

УДК 519.7

Ю.Г.Ефремов (6 курс, каф. ИУС),
Т.В.Леонтьева, к.т.н., доц., И.Г.Черноруцкий, д.т.н., проф.

ПРОБЛЕМЫ ЗАПОЛНЕНИЯ БАЗЫ ЗНАНИЙ И ЕЕ НАСТРОЙКА

Создание экспертной системы (ЭС) включает в себя нескольких этапов: разработка структуры экспертной системы, заполнение базы знаний, отладка и тестовое диагностирование. Остановимся подробнее на этапе заполнения базы знаний.

Наша экспертная система построена на байесовском подходе и принципах, предложенных К. Нейлором. Процесс заполнения базы знаний сводится к определению гипотез, свидетельств, а также подбору «правильных» вероятностей, которые позволят работать экспертной системе корректно.

Рассмотрим медицинскую экспертную систему, где к гипотезам относится список болезней, из которого затем и получаем диагноз – например, грипп. К свидетельствам относятся симптомы, например, высокая температура, кашель. Для заполнения базы знаний необходимо каждой гипотезе сопоставить какую-либо вероятность, которая будет являться априорной вероятностью данной гипотезы. Также необходимо ввести все вероятности, осуществляющие связь каждого свидетельства с какой-либо одной или несколькими гипотезами.

Основная проблема при заполнении базы знаний – в том, откуда взять все необходимые вероятности; при этом занесенные в базу вероятности должны корректно работать. С этой проблемой столкнулись студенты четвертого курса, которым было задание создать собственную экспертную систему. Решалась эта проблема разными путями. Для какой-либо глобальной экспертной системы значительно упрощает подбор начальных вероятностей уже имеющаяся статистическая литература, например, для создания медицинской ЭС. Однако, следует учесть возможность того, что в литературе могут встретиться и недостоверные данные. Для более простых экспертных систем или в случае отсутствия литературы возможен подбор вероятностей на “логико-интуитивной” основе. Как и в первом случае, здесь очень высока вероятность возникновения ошибочно подобранных вероятностей, так как играет роль не только человеческий фактор, но и сложность логического подбора при возрастании количества свидетельств и гипотез.

Различного рода ошибки, будь то опечатка или ошибочно подобранные данные, выявляются на этапе отладки базы знаний. Отладка осуществляется на модельных примерах, обладающих не очень высокой сложностью. При помощи диагностирующей системы проигрывается тот или иной сценарий диалога с пользователем. Тем самым определяются и исправляются грубые ошибки, после чего модельный пример прогоняется через диагностирующую систему заново. После исправления грубых ошибок запускаются более длинные и сложные примеры, а затем тестирование проходит на “реальных” людях в присутствии эксперта, который после ряда тестов вносит все необходимые изменения в базу знаний.

В результате получаем отлаженную на необходимом нам уровне экспертную систему. Причем, чем меньше глубина и количество модельных примеров, тем больше ошибок содержит в себе созданная база знаний.

Данная работа поддерживается грантом для молодых кандидатов наук вузов Санкт-Петербурга, предоставленным Администрацией СПб.