

лимнологических характеристик водных экосистем Центрального Ямала // Высокопроизводительные вычислительные системы и технологии, 2021, Т. 5, №1.

5. Efanov A.V., Kovalevskaya N.M., Kulshin A.V., Khvorova L.A. Monitoring of lake dynamics and assessment of limnological characteristics of aquatic ecosystems of the Central Yamal // Journal of Physics: Conference Series (JPCS), 2021.

УДК 004.85.032.26

Реализация программных агентов для поиска неочевидных игровых стратегий на основе нейронной сети с подкреплением

Д.С. Козлов, О.Н. Половикова

Алтайский государственный университет, г.Барнаул

На современном этапе развития науки технологий нейронные сети активно используются в различных прикладных областях от разработки игр до имитации поведения живых существ. Целью данного исследования является реализация двух программных агентов для генерации неочевидных для человека стратегии на примере одной игры (игра крестики-нолики пять в ряд). Такие стратегии в дальнейшем позволят существенно обогатить банки игрового поведения.

Ключевые слова: *игровая стратегия, нейронная сеть с подкреплением, программный агент*

На сегодняшний день методы машинного обучения с подкреплением являются одним из самых перспективных направлений в области искусственного интеллекта. Обучение с подкреплением применяется в тех случаях, когда нужно выбрать лучший вариант среди многих или достичь сложной цели за множество ходов. Алгоритмы подкрепления, которые включают в себя глубокое обучение, могут победить чемпионов мира в игре Go, начиная с базового понимания правил игры и тренируясь от партии к партии. Таким образом, это уже искусственный интеллект в действии: машина пытается решить задачу разными способами, ошибается, учится на своих ошибках, улучшает показатели. Типичным примером подобного способа компьютерного самообучения являются системы поддержки принятия решений в реальном времени.

Целью данного исследования является реализация двух программных агентов для генерации неочевидных для человека стратегии на примере одной игры (игра крестики-нолики пять в ряд).

Одна из основных идей подобных систем состоит в том, что изначально агенты обучения не имеют представления о правилах, действующих в той среде, в которой существует агент.

На первом этапе необходимо создать саму среду, в которой будут взаимодействовать агенты.

На втором этапе создателю системы необходимо задать целевую функцию для агента, которую необходимо максимизировать (или минимизировать). Первая нейронная сеть будет обучаться практически вслепую. Ей будет известно только то, что составление пяти своих фигур в ряд ведет к победе, для ускорения процесса обучения нейронная сеть будет получать небольшой штраф за каждый сделанный ход.

Ниже представлены основные принципы, по которым будет строиться целевая функция обучения:

- Чем быстрее будет завершена партия – тем выше будет значение целевой функции.
- За победу в партии нейронная сеть будет получать +1000 очков.
- За создание ситуации с постановкой четырех фигур в ряд после своего хода агент получит +800 очков.
- За каждый свой ход агент будет терять по 10 очков.
- За каждый ход соперника агент будет терять 20 очков.
- Начальное значение целевой функции равно 0 очков.
- За поражение в партии агент будет терять 1000 очков.
- За допущение постановки паттерна из четырех открытых фигур соперника в ряд агент будет получать штраф в 800 очков.

Задача агента – максимизировать целевую функцию по окончании партии.

На третьем этапе необходимо создать события, за которые мы будем награждать или наказывать искусственный интеллект. В ходе процесса обучения агенты, достигшие наилучших результатов, будут выбраны в качестве основы для следующего поколения обучения. Схема работы агентов представлена на рисунке 1.

