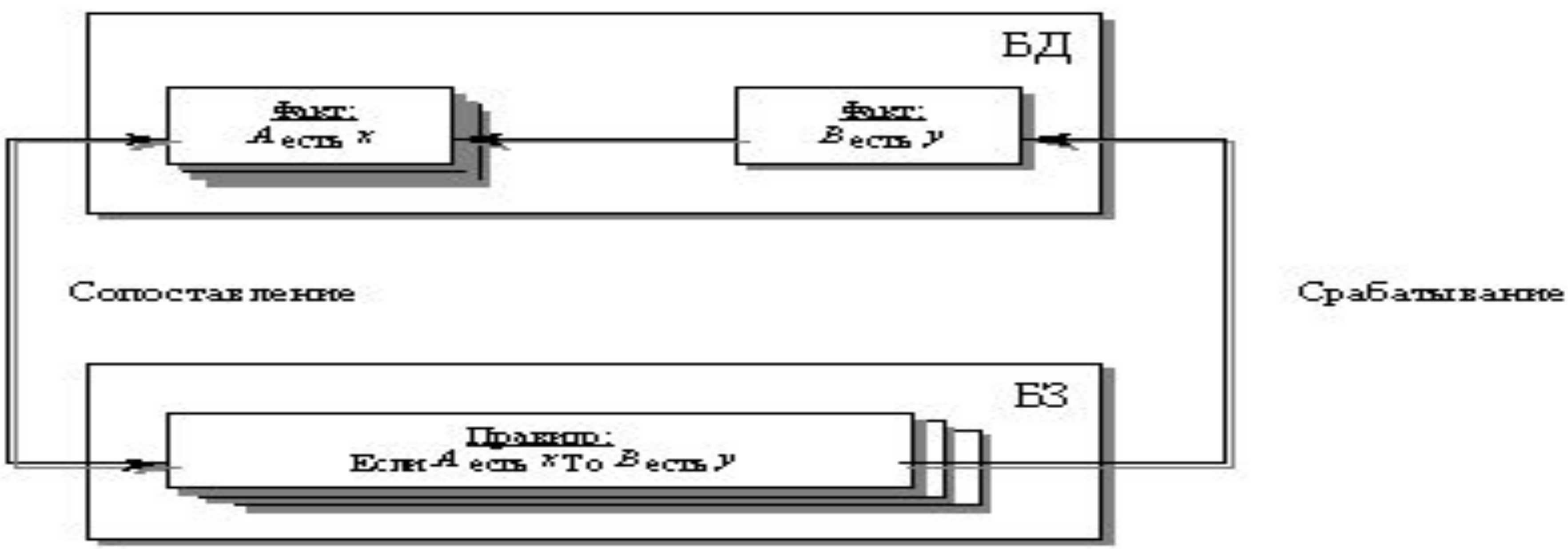
Глава 4. Обработка знаний и вывод решений в интеллектуальных системах

4.1. Методы вывода и поиска решений в продукционных системах. Методы вывода на основе прямой и обратной цепочек

Область знаний представляется мн-вом продукционных правил ЕСЛИ-ТОГДА, а данные представляются мн-вом фактов о текущей ситуации. Механизм вывода сопоставляет каждое правило, хранящееся в БЗ с фактами, содержащимися в БД. Когда часть правила если(условие) подходит факту, правило срабатывает и его часть тогда(действие) исполняется.



Пример цепочки вывода

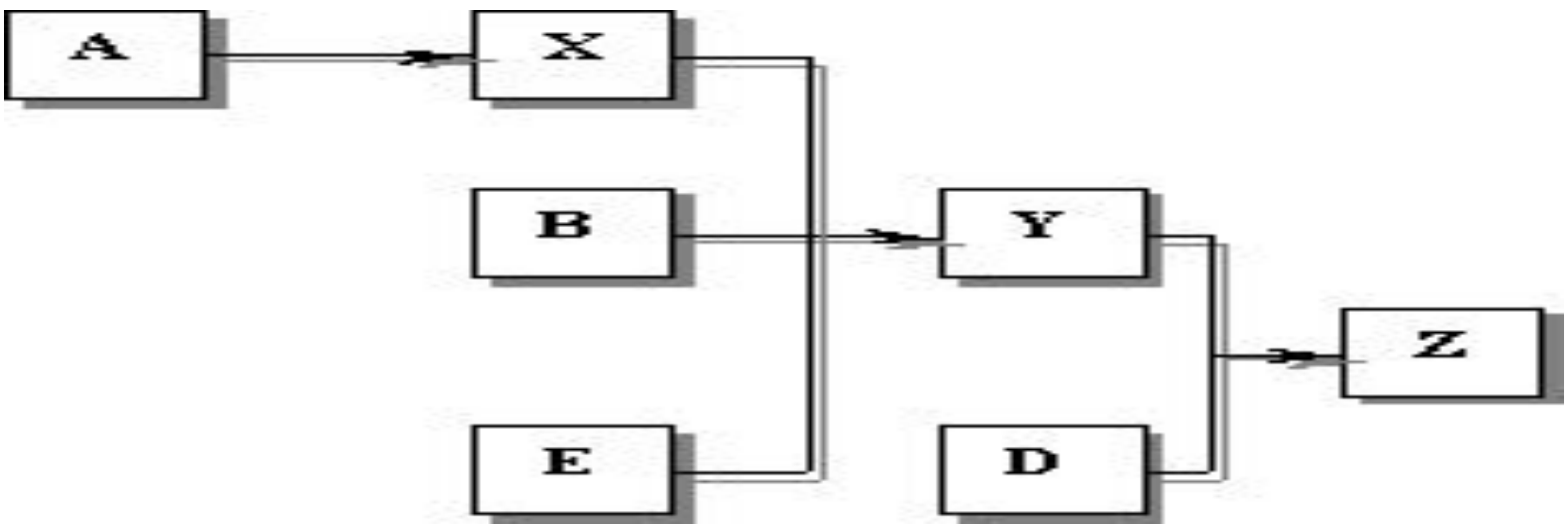
Допустим, БД первоначально включает факты A,B,C,D и E, а БЗ содержит только 3 правила:

Правило 1. Y&D → Z

Правило 2. X&B&E→Y

Правило 3. A→X

Цепочка вывода показывает, как ЭС применяет правила для вывода факта Z.



Целе-выводимые (действие-выводимые или консеквент-выводимые) продукционные системы

Правило вида A&B&C→D
может быть интерпретировано, как
«Логическая конъюнкция A, B и C влечет D» или
«Чтобы доказать D, необходимо установить A, B, C».

Цель должна быть достигнута дедуктивным выводом. Для этого исследуются консеквенты правил для нахождения такого правила, которое позволило бы достичь цели. Когда такое правило найдено, проверяются на истинность все его условия. Если условия истинны, продукция активируется. В противном случае продолжается поиск подходящей продукции.

Пример продукционной системы с консеквент-выводимой архитектурой

Дано: БД: AF

Правило 1: A&B&C→D

Правило 2: D&F→G

Правило 3:A&J→G

Правило 4: B→C

Правило 5: F→B

Правило 6: L→J

Правило 7: G→H

Цель: вывести истинность H.

Решение:

H не присутствует в БД, но есть в П7: G→H,
Теперь требуется вывести истинность G.
G присутствует в П2 и П3, выбираем 2: D&F→G
Далее выводим истинность D&F:
F - истинно (условие задачи), а A&B&C→D.
A - истинно (условие задачи), B - истинно (П5).
Следовательно и C - истинно(П4).
Т.о. истинность D и F доказана. Следовательно, G – истинно (П2).
Из истинности G следует истинность H. Цель достигнута.
Элементы B, C, D, G и H добавляются в БД.

Общие методы поиска решений в пространстве состояний. Методы перебора.

