複雑システムの安全設計のための発想法

Safety thinking for complex system design based on STAMP/STPA

2016.12.6

会津大学　コンピュータ理工学部
兼本　茂

背景(2)　IoT時代の環境変化

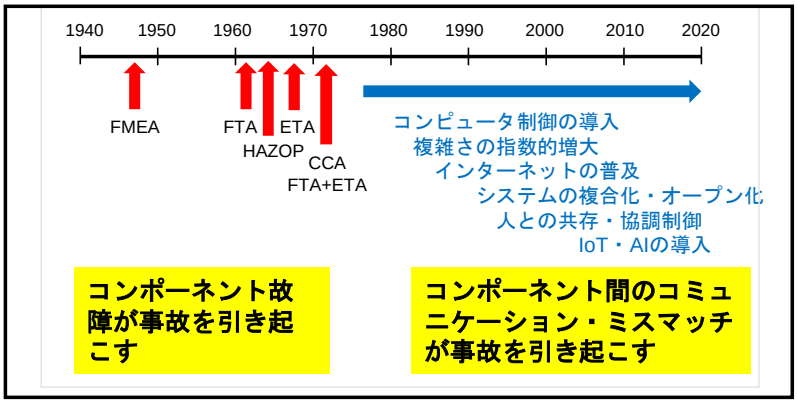
- **人、モノ、環境の相互作用がより緊密・大規模**
  - 事前に想定した環境での安全設計から、変化する環境下での安全設計へ （外部環境の変化、部分的なシステム更新、新たな利用者の登場など）
- **特別な訓練を受けていない人がシステムを使う**
  - 想定外の使い方をされることを想定した安全設計
- **AI技術など進展で機械の知能化が進み、機械と人間の制御指示のコンフリクトが日常化**
- **自動化の進展によって人間の能力が低下し、想定外での危機対応能力が低くなる。**
- **システムの開発体制のアジャイル化、オープン化し安全機能に影響する可能性**
  - 複数の独立企業による共同開発、仕様に記載されない暗黙知、不完全な調達仕様など
- **インターネットを介した悪意のある攻撃が日常化。セキュリティとセーフティの境界がなくなってくる。**
- **マスコミやネットを介した批判に応えるという社会への説明責任の増大。第三者による客観的な安全性の評価が重要になってくる。**

→人・機械・環境の複雑な相互作用を伴うシステムの安全設計はどうあるべきか？

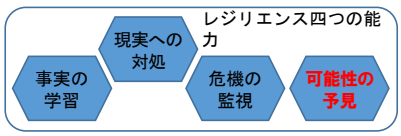
背景(3)

Sec-Seminar(2015.6.18) Lecture by Nancy Leveson

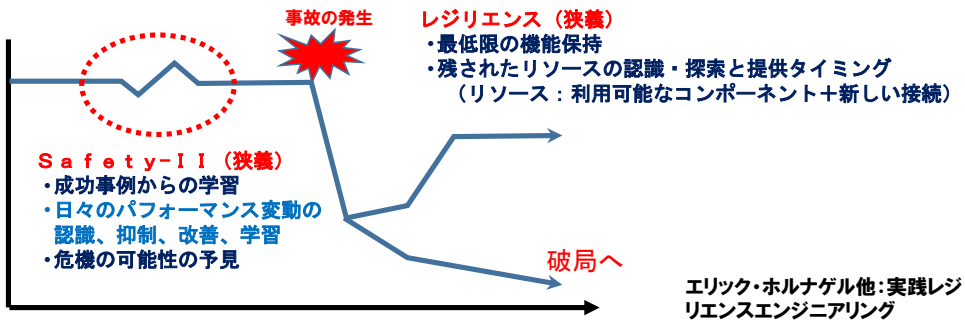
現状の安全分析ツールは、40-65年も昔に開発されたものであり、現代の新しい技術の入った複雑な工学システムの安全分析には限界がある



