

機関番号	研究種目番号	審査区分番号	細目番号	分割番号	整理番号
32657	11	-	1204		0001

平成28年度 (2016年度) 挑戦的萌芽研究 研究計画調書

平成 27 年10月27日

1版

新規

研究種目	挑戦的萌芽研究						
分野	情報学						
分科	人間情報学						
細目	知能情報学						
細目表 キーワード	知的システムアーキテクチャ						
細目表以外の キーワード	学習知識の選択，認知心理学						
研究代表者 氏名	(フリガナ)	スズキ ツヨシ					
	(漢字等)	鈴木 剛					
所属研究機関	東京電機大学						
部局	工学部						
職	教授						
研究課題名	認知心理学におけるプライミング効果を基にした強化学習ロボットによる知識選択						
研究経費 〔千円未満の 端数は切り 捨てる〕	年度	研究経費 (千円)	使用内訳(千円)				
			設備備品費	消耗品費	旅費	人件費・謝金	その他
	平成28年度	2,250	1,000	1,000	150	0	100
	平成29年度	2,750	0	2,000	650	0	100
	平成30年度	0	0	0	0	0	0
	総計	5,000	1,000	3,000	800	0	200
開示希望の有無	審査結果の開示を希望する						

研究組織（研究代表者、研究分担者及び連携研究者）

	氏名（年齢）	所属研究機関 部局 職	現在の専門 学位 役割分担	平成28年度 研究経費 （千円）	エフ ォ ー ト （％）
研究代表者	00349789 （47） スズキ ツヨシ 鈴木 剛	（32657）東京電機大学 （404）工学部 （20）教授	ネットワークロボティクス，情報通信 工学 博士（工学） 知識の保存手法の開発	900	10
研究分担者	50646601 （33） オン ブン 温 文	（12601）東京大学 （885）工学（系）研究科（研究院） （24）研究員	認知科学 博士（心理学） 知識を活性化，想起させる手法の実装	500	5
研究分担者	70758367 （29） コウノ ヒトシ 河野 仁	（12601）東京大学 （885）工学（系）研究科（研究院） （24）研究員	知能ロボット 博士（工学） 知識間の関係性記述手法の開発	850	2
合計 3 名			研究経費合計	2,250	

研究目的

本欄には、研究の全体構想及びその中で本研究の具体的な目的について、冒頭にその概要を簡潔にまとめて記述した上で、適宜文献を引用しつつ記述し、特に次の点については、焦点を絞り、具体的かつ明確に記述してください（記述に当たっては、「科学研究費助成事業における審査及び評価に関する規程」（公募要領 75 頁参照）を参考にしてください。）。

- ① 研究の学術的背景（本研究に関連する国内・国外の研究動向及び位置づけ、応募者のこれまでの研究成果を踏まえ着想に至った経緯、これまでの研究成果を発展させる場合にはその内容等）
- ② 研究期間内に何をどこまで明らかにしようとするのか
- ③ 当該分野における本研究の学術的な特色及び予想される結果と意義

研究目的（概要）※ 当該研究計画の目的について、簡潔にまとめて記述してください。

本研究の目的は、**強化学習するロボットが獲得・保存した複数の知識の再利用における知識選択手法の確立**である。本研究では強化学習アルゴリズムに認知心理学で探求されてきた成果、特に**プライミング効果の発現を目的として、活性拡散モデルを用いた知識間接続関係を記述する手法を実現**する。また、申請者らの専門性を融合し、具体的には以下の課題を解決する。

- (1) 強化学習アルゴリズムによる知識の保存方法の開発（工学的システム開発）
- (2) 保存された知識を活性化・想起させる手法の開発（認知心理学の知見の応用）
- (3) 知識間の関係性を記述する手法の開発（工学的&認知心理学的アプローチ）

