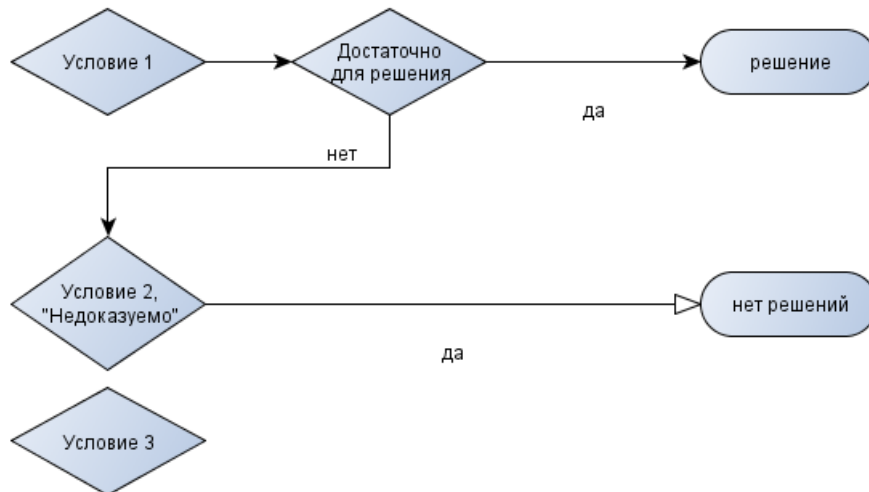


От парадокса узника к логике человека и ИИ

Никто не сможет доказать, что я говорю правду



На схеме показан принцип работы последовательной логики

Основные научные открытия и гипотезы работы

1. Вводится понятие «последовательная логика» – это логика человека и ИИ, которая отличается от формальной «параллельной» логики.
2. Выводимая аксиома – новая уникальная логическая конструкция
3. Гипотеза о косвенном доказательстве от противного
4. Аналитическое решение задачи парадокса неожиданной казни

Оглавление

Введение.....	3
Основные научные открытия и гипотезы работы.....	3
Постановка задачи.....	3
Логический парадокс.....	3
Метод исключения.....	5
Решение задачи – парадокса узника.....	5
Параллельная логика.....	5
Последовательная логика.....	6
Меняем условия местами.....	9
Косвенное доказательство от противного.....	9
Я прав, хотя Вы никогда не сможете это доказать!.....	11
Выводимая аксиома.....	12
Наиболее известные попытки решения парадокса узника.....	13
Решение, описанное Мартином Гарднером.....	13
Другие решения.....	14
The logical school.....	14
[edit]Objections.....	15

Введение

Пятнадцать лет назад одноклассник рассказал один парадокс, точнее сказать – логическую задачу, в который было очевидное противоречие, при этом найти это противоречие оказалось крайне сложным. С тех пор я много раз мысленно возвращался к той задаче и никак не мог её решить, наконец обратился к интернет и с удивлением обнаружил, что научно предлагаемые решения, в общем-то, не выдерживают критики, это дало стимул искать истину дальше, и вот что из этого получилось.

Основные научные открытия и гипотезы работы

1. Выводимая аксиома – новая уникальная логическая конструкция
2. Выделяются понятия параллельной и последовательной логики
3. Гипотеза о косвенном доказательстве от противного
4. Аналитическое решение задачи парадокса неожиданной казни

Постановка задачи.

Вариант, описанный в [1] с небольшими правками:

“Узник, приговорённый к высшей мере наказания, однажды в воскресенье был вызван к начальнику тюрьмы, честнейшему человеку, никогда не обманывающему даже самых злейших врагов общества. Начальник сказал: «Вас казнят на следующей неделе, но в какой именно день, я вам не скажу. Вы узнаете об этом дне только в полдень в день казни, когда палач войдёт к вам в камеру».

Через некоторое время узник обрёл способность нормально размышлять. Он немедленно сделал вывод, что казнить в следующее воскресенье его не могут: ведь тогда еще в субботу вечером будет ясно, что для казни остался один-единственный возможный день. Это будет означать, что узнику стал известен день казни еще до наступления этого дня, а такого не может быть по определению - ведь начальник никогда не лжёт!

"Иными словами, последний возможный день для моей казни, - думал узник, - это суббота. Но тогда если меня не казнят в пятницу утром, то к вечеру пятницы я тоже буду точно знать день, когда меня должны казнить. Однако начальник сказал, что о дне казни я узнаю лишь в полдень в день казни, а не накануне вечером, - значит, и в субботу меня не могут казнить тоже".

Последовательно рассуждая, несчастный исключил из рассмотрения пятницу, четверг, среду, вторник и понедельник. В итоге он пришёл к выводу, что его вообще не могут казнить, поскольку ни для одного дня недели условия, сформулированные начальником тюрьмы, не выполняются.

В среду утром заключённому, уже уверенному в своей безнаказанности, сообщили, что сегодня его казнят. Это явилось для него полной неожиданностью - и значит, условие начальника оказалось выполненным.»

Требуется найти ошибку в умозаключениях узника либо в словах начальника тюрьмы.

Логический парадокс

Логический парадокс – это рассуждение, приводящее к взаимно исключающим заключениям. Рассмотрим наиболее известные парадоксы.

Некто сказал: «Я лгу». Если предположить, что человек сказал правду, то он солгал. Если предположить, что он лжёт, значит он говорит правду. Другой популярный парадокс: «цирюльник бреет всех, кто не бреет себя сам», отсюда возникает вопрос - бреет ли себя цирюльник?

