МИНОБРНАУКИ РОССИИ

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ**

**ВЫСШЕГО ОБРАЗОВАНИЯ**

**“ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ”**

Факультет *компьютерных наук*

Кафедра *цифровых технологий*

*Контроль и управление квантовыми системами с использованием алгоритмов машинного обучения*

*ВКР Бакалаврская работа*

02.03.01 *Математика и компьютерные науки*

*Распределенные системы и искусственный интеллект*

Допущен к защите

Зав. кафедрой \_\_\_\_\_\_\_\_\_\_\_\_\_\_*С.Д. Кургалин, д.ф.- м.н., профессор* \_\_.\_\_.20\_\_

Обучающийся \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*Е.Д. Петряев, 4 курс, д/о*

Руководитель \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ *А.Ф. Клинских, д.ф.- м.н., профессор*

Воронеж 2019

**Содержание**

Управление квантовыми системами

* Временная динамика (временное уравнение Шредингера)
* Понятие управления квантовой системой
* Пример (двухуровневая система)

Контроль как процедуры измерения по квантовой теории

* Понятие измерения
* Примеры

Измерение как процедура машинного обучения

* Машинное обучение
* Нейронные сети
* Примеры

Модель «кубит+нейрон»

* Описание модели
* Структура модели
* Программная реализация
* Моделирование и анализ результатов

Выводы

Список использованных источников

**Введение**

Первая глава содержит информацию о структуре искусственных нейронных сетях, поэтому если вы знакомы с данной технологией ее можно пропустить. Вторая глава содержит информацию о квантовых информационных системах. В этой работе рассматривается квантовая теория обучения: теоретические аспекты машинного обучения

использование квантовых компьютеров. Мы описываем основные результаты, известные для трех моделей обучения: точное обучение по запросам членства и, вероятно, приблизительно правильное (PAC) и

агностик учится на классических или квантовых примерах.

Машинное обучение вошло в теоретическую информатику в 1980-х годах с работой Лесли

Valiant [Val84], который представил модель” вероятно, приблизительно правильного " (PAC) обучения, основываясь на более ранней работе Вапника и других в статистике, но добавляя вычислительный комплекс-ити аспекты. Это обеспечило математически строгое определение того, что это означает (эффективно)

изучите целевую концепцию на приведенных примерах. За прошедшие три десятилетия проделана большая работа в теории вычислительного обучения: некоторые эффективные результаты обучения, многие результаты твердости и еще много моделей обучения. Мы ссылаемся на [KV94b, AB09, SB14] для общих введений к этому область. В последние годы практическое машинное обучение получило огромный импульс от успеха глубокого обучения в важных задачах больших данных, таких как распознавание изображений, обработка естественного языка и многие другие области; это теоретически еще не очень хорошо понято, но часто работает удивительно хорошо. Квантовые вычисления начались в 1980-х годах, а также с предложениями для аналогового квантового компьютера на Манин [Man80, Man99], Фейнман [Fey82, Fey85], и Бениофф [Ben82], и достиг более цифровая основа с определением Дойча универсальной квантовой машины Тьюринга [Deu85].

Поле получило импульс с эффективными квантовыми алгоритмами Шора [Sho97] для факторинга целых чисел и вычисления дискретных логарифмов (которые между ними нарушают большую часть сегодняшнего открытого ключа криптография), и с тех пор расцвела в крупную область на перекрестке физики, математики и информатики. Учитывая успехи как машинного обучения, так и квантовых вычислений, объединение этих двух

направления исследований очевидны. Действительно, вскоре после того, как алгоритм Шора, Bshouty и Джексон [BJ99] представил версию обучения на квантовых примерах, которые являются квантовыми суперпозициями, а не случайными выборками. Они показали, что Дизъюнктивная нормальная форма (DNF) может эффективно учиться на квантовых примерах при равномерном распределении; эффективное изучение DNF из единообразных классических примеров (без запросов членства) было и является важным

открытая проблема в классической теории обучения. Servedio и другие [AS05, AS09, SG04] изучали верхний и нижние границы количества запросов квантового членства или необходимых квантовых примеров для обучения, а в последнее время авторы настоящего исследования получили оптимальные оценки квантовая сложность выборки [AW17].

Фокусируясь на конкретных проблемах обучения, где квантовые алгоритмы могут помочь, Äımeur et

Эл. [ABG06, ABG13] показали квантовое ускорение в контекстах обучения, таких как кластеризация через минимальное связующее дерево, делительная кластеризация и k-медианы, используя варианты алгоритма поиска Гровера [Gro96]. В последние несколько лет был шквал интересных результатов применения различные квантовые алгоритмы (алгоритм Гровера, а также оценка фазы, амплитудный усиление [BHMT02], и алгоритм HHL для решения систем линейных уравнений с хорошим поведением [HHL09]) для конкретных задач машинного обучения. Примеры включают основной компонент Анализ [LMR13b], поддержка векторных машин [RML13], кластеризация k-средних [LMR13a], квантовая рекомендательные системы [KP17], и работы, связанные с нейронными сетями [WKS16b, WKS16a]. Некоторые эта работа, как и большинство прикладного машинного обучения в целом, носит эвристический характер

а не математически строгим. Некоторые из этих новых подходов наводят на мысль об экспоненциальных ускорениях по сравнению с классическим машинным обучением, хотя следует быть осторожным в отношении лежащих в их основе предположения, необходимые для эффективного квантового машинного обучения: в некоторых случаях это также сделайте возможным эффективное классическое машинное обучение. Aaronson [Aar15] дает краткое, но четкое описание проблемы. Этим событиям хорошо послужил ряд недавних обследований.

документы [SSP15, AAD15, BWP16] и даже книгу [Wit14]. Напротив, в этом обзоре мы фокусируемся на теоретической стороне квантового машинного обучения: квантовой теории обучения.1 мы опишем (и набросаем доказательства) основных результатов, которые были получены в трех основных моделях обучения. Они будут описаны более подробно в следующем разделы, но ниже мы даем краткий обзор.

Точное обучение. В этой настройке цель состоит в том, чтобы узнать целевую концепцию из способности взаимодействовать с ним. Для конкретности мы фокусируемся на изучении целевых концепций, которые являются булевыми функциями: цель неизвестной с : {0, 1}n→ {0, 1} из известного класса понятий C функций,2 и наша цель-выявить то, с высокой вероятностью, используя запросы членства (которые позволяют ученик, чтобы узнать c (x) для x по своему выбору). Если мерой сложности является только количество запросов,

основные результаты заключаются в том, что квантово точные учащиеся могут быть полиномиально более эффективными, чем классические,

но не больше. Если мерой сложности является время, то при разумной сложности-теоретическая предположения некоторые классы концепций можно узнать гораздо быстрее из запросов квантового членства

(т.е., где учащийся может запросить c на суперпозицию x), чем это возможно классически.

PAC обучение. В этом параметре также требуется узнать неизвестный c: {0, 1}n→ {0,1} из a известная концепция класса C, но более пассивным способом, чем с запросами членства: учащийся получает несколько помеченных примеров (x,c(x)), где x распределяется в соответствии с некоторой неизвестной проблемой- распределение способностей D над {0, 1}n. Учащийся получает несколько независимые одинаково распределенные случайные величины помечены примеры. Из этого ограниченный "взгляд" на c, учащийся хочет обобщить, выдвинув гипотезу h, которая, вероятно, согласна с c на “большинстве " x, измеренных в соответствии с тем же D . Это классический, вероятно, приблизительно нужной модели (ПАК). В квантовой модели PAC [BJ99] пример не является случайным образец, но суперпозиция |x, c(x)>. Такие квантовые примеры могут быть полезны для некоторых задачи обучения с фиксированным распределением D (например, равномерным D), но оказывается, что в обычном независимая от распределения модель PAC, квантовая и классическая сложность выборки равны до постоянные факторы, для каждого понятия C-класса . Когда мерой сложности является время, под разумные теоретико-сложные предположения, некоторые классы концепций могут быть изучены PAC намного быстрее квантовыми учениками (даже на классических примерах), чем это возможно классически.

Агностическое обучение. В этом параметре требуется аппроксимировать распределение на {0, 1}n+1, найдя хорошая гипотеза h для предсказания последнего бита из первых n битов. "Хорошей" гипотезой является то, что не намного хуже, чем лучший показатель в данном классе с существующих гипотез. Агностическая модель имеет больше свободы, чем модель PAC, и позволяет моделировать более реалистично ситуации, например, когда данные шумные или когда не существует” идеальной " целевой концепции. Как в модель PAC, оказывается, квантовая сложность выборки не значительно меньше классической.

Организация. Обследование организовано следующим образом. В разделах 2 и 3 мы сначала представим основные

понятия квантовой теории и теории обучения, соответственно. В разделе 4 описаны основные результаты

получено для теоретико-информационных мер сложности обучения, а именно сложности запроса

точное обучение и примеры сложностей PAC и агностического обучения. В разделе 5 мы исследуем

основные результаты, известные о временной сложности квантовых учащихся. В разделе 6 Мы заключаем:

резюме результатов и некоторые открытые вопросы для дальнейшего исследования.

**2 Введение в квантовую информацию**

2.1 Обозначения

Для общего введения в квантовую информацию и вычисления мы ссылаемся на [NC00]. В этом опрос, предполагаем знакомство со следующей нотацией. Пусть |0>= и |1> = быть стандартные базисные состояния для C2, пространство, в котором “живет”один кубит. Мульти-кубитовые базисные состояния полученный путем взятия тензорных произведений однокубитовых базисных состояний; например, |0> ⊗ |1> ∈ C4 обозначает

базисное состояние 2-кубитной системы, в которой первый кубит находится в состоянии |0i, а второй - в состоянии |1>. Для b∈{0,1}k, то часто стенография|b1>⊗···⊗|bk> как|b1 ···bk>. Чистое состояние k-кубита |φ> может записывается как |φ> = |i>, где ai-комплексные числа (называемые амплитудами), которые имеют удовлетворять = 1. Мы рассматриваем |φ> как 2k-мерный вектор столбца. Вектор строки, который его комплексный конъюгат обозначается <φ|. R-мерное квантовое состояние ρ (также называемое плотностью матрица)-r × r положительная полуопределенная (psd) матрица ρ со следом 1; это может быть записано (часто не-однозначно) как ρ = | hи, следовательно, можно рассматривать как распределение вероятностей по сравнению с чистым

состоянием |φi>.

Не измерять квантовые операции соответствуют унитарные матрицы U , которые действуют на левым умножением на чистое состояние |ψ>(уступая по U |ψ>), и путем конъюгации на смешанных состояний ρ (выход UpU-1). Например, преобразование 1-кубита Адамара H = соответствует унитарная карта H: |a> → (|0> + (-1)a|1>)/, для ∈ {0,1}.

Чтобы получить классическую информацию из квантового состояния ρ, можно применить квантовое измерение к ρ. Квантовое измерение m-результата, также называемое POVM (положительное операторное значение

мера), описывается множеством положительных полуопределенных матриц {Mi}i∈[m], удовлетворяющих Mi=Id. При измерении ρ с использованием этого POVM вероятность результата j задается Tr (Mjρ).

2.2 Модель запроса

В модели запросов вычислений целью является вычисление булевой функции f: {0, 1}N → {0, 1} на некоторые входные данные x ∈[0, 1]N . Нам не дано X явно, вместо этого нам разрешено запрашивать оракула, который

кодирует биты x, т.е., учитывая i ∈ [N], оракул возвращает xi . Стоимость алгоритма запроса количество запросов, которые алгоритм делает к оракулу. Для простоты мы часто будем считать, что N степень 2, N =2n северный

, поэтому мы можем идентифицировать индексы i с их двоичным представлением i1 . . . in ∈ {0, 1}n .

Формально квантовый запрос соответствует следующему унитарному отображению на n + 1 кубитах:

Ox : |i,b>→ |i, b⊕xi >,

Где i∈{0,...,N -1} и b∈{0,1}. Учитывая доступ к оракулу вышеуказанного типа, мы можем сделать фазовый запрос формы Ox,±: |i> → |i> следующим образом: Начните с |i, 1> и примените Адамарово преобразования до последнего кубит получить |i>|−>, где |-> = (|0> − |1>)/. Применить Ox к |i>|-> i для получения |i>|−> . Наконец, примените преобразование Адамара к последнему кубиту, чтобы отправить его обратно в |1>. Мы кратко выделим несколько алгоритмов квантовых запросов, которые мы будем вызывать позже.

2.2.1 Алгоритм Гровера

Рассмотрим следующую (неупорядоченную) проблему поиска. База данных размера N моделируется как двоичная строка x ∈ [0, 1]N . Решением в базе данных является индекс i такой, что xi = 1. Цель задача поиска состоит в том, чтобы найти решение для запроса доступа к x. Нетрудно заметить, что каждый классический алгоритм, решающий задачу поиска, должен в худшем случае делать Ω(N) запросов. Grover [Gro96, BHMT02] придумал квантовый алгоритм, который находит решение с высоким вероятность с использованием O() запросы (это также известно как оптимальное [BBBV97]).

Для N=2 северный, пусть Dn = 2|0nx0n| - Id является унитарным, который ставит " -1 " перед всеми базовыми состояниями, кроме |0n> все. Гровер итерации G = H⨂nDnH⨂n Ox,± является унитарным, что делает один квантовый запрос. Мы сейчас опишите алгоритм Гровера (предполагая количество решений |x | = 1).

1. Начать с |0n> все.

2. Применить преобразования Адамара ко всем n кубитам, получив .

3. Применить итерацию Гровера G раз.

4. Измерьте конечное состояние, чтобы получить индекс i ∈ [N].

Можно показать, что с высокой вероятностью результат измерения является решением. Если число из решений |x>≥ 1 неизвестно, тогда вариант этого алгоритма из [BHMT02] может быть использован для найдите решение с высокой вероятностью, используя O(запросов. Позднее мы будем ссылаться на следующее: более недавнее применение алгоритма Гровера.

Теорема 2.1 ([Kot14], [LL16, Теорема 5.6]). Предположим, x ∈{0, 1}N . Существует квантовый алгоритм

это удовлетворяет следующим свойствам:

• если x≠0N, то пусть d-первый (т. е. наименьший) индекс, удовлетворяющий xd = 1; алгоритм использует ожидаемое число O() запросы к x и выходы d с вероятностью не менее 2 / 3;

• если x = 0N, то алгоритм всегда выводит “нет решения " после O( запросов).

2.2 .2 Выборка Фурье

Очень простой, но мощный квантовый алгоритм-это выборка Фурье. Чтобы объяснить это, давайте сначала познакомимся с основами Фурье-анализа функций на Булевом Кубе (см. [Wol08 ,O'D14] подробнее). Рассмотрим функцию f : {0, 1}n → R. Ее коэффициенты Фурье равны f^ (S) = Ex [f (x)xS(x)], где S ∈ {0, 1}n , ожидание равномерно по всем x∈{0, 1}n , и XS(x) = (-1)xS-символ функция, соответствующая S . Разложение Фурье f равно f = (S)XS. Тождество Парсеваля сказать это (S)2 = Ex[f (x)2]. Заметим, что если f имеет диапазон {±1} , то Парсеваль подразумевает, что квадрат

Коэффициенты Фурье f^(S)2 суммируется в 1 и, следовательно, образует распределение вероятностей. Выборка Фурье означает выборку S с вероятностью f^(S)2. Классически это трудная проблема, потому что

вероятности зависят от всех 2n значения f . Однако, следующий квантовый алгоритм, из-за Бернштейн и Вазирани [BV97] делают это точно, используя только 1 запрос и O(n) ворот.

