

**Крылов В.С.**

**Компьютерная паралингвистика: событийно ориентированный подход моделирования взаимосвязи вербальных и невербальных каналов коммуникации**

*Крылов Владимир Сергеевич, кандидат биологических наук,  
кафедра информационно-компьютерных технологий*

*Республиканское высшее учебное заведение Крымский инженерно-педагогический университет (РВУЗ КИПУ),  
г. Симферополь, Крым*

**Аннотация.** Информационные технологии (ИТ) сформировали глобальную инфраструктуру, которая обеспечивает передовые услуги за счет организации связи между людьми, и даже между вещами физическими или виртуальными. В сетях интернета возникли и активно развиваются различного рода контактные языки, подобные пиджин и креольским языкам. Синтаксис и грамматика таких языков отличается от синтаксиса и грамматики базового языка. Тексты сообщений дополняются различного рода знаками, символами, смайликами. Сообщения представлены двумя потоками с различными знаковыми системами. Эти потоки объединяются в один поток событий с вербальным и невербальным поведением. Предлагаемый глобальной инфраструктурой набор инструментов и технологий позволяет разрабатывать и экспериментировать как с моделями таких интегрированных потоков событий, так и с реальными реализованными текстами. В учебном процессе, в работе с этими моделями будущий ИТ-специалист приобретает навыки и умения событийно-ориентированного проектирования программного обеспечения, одного из самых сложных в освоении.

**Ключевые слова:** компьютерная паралингвистика, событийно-ориентированный подход, реализованный язык

Развитие интернета повлекло за собой появление новых языковых особенностей, лингвистические характеристики которых определяются соответствующим каналом коммуникации в информационном пространстве. Синтаксис, грамматика таких языков значительно отличается от синтаксиса, грамматики базового языка коммуникантов. Они определяются как контактные языки подобные пиджин и креольским языкам [8, 9].

Стремительное совершенствование информационных технологий (ИТ) предоставило широкому кругу исследователей доступные аппаратные и программные средства сбора и обработки исходных данных, их анализа. Ранее недоступные инструменты и технологии активно используются компьютерной паралингвистикой для экспериментальных и теоретических исследований интегрированных вербально-невербальных моделей поведения. Результаты этих обеспечивают дальнейшее развитие ИТ в различных направлениях, например, в совершенствование систем анализа и обработки естественной речи и текстов, совершенствование поисковых системы.

**Цель настоящей** статьи предложить событийно-ориентированную модель анализа текстов сообщений как потока событий на основе опыта разработки обработчиков событий в интегрированных информационных системах, разного уровня сложности, опыта разработки платформ программирования, компьютерных игр. В учебном процессе в работе с такими моделями, будущий ИТ-специалист приобретает навыки и умения работы с методологией событийно-ориентированного проектирования программного обеспечения, одной из самых сложных в практическом программировании.

В каналах мгновенных текстовых сообщений интернет чатов, твиттере, форумов, СМС и других идет активное словообразование через образование новых аббревиатур. В тексты обменов активно включаются смайлики, визуальные и аудио элементы. Текст сообщения становится поликодовым, в котором все элементы не являются просто суммой, объединением, а интегрированы единое целое для передачи смысла. Такие тексты определяются как реализованные [8]. То есть тексты, состоящие из двух объединённых в единую структуру частей: вербальной (речевой) и невербаль-

ной, которая относится к другим знаковым системам. Обе эти части зависят друг от друга. Вербальную часть невозможно отделить от невербальной. Невербальные средства в языковой коммуникации изучаются в специальном разделе паралингвистики [9, 11].

Например, социальная сеть Инстаграм предназначена для размещения собственных фотографий о каких-либо событиях, пейзажах и так далее, а так же личных фотографий на этом фоне [1]. В конечном итоге фото размещают для получения отзывов или комментариев, как от всех возможных участников сети, так и для определенного круга пользователей. Комментарии являются откликами, и вполне соответствуют определению понятия события, и составляют поток событий. Они сами по себе могут вызывать отклики, то есть являются событием, которое представляет определенную реакцию на данный комментарий.