Общий вид задачи: (*S, F, T*),

где *S* – мн-во начальных состояний;

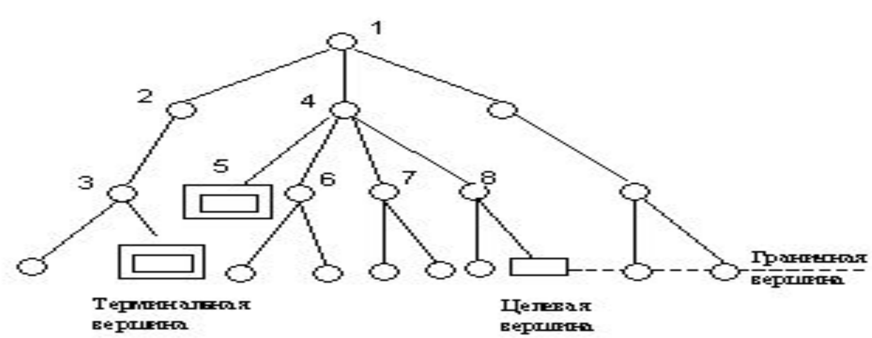
***F* – мн-во операторов,**

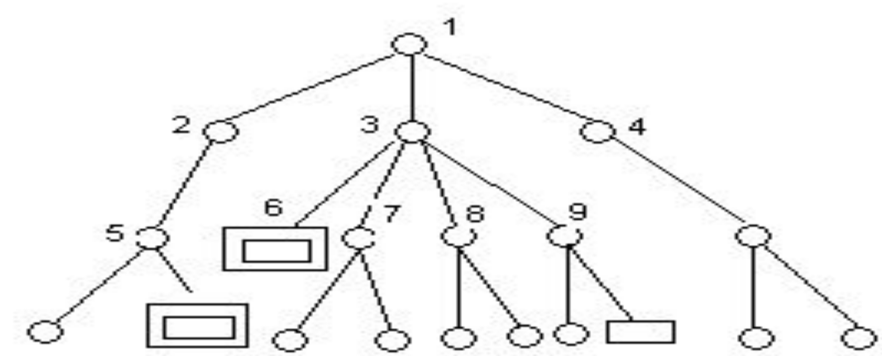
***T* – мн-во целевых состояний.**

Решение задачи: нахождение последовательности операторов *f1,f2,…,fk* (*fi* ∈*F*), которые преобразуют начальные состояния в конечные.

Поиск на графе состояний – это процесс построения графа *G*, содержащего целевую вершину.

а) Поиск в глубину б) Поиск в ширину





Эвристические методы поиска

Используются при наличии некоторых эмпирических правил, которые позволяют <u>сокращать объем вариантов решений.</u>

Открытие вершины стремятся упорядочить таким образом, чтобы процесс поиска распространился в <u>наиболее перспективных направлениях.</u>

Оценочная функция *f(n)* определяет направления поиска и является оценкой стоимости кратчайшего пути из начальной вершины в целевую при условии, что он проходит через вершину *n*.
При определении пути выбирается вершина с минимальным *f(n)*

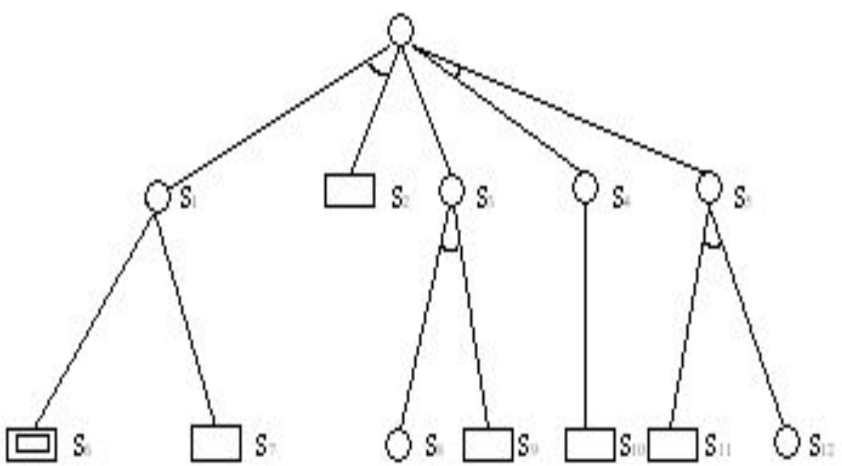
Специалисты, принимающие решения, используют:
* монотонные рассуждения
* рассуждения здравого смысла

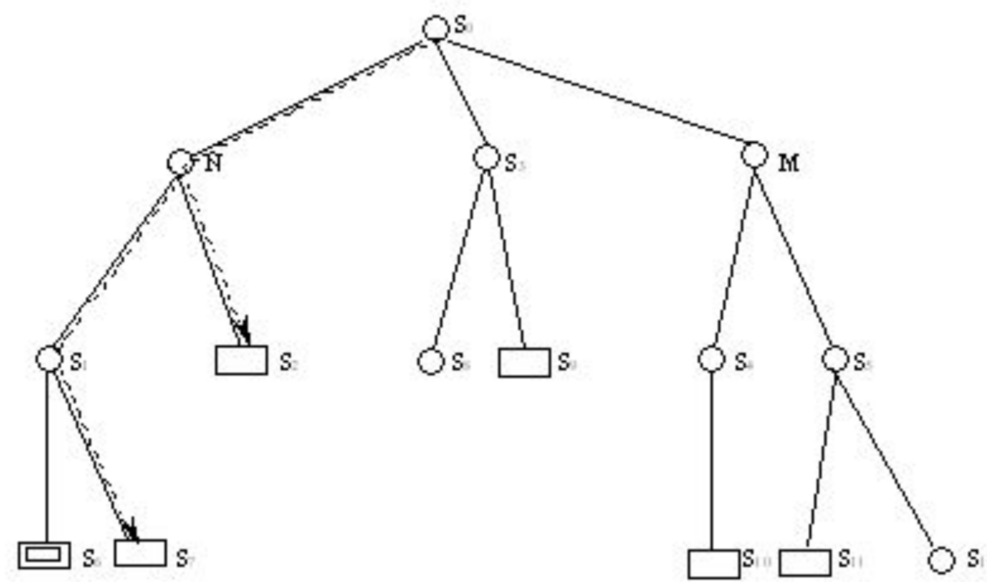
Метод редукции

Поиск необходимой совокупности данных для решения задачи сводится к решению составляющих подзадач

Процесс поиска решения задачи представляет собой граф И/ИЛИ. Граф содержит вершины типа “И” и типа “ИЛИ”.

Пример. Исходная задача *S0* разбивается на группы подзадач. Она может быть решена путем решения подзадач либо *S1* и *S2*; либо *S3*; *S4* и *S5*. Вершины *N*, *S3*, *M*, *S5* – это вершины типа “И”. Граф преобразуется и вводятся дополнительные вершины N и M, которые служат отдельными родительскими вершинами для подзадач *{S1, S2}* и *{S4, S5}.* Т.о., вершина *S0* преобразуется в вершину ИЛИ





Методы поиска решений в больших пространствах состояний

- **Метод порождения и проверки**

  Генератор, настроенный на проблемную область, порождает ряд неполных решений, осуществляется проверка неполных решений, и если решение признается недопустимым, то из дальнейшего рассмотрения исключается целый класс порождаемых им полных решений данного подпространства
- **Метод абстракции путем последовательного уточнения**

  Основная задача редуцируется на фиксированную совокупность описаний подзадач
- **Метод последовательного уточнения сверху**

  Решение задачи реализуется сверху вниз, от поиска и определения решения в абстрактном пространстве к преобразованию этого решения и его конкретизации на более низких (т. е. более подробных) уровнях описания
- **Эвристические методы и процедуры**

  Часто встроены в какие-либо другие методы

4.2. Выводы на фреймах и в семантических сетях
4.2.1. Вывод на фреймах

Структура данных фрейма

- Имя фрейма
- Имя слота
- Указатели наследования
- Указатель типа данных (атрибутов слотов)
- Значение слота
- Процедура – демон
- Присоединенная процедура (процедура – слуга)

Процедуры – демоны и присоединенные процедуры

Демон – процедура, автоматически запускаемая при выполнении некоторого условия.

