

Посвящается Лой, Гордону и Люси

С.Дж. Рассел

Посвящается Крису, Изabelle и Джульетте

П. Норвиг

ПРЕДИСЛОВИЕ

Искусственный интеллект (ИИ) — широкая область знаний; именно поэтому данная книга имеет такой большой объем. Авторы попытались достаточно полно описать теоретические основы искусственного интеллекта, включая математическую логику, теорию вероятностей и теорию непрерывных функций, раскрыть суть таких понятий, как восприятие, рассуждение, обучение и действие, а также описать все технические средства, созданные в рамках этого научного направления, начиная с микроэлектронных устройств и заканчивая межпланетными автоматическими зондами. Большой объем данной книги обусловлен также тем, что авторы стремились представить достигнутые результаты достаточно глубоко, хотя в основной части каждой главы они старались охватить только самые важные идеи, касающиеся рассматриваемой темы. Указания, позволяющие получить более полные сведения, приведены в библиографических заметках в конце каждой главы.

Эта книга имеет подзаголовок “Современный подход”. С помощью этого довольно-таки малосодержательного названия авторы хотели подчеркнуть, что пытались представить в ней в рамках единого способа изложения все современные достижения в области искусственного интеллекта, а не описать каждое отдельное его направление в его собственном историческом контексте. Авторы приносят свои извинения представителям тех направлений, которые стали выглядеть не столь значимыми, как они заслуживают, лишь из-за того, что для их описания принят такой подход.

Главной объединяющей темой этой книги является идея *интеллектуального агента*. Авторы определяют *искусственный интеллект* как науку об агентах, которые получают результаты актов восприятия из своей среды и выполняют действия, причем каждый такой агент реализует функцию, которая отображает последовательности актов восприятия в действия. В данной книге рассматриваются различные способы представления этих функций, в частности продукционные системы, реактивные агенты, условные планировщики в реальном масштабе времени, нейронные сети и системы, действующие на основе теории решений. Авторы трактуют роль *обучения* как распространения сферы деятельности проектировщика на неизвестную среду и показывают, какие ограничения налагает указанный подход к обучению на проект агента, способствуя применению явного представления знаний и таких же способов формирования рассуждений. Кроме того, авторы рассматривают робототехнику и системы технического зрения не как независимо определяемые научные направления, а как области знаний, позволяющие обеспечить более успешное достижение целей, стоящих перед агентами, и подчеркивают важность учета того, в какой среде агент решает поставленные перед ним задачи, при определении соответствующего проекта агента.

Основная цель авторов состояла в том, чтобы изложить идеи, которые были сформулированы в исследованиях по искусственному интеллекту, проводившихся в течение последних пятидесяти лет, а также собраны на протяжении последних двух тысячелетий в тех областях знаний, которые стали стимулом к развитию искусственного интеллекта. Мы старались избегать чрезмерного формализма при изложении этих идей, сохраняя при этом необходимую точность. При любой возможности мы приводили алгоритмы на псевдокоде, чтобы конкретизировать излагаемые идеи;

краткие сведения о применяемом нами псевдокоде содержатся в приложении Б. Реализации этих алгоритмов на нескольких языках программирования можно найти на сопровождающем Web-узле книги (aima.cs.berkeley.edu).

Настоящая книга прежде всего предназначена для использования в базовом университетском курсе или в последовательности курсов. Она может также использоваться в курсе по специальности (возможно, с добавлением материала из некоторых основных источников, предложенных в библиографических заметках). Кроме того, данная книга характеризуется всесторонним охватом тематики и большим количеством подробных алгоритмов, поэтому применима в качестве основного справочника для аспирантов, специализирующихся в области искусственного интеллекта, а также будет небезынтересна профессионалам, желающим выйти за пределы избранной ими специальности. При этом единственным требованием является знакомство с основными понятиями информатики (алгоритмы, структуры данных, классы сложности) на уровне студента-второкурсника. Для понимания материала по нейронным сетям и подробного ознакомления со сведениями о статистическом обучении полезно освоить исчисление на уровне студента первого курса. Часть необходимых сведений из области математики приведена в приложении А.

