

ゼロ和ゲームにおける嘘の頻度と競合状況が与える影響

安井健太^{1*} 寺田和憲¹ 山田誠二² Jonathan Gratch³

¹ 岐阜大学

² 国立情報学研究所

³ University of Southern California

Abstract: 協力者と裏切り者が混在する非ゼロ和社會において、嘘発見能力は必須である。嘘の認識は相手の事前の発話とその後の行動が一致するか否かを認識することで行われる。狼少年に対する村人のように、相手が嘘をつき続けていれば、信じない（無視する）戦略を取ることで嘘つきに対抗することができる。本研究では、ゲーム前に自分の出す手を宣言する「宣言つきじゃんけん」を用いて、宣言と実際の手の乖離（嘘の割合）がどの程度であれば信じる戦略を採用するかどうかを調査した。また、対人競合ゲーム状況が戦略に与える影響について調査するために、宣言じゃんけんと等価な選択課題を用いて結果の比較を行った。実験によって、嘘の割合が5割を超えると相手の宣言を信用しなくなる傾向にあるが、嘘の割合が高い場合であっても、必ずしも、合理的戦略である、宣言を全く信じない戦略を採用しないことが示された。また、対人競合ゲーム状況では宣言を信じやすくなる傾向が示された。

1 はじめに

一日に何回嘘をついていたかを調査した研究によると、アメリカの大学生は1日に1.94回[1]、日本の大学生は1.77回であった[2]。嘘が日常的に使われている社会では嘘発見能力は必須能力である。嘘の認識は相手の事前の発話と結果が一致するか否かを識別することで行われる。例えば、童話狼少年では、村人は少年の「牧場に狼が来た」という発言に対して、実際に牧場に行って狼がいないことを確認することで少年が嘘をついたことを認識した。当初の村人の戦略は少年の発言を「信じる」であった。この戦略は少年が真実を述べているという前提に立つものであり、牧場の羊を失う可能性を低減させられるため合理的である。しかし、村人は何度も牧場に行って狼が居ないことを確認するという体験を繰り返すことで、少年が真実を述べていないという前提に立つようになり、「信じない」戦略に転じた。この戦略は牧場に確認に行くという労力を低減させられるために合理的である。嘘の頻度は情報受信者の戦略（信じる・信じない）に影響を与えるが、発話と結果の祖語の程度と受信者の戦略の関係はこれまでに調査されていない。また、嘘の割合推定は、時間的に隣接する事象間の真偽推定という統計的学習問題であるが、じゃんけんなどの対人競合ゲームという状況であれば、勝って嬉しい、負けて悔しいなどの心理的な影響が生じるため、推定値に影響を与え

ると考えられる。

ゲーム理論におけるゲームの利得に影響を与えないプレイヤー間のコミュニケーションをチーブトークという。チーブトークによって意図的に偏ったあるいは誤った情報を送ることで相手の意思決定を操作できる可能性がある。チーブトークは経済学での研究が活発であり、チーブトークによる消費者への購買意欲の向上や競争相手の市場への参加、不参加の操作可能性などが示されている[3-6]。

チーブトークを用いた非ゼロ和ゲームの研究は多い。一対一の協力ゲームである調整ゲームでチーブトークを用いた結果、チーブトークありの方が両者ともより多くの利得を得られた[7]。一方で一対一の競争ゲームである参入ゲームで受信者と送信者とのチーブトークを行った際は、チーブトークを用いた場合受信者は挑戦意欲が低下し、送信者は挑戦意欲が高まったため送信者がより多くの利得を得ることができた[6]。

意思決定問題において情報は重要な役割を果たす。自分の戦略を隠したり、誤った情報を与えることで受信者の利得を下げられる可能性があり、ゲーム理論における詐欺や嘘の影響を調査する研究も多い。ゼロ和ゲームにおける情報は対戦相手の行動を予測しやすくなるため、受信者の利得を高める効果がある。また、情報のノイズの影響は、受信者の情報に対する依存度と発信者の嘘をつく能力に依存するとされている[8]。

