

## プラント運転員の認知行動過程シミュレーション

Simulation of cognitive behavior of a plant operator

代表研究者	東京大学工学部教授 Prof., Fac. of Eng., Univ. of Tokyo Shunsuke KONDO	近藤 駿 介
協同研究者	東京大学工学部助教授 Assoc. Prof., Fac. of Eng., Univ. of Tokyo Kazuo FURUTA	古 田 一 雄
	現在(株)新日本製鐵 Nippon Steel Corporation Eiji ORIHASHI	折 橋 英 治

A methodology to simulate the cognitive behavior of a plant operator was studied based on the current technologies of artificial intelligence. The truth maintenance system and the black-board control architecture have been selected for the simulation system in order to model opportunistic and nonmonotonic features of human behavior. The effectiveness of the system was verified by coupling it with a plant simulator of BWR (Boiling Water Reactor) and analyzing its response in several cases of emergency situation. The results indicated that this methodology would be useful for the design and evaluation of man-machine interface or operator support system of such large scale systems as nuclear power plant.

### 研究目的

米国スリーマイル島原子力発電所2号機(TMI-2)の事故により原子力発電所の安全にとってヒューマンファクターが極めて重要であることが認識され、以来マンマシンインタフェースの適切な設計を行うために多くの努力が払われてきている。現在では実規模を含む様々なプラントシュミレータが訓練のみならず運転員の事故、故障への対応振りを分析するために用いられており、その結果は次世代プラントの制御室設計などに反映されつつある。こうしたプラントシュミレータと被験者による認知実験は、実際と近い環境でマンマシンインタフェースの評価を行うのに極めて有力な方法であるが、設計の初期段階から行うにはコストも時間もかかる方法である。また、人間は非常に適応性に優れているために、限られた数の被験者から客観的かつ信頼性のある

データを得ることはむずかしいことも指摘されている。

一方、人間信頼性を評価するために、THERPに代表される確率論的手法がこれまでにいくつも提案されてきた。これらの手法は、運転員が誤った行動を起こす確率やそのシステムの信頼性に与える影響を定量的に評価するのに非常に有効である。しかし、それらは主に人間の振舞いの現象面に着目しており、ヒューマンエラーの発現機構を解明するものではない。ヒューマンエラーに存する因果関係を調べるためには、やはり、運転員の認知過程を機構論的に模擬できる何らかの手法が必要と考えられている。

以上の観点から、マンマシンインタフェース設計やマニュアル等も含めた運転支援用ソフトウェアの評価を、被験者の代わりに計算機シミュレーションで行うシステムの開発を試みた。このよう

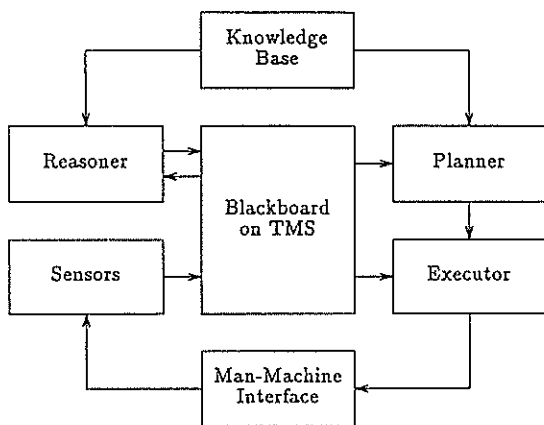


図1. システムの基本構成

な目的に用いられるシステムは、現実の人間のようには振舞うことが必要であり、合理的な思考を行う人間が犯すような誤りも再現できなければならない。本研究は、こうした人間的側面のシュミレーションを目的としてシステム開発を行っているものである。

#### 研究経過

運転員の振舞いをモデル化するための基本的枠組としては図1に示す黑板モデルを採用した。黑板には、運転員の認知的世界を表すプリミティブが書き込まれ、四つの問題解決機が黑板上の関連する記録を参照しつつ問題を解決する。

システムにおいて、観測の集合は運転員が事実として認識した外界での出来事を表し、信念の集合は運転員の思考における世界の解釈を表している。合理的な思考をする運転員においては両者は一致するはずであるが、過渡的な状態においては必ずしも両者が一致しているとは限らない。両者の間の不一致は、まだ信念の集合に受容されていない観測事実は何らかの説明を加えるか、まだ観測していない信念を確認することによって解消される。