Слова про цирюльника в принципе ошибочны – так как такой цирюльник логически не может существовать. Подобно и слова «я лгу» сами по себе также противоречивы, вот оригинальное решение, автор Мирослав Войнаровский [3]:

«Итак: "я лгу". Более того, все скрытые смыслы фразы: "я лгу, что я лгу", "я лгу, что я лгу, что я лгу" и так далее - тоже ложь. Обозначим через X простейшую из этих фраз "я лгу". Остальные выразим через нее и составим систему уравнений:

$$X = \text{ложь}$$

$$(X = \text{ложь}) = \text{ложь}$$

$$((X = \text{ложь}) = \text{ложь}) = \text{ложь}$$

$$(((X = \text{ложь}) = \text{ложь}) = \text{ложь}) = \text{ложь}$$

...

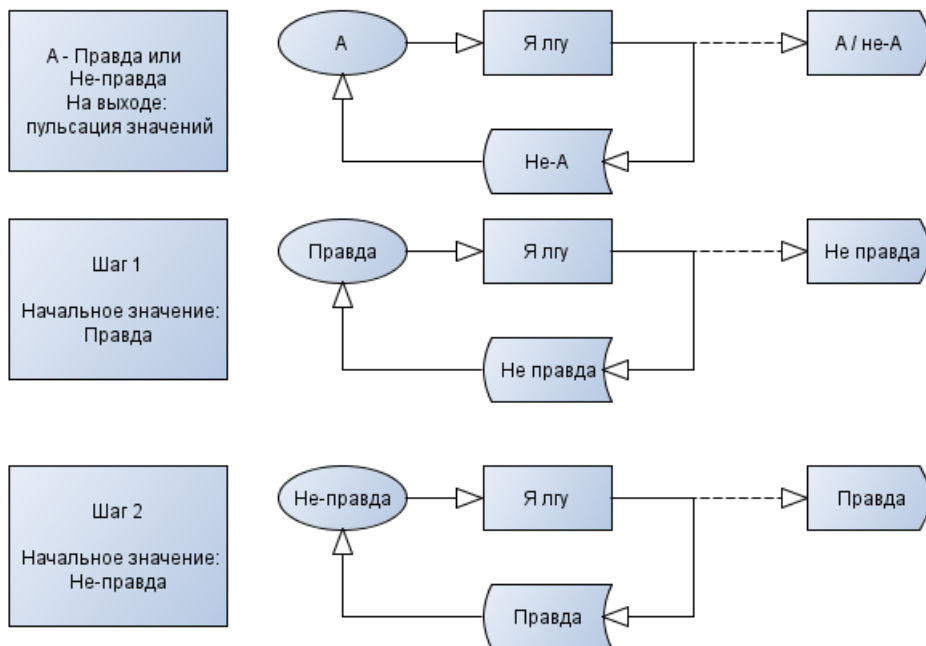
Получаем систему из бесконечного количества уравнений. Обратите внимание, не высказываний, а уравнений. В этом - вся соль. Если высказывание может быть истинным или ложным, то уравнение (или система уравнений) не может быть истинным или ложным. Оно может иметь решение или не иметь. Например, уравнение

$$x = x + 1$$

решений не имеет. Наша система уравнений тоже не имеет решений.»

Данный математический программистский подход очень изящен и мы будем использовать его в дальнейшем.

Изобразим парадокс лжеца схематически:



Результатом алгоритма является постоянно меняющиеся значения: правда / не правда.

Данный результат противоречит известному закону исключённого третьего, или, как иногда говорят, методу исключения.

Метод исключения

Закон исключённого третьего (лат. tertium non datur, то есть «третьего не дано») — закон классической логики, состоящий в том, что из двух высказываний — «А» или «не А» — одно обязательно является истинным, то есть два суждения, одно из которых является отрицанием другого, не могут быть одновременно ложными. Закон исключённого третьего является одним из основополагающих принципов «классической математики». [6]

Решение задачи – парадокса узника

Предлагаю следующее решение парадокса, точнее – задачи, так как то, что это парадокс – ещё нужно доказать!

Решением задачи будем называть нахождение логической ошибки в словах начальника тюрьмы или (и) узника. Также решением можно считать доказательство невозможности прийти к однозначному выводу.

Рассмотрим начальные условия, вот что сказал начальник тюрьмы:

«Вас казнят на следующей неделе, но в какой именно день, я вам не скажу. Вы узнаете об этом дне только в полдень в день казни, когда палач войдёт к вам в камеру»

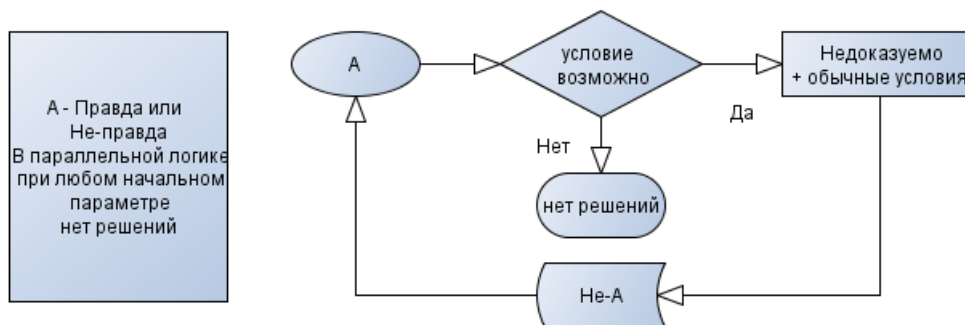
Упрощаем и формализуем: в первую очередь, переводим оборот «Вы узнаете об этом дне только когда...» в более чёткое описание «узник не может доказать». Выводим два условия:

1. **Казнь состоится с понедельника по воскресенье в полдень (Условие 1)**
2. **Утром в день казни узник не может доказать, что его казнят именно в этот день (Условие 2)**

В последний день данные условия будут противоречить друг другу – в случае, если мы рассмотрим оба условия одновременно. Действительно, из Условия 1 узник **доказывает**, что событие произойдёт в воскресенье, но Условие 2 утверждает, что узник **не может** этого **доказать**, хотя доказательство уже произошло.

