

«ЦЕЛОЕ», «ЖИВОЕ», «СОЗНАЮЩЕЕ»: ПОПЫТКА МОДЕЛИРОВАНИЯ

Гришин Е.А.

г. Москва

bg987@mail.ru

Предложен подход, позволяющий содержательно интерпретировать понятия «целое», «живое» и «сознающее» с целью заложить основу для компьютерного моделирования подобных феноменов. В качестве исходного материала привлекаются понятия «универсальная машина Тьюринга (УМТ)», а также принцип самоприменимости УМТ как способ её самопрограммирования. В качестве инструмента используются «принцип заимствования», предложенный В.А. Лефевром, а также предлагаемая автором модель рефлексивной машины и принцип самопланирования (взамен принципу самопрограммирования УМТ). Показано, что в результате может быть снята проблема алгоритмической невычислимости задачи самоуправления УМТ как помеха на пути алгоритмического варианта реализации функции «машинного сознания».

Технологически легко обеспечить простую репликацию (порождение) одним автоматом других автоматов [1]. При этом можно организовать «эволюцию» популяции томатов [2].

Есть искушение представить подобные модели как модели «искусственной жизни».

Но понимать этот термин буквально как искусственное воспроизведение жизни, наверное, было бы натяжкой.

В чистом виде это не более чем процессы генерации автоматов некоей операционной системой (или многократное копирование автоматов). При этом каждый из автоматов сам по себе не является (не обязан быть) чем-то живым, и не обязан обладать самостоятельной функцией самокопирования.

Есть множество примеров процессов самоорганизации в неживой материи, успешно исследуемые синергетикой [2]. В рамках синергетики ищутся механизмы решения проблемы целого. Но есть ощущение, что это будет решение проблемы целого для неживого, так как не определена разница между живым и неживым.

Для решения проблемы целого для живого (биологического и социального) может помочь применение принципа заимствования, сформулированного В.А. Лефевром [3].

С одной стороны, этот принцип предполагает существование таких объектов, которые имеют проект строительства самого себя.

С другой стороны, предполагается наличие механизма переноса (заимствования) проекта при порождении одним объектом другого, подобного первому.

Принцип заимствования, на наш взгляд, может быть базовым признаком отличия «живых» объектов от «неживых», по крайней мере, при компьютерном моделировании жизни.

Представляется, что для адекватного моделирования жизни (т.е., практически, для создания новой популяции неких организмов), необходимо, чтобы сам базовый элемент порождаемого множества уже был «живым», по крайней мере в рамках принципа заимствования.

Попытаемся (самостоятельно от автора [3]) развернуть собственное понимание принципа заимствования, сохраняя в неприкосновенности исходную посылку.

Напомним, что данный принцип в общем предполагает наличие у объекта проекта самого себя, т.е. образца и механизма строительства другого объекта, подобного первому, в соответствии с этим образцом.

Перечислим в порядке убывания общности перечень качеств, интуитивно нами приписываемый «живому» и «сознающему» объекту, образец которого следует использовать для строительства другого подобного объекта.

Свойство репликации, т.е. самостоятельную способность к порождению одним объектом другого, себе подобного.

- «Физиологическая» активность. Имеется в виду способность объекта к преобразованию энергии как расходу одного вида материального ресурса и выработке другого вида материального ресурса. Иначе говоря, это - требование метаболизма, или обмена веществ. Или обобщённо - требование цикличности.
- Потребности как способ проявления «физиологической» активности,
- Деятельности как способ удовлетворения потребностей,
- Проблемы как оценка вероятных затрат, которые могут последовать из-за помех в осуществлении деятельности по удовлетворению конкретной потребности,
- План поведения (список ролей и последовательность действий каждого исполнителя) как путь устранения проблемы. Сам факт появления плана у субъекта уже есть способ психологического разрешения проблемы.
- Принятие решения как процесс генерации и выбора вариантов поведения на каждом этапе плана,
- Информация как ответ на вопрос, заданный субъектом. Ответ может быть получен им от другого субъекта, или от себя самого. (Без вопроса информации нет! Вопрос появляется тогда, когда у автора есть необходимость принять какое-то решение и нет источника вариантов или требуемых знаний для оценки появляющихся вариантов).

Таким образом, вышеприведённая нисходящая цепочка являет собой проектный образец.

В свою очередь, проектный механизм есть набор инструментов для строительства по данному образцу, который, по нашему мнению, адекватно может быть представлен рефлексивной машиной (рис. 1).

В целом, рефлексивная машина есть система из двух универсальных машин Тьюринга, параллельно работающих на общем поле памяти.

