

# 大脳基底核の機能

# 大脳基底核

- 随意運動の制御系  
主に時系列情報の処理を担う
- 報酬予測系  
報酬に動機づけられた学習に関わる

# 大脳基底核

大脳皮質下に位置する神経核群

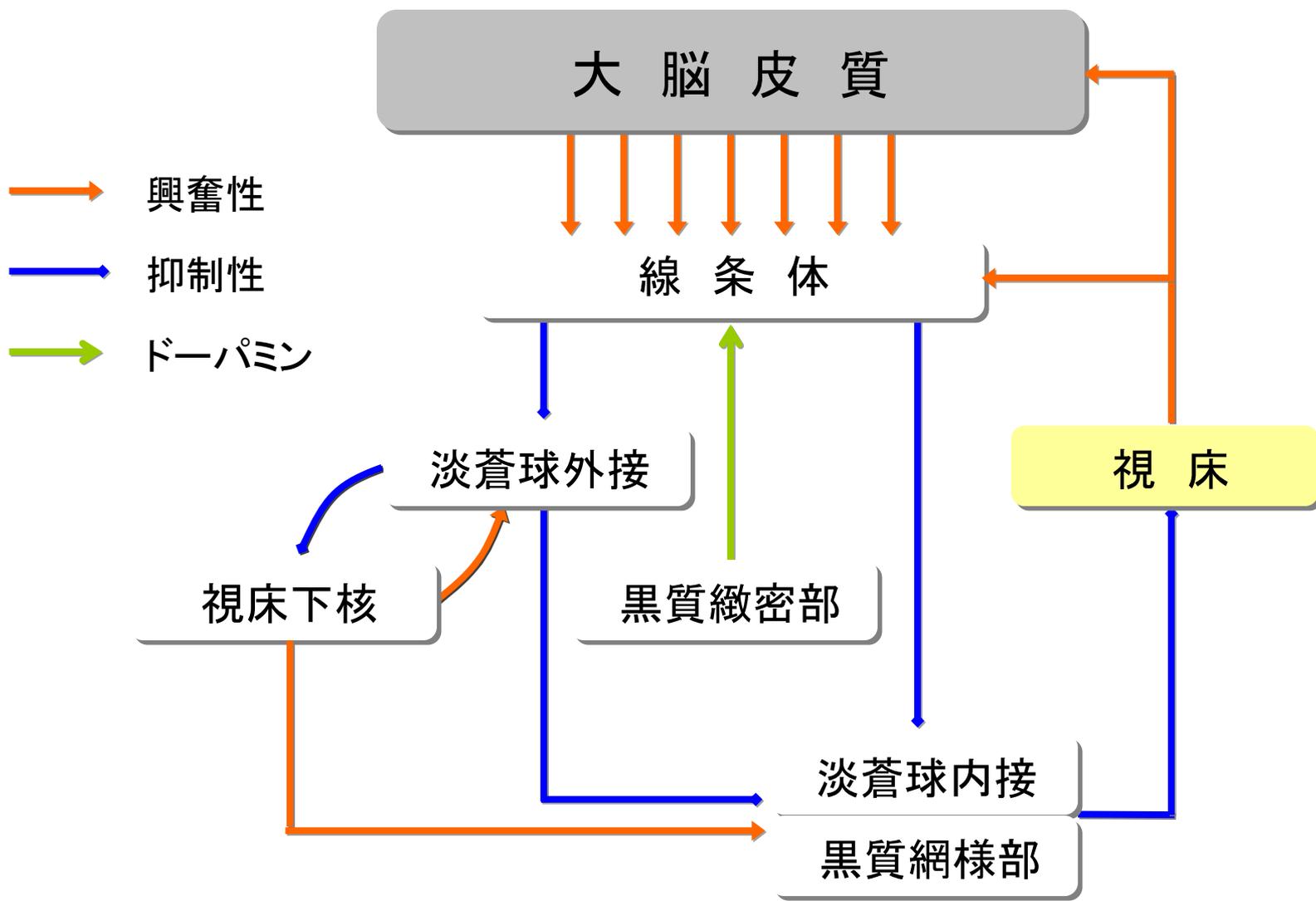
- 線条体(被殻、尾状核)
- 淡蒼球(外接、内接)
- 視床下核
- 黒質(緻密部、網様体)

図は省略: 下記の本を参照してください。

Kandel et al., "Principles of Neural Science", McGraw-Hill, 4th Edition, 2000, pp.855

# 大脳基底核

- 入力: 線条体に入る
  - 大脳皮質のあらゆる領野
  - 視床
- 出力: 淡蒼球内接・黒質網様体から出る
  - 視床
- 中継
  - 直接路(線条体→淡蒼球内接・黒質網様体)
  - 間接路(線条体→淡蒼球外接・視床下核→淡蒼球内接・黒質網様体)



