

産総研人工知能セミナー
「脳全体の情報処理と人工知能」
2017. 3. 15

脳と人工知能のギャップとリンク



銅谷 賢治

doya@oist.jp

沖縄科学技術大学院大学 神経計算ユニット

沖縄科学技術大学院大学 (OIST)

国際的・学際的な研究教育
沖縄の自立的発展に寄与



■ 設立法人理事長
シドニー・ブレナー



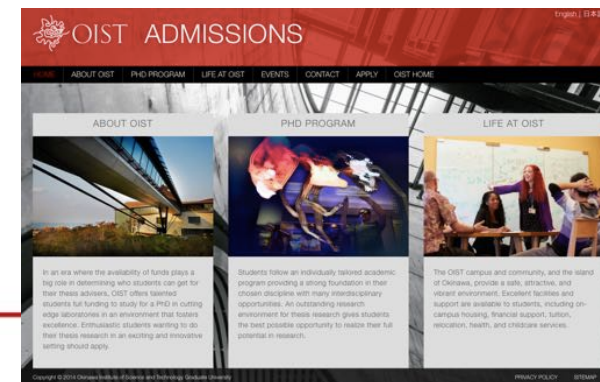
■ 初代学長
ジョナサン・ドーファン

■ 現学長
ピーター・グルース



5年制博士課程 2017年9月入学生募集中

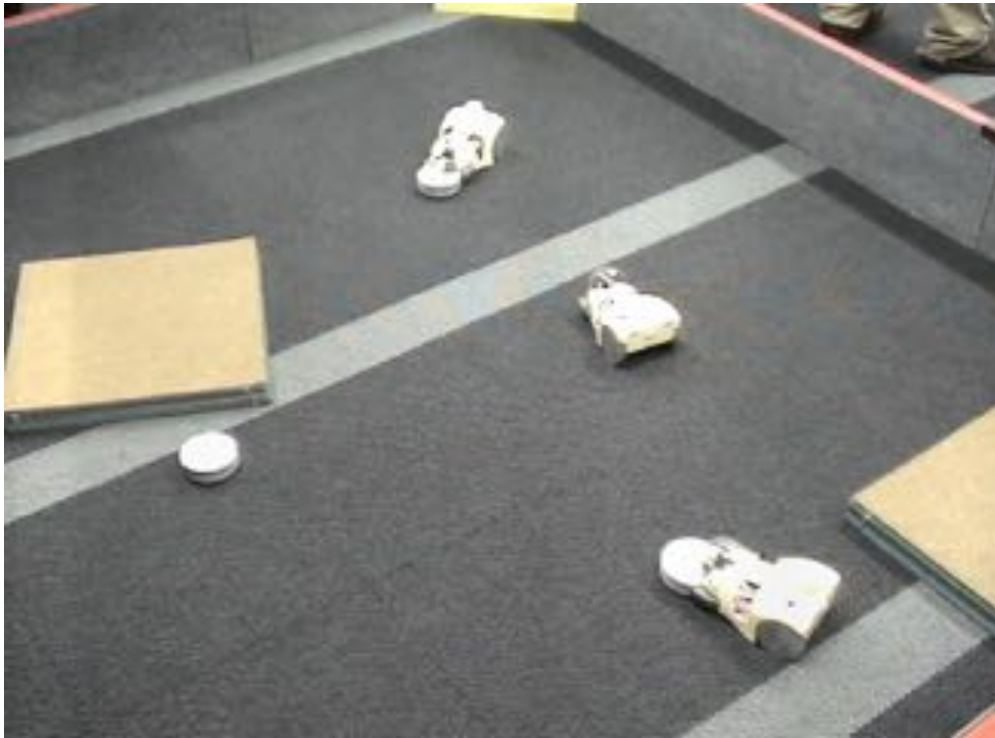
admissions.oist.jp



OIST 神経計算ユニット

柔軟な適応システムをつくる

■ ロボット工学



脳の柔軟な学習機構は？

■ 神経科学



本日のテーマ

脳と人工知能

強化学習

学習し進化するロボット

脳内シミュレーション



脳と人工知能

電子回路で知能を実現するために
脳のしくみにとらわれる必要はない。

脳のような高度な知能の実現例がある
のだからそれに学ばない手はない。

前世紀の人工知能：専門家の知識をプログラム化

今日の人工知能：ビッグデータからの統計的機械学習

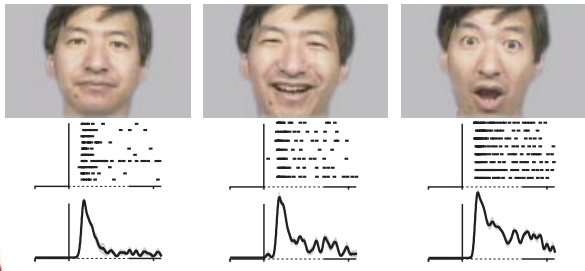
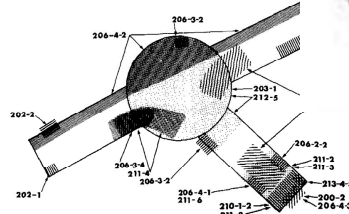
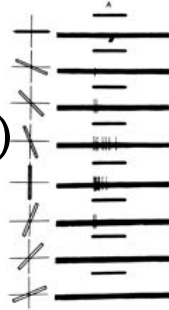
高性能を追い求めた結果、ディープラーニング
のような脳を模した学習の強みが再認識された。



脳科学と人工知能の共進化：視覚

脳科学

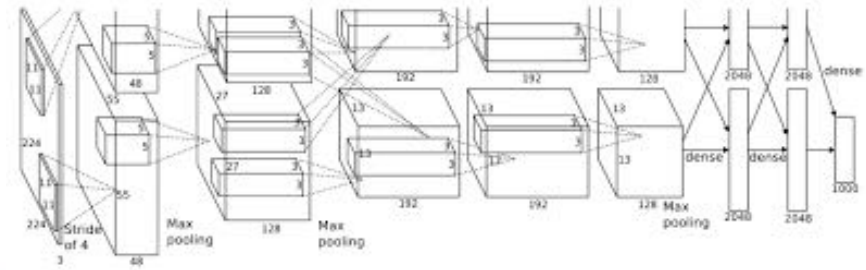
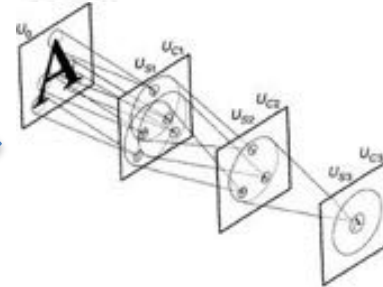
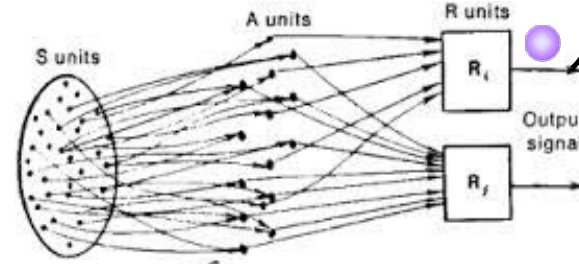
- 特徴抽出細胞 (Hubel & Wiesel 1959)
- 経験依存学習 (Blakemore & Cooper 1970)
- 場所細胞 (O'Keefe 1976)
- 顔細胞 (Bruce, Desimone, Gross 1981)



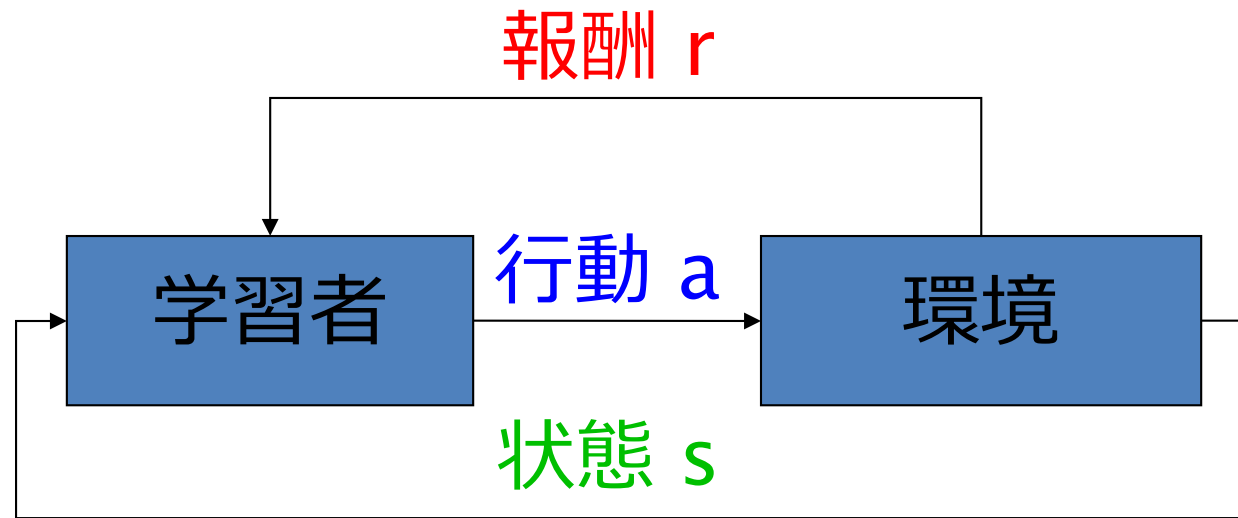
(Sugase et al. 1999)

人工知能

- パーセプトロン (Rosenblatt 1962)
- 多層教師あり学習 (Amari 1967)
- ネオコグニトロン (Fukushima 1980)
- ConvNet (Krizhevsky, Sutskever, Hinton, 2012)
- GoogleBrain (2012)



強化学習



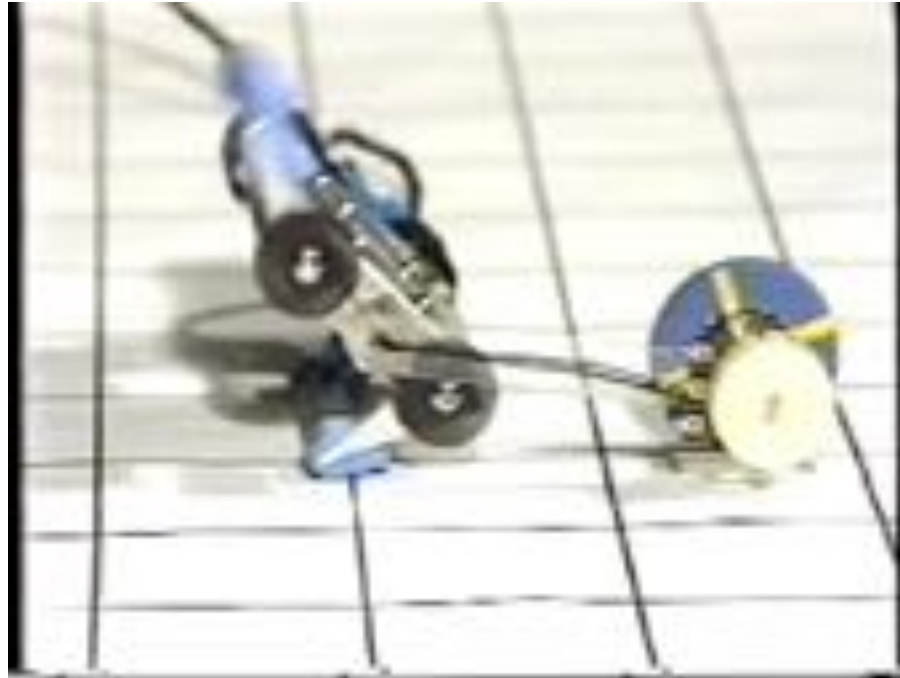
より多くの「報酬」が得られる行動を探索して学習する

- これをいかに確実に効率よく実現するか？
- これがいかに脳で実現されているか？



歩行運動学習ロボット (銅谷, 中野, 1985)

- 行動：4つの目標姿勢 + 周期
- 報酬：歩行スピード



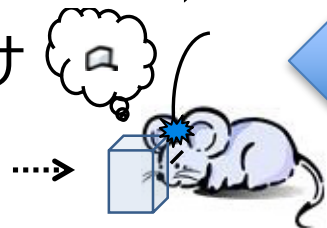
- はい歩き / かけ足 / ジャンプなどのパターンを学習



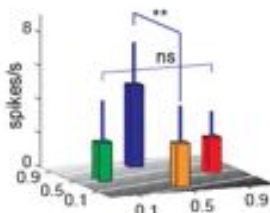
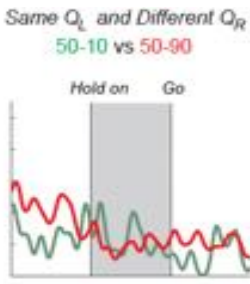
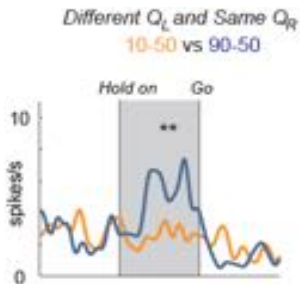
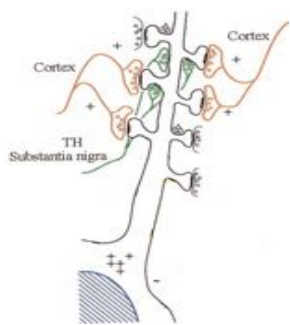
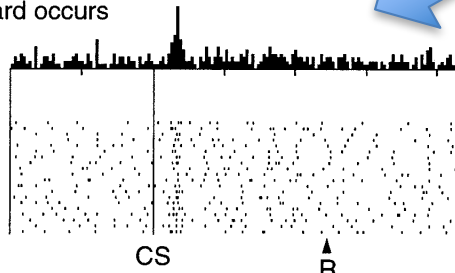
強化学習