本論文ではチーブトークを用いた繰り返しRock-Scissor-Paper (RPS) ゲームにおけるプレイヤーの意思決定問題を調査する。RPSゲームとは2人用の非協力ゼロ和

*連絡先：岐阜大学 自然科学技術研究科 知能理工学専攻
〒501-1193 岐阜県岐阜市柳戸1番1
E-mail: yasui@ai.info.gifu-u.ac.jp

ゲームであり、競争環境における意思決定問題を調査するための基本的なモデルである。弱肉強食のサイクルなどといった、生物学における競争のモデルにもよく使用されている [9-11]。このゲームの3つのアクションである Rock, Paper, Scissor はそれぞれ等価なアクションであり、複雑な意思決定のない単純なコミュニケーションモデルにもなるため、人工知能に RPS ゲームをさせることで、人工知能の性能や人対人工知能のコミュニケーションの調査する研究も多い [12-14]。

RPS ゲームのような対戦状況において、相手が CPU であると感じた場合、対戦相手はプログラムに従って動いていると認識するが、相手が人間であると感じた場合、対戦相手は明確な信念、意図、欲求をもつエージェントであると認識し、対戦相手の反応から行動パターンを読み取ろうとすることが示されている。また、CPU を相手にした場合よりも人間を相手にしたときの方が負けたときにがっかりしたと説明した参加者もいた [15]。他にも、CPU が対戦相手である場合よりも人間が対戦相手であった場合の方が行動エントロピーが高くなることが示されている [16]。実際に対戦相手が CPU であるかどうかに関わらず対戦状況では、対戦相手が何らかの戦略を持っていると感じたり、「勝って嬉しい」「負けて悔しい」などの心理的な変化が意思決定に影響する可能性がある。本論文では RPS ゲームと等価な選択課題と比較して対人競合ゲームが意思決定に与える影響について調査する。

騙しとは相手に対して嘘や偽物を使い、それを本物であると思わせることである。人間同士のコミュニケーションでは主に嘘が用いられる。嘘の種類は事実と全く異なることを伝える嘘、事実よりも大きくまたは小さく表現する嘘、事実には基づいているが一部だけに注目したり、事実を隠したりする嘘の3つに分けられる [1]。また、人間は嘘に対して真実バイアスという特性を持つことが示されている。真実バイアスとはメッセージの受信者がメッセージの真偽性に関わらずメッセージが真実であると判断する傾向のことである [17]。また、メッセージの送信者が一貫して正直または不正直に見えるという態度バイアスという概念も存在する [18]。このバイアスは信じる・信じないの戦略採用に寄与するものと思われる。

自然界の競争でも嘘が存在し、嘘をつく能力を持つ生物が多く存在する環境において、敵対者は必然的にを見抜く能力を得る。そのため、嘘の頻度が高いほどの成功率は低くなる。例えば、キタノトミヨのオスは別のオスの巣にある未受精卵に授精させるために、メスに擬態してその巣に侵入する。これを忍び寄りという。忍び寄りをする者が少ないときには成功率が高いが、多くなると無警戒のオスを見つけるのが難しくなるため忍び寄りは失敗する [19]。また、シロオビアゲハは捕食者への対策として毒蝶であるベニモンアゲハ

に擬態する能力を持つ。しかし、擬態するシロオビアゲハが多い環境においては擬態に成功する確率は低く、逆に擬態する能力が低いものの方が生存率が高くなる [20]。このように、自然界において、敵対者は行為者側の嘘の頻度が高い場合には嘘を信じない戦略を採用することで適応度を向上させているものと考えられる。

