**Классификация и обзор методов извлечения знаний. Краткая характеристика методов.**

Проблема приобретения знаний включает две основные задачи: получение информации и ее систематизацию. Процесс получения знаний от экспертов является ключевым.

*Извлечением знаний называют процесс получения знаний от экспертов.*

Каждый когнитолог сам изобретает язык для описания полученных от эксперта знаний путем пополнения сложившегося языка конкретной науки специальными терминами и знаками. Стандарта таких языков пока не существует. Разработка языков инженерии знаний ведется в различных направлениях. Известны языки классификации, логико-конструкторские языки, структурно-логические языки и т.п. Перспективный подход к созданию подобных языков открывает семиотика-наука о знаковых системах(прикладная семиотика связана с применением знаковых систем для представления и обработки знаний).

Классификация методов извлечения знаний, приведенная ниже, связана с источником знаний и разделяет все методы на два класса: коммуникативные и текстологические.

*Коммуникативные методы извлечения знаний охватывают методы и процедуры контактов инженера по знаниям с экспертом.*

*Текстологические включают методы извлечения знаний из документов (методик, пособий. Руководств) и специальной литературы (статей, монографий, учебников).*

Коммуникативные методы подразделяются на пассивные и активные.

Пассивные методы подразумевают, что ведущая роль в процедуре извлечения как бы передается эксперту, а инженер по знаниям только протоколирует рассуждения эксперта во время его реальной работы по принятию решения или записывает то, что эксперт расскажет в форме лекции.

В активных методах, напротив, инициатива находится в руках инженера по знаниям, который активно контактирует с экспертом различными способами – в играх, диалогах, беседах за круглым столом и т.п.

Активные методы разделяются на две группы в зависимости от числа экспертов: групповые и индивидуальные.

Особо следует отметить игровые методы, которые широко используются в социологии, экономике, менеджменте, педагогике для подготовки руководителей, учителей, врачей и других специалистов.

**Классификация методов извлечения знаний**

|  |  |  |  |
| --- | --- | --- | --- |
| **Методы извлечения знаний** | | | |
| **Коммуникативные** | | | **Текстологические** |
| **Пассивные** | **Активные** | | **Анализ литературы** |
| **Групповые** | **Индивидуальные** | **Анализ учебников** |
| Наблюдение | **Круглый стол** | **Анкетирование** | **Анализ документов** |
| **Протокол "мыслей вслух"** | **Мозговой штурм** | **Интервью** |  |
| **Лекции** | **Ролевые игры** | **Свободный диалог** |
|  |  | **Экспертные игры** |
|  |  |  |

**Коммуникативные методы**

**1. Пассивные методы** извлечения знаний включают методы, где ведущая роль в процедуре извлечения фактически передается эксперту, а инженер по знаниям только фиксирует рассуждения эксперта во время его работы по принятию решения.

*К пассивным методам относятся: наблюдения на рабочем месте; анализ протоколов «мыслей вслух».*

*1.1 В процессе наблюдения* инженер по знаниям находится непосредственно рядом с экспертом во время его профессиональной деятельности или имитации этой деятельности. Он записывает все действия эксперта, его реплики и объяснения. Может использоваться видео запись в реальном масштабе времени. Обязательное условие –это невмешательство инженера по знаниям и навязывание им каких-то своих структур представлений. Существуют две основные разновидности наблюдений: наблюдение за реальным процессом и наблюдение за имитацией процесса. Наблюдение за имитацией процесса применяется в тех случаях, когда наблюдения за реальным процессом невозможны. Например, врач-психиатр не может допускать присутствие постороннего человека.

Затем протоколы наблюдений тщательно расшифровываются и обсуждаются с экспертом.

**1.2 Анализ протоколов «мыслей вслух»** отличается от наблюдений тем, что эксперта просят не просто прокомментировать свои действия, но и объяснить как данное решение было найдено. Весь поток комментарий и объяснений протоколируется инженером по знаниям. Спорным остается вопрос использования диктофонов, так как иногда парализующе действует на эксперта. Затем протоколы наблюдений тщательно расшифровываются и обсуждаются с экспертом.