脳科学

- 古典的条件付け (Pavlov 1903)
- オペラント条件付け (Thorndike 1898)
- ドーパミン細胞の報酬予測誤差応答 (Schultz et al. 1993, 1997)
- ドーパミン依存シナプス可塑性 (Wickens et al. 2000)

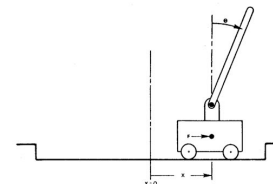
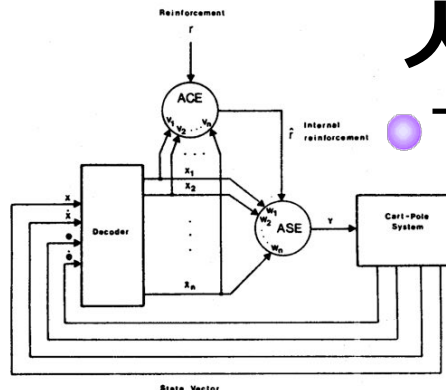


Reward predicted
Reward occurs

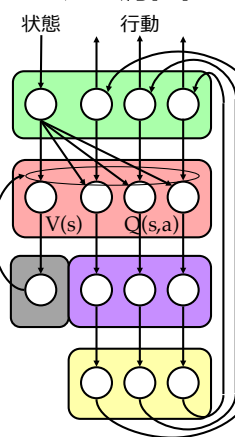


人工知能

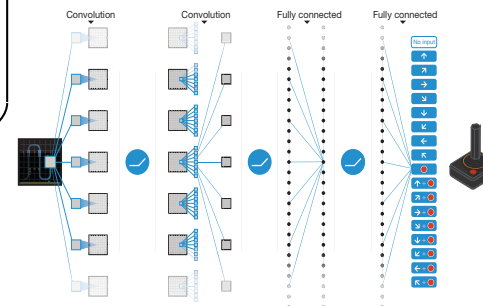
- TD学習 (Sutton et al. 1983)



- 大脳基底核TD学習仮説 (Barto et al. 1995, Montague et al. 1996)



(Barto et al. 1995, Montague et al. 1996)



- 線条体の行動価値表現 (Samejima et al. 2005)

- Deep Q network (Mnih et al. 2015)

AlphaGo

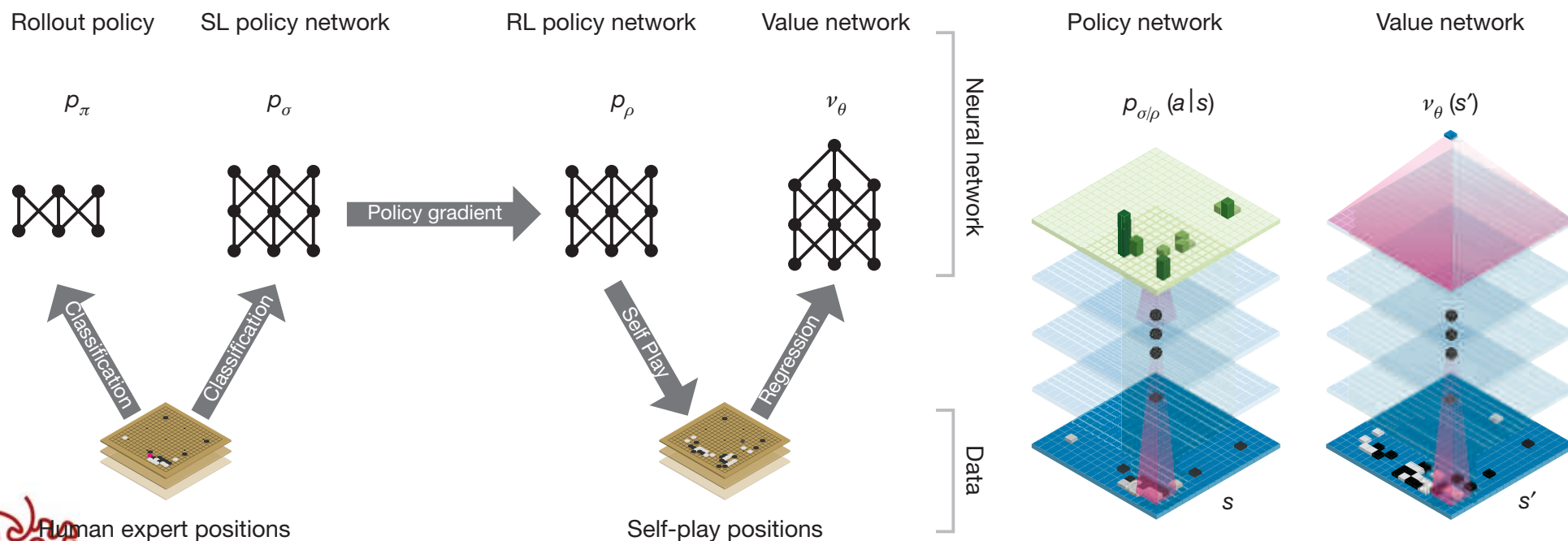
(Silver et al., 2016, Nature)

ARTICLE

Mastering the game of Go with deep neural networks and tree search

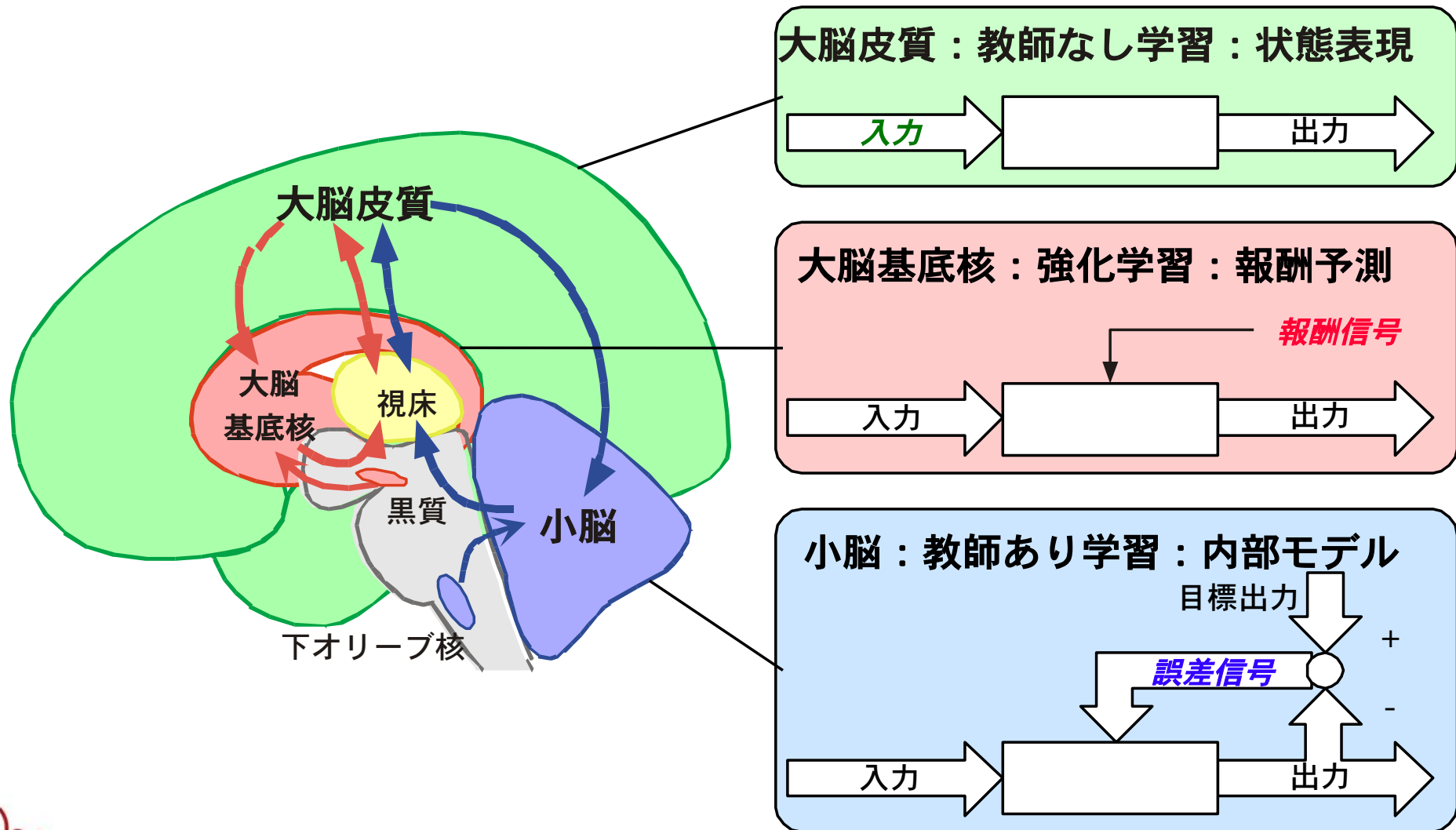
David Silver^{1*}, Aja Huang^{1*}, Chris J. Maddison¹, Arthur Guez¹, Laurent Sifre¹, George van den Driessche¹, Julian Schrittwieser¹, Ioannis Antonoglou¹, Veda Panneershelvam¹, Marc Lanctot¹, Sander Dieleman¹, Dominik Grewe¹, John Nham², Nal Kalchbrenner¹, Ilya Sutskever², Timothy Lillicrap¹, Madeleine Leach¹, Koray Kavukcuoglu¹, Thore Graepel¹ & Demis Hassabis¹

- 棋譜からの教師あり学習
- 自己対戦での強化学習
- 深層ネットワークの盤面表現学習
- 盤面評価による効率良い探索



学習アルゴリズムによる機能分化

(Doya, 1999)