1. Начать с |0n>.

2. Применить преобразования Адамара ко всем n кубитам, получив .

3. Запрос Of, Получение .

4. Применить преобразования Адамара ко всем n кубитам, чтобы получить

)|S>.

5. Измерьте состояние, получив S с вероятностью f^(S)2.

2.3 Довольно Хорошее Измерение

Рассмотрим ансамбль m d-мерных чистых квантовых состояний, E = {(pi|ψi>)}i∈[m], где pi ≥ 0 и = 1. Предположим, нам дано неизвестное состояние |ψj>, отобранное в соответствии с проблемой- способности {pi} и мы заинтересованы в максимизации средней вероятности успеха для идентификации данное состояние (т. е. найти j). Для POVM M = {Mi }i∈[m] средняя вероятность успеха равна PM (E) = <ψi|Mi|ψi>.

Пусть Popt(E) = maxMPM(E) обозначает оптимальную среднюю вероятность успеха E, максимизированную над множеством всех M-исходов POVMs. Так называемое довольно хорошее измерение (PGM) является специфическим

POVM (в зависимости от E ), что делает достаточно хорошо против E . Мы опускаем детали PGM и укажите только те важнейшие свойства, которые нам нужны. Предположим, Ppgm(E) - средний успех вероятность идентификации состояний в E с использованием PGM, то

Popt(E) ≥ Ppgm(E) ≥ Popt(E)2,

где второе неравенство свидетельствует Барнум и Нилл [BK02]. Для ансамбля E = {(pi,|ψi>)}i∈[m], пусть |ψ'>|ψi> для i∈[м]. Пусть G-матрица M×M Грамма для {|ψ'i>}i∈[m], т. е., G (i,j)=<ψ’i |ψ’j> ji для i, j ∈[m]. Тогда можно показать, что Ppgm(E) =(см.,например, [Mon07] или [AW17, раздел 2.6]).

**3 Модели обучения**

В этом разделе мы определим три основные модели обучения, на которых мы сосредоточимся: точную модель обучения, введенную Англуином [Ang87], модель обучения PAC, введенную Валиантом [Val84], и агностическую модель обучения, введенную Хаусслером [Hau92] и Кирнсом и др. [KSS94]. Ниже концептуальный класс C обычно представляет собой набор функций c: {0, 1}n → {0, 1}, хотя мы также можем разрешить функции c : [N ] → {0, 1} или рассматривать такой c как N-разрядную строку, заданную ее таблицей истинности.

3.1 Точное обучение

Классическое точное обучение. В точной модели обучения учащийся A для класса концепции C получает доступ к членству оракула MQ (c) для целевой концепции c ∈ C, которую a пытается изучить. Учитывая вход x ∈{0, 1}n, MQ(c) возвращает метку c (x). Алгоритм обучения A является точным учеником для C, если:

для каждого c ∈ C, учитывая доступ к оракулу MQ(c): с вероятностью не менее 2 / 3, a выводит h таким образом, что h (x) = c (x) для всех x ∈{0, 1}n .4

Эта модель также иногда известна как” идентификация оракула": идея заключается в том, что C-это набор возможных оракулы, и мы хотим эффективно определить, какой c ∈ C является нашим фактическим оракулом, используя членство запросы к c. Сложность запроса A-это максимальное количество вызовов оракула MQ (c), которое ученик делает над всеми понятиями c ∈ C и над внутренней случайностью ученика. Этот сложность запроса именно обучения C-минимальная сложность запроса над всеми точными учащимися для C. 5 Каждое понятие c: {0, 1}n → {0, 1} также может быть задано его N-разрядной таблицей истинности (с N = 2n),

следовательно, можно рассматривать класс понятия C как подмножество {0, 1}N . Для заданных N и M определите (N , M )-запрос, сложность точной обучения в качестве максимальной сложности запроса обучение именно C,

максимизируется над всеми C ⊆ {0, 1}N такими, что |C| = M .

Квантово-точное обучение.

В квантовой обстановке, вместо того, чтобы иметь доступ к МQ(с) оракул, а

квантово-точный ученик получает доступ к оракулу QMQ(c), который соответствует карте QMQ (c) :

|x, b>→ |x, b⊕c (x)> для x ∈{0,1}n, b∈{0,1}. Для заданного C,N ,M можно определить квантовый запрос сложность точного обучения C, а (N , M )-квантовая сложность запроса точного обучения как квантовые аналоги классических мер сложности.

3.2 Вероятно, приблизительно правильное (PAC) обучение

Классическая модель PAC. В модели PAC учащемуся A предоставляется доступ к случайному примеру оракула PEX (c, D), где c ∈ C-целевая концепция, которую A пытается изучить, и D: {0,1}n → [0,1] является

неизвестное распределение вероятности. При вызове PEX (c, D) возвращает помеченный пример (x, c(x)) где x берется из D . Алгоритм обучения A - это (ε, δ)-обучаемый PAC для C, если:

Для каждого c ∈ C и дистрибутива D, учитывая доступ к PEX (c, D) оракул:

с вероятностью не менее 1-δ a выводит h такой, что Prx~D[h(x) ≠ c (x)] ≤ ε.

Обратите внимание, что учащийся имеет свободу выводить гипотезу h, которая сама по себе не входит в класс концепции C. Если ученик всегда производит h ∈ C, то это называется правильным учеником PAC. Сложность выборки A-это максимальное количество вызовов PEX (c, D) oracle

что ученик делает, над всеми понятиями c ∈ C, распределениями D и внутренней случайностью ученика. Сложность выборки (ε, δ) - PAC класса концепта C является минимальной выборкой com-

плексити над всеми (ε, δ)-PAC учащимися для C .

Квантовая модель PAC.

Квантовая модель PAC была введена Bshouty и Jackson [BJ99]. Вместо того чтобы иметь доступ к оракулу PEX(c , D), учащийся quantum PAC имеет доступ к квантовому примеру oracle QPEX (c, D), который создает квантовый пример

Таким квантовым примером является естественное квантовое обобщение классической случайной выборки.6

Хотя не всегда реалистично предполагать доступ к таким (хрупким) квантовым состояниям, можно, конечно, представить ситуации обучения, когда данные предоставляются когерентным квантовым процессом.

Учащийся quantum PAC получает доступ к нескольким копиям квантового примера и выполняет POVM, где каждый результат связан с гипотезой. Его сложность выборки - это максимальное количество вызовов оракула QPEX(c, D), которое учащийся делает, по всем распределениям D и по внутренней случайности учащегося. Определим (ε, δ)-квантовый образец PAC

сложность C как минимальная сложность выборки по всем (ε, δ)-квантовым PAC-учащимся для C .

Заметим, что из квантового примера |x, c(x)> мы можем получить (-1)c(x)|x> с вероятностью 1/2: применить преобразование Адамара к последнему кубиту и измерить его. С вероятностью 1/2 получаем результат 1, в этом случае оставшееся состояние равно (-1)c(x)|x>. Если D-равномерное распределение, то полученное состояние является именно тем состоянием, которое необходимо на Шаге 3 Алгоритм выборки Фурье, описанный в разделе 2.2 .2.

Как модель квантовых примеров сравнивается с моделью квантовых запросов членства? Если известно распределение D, запрос членства можно использовать для создания квантового примера: учащийся может создать суперпозицию |x|0> и применить запрос членства к целевой концепции c для получения квантового примера. С другой стороны, как уже заметили Bshouty и Jackson [BJ99], запрос членства не может быть смоделирован с использованием небольшого числа квантовые примеры. Рассмотрим, например, задачу обучения, соответствующую поиску Гровера, где класс понятия C ⊆ {0, 1}N состоит из всех строк веса 1. Мы знаем, что Θ() квантовые запросы членства необходимы и достаточны для точного изучения целевой концепции с высокой вероятностью. Однако нетрудно показать, что при равномерном распределении нужны квантовые примеры Ω(N), чтобы точно изучить концепцию цели с высокой вероятностью. Следовательно, моделирование одного запроса членства требует по крайней мере Ω() квантовых примеров.

3.3 Обучение агностик

Классическая агностическая модель. В модели PAC предполагается, что помеченные примеры генерируются идеально в соответствии с целевой концепцией c ∈ C, что часто не является реалистичным предположением. В агностической модели для неизвестного распределения D: {0, 1}n+1→ [0, 1] учащемуся предоставляется доступ к оракулу AEX(D). Каждый вызов AEX (D) создает помеченные примеры (x, b), взятые из распределение D (где x ∈{0, 1}n и b ∈{0, 1}). Определите ошибку h: {0, 1}n→ {0,1} под D как errD (h) = Pr(x,b)~D [h(x)≠b]. Когда h ограничен, чтобы исходить из класса концепции C , минимальная достижимая ошибка-optD (C) = minc∈C{errD(c)}. Алгоритм обучения A является (ε, δ)-агностическим учеником для C, если он может создать гипотезу h ∈ C, ошибка которой не намного хуже: для каждого распределения D на {0, 1}n+1 , учитывая доступ к оракулу AEX(D): с вероятностью не менее 1 − δ, a выводит h ∈ C такой, что errD (h) ≤ optD(C) + ε. Если существует c ∈ C, который идеально классифицирует каждый x с меткой b для всех (x, b), так что D (x, b) > 0,

тогда optD (C) = 0, и мы находимся в настройке правильного обучения PAC. Сложность выборки A-это максимальное число вызовов оракула AEX(D), которое учащийся делает, по всем распределениям D и по внутренней случайности учащегося. (Ε, δ)-агностическая сложность выборки класса концепта C является минимальной сложностью выборки для всех (ε, δ)-агностических учащихся для C.

Квантовая агностическая модель.

Модель квантового агностического обучения была впервые изучена в [AW17].

Для распределения D: {0, 1}n+1→ [0, 1] квантовый агностик имеет доступ к оракулу QAEX(D), который создает квантовый пример . Квантовый агностик получает доступ к нескольким копиям квантового примера и выполняет POVM в конце. Аналогично классические сложности, можно определить (ε, δ)-квантовую агностическую сложность выборки как минимальную сложность выборки над всеми (ε, δ)-квантовыми агностическими учащимися для C .

**4 Результаты по сложности запроса и сложности выборки**

4.1 Сложность запроса точного обучения

В этом разделе мы начнем с доказательства границ квантовой сложности запроса точного изучения класса концептов C в терминах комбинаторного параметра γ (C), который мы определяем вкратце, а затем набросаем доказательство оптимальных границ (N , M )-квантовой сложности запроса точного обучения. В этом разделе мы определим понятие с : {0, 1}n → {0, 1} по N-разрядного правду-таблицы (N=2N), Поэтому c ⊆ {0,1}N. Для множества s ⊆ {0,1}N, мы будем использовать “N-разрядный большинства строку” MAJ(S)∈{0,1}N определяется как: MAJ(S)i=1iff|{s∈S:si=1}|≥|{s∈S:si=0}|.

Определение 4.1 . (Комбинаторный параметр γ (C)) зафиксируйте класс понятий C ∈{0, 1}N размера |C| > 1 и пусть C ' ⊆C. Для i∈[N] и b ∈{0,1} определите

γ'(C',i, b)=

как доля понятий в C', удовлетворяющих ci= b. Пусть γ'(C' , i) = min{γ'(C' , i,0),γ '(C' ,i,1)} - минимальная доля понятий, которые могут быть устранены путем обучения ci . Позволять

γ'(C') =

обозначьте наибольшую долю понятий в C ', которые могут быть устранены запросом. Окончательно определиться

γ(C) =

Это сложное определение мотивировано следующим алгоритмом обучения. Предположим, что ученик хочет точно выучить c ∈ C. жадно, ученик будет запрашивать c о "лучшем" входе i ∈ [N ], т. е. i, который устраняет наибольшую часть понятий из C независимо от значения

ci . Предположим, что j - "лучший" вход (т. е. i = j максимизирует γ '(C, i)), и учащийся запрашивает c по индексу j : по крайней мере, γ (C)-часть понятий в C будет несовместима с результатом запроса, и теперь их можно исключить из C . Назовем множество оставшихся понятий C' и отметим, что |C' | ≤ (1 − γ (C))|C|. Самый внешний min в γ (C) гарантирует, что будет другой запрос, что

учащийся может заставить исключить хотя бы γ (C)-часть оставшихся понятий из C ' и так далее. Мы останавливаемся, когда остается только одна концепция. Поскольку каждый запрос будет сжимать набор оставшихся понятий в коэффициент не менее 1-γ (C), то для сокращения C до {c} достаточно сделать T = O ((log |C|)/γ (C)) запросов.

4.1 .1 Сложность запроса в точности обучения с точки зрения γ (С)

Bshouty et al. [BCG + 96] показал следующие оценки классической сложности точного изучения класса концепции C (мы уже набросали верхнюю границу выше).

Теорема 4.2 ([BCG+96, SG04]). Каждый классический точный ученик для concept class C должен использовать

Ω (max{1/ γ (C), log |C|}) запросы членства. Для каждого C есть классический точный ученик, который изучает C использование O () запросы на членство.

Чтобы показать полиномиальную связь между квантовым и классическим точным обучением, Servedio и Gortler [SG04] показали следующие нижние оценки.

Теорема 4.3 ([SG04]). Пусть N = 2n .Каждый квантово-точный ученик для концептуального класса C ⊆ {0, 1}N должен сделать Ω (max{, }) запросы членства.

Эскиз доказательства. Сначала докажем нижнюю границу Ω(1/ ). Мы будем использовать невесомого противника, связанного с Амбейнисом [Amb02]. Одна из версий этой границы гласит следующее: предположим, что у нас есть квантовый алгоритм с возможными входами D ⊆{0, 1}N и отношением R ⊆ D × D (эквивалентно , двудольный граф) со следующими свойствами:

1. Каждая левая вершина v связана по крайней мере с m правыми вершинами w (т. е., |{w ∈ D : (v , w) ∈ R}| ≥ m).

2. Каждая правая вершина w связана по крайней мере с m' левыми вершинами v (т. е., |{v ∈ D : (v , w) ∈ R}| ≥ m' ).

3. Для каждого i ∈ [N] каждая левая вершина v связана не более чем с l правыми вершинами w , удовлетворяющими vi, wi .

4. Для каждого i ∈ [N] каждая правая вершина w связана не более чем с l' левыми вершинами v , удовлетворяющими vi, wi .

Предположим , что для каждого (v, w) ∈ R конечные состояния нашего алгоритма на входах v и w составляют Ω(1) в норме следа. Тогда квантовый алгоритм делает Ω() запросы.