Краткий обзор книги

Данная книга разделена на восемь частей. В части I, “Искусственный интеллект”, предлагается общий обзор тематики искусственного интеллекта, базирующийся на идее интеллектуального агента — системы, которая способна принять решение о том, что делать, а затем выполнить это решение. В части II, “Решение проблем”, изложение сосредоточено на методах принятия решений по выбору оптимальных действий в тех условиях, когда необходимо продумывать наперед несколько этапов, например при поиске маршрута проезда через всю страну или при игре в шахматы. В части III, “Знания и рассуждения”, обсуждаются способы представления знаний о мире — как он функционирует, каковы его основные особенности в настоящее время и к чему могут привести те или иные действия, а также способы формирования логических рассуждений на основе этих знаний. В части IV, “Планирование”, описывается, как использовать эти способы формирования рассуждений для принятия решений по выбору дальнейших действий, особенно при составлении планов. Часть V, “Неопределенные знания и рассуждения в условиях неопределенности”, аналогична частям III и IV, но в ней изложение в основном сосредоточивается на способах формирования рассуждений и принятия решений в условиях неопределенности знаний о мире, с чем обычно приходится сталкиваться, например, в системах медицинской диагностики и лечения.

Части II–V, вместе взятые, содержат описание тех компонентов интеллектуального агента, которые отвечают за выработку решений. В части VI, “Обучение”, описаны методы выработки знаний, необходимых для этих компонентов, которые обеспечивают принятие решений. В части VII, “Общение, восприятие и осуществление действий”, описаны способы, с помощью которых интеллектуальные агенты могут получать результаты восприятия из своей среды, чтобы узнать, что в ней происходит, либо с помощью систем технического зрения, осязания, слуха, либо на основе понимания языка, а также способы, с помощью которых интеллектуальные агенты могут претворять свои планы в реальные действия, такие как выполнение движений

робота или произнесение фрагментов речи на естественном языке. Наконец, в части VIII, “Заключение”, анализируется прошлое и будущее искусственного интеллекта и рассматриваются философские и этические последствия его развития.

Отличия от первого издания

Со времени публикации первого издания этой книги в 1995 году в искусственном интеллекте многое изменилось, поэтому внесены значительные изменения и в саму книгу. Каждая глава была в значительной степени переработана, чтобы в ней можно было отразить новейшие достижения в рассматриваемой области, дать иное толкование старым работам с той точки зрения, которая более согласована с новыми результатами, а также улучшить качество изложения рассматриваемых идей в соответствии с принципами педагогики. Активных пользователей методов искусственного интеллекта должно порадовать то, что представленные в настоящем издании методы стали намного более эффективными по сравнению с теми, которые описывались в издании 1995 года; например, алгоритмы планирования, которые рассматривались в первом издании, позволяли формировать планы, состоящие всего лишь из нескольких шагов, тогда как масштабы применения алгоритмов, описанные в настоящем издании, увеличились до десятков тысяч шагов. Подобные усовершенствования, измеряемые несколькими порядками величин, достигнуты и в областях вероятностного вывода и обработки лингвистической информации, а также в других вспомогательных областях. Наиболее существенные изменения, внесенные во второе издание, описаны ниже.