Событие, как понятие, имеет различные определения в различных отраслях деятельности, ситуациях, обстоятельствах. В данном случае событие представляет собой сообщение, посылаемое объектом, чтобы сигнализировать о совершении какого-либо действия. Любой объект, способный вызывать события, является *отправителем событий*, также называемым *источником событий*. Отправителю события зачастую не известен объект, который будет получать, и обрабатывать сформированные отправителем события. Получатель должен обладать *обработчиком событий* - процедурой, исполняемой при получении соответствующего события [2, 10].

Поток комментариев представляет собой поток событий. Во всяком случае, участники сети воспринимают их как события, и принимают решение как на них реагировать. Ответ или отклик определяется в соответствии с представлениями об объекте и внутренним состоянием на момент ответа. Он может быть представлен текстом, фото, видео, смайликом, а так же какой-либо их комбинацией. Сообщение комментария представляет собой некоторый реализованный текст, отображающий вербальную и невербальную оценку события. Каждая из этих компонент формирует сообщение по своим собственным правилам. Синтаксис и грамматика вербальной части, как правило,

отличаются от синтаксиса и грамматики базового языка. Существует определенная зависимость между получаемым, например, визуальным событием и структурой предложения отклика на событие.

Для изучения взаимосвязи особенностей предъявляемых изображений, видео сюжетов и структуры предложения, описывающего показанную сцену, в середине 90-х годов Томлин Р. предложил исследовательскую парадигму, которую обозначил как фишфильм [5, 6]. В этом исследовании англоговорящие участники эксперимента, по результатам просмотра фильма анимации, должны были описать сюжет, в котором одна рыба съедала другую. Визуальная метка фокусировала внимание участника на возможного пациента или агенса, представленных двумя рыбами приближающихся друг к другу. Так же фокусировалось внимание и на то, что происходит перед событием поедания. Когда внимание фокусировалось на возможного агенса, участники эксперимента описывали событие как предложение в активном залоге. Когда это был пациент, они составляли описание по большей части в пассивном залоге. Результаты экспериментов Томлина Р. подвергались серьезной критике. Тем не менее, методическая часть парадигмы была усовершенствована и проведены исследования для разных языков, в том числе и для русского [5, 6].

Следует отметить, что организация эксперимента во многом соответствует тому, что происходит в социальных сетях. Например, в русской сети Инстаграм можно отобрать обсуждения фото и видео сюжетов, практически совпадающие с тем, как организован эксперимент с усовершенствованной методикой фишфильма. Как это представлено в [6]. Предварительный анализ отобранных сюжетов, дал результат в целом соответствующий выводам относительно места страдательного залога в грамматике русского предложения, редкости его использования в обычной беседе и связи с особенностями описываемой сцены.

Анализ структуры текстов откликов или комментариев, на отобранной ветке обсуждений в социальной сети, проводился с помощью программ парсеров. Эти программы разрабатываются для автоматизированного поиска, сбора и анализа различной информации: текстов кодов программ, фильмов, музыки, тематических архивов документов, отдельных документов. Процесс поиска и анализа такой программой определяется как парсинг. Поэтому в основном под парсингом обычно понимается визуальный, автоматический синтаксический и лексический анализ текстов, разбор какого-либо документа для извлечения из него определенных данных. Часто востребованный результат парсинга заключается в поиске контента, содержащего не только требуемые ключевые слова, но и определенную синтаксическую организацию текста.

Особенность исполнения парсинга определяется заданием определенной информации, например, со страниц сайтов. Особенность исполнения парсинга определяется заданием схемы отбора информации. Параметры поиска и анализа задаются заранее. Собранный информацию парсер предоставляет в определенном виде. Например, результатом работы парсера может быть база данных. Обычно процесс парсинга представлен тремя стадиями:

- сбор информации, например некоторого кода интернет страниц, перечень фильмов, размещенных на определенных порталах и так далее;
- анализ данных и преобразование в определенный формат;
- предоставление данных и их вывод.

