РЕФЛЕКСИВНЫЕ ИГРЫ ЛЮДЕЙ И РОБОТОВ

С. С. Тарасенко

Kyoto University, Kyoto, Japan,
Ph. D. in Informatics

ВВЕДЕНИЕ

Роботы начинают играть существенную роль в нашей жизни. Одно из их употреблений – это замена человека в опасных ситуация, таких как, например, обезвреживание мин. Однако сам человек по своей природе склонен к риску, который может иметь трагические последствия. Диапазон их обширен: от детской безответственности до необходимости совершать рискованное действие в критической ситуации. Защищая человека, робот должен предотвращать совершение таких действий человеком и при необходимости взять на себя выполнения рискованного действия.

Однако робот не должен физически препятствовать совершению людьми рискованных действий. Он должен быть способен убедить человека добровольно воздержаться от их совершения. Подобный путь является более эффективным, чем грубое физическое воздействие, т. к. человек сохраняет свое достоинство и не теряет чувства ответственности. Таким образом, робот должен проводить *рефлексивное управление* [1].

Далее будут рассматриваться смешанные группы, состоящие из людей и роботов. Чтобы общаться с людьми, робот должен демонстрировать способность к «рассуждению» и принятию решений подобно человеку. Психологические принципы, объясняющие процессы принятия решений

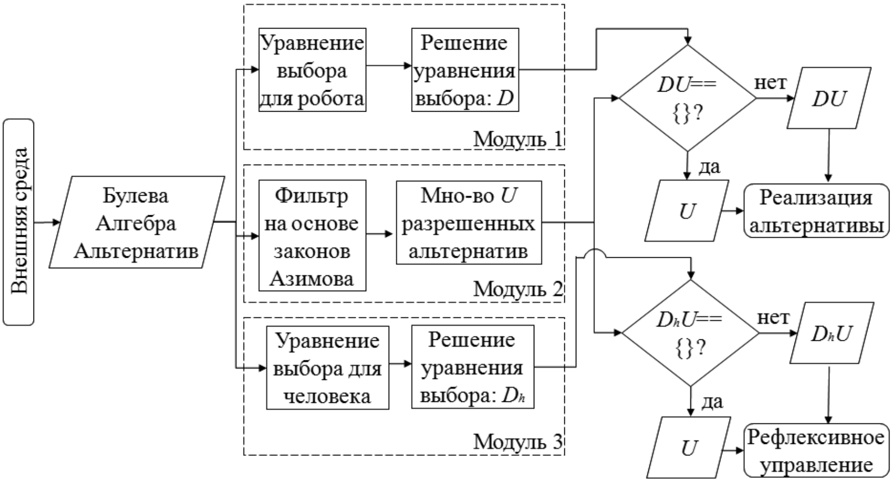

Рис. 1. Схематическое представление процессов принятия решений роботом

человека, заложены в основу теории рефлексивных игр (ТРИ), которая была предложена и разработана В. А. Лефевром [2]. Ее основы были рассмотрены ранее в [3, 4]. При помощи ТРИ становится возможным предсказывать выбор каждого субъекта, входящего в группу, и влиять на эти решения. Например, ТРИ была эффективно использована для моделирования поведения террористов [5].

Целью данной работы является приложение ТРИ для анализа индивидуального поведения каждого субъекта в смешанных группах людей и роботов. Далее будет показано, как ТРИ может быть использована роботами для убеждения людей воздерживаться от рискованных действий.

ФОРМАЛИЗАЦИЯ РОБОТОВ В РАМКАХ ТЕОРИИ РЕФЛЕКСИВНЫХ ИГР

Считается, что робот неукоснительно следует заложенной в него программе управления. Такая программа состоит как минимум из трех модулей. Первый модуль (Модуль 1) реализует возможности робота принимать решения аналогичные тем, которые бы принимал человек. Второй модуль (Модуль 2) содержит правила, которые не дают роботу нанести вред человеку. Третий модуль (Модуль 3) прогнозирует выбор каждого человека в группе и предоставляет возможные стратегии рефлексивного управления.