脳と人工知能のちがい

自主性：自ら生きるものと作られたもの

情報選択：必要な情報を自ら見つけるか、与えられるか

能動性：仮説生成と検証による理解か、データ統計か

モジュール自己組織化：必要に応じた選択、結合、生成



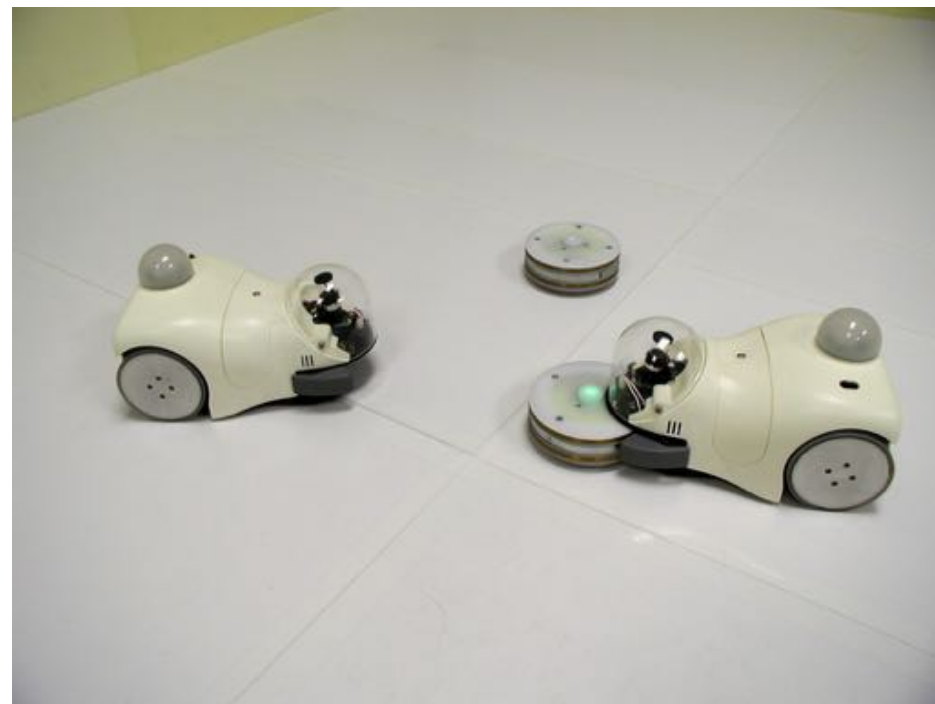
ロボットの報酬系の設計

報酬とは何か？ ... 生物にとって最も重要な機能のため

- 食欲、痛み：自己保存
- 愛情、性欲：自己複製

Cyber Rodents

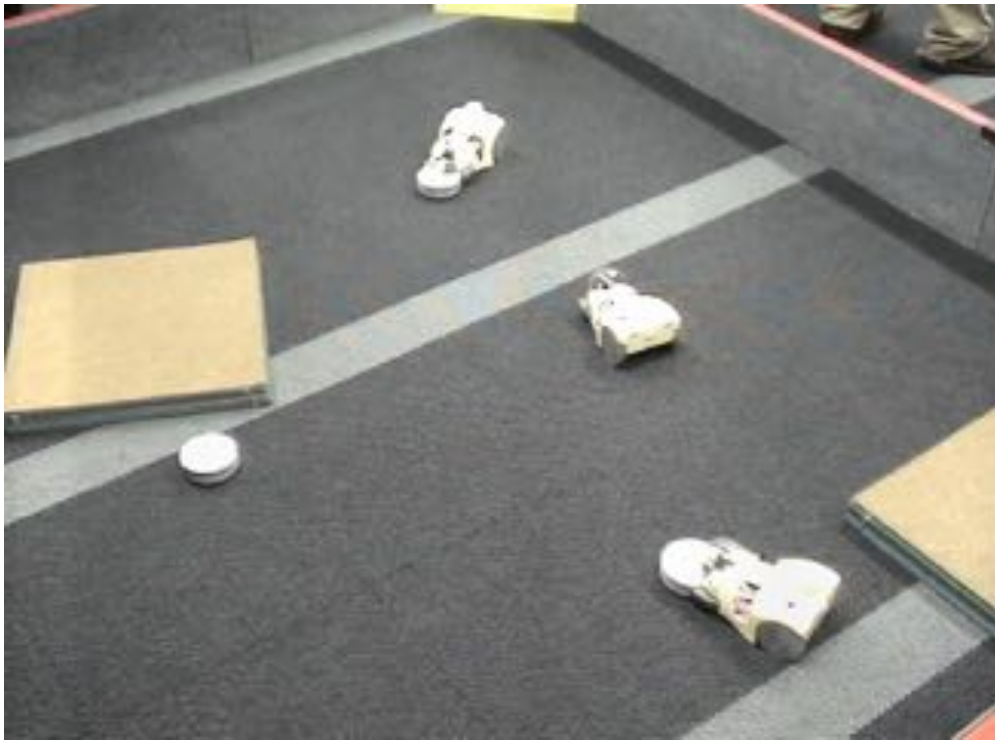
- 電池パックの捕獲と充電
- 赤外線でのソフトのコピー



学習し進化するロボット集団

■ 生存

- バッテリー捕獲



■ 複製

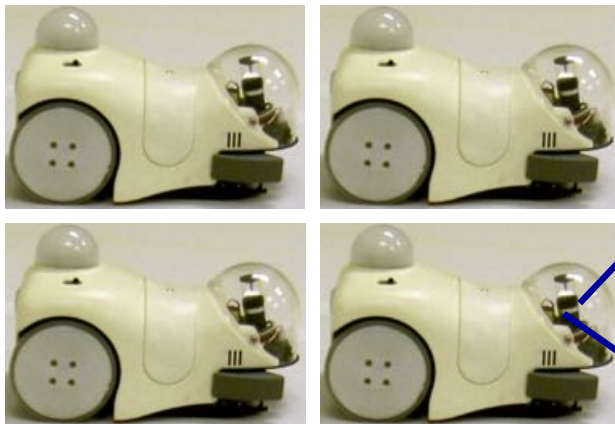
- “遺伝子”のコピー



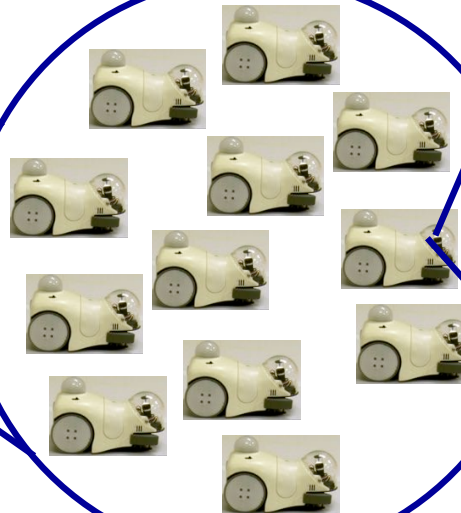
ロボットの学習のしかたの進化

Population

Robots



Virtual agents 15-25



Genes

Weights for top layer NN

$$W_1, W_2, \dots, W_n$$

Weights shaping rewards

$$V_1, V_2, \dots, V_n$$

Meta-parameters

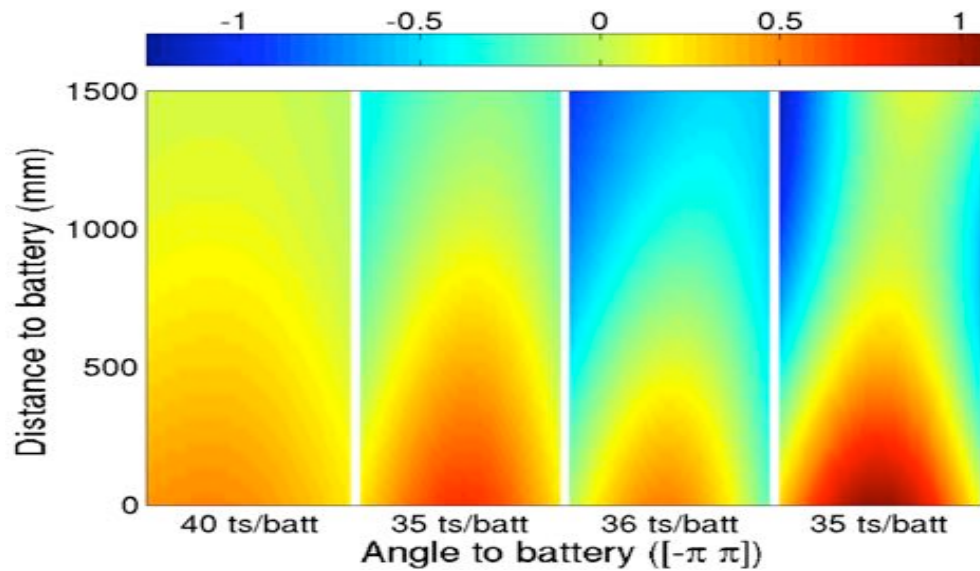
$$\alpha \gamma \lambda \tau_k \tau_0$$

- 一生(約5分)の間に交配して遺伝子を残す
 - 交配の成功は、充電レベルに比例

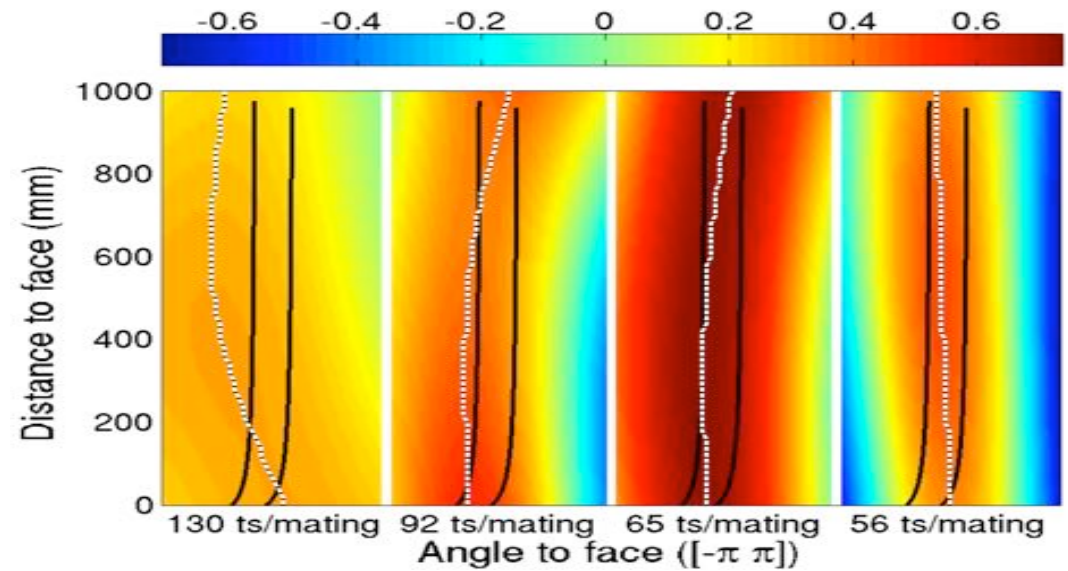


進化により獲得された視覚報酬

■ バッテリー



■ 他のロボットの顔



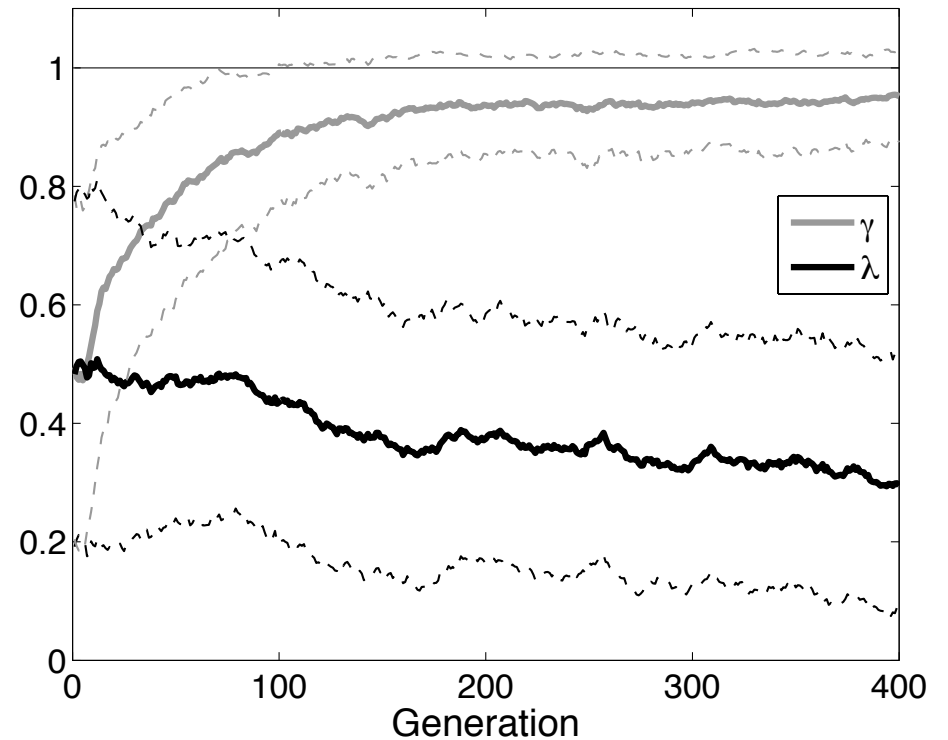
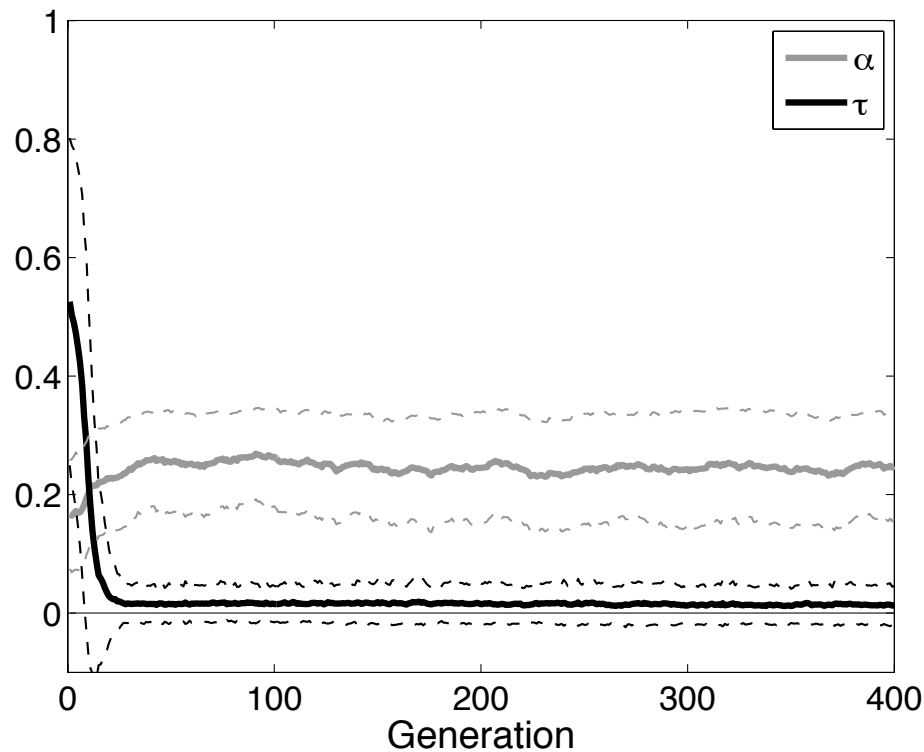
メタパラメタの進化