① 研究の学術的背景

ロボットシステムに実装される学習アルゴリズムとして、Q 学習やニューラルネットワークなどの様々な手法が提案されている。通常、学習した知識は図 1 のように 1 つだけ生成される。近年ではタスクや環境ごとに知識を個別に保存してネットワーク構造として知識を保存・再利用し、環境適応性を高める研究が行われている。しかし、保存された**知識間のネットワークを記述する効果的な手法は確立されていない**。一方、人間では概念や意味がネットワーク構造として記憶されているという仮説が様々な心理学実験により確認され、**プライミング効果の発現がそれを支持している**。プライミング効果とは、先行的に得られる刺激が、後続の処理に無意識的に影響を及ぼす認知心理学の知見である。このプライミング効果が発現するように、ロボットシステム内に保存される獲得知識間のネットワーク構造を記述すれば、知識の検索と選択の正確性と効率を促進できると考える。

そこで本研究では、**強化学習における複数の知識をネットワーク構造で記述・記憶し、効率的な知識選択のために知識ごとに活性レベルを採用する。また、活性拡散モデルに基づいた知識間ネットワークの記述手法と、プライミング効果の発現による知識選択の正確性の向上**を目的とする。

② 何をどこまで明らかにしようとするのか

本申請課題は、学習メカニズム自体が研究対象ではなく、知識の保存方法や活性化・想起方法など、**「学習した後」の知識選択手法の確立が目的**である。

具体的には、ロボットシステムが様々な環境やタスクを学習し、環境やタスクごとに知識を保存する。それらの**知識に活性値を付加し、センサから先行して得られる情報から知識を活性化させることで有効な知識を想起（選択）させるメカニズムの開発**を行う。

③ 当該分野における本研究の学術的な特色および予想される結果と意義

強化学習の分野において学習の高速化や動的環境への適応などの研究が行われ、学習アルゴリズム自体の改良は活発に行われている。しかし、活性拡散モデルやプライミング効果などの**具体的な認知心理学的知見を導入した研究はほとんど例が無く**独自性の高い内容である。

また、本申請課題では基礎的検討として学習アルゴリズムに Q 学習を採用するが、**他の知識ベースシステムやニューラルネットワークなどの機械学習アルゴリズムへの応用が可能**であり、学習システムの実用化に大きく貢献できる。

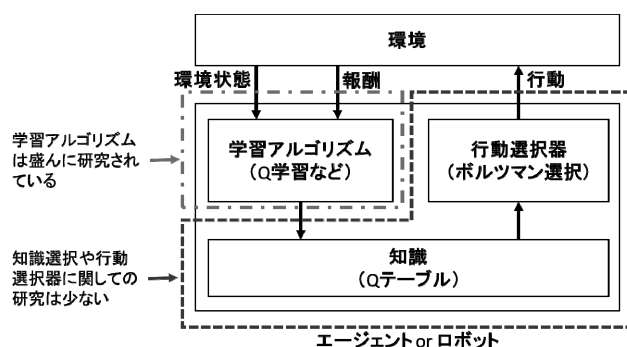


図 1 学習システムの構成

研究の斬新性・チャレンジ性

本欄には、次の点について、焦点を絞り具体的かつ明確に記述してください。

- ① 本研究が、どのような点で斬新なアイデアやチャレンジ性を有しているか
- ② 本研究が、新しい原理の発展や斬新な着想や方法論の提案を行うものである点、または成功した場合に卓越した成果が期待できるものである点等

① 本研究がどのような点で斬新なアイデアやチャレンジ性を有しているか

これまでの学習アルゴリズム関連分野において、強化学習は行動心理学や動物行動学に着想を受けて提案されている。ニューラルネットワークなどは神経科学、知識活用の先端技術である転移学習においては教育学などを基盤として発展を遂げている。これに対し、行動の選択などは他の学問分野から影響を受けている物はなく、 ϵ -greedy 方策や、ボルツマン選択など脈絡ない手法が採用されている。さらに、知識選択においては検討がほとんどなされていない。本申請課題の斬新なアイデアは認知心理学から得られた知見である、人間の行動や知識の無意識的な選択手法であるといえる活性拡散モデルやプライミング効果を知識選択に応用することである。

図2に示すように、本申請課題ではプライミング効果を知識選択と行動選択に応用する。このような、現在の環境状態ではない先行して得られる環境状態を知識選択や行動選択に用いる手法はこれまでに無く、認知心理学的知見を強化学習で利用できるように、具体的なメカニズムを開発することがチャレンジングである。