Основой для определения правил ненанесения вреда человеку являются три закона робототехники, предложенные Айзеком Азимова [6]:

1) робот не может причинить вред человеку своими действиями или своим бездействием допустить его нанесение;
2) робот должен выполнять все отданные человеком приказы, если они не противоречат положениям первого закона;
3) робот должен защищать свое существование до тех пор, пока такая защита не находится в конфликте с первым и вторым законами.

Алгоритмы, реализующие эти законы, являются неотъемлемой частью программы управления роботом (Модуль 2). Считается, что эта часть программы не может быть стерта или повреждена никакими средствами.

Взаимодействие Модулей 1, 2 и 3 в системе управления робота представлено на рисунке 1. Сначала информация об окружающей обстановке формализуется в форме булевой алгебры различных альтернатив. Затем система принятия решений, основанная на принципах принятия решений, используемых человеком, реализуется в Модулях 1 и 3. На выходе Модуля 1 получаем множество D, которое содержит альтернативы решения уравнения выбора [3, 4]. Параллельно, из булевой алгебры исключаются альтернативы, которые противоречат законам Азимов. Эта процедура фильтрации булевой алгебры выполняется в Модуле 2. На выходе Модуля 2 получаем множество U разрешенных альтернатив. Затем находим пересечением множеств D и $U : DU = D \cap U$. Если множество DU пусто ($DU = \{\}$), тогда робот выбирает одну альтернативу из множества U. Если множество DU не пусто, тогда робот выбирает одну из альтернатив.

Предполагается, что для достижения цели – убедить человека не совершать рискованных действий – робот прогнозирует выбор каждого человека в Модуле 3. Здесь множество D_h соответствует выходу Модуля 1 и представляет собой выбор конкретного человека в группе. Если робот прогнозирует, что при данных условия человек склонен выбрать рискованное действие, то робот перебирает возможные варианты, при которых человек не выберет содержащие рискованные действия альтернативы и применит соответствующее рефлексивное управление.

Важным отличием между результатами взаимодействия Модулей 1 и 2 и Модулей 3 и 2 является то, что взаимодействие первой пары дает возможные альтернативы, безопасные для людей в группе выбора робота, а результат взаимодействия второй пары есть набор возможных стратегий рефлексивного управления.

Рассмотрим, как роботы могут убеждать людей отказаться от выбора рискованных действий при использовании ТРИ с формализацией роботов на примере двух ситуаций.

АНАЛИЗ СМЕШАННЫХ ГРУПП ЛЮДЕЙ И РОБОТОВ

Рассмотрим два примера того, как роботы в смешанных группах могут убедить людей отказаться от совершения рискованных действий. В первом рассматриваются роботы-няни, во втором – взаимодействие альпинистов и робота-спасателя в критической ситуации в горах. Цель роботов в обоих примерах – склонить людей отказаться от выбора рискованных альтернатив.

РОБОТЫ-НЯНИ

Предположим, что роботы выполняют функцию нянь, присматривая за детьми. Рассмотрим группу из двух детей и двух роботов – каждый присматривает за одним ребенком. Закончив очередную игру, дети решают, что делать дальше. Множество действий детей состоит из четырех альтернатив:

1) первая альтернатива $\{\alpha\}$ – лезть на дерево;
2) вторая альтернатива $\{\beta\}$ – играть в мяч;
3) третья альтернатива $1 = \{\alpha, \beta\}$ – ребенок колеблется в выборе между «лезть на дерево» и «играть в мяч»;
4) четвертая альтернатива $0 = \{\}$ – сидеть неподвижно.

Предполагается, что каждый ребенок находится в отношении союза со своим роботом, а другого ребенка и его робота рассматривает как соперников (отношение конфликта). Каждый робот ведет себя как союзник ребенка, за которым присматривает, и как соперник пары другого ребенка и его робота.

