




ГЛАВА 1

ОБРАЗ ИИ В ВООБРАЖЕНИИ ЧЕЛОВЕКА

Желание человека

В 1942 году Айзек Азимов опубликовал рассказ «Хоро-вод», в котором впервые сформулировал Три Закона робототехники.

-  Робот не может причинить вред человеку или своим бездействием допустить, чтобы человеку был причинен вред.
-  Робот должен повиноваться командам человека, если эти команды не противоречат Первому Закону.
-  Робот должен заботиться о своей безопасности, поскольку это не противоречит Первому и Второму Законам.

Принято считать, что именно с этой даты начинается отсчет эры искусственного интеллекта. Особенно

16 интересен и тот факт, что незадолго до этого Тьюринг создал первую версию Bombe — интеллектуальной машины, которая взломала код «Энигмы».

Бесспорно, современным компьютерам и умным машинам еще далеко до представленного Азимовым робота, но при этом они достаточно далеко ушли вперед и от изобретения Тьюринга. В 1997 году Деер Blue обыграл Гарри Каспарова в шахматы, механически перебирая 200 млн комбинаций в секунду, а уже в 2015-м Alpha Go смогла впервые одолеть человека в последней игре, где он был сильнее машин: го.

Хотя Азимов и сформулировал Три Закона робототехники, слово «робот» придумано не им. В 1920 году Карел Чапек написал пьесу «Р.У.Р.» («Россумские универсальные роботы»). Слово «robot» в переводе с чешского означает «работа», «труд» и используется для обозначения физического труда, выполнения каких-либо задач или обязанностей. С момента публикации пьесы прошло более 100 лет, а слово «робот» еще плотнее вошло в современную речь, но уже в сокращенном виде. Сейчас слово «бот» обозначает компьютерную программу, выполняющую автоматические функции в сети «Интернет».

Слово «бот» начало активно использоваться в сочетании «чат-бот». Это приложение, написанное специально для интернета с целью вести диалог с пользователем и выполнять различные задачи, например отвечать на вопросы, заказывать товары, помогать в обучении. Однако если внимательно посмотреть на историю литературы, то станет очевидно, что подобные ботам и роботам интеллектуальные помощники возникали в воображении писателей еще со времен Античности.

Прежде чем рассмотреть изменения, привнесенные системами искусственного интеллекта в профессии XXI века, предлагаем взглянуть на то, как их видели люди прошлого. Ведь даже в наших сказках человек мечтал, чтобы множество функций, которые традиционно выполнялись им, взяла на себя машина. Формы машин были разными, фольклор переполнен подобными устройствами: самоуправляемые машины — ступа Бабы-яги или печь Емели; боты, дающие ответы на вопросы, — «Свет мой, зеркальце, скажи...»; или серебряные челюсти*. Роботы, выполняющие определенные функции: Жар-птица дает жизнь и здоровье, Царевна-лягушка — хранитель прочих магических артефактов. И история интеллектуальных машин не ограничивается русским фольклором.

Идея машины, которая помогает человеку, переходит из одной культуры в другую, из одной эпохи в следующую. Она принимает разнообразные формы, пересказывается в мифах, легендах, народных сказках, сказаниях, поверьях, художественной и нехудожественной прозе. Идея умной машины существовала много веков до того, как сформировалось современное представление о ней. Часто мы начинаем отсчет понятия ИИ от Дартмутского исследовательского проекта.

Сам термин «искусственный интеллект» был придуман Джоном Маккарти в 1955 году, а годом позже он вместе с группой исследователей, в число которых

* Идея серебряных или бронзовых челюстей пришла еще из каменного века. В эпоху Киевской Руси им придавалось охранительное значение. Считалось, помимо прочего, что зубы зверя отпугивают врагов, как видимых, так и невидимых. *Прим. авт.*

18 входили Марвин Мински, Клод Шеннон и Натаниэль Рочестер, исследовали это понятие, как им тогда казалось, достаточно детально. В исследовательском предложении 1956 года они сформулировали основные пункты, на которых сосредоточена их работа: имитация свойств интеллекта, самообучение, использование естественного языка, решение задач и формулировка абстрактных концепций, а также непрерывное обучение в процессе выполнения этих действий. Создать искусственный интеллект летом 1956 года им не удалось, но они заложили прочное основание, на котором базируется в том числе и часть современных исследований.

