

МАТЕМАТИЧЕСКИЕ МОДЕЛИ СТУПОРА И ПРИНЯТИЯ РЕШЕНИЯ РОБОТОМ

Черников К.В.

ФГБОУ ВПО «Пермский государственный национальный исследовательский университет»,
Пермь, e-mail: kirillperm@yandex.ru

Коротко рассмотрена и проанализирована проблема исследования человеческой привязанности в современной науке и, в частности, в современной робототехнике. Формально описаны основные псевдоэмоциональные характеристики робота, лежащие в основе предлагаемой модели привязанности робота: сюжет, функция внутренних переживаний робота, псевдоэмоция робота, элементарное псевдовоспитание и псевдовоспитание. Рассмотрена упрощенная модель псевдовоспитания, основанная на понятиях равномерно забывчивого робота и равноценных псевдоэмоциях. Описана общая модель принятия решения роботом в условиях привязанности, ассоциирующей с псевдовоспитанием робота. Рассмотрена ситуация ступора в рамках предложенной модели принятия решения роботом. Сформулировано общее условие наступления ступора. Показано, что для упрощенной модели псевдовоспитания наступление ступора можно спрогнозировать или даже избежать. Коротко рассмотрены преимущества и недостатки предложенной модели принятия решения роботом в условиях привязанности.

Ключевые слова: робот, привязанность, сюжет, псевдоэмоция, псевдовоспитание, принятие решений, ступор

MATHEMATICAL MODELS OF THE STUPOR AND DECISION-MAKING BY THE ROBOT

Chernikov K.V.

National Research Perm State University, Perm, e-mail: kirillperm@yandex.ru

The problem of research of human attachment in a modern science and, in particular, in a modern robotics is shortly considered and analyzed. The main pseudo-emotional characteristics of the robot underlying offered model of attachment of the robot are formally described: plot, function of internal experiences of the robot, pseudo-emotion of the robot, elementary pseudo-education and pseudo-education. The simplified model of pseudo-education based on concepts of evenly forgetful robot and equivalent pseudo-emotions is considered. The general model of decision-making robot in the conditions of the attachment associating with pseudo-education of the robot is described. The stupor situation within the offered model of decision-making by the robot is considered. The general condition of stupor is formulated. It is shown that for the simplified model of pseudo-education, stupor can be predicted or it is even possible to avoid it. Advantages and limitations of the offered model of decision-making robot in the conditions of attachment are shortly considered.

Keywords: robot, attachment, plot, pseudo-emotion, pseudo-education, decision-making, stupor

Почти каждый день в жизни любого человека возникает проблема выбора. Например, пойти или не пойти куда-либо, сделать или не сделать что-либо, сказать или промолчать и т.д. Вопросы такого характера, несомненно, присутствуют в нашей жизни. При этом каждый отвечает на них по-своему, руководствуясь различными фактами, имеющимися знаниями и опытом, и, что немаловажно, мнением других людей, которым доверяют или к которым привязаны. Примером влияния одного человека на принятие решения другим человеком может служить классическая ситуация взаимодействия любого родителя со своим ребенком, когда ребенок попросту не в состоянии принять решение и, как следствие, всецело опирается на мнение родителей.

Проблема осуществления выбора может быть неразрывно связана с проблемой ступора. Ситуация ступора может возникнуть тогда, когда ни одна из альтернатив еще не принята в качестве окончательного выбора. При этом у человека может появиться скованность или замедление мыслей, чувств

и движений, что можно считать негативными последствиями.

Безусловно, в некотором будущем в процессе развития современных технологий и науки неотъемлемой частью человеческого общества станут и человекоподобные роботы. При этом почти со стопроцентной долей уверенности можно говорить, что и у человекоподобных роботов по мере их совершенствования и усложнения их функционала возникнет проблема выбора в условиях социальной привязанности, характерная в настоящее время для всех людей.

Согласно современным определениям, привязанность – это чувство близости, основанное на глубокой симпатии, преданности кому или чему-либо; поиск и установление эмоциональной близости с другим человеком [2, 7].

