DOI: 10.25728/mlsd.2023.0063

ИЕРАРХИЧЕСКИЕ ИГРЫ В ГЛУБОКОМ ОБУЧЕНИИ

Горелов М.А., Ерешко Ф.И.
Федеральный исследовательский центр «Информатика и управление» РАН, Москва, Россия
griefer@ccas.ru, fereshko@yandex.ru

Аннотация. Рассматриваются модели иерархических систем в многошаговом случае, при разной степени информированности центра и подсистем. Исследуются процедуры управления системой игроков в своеобразной архитектуре связей, характерной для искусственных нейронных сетей. Сформулирована принципиально новая для иерархических игр задача вычисления наибольшего гарантированного результата центра.

Ключевые слова: теория иерархических игр, игры с неопределенными факторами, максимальный гарантированный результат, процедуры глубокого обучения.

Введение

Рассматривается новый класс многошаговых иерархических игр, в которых игроки распределены по слоям и взаимодействуют только с игроками предыдущего слоя, получая от них данные, суть решения на предыдущем слое, и со следующим слоем, поставляя им свои данные, свои решения. Такой иерархизированный вариант организации взаимодействий позволяет использовать схему динамического программирования для записи процедуры получения решения выделенного игрока в сформулированной игре.

Обсуждается запись функций целей игроков в стремлении обеспечить несколько условий: «естественный» принцип оптимальности, чтобы не было произвола в выборе этого принципа; построенная модель должна была быть достаточно общей, чтобы описывать частные случаи; при выполнении первых двух условий модель должна быть максимально простой. Построенная игра допускает несколько интерпретаций, в частности, в виде процедуры глубокого обучения и стратегического планирования, что, несомненно, привлекательно с точки зрения принятия решений.

Здесь остановимся на задачах глубокого обучения.

1. Иерархическая игра многих лиц

Рассмотрим следующее обобщение иерархической игры Гермейера Γ_1.

Опишем модель иерархической системы, состоящей из Центра и нескольких агентов. Будем считать, что каждый агент может быть отнесен к одному из K уровней иерархии.

Для упрощения формул предположим, что на каждом уровне иерархии имеется N агентов. Это ничего не меняет по существу: в систему всегда могут быть добавлены фиктивные агенты, множества управлений которых состоят из одной точки. При таком предположении агента можно идентифицировать парой kn, где $k = 1,\dots,K$ – уровень иерархии, $n = 1,\dots,N$ – номер агента.

Центр вправе выбирать свои управления u из множества $U = \prod_{k=1}^{k}\prod_{n=1}^{N} U^{kn}$. Агент kn выбирает свое управление v^{kn} из множества V^{kn}. Интересы Центра описываются стремление к максимизации значения функции $g : U \times \prod_{k=1}^{K}\prod_{n=1}^{N} V^{kn} \to \mathbb{R}$. Интересы агента kn задаются функцией выигрыша $h^{kn} : U^{kn} \times \prod_{i=1}^{N} V^{(k-1)i} \times V^{kn} \to \mathbb{R}$ (формально полагаем $V^{0n} = \{0\}$, $n = 1,\dots,n$).

Для простоты считаем, что все множества управлений наделены топологиями и компактны, а функции выигрыша непрерывны по всем своим аргументам.

Рассмотрим следующий порядок принятия решений. Центр первым выбирает свое управление $u \in U$ и доводит свое решение для всех агентов. Затем одновременно и независимо принимают свои решения агенты первого уровня, и их выборы $v^{11}, v^{12},\dots, v^{1N}$ становятся известными агентам второго уровня. Следом принимают решения агенты второго уровня и т.д. Последними принимают решения агенты уровня K.

В таком случае Центру естественно ориентироваться на свой максимальный гарантированный результат, определяемый следующим образом. Положим

$$BR^{kn}(u^{kn}, v^{k-1}) = \operatorname{Arg}\max_{w^{kn}\in V^{kn}} h^{kn}(u^{kn}, v^{k-1}, w^{kn}),$$

(здесь и далее $u^k = (u^{k1},\ldots,u^{kN})$, $v^k = (v^{k1},\ldots,v^{kN})$, $k = 1,\ldots,K)$, $u = (u^1,\ldots u^K)$),

$$BR^k(u^k, v^{(k-1)}) = \prod_{n=1}^{N} BR^{kn}(u^{kn}, v^{(k-1)}).$$

Тогда максимальный гарантированный результат Центра равен

$$\sup_{u\in U}\ \min_{v^1\in BR(u^1,v^0)} \ldots \min_{v^1\in BR(u^K,v^{K-1})} g(u, v^1, \ldots, v^K).$$

Классическая иерархическая игра двух лиц Γ_1 получится, если положить $K = 1$ и $N = 1$.