背景(4) Safety-II／レジリエンス



現実への対処、事実の学習、危機の監視までは、多くの組織で行われている。**一番難しいのは「可能性の予見」である**（ホルナゲル）



エリック・ホルナゲル他：実践レジリエンスエンジニアリング

背景(5)　現状の安全分析と将来の期待

- 定量的ハザード分析・現状
  - FTA・ETAと故障率によるPRA（確率論的リスク評価）
  - 安全解析コードによる予測（構造解析コードによる耐震性評価など）
    - 課題1：故障率の確からしさ、境界条件の確からしさ
    - 課題2：設計を熟知した専門家が必要
    - 課題3：人間・組織を含んだ複雑な挙動の分析が難しい
- 定性的ハザード分析・現状
  - FTA・ETA・FMEA・HAZOP・CCA
    - 課題1：定性的因果関係の表現の限界、定義したモデルでハザードが決まる
    - 課題2：人間・組織を含んだ複雑な挙動の分析が難しい
- 定性的ハザード分析・STAMP・STPAへの期待
  - 抽象化した制御構造図のもとで柔軟な発想ができうる→人間・組織、ソフトウェアを含んだ複雑システムのハザード分析が可能
  - より少ないドメイン知識での分析が可能、多様な発想を入れ込める→多様な目でハザードを予見できる（第三者によるレビュー）

目的

- 人・機械・環境の複雑な相互作用を伴うシステムの安全設計の一助となる具体策を探る
- STAMP/STPAの可能性を、従来法との比較の事例で評価

STAMPとは？

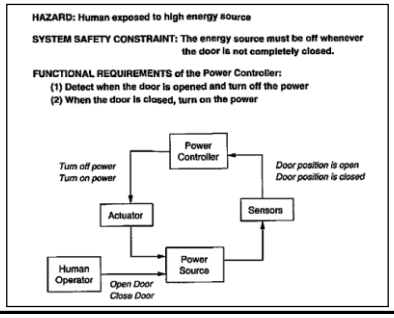
■STAMP （システム理論に基づく事故モデル）: Explanatory Model
- ●System-Theoretic Accident Model and Processes
- ●Three basic constructs: (1) Safety constraints, (2) hierarchical safety control structure, (3) process models

■STPA （STAMPによる安全解析法）
- ●System-Theoretic Process Analysis

■CAST （STAMPによる事故分析法）
- ●Causal Analysis based on STAMP

対象システムを抽象化、階層化した制御構造図と安全制約で、その安全制御構造を理解する

現在の複雑な組込みシステムでは、ソフトウェアや人間・組織を含むサブシステムやコンポーネントで構成されており、そこに不具合がなくとも，サブシステムやコンポーネントの相互作用によってハザードが発生する。従って、従来型のリスク分析手法では限界がある。

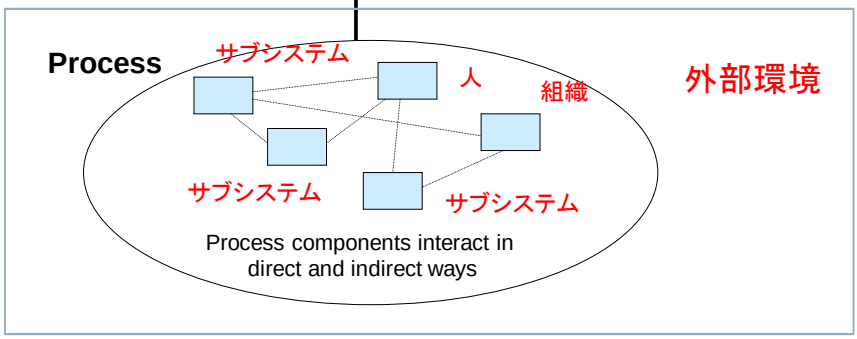


複雑システムの安全性とセキュリティ: 創発特性

Emergent properties
(arise from complex interactions)

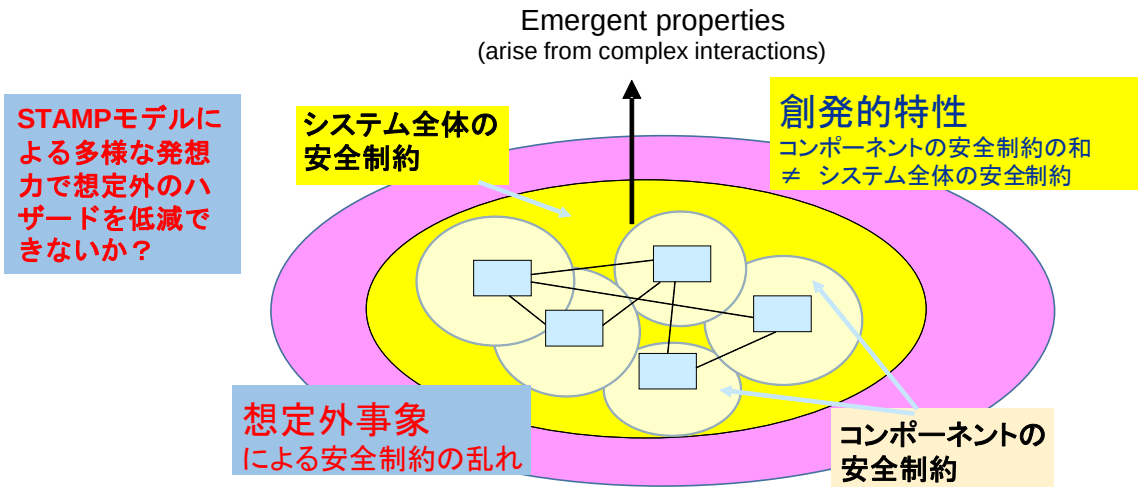
Emergent Property:
“The whole is greater than the sum of the parts”

