

# Коллективная модель мира мобильного робота. Языковой анализ сцены и управление пешеходным переходом

**Харламов А.А.,**

*доктор технических наук старший научный сотрудник Института высшей нервной деятельности и нейрофизиологии РАН, Москва, профессор Департамента программной инженерии НИУ ВШЭ, Москва, профессор Кафедры прикладной и экспериментальной лингвистики МГЛУ, Москва*

**Жаркой Р.М.,**

*Президент ISS («Интеллектуальные системы безопасности»)*

**Арзуманов В.В.**

*Научный сотрудник ISS («Интеллектуальные системы безопасности»)*

## Аннотация

В работе представлена оригинальная архитектура модели мира мобильного робота. И ее сенсорные входы и принятие решения вынесены за пределы собственно мобильного робота. Модель мира (модель предметной области) обеспечивает процесс понимания, без которого невозможна никакая ответная реакция. Она состоит из двух взаимосвязанных частей: языкового и многомодального компонентов. В работе рассматривается задача управления пешеходным переходом с точки зрения предотвращения критических ситуаций (столкновений) участников движения. Участников движения можно в этом случае рассматривать как мобильные роботы, движение которых моделируется с использованием коллективной модели мира. Языковой компонент модели мира представлен шаблонами предложений естественного языка, описывающих стандартные ситуации движения. Система продукционных правил представляет многомодальный компонент модели мира. Модель системы управления пешеходным переходом была реализована на основе языка Clips. Средство визуализации модельной реализации сценариев представляет собой презентацию Powerpoint с программой-макросом и элементами управления ActiveX. Эксперименты проводились с целью выявления областей параметров, соответствующих стандартным и нестандартным ситуациям.

**Ключевые слова:** коллективная модель мира, языковой и многомодальный компоненты модели мира, управление движением на пешеходном переходе, модель системы управления на Clips.

## ВВЕДЕНИЕ

В работе использована оригинальная архитектура модели мира мобильного робота. И сенсорные входы модели мира и принятие решения в ней вынесены за пределы собственно мобильного робота. Сенсоры — это стационарные камеры, которые установлены в местах, наиболее интересных с точки зрения возможного взаимодействия различных участников движения [1]. Будучи реализованной в серверном варианте, такая модель мира становится как бы общей для всех участников движения.

Модель мира (модель предметной области) является необходимым элементом любой системы, принимающей решения. Она обеспечивает процесс понимания, без которого невозможна никакая ответная реакция, от управления движением до синтеза текста [2]. Модель мира может быть представлена в разных формах. В настоящее время наиболее привычными формами представления знаний о мире являются семантические сети [3] и онтологии [4]. Представление модели мира может быть и более традиционным, например, в виде системы продукционных правил, и даже в виде текстового описания.

Решая задачи, подобные созданию модели мира, полезно посмотреть, как эта задача решается человеком. Модель мира человека строится как две взаимосвязанные части: языковой и многомодальный компоненты [3].

Мы рассмотрим задачу управления пешеходным переходом с точки зрения предотвращения критических ситуаций (столкновений) участников движения. Участников движения можно в данном случае рассматривать как мобильные роботы, движение которых моделируется с использованием коллективной модели мира. Языковой компонент модели мира представлен шаблонами предложений естественного языка, описывающих стандартные ситуации дорожного движения. Система продукционных правил представляет многомодальный компонент модели мира.

Модель системы управления пешеходным переходом была реализована на основе языка Clips. Средство визуализации модельных сценариев представляет собой презентацию Powerpoint с программой-макросом и элементами управления ActiveX. Эксперименты проводились с целью выявления областей параметров, соответствующих стандартным и нестандартным ситуациям.

## ЗАДАЧА УПРАВЛЕНИЯ ПЕШЕХОДНЫМ ПЕРЕХОДОМ

Задача автоматизации управления дорожным движением естественным образом вырастает на фундаменте отображения разнообразных ситуаций дорожного движения, начиная с трекинга всех субъектов движения, с последующим переходом к отдельным видам автоматизации учета представленных в видеоряде (треке) ситуаций. И в конечном итоге может быть переведена в разряд автоматически решаемых задач путем использования языкового описания отдельных участников движения и их действий.