■ 学習速度 α

■ 探索の温度 τ

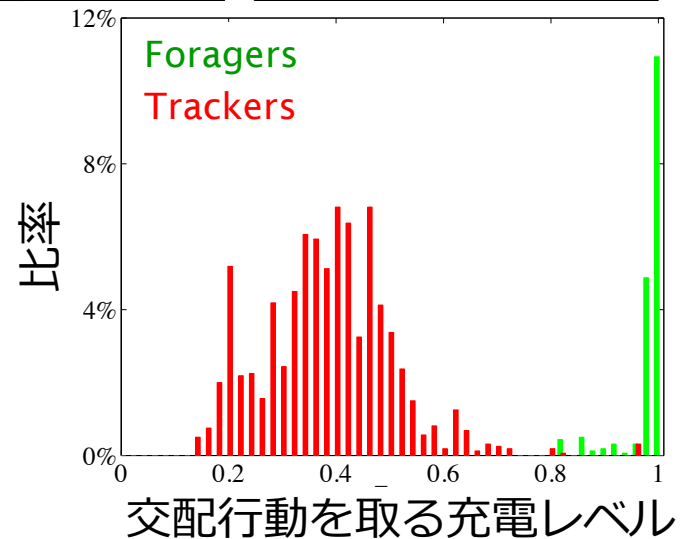
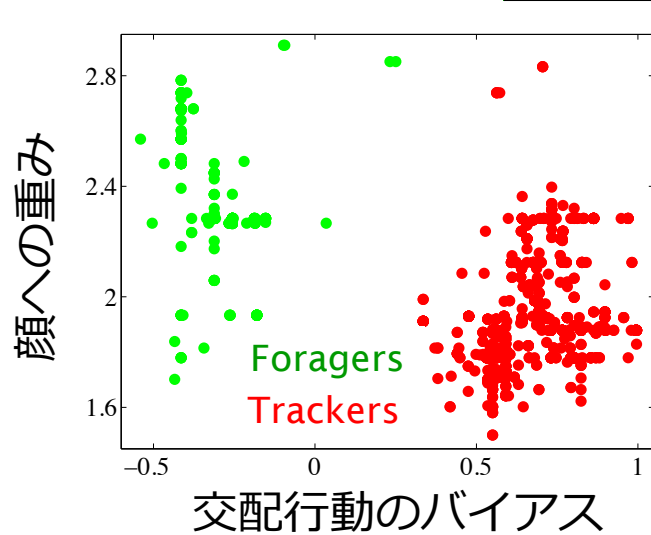
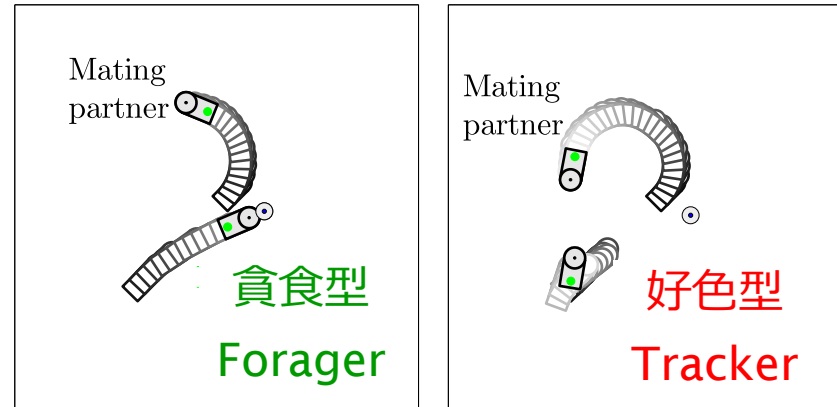
■ 遅延報酬の割引率 γ

■ 状態-行動の履歴 λ

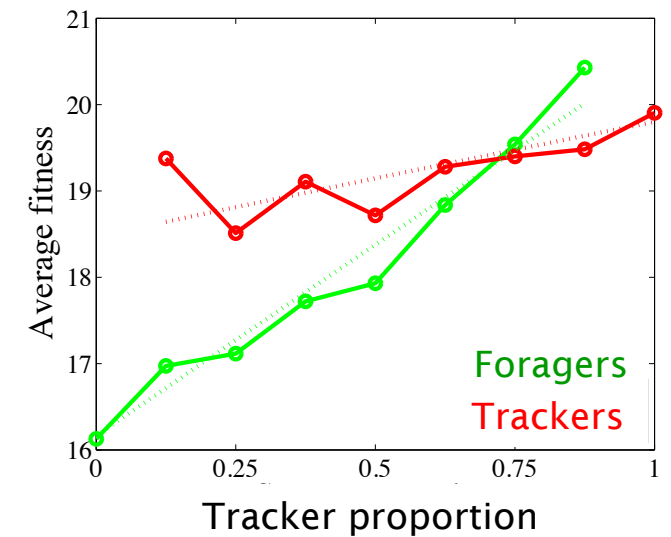


行動戦略の多型性

■ 貪食ロボと好色ロボ



■ 進化的安定性



脳内シミュレーション

「意識」の脳機構：面白いけど定義と検証が困難

脳内シミュレーション：

環境の状態遷移モデル $P(s'|s,a)$ を使った予測

- 過去の状態と行動から、現在の状態を推定
 - 多義感覚識別、場所細胞、...
- 現在の状態から、想定した行動の結果を予測
 - モデルベース意思決定、行動計画、...
- 想定した状態から、行動の結果や原因を予測
 - 思考、推論、言語、科学、...



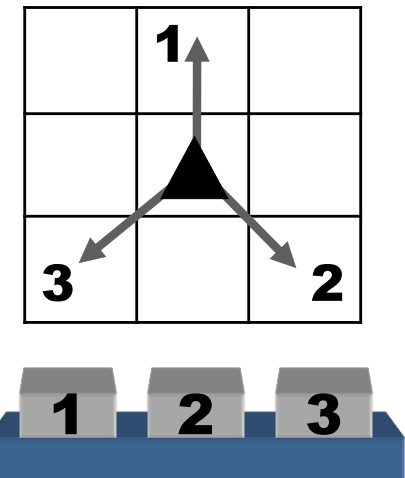
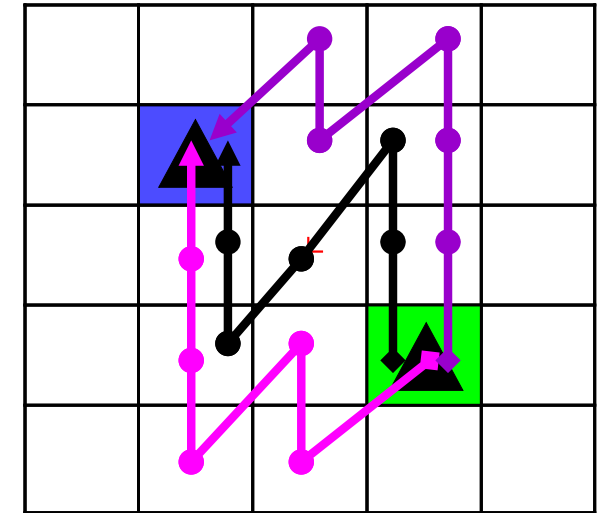
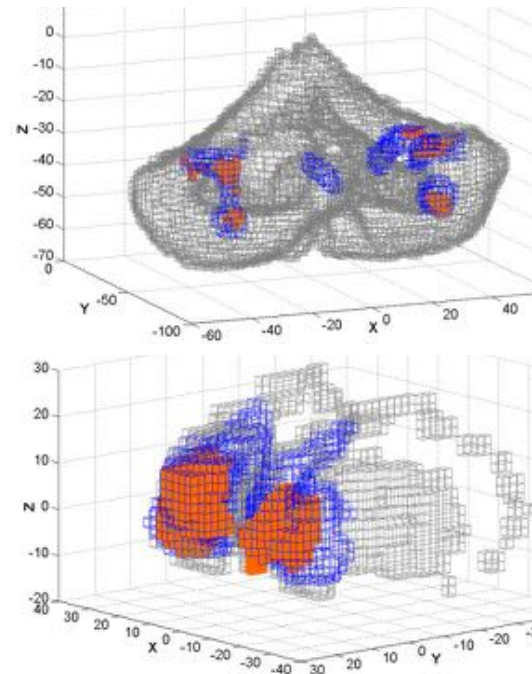
OPEN

Model-based action planning involves cortico-cerebellar and basal ganglia networks

Received: 16 February 2016

Accepted: 19 July 2016

Alan S. R. Fermin^{1,2,3}, Takehiko Yoshida^{1,2}, Junichiro Yoshimoto^{1,2}, Makoto Ito², Saori C. Tanaka⁴ & Kenji Doya^{1,2,3,4}



動的ベイズ推定の神経回路

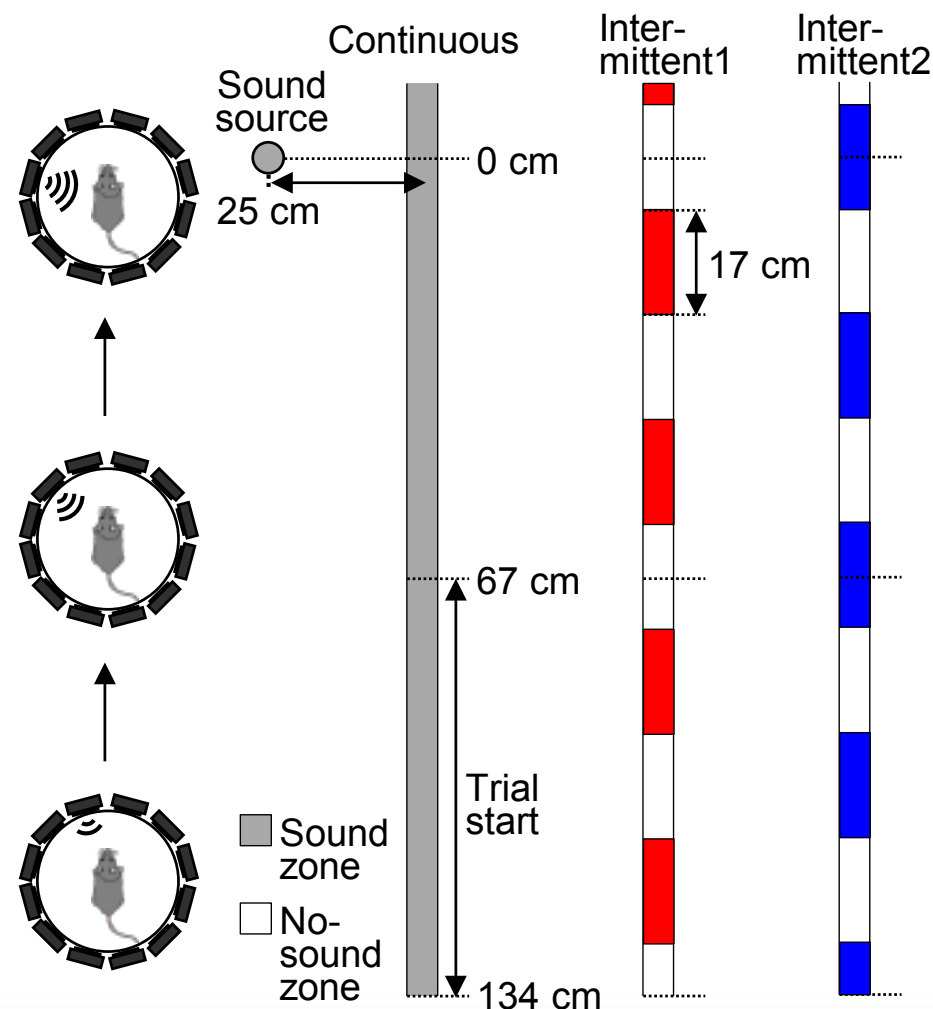
ARTICLES

nature
neuroscience

Neural substrate of dynamic Bayesian inference in the cerebral cortex

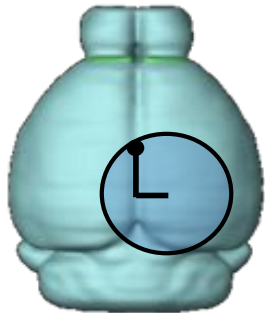
Akihiro Funamizu^{1,2}, Bernd Kuhn² & Kenji Doya¹

■ 二光子顕微鏡下の仮想空間

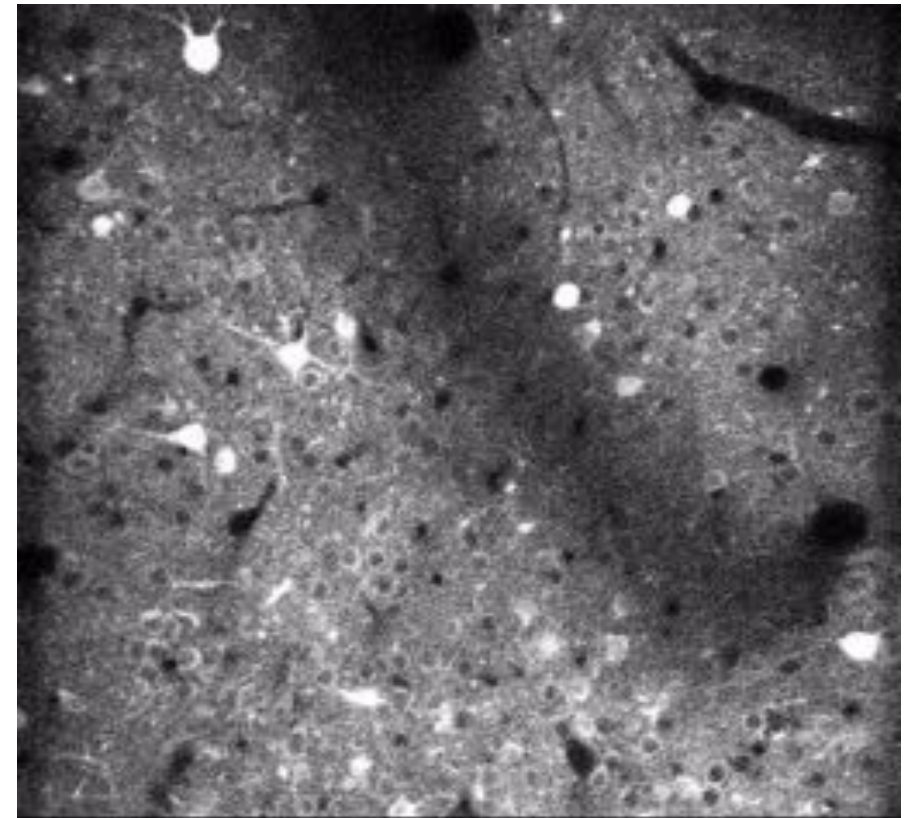
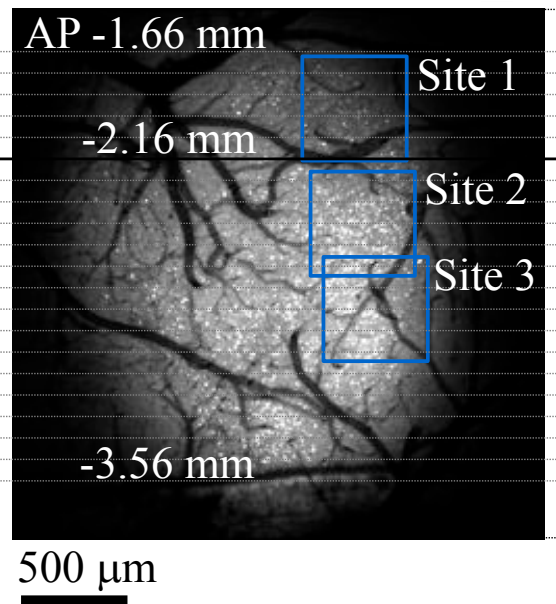


