

Анализ модели эволюции популяции самообучающихся агентов на базе нейросетевых адаптивных критиков

Выполнила: Овчинникова Т.С., 210 гр., ВНД

Основные понятия

- | Анимат (агент)
- | Модель
- | Критик
- | Нейронные сети (НС)
- | Функциональные системы (ФС)
- | Метод обучения с подкреплением
- | Генетический алгоритм как метод эволюционной оптимизации

Параметры модели

$S(t)$ – внешнее воздействие

$r(t)$ – подкрепление

$x(t)$ – сила тока

Переменные

$a(t)$ – реакция, действия анимата

$C(t)$ – суммарное количество

ресурсов анимата

Оценка суммарной награды

$$U(t) = \sum_{k=0}^{\infty} \gamma^k r(t+k),$$

Модель

$$x^M = \{\Delta X(t-m+1), \dots, \Delta X(t)\},$$

$$y_j^M = \tanh \sum_i w_{ij}^M x_i^M,$$

$$\Delta X^{pr}(t+1) = \sum_j v_j^M y_j^M$$

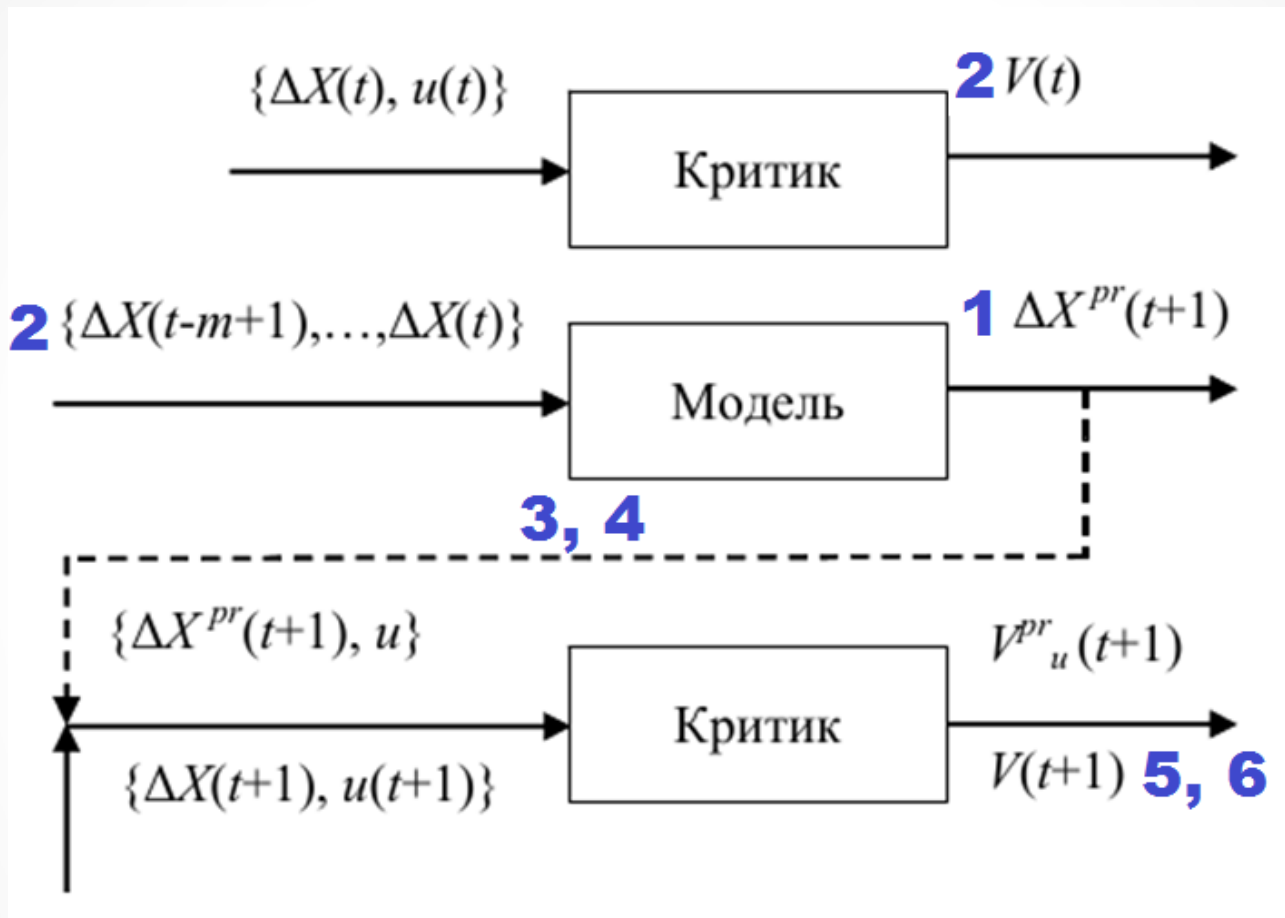
Критик

$$x^C = S(t) = \{\Delta X(t), u(t)\},$$

$$y_j^C = \tanh \sum_i w_{ij}^C x_i^C,$$

$$V(t) = V(S(t)) = \sum_j v_j^C y_j^C$$

Схема системы управления агента

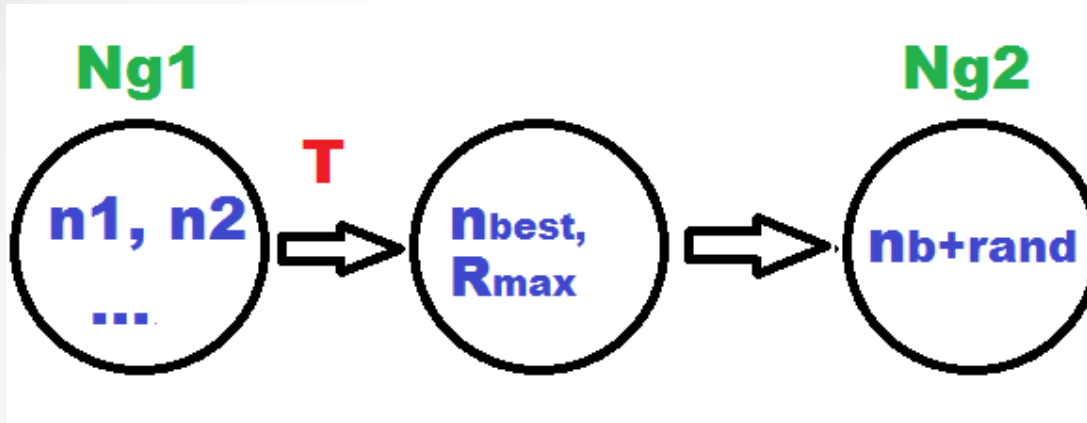


- **1. Ввод в Critic** — строится поток значений $V(t)$ (сигнал ошибки) для критической функции. **2. Ввод в Model** — строится поток значений $\Delta X^{pr}(t+1)$ (прогнозы) для критической функции. **3. Ввод в Critic** — строится поток значений $V(t+1)$ (сигнал ошибки) для критической функции. **4. Ввод