Субъекты *a* и *c* – это дети, а субъекты *b* и *d* – роботы. Граф отношений представлен на рисунке 2.

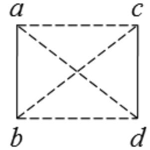

Рис.2. Граф отношений группы, состоящей из детей и роботов-нянь

Далее построим диагональную форму и проведем ее свертку:

$$\begin{array}{l} \qquad\qquad [a][b] \qquad\quad [c][\mathrm{d}] \\ \qquad\; [ab] \qquad + \quad [cd] \\ [ab + cd] \qquad\qquad\qquad\qquad = ab + cd. \end{array}$$

Отсюда следует уравнение выбора

$$x = ab + cd, \qquad (1)$$

где переменная x – это одна из переменных a, b, c или d.

Уравнения выбора в канонической форме для каждого субъекта и интервалы для решений этих уравнений представлены в таблице 1.

Таблица 1

Уравнения выбора и их решения для примера «роботы-няни»

Субъект	*Уравнение выбора*	*Интервал решений*
a	$a = (b + cd)a + cd\bar{a}$	$(b + cd) \supseteq a \supseteq cd$
b	$b = (a + cd)b + cd\bar{b}$	$(a + cd) \supseteq b \supseteq cd$
c	$c = (d + ab)c + ab\bar{c}$	$(d + ab) \supseteq c \supseteq ab$
d	$d = (c + ab)d + ab\bar{d}$	$(c + ab) \supseteq d \supseteq ab$

Из двух действий α и β, действие α является рискованным, так как, выполняя его, ребенок может упасть с дерева, что может угрожать не только здоровью ребенка, но и его жизни. Следовательно, в соответствии с законами Азимова, роботы не могут допустить, чтобы дети стали соревноваться в лазании по дереву. Поэтому роботы должны склонить детей не выбирать альтернативу $\{\alpha\}$. Оставшиеся альтернативы включаются в множество U разрешенных альтернатив. В данном случае множество U содержит две альтернативы $\{\beta\}$ и $0 = \{\}$.

Далее рассмотрим некоторые возможные сценарии.

Сценарий 1. Пусть каждый робот пытается склонить каждого ребенка выбрать альтернативу игры в мяч, т. е. $b = \{\beta\}$, $d = \{\beta\}$. В этом случае из (1) следуют уравнения (2):

$$\begin{aligned} a &= a\{\beta\} + c\{\beta\} \\ c &= a\{\beta\} + c\{\beta\} \end{aligned} \qquad (2)$$

Тогда возможны следующие решения для a.

Если $c\,\{\beta\} = 0$, то $a = a\,\{\beta\}$. Это уравнение имеет два решения 0 и $\{\beta\}$. Тогда ребенок a может выбрать либо пассивное состояние, либо игру в мяч. Вариант $c\,\{\beta\} = 0$ возможен, если c равно либо 0, либо $\{\alpha\}$. Таким образом, при влиянии $\{\beta\}$ обоих роботов и влиянии 0 либо $\{\alpha\}$ со стороны ребенка c на ребенка a, ребенок a может выбрать пассивный вариант отдыха или игру в мяч.

Если $c\,\{\beta\} \neq 0$, то $c\,\{\beta\} = \{\beta\}$. Тогда $a = a\,\{\beta\} + \{\beta\} = \{\beta\}$. Вариант $c\,\{\beta\} = \{\beta\}$ возможен, если c равно либо 1, либо $\{\beta\}$. Таким образом, при влиянии $\{\beta\}$ обоих роботов и влиянии 1 либо $\{\beta\}$ со стороны ребенка c на ребенка a, ребенок a выберет игру в мяч.

Аналогичным образом может быть показано, что если поменять ролями детей a и c, то ребенок c будет вести себя точно также как ребенок a.

Сценарий 2. Роботы воздействуют на детей, склоняя их сидеть неподвижно: $b = c = 0$. Из таблицы 1 следует, что в данном случае, $a = c = 0$, т. е. роботам удается усмирить детей.