1.3 Лекции являются самым старым методом передачи знаний. Если эксперт имеет опыт преподавательской работы, то он может этим методом воспользоваться. От инженера по знаниям требуется грамотно законспектировать лекцию, что не всегда удается.

1. **Активные индивидуальные методы** включают следующие:
   * Анкетирование;
   * Интервью;
   * Свободный диалог;
   * Игры с экспертом.

**2.1 Анкетирование** – наиболее жесткий метод. Инженер по знаниям заранее составляет анкету, размножает ее и использует для опроса нескольких экспертов (ударение на букву «е»). Сама процедура выполняется двумя способами: аналитик вслух задает вопросы и сам заполняет анкету; эксперт самостоятельно заполняет анкету. Выбор способа зависит от конкретных условий. Второй способ предпочтительнее, так как у эксперта появляется неограниченное время на обдумывание вопроса. Рекомендации по составлению анкет:

* анкета не должна быть монотонной и однообразной;
* анкета должна быть приспособлена к языку эксперта;
* последовательность вопросов должна быть продумана;
* анкета не должна быть избыточной.

2.2 Под **интервью** понимается форма общения инженера по знаниям и эксперта, в которой инженер по знаниям задает эксперту серию заранее подготовленных вопросов с целью извлечения знаний. Основное отличие интервью состоит в том, что оно позволяет аналитику опускать ряд вопросов в зависимости от ситуации, вставлять новые вопросы в анкету, изменять темп, разнообразить ситуацию общения. Задание вопросов волнует математиков и философов. Существует специальная логика вопросов в работах Белнапа.  
Вопросы классифицируются:

* по форме (открытые, закрытые, личные, безличные, прямые, косвенные, вербальные, с использованием наглядного материала);
* по функции (основные, зондирующие, контрольные);
* по воздействию (нейтральные, наводящие).

2.3 **Свободный диалог** – это метод извлечения знаний в форме беседы инженера по знаниям и эксперта, в которой нет жестко регламентированного плана и вопросника. Но это не означает, что к диалогу не надо готовиться. Наоборот, внешне свободная форма этого метода требует профессиональной и психологической подготовки. Подготовка занимает достаточно большое время. Общая подготовка включает повышение научной эрудиции, обладание общей культурой, знакомство с системной методологией. Специальная подготовка предусматривает овладение теорией и навыками интервьюирования. Конкретная подготовка включает изучение предметной области, подготовка ситуации общения, знакомство с экспертом, тестирование эксперта. Психологическая подготовка предусматривает знакомство с теорией общения, изучение когнитивной психологии.

2.4 **Экспертные игры.** Игрой называется вид человеческой деятельности, который отражает (воссоздает) другие ее виды. Все игры можно разделить на индивидуальные и групповые. Понятие экспертной игры можно классифицировать:

* на деловые игры;
* диагностические игры;
* компьютерные игры.

Под деловой игрой понимается эксперимент, где участникам предлагается производственная ситуация, а они на основе жизненного опыта, своих специальных знаний и представлений принимают решения. Решения анализируются, и вскрываются закономерности мышления участников эксперимента.

Диагностическая игра – это та же деловая игра, но применяемая конкретно для диагностики методов принятия решения в медицине ( закрепленная предметная область). Эти игры возникли при передаче опыта от опытных врачей новичкам. По использованию ЭВМ игры делятся на компьютерные и ручные. Компьютерные игры вызывают интерес у эксперта, привлекают дизайном и динамикой. К недостаткам компьютерных игр можно отнести сложность и высокая цена создания специализированных игр.

**3. Активные групповые методы.**

К групповым методам извлечения знаний относят:

* + дискуссии за круглым столом, с участием нескольких экспертов;
  + мозговой штурм;
  + ролевые игры.

Основное преимущество:

Использование знаний одновременно от нескольких экспертов; эти методы более популярны, чем индивидуальные.