- В части I изложены факты, которые свидетельствуют о признании исторического вклада в развитие искусственного интеллекта со стороны теории управления, теории игр, экономики и неврологии. Это позволяет создать основу для более целостного описания идей, заимствованных из этих научных областей, в последующих главах.
- В части II описаны алгоритмы оперативного поиска и введена новая глава по удовлетворению ограничений, которая позволяет установить естественную связь между вычислительными методами и приведенными в данной книге материалами по логике.
- Теперь в части III пропозициональная логика, которая в первом издании была рекомендована читателям как промежуточная ступенька на пути к логике первого порядка, рассматривается как полезный сам по себе язык представления, для которого предусмотрены быстродействующие алгоритмы логического вывода и эффективные проекты агентов на основе схемы. Главы по логике первого порядка были реорганизованы для более наглядного изложения материала, а в качестве примера проблемной области приведено описание процесса осуществления покупок в Internet.
- В части IV приведены сведения о более новых методах планирования, таких как Graphplan и планирование на основе выполнимости. Кроме того, увеличен объем изложения, касающегося составления расписаний, условного планирования, иерархического планирования и мультиагентского планирования.
- В часть V включен дополнительный материал по байесовским сетям, в котором описаны новые алгоритмы, в частности алгоритмы устранения перемен-

ных и алгоритмы Монте-Карло на основе марковской цепи, а также введена новая глава по формированию неопределенных рассуждений с учетом времени и созданию покрытий скрытых марковских моделей, а также по применению фильтров Калмана и динамических байесовских сетей. Описание марковских процессов принятия решений стало еще более глубоким; введены новые разделы по теории игр и проектированию механизма.

- В части VI связаны воедино все результаты, достигнутые в области статистического и символического обучения, а также обучения нейронных сетей; кроме того, введены разделы, содержащие сведения об увеличении производительности алгоритмов, алгоритме EM, обучении на основе экземпляра и о ядерных методах (о машинах поддерживающих векторов).
- В части VII к общему объему материала об обработке лингвистической информации добавлены разделы, касающиеся обработки речи и индуктивного вывода грамматики, а также глава по вероятностным языковым моделям, с учетом того, что областью применения этих сведений должны стать информационный поиск и машинный перевод. В ходе изложения вопросов робототехники подчеркнута необходимость применения методов обработки неопределенных сенсорных данных, а в главе по системам технического зрения приведены уточненные сведения по распознаванию объектов.
- В части VIII предусмотрен дополнительный раздел, касающийся этических последствий развития искусственного интеллекта.

Как использовать эту книгу

Книга состоит из 27 глав, причем для изучения каждой из них требуется примерно недельный объем лекций. Таким образом, для учебной проработки всей книги требуется последовательность курсов лекций, рассчитанная на два семестра. Еще один вариант состоит в том, что может быть составлен выборочный курс, удовлетворяющий интересы преподавателя и студента. Благодаря тому что в ней охвачена широкая тематика, эта книга может использоваться в качестве основы для многих курсов, начиная с коротких, вводных циклов лекций для начинающих и заканчивая специализированными курсами с углубленным изучением избранной темы для студентов последних лет обучения. На Web-узле, находящемся по адресу `aima.cs.berkeley.edu`, приведены программы курсов лекций, проводимых более чем в 600 университетах и колледжах, в основу которых было положено первое издание настоящей книги, а также даны рекомендации, позволяющие читателю найти программу курсов лекций, в наибольшей степени соответствующую его потребностям.

Книга включает 385 упражнений. Упражнения, требующие существенного объема программирования, отмечены значком в виде клавиатуры () . Проще всего эти упражнения можно выполнить, воспользовавшись архивом кода, который находится по адресу `aima.cs.berkeley.edu`. Некоторые из упражнений настолько велики, что их можно рассматривать как проекты с заданными сроками. Многие упражнения требуют проведения определенных исследований с помощью доступной литературы; они отмечены значком в виде книги () . Важные примечания отмечены значком в виде «указующего перста» () и выделены курсивным шрифтом. В книгу

включен обширный предметный указатель, состоящий из нескольких тысяч элементов, который поможет читателю найти нужную тему. Кроме того, значком с изображением руки, держащей карандаш (✍), и полужирным шрифтом отмечаются все новые термины, везде, где впервые приведено их определение.