RPS ゲームのような対戦状況において、プレイヤーは対戦相手の行動履歴や勝敗の結果を学習して相手の戦略に合わせた意思決定を行うため、嘘の頻度が高い対戦相手と戦う場合相手を信じないという戦略を立てる可能性が高いと考えられる。サルを使用した実験では3種類のアルゴリズムを持つ CPU とマッチングペニーズゲームをしたところ、サルは相手のアルゴリズムに応じて自分の戦略を変更していたことが分かり、いずれのアルゴリズムで対戦した場合も勝率は比較的高かった [21]。また、人間が人間の行動の規則性を利用する3種類のアルゴリズムを持つ CPU とゼロ和ゲームをしたところ、どのアルゴリズムで対戦した場合でも対戦相手の行動を高いパフォーマンス学習でしていたことが示された [22]。

本研究ではチーブトークを用いた繰り返しゼロ和ゲームによって、嘘の受信者が合理的な行動を取れるかどうかを調査した。我々の知る限りではチーブトークを用いた繰り返しゼロ和ゲームを調べた研究はない。本研究から、嘘の頻度との成功率の関係を数理モデル化することで、認知科学、行動経済学における理論発展に貢献できる可能性がある。

2 実験方法

2.1 実験参加者

実験参加者は Yahoo!クラウドソーシングで募集した、18 歳以上の男性 243 人、女性 101 人であった (Mage = 43:19, SDage = 10:06)。実験計画は2要因(ゲームの種類:対人競合ゲーム,非対人競合ゲーム)×(嘘の割合:0.05,0.25,0.5,0.66,0.75,0.95)参加者間要因配置であった。

2.2 実験装置

データ収集は PHP と Javascript で作成した対人競合ゲームと非対人競合ゲームの2種類のゲームを用いて行った。

対人競合ゲーム水準では実験参加者は宣言じゃんけんゲームを行う。「宣言じゃんけん」はチーブトークを用いた Rock-Paper-Scissors(RPS) ゲームである。図1に「宣言じゃんけん」Web インターフェイスを示す。画面構成は対戦相手(中央)、参加者が出した手(左上)、

ラウンド数と得点 (右上), クリック可能なグー, チョキ, パーのアイコン (下) である (図 1a 参照). 参加者はマウスを使ってグー, チョキ, パーの 3 つのアイコンのうち一つをクリックする. アイコンをクリックすると短い待機時間の後, 対戦相手の選んだ手とラウンドの勝敗が表示される (図 1b 参照). 参加者が 1 ラウンドが終わったときに表示される画面下部の確認ボタンをクリックすることで次のラウンドが開始する. 対戦相手は「宣言通りの手を出す」「宣言以外のランダムな手を出す」の 2 つの行動パターンを持つ.

非対人競合ゲーム水準では実験参加者は時間隣接事象割合推定 (記号合わせ) ゲームを行う「記号合わせ」とは「第一状態」で表示される記号から「第二状態」で表示される記号を予測するゲームである. 第一状態と第二状態で表示される記号はクローバー, ハート, スペードの 3 種類である. 「記号合わせ」の Web インタフェース図 2 に示す. 画面構成は第一状態の記号 (中央), ラウンド数 (右上), クリック可能なクローバー, ハート, スペードのアイコン (下) である (図 2a 参照). 第一状態のとき, 参加者はマウスを使ってグー, チョキ, パーの 3 つのアイコンのうち一つをクリックする. アイコンをクリックすると短い待機時間の後, 第二状態の記号が表示される (図 2b 参照). 参加者が 1 ラウンドが終わったときに表示される画面下部の確認ボタンをクリックすることで次のラウンドが開始する. 第二状態で表示される記号は「第一状態で表示された記号」「第一状態で表示されなかった記号」の 2 つのパターンである. 2 つのゲームはルールが同じであり, 違いはインターフェイスのみである.

