

# Содержание

<b>Предисловие .....</b>	<b>5</b>
<b>Глава 1. Задача тысячелетия .....</b>	<b>7</b>
Интеллект – это, собственно, что такое? .....	9
Критерий Тьюринга .....	11
История проблемы .....	14
Формализация мышления .....	15
Психология мышления .....	23
Эвристические алгоритмы .....	26
В заключение .....	30
<b>Глава 2. Вся жизнь – игра .....</b>	<b>32</b>
Компьютер против человека. Как это выглядит в принципе? .....	34
Как эту работу выполняет человек .....	34
Первая базовая идея – дерево перебора .....	36
Вторая базовая идея – оценочная функция .....	38
Оптимизация минимаксной процедуры. Альфа-бета-алгоритм .....	45
Этапы игры .....	48
Дебют .....	49
Эндшпиль .....	50
Оценка, основанная на приоритетах факторов .....	52
Интересная гипотеза. Интегральный признак .....	53
В заключение .....	56
<b>Глава 3. Интеллект искусственный и обучаемый .....</b>	<b>58</b>
Проблемы построения обучаемых систем .....	59
Игра как проблема обучения .....	62
Как научить машину учиться игре .....	63
Доминирующие факторы .....	65
Метод корректировки оценки с опорой на доминанту .....	66
Изменение состава оценочной функции .....	67
Несколько идей общего характера. Короткий опыт .....	68
Несколько идей общего характера. Фреймы .....	71
Несколько идей общего характера. Причинно-следственные связи .....	73
В заключение .....	75

<b>Глава 4. Сетевая архитектура .....</b>	<b>76</b>
Нейрон .....	79
Обучаемость нейронной сети .....	80
Сеть нелинейной геометрии.....	88
Персептрон Розенблатта .....	90
Сети Кохонена .....	93
Звезды Гроссберга.....	96
В заключение.....	97
 <b>Глава 5. Распознавание образов .....</b>	 <b>98</b>
Выделение объекта из среды.....	100
Более тонкая процедура отделения .....	102
Отделение цветом .....	104
Отделение областей пространства.....	105
Идентификация образа по шаблону.....	108
Метод малых преобразований.....	109
Идея преобразования в общем виде .....	111
Распознавание объекта по набору признаков .....	112
В заключение.....	115
 <b>Глава 6. Искусственное познание .....</b>	 <b>118</b>
Камни преткновения на пути искусственного интеллекта .....	118
Логический вывод и доказательство теорем .....	124
Набор правил вывода машины «Логик-теоретик» .....	128
Эвристические механизмы машины «Логик-теоретик» .....	129
Базы знаний .....	131
Знание как предложение естественного языка .....	133
Картина мира как конструкция классов, объединяющих родственные объекты.....	134
Операционное знание.....	138
Самоорганизующийся искусственный интеллект .....	140
Последнее замечание .....	143
 <b>Глава 7. Интеллект, равный человеческому?! .....</b>	 <b>145</b>
Не просто техническая проблема.....	145
Можно ли улучшить тест Тьюринга .....	147
Проблема номер один .....	149
Минимальный разум .....	150
Экспертное мнение.....	153
 <b>Литература.....</b>	 <b>169</b>

# Предисловие

Эта книга о проблеме, которую очень многие ученые считают ключевой для развития человеческих технологий, – проблеме искусственного интеллекта. Таковой она является по той причине, что создание полноценного разума, равнозначного человеческому или даже его превосходящего, может означать, что дальнейшее развитие земных технологий уже может пойти без приставки «человеческие». Само-развивающийся разум станет, в лучшем для людей случае, нашим полноправным партнером в развитии цивилизации на планете.

Впрочем, пока это область фантастики. Самые лучшие представители систем искусственного интеллекта пока совершенно не обнаруживают ключевых возможностей разума: интуиции, творчества, свободы мышления. Поэтому пока термин «искусственный интеллект» именуется собой инженерную дисциплину, цель которой – лишь имитировать, в частных задачах, поведение, которое мы, люди, считаем разумным. Каким образом человек, рождаясь практически чистым от знаний, получает способность обучаться и развивать свой разум, почему мы способны принимать решения при недостатке информации, каковы механизмы творчества – это все вопросы, пока слишком сложные для теории искусственного интеллекта.

Ряд крупных ученых даже полагает, что эта теория так всегда и останется чисто инженерной областью, впрочем, может быть, и наш интеллект – просто сверхсложная инженерная конструкция, но тогда, как говорит Дэвид Вернон, мы, может быть, и не заметим, что полноценный интеллект уже живет среди нас.