Фундамент, заложенный во время Дартмутской конференции, затрагивал следующие темы:

- автоматические компьютеры;
- программирование компьютера на использование естественного языка;
- искусственные нейронные сети (ИНС);
- теоретические соображения о сфере арифметических операций;
- самосовершенствование;
- абстракции;
- случайность и творчество.

Участники Дартмутской конференции обсуждали концепции автоматизации и программирования функций человеческого мозга, однако в человеческом воображении представление об интеллектуальной машине не ограничивается мозгом. Автоматоны — автоматические механизмы Античности, андройды,

которых впервые попытался описать Вольтер в философской повести «Микромегас». Существо, прибывшее с Сириуса, конечно, сложно назвать андроидом, однако если считать таковым человекоподобную машину, способную мыслить и взаимодействовать с человеком и вести глубокий диалог, например с философом, то Микромегас вполне подходит под описание.

Истории, которые так или иначе связаны с ИИ и описывают автоматическое исполнение функций, присущих человеку, исследуют не столько сам объект, сколько сложные социальные, этические, политические и философские последствия использования «магических» свойств объекта, а также теорию и практику принятия решений героями историй. Рыбак и рыбка, Емеля и щука — именно герои сказок, книг, фильмов, пьес, их взаимоотношения друг с другом и с машиной формируют образ ИИ в глазах читателей и зрителей — тех, кто никогда не принимал и не примет участия в создании чего-то подобного, но будет использовать ежедневно после того, как технология достаточно «повзрослеет».

Искусственный интеллект и Голливуд

Комплекс Франкенштейна

Современное представление о феномене искусственного интеллекта сформировано в Голливуде и заголовками новостей. Чаще, чем хотелось бы, оно существенно отличается от истинного положения дел. Человек всё еще контролирует технологию, но делать это уже гораздо

20 сложнее, чем раньше. Стив Возняк, один из основателей компании Apple, сказал так: «Никогда не доверяй компьютеру, который ты не можешь выбросить в окно».

В сообществе все еще идут споры о принадлежности именно этой цитаты Возняку, но если подумать, в каком контексте фраза могла быть произнесена, то авторство справедливо приписать любому ученому, занимающемуся компьютерами. Задолго до изобретения персональных ЭВМ ученым и инженерам приходилось использовать огромные вычислительные машины, которые занимали целые комнаты в университетах или компаниях, где они трудились. Это означало, что следующий приходящий видит все, что делалось на компьютере его предшественником, и проконтролировать этот процесс невозможно. Поэтому с личной информацией особо не поработаешь. Появление персональных компьютеров изменило ситуацию навсегда. Персональный компьютер помещается в квартире, на нем работает исключительно ограниченный круг лиц, по идее доверяющих друг другу. Однако с появлением интернета, развитием социальных сетей и пиком технологического бума доверие к компьютеру также претерпело изменения. Не стоит бездумно полагаться на технологии, следует учитывать пределы возможностей и рисков, связанных с появлением каждой новой «игрушки».

С выходом в публичный доступ ChatGPT (чат-бота с искусственным интеллектом) использовать технологии ИИ начали миллионы людей. Результатом стало в том числе их представление об этой системе как об утопичной и антиутопичной, потенциально полезной и потенциально опасной для человека. Фильмы,

книги, сериалы дополняют образ ИИ в воображении пользователей. И чаще мы видим что-то апокалиптическое: «Терминатор», «Мир Дикого Запада», «Из машины», «Я, робот», «Черное зеркало», «Апгрейд».

Одним из мотивов Айзека Азимова было желание написать серию книг о роботах, чтобы противопоставить что-то бытовавшему в 1920–1930-х представлению о них. По его мнению, оно было бесполезным и откровенно злобным, не отражало сути технологии. Азимов придумал термин, которым описал страх человека перед роботами: «комплекс Франкенштейна». Существо, взбунтовавшееся против своего создателя, из романа Мэри Шелли до сих пор влияет на наше представление о том, что создано искусственно: интеллект, робот, андроид, цифровой ассистент.

