

社会の縮図としての  
 “人狼がいないかもしれない3人狼”パラダイムの提案と  
 ロボットによる行動変容の試み  
 Modified werewolf game for a microsome of human society  
 and behavioral modification induced by a robot in the game

高橋英之<sup>1</sup>, 伴碧<sup>2</sup>, 大澤博隆<sup>3</sup>  
 Hideyuki Takahashi<sup>1</sup>, Midori Ban<sup>2</sup>, Hiroataka Osawa<sup>3</sup>

1. 大阪大学, 2. 同志社大学, 3. 筑波大学

1. Osaka University, 2. Doshisha University, 3. Tsukuba University  
 takahashi@irl.sys.es.osaka-u.ac.jp

### Abstract

We proposed a modified version of *Werewolf*, a conversation-based game, designed to evaluate trust and various human social behaviors. Although there are always winners and losers in the regular version of *Werewolf*, the rules were modified to generate a possible scenario in which no enemy was present among the players and all players would win if they trusted each other. In the presentation, we would like to introduce several cases of the effect of using robots while playing this *Werewolf* on players' social behaviors and trust.

**Keywords** — *Werewolf* game, trust, robot

### 1. 本発表の概要

“魔女狩り”という言葉が代表するように、古くから人類は疑心暗鬼に囚われることで、時に根拠が無い情報を妄信して、無思慮な行動をしばしばとる傾向がある。特にメディアの媒体が多様化して個人が自由に情報を発信することが可能になった今日においては、SNSなどを通じて信憑性が希薄の情報が独り歩きしてあたかも真実であるかのように爆発的に広まっていく危険性が増してきた。集団が疑心暗鬼に囚われることで、根拠の不確かな情報が真実であるかのように実体化しないためにはどのような方策が有効なのであろうか？

集団における疑心暗鬼の発生を考える上での一つの題材として、我々は人狼ゲームに注目した。人狼ゲームとは、複数のプレイヤーで行う会話ゲームである。まずゲームの開始時に、それぞれのプレイヤーは他のプレイヤーには分からないように自分の役職をゲームマスターから告知される。役職には村人と人狼に加えて様々なものが存在するが（占い師、

狂人など）、大きくは村人側の役職と人狼側の役職にプレイヤーは分けられる。各プレイヤーの役職が決定した後、一定時間の自由会話が行われる。村人側の役職のプレイヤーは人狼のプレイヤーが誰かをこの会話の中から推理する。一方で、人狼側のプレイヤーは人狼のプレイヤーがあたかも人狼ではないかのように会話の流れを誘導する必要がある。そして一定時間の会話が終わったのち、誰が人狼と思うのかを各プレイヤーが投票し、もっとも多く投票がなされたプレイヤーがゲームから追放される。適切に人狼が追放することができたら村人側のプレイヤーの勝利、一方で人狼が追放されないで残ったら人狼側のプレイヤーの勝利となる。人狼ゲームは、他のプレイヤーの気持ちをうまく読み取るスキルや違う役職を演じるスキルなど様々な社会的能力が必要であり、近年では娯楽としてだけではなく、人間の社会性を研究する題材として認知科学や人工知能の研究者の間でも注目が集まっている [1] [2]。

我々は疑心暗鬼を調べるための改変をした上でこの人狼ゲームを研究のパラダイムとして用いている。具体的には、プレイヤーの数を三名とし、役職決めの際に、人狼が誰にも割り振られないケースを一定確率で生じさせた（図1）。そして三名すべてのプレイヤーが、会話を通じてプレイヤーの中に人狼がないという結論になった場合、人狼を追放する投票を行わないという選択肢を新たに加えた。この改変の結果、人狼がプレイヤーの中に存在しない場合に限り、すべてのプレイヤーがお互いを信じることで誰も敗者がいない形でゲームを終わらせることが可能になった。その一方で、本当は人狼が存在しているのに、人狼を追放する投票を行わない場合は無条

件で人狼側の勝利とした。このような改変を行うことで、お互いを信じすぎても、信じなさ過ぎても、村人側の勝利には結びつかないという疑心暗鬼が発生しやすい不確定な状況をつくり出した。実際に予備実験において、同一メンバーでこのゲームを繰り返し行うことで、プレイヤーの中に人狼がないという合意の頻度が低下することが確認されている。

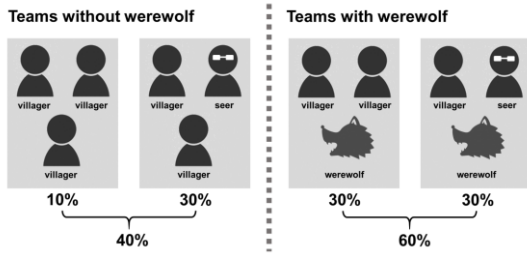


図1. 改造版人狼ゲーム



図2. 実験設定

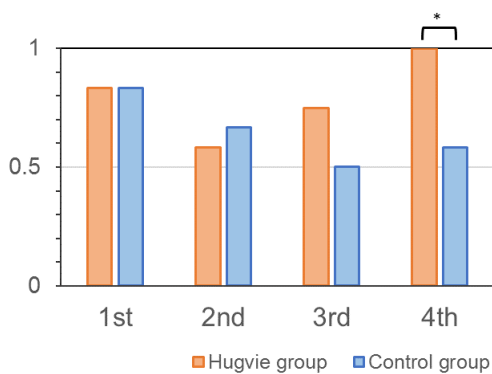


図3. 他のプレイヤーに対する信頼度（村に人狼がないと判断する割合）の試行に応じた変化

このようなパラダイムを用いて、どのような介入をプレイヤーにすれば、村人側の勝率が向上するのかについて検討することは、社会における疑心暗鬼の発生を抑制する一つのテストベッドとなるのでは

ないかと期待している。これまでに行った検討として、ハグビー[3][4]という人の形をした抱き枕型の通信デバイス（図2）を通じて各プレイヤーがこのゲームを行った場合、プレイヤーの中に人狼がないという合意の頻度がゲーム後半まで下落しないことが示されている（図3）。この結果は、抱き枕を抱くという行為が他者を信じる態度を高めることを示唆している。

一方で、ハグビーを用いた条件における村人側の勝率は非常に低く、抱き枕によって高められる他者を信じる態度は根拠が無い妄信ともいえる。村人側の勝率を引き上げるためには、単に根拠なき信頼を高めるだけでは駄目で、個人のプレイヤーが頭を使ってこのゲームに主体的に取り組むようになることが重要である。現在はロボットをこのゲームに介入させることで、参加プレイヤーの“自分の頭で考える力”を引き出し、村人側の勝率を引き上げることができないか検討中である。

引用文献

[1] Hung, H., & Chittaranjan, G. (2010, October). The idiap wolf corpus: exploring group behaviour in a competitive role-playing game. In Proceedings of the 18th ACM international conference on Multimedia (pp. 879-882). ACM.

[2] Kobayashi, Y., Osawa, H., Inaba, M., Shinoda, K., Toriumi, F., & Katagami, D. (2014, October). Development of werewolf match system for human players mediated with lifelike agents. In Proceedings of the second international conference on Human-agent interaction (pp. 205-207). ACM.

[3] Kuwamura, K., Sakai, K., Minato, T., Nishio, S., & Ishiguro, H. (2014, February). Hugvie: communication device for encouraging good relationship through the act of hugging, Lovotics, Vol. 1, Issue 1, pp. 10000104.

[4] Sumioka, H., Nakae, A., Kanai, R., & Ishiguro, H. (2013). Huggable communication medium decreases cortisol levels. Scientific reports, 3, 3034.