Автоматическое управление пешеходным переходом является стереотипной задачей управления дорожным движением, в которой представлены движущиеся транспортные средства и пешеходы. На ней можно представить практически все задачи управления дорожным движением.

Для решения этой задачи она разбивается на ряд подзадач, решение каждой из которых естественным образом вытекает из постановки предыдущей задачи. Она включает в свой состав: (1) сенсорную часть, представляющую сцену (видеокадр) как совокупность объектов с различными значениями параметров; (2) языковое описание сцены, которое кодирует сцену в терминах естественно-языковых шаблонов; (3) переописание сцены в терминах системы продукционных правил. Добавление к системе решения служебной задачи (4) моделирования всего множества возможных комбинаций параметров описания объектов на сцене позволяет провести неумозрительные эксперименты, выявляющие границы между комбинациями состояний объектов на сцене от допустимых до критических.

Полученная в результате моделирования информационно-аналитическая система годится для последующего воплощения в виде реальной системы управления движением на пешеходном переходе, которая позволит учитывать параметры (состояния) участников движения в пределах некоторой зоны дороги, включающей переход, предсказывать развитие событий и влиять на это развитие с целью предотвращения критических ситуаций (дорожно-транспортных происшествий).

### Архитектура коллективной модели мира

Для управления движением на пешеходном переходе используется модель пешеходного перехода (модель предметной области) как часть модели мира, например, в пределах одного города. Если ранее такая модель мира предполагала быть сформированной у каждого участника движения (каждого движущегося средства), в нашей постановке она формируется как общая для всех участников движения в пределах одной акцентированной зоны (в данном случае — конкретного пешеходного перехода).

Такой подход позволяет стандартизовать все участвующие в движении объекты по всем техническим средствам управления движением, так как они (средства) являются общими для всех участников, чего невозможно добиться в случае использования индивидуальных искусственно-интеллектуальных средств (принцип сложности В.В. Солодовникова гласит, что наихудшее звено является решающим в системе). Позволяет иметь преемственность результатов анализа различных акцентных зон (учитывать историю каждого отдельного участника движения в пределах разных локальных зон). И таким образом реализовать интегральную модель мира в рамках всей управляемой зоны (например, целого города). Это не исключает использования индивидуальных управляющих искусственно-интеллектуальных средств отдельных участников движения.

Архитектура коллективной модели мира включает в свой состав: (1) сенсорные средства — например, стационарные видеокamеры; (2) локальные модели предметных областей акцентных единиц, реализованные на локальных вычислительных средствах; (3) сервер верхнего уровня, осуществляющий интегральный мониторинг всей ситуации в рамках всей управляемой зоны (в данной работе не рассматривается). Локальные модели предметных областей реализуют два компонента модели мира:

(1) многомодальный компонент, представленный системой продукционных правил, позволяющий управлять ситуацией в пределах акцентной единицы (в данном случае — пешеходным переходом); и (2) языковой компонент, который, с одной стороны, позволяет перейти к модели многомодального компонента, кодируя в терминах шаблонов естественного языка сенсорный слепок внешнего мира, а с другой, — является средством общения с оператором этой локальной акцентной единицы.

### Сенсорное описание сцены

Сцена на пешеходном переходе фиксируется (одной или несколькими) камерой, и описывается в терминах выходов первичных сенсоров, набор которых может быть различным, например, таким (в случае трекинга): (1) детектор бега; (2) детектор оставленных предметов; (3) детектор праздничества; (4) детектор проникновения в запрещенную зону; (5) детектор скопления людей; (6) счетчик объектов [1]. Список акцентных зон в таком представлении: площади, парки, улицы города, вокзалы, аэропорты, административные здания, территории возле памятников и мемориалов, территории расположения специальных объектов, контроля входа-выхода в здании.