Использование Web-узла

На Web-узле `aima.cs.berkeley.edu` приведено следующее:

- реализации алгоритмов, описанных в книге, на нескольких языках программирования;
- список более чем 600 учебных заведений, в которых используется данная книга, сопровождающийся многочисленными ссылками на материалы курсов, доступные в оперативном режиме;
- аннотированный список более чем 800 ссылок на Web-узлы с полезными сведениями по искусственному интеллекту;
- списки дополнительных материалов и ссылок, относящихся к каждой главе;
- инструкции с описанием того, как присоединиться к дискуссионной группе, посвященной данной книге;
- инструкции с описанием того, как обратиться к авторам, чтобы передать им свои вопросы или комментарии;
- инструкции с описанием того, как сообщить об ошибках, обнаруженных в книге;
- копии рисунков из оригинала книги, а также слайды и другие материалы для преподавателей.

Благодарности

Основная часть главы 24 (по системам технического зрения) написана Джитендрой Маликом (Jitendra Malik). Глава 25 (по робототехнике) в основном написана Себастьяном Граном (Sebastian Thrun) для настоящего издания и Джоном Кэнни (John Canny) для первого издания. Дуг Эдвардс (Doug Edwards) провел исследование, на основании которого написаны исторические заметки для первого издания. Тим Хуанг (Tim Huang), Марк Паскин (Mark Paskin) и Синтия Бруинс (Cynthia Bruyns) оказали помощь при оформлении диаграмм и алгоритмов. Алан Апт (Alan Apt), Сондра Чавес (Sondra Chavez), Тони Холм (Toni Holm), Джейк Вард (Jake Warde), Ирвин Zucker (Irwin Zucker) и Камилла Грантакост (Camille Trentacoste), сотрудники издательства Prentice Hall, приложили большие усилия, чтобы помочь нам соблюсти намеченный график подготовки книги, и внесли много полезных предложений по оформлению и содержанию книги.

Стюарт хотел бы поблагодарить своих родителей за их постоянную помощь и поддержку, а также свою жену, Лой Шефлотт (Loy Sheflott), за ее бесконечное терпение и безграничную мудрость. Он надеется, что скоро эту книгу прочитают Гордон и Люси. Исключительно полезной для него была работа с RUGS (Russell's Unusual Group of Students — необыкновенная группа студентов Рассела).

Питер хотел бы поблагодарить своих родителей, Торстена и Герду, за то, что они очень помогли ему на первых порах, и свою жену Крис, детей и друзей за то, что

подбадривали его и терпели его отсутствие в течение тех долгих часов, когда он писал эту книгу, и тех еще более долгих часов, когда он снова ее переписывал.

Мы очень обязаны библиотекарям, работающим в университете г. Беркли, Стэнфордском университете, Массачусетском технологическом институте и агентстве NASA, а также разработчикам узлов CiteSeer и Google, которые внесли революционные изменения в сам способ проведения исследований.