参加者はどちらか片方のゲームを 100 ラウンドプレイする. また, 宣言じゃんけんであれば「宣言以外のランダムな手を出す」記号合わせであれば「第一状態で表示されなかった記号」といった, 事前に提示された情報と一致しない結果を出すパターンがある. これを本研究では嘘の宣言とした. そして, 100 回行われる宣言のうち, 嘘の宣言をする割合は 6 通り (0.05, 0.25, 0.50, 0.66, 0.75, 0.95) のパターンがある. 本研究ではこれを嘘の割合とした.

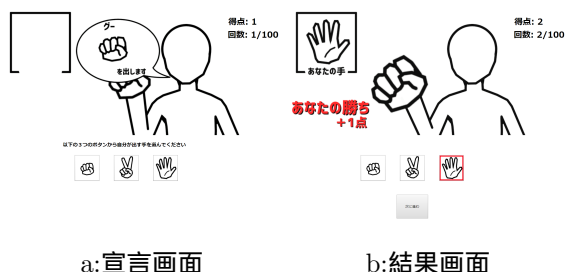


図 1: 宣言じゃんけん



図 2: 記号合わせ

2.3 測定方法

実験参加者が対戦相手の宣言を信じていたか調査するため, ゲームプレイ中における実験参加者の行動履歴 (宣言, 出し手) と勝敗を記録し, 各ラウンドの勝敗から正答率を算出した. また, 各条件の正答率から平均正答率を算出した. 実験参加者の平均正答率は理論値と比較して, 対人競合の有無が実験参加者の意思決定に影響を与えるか調査した.

また, 対戦相手が滅多に嘘をつかない性格だと判断した場合対戦相手の宣言を信じる, 対戦相手がよく嘘をつく性格であると判断した場合対戦相手の宣言を信じない, という戦略をとるのが合理的である. そのため, 実験参加者が対戦相手を信じていたか信じていなかったかを調べるために, 対戦相手の宣言と参加者の選んだ手が連続で 10 ラウンド以上一致した場合, その区間は実験参加者が対戦相手を信じるという戦略を採用していたとみなした. 例えば, 15 ラウンド連続で対戦相手の宣言と参加者の選んだ手が一致した場合, その参加者は 15 ラウンドの間信じる戦略を採用していたとする. そして, 100 ラウンド中 50 ラウンド以上対戦相手を信じる戦略を採用していた実験参加者は対戦相手が信用できる人物であると判断し, 対戦相手を信じるという戦略をとっていたとみなした.

2.4 手順

実験参加者にプレイするゲームの種類と嘘の割合を割り当て, 宣言じゃんけんまたは記号合わせのルール説明を行った. 宣言じゃんけんにおいては対戦相手が人工知能であると説明し, どちらのルールでも嘘の宣言が存在することを伝えなかった. その後実験参加者にゲームをプレイするように求め, 実験参加者がゲームを 100 ラウンドプレイし終えたら実験を終了した.

3 実験結果

3.1 正答率

図3に各条件ごとの参加者の平均正答率を示す。対人競争ゲームと非対人競争ゲームの間で多重比較検定を行った。結果、嘘の割合が0.25であるとき記号合わせよりも宣言じゃんけんの方が正答率が有意に高かった ($F(1, 332) = 7.787, p < 0.01$)。嘘の割合が0.50であるとき記号合わせよりも宣言じゃんけんの方が正答率が有意に高かった ($F(1, 332) = 3.935, p < 0.05$)。嘘の受信者は、嘘の割合が0.66以下であれば発信者の宣言を全て信じる、0.66以上の場合相手の宣言を全て信じない選択をすることで期待値を最大化し、合理的な戦略となる(表3参照)。このときの期待値を本研究では理論値とした。平均正答率と理論値の差を各条件ごとに比較すると、記号合わせの場合は嘘の割合が0.05であるときと比べて、嘘の割合が0.25であるときの方が平均正答率と理論値との差は有意に高かった ($F(5, 332) = 14.360, p < 0.05$)

