

三次元空間におけるサルとバナナ問題のヒューリスティックな学習実験

旭川工業高等専門学校 ○中澤 大輔 渡辺 美知子 古川 正志

要旨

サルとバナナ問題は、AI分野でテスト問題とされ、最近では久保や鈴木らにより「山登り動機付け法」及び、フェロモンを用いた解法が示されている。本研究では、問題を三次元に拡張し、2匹のサル間の好意度を制御し集合の行動を予測するヒューリスティックな方法を提案し、その実験結果を示す。

1. はじめに

サルとバナナ問題は、AI分野では計算機の演算速度の測度となるテスト問題として扱われている。この問題は、地面にサルが存在し、同時に天井からバナナがぶら下がっている限られた空間を想定し、その空間内において、サルがバナナを入手しようとするものである。バナナはサル一匹よりも高い位置に存在し、サルは他のサルに登らねば(協力しなければ)バナナを入手することができない。各サルはそれぞれの目標とするバナナを手に入れようと自律的に行動する。これは、複数のエージェントが各々の目的地へいかに到達するかという競合と協調の問題と考えられ、問題は個々のエージェントの自律行動を仮定する事により複雑化する。一つの単純化した問題としては、すべてのサルが同一のバナナを目標とすることである。しかし、サルがどのバナナを目標とするか、また、各サルはどのように目標の情報を得るか、という問題が生ずる。本研究では、あるエージェントが別のエージェントに協力する、つまり、二匹で一つのバナナを目標とすると全体としてはどのような傾向が現れるかを検討し数値計算実験を行ったので、その結果を報告する。