# 大脳基底核の破壊症状

直接路と間接路のバランスが崩れることが原因

- 運動不全 (hypokinetic disorder)
  - 運動をしにくくなる
  - パーキンソン病など
- 多動障害 (hyperkinetic disorder)
  - 不随意運動が現れる
  - ハンチントン病など

# 大脳基底核の破壊症状

直接路と間接路のバランスが崩れる原因

- 黒質緻密部のドーパミン細胞の縮退  
治療法：L-DOPAの投与、脳深部刺激療法
- 線条体神経細胞の縮退

# 大脳基底核と報酬予測

## 神経修飾物質ドーパミン

- 黒質緻密部、腹側被蓋野(中脳)から投射される
  - ・黒質緻密部 → 線条体
  - ・腹側被蓋野 → 大脳皮質・前頭前野部、側座核
- 報酬に関連する情報を伝達する  
“快樂物質”と呼ばれる所以

# 大脳基底核と報酬予測

## Schultzらの実験(1997)

- オペラント条件付け学習を利用
  1. モニタ上のスポット刺激(条件刺激; CS)を呈示
  2. 刺激提示後、遅延期間は待機
  3. 遅延期間後、運動指示刺激が呈示されれば、できるだけ早くボタンを押す
  4. 一定時間以内に反応できれば、報酬が与えられる

図は省略：下記の本を参照してください。

**W. Schultz et al., Science 275, 1593 -1599 (1997)**

# 大脳基底核と報酬予測

## 実験結果の解釈

- 報酬:  $r(t)$
- 期待報酬:  $V(t)$
- 報酬予測誤差:  $\delta(t)$
- TD (temporal difference) 学習モデル:  
$$\delta(t) = r(t) + \gamma V(t) - V(t-1)$$
  
※ 強化学習の1モデル

# 強化学習

- ある環境下におかれたエージェントが、環境に対し起こした行動の結果として報酬(あるいは罰)を得、その環境からのフィードバックに基づき、より多くの報酬が得られるように行う学習
- 時刻 $t$ におけるエージェントの状態:  $s_t$
- 時刻 $t$ に環境から受ける報酬:  $r_t$
- 状態 $s_t$ の評価値:  $V(s_t)$   
(状態 $s_t$ を取ることにより、将来に期待できる報酬量)

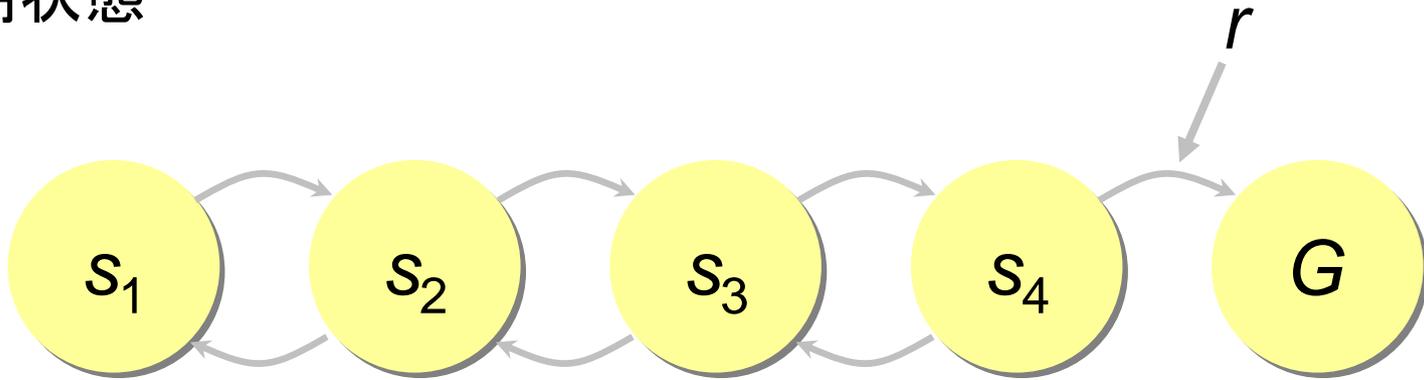
# TD学習のアルゴリズム

1. 時刻 $t$ において、エージェントが状態 $s_t$ を取る
2. ある行動を選択して状態 $s_{t+1}$ に移り、同時に報酬 $r_t$ を得る
3. TD誤差  $\delta_t$ を次の通りに計算  
$$\delta_t = r_t + \gamma V(s_{t+1}) - V(s_t)$$
4. 状態の評価値  $V(s_t)$ を次のとおり更新  
$$V(s_t) = V(s_t) + \alpha \delta_t$$
5. 更新された $V(s_t)$ に基づき、次の行動を選択