**3.1 Круглый стол**. Этот метод заимствован из журналистики. Он предусматривает обсуждение какой-либо проблемы из области с несколькими экспертами, в котором принимают участие с равными правами несколько экспертов.

В начале участникам предлагают высказаться в определенном порядке, а затем переходят к живой дискуссии. Число участников от3 до 5 человек.

Инженер по знаниям должен тщательно подготовить «круглый стол» (место, время, обстановка, чай, кофе).

Инженер по знаниям должен уметь вставлять реплики, иметь чувство юмора, отличную память на имена, отчества, фамилии, способность гасить конфликтные ситуации.

Большинство участников выступает под воздействие эффекта фасада, т.е. желание произвести впечатление на других экспертов. Ход беседы можно записывать на магнитофон, с последующей расшифровкой. Цель дискуссии – это коллективно с разных точек зрения исследовать спорные вопросы предметной области. Для остроты на «круглый стол приглашают экспертов разных направлений».

Рекомендации:

1. Убедитесь, что все правильно понимают задачу;
2. Четко сформулировать тему;
3. Установить регламент.

**3.2 «Мозговой штурм».** «Мозговой штурм» или «мозговая атака» был предложен в 1939 г. Осборном (США). Этот метод используется как источник генерации новых идей в условиях запрещения критики. Штурм длится около 40 минут. Число участников до 10 человек. Им предлагается высказывать любые идеи на заданную тему. Регламент ограничен двумя минутами. Обычно высказывается около 50 идей, полезными оказываются 10-15%. Самый интересный момент штурма – наступление пика, ажиотажа, когда идеи начинают «фонтанировать», т.е. происходит произвольная генерация гипотез. Это пик имеет теоретическое обоснование в работах швейцарского психолога и психиатра З. Фрейда о бессознательном.

Ведущий «штурма» - инженер по знаниям. Он должен свободно владеть аудиторией, подобрать активную группу экспертов – генераторов и должен обеспечить «подогрев» генераторов. Фиксация производится на магнитофон или протокол.

Недостаток:

Используется только для новых исследовательских проблем; низкий процент продуктивных идей.

Достоинство:

Возможность получения новых знаний и выявить глубинные пласты знаний.

**3.3 «Ролевые игры» п**редусматривают участие в игре нескольких экспертов. Перед игрой разрабатывается сценарий, распределяются роли и к каждой роли готовится портрет (описание). Существует несколько способов проведения ролевых игр. В одних методах игроки меняют имена, в других роль вытягивают по жребию.

Роль – комплекс образцов поведения. Она связана с другими ролями. От 3 до 6 экспертов в игре. «Ролевая игра» как правило, требует несколько простейших заготовок. Пример извлечения знаний из специалистов предприятия, разрабатывающих производственные планы выпуска для цехов и принимающих различные решения по управлению производством. В игре экспертов разбили на 3 группы: ЛПР1 – группа планирования; ЛПР2 – группа менеджеров; Э – группа экспертов по оцениванию действий ЛПР1 и ЛПР2. Группам ЛПР1 и ЛПР2 раздавались различные производственные ситуации и тщательно протоколировались споры, рассуждения, аргументы по принятию решений. В результате был создан прототип базы знаний ЭС планирования выпускаемой продукции.

**4. Текстологические методы:**

Группа текстологических методов объединяет методы извлечения знаний, основанные на изучении специальных текстов из учебников монографии, нормативной литературы. Задачу извлечения знаний из текстов можно сформулировать как задачу понимания и выделения смысла текста. При этом можно выделить следующие смысловые структуры:

М1 – смысл, который пытался изложить автор.

М2 – смысл, который постигает читатель (инженер по знаниям);

Сложность процесса заключается в принципиальной невозможности совпадения знаний М1 и М2, т.е. М1 образуется за счет своей совокупности представлений, а М2 – своей.

Здесь встает задача выяснить за счет чего можно достичь максимальной адекватности знаний М1 и М2.

**5. Аспекты приобретения знаний**

Процессы извлечения знаний рассматривают в трех основных аспектах: психологическом, лингвистическом и гносеологическом ( рисунок 1)