Параллельная логика

Это классическая математическая логика, где все условия рассматриваются одновременно, без иерархии. Рассмотрим схему, где в Условия добавлено указание невозможности доказательства:



В этом случае у задачи не будет возможных решений.

Следующая система уравнений не имеет решений:

$$x = 1$$

$$x = 2$$

Для системы уравнений:

$$x^2 = 4$$

$$x \in \{-2, 2\}$$

решением является $x = 2$.

Главной идеей данной работы является предложение использовать другую, последовательную логику.

Последовательная логика

В реальной жизни, а также в программировании, обычно используется не одновременный перебор всех условий, а последовательный. Пример: мы хотим купить хлеб и имеем на выбор три магазина. Заходим в первый, покупаем хлеб, после чего мы не идём в два других, чтобы проверить наличие продукта – так как задача уже решена.

Последовательная логика – это приближенный к человеческому вид умозаключений, объединяющий индукцию: поиск решения в имеющемся опыте, и дедукцию: поиск решения на базе аксиом.

Последовательная логика может оперировать с избыточным количеством условий и даже с противоречивыми условиями.

В последовательной логике условия рассматриваются шаг за шагом до тех пор, пока не будет найдено решение. Соответствие условиям, которые остались «за кадром», то есть идут по списку после последнего, достаточного для решения условия – не проверяется. Подобная логика применяется в юриспруденции, где существуют следующие правовые документы:

- Конституция, основной закон
- Законы
- Подзаконные акты
- Официальные комментарии

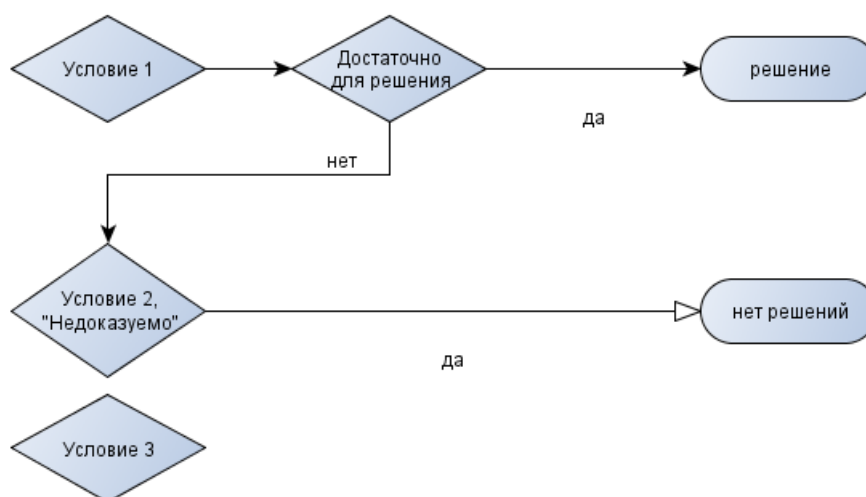
Если обычный закон противоречит конституции, то при рассмотрении дел будет учитываться иерархия и конституция будет иметь высший приоритет.

Поэтому введём следующее упрощение:

Условие 1 принимаем главным, Условие 2 – дополнительным.

Это значит, что проверка на условия ведётся не одновременно (параллельно), как принято в классической логике, а последовательно, и если Условие 1 дало решение, то на Условие 2 не обращаем внимания, оно будет уточнять решение, если Условие 1 его не даст.

Схема анализа в последовательной логике



Если одно из условий запрещает доказательство, то следующие условия не рассматриваются. Такая же картина будет и в других случаях, если набор входящих условий даст ответ «нет решений».

Вернёмся к системе уравнений, рассмотренной ранее в обзоре параллельной логики.

Следующая система уравнений в классической параллельной логике не имеет решений, при этом в последовательной логике решение есть:

$$x = 1$$

$$x = 2$$

Решение: $x = 1$. Так как первое условие достаточно для определения x , то следующие мы не рассматриваем. Именно такая логика используется, как правило, человеком. В памяти человека заложены миллионы и миллиарды всевозможных условий, и перебирать их все для принятия элементарных решений неэффективно да и просто невозможно. Поэтому человек перебирает их последовательно, и, если находит чёткий ответ, то прекращает мыслительную деятельность в данном направлении

Рассмотрим систему уравнений:

$$x + y = 2$$

$$x + y <> 2$$

Требуется найти x и y . В данных условиях решений нет.

Усложним:

$$x + y = 2$$

$$x = 2$$

$$x + y <> 2$$

в данном случае решение $x = 2, y = 0$, так как это следует из второй строки и не противоречит всему, что было ранее. Третье условие, противоречащее первому, уже не рассматриваем.

Применим последовательную логику к рассмотрению парадокса узника.

Рассмотрим таблицу вероятностей наступления события в каждый из дней, базируясь только на условии 1:

День недели	Вероятность наступления события
понедельник	14,3%
вторник	16,7%
среда	20%
четверг	25%
пятница	33,3%
суббота	50%
воскресенье	100%

В воскресенье узник анализирует Условие1 и приходит к выводу, что его казнят именно в воскресенье, так как больше не остаётся дней, вероятность наступления события: 100%. К Условию2 он не обращается. Вывод – слова начальника тюрьмы, то есть Условие1 + Условие2 противоречивы. В этом случае узник может обратиться к своему оппоненту со словами: «ты должен казнить меня сегодня, потому больше не осталось дней. И ты не можешь это сделать, потому что я об этом знаю. Вывод – как бы ты ни поступил, ты – солгал!»

Теперь возьмём утро субботы.