В общем, эта книга об интеллекте, проблемах его разработки и критериях его распознавания. Таковой критерий сам по себе представляет собой отдельную задачу. Действительно, если мы видим перед собой автомобиль, или воздушный шар, или телескоп, нет необходимости в специальных процедурах высокой степени сложности, чтобы убедиться в том, что мы видим автомобиль или воздушный шар. А вот сложность вопроса, что перед нами: машина, работающая по алгоритму, или мыслящее существо, – неимоверно велика. Это вопрос философский, мировоззренческий, ответ на который сопоставим с созданием интеллекта.

В книге некоторые идеи даны лишь в виде набросков, сложные теории нейронных сетей и распознавания образов излагаются на уровне общих положений, но вы читаете не научную монографию. Я ставил перед собой цель ввести своего читателя в проблему, не требуя специальных знаний и хорошей математической подготовки, поэтому перед вами – лишь популярное изложение. Но я надеюсь, что, прочитав эту книгу, вы получите хорошее представление о том, чем занимаются специалисты по искусственному интеллекту, что это дает человеческой науке и технике и насколько до сих пор на самом деле неясен сам этот термин «интеллект».

# Глава 1

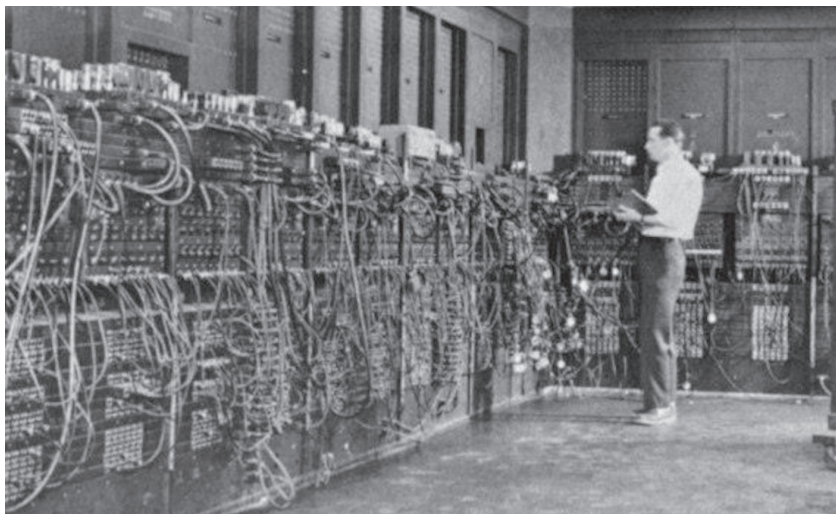
## Задача тысячелетия

Природа создала человека очень слабым. Мы медленнее всех бегаем, у нас нет острых зубов и когтей, мускульная сила оставляет желать лучшего, и даже способность наблюдать окружающий мир с помощью органов зрения и слуха сомнительна в сравнении с возможностями животных. Единственное наше преимущество, которое, впрочем, оказалось в эволюционной борьбе решающим, — это способность мыслить, делать выводы, накапливать знания. Разум дал возможность компенсировать все недостающее. Человек начал развивать науку в самом общем смысле этого слова, наука дала не только понимание окружающего мира, но и способы создания искусственной среды, комфортной для человеческого существования. Мы придумали технику, заменившую мускульную силу, появились устройства и технологии, не просто способные сдвинуть большой каменный блок или повалить дерево, но могущие выполнить тонкую работу с глубоко дифференцированными движениями. Машины стали делать работу, в принципе непосильную человеку, не по энергетике, а по сложности движений.

Все это было здорово, но интересно другое. В то время как люди уже сотни лет пользовались сложнейшими машинами в помощь своему телу, устройства, используемые в помощь интеллекту, были на удивление примитивны. Мы уже летали в небе своей планеты, строили железные дороги, высотные дома, в общем, много чего умели, и в это же время наши интеллектуальные помощники — это арифмометр, логарифмическая линейка и т. д., в общем, устройства, принципиально не отличающиеся от бухгалтерских счетов.

Наконец, во второй половине XX века произошло событие, качественно меняющее ход развития человеческой цивилизации. Наш разум впервые получил очень серьезного помощника в лице ЭВМ (элек-

тронно-вычислительной машины). И хотя прабабушка современного компьютера – ламповая ЭВМ – занимала большую комнату, было ясно, что это не просто большой арифмометр. Ее фотография представлена на рис. 1.1.



**Рис. 1.1** ❖ «Эниак», ламповая ЭВМ

Даже первые машины могли перестраивать свою работу в соответствии с выполняемой программой. Правда, первые программы вводили в память машины не с клавиатуры, которой еще не было, но это уже были программы.

Надо сказать, однако, что первые разумозаменители, как и их потомки с терабайтами памяти и многоядерными процессорами, обладают принципиальным отличием от машин, заменяющих мускульное усилие.