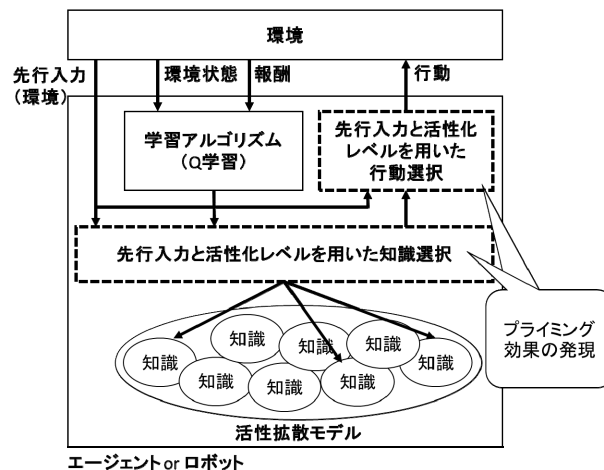


図2 プライミング効果を導入した強化学習

② 新しい原理の発見や斬新な着想や方法論の提案、成功した場合に卓越した成果の期待

本申請課題は、言わば「直感」や「勘」のような、人間が無意識的に行うメカニズムを、ロボットシステムに実装することが最終的な目標である。本申請課題は、その第一歩として強化学習の知識の再利用に着目し、認知心理学的知見を応用する。

本申請課題で期待される成果の1つは、強化学習で獲得した知識の効果的な選択によるロボットやエージェントの振る舞いの多様化である。

これまでの学習アルゴリズムでは単一タスクを学習することが主な目的となり、本来人間や動物が持つ学習した結果の多様性が実現できなかった。ロボットシステムが環境とインタラクションする場面では、学習だけが重要ではなく、学習した結果（知識）を適切に再利用することも重要である。そこで、本申請課題では基礎的検討として、認知心理学分野で研究されてきた活性拡散モデルを参考に知識間の関係を記述し、先行入力と活性化レベルという基準を用いて知識を選択することで、強化学習の実装されたロボットシステムの振る舞いの多様化へのブレークスルーを狙う研究である。また図3に示すように、認知心理学における知見の活用を通じて、ロボットシステムへの実装を行うことでプライミング効果の新たな仮説の発見や、認知心理学分野への知見のフィードバックを行えると考えられ、各分野の共創を実現する。

本研究は、学習プロセスと獲得する知識が時空間的に分離可能なメカニズムであれば、その学習アルゴリズムへの応用も可能であり、例えばニューラルネットワークや他の機械学習手法へ研究成果・技術を転移することも可能である。これにより、本研究で取り上げる強化学習以外にも知識の選択が可能となり、学習ロボットや学習エージェントの実用化・普及に大きな貢献が期待される。

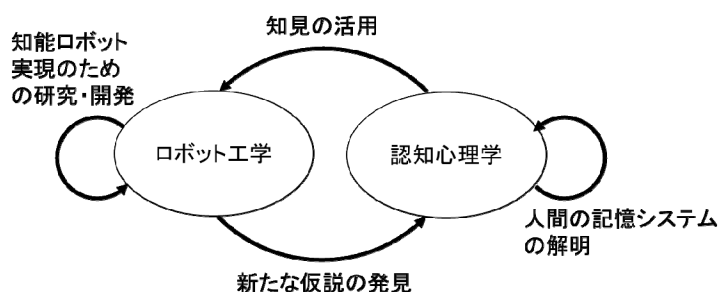


図3 本申請課題における新しい原理の発見

研究計画・方法

本欄には、研究目的を達成するための具体的な研究計画・方法について、冒頭にその概要を簡潔にまとめて記述した上で、平成28年度の計画と平成29年度以降の計画に分けて、適宜文献を引用しつつ、焦点を絞り、具体的かつ明確に記述してください。また、研究計画を遂行するための研究体制について、研究分担者とともに行う研究計画である場合は、研究代表者、研究分担者の具体的な役割（図表を用いる等）、学術的観点からの研究組織の必要性・妥当性及び研究目的との関連性についても述べてください。さらに、研究体制の全体像を明らかにするため、連携研究者及び研究協力者（海外共同研究者、科研費への応募資格を有しない企業の研究者、その他技術者や知財専門家等の研究支援を行う者、大学院生等（氏名、員数を記入することも可））の役割についても記述してください。