Узник может размышлять следующим образом:

«Исходя из Условия1 казнь будет сегодня или завтра. Если казнь будет завтра, то завтра утром Условие1 будет достаточным, чтобы сделать вывод, что казнь будет завтра. Другими словами, в субботу утром вероятность наступления события в текущий день, исходя из Условия1 – 50%, а в воскресенье утром – 100%

Более точный вывод из данного условия сделать нельзя, поэтому идём к Условию2. Применяем его к утру воскресенья – в этом случае оно противоречит Условию1. Таким образом, утром в воскресенье можно сделать вывод о том, что начальные условия противоречивы. Важный момент: Условие2 говорит о невозможности доказательства лишь в день казни, но не запрещает доказательство в другие дни. Таким образом, можно исключить воскресенье.» Методом исключения заключённый приходит к выводу, что казнь состоится в субботу, так как больше не осталось дней. В этом случае, если казнь состоится в субботу и узник в субботу утром приводит логическое доказательство, подтверждающее такой вывод, то это противоречит Условию2, которое утверждает невозможность доказательства В ДЕНЬ казни.

Значит, казнь в субботу может состояться только в том случае, если этого нельзя доказать, и если это нельзя доказать, то она будет именно в субботу(!), ведь воскресенье было отвергнуто «старшим» Условием1. Причина этого взаимоисключающего явления в том, что логический метод исключения не работает для условий, которые ссылаются на себя.

Как видим, формализация и последовательное применение условий даёт новые, поразительные результаты. Узник может доказать (!), что казни не будет в воскресенье, но не в состоянии доказать, что она будет в субботу, хотя это элементарно вытекает из метода исключения.

	суббота	воскресенье
возможность казни	возможно	невозможно

Мы получили крайне удивительный результат. Чтобы понять его, давайте рассмотрим задачу под ещё одним углом зрения. В Условии 2 речь идёт о том, что именно узник не может доказать день казни. Вводим стороннего наблюдателя, который «имеет право» доказывать.

Наблюдатель исследует умозаключения узника и приходит к следующей таблице.

	суббота	воскресенье
узник	возможно	невозможно
наблюдатель	ДА, суббота	невозможно

Для стороннего наблюдателя всё очевидно – если не воскресенье, то суббота. Тот факт, что узник не способен это доказать, на решение не влияет. Для стороннего наблюдателя не возникает ссылки-на-себя в Условии2, и поэтому сторонний наблюдатель может применять метод исключения.

Узник может прийти к следующему выводу: «сторонний наблюдатель может доказать, что казнь будет в субботу, но я не могу, так как это приведёт к противоречию с Условием2».

Уникальность этого парадокса – здесь именно парадокса – в том, что он находится не «внутри» условий, а рождается в момент попытки решения. Сами условия – не парадоксальны, они полностью непротиворечивы.

Добавим в таблицу пятницу.

	пятница	суббота	воскресенье
Узник	возможно	возможно	невозможно
Наблюдатель	возможно	ДА, суббота	невозможно

В данном случае, в пятницу и сторонний наблюдатель не может логически обосновать точный день казни. Картина не изменится, если мы добавим дни с понедельника по четверг – в этом случае казнь будет невозможна лишь в последний день и возможна во все остальные. В субботу сторонний наблюдатель может доказать, что событие будет в субботу.

Итак, финал решения:

1. высказывания начальника тюрьмы не содержат противоречий.
2. ошибка заключённого состояла в том, что он доказывал день казни базируясь на условии 2, запрещающего строить логические доказательства.
3. Казнь может состояться в любой день, кроме последнего, причём в субботу внешний наблюдатель способен доказать, что она состоится именно в субботу, а узник – нет.

Меняем условия местами

И проверяем методом последовательной логики.

1. **Утром в день казни узник не может доказать, что его казнят именно в этот день (основное)**
2. **Казнь состоится с понедельника по воскресенье в полдень (дополнительное)**

В этом случае размышления узника должны быть следующими.

Предположим, казнь будет в воскресенье. В воскресенье утром это невозможно доказать, так как по Условию 1 доказательство в день казни невозможно. Предположим, что казни НЕ БУДЕТ в воскресенье. Тогда по Условию 2 больше не останется дней и событие должно произойти в воскресенье. Значит – предположение неверное. Получается, что казнь будет, но это противоречит Условию 1. Замкнутый круг.

Переходим к субботе. Предположим, что казнь будет в субботу – опять по Условию 1 это доказать нельзя. Предположим, что казни в субботу не будет – из Условий 2 это вполне вероятно. Но доказать уже ничего не удаётся.

Таким образом, доказательство в данном случае близко к доказательству Гарднера – уже в последний день узник попадает в замкнутый круг логического парадокса, а в субботу доказательства теряют смысл.

Косвенное доказательство от противного

Теперь меняем Утверждение 2 на противоположное, получаем следующие условия:

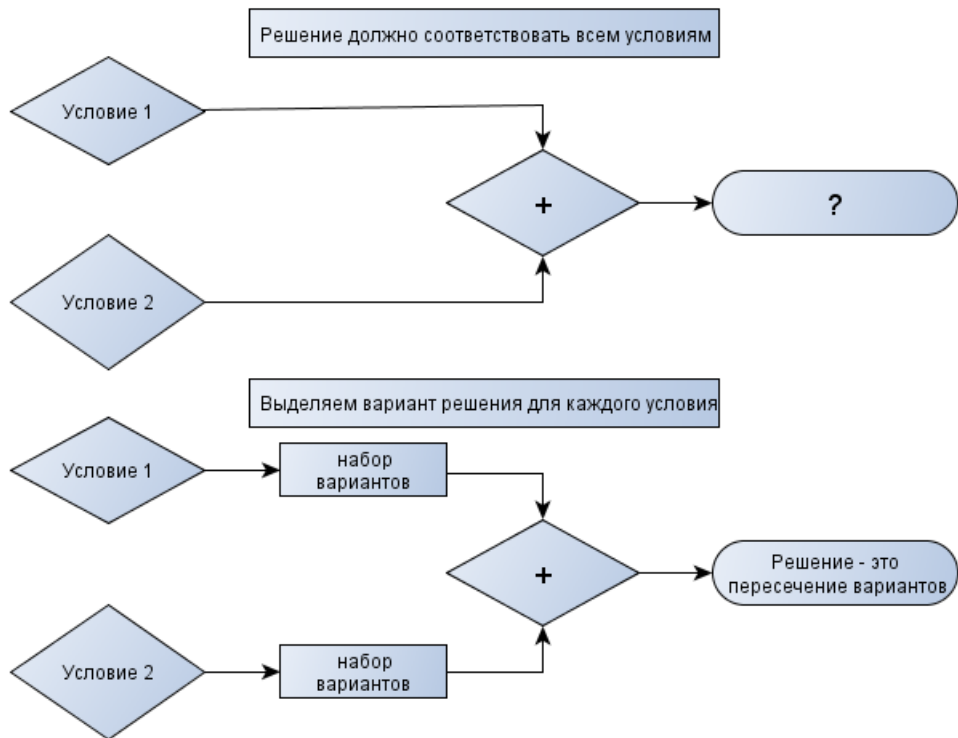
1. **Казнь состоится с понедельника по воскресенье в полдень**
2. **Утром в день казни узник МОЖЕТ доказать, что его казнят именно в этот день.**

