operateur*, das *Prozessführungssystem* und den *Prozess*. Die Leitwarte bildet den Prozess ab. Sie soll ihn in eine für den Menschen hoffentlich verständliche und beeinflussbare Form übersetzen. Die Leitwarte wirkt auf diese Weise als *Medium* zum Prozess.

Bei diesem *Transformationsprozess* gibt es teils zwangsläufig, teils unbeabsichtigt Deformationen des realen Prozesses (Abbildung 1):

1. Reduktion durch maschinelle Sensorik erzeugt Lücken in der Abbildung

2. Artefakte durch maschinelle Sensorik lassen nicht Vorhandenes erscheinen

3. Transformation durch maschinelle Funktionen verfälscht die Sensordaten

4. Aggregation durch maschinelle Funktionen fasst mehrere Komponenten zu einer Komponente zusammen

5. Fokussierung durch maschinelle Funktionen reduziert auf einen Ausschnitt des ganzen Prozesses

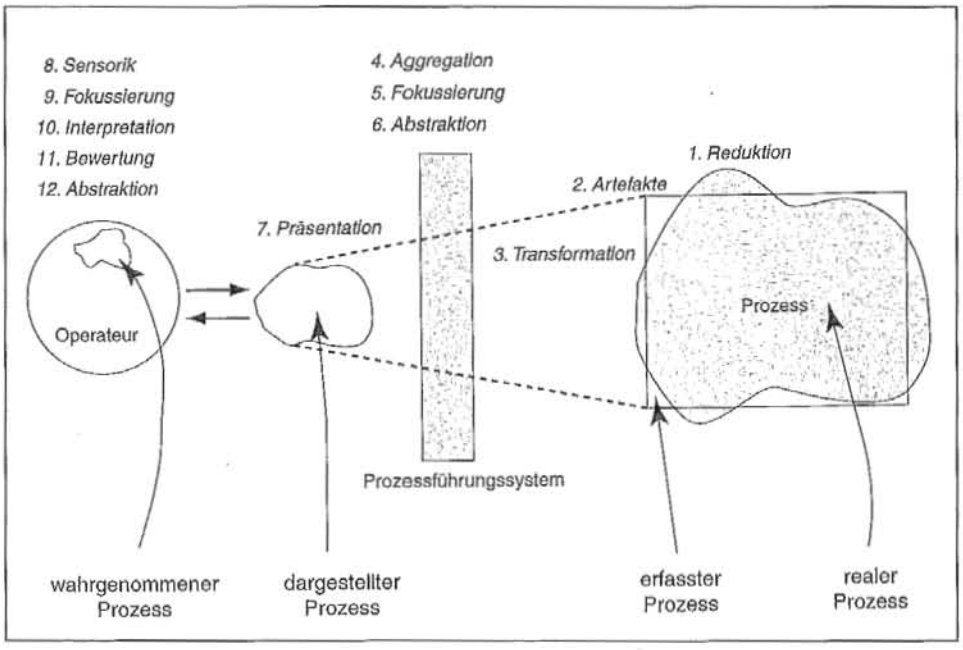

Abbildung 1: Transformationsprozess

6. Abstraktion durch maschinelle Funktionen vereinfacht die Realität durch Bildung abstrakter Prozessgrößen

7. Präsentation durch maschinelle Funktionen visualisiert auch Nicht-Visuelles mit unklaren Konsequenzen hinsichtlich mentaler Modellbildungen

8. Sensorik und Wahrnehmung des Menschen mit ihren Beschränkungen erfasst nur einen Teil des Präsentierten

9. Fokussierung durch den Menschen reduziert den präsentierten Ausschnitt durch weitere Ausschnittsbildung

10. Interpretation durch den Menschen versucht die Dekodierung des Wahrgenommenen zur Extraktion von Information und damit zur Erkennung von Systemzuständen