Развитие техники медленно, но верно идет по пути исключения человека из выполняемого действия. Кое-где это уже произошло. Достигли своего идеала станки с ЧПУ, самолеты-беспилотники, системы управления атомными станциями и т. д. Человек здесь опустился (или поднялся, это кому как нравится) до функции чистого контроля. Компьютер же для своей работы потребовал новой касты высококвалифицированных специалистов, называемых программистами, которые и обеспечивают работу интеллектуальных машин, и это в то

время, когда мускульное усилие в производстве неотвратимо теряет свои позиции. Стало ясно, что машины могут заменить человека во всех видах активной, но не творческой деятельности. Большая же и важнейшая часть интеллектуальной работы для компьютера сегодня так же недоступна, как и для первых ЭВМ. Осмелюсь даже предположить, что и появление в некоем абстрактном будущем квантовых компьютеров с работоспособностью, на порядки превышающей все, что мы имеем сейчас, и практически безграничной памятью в этом смысле ничего не изменит.

И это «НО» означает, что появление компьютера – только преддверие к революции. Умение быстро обрабатывать большие объемы информации и выполнять непосильные для человека вычисления – это хорошо, но не так уж принципиально. А вот настоящая революция, способная изменить самое лицо нашей цивилизации, произойдет с появлением на планете искусственного интеллекта. Осталось только понять, что это такое, и затем реализовать его в осязаемом устройстве.

## Интеллект – это, собственно, что такое?

Для создания искусственного разума надо понять, что такое интеллект вообще. Вопрос очень непрост. Например, если разум определить как способность разрешать проблемы теоретической физики, то окажется, что мало кто на планете обладает интеллектом. Определить же его как качественную характеристику именно человека означает ограничить понятие его человеческой разновидностью, и, кроме того, в таком определении не содержится никакой полезной информации, а значит, придется отвечать на вопрос: что же такого имеет человек, что свойственно только ему?

***Как говорят мудрецы, чтобы вещью овладеть,  
ей надо дать имя***

Поэтому получить точное определение все же хотелось бы. Представьте себе, что вам принесли некую машину (или привели внеземную зверюшку), и утверждается, что она обладает интеллектом. Как это проверить? Нужны определение интеллекта (критерий) и точная процедура проверки этого критерия. Рассмотрим такое рабочее определение:

Интеллект – качество психики, состоящее из способности адаптироваться к новым ситуациям, способности к обучению на основе опыта, пониманию и применению абстрактных концепций и использованию своих знаний для управления окружающей средой.

Оно опирается на понятия, которые, будучи интуитивно ясны, все же неточны и нуждаются в доопределении. К этому определению можно задать слишком много вопросов:

- Что такое знание?
- Что такое абстрактные концепции?
- Что значит обучение на основе опыта? Обучить дрессировкой можно очень многие виды животных. Означает ли это наличие у них интеллекта? А что такое обучение без дрессуры? В наших, человеческих школах очень многое делается именно дрессурой, означает ли это, что люди – школьники не обладают интеллектом?
- Что считать новой ситуацией и что считать успешной адаптацией? Очень часто люди теряются в необычной ситуации, означает ли это, что они не разумны?

Можно привести и другие, столь же мудреные и ненадежные определения. В чем их недостаток? Они для определения интеллекта используют сложные понятия. Это грубая логическая ошибка. Хорошее определение может свести понятие к более общему с добавлением достаточного уточняющего признака. Например, белая береза – это дерево с белой корой. При условии что понятие «дерево» уже определено и больше деревьев с белой корой не существует, это определение вполне удачно. Причем черную березу (а такая тоже есть) мы в определении отменили, так как ограничились именно белой.

Еще хорошее определение может свести термин к комбинации более простых. Например, окружность – это множество точек плоскости, равноудаленных от одной точки, называемой центром. Понятия точки, множества, плоскости проще, нежели понятие окружности. Кроме того, этим определением мы уходим от необходимости пояснять такие сложные вещи, как непрерывность и кривизна линии.

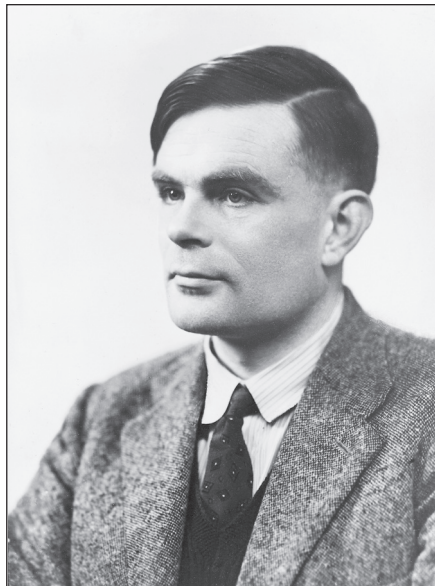
А еще хорошее определение может воспользоваться понятием эквивалентности. Например, синий цвет – это цвет неба. Круглый – это форма футбольного мяча. В общем, способов дать хорошее определение достаточно много. Заметим еще, что хорошее определение должно быть рабочим, то есть применимым для идентификации объектов. Это тоже не всегда просто. Например, можно определить километр как расстояние, которое способен пройти взрослый человек по прямой, не торопясь, за 10 минут. Допустим, что любой взрослый человек идет, не торопясь, с одной и той же скоростью (это не так, но допустим), то есть это точное определение. Но оно все равно не рабочее, так как измерять таким способом расстояния в сотни километров нереально.



