

*Xudoyorov Shaxriyor*

*Termiz iqtisodiyot va xizmat ko'rsatish universiteti, Iqtisodiyot va axborot  
texnologiyalari fakulteti, Axborot tizimlari va texnologiyalari kafedrası, 3-kurs  
talabasi*

*Ilmiy rahbar: Ernazarov Mirzohid*

*E-mail: [xudoyorovshaxriyor57@gmail.com](mailto:xudoyorovshaxriyor57@gmail.com)*

**Annotatsiya:** Ushbu ilmiy maqolada sun'iy intellekt (SI)ning eng kuchli yo'nalishlaridan biri bo'lgan Mustahkamlovchi o'rganish (Reinforcement Learning, RL) va uning o'yinlardan to real dunyo tizimlarini boshqarishgacha bo'lgan qo'llanilishi har tomonlama tahlil qilinadi. Boshqa mashinaviy o'rganish usullaridan farqli o'laroq, mustahkamlovchi o'rganishda agent o'z muhiti bilan sinov va xato orqali o'zaro ta'sirlashib, mukofotlar tizimi rahbarligida optimal xulq-atvorni o'rganadi. Tadqiqotda mustahkamlovchi o'rganishning nazariy asoslari, uning o'yinlardagi (Go, shaxmat, video o'yinlar) hayratlanarli muvaffaqiyatlari va real dunyo masalalariga (robototexnika, energiya boshqaruvi, avtonom tizimlar) tatbiq etilishining istiqbollari va muammolari ilmiy asoslanadi. Maqolaning ilmiy yangiligi mustahkamlovchi o'rganish mustaqil qaror qabul qila oladigan tizimlarni yaratish uchun asosiy texnologiya ekanligini ko'rsatishdan iborat. Tahlillar natijasida ushbu yo'nalishning istiqbollari bo'yicha tavsiyalar ishlab chiqiladi.

**Kalit so'zlar:** Mustahkamlovchi o'rganish, sun'iy intellekt, agent, mukofot, siyosat, chuqur mustahkamlovchi o'rganish, robototexnika, boshqaruv tizimlari, AlphaGo, mashinaviy o'rganish, qaror qabul qilish, neyron tarmoqlar.

## **Kirish**

Mashinaviy o'rganishning aksariyat usullari tayyor ma'lumotlar asosida o'rganadi: ularga misollar va to'g'ri javoblar ko'rsatiladi, model esa shulardan o'rganadi. Biroq, real hayotda o'rganish ko'pincha boshqacha tarzda — sinov va xato orqali sodir bo'ladi. Bola ko'p marta yiqilib, yurishni o'rganadi; inson velosipedda uchishni ko'p marta urinish orqali o'rganadi. Bu jarayonda oldindan to'g'ri javobni beradigan "o'qituvchi" yo'q; aksincha, o'rganish o'z harakatlarining natijalari — muvaffaqiyat yoki muvaffaqiyatsizlik — bilan boshqariladi. Aynan shu turdagi o'rganishni sun'iy

intellektning Mustahkamlovchi o'rganish (Reinforcement Learning, RL) deb ataluvchi yo'nalishi taqlid qiladi.

Mustahkamlovchi o'rganishda "agent" (SI tizimi) "muhit" bilan o'zaro ta'sirlashadi. Agent harakatlarni amalga oshiradi, javobida esa muhit o'z holatini o'zgartiradi va agentga "mukofot" yoki "jazo" beradi. Agentning maqsadi — uzoq muddatda umumiy mukofotni maksimal darajaga ko'taradigan harakatlarni tanlashdir. Sinov va xato orqali agent qaysi vaziyatda qaysi harakatlar yaxshi natijaga olib kelishini asta-sekin o'rganadi — ya'ni u optimal "siyosat"ni ishlab chiqadi. Bu yondashuv, ayniqsa, ketma-ket qarorlar qabul qilishni talab qiladigan masalalar uchun juda mos keladi.

Ushbu maqolaning dolzarbligi shundaki, mustahkamlovchi o'rganish so'nggi yillarda hayratlanarli muvaffaqiyatlarga erishdi. Ushbu usulga asoslangan SI tizimlari Go va shaxmat kabi murakkab o'yinlarda dunyoning eng yaxshi o'yinchilarini mag'lub etdi va video o'yinlarni insondan yuqori darajada o'ynashni o'rgandi. Endi bu muvaffaqiyatlarni real dunyo masalalariga — robototexnika, energiya boshqaruvi va avtonom tizimlarga tatbiq etish bo'yicha sa'y-harakatlar olib borilmoqda. Ushbu maqolada mustahkamlovchi o'rganishning asoslari, uning o'yinlardagi muvaffaqiyatlari va real dunyo masalalariga tatbiq etilishining istiqbollari tahlil qilinadi.