2. エージェントによる問題のモデル化

いま各サルをエージェント1個体とみなし

$$M = \{M_i : i = 0, 1, 2, \dots, L_{max}\} \quad (1)$$

と表す。各エージェントの構成を

$$M_i = \{I, O, S, F, G\} \quad (2)$$

と表す。I, O, S, F, Gはそれぞれ入力、出力、状態、状態遷移関数、出力関数である。

また、バナナを

$$B = \{B_i : i = 0, 1, 2, \dots, B_{max}\} \quad (3)$$

と表し、バナナの位置を

$$B_i = [X, Y, Z] \quad (4)$$

で表す。

ここで、エージェントの各項は以下の様に定義する。

a) 入力 I

サルへの入力は目標とするバナナまでの距離dとする。

b) 出力 O

サルの出力Oはサルの移動するベクトル $[O_x, O_y]$ とする。

c) 状態 S

サル M_i の状態を

$$S = (p, M_p, p_t, B_k, h, \Delta p, \Delta g, x, y, z) \quad (5)$$

と表す。ここで、各項は以下に定義される。

p : 協力すべきサルへの好意度

M_p : 協力すべきサル

$$M_p = \{M_j : j \neq i | j = 0, 1, 2, \dots, L_{max}\}$$

p_t : 協力する度合い

B_k : 自分の目標とするバナナ $B_k = \{B_k \in B\}$

h : 自分の目標とするバナナへの執着度

Δp : 好意度の変化量

Δg : 執着度の変化量

x : サルのX座標

y : サルのY座標

z : サルのZ座標

d) 状態の遷移関数 $F: S \times I \rightarrow S$

各サルの協力すべきサルへの好意度p、自分の目標とするバナナへの執着度gは目標とするバナナとの距離dによって以下の様に定める。

$$p = p + d \cdot \Delta p \quad (p \text{ が充分大きいとき})$$

$$g = g + d \cdot \Delta g \quad (p \text{ が少ないとき})$$

目標とするバナナ B_k は以下の様に遷移する。ここで

P, Hは、ある定数とする。

(1) $p \cdot p_t > P$ のとき

協力すべきサルの目標のバナナに自分の目標のバナナを変更する。

(2) $p \cdot p_t < P$ のとき

$h > H$ ならば、自分の目標とするバナナを変更しない。

$h < H$ ならば、自分の目標とするバナナをランダムに変更する。

協力するサル M_p は以下の様に遷移する。

(1) $p < 0$ のとき

協力するサルを変更する。

(2) $p > 0$ のとき

協力するサルを変更しない。

e) 行動決定法 $G: S \rightarrow O$

サルの出力は移動する方向で与えられるが、移動する方向は各サルの目標とするバナナへの方向ベクトルで与えられる。出力Oは

$$[O_x, O_z] = [X - x, Z - z]$$

で与えられる。

f) 制約条件

以上a)~e)までの各エージェントの構成に以下のような制約条件を設定する。

(1) サルは自分の目標とするバナナまでの距離以外知ることとはできない。

(2) サルはXYZ空間においてXZ方向以外の移動は不可能である。(Y方向については後述する。)

(3) サルが目標とするバナナを変えられるのは

・ 他のサルに協力するとき。

・ 他のサルに協力せず、且つ、自分の目標とするバナナへの執着度が小さいときである。

(4) サルの移動方向は、目標とするバナナへの方向のみである。

3. シミュレーションアルゴリズム

(1)初期設定

サルは $Y=0$ のXZ平面上にランダムに発生し、バナナは天井から高さの3/4までのXZ平面上にランダムに発生する。 $\Delta p \cdot \Delta g$ はそれぞれ $[0,1]$ に発生し、最初の目標バナナ B_k は発生時にランダムに決定するものとする。

(2)Y方向への移動

サルのY方向へ移動する「登る」という行動は、サルが隣り合ったとき無条件に発生する。

(3)行動

総てのサルが移動を終えたとき1ステップ終了とみなしエージェントの行動を前章のモデルによって決定する。

(4)終了

指定したステップ内のとき(2)~(3)を繰り返す。そうでないとき終了する。

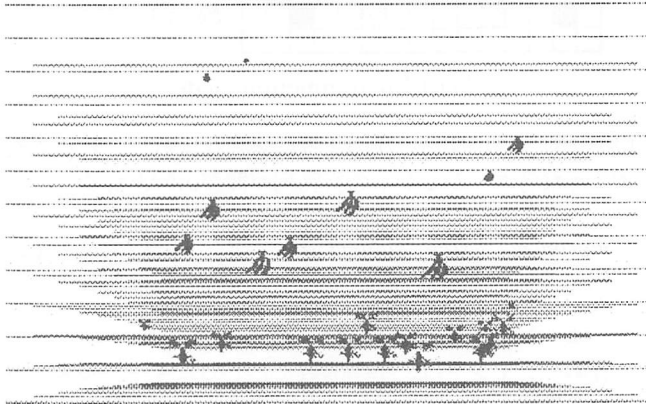


Fig.1 Monkeys and Bananas in the 3-dimensional space

4. 数値計算実験及び結果

4.1 実験条件

$20 \times 20 \times 20$ のフィールド内に16匹のサルと10個のバナナを発生させる。各サルの協力度合い p_t は、その環境全体の協力度合いを P_t とし、 $p_t = P_t + \Delta P_t$ と発生する。この環境下で、 P_t を0~2および0.5~1.5まで0.01づつ増やし、協力度合いによるバナナの入手ステップ数を実験した。ステップ数は10000までとし、これを越えるものは入手できないものとした。

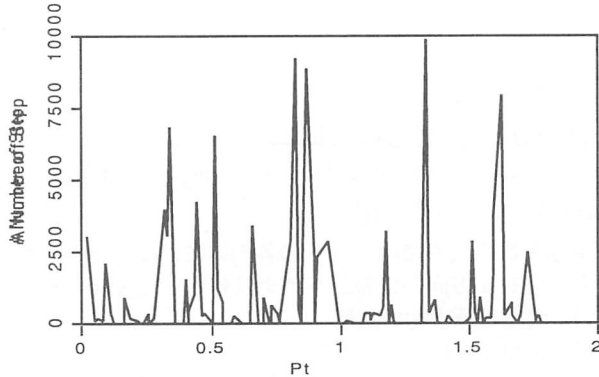


Fig.2 Preference vs steps ($\Delta p_t=0.1$)