11. Bewertung durch den Menschen zur Erfassung der Bedeutung von Systemzuständen

12. Abstraktion durch den Menschen führt zu einer weiteren Vereinfachung des Wahrgenommenen

Durch diesen komplexen Transformationsprozess wird die Sicht und die Möglichkeit zur Einflussnahme auf den realen Prozess in vielfältiger Weise eingeschränkt und verändert. Dies ist auch notwendig, da die menschliche Wahrnehmung, Informationsverarbeitung sowie Handlungsfähigkeit in Relation zum Prozessgeschehen stark begrenzt ist. Die Kunst der Entwicklung eines aufgaben- und benutzergerechten Mensch-Maschine-Systems besteht nun offenbar darin, die richtigen Funktionen seitens der Maschine zu realisieren und diese mit den vorhandenen oder trainierten Wahrnehmungs- und Handlungsprozessen seitens des Menschen zu verschränken. Daraus besteht die gegenseitige Anpassung von Mensch und Maschine.

Ein Mensch, der eine solche Prozessleitwarte bedient, lässt sich ungefähr mit dem Modell von Rasmussen beschreiben (Abbildung 2). In diesem Modell stecken Konzepte menschlichen Wissens und Handelns sowie Prozesse, wie sich Wissen verändert.

In dem Modell werden 3 Ebenen unterschieden:

Wissensebene (Knowledge-Based Behaviour)

Auf der Wissensebene werden Situationen und Ereignisse unter Verwendung des zugreifbaren Wissens untersucht und Handlungen zielorientiert geplant.

Regelebene (Rule-Based Behaviour)

Auf der Regelebene werden in Abhängigkeit von erkannten Mustern (Bedingungen) erlernte Handlungen (Aktionen) ausgeführt.

Automatisierungsebene (Skill-Based Behaviour)

Auf der Automatisierungsebene werden sensorische Ereignisse erfasst und im Sinne von wahrgenommenen Mustern in die höheren Handlungsschichten weitergeben oder in Form von automatisiertem Wissen direkt in motorische Handlungen umgesetzt (Stimulus-Response- Reaktionen).

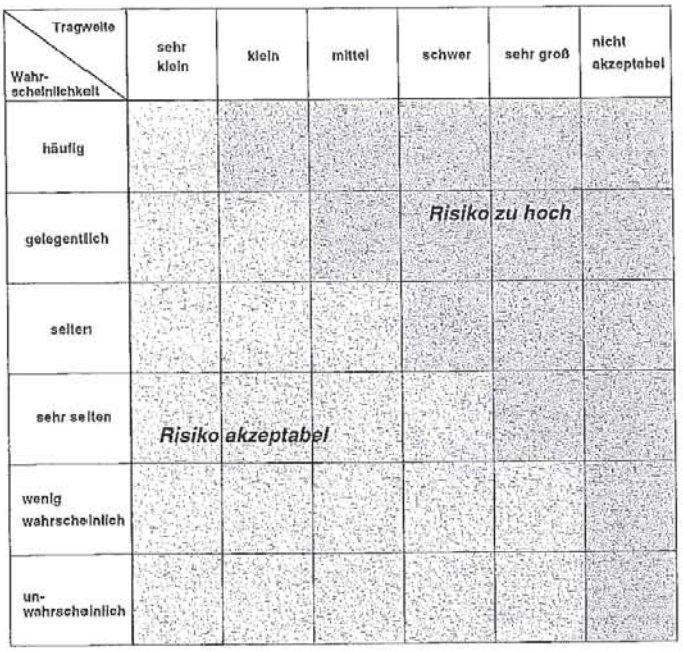

Abbildung 2: Prozessführungsmodell nach Rasmussen

Mentale Modelle

Bei Prozessführungsaufgaben ist es wichtig, dass Mensch und Maschine Wissensstrukturen besitzen, die aufeinander eindeutig abbildbar sind und die sich auch mit der realen Welt decken. Die Wissensmodelle nennen wir auch *Mentale Modelle* seitens des Menschen und *Konzeptuelle Modelle* seitens der Maschine. Die möglichst hohe Kompatibilität der mentalen mit den konzeptuellen Modellen ist die wichtigste Grundlage, die gegenseitigen Intentionen und Vorgehensweisen zu verstehen und sich dabei gegenseitig geeignet zu unterstützen oder zu korrigieren.