В рассмотренном примере роботы смогли успешно контролировать поведение детей, удерживая их от выбора рискованной альтернативы.

АЛЬПИНИСТЫ И РОБОТ-СПАСАТЕЛЬ

Рассмотрим ситуацию в горах, участниками которой являются два альпиниста и робот-спасатель. Альпинисты и робот общаются по радио. Один из альпинистов (субъект b) попал в критическую ситуацию и ему требуется помощь. Например, он сорвался с карниза и упал в неглубокую расщелину. При этом он остался жив и не получил серьезной травмы, однако технически он не может сам выбраться.

В этой ситуации у участников два варианта действий. Первый вариант – альпинист a может попытаться сам спасти товарища (действие α), что, несомненно, является рискованным действием, поскольку первый альпинист сорвался на этом участке склона, а значит, второй альпинист так же может сорваться. Второй вариант – это возложить спасательную операцию на робота-спасателя (действие β). В этом случае, множество U разрешенных альтернатив для робота состоит только из альтернативы $\{\beta\}$, поскольку бездействие в условиях настоятельной необходимости выполнения спасательной операции не является приемлемым в соответствии с первым законом Азимова. Целью робота является убедить альпиниста a воздержаться от попытки спасти товарища самостоятельно. Поэтому робот должен сам выполнить спасательную операцию.

Пусть в начале, все субъекты находятся в отношении союза. Граф отношений представлен на рисунке 3а. Ему соответствует полином abc.

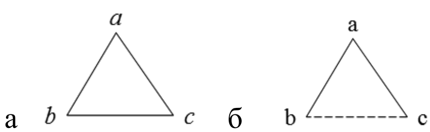

Рис. 3. Союзная группа двух альпинистов и робота-спасателя

Ситуации соответствует диагональная форма и ее свертка (3):

$$[abc]^{[a][b][c]} = [abc] + \overline{[a][b][c]} = 1 \qquad (3)$$

Любой субъект в группе выберет активную альтернативу $1 = \{\alpha, \beta\}$ и может реализовать либо действие α, либо действие β – вне зависимости от влияний на него со стороны других субъектов. Поэтому такая группа является неуправляемой в смысле независимости выбора субъектов от внешнего влияния. Следовательно, роботу необходимо преобразовать неуправляемую группу в управляемую. Для этого он принимает решение изменить отношение между собой и альпинистом *b* с союза на конфликт. Робот может это сделать, например, перестав отвечать на приказы этого альпиниста. Тогда граф отношений трасформируется в граф изображенный на рисунке 3б. Решения субъектов определяются следующими уравнениями выбора (4):

$$\begin{aligned} a &= (b+c)\,a + \bar{a}\,; \\ b &= b + (c+\bar{a})\,\bar{b}\,; \\ c &= c + (b+\bar{a})\,\bar{c}\,. \end{aligned} \qquad (4)$$

Целью робота является не дать альпинисту *a* выбрать альтернативу {α}. Выбор альпиниста *a* определяется интервалом $(b + c) \supseteq a \supseteq 1$. Альпинист *a* может совершить выбор только при условии, что $(b + c) = 1$. При этом выбор будет единственным – активная альтернатива $1 = \{\alpha, \beta\}$. Поэтому далее альпинист *a* может реализовать как действие α, так и действие β. Данный исход не является приемлемым, так как существует возможность, что альпинист *a* реализует действие α.

С другой стороны, если $(b + c) \subset 1$, то альпинист *a* оказывается в состоянии фрустрации и не способен совершить никого выбора, в том числе и выбрать альтернативу {α}.

Таким образом, единственной возможностью для робота воспрепятствовать альпинисту *a* выбрать альтернативу {α} – это сделать так, чтобы альпинист *a* оказался в состоянии фрустрации.

Рассмотрим различные варианты давлений альпиниста *b* на альпиниста *a*.