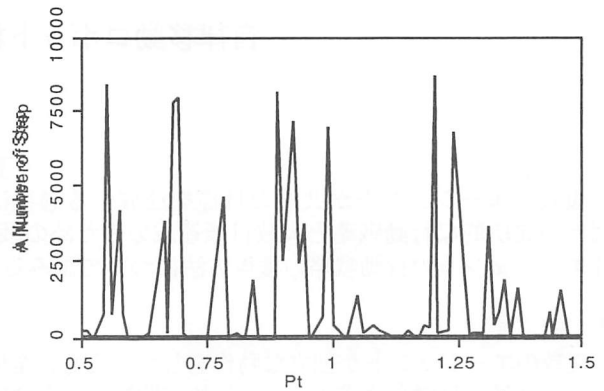


Fig.3 Preference vs steps ($\Delta P_t = \pm 0.5$)

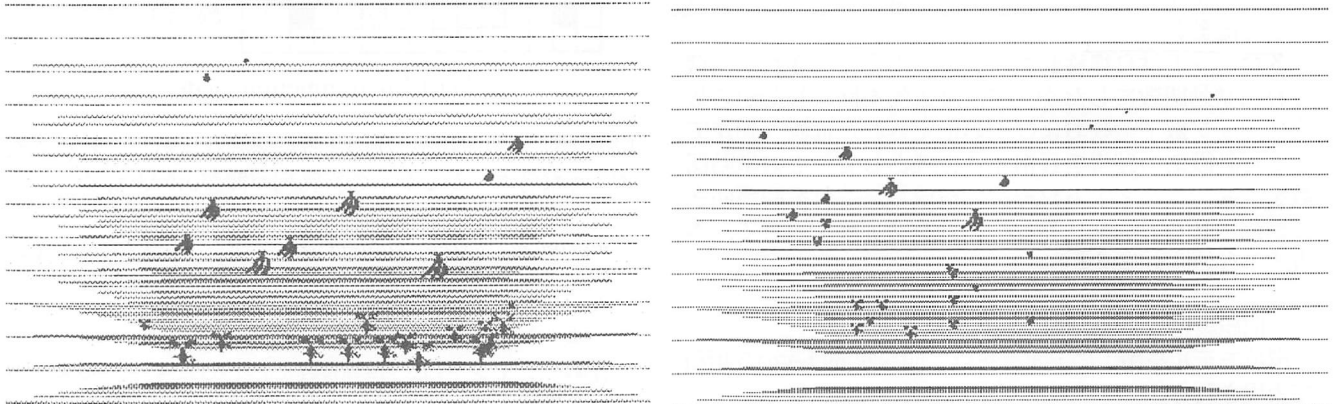


Fig.4 Scene when Monkeys reach Banana

4.2 実験結果

ΔP_t を $\pm 0.1, \pm 0.5$ としたときの実験結果を図2,3に示す。

4.3 結果からの考察

実験の結果から、 P_t の単純増加による収束は見られないが $P_t=1.1 \sim 1.2$ 付近で安定してバナナを入手している。また、 P_t の増分が大きいときある程度到達回数を減少できることが解った。

5. おわりに

「山登り動機付け法」¹⁾と比べ、よりヒューマンライクなアルゴリズムであるこの実験から、ほぼ近似的な結果を得る事ができた。これは、このアルゴリズムが競合と協調の問題を解決していると考えられる。今後、久保・鈴木らの示すように、このアルゴリズムにフェロモン・GA等を用い、拡張した好意度アルゴリズムを検討したい。

参考文献

- 1) Keiji SUZUKI and Azuma OHUCI: "Reorganization of Agents with Pheromone Style Communication in Multiple Monkey Banana Problem" Intelligent Autonomous Systems IOS Press,(1998)