Кое-что из описанных выше конструкций требует объяснения. Главным образом это относится к весьма специальному виду функций выигрыша агентов. При построении данной модели преследовалось три цели:

- хотелось построить игру, для которой существует «естественный» принцип оптимальности, чтобы не было произвола в выборе этого принципа;
- построенная модель должна была быть достаточно общей, чтобы описывать частные случаи, рассмотренные в разделах 4 и 5 далее;
- при выполнении первых двух условий модель должна быть максимально простой.

Фиксированный порядок ходов – достаточно распространенный способ моделирования иерархии. А если такой порядок фиксирован, то для выполнения первой из сформулированных выше целей необходимо, чтобы выигрыш агента ни прямо, ни косвенно не зависел от действий агентов, принимающих решения одновременно с ним. В противном случае для агентов одного уровня получим игру достаточно общего вида, для которой, как известно, «естественного» принципа оптимальности не существует (в разных случаях рассматривают равновесия по Нэшу, оптимальность по Парето и много что еще). Но тогда функция выигрыша агента kn не должна зависеть явно от управлений агентов kl при $l \neq k$. А, кроме того, она не должна зависеть от выборов агента ij на уровне $i > k$, выбор которого как-то зависит от выбора агентов kl с $l \neq k$. Здесь остается, по сути две возможности. Или любой агент «связан» с единственным агентом на каждом из более высоких уровней, или выигрыш агента kn не зависит от выборов агентов на более низких уровнях иерархии (разумеется, можно рассматривать «гибридные» варианты, но они заметно сложнее). Два описанных класса моделей пересекаются (выигрыш агента может не зависеть от управлений агентов более низкого уровня и зависеть явно только от управления одного агента более высокого уровня). Именно такая модель рассмотрена в разделе 5.

Идя по первому пути, мы придем к так называемому «веерному» случаю. Он детально изучен, но не годится для описания искусственных нейронных сетей, т.е. не достигается вторая из сформулированных целей. Поэтому выбран второй путь.

Если соответствующая гипотеза принята, то первую цель можно считать достигнутой. В самом деле, в момент выбора своего решения агент kn уже знает стратегии, выбранный агентами более высокого уровня. Следовательно, его выигрыш зависит, по существу, только от его собственного решения. Но тогда единственный разумный путь – это выбор стратегии из множества $BR^{kn}(u^{kn},v^{k-1})$. Собственно, это и заложено в модель вместе с классическим принципом гарантированного результата.

О том, как достигается вторая цель рассказано в соответствующих разделах. А третья цель приводит к тому, что функция h^{kn} не зависит от управлений агентов ij с $i < k - 1$.

Та же первая цель приводит к отказу от использования стратегий с обратной связью. А именно, не получается построить модель, в которой стратегия агента kn – это функция от управлений, выбранных агентами более низких уровней. А при отказе от рассмотрения стратегий с обратной связью сделанные выше предположения о непрерывности и компактности выглядят вполне естественно.

По существу, сделанные предположения описывают некую специальную иерархическую систему, аналогичную, например, веерной (кстати говоря, эту структуру трудно адекватно описать на языке теории графов). Возможно, было бы интересно поискать еще какие-то структуры, для которых достигается первая из сформулированных целей

2. Вычисление максимального гарантированного результата

Сделанные предположения о структуре функций выигрыша позволяют использовать стандартную попятную процедуру для вычисления максимального гарантированного результата Центра.

Положим

$$L^K(u, v^1, ..., v^K) = g(u, v^1, ..., v^K),$$

$$L^k(u, v^1, ..., v^k) = \min_{w^{k+1} \in BR^{k+1}(u^{k+1}, v^k)} L^{k+1}(u, v^1, ... v^k, w^{k+1}), \; k = K-1, ..., 1,$$

$$M^K(u^1, ..., u^K) = M^K(u) = \min_{w^1 \in BR^1(u^1, v^0)} L^1(u, w^1),$$

$$M^k(u^1, ..., u^k) = \sup_{\omega^{k+1} \in U^{(k+1)1} \times ... \times U^{(k+1)N}} M^{k+1}(u^1, ..., u^k, \omega^{k+1}), k = K-1, ..., 1.$$

Тогда максимальный гарантированный результат Центра равен $\sup\limits_{u^1 \in U^{11} \times ... U^{1K}} M^1(u^1)$.

