

令和 2 年 7 月 1 日現在

機関番号：17104

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K00345

研究課題名(和文) エージェントとトレーナの双方向強化学習による介入最適化手法と実データでの検証

研究課題名(英文) Interactive Optimization Method based on Bidirectional Reinforcement Learning of Agent and Trainer and Verification with Practical Data

研究代表者

堀尾 恵一 (HORIO, KEIICHI)

九州工業大学・大学院生命体工学研究科・教授

研究者番号：70363413

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：本研究では、対象となるエージェントが強化学習に基づく行動を獲得するということを仮定し、エージェントの特性を分類し、それに応じた適切な報酬の与え方を設計する。まず、エージェントの特性を分類することのために、スポーツ選手のメンタル状態を計測するアプリを開発し、収集したデータに基づいて選手のメンタル状態とパフォーマンスの関係性を解析した。また、選手個々に対する適切な報酬の与え方として、アプリを介して選手にフィードバックするコメントに関し、ポジティブパターン、率直パターンなどを定義し、選手により適切なフィードバックコメントの傾向が異なることを確認した。

研究成果の学術的意義や社会的意義

ヒトが社会で生活を営むにおいて、他者からの介入は非常に重要な意味を持つ。適切な介入はヒトのメンタル状態を改善し、結果としてその行動を変えることができる。本研究は、多数のスポーツ選手を例として、選手のメンタル状態とパフォーマンスの関係性を解析することで、各選手が良いパフォーマンスを発揮するメンタル状態を推定し、また、選手への介入として状態を解析結果を様々なパターンで提示することで、各選手にとって適切なコメントパターンを推定することを実現した。これらは、スポーツ現場のみならず、教育を始めとする様々なシーンに適用可能である。

研究成果の概要(英文)：In this research, we assume that the target agent acquires the behavior based on reinforcement learning, classify the characteristics of the agent, and design an appropriate reward method accordingly. First, in order to classify the characteristics of agents, we developed an application that measures the mental status of athletes and analyzed the relationship between the mental status of athletes and performance based on the collected data. In addition, as an appropriate way to give rewards to each player, we defined positive patterns, candid patterns, etc. regarding comments that are fed back to the players via the application, and confirmed that the tendency of appropriate feedback comments differs depending on the players.

研究分野：データ解析

キーワード：メンタル状態 フィードバックコメント 強化学習 報酬設計

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

学校やスポーツの現場では、教師やコーチ、トレーナが生徒や選手に適宜指導を行い、その成長を促すことが期待される。ここでは、教師やコーチの指導を介入と呼ぶこととする。介入は、その内容もさることながらタイミングもまた重要である。さらに、生徒や選手は様々な個性を有しており、同じように介入しても同様の成長が期待できるわけではない。教師やコーチは、介入方法に関して、予め教育を受けているが、多様な個性に対し、それぞれに適切な介入を実現することは容易ではなく、教師やコーチも多くの試行錯誤を経て、適切な介入を獲得することになる。ここでは、教師やコーチをトレーナと呼び、生徒や選手をエージェントと呼ぶ。エージェントの行動学習の枠組みとして強化学習が広く知られ、ロボティクスなど様々な分野へ応用されて成果を挙げている。強化学習では、目的を達成した際に報酬が与えられ、エージェントは将来得られる報酬に基づき行動を決定する。一方、理想的な行動からどのような報酬を与えることが良いかという報酬設計に関する研究も近年注目されており、これは逆強化学習と呼ばれている。しかし、逆強化学習では、理想的な行動を行うエージェントがいることを仮定しているが、本研究の目的は、理想的なエージェントは仮定せず、トレーナもまた試行錯誤を繰り返すことで報酬設計を実現するという点でこれまでにない取り組みである。図1に本研究の概要を示す。種々の特性を有するエージェント集団に対し、各エージェントの学習過程の行動をモデル化し、それらを分類する。分類された各エージェントクラスに対し、トレーナは試行錯誤を繰り返しながら最適な報酬を設計するものである。

