ムカデゲームでの意思決定の正当化・定式化について

[目的]

ムカデゲームでの後ろ向き帰納法による解は最初に S を選択することである。しかし実際ゲームを行うと C がいくらか続くことが多い。しかし最後の意思決定では S を選択することが殆どであり、C はありえない（100% S を選ぶ）と言える。この現象をうまく表すモデルについて考察する。

[問題分析]

ゲームが進めば進むほどお互いの利得が多くなるのは事実であり、即ちお互いが C を選びにいくという可能性は十分にあるから、最初に S を選ぶことは正当ではないということをまず考えた。
その上で、ゲームが進むごとにだんだん S を選びたくなる確率が大きくなる(経験則)から、「自分の合理性」というパラメータを考えた：ここでの合理とは後ろ向き帰納法における最適解である。
即ち、後ろ向き帰納法にどれだけあやかるかというパラメータである。
しかしこの場合最後に S を必ず選択することをうまくあわらす事ができない(最後ですらも自分が完全に合理的でない可能性がある)。
そこで考えたのが、「相手の合理性への信頼性」である。

[考えたモデル]

相手の合理性への信頼性 p $(0 \leqq p \leqq 1)$を各プレイヤーが持つとする。
基本的には後ろ向き帰納法を用いるが、相手の意思決定を
・p の確率で後ろ向き帰納法
・$1-p$ の確率でそうではない選択
をするモデルを考える。
相手の意思決定を予想するとき、各プレイヤーの
・後ろ向き帰納法に基づく場合の利得を a
・基づかない場合の利得を b としたとき
プレイヤーの確率的利得を $ap + b(1-p)$とする。
このモデルの重要なところは最後の選択において相手の選択が絡まないため、確実に自分の利得が大きくなる方：100% S を選択するということが合理化される。

[n 期ムカデゲームの解析]

i 番目$(1 \leqq i \leqq 2)$プレイヤーの j 番目$(1 \leqq i \leqq n)$の選択が S となる確率を a_{ij} とする。
i 番目$(1 \leqq i \leqq 2)$プレイヤーの相手の合理性の信頼性を p_i とする。
i 番目$(1 \leqq i \leqq 2)$プレイヤーが j 番目$(1 \leqq i \leqq n)$に C を選択したときの k 番目$(1 \leqq k \leqq 2)$プレイヤーの利得を u_{ijk} とする。
この時、a_{ij} を次の方法で定める：

・$a_{2,n} = 1$ である。(この時 $u_{2,n,1} = n-2$, $u_{2,n,2} = n$)
・$a_{1,n}$ に於いて、相手は確率 p_1 で S、$1-p_1$ で C と考えられるため、自分が C を選んだときの確率的利得 $u_{1,n,1}$ は $p_1(n-2)+(1-p_1)(n+1) = 1+n-3p_1$ と表せる。
・これに基づき $1+n-3p_1$ が自分が S を選択した場合の利得 n-1 よりも大きい場合は S ではなく C を選択することは合理的であると言える。(相手が一定以上合理的でないと自分が判断している場合 ： この時は $p_1 < 2/3$ のとき)
・この方法を数学的帰納法的に適応することで最初からの決定木を構築する。

この通り **p** にかなり依存するためいくつかの具体的な **p** の値に関する解析結果を表示する。
また n=**5**(授業でも行った 5 期ムカデゲーム)とする。

1. **p** = (1,1)のとき ： お互いが合理的だと互いに思っている場合
   これは後ろ向き帰納法と一致する。よって最初の選択で S が選ばれ、利得は 1, -1 である。
2. **p** = (0,0)のとき ： お互いが常に協力しようと考える場合
   相手が C を取るということは、自分の利得が増えることにつながるので自分も C を取る。
   よって最後の選択でプレイヤー2 が S を選択し、利得は 8, 10。
3. **p** = (1,0)のとき ： プレイヤー1 が「信頼」(=協力的だと信じる)していない場合
   最初にプレイヤー1 が「合理的選択」を行う。よって利得は 1, -1 である。
4. **p** = (0,1)のとき ： プレイヤー2 が「信頼」(=協力的だと信じる)していない場合
   最初のプレイヤー2 の番で「合理的選択」を行うので利得は 0, 2。

また、コンピュータシミュレーションを行うと、**p** = (0.3, 0.6)のような間の例では利得は 6, 8 となり、絶妙な位置ではあるが相手を疑っている側のほうが利得が高いという結果を得た。

[結論]
ムカデゲームでの「最後の選択では 100% S を選択する」「後ろ向き帰納法を少し変更した」意思決定モデルを考えた。実際にこのモデルはこれらの条件を満たしている。

なお、このモデルでは常に同じ確率を仮定している ： ゲームが進むことで相手への信頼性が減ることを考慮に入れていない。これは時間の概念をモデルに取り込んでないことから生じる。しかし目的自体は満たしているため、今回では問題ではない。

[参考]
今回作成したプログラムである。http://ideone.com/nVLs59