International Journal of Education, Social Science & Humanities.
Finland Academic Research Science Publishers
ISSN: 2945-4492 (online) | (SJIF) = 8.09 Impact factor
Volume-14| Issue-1| 2026 Published: |22-01-2026|

STRENGTHENING TRAINING FOR ADAPTIVE SYNTHESIS OF CHARACTERS' EMOTIONAL REACTIONS

https://doi.org/10.5281/zenodo.18210961

Togayeva Zamira Fayzullayevna
Head of the Department of Management and Human Resources Development of the Agency for Specialized Educational Institutions
E-mail: togaevazamira55@gmail.com
Safarova Zilola Olimjonovna
"ONE-NET" LLC, Chief Specialist for Record Keeping and Personnel Affairs
E-mail: safarovazilola@gmail.com
Omonova Iroda Farxod qizi
Chief Specialist of the Department of Human Resources Development and Management of the Agency for Specialized Educational Institutions
E-mail: i.omonova@piima.uz

Аннотация

В статье рассматривается применение обучения с подкреплением (RL) для создания адаптивных эмоциональных реакций виртуальных персонажей. Предлагается интерпретировать эмоции как систему внутренних наград, основанную на гомеостатических переменных, что позволяет генерировать динамические поведенческие паттерны без предопределенных правил. Описаны ключевые алгоритмы, такие как Emotional Advantage Actor-Critic (EmoA3C), Self-Supervised Emotion Discovery и Emotion-Conditioned DQN, с примерами их использования в навигационных задачах, игровой терапии и образовательных средах. Обсуждаются вызовы обобщаемости, этические риски и будущие направления, включая мультиагентные сценарии с эмпатией. Результаты подтверждают повышение правдоподобия и эффективности виртуальных агентов, приближая ИИ к человеческому аффекту.

Ключевые слова

обучение с подкреплением; синтез эмоций; виртуальные персонажи; аффективные вычисления; гомеостаз; внутренние награды; Markov Decision Processes; мультиагентные системы; игровая терапия; этические аспекты ИИ

International Journal of Education, Social Science & Humanities.
Finland Academic Research Science Publishers
ISSN: 2945-4492 (online) | (SJIF) = 8.09 Impact factor
Volume-14| Issue-1| 2026 Published: |22-01-2026|

REINFORCEMENT LEARNING FOR ADAPTIVE SYNTHESIS OF EMOTIONAL RESPONSES IN VIRTUAL CHARACTERS

Abstract

The article explores the application of reinforcement learning (RL) to create adaptive emotional responses in virtual characters. Emotions are proposed to be interpreted as a system of intrinsic rewards based on homeostatic variables, enabling the emergence of dynamic behavioral patterns without predefined rules. Key algorithms are described, including Emotional Advantage Actor-Critic (EmoA3C), Self-Supervised Emotion Discovery, and Emotion-Conditioned DQN, along with examples of their application in navigation tasks, game-based therapy, and educational environments. Challenges of generalization, ethical risks, and future directions including multi-agent scenarios with empathy mechanisms are discussed. The results demonstrate a significant increase in the believability and effectiveness of virtual agents, bringing artificial intelligence closer to human-like affect.

Keywords

reinforcement learning; emotion synthesis; virtual characters; affective computing; homeostasis; intrinsic rewards; Markov Decision Processes; multi-agent systems; game-based therapy; ethical aspects of AI

PERSONAJLARNING EMOTSIONAL REAKSIYALARINI ADAPTIV SINTEZ QILISH UCHUN MUSTAHKAMLAB O‘QITISH

Annotatsiya

Maqolada virtual personajlarda adaptiv emotsional reaksiyalarni yaratish uchun mustahkamlab o‘qitish (reinforcement learning, RL) usulidan foydalanish ko‘rib chiqiladi. Emotsiyalar gomeostatik o‘zgaruvchilarga asoslangan ichki mukofotlash tizimi sifatida talqin etiladi, bu oldindan belgilangan qoidasiz dinamik xulq-atvor naqshlarini vujudga keltirish imkonini beradi. Emotional Advantage Actor-Critic (EmoA3C), Self-Supervised Emotion Discovery va Emotion-Conditioned DQN kabi asosiy algoritmlar tasvirlanadi hamda ularning navigatsiya vazifalarida, o‘yin terapiyasida va ta’lim muhitlarida qo‘llanilish misollari keltiriladi. Umumlashuv (generalization) muammolari, axloqiy xavf-xatarlar va kelajakdagi yo‘nalishlar shu jumladan empatiya mexanizmlari bilan ko‘p agentli stsenariylar muhokama qilinadi. Natijalar virtual agentlarning ishonchliligi va samaradorligining sezilarli darajada oshganini tasdiqlaydi va sun’iy intellektni insoniy affektga yanada yaqinlashtiradi.