複雑な相互作用に起因する創発的特性



Safety and security are emergent properties

複雑システムの安全性とセキュリティ：創発特性



Safety and security are emergent properties

STAMP/STPAの手順

Step-0：制御構造図とアクシデント・ハザード・安全制約の定義。抽象的な機能に着目してトップダウンで階層的な制御モデル、プロセスモデルを作成

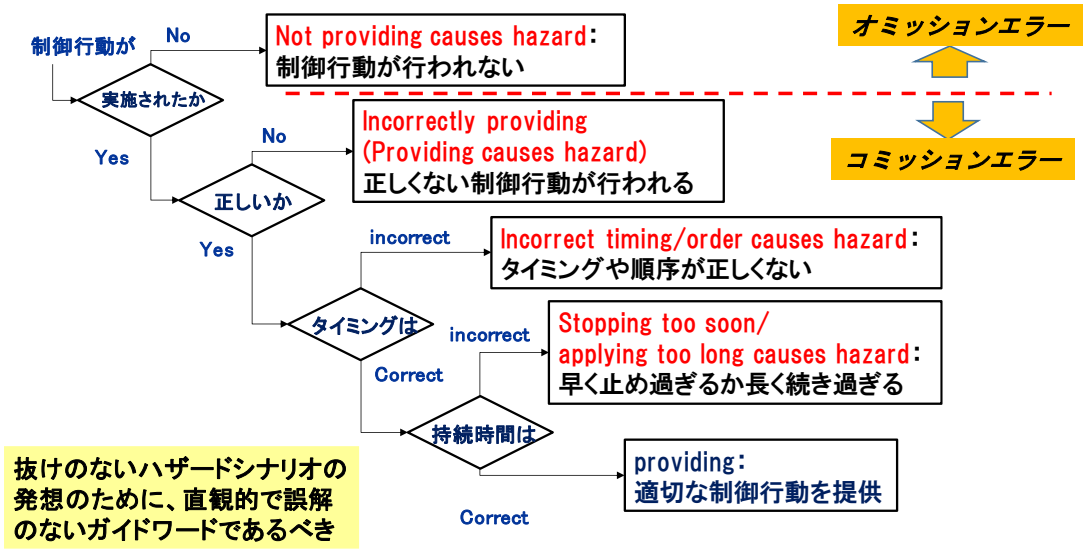
Step-1：四つのタイプの非安全制御行為（UCA：Unsafe Control Action)を抽出

(1)Not Provided、 (2)Incorrectly Provided、
(3)Provided Too Early, Too Late, or Out of Sequence、(4)Stopped Too Soon, Too Late

Step-2：Control Loopのガイドワードを用いて、UCAごとに、ハザード誘発要因（HCF:Hazard Causal Factor)を分析し、ハザードシナリオを導出