Итак, какой способ нам выбрать? Свести к более общему с указанием определяющего свойства не получится. К сожалению или к счастью, но нам не известен ни один пример инопланетного разума. Животным, живущим рядом с нами, мы в разуме отказываем. Это означает, что у человечества нет возможности определить интеллект «вообще», без привязки к его человеческой разновидности. Разложить на более простые вещи не получится, для этого надо хорошо понимать, как он работает и из чего состоит. Этот метод фактически работает только на простых объектах. *Единственное, что нам остается, – это признать интеллект свойством, присущим каждому человеку (здоровому в медицинском смысле), и выделить некие вещи, которые может делать только человек, и только потому, что он разумен. Таким образом, мы для построения теста на интеллектуальность вынуждены выбрать человека как эталон.*

## Критерий Тьюринга

Европейский математик Алан Тьюринг предложил следующий критерий:



**Рис. 1.2** ❖ Алан Мэтисон Тьюринг

**Для справки.** Изображенный на рис. 1.2 Алан Мэтисон Тьюринг – английский математик, логик, криптограф, оказавший существенное влияние на развитие информатики. Кавалер Ордена Британской империи, член Лондонского королевского общества. Предложенная им в 1936 году абстрактная вычислительная «машина Тьюринга», которую можно считать моделью компьютера общего назначения, позволила формализовать понятие алгоритма и до сих пор используется во множестве теоретических и практических исследований.

*«Человек взаимодействует с одним компьютером или одним человеком (с кем он говорит, ему неизвестно). На основании ответов на вопросы он должен определить, с кем он разговаривает: с человеком или компьютерной программой. Если человек не сможет сделать правильный выбор, значит, компьютерная программа обладает интеллектом».*

Предложенный Тьюрингом критерий основан на том убеждении, что человеческая речь, имеется в виду, конечно же, осмысленная речь, есть неотъемлемая часть разума. Поэтому если я – разумное существо и мой собеседник может неограниченно со мной беседовать, и я не могу уличить его в бессмысленности, то он тоже разумен.

Критерий Тьюринга выглядит очень солидно, но необходимо сделать важное уточнение. Например, если я – специалист по переработке нефти, то я не имею права требовать от партнера знаний о процессе переработки. А это значит, что возможные бессмыслицы, которые он скажет, если побоится сознаться в своем незнании, не в счет. А утайка некомпетентности для разумного человека – дело, достаточно обычное. Это означает, что в тесте допустимы только темы общего порядка. Но и здесь мы пойдем по тонкому льду. Я знаю, что Волга впадает в Каспийское море. Мой партнер может этого не знать, следовательно, даже в бытовых темах нужно делать скидку на уровень образования. То есть незнание не эквивалентно бессмыслице.

Анализируя эти затруднения, легко прийти к выводу, что тест Тьюринга содержит в себе серьезное противоречие. С точки зрения теста, мы обязаны для обеспечения чистоты теста определить, что такое осмысленная речь. Фактически требуется понятие интеллекта определить через очень сложное понятие осмысленной речи. Это своего рода тавтология.

Единственный выход из создавшегося положения – отбросить все наукоемкие и философски солидные рассуждения, признать человеческую речь осмысленной по определению, на всякий случай отбросить все умные темы, ограничившись общими понятиями.

ми, и поговорить с программой. Однако в этом пункте разработчиков искусственного разума ожидало самое главное разочарование. Оказывается, так называемая осмысленная речь достаточно легко алгоритмизируется. Первый виртуальный собеседник появился уже в 1966 году – программа Элиза, созданная Джозефом Вейзембаумом. Элиза моделировала речевое поведение, используя технику активного слушания, например используя фразы «Продолжайте, пожалуйста». Такие собеседники стали создаваться регулярно программистами и лингвистами, интересующимися проблематикой искусственного интеллекта, и с учетом того, что тест Тьюринга не имеет математически строгой формулировки и не вполне понятно, о какой ситуации можно сказать определенно «Программа тест прошла», то ценность самого теста таяла с появлением каждого нового виртуального собеседника.