Sonst passiert das, was einem Autofahrer zustieß, der mit seinem Wagen an einem Fähranleger ins Wasser gefahren war, weil ihm sein Navigationssys-tem eine vermeintliche Brücke angezeigt hatte. Er hat den Hersteller des Navigationssystems dafür verantwortlich gemacht, worauf ihn dieser darauf aufmerksam machen musste, dass ihn ein Navigationssystem nicht davon entbindet, aus dem Fenster zu sehen. Dies wäre sogar der Betriebsanleitung zu entnehmen.

Aus dem „Fenster" zu sehen, ist allerdings nicht immer so einfach. Der Schichtleiter von Tschernobyl hatte kein angemessenes mentales Modell vom Prozess. Ihm war offenbar beispielsweise nicht bewusst, dass auch ein in Unterlast betriebener Reaktor aufgrund seiner dann vorhandenen kerntechnischen Instabilität gefährlich werden könnte. Ihm war auch nicht klar, dass das Abschalten dieses Reaktortyps in einem solchen Zustand zu einem dramatischen Leistungsanstieg führen

kann. Außerdem hatte er kein geeignetes „Fens-ter" in den Prozess des Kernkraftwerkes, das ihm geholfen hätte, sein fehlerhaftes mentales Modell zu korrigieren. Die Instrumente stellten ihm Tausende von Sensordaten aus dem Prozess gleichzeitig zur Verfügung. Der Gefahrenstatus des Kraftwerks hinsichtlich der Ursache der Instabilität war daraus nicht direkt zu entnehmen. Nur die Instabilität selbst war sichtbar.

Im Falle einer gefährlichen Störung beginnt eine Kernkraftwerksleitwarte sehr „lebendig" zu werden. Bei der Beinahekatastrophe von Three Mile Island (Harrisburg) liefen 3 Alarmsirenen, leuchteten Hunderte von Lämpchen und klingelten mehrere Telefone. Der Computer schaffte es erst mit einigen Stunden Verzögerung, alle Alarmmeldungen auszugeben. Außerdem befanden sich etwa 40 Personen in der Leitwarte, die alle helfen wollten.

Auch die Operateure des Therac-25 hatten kein klares Modell von der Funktion des Systems, sonst hätten sie die Möglichkeit einer Verstrahlung ihrer Patienten nach deren Klagen über brennende Schmerzen durchdacht und in Betracht gezogen. Sie haben die Anlage nur hochautomatisiert ohne tieferes Verständnis bedient. Eine Situation, in der aus Störfällen leicht Unfälle entstehen können.

Korrekte mentale Modelle vom Prozess, seinen Strukturen, seinem Zustand, seinen kausalen Bezügen und den Möglichkeiten der Steuerung sind unabdingbare Voraussetzungen für den sicheren und effizienten Betrieb von Mensch-Maschine-Systemen. In anderen Fällen ist das Risiko noch nicht einmal in erster Näherung kalkulierbar. Hierzu ist intensive Ausbildung, regelmäßige Schulung und Training eine wichtige Voraussetzung. So haben Verkehrspiloten ein intensives Auswahlverfahren und eine lange Ausbildung hinter sich, bevor sie zum ersten Mal eine Verkehrsmaschine fliegen und müssen halbjährlich ihre Fähigkeiten in einem anspruchsvollen Simulatorflug unter Beweis stellen, um ihre Lizenz zu erhalten. Die Fahrer von Kernkraftwerken werden im Allgemeinen in einem einfachen Bewerbungsgespräch ausgewählt, danach eingearbeitet, gelegentlich weitergebildet, aber nach der Ausbildung nie wieder einer grundsätzlichen Eignungsprüfung unterzogen. Letzteres gilt auch für Schiffskapitäne und in vielen Ländern übrigens auch für Autofahrer.

