

B3-3 自律エージェントのためのストレス反応を有する情動行動学習システム

福井大学 工学部 知能システム工学科 進化ロボット研究室
加藤 進 (指導教員：前田 陽一郎)

1. 緒言

近年のヒューマノイドロボット開発は急速な進歩を遂げており、人間との間でコミュニケーションがとれるような人間らしさをもつロボットの研究が重要なテーマとなっている。人間同士のコミュニケーションで大切なことは心の触れ合いであるように、ロボットが人間とコミュニケーションをするためには、ロボットにも心にあたる知的機能を持たせる必要があるといえる。そのような知的機能の研究として、自己保存評価関数を用いた人工感情モデルを有する自律ロボット [1] や、振動子を用いた情動モデル [2] などが挙げられる。

これらのようにロボットの人工感情モデルは工学分野、医学分野などさまざまな分野からアプローチが行われている。本研究室においても、人工感情モデルにおける感情認識として人間の脳波情報を用いた情動推論実験 [3] や感情生成と感情表現を実現するために、情動、神経修飾物質系、学習システムのメタパラメタの関連性を考慮した情動行動学習システム構築 [4] などを行っている。

そこで本研究では、この情動行動学習システムの動作をより多様化させるために、ストレス反応とそれに関わる神経修飾物質系を加えたシステムを提案する。従来のシステムはシミュレータ上のエージェントが獲得したネガティブな行動に関する感性評価が低かった。このシステムは学習システムの罰をストレス刺激と判断した場合にネガティブな行動や情動を強めるため、それらを改善できると考えられる。システムの有効性を示すためにシミュレータ実験と感性評価実験を行ったので、この結果についても報告する。

2. ストレス反応と神経修飾物質系

ストレスという概念はいまいで不明確な要素があるが、ここでは個体のもつ通常の対処能力を超えた外的刺激に対する反応と考える。ストレスとなる刺激をストレス刺激、反応をストレス反応として説明する。

ストレス刺激になりうる要因は生物が生活する上で至るところに存在しているが、それらすべてに対してストレス反応を起こしているわけではない。これは脅威に対して生まれたときから持っている反応、言い換えれば遺伝的にプログラムされた反応と、生まれてから学習した脅威に対する反応があるからである。ストレス刺激に反応するかどうかを決める内部状態には、情動、欲求、身体状態、そして学習から得た記憶などがあげられる。このような内部状態によってストレス反応はその時々において心理的、本能的評価を受けて同じ刺激に対しても異なった反応を示すことが明らかになっている。

ストレス反応時の体内での変化には明らかにされていないことが多々あるが、脳から副腎皮質に向けてコルチゾールを分泌させる命令が出ることが分かっている。

そして命令を受けた副腎皮質ではコルチゾールが分泌され、血糖値の上昇、免疫力の低下などの作用を引き起こす。またコルチゾールは脳の海馬にある受容体にも到達する。海馬は主に記憶に関する機能を持っていることで有名だが、コルチゾールが結合する受容体を持ち、コルチゾールを抑制する命令を出すことが明らかになっている。

3. ストレス反応を有する情動行動学習システムの提案

秋口らの提案した情動行動学習システム [4] をもとにストレス反応を組み込んだ情動行動学習システムを提案する。図 1 においてコルチゾール、蓄積、抑制命令が本研究で新たに組み込まれた箇所である。

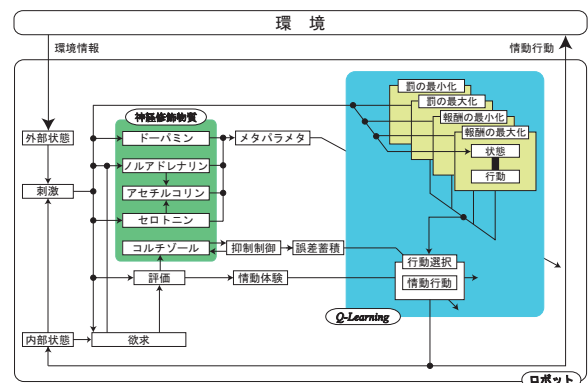


図1 本研究で提案する情動行動学習システム

本研究におけるストレス反応を起こすストレス刺激は、秋口らが提案した目標選択型 Q-Learning [4] で提案された罰に限定して用いた。この「罰」に対してストレス反応を起こすかどうかを決める内部状態に情動を用い、怒り、恐れが他の情動より強く出ているれば刺激に対して反応すると考える。これらをふまえたストレス反応値 R_t の式を以下に示す。