Верхние грани в последней формуле могут не достигаться даже при выполнении сделанных предположений о непрерывности и компактности (это так уже в классической игре Γ_1). Но “почти оптимальное” управление Центра можно найти, если выбирать его компоненты достаточно точно реализующими соответствующие верхние грани.

Разумеется, задачи вычисления минимумов в формуле и супремумов в формулах допускают дальнейшую декомпозицию с учетом того, что $w^k = (w^{k1}, \dots w^{kN})$, $\omega^k = (\omega^{k1}, \dots, \omega^{kN})$ а множества значений этих переменных являются декартовыми произведениями множеств $BR^{kn}(u^{kn}, v^{k-1})$ и U^{kn} соответственно.

Обычно решить много простых задач проще, чем одну сложную. Поэтому отмеченные возможности декомпозиции задачи вычисления максимального гарантированного результата в данной игре свидетельствуют о том, что рассматриваемая модель доступна для качественного и количественного исследования. Это – определенный довод в ее пользу.

3. Искусственные нейронные сети

Частным случаем описанной выше задачи является задача обучения искусственной нейронной сети.

Рассмотрим вариант полносвязной ИНС.

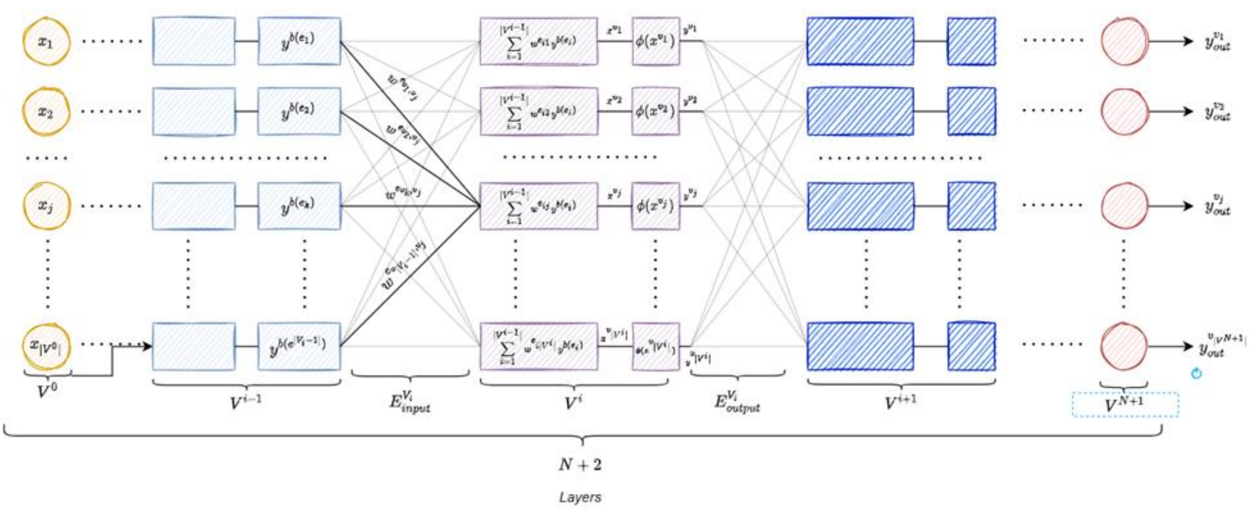

Рис.1. Полносвязная ИНС

Для конкретности рассмотрим задачу обучения с учителем нейронной сети, состоящей из K слоев, на каждом слое которой находится N нейронов. Каждый нейрон преобразует набор входных сигналов