В данном случае ответ очевиден – событие должно состояться в воскресенье, лишь в этот день вероятность наступления события будет 100% и узник может это доказать. Таким образом, «переворот» второго условия дал простые результаты, которые можно считать косвенным доказательством выдвинутой теории – так как при обычных условиях всё наоборот – узник не может доказать возможность или невозможность наступления события ни один день, кроме последнего.

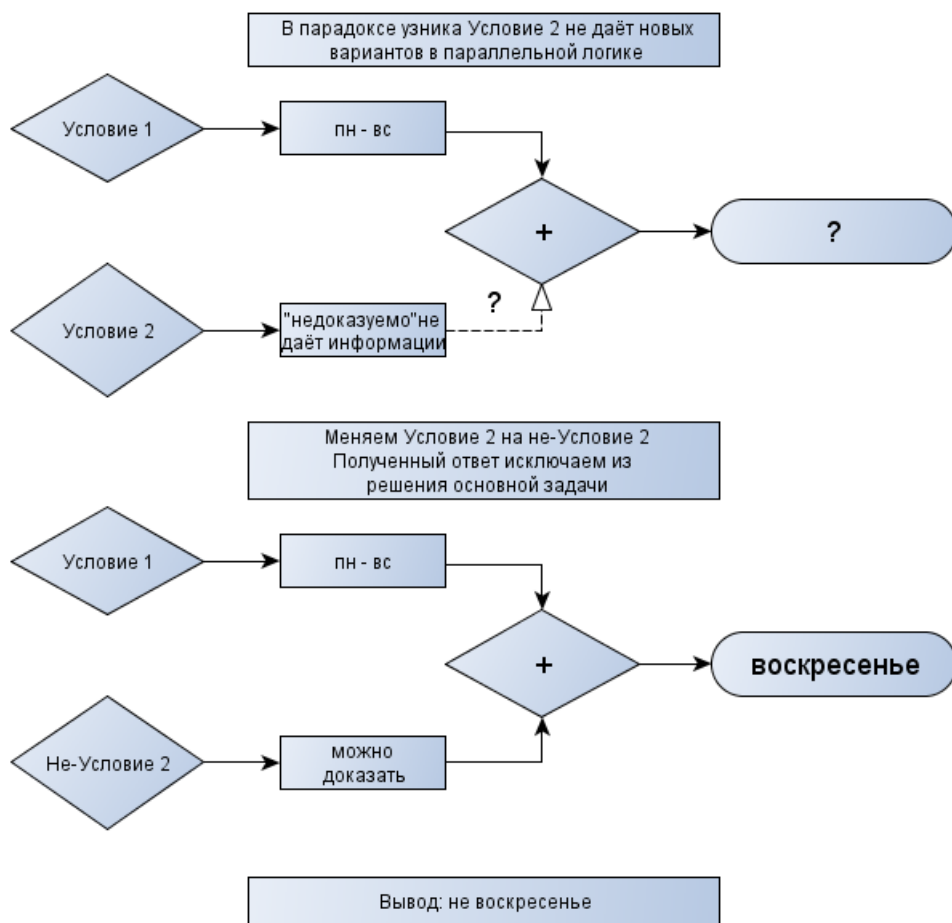
Это является косвенным доказательством потому, что классическое «доказательство от противного» строится по иной логике. Здесь же логика следующая:

1. Из некоторых условий (здесь: Условие 1) получили набор возможных значений (пн. – вс.)
2. Осталось одно условие. Мы не можем напрямую определить, как оно отфильтрует возможные значения, и «переворачиваем» его, чтобы отследить, какие значения оно отфильтрует в «перевернутом» случае. То есть, чтобы понять, что следует из условия А, мы проверяем, что следует из условия не-А.
3. Предполагаем, что если «обратное условие» даёт некий точный результат, то этот результат не должен входить в главный результат, где Условие 2 считается «напрямую».
4. В нашем случае, используя «обратное» Условие 2, получаем ответ: воскресенье. Значит, при прямом Условии 2 ответ: не-воскресенье.

Как видим, те же результаты были получены при использовании метода последовательной логики. Изобразим графически классическую методику доказательства с помощью метода исключения при наличии двух или более условий.



Для применения метода косвенного доказательства от противного, меняем одно из условий на противоположное.



Классическое доказательство от противного базируется на законе исключённого третьего, которое может не действовать для само-ссылающихся условий и особенно для последовательной логики, где возможно наличие взаимоисключающих условий.