*И наконец, самая большая проблема. Человек обладает способностью осознавать то, что он делает. В отношении машины, отрабатывающей успешно тест Тьюринга, остается вопрос, а понимает ли машина, что она ведет беседу. В этой книге мы еще вернемся к этому вопросу, а пока заострите свое внимание на том факте, что обнаружение рефлексии на себя представляет задачу особой сложности.*

Конечно, хотелось бы завершить этот текст на позитиве. Вроде того: «И наконец, они придумали идеальное определение и построили разумную машину, начав новую историю человечества». Но ничего подобного не произошло. Сегодня, что такое интеллект, точно так же не понятно, как и на заре создания теории ИИ, а значит, как и прежде, не ясно, как нам реагировать на появление еще одной умной машины. Мы уже знаем, что выигрыш машиной у шахматного мастера, даже гроссмейстера, не означает разума, точно так же мы не обнаруживаем интеллект у программ, управляющих в режиме реального времени сложнейшими техническими системами. С каждой новой разработкой мы просто становимся свидетелями еще одного доказательства, что превосходство в еще одной частной сфере не означает возникновения интеллекта у машины. Проблема тысячелетия по-прежнему остается открытой.

Однако теория искусственного интеллекта существует, существуют и одноименные технические системы, используемые в реальных задачах. Искусственный интеллект, например, управляет системами защиты боевого корабля, позволяя исключить человеческий фактор

в условиях динамичного и сверхманевренного боя, в котором есть необходимость отбиваться сразу от большого количества нападений ракетами, самолетами, торпедами, движущимися очень быстро и способными маневрировать. Это реально работает. Так что такое искусственный интеллект?

В истории развития теории разума произошла следующая метаморфоза. Ученые, довольно основательно побившись головой о стенку философского осмысления понятия интеллект, обнаружили, что для решения конкретных задач нет необходимости во всеобщем, универсальном, самообучаемом и т. д. интеллекте. Это примерно то же самое, как нет необходимости моделировать кузнеца-человека, чтобы построить машину, штампующую из металла конкретные детали. Интеллектуальные процессы тоже можно моделировать и создавать конкретные программы под конкретные задачи. Таким образом, современная теория искусственного интеллекта перешла к созданию технических систем, работающих так, как будто они разумны, но только в рамках вполне определенной задачи.

Это направление человеческой мысли сейчас доминирует, но, конечно, задача разобраться в том, как работает интеллект вообще, осталась, и, может быть, и она когда-нибудь будет разрешена, но сейчас мы уже понимаем, насколько это сложно в действительности, и сколько уйдет на решение времени, совершенно не понятно.

## История проблемы

К задаче определения искусственного интеллекта можно подойти с разных сторон. На мой личный взгляд, наиболее сложный способ заключается в том, чтобы попытаться дать определение без привязки к нашему человеческому мышлению и дать строгое, математически точное описание. Почему это сложно? Да потому, что надо дать исчерпывающее определение в терминах, не включающих информацию о человеческом разуме. Может быть, это было бы проще, если бы удалось поговорить с парой-тройкой представителей инопланетных цивилизаций. Но такой возможности нет. Другие подходы более реальны. Можно попытаться создать математическую теорию, описывающую именно наше, человеческое мышление, и создать устройство, работающее по этой теории, но не копируя человеческий мозг. И можно исследовать мозг в деталях и сделать его точную копию.

Нейрофизиологические исследования головного мозга дали огромную груду информации. Сегодня мы знаем, какие функции локали-

зованы в специфических отделах головного мозга, а какие размазаны по всему человеческому мыслительному аппарату. Выяснен главный системотехнический принцип. Оказывается, сверхвысокая эффективность работы мозга обеспечивается очень примитивными элементами – нейронами (нервными клетками). Каждый нейрон в отдельности практически ничего не умеет, но именно в этом ничегонеумии и заключается секрет успеха. Их примитивность на самом деле проявляется как универсальность. Действительно, если некий исполнитель способен выполнять только простую операцию и не желает вникать в общую постановку задачи, то такого исполнителя можно вставить в любую схему, большое количество элементов которой за счет специальной организации уже будет способно на многое. Нейронная идея даже породила целое направление в теории искусственного интеллекта – так называемые нейронные сети. И хотя пока самая сложная нейронная сеть не способна приблизиться по своим возможностям к человеческому мозгу, это направление считается очень перспективным.

Что же касается математически точной теории разума, то здесь положение дел очень туманно. Пока хорошая математика присутствует только в описании нейронных сетей и в теории эвристических алгоритмов, немного, конечно, но пока так. И это несмотря на то, что попытки математического осмысления проблемы имеют очень долгую историю. Здесь ожидалось большие прорывы в силу ошибочного сведения мышления к одной из его форм – строгому логическому выводу. Впрочем, может быть, никакой особой ошибки в этом и нет, просто логика – такая вещь, которая наиболее просто формализуется и поддается исследованию, а пытаться пройти простым путем свойственно для нашего разума.