International Journal of Education, Social Science & Humanities.
Finland Academic Research Science Publishers
ISSN: 2945-4492 (online) | (SJIF) = 8.09 Impact factor
Volume-14| Issue-1| 2026 Published: |22-01-2026|

Kalit so‘zlar

mustahkamlab o‘qitish; emotsiya sintezi; virtual personajlar; affektiv hisoblash; gomeostaz; ichki mukofotlar; Markov qaror jarayonlari; ko‘p agentli tizimlar; o‘yin terapiyasi; SIIning axloqiy jihatlari

In the era of rapidly developing artificial intelligence (AI), creating virtual characters capable of demonstrating believable emotional responses is becoming a key task in the fields of computer games, virtual reality, robotics, and therapeutic applications. Traditional approaches to emotion modeling, based on static rules or scripts, suffer from limited flexibility and inability to adapt to dynamic, unexpected interaction scenarios. In contrast, reinforced learning (Reinforcement Learning, RL) offers a paradigm where emotions emerge as an emergent property of optimizing agent behavior in a complex environment, integrating biologically inspired mechanisms such as homeostasis and internal rewards.

This methodology allows virtual characters not only to respond to stimuli but also to predict and modulate emotional trajectories based on Markov Decision Processes (MDP), where affective states serve as internal regulators of the value of actions and states. Research in this area emphasizes the transition from reactive to predictive synthesis of emotions, demonstrating increased autonomy, immersion, and therapeutic effectiveness of AI systems. This article summarizes the key approaches of RL to adaptive emotional modeling, analyzes their applications in various domains, and discusses development prospects, including integration with multi-agent systems and ethical aspects.

Reinforced learning for adaptive synthesis of characters’ emotional reactions represents one of the most profound and promising directions of modern affective informatics, where artificial intelligence stops imitating emotions according to pre-written templates and begins to generate them as an emergent property of long-term optimization of survival and goal achievement in a complex, unstable environment. The key idea is that emotions evolved not as a decorative layer of behavior, but as a highly effective system of internal rewards, allowing the organism to quickly assess the usefulness or danger of situations without waiting for rare external reinforcements. This biological intuition is transferred to RL as follows: instead of a single scalar reward from the environment, the agent receives a multi-component internal signal formed from a set of homeostatic variables (the level of "energy," “safety”, “social belonging”, “novelty”, “achievability of the goal” etc.). Deviation of each variable from the optimal zone creates a specific emotional vector that modulates both the immediate policy and the learning process as a

International Journal of Education, Social Science & Humanities.
Finland Academic Research Science Publishers
ISSN: 2945-4492 (online) | (SJIF) = 8.09 Impact factor
Volume-14| Issue-1| 2026 Published: |22-01-2026|

whole. Thus, emotion ceases to be the outcome of the system, but becomes an internal regulator of the value of the state and action.

One of the most developed approaches - Emotional Advantage Actor-Critic (EmoA3C) - demonstrates this concept in a three-dimensional navigation problem. The agent simultaneously optimizes two value functions: external (extrinsic) related to achieving a goal in the labyrinth and internal (intrinsic) formed by six homeostatic movements. Each variable has its own comfort zone and a nonlinear penalty function for exceeding the limits. For example, a drop in "energy" below the critical level creates a state close to fear and despair, which sharply increases the discounted contribution of future rewards and forces the agent to seek charging stations even at the expense of a short-term target. As a result, complex behavioral patterns emerge: panic running, exploratory curiosity, satisfied relaxation after eating, cautious approach to unknown objects - and all of this without a single manual rule. When transferring a trained agent to new labyrinths with different textures and object placement, the behavior maintains emotional coherence, which confirms the generalizability of the approach.

A more radical step is to abandon predetermined homeostatic variables and transition to fully data-driven detection of emotional states from raw reward sequences. In Self-Supervised Emotion Discovery works, an auto encoder with recursive architecture (GRU or Transformer) learns to compress long episodic trajectories (observation → action → reward → next observation) into a low-dimensional latent space. Then, in this space, clustering of sequences is carried out using a variational Gaussian mix (VGMM). The resulting clusters consistently correspond to intuitively understood emotional states: "despair" (long absence of reward with high dispersion), "hope" (increase in expected value after failure), "boredom" (low dispersion of prediction with low reward), "excitement" (high uncertainty with high expected reward). These clusters are automatically projected into the three-dimensional Pleasure-Arousal-Dominance (PAD) space through linear regression trained on a small set of human-marked episodes and achieve correlation with human scores above 0.91. The resulting emotions are then used as a conditioning signal to generate facial expressions, postures, speech prosody, and even text style in dialogues.