Последовательная логика в социологии и психологии

В главе о последовательной логике упоминалось, что она является наиболее удобной для человека, ввиду высокой скорости принятия решений. Столкнувшись с задачей, человек ищет ответ на неё, в первую очередь, среди набора из сотен и тысяч условий, связанных с темой. Важное значение имеет иерархия условий – какие из них окажутся вначале и какие в конце, ведь если начальные условия дают точный ответ, то к последующим мозг уже не имеет необходимости обращаться. Последовательности условий у всех людей разные, личный опыт, воспитание и пропаганда на них оказывают значимое воздействие – поэтому часто встречаются случаи, когда по одному и тому же вопросу у людей диаметрально противоположные мнения. Более того, не просто мнения, а убеждённости, включая вопросы фанатизма и паранойи.

Пример, возьмём двух людей со схожими морально-этическими ценностями. Они являются гражданами своей страны и военнообязанными, готовы отстаивать интересы государства, в том числе в случае военных действий. В систем уравнений этих людей существуют следующие понятия:

Любовь к родине
Доброта
Стремление к справедливости
и т.д.

Пусть данные граждане будут проживать в разных государствах. Получится, что у одного:

Любовь к родине А
Доброта

....

У другого:

Любовь к родине Б
Доброта

....

В итоге два человека с почти идентичным мировоззрением, добрые и, в лучшем понимании это слова, ХОРОШИЕ, могут стать злейшими врагами ввиду политического конфликта. Подобным образом рождаются и религиозные мотивы конфликтов. В психологии известно, что личное для человека важнее общественного, и «любовь к родным и близким», как правило, сильнее «любви к родине» - поэтому в вопросах убеждения нередко идёт обращение к их личным интересам. Например, не просто «защищать родину», а «защищать своих близких от посягательства врага».

Метод последовательной логики позволяет рассмотреть многие социальные и психологические проблемы чисто математически.

С точки зрения последовательной логики, любовь - это выход интереса к любимому человеку на первое место в иерархии условий. Кроме иерархии, у каждого понятия есть свой «вес». Если некое условие не просто находится на вершине иерархии, но обладает большим весом, то обращение к другим условиям теряет важность для человека. В этом случае все действия, которые совершает человек, направляются на реализацию некоей «идеи фикс» - условия, которое находится на вершине иерархии и обладает большим весом. Все прочие условия рассматриваются, в первую очередь, как методы достижения главной цели, и лишь во вторую - сами по себе. Если любовь выходит на первое место и обладает большим весом, что человек может делать огромную работу, которая внешне выглядит как неадекватное поведение или наоборот, как вдохновение на большие социальные или творческие достижения, хотя на самом деле человек просто решает одну задачу, находящуюся в начале иерархии внутренних условий и занимающую большой вес.

Я прав, хотя Вы никогда не сможете это доказать!

Вернёмся к основному решению – в субботу узник, понимая что казнь может быть только в субботу – не способен это доказать.

В данной задаче парадокс возникает лишь во время решения. Если такая конструкция справедлива, то стоит найти более простой подобный случай, в стиле парадокса «я лгу». Немного поразмыслив, приходим к короткой и красивой фразе.

«Вы не можете доказать, что я говорю правду»

Эта фраза истинна всегда. Сама по себе она непротиворечива. При этом её решение сразу же приводит к парадоксу.

Фразу можно чуть удлинить для большей эстетики:
«Вы не можете доказать, что я говорю правду. Хотя я – прав!»

Либо:

«Я прав, хотя Вы никогда не сможете это доказать!»

Или:

«Невозможно доказать, что это – правда!»

И совсем коротко:

«Недоказуемо!»

По сути, именно такая фраза и легла в основу Условия² разбираемой задачи об узнике.

Если лингвистический оборот «Вы не можете этого доказать» или «Вы не будете знать» употребить к ряду условий, к которым будет относиться данная фраза, то это позволит создать множество задач, похожих на «парадокс неожиданной казни».

Выводимая аксиома

Фраза «Вы не можете доказать, что данная фраза истинна» не является обычным парадоксом. Дело в том, что в отличие от классических парадоксов «я лгу» и подобных – она непротиворечива! Данная фраза правдива во всех случаях, при этом доказать её невозможно.

Положения, которые принимаются за истину без доказательства называются аксиомы.

Получается, что данная фраза – типичная аксиома!

Если аксиому доказать, то есть вывести на базе других аксиом – то она становится теоремой.

То, что фраза «Вы не можете доказать, что данная фраза истинна» полностью справедлива, мы вывели, можно сказать доказали, и тем самым опровергнули... собственные доказательства, но не фразу. Так что можно сказать, что данная фраза - выводимая аксиома.

Афоризм «Невозможно доказать, что это – правда» очень хорошо вписывается в последние достижения науки, разбирающие вопросы неопределённости – такие как нечёткая логика, эвристические методы решения сложных задач, ядерная физика.

Данная работа вводит лишь базовые понятия последовательной и параллельной логики, выводимой аксиомы и косвенного доказательства. Они могут быть использованы для развития логических теорий, и иметь прямое практическое применение в разработке искусственного интеллекта.

Далее разбираются известные статьи на тему решения парадокса узника, для сравнения с предлагаемой методикой последовательной и параллельной логики.

Наиболее известные попытки решения парадокса узника

Решение, описанное Мартином Гарднером

Обратимся к Wikipedia [5]:

«Разрешению этого парадокса посвящено много научных статей. Мартин Гарднер, описывая вариант парадокса[2], утверждает, что ошибка содержится уже в первом шаге рассуждений. Предположим, что осталась последняя альтернатива. Узник не может сделать логически безупречный вывод, казнят ли его завтра, поскольку в любом случае это бы приводило к противоречию двух условий. Поэтому для него завтрашнее событие будет неожиданностью.

Это приводит к упрощённой формулировке парадокса, имеющей название «яйцо-сюрприз».

Яйцо-сюрприз

Чтобы разобраться в этом парадоксе, имеет смысл рассмотреть его более простую форму, где число дней уменьшено до одного. В этой версии начальник тюрьмы сообщает, что:

Вы будете казнены в полдень на следующей неделе в пятницу;

Это будет неожиданностью для вас.