## Формализация мышления

Отметим сразу, что единственное устройство, позволяющее моделировать искусственный интеллект, – это компьютер, работающий под управлением алгоритмов, представляющих собой последовательность команд, каждая из которых должна быть однозначно понимаема. Плюс к этому компьютер способен выполнять в одно и то же время только одну команду алгоритма. Существование параллельных вычислений в этом смысле мало что меняет. Возможность параллельных алгоритмов означает существование внутри алгоритма независимых частей, что-то вроде более простых алгоритмов. Эти технические ограничения мы обязаны иметь в виду при всех дальнейших рассуждениях. Системотехнические ограничения очерчивают жесткие гра-

ницы возможного. А надо сказать, что архитектура вычислительной системы – на самом деле главный фактор эффективности. Настолько важный, что суперкомпьютеры стали таковыми не столько за счет высокой скорости работы процессоров, сколько за счет усложнения конструкции. Но пока даже нейронные сети – это не всегда реальные технические устройства, а лишь модели на базе традиционных компьютеров. И, несмотря на то что современная наука, электроника и теория алгоритмов уже уверенно видят новые горизонты производительности, старая добрая архитектура фон Неймана является основным техническим решением.

Первую попытку формализации мышления следует признать за Аристотелем. Конечно, вряд ли древний грек формулировал задачу построения искусственного интеллекта. В античности такая задача не являлась актуальной хотя бы потому, что для древних Земля была населена массой различных мыслящих существ. Современное желание разобраться в вопросе разума, как мне кажется, произошло от осознания уникальности человеческого мышления. И логику силлогизмов, созданную Аристотелем, следует признать попыткой математически точного описания мыслительных процессов. И попыткой, достаточно успешной для того времени.



Рис. 1.3 ❖ Аристотель

**Для справки.** Аристотель – древнегреческий философ. Ученик Платона. С 343 г. до н. э. – воспитатель Александра Македонского. В 335 г. до н. э. основал Ликей (др.-греч. *Λύκειο* – Лицей), или перипатетическую школу. Натуралист классического периода. Наиболее влиятельный из философов древности; основоположник формальной логики. Создал понятийный аппарат, который до сих пор пронизывает философский лексикон и стиль научного мышления. Аристотель был первым мыслителем, создавшим всестороннюю систему философии, охватившую все сферы человеческого развития: социологию, философию, политику, логику, физику. Скульптура на рис. 1.3 естественно представляет предполагаемый облик философа.

Логика Аристотеля еще называется логикой силлогизмов. Силлогизм есть дедуктивное доказательство, состоящее из трех частей: большой посылки, меньшей посылки и заключения. Часто, иллюстрируя логику Аристотеля, в качестве примера приводят следующий силлогизм:

Все люди смертны (большая посылка).  
Сократ – человек (меньшая посылка).  
Следовательно, Сократ смертен (умозаключение).

Силлогистическая форма вывода истинных заключений присутствует во многих логических выкладках в более сложной форме. Мы не будем углубляться в развитие теории, заметим только следующее.

Интуитивно ясно, что силлогизмы дают слишком бедный аппарат для описания мышления. Чтобы в этом убедиться, достаточно посмотреть внимательно на доказательство любой сложной математической теоремы. Не так уж часто можно сделать вывод простым переходом от общего (Все люди смертны) к частному (Сократ смертен). В эту схему, например, не укладываются индуктивные рассуждения (переход от частного к общему). Совершенно не понятно, как логику силлогизмов использовать для доказательства простого утверждения «Ряд натуральных чисел бесконечен» или « $\sqrt{2}$  является иррациональным числом». Математика, по своей сути, не силлогистична. Еще менее полезна логика силлогизмов для естествоиспытателя, в чьей деятельности, например, есть такой внелогичный метод, как метод проб и ошибок, при использовании которого выдвигаются гипотезы, проверяемые экспериментом.

Логика силлогизмов не была успешна в полной формализации человеческого мышления, но начиная с Аристотеля стало ясно, что интеллект – не совсем целостная вещь. В нем есть различные компонен-

ты, например способность к логическому выводу. Для изучения этой способности начала развиваться наука – логика. Со временем логика смогла объяснить многое и выработать точные методы и принципы правильного логического мышления. Вот некоторые из них.

1. Из утвердительных (не путайте с истинными) суждений не может быть сделан отрицательный (не путайте с ложным) вывод.
2. Если одно из суждений отрицательно, то общий вывод будет отрицательным.
3. Закон тождества. Всякое суждение тождественно самому себе.
4. Закон непротиворечия. Два противоречащих суждения не могут быть одновременно истинными, одно из них обязательно ложно.
5. Закон исключенного третьего. Два противоречащих суждения не могут быть одновременно ложными. Одно из них необходимо истинно, другое ложно, третье исключено.
6. Закон достаточного основания. Всякая истина имеет достаточное основание.