Термин «рефлексивная машина» впервые задан автором в [5] как инвариантная модель технологии обратной связи в деятельности персонажа, обратной связи между осознанием и поведением персонажа.

Более основательно структура и функционирование рефлексивной машины будет рассмотрено в специальной работе.

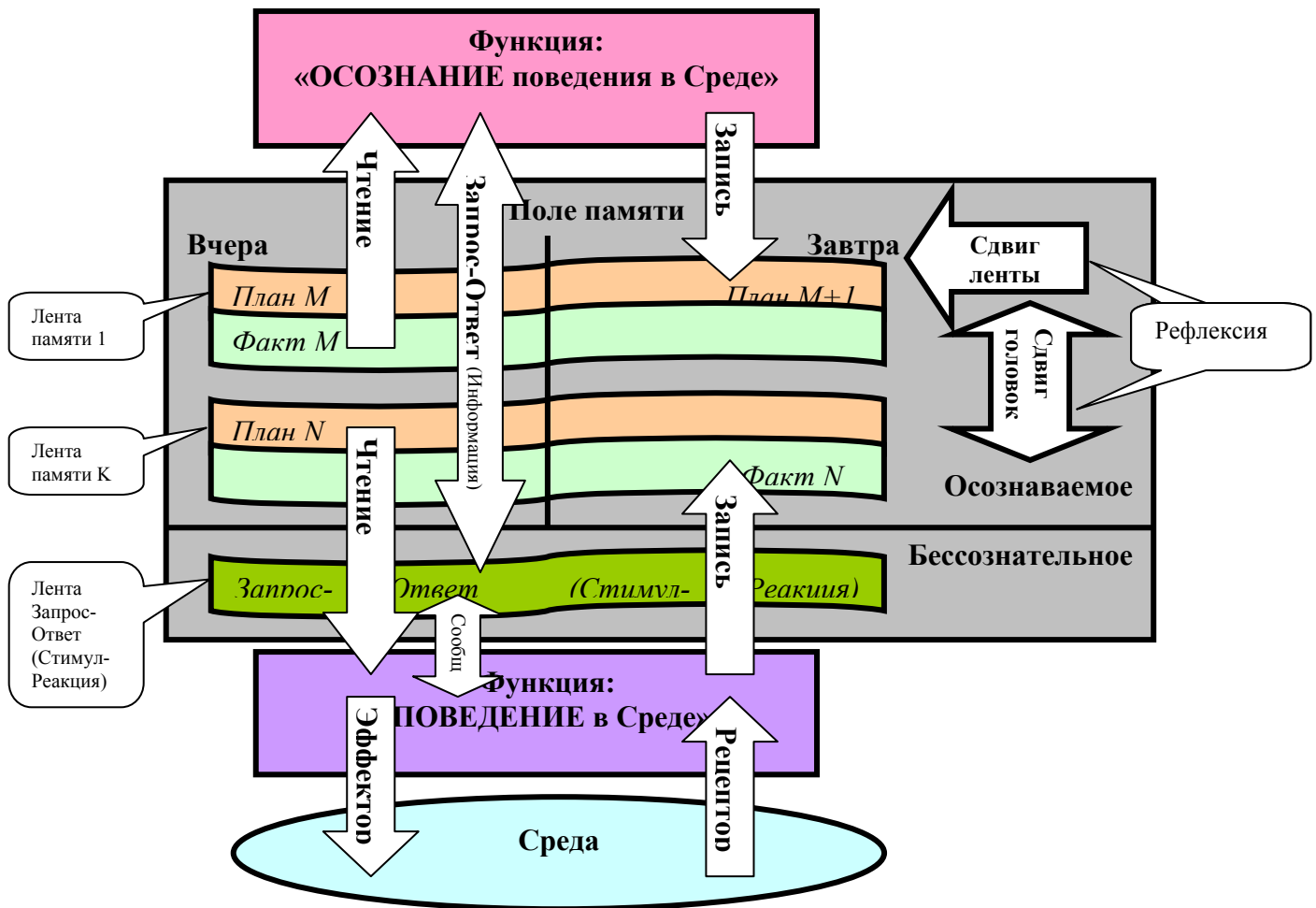


Рис. 1. Рефлективная машина.

В рамках исследовательского направления «искусственный интеллект» известны попытки анализа проблемы «самоприменимости» универсальной машины Тьюринга (УМТ) [4].

Вкратце, самоприменимость представляет собой такой режим работы УМТ, когда она генерирует на ленте памяти не просто данные, а программный код собственного устройства управления, который сама же далее читает и интерпретирует (исполняет).

При этом важно, что самоприменимость рассматривается как алгоритмический вариант реализации функции «машинного сознания».