Intensives Training und Routine bergen bei sicherheitskritischen Systemen aber auch unerwartete Gefahren. Dies soll zunächst am Modell der drei Wissensebenen (Abbildung 2) erklärt werden. Durch Routine werden *Wissensstrukturen* in Regeln umgeformt. Wenn wir immer demselben Stau ausweichen, müssen wir nicht jedesmal von neuem Planen, sondern wir greifen im Falle einer bekannten auftretenden Bedingung auf einen vorbereiteten Plan, ein Verfahren, eine komplexe *Regel* zurück. Führen wir solche Regeln oft genug aus, werden Regeln zu Automatismen, die wir, obwohl wir sie ausführen, nicht mehr bewusst wahrnehmen. Regeln sind, informationstechnisch gesprochen, algorithmisiertes Wissen. *Automatismen* sind kompiliertes, in die menschliche sensomotorische Sprache übersetzte Regeln. Der Verschiebungsprozess von Wissen zu Regeln und von Regeln zu Automatismen ist Teil der Ausbildung, des Trainings, der Routine. Der Zweck ist die Fähigkeit zur zeitgerechten, ökonomischen Überwachung und Steuerung von Prozessen. Es wäre ohne diese Regeln und Automatismen nicht möglich, im Straßenverkehr zu bestehen. Genauso wenig wäre es möglich, zeitgerecht und effizient ein Flugzeug zu starten oder zu landen oder einen Zug zu führen.

Leider werden Automatismen und Regeln in manchen Situationen ausgeführt, in denen sie nicht angemessen sind. Die Zeit- und Aufwandsvorteile gehen zu Lasten der Korrektheit der Anwendung. Die Situation wird nicht mehr bewusst analysiert.

Die Bediener des Therac-25-Systems haben die Voreinstellungen des angezeigten Formulars bestätigt, ohne die Darstellungen zu lesen und zu beurteilen. Ihre routinierte Bedienung war Teil der Unfallursache.

Man nennt die Anwendung von im Prinzip richtigen Regeln aufgrund falsch wahrgenommener Zustände auch *Shortcuts (Abkürzungen)* im Falle daraus entstehender Fehler auch *Action Slips (Ausrutscher)*. Dabei werden sensorische Informationen auf einer zu niedrigen Wahrnehmungsebene verarbeitet und gelegentlich in Fehlhandlungen umgesetzt. Diese Shortcuts sind ein Effekt unzulänglicher und einseitiger Automatisierungsprozesse.

Shortcut-Effekte sind allgegenwärtig. Jeder von uns produziert täglich eine Vielzahl solcher Shortcuts, von denen viele leicht korrigierbar sind und oft nur selten bewusst werden. Das wirre Brems- und Lenkverhalten von Autofahrern in kritischen Situationen ist typisch dafür. Auch die Reaktionen der Tschernobyl-Operateure waren von falschen Regelbildungen und Automatisierungen gekennzeichnet.

Was können wir dagegen tun? Es ist nicht ausreichend, die Sicherheitsberechnungen von Mensch-Maschine-Systemen einseitig auf die Fehlerausfallwahrscheinlichkeit von Bauteilen abzubilden. Die Wahrscheinlichkeit von menschlichen Fehlhandlungen ist dabei einzurechnen, auch wenn sie ungleich schwerer zu messen ist.

Design for Error

Wenn wir schon wissen, dass menschliche Operateure Fehlhandlungen begehen könnten, müssen Mensch-Maschine-Systeme darauf vorbereitet sein. Wir nennen dies auch *Design-for-Error*. Das Prinzip besteht darin,

als Konstrukteur mit technischen Fehlern und Fehlhandlungen zu rechnen und dafür zu sorgen, dass das Gesamtsystem dadurch nicht in Gefahr gerät. Mensch und Maschine unter Einbezug ihres potentiellen Fehlverhaltens müssen hochgradig aufeinander abgestimmt werden.

Mensch-Maschine-Systeme werden komplexe Probleme erst dann geeignet gemeinsam lösen können, wenn sie gegenseitig Intentionen austauschen und abstimmen können. Wir nennen dies *Intention-Based-Supervisory-Control*, sicher einer der schwierigsten, aber auch vielversprechendsten Ansätze für komplexe arbeitsteilige Mensch-Maschine-Systeme.