Мы буквально не с состоянием выразить свою признательность всем тем, кто использовал данную книгу и внес свои предложения, но хотели бы поблагодарить за особо полезные комментарии следующих: Кшиштофа Апта (Krzysztof Apt), Эллери Эзиела (Ellery Aziel), Джефа Ван Баалена (Jeff Van Baalen), Брайена Бейкера (Brian Baker), Дона Баркера (Don Barker), Тони Баррета (Tony Barrett), Джеймса Ньютона Баса (James Newton Bass), Дона Била (Don Beal), Говарда Бека (Howard Beck), Вольфганга Бибеля (Wolfgang Bibel), Джона Биндера (John Binder), Лэрри Букмана (Larry Bookman), Дэвида Р. Боксолла (David R. Boxall), Герхарда Бревку (Gerhard Brewka), Селмера Бринсйорда (Selmer Bringsjord), Карла Бродли (Carla Brodley), Криса Брауна (Chris Brown), Вильгельма Бургера (Wilhelm Burger), Лорен Берка (Lauren Burka), Жоао Кашпоро (Joao Cachopo), Меррея Кэмпбелла (MurRAY Campbell), Нормана Карвера (Norman Carver), Эммануэля Кастро (Emmanuel Castro), Анила Чакраварти (Anil Chakravarthy), Дэна Чизарика (Dan Chisarick), Роберто Сиполлу (Roberto Cipolla), Дэвида Коэна (David Cohen), Джеймса Коулмэна (James Coleman), Джули Энн Компарини (Julie Ann Comparini), Гэри Котрелла (Gary Cottrell), Эрнеста Дэвиса (Ernest Davis), Рину Дехтер (Rina Dechter), Тома Диттерика (Tom Dietterich), Чака Дийера (Chuck Dyer), Барбару Энгельхардт (Barbara Engelhardt), Дуга Эдвардса (Doug Edwards), Кутлухана Эрола (Kutluhan Erol), Орена Этциони (Oren Etzioni), Хану Филир (Hana Filip), Дугласа Фишера (Douglas Fisher), Джеффри Форбса (Jeffrey Forbes), Кена Форда (Ken Ford), Джона Фослера (John Fosler), Алекса Франца (Alex Franz), Боба Фатрелла (Bob Futrelle), Марека Галецки (Marek Galecki), Штефана Гербердинга (Stefan Gerberding), Стюарта Джилла (Stuart Gill), Сабину Глеснер (Sabine Glesner), Сета Голуба (Seth Golub), Гошту Гранье (Gosta Grahne), Расса Грейнера (Russ Greiner), Эрика Гримсона (Eric Grimson), Барбару Грош (Barbara Grosz), Лэрри Холла (Larry Hall), Стива Хэнкса (Steve Hanks), Отара Хэнссона (Othar Hansson), Эрнста Хайнца (Ernst Heinz), Джима Эндлера (Jim Hendler), Кристофа Херманна (Christoph Herrmann), Вазанта Хонавара (Vasant Honavar), Тима Хуанга (Tim Huang), Сета Хатчинсона (Seth Hutchinson), Джуста Джейкоба (Joost Jacob), Магнуса Йоханссона (Magnus Johansson), Дэна Джурафски (Dan Jurafsky), Лесли Кэлблинга (Leslie Kaelbling), Кейдзи Канадзава (Keiji Kanazawa), Сурекха Касибхатла (Surekha Kasibhatla), Саймона Казифа (Simon Kasif), Генри Каутца (Henry Kautz), Гернота Кершбаумера (Gernot Kerschbaumer), Ричарда Кирби (Richard Kirby), Кевина Найта (Kevin Knight), Свена Кёнига (Sven Koenig), Дафну Коллер (Daphne Koller), Рича Корфа (Rich Korf), Джеймса Керина (James Kurien), Джона Лафферти (John Lafferty), Гуса Ларссона (Gus Larsson), Джона Лаззаро (John Lazzaro), Джона Лебланка (Jon LeBlanc), Джейсона Литермана (Jason Leatherman), Фрэнка Ли (Frank Lee), Эдварда Лима (Edward Lim), Пьера Луво (Pierre Louveaux), Дона Лавленда (Don Loveland), Сридхара Махадевана (Sridhar Mahadevan), Джима Мартина (Jim Martin), Энди Мейера (Andy Mayer), Дэвида Мак-Грэйна (David McGrane), Джей Менделсон (Jay Mendelsohn), Брайена Милча (Brian Milch), Стива Майнтона (Steve Minton), Вибху Миттала (Vibhu Mittal), Леору Моргенстерн (Leora Morgenstern), Стивена Магглтона (Stephen Muggleton), Кевина Мэрфи (Kevin Murphy), Рона Мьюзика (Ron Musick), Санга Миаэнга (Sung Myaeng), Ли Нэйша (Lee Naish), Панду Найака (Pandu