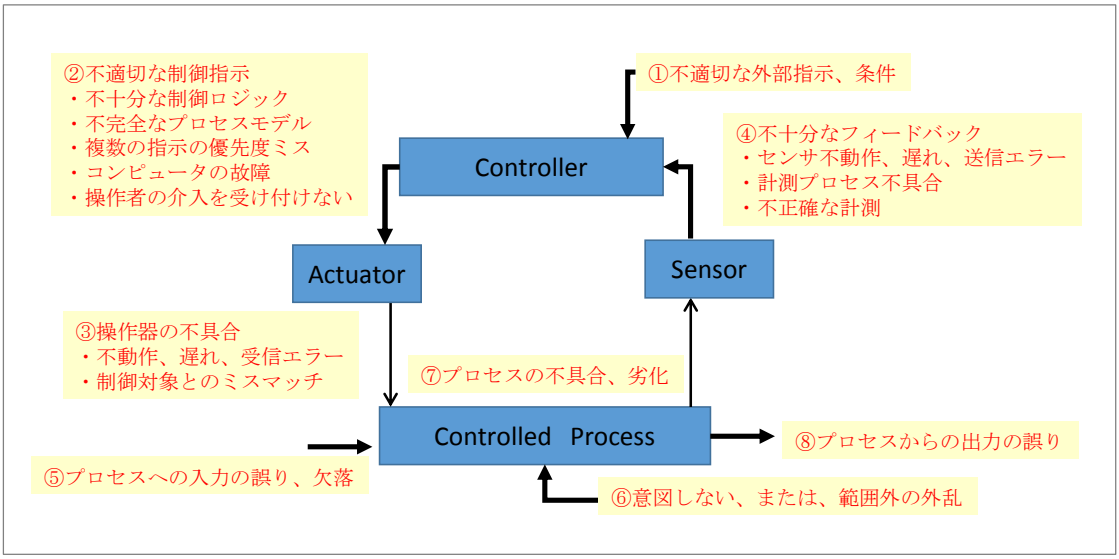
Step-3：安全制約の識別：HCFを制御・除去するためのコンポーネント安全制約を明示化する（階層的な分析）

Unsafe Control Actionsの分類 (網羅的・排他的)

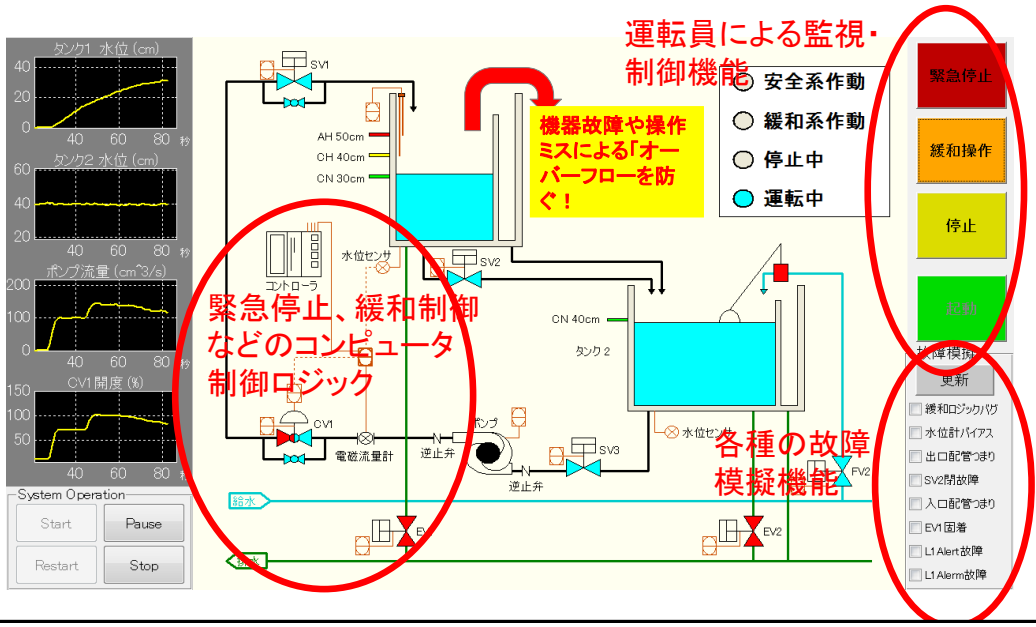


抜けのないハザードシナリオの発想のために、直観的で誤解のないガイドワードであるべき

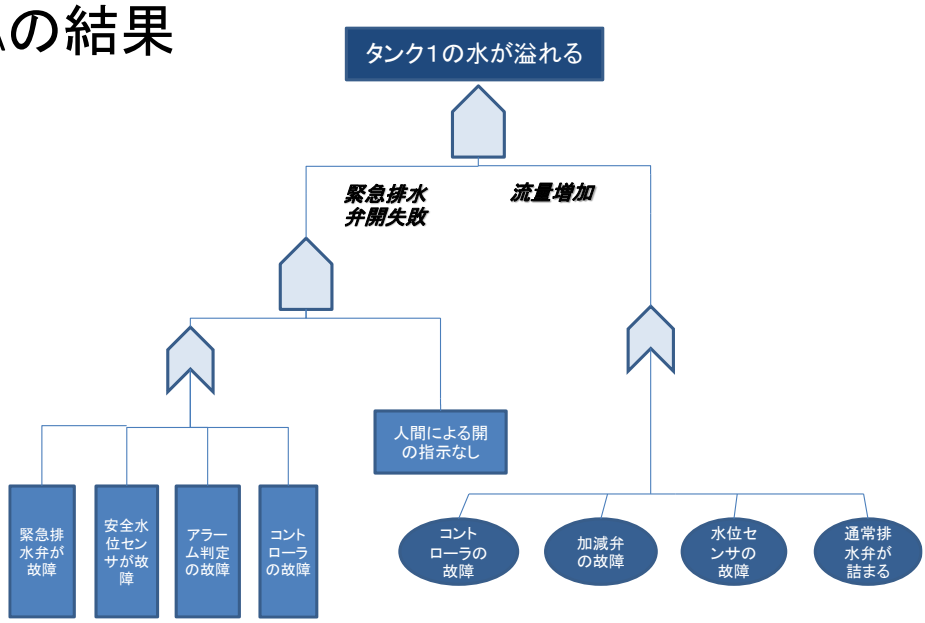
ハザード誘発要因(HCF)のガイドワード



検証用サンプルシステム（化学プラントシミュレータ）



FTAの結果



Step0: アクシデント、ハザード、安全制約の定義

「アクシデント」

タンクー1からの溢水（Water overflow）

ここでは、Tank-2からの溢水は、アクシデントやハザードとは考えない

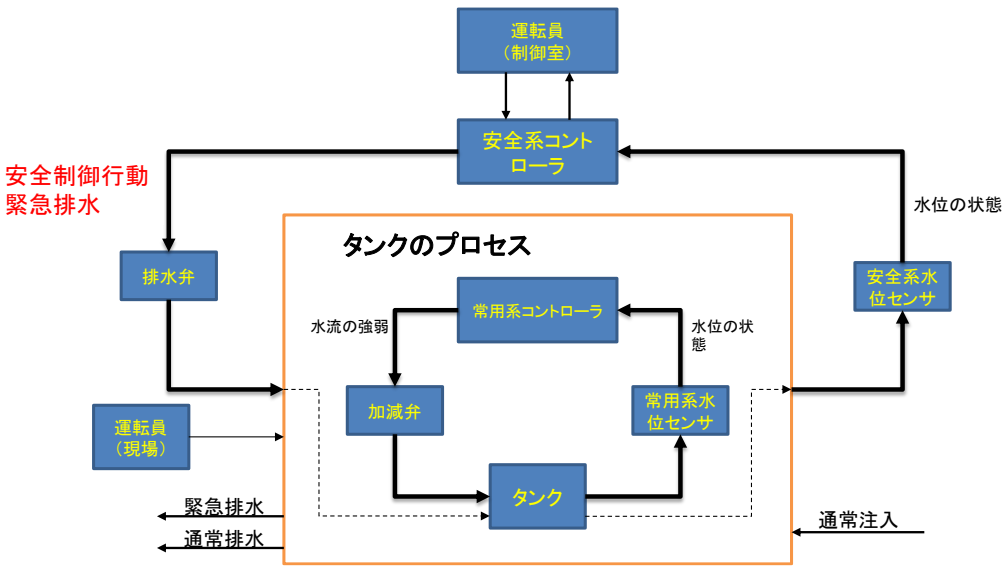
「ハザード」

水位が、アラームレベル（50cm)を超えた状態

「安全制約」

水位がアラームレベルを超えた状態にならないこと

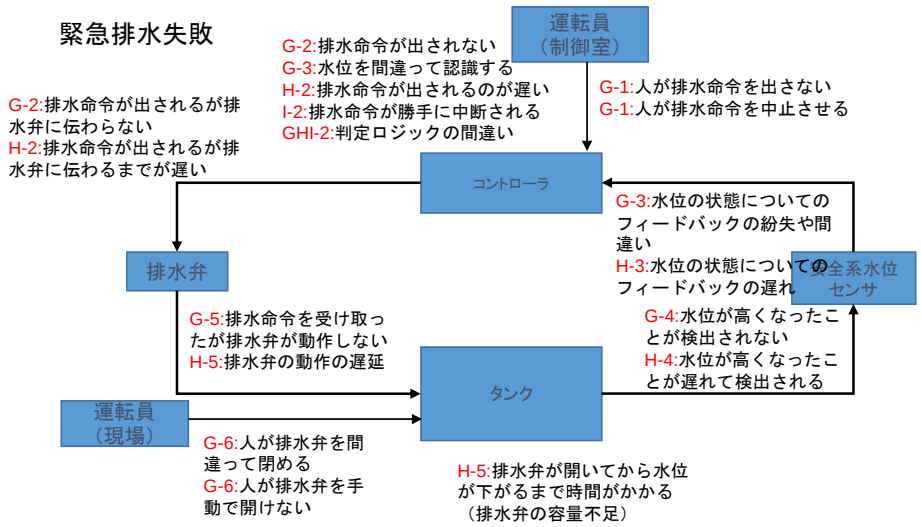
Step0: 制御構造図の作成



Step1: ハザードシナリオとUCAの抽出

制御行動	Not providing causes hazard	Providing causes hazard	Incorrect Timing / Order	Stopped Too Soon / Applied too long
緊急排水をする	UCA-G 緊急時に排水が行われない	ハザードなし （水位が下がる）	UCA-H 水位が高くなってから排水が始められるまでの時間が長い	UCA-I 水位が十分に下がり切っていないのに排水を中止する

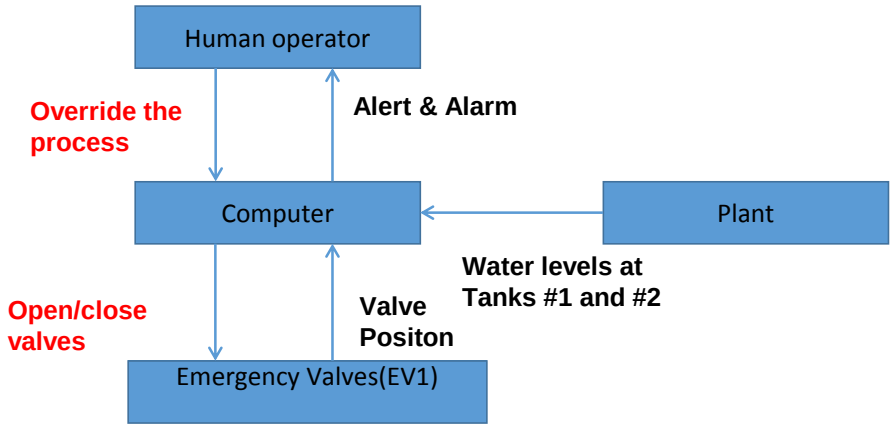
Step2:ハザード誘発要因（HCF）の抽出



STAMP/STPAによるHCFとFTAの比較

FTAの要因	STPAの要因
安全水位センサの故障	G-4:水位が高くなったことが検出されない H-4:水位が高くなったことが遅れて検出される
アラーム判定の故障	GHI-2:判定ロジックの間違い
コントローラの故障	G-2:排水命令が出されない G-3:水位を間違って認識する H-2:排水命令が出されるのが遅い I-2:排水命令が勝手に中断される(制御ロジックエラー)
人間による開の指示なし	G-1:人が排水命令を出さない G-1:人が排水命令を中止させる G-6:人が排水弁を間違って閉める G-6:人が排水弁を手動で開けない
緊急排水弁の故障	G-5:排水命令を受け取ったが排水弁が動作しない H-5:排水弁の動作の遅延 H-5:排水弁が開いてから水位が下がるまで時間がかかる(排水弁の容量不足)
その他(通信系)	G-2:排水命令が出されるが排水弁に伝わらない H-2:排水命令が出されるが排水弁に伝わるまでが遅い G-3:水位の状態についてのフィードバックの紛失や間違い H-3:水位の状態についてのフィードバックの遅れ

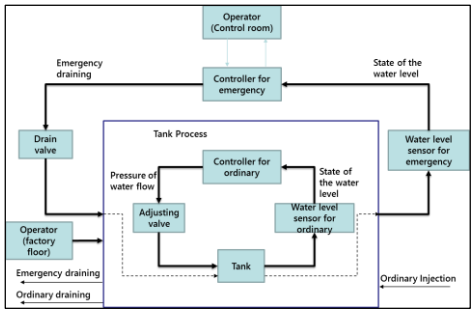
Step0:制御構造図(MIT分析)