四つの問題解決機によって実行されるメタタスクは以下に述べる7種類であり、このうち実行可能なタスクが複数ある場合には優先順位の最も高いタスクが選択され、実行される。各タスクに割り当てられる優先順位は問題解決の基本戦略を決

定する。優先順位は運転員の心理的状态を表す変数によって変わるようにできるが、現在は不安因子が高まると信念の確認に対する優先順位が高くなるだけになっている。

1) 無作為の監視：感覚器によって無作為に環境を監視し、観測したことを黑板に書き込む。本タスクの優先順位は最も低く、他に実行可能なタスクが何もないときに実行される。

2) 信念の確認：推論によって新たに信念の集合に加えられたもののまだ確認されていないことを確認するための行動を開始する。

3) 新たな信念の生成：現在の信念に知識ベース中のルールあるいはデフォルトを前向きに適用し、新たな言明を生成して信念の集合に加える。

4) 観測事実の受容：ルールあるいはデフォルトを後ろ向きに適用することによって、まだ信念の集合に受容されていない観測事実の説明を加える。ここで、もし観測事実が知識ベースと現在の信念とから証明可能であるならば、証明の過程に現れるすべての言明は信念の集合に加えられる。また、推論機は証明を完成させるために仮説を必要とする場合には、必要な仮説を生成する。観測事実が現在の信念と矛盾している場合には、矛盾の原因を突き止めてこれを取り除く。

5) 質問に対する応答：推論機は推論機以外の問題解決機が生成した質問事項に、知識ベースと現在の信念に基づく回答を与える。もし、質問事項に対する真理値が不定の場合には、それを確かめるための行動を喚起する。

6) 行動計画：計画機により、必要な行動目標を達成するための手続きの系列を生成する。ここで用いる計画機は、異なるステップの間の相互作用を考慮することが可能な非線形計画機である。

7) 行動の実行：完成した行動計画の各ステップを実行機によって実行し、行動目標を達成する。本タスクは最大の優先順位を持っている。

本シミュレーションシステムにおける情報処理は通常、センサによる環境の認知、推論機による観測事実の説明と理解、推論機による自発的行動の喚起、計画機による行動計画の作成、実行機による行動計画の実行の順序で進んでいく。しか

し、黒板制御を採用しているので、知識ベースの内容如何ではこのような順序に従わないタスクの実行、たとえば計画を飛び越した行動の実行や、理解を欠いたままの行動喚起などが考えられる。特に、運転員が十分に習熟した作業については、上に述べたようなすべての経路を経て進むよりも、むしろ短絡的なタスクの実行が行われる可能性が高い。このような現象は、J. Rasumussen によって提案された運転員の認知モデルともよく一致していると考えられる。

本システムの知識ベースに含まれる主要な知識のうち、ルールと正規デフォルトは、推論機がゴールの証明を行うのに使用される。ここで、各ルールはその体部が現在の信念で充足可能でさえあれば適用可能であるが、正規デフォルトはさらにそれを適用することによって信念の集合に矛盾を生ぜしめないときに限り適用可能となる。仮説はゴールの証明の際に必要なに応じて使うことができる仮説の候補を記述したものである。オペレータは運転員が行う基本的な動作を定義したものであり、計画機はこれを素材にして行動計画を作成する。スクリプトは行動計画の完成品あるいは半完成品で、習熟した行為を表しており、スクリプトが存在する行動に対しては計画作成が省略される。

信念の集合の一貫性は TMS (Truth Maintenance System) によって維持されている。運転員の各信念はそう信ずるに至った根拠であるところの正当化とともに黒板に記録されている。正当化は他の信念の二つのリスト、in リストと out リストから成っており、in リストのすべての信念が信じられており、out リストのすべての信念が信じられていないときに限り正当な根拠があると見なされる。推論機が新たな信念を生成するときには、推論の履歴を正当化として TMS に引き渡すようになっている。ただし、数値比較のように条件判定は表すものの信念を生成しない組み込み述語は、正当化のリストには含まない。デフォルトに基づいて生成された信念は自分と背反な信念を out リストに持っており、このような信念は翻意可能である。信念体系に何らかの変更が加えられ

たときには、TMS が関係する各信念の正当化の有効性を調べ、正当な根拠を失った信念を信念の集合から取り除く。

もし、観測の集合と信念の集合の間に矛盾を生じた場合には、観測を受容するように信念を修正する必要がある。そのような場合には依存関係によるバックトラックが起こり、TMS は矛盾の原因となっている仮定を捜し出し、これを否定するような新たな信念を加える。ここで、仮定とは空でない out リストを持つ信念のことで、そのような信念に対して疑念を生じた場合には、out リストに含まれる言明のうちのいずれかを信じることによって翻意させることが可能である。デフォルトに基づく信念は仮定として作られ、翻意可能である。

