Бесхлебнова Г.А., Редько В.Г. Модель формирования обобщенных понятий автономными агентами. Четвертая международная конференция по когнитивной науке: Тезисы докладов: В 2 т. Томск: ТГУ, 2010. Т. 1. С. 174-175.

Модель формирования обобщенных понятий автономными агентами

Бесхлебнова Г.А., Редько В.Г.
gab19@list.ru
НИИ системных исследований РАН (Москва, Россия)

Одно из важных когнитивных свойств живых организмов – свойство формирования обобщенных понятий. Использование понятий приводит к сокращению требуемой памяти и времени обработки информации. Но как происходит формирование понятий, можно ли представить процессы формирования понятий с помощью компьютерного моделирования? Ранее в [1] моделировалось поведение автономных агентов, система управления которых представляла собой набор логических правил, и было продемонстрировано, что в процессе обучения агент выбирает определенные правила, использование которых можно рассматривать как поведение в соответствии с обобщающими эвристиками. Однако эти эвристики были обнаружены путем наблюдения за работой компьютерной программы, у самих же агентов эти эвристики в явном виде не формировались. В настоящей работе развиваются исследования, выполненные в [1], и строится компьютерная модель, в которой агент самостоятельно проводит обобщения, в результате которых формируются понятия.

Описание модели. Рассматривается поведение одного автономного агента в двумерной клеточной среде в дискретном времени. Каждый такт времени t агент выполняет одно из следующих пяти действий: питание, перемещение на одну клетку вперед, поворот направо или налево, отдых. В половине клеток случайным образом размещены порции пищи. Агент обладает ресурсом $R(t)$, который увеличивается при питании и уменьшается при выполнении агентом других действий. При выполнении действия питания агент съедает всю порцию пищи в той клетке, в которой он находится. После этого новая порция пищи добавляется в случайную клетку внешней среды.

Выбор действий агента обеспечивается имеющейся у него системой управления. Система управления агента представляет собой набор правил вида:

$$\mathbf{S}_k \rightarrow A_k ,$$

где $\mathbf{S}_k$ – ситуация, A_k – действие, k – номер правила. Каждое правило имеет свой вес W_k, веса правил модифицируются при обучении агента. Компоненты вектора $\mathbf{S}_k$ принимают значения 0 или 1; они соответствуют наличию или отсутствию порции пищи в определенной клетке в «поле зрения» агента. Поле зрения включает в себя четыре клетки: ту клетку, в которой агент находится, клетку впереди агента и клетки справа и слева агента.

Каждый такт времени агент осуществляет выбор действия и обучается. Выбор действия агентом осуществляется следующим образом. Если имеются правила, для которых все компоненты вектора $\mathbf{S}_k$ совпадают с компонентами вектора текущей ситуации $\mathbf{S}(t)$, то из этих правил с вероятностью $1-\varepsilon$ выбирается то правило, для которого вес W_k максимален, и выполняется действие A_k, входящее в это правило, а с вероятностью ε выполняется случайное действие. Случайное действие выполняется также и тогда, когда нет правил, для которых $\mathbf{S}_k = \mathbf{S}(t)$. Если при случайном выборе действия A у рассматриваемого агента правило $\mathbf{S}(t) \rightarrow A$ отсутствует, то это новое правило добавляется к имеющимся, вес его полагается равным 0. При моделировании использовался «метод отжига»: на начальных тактах моделирования, когда логические правила еще не сформированы, полагалось $\varepsilon \sim 1$, со временем величина ε по экспоненте уменьшалась до нуля, характерное время уменьшения составляло 1000 тактов. При обучении веса правил W_k модифицируются методом обучения с подкреплением [2]. В

результате обучения увеличиваются веса правил, применение которых приводит к росту ресурса агента.

Дополнительно в компьютерную программу вводилась процедура усреднения. Агент для каждого действия проводил усреднение по времени, а именно, вычислялось среднее число применений данного действия по всему времени расчета для той или иной текущей ситуации $\mathbf{S}(t)$. Это позволяло агенту производить обобщение ситуаций и формировать понятия, характеризующие внешнюю среду. Причем эти понятия важны для агента, так как они связаны с его действиями и с приростом его ресурса.

Результаты моделирования. Пример полученной при моделировании зависимости ресурса R агента от времени t показан на рис. 1.

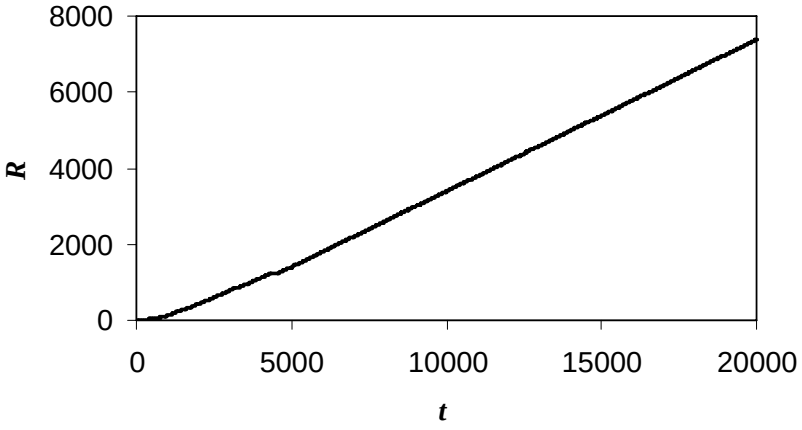

Рис. 1. Зависимость $R(t)$ для самообучающегося агента.

Видно, что в начальные такты времени ($t \sim 1000$), когда правила только формируются, рост ресурса R замедленный, затем динамика $R(t)$ становится линейной.

Проведенное усреднение показало, что оцененные в конце расчета средние по времени частоты действий определяют следующие преимущественные действия агента. Действие *питание* выполняется, если имеется пища в той клетке, в которой находится агент (независимо от того, имеется ли пища в других клетках поля зрения агента). Действие *перемещение на одну клетку вперед* выполняется, если нет пищи в той клетке, в которой находится агент, и имеется пища в клетке впереди агента . Действие *поворот направо/налево* выполняется, если нет пищи в той клетке, в которой находится агент, и в клетке впереди агента, но имеется пища в клетке справа/слева от агента. Частота действия *отдых* пренебрежимо мала. Тем самым формировались цепочки действий, приводящие к нахождению пищи и увеличению ресурса агента.

В результате усреднения агент формирует внутренние понятия «*имеется пища в моей клетке*», «*имеется пища в клетке впереди меня*», «*имеется пища в клетке справа/слева от меня*». Итак, наблюдая за ситуациями и выполняемыми действиями, агент способен самостоятельно формировать понятия, обобщающие сенсорную информацию.

В дальнейшие исследования целесообразно включать моделирование использования понятий и возникновения логических выводов при планировании поведения автономных агентов.

Литература

1. Редько В.Г., Бесхлебнова Г.А. Модель формирования адаптивного поведения автономных агентов // Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сборник научных трудов V-й Международной научно-практической конференции. В 2-х томах. Т.1. М.: Физматлит, 2009. С. 70-79.
2. Sutton R.S., Barto A.G. Reinforcement Learning: An Introduction. MIT Press, 1998.