代表研究者の堀尾は、分担研究者である磯貝および連携研究者である古川とスポーツチームにおける選手間あるいは選手とコーチ間のコミュニケーションの解析、また、心理状態がパフォーマンスに及ぼす影響の解析に既に取り組んでいる。この取り組みの中で、選手とコーチ間のコミュニケーションがチームワークやチームのパフォーマンスに影響を及ぼすことが示唆されている。この知見をより詳細に調査するためには、コーチの介入方法を詳しく定量的に評価する必要があると考えている。一方で、堀尾と磯貝は、一般社団法人行動評価システム研究所を立ち上げ、その事業の一環として、メンタルトレーナ向けのアプリを平成28年11月にリリースする。この事業は、トレーナが選手に対して行う指導を補佐するもので、選手は日々のフィジカルコンディショニングや心理状態などを朝晩の2回アプリを介して入力し、それらの情報がトレーナに届けられる。トレーナは必要に応じて、選手に指示を出したり、注意を促したりする。国内のトップレベルで活動する選手を担当するトレーナへの普及活動を開始しており、これらのデータを活用して、トレーナの選手に対する介入の妥当性や問題点を客観的に評価・検証する必要性を感じている。そのためには、綿密なシミュレーションによる介入モデルを構築する必要があると考えるに至り、本研究を計画するに至った。

### 2. 研究の目的

本研究では、以下の各項目について明らかにする。

(1) シミュレーションによるエージェントとトレーナの双方向型強化学習の確立

(2) トレーナアプリのデータ解析に基づくトレーナの介入を評価・検証

(1)を実現するために

(a) 学習過程を観察することによるエージェントの内部パラメータ推定および分類

(b) 双方向型強化学習による各エージェントクラスに対する報酬の最適化

(c) 双方向型強化学習による行動獲得時の学習安定性の評価

について取り組む。この過程で、エージェントの分類に必要な学習過程を表現する変数の選定、双方向型強化学習におけるエージェントクラス間の協調による学習効率化、エージェントとトレーナが互いに強化学習により行動を獲得していく過程の安定性を決定する指標、について明らかにしていく。

収集したデータから、トレーナの介入が選手の行動にいかに関与しているかを調査することで、選手の分類を行う。各選手クラスに対し、トレーナがどのような介入を行っているのか、その相違点を解析することで、双方向型強化学習の確立に資する知見を得る。

### 3. 研究の方法

本研究では、シミュレーションによる双方向強化学習の基礎的な理論構築を行い、スポーツにおける選手に対するメンタルトレーナの介入を解析することによるその有用性を検証する。シミュレーションにおいては、種々の特性を持つエージェント集団をその学習課程の行動をもとに内部パラメータを推定し、エージェントの分類を行う。さらに、分類された各エージェントに関し、効果的な学習を実現する報酬の与え方を実現する。これにより双方向強化学習の枠組みの確立を行い、その安定性の解析も併せて行う。一方で、メンタルトレーナと選手のコミュニケーションを解析し、トレーナの介入の妥当性を検証する。

【平成29年度】

本研究は、様々なタイプのエージェントが存在する状況下で双方向強化学習を実現可能であることを計算機上のシミュレーションにより示すことと、メンタルトレーナのアプリで得られたデータからトレーナの介入が強化学習の枠組みで説明できることを示すこと、が目的となるので、シミュレーションと実データの解析に分けて計画を立てる。平成29年度は具体的に以下に取り組む。各項目の分担を図2、各項目のスケジュールを図3に示す。

## シミュレーションに基づく双方向強化学習のためのエージェントの分類

(A) シミュレーション対象とする問題設定の決定 (担当: 堀尾)

当面は、強化学習の典型的なタイプロblemと用いられる迷路問題を取り扱う。強化学習に基づき行動を獲得するエージェントには、報酬の減衰率やグリーディ戦略に関するパラメータ、行動決定におけるどの程度の将来を考慮して行動を決定するかなど、種々のパラメータが含まれる。様々なパラメータのエージェントを用意し、それらの学習課程の差異を観測する。

(B) エージェントの学習課程から抽出する特徴量の選定 (担当: 堀尾)

エージェントの学習課程は、そのエージェントが従うパラメータによって異なるが、できる限り学習の早期に適切な行動の特徴を抽出することができれば、実応用において非常に有用である。本研究では、学習の初期において、エージェント間の行動の差異を適切に表現できる特徴を選定する。

(C) 特徴量をもとにエージェントの特性を分類する手法の確立 (担当: 堀尾, 古川)

上記(B)で選定した特徴量に基づきエージェントの分類を行う。分類手法としては、詳細は-(A)で述べるが、クラス間の類似性を保持することで、データを学習に有効に活用することが可能となることや、実問題への適用ではトレーナや教師が理解しやすい形式で表現することが望ましいと考えられるので、堀尾や古川がこれまで数多く取り組んできた自己組織化マップの拡張手法を適用することを考えている。