なお、研究期間の途中で異動や退職等により研究環境が大きく変わる場合は、研究実施場所の確保や研究実施方法等についても記述してください。

研究計画・方法（概要）※ 研究目的を達成するための研究計画・方法について、簡潔にまとめて記述してください。

本申請課題では、強化学習するロボットが獲得した知識を環境やタスクごとに保存・蓄積し、先行入力や活性拡散モデルを用いて、適切な知識の選択・再利用を実現する手法の確立を目指す。

平成28年度では、(1)強化学習により獲得した知識の再利用可能形式での保存方法の開発や、(2)保存されている知識に活性値を記述する方法、プライミング効果により知識を想起させる手法の開発を行う。また、(3)保存される知識は単純に保存されるだけでなく、知識同士の関係性を記述する手法の開発も行う。

平成29年度においては、上記3つの手法を統合し、自律移動ロボットによる評価実験を行う。

以上について、図4に示す役割分担で有機的に連携して遂行する。

【平成28年度】

(1)強化学習により獲得した知識の保存方法の開発（鈴木・大学院生）

本申請課題では、強化学習するロボット（以下、強化学習ロボット）が獲得した知識を再利用可能な形式で保存する必要がある。これは、1つの知識（強化学習の方策）で全ての学習結果を記録することは現実的でなく、タスクや環境ごとに知識を分割し、選択させる形で知識の再利用を行うからである。

強化学習の知識の保存方法に関しては、いくつか先行研究が存在する。例えば、基本的な Look-up table として保存する方法やニューラルネットワークとして記述する関数近似を用いた方法などがある。しかし、関数近似を用いた手法では関数の近似誤差により獲得知識の品質が劣化する可能性がある。そのため、本項目では Look-up table 型の知識保存方法を採用し、強化学習ロボットがリロードしやすい形式を開発する。さらに、活性値などの直接知識内には記述できないパラメータの保存方法も開発する。

(2)保存された知識を活性化・想起させる手法の開発（温）

強化学習ロボットにより複数の知識が獲得され、それらを再利用するとき各知識に記述された活性値により知識の選択を行う。知識を再利用する際、各知識は次項で述べるネットワーク構造で互いの知識の接続関係を持っているとする。このとき、環境の先行入力（本研究ではカメラ情報から得られる色など）を用いて活性値が高い再利用知識の候補を上げる（図5）。さらに、それらの活性値を比較し任意の閾値（活性レベル）を超えた知識を再利用する。さらに、再利用された知識は活性化を受け、知識と共に保存されている活性値の上昇がなされる。これらの繰り返しのにより、環境に対する適切な知識の選択を獲得していく。

しかし、先行入力による知識選択が必ずしも環境に対して正しい知識であるとは限らない。そのため、本項の「知識の活性化・想起」手法では想起が正しかったのかどうか判断する評価関数や想起プロセスにまで検討を深める。

