

КОМПЛЕКСНЫЙ ПОДХОД К РЕАЛИЗАЦИИ ИНТЕРФЕЙСА НА ЕСТЕСТВЕННОМ ЯЗЫКЕ В ИНФОРМАЦИОННЫХ СИСТЕМАХ

Взаимосвязанное развитие экономики, науки, техники, технологий привело к возрастанию информационных потоков и усложнению информационных процессов, которые обеспечивают это развитие. Тенденция возрастания роли информации наблюдается во всех сферах деятельности человека и проявляется в разнообразных формах, от управления сложными техническими средствами и технологическими процессами, до управления в экономике и социальной сфере. Безусловно, информационные процессы всегда имели место в истории развития общества, но сейчас их роль усиливается, переходит на качественно новый уровень.

Вышеперечисленные факты привели к появлению разнообразных информационных систем (ИС) с их многочисленными пользователями. Вместе с этим возникла проблема создания удобного и эффективного взаимодействия между ними, которое обычно реализуется на формализованных языках. Для пользователей ИС формулирование своих запросов, заданий, команд, тем более выражение своих целей, намерений, желаний на таких языках представляется затруднительным. Применение формализованных языков сокращает число потенциальных пользователей ИС, заставляет их прибегать к посредническим услугам специалистов. Решением этой проблемы является использование естественного языка (ЕЯ), особенно в разнородной информационной среде, при наличии общечеловеческого контекста, для облегченного доступа к информационным ресурсам массового использования и необходимости диалога в процессе общения.

Современные эффективные реализации ЕЯ-интерфейса и обработки текста используют лингвистические, статистические и эвристические методы. Эффективность их базируется на узости предметной области, функциональной ограниченности и простоте реализации. Лингвистические методы основаны на упрощенных моделях языка, в которых априорно закладываются знания для осуществления синтаксического анализа. Для извлечения семантики используются жесткие, плохо адаптируемые модели предметных областей. Прагматический момент либо отсутствует, либо упрощен и опирается на формальные правила, что приводит к невозможности осуществления содержательного диалога. Примером использования таких методов может служить подход к синтаксическому анализу, описанный в [1]. Статистический анализ в основном используется для специализированной обработки текстов и основан на вероятностной оценке связей лексических единиц. Лингвистические и статистические методы являются формализованными, легко реализуются и могут использоваться для первичной обработки в более гибких системах интерпретации ЕЯ. Эвристические методы представляют собой формализацию интуитивно найденных человеком способов решения некоторых частных задач при создании ЕЯ-интерфейсов.

Общими недостатками существующих систем ЕЯ-интерфейса являются отсутствие универсальности, адаптивности, возможности поддержания диалога и обучаемости в процессе работы. Эти недостатки объясняются большой степенью формализации этих систем, отсутствием единой концепции и возможности развития. При создании ИС, и ЕЯ-интерфейсов, в частности, применяется традиционный подход, который ставит во главу угла априорный опыт создателей ИС, детализацию всех возможных процессов в системе, делающий их жесткими на всех уровнях иерархии, не автономными. Такой подход не приемлем при создании перспективных разработок, особенно при росте сложности ИС. Надо отметить, что ограничением в создании полноценных ЕЯ-интерфейсов, в большой степени являются принципы устройства и возможности современной вычислительной техники.

Таким образом, возникла необходимость поиска новых подходов для попытки создания перспективных ЕЯ-интерфейсов, дающих возможность комфортного и эффективного взаимодействия широкого круга пользователей с ИС. Предлагаемый подход, имеющий комплексный характер, можно выразить в четырех пунктах.

Первый пункт заключается в утверждении о том, что язык базируется на модели предметной области и сам является ее неотъемлемой частью и источником ее развития. Эта посылка подробно рассмотрена в [2], где доказывается, что языковую деятельность следует считать основой мышления. Тандем языка и модели действительности является основой интеллекта человека и как предполагается, интеллектуальности будущих систем ИС. Для разработки эффективных ЕЯ-интерфейсов следует опираться на достаточно сложную и развитую модель предметной области, сопоставимую по сложности с моделью действительности у человека и пересекающуюся с ней. Такие же предложения выдвигаются по отношению к разработке сложных систем управления в [3].