Теперь мы хотим применить эту нижнюю границу к квантово точному ученику для концептуального класса C. Мы можем думать об алгоритме обучения как о выполнении запросов к N-разрядной входной строке и создании имени понятия c ∈ C в качестве вывода. Предположим, что C ' ⊆ C является минимизатором в определении γ (C) (i.e ., γ'(C') = γ (C)). Определить c = MAJ (C'). Обратите внимание, что c не обязательно должен быть в C' или даже в C, но мы все еще можем рассмотреть, что делает наш ученик на входе c . Рассмотрим два случая:

Случай 1: Для каждого c∈C' , вероятность того, что учащийся выводит c при запуске на типичном понятии c^, составляет < 1/2. В этом случае мы выбираем наше отношение R = {c^} × C' . Вычисляя параметры для противника, мы имеем m = |C'|, m ' = 1, l ≤ γ'(C')|C'| (потому что для каждого i, c^i≠ ci, для γ'(C', i) - доля c ∈ C' и γ '(C' , i) ≤ γ'(C') по определению) и l' = 1. Поскольку для каждого C∈C' обучаемый выводит c с высокой вероятностью на вход c, конечные состояния на каждой паре R-связанные понятия будут Ω (1) друг от друга. Следовательно, количество запросов, которые делает наш ученик, равно Ω () = Ω() = Ω(1/ py (C)) (потому что C' минимизировал γ (C)).

Случай 2: существует определенный c ∈ C ', который учащийся дает в качестве выхода с вероятностью ≥ 1 / 2 при запуске на входе c^ . В этом случае мы выбираем R = { c^} × (C' \{c}), гарантируя, что конечные состояния на каждой паре R-связанных понятий будут Ω (1) друг от друга. Теперь у нас есть m = | C'|-1, m '= 1, l ≤ γ'(C')|C'| (для той же причине, что и в случае 1), и l' = 1. Поскольку (|C'| - 1)/|C'| = Ω(1), граница противника снова дает Ω ().

Теперь мы докажем Ω ((log|C/) / n) нижнюю границу теоретико-информационного аргумента следующим образом.

Просмотрите целевую строку c ∈ C как равномерно распределенную случайную величину. Если наш алгоритм может точно идентифицировать c с высокой вероятностью успеха, он узнал Ω(log |C|) бит информации о c (формально взаимная информация между c и выходом учащегося равна Ω(log |C|)). Из теоремы Холево [Hol73], поскольку квантовый запрос действует только на n + 1 кубитов, один квантовый запрос может дать не более O(n) бит информации о c. Следовательно, Ω ((log |C|)/n) квантовые запросы необходимы.

Оба указанных выше нижней границы являются на самом деле индивидуально оптимальным. Во-первых, если принять , что C ⊆ {0, 1}N состоит из N функций c, для которых c(i) = 1 Для ровно одного i, то точное обучение соответствует неупорядоченной задаче поиска с 1 решением. Здесь γ (C) = 1 / N и Θ() запросы необходимы и достаточно благодаря алгоритму Гроувера. Во-вторых, если C-класс N = 2n линейные функции на {0 , 1}n, C = {c(x) = a · x : a ∈ {0, 1}n }, затем выборка Фурье дает алгоритм O(1)-запроса (см. раздел 5.3). В дополнение к этим квантово-классическим разделениям, основанным на выборке Гровера и Фурье, в разделе 5.3 мы также упоминаем разделение четвертой степени между Q (C) и D(C) из-за

Белову [Bel15], для задачи изучения некоторых k-Хунт.

Комбинируя теоремы 4.2 и 4.3, Servedio и Gortler [SG04] показали, что классическая и квантовая сложность точного обучения существенно полиномиально связаны для каждого C.

Следствие 4.4 ([SG04]). Если концепт класса C имеет классическую и квантовую сложности членства D(C) и Q(C) соответственно, то D(c) = O(nQ(C)3 ).

4.1 .2 (N, M) - сложность запроса точного обучения

В этом разделе мы сосредоточимся на (N , M )-квантовой сложности запроса точного обучения. Классически, следующую характеристику легко доказать.

Теорема 4.5 (Фольклор). (N , M )-сложность запроса точного обучения равна Θ (min{M , N }).

В квантовом контексте сложность (N, M )-запроса точного обучения была полностью охарактеризована Kothari [Kot14]. Улучшая [AIK + 04, AIK + 07, AIN+09], он показал следующую теорему.

Теорема 4.6 ([Kot14]). (N , M )-квантовая сложность запроса точного обучения равна Θ () для m≤N и Θ() для N<M≤2N.

Эскиз доказательства. Рассмотрим сначала нижнюю границу для случая m ≤ N . Предположим, что C ⊆{c ∈{0, 1}N :| c | = 1} удовлетворяет|C| = M . Тогда точно узнать C так же сложно, как неупорядоченная задача поиска на M битах, которая требует Ω() квантовых запросов. Нижняя граница для случая N < M ≤ 2N довольно техническая, и мы отсылаем читателя к [AIN+09].

Теперь мы набросаем доказательства верхней границы. Мы используем следующее обозначение: для u ∈{0, 1}n и S ⊆ [n], пусть uS ∈ {0, 1}|S| - строка, полученная путем ограничения u индексами в S .

Сначала мы описываем квантовый алгоритм, который дает худшую верхнюю границу, чем обещано, но легко объяснить. Предположим, что c ⊆{0, 1}N удовлетворяет |C / = M . Пусть c ∈ C-неизвестная целевая концепция, которую пытается изучить алгоритм. Основная идея алгоритма заключается в следующем: используйте алгоритм теоремы 2.1, чтобы найти первый индекс p1 ∈ [N], при котором C и MAJ(C) отличаются. Это использует ожидаемые o () запросы к c (если нет никакой разницы, то есть c = MAJ (C), то алгоритм сообщит нам об этом после O () запросов, и мы можем остановиться). Теперь мы узнали первые биты p1 из c. Пусть С1 = {Z[N]/[ : z ∈ С, MAJ(C, = MAJ(C } ⊆ {0, 1-множество суффиксов понятий в C, что согласен с MAJ(С) на первом p1 − 1 индексы и несогласие с MAJ(С) на p1 th индекса. Аналогично, пусть c1 = - ” обновленная " неизвестная целевая концепция после ограничения c координатами {p1+1, . . . ,N.} Затем мы используем ту же идею , чтобы найти первый индекс p2 ∈ [N−p1] такой, что (c1, MAJ(C1. Повторяйте это до тех пор, пока не останется только одно понятие, и пусть r-число повторений (т. е. до |Cr| = 1).

Чтобы проанализировать сложность запроса, сначала отметим, что для k ≥ 1 k-я итерация процедуры дает нам pk бит c. Поскольку процедура повторяется r раз, имеем p1 + … + pr ≤ N . Во-вторых, каждое повторение в алгоритме уменьшает размер Ci не менее чем наполовину, т. е. для i ≥ 2, |Ci | ≤ |Ci-1 |/2. Следовательно, необходимо повторить процедуру не более r ≤ O (log M ) раз. Последний запуск будет использовать запросы O() и скажет нам, что мы узнали все биты c. Из этого следует, что общее количество запросов, которые алгоритм делает к c

+ О()≤О() +О()≤O(),

где мы использовали неравенство Коши-Шварца и наши верхние оценки на r и .7 этот алгоритм является O(-фактором от обещанной верхней границы. Настройки алгоритма, чтобы сохранить логарифмического коэффициента используется следующая Лемма из [Heg95]. Он показывает, что существует явное упорядочение и строка si такая, что замена MAJ(Ci ) в базовом алгоритме приводит к более быстрому уменьшению |Ci|.

Лемма 4.7 ([Heg95, Lemma 3.2]). Пусть L ∈ N и C ⊆{0, 1}L . Существует S ∈ {0, 1}Lи перестановка π: [L]→[L], такая, что для любого p ∈[L] имеем |Cp|≤, где Cp = {c ∈ C: = } - множество строк в C, согласующихся с s при π (1),. . . , π (p-1) и не согласны с π(p).

Теперь опишем окончательный алгоритм.

1.Установите C1:=C, N1:=N и c1:=c.

2. Повторять до |Ck| = 1

• Пусть sk ∈ {0, 1-строка, а πk: [Nk] → [Nk] - перестановка, полученная путем применения леммы 4.7 к Ck (с L = Nk )

• Поиск первого (по πk ) расхождения между sk и ck с использованием алгоритма теоремы 2.1 . Предположим, мы находим расхождение в индексе πk (pk ) ∈ [Nk ], т.е., sk и ck согласуются по показателям={πk(1),..., πk (pk-1)}

• Множество Nk+1:=Nk−pk,ck+1:= и Ck+1 := {: u ∈ Ck, =, ≠}

3. Выведите уникальный элемент Ck .

Пусть r-число повторов цикла на Шаге 2 и предположим, что на k-й итерации мы узнали pk бит c. Тогда имеем ≤ N . Общая сложность запроса равна T = O (). Ранее мы имели |Ck+1| ≤ |Ck|/2 и, следовательно, r ≤ O(log M). Но теперь из леммы 4.7 мы имеем |Ck+1| ≤ |Ck|/ max{2, pk}. Поскольку каждая итерация уменьшает размер Ck в max{2, pk}, мы имеем ≤ M . Решение этой задачи оптимизации (т.е. min Т такие что , ≤ N ), Kothari показал

T=O() если M≤N, и T=O( если M>N.

Котари [Kot14], улучшая [SG04, AS05], разрешил гипотезу Хунцикера и др. [HMP+10], показав следующую верхнюю границу сложности квантового запроса точного обучения. Это именно тот алгоритм, который анализируется в терминах γ (C).

Теорема 4.8 ([Kot14]). Для каждого концептуального класса C существует квантово-точный ученик для C с использованием запросов O() квантовое членство.

Мошкин [Mos83] ввел еще один комбинаторный параметр, который мы [Heg95] назвали расширенной размерностью обучения EXT-TD(C) класса концепта C (мы не будем определять EXT-TD(C) здесь, см. [Heg95] для точного определения). Основываясь на работе [Mos83], Hegedus доказал

следующая теорема.

Теорема 4.9 ([Mos83], [Heg95, Теорема 3.1]). Каждый классический точный ученик для класса концепции C должен использовать Ω(max{EXT-TD (C), log |C|}) запросы членства. Для каждого C, есть классическая точное ученика, который учится с помощью с запросами O()членство.

Мошкин [Mos83] ввел еще один комбинаторный параметр, который мы [Heg95] назвали расширенной размерностью обучения EXT-TD(C) класса концепта C (мы не будем определять EXT-TD(C) здесь, см. [Heg95] для точного определения). Основываясь на работе [Mos83], Hegedus доказал

следующая теорема.

Теорема 4.9 ([Mos83], [Heg95, Теорема 3.1]). Каждый классический точный ученик для класса концепции C должен использовать Ω(max{EXT-TD (C), log |C|}) запросы членства. Для каждого C, есть классическая точное ученика, который учится с помощью O() с запросами| членство.

Сравнивая это с теоремой 4.2, заметим, что как 1/γ (C), так и EXT-TD(C) дают нижние оценки классической сложности запроса, но верхняя граница в терминах EXT-TD(C) лучше логарифмическим фактором. Также для анализировать сложность кванта, EXT-TD (C) может быть главным параметром.

4.2 пример сложности обучения PAC

Одним из наиболее фундаментальных результатов в теории обучения является то , что сложность выборки C жестко определяется комбинаторным параметром, называемым размерностью VC C, названной в честь Вапника и Червоненкиса [VC71] и определенной следующим образом.

Определение 4.10 . (Измерение VC) зафиксируйте класс концепции C над {0, 1}n. Множество S={s1,..., st}⊆{0,1}N называется разрушенным классом понятий C, если {(c(s1) · · ·c (st)) : c ∈ C} = {0,1}t. Другими словами, для каждой маркировки l ∈{0,1}t существует c ∈ C такой, что (c(s1)···c(st)) =l. Размерность VC C - это размер наибольшего S ⊆ {0,1}n, который разбивается на C.

Blumer и другие [BEHW89] доказали, что сложность выборки (ε , δ)-PAC класса концепта C с размерностью VC d ниже, ограничена Ω(d/ε + log(1/ δ)/ε), и они доказали верхнюю границу, которая была хуже только log(1/ε)-фактором. В недавней работе Ханнеке [Han16] (улучшение на Simon [Sim15]) избавился от этого логарифмического фактора, показав, что нижняя граница Блюмера и др. на самом деле оптимально. Объединяя эти границы, мы имеем следующую теорему.

Теорема 4.11 ([BEHW89, Han16]). Позвольте c быть концепция класса с VC-dim(С) = d + 1. Тогда Θ (d/ ε + log(1/ δ)/ε) примеры необходимы и достаточны для ученика (ε, δ)-PAC для C .

Это характеризует количество образцов, необходимых и достаточных для классического обучения PAC с точки зрения измерения VC. Сколько квантовых примеров необходимо для изучения класса концептов C размерности VC D ? Тривиально, верхние границы классической сложности выборки подразумевают верхние границы квантовой сложности выборки. Для некоторых фиксированных распределений, в частности равномерного, мы увидим в следующем разделе, что квантовые примеры могут быть более мощными, чем классические примеры. Тем не менее , обучение PAC требует, чтобы учащийся мог изучать c во всех возможных распределениях D, а не только единообразно. Мы показали, что квантовые примеры не более мощные, чем классические примеры в модели PAC, улучшаясь по сравнению с результатами [AS05, Zha10].

Теорема 4.12 ([AW17]). Позвольте c быть концепция класса с VC-dim(С) = d + 1. Тогда для каждого δ ∈ (0, 1/ 2) и ε ∈ (0,1/20) Ω (d/ε +1/ε log(1/δ) необходимы примеры для (ε, δ)-квантового PAC для C .

Эскиз доказательства. d-независимая часть нижней границы имеет легкое доказательство, которое мы опускаем. Чтобы доказать Ω (d/ε), сначала определим распределение D на разбитом множестве S = {s0, . . . , sd } ⊆ {0,1}n следующим образом: D(s0) =1 −20ε и D(si) = 20ε/d для всех i ∈ [d].

Квантовый PAC учащийся получает t копий квантового примера для неизвестного понятия c и должен вывести гипотезу H, которая ε-близка к c. Мы хотим связать это с проблемой идентификации государства в разделе 2.3. Чтобы сделать ε-аппроксимацию c эквивалентной идентификации c, мы используем [d, k, r ]2 линейный код коррекции ошибок для K ≥ d/4, расстояние r ≥ d/8, с генераторной матрицей M ∈ FDxk2 (мы знаем, что такие коды существуют, если d-достаточно большая константа). Позволять

{M z: z ∈ {0, 1}k} ⊆ {0, 1}d-множество кодовых слов 2k в этом линейном коде; они имеют расстояние Хэмминга dH(Mz,My) ≥ d/8 всякий раз , когда z≠ y. Для каждого z ∈{0,1}k рассмотрим понятие cz, определенное на разбитом множестве как: cz(s0) = 0 и cz(si) = (Mz)i для всех i ∈ [d]. Такие понятия существуют в C, потому что S разбивается на C. Кроме того, поскольку r ≥ d/8 , мы имеем [cz(s) ≠ cy(s)] ≥ 5ε/2 всякий раз, когда z≠ y.