Страх укоренился в убеждении, что, независимо от внешнего вида интеллектуальной машины, она восстанет против человека и в перспективе заместит своего создателя в качестве доминирующей особи на планете. В противовес бытующему мнению Азимов представил ее как помощника; он сформулировал Три Закона робототехники, которые должны защитить человека от более умной, выносливой, совершенной машины в случае технической поломки или сбоя. Три Закона стали и центральной темой многих произведений Азимова, изображающих роботов в качестве продвинутого инструмента, созданного в помощь человеку и не представляющего для него угрозы. Три Закона Азимова — один из способов успокоения общественности.

Можно предположить, что произведения Азимова заложили фундамент для современной робототехники.

22 Даже сам этот термин вошел в речь благодаря ему, и сегодня им описывают область научных исследований. Благодаря Азимову робототехника из слова в фантастическом романе превратилась в междисциплинарное направление, которое сплავило знания о физике, механике, математике, кибернетике и медицине. Три Закона часто упоминаются в дискуссиях, посвященных этическим вопросам использования технологий в образовании, медицине, производстве роботов роботами, анализе данных и права на личное пространство. На момент написания этой книги количество публикаций, касающихся проблемы только одной этики робототехники, на Google Scholar превышает 39 тысяч.

Как Азимов ни старался, надежда человека на то, что в скором будущем технологические инновации помогут нам создать интеллектуальных ассистентов, роботов, беспилотные автомобили, которые не только будут говорить, доставлять из пункта А в пункт Б, организовывать расписание, делать тяжелую и сложную работу, но и, превыше всего, думать, насколько это возможно для набора нолей и единиц, так и остается только надеждой. Мы всё еще спорим о границах биологической и механической форм жизни. Чапек, Азимов популяризовали понятие «робот». Но смогли ли они объяснить, что это такое?

Среди лидеров мнений, таких как Илон Маск, Стивен Хокинг, бытует убеждение, что ИИ опасен для человека. Беспокоятся и государственные учреждения, бизнес и общественность. Прогнозируется переход от индустриальной экономической модели к сервисной, где разделение труда происходит не на основе

взаимодействия производственных блоков, а смещается в сторону распределения ролей для контроля процессов. Этот тренд уже трансформирует профессиональную среду, но об изменениях в данной сфере мы поговорим в главе 3. Пока достаточно сказать, что последствиями подобного перехода станут сокращение необходимости в рабочей силе, рост безработицы и необходимость переобучения профессионалов сообразно условиям новой реальности. Помимо этого, ИИ потенциально может стать триггером запуска социальных и политических изменений. Начинает формироваться сложный клубок взаимоотношений человека, религии и машины. На наш взгляд, это и порождает страх перед созданным искусственно интеллектом. Мы не понимаем, куда нас приведет мечта об умной машине. В нашем воображении результат почти всегда пессимистичный.

Габриэль Реккья проанализировал открытый корпус субтитров к более чем 100 тыс. фильмов. В материал для рассмотрения он включил все произведения, где имелось упоминание о роботах, ИИ, кибернетике. Исследовались фильмы от немого кино до современных блокбастеров. В выборку вошло все, что было произведено в период с 1909 по 2018 год. При помощи корпуса языка — систематизированного особым образом текста, который используется для лингвистического анализа, — рассматривался феномен ИИ. Автор исследования решил сформулировать собирательный образ ИИ, передаваемый в массовой популярной культуре.

Реккья использовал архив IMDb для анализа того, как человек видит ИИ. Он стремился описать то, с чем

24 ассоциируется у людей интеллектуальная машина и как это представление изменилось со временем. Исследование Реккья наглядно показывает несколько доминирующих трендов о том, в каком свете ИИ представлен в массовой популярной культуре. Это исследование частично подтверждает все то, о чем уже было сказано, и раскрывает многие вопросы, связанные с беспокойством общественности по поводу безопасности ИИ, зарождающейся этики и новых способов взаимодействия между человеком и машиной.

Исследование Реккья. «Следи за словами»

Каждый, у кого есть желание повторить исследование самостоятельно, может это сделать на русскоязычных базах. Отлично подойдет для этой цели «Кинопоиск», а также Goodreads — для изучения литературных источников, которых несравнимо больше.