$$R_t = \frac{(I_{P-} + I_{N+}) - (I_{P+} + I_{N-})}{I_{P+} + I_{P-} + I_{N+} + I_{N-}} p_t \quad (R_t \geq 0) \quad (1)$$

I_{P-} 、 I_{N+} 、 I_{P+} 、 I_{N-} はそれぞれ怒り、恐れ、喜び、安らぎをあらわし、 p_t が罰の値である。ストレス反応によって分泌されるコルチゾールの量は任意の定数 k を用いて kR_t とする。またこのコルチゾールを抑制する海馬の役割をする機構としてフィードバック機構のPID制御を用いる。これは文献 [4] の手法における神経修飾物質系の分泌にも使われているので、同じ内分泌系の制御として統一した。この制御で取得される抑制値 X_t によって抑制できなかった誤差 D_t を、エージェントが

ストレス反応によって受ける身体的なストレス負担 S_t として蓄積させる。

$$D_t = kR_t - X_t \quad (2)$$

$$S_t = \int_{t-m}^t D_t dt \quad (3)$$

S_t は過去 m ステップから現在のステップ数 t までの D_t の積分で得られる。このストレス負担には身体的な許容量としての閾値 b_1 、 b_2 を設定し、閾値を越えた場合の行動、内部状態に変化をつける。変化は3段階用意し、負担にならない状態、負担によって障害が発生し始める状態、負担が身体に害を及ぼしている状態とした。以下にそれらの状態による主な動作変化を示す。

表1 ストレス負担による行動変化

ストレス負担	行動変化
$0 \leq S_t \leq b_1$	すくみ、逃避、攻撃、ノルアドレナリンの一時的な上昇
$b_1 < S_t \leq b_2$	パニックによるランダム行動、セロトニンの分泌量の低下
$b_2 < S_t$	動作の遅滞、逃避、ランダム行動、セロトニン分泌量、欲求の低下

4. シミュレーションおよび感性評価実験

前章で提案したストレス反応を有する情動行動学習システムを組み込んだシミュレータを用いて感性評価実験を行った。

4.1 実験方法

実験で用いたシミュレータ内環境を図2に示す。このような環境を用意し、エージェントに情動行動を獲得させた。実験の結果を被験者に提示してアンケートを行った。アンケートには16の情動行動パターンを用意し、11人の成人男性に動作を見てもらい、感じた情動4つに対して1~7で評価してもらった。アンケート結果は κ 検定を用い、用意した情動行動と被験者の評価との一致を判定した。 κ の一致度は Landis と Koch らによると、 $0 \leq \kappa \leq 0.2$ はごく軽度の一致、 $0.2 < \kappa \leq 0.4$ は軽度の一致、 $0.4 < \kappa \leq 0.6$ は中等度の一致、 $0.6 < \kappa \leq 0.8$ は高度の一致、 $\kappa > 0.8$ はほぼ完璧な一致とされている。

