

Экспертные системы. Технология разработки. Этап концептуализации.

На данном этапе проводится содержательный анализ проблемной области, выявляются используемые понятия и их взаимосвязи, определяются методы решения задач.

Этот этап завершается созданием модели предметной области (ПО), включающей основные концепты и отношения.

На этапе концептуализации определяются следующие особенности задачи:

- типы доступных данных;
- исходные и выводимые данные,
- подзадачи общей задачи;
- используемые стратегии и гипотезы;
- виды взаимосвязей между объектами ПО, типы используемых отношений:
 - ❖ иерархия,
 - ❖ причина — следствие,
 - ❖ часть — целое и т.п.;

- процессы, используемые в ходе решения;
- состав знаний, используемых при решении задачи;
- типы ограничений, накладываемых на процессы, используемые в ходе решения;
- состав знаний, используемых для обоснования решений.

Существует два подхода к процессу построения модели предметной области, которая является целью разработчиков ЭС на этапе концептуализации.

- Признаковый или атрибутивный подход предполагает наличие полученной от экспертов информации в виде троек объект — атрибут — значение атрибута, а также наличие обучающей информации. Этот подход развивается в рамках направления, получившего название формирование знаний или "машинное обучение" (machine learning).
- Второй подход, называемый структурным (или когнитивным), осуществляется путем выделения элементов предметной области, их взаимосвязей и семантических отношений.

Для атрибутивного подхода характерно наличие наиболее полной информации о предметной области: об объектах, их атрибутах и о значениях атрибутов.

Кроме того, существенным моментом является использование дополнительной обучающей информации, которая задается группированием объектов в классы по тому или иному содержательному критерию.

Тройки объект-атрибут-значение атрибута могут быть получены с помощью так называемого метода реклассификации, который основан на предположении что задача является объектно-ориентированной и объекты задачи хорошо известны эксперту.

Идея метода состоит в том, что конструируются правила (комбинации значений атрибутов), позволяющие отличить один объект от другого.

Обучающая информация может быть задана на основании прецедентов правильных экспертных заключений, например, с помощью метода извлечения знаний, получившего название "анализ протоколов мыслей вслух".

Структурный подход к построению модели предметной области предполагает выделение следующих когнитивных элементов знаний:

1. Понятия.
2. Взаимосвязи.
3. Метапонятия.
4. Семантические отношения.

Выделяемые понятия предметной области должны образовывать систему, под которой понимается совокупность понятий, обладающая следующими свойствами:

- уникальностью (отсутствием избыточности);
- полнотой (достаточно полным описанием различных процессов, фактов, явлений и т.д. предметной области);
- достоверностью (валидностью — соответствием выделенных единиц смысловой информации их реальным наименованиям)
- непротиворечивостью (отсутствием омонимии).

При построении системы понятий с помощью "метода локального представления" эксперта просят разбить задачу на подзадачи для перечисления целевых состояний и описания общих категорий цели.

Далее для каждого разбиения (локального представления) эксперт формулирует информационные факты и дает им четкое наименование (название).

Считается, что для успешного решения задачи построения модели предметной области число таких информационных фактов в каждом локальном представлении, которыми человек способен одновременно манипулировать, должно быть около семи.

"Метод вычисления коэффициента использования" основан на следующей гипотезе.

Элемент данных (или информационный факт) может являться понятием, если:

1. Он используется в большом числе подзадач.
2. Используется с большим числом других элементов данных.
3. Редко используется совместно с другими элементами данных по сравнению с общим числом его использования во всех подзадачах (это и есть коэффициент использования).

Полученные значения могут служить критерием для классификации всех элементов данных и, таким образом, для формирования системы понятий.

"Метод формирования перечня понятий" заключается в том, что экспертам (желательно, чтобы их было больше двух) дается задание составить список понятий, относящихся к исследуемой предметной области.

Понятия, выделенные всеми экспертами, включаются в систему понятий, остальные подлежат обсуждению.

"Ролевой метод" состоит в том, что эксперту дается задание обучить инженера по знаниям решению некоторых задач предметной области.

Таким образом, эксперт играет роль учителя, а инженер по знаниям — роль ученика.