Doyle の TMS の一つの欠点は、状態の遷移と仮定の変更を区別できない点であるが、状態の遷移はかつて正当であった事実の否定を意味するものではないので、両者は明確に区別される必要がある。この問題を解決するため、時間的制約条件を考慮した新しい TMS を開発した。TMS 中の各信念はその有効時間によって指標付けされており、たとえば、

$\text{hold}(P, [X, Y], [(Q \cdot Z), [\sim P \cdot X]])$

といった信念は、時刻 Z から始まる言明 Q が信じられており、時刻 X から始まる言明  $\sim P$  が信じられていないことを根拠に、X から Y までの間言明 P が成り立つと信じられている、ということを表す。

状態の時間的な遷移を認識するためには、いったん正当化された言明はこれと矛盾する何らかの言明が信念の集合に加えられる時点までは有効である、という一種のデフォルト推論がつかわれる。もし、新たに加えられた信念が既存の信念と矛盾し、しかも前者が後者の始まった後から始まるようならば、時間的な状態の遷移が起こったと見なされ、TMS は前の状態を新たな信念の開始時刻で終了させる。この場合、信念の終了時刻は変更されるが、無矛盾性の維持管理プロセスは実行されない。

時間的遅れを伴う現象についての推論を行うた

め、delay という特殊な述語を導入した。この述語は、ルールあるいは正規デフォルトの体部に

$P < \cdot Q \text{ delay } X \cdot Y$

のように使うことができ、状態の Q が実現してから X 以上 Y 以下の時間遅れを伴って状態 P が実現する、という意味である。もしも、時間遅れが Y よりも短いときには時間遅れに関する信念は翻意可能であるが、時間遅れが Y に達するや P の実現にとって十分な時間経過があったとして、時間遅れに関する信念は翻意不可能となる。この述語により、シミュレーションシステムは状態の時間的遷移について人間の運転員が行うような自然で尤もらしい推論を行うことができる。

知識ベースには運転員が機械系に対して能動的に働きかけるべき状況を検知するためのルールあるいはデフォルトが含まれており、これらが適用された場合には行動計画の作成とその実行が起こる。

行動計画は手続きネットワークと呼ばれるデータ構造によって表されている。手続きネットワーク中の各ステップは運転員が行う特定の基本操作に対応し、計画機によって各テスツには適切なオペレータが選択され、割り当てられる。複数のステップは行動目標の達成にとって適切な順序に並べられるが、本システムで用いた非線形計画機はステップ間に生じる相互作用を考慮しながらステップの最終的な実行順序を決定する。

並列に展開されたステップは順序付けのための情報が得られるまでは並べられず、ステップの順序はテスツ間に生ずる相互作用を考慮に入れて決定される。他のステップの実行を妨害するようなステップはそのテスツの後に配置されるが、このような相互作用は同時性制約を用いて検出される。また、冗長なステップも検出され、取り除かれる。

よく習熟した動作に対してはオペレータを選択して順序付けるといったようなことはせずに、あらかじめ知識ベースに容易されたスクリプトを参照して使用する。スクリプトは手続きネットワークの半完成品であり、いわゆるスキルベースの行動を表している。

行動計画が完了すると、その中のステップが実行機によってつぎつぎに実行される。実際が実行機によって行われるのは、インターフェイスに対する何らかのコマンドの出力か、あるいはセンサ機能の実行である。

以上に説明したシステムを実動のプログラムとしてワークステーション上に構築した。使用言語には Prolog を用いた。また、システムのテストのために用いる簡易型 BWR プラントシミュレータの開発も行った。このシミュレータは日本原子力研究所から導入した BWRDYN コードを参考に、Fortran 言語を用いて新たに開発した。

次に、本システムのテストを行った。プラントシミュレータと運転員シミュレータとは、各々別の計算機上で非同期に実行された。ただし、2 台の計算機は 1 本の直列通信回線で結ばれており、運転員側からプラント側へは操作あるいは問い合わせ用のコマンドが、逆方向には各種のプラント情報が渡されるようになっている。