3.2 戦略

各参加者がどの戦略 (trust, no strategy, not trust) を選んだかを調べて、各条件における戦略の割合を図4に示した。嘘の割合が0.25, 0.50であるとき記号合わせよりも宣言じゃんけんの方が相手の宣言を信じるという戦略を使っている参加者の人数が多かった。(表2参照) また、嘘の割合が0.05であるときと比べて、嘘の割合が0.95であるときのほうが、合理的な戦略を採用した参加者の人数の割合は宣言じゃんけん ($\chi^2 = 25.657, df = 1, p < .001$) と記号合わせ ($\chi^2 = 23.478, df = 1, p < .001$) のどちらのゲームにおいても有意に少なかった。

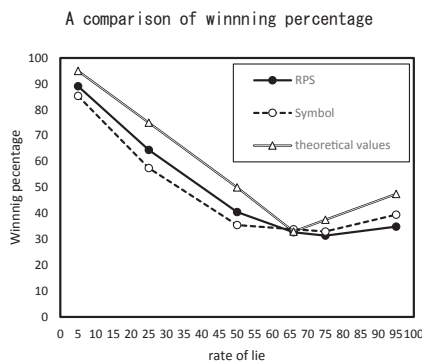


図3:平均正答率の比較

Lie rate	0.05	0.25	0.50	0.66	0.75	0.95
RPS	89.1	64.5	40.5	32.6	31.3	34.8
Symbol	85.3	57.4	35.5	33.8	33.0	39.4

表1:正答率

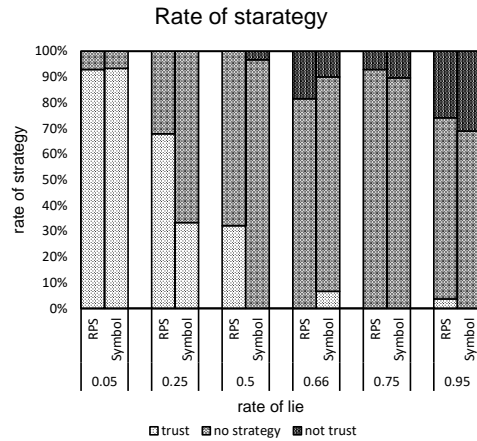


図4:戦略の比較

4 議論

平均正答率と理論値を比較すると、記号合わせの場合嘘の割合が0.05から0.25に上がったとき平均正答率は理論値から大きく離れた。これは嘘の割合が極めて低い場合(嘘の割合が0.05)でなければ、人間は合理的な行動を取りづらくなるといえる。一方、記号合わせと比べると宣言じゃんけんは嘘の割合が0.25の場合でも、比較的合理的な行動を取ることができていた。また、嘘の割合が0.05のとき9割以上の参加者が信じる戦略を採用したが、嘘の割合が0.95のとき信じない戦略を使った参加者は3割以下であった。これは参加者は相手の宣言を全て信じるという合理的戦略は採用しやすいが、相手の宣言を全て信じないという合理的戦

	Lie rate	0.05	0.25	0.50	0.66	0.75	0.95
RPS	trust	92.8	67.8	32.1	0.0	0.0	3.7
	no strategy	7.1	32.1	67.8	81.4	92.8	70.3
	not trust	0.0	0.0	0.0	18.5	7.1	25.9
Symbol	trust	93.3	33.3	0.0	6.6	0.0	0.0
	no strategy	6.6	66.6	96.6	83.3	89.6	68.9
	not trust	0.0	0.0	3.3	10.0	10.3	31.0

表2:戦略の割合

Lie rate	0.05	0.25	0.50	0.66	0.75	0.95
Theoretical values	0.95	0.75	0.50	0.33	0.375	0.475

表 3: 理論値

略は採用されにくいということを示している。また、実際の嘘の割合は同じであるにもかかわらず、嘘の割合が 0.25 と 0.5 であるときは記号合わせよりも宣言じゃんけんをプレイした参加者の方が信じる戦略の採用率が非常に高く、正答率も宣言じゃんけんの方が有意に高かった。これは記号合わせよりも宣言じゃんけんをプレイした参加者の方が対戦相手の宣言を信じる傾向があったことを示している。