### トレーナアプリによるデータ収集および基礎的解析

(A) トレーナの募集・選定およびデータ収集 (担当: 磯貝, 堀尾)

アプリは既に改正済みであり、11月に開催される日本スポーツ心理学会の展示ブースにおいてアプリの紹介を行い、賛同者をつくる予定である。当面は、トップクラスの選手を担当するトレーナにアプリを使用してもらい、選手の日々のコンディションデータおよびトレーナの介入に関するデータを収集する。これは、トップクラスの選手は、コンディション管理に熱心であるのみならず、トレーナも訓練されているので、質の高い、つまり安定したデータが得られることを期待するものである。

(B) 収集した小規模データを用いた基礎的な解析 (担当: 古川, 堀尾)

平成29年度後半からは、それまでに収集されたデータをもとに、選手の変化を表現する特徴量の選定および分類に取り掛かる。データが小規模であるので、心理学的知見も参考にしながら、詳細な解析を行う。

【平成30年度以降】

平成29年度に引き続き、シミュレーションおよび実データの解析を行う。具体的には、以下に取り組む。

### シミュレーションに基づく双方向強化学習のアルゴリズム構築および安定性解析

(A) 各エージェントクラスに対する報酬の最適化手法の確立 (担当: 堀尾, 古川)

上記-(C)で分類された各エージェントクラスに対し、報酬を設定しエージェントの学習行動の観測する。報酬の与え方の善し悪しによりエージェントの学習効率が変化したが、それに対して試行錯誤を行い、結果として適切な報酬の与え方を獲得する。これは、教師やトレーナもまた介入方法について強化学習を通じて獲得していることを仮定している。

(B) 双方向強化学習の安定性解析 (担当: 堀尾)

エージェントとトレーナが相互に強化学習で行動を獲得していく過程において、フィードバックが存在することによる学習の不安定性が懸念される。不安定性を回避するために、報酬設計に制約を設ける必要があると考えている。学習の効率化と安定性の向上という相反する目的を調整する制約の設計を行う。

### トレーナアプリの大規模データの解析および介入妥当性の評価

(A) 平成29年度までに収集した大規模なデータを用いたデータ解析 (担当: 堀尾, 古川)

平成29年度までは、主にトップクラスの選手を担当するトレーナからデータを収集することを行っていき、平成30年度からは、より幅広いレベルのトレーナ、選手からデータを収集する。これにより多様なデータを取りそろえ、様々な状況において、様々な介入に関するデータが得られることを期待するものである。

(B) 解析結果の検証によるトレーナ介入の妥当性評価 (担当: 磯貝, 堀尾)

上記解析結果を堀尾と磯貝で評価する。磯貝はスポーツ心理の専門家であり、メンタルトレーナの講師も行っており、各トレーナがいつどのような介入を行ったかを客観的に評価可能である。解析結果と磯貝の知見を総合的に判断し、トレーナの介入の妥当性を評価する。また、双方向強化学習のシミュレーション結果との比較検討を行い、双方向強化学習の妥当性もまた検証する。

## 4. 研究成果

平成28年度にリリースしたアプリでは、毎朝夜にフィジカル状態およびメンタル状態に関する約20項目について、5段階評価や100点満点のスコアとして入力可能である。また、夜入力では、当日の練習もしくは試合におけるパフォーマンスに関して主観的な評価が入力される。本研究では、(1)各選手のメンタル状態とパフォーマンスの関係性の解析、(2)選手全員を直感的に把握するための状態解析、および(3)フィードバックコメントがメンタル状態とパフォーマンスに及ぼす影響の解析を行った。

(1) 各選手のメンタル状態とパフォーマンスの関係性解析

メントレアプリの入力項目のうち、12項目とパフォーマンスに関する2項目を用いて、各選手のメンタル状態（一部のフィジカル状態も含む）とパフォーマンスとの関係性を解析した。解析では、K大学野球部20名とS大学剣道部27名を対象とし、各チーム内でのバラツキやチーム（競技）間での特徴の相違などを考察した。