### Языковой анализ сцены

Оператор, наблюдающий сцену на экране монитора, видит последовательность видеок кадров (см. рис. 1), которая может быть описана естественно-языковым текстом, например, так: «Камера фиксирует происходящее у арки входа в Кремлевский сад в Вологде в зимний день. Ближе ко второй колонне слева находится мужчина, у которого в правой руке предмет. У левого края арки виден рюкзак. Его содержимое неизвестно. Судя по поведению мужчины, он кого-то (или чего-то) ждет».



Рис. 1. Кадр видеоряда, представляемого на мониторе оператора.

Это описание может быть записано в виде последовательности естественно-языковых предложений, которые описывают происходящее на сцене более формально.

00:00 — 00:04 — Мужчина движется от правого края арки к левому. Мужчина останавливается.

00:04 — 00:05 — Мужчина начинает движение. Мужчина движется от левого края арки к правому.

Такое представление предполагает классификацию ситуаций и формирование системы естественно-языковых шаблонов для описания классов ситуаций. Шаблонов формируется достаточно много, и их можно представить на примере нескольких, описывающих разные типы ситуаций: «Появление объекта», «Движение объекта», «С учетом зонирования», «С учетом движения других объектов».

Появление / исчезновение («4 координаты» / с применением зонирования)

Человек 1 появляется слева.

Человек 1 появляется у камеры.

Человек 1 появляется в зоне «Тротуар 1».

Движение: через координаты

Человек 1 движется вправо.

Человек 1 движется влево.

Движение: с применением зонирования

Человек 1 движется к Пути 1.

Человек 1 пересекает границу Запрещенной Зоны 1.

Человек 1 движется по Тротуару 1.

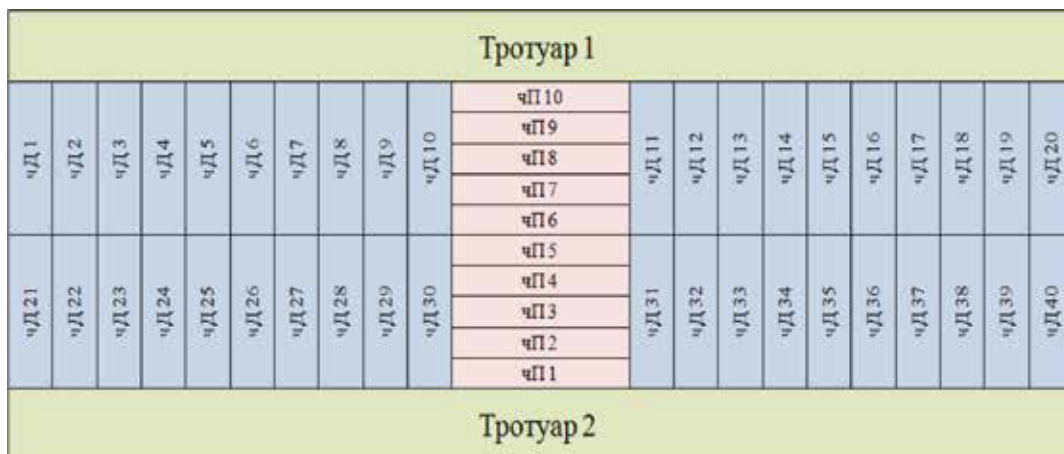


Рис. 2. Модель пространства реализации типового сценария (чП = «часть перехода», чД = «часть дороги»).

Тогда последовательность видеокладов, полученная с видеокамеры, описывается следующим образом (зона «Сквер»):

00:00 — 00:01 — Человек 1 появляется вдали. Человек 1 движется влево к камере. Человек 2 появляется вдали. Человек 2 движется влево к камере.

00:01 — 00:02 — Человек 1 движется влево к камере. Человек 2 движется влево к камере. Человек 1 поворачивает. Человек 2 поворачивает. Человек 3 появляется слева в Сквере. Человек 3 движется влево.