О чем говорит, например, первый закон? Сколько бы у вас ни было о некоей вещи или ситуации утвердительных высказываний, ни одно из них не даст основания для отрицания. Например:

- Это яблоко красное.
- Это яблоко сладкое.
- Это яблоко кубанское.
- Это яблоко весит 200 граммов.

Ясно, что мы не можем на основании сказанного утверждать, что яблоко не круглое, но, может быть, можно утверждать, что оно не желтое? Не это ли означает красный цвет яблока, о чем утвердительно заявляется? С позиции обычной бытовой логики да, так, но с позиции строгой формальной силлогизм тогда должен выглядеть следующим образом:

Это яблоко красное (большая посылка).

Красный цвет не есть желтый (малая посылка).

Следовательно, это яблоко не желтое (следствие).

Как видите, для необходимого отрицательного вывода в исходном наборе утверждений не хватает посылки «Красный цвет не есть желтый», а значит, построенный силлогизм не законен. Можно ли алгоритмизировать первый принцип из списка? Очевидно, да. Если в результате цепочки логического вывода получилось отрицательное суждение, то достаточно проверить набор исходных суждений, и если



среди них нет отрицательного, то логический вывод следует признать ошибочным. Единственно, заметим, что наличие отрицательных суждений в исходном наборе посылок еще не гарантирует истинности отрицательного заключения.

Разберем еще один закон. Последний, шестой – закон достаточного основания. Его можно интерпретировать следующим образом: если некоторое суждение истинно, то существуют набор истинных суждений и логическая цепочка, приводящая к искомому суждению от исходного набора. Этот закон также достаточно легко алгоритмируется. Множество исходных суждений конечно. Следовательно, множество возможных логических цепочек (разумной конечной длины), которые можно построить на данном наборе суждений, так же конечно, а значит, достаточно построить все возможные логические цепочки и посмотреть, появится ли искомое суждение среди результатов. Если количество исходных суждений велико, то вычислительный процесс может занять время, столь длительное, что реально эта проверка окажется бессмысленной, но мы сейчас рассматриваем лишь теоретическую возможность, а вообще процесс когда-нибудь закончится, и мы получим вполне определенный результат.

Безусловно, современная формальная логика не исчерпывается шестью законами, это довольно развитая и сложная наука. Но раз нам удалось показать алгоритмируемость двух законов, то можно надеяться, что алгоритмизация всей формальной логики – скорее дело большого труда, нежели принципа. Но вот беда – формальной логики для описания интеллектуальной деятельности явно недостаточно.

### **Во-первых, есть проблема целеполагания.**

#### **Как поставить правильную цель, и что такое вообще правильная цель?**

Рассмотрим простую ситуацию. Пусть процесс логического вывода имеет в своем начале только пять суждений. Для упрощения положим, что вывод осуществляется лишь в форме силлогизмов, и каждое исходное суждение может быть как малой, так и большой посылкой. Тогда имеем  $2^5 = 32$  следствия. Теперь добавим эти следствия как возможные посылки к исходным и получим на втором шаге  $2^{32+5}$  логических следствий. Это уже астрономическое число. Вывод неутешителен. Развивать любую науку во всех возможных и мыслимых направлениях невозможно. Процесс очень быстро потребует ресурсов, которых нет и никогда не будет у человечества.

Наша же наука способна развиваться и получать результаты за осмысленное время потому, что люди умеют ставить перед собой конкретную цель и определять направление исследований, продвигающее к поставленной цели. Ясно, что формальная логика не дает никаких средств для постановки цели, для выделения промежуточных целей, оценки полученного результата. Постановка цели – задача внелогическая, выполняемая какими-то другими механизмами, возможно, находящимися за пределами чистого мышления.

### **Во-вторых, не любая задача логически разрешима**

Методы формальной логики ограничены в своем применении даже в очень простых задачах. Для иллюстрации рассмотрим ставшую уже классической проблему парикмахера. Эта задача достаточно сложно излагается в терминах теории множеств, но для ее популяризации придумана очень интересная и простая формулировка. Итак.



**Рис. 1.4** ❖ Деревенский парикмахер

*Условие* (рис. 1.4 – внешний вид предполагаемого парикмахера). В некоей деревне живут мужчины. Женщины и дети там тоже живут, но нас интересуют только мужчины. Все мужчины делятся строго на две категории: мужчины, которые бреются сами, и мужчины, которых бреет парикмахер, других видов мужчин нет. Парикмахер – тоже мужчина, он тоже живет в этой деревне, и он один. Вопрос: кто бреет парикмахера?