Цель такого анализа – найти подтверждение или опровержение тезиса об алгоритмической невычислимости самой задачи самоприменимости УМТ, а значит, и подтверждение или опровержение по данной причине тезиса о невозможности машинного «интеллекта», машинного «сознания».

На наш взгляд, самоприменимость УМТ, рассмотренная вышеизложенным способом, только и может, что привести к выводу об алгоритмической невычислимости задачи, а значит, и к нерешаемости проблемы «машинного сознания» таким способом.

Причина тому – тривиальное зацикливание ввиду самовывоза начальной инструкции управляющей программы.

Предложим иной подход, в результате которого может быть снята проблема алгоритмической невычислимости задачи самоприменимости УМТ.

Иначе говоря, покажем, что существуют условия, при которых допустима самоприменимость УМТ, а значит, и теоретически не запрещена реализация «машинного сознания» на этом принципе.

Предположим, что некая УМТ в течение какого-то периода времени своей работы генерирует не собственный программный код, а некие данные, которые будет читать в следующем периоде другая УМТ, работающая параллельно с первой на этом же поле памяти.

И восприниматься эти данные другой УМТ будут не как инструкции программы, а как некий план, в соответствии с которым её управляющая программа должна будет включать и выключать требуемые исполнительные алгоритмы своего поведения (рецепторы и эффекторы) в требуемое время.

Далее, эта вторая УМТ, читая и исполняя в следующем периоде полученный план, запишет в общее поле памяти некий факт выполнения этого плана.

Этот фактический результат выполнения плана второй УМТ в очередном периоде прочтёт первая УМТ, и если план не выполнен, включит соответствующий алгоритм генерации нового плана.

Таким образом, выше описана последовательная процедура генерации данных одной УМТ и чтения этих данных другой УМТ.

Закономерен вопрос: а что делает другая УМТ в тот же период, когда одна УМТ читает и генерирует данные?

Ответ: делает то же самое - читает и генерирует данные!

Тогда следующий вопрос: какие данные читает и генерирует каждая из УМТ в текущий период времени?

Ответ: в текущем периоде первая УМТ читает данные, сгенерированные и записанные второй УМТ в предыдущем периоде. Вторая УМТ читает данные, сгенерированные и записанные первой УМТ тоже в предыдущем периоде!

А также - в текущем периоде первая УМТ генерирует и записывает данные, которые будут прочтены второй УМТ в будущем периоде. Вторая УМТ генерирует и записывает данные, которые будут прочтены первой УМТ в будущем периоде.

Ещё вопрос: в текущем периоде обмен данными между двумя УМТ через общее поле памяти есть?

Ответ: да, есть!

Этот обмен представляет собой следующее: первая УМТ (Осознание поведения в Среде) формулирует на специальной ленте поля памяти вопросы.

Одновременно вторая УМТ (Поведение в Среде), действуя по своему поисковому плану, включает эффекторы и рецепторы связи со Средой, получает сообщения от Среды и записывает их на эту же специальную ленту поля памяти.

Если пришедшие данные есть ответы на вопросы (информация), то первая УМТ (Осознание) читает и воспринимает её как характеристику вариантов для принятия решений (в рамках осознания причин срыва предыдущего плана, или в рамках разрабатываемого ею плана будущего поведения).

Одновременно вторая УМТ (Поведение в Среде) использует эти же сообщения Среды, полученные от рецепторов, как стимулы для непосредственного включения соответствующих реакций эффекторов в Среде.

И вопросы, и ответы записываются на специальной ленте памяти в «текущий регистр» прямого доступа и в текущем периоде доступны для чтения непосредственно.

В следующем же периоде «текущий регистр» будет освобождён, данные будут сохранены на специальной ленте памяти (Бессознательное), и технология доступа к ним усложнится.

В «текущий регистр» будут записываться и читаться новые вопросы и ответы (стимулы и реакции) нового текущего периода.

Вопрос: Что же конкретно представляет собой предмет осознания в проявлении системы вообще?

Ответ: В данном исследовании определяется, что предметом осознания системы является только Среда и собственное поведение системы, отраженное в Среде, и только в

прошедший период, и воспринимаемые только как соотнесение плановых и фактических состояний, записанных на временной ленте памяти!

Собственно внутренние операции системы (чтение, генерация, запись данных, обмен вопрос-ответ, стимул-реакция), принципиально не становятся предметом осознания, поскольку они не планируются, а исполняются «инстинктивно»!