また、知識選択後の行動選択部分においても先行入力による活性値を活用した手法を検討する。

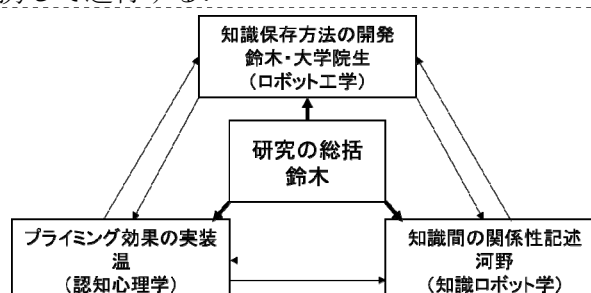


図4 役割分担

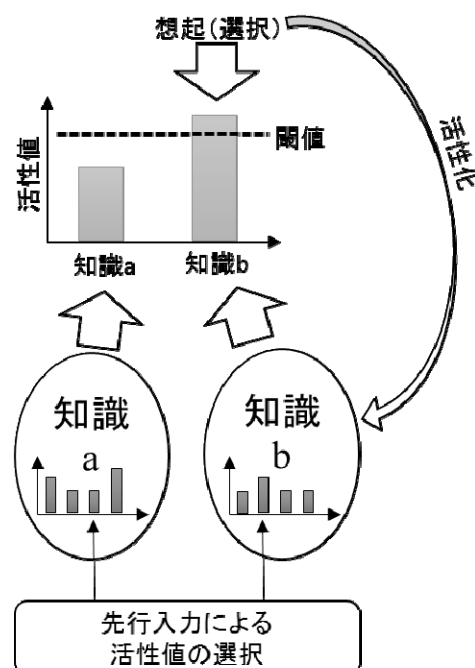


図5 活性値による知識選択の例

研究計画・方法（つづき）

(3)知識間の関係性を記述する手法の開発（河野）

強化学習ロボットが獲得した知識は、システマ的に個別に保存されているだけでは再利用するためのラベルがなく、また、ラベルを付加する場合でも知識獲得毎にラベル定義作業が必要である。そのため、知識間での関係性を知識獲得時に自動設定するようにしておき、知識選択時の想起頻度（選択頻度）に応じて知識間の関連度合いを高めるアルゴリズムが必要である。

そこで、本項目では活性拡散モデルや意味ネットワークなどのモデルを参考に、強化学習ロボット内に知識間の関係性を記述する手法を開発する。具体的には図 6 に示すような、知識間の接続関係を記述する手法と想起頻度に応じて知識間距離を調整するアルゴリズムを開発する。

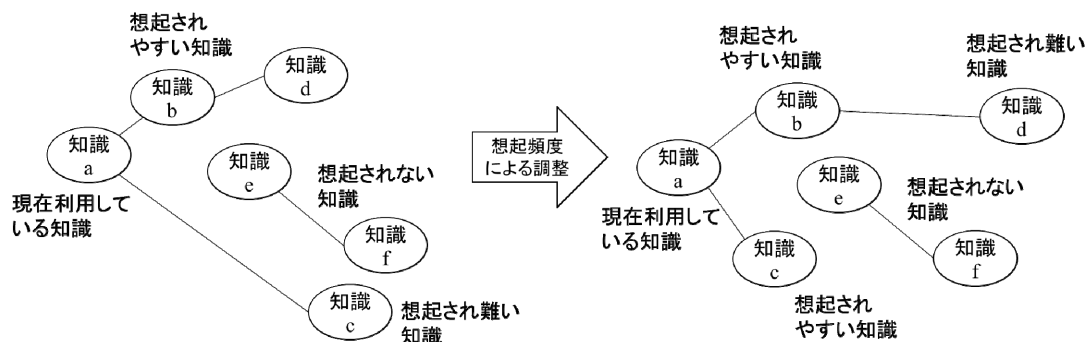


図 6 知識間の接続関係と想起頻度による接続関係の調整の例

【平成 29 年度】

(4)提案手法の実験用プラットフォームへの実装（鈴木・河野）

平成 28 年度に開発した 3 つの機能を、実際の移動ロボットへ実装する。ロボットシステムには鈴木・河野らが開発し、取り扱いに精通している卓上実験サイズの自律全方向移動ロボット（以下、実ロボット）を用いる（図 7）。また、現在このロボットは試作機 1 台のみであり、他の研究用のため本申請課題では新たに 2 台の実ロボットを製作し、研究代表者と研究分担者の各大学で実験が行えるようにする。

ロボットが蓄積した知識の活性値を調整、活性化による想起を発生させるために、本年度では図 2 の先行入力にカメラから得られる色情報を用いる。そのため、本項目ではカメラシステムのロボットへの実装も行う。

(5)自律移動ロボットを用いた評価実験（鈴木・温・河野）

最終年度の実験として、上記項目(4)にてプライミング効果を実装した強化学習ロボットに評価実験を行う。実験環境としては萌芽的本研究の性質から、図 8 に示すようなシンプルな最短経路探索問題で評価を行うことを検討している。