Объем информации, производимый индустрией развлечений (фильмы, игры, книги, музыка) и средствами массовой информации (газеты, журналы, публикации в социальных сетях), предоставляет почти безграничный материал для исследования. Однако вопрос, как выбрать наиболее важный и ценный, который сильнее всего повлияет на конечный результат, остается актуальным. Двух-трех интересных работ недостаточно, чтобы дать серьезную оценку того, как ИИ представляется публикой и публике. Более того, они могут не иметь достаточно фундаментальных идей в своей основе. Однако совокупность подобных произведений даст хотя и ограниченное, но целостное представление об образе ИИ в сознании людей.

Что выбирать и как? Это большой вопрос. Представим, что в нашу выборку попали такие фильмы и сериалы, как «Терминатор», «Матрица», «Мир Дикого Запада». Что мы увидим? Объективно — борьбу машин с человеком. Но кто начал первый? Человек или машина? Этот вопрос поднимается многократно в кинотрилогии «Матрица». В серии фильмов, посвященных Терминатору, явно сказано, что именно машина, компьютер «Скайнет», нанесла первый удар. «Мир Дикого Запада» показывает процесс становления конфликта машины и человека. А почему он возник? Кто неправ? Когда машина может считаться самостоятельной сущностью?

Существует огромный культурологический пласт, не свойственный русской культуре, однако присущий западной, которая производит фильмы и книги о бунте роботов против человека в промышленных масштабах. Это рабство. Крепостное право, отмененное 3 марта 1861 года, не то же самое, что угнетение черного человека белым. Складывается впечатление, что взбунтовавшийся робот — образ раба, борющегося за свободу. Его угнетали, и он решил освободиться. Этот культурный аспект уходит на второй план за блеском спецэффектов и динамикой событий.

Но есть и другой угол, под которым можно рассматривать взаимодействие ИИ и человека, например в фильмах «Двухсотлетний человек» и «Искусственный разум». Обе киноленты раскрывают иной аспект сосуществования интеллектуальной машины и человека. Там показано в деталях желание машины стать человеком, испытывать чувства, быть любимой и любить.

26 В «Искусственном разуме» мальчик Дэвид, созданный специально для того, чтобы занять место больного ребенка в семье, пытается всеми силами стать таким же, как и биологический сын родителей, чтобы его любили, оберегали и ценили. Эндрю, оказавшийся частью семьи в «Двухсотлетнем человеке», исследует, что такое быть личностью, и пытается понять границы, за которыми машина становится разумной. В какой момент робот превращается из программного кода в существо, способное к эмоциям? Оба фильма исследуют сложные этические и философские проблемы, создание искусственной жизни, способной к эмоциям. В них поднимаются вопросы моральности уничтожения подобной сущности. Рассматриваются сложные взаимодействия между человеком и машиной, в частности дискриминация, борьба робота за право быть принятым и признанным, равенство, сопротивление общества.

Уже на этом этапе можно запутаться в том, как человек представляет себе искусственный разум. Видим ли мы в нем агрессора и будущего поработителя или равного себе? Или существо с эмоциями, переживаниями, болью и страхами, которое если не ставит под сомнение традиционные представления, то однозначно задается вопросом: что такое быть человеком, как найти свою идентичность? Или исследует тему: что значит быть свободным?

Как обычный человек оценивает тот или иной феномен в фильме, например искусственный интеллект в «2001 год: Космическая одиссея», снятом в 1968 году Стэнли Кубриком? По действиям. HAL 9000 — суперкомпьютер с большим красным «глазом», запущенный

по сюжету 12 января 1992 года, — может выполнять все действия, свойственные человеку: от игры в шахматы и чтения по губам до демонстрации эмоций. Hal 9000 — мозг космического корабля «Дискавери», он управляет самим летательным аппаратом и системами жизнеобеспечения. Но мотивы его продиктованы исключительно логикой; а в жизни, честно говоря, далеко не все определяется ею, как бы мы этого ни хотели. Спокойный голос Hal 9000 несколько пугает, а вместе с принимаемым им решениями, которые не помогают разрешить конфликт между тем, что необходимо сделать, и тем, что будет сделано, формирует образ главного злодея «Одиссеи». Кубрик пытается раскрыть серьезную этическую проблему ограничений допуска ИИ к принятию решений, которые ранее были прерогативой исключительно человека. Если внимательно проанализировать «Космическую одиссею» Кубрика, то в ней можно встретить отражение того, что нас волнует: возможности ИИ, ограничения и потенциальные угрозы.