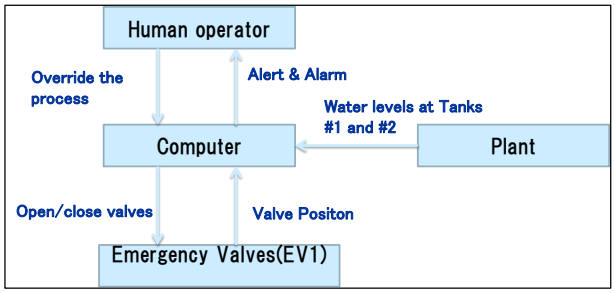


二つの制御構造図の比較

WG作成(2014)
（水位の制御という視点で書いている）
（運転員の指示は強調していない）

ナンシー先生が作成したもの
（排水弁の制御という視点で書いている）
（運転員の指示を強調している）
（抽象度が高い）





分析者の視点で表現できる柔軟さがある。同時に、抽象化（汎化）能力が問われる

(Step 1) Unsafe Control Action

	Not providing causes hazard	Providing causes hazard	Incorrect Timing / Order	Stopped Too Soon / Applied too long
Computer Action 給水弁(SV1)開とドレン弁(EV1)閉		水位アラームレベル状態で、コンピュータが給水弁開とドレン弁閉指示を出す（UCA1）	水位アラームレベル以下になる前に、コンピュータが給水弁開とドレン弁閉指示を出す（UCA1）	
Computer Action 給水弁(SV1)閉とドレン弁(EV1)開	水位アラームレベルに達した際、コンピュータが給水弁を閉しない、または、ドレン弁を開しない（UCA1）		水位アラームレベルに達した後X秒以内に、コンピュータが給水弁を閉しない、または、ドレン弁を開しない（UCA1）	コンピュータが給水弁が完全に閉まる前に指示をやめる、または、ドレン弁が完全に開く前に指示をやめる（UCA2）
Human Action 運転員の手動操作の介入（コンピュータ指示に優先）	コンピュータ操作が溢水を引き起こしそうな時に、手動操作で介入しない（UCA3）	コンピュータが意図通りに動いている時に、運転員が介入して溢水を引き起こす（UCA4）	コンピュータ操作が溢水を引き起こしそうな時に、手動操作の介入がX秒以上遅れる（UCA3）	誤解により運転員が緊急排水を中断する、または、給水を継続する（UCA5）

(Step 2) ハザード誘発要因(1)

[UCA2]
✓コンピュータが給水弁が完全に閉まる前に指示をやめる、または、ドレン弁が完全に開く前に指示をやめる

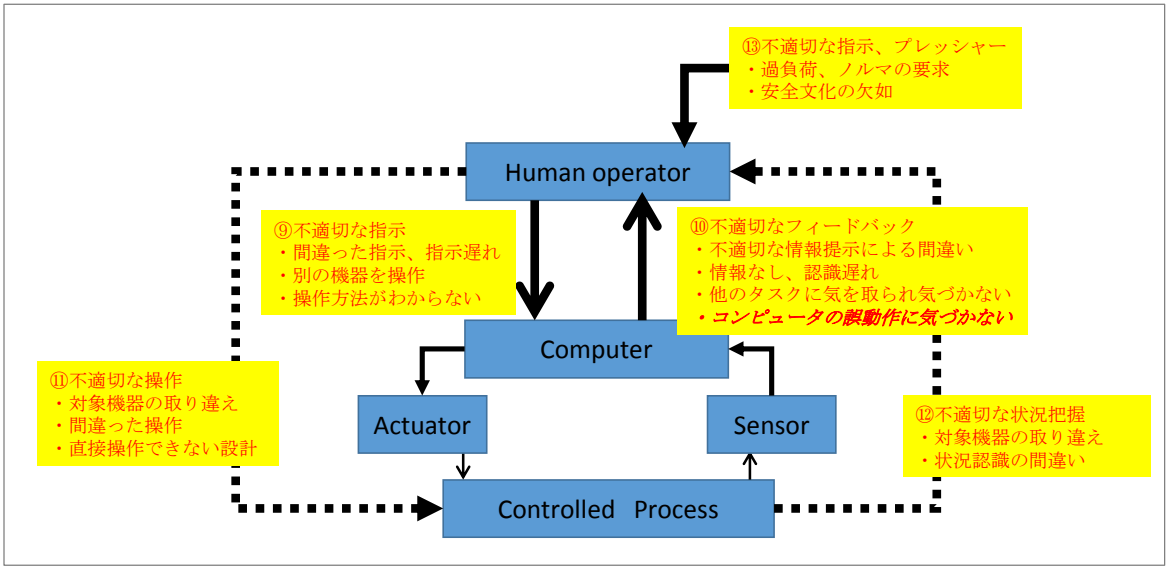
[Scenario 2-1]
✓コンピュータはバルブ開閉信号を出したので、バルブは指示どうりの状態にあると、コンピュータは思っている
  ✓バルブの位置情報のフィードバックがないか、間違っていたため、開き終わったと勘違いして、開閉指示をやめる