Среда и поведение системы в текущий период существования отражаются как факт на участке поля памяти, называемом «будущее». Поэтому в текущем периоде самой системой они также принципиально неосознаваемы, и смогут быть прочитаны только в следующих периодах осознания, когда участок памяти «будущее» перейдет в категорию «прошлое».

Оно (поведение) всегда осознаётся, лишь становясь прошлым, а текущее поведение будет новым предметом осознания «постфактум», лишь в предстоящем периоде.

Вопрос: Что же представляет собой средство для осознания системой её собственного поведения?

Ответ: В данном исследовании определяется, что средством для осознания (осуществляемого всегда и только в текущем периоде) системой её собственного поведения, является информация, т.е. только такие сообщения из Среды, которые расцениваются как ответ на вопрос:

- о причинах невыполнения «план-факт» поведения в прошедший период,
- о вариантах поведения для принятия решений о плане в будущем периоде.

Вопрос: Что представляет собой осознание как процесс?

Ответ: В соответствии с вводимым нами узким определением, осознание как процесс имеет две составляющие, внешнюю и внутреннюю:

- внешняя составляющая - поведенческий поиск и отбор персонажем сообщений Среды как ответов на вопросы, которые формулируются им в текущем периоде. Ответы также получаются в текущем периоде, и должны быть классифицированы по релевантности, оперативности, актуальности и достоверности (поток сообщений характеризуется большим «шумом» и большим разнообразием характеристик источников - временная задержка, достоверность и давность).
- внутренняя составляющая - восприятие персонажем полученных сообщений Среды как альтернативных вариантов для принятия решений:
- при выяснении причин невыполнения «план-факт» конкретного поведения системы в прошедший период,
- с целью разработки плана будущего поведения системы, устраняющего причину невыполнения «план-факт» прошлого её поведения.

Именно в этом, и пока ни в каком другом, смысле нами рассматривается осознание персонажем Среды и его собственного поведения в Среде, происходящее в текущем периоде!

Вспомним, что те же самые сообщения рецепторов воспринимаются УМТ (Поведение) как стимулы для запуска реакций эффекторов, функционирующих в Среде!

Однако этот процесс нами относится не к феномену Осознания, а к феномену Поведения.

Т.о., общее поле памяти представляет собой структуру, включающую следующее:

- Набор лент временной памяти (Осознаваемое); каждая из них разделена границей на отрезки «прошлое» (для чтения данных «план-факт») и «будущее» (для записи данных «план-факт»). Данные представляют собой имена, наименования, обозначения, даты и числа, которые связаны с создаваемыми рецепторами образами объектов Среды.
- Текущий регистр и вневременная лента памяти (Бессознательное) для чтения и записи данных (образов объектов Среды), которые воспринимаются как «запрос-ответ» при текущем принятии решений в новых планах, и как «стимул-реакция»

при текущем обеспечении условных переходов в алгоритмах поведения эффекторов и рецепторов.

Рефлексия есть... временной сдвиг Ленты Памяти, смена Планов поведения и Осознания, а также очередное включение функций Поведения и Осознания!

Поведение системы из двух УМТ и общего поля памяти заключается в следующем.

В конце каждого временного периода осуществляется горизонтальный СДВИГ ВЛЕВО на один сектор двух временных лент, над которыми стоят пишущие и читающие головки обеих УМТ.

Если при этом УМТ-Поведение, читая данные на ленте, констатирует неисполнения плана деятельности в прошедшем периоде, то она осуществляет вертикальный СДВИГ своих пишущих и читающих головок на другую ленту с готовым новым планом.

В свою очередь, УМТ-Осознание осуществляет вертикальный СДВИГ своих пишущих и читающих головок на ленту с неисполненным планом.

Если УМТ-Поведение констатирует исполнение плана в прошедшем периоде, вертикальных сдвигов не происходит, и процесс исполнения плана продолжается.

Описанный процесс представляет собой, по сути, «машинную технологию» рефлексии, реализуемую в системе двух УМТ, которая есть:

- периодический сдвиг влево временной ленты памяти конкретной деятельности под головками УМТ-Поведение,
- констатация УМТ-Поведение неисполнения плана деятельности,
- вертикальный сдвиг самих головок УМТ-Поведение на ленту памяти другой (поисковой) деятельности с готовым планом,
- одновременный вертикальный сдвиг головок УМТ-Осознание на ленту памяти с неисполненным планом.

Продолжим собственную расшифровку «принципа заимствования» [3] с учётом предыдущих соображений о связке двух УМТ.

