

**ПОДХОД К СИНТЕЗУ ЕСТЕСТВЕННО - ЯЗЫКОВЫХ СООБЩЕ-
НИЙ
ПО ФОРМАЛИЗОВАННОМУ ПРЕДСТАВЛЕНИЮ
БАЗЫ ЗНАНИЙ**

к.т.н. Б.Н. Судаков, Омар А.Х. Авадала, А.А. Фоменко
(представил д.т.н., проф. Е.И. Бобыр)

В статье раскрывается актуальность и необходимость синтеза естественно-языковых (ЕЯ) конструкций при общении пользователя с ЭВМ. Подробно рассмотрен этап логического синтеза структуры ЕЯ сообщений. Показано, что после выполнения этого этапа осуществляется семантическая интерпретация заданного фрагмента.

В последнее время в различных сферах человеческой деятельности широко применяются интеллектуальные экспертные системы (ЭС). При этом под интеллектуализацией понимается развитие возможностей ЭС в направлении обеспечения совместного с пользователем решения задач, упрощение процесса общения человека и ЭВМ, постоянное расширение доли машины в совместной с человеком деятельности по решению трудно формализуемых задач.

При этом экспертная система в составе человеко - машинной системы превращается из пассивного звена в активное целенаправленное, которое обеспечивает получение требуемого в постановке задачи знания, за счет автоматического формирования и выполнения оптимальной в определенном смысле последовательности вычислительных, логических и поисковых операций.

Структура интеллектуальной ЭС состоит, по крайней мере, из трех компонентов. Первый компонент представляет собой совокупность средств, выполняющих программы. Второй – это совокупность средств интеллектуального интерфейса, имеющих гибкую структуру, которая обеспечивает возможность адаптации в широком спектре интересов конечного пользователя. Третий компонент обеспечивает взаимодействие первых двух и представляет собой базу знаний об объектах реального мира и связях между ними.

Структура базы знаний интеллектуальной ЭС иллюстрируется рис.1. Здесь можно выделить знания о предметной области (ПО) и знания о языках, используемых в системе. При этом вводятся перечисленные ниже обозначения.

Зн.1 – знания пользователей об основных понятиях ПО и связях

между ними.

Зн.2 – материализованные знания (руководства, инструкции и т.д.).

Зн.3 – поле знаний (полуформализованное описание Зн.1 и Зн.2 в виде рисунков, таблиц, схем, сетей, диаграмм, представленных на естественном языке (ЕЯ).

Зн.4 – лингвистические знания, которые разделяются на следующие типы:

- морфологические знания, описывающие структуру слов языка;
- синтаксические знания, описывающие структуру предложений и текстов;
- семантические знания, указывающие всю информацию, связанную со значениями и смыслом единиц языка и связей между ними;
- прагматические знания описывают совокупность конструкций языка, обеспечивающих выражение целей пользователя в процессе общения.

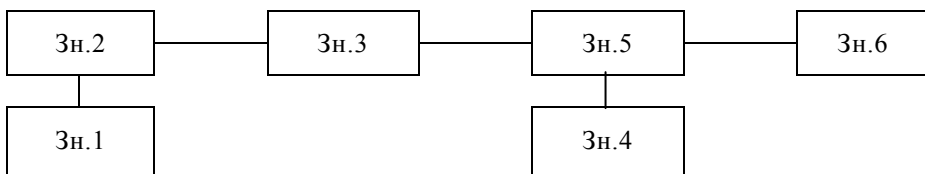


Рис.1. Структура базы знаний интеллектуальной ЭС

В дальнейшем структура знаний представляется в виде формализованного описания, которое обозначено как Зн.5.

Зн.6 – база знаний в ЭВМ на машинных носителях информации.

Таким образом, база знаний представляет собой иерархическую многоуровневую систему. На каждом из перечисленных уровней знаний используются свои языки. На уровнях Зн.1 и Зн.2 используются, как правило, профессиональный естественный язык – единственный мета-язык, позволяющий описывать всю предметную область.

Концепция интеллектуализации процесса сбора и обработки информации предполагает совместное с широким кругом пользователей (операторы, эксперты, когнитологи) решение задач ЭВМ. Это возможно осуществить в том случае, если будет не только формализован процесс принятия решения, но и предоставлены средства взаимодействия, позволяющие пользователям более эффективно использовать возможности ЭС.

