

Г.В. РЫБИНА, В.В. СМИРНОВ, Т.А. ЛУПИЛЬЦЕВА  
 Московский государственный инженерно-физический институт  
 (технический университет)

## МОДЕЛИ ВЕРИФИКАЦИИ ПОЛЯ ЗНАНИЙ И БАЗЫ ЗНАНИЙ ДЛЯ ПОСТРОЕНИЯ ИНТЕГРИРОВАННЫХ ЭКСПЕРТНЫХ СИСТЕМ СРЕДСТВАМИ КОМПЛЕКСА АТ-ТЕХНОЛОГИЯ

Работа посвящена описанию моделей верификации поля знаний и базы знаний, а также средствам их реализации в комплексе АТ-ТЕХНОЛОГИЯ.

Средства автоматизированного построения баз знаний (БЗ) комплекса АТ-ТЕХНОЛОГИЯ [1] предназначены для извлечения, структурирования и формализации знаний из различных источников. Добавление нового источника знаний - базы данных (БД) [2] привело к необходимости решения задач, связанных с объединением и верификацией фрагментов поля знаний (ПЗ), формируемых при использовании различных методов, а также формированием БЗ на языке представления знаний продукционного типа путем преобразования единого ПЗ. В данной работе рассматриваются предложенные модели верификации, а основные средства верификации ПЗ были рассмотрены в работе [1].

Модель верификации ПЗ может быть представлена в следующем виде:

$M_{v1} = \langle K, I, P_1, R_1, F_1 \rangle$ , где  $K$  - сеансы извлечения знаний (ИЗ) из различных источников различными методами;  $I$  - единое ПЗ;  $P_1 = \{P_i\}$  - план верификации, состоящий из множества выбранных ошибок, где  $P_i \in E_1$ ,  $E_1$  - множество всех предусмотренных в ПЗ ошибок;  $R_1 = \{R_1^1, R_1^2\}$  - протокол верификации, где  $R_1^1$  и  $R_1^2$  - обнаруженные и устраненные ошибки;  $F_1 = \{F_1^1, F_1^2\}$  - функции формирования  $R_1$  на основе  $K, I, P_1, F_1^1$  и  $F_1^2$  - множества функций обнаружения и устранения ошибок.

Источниками  $K$  является  $S = \{S_1, S_2, S_3, S_4, S_5\}$ , где  $S_1$  - эксперты;  $S_2$  - проблемно-ориентированные тексты;  $S_3$  - БД;  $S_4$  - ПЗ;  $S_5$  - БЗ.

Модель верификации БЗ может быть представлена в следующем виде:

$M_{v2} = \langle KB, P_2, R_2, F_2 \rangle$ , где  $KB$  - БЗ;  $P_2 = \{P_2i\}$  - план верификации, состоящий из множества выбранных ошибок, где  $P_2i \in E_2$ ,  $E_2$  - множество всех предусмотренных ошибок;  $R_2 = \{R_2^1, R_2^2\}$  - протокол верификации, где  $R_2^1$  и  $R_2^2$  - обнаруженные и устраненные ошибки;  $F_2 = \{F_2^1, F_2^2\}$  - функции формирования  $R_2$  на основе  $KB, P_2$ , путем обнаружения ( $F_2^1$ ) и устранения ( $F_2^2$ ) ошибок.

Единос ПЗ I включает в себя объекты и правила, поэтому  $F_1$  являются реализацией двух групп методов: 1) методы тестирования объектов: анализ и корректировка протоколов интервьюирования экспертов (ПИЭ); анализ граничных значений атрибутов объектов (для выявления несогласованности фрагментов БЗ); 2) методы тестирования правил: анализ таблиц соответствий посылок и действий правил (для выявления противоречий, избыточности, циклов, пересечений в правилах); генерация тестов по методу покрытия операторов (для выявления неполноты);

План верификации  $P=\{P_1, P_2\}$  предназначен для настройки средств верификации и формируется на основе  $E=\{E_1, E_2\}$ . Ошибки учитываются в  $R^1=\{R_1^1, R_2^1\}$ , чтобы инженер по знаниям принимал решение о методе их устранения. Результаты устранения ошибок (редактированием ПЗ, ПИЭ или БЗ; автоматической коррекцией ПЗ; путем выбора метода ИЗ и запуском соответствующего сеанса ИЗ) помещаются в  $R^2=\{R_1^2, R_2^2\}$ .

Метод анализа граничных значений применяется к значениям атрибутов объектов, факторам уверенности и точности для проверки соответствия ограничениям, устанавливаемым по умолчанию по моделям решения типовых задач [1] и уточняемым в каждом сеансе ИЗ. Метод анализа таблиц соответствий является развитием метода эквивалентного разбиения, из традиционного программирования. Таблицы заполняются признаками пересечения, вхождения, отрицания и т.д. антецедентов и/или консеквентов, чтобы выявлять различные классы ошибок. Основой для генерации тестов служат фрагменты БЗ, сформированные при ИЗ из БД[2]. Для сокращения перебора при генерации тестов применяется группировка правил по использованию атрибутов и объектов. Для независимости процессов извлечения, структурирования (включая верификацию) и формализации знаний, разработан набор функций обработки единого ПЗ, элементы которого, имеют вид "предикат-объект-атрибут-значение-коэффициенты", соответствуют утверждениям, извлекаемым из эксперта или БД, и связаны друг с другом с помощью правил.

#### *Список литературы*

1. Рыбина Г.В. Автоматизированное построение баз знаний для интегрированных экспертных систем: Изв. РАН. Теория и системы управления, 1998, №56. С.152-166.
2. Калинин Е.А., Рыбина Г.В. Применение технологии Data Mining для автоматизированного построения баз знаний интегрированных экспертных систем // В кн. КИИ-2000. Седьмая все. конференция с межд. участием. Тр. конф. М.: Физматлит, 2000. Т.2. С. 119-127.