Следовательно, с вероятностью не менее 1-δ, квантовый ученик (ε, δ)-PAC, пытающийся ε-аппроксимировать понятие из {cz : z ∈ {0, 1}k}, точно идентифицирует это понятие.

Рассмотрим следующую проблему идентификации состояния:

для z ∈{0,1}k пусть |ψi>= |si,cz(si)>, и E = {(2-k |). Пусть G - Грамм матрица 2K × 2k Для этого E. Из раздела 2.3 мы знаем, что средняя вероятность успеха PGM равна ,z)2.

Прежде чем вычислить (z, z), заметим, что (z, y)-й вход G является функцией z ⊕ y:

G(z, y) = <ψz|ψy>T = (1− |M(z ⊕ y)|)T .

Следующее утверждение будет полезно при анализе

(z, z), запись матрицы грамм.

Теорема 4.13 ([AW17, Теорема 17]). Для m ≥ 10 пусть f: {0,1}m →R определяется как f(w)=(1−β|w|/m)T для некоторых β ∈(0,1] и T∈[1,m/(e3ß)]. Для K≤m пусть - матрица ранга k. Предположим, что матрица A ∈ R2kx2k определена как A(z,y)=(f ◦M) (z⊕y)для z,y ∈{0,1}k , тогда

(z,z)≤ для всех z ∈ {0,1}k.

Мы не будем это доказывать, но упомянем, что доказательство теоремы принципиально использует тот факт, что (z, y)-вход матрицы A является функцией z ⊕ y, что позволяет легко диагонализировать A. Используя теорему и определение Ppgm(E) из раздела 2.3, мы имеем

Thm.4 .13

Ppgm(E) = ≤ .

Существование an(ε, δ)-обучаемого подразумевает Popt(E) ≥ 1 − δ . Поскольку Popt(E)2 ≤ Ppgm(E), приведенная выше величина равна Ω(1), Что означает T ≥ Ω(d/ε).

4.3 Пример сложности агностического обучения

Следующая теорема характеризует классическую сложность выборки агностического обучения в терминах размерности VC.

Теорема 4.14 ([VC74, Sim96, Tal94]). Позвольте c быть концепция класса с VC-dim(С) = d . Тогда Θ( + log(1/ δ)/ε2 примеры необходимы и достаточны для (ε, δ)-агностического ученика для C .

Нижняя граница была доказана Вапника и Червоненкиса [VC74] (см. Также Саймон [Sim96]), а верхняя была доказана Talagrand [Tal94]. Шалев-Шварц и Бен-Давид [SB14, раздел 6.4] называют теоремы 4.11 и 4.14 “фундаментальной теоремой обучения PAC”. Оказывается, квантовая сложность выборки агностического обучения равна (до постоянных факторов) классической сложности выборки. Доказательство нижней границы аналогично доказательству дело Пака.

Теорема 4.15 ([AW17]). Позвольте c быть концепция класса с VC-dim(С) = d. Тогда для каждого δ ∈ (0, 1/2) и ε ∈ (0,1/10) Ω(d/ε2+1/ ε2+log(1/δ) необходимы примеры для (ε, δ)-квантового агностика для C .

Эскиз доказательства. Мы опускаем легкое доказательство d-независимой части в нижней границе. Чтобы доказать часть Ω (d/ε2), аналогичную доказательству теоремы 4.12, рассмотрим линейный код [d , k , r ]2 (для K ≥ d/4, r ≥ d/8) с генераторной матрицей M ∈ . Пусть {Mz: z ∈ {0,1}k} - множество кодовых слов 2k, они имеют расстояние Хэмминга dH(Mz, My) ≥ d/8 всякий раз , когда z≠y. Каждому z ∈ {0,1}k мы связываем a

распределение Dz :

Dz(si,b)= 1/d (1/2 + 10(−1ε), для (i,b)∈[d]×{0,1},

где ={s1,...,sd} разрушается. Пусть cz∈C-понятие, которое помечает S в соответствии с Mz ∈{0, 1}d . Легко видеть, что cz-это концепция минимальной ошибки в C со ссылкой на распределение Dz. Кроме того, любой учащийся, который помечает S в соответствии с l ∈{0, 1}d, имеет дополнительную ошибку dH (M z, l) · 20ε/d по сравнению с cz. Следовательно, с вероятностью не менее 1-δ, an (ε, δ)-квантовый агностик найдет обозначение l такое, что dH (M z, l) ≤ d /20. Как и в доказательстве теоремы 4.12, поскольку Mz является кодовым словом, нахождение l, удовлетворяющего dH (M z, l) ≤ d/ 20, эквивалентно идентификации Mz (и, следовательно, z).

Теперь рассмотрим следующую задачу идентификации состояния: пусть |ψz> = для Z∈ {0,1}k E={(2-k |>) . Пусть G-Матрица грамм для этого E . У нас есть

G(z, y) = <ψz|ψy>T = (1- |M(z ⊕ y)|)T.

Следовательно, (z, y)-запись G зависит только от z ⊕ y, и мы можем использовать теорему 4.13. Подобно доказательству теоремы 4.12, получаем

Thm.4.13

Ppgm(E) = ≤ .

Это означает, что T = Ω (d/ ε2 ) и доказывает теорему.

Мы только что видели, что в сложности выборки для моделей PAC и агностические квантовые примеры не дают преимущества. Гавинский [Gav12] представил модель обучения под названием “прогностический Квант” (PQ), вариацию квантовой модели PAC. Он показал класс реляционной концепции, который полиномиально изучается в PQ, в то время как любая “разумная” классическая модель требует экспоненциального числа помеченных примеров для изучения класса.

4.4 Обучаемость квантовых состояний

Помимо изучения классических объектов, таких как булевы функции, можно также рассмотреть обучаемость квантовых объектов. Ааронсон [Aar07] изучил, насколько хорошо квантовое состояние ρ можно узнать из результатов измерений. Здесь мы предполагаем, что каждое измерение применяется к самому ρ, поэтому нам требуется столько же свежих копий ρ, сколько и количество используемых измерений. Цель состоит в том, чтобы в конечном итоге получить классическое описание квантового состояния σ, которое в некотором смысле близко к ρ—и какое чувство “близости” нам требуется, имеет огромное значение. Изучение такого хорошего приближения ρ в следовом расстоянии называется государственной томографией.

В общем случае n-кубитовое состояние ρ является Эрмитовой матрицей 2n × 2n следа 1 и, следовательно, описывается примерно вещественными параметрами. Для простоты давайте ограничимся тем, что допустим только два результата измерения состояния (Ааронсон обсуждает также более общий случай). Такое измерение задается двумя положительными полуопределенными операторами E и Id −E , и вероятность того, что измерение даст первый результат, равна Tr(Ep). Поскольку измерение с двумя результатами дает не более одного бита информации о ρ, результаты измерения Ω () необходимы для изучения σ, который очень близок к ρ на расстоянии трассировки или норме Фробениуса. Недавно было показано, что такого количества копий также достаточно [OW16, HHJ+16].  
Из-за экспоненциального масштабирования числа кубитов количество измерений, необходимых для томографии произвольного состояния, скажем, на 100 кубитах, уже непомерно велико. Однако Ааронсон показал интересный и удивительно эффективный PAC-подобный результат: из результатов измерений O(n), с измерениями, выбранными независимый и одинаково распределенный согласно неизвестному распределению D на

множество всех возможных измерений с двумя результатами, мы можем построить N-кубитовое квантовое состояние σ, которое имеет примерно такое же ожидаемое значение, как ρ для” большинства " измерений с двумя результатами. В последнем случае "большинство “снова измеряется под тем же D, который генерировал измерения, как и в обычной настройке PAC, где оценивается” приблизительная правильность" гипотезы учащегося

в том же дистрибутиве D, который создал примеры учащегося. Тогда выход государства σ могут быть использованы для прогнозирования поведения ρ на два-результат измерений, и это даст хороший

прогнозирование для большинства измерений. Соответственно, o(n), а не exp (n) результаты измерений

достаточно для “довольно хорошей томографии": приблизительно узнать состояние N-кубита, которое, возможно, нет

близко к ρ на расстоянии трассировки, но все же достаточно хорошо для большинства практических целей. Точнее,

Результат Ааронсона следующий.

Теорема 4.16 ([Aar07]). Для каждого δ, ε, γ > 0 существует обучаемый со следующим свойством: для каждого распределения D на множестве измерений с двумя результатами, учитывая T = n \* poly(1 / ε, 1 / γ, log(1 / δ)) результаты измерений (E1,b1), . . . , (ET, bT) где каждый Ei рисуется независимый и одинаково распределенный из D и bi бит с Pr[bi = 1] = Tr (Eiρ), с вероятностью ≥ 1 − δ обучаемый производит классическое описание состояния σ такое, что

[|Tr(Eσ) − Tr(Eρ)| > γ] ≤ ε.

Отметим, что “примерно правильный” мотивация оригинального Pac модель теперь определяется двумя параметрами ε и γ , а не только по одному параметру ε, как и раньше: выход состояния σ считается приблизительно правильным, если значение Tr(Eσ ) и аддитивной погрешностью не более γ (По сравнению с правильным значением Tr(Еρ)), только с вероятностью ε над выбором E. Мы потом хотите, чтобы на выходе было примерно правильно, за исключением с вероятностью δ, как и раньше. Заметим также, что теорема говорит только о сложности выборки обучаемого (т. е. о числе t результатов измерений, используемых для построения σ ), а не о сложности времени, что может быть довольно плохо в целом.

Эскиз доказательства. Доказательство вызывает общие результаты из-за Энтони и Бартлетта [AB00] и Бартлетта и Лонга [BL98] об изучении классов вероятностных функций8 в терминах их γ-fat-shattering размерности. Это обобщает размерность VC от булевых до вещественных функций следующим образом. Для некоторого множества E пусть C-класс функций f: E→[0,1].Мы говорим, что множества={E1,...,Ed}⊆E и γ- fat-shattering C, если существует α1,. . . , αd ∈ [0,1] такое, что для всех Z ⊆ [d] существует F ∈ C, удовлетворяющее:

1. Если i∈Z, то f (Ei)≥ αi+γ.

2. Если i<Z, то f (Ei)≤ αi−γ .

Размер γ-fat-shattering C-это размер наибольшего S, который разрушается C.9 для приложения к изучению квантовых состояний, пусть E-множество всех операторов измерения n-кубитов. Соответствующий класс вероятностных функций соответствует матрицам плотности n-кубитов:

C={f: ε →[0,1] | n-кубит ρ с учетом ∀E∈ ε,f(E)=Tr(Ep)}.

Предположим, что множество S = {E1,. . . , Ed} γ-fat-shattered C. Это означает, что для каждой строки z ∈ {0,1}d существует N-кубитовое состояние pz , из которого бит zi может быть восстановлен с помощью измерения Ei, с γ-преимуществом над просто выводом 1 с вероятностью ai . Такие кодировки Z → pz классических строк в квантовые состояния называются квантовыми кодами произвольного доступа. С использованием известных ограничений на такое коды [ANTV02], Aaronson показывает, что d = O(n/γ2). Эта верхняя граница на размерности γ-fat-shattering C затем может быть подключена к [AB00, BL98], чтобы получить теорему.

Совсем недавно, в аналогичном духе изучения квантовых объектов, Ченг и др. [CHY16] изучил, сколько состояний достаточно, чтобы узнать неизвестное квантовое измерение. Здесь ответ оказывается линейным в размерности пространства, поэтому экспоненциальным в количестве кубитов. Однако изучение неизвестного квантового состояния становится двойной проблемой для их вопроса и использования этого связь они могут по-другому порицать результаты Ааронсона [Aar07].

**5 Сложность**

Во многих отношениях лучшим показателем эффективного обучения является низкая временная сложность. При низкой пробы сложность является необходимым условием эффективного обучения, информационно-теоретической достаточности небольшая выборка не очень помогает на практике, если поиск хорошей гипотезы все еще занимает много времени. Десять В этом разделе мы опишем ряд результатов, где лучший квантовый ученик имеет гораздо меньше сложность времени, чем самый известный классический ученик.

5.1 Эффективное по времени квантовое обучение PAC

При попытке найти примеры квантовых ускорений для обучения имеет смысл начать с самого известного примера квантового ускорения, который у нас есть: алгоритм Шора для факторинга целых чисел в полиномиальное время [Sho97]. Широко предполагается, что классические компьютеры не могут эффективно множить целые числа Blum (т. е. целые числа, являющиеся произведением двух различных простых чисел равной длины бита, каждое из которых конгруэнтно 3 mod 4). До открытия Шора Кирнс и Валиант [KV94a] уже построили класс концептов C, основанный на факторинге, в качестве примера простого и эффективно представимого класса концептов с небольшим измерением VC, которое невозможно эффективно изучить. Грубо говоря , каждое понятие C ∈ C соответствует целому числу Blum N, и положительно помеченный пример для понятия показывает N . Краткое описание c, однако, зависит от факторизации N , которая, как предполагается, трудно вычисляется классическими компьютерами. Servedio и Gortler [SG04] заметили, что благодаря алгоритму Шора этот класс эффективно изучается квантовыми компьютерами. Они также заметили, что класс концептов на основе факторинга, разработанный Англуином и Харитоновым [AK95], чтобы показать-

Несс обучения даже с запросами членства, легко узнать с помощью квантовых компьютеров.

Несс обучения даже с запросами членства, легко узнать с помощью квантовых компьютеров.

Теорема 5.1 ([SG04]). Если нет эффективного классического алгоритма факторинга целых чисел Blum, то

1. существует концептуальный класс, который эффективно PAC изучается квантовыми компьютерами, но не классическими компьютерами;

2. существует класс concept, который эффективно точно изучается из запросов членства квантовыми компьютерами, но не классическими компьютерами.

Можно построить классические односторонние функции, основанные на предположении, что факторинг сложен.

Эти функции могут быть нарушены (т. е. эффективно инвертированы) с помощью квантовых компьютеров. Однако, есть и другие классические односторонние функции, которые мы не знаем, как порвать с квантовым компьютером. Удивительно, Servedio и Gortler [SG04] удалось построить концепцию классов с квантово-классическое разделение на основе любой классической односторонней функции—независимо от того, эта односторонняя функция может быть нарушена квантовым компьютером! Конструкция строит понятия, объединяя экземпляры задачи Саймона [Sim97] с семейством псевдослучайных функций, которые

можно получить из односторонней функции.

Теорема 5.2 ([SG04]). Если существуют классические односторонние функции, то существует концептуальный класс C, который эффективно точно изучается из запросов членства квантовыми компьютерами, но не классическими компьютерами.