(Step 2)ハザード誘発要因(2)

[UCA4]
✓コンピュータが意図通りに動いている時に、運転員が介入して溢水を引き起こす

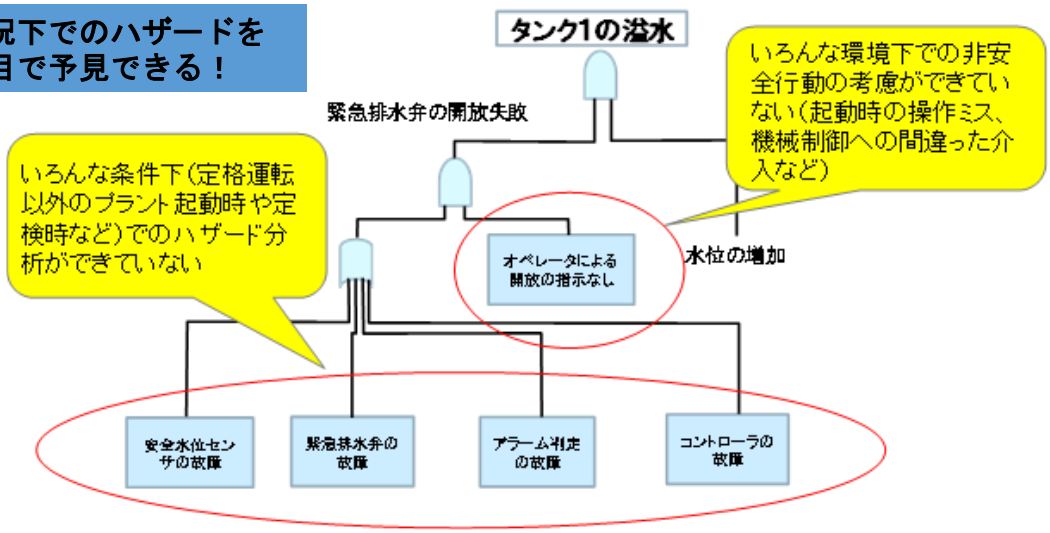
[Scenario 4-1]
✓コンピュータが適切に制御しているのに、運転員がこれを不十分と誤解する
  ➢運転員への、水位に関する間違った情報提示、または、不適切な操作ガイドの提示
  ➢コンピュータの動作に関する設計を理解していない
[Scenario 4-2]
✓不注意によるコンピュータへの間違った操作介入
  ➢操作介入機能の設計ミス（間違えやすい設計）
[Scenario 4-3]
✓運転員の意図的な操作介入（安全マージンを犠牲にした生産効率の向上など）
  ➢過負荷やノルマなどのプレッシャー
  ➢安全文化の不備

ハザード誘発要因(HCF)のガイドワード(人間と機械の干渉)

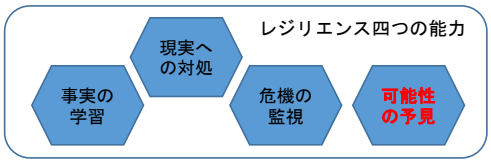


複雑な状況下でのハザード分析における従来の故障解析法の限界とSTAMP/STPAの可能性

複雑な状況下でのハザードを第三者の目で予見できる!



まとめ



現実への対処、事実の学習、危機の監視までは、多くの組織で行われている。一番難しいのは「可能性の予見」である（ホルナゲル）
→発想法としてのSTAMPの事例研究

■複雑システムの危機の可能性の予見
- ■想定外の事象を予見できる論理的な発想力
- ■異なる立場で見る発想の転換
- ■過去の事例の抽象化（汎化・正則化）による類推

■STAMPモデルによる事象の分析は、発想力、論理的な説明力で、従来のリスク分析法にない優れた視点を与えてくれる