Из условия задачи ясно, что есть возможность применить закон исключенного третьего. Действительно, для парикмахера есть только две возможности: либо он бреется сам, либо он не бреется сам. Это взаимоисключающие суждения, поэтому с необходимостью одно из них ложно, а другое истинно, третьего не дано. Так нам говорит закон исключенного третьего. Однако проведем рассуждения:

*Суждение первое. Парикмахер бреется сам.*

В этом случае парикмахер – это мужчина, который бреется сам, а таких мужчин не бреет парикмахер, а так как он и есть парикмахер, то, следовательно, он сам себя брить не может, следовательно, это суждение ложно.

*Суждение второе. Парикмахер не бреется сам.*

В этом случае парикмахер – это мужчина, которого бреет кто-то другой, но это означает, что его бреет парикмахер, а так как он и есть парикмахер, то получается, что он бреется сам, и мы получили противоречие. Следовательно, и это суждение ложно.

Итак, два суждения взаимоисключающие, оба ложны, а третьего не дано. Как быть в такой ситуации, формальная логика ничего сказать не может. Теоретически проблема решена Давидом Гильбертом, но способом, который просто ограничивает сферу деятельности формальной логики, а значит, решает задачу, убивая окончательно наши надежды положить формальную логику в основу искусственного интеллекта.

Вообще, вопрос, что делать с задачей, которая не решается, – возможно, один из самых интересных в истории и философии человеческой науки. В этом отношении очень показателен разговор двух героев братьев Стругацких из романа «Понедельник начинается в субботу» – двух магов: Федора Симеоновича Киврина и Кристобала Хозевича Хунты:

– Г-голубчики, – сказал Федор Симеонович озадаченно, разобравшись в почерках. – Это же проблема Бен Б-бецалея. К-калиостро же доказал, что она н-не имеет решения.

– Мы сами знаем, что она не имеет решения, – сказал Хунта, медленно ошетинаясь. – Мы хотим знать, как ее решать.

– К-как-то ты странно рассуждаешь, К-кристо... К-как же искать решение, к-когда его нет? Б-бессмыслица какая-то...

– Извини, Теодор, но это ты очень странно рассуждаешь. Бес-смыслица – искать решение, если оно и так есть. Речь идет о том, как

поступать с задачей, которая решения не имеет. Это глубоко принципиальный вопрос, который, как я вижу, тебе, прикладнику, к сожалению, не доступен. По-видимому, я напрасно начал с тобой беседовать на эту тему.

И на самом деле это глубоко принципиальный вопрос. Хочу заметить, что самые большие открытия человеческая наука совершала, перескакивая через нерешаемые и не понимаемые здравым смыслом задачи. Пример тому – борьба с аксиомой параллельных. Есть два противоречащих суждения: параллельные прямые существуют, и параллельные прямые не существуют, – и это та самая ситуация, когда взаимоисключающие утверждения могут быть истинными. Евклид положил, что да, через точку, не принадлежащую данной прямой, можно провести одну и только одну прямую, не пересекающуюся с данной прямой.

Это утверждение с точки зрения Евклида является аксиомой, но уж больно по своей сложности оно похоже на теорему. Поэтому люди две тысячи лет пытались его либо доказать, либо опровергнуть. В XIX веке трое ученых: Гаусс, Лобачевский и Риман – догадались отбросить логические законы и положить, что любое суждение о параллельных истинно, если на его базе можно развить геометрию. Так появились неевклидовы геометрии и совершенно новое понимание свойств пространства и заодно ограниченности формальной логики.

Вернемся к критике формальной логики. Еще древние обнаружили существование парадоксов. Парадокс – это ситуация, когда вроде бы посылки для логического вывода безупречны, сам вывод проведен строго, в полном соответствии с законами логики, но полученный результат откровенно ложен, до нелепости ложен. Одним из первых логиков, описавших такие ситуации, был древнегреческий философ Зенон. Его умозаключения называются апории Зенона. Приведем для примера один из них.

## **Ахиллес и черепаха**

Ахиллес – это древнегреческий воин, могучий, как все мифологические герои. Соответственно, он и бежит быстро. Что такое черепаха, думаю, объяснять нет необходимости. И вот эти двое решили, по Зенону, побегать наперегонки. Ахиллес, понимая, что черепаха бежит несколько медленнее, дал ей фору. То есть сначала стартует черепаха, и лишь спустя некоторое время Ахиллес. А теперь, как говорят фокусники, следите за руками (рис. 1.5 – иллюстрация к задаче).

Конец ознакомительного фрагмента.

Приобрести книгу можно  
в интернет-магазине «Электронный универс»  
([e-Univers.ru](http://e-Univers.ru))