

## デジタル危険予知訓練システム開発のための回答文の正否判定システムの開発 Development of a system to determine whether answer is right or not to realize a digital hazard prediction training system

箕輪弘嗣<sup>†</sup> 藤本宏美<sup>‡</sup> 竹内孔一<sup>§</sup>  
Hirotsugu Minowa<sup>†</sup> Hiromi Fujimoto<sup>‡</sup> Koichi Takeuchi<sup>§</sup>

### 1. 緒言

化学プラントでは、石油類や化学物質が多量に扱われており、引火性、爆発性など危険性を持つ物質も多い。そのため、化学プラントで、事故が発生した場合、企業の責任問題、重大な環境汚染の発生や、近隣の住民に対しても被害が及ぶ恐れがある。また、作業員に対しては死傷者が出てしまう恐れもある。このような惨事を起こさないためにも、企業にとって事故防止、保安力の向上は重要であり、作業員は危険に対して予測し対応する危険予知能力が必要不可欠である。現場では危険予知能力向上のため、実地訓練(OJT)やモックアッププラントを用いて、危険に対する知識や経験を鍛える取り組みがなされている。

その中でも、4R法による訓練が幅広く行われている。4R法は、教示者の指示の下で、訓練者がイラスト中に潜む危険を予測し、対策を提案する事を求められる。この自発的な思考過程を通して、危険の予知力や対応力を育成する簡易な危険予知訓練方法である。

4R訓練法の問題は、教示者、他の作業員が要る必要があり、作業員一人では学習できず、訓練の機会を大きく制約されてしまう点である。また、現場の負担から新規コンテンツを作成するのは難しい点も問題である。変化が乏しくなるために、訓練効果を低下させてしまう危惧がある[1-2]。以上、4R法は危険予知能力の向上のため、危険の抽出および対応策を考える思考力育成に有用であるが、環境が整っていないと実施できない問題がある。

この問題を解決するため本研究では、デジタル4R訓練システムを提案する。4R訓練における教示者を電算機(PC)に代替わりさせる事で、既存の4R法における教示者が必要である問題を解決し、学習における時間的・空間的制約を緩和できると期待する。デジタル4R訓練システム実現のためには、訓練者の回答の正否を判定するサブシステムが必要である。そこで、本論文では、SVMといった教師付き機械学習を用いて統計的データに基づき、回答の正否を判定するサブシステムの開発、評価を実施した結果について報告する。

#### 1.1 4R法の訓練の必要性

近年、拡張現実感(Virtual Reality)による訓練(以下、VR訓練)が提案されており、より実環境に近い環境で訓練できる利点がある。しかし、工業プラント下における非スキルベースなVR訓練に関して言えば、広大な環境を動きまわる行動や作業が大半を締め、訓練に時間を要してしまい、沢山の訓練をするのは実質難しい。そのため、4R訓

練法といった短時間で、自発的に回答しないと訓練を終了できない学習法と併用する事が、幅広く危険を察知・対応できる危険予知の知識基盤を訓練者の中に作る事に繋がるだろう。2つの知識は別ものであるため、知識や能力の住み分けができるものとする。

#### 1.2 4R訓練法の流れ

4R訓練は、表1に示す4ラウンド(段階)で訓練を実施する。KYT(Kiken Yochi Training)シート内に描かれた、潜在的な危険が在る内容の1枚のイラスト図から、訓練者が予測される危険要因を抽出(1R, 2R)し、改善策を提案(3R, 4R)する訓練法である。KYTシート内のイラストには、作業員の専門的な作業の状況や、日常の行動や作業の情景の一コマが描かれている。KYTシート内のイラスト例を図1に示す。

表1 4R法の各ラウンドの試行内容

ラウンド	既存の4R法
1R	みんなの話し合いで、1枚のKYTシートに描かれたイラストの状況に潜む危険を発見し、危険要因と、要因が引き起こす現象を想定し発表しあう。
2R	第1Rに出た意見に対して、監督者と一緒に重要だと思われる項目に○印をつける。さらに、絞り込んで◎印とアンダーラインを付け、“危険ポイント”とする。
3R	“危険のポイント”を解決するにはどうしたらよいかを考え、具体的な対策をたてる。
4R	第3Rに出た意見を統一し、作業員全員でどのように行動して危険、事故を防ぐかの目標を決定する。
確認	目標を指さし、声出し確認する。