Процесс обучения записывается на магнитофон. Затем третий участник прослушивает магнитофонную ленту и выписывает на бумаге все понятия, употребленные учителем или учеником.

При использовании метода "составления списка элементарных действий" эксперту дается задание составить такой список при решении задачи в произвольном порядке.

В методе "составление оглавления учебника" эксперту предлагается представить ситуацию, в которой его попросили написать учебник. Необходимо составить на бумаге перечень предполагаемых глав, разделов, параграфов, пунктов и подпунктов книги.

"Текстологический метод" формирования системы понятий заключается в том, что эксперту дается задание выписать из руководств (книг по специальности) некоторые элементы, представляющие собой единицы смысловой информации.

Группа методов установления взаимосвязей предполагает установление семантической близости между отдельными понятиями.

В основе установления взаимосвязей лежит психологический эффект "свободных ассоциаций", а также фундаментальная категория близости объектов или концептов.

Эффект свободных ассоциаций заключается в следующем. Испытуемого просят отвечать на заданное слово первым пришедшим на ум словом. Как правило, реакция большинства испытуемых (если слова не были слишком необычными) оказываются одинаковой.

Количество переходов в цепочке может служить мерой "смыслового расстояния" между двумя понятиями.

Многочисленные опыты подтверждают гипотезу, что для двух любых слов (понятий) существует ассоциативная цепочка, состоящая не более чем из семи слов.

"Метод свободных ассоциаций" основан на психологическом эффекте, описанном выше. Эксперту предъявляется понятие с просьбой назвать как можно быстрее первое пришедшее на ум понятие из сформированной ранее системы понятий. Далее производится анализ полученной информации.

В методе "сортировка карточек" исходным материалом служат выписанные на карточки понятия.

Применяются два варианта метода.

- В первом эксперту задаются некоторые глобальные критерии предметной области, которыми он должен руководствоваться при раскладывании карточек на группы.
- Во втором случае, когда сформулировать глобальные критерии невозможно эксперту дается задание разложить карточки на группы в соответствии с интуитивным пониманием семантической близости предъявляемых понятий.

"Метод обнаружения регулярностей" основан на гипотезе о том, что элементы цепочки понятия, которые человек вспоминает с определенной регулярностью, имеют тесную ассоциативную взаимосвязь.

Для эксперимента произвольным образом отбирается 20 понятий. Эксперту предъявляется одно из числа отобранных.

Процедура повторяется до 20 раз, причем каждый раз начальные концепты должны быть разными.

Затем инженер по знаниям анализирует полученные цепочки с целью нахождения постоянно повторяющихся понятий (регулярностей).

Внутри выделенных таким образом группировок устанавливаются ассоциативные взаимосвязи.

Кроме рассмотренных выше неформальных методов для установления взаимосвязей между отдельными понятиями применяются также формальные методы.

Сюда в первую очередь относятся методы семантического дифференциала и репертуарных решёток.

Выделенные понятия предметной области и установленные между ними взаимосвязи служат основанием для дальнейшего построения системы метапонятий — осмысленных в контексте изучаемой предметной области системы группировок понятий.

Для определения этих группировок применяют как неформальные, так и формальные методы.

Интерпретация, как правило, легче дается эксперту, если группировки получены неформальными методами. В этом случае выделенные классы более понятны эксперту. Причем в некоторых предметных областях совсем не обязательно устанавливать взаимосвязи между понятиями, так как метапонятия, образно говоря, "лежат на поверхности".

Последним этапом построения модели предметной области при концептуальном анализе является установление семантических отношений между выделенными понятиями и метапонятиями.

Установить семантические отношения — это значит определить специфику взаимосвязи, полученной в результате применения тех или иных методов. Для этого необходимо каждую зафиксированную взаимосвязь осмыслить и отнести ее к тому или иному типу отношений.

Существует около 200 базовых отношений, например, "часть — целое", "род — вид", "причина — следствие", пространственные, временные и другие отношения.

Для каждой предметной области помимо общих базовых отношений могут существовать и уникальные отношения.

"Прямой метод" установления семантических отношений основан на непосредственном осмыслении каждой взаимосвязи.

В том случае, когда эксперт затрудняется дать интерпретацию выделенной взаимосвязи, ему предлагается следующая процедура.