過去のチープトークを用いた繰り返しではないゼロ和ゲームの研究では、情報の送受信がある対戦状況において情報のノイズが小さいほど受信者の利得は高くなるとされている [8]。本研究でも嘘の割合が最も低いときに最も正答率が高くなったため、同様の結果を得られたと言える。そのため、繰り返しゲームであってもこの傾向は不変である。また、過去の研究では繰り返しゼロ和ゲームにおいて単純な戦略よりも混合戦略の方が相手に予測不可能にさせることができるため効果的であることを示した [23]。本研究における混合戦略とは発信者が真実と嘘を織り交ぜることであり、嘘の割合が中程度である 0.50 や 0.66 であるときに対戦相手は混合戦略をしていたと言える。そして、嘘の割合が極端に高いまたは低いときよりも嘘の割合が中程度であるときのほうが正答率が低くなっていることから、チープトークを用いた場合でも単純な戦略よりも混合戦略の方が効果的であったと言える。対戦状況において人間が対戦相手の行動履歴から対戦相手の行動パターンを学習していくことが過去の実験によって示されている [22]。本研究の参加者も前半から後半にかけて信じる戦略や信じない戦略を採用する参加者が増えていったため、対戦相手の行動履歴から嘘つきであるか否かを判断していたと考えられる。過去の研究では対戦ゲームにおいて対戦相手が CPU である場合と人間である場合を比較したところ、実際に対戦相手が人間であるかどうかに関わらず対戦相手が人間であると認識したときの方が実験参加者は戦略的な思考、行動をすることが示されていた [15, 16]。どちらのゲームの参加者も同等に戦略的な行動をとっていたため、先に述べた対戦相手が人間であるか CPU であるかの差とは異なるものであったが、本研究では同じアルゴリズムの対戦相手であるにも関わらず対人競合ゲームであるか否かによって参加者の意思決定に差が生じることを示した。

本研究では対人競合ゲーム状況では宣言を信じやすくなる傾向を示したが、その影響を与えた要因につい

てはまだ調べられていない。今後は対人競合ゲームをプレイした参加者と非対人競合ゲームをプレイした参加者にそれぞれゲームプレイ中の意識や感情に関するアンケート調査を行い、調査結果の差から要因を推測、特定する必要がある。

謝辞

本研究は JSPS 科研費 16KK0004 の助成を受けたものである。

参考文献

- [1] Bella M. DePaulo, Deborah A. Kashy, Susan E. Kirkendol, Melissa M. Wyer, Jennifer A. Epstein.: Lying in everyday life., *Journal of Personality and Social Psychology*, Vol. 70, No. 5, pp. 975–995 (1996)
- [2] 村井潤一郎：青年の日常生活における欺瞞, *性格心理学研究*, No. 9, pp. 56–57 (2000)
- [3] Maurice Doyon, Laure Saulais, Bernard Ruffieux, Denise Bweli : Hypothetical Bias for Private Goods: Does Cheap Talk Make a Difference?, *Theoretical Economics Letters*, Vol. 5, No. 6, pp. 749–756 (2015)
- [4] Matthew Backus., Thomas Blake., Steven Tadelis.: On the Empirical Content of Cheap-Talk Signaling: An Application to Bargaining, *Journal of Political Economy*, University of Chicago Press (2019)
- [5] Xinyu Li, Ronald Peeters.: Cheap Talk with Multiple Strategically Interacting Audiences: An Experimental Study, *PLOS ONE*, Vol. 11, No. 10, e0163783 (2016)
- [6] Dustin H. Tingley, Barbara F. Walter.: Can Cheap Talk Deter?, *Journal of Conflict Resolution*, Vol. 55, No. 6, pp. 996–1020 (2011)
- [7] Gary Charness, Brit Grosskopf.: Cheap Talk, Information, and Coordination - Experimental Evidence, *SSRN Electronic Journal*, Elsevier BV (2001)
- [8] Dongxu Li, Jose B. Cruz Jr.: Information, decision-making and deception in games, *Decision Support Systems* (2009)