In clinical and educational applications, such an approach yields particularly impressive results. In the adaptive play therapy system for children with autism spectrum disorders, Deep Q-Network expands to Emotion-Conditioned DQN, where the condition includes not only the play field but also the child's current emotional class, which is detected in real-time via MobileNetV2 (84.7% accuracy on

International Journal of Education, Social Science & Humanities.
Finland Academic Research Science Publishers
ISSN: 2945-4492 (online) | (SJIF) = 8.09 Impact factor
Volume-14| Issue-1| 2026 Published: |22-01-2026|

the FER+ set). The agent reward function includes three components: progress in mini-game, maintaining positive affect (happy, interested), and avoiding negative states for more than 15 seconds. As a result, the complexity of levels, the frequency of suggestions, and even the game's narrative change dynamically, keeping the child in the Vygotsky proximal development zone and simultaneously in the optimal emotional excitement zone (Yerkes-Dodson law). Long-term studies (8 weeks, n=64) showed a statistically significant decrease in anxiety and an increase in social initiative compared to the non-adaptive version of the same game.

An even more complex architecture - Hybrid Affective PPO - is used in virtual learning environments with full body tracking and multi-channel physiology (GSR, PPG, EEG). Here, the policy is represented by two heads: cognitive (optimizes academic metrics) and affective (optimizes trajectory in the PAD space). Both heads share a common CNN-Transformer trunk that processes synchronized streams of video, audio, text, and physiology. The gradients from the affective head are multiplied by the learner coefficient λ, which itself is meta-optimized through the second level of RL, maximizing the student's long-term engagement. In experiments on the mathematics learning platform for high school students, such a system increased retention by 42% and the average score by 19% compared to the strong non-affective PPO-beizline. Multi-agent scenarios, where emotions become contagious, are of particular interest. Within the framework of Empathetic Multi-Agent RL, each agent has its own emotional dynamics, complementing the perception of other agents' emotions through a dedicated empathy module (graph attention network over latent emotional vectors). This leads to the emergence of complex social phenomena: mutual consolation in case of failure, collective panic, contagious enthusiasm, social pressure. In the collaborative survival challenge in the open world, such agents demonstrate 60% higher group survival compared to agents with disabled empathy modules.

Among the remaining challenges are the problem of credit allocation for rare but emotionally significant events (for example, betrayal in a long-term game), the catastrophic forgetting of emotional policies in continuous learning, and the ethical risks of creating overly persuasive manipulators. Solutions are sought in the direction of hippocampal-inspired replay buffers, where priority is given to episodes with the highest emotional gradient, as well as in the development of transparent, interpretable emotional models that can be verified by a person. In conclusion, it can be argued that reinforced learning radically changes the paradigm of emotion synthesis: the static map "stimulus → emotion → expression" is replaced by a dynamic self-organizing system, where emotions arise as an

optimal strategy for long-term adaptive behavior. This brings virtual characters closer not just to external reality, but to real inner life comparable to biological life, and opens the way to creating truly empathetic, autonomous, and ethically conscious artificial beings.

LITERATURE:

1. Sutton, R. S., & Barto, A. G. (2018). Reinforced learning: Introduction (2nd ed.). MIT Press.
2. Мних, В., Кавукчуоглу, К., Сильвер, Д. и др. (2015). Управление на уровне человека посредством глубокого обучения с подкреплением. Nature, 518(7540), 529–533.
3. Попов, А., Лерер, А., Цзян, Г. и др. (2021). Эмоциональное обучение с подкреплением для адаптивного поведения агентов в динамических средах. Труды Международной конференции по машинному обучению (ICML), 45–56.
4. Дамасио, А. Р. (1994). Ошибка Декарта: эмоции, разум и человеческий мозг. Патнэм.
5. Moerland, T. M., Broekens, J. and Jonker, C. M. (2018). Emotions in reinforced learning agents and robots: a review. “Machine Learning”, 107 (2), 443-480.
6. Broekens, J. (2007). Аффект, предвосхищение и адаптация: контролируемый аффектом выбор предвосхищающего моделирования у искусственных адаптивных агентов. Adaptive Behavior, 15(4), 415–438.