二光子顕微鏡による神経活動記録

■ AAウイルスによるカルシウム感受性蛍光タンパク発現

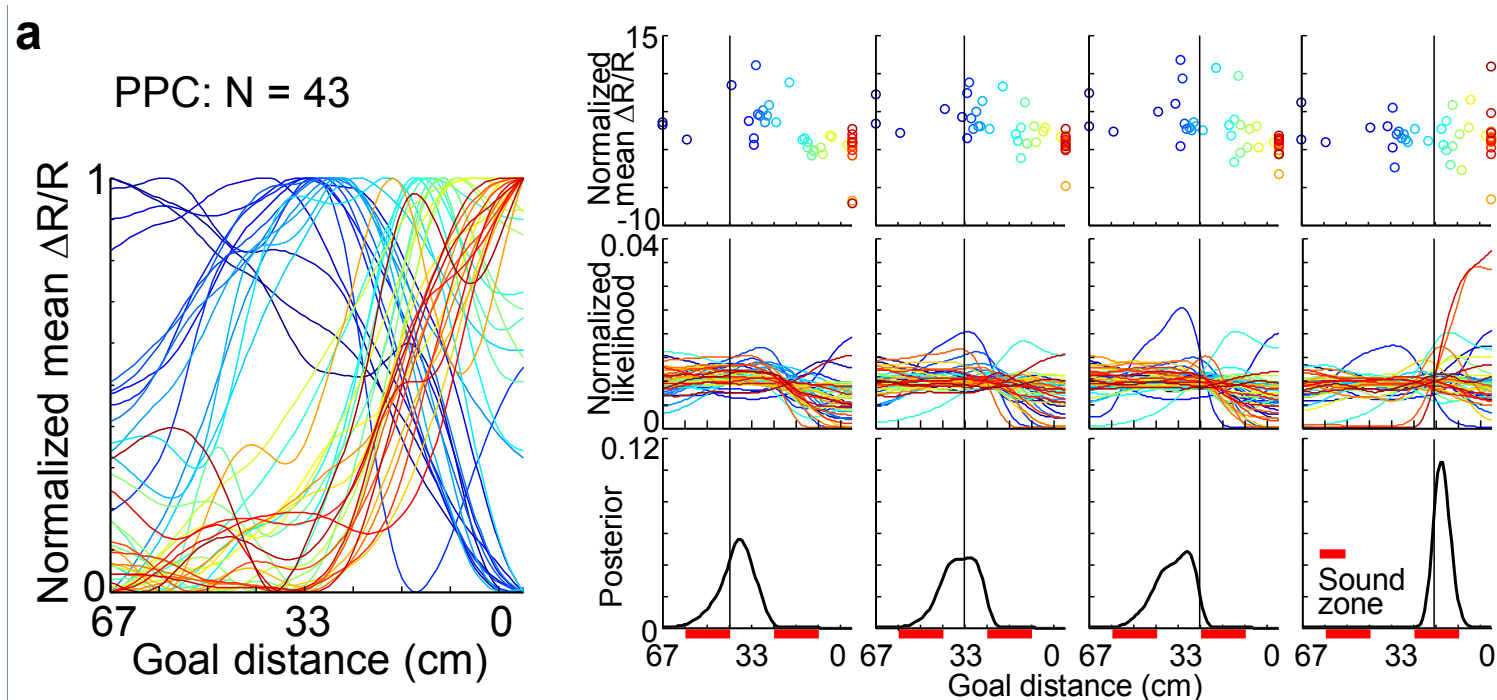


- 後頭頂葉 (PPC)
- 高次聴覚野 (PM)



ゴール距離のデコーディング

- 各ニューロンのゴール距離 x での活動 f_i
 - 応答関数 $p(f_i|x)$
- 距離 x のベイズ推定: $p(x|f_1, \dots, f_N) \propto \prod_i p(f_i|x)p(x)$



- ニューロンの示すゴール距離は無音区間でも変化

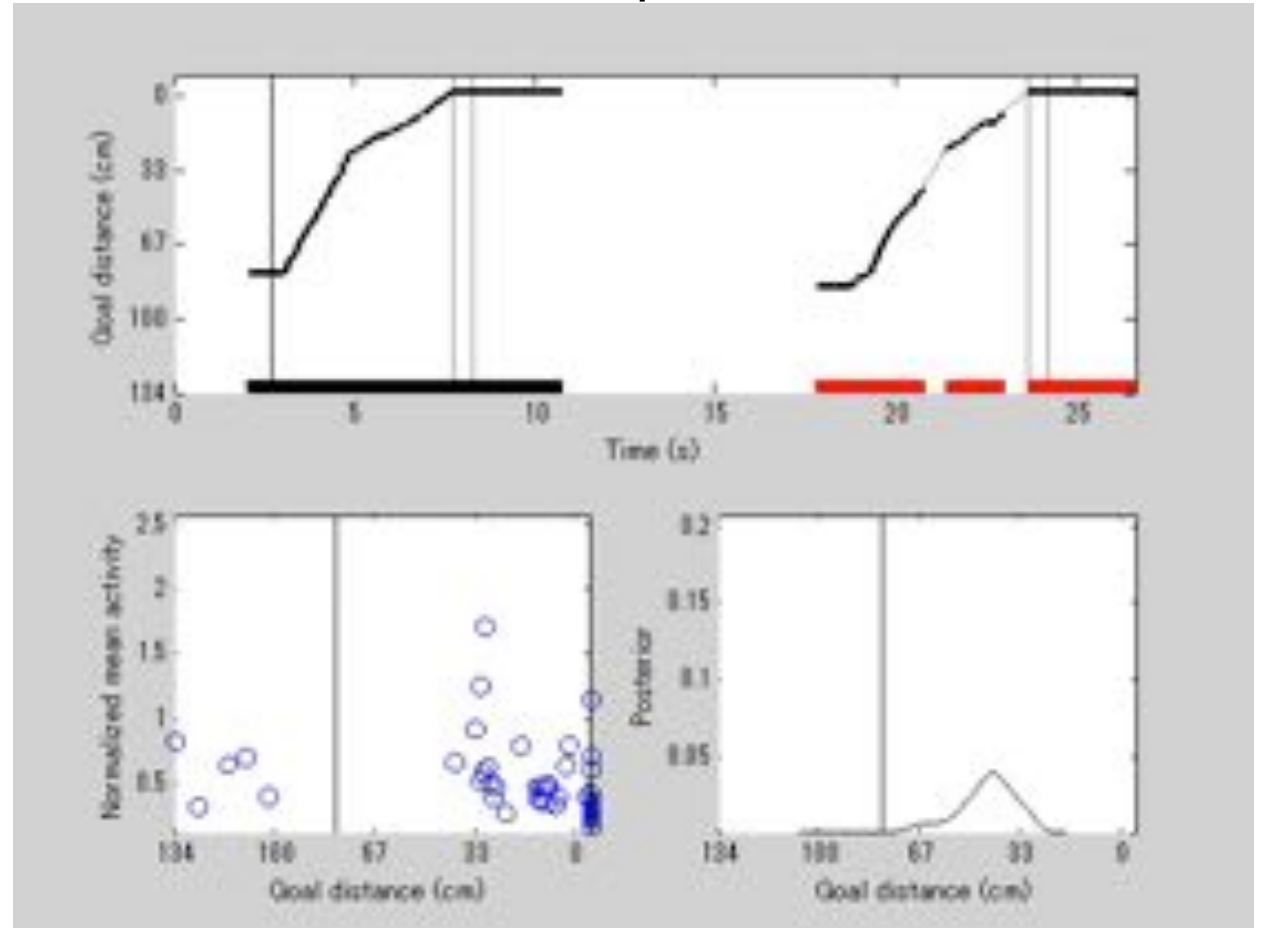
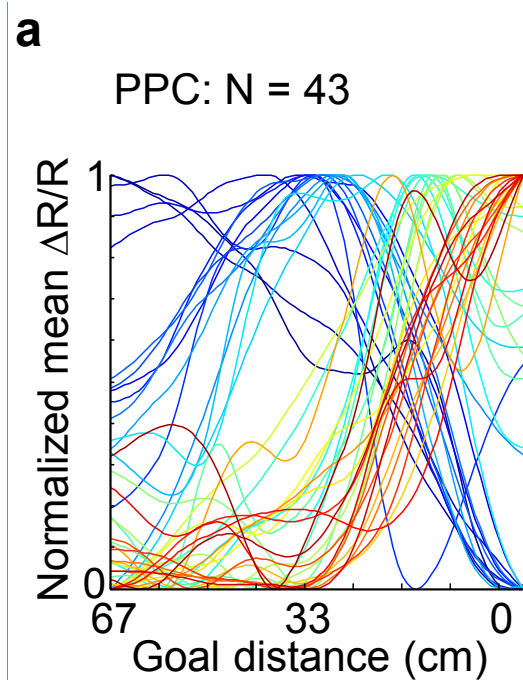


ゴール距離のデコーディング

■ 各ニューロンのゴール距離 x での活動 f_i

● 応答関数 $p(f_i|x)$

■ 距離 x のベイズ推定:



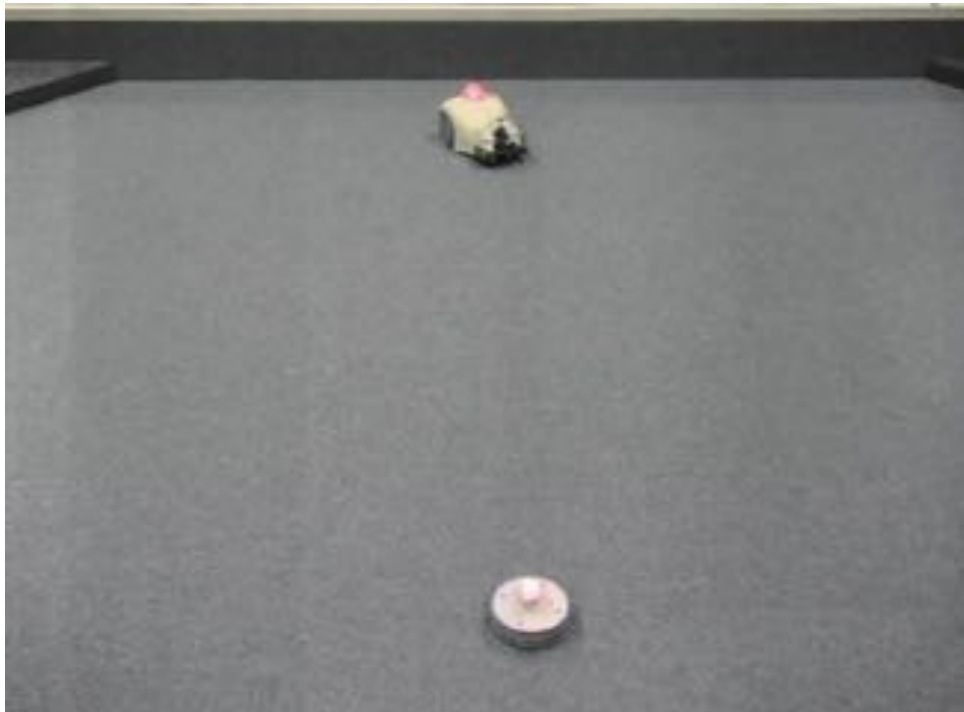
● ニューロンの示すゴール距離は無音区間でも変化



将来報酬の割引率 γ

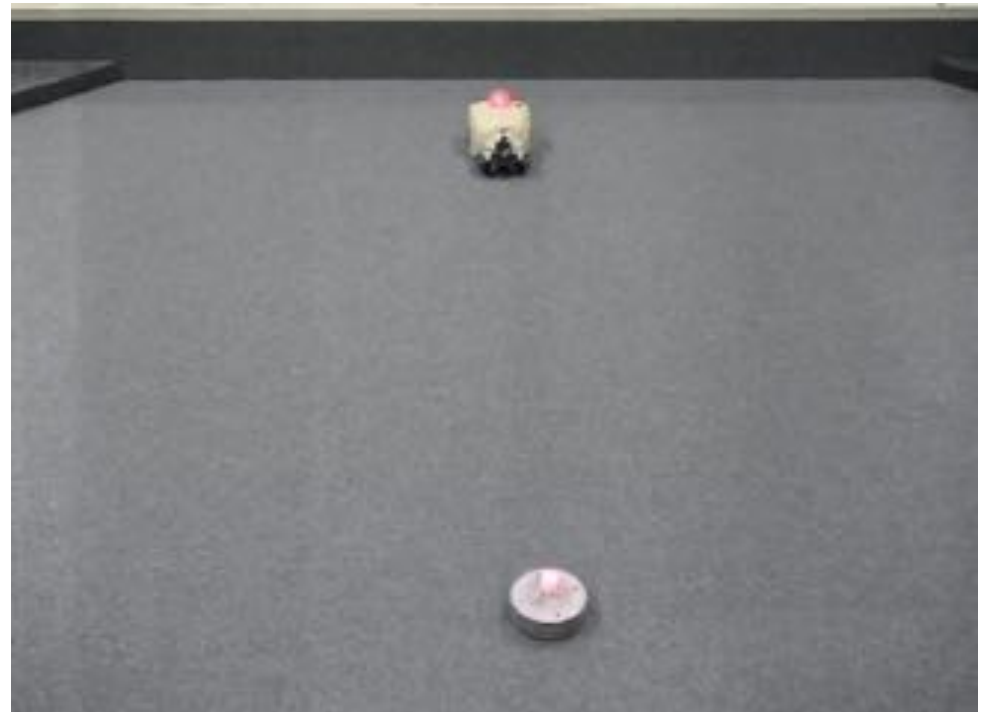
■ 大きな γ

- 遠くの電池も取りに行く



■ 小さな γ

- 近くの電池しか取らない

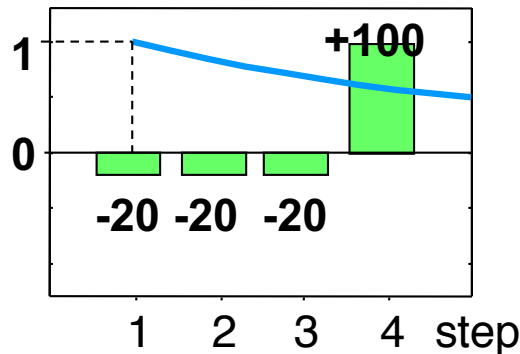


報酬予測の時間スケール

$$V(t) = E[r(t) + \gamma r(t+1) + \gamma^2 r(t+2) + \dots]$$

割引係数 γ : どれぐらい先の報酬まで予測するか

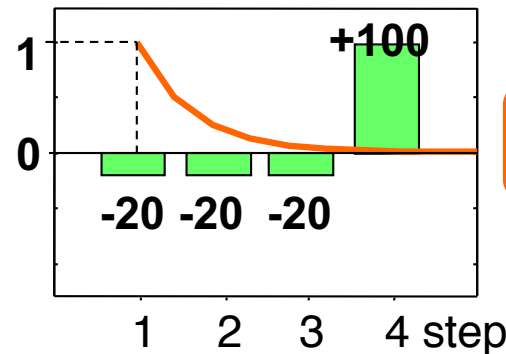
γ 大



苦勞してもやる!

$$V = 18.7$$

γ 小



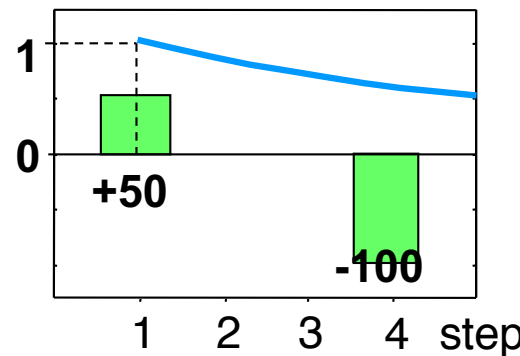
うつ病?

やらないほうがまし

$$V = -25.1$$

危ないことはやらない

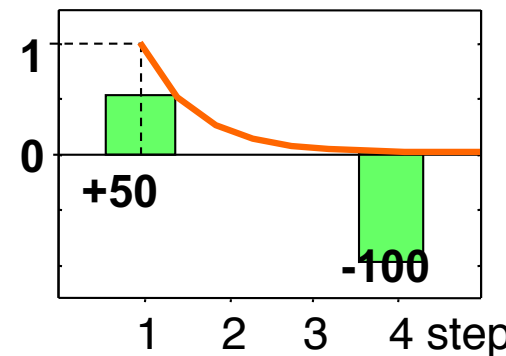
$$V = -22.9$$



衝動性?

ついやってしまう

$$V = 47.3$$



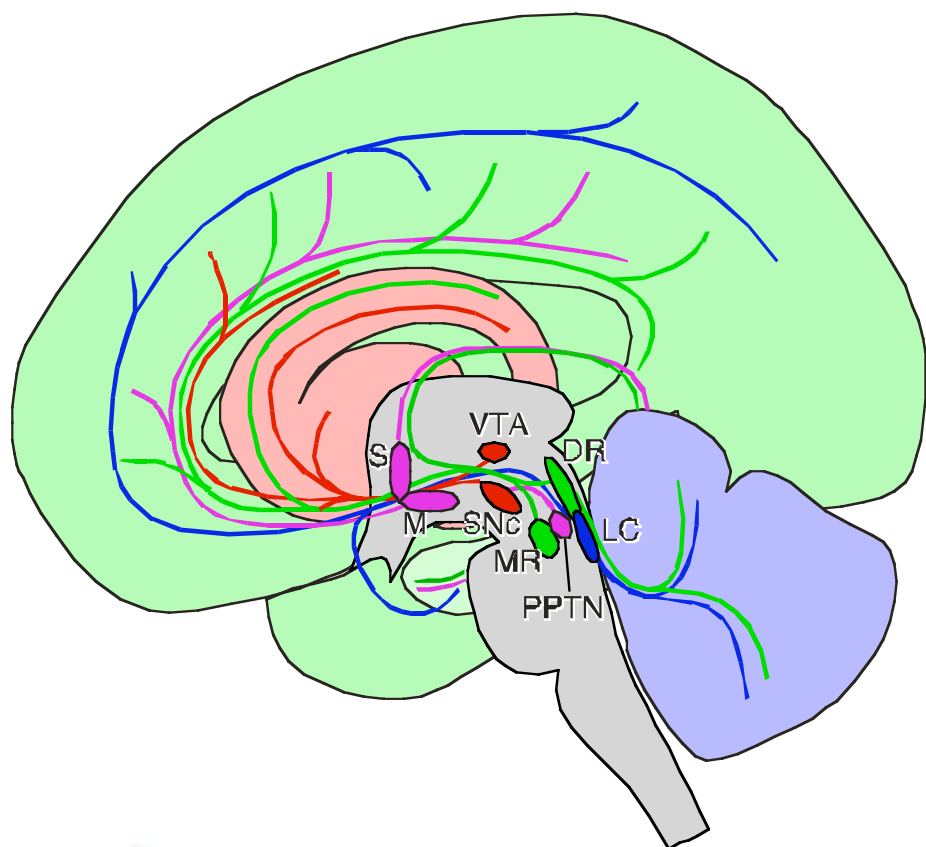
セロトニン?



神経修飾物質系とメタ学習

(Doya, 2002)

- 脳幹から小脳，基底核，大脳皮質に広範に投射
 - 学習の大域変数／パラメタを制御？



ドーパミン: 報酬予測誤差 δ

アセチルコリン: 学習速度 α

ノルアドレナリン: 探索の絞り β

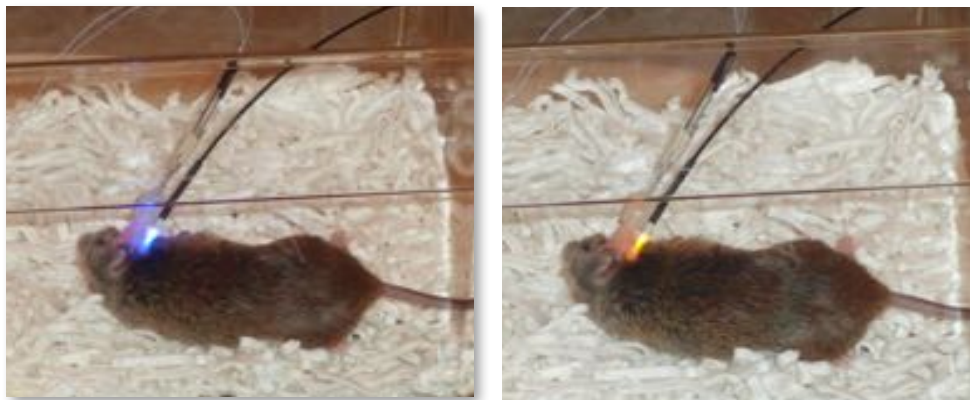
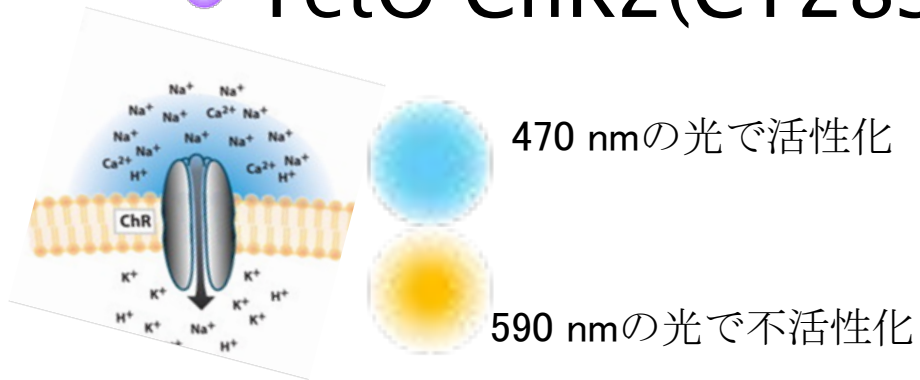
セロトニン: 予測の時間スケール γ



セロトニンニューロン選択的ChR2マウス

(生理研 山中章弘先生、田中謙二先生)

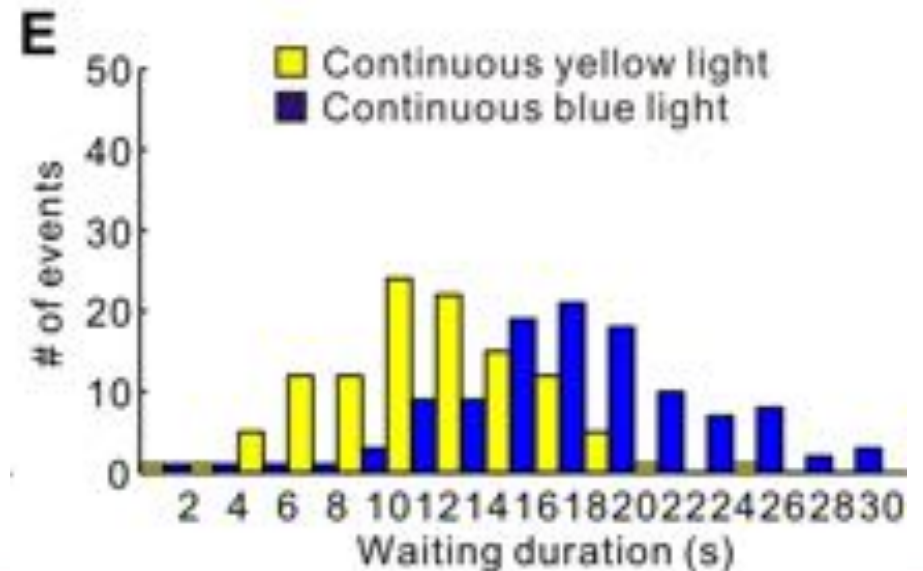
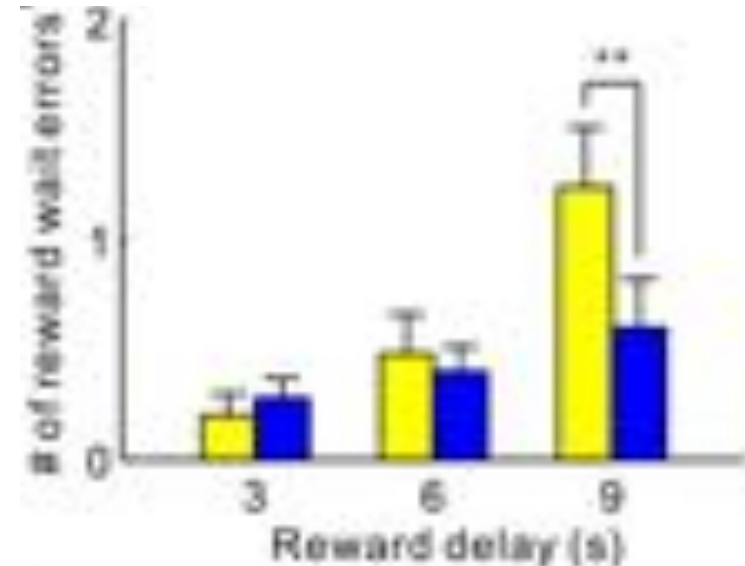
- tryptophan hydroxylase-2 (Tph2) promoter-tTA
- TetO-ChR2(C128S)-EYFP (Tanaka et al., 2012, Cell Reports)