- [9] Marcus Frean, Edward R. Abraham.: Rock-scissors-paper and the survival of the weakest, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, Vol. 268, No. 1474, pp. 1323–1327 (2001)
- [10] Mauro Mobilia.: Oscillatory dynamics in rock-paper-scissors games with mutations, *Journal of Theoretical Biology*, Vol. 264, No. 1, pp. 1–10 (2010)
- [11] Duncan D. Cameron, Andy White, Janis Antonovics.: Parasite-grass-forb interactions and rock-paper-scissor dynamics: predicting the effects of the parasitic plant *Rhinanthus minor* on host plant communities, *Journal of Ecology*, Vol. 97, No. 6, pp. 1311–1319 (2009)
- [12] Tjeng Wawan Cenggoro, Awang Harsa Kridalaksana, Eka Arriyanti, M. Irwan Ukkas.: Recognition of a human behavior pattern in paper rock scissor game using backpropagation artificial neural network method, *2014 2nd International Conference on Information and Communication Technology (ICoICT)*, IEEE (2014)
- [13] Franco Salvetti, Paolo Patelli, Simone Nicolo.: Chaotic time series prediction for the game, Rock-Paper-Scissors, *Applied Soft Computing*, Vol. 7, No. 4, pp. 1188–1196 (2007)
- [14] Elaine Short, Justin Hart, Michelle Vu, Brian Scassellati.: No fair!! An Interaction with a Cheating Robot, *Proceeding of the 5th ACM/IEEE international conference on Human-robot interaction - HRI10*, ACM Press (2010)
- [15] Helen L. Gallagher, Anthony I. Jack, Andreas Roepstorff, Christopher D. Frith.: Imaging the Intentional Stance in a Competitive Game, *NeuroImage*, Vol. 16, No. 3, pp. 814–821 (2002)
- [16] Hideyuki Takahashi, Keise Izuma, Madoka Matsumoto, Kenji Matsumoto, Takashi Omori.: The Anterior Insula Tracks Behavioral Entropy during an Interpersonal Competitive Game, *PLOS ONE*, Vol. 10, No. 6, e0123329 (2015)
- [17] Charles F. Bond, Bella M. DePaulo.: Accuracy of Deception Judgments, *Personality and Social Psychology Review*, Vol. 10, No. 3, pp. 214–234 (2006)
- [18] Miron Zuckerman, Bella M. DePaulo, Robert Rosenthal.: Verbal and Nonverbal Communication of Deception, *Advances in Experimental Social Psychology*, Elsevier, pp. 1–59 (1981)
- [19] P. M. Sheppard.: The Evolution of Mimicry a Problem in Ecology and Genetics, *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. 24, No. 0, pp. 131–140 (1959)
- [20] Desmond Morris.: Homosexuality in the Ten-Spined Stickleback (*Pygosteus pungitius* L.) 1), *Behaviour*, Vol. 4, No. 1, pp. 233–261 (1951)
- [21] Daeyeol Lee, Michelle L. Conroy, Benjamin P. McGreevy, Dominic J. Barraclough.: Reinforcement learning and decision making in monkeys during a competitive game, *Cognitive Brain Research*, Vol. 22, No. 1, pp. 45–58 (2004)
- [22] Leonidas Spiliopoulos.: Strategic adaptation of humans playing computer algorithms in a repeated constant-sum game, *Autonomous Agents and Multi-Agent Systems*, Vol. 27, No. 1, pp. 131–160 (2012)
- [23] Jon von Neuman, Oskar Morgenstern.: *Theory of Games and Economic Behavior*, Princeton University Press (1944)