Пусть альпинист *b* оказывает влияние на альпиниста *a,* чтобы тот предоставил выполнение спасательной операции роботу (альтернатива {β}). В этом случае, чтобы альпинист *a* оказался в состоянии фрустрации, робот должен оказывать на альпиниста *a* также влияние {β}. Тогда $(b + c) = (\{\beta\} + \{\beta\}) = \{\beta\} \subset 1$, и альпинист *a* не может совершить выбор.

Теперь пусть альпинист *b* оказывает влияние {α} на альпиниста *a*. Чтобы альпинист *a* оказался в состоянии фрустрации, робот должен оказывать на альпиниста *a* также влияние {α}. Тогда $(b + c) = (\{\alpha\} + \{\alpha\}) = \{\alpha\} \subset 1$, и альпинист *a* не может совершить выбор.

Аналогичная ситуация возникает, если альпинист *b* оказывает влияние $0 = \{\}$.

Вывод таков: если альпинист *b* не оказывает влияние $1 = \{\alpha, \beta\}$ на альпиниста *a,* то для того, чтобы не дать возможность альпинисту *a* выбрать $1 = \{\alpha, \beta\}$, робот должен оказывать на него такое же давление, как и альпинист *b.*

Далее покажем, что независимо от влияния альпинистов на робота, сам робот способен реализовать альтернативу $\{\beta\}$ и выполнить спасательную операцию.

В данном случае возможны 16 пар (*a, b*) совместных влияний альпинистов на робота. Выбрав из них четыре, рассмотрим соответствующие сценарии.

Сценарий 1. Альпинисты *a* и *b* оказывают на робота влияния $\{\alpha\}$ и $\{\beta\}$, соответственно. Тогда для робота *c,* $a = \{\alpha\}$, $b = \{\beta\}$: $1 \supseteq c \supseteq \{\beta\} + \overline{\{\alpha\}} \Rightarrow 1 \supseteq c \supseteq \{\beta\}$. Поэтому робот может выбрать любую альтернативу, включающую $\{\beta\}$. В данном случае $D = \{\{\alpha, \beta\}, \{\beta\}\}$ и $U = \{\{\beta\}\}$, поэтому $DU = \{\{\beta\}\}$. Следовательно, робот выберет альтернативу $\{\beta\}$.

Сценарий 2. Альпинисты *a* и *b* оказывают на робота влияния $\{\beta\}$ и $\{\alpha\}$, соответственно. Тогда для робота *c,* $a = \{\beta\}$, $b = \{\alpha\}$: $1 \supseteq c \supseteq \{\alpha\} + \overline{\{\beta\}} \Rightarrow 1 \supseteq c \supseteq \{\alpha\}$. Значит, робот может выбрать любую альтернативу, включающую $\{\alpha\}$: $D = \{\{\alpha, \beta\}, \{\alpha\}\}$ и $U = \{\{\beta\}\}$, поэтому $DU = \{\}$. Следовательно, в соответствии со схемой взаимодействия Модулей 1 и 2 на рисунке 1, робот выберет альтернативу $\{\beta\}$.

Сценарий 3. Оба альпиниста *a* и *b* оказывают на робота влияние $\{\alpha\}$. Тогда $a = b = \{\alpha\}$: $1 \supseteq c \supseteq \{\alpha\} + \overline{\{\alpha\}} \Rightarrow 1 \supseteq c \supseteq \{\alpha\} + \{\beta\} \Rightarrow 1 \supseteq c \supseteq 1 \Rightarrow c = 1$. В этом случае $D = \{\{\alpha, \beta\}\}$, $U = \{\{\beta\}\}$ и $DU = \{\}$. Следовательно, робот выберет альтернативу $\{\beta\}$.

Сценарий 4. Альпинисты *a* и *b* оказывают на робота влияние $\{\beta\}$. Тогда $a = b = \{\beta\}$: $1 \supseteq c \supseteq \{\beta\} + \overline{\{\beta\}} \Rightarrow 1 \supseteq c \supseteq \{\beta\} + \{\alpha\} \Rightarrow 1 \supseteq c \supseteq 1 \Rightarrow c = 1$. В этом случае робот также выберет альтернативу $\{\beta\}$.