Демоны обычно ограничены утверждениями ЕСЛИ-ТОГДА

Метод – это процедура, присоединенная к атрибуту фрейма, которая выполняется всякий раз, когда к ней обращаются.

Существует 2 типа методов:

* КОГДА-ИЗМЕНЕНО
* КОГДА-НЕОБХОДИМО

Методы подходят, если надо описать сложные процедуры.

Вывод во фреймовой системе

3 основных процесса, происходящие во фреймовых системах:

- Создание экземпляра фрейма
- Активация фреймов
- Организация вывода

Процедуры могут хранить знания, позволяющие давать ответ на следующие вопросы:

- Когда активировать фрейм? Подобно «демонам» фреймы могут активировать сами себя в случае, если распознана соответствующая ситуация.
- В каком случае считать, что данный фрейм неадекватен ситуации и что в этом случае делать? Фрейм мог бы, например, автоматически передать управление другому фрейму или деактивировать себя.
- Когда осуществлять заполнение слотов — в момент вызова или позднее, по мере необходимости?

4.2.2. Вывод в семантических сетях. Структурирование знаний в семантической сети

- Использование предикатов отношений двух типов:
  *является (IS-A)
  *часть (PART-OF)
- Семантическими сетями можно также представлять знания, касающиеся атрибутов объекта
- Вершины семантической сети обычно показывают объект проблемной области, концепт, ситуацию и т.п., а дуги – отношения между ними.
- Большинство систем с семантическими сетями имеет унифицированную структуру применительно к факторам действия и объекта по отношению к некоторому концепту (возможности наследования ожидаемых значений и значений по умолчанию, которые являются значениями атрибута в вершине экземпляра.)
- Проблемой, характерной для семантических сетей, является наследование атрибутов между иерархическими уровнями. (т.е. результат вывода, получаемого с помощью семантической сети, не гарантирует достоверность как логический формализм.)

Процедурные семантические сети

Процедурные семантические сети используются в целях введения единой семантики в семантической сети. Сеть строится на основе класса (понятия), а вершины, дуги (отношения) и процедуры представлены как объекты.

Процедурами определяются следующие действия над дугами:

- установление связи;
- аннулирование связи;
- подсчет числа вершин, соединенных заданной дугой;
- проверка наличия – отсутствия связи между заданными вершинами.

Существуют также процедуры, определяющие основные действия над вершинами, например:

- определение экземпляра класса;
- аннулирование экземпляра;
- подсчет числа экземпляров, принадлежащих к классу;
- проверка принадлежности экземпляр к некоторому классу.

Вывод в семантических сетях

Особенностью и недостатком семантической сети является целостность системы, не позволяющей разделить БЗ и механизм вывода.

Интерпретация семантической сети определяется с помощью использующих ее процедур, которые основаны на нескольких способах(например, способ сопоставления частей сетевой структуры).

Выводы в семантических сетях отличаются значительной полнотой, они сравнимы с нестандартными выводами процедурного представления и имеют ясную концептуальную интерпретацию.

Особый тип генерации вывода, используемый в семантических сетях, - это метод «распространяющейся активности и техники пересечений».

4.3. Дедуктивные методы вывода Вывод на предикатах

Основные формы логического вывода:

- **Индукция** (лат. наведение)
- **Дедукция** (лат. выведение)

Две формулы *P* и *G* эквивалентны или *P* эквивалентна *Q* (*P*=*Q*), когда истинные значения *P* и *Q* совпадают при каждой интерпретации *P* и *Q.*

Литера – это атомарная формула или ее отрицание.

Ф-ла *P* находится в конъюнктивной нормальной форме (КНФ), тогда и только тогда, когда *P* имеет вид $P \Leftrightarrow P1 \wedge \ldots \wedge Pn$, $n \geq 1$

Где каждая *P1, P2,…, Pn* есть дизъюнкция литер.

Ф-ла *P* находится в дизъюнктивной нормальной форме (ДНФ), тогда и только тогда, когда имеет вид $P \Leftrightarrow P1 \vee \ldots \vee Pn$, $n \geq 1$

Где каждая *P1, P2,…, Pn* есть конъюнкция литер.

Законы эквивалентных преобразований

4.1 $P \leftrightarrow Q = (P \to Q) \wedge (Q \to P)$

4.2 $P \to Q = \neg P \vee Q$

4.3а $P \vee (Q \wedge H) = (P \vee Q) \wedge (P \vee H)$

4.3б $P \wedge (Q \vee H) = (P \wedge Q) \vee (P \wedge H)$ – дистрибутивные законы

4.4 $\neg(\neg P) = P$- закон двойного отрицания

4.5а;4.5б - законы де Моргана

$$\neg(P \vee Q) = \neg P \wedge \neg Q \qquad \neg(P \wedge Q) = \neg P \vee \neg Q$$

4.6 P∨¬P=□ - закон исключенного третьего

4.7 P∧¬P=□ - закон противоречия

4.8 P ∨ □ =P; P ∧ □ =P

4.9 P ∨ □=□; P∧□

Правила вывода в логике высказываний

Правило подстановки. Пусть *P* – ППФ, содержащая атомарную ф-лу *X*. Тогда если *P* – тавтология, то, заменяя в ней *X* всюду, где она входит, произвольной ППФ *B*, получают также тавтологию.

Modus Ponens **Modus tollens**

$P \to Q$; $P \vdash Q$; $P \to Q$ $\neg Q \vdash \neg P$

Modus tollendo Ponens, **Modus ponendo tolleus**

$P \vee Q$ $\neg P \vdash Q$ $P \vee Q$ $P \vdash \neg Q$

Правило силлогизма, **Правило контрапозиции**

$P \to Q$ $Q \to R \vdash P \to R$ $P \to Q \vdash \neg Q \to \neg P$.

Теорема дедукции. Пусть даны формулы *P1,…,Pn* и формула *Q*. Тогда *Q* есть логическое следствие *P1,…,Pn* тогда и только тогда, когда формула общезначима.

Теорема о противоречии. Пусть даны формулы *P1,…,Pn* и формула *Q*. Тогда *Q* есть логическое следствие *P1,…,Pn* тогда и только тогда, когда формула противоречива.

Пример использования правил вывода

Пусть в некоторой ПО справедливы следующие утверждения:

1. f1: если скорость движения конвейера непостоянна(S), То точность захвата роботом заготовки и установки ее под пресс уменьшается(P).

2. f2: если точность установки роботом заготовки под пресс уменьшается (*P*), то увеличивается процент бракованных изделий(*U*)

3. f3: скорость движения конвейера непостоянна (S).

4. f4: увеличивается процент бракованных изделий (U).

$f1 : S \rightarrow P,\ F2{:}P \rightarrow U;\ F3{:}S;\ F4{:}U.$

f1 : S→P, F2:P→U; F3:S; F4:U.

Покажем, что F4 истинно, как только F1∧ F2∧F3 истинно.