Основоположником теории привязанности является известный английский психиатр и психоаналитик, специалист в области психологии развития, психологии семьи, психоанализа и психотерапии Джон Боулби [1]. В своей книге «Эмоциональная ма-

шина» [8] Марвин Минский анализирует работы Боулби и ряда других психологов и строит ряд абстрактных моделей того, как работает человеческий мозг, включая то, как привязанность может влиять на цели и действия людей. Несомненным преимуществом анализа, проведенного Минским, является максимальная общность предлагаемых моделей. В то же время Минский не дает каких-либо конкретных рекомендаций по применению данных моделей.

Современная робототехника стоит на пути моделирования привязанности роботов к их владельцам и определения поведения роботов в зависимости от этой привязанности. В этом направлении можно выделить следующие известные проекты: японский робот-пес Aibo, робот-тюлененок PARO, робот-динозаврик PLEO [3-6].

В настоящей статье предлагаются математическая модель привязанности робота, алгоритм принятия решения роботом в условиях привязанности, а также рассмотрена ситуация ступора в рамках предложенной модели. Модель и алгоритм строятся на основе следующих псевдоэмоциональных характеристик робота.

Псевдоэмоциональные характеристики робота и привязанность

Из теории психологии человека известно, что эмоции являются ответной реакцией организма на некий раздражитель-стимул. Для роботов этот стимул будем называть сюжетом и пусть он имеет следующее определение.

Пусть t – время.

Функцию $S(t)$ будем называть **сюжетом**, если она обладает следующими свойствами:

1. Область определения $S(t)$: $t \in [t_0, T_0]$, $0 < t_0 < T_0 < \infty$.
2. $0 < S(t) < \infty$ для любого $t \in [t_0, T_0]$.
3. $S(t)$ – непрерывная и взаимнооднозначная функция на $[t_0, T_0]$.

Функцию $f(t)$, удовлетворяющую соотношению $f(t) = a(S(t), t)S(t)$, где $a(S(t), t)$ – произвольная функция, назовем *функцией внутренних переживаний робота* (область определения $f(t)$, $a(S(t), t)$ и $S(t)$ совпадают).

Функцию внутренних переживаний робота $M(t)$ назовем *псевдоэмоцией* робота, если она удовлетворяет условиям:

1. Область определения $M(t)$: $t \in [t_0, T_0]$, $0 < t_0 < T_0 < \infty$.
2. $M(t)$ – дифференцируемая на (t_0, T_0) , непрерывная и однозначная функция на $[t_0, T_0]$.
3. $M(t_0) = 0$ и $M(T_0) = 0$.
4. В области определения существует единственная точка z , такая, что: $z \neq t_0$, $z \neq T_0$ и $\frac{dM(z)}{dt} = 0$.

Под *псевдовоспитанием робота* или просто *псевдовоспитанием* будем понимать относительно устойчивое отношение робота к сюжету.

Элементарным псевдовоспитанием робота $r(t)$ на сюжете $S(t)$ назовем функцию вида:

$$r(t) = \int_{t_0}^t a(S(\tau), \tau)S(\tau)d\tau = \int_{t_0}^t M(\tau)d\tau.$$

Будем считать, что с течением времени робот забывает псевдоэмоции, которые он когда-то испытывал. Прошлые псевдоэмоции все меньше и меньше сказываются на его текущем псевдовоспитании. А вместе с тем забываются и прошлые элементарные псевдовоспитание, порождаемое испытываемыми ранее псевдоэмоциями робота.

Исходя из этого введем следующее определение:

Псевдовоспитанием робота $R(t)$ (псевдовоспитание робота во время действия псевдоэмоции) на сюжете $S(t)$ назовем функцию вида:

$$R_{i+1}(t) = r_{i+1}(t) + \theta_{i+1}(t)R_i(t),$$

где t – текущее время, $t > t_i$, $0 \leq \theta_{i+1}(t) \leq 1$. Текущее время удовлетворяет соотношению $t = \tau + t_i$, где τ – время действия текущей псевдоэмоции от начала ее проявления, t_i – время функционирования (псевдовоспитания) робота до появления текущей псевдоэмоции, $R_i(t_i)$ – псевдовоспитание, полученное роботом ко времени t_i . Нижний индекс i означает, что до момента действия текущей псевдоэмоции робот уже испытал i псевдоэмоций. Кроме этого, будем подразумевать, что процесс псевдовоспитания является непрерывным в том смысле, что на смену одной псевдоэмоции приходит сразу же другая псевдоэмоция, порожденная очередным сюжетом с порядковым номером $i + 1$.

