A Computational Model of Achievement Motivation for Artificial Agents (Extended Abstract)

Kathryn E. Merrick
University of New South Wales, Australian Defence Force Academy
School of Engineering and Information Technology
k.merrick@adfa.edu.au

ABSTRACT

Computational models of motivation are tools that artificial agents can use to autonomously identify, prioritize, and select the goals they will pursue. Previous research has focused on developing computational models of arousal-based theories of motivation, including novelty, curiosity and interest. However, arousal-based theories represent only one aspect of motivation. In humans, for example, curiosity is tempered by other motivations such as the need for health, safety, competence, a sense of belonging, esteem from others or influence over others. To create artificial agents that can identify and prioritize their goals according to this broader range of needs, new kinds of computational models of motivation are required. This paper expands our 'motivation toolbox' with a new computational model of achievement motivation for artificial agents. The model uses sigmoid curves to model approach of success and avoidance of failure. An experiment from human psychology is simulated to test the new model in virtual agents. The results are compared to human results and existing theoretical and computational models. Results show that virtual agents using our model exhibit statistically similar goal-selection characteristics to humans with corresponding motive profiles. In addition, our model outperforms existing models of achievement motivation in this respect.

Categories and Subject Descriptors

I.2 [**Artificial Intelligence**]: I.2.0 [**General**]: Cognitive simulation; I.2.11 [**Distributed Artificial Intelligence**]: Intelligent agents.

General Terms

Algorithms, Experimentation, Human Factors.

Keywords

Computational models of motivation, achievement motivation, cognitive agents, virtual agents, autonomous mental development.

1. ACHIEVEMENT MOTIVATION

Achievement motivation drives humans to strive for excellence by improving on personal and societal standards of excellence [1]. In artificial agents, achievement motivation has potential roles in driving the acquisition of skills and competencies in a domain-independent manner. Some existing work has been done with competence-based computational models of motivation [2, 3], but this work has focused on modeling competence in terms of learning error and identifying situations where there is an optimal potential for learning. In contrast, achievement motivation is based on estimations of success probabilities and task difficulty. This suggests an approach to goal-selection that is independent of learning. Other work has developed competence-based models specifically concerned with achievement motivation, but experimental results have indicated that these models cannot accurately reproduce the characteristic achievement-related responses seen in humans [4]. This paper takes a different approach to developing such a model, with more accurate results.

The foremost psychological model of achievement motivation is Atkinson's Risk-Taking Model (RTM) [5]. The RTM was designed to predict individual preferences for task difficulty. It defines achievement motivation in terms of conflicting person-specific desires to approach success M_s or avoid failure M_f and a situation-specific component for probability of success P_s:

$$T_r = (M_s - M_f)\,(P_s - P_s^2) \qquad (1)$$

The RTM has been an influential and successful aid to understanding achievement motivation in humans. However, to capture the subtleties of human behavior in an artificial system a more sensitive model is required. Thus, this paper draws on the ideas of probability of success and approach-avoidance motivation proposed by Atkinson, but uses sigmoid rather than quadratic functions to model the resultant tendency T_r for achieving a goal with a given probability of success P_s. Using sigmoid representations, approach motivation grows stronger as the probability of success increases, until a certain threshold probability is reached and approach motivation plateaus. Conversely, avoidance motivation is initially zero, and becomes a large negative number as probability for success increases. This means that failure at a very easy task is punished the most. The resultant tendency for achievement motivation is the sum of these hypothetical curves as follows:

$$T_r = \frac{1}{1 + e^{\rho^{+}(M^{+} - P_s)}} - \frac{1}{1 + e^{\rho^{-}(M^{-} - P_s)}} \qquad (2)$$

The model has five parameters M^+, M^-, ρ^+, ρ^- and P_s. M^+ and M^- are the turning points of the sigmoids for approach and avoidance motivation respectively. $\rho^+ > 0$ is the gradient of approach to success and $\rho^- > 0$ is the gradient of avoidance of failure.

Cite as: A Computational Model of Achievement Motivation for Artificial Agents (Extended Abstract), Merrick, K, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 1067 -1068.

2. THE RING-TOSS EXPERIMENT

The ring-toss game involves throwing a ring over a set distance to land over a spike. In psychology, the ring-toss experiment was originally designed to verify theories of achievement motivation in humans [6]. Because a player can stand different distances from the spike, the game defines a series of goals of different difficulty (and thus different probability of success). Psychologists hypothesize that individuals with different levels of achievement motivation will choose different distances from which to toss their ring. Atkinson and Litwin [6] conducted an investigation of the effects of achievement motivation in a ring-toss experiment. Individuals' tendency to approach success or avoid failure was gauged using the projective test of need achievement and Mandler-Sarason test. Individuals were then broken into four groups corresponding to four motivation types as follows:

- **H-L** high approach motivation and low avoidance motivation,
- **H-H** high motivation to approach success and avoid failure,
- **L-L** low motivation to approach success and avoid failure,
- **L-H** low approach motivation and high avoidance motivation.

Atkinson and Litwin [6] had forty-five human participants in their experiment and each was allowed ten opportunities to toss a ring at a peg from a distance of their choice in the range of 0 to 15 feet (approx 4.57 meters). Results were collated for each motivation type in three range-brackets for 'easy', 'moderate' and 'hard' goals. These brackets are shown in the first row of Table 3. When multiplied by the four motivation types, this gives a total of twelve experimental categories. Atkinson and Litwin's human experimental results are shown in the next four rows of Table 3.