解析には、自己組織化マップ（SOM）を拡張した手法を用い各選手の最も良いパフォーマンスを発揮する際のメンタル状態を抽出した。

ここで抽出したメンタル状態は、スポーツ心理学の分野でパフォーマンスと関連性が示されているもので、ここではPOES（Profile of Mental State）と呼ぶ。POESは、「緊張」、「抑鬱」、「怒り」、「活気」、「疲労」、「混乱」の6つの項目であり、一般的には、活気のみが高く他は低い方が良いとされる。図1および図2に、K大学野球部20名とS大学剣道部27名の各選手の最もパフォーマンスが良い際のPOESを示している。各図で左は全選手を折れ線で表示し、右は箱ひげ図で表示している。図から、K大学野球部では、ほとんどの選手が「活気」のみ高い際にパフォーマンスが良いことが分かる。この状態は、POESのグラフでの形状から冰山型と呼ばれている。「疲労」に関しては選手により傾向が大きくことなりバラツキが大きいことが分かる。一方で、S大学剣道部では、多くの項目でバラツキが大きいが、全体としては冰山型を示している。野球と剣道という競技の違いにより「疲労」に関する傾向が大きく異なっていると考えられる。これまで冰山型が良いとされてきたメンタル状態について、個人差も大きいことを示すことができた。このような個人差を選手本人また指導者が適切に把握することは、選手自身のコンディション調整や指導者の介入に非常に有用である。

(2) 選手全員を直感的に把握するための状態解析

監督、コーチらが所属する全選手のコンディションを容易に確認でき、また、過去の練習メニューの考察を容易にする可視化手法を検討した。使用したデータは、K大学野球部20名のメントレで収集した17項目データであり、61日間のデータである。このデータは、3つのモード（選手×日付×コンディション項目）のテンソルであり、このデータに対し、テンソル分解型自己組織化マップを適用し、各モードに関してマップを作製した。その結果を図3に示す。図3の各マップは、配置で類似関係を表現し、色で周辺との距離を表しており、図3の各図から、選手間や日付間、またコンディション項目間の類似性が分かる。また、この表示法は、あるモードの特定の箇所を選択すると、その箇所に対応する他のモードの状況を確認することができる。例えば、図4は、それぞれ選手マップ上の特定の選手を指定した際に、その選手が(a)どの日付にパフォーマンスが高かったのか、(b)どのコンディション項目が値が高いのか、を赤・青の色で示している。この図を見ると、全ての選手において、日付マップの下方に対応する日にパフォーマンスが良いことが分かる。これらの情報に基づき、指導者は練習メニューなどの振り返りを行うことで、改善を図ることが可能となる。同様に、図5は、日付を指定した際に、その日に(a)ど