5.2 Изучение DNF из однородных квантовых примеров

Как мы видели в разделе 3, Бшаути и Джексон [BJ99] представили модель обучения на квантовых примерах. Их главный положительный результат-показать, что дизъюнктивные формулы нормальной формы (DNF) можно изучать в полиномиальное время из квантовых примеров при равномерном распределении. Для изучения DNF при равномерном распределении из классических примеров лучшей верхней границей является квазиполиномиальное время [Ver90]. С добавленной силой запросов членства, где учащийся может активно запрашивать метку любого x по своему выбору, формулы DNF, как известно, могут быть изучены в полиномиальное время при равномерном D [Jac97], но полиномиальная обучаемость без запросов членства является давней открытой проблемой. Классический алгоритм полиномиального времени для обучения DNF с использованием запросов членства-алгоритм гармонического сита Джексона [Jac97]. Грубо говоря, он делает следующее. Во-первых, можно показать, что если целевое понятие c : {0, 1}n → {0, 1} является s-термом DNF (т. е. дизъюнкцией не более s конъюнкций переменных и отрицаемых переменных), то существует n-разрядная функция четности, которая согласуется с c на 1/ 2 + Ω(1/s) фракции 2n входов. Кроме того, алгоритм Голдрейха-Левина [GL89] может быть использован для эффективного поиска такой функции четности с помощью запросов членства. Это представляет собой “слабый ученик”: алгоритм поиска гипотезы, которая согласуется с целевой концепцией с вероятностью не менее 1/2 + 1/poly(s). Во-вторых, существуют общие методы, известные как “форсирование” [Fre95], которые могут превратить слабого ученика в "сильного" ученика, т. е. тот, который производит гипотезу, которая согласуется с целью с вероятностью 1-ε, а не с вероятностью 1/ 2 + 1/poly(s). Как правило, такие алгоритмы повышения допускают доступ к слабому ученику , который может создать слабую гипотезу при каждом возможном распределении D, а не только равномерное D . Идея в том, чтобы начать с

распределение D1 = D, и использовать слабого ученика, чтобы узнать слабую гипотезу h1 в отношении D1. Затем определите новое распределение D2, фокусируясь на входах, где предыдущая гипотеза потерпела неудачу; используйте слабого ученика для создания слабой гипотезы h2 в отношении D2 и так далее. После R = poly(s) таких шагов общая гипотеза h определяется как мажоритарная функция, применяемая к (h1,. . . , hr.)11 обратите внимание, что при обучении под фиксированной равномерной D мы можем непосредственно выбирать только первое распределение D1 = D. К счастью, если посмотреть на последующие распределения D2, D3,. . . , Dr производится путем форсирования в этом конкретном случае выборка этих распределений Di может быть эффективно "смоделирована" с использованием образцов из равномерного распределения. Объединение этих идей дает классический полиномиальный ученик для DNF при равномерном распределении, используя запросы членства.

Частью классического гармонического сита, использующего запросы членства, является алгоритм Голдрейха-Левина для нахождения четности (т.е., символьная функция χ S), что является слабой гипотезой. Ключом к квантовому ученику является наблюдение, что можно заменить Голдрайха-Левина выборкой Фурье из однородных квантовых примеров (см. раздел 2.2 .2). Пусть f = 1-2c, что равно c в ±1-нотации. Если χ S имеет корреляцию Ω (1/s) с целью, то b f^(S) = Ω(1/s) и выборка Фурье выводит S с вероятностью Ω (1/ s2). Следовательно, Poly(s) пробеги выборки Фурье с высокой вероятностью дадут нам слабую гипотезу. Поскольку состояние на Шаге 3 алгоритма выборки Фурье может быть получено с вероятностью 1/2 из однородного квантового примера, мы больше не требуем использования запросов членства. Описание этого алгоритма (и лежащей в его основе классической гармоники сито) в полной мере выходит за рамки этого обзора, но приведенный выше эскиз, надеюсь, дает основные идеи результата [BJ99].

Теорема 5.3 ([BJ99]). Концептуальный класс S-term DNF эффективно изучается PAC при равномерном распределении на квантовых примерах.

5.3 изучение линейных функций и Хунт на однородных квантовых примерах

Однородные квантовые примеры можно использовать и для изучения других вещей. Например, предположим, что f (x) = A · X mod 2 является линейной функцией над F2. Тогда спектр Фурье f, рассматриваемый как ±1-значная функция, имеет весь свой вес на xa . Следовательно, путем выборки Фурье мы можем отлично восстановить A с O(1) квантовой сложностью выборки и O(n) временной сложностью. Напротив, классические учащиеся нуждаются в Ω (n) примерах для изучения f по той простой причине , что каждый классический пример (и даже каждый запрос членства, если они доступны учащемуся) дает не более одного бита информации о целевой концепции.

Более сложным и интересным примером является изучение функций, которые зависят (возможно, нелинейно) не более чем от k из n входных битов, с k ≪ n. Такие функции называются k-juntas, так как они “управляются” небольшим подмножеством входных битов. Мы хотим узнать такое f до ошибки ε из равномерных (квантовых или классических) примеров. Тривиальный ученик будет пробовать o (2klog n) классических примеров, а затем пройти через все возможных наборов из k переменных, чтобы найти тот, который соответствует образцу. Это дает временную сложность O (nk ). Наиболее известная верхняя граница сложности по времени [MOS04] лишь немного лучше: O), где ω ∈ [2, 2.38] - оптимальный показатель для умножения матрицы.

Эффективное по времени обучение k-Хунт при равномерном распределении для k = O(log n) является печально известным узким местом в классической теории обучения, поскольку это частный случай обучения DNF: каждая k-хунта может быть записана как S-терм DNF с s < 2k , просто взяв или над 1-входами базовой K-битной функции. В частности, если мы хотим эффективно изучать poly(n)-term DNF из единообразных примеров (все еще открытая проблема, как упоминалось в предыдущем разделе), то мы должны, по крайней мере, быть в состоянии эффективно изучать O(log n)-juntas (также все еще открытые). Ученик DNF бшаути и Джексона из однородных квантовых примеров подразумевает , что мы можем изучать k-хунты, используя Поли(2k, n) квантовые примеры и время (для фиксированных ε, δ). Atıcı и Servedio [AS09] дали более точную верхнюю границу.

Теорема 5.4 ([AS09]). Существует квантовый алгоритм для обучения к-хунтас под равномерное распределение, которое использует O(k log(k)/ε) единого квантового примеры, O(2k) форма и классические образцы, и O(nk log(k)/ε + 2k log(1/ε)) время.

Эскиз доказательства. Идея состоит в том, чтобы сначала использовать выборку Фурье из квантовых примеров, чтобы найти K переменных (по крайней мере, те, которые имеют несущественное влияние), а затем использовать O(2k) равномерные классические примеры, чтобы узнать (почти все) таблицу истинности функции по этим переменным. Просмотр целевой k-хунты f как функции с диапазоном ±1. Пусть влияние переменной xi на f

Infi(f ) = =Ex [)2] =Prx[f(x)≠f(x⊕ei)],

где x⊕ei ,х после перевертывания i-й бит.Если Si=1 для i, который не находится в хунте, то b f^ (S)=0. Следовательно, выборка Фурье возвращает S такой, что Si = 1 только для переменных в хунте. Infi (f) - это именно вероятность того, что Si = 1. Следовательно , для фиксированного i вероятность того, что i не появится в T образцах Фурье, равна

(1−Infi(f))T ≤.

Если мы зададим T = O(K log (k)/ε) и пусть V-объединение опор образцов T Фурье, то с большой вероятностью V содержит все переменные хунты, кроме тех, у которых Infi (f ) ≪ ε/ k (последние можно игнорировать, так как даже их совместное влияние пренебрежимо мало). Теперь используйте o(2K log (1/ε)) равномерные классические примеры. С большой вероятностью появится по крайней мере 1 − ε/ 2 из всех 2|V | возможных настроек переменных в V, и мы используем их для формулирования нашей гипотезы h (скажем, со случайными значениями для нескольких входов таблицы истинности, которые мы не видели в нашей выборке, и для тех, которые появились дважды с непоследовательными значениями f). Можно показать, что,

с большой вероятностью h будет не согласен с f не более чем на ε-долю {0, 1}n.

В связанном результате Belovs [Bel15] дает очень жесткий анализ количества запросов квантового членства (хотя и не временной сложности), необходимых для точного изучения k-Хунт, чья базовая K-разрядная функция симметрична. Например, если K-разрядная функция является OR или большинством, то O (k1 / 4) запросов квантового членства. Для случая большинства Θ (k) классические запросы членства требуются, давая разделение в четвертой степени между квантовым и классическим запросом членства сложность точного обучения (см. следствие 4.4).

**6 Заключение**

Квантовая теория обучения изучает теоретические аспекты квантового машинного обучения. Мы изучили то, что известно об этой области. Конкретно

• Сложность запроса точного обучения. Количество запросов квантового членства, необходимых для точного изучения целевой концепции, может быть полиномиально меньше, чем количество классических запросов членства, но не намного меньше.

• Сложность выборки. Для независимых от распределения моделей PAC и агностического обучения квантовые примеры не дают существенного преимущества перед классическими случайными примерами: для каждого класса концептов классические и квантовые сложности выборки одинаковы до постоянных факторов. Напротив, для некоторых фиксированных распределений (например, равномерных) квантовые примеры могут быть намного лучше классических примеров.

• Временная сложность. Существуют концептуальные классы, которые могут быть изучены суперполиномиально быстрее квантовыми компьютерами, чем классическими компьютерами, например, на основе Шора или Саймона

алгоритм. Это справедливо как в модели точного обучения с запросами на членство, так и в модель PAC-обучения. Если допустить однородные квантовые примеры, DNF и хунты могут быть изучены гораздо эффективнее, чем мы знаем, как это сделать классически.

Заканчиваем ряд направлений для будущих исследований.

• Bshouty и Jackson [BJ99] показали, что DNF (то есть дизъюнкции конъюнкций переменных и отрицания переменных) можно эффективно изучать на однородных квантовых примерах. То же самое верно для схем глубины-3? А как насчет схем постоянной глубины с неограниченными вентиляторными и/или даже пороговыми воротами, то есть концептуальными классами AC0 и TC0-могут ли они быть эффективно изучены из единообразных квантовых примеров или даже PAC —изучаемыми? Этот

последний является одним из “десяти полукруглых вызовов для теории квантовых вычислений” Скотта Ааронсона [Aar05]. Классически лучшие верхние оценки временной сложности обучения AC0 являются квазиполиномиальными при равномерном распределении [LMN93] и примерно exp(n1/3) в модели PAC (т. е. при всех возможных распределениях) [KS04]; см. [DS16] для недавнего повторения твердости.

• Atıcı и Servedio [AS05] спросили, Может ли для каждого C верхняя граница в следствии 4.4 быть улучшена до D(C) ≤ O(nQ(C) + Q(C)2)?

• Можем ли мы охарактеризовать классическую и квантовую сложность запроса точного изучения класса концептов C в терминах комбинаторного параметра γ (C) или в терминах расширенной размерности обучения C?

• Можем ли мы найти больше примеров концептуальных классов, где квантовые примеры полезны при изучении со ссылкой на некоторое фиксированное распределение (равномерное или иное) или некоторый ограниченный набор распределений?

• Можем ли мы найти примеры квантового ускорения в англуинской [Ang87] модели запросов эквивалентности плюс запросы членства?

• Большинство исследований в квантовой теории обучения (и, следовательно, этот обзор) сосредоточены на концептуальных классах булевых функций. Как насчет изучения классов вещественных или даже векторных функций?

• Можем ли мы найти правильного ученика quantum PAC с оптимальной сложностью выборки, т.е., чья выходная гипотеза лежит в самом C? Или правильный эффективный квантовый ученик для DNF с использованием однородных квантовых примеров?

\* Можем ли мы найти практические проблемы машинного обучения с большим доказуемым квантовым ускорением?

\* Можем ли мы использовать квантовое машинное обучение для “квантового превосходства”, т.е., для решения какой-то задачи с использованием 50-100 кубитов убедительно быстрее, чем это возможно на больших классических компьютерах? (См., например, [AC16] для некоторых результатов сложности, касающихся квантового превосходства.)

Подтверждения.

Мы благодарим Лейна Хемаспаандру за заказ этого исследования для колонки теории сложности новостей SIGACT и полезных комментариев, а также Робина Котари за полезные комментарии и указатели (в том числе на [Aar05]) и за отправку нам его рукописи 2012 года [Kot12].

**Постановка задачи**

Состояние частицы задается двумя величинами: координатами (радиус-вектором) и [импульсом](http://ru.solverbook.com/spravochnik/mexanika/dinamika/zakon-soxraneniya-impulsa/). В рамках квантовой механики ставить вопрос о точном местоположении, [траектории](http://ru.solverbook.com/spravochnik/mexanika/kinematika/traektoriya/) частицы не корректно. Для квантовой частицы координаты и импульс могут быть неопределёнными. Поэтому ее состояние задается двумя вероятностными функциями:

W(x, y, z), V(px, py, pz)

Первая характеризует неопределённые координаты частицы, вторая — неопределённые импульсы. Вместо двух указанных функций W и V в квантовой механике вводится одна, комплексная функция, называемая волновой функцией. (Комплексная функция равносильна двум функциям, т.к. состоит из двух частей: действительной и мнимой.) Достоинством такого метода является в первую очередь то, что действительная и мнимая части волновой функции являются функциями не различных переменных (х и px) а переменных одного pода: либо только координат, либо только импульсов. Итак, состояние квантовой частицы можно характеризовать волновой функцией (комплексной), в двух представлениях — либо в координатном: ψ(x, y, z, t) , либо в импульсном:  Y(px, py, pz, t). [Уравнение движения](http://ru.solverbook.com/spravochnik/uravneniya-po-fizike/uravnenie-dvizheniya/) свободной частицы особенно просто выглядит в импульсном представлении, т.к. импульс свободной частицы сохраняется. Это означает на квантовом языке, что функция Y(px, py, pz) не зависит от времени.

Уравнение же связанной частицы, на которую действуют силы, удобнее получить в координатном представлении. Нужно сказать, что в квантовой механике, строго говоря, нельзя ввести понятие силы, как нельзя ввести понятие [скорости](http://ru.solverbook.com/spravochnik/mexanika/kinematika/skorost/). И это ясно, если вспомнить, что по определению сила есть [производная](http://ru.solverbook.com/spravochnik/proizvodnye/) от импульса частицы по времени. Импульс же квантовой частицы является неопределённым, и его невозможно продифференцировать по времени. Поэтому взаимодействие частиц в квантовой механике характеризуют не силой, а [потенциальной энергией](http://ru.solverbook.com/spravochnik/mexanika/dinamika/zakon-soxraneniya-energii/).

Движение связанной частицы [массы](http://ru.solverbook.com/spravochnik/mexanika/dinamika/massa-plotnost/) m будет задаваться уравнением следующего вида:

ih=-Δ+U(x, y, z, t)

где Δ= + + -оператор Лапласса, x,y,z – координаты,h- постоянная Планка, деленная на 2π.

Это уравнение называется временным уравнением Шредингера.Если U(x, y, z) не зависит от времени, то решение уравнения Шредингера можно представить как:

Ψ(x, y, z, t) = exp(-

где E-полная энергия квантовой системы, а удовлетворяет [стационарному уравнению Шредингера](http://ru.solverbook.com/spravochnik/uravneniya-po-fizike/stacionarnoe-uravnenie-shredingera/):

-

Уравнение Шредингера является основным уравнением движения частицы в квантовой механике. Оно не может быть выведено из других соотношений.

Его следует рассматривать как исходное основное предположение, справедливость которого подтверждается тем, что все следствия из него вытекающие, подтверждаются опытами.

**2 Анализ**

**2.1 Анализ предметной области**

## Решение уравнения Шредингера

С математической точки зрения — это [дифференциальное уравнение](http://ru.solverbook.com/spravochnik/differencialnye-uravneniya/) в [частных производных](http://ru.solverbook.com/spravochnik/proizvodnye/chastnye-proizvodnye/). Уравнение в частных производных имеет множество решений. В каждой конкретной задаче из этого множества следует выбрать одно решение, отвечающее условиям задачи.

С физической точки зрения нужно отметить, что согласно уравнению Шредингера волновая функция изменяется детерминировано, то есть совершенно однозначно. В этом смысле квантовая механика напоминает классическую, в которой движение системы заранее предопределено начальными условиями. Однако сама волновая функция имеет вероятностный смысл. Можно сказать, в квантовой механике детерминировано изменяются вероятности, а не сами физические события. События же всегда случайны и совершаются непредсказуемо.

Наконец, необходимо отметить еще одну очень важную особенность уравнения Шредингера: оно линейно. Волновая функция и ее производные входят в него в первой степени и для волновых функций справедлив принцип суперпозиции. Он в квантовой механике играет очень важную роль, так как позволяет сложные движения раскладывать на более простые движения. Например, движение свободной частицы выражается отнюдь не только волнами де-Бройля. Возможны более сложные выражения для результирующих волновых функций той же свободной частицы. Вместе с тем согласно принципу суперпозиции любое сложное движение свободной частицы можно представить как сумму волн де-Бройля.

Уравнение Шредингера является математическим выражением корпускулярно-волнового дуализма микрочастиц. В предельном случае, когда длины волн де Бройля значительно меньше размеров рассматриваемого движения уравнение Шредингера позволяет описывать движение частиц по законам классической [механики](http://ru.solverbook.com/spravochnik/mexanika/).

Тогда как с точки зрения математики уравнение Шредингера – это волновое уравнение, по структуре подобно уравнению колебания струны. Однако, решения уравнения Шредингера Ψ(x, y, z, t)  прямого физического смысла не имеют.

Физический смысл имеет модуль произведения | ψ (x, y, z, t)\*ψ\*(x, y, z, t)| =|Ψ(x, y, z, t)|2=w, w — определяется как плотность вероятности нахождения частицы в точке пространства, где ψ\*(x, y, z, t) -комплексно сопряженная функция с ψ (x, y, z, t).

W==

где W – вероятность нахождения частицы в объеме V.

Из вероятностного смысла волновой функции следует, что квантовая механика имеет статистический характер. С помощью волновой функции, которая является решением уравнения Шредингера нельзя точно описать траекторию движения квантовой частицы, можно лишь сказать какова вероятность обнаружить эту частицу в разных областях пространства.

**1 Искусственные нейронные сети**

**1.1 Биологические основы функционирования нейрона**

Тематика искусственных нейронных сетей относится к междисциплинарной сфере знаний, связанных с биокибернетикой, электроникой, прикладной математикой, статистикой, автоматикой и даже с медициной. Искусственные нейронные сети возникли на основе знаний о функционировании нервной системы живых существ. Они представляют собой попытку использования процессов, происходящих в нервных системах, для выработки новых технологических решений.

Нервная клетка, сокращенно называемая нейроном, является основным элементом нервной системы. Изучение механизмов функционирования отдельных нейронов и их взаимодействия принципиально важно для познания протекающих в нервной системе процессов поиска, передачи и обработки информации. С этой точки зрения представляется необходимым построить и изучить модель биологического нейрона.

Как и у любой другой клетки, у нейрона имеется тело со стандартным набором органелл, называемое сомой, внутри которого располагается ядро. Из сомы нейрона выходят многочисленные отростки, играющие ключевую роль в его взаимодействии с другими нервными клетками. Можно выделить два типа отростков: многочисленные тонкие, густо ветвящиеся дендриты и более толстый, расщепляющийся на конце аксон (рис 1.1)

Синапсы отличаются друг от друга размерами и возможностями концентрации нейромедиатора вблизи своей оболочки. По этой причине импульсы одинаковой величины, поступающие на входы нервной клетки через различные синапсы, могут возбуждать ее в разной степени. Мерой возбуждения клетки считается уровень поляризации ее мембраны, зависящий от суммарного количества нейромедиатора, выделенного на всех синапсах. Из сказанного следует, что каждому входу клетки можно сопоставить численные коэффициенты (веса), пропорциональные количеству нейромедиатора, однократно выделяемого на соответствующем синапсе. В математической модели нейрона входные сигналы должны умножаться на эти коэффициенты для того, чтобы корректно учитывать влияние каждого сигнала на состояние нервной клетки. Синапсические веса должны быть натуральными числами, принимающими как положительные, так и отрицательные значения.

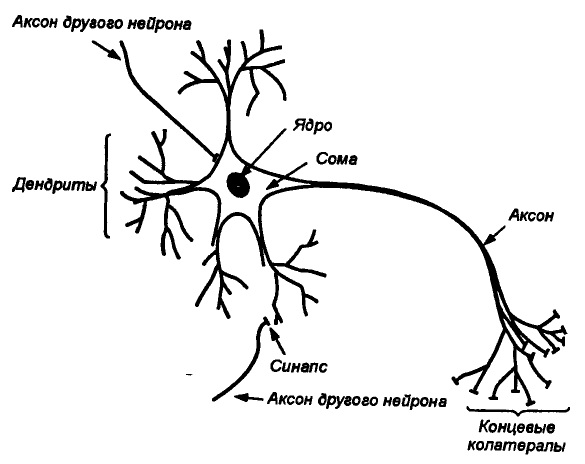


Рис 1.1 Упрощенная структура биологической нервной клетки

В первом случае синапс оказывает возбуждающие, а во втором – тормозящее действие, препятствующее возбуждению клетки другими сигналами. Таким образом, действие возбуждающего синапса может моделироваться положительным значением синапсического веса, а действие тормозящего синапса – отрицательным значением.

В результате поступления входных импульсов на конкретные синапсы и высвобождения соответствующих количеств нейромедиатора происходит определенное электрическое возбуждение нервной клетки. Если отклонение от состояния электрического равновесия невелико, либо если баланс возбуждений и торможений является отрицательным, клетка самостоятельно возвращается в исходное состояние, и на ее выходе какие-либо изменения не регистрируются. В этом случае считается, что уровень возбуждения клетки был ниже порога ее срабатывания. Если же сумма возбуждений и торможений превысила порог активации клетки, значение выходного сигнала начинает лавинообразно нарастать, принимая характерный вид нервного импульса (рис. 1.2), пересылаемого аксоном на другие нейроны, подключенные к данной клетке. Величина этого сигнала не зависит от степени превышения порога. Клетка действует по принципу «все или ничего». После выполнения своей функции нейромедиатор удаляется. Механизм удаления заключается либо во всасывании этой субстанции клеткой, либо в ее разложении, либо в удалении за пределы синапса. Одновременно с генерацией нервного импульса в клетке запускается процесс рефракции. Он проявляется как стремительное возрастание порога активации клетки до значения «плюс бесконечность», в результате чего сразу после генерации импульса нейрон теряет способность вырабатывать очередной сигнал даже при сильном возбуждении. Такое состояние сохраняется в течение времени Δtr , называемого периодом абсолютной рефракции. По окончании этого срока наступает период относительной рефракции Δtw, за который порог срабатывания возвращается к первоначальному значению. В это время клетку можно активировать, но только с приложением более сильных возбуждений. В естественных процессах, как правило, выполняется отношение Δtw>> Δtr.

Количество взаимодействующих друг с другом нервных клеток чрезвычайно велико. Считается, что человеческий мозг содержит около 1011 нейронов, каждый из которых выполняет относительно примитивные функции суммирования весовых коэффициентов входных сигналов и сравнения полученной суммы с пороговым значением. Каждый нейрон имеет свои веса и свое пороговое значение. Они определяются местонахождением нейрона и решаемой им задачи и могут интерпретироваться аналогично содержимому локальной памяти процессора. Громадное количество нейронов и межнейронных связей (до 1000 входов в каждый нейрон) приводит к тому, что ошибка в срабатывании отдельного нейрона остается незаметной в общей массе взаимодействующих клеток. Нейронная сеть проявляет высокую устойчивость к помехам – это «стабильная» сеть, в которой отдельные сбои не оказывают существенного влияния на результаты ее функционирования. Таково главное отличие нейронных систем от обычных электронных систем, созданных человеком. Следует подчеркнуть, что ни одна современная технология не позволяет построить искусственную нейронную сеть, близкую по масштабам к нейронной сети мозга.

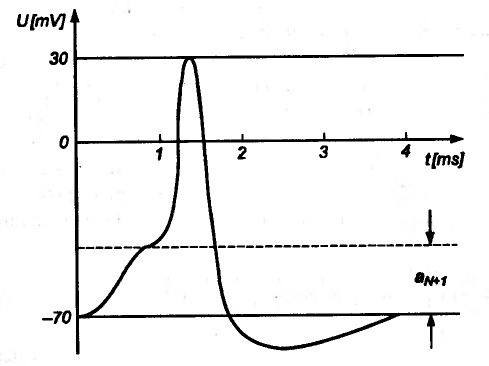


Рис. 1.2 Типичная форма нервного импульса

Однако изучение и копирование биологических нервных систем позволяют надеяться на создание нового поколения электронных устройств, имеющих аналогичные характеристики.

Другая важная особенность нервных систем – высокая скорость их функционирования, несмотря на относительно длительный цикл срабатывания каждой отдельной клетки, измеряемый в миллисекундах и показанный на рис. 1.2. Она достигается благодаря параллельной обработке информации в мозге огромным количеством нейронов, соединенных многочисленными межнейронными связями. Такие операции, как распознавание образов и звуков либо принятие решений, выполняется человеческим мозгом за промежутки времени, измеряемые миллисекундами. Достижение такого результата при использовании полупроводниковых технологий все еще выходит за границы современных технических возможностей, хотя цикл срабатывания каждой отдельных исполнительных элементов является достаточно коротким и имеет порядок 10-8 секунды. Если удастся, взяв за образец нервную систему, создать устройство с высокой степенью параллельности выполнения независимых операций, то скорость его функционирования может быть существенно увеличена и приближена к уровню, наблюдаемому в процессах обработки информации биологическими объектами.

**1.2 Первые модели нейронной сети**

Из приведенных выше рассуждений следует, что каждый нейрон можно считать своеобразным процессором: он суммирует с соответствующими весами сигналы, приходящие от других нейронов, выполняет нелинейную (например пороговую) решающую функцию и передает результирующее значение связанным с ним нейронам. В соответствии с действующим правилом «все или ничего» в простейших моделях нейронов выходной сигнал принимает двоичные значения: 0 или 1. Значение 1 соответствует превышению порога возбуждения нейрона, а значение 0 – возбуждению ниже порогового уровня.

В одной из первых моделей нейрона, называемой моделью МакКаллока-Питса ( предложенной в 1943 г.) нейрон считается бинарным элементом. Структурная схема этой модели представлена на рис 1.3

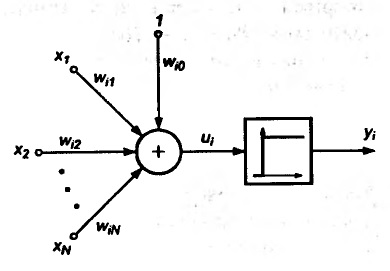


Рис. 1.3 Модель нейрона МакКаллока-Питса

Входные сигналы xj (j =1, 2, .., N) суммируются с учетом соответствующих

весов wij (сигнал поступает в направлении от узла i к узлу j) в сумматоре, после чего результат сравнивается с пороговым значением wi0. Выходной сигнал нейрона yi определяется при этом зависимостью

yi=f( (1.1)

Аргументом функции выступает суммарный сигнал ui=. Функция f(ui) называется функцией активации. В модели МакКаллока-Питса это пороговая функция вида

f(u)= (1.2)

Коэффициенты wij, присутствующие в формуле (1.1), представляют веса синоптических связей. Положительные значение wij соответствует возбуждающим синапсам, отрицательное значение wij тормозящим синапсам, тогда как wij=0 свидетельствует об отсутствии связи между i-м и j-м нейронами. Модель МакКаллока-Питса – это дискретная модель, в которой состояние нейрона в момент (t+1) рассчитывается по значениям его входных сигналов в предыдущий момент t. Построение дискретной модели обосновывается проявлением рефракции у биологических нейронов, приводящей к тому, что нейрон может изменять свое состояние с конечной частотой, причем длительность периодов бездействия зависит от частоты его срабатывания.

Через несколько лет Д. Хебб в процессе исследования ассоциативной памяти предложил теорию обучения (подбора весов wij) нейронов. При этом он использовал наблюдение, что веса межнейронных соединений при активации нейронов могут возрастать. В модели Хебба приращение веса Δwij в процессе обучения пропорционально произведению выходных сигналов yi и yj нейронов, связанных весом wij:

wij(k+1) = wij(k)+η\*yi(k)\*yj(k), (1.3)

где k означает номер цикла, а η – это коэффициент обучения.

В начале 60-х годов Б. Видроу предложил теоретическое обоснование и сформулировал принципы практической реализации адаптивных устройств обработки сигналов, что стало существенным вкладом в развитие нейронных сетей, функционирующих в режимах «онлайн» и «оффлайн».

В 1962 г. была опубликована книга Ф. Розенблатта, в которой представлена теория динамических нейронных систем для моделирования мозговой деятельности, основанная на персептронной модели нервной клетки. В этой теории использовались представление нейрона моделью МакКаллока-Питса, в которой функция активации принимала двоичные значения 0 и 1.

Ограниченные возможности одиночного персептрона и составляемых из таких элементов одноуровневых сетей подверглись критике в книге М. Минковского и С. Пейперта, что вызвало резкое снижение финансирования этой сферы научных исследований и привело в результате к замедлению развития искусственных нейронных сетей. Только отдельные

научные группы, сконцентрированные вокруг таких ученых, как Гроссберг, Видроу, фон дер Мальсбург, Амари, Фукушима и Кохонен, продолжали работу в этой области.

**1.3. Структура нейронной сети**

Многослойная сеть состоит из нейронов, расположенных на разных уровнях, причем помимо входного и выходного слоев, имеется еще, как минимум, один внутренний, т.е. скрытый слой. Как уже отмечалось в литературе, посвященной проблематике нейронных сетей, такая нейронная система называется многослойным персептроном. На рис. 1.4 представлена сеть с одним скрытым слоем. Все последующие рассуждения относятся к сетям именно такого типа. Обозначения сигналов и весов также будут соответствовать этому рисунку. Веса нейронов скрытого слоя пометим верхним индексом (1), а выходного слоя верхним индексом (2). Выходные сигналы нейронов скрытого слоя обозначим vj (j=1, 2, .., K), а выходного слоя – yj (j=1, 2, .., M). Примем, что функция активации нейронов задана в сигмоидальной униполярной или биполярной форме. Для упрощения описания будем использовать расширенное обозначение входного вектора сети в виде x=[x0, x1, .., xN]T, где x0=1 соответствует единичному сигналу поляризации. С вектором x связаны два входных вектора сети: вектор фактических выходных сигналов y=[y0, y1, .., yM]T и вектор ожидаемых выходных сигналов d=[d0, d1, .., dM]T. Цель обучения состоит в подборе таких значений весов wij(1) и wij(2) для всех сигналов сети, чтобы при заданном входном векторе x получить на выходе значения сигналов yi,которые с требуемой точностью будут совпадать с ожидаемыми значениями di для i = 1, 2, .., M. Если рассматривать единичный поляризационный сигнал как один из компонентов входного вектора x, то веса поляризации можно добавить в векторы весов соответствующих нейронов обоих слоев.

При таком подходе выходной сигнал i-го нейрона скрытого слоя удается описать функцией

vj=f(, (1.4)

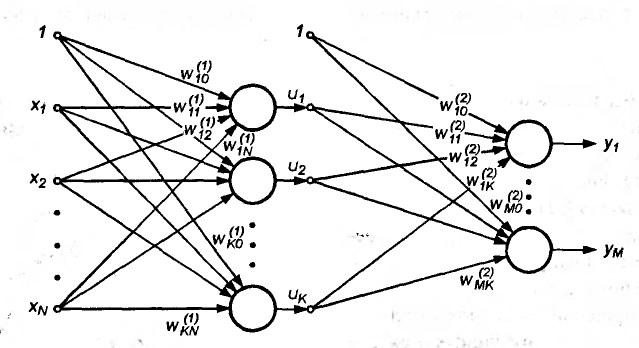


Рис. 1.4 Обобщенная структура двухслойной сигмоидальной нейронной сети

в которой индекс 0 соответствует сигналу и весам поляризации, причем v0=1,x0=1. В выходном слое k-й нейрон вырабатывает выходной сигнал,

определяемый как

yk=f(=f(. (1.5)

Из формулы (1.5) следует, что на значение выходного сигнала влияют веса обоих слоев, тогда как сигналы, вырабатываемые в скрытом слое, не зависят от весов выходного слоя.

**2. Квантовые системы**

**2.1 Кубит**

Исходным понятием является кубит (quantum bit) — элементарный носитель информации. В качестве физической реализации в принципе может выступать любой квантовый объект, имеющий два базовых состояния, которые обозначаются |0⟩ и |1⟩. На роль кубита, например, годится фотон с одной из двух перпендикулярных поляризаций или электрон с одним из двух противоположных направлений спина. С математической точки зрения состояния являются векторами, которые можно умножать на комплексные числа, а также складывать между собой. Таким образом, помимо базовых состояний |0⟩ и |1⟩, которые аналогичны 0 и 1 в обычном бите, кубит может пребывать в квантовом состоянии

|x⟩=c0\*|0⟩+c1\*|1⟩ (2.1)

где c0,c1 — любые комплексные числа (в частности действительные). При этом физическое состояние кубита не изменится, если коэффициенты c0,c1 умножить на одно и то же число a≠0. Поэтому вектор |x⟩ можно нормировать, т.е., выбрать множитель a∈C так, чтобы новые коэффициенты cj′=acj удовлетворяли условию |c0′|2+|c1′|2=1. Тогда вектор |x′⟩=c0′⋅|0⟩+c1′⋅|1⟩ называется нормированным или единичным.

Физический смысл состояния (2.1), которое называется суперпозицией базовых состояний, заключается в следующем. Если вектор |x⟩ суть единичный, то числа |c0|2 и |c1|2 дают вероятности того, что при измерении состояния кубита будет получено |0⟩ и |1⟩ соответственно. После измерения кубит останется в том базовом состоянии, которое оказалось измеренным. Вывести из него может только внешнее воздействие. Таким образом можно сказать, что кубит в нормированном состоянии (2.1) с вероятностью |c0|2 равен 0 и с вероятностью |c1|2 равен 1. Ничего подобного с обычным (классическим) битом происходить не может. Суперпозиция — существенно квантовый эффект! Термин «базовые» применительно к состояниям |0⟩ и |1⟩ означает, что любое другое состояние кубита может быть выражено их суперпозицией в смысле (2.1) для некоторых чисел c0,c1 (определенных с точностью до пропорциональности).

Рабочий регистр квантового компьютера мыслится набором из n кубитов, которые каким-то образом взаимосвязаны между собой — *запутаны*. Для того, чтобы реализовать его грандиозные возможности, число n должно быть достаточно большим, скажем n>100. Пусть каждый кубит номер j в регистре находится в своем состоянии |xj⟩, где xj∈{0,1}. Если рассматривать набор из n кубитов, как квантовый объект, то его состояние можно описать набором векторов |x1⟩|x2⟩...|xn⟩, который кратко обозначается |x1x2...xn⟩. Часто используется термин «тензорное произведение» и обозначения вроде |x1⟩⊗...⊗|xn⟩, способные смутить многих читателей статей о квантовых компьютерах. Им можно посоветовать просто игнорировать значок ⊗, полагая

|x1⟩⊗|x2⟩⊗...⊗|xn⟩=|x1x2...xn⟩ (2.2)

Пока нет никакой запутанности — просто набор независимых кубитов, хотя и считающихся единым объектом. Запутанность появится, если мы введем в рассмотрение суперпозиции состояний (2.2), т.е., векторы (точнее тензоры) состояний регистра, имеющие вид

(2.3)

где cj — комплексные числа, |xkj⟩ — вектор состояния k — го кубита, xkj∈{0,1}. Множество всевозможных векторов вида (2.3) называется тензорным произведением n пространств состояний одиночных кубитов, хотя без слова «тензор» вполне можно обходиться (в фундаментальной книге Дирака «Принципы квантовой механики» оно ни разу не встречается).

Само собой разумеется, что интерпретации суперпозиции состояний, как и вероятностной природы наблюдений над квантовыми системами, уделялось много внимания. Одной из схем, полезных при рассмотрении кубитов, является следующее геометрическое представление. Поскольку |α|2+|β|2=1, можно переписать формулу (2.4) в виде

|ψ⟩ = α|0⟩+β|1⟩ (2.4)

|ψ⟩=eiγ(cos(θ/2)|0⟩+eiϕsin(θ/2)|1⟩) (2.5)

где θ,ϕ и γ – действительные числа. Множитель eiγ можно игнорировать,поскольку он не приводит к наблюдаемым эффектам, и по этой причине формула (2.5) фактически сводится к

|ψ⟩=cos(θ/2)|0⟩+eiϕsin(θ/2)|1⟩ (2.6)

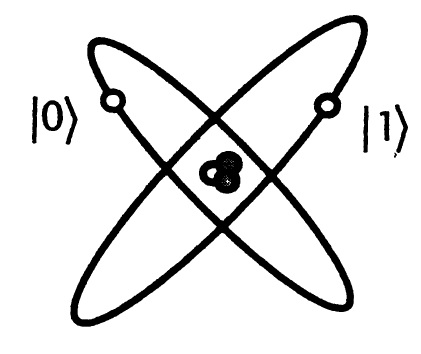


Рис. 2.1 Кубит, представленный двумя электронными уровнями в атоме.

Числа θ и ϕ задают точку на единичной трехмерной сфере, как показано на рис. 2.2. Эта сфера часто называется сферой Блоха; она позволяет наглядно представлять состояние одиночного кубита и часто служит в качестве превосходного «испытательного стенда» для идей из области квантовых вычислений и квантовой информации. Многие операции над одиночными кубитами рассматриваемые далее в этой главе, изящно описываются с использованием сферы Блоха. Однако, нужно иметь в виду, что возможности этого представления ограничены, так как не известно простого обобщения сферы Блоха на случай нескольких кубитов.

Сколько информации представляет кубит? Как ни парадоксально, на единичной сфере бесконечное количество точек, поэтому в принципе можно хранить все тексты Шекспира в бесконечном двоичном разложении θ.

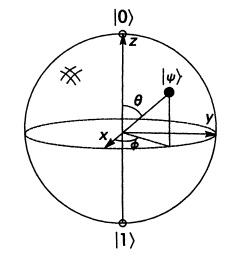


Рис 2.2 Представление кубита при помощи сферы Блоха

**2.2 Квантовая запутанность**

По определению состояние (2.3) является запутанным, если этот вектор *нельзя* разложить в произведение |A1⟩|A2⟩...|An⟩ векторов состояний одиночных кубитов. В этом случае воздействие на любой из кубитов может отражаться на состояниях каких-то других кубитов регистра. Заметим, что каждый вектор |Aj⟩, вообще говоря, является суперпозицией базовых, так что |Aj⟩=cj0|0⟩+cj1|1⟩для некоторых чисел cj0,cj1.   
  
Для иллюстрации рассмотрим случай двух кубитов. Их общее состояние |01⟩ не является запутанным, т.к. |01⟩=|0⟩|1⟩. Измерив, скажем, второй кубит мы найдем его в состоянии |1⟩. При этом первый останется в то же состоянии |0⟩, т.е., измерение второго на него не повлияло. Пусть теперь пара кубитов находится в состоянии |01⟩+|10⟩. Оно является запутанным, т.к. этот вектор нельзя представить в виде произведения |A1⟩|A2⟩ (легко проверить).   
  
При измерении второго кубита мы с равной вероятностью 0.5 найдем его в состоянии |0⟩ или |1⟩. Если второй кубит обнаружен в состоянии |0⟩, то это означает, что запутанная пара оказалась в |10⟩. Соответственно, первый кубит автоматически попал в состояние |1⟩. Если же второй кубит измерен в состоянии |1⟩, то пара оказалась в |01⟩. Следовательно, первый кубит оказался в состоянии |0⟩ в тот момент, как мы измерили второй. Таким образом, измерение состояния одного из двух запутанных кубитов мгновенно влияет на состояние второго. При этом исходное, общее состояние пары кубитов разрушается, что драматически называют коллапсом волновой функции (термин «волновая функция» можно считать синонимом для «вектор состояния», хотя между ними все же есть формальное различие).   
  
Примером запутанных кубитов являются электроны одного атома или одной орбитали, рассматриваемые в спиновых состояниях. Принцип Паули запрещает двум электронам иметь общий энергетический уровень, орбитальный и спиновой момент. Предположим, что для одного электрона удалось измерить спин, а до этого он был в суперпозиции спиновых состояний. Тогда второй электрон на той же орбитали сразу приобретает спин ему обратный, хотя до этого он также был в суперпозиции. Даже если при измерении первого электрона второй не был затронут!

Предположим, что у нас есть два кубита. Будь это классические биты, для них существовало бы четыре возможных состояния: 00, 01, 10 и 11. Подобно этому, система из двух кубитов имеет четыре состояния вычислительного базиса, обозначиваемых как |00⟩, |01⟩ , |10⟩  и |11⟩ . Пара кубитов также может находится в суперпозициях этих четырех состояний, поэтому для описания квантового состояния такой системы требуется сопоставить каждому состоянию вычислительного базиса комплексный коэффициент, иногда называемый амплитудой. В итоге вектор состояния, описывающий два кубита, имеет вид

|ψ⟩ =α00|00⟩ +α01|01⟩ +α10|10⟩ +α11|11⟩  (2.7)

Подобно случаю одиночного кубита, результат измерения x=(00, 01, 10 или 11) встречается с вероятностью |αx|2 и после измерения кубиты остаются в состоянии |x⟩. Требование, чтобы сумма вероятностей равнялась единице, выражается условием нормировки где «[0,1]2» обозначает множество строк из двух символов, где каждый символ является либо нулем, либо единицей. Для системы двух кубитов мы могли бы измерять только подмножество кубитов, скажем первый кубит. При измерении только первого кубита получается 0 с вероятностью |α00|2+| α01|2, а система переходит в состояние

|ψ'⟩ =. (2.8)  
Обратите внимание, что состояние после измерения перенормировано на коэффициент , поэтому оно по прежнему удовлетворяет условию нормировки, как и следует ожидать для допустимого квантового состояния.

Важным частным случаем состояния двух кубитов является состояние Белла, или ЭПР-пара,

(2.9)

Это безобидное на первый взгляд состояние ответственно за многие сюрпризы в области квантовых вычислений и квантовой информации. Оно играет ключевую роль в квантовой телепортации и сверхплотном кодировании и является прототипом многих других интересных квантовых состояний. Для состояния Белла характерно то, что при измерении первого кубита возможны два результата: 0 с вероятностью ½ и конечным состоянием |ϕ’, и 1 с вероятностью ½ и конечным состоянием |ϕ’|11 . Как следствие, измерение второго кубита всегда дает тот же результат, что и измерение первого кубита, т.е. данные измерений оказываются коррелированными. Над состоянием Белла можно выполнять измерения и других типов, применяя сначала некоторые операции к первому или второму кубиту, и эта любопытная корреляция между результатами измерения первого и второго кубитов по-прежнему будет существовать. Эти корреляции вызывают большой интерес с момента появления известной работы Эйнштейна, Подольского и Розена(ЭПР), в которой было впервые указано на странные свойства состояний наподобие белловского. Идеи ЭПР были подхвачены и значительно развиты Джоном Беллом, доказавшим потрясающий факт: корреляция измерений в состоянии Белла сильнее любой корреляции, которая может существовать между какими-либо классическими системами. Эти результаты были первым указанием на то, что квантовая механика позволяет обрабатывать информацию принципиально иначе, чем в классическом мире.

Литература

1. Осовский С. «Нейронные сети для обработки информации» (2004 г.)
2. И. Чанг, М. Нильсен «Квантовые вычисления и квантовая информация»(2006 г.)
3. Статья «Не все так просто с квантовым компьютером» 4 февраля 2017 <https://habr.com/ru/post/401315/>
4. Статья Уравнение Шредингера <http://ru.solverbook.com/spravochnik/uravneniya-po-fizike/uravnenie-shredingera/>
5. [AAD+15] J. Adcock, E. Allen, M. Day, S. Frick, J. Hinchliff, M. Johnson, S. Morley-Short, S. Pallister, A. Price, and S. Stanisic. Advances in quantum machine learning, 9 Dec 2015. arXiv:1512.02900. 2
6. [Aar05] S. Aaronson.Ten semi-grand challenges for quantum computing theory. http://www.scottaaronson.com/writings/qchallenge.html , 2005. 20,21
7. [Aar07] S. Aaronson. The learnability of quantum states. Proceedings of the Royal Society of London, 463(2088), 2007. quant-ph/0608142. 15,16
8. [Aar15] S. Aaronson. Quantum machine learning algorithms: Read the fine print. Nature Physics, 11(4):291–293, April 2015. 2
9. [AB00] M. Anthony and P. Bartlett. Function learning from interpolation. Combinatorics, Probability, and Computing, 9(3):213–225, 2000. Earlier version in EuroCOLT’95. 16
10. [AB09] M. Anthony and P. L . Bartlett. Neural network learning: Theoretical foundations. Cambridge University Press, 2009. 1
11. [ABG06] E. Äımeur, G. Brassard, and S. Gambs. Machine learning in a quantum world. In Proceedings of Advances in Artificial Intelligence, 19th Conference of the Canadian Society for Computational Studies of Intelligence, volume 4013 of Lecture Notes in Artificial Intelligence, pages 431–442, 2006. 2
12. [ABG13] E. Äımeur, G. Brassard, and S. Gambs. Quantum speed-up for unsupervised learning. Machine Learning, 90(2):261–287, 2013. 2 21
13. [AC16] S. Aaronson and L. Chen. Complexity-theoretic foundations of quantum supremacy experiments. arxiv:1612.05903, 2016. 21
14. [AIK+04] A. Ambainis, K. Iwama, A. Kawachi, H. Masuda, R. H. Putra, and S. Yamashita. Quantum identification of Boolean oracles. In Proceedings of 30th Annual Symposium on Theoretical Aspects of Computer Science (STACS’04), pages 105–116, 2004. arXiv:quant-ph/0403056. 10
15. [AIK+07] A. Ambainis, K. Iwama, A. Kawachi, R. Raymond, and S. Yamashita. Improved algorithms for quantum identification of Boolean oracles. Theoretical Computer Science, 378(1):41–53, 2007. 10
16. [AIN+09] A. Ambainis, K. Iwama, M. Nakanishi, H. Nishimura, R. Raymond, S. Tani, and S. Yamashita. Average/worst-case gap of quantum query complexities by on-set size. 2009. Preprint at arXiv:0908.2468v1. 10
17. [AK95] D. Angluin and M. Kharitonov. When won’t membership queries help? Journal of Computer and System Sciences, 50(2):336–355, 1995. Earlier version in STOC’91. 17
18. [Amb02] A. Ambainis. Quantum lower bounds by quantum arguments. Journal of Computer and System Sciences, 64(4):750–767, 2002. Earlier version in STOC’00. quant-ph/0002066. 9
19. [Ang87] D. Angluin. Queries and concept learning. Machine Learning, 2(4):319–342, 1987. 6,20
20. [ANTV02] A. Ambainis, A. Nayak, A. Ta-Shma, and U. V. Vazirani. Dense quantum coding and quantum finite automata. Journal of the ACM, 49(4):496–511, 2002. Earlier version in STOC’99. 16
21. [AS05] A. Atıcı and R. Servedio. Improved bounds on quantum learning algorithms. Quantum Information Processing, 4(5):355–386, 2005. quant-ph/0411140. 1,12,13,20
22. [AS09] A. Atıcı and R. Servedio. Quantum algorithms for learning and testing juntas. Quantum Information Processing, 6(5):323–348, 2009. arXiv:0707.3479. 1,19
23. [AW17] S. Arunachalam and R. de Wolf. Optimal quantum sample complexity of learning algorithms. 2017. To appear in Proceedings of CCC’17. arXiv:1607.00932. 2,5,8,13,14
24. [BBBV97] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. SIAM Journal on Computing, 26(5):1510–1523, 1997. quant- ph/9701001. 4
25. [BCG+96] N. H. Bshouty, R. Cleve, R. Gavald `a, S. Kannan, and C. Tamon. Oracles and queries that are sufficient for exact learning. Journal of Computer and System Sciences, 52(3):421– 433, 1996. Earlier version in COLT’94. 9
26. [BEHW89] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. Journal of the ACM, 36(4):929–965, 1989. 12,13
27. [Bel15] A. Belovs. Quantum algorithms for learning symmetric juntas via the adversary bound. Computational Complexity, 24(2):255–293, 2015. Earlier version in Complexity’14. arXiv:1311.6777. 10,19 22
28. [Ben82] P. A . Benioff. Quantum mechanical Hamiltonian models of Turing machines. Journal of Statistical Physics, 29(3):515–546, 1982. 1
29. [BHMT02] G. Brassard, P. Høyer, M. Mosca, and A. Tapp. Quantum amplitude amplification and estimation. In Quantum Computation and Quantum Information: A Millennium Volume, volume 305 of AMS Contemporary Mathematics Series, pages 53–74. 2002. quantph/0005055. 2,4
30. [BJ99] N. H. Bshouty and J. C. Jackson. Learning DNF over the uniform distribution using a quantum example oracle. SIAM Journal on Computing, 28(3):1136–1153, 1999. Earlier version in COLT’95. 1,2,7,17,18,20
31. [BK02] H. Barnum and E. Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. Journal of Mathematical Physics, 43:2097–2106, 2002. quant- ph/0004088. 5
32. [BL98] P. Bartlett and P. M. Long. Prediction, learning, uniform convergence, and scale-sensitive dimensions. Journal of Computer and System Sciences, 56(2):174–190, 1998. 16
33. [BV97] E. Bernstein and U. Vazirani. Quantum complexity theory. SIAM Journal on Computing, 26(5):1411–1473, 1997. Earlier version in STOC’93. 5
34. [BWP+16] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. Quantum machine learning, 28 Nov 2016. arXiv:1611.09347. 2
35. [CHY16] H. C. Cheng, M. H. Hsieh, and P. C. Yeh. The learnability of unknown quantum measurements. Quantum Information and Computation, 16(7&8):615–656, 2016. arXiv:1501.00559 . 16
36. [Deu85] D. Deutsch. Quantum theory, the Church-Turing principle, and the universal quantum Turing machine. In Proceedings of the Royal Society of London, volume A400, pages 97– 117, 1985. 1
37. [DS16] A. Daniely and S. Shalev-Shwartz. Complexity theoretic limitations on learning DNF’s. In Proceedings of the 29th Conference on Learning Theory (COLT’16), 2016. 20
38. [Fey82] R. Feynman. Simulating physics with computers. International Journal of Theoretical Physics, 21(6/7):467–488, 1982. 1
39. [Fey85] R. Feynman. Quantum mechanical computers. Optics News, 11:11–20, 1985. 1
40. [Fre95] Y. Freund. Boosting a weak learning algorithm by majority. Information and Computation, 121(2):256–285, 1995. Earlier version in COLT’90. 18
41. [Gav12] D. Gavinsky. Quantum predictive learning and communication complexity with single input. Quantum Information and Computation, 12(7-8):575–588, 2012. Earlier version in COLT’10. arXiv:0812.3429. 15
42. [GL89] O. Goldreich and L. Levin. A hard-core predicate for all one-way functions. In Proceedings of 21st ACM STOC, pages 25–32, 1989. 18 23
43. [Gro96] L. K. Grover. A fast quantum mechanical algorithm for database search. In Proceedings of 28th ACM STOC, pages 212–219, 1996. quant-ph/9605043. 2,4
44. [Han16] S. Hanneke. The optimal sample complexity of PAC learning. Journal of Machine Learning Research, 17(38):1–15, 2016. arXiv:1507.00473. 12,13
45. [Hau92] D. Haussler. Decision theoretic generalizations of the PAC model for neural net and other learning applications. Information and Computation, 100(1):78–150, 1992. 6
46. [Heg95] T. Heged ̋us. Generalized teaching dimensions and the query complexity of learning. In Proceedings of the 8th Conference on Learning Theory (COLT’95), pages 108–117, 1995.11,12
47. [HHJ+ 16] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yi . Sample-optimal tomography of quantum states. In Proceedings of 48th ACM STOC, pages 913–925, 2016. arXiv:1508.01797. 15
48. [HHL09] A. Harrow, A. Hassidim, and S. Lloyd. Quantum algorithm for solving linear systems of equations. Physical Review Letters, 103(15):150502, 2009. arXiv:0811.3171. 2
49. [HMP+10] M. Hunziker, D. A . Meyer, J. Park, J. Pommersheim, and M. Rothstein. The geometry of quantum learning. Quantum Information Processing, 9(3):321–341, 2010. quant-ph/0309059. 12
50. [Hol73] A. S . Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. Problemy Peredachi Informatsii, 9(3):3–11, 1973. English translation in Problems of Information Transmission, 9:177–183, 1973. 10
51. [Jac97] J. C. Jackson. An efficient membership-query algorithm for learning DNF with respect to the uniform distribution. Journal of Computer and System Sciences, 55(3):414–440, 1997. Earlier version in FOCS’94. 18
52. [Kot12] R. Kothari. Quantum computing and learning theory. Unpublished manuscript, 2012. 2,21
53. [Kot14] R. Kothari. An optimal quantum algorithm for the oracle identification problem. In 31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), pages 482–493, 2014. arXiv:1311.7685. 4,10,12
54. [KP17] I. Kerenidis and A. Prakash. Quantum recommendation systems. In Innovations in Theoretical Computer Science (ITCS’17), 2017. arXiv:1603.08675. 2
55. [KS04] A. Klivans and R. Servedio. Learning DNF in time Journal of Computer and System Sciences, 68(2):303–318, 2004. Earlier version in STOC’01. 20
56. [KSS94] M. J. Kearns, R. E . Schapire, and L. Sellie. Toward efficient agnostic learning. Machine Learning, 17(2-3):115–141, 1994. Earlier version in COLT’92. 6
57. [KV94a] M. J. Kearns and L. G . Valiant. Cryptographic limitations on learning Boolean formulae and finite automata. Journal of the ACM, 41(1):67–95, 1994. 17
58. [KV94b] M. J. Kearns and U. V. Vazirani. An introduction to computational learning theory. MIT Press, 1994. 1 24
59. [LL16] C. Y. -Y. Lin and H. Lin. Upper bounds on quantum query complexity inspired by the Elitzur-Vaidman bomb tester. Theory of Computing, 12:1–35, 2016. Earlier version in CCC’15. arXiv:1410.0932. 4
60. [LMN93] N. Linial, Y. Mansour, and N. Nisan. Constant depth circuits, Fourier transform, and learnability. Journal of the ACM, 40(3):607–620, 1993. Earlier version in FOCS’89. 20
61. [LMR13a] S. Lloyd, M. Mohseni, and P. Rebentrost. Quantum algorithms for supervised and unsupervised machine learning, 1 Jul 2013.arXiv:1307.0411. 2
62. [LMR13b] S. Lloyd, M. Mohseni, and P. Rebentrost. Quantum principal component analysis. Nature Physics, 10(631–633), 2013. arXiv:1307.0401. 2
63. [Man80] Y. Manin. Vychislimoe i nevychislimoe (computable and noncomputable). Soviet Radio, pages 13–15, 1980. In Russian. 1
64. [Man99] Y. Manin. Classical computing, quantum computing, and Shor’s factoring algorithm. quant-ph/9903008, 2 Mar 1999. 1
65. [Mon07] A. Montanaro. On the distinguishability of random quantum states. Communications in Mathematical Physics, 273(3):619–636, 2007. quant-ph/0607011. 5
66. [Mos83] M. Yu. Moshkov. Conditional tests. Problemy Kibernetzkt, 40:131–170, 1983. In Russian. 12
67. [MOS04] E. Mossel, R. O’Donnell, and R. Servedio. Learning functions of k relevant variables. Journal of Computer and System Sciences, 69(3):421–434, 2004. Earlier version in STOC’03. 19
68. [NC00] M. A. Nielsen and I. L . Chuang. Quantum Computation and Quantum Information. Cambridge University Press, 2000. 3
69. [O’D14] R. O’Donnell. Analysis of Boolean Functions. Cambridge University Press, 2014. 5
70. [OW16] R. O’Donnell and J. Wright. Efficient quantum tomography. In Proceedings of 48th ACM STOC, pages 899–912, 2016. arXiv:1508.01907. 15
71. [RML13] P. Rebentrost, M. Mohseni, and S. Lloyd. Quantum support vector machine for big data classification. Physical Review Letters, 113(13), 2013. arXiv:1307.0471. 2
72. [SB14] S. Shalev-Shwartz and S. Ben-David. Understanding machine learning: From theory to algorithms. Cambridge University Press, 2014. 1,14
73. [SG04] R. Servedio and S. Gortler. Equivalences and separations between quantum and classical learnability. SIAM Journal on Computing,33(5):1067–1092, 2004. Combines earlier papers from ICALP’01 and CCC’01. quant-ph/0007036. 1,9,10,12,17
74. [Sho97] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM Journal on Computing, 26(5):1484–1509, 1997. Earlier version in FOCS’94. quant-ph/9508027. 1,17
75. 25
76. [Sim96] H. U. Simon. General bounds on the number of examples needed for learning probabilistic concepts. Journal of Computer and System Sciences, 52(2):239–254, 1996. Earlier version in COLT’93. 14
77. [Sim97] D. Simon. On the power of quantum computation. SIAM Journal on Computing, 26(5):1474–1483, 1997. Earlier version in FOCS’94. 17
78. [Sim15] H. U. Simon. An almost optimal PAC algorithm. In Proceedings of the 28th Conference on Learning Theory (COLT), pages 1552–1563, 2015. 12
79. [SSP15] M. Schuld, I. Sinayskiy, and F. Petruccione. An introduction to quantum machine learning. Contemporary Physics, 56(2):172–185, 2015. arXiv:1409.3097. 2
80. [Tal94] M. Talagrand. Sharper bounds for Gaussian and empirical processes. The Annals of Probability, pages 28–76, 1994. 14
81. [Val84] L. Valiant. A theory of the learnable. Communications of the ACM, 27(11):1134–1142, 1984. 1,6
82. [VC71] V. Vapnik and A. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. Theory of Probability & Its Applications, 16(2):264–280,
83. 1971. English translation of 1968 Russian paper in Dokl. Akad. Nauk. 181(4). 12
84. [VC74] V. Vapnik and A. Chervonenkis. Theory of pattern recognition. Nauka, USSR, 1974. In Russian. 14
85. [Ver90] K. A . Verbeurgt. Learning DNF under the uniform distribution in quasi-polynomial time. In Proceedings of the 3rd Annual Workshop on Computational Learning Theory (COLT’90), pages 314–326, 1990. 17
86. [Wit14] P. Wittek. Quantum Machine Learning: What Quantum Computing Means to Data Mining. Elsevier, 2014. 2
87. [WKS16a] N. Wiebe, A. Kapoor, and K. Svore. Quantum deep learning. Quantum Information and Computation, 16(7):541–587, 2016. arXiv:1412.3489 . 2
88. [WKS16b] N. Wiebe, A. Kapoor, and K. M . Svore. Quantum perceptron models, 2016. Preprint at arXiv:1602.04799. 2
89. [Wol08] R. de Wolf. A brief introduction to Fourier analysis on the Boolean cube. Theory of Computing, 2008. ToC Library, Graduate Surveys 1. 5
90. [Zha10] C. Zhang. An improved lower bound on query complexity for quantum PAC learning. Information Processing Letters, 111(1):40–45, 2010. 13