Ring-toss experiments can also be designed for artificial agents that use a computational model of achievement motivation to compute a resultant tendency for each available goal, assuming that the probability of success of the goal is known. This paper compares the results of three such experiments to human results:

- EXPT 1: Agents using the RTM in Equation 1;
- EXPT 2: Agents using the Simkins et al. [4] model;
- EXPT 3: Agents using the new model in Equation 2.

By creating multiple agents of each model and randomizing their parameter values within limited ranges, agents with the four motivation types can be created. We used the parameter ranges in Tables 1 and 2 for EXPTs 1 and 3 respectively. Further details of the experimental setup for EXPT 3 are reported in [7]. Details of the experimental setup for EXPT 2 are reported in [4].

Table 1. Parameters and their value ranges for EXPT 1.

Param	H-L	H-H	L-L	L-H
M_s	[0.9, 1]	[0.9, 1]	[0.8, 0.9]	[0.8, 0.9]
M_f	[0, 0.1]	[0.2, 0.3]	[0, 0.1]	[0.2, 0.3]

Table 2. Parameters and their value ranges for EXPT 3.

Param	H-L	H-H	L-L	L-H
M^+	[0.1, 0.2]	[0.1, 0.2]	[0, 0.1]	[0, 0.1]
M^-	[0.8, 0.9]	[0.9, 1.0]	[0.8, 0.9]	[0.9, 1.0]
ρ^+	[0, 80]	[0, 100]	[0, 100]	[0, 100]
ρ^-	[0, 40]	[0, 50]	[0, 50]	[0, 90]

Table 3 reports the percentage of each type of agent to assign a maximal resultant tendency to goals in each bracket, and shows the z-value for all agent-human comparisons at the 95% confidence interval. The critical z-value for a two-tailed z-test of two proportions at the 95% confidence level is ±1.96. Results for EXPT 1 show that agents using the RTM have a maximum motivational tendency at $P_s = 0.5$ regardless of the values of other parameters. Thus all these agents choose 'moderate' goals. This experiment confirms that the RTM is inappropriate for use in artificial agents. Results for EXPT 2 summarize those reported by Simkins et al. [4] using z-values rather than confidence intervals. Using confidence intervals, their agents have statistically different performance to humans in eight of the twelve experimental categories. However z-values still indicate a statistical difference in five of the twelve categories. This result also supports the need for a more accurate model of achievement motivation, such as the one proposed in this paper. Finally, Table 3 shows that agents using the new sigmoid model in EXPT 3 produce statistically similar results to human studies in all twelve categories.

Table 3. Comparison of humans to agents using the RTM, Simkins' model and the new sigmoid model of achievement motivation. *indicates a statistically significant difference in results between humans and agents.

		0.00 – 2.00m (Easy)	2.25 – 3.50m (Moderate)	3.75 – 4.50m (Hard)
Human	H-L	11%	82%	7%
	H-H	26%	60%	14%
	L-L	18%	58%	24%
	L-H	32%	48%	20%
EXPT 1	H-L (Z)	0% (–3.890*)	100% (5.071*)	0% (–4.591*)
	H-H (Z)	0% (–5.467*)	100% (7.071*)	0% (–3.880*)
	L-L (Z)	0% (–4.219*)	100% (6.917*)	0% (–4.954*)
	L-H (Z)	0% (–7.037*)	100% (9.58*)	0% (–5.375*)
EXPT2	H-L (Z)	7.7% (–0.914)	75.4% (–1.300)	16.9% (0.418)
	H-H (Z)	14.0% (–2.121*)	69.0% (1.330)	17.0% (0.586)
	L-L (Z)	5.6% (–2.578*)	74.4% (2.326*)	20.0% (–0.648)
	L-H (Z)	8.5% (–4.714*)	80.0% (5.375*)	11.5% (–1.881)
EXPT 3	H-L (Z)	13.7% (0.850)	76.0% (–1.522)	10.3% (1.184)
	H-H (Z)	22.3% (–0.843)	67.6% (1.540)	10.1% (–1.215)
	L-L (Z)	11.5% (–1.815)	56.7% (–0.238)	31.8% (1.530)
	L-H (Z)	33.3% (0.296)	51.6% (0.772)	15.1% (–1.446)

3. REFERENCES

[1] Heckhausen, J. and Heckhausen, H., 2008. Motivation and action. New York: Cambridge University Press.

[2] Oudeyer, P-Y., Kaplan, F., and Hafner,V., 2007. Intrinsic motivation systems for autonomous mental development, IEEE Trans on Evolutionary Computation, 11(2):265-286.

[3] Merrick, K. and Maher, M. L., 2009. Motivated reinforcement learning: curious characters for multiuser games, Berlin: Springer.

[4] Simkins, C., Isbell, C., and Marquez, N., 2010. Deriving behavior from personality: a reinforcement learning approach. In Proceedings of the International Conference on Cognitive Modeling, Philadelphia, PA, pp. 229-234.

[5] Atkinson, J. W., 1957. Motivational determinants of risk-taking behavior, Psychological Review, 64:359-372.

[6] Atkinson, J. W. and Litwin, G. H., 1960. Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure, Journal of Abnormal and Social Psychology, 60:52-63.

[7] Merrick, K., Shafi, K., 2011, Achievement, affiliation and power: motive profiles for artificial agents, Ezequiel Di Paolo (Ed), Adaptive Behavior, (to appear)