Grundsätzliche Konzepte für Design-for-Error sind:

Maßnahmen gegen menschliche Fehler

- Redundanzen in der Syntax (kleine formale Fehler wirken sich nicht aus)
- Confirmations (Bestätigungsfunktionen)
- Checklisten (systematisches Abarbeiten mit einer vorliegenden Liste von Prüfschritten)
- aktiver Einbezug des Menschen in den Prozessverlauf (Gewährleistung von Vigilanz)
- Alive-Funktion („Tot-Mann-Taste")
- Interlocks (Technik als funktionale Redundanz)

Maßnahmen gegen technische Fehler

- Redundanzen in der Funktionalität (mehrere unabhängige Funktionen mit gleicher Wirkung)
- Entkopplung (Verhinderung von Fehlerausbreitung durch getrennte Teilsysteme)
- Accept-Funktion (Bestätigung der Funktionsausführung durch menschlichen Operateur)
- Fail-Safe (Fehler führen in sichere Zustände)
- Overruling (Mensch als funktionale Redundanz)

Maßnahmen gegen Interaktionsfehler

- Erstellen von Aufgabenmodellen und Szenarien (Analyse von Arbeitsabläufen)
- Abgleich der mentalen und konzeptuellen Modelle (Abstimmung von Mensch und Maschine)
- zeitgerechte Automatisierungsfunktionen (Automatisierung zeitkritischer Abläufe)
- gestufte, abschaltbare Automatisierungsfunktionen (Möglichkeit der menschlichen Einflussnahme auf Automatisierungsfunktionen)
- Incident-Reporting, dessen Auswertung und Umsetzung (Lernen aus Fehlern)
- Mensch-Technik-Redundanz (Mensch und Maschine überwachen und ersetzen sich bei Bedarf gegenseitig)
- Intention-Based-Supervisory-Control (Mensch und Maschine stimmen ihre Ziele und Arbeitsabläufe ab)

Die wichtigste Grundlage ist, dass die Begegnung von Mensch und Maschine in sicherheitskritischen Anwendungen als komplexes System verstanden werden muss und nicht als zufällige Begegnung, die sich durch Praxis und Gewöhnung fügen wird.

Unfälle und Beinaheunfälle sind in den meisten Fällen nicht reduzierbar auf technisches oder menschliches Versagen. So ist das Versagen des Mensch-Maschine-Systems eher das Unvermögen der Konstrukteure in der Konzeption, Beobachtung und dem Verständnis dieser Mensch und Maschine umfassenden Systeme.

Restrisiko und Ausblick

Unabhängig vom Aufwand für die Konzeption und Realisierung der Mensch-Maschine-Systeme, wird immer ein Restrisiko des Versagens des Gesamtsystems übrig bleiben. So bleibt die Verantwortung der Entscheidung, ob bestimmte Technologien verantwortbar sind, letztlich bei den hoffentlich informierten Bürgern demokratischer Gesellschaften.

Man sollte davon ausgehen, dass die Risikofunktion nicht linear verlaufen darf. Das heisst, dass die Funktion, die das akzeptable Risiko eingrenzt, eine Abbruchstelle dort hat, wo auch bei kleinster Eintrittswahrscheinlichkeit eines Unfalls eine Technologie und die Folgen ihres Versagens nicht verantwortet werden darf (Abbildung 3). Hinzu kommt, dass die realen die berechneten Eintrittswahrscheinlichkeiten bei Mensch-Maschine-Systemen aus den genannten Gründen systematisch übersteigen. Ein GAU in einem KKW, z.B. der Eintritt der Kernschmelze, dürfte nach Berechnung der Hersteller nur etwa einmal in einigen Millionen Jahren eintreten. In den letzten 25 Jahren hatten wir mindestens 3 dokumentierte Fälle. 2 dieser Fälle haben durch reine Glücksache und nicht etwa durch geplante Schutzmechanismen in der letzten Phase nicht zur Katastrophe geführt.

Es bedarf einer neuen Art von Auseinandersetzung mit der Frage einer besseren Verzahnung von Mensch und Maschine in sicherheitskritischen interaktiven Systemen. Dies betrifft insbesondere Techniker, Ingenieure und Entscheidungsträger, die Einfluss auf die Entwicklung dieser Systeme haben. Aber auch die Systemoperateure sind hinsichtlich der Fragen von ständiger Qualifizierung und aktiver Auseinandersetzung mit den Maschinen und Prozessen gefordert, die sie bedienen und steuern. Sicherheit in Mensch-Maschine-Systemen lässt sich nicht einfach erkaufen. Selbst das teuers-te Auto ist den Gesetzen der Physik und vor allem des menschlichen Geistes unterworfen. Dies sollte das System ständig widerspiegeln. Ein Autofahrersitz darf nicht zum Wohnzimmersessel werden.

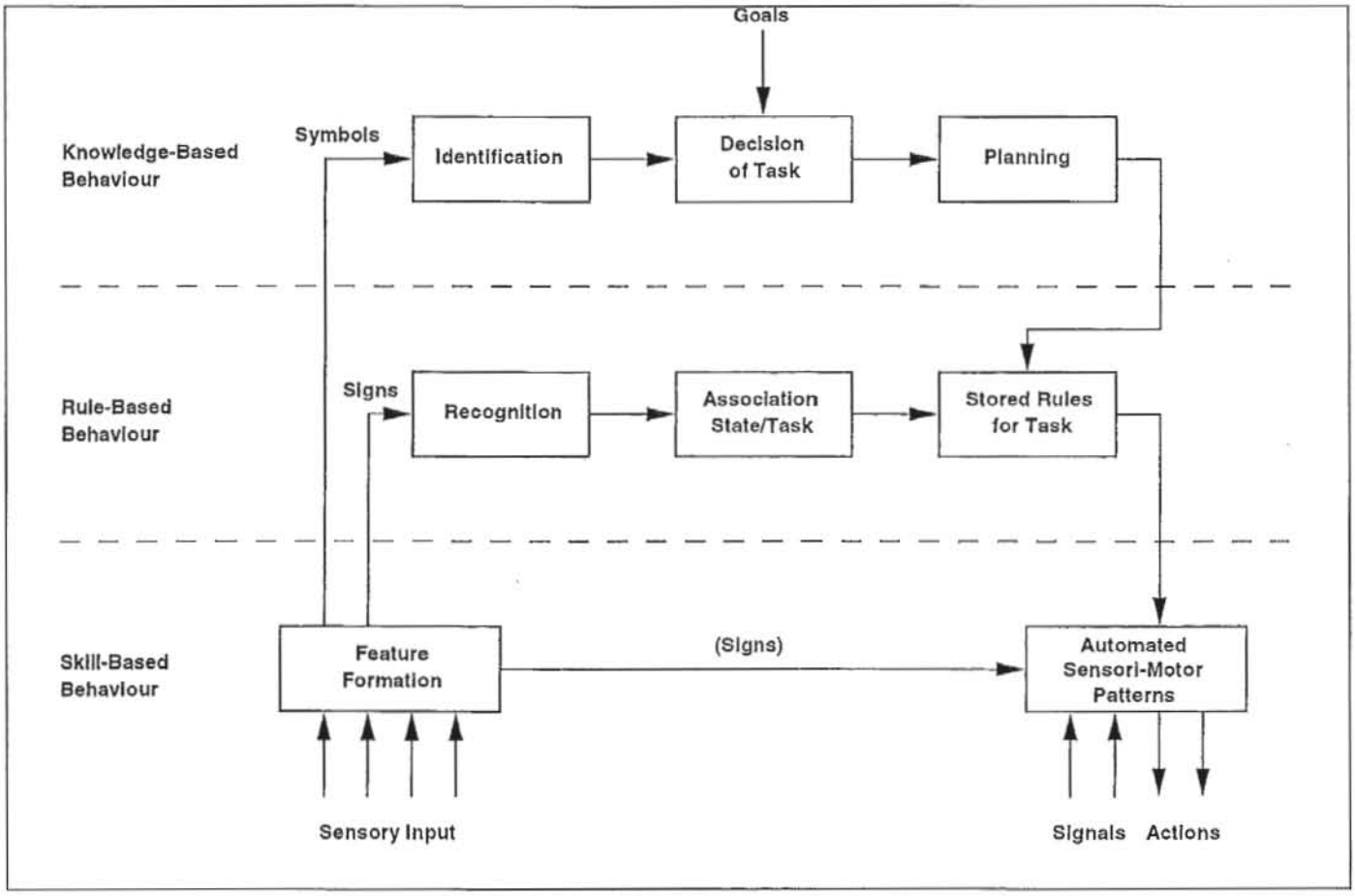

Abbildung 3: Risikomatrix

In den diversen Anwendungsbereichen, in denen sicherheitskritische Mensch-Maschine-Systeme realisiert werden, existieren sehr unterschiedliche Sensibilitäten, Erfahrungen und Rationalitäten, aus denen unterschiedlichste Entwicklungs- und Nutzungskonzepte entspringen. Auf der Straße verunglücken jährlich allein in Deutschland etwa 40.000 Menschen. Das entspricht der Einwohnerzahl einer deutschen Kleinstadt. Im Luftverkehr sind es etwa 1.000 Menschen jährlich weltweit. Durch verunglückte großtechnische Systeme können es viele Millionen in wenigen Minuten werden.

Welche dieser Systeme mit ihren realen Restrisiken jenseits der Akzeptanzschwelle liegen, ist allerdings letztlich eine gesellschaftliche Entscheidung.

Literatur

1. T. van Beveren. Runter kommen sie immer – Die verschwiegenen Risiken des Flugverkehrs, Campus Verlag, 1995
2. M. Czakainski. Tschernobyl – Der Reaktorunfall, Informationskreis Kernenergie, Bonn, 1996