## Muhokama

Mustahkamlovchi o'rganish SI hamjamiyatida eng faol muhokama qilinayotgan mavzulardan biridir. Asosiy muhokama predmeti — o'yinlarda erishilgan muvaffaqiyatlarni real dunyoga o'tkazishdagi qiyinchilikdir.

Birinchidan, o'yinlar va real dunyo o'rtasidagi farq. O'yinlarda qoidalar aniq, muhit to'liq kuzatiladigan va millionlab urinishlarni tez hamda xavfsiz tarzda amalga oshirish mumkin. Real dunyoda esa muhit murakkab va bashorat qilib bo'lmaydigan bo'lib, xato qilish qimmatga tushishi yoki xavfli bo'lishi mumkin (masalan, robotning buzilishi yoki avariya sodir bo'lishi). Shu sababli, o'yinlarda ishlaydigan usullarni real dunyoga o'tkazish murakkab vazifadir.

Ikkinchidan, namuna samaradorligi muammosi. Mustahkamlovchi o'rganish ko'pincha o'rganish uchun millionlab urinishlarni talab qiladi. Bu kompyuterda tez o'ynaladigan o'yinda maqbul, biroq real dunyoda millionlab urinishlarni amalga oshirish imkonsiz yoki juda qimmatga tushadi. Shu sababli, kam sonli urinishlardan o'rganish usullari va simulyatsiyalardan foydalanish muhokamaning markazida turibdi.

Uchinchidan, xavfsizlik masalasi. Mustahkamlovchi o'rganish agenti o'rganish jarayonida turli harakatlarni sinab ko'radi, ularning ba'zilari xavfli bo'lishi mumkin.

Real dunyoda bu qabul qilinishi mumkin emas. Shu sababli, "xavfsiz mustahkamlovchi o'rganish" sohasi — agent o'rganish jarayonida xavfli harakatlarni amalga oshirmasligini ta'minlash — rivojlantirilmoqda.

## Asosiy qism

**1. Mustahkamlovchi o'rganishning asoslari.** Mustahkamlovchi o'rganish bir nechta asosiy tushunchalarga asoslanadi. Agent — o'rganadigan va qaror qabul qiladigan SI tizimidir. Muhit — agent o'zaro ta'sirlashadigan dunyodir. Holat — muhitning hozirgi vaziyatidir. Harakat — agentning bajaradigan ishidir. Mukofot — agentning harakatiga javoban oladigan signalidir. Siyosat — agentning qaysi holatda qaysi harakatni tanlash strategiyasidir. Agentning maqsadi — uzoq muddatda umumiy mukofotni maksimal darajaga ko'taradigan optimal siyosatni topishdir.

**2. Chuqur mustahkamlovchi o'rganish.** Mustahkamlovchi o'rganish chuqur neyron tarmoqlar bilan birlashtirilganda, chuqur mustahkamlovchi o'rganish vujudga keladi. Bu birlashma agentga murakkab, yuqori o'lchamli ma'lumotlardan (masalan, tasvirning pikselleridan) bevosita o'rganish imkonini beradi. Aynan shu yondashuv SI'ga video o'yinlarni ekrandagi tasvirdan bevosita o'rganib o'ynashga va Go kabi murakkab o'yinlarni egallashga imkon berdi. Chuqur neyron tarmoqlar agentga murakkab holatlarni tushunish va mos harakatlarni tanlashga yordam beradi.

**3. O'yinlardagi muvaffaqiyatlar.** Mustahkamlovchi o'rganish o'zining eng yorqin muvaffaqiyatlariga o'yinlarda erishdi. AlphaGo tizimi uzoq vaqt davomida eng murakkab o'yinlardan biri va kompyuterlar tomonidan yengib bo'lmaydi deb hisoblangan Go o'yinida dunyoning eng yaxshi o'yinchisini mag'lub etdi. Keyinroq, AlphaZero tizimi hech qanday inson bilimisiz, faqat o'zi bilan o'ynash orqali shaxmat, Go va shogi o'yinlarini insondan yuqori darajada egalladi. SI tizimlari shuningdek, murakkab video o'yinlarni yuqori darajada o'ynashni ham o'rgandi. Bu muvaffaqiyatlar mustahkamlovchi o'rganishning kuchini namoyish etdi.