Второй пункт заключается в утверждении о необходимости развивающейся иерархии в устройстве ИС, способной взаимодействовать на ЕЯ, что указывается в работах [4, 5] и о необходимости разбиения процесса создания ИС на два последовательных этапа. На первом этапе производится конструирование базы ИС, состоящей из начальных уровней иерархии. Эти уровни должны содержать механизмы порождения следующих уровней иерархии, не заложенных в ИС, но предполагаемых у нее в виде когнитивных структур. На втором этапе производится обучение ИС с использованием ЕЯ-интерфейса для формирования у нее требуемых функциональных свойств. В момент обучения должно происходить формирование иерархии когнитивных структур – образов, понятий, прототипов, категорий, на основании которых система будет распознавать и моделировать ситуации, возникающие в процессе взаимодействия с пользователем. На основании интерпретации этих ситуаций ИС будет действовать, выполняя задания пользователя. При этом предполагается параллельное обучение самому языку и наполнение содержанием концептуальной модели предметной области у ИС. Базовая, минимальная конфигурация ЕЯ и концептуальная модель предметной области должны быть взаимосвязаны и реализованы на первом этапе, на нижних уровнях иерархии ИС. Под минимальной конфигурацией ЕЯ здесь понимается ограниченный набор лексических единиц, на которые будет реакция, заключающаяся в запуске когнитивных механизмов, заложенных в концептуальной модели предметной области. Эти механизмы должны осуществлять начальный анализ и генерацию сообщений на ЕЯ и формирование начальных когнитивных структур в модели предметной области. В выстраиваемой модели предметной области должны быть отражены и ИС и пользователь.

Третий пункт заключается в утверждении о том, что при построении когнитивных механизмов ИС, без которых невозможно осуществление полноценного ЕЯ-интерфейса, надо пользоваться универсальными принципами организации информационных процессов, которые присутствуют в природе. Среди них можно отметить самые основные: принцип иерархии, принцип обратной связи, принцип неопределенности, принцип оптимальности, принцип целенаправленности. Необходимо отметить, что эти принципы взаимосвязаны и проявляются в совокупности, тем более это касается сложных ИС. Следовательно, в процессе их создания надо пропорционально интегрировать механизмы, реализующие все эти принципы и распространять их на все уровни иерархии ИС.

Четвертый пункт заключается в рекомендации о том, что при создании ИС и ЕЯ-интерфейсов в частности, надо использовать широкий круг знаний и достижений из всех областей науки, прямо или косвенно относящихся к данной проблеме. Благодаря этому увеличится вероятность появления новых идей при решении каких-либо задач, связанных с проблемой создания ЕЯ-интерфейса.

Можно предположить, что предложенный подход по реализации взаимодействия ИС с человеком посредством естественного языка не только повысит уровень общения, но и неизбежно приведет к интеллектуализации ИС, формированию в них механизмов, не аналогичных, но приближенных по функциональным возможностям к механизмам, реализующим интеллект у человека. В свою очередь интеллектуализация, как следствие повышения уровня взаимодействия с ИС на естественном языке, приведет к расширению функциональных возможностей этой системы. К этим возможностям можно отнести: поддержание диалога с пользователем, участие в принятии решений, консультирование по широкому спектру вопросов, решение задач, имеющих большую степень неопределенности, обучение и самообучение, прогнозирование каких-либо ситуаций, существенных для пользователя, идентификация пользователя по лексике, тематике и другим параметрам и т.д.

СПИСОК ЛИТЕРАТУРЫ

- 1 Волкова И.А., Головин И.Г. Об одном подходе к построению синтаксического модуля в системе распознавания устной речи // ДИАЛОГ-97: Труды междунар. семинара. М., 1997.
- 2 Турчин В.Ф. Феномен науки. Кибернетический подход к эволюции. М.: Наука, 1993.
- 3 Лачинов В.М., Поляков А.О. Информодинамика или путь к миру открытых систем. СПб.: СПбГТУ, 1999.
- 4 Сергеев В.М. Искусственный интеллект как метод исследования сложных систем // Системные исследования: методологические проблемы (ежегодник). М.: Наука, 1984.
- 5 Сергеев В.М. Искусственный интеллект: Опыт философского осмысления // Будущее искусственного интеллекта. М.: Наука, 1991.

Кафедра «Информационные системы»