今回のテストでは、定格出力で運転中に原子炉給水ポンプ（タービン駆動）が何らかのトラブルによりトリップする場合を想定し、事象のシナリオ、知識ベース、推論戦略を変えながら単一故障の標準的な場合、多重故障の場合、先入観による固執を起こす場合、視野狭窄に陥りやすい場合、外部からの助言を得られる場合の五つのケースについて運転員シミュレータの振舞をし調べた。以下、運転員とは運転員のシミュレーションを意味するものとする。

#### 1) 単一故障の場合

まず、故障が起因事象だけの単純なシナリオで、かつ必要な知識と合理的な推論戦略をシミュレータに与えた場合にえられる振舞の概要を表 1 に示す。

運転員はアラームの発報直後、アナンシエータの確認によって直ちに給水ポンプが何らかのトラブルによりトリップしたことを認識し、給水流量の減少によってこれを確認する。次に、定常運転時には待機状態にある別系統の給水ポンプ（モータ駆動）を手動で起動し、給水流量を確保しようとする。炉内の水位が規定値以下になるために、

表 1. システムの振舞いの概要 (単一故障)

時刻 (秒)	プラントの振舞い	運転員の振舞い
0.0	TD-RFP トリップ, アラーム発報	アラーム受信
1.5	給水流量低下	アナンシエータ確認, TD-RFP 故障を仮定
2.3		アラームリセット
5.4	水位低下	給水流量低下確認
7.4	MD-RFP 起動, 給水流量増加	MD-RFP 手動起動
9.3		MD-RFP 起動確認
11.2	アラーム発報	アラーム受信
	水位低, スクラム, RCP トリップ	
12.9	炉出力低下, 再循環流量低下 炉停止, 水位上昇	水位低, スクラム, RCP ト確認
15.9		アラームリセット
19.0		炉停止確認
28.0		再循環流量低下確認
34.5		水位上昇確認
51.5		給水流量低下確認
57.5		水位回復確認

スクラムがかかって原子炉は停止するが、モータ駆動給水ポンプによって給水が回復されるために水位はほどなく回復する。この間に運転員は炉停止の確認や水位の監視などの必要な行動をとる。

このケースでは運転員が必要な知識を備えていたために適切な行動がとられたが、必要な知識の一部を欠いているような場合には当然これとは異なった振舞を示す。例えば、モータ駆動給水ポンプの起動に係わる知識を欠いた場合には、給水が回復されずに炉内水位の低下が続き、主蒸気隔離弁の閉止と高圧炉心スプレー系の起動に至る。炉内水位は結局は回復されるが、運転員がモータ駆動給水ポンプの手動による起動の必要性に気づいた場合よりは大幅に遅れることになる。

## 2) 多重故障の場合

タービン駆動給水ポンプのトリップに加えモータ駆動給水ポンプも起動せず、さらに炉内圧力を低下させるために開いた逃し安全弁が開固着するという。多重の故障を想定したケースでシミュレーションを実施した。

モータ駆動給水ポンプを手動で起動するまでは前のケースと同じであるが、運転員はポンプが起動しないのに気づいてモータ駆動給水ポンプにも何らかのトラブルがあったと想定する。その後、

スクラム、再循環ポンプトリップ、主蒸気隔離弁閉止、高圧炉心スプレー系起動等の安全保護動作がつつぎに起こり、運転員はアラームの処理や炉停止の確認に追われる。主蒸気隔離後に炉内圧力は上昇し、規定値に達したところで逃し安全弁が開き圧力は低下し始める。吹止り圧力まで低下したところで閉まるはずの逃し安全弁は閉らないが、圧力を監視していた運転員はこれに気づいて弁の開固着であると正しく認識する。炉内水位は弁の開固着によって若干回復が遅れるものの、結局は高圧炉心スプレー系の働きによって回復される。

最初に運転員はモータ駆動給水ポンプや逃し安全弁の故障は意識しておらず、それらが正常に動作することを想定して推論を行っていく。しかし、期待した結果が観測された事実と異なるに及んで、これらの機器が正常に動作しなかったことを認識している。ここで、知識ベースにはモータ駆動給水ポンプあるいは逃し安全弁の動作に関する知識が正規デフォルトとして与えられているが、これらが故障する可能性は仮説として陽に与えていない。にもかかわらずシステムが正しい異常原因に到達することに成功しているのは、信念と観測事実の間の矛盾を解消する過程を一種の推