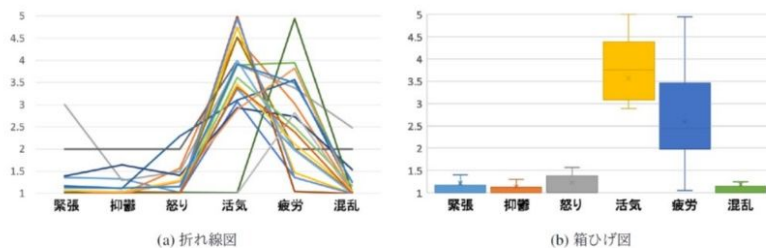


図1: K大学野球部20名の各選手の最良のPOES

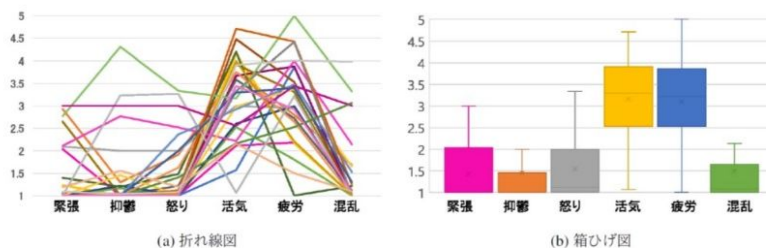


図2: S大学剣道部27名の各選手の最良のPOES

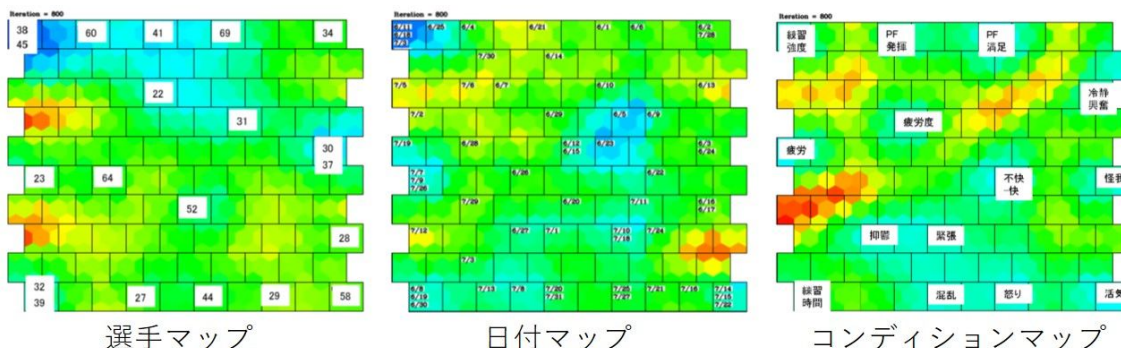


図3: K大学野球部の選手-日付-コンディション同時解析の結果

の選手のパフォーマンスが高かったのか、(b)どのコンディション項目が高かったのかを赤・青



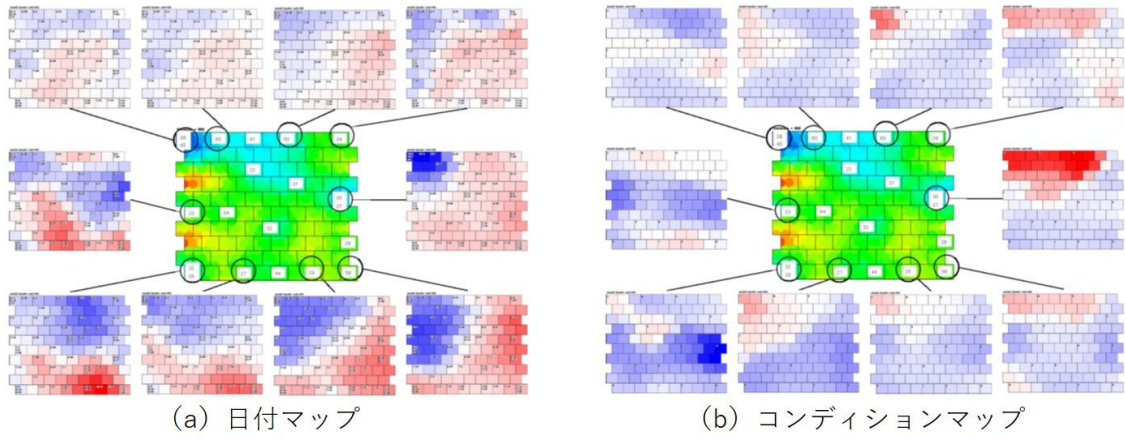


図4: 選手を指定した際の (a) 日付マップと (b) コンディションマップ

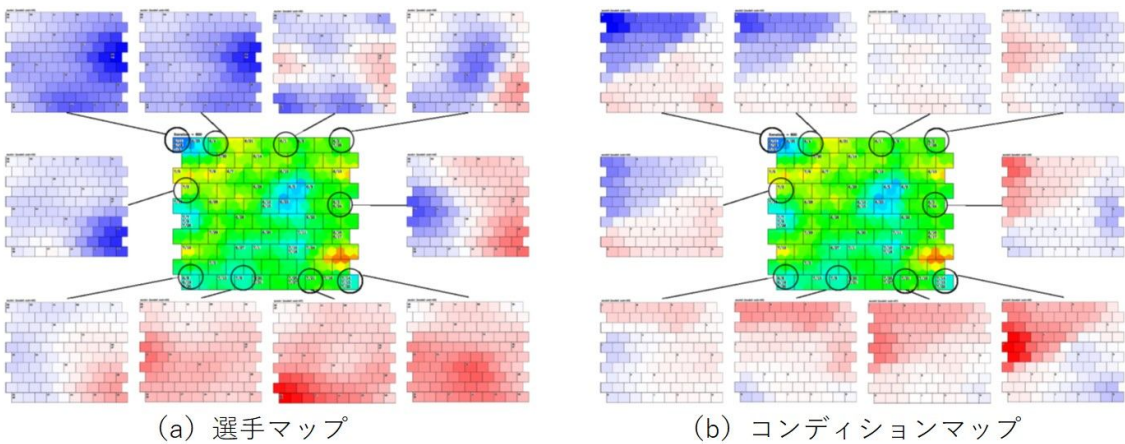


図5: 日付を指定した際の (a) 選手マップと (b) コンディションマップ

の色で示している。この結果から、日付マップの左上に位置する日には、全ての選手のパフォーマンスが低く、これらの日には何かしらの問題が潜んでいることが分かる。

### (3) フィードバックコメントがメンタル状態とパフォーマンスに及ぼす影響の解析

ここで用いるデータは、S大学剣道部においてアプリを介して収集したものである。選手がアプリに状態を入力した直度に、フィードバックのコメントを表示している。その効果を検討するにあたり、アプリケーションへの入力項目の1つであり、選手がその日にどの程度のパフォーマンスを発揮することができたのかを主観評価した「パフォーマンス発揮度 (PF 発揮度)」を用いた。また、フィードバックコメントの内容に関しては、我々が独自に考案した、ポジティブパターン、率直パターン、平均との比較パターンの3パターンのフィードバックコメントを選手に適用した。調査期間は、2018年7月23日から10月14日までの84日間と、2019年7月23日から10月14日までの84日間の計168日間である。各年の84日間を4週間(28日間)ずつの3タームとして、1タームごとに各選手へのフィードバックコメントのパターンを変更しながら、PF 発揮度やその他の項目にどのような影響を与えるか検討する。集団解析として、1つのチームを対象に、そのチームの傾向を明らかにすることを目的とした。各年で選手を3グループに分け、各グループでタームごとにフィードバックコメントパターンを変更し、グループ間での解析、ターム間での解析、そしてコメント内容の違いによるPF 発揮度の差異を調べた。個人解析として、コメントパターンの違いによるPF 発揮度の差異を調べた。集団解析の結果、グループ間で有意差を確認することができたが、別のグループに属していた同程度の競技レベルの選手について、実験期間内のPF 発揮度に大きな違いが見られたことから、主観評価によって生じた有意差である可能性が考えられる。また、ターム間、コメント内容の違いによる有意差は見られなかった。個人解析の結果、計7名の選手で有意差を確認することができ、フィードバックコメントの内容の違いによって選手ごとに異なる傾向を示す可能性が示された。