Nayak), Бернхарда Небеля (Bernhard Nebel), Стюарта Нельсона (Stuart Nelson), Шуан-лонг Нгуэн (XuanLong Nguyen), Иллаха Нурбакша (Illah Nourbakhsh), Стива Омохандро (Steve Omohundro), Дэвида Пейджа (David Page), Дэвида Палмера (David Palmer), Дэвида Паркса (David Parkes), Рона Парра (Ron Parr), Марка Паскина (Mark Paskin), Тони Пассера (Tony Passera), Майкла Паззани (Michael Pazzani), Вима Пейлса (Wim Pijls), Иру Пол (Ira Pohl), Марту Поллак (Martha Pollack), Дэвида Пула (David Poole), Брюса Портера (Bruce Porter), Малкома Прадхана (Malcolm Pradhan), Билла Прингла (Bill Pringle), Лоррэн Прайор (Lorraine Prior), Грэга Прована (Greg Provan), Уильяма Рапапорта (William Rapaport), Филипа Ресника (Philip Resnik), Франческу Росси (Francesca Rossi), Джонатана Шеффера (Jonathan Schaeffer), Ричарда Шерла (Richard Scherl), Ларса Шустера (Lars Schuster), Сохейль Шамс (Soheil Shams), Стюарта Шапиро (Stuart Shapiro), Джюд Шавлик (Jude Shavlik), Сатиндера Сингха (Satinder Singh), Дэниела Слитора (Daniel Sleator), Дэвида Смита (David Smith), Брайена Соу (Bryan So), Роберта Спрула (Robert Sproull), Линн Стейн (Lynn Stein), Лэрри Стивенса (Larry Stephens), Андреаса Штолке (Andreas Stolcke), Пола Страдинга (Paul Stradling), Девику Субраманиан (Devika Subramanian), Рича Саттона (Rich Sutton), Джонатана Тэша (Jonathan Tash), Остина Тэйта (Austin Tate), Майкла Тилшера (Michael Thielscher), Уильяма Томпсона (William Thompson), Себастьяна Траны (Sebastian Thrun), Эрика Тидеманна (Eric Tiedemann), Марка Торранса (Mark Torrance), Рэндалла Уфама (Randall Upham), Пола Утгоффа (Paul Utgoff), Питера ван Бека (Peter van Beek), Хала Вариана (Hal Varian), Сунила Вемури (Sunil Vemuri), Джима Уолдо (Jim Waldo), Бонни Веббер (Bonnie Webber), Дэна Вэлда (Dan Weld), Майкла Веллмана (Michael Wellman), Майкла Дина Уайта (Michael Dean White), Камина Уайтхауза (Kamin Whitehouse), Брайена Уильямса (Brian Williams), Дэвида Уолфа (David Wolfe), Билла Вудса (Bill Woods), Олдена Райта (Alden Wright), Ричарда Йэна (Richard Yen), Вейшионг Джанг (Weixiong Zhang), Шломо Зильберштейна (Shlomo Zilberstein), а также анонимных рецензентов, привлеченных издательством Prentice Hall.

Об обложке

Изображение на обложке было спроектировано авторами и выполнено Лайзой Мэри Сарденья (Lisa Marie Sardegna) и Мэриэнн Симмонс (Maryann Simmons) с использованием программ SGI Inventor™ и Adobe Photoshop™. На обложке показаны перечисленные ниже предметы, иллюстрирующие историю искусственного интеллекта.

1. Алгоритм планирования Аристотеля из книги *De Motu Animalium* (ок. 400 до н.э.).
2. Генератор понятий Раймунда Луллия из книги *Ars Magna* (ок. 1300).
3. Разностная машина Чарльза Бэббиджа, прототип первого универсального компьютера (1848).
4. Система обозначений Готтлоба Фреге для логики первого порядка (1789).
5. Диаграммы Льюиса Кэрролла для формирования логических рассуждений (1886).
6. Система обозначений вероятностной сети Сьюэлла Райта (1921).
7. Алан Тьюринг (1912–1954).
8. Робот Shakey (1969–1973).
9. Современная диагностическая экспертная система (1993).