Широкое применение экспертных систем, использующих формализованное представление знаний о предметной области, тормозится из-за отсутствия развитого ЕЯ интерфейса. Это обусловлено в первую очередь, трудностями, связанными с анализом и синтезом ЕЯ конструкций. В настоящее время при организации диалога ЭС–пользователь широко

используется метод интерпретации результатов логического вывода в виде, так называемой, "Семантической компоненты". Сущность этого метода заключается в том, что для каждой результирующей вершины в ЭВМ хранится заранее записанная ее семантическая интерпретация, которая в конечном итоге и выдается пользователю в виде сообщения. Такой подход оправдывает себя в системах реального времени, где его применение наиболее эффективно при решении прикладных задач. Однако, в ряде систем, например, планирования, прогнозирования, данный подход не позволяет решить ряд важных задач, в частности, выдачи пользователю пояснительной компоненты на понятном ему языке, проверки фрагментов базы знаний на семантическую корректность при ее пополнении и изменении. Для решения этих прикладных задач необходимо разработать методы алгоритмического синтеза ЕЯ сообщений по формализованному фрагменту базы знаний.

Решение задачи синтеза ЕЯ сообщений во многом определяется выбранными методами логической обработки знаний и средствами их формализованного представления. Так, например, при синтезе пояснительной компоненты языковыми средствами, важно представлять реализацию логического вывода. Это влияет, в свою очередь, на структуру алгоритма отбора фрагментов базы знаний, необходимых для формирования пояснительной компоненты. В этом случае, исходя из особенностей логической обработки, в соответствии с алгоритмом синтеза, производится обратный просмотр от результирующей вершины до исходных данных, анализируя промежуточные узлы, и принимается решение, какие из промежуточных результатов необходимы для формирования пояснительной компоненты. В некоторых ситуациях достаточно, когда такой фрагмент включает лишь исходные и результирующие данные. Например, для пояснения сообщения "Покинуть помещение", пользователю достаточно знать первопричину (например, уровень радиации повышен, отключение электроэнергии и т.д.).

Для других задач важно выбрать некоторые переходы в дереве доказательства, например, по имени результирующей ситуации, исключая дополнительные ветви по пересчету вероятностных оценок, определения условий выполнимости ситуации и других средств, участвующих в выводе. При построении полного дерева обратного вывода и его последующей интерпретации естественно – языковыми средствами, можно решать задачу тестирования логической обработки на корректность. Этот этап формирования необходимого фрагмента базы знаний будем называть логическим синтезом. После выполнения этого этапа осуществляется семантическая интерпретация заданного фрагмента, например, для решения задачи проверки на семантическую корректность пополняемых фрагментов, записанных на формализованном языке.

В соответствии с алгоритмом синтеза ЕЯ сообщений в данном случае начинается непосредственно с этапа семантического синтеза, сущность

которого заключается в адекватной семантической интерпретации отношений и понятий, представленных на формализованном языке ЕЯ средстами.

Например: пусть выражению Ф: < ИО , ТО >, где ИО – имя объекта, ТО – тип объекта соответствует запись на внутреннем языке

Ф: < РАЗГРУЗКА , объект – процесс >.

Этой записи соответствует

< Производить разгрузку / В.п. > ,

где В.п. – винительный падеж.

Результаты выполнения этапа семантической интерпретации в соответствии с общим алгоритмом передаются на этап синтаксического и морфологического синтеза. В его задачу входит:

- формирование границ предложений, в соответствии с заданными синтаксическими правилами;
- грамматическое согласование словоформ в пределах предложения;
- синтез словоформ на основе сформированной грамматической информации.

Таким образом, подход к решению задачи синтеза естественно – языкового сообщения по формализованному представлению базы знаний, как интерпретации внутреннего представления естественно – языковыми средствами позволят не только значительно повысить эргономические показатели системы, но и контролировать машинное "понимание" выполнения прикладной задачи на принципиально новом качественном уровне.

ЛИТЕРАТУРА

1. Материалы II Всесоюзной конференции "Искусственный интеллект – 90", 21-24 октября 1990 г., т.2. – Минск, 1990. – 175 с.
2. Попов Э.В. Общение с ЭВМ на естественном языке. – М.: Наука, 1982. – 360 с.
3. Справочник. Искусственный интеллект. Книга 2. Модели и методы. – М.: Наука, 1992. – 304 с.

Поступила в редколлегию 15.10.2001