Преобразуем флу ((S→P) ∧ (P→U) ∧S) в нормальную форму:

$$((S \to P) \land (P \to U) \land S) = ((\neg S \lor P) \land (\neg P \lor U) \land S)$$

$= (S \land (\neg S \lor P) \land (\neg P \lor U)) = (((S \land \neg S) \lor (S \land P)) \land (\neg P \lor U))$

$= ((\square \lor (S \land P)) \land (\neg P \lor U)) = (S \land P) \land (\neg P \lor U)$

$= (S \land P \land \neg P) \lor (S \land P \land U) = (S \land \square) \lor (S \land P \land U)$

$= \square \lor (S \land P \land U) = S \land P \land U$

Следовательно, если ((S→P) ∧ (P→U) ∧S) истинна, то (S∧P∧U)

истинна.Так как (S∧P∧U) истинна, только если S, P и U все

истинны, заключаем, что U истинна.

Здесь U есть логическое следствие ф-л f1, F2 и f3.

Типы символов, использующиеся для построения атомарных формул :

- индивидные символы (имена объектов), или константы (строчные буквы *a, b, c,*...);
- символы предметных переменных (строчные буквы *x,y,z);*
- функциональные символы (строчные буквы *f, g, h,...* или осмысленные слова из строчных букв);
- предикатные символы (прописные буквы *P, Q, R, ...* или осмысленные слова из прописных букв).

Основные правила в логике предикатов

- **Правила обобщения** (правило связывания квантором общности). Пусть Q — ф-ла, не содержащая свободных вхождений x; $P(x)$—ф-ла. Тогда если ф-ла $Q \to P(x)$ выводима, то ф-ла $Q \to \forall x P(x)$ также выводима.
- **Правило связывания квантором существования.** Пусть Q — ф-ла, не содержащая свободных вхождений x; $P(x)$—ф-ла. Тогда если ф-ла $P(x) \to Q$ выводима, то ф-ла $\exists x\, P(x) \to Q$ также выводима.
- **Правило универсальной конкретизации.** Пусть $P(y)$—ф-ла, свободная для y. Тогда из $\forall x P(x)$ выводится $P(y)$ подстановкой в $P(x)$ вместо переменной x переменной y, т. е. если предикат P выполняется для всех x, то он выполняется также для любого y.
- **Правило специализации.** Это правило применяется для определения конкретного значения квантора общности, т.е. если некоторому классу объектов присуще какое-либо свойство, то любой объект этого класса будет обладать этим свойством: $\forall x\ P(x)$, а $\vdash x P(a)$.
- **Правило конкретизации для квантора существования.** Это правило позволяет перейти от $\exists x\ P(x)$ к $P(a)$. Пусть а — определенный элемент, такой, что, если $\exists x\ P(x)$ истинно, то $P(a)$ также истинно. Тогда $\exists x\ P(x) \vdash P(a)$.

Законы, содержащие кванторы

4.10а $xP(x) \vee Q = x\{P(x) \vee Q\}$,

4.10б $xP(x) \wedge Q = x\{P(x) \wedge Q\}$,

4.11а $\forall xP(x) = \neg\exists x\neg P(x)$,

4.11б $\exists xP(x) = \neg\forall x\neg P(x)$,

4.12а $\neg\forall xP(x) = \exists x\neg P(x)$,

4.12б $\neg\exists xP(x) = \forall x\neg P(x)$,

4.13а $\forall xP(x) \wedge \forall xH(x) = \forall x\{P(x) \wedge H(x)\}$,

4.13б $\exists xP(x) \vee \exists xH(x) = \exists x\{P(x) \vee H(x)\}$,

4.14а $xP(x) \vee {}_2xH(x) = x\,y\{P(x) \vee H(y)\}$,

4.14б ${}_1xP(x) \wedge {}_2xH(x) = {}_1x\;{}_2y\{P(x) \wedge H(y)\}$

Процедура стандартизации - преобразования формул в предложения

(необходима для использования метода Эрбрана и метода резолюций)

Предложение – это множество дизъюнктов.
Дизъюнкт – это дизъюнкция литер.

- Удаление операторов $\leftrightarrow$ и $\rightarrow$ использованием законов (4.1), (4.2).
- Уменьшение области действия оператора $\neg$, введение его внутрь ф-лы. Законы (4.4), (4.5) и (4.12).
- Стандартизация переменных, т. е. переименование связанных переменных; если это необходимо.
- Приведение к ПНФ. Используют законы (4.10), (4.13) и (4.14) для вынесения кванторов, получения префикса и матрицы ф-лы.
- Приведение матрицы к КНФ. Закон (4.3а).
- Удаление квантора существования. Для этого производят сколемизацию ПНФ.
- Удаление квантора общности.
- Удаление оператора $\wedge$. Полученная форма является конъюнкцией множества дизъюнктов.

В результате этих преобразований можно получить стандартную форму в виде множества дизъюнктов.

Метод Эрбрана

Применяется процедура поиска опровержения, т.е. вместо доказательства общезначимости ф-лы доказывается, что опровержение ф-лы противоречиво.

Пусть ***S*** **— стандартная форма ф-лы *F,* представленная в виде мн-ва дизъюнктов.**

Тогда *F* противоречива в том случае, когда противоречива *S*.

Под *F* подразумевается отрицание исходной теоремы.

Исходная теорема общезначима, когда *S* противоречива.

Мн-во дизъюнктов невыполнимо тогда, когда оно ложно при всех интерпретациях на всех областях.

***H*(*S*) – универсум Эрбрана** – такая область, что если не существует удовлетворяющей интерпретации в этой области, то ее вообще не существует,т. е. *S* невыполнимо тогда, когда оно ложно при всех интерпретациях в этой области.

t1,…, tn - элементы универсума Эрбрана *H*(*S*).

Метод Эрбрана (продолжение)

***Pn*(*t1,…, tn*) - *n*-местная атомарная формула в *S*.**

Эрбрановская база для мн-ва *S* - мн-во основных атомарных ф-л вида *Pn*(*t1,…, tn*), встречающихся в *S*.

Задание интерпретации в области *H*(*S*) заканчивается для мн-ва *S* тогда, когда каждой атомарной ф-ле эрбрановской базы приписано значение истинности.

Для доказательства невыполнимости мн-ва дизъюнктов S необходимо порождать мн-ва S'i, …, S'n основных примеров дизъюнктов из S и последовательно проверять их на ложность.

Эта процедура согласно теореме Эрбрана обнаружит такое конечное *n*, что S'n невыполнимо.

Недостаток метода Эрбрана - экспоненциальный рост мн-ва *Si*- основных примеров дизъюнктов при увеличении *i.*

Для ограничения порождения мн-ва основных примеров эффективным является метод резолюций.

Метод резолюций

Цель метода- проверка невыполнимости мн-ва дизъюнктов.

Метод резолюции является правилом вывода, при использовании которого порождаются новые дизъюнкты из мн-ва *S*.

Идея метода резолюций заключается в проверке наличия в мн-ве *S* пустого (ложного) дизъюнкта •. Если *S* содержит •, то оно невыполнимо; если не содержит — то выводятся новые дизъюнкты до тех пор, пока не будет получен • (это всегда имеет место для невыполнимого *S*).

Правило резолюции: если в любых двух дизъюнктах *C1* и *C2* существует контрарная пара литер (L и ¬L), то, вычеркнув их, можно построить дизъюнкцию из оставшихся частей дизъюнктов *C1*и *C2*. Новый дизъюнкт есть резольвента дизъюнктов *C1* и *C2.*

Мн-во **{L, ¬L}** называется **контрарной парой,** если две литеры L и ¬L контрарны друг другу (L — атомарная ф-ла).

Метод резолюций (продолжение)

Если имеются два однолитерных дизъюнкта, образующих контрарную пару, то их резольвента есть пустой дизъюнкт •. Для невыполнимого S, применив правила резолюций, можно породить •.

Вывод пустого дизъюнкта • из *S* есть такая конечная последовательность *C1,C2,…,Ck* дизъюнктов, что каждый *Ci* или является дизъюнктом из *S*, или есть резольвента предыдущих дизъюнктов, полученная методом резолюции, и *Ck* = •.