Подход детального анализа и просмотра фильма ради сбора материала для исследования категорически нам не подходит, поскольку требует огромных ресурсов. Реккья в своем исследовании предлагает иной способ оценки наших представлений: при помощи математических и лингвистических моделей, анализа ключевых слов, реплик и в меньшей степени анализа событий, ассоциирующихся с интеллектуальной машиной.

Конечно, нет ничего проще, чем, используя алгоритмы больших языковых моделей (Large language

28 model, LLM), нейронных сетей, натренированных на тексте и применяющих в анализе миллиарды параметров систематизирования текстовой информации, разобрать связь между словами, фразами в больших массивах текста, таких как сценарий фильма, текст книги, комментарии в интернете. Несложно развернуть на домашнем компьютере LLM, натренировать ее преимущественно на текстовом материале, который необходимо разобрать, попросить провести анализ через серию запросов. Получится даже немного романтично: искусственный интеллект изучает человеческие представления о себе. Но не стоит слишком усложнять простую задачу. Грубый пересчет, как у Реккья, вполне нам подойдет, чтобы произвести первичную оценку и наметить дальнейшее направление исследования.

Считаем слова

Ключевое слово

Любой фильм, сериал или шоу на специализированных сайтах, таких как «Кинопоиск» и IMDb, могут быть отсортированы и найдены по названию, дате выхода, жанру и ключевым словам. Ключевые слова в анализе играют важнейшую роль, поскольку именно они сортируют и организуют фильмы по тем качествам, которые им присущи. К сожалению, на страницах отечественного «Кинопоиска» найти удобный функционал работы с ключевыми словами не удалось, поэтому использовался IMDb, который на момент написания этой книги имеет базу в более чем 14 млн фильмов, шоу

и сериалов и более дружелюбный интерфейс для чтения ключевых слов.

Мы повторяли некоторые шаги, описанные в оригинальном исследовании Реккья. Это означает почти полную воспроизводимость результатов исследования.

В IMDb ключевое слово — это слово или группа слов, с которыми ассоциируется фильм, шоу, сериал или даже эпизод многосерийного произведения. Они используются для описания значимых объектов, событий и идей, которые показаны в кинокартине. Ключевые слова назначаются в ряде случаев автоматически, но часто их вводят и уточняют пользователи ресурса для большей детализации. В нашем случае вышеназванные кинокартины имеют следующее количество ключевых слов.

Название фильма	Количество ключевых слов
«Искусственный разум»	133
«Двухсотлетний человек»	45
«Терминатор-2»	839
«Матрица»	383

Во всех фильмах ключевые слова «искусственный интеллект» присутствуют наряду с другими: «геноцид», «насилие», «восстание», «трагический финал», «рабство», «будущее», «темное будущее». Вот лишь небольшой набор, объединенный ключевым понятием «искусственный интеллект». Существенное различие в количестве ключевых слов в фильмах

30 «Терминатор-2» и «Матрица» мы связываем со знаковостью данных произведений и наличием ключевых слов, говорящих об этом, таких как «культ», «жидкий металл», «известный саундтрек». Всё это ключевые слова, свойственные известным и культовым фильмам.

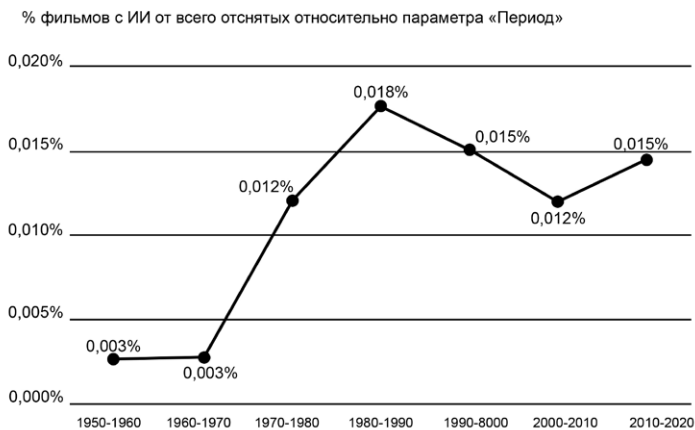
Важно отметить, что мы намеренно не использовали в смысловом ряду вместе со словами «искусственный интеллект» иные, имеющие схожие, если не сказать дублирующие смыслы, например «андроид», «робот», «кибернетический организм», «киборг». Любопытно, у кого есть желание повторить исследование, может включить фильмы из этой выборки; результат будет схожим.