Для удобства квантования времени вводится зонирование отдельных частей перехода. В такой постановке семантика перехода (см. рис. 2) описывается следующим образом.

Таблица 1

Семантика пешеходного перехода с учетом зонирования

ЕЯ-шаблоны	Сенсорика + правила
(1) 00:00:00 — 00:00:02 — Автомобиль 1 равномерно движется вправо [↑] / влево [↓]	<p>Детектор ведет движущийся объект типа «Автомобиль»</p> <p>Рассматриваем два кванта времени:</p> <p><math>t_1 \{00:00:00 — 00:00:01\}</math> — объект проходит <math>\Delta S_1</math></p> <p><math>t_2 \{00:00:01 — 00:00:02\}</math> — объект проходит <math>\Delta S_2</math></p> <p><math>\Delta S_1 = \Delta S_2 \neq 0</math>,</p> <p>где <math>(\Delta S)</math> — модуль разности <u>номеров зон</u> — то есть, скорость</p>

### Система продукционных правил управления пешеходным переходом

Далее в соответствии с этим множеством шаблонов в среде Clips формируется система продукционных правил, описывающая формально ситуации на переходе. Пример представляет часть продукционных правил для случая пешеходного перехода через двухполосную дорогу с двухсторонним движением, один пешеход, одно движущееся транспортное средство.

Осуществляется «разбор сцены», представляющий собой обработку информации в зрительной модальности (в данном случае это видеоряд). В рамках этого процесса работают определенные сенсоры, предоставляющие некоторые данные об объектах сцены.

Несмотря на то, что вопрос входной информации в рамках постановки не является первостепенным, хотелось бы дать некоторые определения. Краткое рассмотрение уже созданных продукционных правил показывает, что ключевыми категориями для необходимого анализа являются объекты «Автомобиль» и «Человек», рассмотрение взаимодействия между которыми порождает соответствующие правила, в которых динамика ситуации, а также ограничения, накладываемые на взаимодействие объектов на сцене (с точки зрения корректности этого взаимодействия) включают следующие понятия:

- зоны (через них определяется скорость, а также «разрешения» и «запреты» — зонирование вводится с целью последующей интерпретации);

- время (через него определяется порядок следования событий — одновременность / предшествование / следование).

В качестве примера рассмотрим фрагмент комплекса продукционных правил по нарушениям Правил дорожного движения и опасным ситуациям на дороге (см. ниже).

Рассматриваются 2 движущихся объекта: типа «**Автомобиль**» и типа «**Человек**».

Если объект типа «**Человек**» остановился в зоне «Зона ожидания 1» [ $\uparrow$ ] / «Зона ожидания 2» [ $\downarrow$ ] (см. условия шаблона 4), и объект типа «**Автомобиль**» движется к переходу (см. условия шаблона 8), а потом — движется от перехода (см. условия шаблона 9), и между двумя последними не фиксируется остановка (см. условия шаблона 4), то фиксировать нарушение (автомобиль не предоставил преимущество / возможность перейти дорогу).

Рассматривается движущийся объект типа «**Автомобиль**». Допущение: зная о том, как расквантовано пространство — т. е. каковы примерные «размеры» зон, — мы можем установить некий предел —  $\Delta S_{max}$ ; это предел по скорости, который запрещено превышать.

Если в каком-либо кванте времени объект типа «**Автомобиль**» проходит  $\Delta S > \Delta S_{max}$ , то фиксировать нарушение (превышение скорости).

И т.д.

Тогда им соответствующие естественно-языковые шаблоны имеют вид:

Детектор ведет движущийся объект типа «Автомобиль»

Рассматриваем два кванта времени:

$t1 \{00:00:00 - 00:00:01\}$  — объект проходит  $\Delta S_1$

$t2 \{00:00:01 - 00:00:02\}$  — объект проходит  $\Delta S_2$

$\Delta S_1 = \Delta S_2 \neq 0$

Детектор ведет движущийся объект типа «Автомобиль»

Рассматриваем два кванта времени:

$t1 \{00:00:00 - 00:00:01\}$  — объект проходит  $\Delta S_1$

$t2 \{00:00:01 - 00:00:02\}$  — объект проходит  $\Delta S_2$

$\Delta S_1 < \Delta S_2 \cap \Delta S_1 \neq 0 \cap \Delta S_2 \neq 0$

И далее.