Свойство полноты: мн-во дизъюнктов *S* невыполнимо тогда и только тогда, когда существует вывод пустого дизъюнкта из *S*.

Недостатки:

***** неограниченное применение правила резолюций может вызывать порождение большого числа дизъюнктов, многие из которых лишние и не относятся к делу

***** методы резолюций неприемлемы в случае сложных проблем, так как пространство поиска, образуемое ими, возрастает экспоненциально числу формул, используемых для описания проблемы.

4.4. Вывод в условиях неопределенности Неопределенность

Любая предметная область реального мира содержит неточные знания и нужно справляться с неполными, противоречивыми или отсутствующими данными.

Интеллектуальные СПР и ЭС должны быть в состоянии управлять неопределенностью.

Определяют четыре источника неопределенных знаний в интеллектуальных СПР:

- неизвестные данные,

- неточный язык,

- неявное смысловое содержание,

- трудности, связанные с сочетанием взглядов

- различных экспертов.

Наиболее известные подходы к управлению неопределенностью:

- Байесовское вероятностное рассуждение и его расширения,
- теорию уверенности,
- нечеткая логика.

4.4.1. Вероятностный вывод
Вероятностный подход
Основные положения байесовского метода и правило Байеса

Условная вер-ть p(A|B)– вер-ть того, что событие A произойдет, если произойдет событие B

p(A|B)= $\frac{\text{количество взможных совместных проявлений } A \text{ и } B}{\text{количество возможных проявлений } B}$

Совместная вер-ть A и B p(A∩B)- кол-во возможных совместных проявлений *A* и *B*, или вер-ть того, что *A* и *B* произойдут совместно

$$p(A \mid B) = \frac{p(A \cap B)}{p(B)}$$

где **p(B)** – вер-ть того, что произойдет событие B.

Условная вер-ть того, что произойдет событие B при условии, что имело место событие A определяется

$$p(B \mid A) = \frac{p(B \cap A)}{p(A)}$$

Отсюда **p(B∩A) =p(B|A)×p(A)**

Совместная вер-ть является коммутативной, т. о. **p(A∩B)= p(B∩ A)**

Следовательно, **p(A∩B) =p(B|A)×p(A)**

Правило Байеса:

$$p(A \mid B) = \frac{p(B \mid A) \times p(A)}{p(B)}$$

Основные положения байесовского метода и правило Байеса (продолжение)

Пусть событие *А* зависит не просто от события В, а от некоторого числа несовместимых событий *B1, B2, …, Bn.*

p(A∩B1) =p(A|B1)×p(B1)

Тогда из $p(A|B) = \frac{p(A \cap B)}{p(B)}$ следует p(A∩B2) =p(A|B2)×p(B2)

. . .

p(A∩Bn) =p(A|Bn)×p(Bn)

Или после объединения: $\sum_{i=1}^{n} p(A \cap B_i) = \sum_{i=1}^{n} p(A | B_i) \times p(B_i)$

Просуммировав ур-е по всему полному перечню событий Bi, получим $\sum_{i=1}^{n} p(A \cap B_i) = p(A)$

Выразим Р(А) с помощью ф-лы полной вер-ти: $p(A) = \sum_{i=1}^{n} p(A | B_i) \times p(B_i)$

Если проявление события *А* зависит только от двух взаимно исключающих событий, например В и НЕ В, тогда получим:

p(A)=p(A|B)×p(B)+p(A\¬B)×p(¬B), где ¬ - логическое НЕ.

Также p(B)=p(B|A)×p(A)+p(B\¬A)×p(¬A)

Подставив это ур-е в ф-лу Байеса, получим: $p(A|B) = \frac{p(B|A) \times p(A)}{p(B|A) \times p(A) + p(B|\neg A) \times p(\neg A)}$

Байесовский вывод

Пусть **правила БЗ** имеют вид: **Если Н истинно,тогда С истинно {с вер. р}**
Подставив Н и С вместо А и В в предыдущее ур-е, получим:

ная вер-ть

$$p(H \mid C) = \frac{p(C \mid H) \times p(H)}{p(C \mid H) \times p(\neg H) + p(C \mid \neg H) \times p(\neg H)}$$

p(C|H) – вер-ть того, что гипотеза Н истинна, будет результатом свидетельства С;
p(¬H) - **априорная вер-ть** того, что гипотеза Н ложна;
p(C|¬H) – вер-ть нахождения свидетельства С даже когда гипотеза Н ложна.
p(C|H) - апостериорная вер-ть гипотезы Н при наблюдаемом свидетельстве С.

Иногда для простого свидетельства С эксперт обеспечивает мн-во гипотез H1,H2,…, Hm:

$$p(H_i \mid C) = \frac{p(C \mid H_i) \times p(H_i)}{\sum_{k=1}^{M} p(C \mid H_k) \times p(H_k)}$$

Или при мн-ве свидетельств *C1, C2, …, Cn* также имеется мн-во гипотез:

$$p(H_i \mid C_1, C_2, \ldots, C_n) = \frac{p(C_1, C_2, \ldots, C_n \mid H_i) \times p(H_I)}{\sum_{k=1}^{m} p(C_1, C_2, \ldots, C_n \mid H_k) \times p(H_k)}$$

Это ур-е требует получения условных вер-тей всех возможных комбинаций свидетельств для всех гипотез, что делает задачу невыполнимой.
Поэтому вместо неосуществимого ур-я

$$p(A \mid B) = \frac{p(A \cap B)}{p(B)}$$

мы получаем:

$$p(H_i \mid C_1, C_2, \ldots, C_n) = \frac{p(C_1 \mid H_i) \times p(C_2 \mid H_i) \times \ldots \times p(C_n \mid H_i) \times p(H_i)}{p(C_1 \mid H_k) \times p(C_2 \mid H_k) \times \ldots \times p(C_n \mid H_k) \times p(H_k)}$$

Байесовский вывод. Пример

Пусть при наличии 3х условно независимых свидетельства ***C1, C2* и *C3*,** Эксперт создает 3 взаимно исключающие гипотезы ***H1, H2* и *H3*** и обеспечивает априорные вер-ти для этих гипотез – ***p(H1), p(H2)* и *p(H3)*,** а также определяет условные вер-ти каждого отмеченного свидетельства для всех возможных гипотез.

Вероятности	Гипотезы		
	i= 1	i= 2	i= 3
P(H)	0.4	0.35	0.25
P(C₁|H)	0.3	0.8	0.5
P(C₂|H)	0.9	0.0	0.7
P(C₃|H)	0.6	0.7	0.9

ла

ИС вычисляет апостериорные вер-ти для всех гипотез по ур-ю **4.28**:

$$p(H_i \mid C_3) = \frac{p(C_3 \mid H_i) \times p(H_i)}{\sum_{k=1}^{3} p(E_3 \mid H_i) \times p(H_k)}, i = 1,2,3$$

Таким образом

$$p(H_1 \mid C_3) = \frac{0{,}6 \times 0{,}4}{0{,}6 \times 0{,}4 + 0{,}7 \times 0{,}35 + 0{,}9 \times 0{,}25} = 0{,}34$$

$$p(H_2 \mid C_3) = \frac{0{,}7 \times 0{,}35}{0{,}6 \times 0{,}4 + 0{,}7 \times 0{,}35 + 0{,}9 \times 0{,}25} = 0{,}34$$

$$p(H_3 \mid C_3) = \frac{0{,}9 \times 0{,}25}{0{,}6 \times 0{,}4 + 0{,}7 \times 0{,}35 + 0{,}9 \times 0{,}25} = 0{,}32$$