エサ待ち 3, 6, 9, ∞ 秒



- 3秒
- エサなし: 12.1 s
- エサなし: 20.8 s

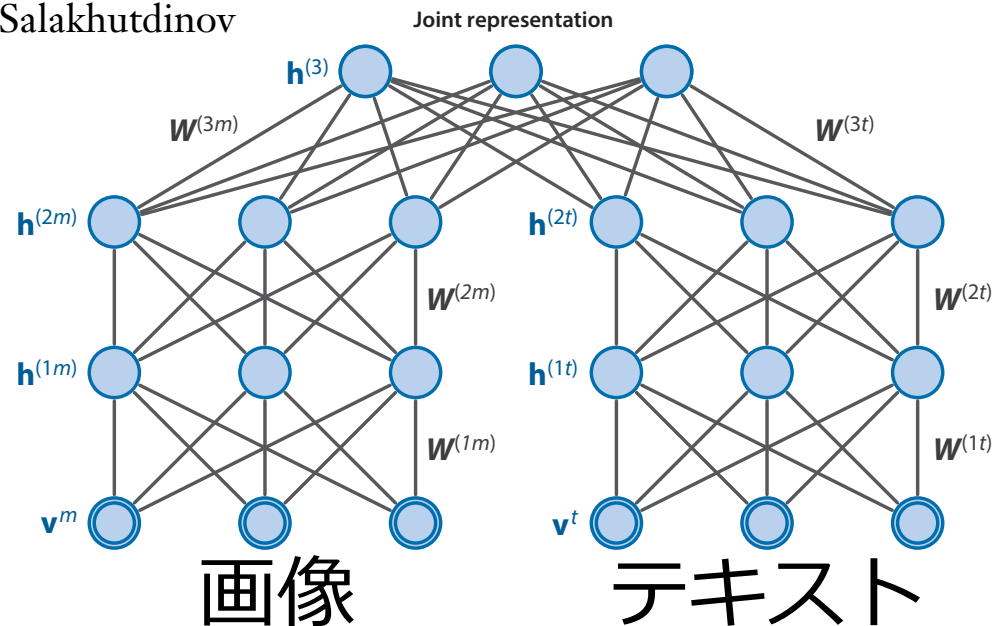


深層生成ネットワーク

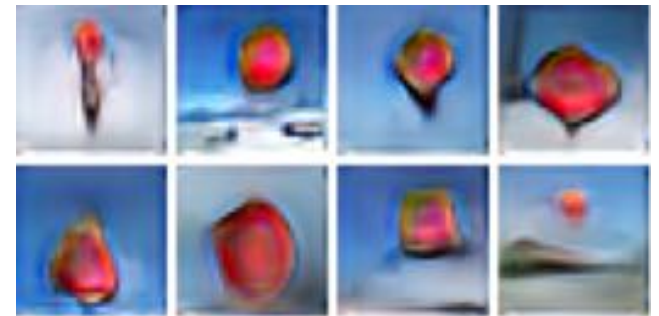
Learning Deep Generative Models

(Salakhutdinov, 2015, Ann. Rev. Stat.)

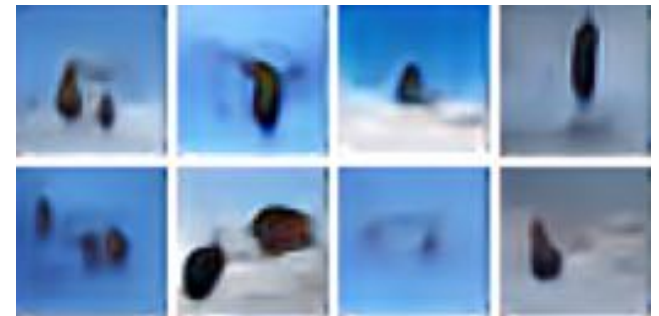
Ruslan Salakhutdinov



A **stop sign** is flying in blue skies



A **herd of elephants** is flying in blue skies



- 画像とテキストに共通の高次表現
- テキストからの画像生成



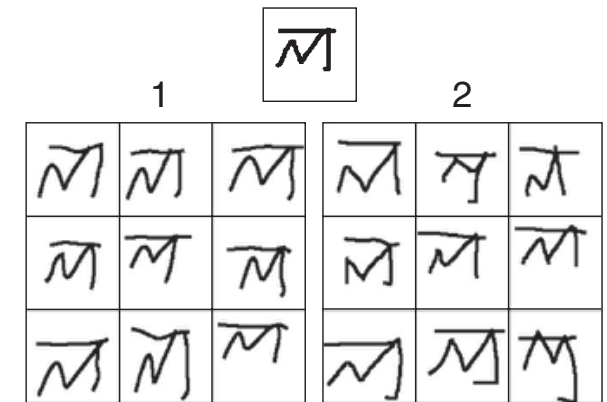
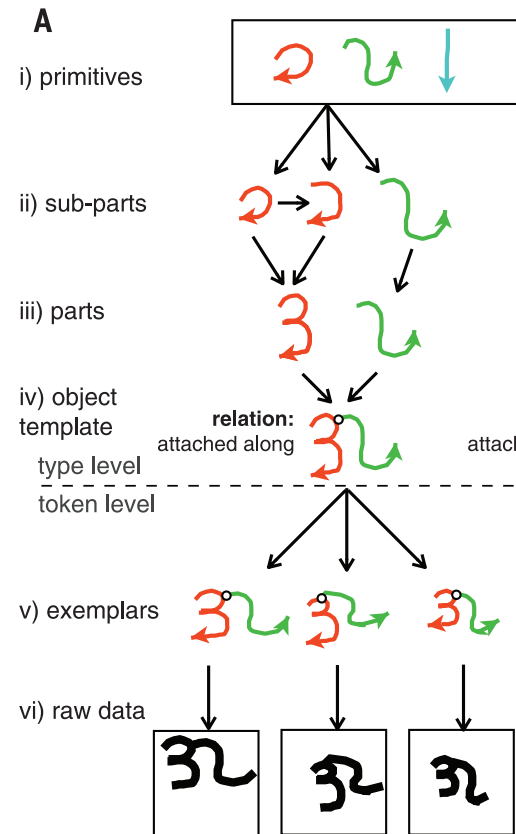
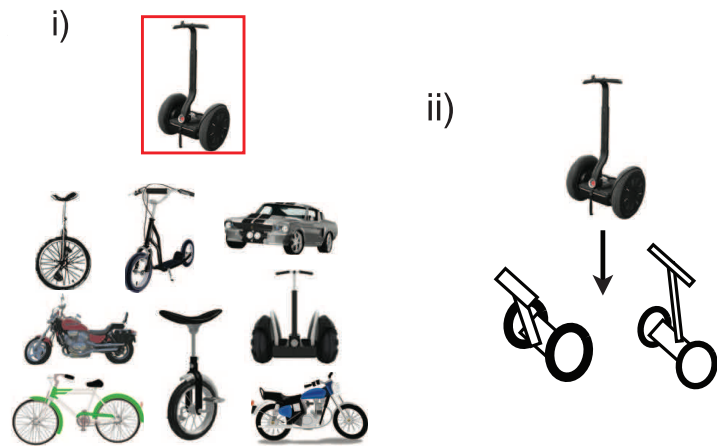
確率的プログラミング

RESEARCH ARTICLES

COGNITIVE SCIENCE

Human-level concept learning through probabilistic program induction

Brenden M. Lake,^{1*} Ruslan Salakhutdinov,² Joshua B. Tenenbaum³



Human or Machine?

■ 合成可能性、因果性、学習しかたの学習

(Lake et al., 2015, Science)



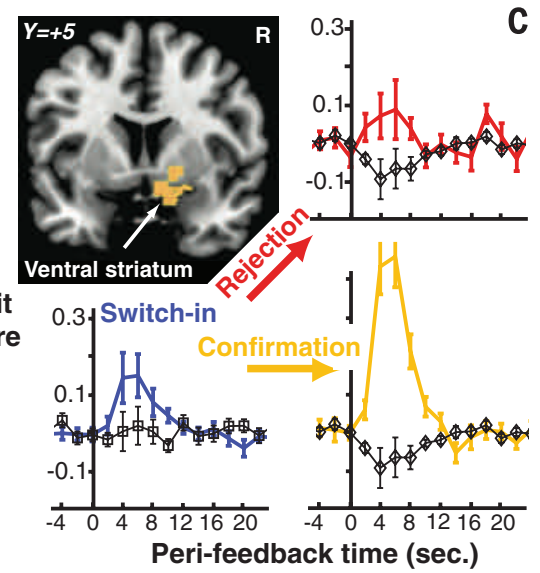
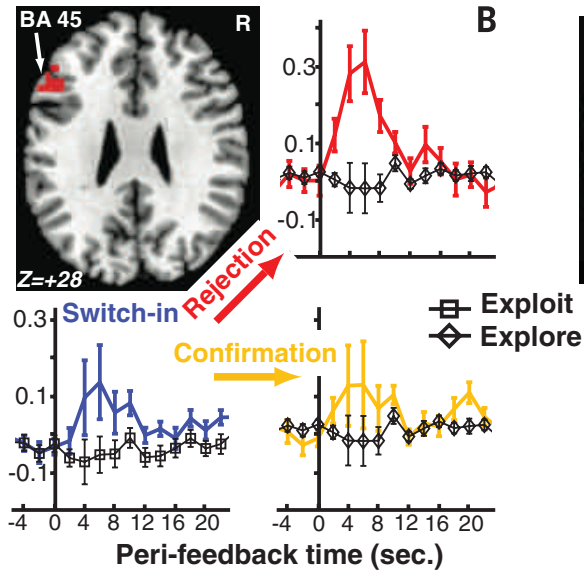
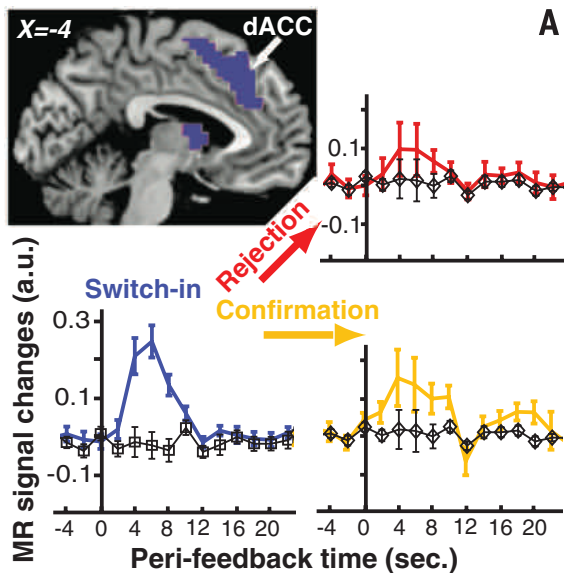
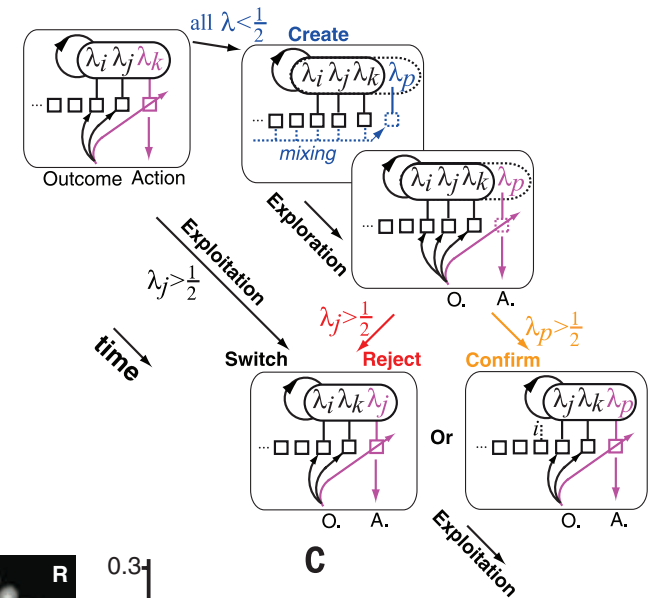
仮説生成／検証／棄却の脳回路

RESEARCH ARTICLE

HUMAN COGNITION

Foundations of human reasoning in the prefrontal cortex

Maël Donoso,^{1,2,3} Anne G. E. Collins,^{2,4} Etienne Koechlin^{1,2,3,*}

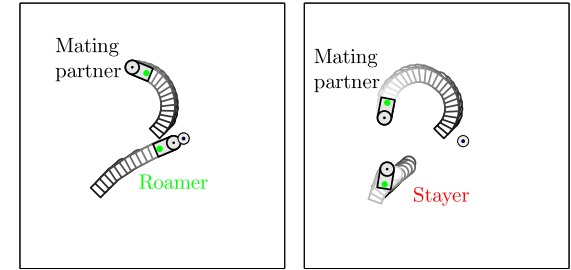
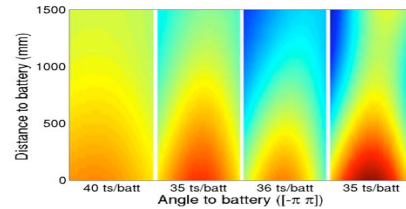


(Donoso et al., 2014, Science)

脳と現在の人工知能のちがい

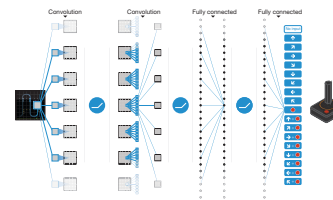
自主性：自ら生きるものと作られたもの

- サイバーローデント



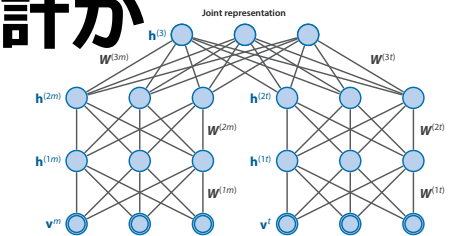
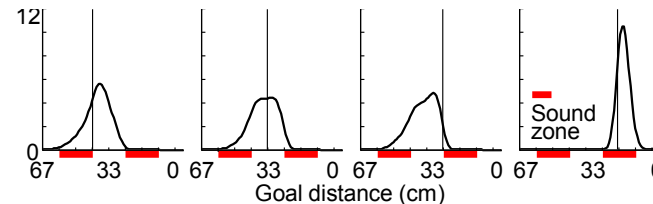
情報選択：必要な情報を自ら見つけるか、与えられるか

- ディープラーニング



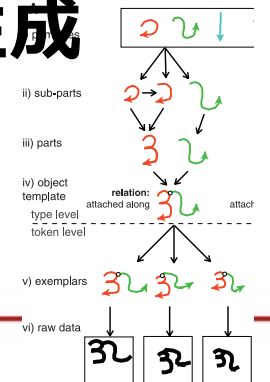
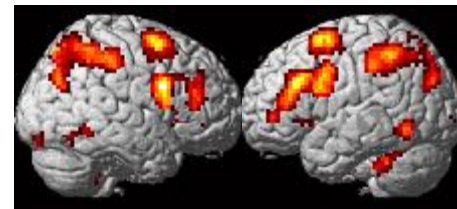
能動性：仮説生成と検証による理解か、データ統計か

- 深層生成ネットワーク



モジュール自己組織化：必要に応じた選択、結合、生成

- 確率的プログラミング



脳のパジュール接続問題

- 必要なパジュールがいかに選ばれるのか？
 - central executive?
 - meta cognition?
 - peer-reviewing?
- 脳のスイッチボード？
 - 視床? 基底核?
 - スパイク／脳波の同期?
- 個々にに学習した表現がなぜつながってしまうのか？
 - 辞書／プロトコルフリーの通信?



新学術領域研究(研究領域提案型) 人工知能と脳科学の対照と融合

Correspondence and Fusion of Artificial Intelligence and Brain Science

ナビゲーション

ホーム

お知らせ

領域の概要

研究組織

イベント案内

ニュースレター

研究成果

コンタクト

Navigation

Home

News

Project Outline

Organization

Events

Newsletter

Publications

Contact

ホーム

科学研究費・新学術領域研究(研究領域提案型)「人工知能と脳科学の対照と融合」(代表・副谷賢治) 領域略称名:人工知能と脳科学 のホームページです。

最新のお知らせ

新学術領域「人工知能と脳科学」キックオフシンポジウム+公募説明会 H26年度にスタートとした新学術領域研究「人工知能と脳科学」のキックオフシンポジウムおよびH29-30年度の公募研究についての説明会を、下記のとおり開催いたします。—日時:9月13日(火)会場:東京大学工学部2号館1階 212号講義室参加登録:参加予定の方はこちらから登録をお願いいたします。—13:30 開会13:35 人工知能と脳科学の共進化:副谷賢治(沖縄科学技術大学院大学)14:00 ディナーラーニングはどこまで行くのか:松尾豊(東京大学…

Posted 17 hours ago by Kenji Doya

新学術領域「人工知能と脳科学」+ポスト京「脳と人工知能」合同ワークショップを開催しました。 新学術領域「人工知能と脳科学」+ポスト京「脳と人工知能」合同ワークショップを開催しました。 —日時:8月11日(木)~8月12日(金)場所:湘南国際村センター、第2研修室 <http://www.shonan-village.co.jp> —8/11(祝)10:15 代表挨拶、参加者紹介(副谷賢治)…

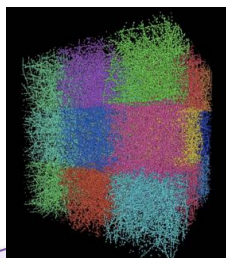
Posted 20 hours ago by Kenji Doya

ホームページの設置 ホームページを設置しました。

Posted 17 Jul 2016, 19:30 by Tadahiro Taniguchi



ポスト「京」研究開発萌芽領域的課題④
 思考を実現する神経回路機構の解明と人工知能への応用
**脳のビッグデータ解析、全脳シミュレーションと
 脳型人工知能アーキテクチャ**



京都大学
大羽成征

サブ課題A: 脳の構造と活動の大規模データ解析

京都大学
石井信

サブ課題E: 脳型人工知能アーキテクチャの開発

理化学研究所
五十嵐潤

サブ課題B: 大脳皮質神経回路のデータ駆動モデル構築

代表機関

沖縄科学技術大学院大学

研究代表者: 銅谷賢治
 サブ課題D: 大脳皮質・基底核・小脳モデル統合による全脳シミュレーション

理化学研究所
高橋恒一

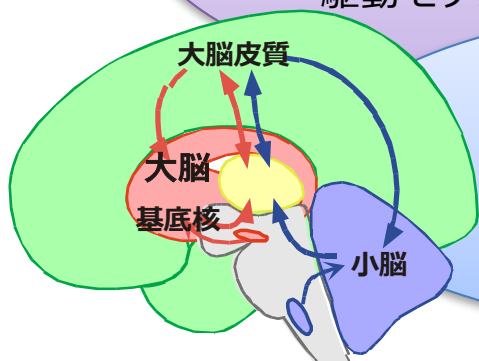
サブ課題F: 脳型人工知能大規模高性能計算プラットフォームの開発

電気通信大学
山崎匡

サブ課題C: ヒト全小脳モデル構築と大脳小脳連関シミュレーション

東京大学
原田達也

サブ課題G: 脳型人工知能の大規模実問題への応用



Juelich Research Center
Markus Diesmann
NEST開発

中部大学
稲垣圭一郎
視覚系モデル

九州工業大学
立野勝巳
海馬系モデル

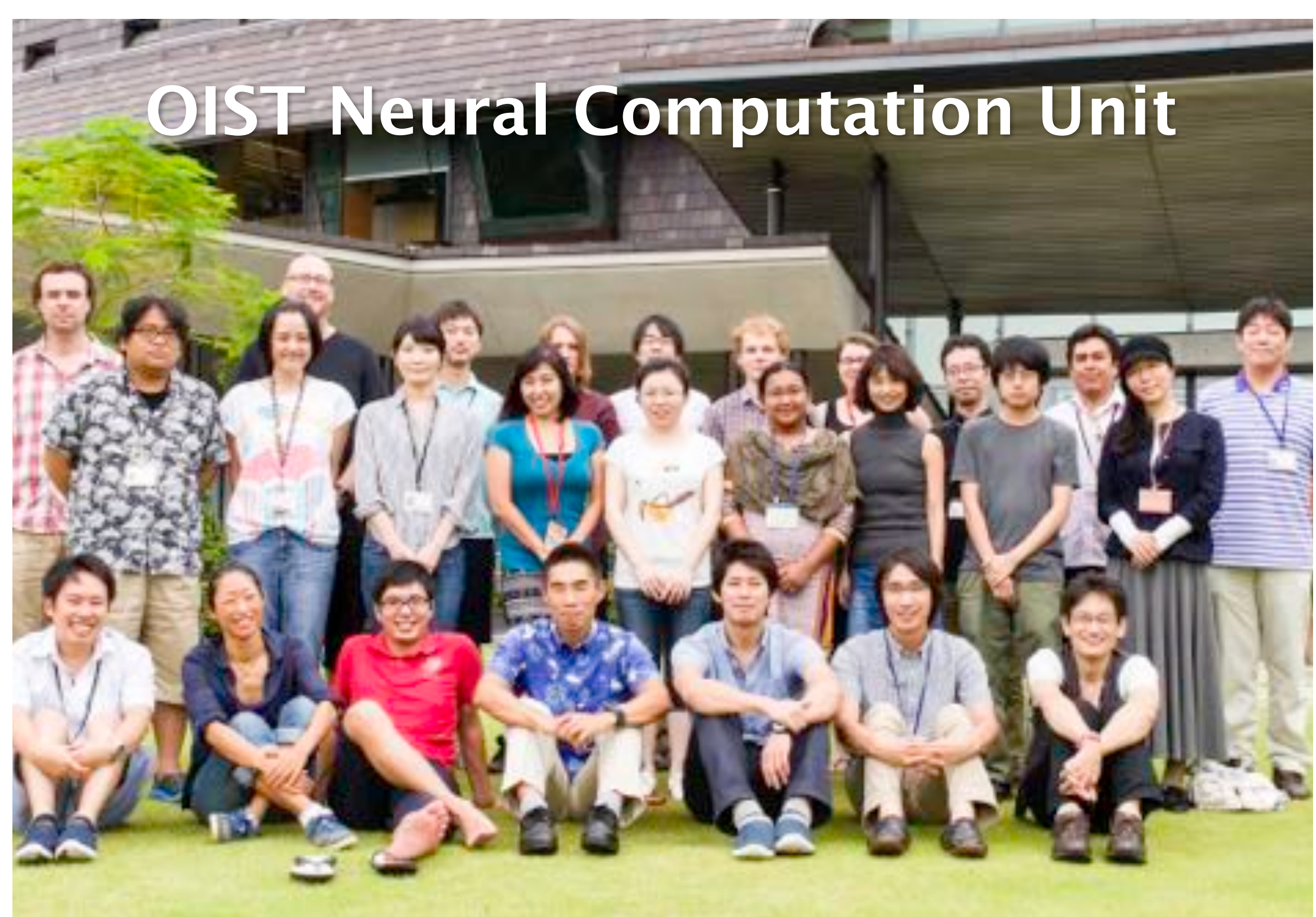
株式会社プリファード
ネットワーク
深層学習の応用

謝辞

- 強化学習ロボット
 - 内部英治
 - Stefan Elfving
 - Jiexin Wang
 - Paavo Parmas
 - 森本淳 (ATR)
- 行動価値/線条体
 - 伊藤真
 - 鮫島和行 (玉川大)
 - 木村實 (玉川大)
 - 上田康雅 (京都府立医大)
- 2光子計測
 - 船水章大
 - Bernd Kuhn
- セロトニン
 - 宮崎勝彦
 - 宮崎佳代子
 - 田中謙二 (慶應大)
 - 山中章宏 (名古屋大)
 - 濱田太陽
- fMRI
 - 清水優
 - 徳田智磯
 - 吉本潤一郎
 - Alan Fermin (玉川大)
 - 吉田岳彦
 - 田中沙織 (ATR)
 - Nicolas Schweighofer (USC)
 - 山脇成人 (広島大)
 - 川人光男 (ATR)
- 大脳基底核-視床-大脳皮質モデル
 - 庄野修 (HRI)
 - 五十嵐潤
 - Jan Moren
 - 高木周 (東大)
- 新学術領域研究「予測と意思決定」
- 新学術領域研究「人工知能と脳科学」
- ポスト京萌芽的課題「全脳シミュレーションと人工知能」
- 脳科学研究戦略推進プログラム
- 革新的技術による脳機能ネットワークの全容解明プロジェクト (革新脳)
- HPC戦略プログラム「予測する生命科学」



OIST Neural Computation Unit



新学術領域研究「人工知能と脳科学の対照と融合」若手サマースクール

人工知能と脳科学の新たな協働が起きています。人工知能と脳科学の研究者の対話と共同作業を深化させることで新発見を生み出すために、新学術領域「人工知能と脳科学の対照と融合」が発足しました（領域HP参照）。

本サマースクールは、人工知能と脳科学の第一線で活躍する研究者による講義とデモンストレーションを通じて、両者を融合した研究の新展開を担う人材育成を目指しています。ふるってご応募ください。

【期間】	2017年8月2日（水）～8月4日（金）；2日は昼ごころ集合、4日は夕方に解散の予定
【会場】	理化学研究所 脳科学総合研究センター 埼玉県和光市広沢2-1 最寄り駅：和光市（東武東上線／東京メトロ有楽町線・副都心線）
【講師】	<p>石井 信（京都大学）</p> <p>内藤英治（(株)国際電気通信基礎技術研究所（ATR））</p> <p>岡野原大輔（株式会社Preferred Networks（PFN））</p> <p>鹿島久嗣（京都大学・理化学研究所 革新知能統合研究センター）</p> <p>神谷之康（京都大学）</p> <p>田中啓治（理化学研究所 脳科学総合研究センター）</p> <p>鮎谷賢治（沖縄科学技術大学院大学）</p> <p>中原雅之（理化学研究所 脳科学総合研究センター）</p> <p>松尾豊（東京大学・理化学研究所 革新知能統合研究センター）</p> <p>森本淳（(株)国際電気通信基礎技術研究所（ATR））</p> <p>（五十音順；敬称略）</p> <p>※講義は日本語となります。</p>
【定員】	25名程度。 ※ 応募者多数の場合、応募資料により審査を行います。なおスクールの参加者の方の参加費は無料。原則、8月2日と3日の宿泊費と旅費のサポートあり（詳細は応募ページの併し書きをご覧ください）。
【応募資格】	主には大学院生、ポストドクター、大学院進学を考えている学部学生。また企業の研究者の方も若干名受入を予定。
【応募方法】	下記、申し込み応募のページから、申し込みください。

http://www.brain-ai.jp/summer_school2017 (4/28締切)



OIST

OKINAWA INSTITUTE OF SCIENCE AND TECHNOLOGY GRADUATE UNIVERSITY

第27回日本神経回路学会全国大会 JNNS2017

2017年9月20日(水)～22日(金)

北九州国際会議場

JR小倉駅から徒歩5分

投稿開始: 5月22日(月)

投稿締切: 6月19日(月)

採択通知: 8月4日(金)

大会長

九州工業大学大学院生命体工学研究科 教授 柴田 智広

基調講演

東京大学 教授、九州工業大学 客員教授 合原 一幸

企画シンポジウム

- ・新学術領域『多様な「個性」を創発する脳システムの統合的理解』シンポジウム
- ・新学術領域『人工知能と脳科学の対照と融合』シンポジウム
- ・新学術領域『スパースモデリングの深化と高次元データ駆動科学の創成』シンポジウム

<https://www.facebook.com/JNNS2017/>