Коэффициент $\theta_i(t)$ будем называть *коэффициентом памяти прошлых событий* или *коэффициентом памяти*.

Под *псевдовоспитательным процессом* будем понимать процесс изменения значения функции *псевдовоспитания* робота. В начальный момент времени значение функции псевдовоспитания робота равно нулю.

Таким образом, будем считать, что величина функции псевдовоспитания $R_i(t)$ будет показывать привязанность робота к некоторому субъекту, который ассоциируется с воздействием на робота последовательностью сюжетов $S_1(t) \dots S_i(t)$. Если на робота оказывается воздействие со стороны n различных субъектов, каждый из которых характеризуется последовательностью раздражителей: $S_1^k(t) \dots S_{m_k}^k(t)$, где k – порядковый номер субъекта, m_k – количество сюже-

тов воздействия со стороны k -го субъекта, то величины функций псевдовоспитания $R_{m_1}^1(t) \dots R_{m_n}^n(t)$ показывают привязанность робота относительно каждого из воздействующих субъектов. В дальнейшем будем предполагать, что коэффициенты памяти робота могут быть различным в зависимости от каждого конкретного воздействующего сюжета.

Рассмотрим упрощение предложенной модели, для этого введем несколько определений.

Забывчивым роботом будем называть робота, для псевдовоспитательного процесса которого характерно:

$$\exists 0 < \partial < 1, \delta = \text{const} : 0 \leq \theta_i(t) < 1 - \partial.$$

При этом, если в момент окончания любой псевдоэмоции коэффициент памяти прошлых событий $\theta_i(t) = \theta$, то робота будем называть *равномерно забывчивым роботом*.

Псевдоэмоции $M_1(t), \dots, M_n(t)$, влекущие одинаковое элементарное псевдовоспитание q , то есть $r_1(T_0) = \dots = r_n(T_0) = q$, будем называть *равноценными псевдоэмоциями*. При этом будем говорить, что осуществляется процесс псевдовоспитания на равноценных псевдоэмоциях с элементарным псевдовоспитанием q .

Таким образом, учитывая вышесказанное, можно получить следующий результат.

Для равномерно забывчивого робота с коэффициентом памяти прошлых событий θ , псевдовоспитательный процесс которого осуществляется на равноценных псевдоэмоциях с элементарным псевдовоспитанием q , значение функции псевдовоспитания в момент окончания ощущения i -й псевдоэмоции вычисляется как

$$R_i = q \frac{1 - \theta^i}{1 - \theta}.$$

Принятие решения роботом в условиях социальной привязанности

Предположим, что перед роботом возникает проблема выбора, которая заключается в следующем: робот должен принять некоторое решение относительно своих дальнейших действий или поведения. При этом робот должен учитывать свои социальные связи и социальную привязанность к владельцу при принятии решения. Например, когда два человека дают роботу одновременно две взаимоисключающие команды, робот должен принять решение в пользу той или иной команды в зависимости от своей привязанности к каждому из людей-владельцев робота.

Пусть на робота действует n различных субъектов с последовательностями раздражителями-сюжетами: $S_1^k(t) \dots S_m^k(t)$. Таким образом, к некоторому моменту времени функционирования робота у него сложи-

лись псевдовоспитания $R_{m_1}^1(t) \dots R_{m_n}^n(t)$ относительно каждого из раздражителей, то есть относительно каждого из воздействующих субъектов.

Принимая $R_{m_1}^1(t) \dots R_{m_n}^n(t)$ в качестве меры привязанности робота относительно каждого из воздействующих субъектов, можем сформулировать простое правило принятия роботом решения в пользу того или иного субъекта: решение принимается в пользу того субъекта, относительно которого сформировалось максимальное псевдовоспитание. При этом возникает вопрос что делать роботу, если максимальное псевдовоспитание сформировалось сразу относительно нескольких субъектов? В какой последовательности реагировать на воздействия или быть может оставить некоторые воздействия попросту без реакции? Данную ситуацию будем называть ситуацией ступора.