1. Формируются тройки: понятие 1 — связь — понятие 2.
2. Рядом с каждой тройкой записывается короткое предложение или фраза, построенное так, чтобы понятие 1 и понятие 2 входили бы в это предложение.

В качестве связок используются только содержательные отношения и не применяются неопределенные связки типа "похож на" или "связан с".

Для "косвенного метода" необязательно иметь взаимосвязи, а достаточно лишь наличие системы понятий.

Формулируется некоторый критерий, для которого из системы понятий выбирается определенная совокупность концептов.

Эта совокупность предъявляется эксперту с просьбой дать вербальное описание сформулированного критерия.

Концепты предъявляются эксперту все сразу (желательно на карточках). В случае затруднений эксперта прибегают к разбиению отобранных концептов на группы с помощью более мелких критериев.

Исходное количество концептов может быть произвольным, но после разбиения на группы в каждой из таких групп должно быть не более десяти концептов.

После того как составлены описания по всем группам, эксперту предлагают объединить эти описания в одно.

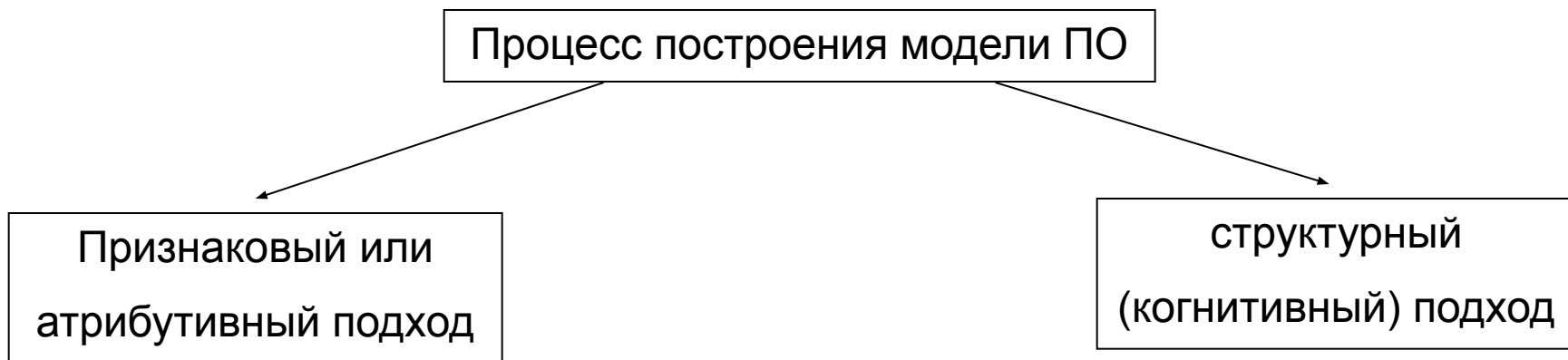
Следующий шаг в косвенном методе установления семантических отношений — это анализ текста, составленного экспертом.

Концепты заменяют цифрами (это может быть исходная нумерация), а связи оставляют.

Тем самым строится некоторый граф, вершинами которого служат концепты, а дугами — связки (например, "ввиду", "приводит к", "выражаясь с одной стороны", "обуславливая", "сочетаясь", "определяет", "вплоть до" и т.д.)

Этот метод позволяет устанавливать не только базовые отношения, но и отношения, специфические для конкретной предметной области.

Рассмотренные выше методы формирования системы понятий и метапонятий, установления взаимосвязей и семантических отношений в разных сочетаниях применяются на этапе концептуализации при построении модели предметной области.



Объект-свойство-значение

1. Понятия.
2. Взаимосвязи.
3. Метапонятия.
4. Семантические отношения.

Построение системы понятий:

- "метод локального представления"
- "Метод вычисления коэффициента использования"
- "Метод формирования перечня понятий"
- "Ролевой метод"
- "метод составления списка элементарных действий"
- "составление оглавления учебника"
- "текстологический метод"

Методы установления взаимосвязей :

□ Неформальные:

- ❖ "Метод свободных ассоциаций"
- ❖ "сортировка карточек"
- ❖ "Метод обнаружения регулярностей"

□ Формальные:

- ❖ "Прямой метод"
- ❖ "косвенного метода"