$x^1,\ldots,x^N$ ($x^n \in \mathbb{R},\ n = 1,\ldots,N$) в выходной сигнал $y = \varphi\left(w^0 + \sum_{i=1}^{N} w^i x^i\right) \in \mathbb{R}$, где $\varphi : \mathbb{R} \to \mathbb{R}$ – заданная функция, а веса w^i – неотрицательные числа, удовлетворяющие условию $\sum_{i=0}^{K} w^i = 1$. Сеть организована так, что на вход каждого нейрона из слоя $k+1$ подаются выходы всех нейронов слоя k ($k = 1,\ldots,K-1$). Заданы наборы $\left(x_t^1,\ldots,x_t^N\right),\ t = 1,\ldots,T$ сигналов, подающихся на вход нейронов первого слоя и соответствующие "эталонные" наборы выходных сигналов нейронов слоя K. Требуется подобрать веса, вообще говоря, свои для каждого нейрона, так, чтобы для выходных сигналов $\left(v_t^1,\ldots,v_t^N\right)$ нейронов слоя K соответствующих входным сигналам $\left(x_t^1,\ldots,x_t^N\right)$, реализовался максимум функции качества, например, функции $\sum_{t=1}^{T}\sum_{n=1}^{N}\left(v_t^n - y_t^n\right)^2$.

Это – классическая задача из теории искусственных нейронных сетей. Покажем, как она вкладывается в модель, описанную в разделе 1.

Сопоставим каждому нейрону, лежащему в слое k элемент иерархической системы на k-ом уровне иерархии с множеством стратегий $V^{kn} = \mathbb{R}^T$ и функцией выигрыша

$$h^{kn}(u^{kn}, v^{(k-1)1},\ldots,v^{(k-1)n}, v^{kn}) = -\sum_{t=1}^{T}\left(v_t^{kn} - \varphi\left(u^{kn,0} + \sum_{i=1}^{N} u^{kn,i} v_t^{(k-1)n}\right)\right)^2$$

(здесь $u^{kn} = \left(u^{kn,0}, u^{kn,1},\ldots,u^{kn,N}\right) \in U^{kn} = \left\{\left(u^{kn,0}, u^{kn,1},\ldots,u^{kn,N}\right) \in \mathbb{R}_+^{N+1} : \sum_{i=1}^{N} u^{kn,i} = 1\right\}$ – элемент стратегии Центра, $v^{kn} = \left(v_1^{kn},\ldots,v_T^{kn}\right)$ – стратегия агента kn). Числа $v_t^{0n} = x_{tn}$, $t = 1,\ldots,T$, $n = 1,\ldots,N$ считаются заданными параметрами задачи.

Центр имеет множество стратегий $U = \prod_{k=1}^{k}\prod_{n=1}^{N} U^{kn}$ и стремится максимизировать значение функции $g(u, v^1,\ldots,v^K) = -\sum_{t=1}^{T}\sum_{n=1}^{N}\left(v_t^{Kn} - y_t^n\right)^2$ (здесь набор чисел y_t^n, $v_t^{0n} = x_{tn}$, $t = 1,\ldots,T$, $n = 1,\ldots,N$ – параметр задачи).

В такой задаче каждое множество $BR^{kn}(u^{kn},v^k)$ состоит из одной точки, поскольку функции выигрыша агентов строго вогнуты по их собственным управлениям. Следовательно, оптимальный ответ на стратегии Центра и агентов более высокого уровня непрерывно зависит от этих стратегий. Поэтому в данной задаче максимальный гарантированный результат Центра достигается и равен наилучшему качеству обучения нейросети, а стратегия, гарантирующая получение этого результата, дают наилучшие веса для данной нейронной сети.

Векторная задача идентификации параметров.

На вход графа последовательно подаётся массив прецедентов, паттернов. Это массив векторов $X_t^k = (x_{t0}^k, x_{t1}^k,\ldots,x_{tl}^k,\ldots,x_{tL}^k),\ k = 1,2,\ldots,K$. Цель идентификации состоит в том, чтобы настроить управления, т. е. параметры ω^v, α^v, так, чтобы по возможности минимизировать форму

$$\sum_{k=1}^{K}\sum_{v\in V^{n+1}}(y^{k,v} - \overline{y}^{k,v})^2 .$$

То есть, минимизировать не для одного отдельного образца k, а на всём множестве прецедентов.

Решение этой задачи в общем случае достаточно трудоёмко. Поэтому разрабатываются методы какого-то приближения к её решению, главным образом, эвристические и итерационные.

Итерации состоят в циклической обработке элементов массива $X_t^k = (x_{t0}^k, x_{t1}^k,\ldots,x_{tl}^k,\ldots,x_{tL}^k),\ k = 1,2,\ldots,K$ и постепенном продвижении значений функции отклонения к минимуму.