Принятие решения роботом и ситуация ступора

Предположим, что из n действующих на робота субъектов, относительно p субъектов у робота в момент принятия решения сформировалось максимальное псевдовоспитание, то есть

$$R_{m_1}^1 = \dots = R_{m_p}^p = \max R.$$

Пусть робот является равномерно забывчивым относительно каждого из действующих субъектов с коэффициентом памяти прошлых событий $\theta_i, i=1, p$, и псевдовоспитательный процесс относительно каждого из воздействующих субъектов происходит на равноценных псевдоэмоциях с элементарным псевдовоспитанием q . Тогда ситуация ступора будет выглядеть следующим образом:

$$R_{m_1}^1 = \dots = R_{m_p}^p = \max R \Rightarrow$$

$$q \frac{1 - \theta_1^{m_1}}{1 - \theta_1} = \dots = q \frac{1 - \theta_p^{m_p}}{1 - \theta_p}.$$

Сократив предложенное соотношение на q , получаем условие ступора для равномерно забывчивого робота с псевдовоспитательным процессом на равноценных псевдоэмоциях:

$$\frac{1 - \theta_1^{m_1}}{1 - \theta_1} = \dots = \frac{1 - \theta_p^{m_p}}{1 - \theta_p}.$$

Как видно из условия ступора, независимо от коэффициентов памяти $\theta_i, i=1, p$ при $m_1 = \dots = m_p = 1$ условие ступора выполняется всегда. Данный случай является тривиальным, поэтому далее при рассмотрении

условия ступора будем предполагать, что m_1, \dots, m_p не равны единице одновременно.

Разумно предположить, что лучше, чем искать решения для преодоления ситуации ступора, необходимо искать возможность для того чтобы вовсе не попадать в данную ситуацию. Очевидно, что робот никогда не придет в ступор, если для $\forall m_i, i = 1, p$ и заданных $\theta_i, i = 1, p$ полученное условие ступора никогда не будет выполняться. Коэффициенты памяти прошлых событий, для которых при любых натуральных значениях $m_i, i = 1, p$, условие ступора не выполняется, назовем *антиступорными коэффициентами памяти*.

Покажем, что антиступорные коэффициенты памяти существуют. Рассмотрим ситуацию, когда ступор возникает относительно двух субъектов. Условие ступора в данном случае выглядит следующим образом:

$$\frac{1 - \theta_1^{m_1}}{1 - \theta_1} = \frac{1 - \theta_2^{m_2}}{1 - \theta_2}.$$

Пусть выполняется: $\theta_1 = \frac{1}{2}$ и $\theta_2 = \frac{1}{3}$. Таким образом, получим соотношение:

$$\left(1 - \left(\frac{1}{2}\right)^{m_1}\right) \left(1 - \frac{1}{3}\right) = \left(1 - \left(\frac{1}{3}\right)^{m_2}\right) \left(1 - \frac{1}{2}\right).$$

$$\text{Преобразовав, получаем: } 2^{m_1} = \frac{4 \cdot 3^{m_2}}{3^{m_2} + 3}.$$

Из полученного соотношения можно сделать вывод, что, при m_1 и m_2 , одновременно не равных единице, условие ступора никогда не выполняется. Докажем это.

Рассмотрим левую часть равенства. Возможные значения, которые будет принимать левая часть, при бесконечном увеличении m_1 – это: 2, 4, 8, 16, ... В то же время правая часть равенства при бесконечном увеличении m_2 , будет стремиться от 2 к 4, при этом никогда не достигнув 4. Таким образом, при m_1 и m_2 одновременно не равных единице, условие ступора никогда не выполнится.

Аналогичным образом можно показать, что для трех субъектов коэффициенты памяти $\theta_1 = \frac{1}{2}$, $\theta_2 = \frac{1}{3}$ и $\theta_3 = \frac{1}{5}$ являются также антиступорными.