### Модель системы управления пешеходным переходом. Эксперименты на модели

Представленные выше соображения были воплощены в виде несложной модели, включающей в свой состав: (1) языковую модель мира, описывающую сцену в терминах предложений естественного языка (естественно-языковых шаблонов); (2) много-модальную модель мира, представляющую мир в виде системы продукционных правил; и (3) аналог видеомонитора — презентация в среде PowerPoint, которая представляет некую модельную ситуацию (параметры задаются извне) на пешеходном переходе (см. рис. 2).



Меняя параметры ситуации можно простым перебором выявить ограничения на систему параметров, которая соответствует критической ситуации («неуступление преимущества пешеходу»).

Средство визуализации PowerPoint включает в себя несколько элементов:

- модель пространства реализации типового сценария (см. рис. 2), включающая в свой состав: (1) разбитый на зоны пешеходный переход; (2) модели появляющихся объектов — автомобили (представлены как желтые прямоугольники), пешеходы (представлены как красные круги);
- элементы управления, в том числе: (1) модуль выбора типового сценария и его инициализация (установка параметров); (2) модуль запуска — пошаговый и непрерывный режим; и (3) модуль управления общими параметрами.

Модуль управления общими параметрами позволяет изменять направления движения на дорожных полосах (по умолчанию движение правостороннее). При выборе сценария и его инициализации становятся доступными режимы непрерывной (кнопка «Непрерывно») и пошаговой (кнопка «Шаг») реализации. По запуску объекты типа «Человек» начинают двигаться от исходно заданных точек к пешеходному переходу, переходят дорогу и затем идут по тротуару вправо или влево вдоль дороги (выбор случаен), а объекты типа «Автомобиль» движутся по направлению к противоположному краю дороги, пересекая переход. При непрерывной реализации ожидается достижение всеми участвующими в сценарии объектами типа «Человек» краев противоположных тротуаров и объектами типа «Автомобиль» противоположных краев дороги (благоприятный исход) или столкновение (неблагоприятный исход), при котором выдается соответствующее сообщение. При пошаговой реализации окончание сценария соответствует тем же правилам, но картинка обновляется не автоматически, а по требованию. По окончании реализации сценария участвующие объекты удаляются из памяти макроса.

Скорости и ускорения объектов измеряются как количество пересекаемых зон на шаг (единицу времени) и указываются как значения соответствующих переменных в коде макроса презентации, обеспечивающего реализацию сценария.

Макрос PowerPoint записывает информацию об объектах и их характеристиках в лог-файл, который на данном этапе представляется связующим звеном между анализатором, использующим правила, описанные на естественном языке, и рассматриваемой сценой. В рамках модельного представления формат данных обладает определенной степенью условности. В реальной системе информация, подаваемая на вход — это информация, снимаемая сенсорной системой. Для анализатора численные значения, записываемые в лог-файл — это совокупность характеристик положения. На их основе и на основе правил на естественном языке выдается сообщение об опасности, в случае если срабатывает соответствующее правило.



### Обсуждение результатов моделирования

Результаты моделирования показали возможность предсказывать возникновение критических ситуаций на пешеходном переходе, и тем самым, воздействуя на движущиеся средства, предотвращать их возникновение.

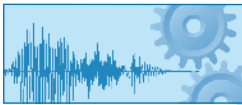
Тематический анализ Правил дорожного движения [5] позволяет выявить различные классы ситуаций дорожного движения, которые могли бы быть описаны подобным же образом. То есть, таким образом, имеется возможность, с одной стороны, выявить различные классы ситуаций в рамках всего множества ситуаций, а, с другой — путем несложных манипуляций сформировать соответствующие фрагменты модели мира. Созданная таким образом модель мира становится основой для автоматизации процесса управления движением.