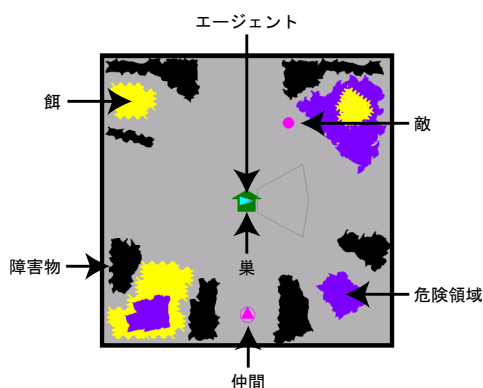


図2 情動行動学習システムのシミュレータ

4.2 実験結果

被験者のアンケート結果の平均値を表2に示す。表中の下線部は、その行動における被験者の平均値の最大値であり、 κ 係数の計算には最大値となる情動を人

間の判定結果として用いる。また、動作番号横に記載されているのは当初動作が表現しようとした情動を表わす。

表2 アンケートによる感性評価結果

動作	アンケート結果平均値			
	喜び	怒り	恐れ	安らぎ
動作1(怒り)	5.0	<u>5.5</u>	3.8	2.6
動作2(恐れ)	4.1	3.5	<u>4.2</u>	4.0
動作3(喜び)	<u>4.0</u>	3.2	3.8	3.8
動作4(安らぎ)	<u>4.0</u>	2.7	3.3	3.8
動作5(喜び)	2.9	1.9	<u>5.6</u>	3.0
動作6(怒り)	3.5	3.3	<u>4.3</u>	2.4
動作7(恐れ)	4.0	1.9	4.0	<u>4.4</u>
動作8(怒り)	2.2	<u>4.6</u>	4.4	3.4
動作9(恐れ)	4.1	1.7	<u>4.2</u>	3.2
動作10(安らぎ)	2.9	2.7	4.3	<u>5.6</u>
動作11(怒り)	4.0	<u>4.1</u>	3.7	3.2
動作12(喜び)	2.6	<u>5.1</u>	3.1	2.6
動作13(安らぎ)	2.9	3.3	<u>4.7</u>	2.0
動作14(恐れ)	1.5	3.5	<u>5.7</u>	3.2
動作15(喜び)	2.5	<u>5.0</u>	4.2	2.4
動作16(安らぎ)	4.0	3.5	3.0	<u>4.9</u>

これらの判定結果より、 κ 係数の計算を行った結果を表3に示す。今回のアンケート結果では κ 係数は0.417となり、中等度の一致を示した。

表3 κ 係数の計算結果

		人間の評価				合計	レート
		喜び	怒り	恐れ	安らぎ		
システム の評価	喜び	1	2	1	0	4	0.25
	怒り	0	3	1	0	4	0.25
	恐れ	0	0	3	1	4	0.25
	安らぎ	1	0	1	2	4	0.25
	合計	2	5	6	3	16	1
レート		0.125	0.3125	0.375	0.1875	1	-

5. 結 言

今回の実験で得た感性評価は秋口らの実験の一致度より低い値であったが、ネガティブな行動に対する評価は一致度が高かった。これらのことより、ネガティブな行動に関してストレス反応の動作が有効であったが、逆にポジティブな動作を不明確なものにしてしまったという問題も残された。

今後の課題としては、ストレス負担を決める閾値を自律的に獲得し、ポジティブな動作とネガティブな動作に明確な差が出せるような情動行動も実験に加える必要があると考えられる。

参考文献

- [1] 尾形哲也, 菅野重樹, " 情動モデルを有する自律ロボット WAMOEBA-2 と人間との情緒交流," 日本機械学会論文集, 第65巻, 第633号C編, pp.166-172 (1999)
- [2] 加藤祥史, 竹内義則, 松本哲也, 工藤博章, 大西昇, " 振動子を用いた情動モデルとそのロボットへの応用," 信学技報, pp.43-50 (2001)
- [3] 本間雄仁, 前田陽一郎, " 脳波情報を用いた情動計測システムの構築およびファジィ情動推論実験," 第23回ファジィシステムシンポジウム, CD-ROM WD3-4 (2007)
- [4] 秋口俊輔, 前田陽一郎, " 神経修飾物質系に基づく情動行動学習システムを用いた動物型エージェントによる感性評価実験," 第23回ファジィシステムシンポジウム, CD-ROM WD3-3 (2007)