以下、繰り返し

# TD学習の流れ

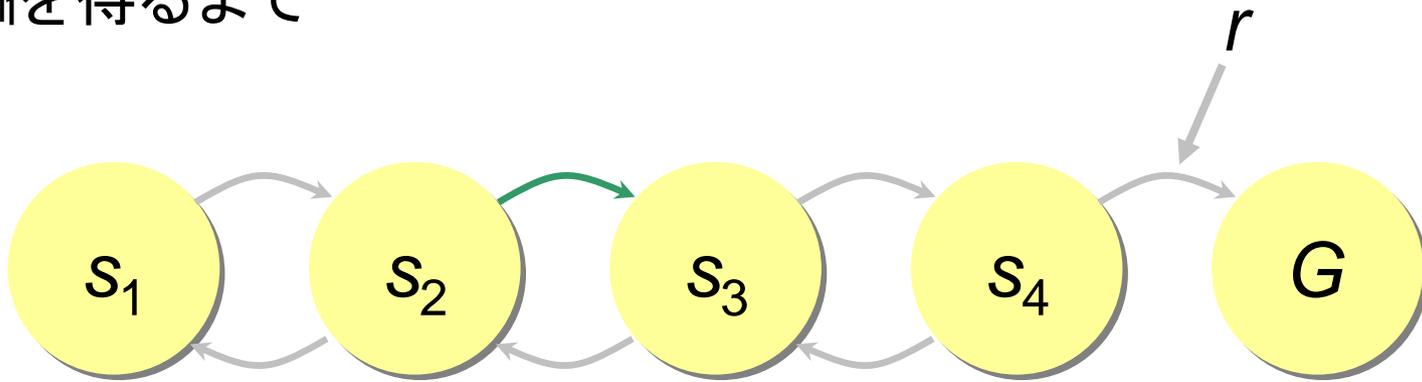
初期状態



$V(s_t)$       0            0            0            0            0

# TD学習の流れ

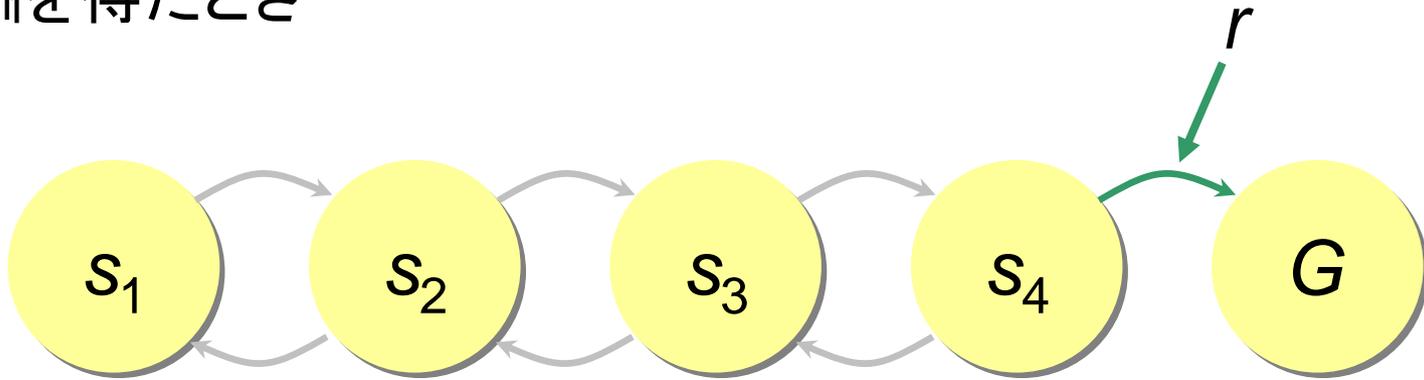
報酬を得るまで



$V(s_t)$     0            0            0            0            0

# TD学習の流れ

報酬を得たとき

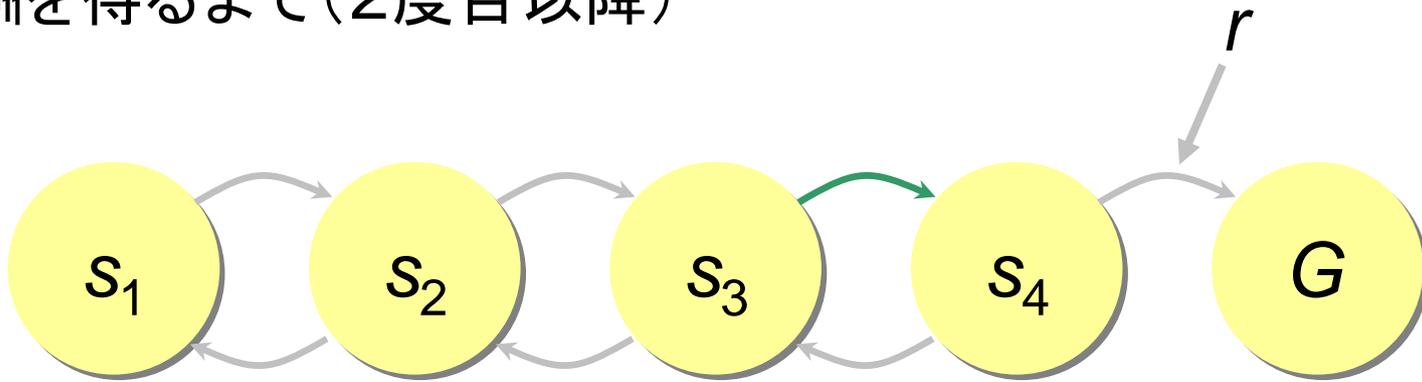


$V(s_t)$       0                  0                  0                   $r$                   0

$$\delta_t = r + \gamma \cdot 0 - 0$$

# TD学習の流れ

報酬を得るまで(2回目以降)

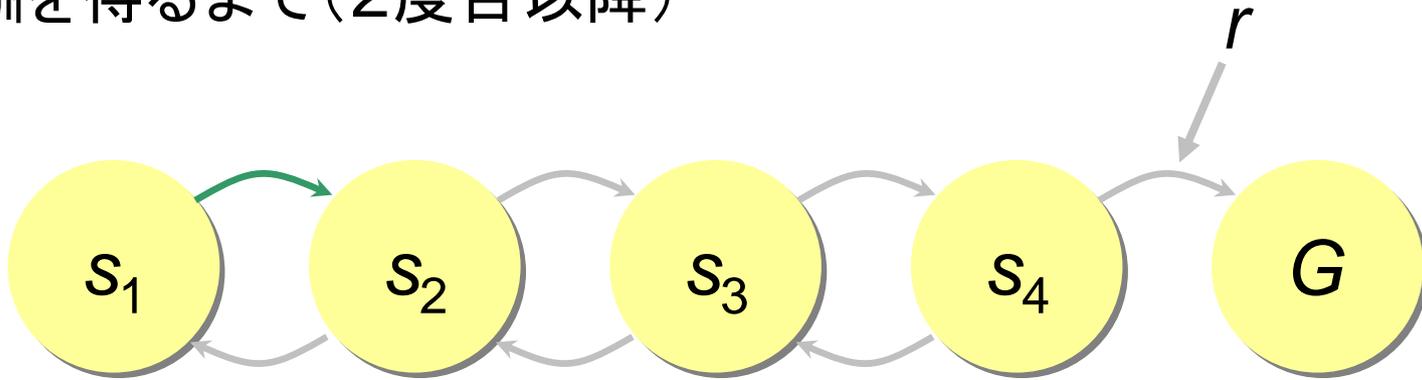


$V(s_t)$       0                  0                   $\gamma r$                    $r$                   0

$$\delta_t = 0 + \gamma \cdot r - 0$$

# TD学習の流れ

報酬を得るまで(2回目以降)