в Critic** — строится поток значений $V_u^{pr}(t+1)$ (прогнозы) для критической функции. **5. Ввод в Critic** — строится поток значений $V(t+1)$ (сигнал ошибки) для критической функции. **6. Ввод в Critic** — строится поток значений $\{\Delta X(t+1), u(t+1)\}$ (сигнал ошибки) для критической функции.

Схема системы управления агента

- 1. **Модель предсказывает** следующее изменение временного ряда (ситуации)
- 2. **Критик оценивает** качество (характер) текущей ситуации и качество ситуаций для обоих возможных действий
- 3. **Модель выбирает действие** (работает ϵ -жадное правило)
- 4. **Выбранное действие выполняется.** Происходит переход к следующему моменту времени. Наблюдаемое значение характера ситуации сравнивается с предсказанным. Веса НС подстраиваются так, чтобы минимизировать ошибку. **Происходит обучение Модели.**
- 5. **Критик подсчитывает** качество ситуации в момент времени $t+1$, рассчитывает ошибку временной разности.
- 6. Веса НС **Критика** подстраиваются так, чтобы минимизировать ошибку. **Происходит обучение Критика**

Схема эволюции



1. Эволюция происходит в ряду поколений Ng1, Ng2...
2. Продолжительность каждого поколения T тактов времени (время жизни агента)
3. В начале каждого поколения ресурс агента равен нулю.
4. Изменение ресурса агента в течение жизни:

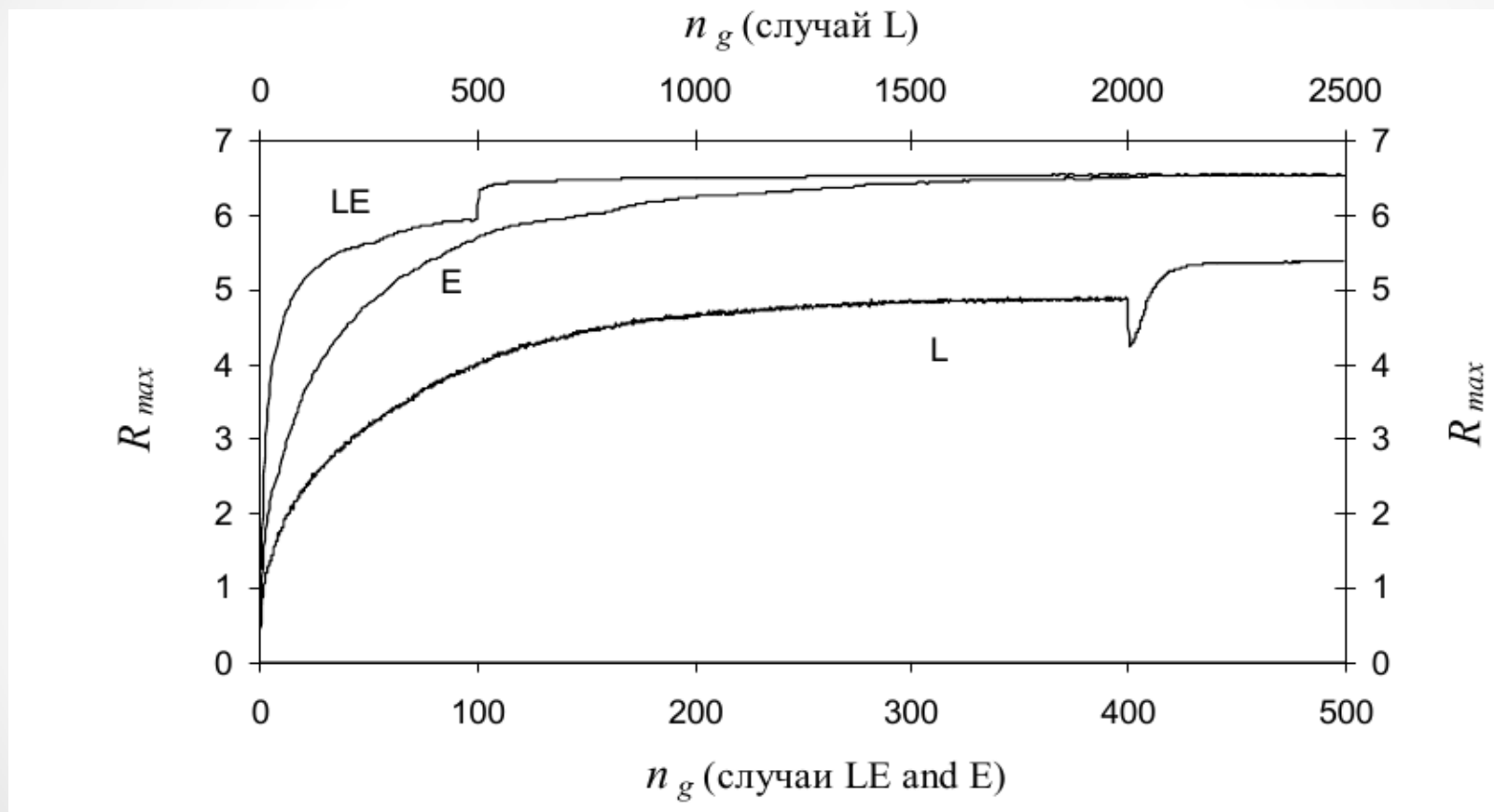
$$R(t+1) = R(t) + r(t)$$

$$r(t) = \log [1 + u(t+1) \Delta X(t+1) / X(t)].$$

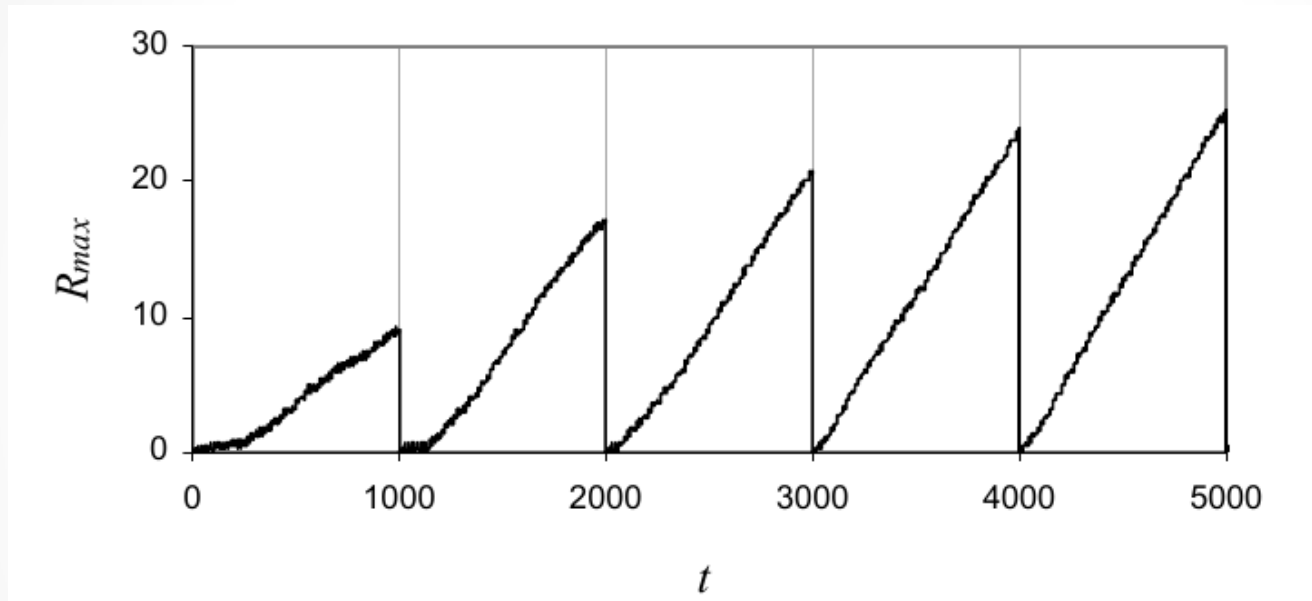
Ход эволюции:

1. Имеется популяция агентов, накапливающих ресурс в ходе времени жизни T. Каждый агент имеет генотип, определяющий качество НС: Модели и Критика. Генотип неизменен в течение жизни, а веса НС меняются в ходе обучения.
2. Агент, имеющий наибольший ресурс, дает жизнь следующему поколению агентов, имеющих генотип предка и случайные мутации. В начале каждого такта времени (поколения), агенты имеют нулевой ресурс и равные веса НС.

Особенности адаптивного поиска, L, L+E, E



Первые пять поколений, LE. Эффект Болдуина



Кривая LE соответствует случаю эволюции, объединенной с обучением, кривая E – случаю чистой эволюции, кривая L – случаю чистого обучения. Временная шкала для случаев LE и E (номер поколения ng) представлена снизу, для случая L (индекс ng) – сверху. Моделирование проведено для синусоиды (10), кривые усреднены по 1000 экспериментам; $n = 10$, $T = 200$.

Выводы

- 3. Представленная модель не является совершенной и имеет свои достоинства, недостатки и «подводные камни». Критерийные примеры, ситуация E и ситуация LE очень похожи, в то время как в реальных биологических системах иерархия обучения (обучение овладения бактериями) и эволюция с обучением (популяция крыс) имеют разную динамику в плане накопления козювото киностатического процесса. Это связано с понятиями «горизонтального» и «вертикального» обучения. В процессе антропогенеза в гораздо большем числе поколений, чем резистентность бактерий к антибиотикам, модельным объектам (животным).
- 4. Настоящая модель не учитывает возможность т.н. «горизонтального» обучения, то есть от анимата к анимату, а не только «вертикального»: Критик-Модель, Экспериментатор-Анимат.

Список литературы

1. Редько В.Г. Модели адаптивного поведения – биологически инспирированный подход к искусственному интеллекту.
2. Лакофф Дж., Нуньес Р. Откуда взялась математика: как разум во плоти создает математику.
3. Риццоллатти Дж., Синигалья К. Зеркала в мозге. О механизмах совместного действия и сопереживания.