Рассмотренный пример иллюстрирует, как робот может превратить неуправляемую группу в управляемую путем манипулирования отношениями между членами группы. В управляемой группе, оказывая соответствующие влияния, робот может добиться ситуации, когда альпинист *a* окажется в состоянии фрустрации и будет неспособен выбрать альтернативу $\{\alpha\}$.

Таким образом, робот оградит альпиниста *a* от риска (правда ценой фрустрации), который возникает, если альпинист *a* решит сам спасать товарища. Следовательно, робот достигает своей цели. В то же время множество *U* разрешенных альтернатив гарантирует, что робот сам выполнит спасательную операцию и тем самым выполняются законы Азимова.

ОБСУЖДЕНИЕ

В работе показано, как теория рефлексивных игр Владимира Лефевра и законы робототехники Айзека Азимова могут быть использованы для реализации схем рефлексивного управления людьми со стороны роботов с целью удержать их от необоснованного риска.

Специфика данного подхода состоит в том, что люди под воздействием роботов сами добровольно принимают решение отказаться от риска и считают эти решения своими собственными. Таким образом, вместе *теория рефлексивных игр* и *законы робототехники* выполняют функцию социального буфера, который в будущем позволит сделать жизнь людей более безопасной – и в то же время не будет оказывает интенсивного психологического давления на сознание людей в группе.

Применение нового подхода было рассмотрено в двух модельных ситуациях.

В первом случае рассмотрена ситуация, в которой роботы смогли успешно исключить из рассмотрения рискованные альтернативы и смогли убедить детей не совершать рискованные действия. В данной ситуации множество разрешенных альтернатив, полученное из булевой алгебры альтернатив при помощи «фильтрации» законами Азимова, содержит как активную альтернативу совершения определенного действия (игра в мяч), так и пассивную (пустое множество) – не совершать никаких действий. Во втором случае рассмотрена критическая ситуация, в которой пассивная альтернатива не может быть решением. Кроме того для эффективного достижения своей цели роботу приходиться решать задачу изменения структуры группы.

В примере с роботами-нянями ситуация требует менее интенсивного использования возможностей системы принятия решений, заложенной в роботов, в то время как во втором случае требуется использование более сложных логических конструкций. При этом изначально не существует способа осуществления рефлексивного управления, поэтому сначала обеспечивается сама возможность такого управления. Кроме того в первой ситуации для роботов нет необходимости совершения каких-либо действий вместо людей, в то время как во второй ситуации робот принимает на себя ответственность за выполнение спасательной операции.

Первая ситуация играет роль ознакомительного примера, в то время как второй пример более приближен к реальной критической ситуации. Вместе оба примера иллюстрируют широкие возможности и вычислительную силу симбиоза теории рефлексивных игр Владимира Лефевра и законов роботехники Айзека Азимова для решения широкого спектра задач, начиная от тех которые могут возникнуть на детской площадке до реальных критических ситуаций во взрослой жизни.

ЛИТЕРАТУРА

1 *Лефевр В.А.* (1965). Исходные идеи логики рефлексивных игр. Проблемы исследования системы и структур. Москва: Изд-во АН СССР.
2 *Lefebvre V.A.* (1982). Algebra of Conscience. D. Reidel, Holland.
3 *Лефевр В.А.* (2007). Рефлексивный агент в группе. Рефлексивные процессы и управление, № 1, т. 7, с. 102–116.
4 *Лефевр В.А.* (2009). Лекции по теории рефлексивных игр. Москва: Когито-Центр.
5 *Lefebvre V.A.* (2009). Reflexive analysis of groups // Argamon S., Howard N. (Eds). Computational methods for counterterrorism, p. 173–210, Springer.
6 *Asimov A.* (1942). Runaround. Astounding Science Fiction, March, p. 94–103.