Как можно видеть, после того, как наблюдается свидетельства ***C3*,** доверие гипотезе ***H2*** и становится равным доверию гипотезе ***H1*.** Доверие гипотезе ***H3*** также возрастает и даже приблизительно достигает доверию гипотезам ***H1*** и ***H2*.**

Байесовский вывод. Пример (продолжение)

Предположим теперь, что мы наблюдаем свидетельство ***С1***. Апостериорные вер-ти рассчитываются по ур-ю **4.30**:

$$p(H_i \mid C_1C_3)=\frac{p(C_1 \mid H_i)\times p(C_3 \mid H_i)\times p(H_i)}{\sum_{k=1}^{3} p(C_1 \mid H_i)\times p(C_3 \mid H_i)\times p(H_k)}, i=1,2,3$$

Таким образом

$$p(H_i \mid C_1C_3)=\frac{0{,}3\times 0{,}6\times 0{,}4}{0{,}3\times 0{,}6\times 0{,}4+0{,}8\times 0{,}7\times 0{,}35+0{,}5\times 0{,}9\times 0{,}25}=0{,}19$$

$$p(H_2 \mid C_1C_3)=\frac{0{,}8\times 0{,}7\times 0{,}35}{0{,}3\times 0{,}6\times 0{,}4+0{,}8\times 0{,}7\times 0{,}35+0{,}5\times 0{,}9\times 0{,}25}=0{,}52$$

$$p(H_3 \mid C_1C_3)=\frac{0{,}5\times 0{,}9\times 0{,}25}{0{,}3\times 0{,}6\times 0{,}4+0{,}8\times 0{,}7\times 0{,}35+0{,}5\times 0{,}9\times 0{,}25}=0{,}29$$

Гипотеза ***Н2*** теперь рассматривается как наиболее вероятная. После наблюдения свидетельства ***С2*** СПР вычисляет апостериорные вер-ти для всех гипотез:

$$p(H_i \mid C_1C_2C_3)=\frac{p(C_1 \mid H_i)\times p(C_2 \mid H_i)\times p(C_3 \mid H_i)\times p(H_i)}{\sum_{k=1}^{3} p(E_1 \mid H_i)\times p(C_2 \mid H_i)\times p(C_3 \mid H_i)\times p(H_k)}, i=1,2,3$$

Таким образом

Первоначальным ранж-ем было

(*С1, С2* и *С3*)

зы ***Н1*** и ***Н3.***

ыть

$$p(H_i \mid C_1C_2C_3)=\frac{0{,}3\times 0{,}9\times 0{,}6\times 0{,}4}{0{,}9\times 0{,}3\times 0{,}6\times 0{,}4+0{,}8\times 0{,}0\times 0{,}7\times 0{,}35+0{,}5\times 0{,}7\times 0{,}9\times 0{,}25}=0{,}45$$

$$p(H_2 \mid C_1C_2C_3)=\frac{0{,}8\times 0{,}0\times 0{,}7\times 0{,}35}{0{,}3\times 0{,}9\times 0{,}6\times 0{,}4+0{,}8\times 0{,}0\times 0{,}7\times 0{,}35+0{,}5\times 0{,}7\times 0{,}9\times 0{,}25}=0$$

$$p(H_3 \mid C_1C_2C_3)=\frac{0{,}5\times 0{,}7\times 0{,}9\times 0{,}25}{0{,}3\times 0{,}9\times 0{,}6\times 0{,}4+0{,}8\times 0{,}0\times 0{,}7\times 0{,}35+0{,}5\times 0{,}7\times 0{,}9\times 0{,}25}=0{,}55$$

Вывод на основе теории Демпстера-Шафера

Проблемами байесовского подхода являются:

* необходимость заранее устанавливать априорные вер-ти каждого свидетельства
* байесовская вер-ть не позволяет эффективно описать незнание.

Теория вер-тей Демпстера-Шафера была введена для представления субъектной ненадежности. Она не фиксирует значения вер-ти, и может представлять и незнание.

Демпстер предложил такие понятия, как **нижняя и верхняя вер-ти.**
Шафер переименовал их в **функцию доверия и меру правдоподобия**

Этот подход вводит различия м/у неопределенностью и незнанием путем создания **функций доверия,** позволяющих использовать знания, при ограниченной возможности присвоения вер-тей.

Теория подходит для объединения экспертных мнений.

Но теория не приспособлена к ситуации, когда знания экспертов частично совпадают, что часто встречается в реальности.

4.4.2. Вывод на основе теории уверенности (ТУ) Коэффициент уверенности (КУ) и доверие

Неопределенность представляется как степень уверенности.

Существуют **2 стадии** в невероятностном методе неопределенности.
1:необходимо иметь возможность выразить степень уверенности.
2: необходимо манипулировать (объединять и т.д) степени уверенности в процессе использования системы, основанной на знаниях.
КУ выражают доверие событию (факту или гипотезе), основанное на свидетельстве (или оценке эксперта).

Пример: 1 или 100 означает абсолютную истину и 0 уверенную неправду. Если мы говорим: «существует 90% возможности, что пойдет дождь, то либо идет дождь (90%) либо нет дождя (10%)». При невероятностном подходе говорят: КУ дождь=90 означает, что очень возможно пойдет дождь.
ТУ предлагает понятия ***доверия*** и ***недоверия*,** объединяющиеся ф-лой: **КУ(В,С)=МД(В,С)-МН(В,С)**,

где ***КУ*** – коэффициент уверенности*,* ***МД*** – мера доверия
МН – мера недоверия*,* ***В*** – вероятность*,**С*** – свидетельство или событие.

Объединение коэффициентов уверенности

Наиболее приемлемым способом объединения КУ является подход, используемый в EMYCIN. При этом подходе существуют 2 случая:

I. Объединение нескольких КУ в одном правиле.

Рассмотрим это правило **с оператором *И***:

Если инфляция высокая, *КУ*=50%, *(A)*, ***И***
Если ур. безработицы больше 7%,КУ=70%,(B), ***И***
Если цены облигации снижаются, *КУ*=100%, *(C)*
Тогда биржевые цены снижаются.

Для того, чтобы заключение было истинным, все *Если* должны быть истинны.

<u>КУ(A,B и C)=minimum[КУ(A), КУ(B), КУ(C)]</u>.

Т.о., ***КУ*** для утверждения «биржевые цены снижаются» будет **50%.**

Теперь посмотрим на это правило **с оператором ИЛИ:**

Если инфляция низкая, *КУ*=70%, ***ИЛИ***
Если цены облигаций высокие, *КУ*=85%;
Тогда биржевые цены будут высокими.

Достаточно, чтобы только одно из Если являлось истинным для того, чтобы заключение было истинным. Т.о., когда <u>обеим</u> Если доверяют как истинным (по их *КУ*), тогда заключение будет иметь КУ максимальное значение из двух:

КУ(A или B)=maximum[КУ(A), КУ(B)].

В нашем случае ***КУ=85%,*** что биржевые цены вырастут.

Объединение КУ (продолжение)

Допустим, что имеется 2 правила:

Правило 1

Если уровень инфляции меньше, чем 5%,

Тогда биржевые цены поднимаются (*КУ*=0.7)

Правило 2

Если уровень безработицы меньше, чем 7%

Для **3го** добавленного правила используется формула:

КУ(П1,П2, П3)=КУ(П1,П2)+КУ(П3)[1-КУ(П1П2)]

Пусть **Правило 3:**

Если цены на облигации возрастают

Тогда биржевые цены возрастают (КУ=0,85).