Наиболее часто парсинг опирается на систему регулярных выражений. Регулярные выражения – система обработки текста, основанная на специальной системе записи образцов для поиска [10]. Образец или паттерн, задающий правило поиска, называют «шаблоном», «маской». Регулярные выражения используются многими текстовыми редакторами и утилитами для поиска и изменения текста на основе выбранных правил. В настоящее время сложно найти язык программирования, который не поддерживает регулярные выражения для работы со строками. Не имеют встроенный в их синтаксис механизм обработки регулярных выражений. Регулярные выражения используются для сжатого описания некоторого множества строк с помощью шаблонов, без необходимости перечисления всех элементов этого множества. При составлении шаблонов используется специальный синтаксис, поддерживающий, обычно, следующие операции:

- Перечисление, то есть допустимые варианты строки.
- Группировка, то есть объединение шаблонов одно выражение, которое воспринимается как целостная структура.– Квантификация Квантификатор после символа или группы определяет, сколько раз предшествующее выражение может встречаться.

Размещенные в текстах откликов различные знаки, символы и смайлики имеют определенные коды, поэтому возможно составлять шаблоны поиска и анализа сообщений, составленных разными знаковыми системами. То есть возможен последовательный поиск и анализ собственно текстов, отдельно последовательностей символов или и знаков, отображающих невербальную компоненту сообщения. И, кроме того, возможен поиск и анализ организации объединенной вербально-невербальной последовательности символов. Следует подчеркнуть, что шаблоны из регулярных выражений образуют некоторый метаязык. Собственно этот метаязык образует модель организации анализируемых сообщений представленных некоторым креолизованным языком. Кроме того, такой шаблон означает некоторую структуру, которая воспринимается как событие. Анализ последовательностей встретившихся единиц сообщения, соответствующих некоторому шаблону, требует обработчика событий. Поэтому анализ структуры сообщений креализованного текста будет являться результатом работы этого обработчика событий.

Модель представления структуры сообщений на креализованном языке, как последовательности некоторых шаблонов, объединяющих разные знаковые системы, дает возможность в едином ключе, на единой основе обеспечивать интегрированную моделирование целостного поведения, моделирование паралингвистических факторов в речевой активности.

Такой подход в определении интегрированных моделей как последовательности шаблонов регулярных выражений дополняет методы компьютерной пара-

лингвистики, и имеет не только научное или прикладное значение для развития и совершенствования существующих информационных систем [12]. Он позволяет формулировать различной сложности задания в учебном процессе подготовки специалистов в области информационных технологий. Выполнение заданий стимулирует развитие навыков и умений приобретать дополнительные знания, осваивать как можно больше востребованных информационных технологий, умение разрабатывать программные продукты для разных отраслей, сфер и направлений практического их применения.

Включение в учебный процесс достижений прикладных исследований новых направлений, таких как компьютерная лингвистика, компьютерная паралингвистика позволяет будущему специалисту приобрести навыки эффективного решения поставленных заказчиком задач. Именно на таком материале, на основе приобретенного опыта, молодой специалист получает необходимый набор профессиональных навыков успешной конкуренции на ИТ-рынке труда [3].

В отличие от заданий для самостоятельной работы студентов, которые заключались в разработке анализатора упрощенного креализованного языка с заданным ограниченным набором слов и смайликов, задание строилось на анализе реальных сообщений взятых из социальной сети [4]. Цель задания состояла в приобретении навыков работы с выбранной платформой программирования, умения на практике применить теоретические знания из теории формальных языков и методов разработки трансляторов, наработка практических навыков приобретения в короткий срок и на достаточном уровне знаний в различных отраслях деятельности для разработки прикладных программ. В дополнение к этому необходимо было освоить методы работы с регулярными выражениями. Построить шаблоны с типичными грамматическими структурами, и на основе этих шаблонов разработать обработчик получаемых событий [10]. Следует отметить, что наибольшие трудности возникли в освоении событийно-ориентированных методов проектирования программ.