Таким образом, зная коэффициенты памяти равномерно забывчивого робота, можно делать прогноз по поводу того, как возможно робот будет себя вести при принятии решений. Попадет ли он в ступор или такая ситуация в принципе невозможна.

Заключение

В статье представлена модель принятия решения роботом в условиях социальной привязанности, основанная на введенных

псевдоэмоциональных характеристиках. Модель является достаточно простой, но в то же время не исключающей общности, что является несомненным преимуществом. Для того чтобы применить модель в конкретной предметной ситуации, достаточно определить всего две функции: функцию сюжета и функцию псевдоэмоции, что является одновременно преимуществом и недостатком. Преимущество заключается в том, что этих функций всего две. С другой стороны, определить конкретный вид данных функций вне зависимости от предметной области не представляется возможным.

Кроме этого, рассмотрена ситуация ступора в рамках предложенной модели принятия решения роботом. Сформулировано условие наступления ступора. Показано, что для определенного класса роботов ситуацию ступора можно достаточно просто предугадать или попросту избежать.

Список литературы

1. Булби Дж. Привязанность. – М.: Гардарики, 2003. – 477 с.
2. Психологический словарь [Электронный ресурс]. – Режим доступа: http://mirslivarej.com/content_psi_privjazannost-1351.html (дата обращения: 01.07.2012).
3. Радужная Анастасия. Станет ли робот когда-нибудь другом человека [Электронный ресурс]? – Режим доступа: <http://robotor.ru/2011/05/24/robot-as-dog/> (дата обращения: 02.07.2012).
4. Робот PARO [Электронный ресурс]. – Режим доступа: <http://roboting.ru/954-robot-paro.html> (дата обращения: 02.07.2012).
5. Робот-динозавр PLEO [Электронный ресурс]. – Режим доступа: http://www.icasies.ru/buy_pleo.html (дата обращения: 02.07.2012).
6. Собака-робот AIBO [Электронный ресурс]. – Режим доступа: <http://www.prorobot.ru/04/robot-dog-aibo.php> (дата обращения: 02.07.2012).
7. Толковый словарь Ушакова [Электронный ресурс]. – Режим доступа: <http://ushakovdictionary.ru/> (дата обращения: 01.07.2012).
8. Marvin Minsky. The emotion machine. Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind. – New York: SIMON & SCHUSTER, 2006. – 372 p.

References

1. Boulbi Dzhon. *Privjazannost'* (Attachment). Moscow, Gardariki, 2003. 477 p.
2. *Psihologicheskij slovar'* (Psychological Dictionary), Available at: http://mirslivarej.com/content_psi_privjazannost-1351.html (accessed 1 July 2012).
3. Raduzhnaja Anastasija. *Stanet li robot kogda-nibud' drugom cheloveka* (Whether there is a robot sometime the friend of the person?), Available at: <http://robotor.ru/2011/05/24/robot-as-dog/> (accessed 2 July 2012).
4. *Robot PARO* (PARO robot), Available at: <http://roboting.ru/954-robot-paro.html> (accessed 2 July 2012).
5. *Robot-dinozavr PLEO* (Robot Dinosaur PLEO), Available at: http://www.icasies.ru/buy_pleo.html (accessed 2 July 2012).
6. *Sobaka robot AIBO* (AIBO robot dog), Available at: <http://www.prorobot.ru/04/robot-dog-aibo.php> (accessed 2 July 2012).
7. *Tolkovij slovar' Ushakova* (Ushakov's explanatory dictionary), Available at: <http://ushakovdictionary.ru/> (accessed 1 July 2012).
8. Marvin Minsky. *The emotion machine. Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York, SIMON & SCHUSTER, 2006. 372 p.

Рецензенты:

Пенский О.Г., д.т.н., доцент, профессор кафедры процессов управления и информационной безопасности, ФГБОУ ВПО «Пермский государственный национальный исследовательский университет», г. Пермь;

Тарунин Е.Л., д.ф.-м.н., профессор кафедры прикладной математики и информатики, ФГБОУ ВПО «Пермский государственный национальный исследовательский университет», г. Пермь.

Работа поступила в редакцию 11.01.2013.