**4. Real dunyo masalalariga tatbiq etish.** Endi mustahkamlovchi o'rganishning muvaffaqiyatlarini real dunyo masalalariga tatbiq etish bo'yicha sa'y-harakatlar olib borilmoqda. Robototexnikada robotlar narsalarni ushlashni, yurishni va murakkab harakatlarni bajarishni o'rganmoqda. Energiya boshqaruvida mustahkamlovchi o'rganish ma'lumotlar markazlarining sovutish tizimlarini optimallashtirib, energiyani tejamoqda. Avtonom tizimlarda u transport vositalari va dronlarni boshqarish uchun qo'llaniladi. Sanoatda u ishlab chiqarish jarayonlarini optimallashtiradi. Bu qo'llanishlar hali rivojlanmoqda, biroq istiqbollar kengdir.

**5. Istiqbollar va O'zbekistondagi vaziyat.** Mustahkamlovchi o'rganish mustaqil qaror qabul qila oladigan tizimlarni yaratish uchun asosiy texnologiyadir. Xavfsizlik va

namuna samaradorligi usullari takomillashgani sari, uning real dunyo masalalariga tatbiq etilishi kengayadi. O'zbekiston uchun ushbu sohani egallash robototexnika, sanoatni avtomatlashtirish va energiya boshqaruvining rivojlanishi uchun muhimdir. Buning uchun ushbu sohada mutaxassislar tayyorlash va tadqiqotlarni rivojlantirish zarur.

## Xulosa

Mustahkamlovchi o'rganish sun'iy intellektning eng kuchli yo'nalishlaridan biri bo'lib, sinov va xato orqali o'rganishning tabiiy jarayonini taqlid qiladi. Go va shaxmat kabi murakkab o'yinlarda hayratlanarli muvaffaqiyatlarga erishish orqali, bu usul o'z kuchini namoyish etdi va endi bu muvaffaqiyatlarni real dunyo masalalariga tatbiq etish bo'yicha sa'y-harakatlar olib borilmoqda.

Bizning tadqiqotimiz shuni ko'rsatadiki, o'yinlarda erishilgan muvaffaqiyatlarni real dunyoga o'tkazish murakkab vazifadir. O'yinlar va real dunyo o'rtasidagi farq, namuna samaradorligi muammosi va xavfsizlik masalasi dolzarb muammolardir. Xavfsiz va samarali o'rganish usullarini ishlab chiqish ushbu sohaning real dunyo masalalariga tatbiq etilishi uchun hal qiluvchi ahamiyatga ega.

Xulosa qilib aytganda, mustahkamlovchi o'rganish mustaqil qaror qabul qila oladigan tizimlarni yaratish uchun asosiy texnologiyadir. O'zbekiston uchun ushbu sohani egallash robototexnika, sanoatni avtomatlashtirish va energiya boshqaruvining rivojlanishi uchun muhimdir.

## Foydalanilgan adabiyotlar

1. Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction (2nd ed.). MIT Press.
2. Mnih, V., et al. (2015). "Human-level control through deep reinforcement learning." Nature.
3. Silver, D., et al. (2016). "Mastering the game of Go with deep neural networks and tree search." Nature.
4. Silver, D., et al. (2017). "Mastering the game of Go without human knowledge." Nature.
5. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
6. Silver, D., et al. (2018). "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play." Science.
7. Levine, S., et al. (2016). "End-to-End Training of Deep Visuomotor Policies." Journal of Machine Learning Research.

8. Schulman, J., et al. (2017). "Proximal Policy Optimization Algorithms." arXiv preprint.
9. Vinyals, O., et al. (2019). "Grandmaster level in StarCraft II using multi-agent reinforcement learning." Nature.
10. Lillicrap, T. P., et al. (2016). "Continuous control with deep reinforcement learning." ICLR.
11. Kober, J., et al. (2013). "Reinforcement learning in robotics: A survey." The International Journal of Robotics Research.
12. Russell, S., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
13. Garcia, J., & Fernandez, F. (2015). "A Comprehensive Survey on Safe Reinforcement Learning." Journal of Machine Learning Research.
14. Haarnoja, T., et al. (2018). "Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning." ICML.
15. O'zbekiston Respublikasi Prezidentining "Raqamli O'zbekiston 2030; strategiyasini tasdiqlash to'g'risida"gi Farmoni (2020-yil 5-oktyabr).
16. Mnih, V., et al. (2016). "Asynchronous Methods for Deep Reinforcement Learning." ICML.
17. UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence.
18. Bostrom, N. (2014). Superintelligence: Paths, Dangers, Strategies. Oxford University Press.