論として利用しているからであり、これは予想外の事象に対して人間がよく用いる推論方法であると考えられる。

### 3) 先入観による固執を起こす場合

上記の多重故障と同じシナリオにおいて、モータ駆動給水ポンプあるいは逃し安全弁の動作に関する知識をデフォルトではなくルールとして与えてシミュレーションを行った。こうすることによりシステムはこれらの機器の動作を疑わなくなり、一種の先入観を導入することができる。このような場合、信念と観測事実との矛盾に気づいてもその原因を同定することが不可能となり、矛盾解消プロセスが無限ループに陥る結果となった。このような振舞は、先入観よりある考えに固執している運転員が正しい異常原因を同定できずに無為に時間を過す結果となるのと一致している。

### 4) 視野狭窄に陥りやすい場合

すでに述べたように、本シミュレーションシステムでは不安因子というパラメータによって信念を確認するタスクの優先順位を変化させている。このメカニズムを使用せずにシミュレーションを行い、その影響を調べた。このケースでは、外界で起こっている出来事とは無関係に推論が瑣末な方向に偏りやすくなり、モータ駆動給水ポンプの起動等の重要な操作のタイミングを逃す結果となった。これは視野狭窄に陥りやすい運転員が、重要な事項を見落としてしまうような場合を表していると考えられる。

### 5) 外部からの助言を得られる場合

本システムを用いて、外部からの助言が得られる場合の運転員の振舞いをシミュレーションでできるかどうかを検討した。上記のケースと同じく視野狭窄に陥りやすい運転員に対して、モータ駆動給水ポンプ起動の必要性に関するルールがあらかじめ決められた適切な時間で発火するようにしてみた。すると、運転員はルール発火後（助言後）まもなく偏った思考から抜け出し、ポンプの起動その他の必要な行動を取るようになった。この結果より、外部からの助言のような効果もシミュレーションで考慮することが可能であることが分かった。

## 研究結果

プラント運転員の認知行動過程のシミュレーションを行うためのモデルが提案された。シミュレーションシステムは黑板モデルに基づいており、黑板には運転員の認知的世界が記録され、その内容と知識ベースを参照しながら問題解決機が7種類のメタタスクを実行する。信念の無矛盾性を管理するためにはTMSが用いられているが、各信念を時間で指標付けすることによって状態の遷移と過程の変更を区別して扱えるようにしている。BWRプラントシミュレータと組み合わせたシミュレーションを行ったところ、人間の運転員に特徴的な機会主義的で非単調な振舞いを模擬できることがわかった。また、黑板制御の柔軟な問題解決戦略とTMSにより、不完全な知識、固執、視野狭窄、外部からの助言などの効果が考慮できる見通しが得られた。

より現実的で複雑な状況での検証がなお必要ではあるものの、現在までに得られた結果によってこの方式による運転員行動の機構論的なシミュレーションの可能性が示された。

### 今後の課題と発展

すでにエディタなどの計算機環境の一部に対してはユーザ・モデルを用いたシミュレーションによる評価が試みられ、かなりの成功を収めている。本研究ではこうしたアプローチが原子力プラントといったより複雑なシステムの場合にも適用できる見通しが得られたが、これを設計評価手法として確立するためにはなお未解決の課題を残している。

まず、認知実験の結果との比較に基づいて、本システムが実際の運転員の行動を予測し得るかどうか、引き続き検証することが必要である。すでに我が国の内外で実規模プラントシミュレータを用いた認知実験がいくつか行われており、こうした実験の結果が入手できれば運転員モデルの改良を行うことも可能であろう。

また、本研究では時間的ストレス、疲労などの心理状態が問題解決戦略に与える影響を考慮していないが、実際はこれらの心理的ファクターがヒューマンエラーの発生にとって重要な意味を持

つと考えられる。シミュレーションシステムで採用された黑板制御モデルは戦略選択において非常に柔軟な拡張性を有し、これら考慮できなかったファクターを今後取り入れていくことは十分可能であろう。

さらに、シミュレーション結果から実際の運転員の作業成績を予測し、マン・マシンインタフェースの評価を行う手法の開発が必要である。特に、作業中の運転員が受ける精神的負荷の客観

的指標を得るための道具として、本システムを発展させることが考えられる。ここで精神的負荷と人間行動の信頼性との関係が明らかになっていれば、人間を含むシステムの総合的信頼性評価に本手法が役立つものと期待される。

#### 発表論文

Furuta, K. and Kondo, S.: Reasoning on Plant Process using TMS with Temporal Constraints, Artificial Intelligence in Manufacturing," Springer-Verlag (1989), pp. 313-321.