図1 KYシートのイラスト例

#### 1.3 提案システムと従来訓練のやりとり

目標とするシステムと従来訓練法におけるやりとりの比

<sup>†</sup>岡山商科大学, Okayama Shoka University

<sup>‡</sup>岡山県立大学(非常勤講師), Okayama Prefecture University

<sup>§</sup>岡山大学, Okayama University

較を図2に示す。PCが教示者の役割を担い、訓練者はPCを通して学習する。このフレームワークのメリットとしては、いつでもどこでも4R法を学ぶことができる点にある。



図2 従来訓練法と提案訓練法のやりとりの比較

## 2. 研究内容

デジタル4R訓練システム実現のためには、訓練者の回答の正否を判定するサブシステムが必要である。本論文では、回答の正否判定サブシステムの開発において、教師付き機械学習にはSVMを用い、2.1節に従い作成した訓練データより生成した正否判定サブシステムの性能を評価し報告する。

### 2.1 訓練データ作成法

正否判定システムの訓練データは、次の手順で作成する。

#### Step 1. 回答文の形態素の抽出

事前に収集した複数の回答文を形態素解析し、形態素を抽出する。

#### Step 2. 出現頻度による除外

Step1で抽出した単語の内、出現頻度の高低が大きいものは除外する。

#### Step 3. Bag of Word(BoW)ベクトルの生成

Step2で抽出されたすべての単語が含まれるBoWベクトルを作成する。

#### Step 4. 正例/負例の情報追加

Step3で作成した各BoWベクトルの各回答に正例、負例であるかの情報を追加する。

### 2.2 回答データの取得方法

交通に関するKYTシート集[3]からNo. 1, 4, 5, 10, 12, 13, 14の7枚のKYTシートを選び、1グループ2人または3人に分かれた高校生21グループ(計43名)に回答してもらった。この時、1シートに対して3グループを割り当て、各グループは約10個ずつ、KYTシートのイラストに対する回答文を作成してもらった。この7枚のシートの選定理由は、本KYTシートの中でも平易な内容で、運転免許を持たない高校生でも回答は困難でないだろうと判断したためである。高校生が抽出した潜在的危険を説明した回答文と筆者らが考えた模範解答を含む33-36個の文章を、筆者らが正例、負例に振り分けた。このデータから2.1節で述べた手順により訓練データを作成した。この訓練データにおける正例、負例のレコード数はほぼ同じであった。

### 2.3 実装

この評価実験にはLibSVM Ver. 3.20ライブラリを使用した。そのパラメータはカーネルの種類はlinearであり、その他のパラメータについてはLibSVMのデフォルトを使用した。トレーニングデータを作る単語の除去条件は3未満のドキュメントまたは、0.2以下の文章の場合は、トータルコーパスのサイズではなく絶対数の割合でデータを除

外した。また、形態素を抽出するためにMeCab[4] Ver. 0.966やCabocha[5] Ver. 0.63を使用した。

## 3. 結果

作成した訓練データからSVMで正否判定システムを作成した。分析のまとめを表2に示す。表2は各KYTシートにおける平均正答率と標準偏差を示す。全てのシートの平均正答率と標準偏差は最高低値を除外したデータから求めた平均の正答率と標準偏差を示している。その結果、本研究の平均精度は63.0±22.9%であった。

シート13の精度は非常に悪くなっている。その理由としては、回答内容が分散してしまう内容であると考えられる。シート13の出題は「あなたは前方の自転車を追い越そうとしています」であり、自転車との位置関係について考える事を促しているものである。回答には、前方の車ではなく、歩道を歩いている歩行者に危険がある、とした回答があり、回答内容の方向性が分散してしまった結果、判定データの不足により正否判断サブシステムが判断を誤ってしまった結果と考えられる。実用化のためには、正答率が低いシートは除外する事も可能であろうが、訓練データ数を増やす事で回答率の向上は期待できると考える。

表2 平均と標準偏差

シート No	平均正答率 [%]	標準偏差 [%]
1	64.7	22.8
4	59.1	23.3
5	60.7	23.2
10	84.8	17.6
12	69.6	22.5
13	27.0	17.2
14	60.8	22.9
平均正答率	63.0	22.9

## 4. 結言

デジタル4R訓練システム開発のため、本稿では、機械学習としてSVMを用いた訓練者の回答の正否判定システム開発の取り組みについて報告した。実験結果は平均精度が63.0±22.9%となった。本正答率は訓練システムに用いるのは性能不足であると考え、今後その精度向上を目指す。

### 参考文献

- [1] 中央労働災害防止協会, “危険予知活動トレーナーのためのゼロ災運動Q&A”, 中央労働災害防止協会 (2003).
- [2] 中央労働災害防止協会, “ゼロ災実践シリーズ危険予知訓練”, 中央労働災害防止協会(2011).
- [3] 中央労働災害防止協会, “危険予知訓練(KYT)イラストシート集交通編”, 中央労働災害防止協会 (2014).
- [4] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto, “Applying Conditional Random Fields to Japanese Morphological Analysis”, Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004), pp.230-237 (2004).
- [5] 工藤 拓, 松本 裕治, “チャンキングの段階適用による日本語係り受け解析”, 情報処理学会論文誌, 43(6), pp.1834-1842, (2002).