本実験では様々な経路の最短経路の環境を用意し、同時に壁の色を数種類用意する。これにより、過去に学習した環境の知識を、カメラシステムを用いて認識した環境中の色により想起させる実験を行い、プライミング効果の発現を観察する。実験の多様性を持たせるために、複数の色が存在する環境や、さらには意図的に色と経路に関係性のない環境においても実験を行う。

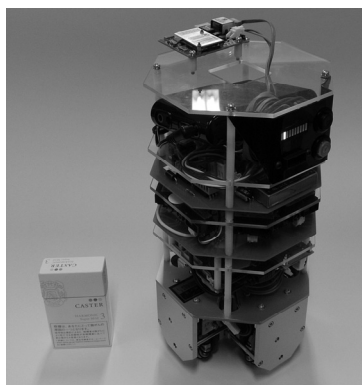


図 7 自律全方向移動ロボット



図 8 最短経路探索問題の例

人権の保護及び法令等の遵守への対応（公募要領 4 頁参照）

本欄には、研究計画を遂行するに当たって、相手方の同意・協力を必要とする研究、個人情報の取り扱いの配慮を必要とする研究、生命倫理・安全対策に対する取組を必要とする研究など法令等に基づく手続が必要な研究が含まれている場合に、どのような対策と措置を講じるのか記述してください。

例えば、個人情報を伴うアンケート調査・インタビュー調査、提供を受けた試料の使用、ヒト遺伝子解析研究、組換え DNA 実験、動物実験など、研究機関内外の倫理委員会等における承認手続が必要となる調査・研究・実験などが対象となります。

なお、該当しない場合には、その旨記述してください。

本申請課題はロボットシステムを対象とした研究・開発であり、倫理委員会等の承認が必要な研究計画には該当しません。

研究経費の妥当性・必要性

本欄には、「研究計画・方法」欄で述べた研究規模、研究体制等を踏まえ、次頁以降に記入する研究経費の妥当性・必要性・積算根拠について記述してください。また、研究計画のいずれかの年度において、各費目（設備備品費、旅費、人件費・謝金）が全体の研究経費の 90% を超える場合及びその他の費目で、特に大きな割合を占める経費がある場合には、当該経費の必要性（内訳等）を記述してください。

本申請課題は、強化学習の内部メカニズムにプライミング効果を導入し、環境に適応した知識選択を実現する研究である。そのため、検討段階では計算機シミュレーションなどで研究作業が可能である。計算機シミュレーションなどは申請者らが所有しているワークステーションを活用する。

しかし、実ロボット実験においては汎用のロボットが必要であるため、新たに 2 台の全方向移動ロボットの製作費用を平成 28 年度に計上した。このロボットは鈴木と河野により開発された汎用卓上実験サイズ全方向ロボットシステムであり、他の研究でも利用していることから、プログラム実装やメンテナンスに精通している。そのため、速やかに実ロボット実験に取り掛かることが可能である。

全方向移動ロボット以外にも、環境からの先行入力を認識するための USB カメラやインタフェースなどの購入、全方向移動ロボットの部品磨耗によるメンテナンス部品の購入のために、電子部品費と機械部品費として平成 28 年度では 1,000 千円、平成 29 年度では 2,000 千円を計上した。

各年度において、本研究によってえられる成果の発表および議論のために、国内の成果発表のための旅費（平成 28 年度 150 千円、平成 29 年度 250 千円）を計上した。また、平成 29 年度では、得られた成果を国外へ積極的に発信するために国際学会のための旅費(400 千円)を計上した。加えて、各年度において査読付き論文誌（特に International journal かつ Open access の論文誌）への論文投稿を積極的に行うため、投稿費用を各年度で 100 千円計上した。

なお研究計画のいずれの年度においても、各費用が全体の 90%を超えることはない。

挑戦的萌芽－6
(金額単位：千円)