#### ЛИТЕРАТУРА (REFERENCES TRANSLATED AND TRANSLITERATED)

1. Все о программе Инстаграм (Instagram) по-русски // [Э-ресурс].  
*All About Instagram (Instagram) in Russian* // [Online]: <http://instagrama.ru/>
2. Концептуальная модель для систем обработки событий // [Э-ресурс]  
*Conceptual model for event processing systems* // [Online]: <http://www.ibm.com/developerworks/ru/library/ws-eventprocessing/index.html>
3. Крылов В.С. Молодые специалисты способны конкурировать на рынке труда информационных технологий. // Крылов В.С. - Science and Education a New Dimension: Pedagogy and Psychology. Vol. 5, 2013 – p. 98 – 101 П  
*Krylov V.S. Young professionals are able to compete on the labor market of information technologies.* // V.S. Krylov - Science and Education a New Dimension: Pedagogy and Psychology. Vol. 5, 2013 - p. 98 - 101 P
4. Крылов В.С. Компьютерная паралингвистика как основа приобретения специальных профессиональных навыков в области информационных технологий/ Крылов В.С. – Science and Education a New Dimension. Philology, II(4), Issue: 24, 2014 с. 41-44  
*Krylov V.S.. Computer paralinguistics as the basis of acquisition of special skills in information technology / VS Krylov - Science and Education a New Dimension. Philology, II (4), Issue: 24, 2014 p.41-44.*
5. Myachykov A. Perceptual priming and structural choice in Russian sentence production // Myachykov A., Tomlin R. University of Edinburgh1 and University of Oregon – [Online]: [file:///C:/Users/Papa/Downloads/MT\\_2008.pdf](file:///C:/Users/Papa/Downloads/MT_2008.pdf)
6. Myachykov A. Visual Cues to Structural Choice in Visually Situated Sentence Production // Myachykov A., Thompson D., Garrod S., Scheepers C. *Frontiers in Psychology*, Published online Jan 18, 2012. – [Online]: <http://journal.frontiersin.org/Journal/10.3389/fpsyg.2011.00396/full>
7. События и обработчики событий // [Э-ресурс]  
*Sobytiya i obrabotchiki sobyiy [Events and Event Handlers]* // [Online]: [http://msdn.microsoft.com/ru-ru/library/2z7x8ys3\(v=vs.90\).aspx](http://msdn.microsoft.com/ru-ru/library/2z7x8ys3(v=vs.90).aspx)
8. Beisenova Z.S. The Study of Creolized Texts in Written Communication/ Beisenova Z.S., Bayekeyeva A.T., Sapina S.M., Dinayeva B.B., Utanova A.K. - INDIAN JOURNAL OF APPLIED RESEARCH, Volume : 4, Issue : 5, May 2014. p. 217
9. Bickerton D. Creole Languages // Bickerton D. – Scientific American, July 1983, Vol. 249, No.1, pp. 116-122.
10. Event-Based Programming. Taking Events to the Limit / [Online]: <http://sharebookfree.com/event-based-programming-taking-events-to-the-limit/>
11. Krämer N.C. Nonverbal Communication // Nicole C. Krämer – Human Behavior in Military Contexts, P. 150. – 188, [Online]: <http://books.nap.edu/catalog/12023.html>
12. Schuller B., Batliner A. Computational paralinguistics: emotion, affect and personality in speech and language processing / Schuller BjoËrn, Batliner Anton. – First Edition, John Wiley & Sons, Ltd , 2014, – p. 321

#### **Krylov V.S. Computer-paralinguistics: event-oriented approach of modeling the relationship of verbal and nonverbal communication channels**

**Abstract.** Information technology (IT) have formed a global infrastructure that provides advanced services through the organization of communication between people and even between things physical or virtual. In the networks of the Internet emerged and are actively developing various kinds of contact languages like Pidgin and Creole. The syntax and grammar of languages differs from the syntax and grammar of the base language. The message texts are complemented by various types of signs, symbols, emoticons. Messages are presented in two streams with different sign systems. These streams are merged into a single stream of events with verbal and nonverbal behavior. The proposed global infrastructure tools and technologies allows us to develop and experiment with models such as the integrated flow of events and real krealizovannymi texts. In the educational process, in working with these models for future IT specialist acquires skills event-oriented software design, one of the most difficult to master.

**Keywords:** computer paralinguistics, event-oriented approach, krealizovanny language