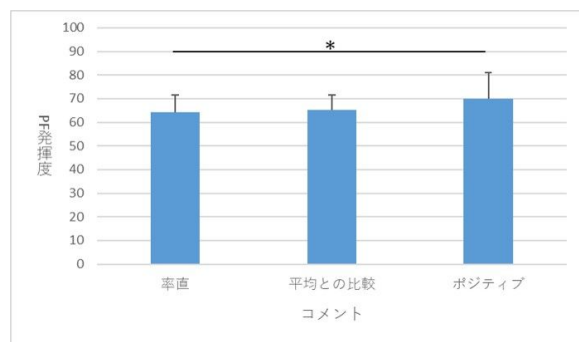


図6: 選手m14の解析結果

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計7件（うち招待講演 1件 / うち国際学会 2件）

1. 発表者名 堀尾 恵一
2. 発表標題 ディスカッション中の幼児の行動分析に基づき役割および個性の推定
3. 学会等名 電子情報通信学会 ヒューマンコミュニケーション基礎研究会
4. 発表年 2018年

1. 発表者名 堀尾 恵一
2. 発表標題 評価値付き入力ベクトルを扱う自己組織化マップを用いたエージェントの学習パラメータに応じた報酬設計手法
3. 学会等名 第34回ファジィシステムシンポジウム
4. 発表年 2018年

1. 発表者名 福田 量
2. 発表標題 ディスカッション中の幼児の行動分析におけるメンバー変更時の行動変化に関する考察
3. 学会等名 電子情報通信学会 ヒューマンコミュニケーション基礎研究会
4. 発表年 2018年

1. 発表者名 森逸 平
2. 発表標題 エージェントの学習過程に基づく学習パラメータの推定及び報酬付与方法に関する考察
3. 学会等名 第33回ファジィシステムシンポジウム
4. 発表年 2017年

1. 発表者名 Yuji Watanabe
2. 発表標題 Childrens' Characteristic Analysis in Group Discussion Using Tensor SOM
3. 学会等名 International Workshop on Smart Info-Media Systems in Asia (国際学会)
4. 発表年 2017年

1. 発表者名 Ippei Mori
2. 発表標題 ESTIMATION OF AGENT'S LEARNING PARAMETERS FROM ITS LEARNING PROCESS AND METHODOLOGY FOR DESIGN OF REWARDS
3. 学会等名 International Conference on Innovative Computing, Information and Control (国際学会)
4. 発表年 2017年

1. 発表者名 堀尾 患一
2. 発表標題 データ解析手法のスポーツ心理学への適用 ~ベイズ統計と可視化からのアプローチ~
3. 学会等名 九州スポーツ心理学会第31回大会 (招待講演)
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	磯貝 浩久  (ISOGAI HIROHISA)  (70223055)	九州産業大学・人間科学部・教授    (37102)	