Один проход этого процесса по всему массиву паттернов $[1,K]$ – это одна итерация.

4. Заключение

Наше исследование возможностей построенной игры показывает, что существует формальный инструмент, который позволяет единообразно генерировать и рассматривать разные постановки задач глубокого обучения и стратегического планирования. Так для задач глубокого обучения актуальна задача поиска решений весовых коэффициентов при условии наличия неопределённых факторов, воздействующих на всю сеть, и игровой подход предоставляет инструментальную возможность записать алгоритмы решения в различных схемах перебора в пространстве состояний или градиентных алгоритмов.

В задачах стратегического планирования весовые параметры свёрток, как правило, задаются технологическими процессами, и наиболее актуальными становятся проблемы формирования функций активации, в данной интерпретации стратегии игроков, и графа взаимодействия, т.е. архитектуры искусственной нейронной сети. Перспективы практической реализации предложенной теоретической схемы определятся в последующих приложениях, и составят предмет вычислительных экспериментов

Литература

1. *Гермейер Ю.Б., Моисеев Н.Н.* О некоторых задачах теории иерархических систем. // Проблемы прикладной математики и механики. М.: Наука, 1971. – С. 30-43.
3. *Ватель И.А., Ерешко Ф.И.* Игры с иерархической структурой. //Математическая энциклопедия. Т.2. М.: Советская энциклопедия, 1979. – С.478-482.
4. *Горелов М.А.* Максимальный гарантированный результат в иерархических играх // Управление большими системами. Вып. 67. М.: ИПУ РАН, 2017. – С. 4-31.
5. *Горелов М.А.* Иерархические игры с неопределенными факторами // Управление большими системами. Вып. 59. М.: ИПУ РАН, 2016. – С.6-22.
6. *Горелов М.А., ЕрешкоФ.И.*О моделях централизации и децентрализации управления в цифровом обществе // Контуры цифровой реальности: Гуманитарно-технологическая революция и выбор будущего/Под ред. В.В.Иванова, Г.Г.Малинецкого, С.Н. Сиренко. М.: Ленанд, 2018. –С. 187-202.
7. *Горелов М.А., Ерешко Ф.И* .Информированность и децентрализация управления // Автоматика и телемеханика. 2019. № 6. – С. 156-172.
10. *Горелов М.А., Ерешко Ф.И.*Информированность и децентрализация управления (стохастический случай) // Автоматика и телемеханика. 2020. № 1. – С. 52-66.
11. *Горелов М.А.*Принятие решений при избытке информации. Управление развитием крупномасштабных систем// Труды тринадцатой международной конференции. М.: ИПУ РАН, 2020. – С. 536-546.
12. *Гудфеллоу Я., Бенджио И., Курвилль А.* Глубокое обучение / пер. с анг. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.: цв. ил. ISBN 978-5-97060-618-6
13. *Шолле Ф.* Глубокое обучение на Python. СПб.: Питер. 2018. 400 с.
14. *Бирюкова Т.К.* Построение нейронных сетей различных типов с использованием параболических интегро-дифференциальных сплайнов как функций активации // Системы высокой доступности. 2020. Т. 16. № 4. С. 40−49. DOI: 10.18127/j20729472-202004-03
15. *Горелов М.А., Ерешко Ф.И.* Информированность и децентрализация управления // Автоматика и телемеханика. 2019. №6. – С. 156–172.
16. *Горелов М.А., Ерешко Ф.И.* Иерархическая структура сетевых моделей в экономике и искусственных нейронных сетей. Тенденции развития Интернет и цифровой экономики: Труды V Всероссийской с международным участием научно-практической конференции, Симферополь-Алушта, 02–04 июня 2022 года. – Симферополь: Крымский федеральный университет им. В.И. Вернадского, 2022. – 204 с. – ISBN 978-5-6047625-0-9. – EDN SFKMEN.
17. *Ерешко Ф.И., Горелов М.А.* Игровое представление искусственных нейронных сетей // Анализ, моделирование, управление, развитие социально-экономических систем: сборник научных трудов XVI Международной школы-симпозиума АМУР-2022, Симферополь-Судак, 14-27 сентября 2022 / ред. совет: А. В. Сигал (предс.) и др. – Симферополь : ИП Корниенко А. А. – 2022. – С. 146–149.