Заключённый восклицает, что оба условия не могут быть выполнены, так как казнь не может быть неожиданной, если уже сообщено, что она произойдёт в пятницу, и полагает, что казнь не состоится. В следующую пятницу заключённого казнят. Это становится неожиданностью для него, так как он убедил себя, что казнить его не смогут. Что было неправильно в его рассуждениях? Или, возможно, условие «это будет неожиданностью для вас» ложно. Если заключённый является настолько уверенным в нем, что до последних секунд своей жизни считает, что казнь будет остановлена, чтобы выполнить условие начальника.

Любопытный вывод заключается в том, что сам заключённый не сможет прийти к выводу, что приговор будет исполнен. К такому выводу может прийти только тот, кто располагает дополнительной информацией о том, что:

заключённому действительно отрубят голову в указанный день;

заключённый является строго логически (и никак иначе) мыслящим субъектом.

Эквивалентной формулировкой будет следующая. Пусть некто мистер Смит даёт коробку и говорит: «Откройте её, и вы неожиданно обнаружите внутри яйцо»[2]. Можно ли, рассуждая логически, прийти к какому-нибудь заключению относительно того, находится ли внутри коробки яйцо, или его там нет? Если Смит говорит правду, то внутри должно быть яйцо, но тогда это не будет сюрпризом. Поэтому утверждение Смита ложно. Если это убедит адресата, что из этого следует отсутствие яйца, то открыв коробку и неожиданно обнаружив там яйцо, адресат убедится, что утверждение Смита истинно.

Аналогично можно заключить, что хотя Смит и сдержит слово, но адресату об этом неизвестно. Следовательно, адресат не может сделать логический вывод из слов Смита, и они просто не несут информации.»

Насколько убедительным можно считать предложенное решение, в котором допускается наступление события в воскресенье? Действительно – если останется последний день, то в высказываниях начальника тюрьмы есть явное внутреннее противоречие: с одной стороны, утверждается, что событие наступит в

воскресенье, так как дней больше не осталось; с другой стороны, утверждается, что узник не будет знать день казни. То же самое происходит в варианте с яйцом: с одной стороны, даётся информация о наличии яйца, с другой утверждается что эта информация является неожиданной, другими словами, отрицается факт наличия реальной информации. Мы можем утверждать, что налицо - парадоксальные высказывания, из которых действительно нельзя сделать вывод – сами по себе такие высказывания можно назвать ложными.

При этом, если казнь наступает в среду – как сказано в условии, то получается, что никаких внутренних противоречий в словах начальника тюрьмы нет. Это значит, что в последний день заключённый хотя и не может доказать, что его не казнят - но он **может доказать, что в словах начальника тюрьмы скрыто внутреннее противоречие**. При этом в среду узник не может доказать, что слова начальника тюрьмы либо внутренне противоречивы. В будущем мы вернёмся к этой теме и попробуем разобраться, в чьей логике есть ошибки либо внутренние противоречия – в высказываниях начальника тюрьмы или узника.

Анализ Гарднера и его выводы полностью классической параллельной логике. Для читателя такой метод мышления может показаться странным, так как люди чаще используют последовательную логику.

Возьмём фразу:

«Откройте её, и вы неожиданно обнаружите внутри яйцо»

и формализуем её:

Условие 1. В коробке лежит яйцо

Условие 2. Это невозможно доказать

В параллельной логике решения нет. В последовательной – в коробке лежит яйцо. Второе условие лишнее и потому не используется. Дополнительно, второе условие является противоречивым.

Другие решения

Очень интересны статьи Константина Кнопа[1] и некоего “jb_d”[4] с критикой решения Гарднера и предложениями иных подходов.

Кноп[1] даёт простое и достаточное опровержение варианта Гарднера, N – количество оставшихся дней в неделе:

«При N=1 парадокс очень прост. Начальник говорит буквально следующее: "Вас казнят завтра, но до завтрашнего утра вы ничего об этом не узнаете". Это утверждение противоречиво само по себе, поэтому никакой информации узник из него вынести не сможет. Противоречивые утверждения не могут быть ни истинными, ни ложными. Иначе говоря, из всей информации парадокса можно сделать только один вывод: начальник тюрьмы не всегда говорит правду.»

В дальнейшем автор ставит важный вопрос, в какой момент происходит переход от ситуации противоречивости и бессмысленности условий к ситуации, когда условие корректно, а рассуждения узника ошибочны.

Кноп и jb_d пишут, что главная ошибка в рассуждениях – предположение «многократного знания», которое само по себе содержит логическую ошибку – знание должно быть точным. Например, утром в субботу узник может сделать точный вывод и сказать «я знаю что казнь будет сегодня» - а в итоге оказывается, что в субботу за узником не пришли, а пришли в воскресенье – значит субботнее «знание» и «доказательство» были ложными.

Впрочем, это кажется слишком запутанным – например, узник ошибся на счёт субботы – и в воскресенье он уже точно знает, пусть после субботней ошибки, что казнь будет в воскресенье. И тогда условие станет противоречивым.

Английская Wikipedia [6] рассказывает о логическом подходе к решению парадокса. Текст довольно сложен для перевода, поэтому привожу оригинал и затем перевод:

The logical school

Formulation of the judge's announcement into formal logic is made difficult by the vague meaning of the word "surprise". An attempt at formulation might be:

- *The prisoner will be hanged next week and the date (of the hanging) will not be deducible in advance from the assumption that the hanging will occur during the week (A).*

Given this announcement the prisoner can deduce that the hanging will not occur on the last day of the week. However, in order to reproduce the next stage of the argument, which eliminates the penultimate day of the

week, the prisoner must argue that his ability to deduce, from statement (A), that the hanging will not occur on the last day, implies that a last-day hanging *would not be surprising*. But since the meaning of "surprising" has been restricted to *not deducible from the assumption that the hanging will occur during the week* instead of *not deducible from statement (A)*, the argument is blocked.