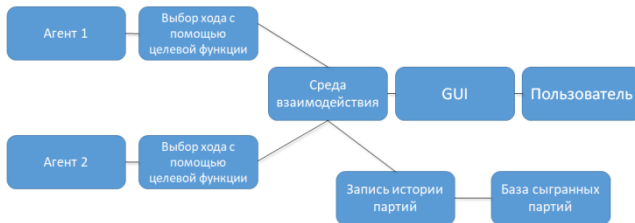


Рисунок 1 – Схема работы агентов

Система «награждения-наказания» программного агента определяет весь процесс обучения нейронной сети. Заранее невозможно предсказать результат обучения с подкреплением ввиду случайности принимаемых решений на ранних этапах обучения. Вероятен следующий исход: на раннем этапе стратегия, наиболее слабая в долгосрочной перспективе, приносит наиболее высокий выигрыш на раннем этапе игры, искусственный интеллект видит, что у данной стратегии наиболее высокий выигрыш, и отбрасывает остальные стратегии, и дальнейшая эволюция обучения происходит на основе заведомо наиболее слабой стратегии в долгосрочной перспективе. Возможные способы решения данной проблемы заключаются в изменении правил игры (как отдельных количественно для отдельных параметров, так и добавление новых параметров, создание новых ключевых событий в игре, переоценка важности целевой функции, применение комбинированных алгоритмов.

Библиографический список

1. Вагин В.Н., Еремеев А.П. Некоторые базовые принципы построения интеллектуальных систем поддержки принятия решений реального времени // Изв. РАН. Теория и система управления. 2001. № 6. С. 114–123.
2. Еремеев А.П., Кожухов А.А. Разработка интегрированной среды на основе методов прогнозирования и обучения с подкреплением для интеллектуальных систем реального времени // IS&IT'16: тр. Конгресса по интелект. сист. и информ. технологиям. Науч. изд. в 3 т. Таганрог: Изд-во ЮФУ, 2016. Т. 1. С. 140–149.
3. Рыбина Г.В., Паронджанов С.С. Технология построения динамических интеллектуальных систем: учеб. пособие. М.: Изд-во НИЯУ МИФИ, 2011.
4. Осипов Г.С. Методы искусственного интеллекта. М.: Физматлит, 2011.

5. Busoniu L., Babuska R., and De Schutter B. Multi-agent reinforcement learning: An overview. Chapter 7 in Innovations in Multi-Agent Systems and Applications-1 (D. Srinivasan and L.C. Jain, eds.), vol. 310 of Studies in Computational Intelligence, Berlin, Germany: Springer, 2010, pp. 183–221.

УДК 591.526:51

Изменение скорости роста популяции северного морского котика о. Тюлений и стационарной возрастной структуры самцов: результаты детерминистской и стохастической модели

**О.Л. Жданова^{1,2}, В.С. Жданов³, А.Е. Кузин⁴,
Е.Я. Фрисман²**

¹ИАПУ ДВО РАН, г. Владивосток; ²ИКАРП ДВО РАН, г. Биробиджан; ³ДВФУ, г. Владивосток; Тихоокеанский филиал ФГБНУ "ВНИРО" ("ТИНРО"), г. Владивосток

В статье на основе матричной модели динамики наблюдаемой части популяции (самцов), проанализирована скорость роста популяции северного морского котика на фоне совокупного изменения рождаемости и выживаемости с учетом стохастических эффектов.

Ключевые слова: *матричная модель, северный морской котик, Callorhinus Ursinus, оценка параметров, стохастическая скорость роста популяции.*

На фоне многолетнего управляемого промысла в популяции северного морского котика о. Тюленьего, как и в других популяциях Серевной Пацифики [1-3], появились признаки депрессии; в частности, численность новорожденных щенков уменьшилась вдвое и практически стабилизировалась на этом низком уровне. Несмотря на последовавшее значительное ограничение промысла, а позднее и его полный запрет, ожидаемого восстановления рождаемости в популяции не произошло [4].

В работе приведен анализ гаремных отношений и динамики рождаемости в популяции; обнаружена связь этого параметра с интенсивностью промысла, что позволило разделить весь ряд наблюдений на периоды различной популяционной динамики. С помощью матричной модели динамики наблюдаемой части популяции (самцов), проанализирована скорость ее роста на фоне совокупного изменения рождаемости и выживаемости с учетом стохастических эффектов. В