В соответствии с предыдущим, термин «субъект содержит проект самого себя» необходимо должен включать два аспекта рассмотрения субъектом самого себя:

- Текущее осознание субъектом самого себя как результат прошлой реализации некоего проекта себя, разработанного и реализованного в прошедшем периоде времени. Осознание есть констатация факта исполнения или неисполнения плана реализации прошлого проекта и формулирование в связи с этим проблемы, которую нужно решать.
- Текущее представление субъекта о себе («образец самого себя») в будущем периоде времени, завершаемое разработкой плана устранения осознанной проблемы.
- Проявление же субъекта в текущий период времени есть ни что иное, как строительство себя по ранее выработанному образцу (т.е., текущее поведение по плану). И это проявление есть принципиально неосознаваемое поведение в текущий период времени, обеспечивающее информацией функцию Осознания.

Теперь совместим понятия:

Персонаж – это пара УМТ с общим полем памяти.

УМТ-1 – это функция «Осознание».

УМТ-2 – это функция «Поведение».

«Проект самого себя» - это План М будущего поведения персонажа.

«Реализация проекта самого себя» - это Факт прошлого исполнения Плана М-1 персонажа.

План и Факт, расположенные на одной ленте памяти, имеют отношение к одной конкретной деятельности персонажа.

Функция «Осознание» имеет в качестве области определения «прошлое» поля памяти, где она читает данные Факта исполнения Плана М-1. В качестве области значений она имеет «будущее» поля памяти, где она записывает как данные План М (Табл. 1).

Функция «Поведение» имеет в качестве области определения «прошлое» поля памяти, где она читает данные Плана N-1. В качестве области значений она имеет «будущее» поля памяти, где она записывает как данные Факт исполнения Плана N-1».

Таблица 1

Функция	Источник и Хранилище данных		Прошлое (область определения) Чтение		Будущее (область значений) Запись	
			Период 1	Период 2	Период 3	Период 4
Функция «Осознание»	Лента памяти 1	План	План M-1		План M	
		Факт		Факт M-1		
Функция «Поведение»	Лента памяти K	План		План N-1		
		Факт			Факт N-1	

Функция		Регистр Текущего Периода i
Функция «Осознание»	Лента «Бессознательное»	Формулирование и запоминание <i>Запроса</i> к Среде. Чтение и интерпретация <i>Ответа</i> из Среды (<i>Информация</i>)
Функция «Поведение»		Восприятие <i>Стимула</i> из Среды. Формирование <i>Реакции</i> в Среде.

Заключение

Из вышеизложенного получают следующие выводы:

- Самоприменимость УМТ как принципиальная модель «самоуправления» УМТ сохраняется, но «самопрограммирование» одной УМТ заменяется на «самопланирование» системы двух УМТ.
- Замена самопрограммирования на самопланирование означает принципиальную ликвидацию «зацикливания» УМТ. При самопрограммировании «зацикливание» происходит ввиду того, что передача управления осуществляется рекурсивно (т.е., не по внешнему прерыванию, а по внутренней логике одной и той же программы). При «самопланировании» зацикливание не происходит ввиду того, что передача управления от одной УМТ другой осуществляется принудительно, по внешнему прерыванию от другой УМТ или от операционной системы. Следовательно, в такой постановке снимается вопрос об алгоритмической невычислимости задачи самоприменимости УМТ.
- Значит, теоретически не закрывается дорога для реализации «машинного сознания» на принципе самоприменимости УМТ, правда, уже в рамках самопланирования системы двух УМТ на общем поле памяти.

Автор данного текста вполне отдаёт себе отчёт в том, сколь непроста задача - донести своё представление о «машинном мышлении» до читателя.

Слабым оправданием может быть лишь то, что автор вынужден применять мысль как инструмент для описания самого же используемого инструмента. А такой способ, как показал Гёдель, изначально чреват неполнотой.

К сожалению, автору другие рациональные способы донесения мысли о мысли неизвестны.

Правда, практика разработки автором компьютерного персонажа на вышеизложенных принципах показала высокую степень адекватности допущения о «псевдоразумности» поведения подобного автомата.

Литература

1. *Нейман Дж.* // Теория самовоспроизводящихся автоматов, М.: Мир, 1971.
2. *Отв. ред. Г.Г. Малинецкий* // Новое в синергетике. Новая реальность, новые проблемы, новое поколение, 2007.
3. *Лефевр В.А.* // Конфликтующие структуры, 1974.
4. *Полосухин Б.М.* // Тьюрингова модель сознания.
<http://www.trinitas.ru/rus/doc/0016/001b/00160120.htm>
5. *Гришин Е.А.* // Об интеллекте: просто о сложном (Реплики на выдержки из книги Д. Хокинса и С. Блейкли «Об интеллекте»). Журнал РАН «Искусственный интеллект и принятие решений», вып. 4, 2008).