設備備品費の明細					消耗品費の明細				
記入に当たっては、挑戦的萌芽研究 研究計画調書作成・記入要領を参照してください。					記入に当たっては、挑戦的萌芽研究 研究計画調書作成・記入要領を参照してください。				
年度	品 名・仕 様 (数量×単価) (設置機関)				金 額	品 名		金 額	
28	全方向移動ロボット ZEN-Q (2×500 千円) (東京電機大学 1 台, 東京大学 1 台)				1,000	電子部品費		500	
						機械部品費		500	
	計				1,000	計		1,000	
29	なし				0	電子部品費		1,000	
						機械部品費		1,000	
	計				0	計		2,000	
旅費等の明細 記入に当たっては、挑戦的萌芽研究 研究計画調書作成・記入要領を参照してください。									
年度	国内旅費		外国旅費		人件費・謝金		その他		
	事 項	金額	事 項	金額	事 項	金額	事 項	金額	
28	研究成果発表	100	なし	0	なし	0	論文投稿料	100	
	打ち合わせ	50							
	計	150	計	0	計	0	計	100	
29	研究成果発表	200	研究成果発表	400	なし	0	論文投稿料	100	
	打ち合わせ	50							
	計	250	計	400	計	0	計	100	

研究費の応募・受入等の状況・エフォート

本欄は、第2段審査（合議審査）において、「研究資金の不合理な重複や過度の集中にならず、研究課題が十分に遂行し得るかどうか」を判断する際に参照するところですので、本人が受け入れ自ら使用する研究費を正しく記載していただく必要があります。本応募課題の研究代表者の応募時点における、（１）応募中の研究費、（２）受入予定の研究費、（３）その他の活動について、次の点に留意し記入してください。なお、複数の研究費を記入する場合は、線を引いて区別して記入してください。具体的な記載方法等については、研究計画調書作成・記入要領を確認してください。

- ① 「エフォート」欄には、年間の全仕事時間を100%とした場合、そのうち当該研究の実施等に必要となる時間の配分率（%）を記入してください。
- ② 「応募中の研究費」欄の先頭には、本応募研究課題を記入してください。
- ③ 科研費の「新学術領域研究（研究領域提案型）」にあつては、「計画研究」、「公募研究」の別を記入してください。
- ④ 所属研究機関内で競争的に配分される研究費についても記入してください。

（１）応募中の研究費

資金制度・研究費名（研究期間・配分機関等名）	研究課題名（研究代表者氏名）	役割（代表・分担の別）	平成28年度の研究経費（期間全体の額）（千円）	エフォート（%）	研究内容の相違点及び他の研究費に加えて本応募研究課題に応募する理由（科研費の研究代表者の場合は、研究期間全体の受入額を記入すること）
【本応募研究課題】 挑戦的萌芽研究 (H28～H29)	認知心理学における プライミング効果を 基にした強化学習ロ ボットによる知識選 択 (鈴木剛)	代表	900 (1,850)	10	(総額 5,000 千円)
基盤研究 B (一般) (H28～H30)	ヘテロジニティを 考慮したマルチエー ジェント転移学習に よる知識の共創 (鈴木剛)	代表	1,880 (4,770)	15	異なる機能を持つ複数台の ロボット間での転移学習 と、解法の異なる学習知識 の統合による洗練知識の生 成を目的とする。手法の一 部に知識選択を含むが、ロ ボットの異種性評価に基づ くものであり、認知心理学 的知見を応用する本応募研 究課題とは目的およびアプ ローチが異なる。 (総額 17,670 千円)

研究費の応募・受入等の状況・エフォート（つづき）					
（２）受入予定の研究費					
資金制度・研究費名（研究期間・配分機関等名）	研究課題名（研究代表者氏名）	役 割 （代表・分担の別）	平成 28 年度 の研究経費 （期間全体の額 （千円）	エ フ ォ ー ト（％）	研究内容の相違点及び他の研究費に加えて本応募研究課題に応募する理由 （科研費の研究代表者の場合は、研究期間全体の受入額を記入すること）
基盤研究（B）（一般） （H27～H29）	サンゴ礁生態系保全支援のための観測情報場構築技術の研究 （川端邦明）	分担	1,100 （3,300）	10	サンゴ礁の環境保全のための、水中状態観測を支援する観測情報場を構築する基盤技術の開発を行う。本応募研究課題とは目的および期待される成果が異なる。
（３）その他の活動 上記の応募中及び受入予定の研究費による研究活動以外の職務として行う研究活動や教育活動等のエフォートを記入してください。				65	
合 計 上記（１）、（２）、（３）のエフォートの合計				100 （％）	