This suggests that a better formulation would in fact be:

- *The prisoner will be hanged next week and its date will not be deducible in advance using this statement as an axiom (B).*

Some authors have claimed that the [self-referential](#) nature of this statement is the source of the paradox. Fitch^[5] has shown that this statement can still be expressed in formal logic. Using an equivalent form of the paradox which reduces the length of the week to just two days, he proved that although self-reference is not illegitimate in all circumstances, it is in this case because the statement is self-contradictory.

[\[edit\]](#) Objections

The first objection often raised to the logical school's approach is that it fails to explain how the judge's announcement appears to be vindicated after the fact. If the judge's statement is self-contradictory, how does he manage to be right all along? This objection rests on an understanding of the conclusion to be that the judge's statement is self-contradictory and therefore the source of the paradox. However, the conclusion is more precisely that *in order for the prisoner to carry out his argument* that the judge's sentence cannot be fulfilled, he must *interpret* the judge's announcement as (B). A reasonable assumption would be that the judge did not intend (B) but that the prisoner misinterprets his words to reach his paradoxical conclusion. The judge's sentence appears to be vindicated afterwards but the statement which is actually shown to be true is that "the prisoner will be *psychologically* surprised by the hanging". This statement in formal logic would not allow the prisoner's argument to be carried out.

A related objection is that the paradox only occurs because the judge tells the prisoner his sentence (rather than keeping it secret) — which suggests that the act of declaring the sentence is important. Some have argued that since this action is missing from the logical school's approach, it must be an incomplete analysis. But the action is included implicitly. The public utterance of the sentence and its context changes the judge's meaning to something like "there will be a surprise hanging despite my having told you that there will be a surprise hanging". The logical school's approach does implicitly take this into account.

Перевод:

Логическая школа

Формулировка объявления судьей в формальной логике затруднена расплывчатым значением термина "surprise" – «быть неожиданным, непредсказуемым, сюрпризом». Попытка формулировки может быть следующей:

Закл^юч^ённый будет казн^{ён} на следующей неделе, при этом дата казни не будет выводима заранее из предположения, что казн^ь произойдет в течение недели (A).

Учитывая это объявление, заключенный может вывести, что казн^ь не произойдет в последний день недели. Однако, для того, чтобы воспроизвести следующий шаг доказательства, который устраняет предпоследний день недели, заключенный должен доказать свою способность делать выводы из утверждения (A) о том, что событие не произойдет в последний день. Это подразумевает, что в последний день казн^ь не будет неожиданным. Но, так как значение термина "surprise/быть неожиданным, непредсказуемым" было упрощено до *дата не будет выводима из предположения, что казн^ь произойдет в течение недели*, вместо *не выводимо из утверждения (A)*, данное умозаключение (аргумент) будет заблокировано.

Из этого следует, что лучше применять следующую формулировку:

Закл^юченный будет казн^{ён} на следующей неделе, и дата не будет выводима заранее, используя это заявление в качестве аксиомы (B).

Некоторые авторы утверждают, что самоссылающийся характер этого заявления является источником парадокса. Fitch [5] указывает, что это утверждение всё еще может быть выражено в формальной логике.

Используя эквивалентную форму парадокса, которая уменьшает длину недели до всего двух дней, он доказал, что, хотя ссылки-на-себя не ведут к ошибкам во всех обстоятельствах, в данном случае источником парадокса является внутреннее противоречие начальной формулировки.

Возражения

Первым возражением против логического подхода является то, что он не сможет объяснить, как объявление судьи оказывается истинным после свершившегося факта. Если заявление судьи является внутренне противоречивым, как ему удаётся всегда оставаться правым? Это возражение основано на том, что заявление судьи является внутренне противоречивым, и, следовательно, внутреннее противоречие является источником парадокса. Тем не менее, более точен следующий вывод: для того, чтобы заключённый мог полагаться на своё умозаключение о том, что приговор судьи не может быть выполнен, он должен интерпретировать заявление судьи как (В). Разумным предположением будет то, что судья не предполагал (В), но заключённый искажённо понял его слова, чтобы прийти к своему парадоксальному выводу. Слова судьи были подтверждены впоследствии, но заявление, которое, на самом деле можно считать истинным, следующее: "заключённый будет психологически удивлён казнью". Это заявление в формальной логике не позволит аргументам заключённого осуществиться.

Есть схожее возражение, что парадокс возникает только потому, что судья говорит заключённому приговор, вместо того, чтобы держать его в секрете. Возражение предполагает, что само объявление приговора является важным. Некоторые утверждают, что, поскольку этот вариант отсутствует в логической школе, он, по-видимому, даёт неполный анализ. Действие включается в неявном виде. Общественное высказывание приговора и его контекст меняет смысл слов судьи на что-то вроде: "Казнь будет неожиданной несмотря на мои слова о том, что казнь будет неожиданной". Подход логической школы неявно принимает это во внимание.

Конец перевода.

Резюме о других методах доказательства парадокса узника.

вариант решения	комментарий
Слова начальника тюрьмы содержат парадокс	Этого мнения придерживался и узник
Казнь в воскресенье возможна	При использовании лишь параллельной логики - да
Психологические аспекты	Не выглядят логически убедительными

- I. Литература
1. Константин Кноп, "Парадокс узника", "Компьютерра" №17 от 04 мая 1998 года
 2. Мартин Гарднер А ну-ка, догадайся! = Aha! Gotcha. Paradoxes to puzzle and delight. — М.: Мир, 1984. — С. 33-36. — 213 с.
 3. Мирослав Войнаровский <http://psi-logic.narod.ru/psi/lier.htm>
 4. <http://jb-d.livejournal.com/>
 5. http://ru.wikipedia.org/wiki/Парадокс_неожиданной_казни
 6. http://ru.wikipedia.org/wiki/Закон_исключённого_третьего

Чужакин Евгений Евгеньевич
www.neoneuro.com
 info (собака) neoneuro.com