3. R. Gerling, O.-P. Obermeier (Hrsg.). *Risiko-Störfall-Kommunikation*, Gerling Akademie Verlag, 1994.
4. R. Gerling, O.-P. Obermeier (Hrsg.). *Risiko-Störfall-Kommunikation 2*, Gerling Akademie Verlag, 1995.
5. M. Herczeg. *Software-Ergonomie - Grundlagen der Mensch-Computer-Kommunikation*, Addison-Wesley-Longman, Bonn und Oldenbourg Verlag, München, 1994.
6. M. Herczeg. *A Task Analysis Framework for Management Systems and Decision Support Systems*, Proceedings of the AoM/IAoM 17th International Conference on Computer Science, San Diego, California, August 6-8, 1999, Journal of Computer Science and Information Management (CSIM), The International Association of Management (IAoM) and Maximilian Press Publisher, 1999.
7. N. G. Levenson, C. S. Turner. An Investigation of the Therac-25 Accidents, J EEE Computer, Vol. 26. No. 7, July 1993
8. C. Perrow. *Normale Katastrophen - Die unvermeidbaren Risiken der Großtechnik*, Campus-Verlag, 1992.
9. J. Rasmussen. *Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models*, IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-13, No. 3, May/June 1983.
10. J. Rasmussen and L.P. Goodstein. *Decision Support in Supervisory Control*, Technical Report M-2525, Risø National Laboratory, Roskilde, Denmark, August 1985.
11. J. Rasmussen and L.P. Goodstein. *Information Technology and Work*, in M. Helander (Hrsg.), Handbook of Human-Computer Interaction, Kap. 9, S. 175-201. Elsevier Science Publishers B.V. (North Holland), Amsterdam, 1988.
12. T.B. Sheridan. *Supervisory Control*, in G. Salvendy (Hrsg.), Handbook of Human Factors, Kap. 9.6, S. 1243-1268. John Wiley & Sons, New York, 1987.
13. T.B. Sheridan. *Task Allocation and Supervisory Control*, in M. Helander (Hrsg.), Handbook of Human-Computer Interaction, Kap. 8, S. 159-173. Elsevier Science Publishers B.V. (North Holland), Amsterdam, 1988.