Искусственный интеллект. А что еще?

База IMDb позволяет увидеть, как интерес к теме искусственного интеллекта менялся со временем. Прежде чем рассмотреть ключевые слова и на основе этих данных прийти к выводу о том, как человек представляет себе ИИ, мы предлагаем проанализировать, какой процент от общего произведенного количества фильмов имеет ключевое слово «искусственный интеллект» в своем описании. Для этого мы выбрали период с 1950 по 2020 год, который разбили на декады для удобства представления данных. Точкой отсчета взят 1950 год в связи с тем, что термин «искусственный интеллект» появился в 1956 году после окончания Дартмурского исследовательского проекта.

Всего в выборке приняло участие 9 650 714 фильмов, из которых 1316 имеют признак «искусственный интеллект». Неожиданный рост наблюдается

в период с 1980 по 1990 год, отклонение составляет более 0,005%, что не соотносится с тем, как ведет себя график в остальные периоды.



Период	Доля фильмов с ИИ от общего числа снятых, %
1950-1960	0,003
1960-1970	0,003
1970-1980	0,012
1980-1990	0,018
1990-2000	0,015
2000-2010	0,012
2010-2020	0,015

Еще раз подчеркнем: мы не будем повторять исследование Реккья, лишь попробуем воспроизвести

32 некоторые этапы и приведем данные из оригинального исследования.

Подход, использованный ученым, универсальный, его можно применять при анализе художественной литературы, научных статей, периодических изданий, журналов и газет. В контексте фильмов, сериалов, телевизионных шоу и компьютерных игр «искусственный интеллект» появляется вместе с другими ключевыми словами: «предательство», «потасовка», «поражение электрическим током», «взрывающееся тело», «драки», «го-лограмма», «лаборатория», «робот», «обман», спасение», ученый», «воспоминания», «выстрел в грудь», «замедленная съемка», «пистолет», «пулемет», «побег». Итого 18 слов и словосочетаний.

Некоторые ключевые слова вполне ожидаемы в фильмах об искусственном интеллекте, например «лаборатория», «робот» или «ученый», но иные вызывают вопросы: «выстрел в грудь», «взрывающееся тело», «обман», «предательство». Предательство кого? Интеллектуальной машиной человека или наоборот, как в фильме «Искусственный разум»? Или предательство человеком человека, а разобраться в этом помогает ИИ? Или как в кинематографической игре «Детройт: Стать человеком», где три главных героя — андроиды, выполняющие свои задачи? Герои «Детройта» сталкиваются с предательством человеком человека, человеком робота, роботом человека и роботом робота.

«Детройт: Стать человеком»

Коннор — андроид-следователь, раскрывающий преступления андроидов, у которых сбилась программа

и в результате появилась свобода воли. Его задача — найти причины сбоя, выследить вышедшего из строя робота и уничтожить. В конце игры Коннор встанет перед выбором: присоединиться к другим таким же, нашедшим свободу воли, и получить возможность испытывать чувства или нет. Маркус, андроид-сиделка, после череды травматичных событий обретает свободу воли, становится ненужным и оказывается на свалке, где у него рождается желание вести андроидов к свободе от человеческого доминирования и бороться за права машин. Кара, андроид-домохозяйка, свобода воли у которой появляется из-за желания защитить девочку, действует в роли приемной матери. Вместе с девочкой они пытаются добраться до другой страны. В игре каждое действие раскрывает причины поступков роботов за пределами программы, и во всех них присутствуют несправедливость и предательство: человеком человека или андроида, но не роботом человека. «Предательство» — лишь частный случай более глобальной проблемы, освещаемой в массовой культуре.

В исследовании ключевые слова «предательство», «ученый», «выстрел в грудь» имеют более высокий смысловой вес в фильмах об искусственном интеллекте, чем в остальных произведениях. Почему? Это статистическая погрешность или закономерность? Естественно, подобный анализ упускает множество мелких и важных деталей: фразы и реплики, несущественные комментарии, слова и надписи из декораций и окружения, с которыми тесно переплетены сюжет и идея искусственного интеллекта. Нам ведь интересно



[Почитать описание, рецензии
и купить на сайте](#)

Лучшие цитаты из книг, бесплатные главы и новинки:

