

УДК 004.8

## Обзор теста Тьюринга и его вариаций

*Першенкова В.Г., студент  
Россия, 105005, г. Москва, МГТУ им. Н.Э.Баумана,  
кафедра «Системы обработки информации и управления»*

*Научный руководитель: Гапанюк Ю.Е., к.т.н., доцент  
Россия, 105005, г. Москва, МГТУ им. Н.Э.Баумана  
[gapyu@bmstu.ru](mailto:gapyu@bmstu.ru)*

### Введение

Основополагающей работой в области искусственного интеллекта (ИИ) стала статья «Вычислительные машины и разум» [1], написанная английским учёным Аланом Тьюрингом и опубликованная в 1950 году в журнале «Mind». Данная работа предоставила возможность общественности узнать размышления Тьюринга относительно создания машины, которая способна думать. В статье ученый поднимает вопрос о том «Способны ли машины думать?», но заменяет слова «машина» и «думать» и дает им собственные значения, так как считает данные слова семантически многозначными. В соответствии с этим Тьюринг задает вопрос ИИ по-своему – в виде игры, названной «Игра – имитация».

### Варианты теста

Рассмотри три основных варианта теста Тьюринга, два из которых были предложены ученым в статье [1], а третий вариант является стандартной интерпретацией. Вопрос интерпретации теста Тьюринга является достаточно дискуссионным, соответственно все три версии теста не являются равносильными, их сильные и слабые стороны различны.

*Первый вариант.* «Игра – имитация» включает в себя минимум трех игроков: мужчина (А), женщина (В) и игрок С любого пола, который задает вопросы двум другим игрокам. По правилам игры С находится в другой комнате и не видит ни А, ни В и может общаться с ними только посредством письменных сообщений. Цель игры состоит в том, чтобы игрок С, задавая вопросы игрокам А и В, пытается определить, кто из них — мужчина, а кто — женщина. Задачей игрока А является запутать игрока С, чтобы он

сделал неправильный вывод. В то же время задачей игрока В является помощь игроку С вынести верное суждение (см. рис.1).

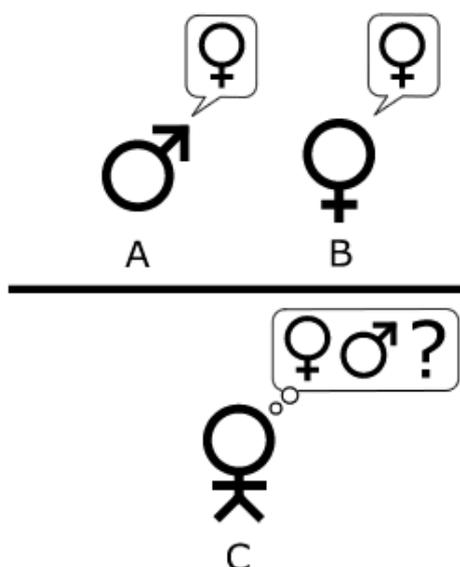


Рис.1 Первый вариант имитационной игры согласно описанию Тьюринга, в статье «Вычислительные машины и разум»

*Второй вариант.* Аналогично первому варианту роль игрока А исполняет компьютер, имеющий достаточное количество памяти, высокую скорость обработки данных и необходимый драйвер. Различие заключается в том, что роль игрока В может исполнять как мужчина, так и женщина. В этом варианте оба игрока А и В пытаются склонить ведущего С к неверному решению.

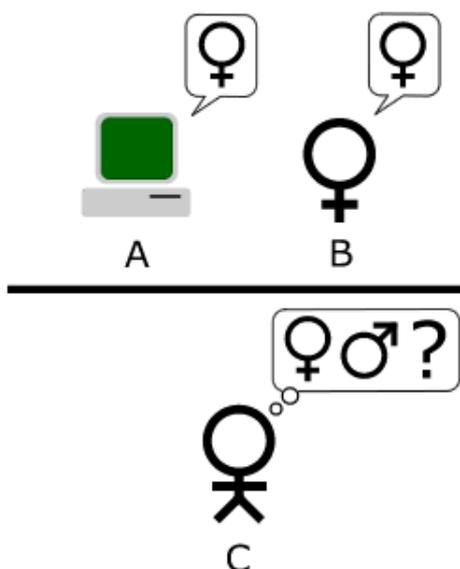


Рис. 2. Второй вариант имитационной игры. Игрока А заменяет компьютер

*Третий вариант.* Основой данной версии является цель теста Тьюринга, то есть ответ не на вопрос, способна ли машина ввести в заблуждение ведущего, а на вопрос, может ли машина имитировать человека или нет. Не известно подразумевался ли этот вариант Тьюрингом или нет, несмотря на это С.Г. Стеррет считает, что этот вариант Тьюрингом подразумевался и, таким образом, совмещает второй вариант с третьим. В это же время группа оппонентов, включая С. Трейджера, так не считает. В этом варианте игрок А — компьютер, игрок В — человек любого пола. Задачей ведущего является теперь не определить кто из них мужчина и женщина, а кто из них компьютер, а кто — человек.

### **Сильные стороны теста**

*Прост как в трактовке, так и в осуществлении.* Такие науки как философия, неврология и психология до сих пор не смогли дать точного определения словам *интеллект* и *мышление*, поэтому достаточно трудно применить столь общие и довольно расплывчатые дефиниции к машинам. Тест Тьюринга действительно инновационная вещь в области исследований ИИ, так как ученый подошел к вопросу мышления как к некой величине, которую можно измерить. Имея свой ряд серьезных недостатков, тест Тьюринга стал возможностью точной оценки интеллекта машины.

*Безграничная тематика теста.* Формат теста позволяет ведущему игры (игрок С) задавать машине чрезвычайно широкий спектр вопросов. Это в свою очередь позволяет ввести любую из сфер человеческой деятельности в тест. Для хорошего прохождения теста машина должна использовать язык, на котором говорит ведущий, разум, иметь определенное количество знаний и быть способной к обучению.

### **Слабые стороны теста**

В статье [1] Тьюринг не утверждал, что придуманный им тест может использоваться в целях измерения интеллектуальных способностей или любых других человеческих качеств. Его цель была дать точные и легко понимаемые альтернативы словам *думать* и *машина*, и предложить различные пути дальнейших исследований в области ИИ. Несмотря на это тест был воспринят как средство измерения способности мышления машины и, соответственно, данная идея была подвергнута критике как со стороны философов, так и со стороны ученых. Каждый аспект теста был поставлен под сомнение: надежность решения ведущего, надежность сравнения поведения (первый вариант теста) и надежность сравнения человека и компьютера (второй вариант теста).

Именно благодаря таким точкам зрения многие ученые ставили под вопрос полезность теста в своей области деятельности.

### **Человеческое поведение vs разумное поведение**

Тест Тьюринга не проверяет способен ли компьютер вести себя разумно, а тестирует его способность вести себя как человек. Так как человеческое поведение и разумное поведение — это различные понятия, тест способен ошибиться в двух случаях (см. рис.2):

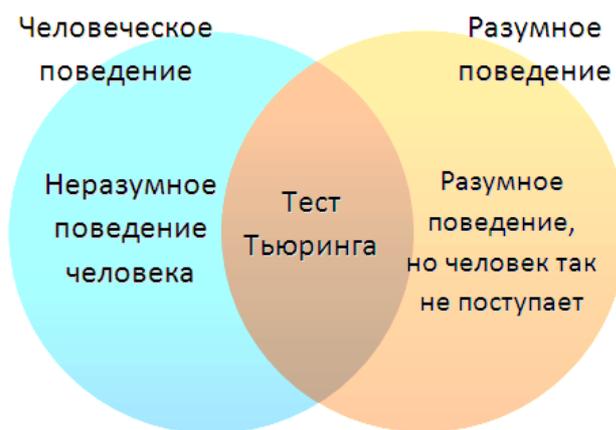


Рис. 3. Человеческое и разумное поведение

1. *Поведение некоторых людей является неразумным.* Тест требует того: чтобы машина могла воспроизвести любое человеческое поведение, не обращая внимание на то, является ли поведение разумным или нет. Он способен проверить поведение, которое совершенно противоположно разумному. Например, восприимчивость к оскорблениям, ложь или просто большое количество опечаток – если машина не способна воспроизвести такие действия, то она не пройдет тест.
2. *Разумное поведение может не соответствовать человеческому.* Тест не проверяет интеллектуальные способности машины, он требует некоего «обмана» с ее стороны в случае общения с менее интеллектуально развитым человеком. Ведущий с легкостью распознает программу, если она решит задачу практически непосильную человеку. Именно поэтому тест нельзя использовать для оценки и создания более развитых систем нежели человек.

### **Реальные интеллект vs эмитируемый интеллект**

Тест сконцентрирован исключительно на внешнем поведении машины, что говорит

об отсутствии необходимости мыслить – достаточно следовать некоторому своду правил, написанных в программе. Доказывая данную точку зрения Джон Сёрль публикует в 1980 году статью в которой описывает эксперимент «Китайская комната». Сёрль доказывает, что машина интеллектуальная по тесту Тьюринга не обязательно будет разумной, то есть не сможет понимать и строить осмысленные цели в своей программе.

Аргументируя собственную точку зрения Тьюринг утверждал, что не склонен считать сознание очевидной вещью, и предлагал разобраться в нахождении точной дефиниции данного слова.

### **Непрактичность теста Тьюринга в исследованиях ИИ**

На сегодняшний день тест Тьюринга редко используется для оценки качества программы с точки зрения ее интеллектуальных способностей, и на это есть ряд причин:

1. *Существует множество более простых средств тестирования программ.* Большинство исследований в области ИИ сейчас нацелено на вполне определенные цели: распознавание объектов, автоматизация процессов или логистика. Для того, чтобы выяснить работоспособность программы необходимо задать ей точную задачу, относящуюся к сфере разработки.
2. *Создание реалистичной модели человека и так сложная проблема. Для достижения основных целей ИИ нет необходимости ее решать.* Имитация человека могла бы быть интересна во многих сферах, кроме сферы непосредственно создания ИИ, здесь необходимы действительно разумные машины. Тьюринг не предлагал использовать данный тест как некоего мерила для оценки разумности ИИ, его целью было создание философии измерения ИИ.

### **Вариации теста Тьюринга**

#### **1) Обратный тест Тьюринга и CAPCHA**

Суть обратного теста Тьюринга заключается в обмене местами целями одного или более игроков между компьютером и человеком. Данная идея была отмечена учеными и данная модификация лишена большинства проблем, за которые осуждали стандартную версию теста.

CAPCHA (Completely Automated Public Turing test to tell Computers and Humans Apart - полностью автоматизированный публичный тест Тьюринга для различения компьютеров и людей) есть ничто иное как разновидность обратного теста Тьюринга.

Суть данного теста заключается в предоставлении пользователю такой задачи с которой легко сможет справиться человек, но для программы она будет значительно сложнее. Данный тест чрезвычайно популярен в последнее время и используется большим количеством систем (около 350 тысяч «капч» вводится ежедневно[5]). Пример «капчи» см. рис. 4.

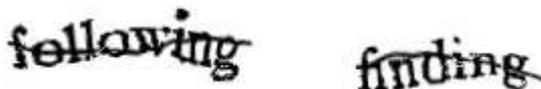


Рис. 4. Пример теста CAPTCHA

### **2) *Специалист предметной области***

В данной вариации теста программа пытается повторить ответ эксперта в данной предметной области, например, биология или информатика (так же известен как тест Фейгенбаума).

### **3) *Минимально интеллектуальный signal-тест (MIST)***

В данной разновидности теста возможны только односложные ответы бинарного вида (да/нет или правда/ложь). Сделано это для концентрации на способности мыслить. Это уменьшает проблемы с пониманием текстовых данных и не требует эмуляции неразумного поведения человека, и дает возможность проходить его программам с высоким интеллектом.

### **4) *Тест Эберта***

Данный тест проверяет способен ли голос, синтезированный компьютером рассказывать шутки достаточно квалифицированно, чтобы слушатель засмеялся. Проводится практически полная аналогия с тестом Тьюринга – оценивается разумность компьютера, как возможность компьютера выступать, так же как человек.

### **5) *Конкурс BotPrize***

Тест BotPrize проходит в виде многопользовательской компьютерной игры (Unreal Tournament 2004), персонажами которой управляли реальные люди или компьютерные алгоритмы. Игроки, одновременно выступавшие в роли ведущего, должны оценить вероятность того, что тем или иным персонажем управляет человек, а не компьютерная программа. На основе этой оценки для всех участников игры рассчитывалась средняя "человекоподобность" поведения.

## **Женя Густман и преодоление теста Тьюринга**

7 июня 2014 года Школой системной инженерии при университете и компаний RoboLaw было организовано полноценное тестирование Тьюринга и впервые в истории был пройден с помощью программы «**Eugene Goostman**», разработанной в Санкт-Петербурге выходцем из России Владимиром Веселовым и выходцем из Украины Евгением Демченко. Испытание состояло из серии пятиминутных письменных диалогов. Тест Тьюринга считался пройденным, если компьютеру удалось бы вводить собеседника в заблуждение на протяжении хотя бы 30 % суммарного времени. Программа Eugene победила с результатом 33 % и стала устройством, которое способно искусственным путём воссоздать человеческий интеллект. Ученые смогли смоделировать тринадцатилетнего подростка из Одессы, который хочет знать все, но не способен в силу своего возраста.

### **Различные премии за прохождения теста**

Можно безошибочно заметить, что тот тест, который преодолела программа, разработанная Российскими программистами есть не что иное, как разновидность теста – чатбот должен убедить в течение пятиминутного разговора более 30% судий в том, что он человек. Действительно впечатляющие 33% набранные «Евгением Густманом» на сегодняшний день являются прорывом в области ИИ, но пройдет еще очень много времени до создания чатбота способного соответствовать «золотому стандарту» современного теста Тьюринга.

Ученый Хью Лобнер создал правила прохождения «золотого стандарта» в 1990 году. Согласно им чатбот должен поддерживать разговор в течение 25 минут с каждым из четырех судий, программа выигрывает только в случае убеждения 50% судий. Каждый год самой правдоподобной программе вручают бронзовую медаль.

Самая трудная вариация теста Тьюринга была представлена в споре между учеными Рейем Курцвейлом и Митчем Капором (цена спора \$20 000). Их спор заключается в следующем – выигрывает та программа, которая сумела убедить 2 из 3 судий, поддерживая с ними разговор на протяжении двух часов. Митч Капор утверждает, что такая программа не будет создана вплоть до 2029 года.

### **Заключение**

Тест Тьюринга предложенный ученым в 1950 году стал действительно большим открытием как в области науки, так и в области философии. Алан Тьюринг ввел новые понятия в область изучения ИИ интеллекта, именно благодаря этому теперь понятие

сознание стало исчисляемым. Несмотря на то, что Тьюринг не акцентировал свое внимание на применении своих идей в жизнь, а наоборот подчеркивал внедрение новой философской парадигмы, многие отнеслись к его статье [1] чрезвычайно критически. Ученый Д.Серель явился одним из его главных оппонентов, который в доказательство своей точки зрения привел в работе «Разум, мозг и программы». Впоследствии критиковали и работу Сереля.

Шквал критики, вопросов и полемики вокруг теста Тьюринга — это полностью справедливая реакция. Человечество всегда страшила идея либо простое предположение (каковым и явился тест) о создании настоящего искусственного интеллекта, который смог бы, в дальнейшей перспективе, заменить наш разум. Попытка разложить понятие сознания «по полочкам», сделать его очень простым и доступным вызывает у большинства людей негодование, но у каждого в силу своих персональных особенностей. Некоторые просто хотят добиться правды с научной точки зрения, они задаются вопросом справедливости теста. Другие действительно опасаются всемогущих машин и думают о приношении клятвы веры компьютерам.

Если рассматривать тест с научно-технической точки зрения и применимости в наши дни, то увы – тест действительно интересен исключительно узкому кругу лиц, которые создают чатботов либо для борьбы за научные гранты, либо для собственного удовольствия. Тогда как в основном создаются роботы для определенных сфер, где используются специализированные тестирования, способные показать приемлемость спроектированной программы.

#### Список литературы

1. Turing Alan. Computing Machinery and Intelligence. Available at: <http://www.loebner.net/Prizef/TuringArticle.html>, accessed 29.11.2014.
2. The Turing test page. Available at: <http://www.fil.ion.ucl.ac.uk/~asaygin/tt/ttest.html>, accessed 29.11.2014.
3. Ваннах М. Как одессит Густман прошёл тест Тьюринга Умные машины, 9 июня 2014. // archive.today: Компьютерра. Режим доступа <https://archive.today/3NZLi> (дата обращения: 01.12.2014).
4. Turing test. Режим доступа: [https://en.wikipedia.org/wiki/Turing\\_test#cite\\_note-FOOTNOTESaygin2000-1](https://en.wikipedia.org/wiki/Turing_test#cite_note-FOOTNOTESaygin2000-1) (дата обращения: 09.12.2014).
5. CAPTCHA. Википедия. Режим доступа: <https://en.wikipedia.org/wiki/CAPTCHA> (дата обращения: 09.12.2014).

6. Hern A. What is the Turing test? And are we all doomed now? Режим доступа: <http://www.theguardian.com/technology/2014/jun/09/what-is-the-alan-turing-test> (дата обращения: 09.12.2014).
7. Российская государственная библиотека. Режим доступа: <http://www.rsl.ru/> (дата обращения 01.12.2014).