$$V(s_t) \quad \gamma^3 r \quad \gamma^2 r \quad \gamma r \quad r \quad 0$$

$$\delta_t = 0 + \gamma \cdot \gamma^3 r - 0$$

# 大脳基底核と報酬予測

## 実験結果の解釈

- 学習前:  $V(t) = V(t-1) = 0 \rightarrow \underline{\delta(t) = r(t)}$
- 学習後、報酬あり:
  - at CS:  $V(t) - V(t-1) > 0, r(t) = 0 \rightarrow \underline{\delta(t) > 0}$
  - at R:  $V(t) - V(t-1) < 0, r(t) > 0 \rightarrow \underline{\delta(t) = 0}$
- 学習後、報酬省略:
  - at CS:  $V(t) - V(t-1) > 0, r(t) = 0 \rightarrow \underline{\delta(t) > 0}$
  - at R:  $V(t) - V(t-1) < 0, r(t) = 0 \rightarrow \underline{\delta(t) < 0}$

# その他の報酬系

- 腹側被蓋野:ドーパミン投射の起点
- 側座核(腹側線条体):「やる気」、薬物依存症に関係する

内側前脳束:腹側被蓋野→側座核を経由する神経線維  
電気刺激により、強い快感をもたらす  
例)ロボラット(Talwarら、2002)