Если все правила истинны в своей левой части *Если*, возможность того, что биржевые цены возрастут определится как

КУ(П1П2П3)=0,88+0,85(1-0,88)=0,88+0,85(0,12)=0,982

Т.е., существует возможность на уровне 98,2%, что биржевые цены возрастут

4.4.3. Нечеткая логика и приближенные рассуждения

Истинность высказывания в нечеткозначной логике опр-ся значениями типа: ***истинно, ложно, очень истинно, абсолютно истинно, не очень истинно, очень ложно и т.п.***

Приближенные рассуждения - процесс получения из нечетких посылок некоторых следствий.

* В классической теории исчисления высказываний выражение **«Если *A*, Тогда *B*»,** где ***A*** **и *B* – пропозициональные пер-ые,** записывается как $A \to B$, где импликация ($\to$) рассматривается как связка, смысл которой определяется таблицей истинности.

Таким образом $\boldsymbol{A \to B \equiv \neg A \vee B}$ **4.31**

Это значит, что $\boldsymbol{A \to B}$ **(А влечет В)** и $\boldsymbol{\neg A \vee B}$ **(не А или В)** имеют идентичные таблицы истинности.

* Более важным в нашем случае является неопределенное высказывание **«Если *A*, Тогда *B*» (*A*→*B*)**, в котором ***A*(антецедент) и *B* (консеквент) – нечеткие мн-ва**, а не пропозициональные пер.

Типичные примеры высказываний:

Если «большой», Тогда «малый»

Если «скользкий», Тогда «опасный;»

Они являются сокращениями предложений:

Если х - «большой», Тогда у - «малый»;

Если дорога «скользкая», Тогда езда «опасна».

Приближенные рассуждения (продолжение)

Пусть ***A*-нечеткое подмн-во области рассуждений *U*** и пусть В - **нечеткое подмн-во другой области рассуждений V**.

Декартово произведение *A* и *B* определяется так:

$$A \times B \stackrel{\Delta}{=} \int_{U \times V} \mu_A(u) \wedge \mu_B(v) / (u,v) \qquad \textbf{4.32}$$

где ***U×V*** означает декартово произведение мн-в *U* и *V*, т.е.

$$U \times V \stackrel{\Delta}{=} \{(u,v) / u \in U, v \in V\}$$

4.32 означает, что *A×B* нечеткое мн-во упорядоченных пар (*u,v*), $u \in U, v \in V$ со степенью принадлежности *(u,v)* к *(A×B)*, задаваемой ф-лой $\mu_A(u) \wedge \mu_B(v)$. То есть *A×B* есть нечеткое отношение *U* и *V*.

Пример.

Пусть ***U=****1+2* ***V=****1+2+3* **A=**1/1+0,8/2 **B=***0,6/1+0,9/2+1/3* ***Тогда***

A×B=0,6/(1,1)+0,9/(1,2)+1/(1,3)+0,6/(2,1)+0,8/(2,2)+0,8/(2,3)

Отношение, определенное **A→B≡¬A∨B** (4.31) можно представить матрицей отношения

$$\begin{array}{cc} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 0,6 & 0,9 & 1 \\ 0,6 & 0,8 & 0,8 \end{bmatrix} \end{array}$$

«Если *A*, Тогда *B*» можно рассматривать как **«Если *A*, Тогда *B*, Иначе *C*»** где *A*, *B* и *C* – нечеткие подмн-ва, возможно, различных областей *U* и *V*.

В терминах декартова произведения последнее предложение опр-ся так:

Если *A*, Тогда *B*, Иначе *C* $\stackrel{\Delta}{=} A \times B + (\neg A \times C)$

Приближенные рассуждения (продолжение)

Чтобы обобщить понятие материальной импликации на нечеткие мн-ва, предположим, что *U* и *V* – 2 возможно различных универсальных мн-ва, а *A, B* и *C* – нечеткие подмн-ва мн-в *U, V* и *V* соответственно.

$$\textbf{\textit{Если A, Тогда B, Иначе C}} = A\times B+\neg A\times C \qquad \textbf{4.40}$$

То есть, ***A, B* и *C*** – унарные нечеткие отношения в ***U, V* и *V*,** тогда **Если *A*, Тогда *B*, Иначе *C*** – бинарное нечеткое отношение в ***U×V*,** которое является объединением декартова произведения *A* и *B* и декартова произведения отрицания *A* и *C*.

Далее высказывание **Если *A*, Тогда *B*** можно рассматривать как частный случай высказывания **Если *A*, Тогда *B*, Иначе *C*** *при допущении,* что ***C* – полное множество *V*.**

Т.о. $\textbf{\textit{Если A, Тогда B}} \overset{\Delta}{=} \textbf{\textit{Если A, Тогда B, Иначе}}\ V = A\times B+\neg A\times V$ **4.41**

Если A,Тогда B равнозначно ***Если A, Тогда B, Иначе безразлично***

Приближенные рассуждения
Иллюстрация 4.40 и 4.41

Пусть, **U=V=1+2+3** **B=большой=0,4/2+1/3**

A=малый=1/1+0,4/2 **C=не большой=1/1+0,6/2**

Тогда

Если *A*, Тогда *B*, Иначе *C*=

=(1/1+0,4/2)×(0,4/2+1/3)+(0,6/2+1/3)×(1/1+0,6/2)=

=0,4/(1,2)+1/(1,3)+0,6/(2,1)+0,6/(2,2)+0,4/(2,3)+1/(3,1)+0,6/(3,2)

Что можно представить в виде матрицы отношения

Если A, Тогда B, Иначе C= $\begin{bmatrix} 0 & 0,4 & 1 \\ 0,6 & 0,6 & 0,4 \\ 1 & 0,6 & 0 \end{bmatrix}$

Аналогично

Если *A*, Тогда *B*=*(1/1+0,4/2)×(0,4/2+1/3)+(0,6/2+1/3)×(1/1+1/2+1/3)=*

=0,4/(1,2)+1/(1,3)+0,6/(2,1)+0,6/(2,2)+0,6/(2,3)+1/(3,1)+1/(3,2)+1/(3,3)

Или эквивалентно

Если A, Тогда B= $\begin{bmatrix} 0 & 0,4 & 1 \\ 0,6 & 0,6 & 0,6 \\ 1 & 1 & 1 \end{bmatrix}$

Композиционное правило вывода (КПВ)

Пусть U и V – 2 универсальных мн-ва с базовыми пер-ми u и v соответственно. Пусть $R(u)$, $R(u,v)$ и $R(v)$ обозначают ограничения на u, (u,v) и v соответственно и представляют собой нечеткие отношения в U, $U \times V$ и V.

Пусть A и F – нечеткие подмн-ва мн-в U и $U \times V$.

Тогда **КПВ утверждает, что решение урав-й назначения**

$R(u)=A$ (унарное нечеткое отношение)

$R(u,v)=F$ (бинарное нечеткое отношение)

Имеет вид $R(v)=A°F$

Где $A°F$ – композиция A и F. В этом смысле мы должны делать вывод $R(v)= A°F$ из того, что $R(u)=A$ и $R(u,v)=F$.

При этом функция принадлежности определяется как

$$\mu_R(v) = \max\left[\min(\mu_R(u), \mu_R(u,v))\right].$$

Для определения КПВ применяют нечеткие отношения.

Например, нечеткое отношение Заде:

$$R_m = (A \times B) \cup (\neg A \times V) = \int_{U \times V} (\mu_A(u) \wedge \mu_B(v)) \vee (1 - \mu_A(u)) / (u,v)$$

Композиционное правило вывода. Иллюстрация

U=V=1+2+3+4   *A=малый=1/1+0.6/2+0./3*

F= примерно равны=1/(1,1)+1/(2,2)+1/(3,3)+
1/(4,4)+0,5/((1,2)+(2,1)+(2,3)+(3,2)+(3,4)+(4,3)).

Т.е., *A*-унарное нечеткое отношение в *U*, названное «малый»;

F – бинарное нечеткое отношение в *U*×*V*, названное «примерно равны».