### ЗАКЛЮЧЕНИЕ

В работе описывается решение задачи управления пешеходным переходом с использованием локального фрагмента коллективной модели мира всех участников движения, включающей языковой и многомодальный компоненты. Управление пешеходным переходом — стереотипная задача управления дорожным движением, так как включает в себя все типы участников дорожного движения и все виды их взаимодействия на дороге. Проведенное моделирование показало эффективность предложенных методов, а, главное, эффективность использования единой для всех участников движения коллективной модели мира. Реализация подхода позволила проверить возможность автоматизации процесса формирования базы естественно-языковых шаблонов языковой части модели мира, которая (в силу изоморфности частей модели мира) позволяет формировать систему продукционных правил многомодальной части модели. Автоматический анализ текстов [6], описывающих предметную область (например, Правил дорожного движения) позволяет извлекать из них, а, затем, кластеризовать множества предложений, которые, описывая типовые ситуации, после незначительной корректировки специалистом становятся такими шаблонами.

### Литература

1. Харламов А.А., Жаркой Р.М. Семантический анализ визуальной сцены. Труды Международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS'2017), стр. 317–320. Минск (2017).
2. Леонтьева Н.Н. Принципиальные особенности понятий «Тексты» и «Смыслы» в прикладных системах. Вестник РГГУ Серия «История. Филология. Культурология. Востоковедение» № 9 (18), 121–140 (2016).
3. Харламов А.А. Ассоциативная память — среда для формирования пространства знаний. От биологии к приложениям. Palmarium Academic Publishing, Дюссельдорф (2017).
4. Ефименко В., Хорошевский В.Ф. Онтологическое моделирование экономики предприятий и отраслей современной России. Изд. дом Высш. шк. экономики, Москва (2011).
5. Правила дорожного движения, <https://экзамен-пдд-онлайн.рф/%D0%BF%D0%B4%D0%B4/>
6. Харламов А.А. Технология автоматического смыслового анализа текстов TextAnalyst. Вестник Московского государственного лингвистического университета. Языкознание. Междисциплинарный подход в теоретической и практической лингвистике. Вып. 13 (699), 234–244 (2014).



## MOBILE ROBOT'S COLLECTIVE MODEL OF THE WORLD. LANGUAGE ANALYSIS OF THE SCENE AND PEDESTRIAN CROSSING MANAGEMENT

***Kharlamov A.A.,***

*Doctor of Technical Sciences, Senior Researcher, Institute of Higher Nervous Activity RAS, Moscow, Professor, Department of Applied and Experimental Linguistics, MSLU, Moscow, Professor, School of Software Engineering HSE, Moscow*

***Jarkoi R.M.,***

*Prezident, ISS ("Intellectual Security Systems")*

***Arzumanov V.V.***

*Researcher, ISS ("Intellectual Security Systems")*

### Abstract

The paper presents the original architecture of a mobile robot's world model. Its sensory inputs and decision making are excluded from the mobile robot itself. The world model (the object domain model) ensures the capability of understanding, without which no response is possible. It consists of two interrelated parts: language and multimodal components. This paper considers the problem of pedestrian crossing management regarding prevention of critical situations (collisions) of the traffic participants. In this case, traffic participants can be viewed as mobile robots, the movement of which is modeled using a collective model of the world. The language component of the world model is represented by templates of natural language sentences describing standard traffic situations. The system of production rules represents a multimodal component of the world model. The pedestrian crossing management model was implemented on the basis of the CLIPS language. The visualization tool for model implementation of scenarios is a PowerPoint presentation with a macro program and ActiveX control elements. In order to identify areas of parameters corresponding to standard and non-standard situations, some experiments were performed.

**Keywords:** collective world model, linguistic and multimodal world model components, pedestrian crossing control, CLIPS-based control system.