Уравнения назначения в этом случае имеют вид:

R(u)= малый   *R(u,v)= примерно равны,* И следовательно

R(v)= малый° примерно равны =

$$[1 \quad 0{,}6 \quad 0{,}2 \quad 0]\circ\begin{bmatrix} 1 & 0{,}5 & 0 & 0 \\ 0{,}5 & 1 & 0{,}5 & 0 \\ 0 & 0{,}5 & 1 & 0{,}5 \\ 0 & 0 & 0{,}5 & 1 \end{bmatrix}=[1 \quad 0{,}6 \quad 0{,}5 \quad 0{,}2],$$

Что можно аппроксимировать сле

R(v)= более или менее малый.

Примечание 1: ур-ие назначения для *X* имеет вид *x=u:R(x)* или эквивалентно *x=u*, *u∈R(X),* и отражает то, что эл-ту *x* *назначается* значение *u* с учетом ограничения *R(X)*.

Т.е., мы вывели, что *R(v)=[1 0.6 0.5 0.2]* точно и *R(v)=«более или менее малый»* -в качестве лингвистического приближения.

Словами этот приближенный вывод записывается так:

u-*«малый»* (предпосылка)*u* и *v – «примерно равны»* (предпосылка)

v – «более или менее малый»   (приближенный вывод)

Напоминание

Лингвистической называют пер-ую, если ее значениями являются слова или фразы естественного языка.

При этом лингвистическая пер-ая может быть описана набором

(*X, T, U, G, M*),

где ***X*** — название пер-ой;

T - терм-множество пер-ой *X,* т. е. мн-во всех названий лингвистических значений пер-ой *x,* причем каждое из таких значений является нечеткой переменной *x* со значениями из универсального мн-ва ***U*** *с* базовой пер-ой *и;*

G *-* синтаксическое правило, порождающее названия *X* значений пер-ой *x;*

M — семантическое правило, которое ставит в соответствие каждой нечеткой пер-ой x смысл *M(x).* Конкретное название *X,* порожденное синтаксическим правилом G, называется термом.

Нечеткое подмн-во *M* - некоторое мн-во в *X* с ф-ей принадлежности μм(*x*), принимающей значения из интервала [0, 1]:

где **μм** : *X* → [0, 1] – ф-ия принадлежности.

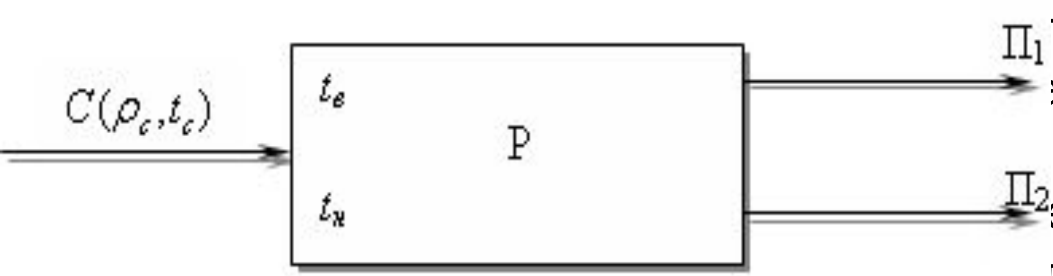
$$M = \int_{x \in X} \mu_M(x) / x$$

Применение приближенных рассуждений и композиционного правила вывода

(на примере управления технологической установкой)

С – сырье, ***tc*** – температура,

ть



авляющие пар-ры:

tн - темп-ра низа

ерха тех. колонны

целевые продукты

tc, pc, P, tн, tв будем рассматривать как **лингвистические пер-ые.**

Из них ***tн, tв, P*** - наиболее информативные для упр-я.

Определим области изменения каждого параметра:

для tн *[90... ...300] °С,* ***для tв*** *[200... 400] °С,* ***для Р*** *[40. ..70] атм.*

Для ***tн, tв, P*** определены (на основе опросов) **терм-множества:**

T (темп-ра *tн*)= низкая + средняя + высокая;

T (темп-ра *tв*)= низкая + средняя + высокая;

T (давление Р) = ниже нормы + норма + выше нормы.

Для **лингв. пер-х,** описывающих выход *П1* и *П2,* области изменения: [90, 300] **для П1** и [120, 250] **для П2.**

Терм-множества: *T* (выход продукта *П1*)= малый + средний + большой;

T (выход продукта *П2*) = малый +средний + большой.

Продолжение примера

Наименование лингв. пер-й	Термы	Носитель нечеткого множества
Температура низа колонны $t_н$	Низкая	90—220
	Средняя	190—250
	Высокая	245—300
Температура верха колонны $t_в$	Низкая	200—250
	Средняя	240—330
	Высокая	321—400
Давление P	Ниже нормы	40—51
	Норма	50—63
	Выше нормы	62—70
Выход продукта $П1$	Малый	90—158
	Средний	147—236
	Большой	230—300
Выход продукта $П2$	Малый	120—156
	Средний	151—210
	Большой	208—250

ма из

троится

носителем,

нове экспертных оценок.

ого мн-ва -

$X' = \{x \mid \mu_M(x) > 0, x \in X'\}$.

бор правил,

Правила имеют вид: **Если *A* = *N*, тогда *B* = *Q*,**
где ***A, B*** — лингв. пер-ые; ***N, Q*** - термы лингв. пер-х.
Например, одно из правил: **Если** темп-ра низа колонны = <низкая>

Тогда выход продукта *П1*= (большой)
На основе набора правил строится матрица нечетких отношений.

Напоминание

Как известно, **нечеткое бинарное отношение *R* есть подмн-во декартового произведения *X* × *Y*:**

$$R = \int_{(x,y)\in X\times Y} \mu_R(x,y)/(x,y).$$

Далее построенные матрицы объединяются $\mu_R = \max(\mu_{R_1},\dots,\mu_{R_n})$.

Для организации логического вывода используется КПВ, частным случаем которого является известный силлогизм Modus Ponens

ПОСЫЛКА 1 ЕСЛИ *x* есть *A,* Тогда *y* есть *B*

<u>ПОСЫЛКА 2 *x* есть *A* .</u>

ВЫВОД *y* есть *B*

Композиционное правило Л. Заде утверждает, что композиция нечетких отношений *R*(*и*) = *A, R*(*u,v*) = *F* есть *R* (*v*) = *A* o *F*,

где o - знак композиции.

При этом ф-ия принадлежности определяется как

$$\mu_R(v) = \max[\min(\ \mu_R(u), \mu_R(u,v))].$$

Окончание примера

Исходя из правила Заде и учитывая, что нечеткие мн-ва определены в виде ***N* = низкая = 1/100 + 0,4/200, *Q* = большая =0,4/200 + 1/300,** получаем:

ЕСЛИ *А* = *N*, Тогда *В* = *Q* = *N* × *Q* + ¬ *N* × *V* = (1/100 + 0,4/200) × (0,4/200 + 1/300) + (0,6/200+ 1/300) × (1/100+ 1/200 + 1/300) = 0,4(100, 200) + 1/(100, 300)+0,6/(200, 100)+0,6/(200, 200) + + 0,6/(200, 300)+ 1/(300, 100)+ 1/(300, 200) + 1/(3ОО, 300).

Последнее можно записать *в* виде матрицы:

$$S = \begin{vmatrix} 0 & 0.4 & 1 \\ 0.6 & 0.6 & 0.6 \\ 1 & 1 & 1 \end{vmatrix}$$

Пусть теперь текущее значение температуры низа колонны

***Ni* = почти низкая =1/100+ 0,4/200+0,2/300.**

Для определения выхода продукта *П* необходимо применить КПВ

$$Q_1 = N \circ S = (1/100 + 0.4/200 + 0.2/300) \circ \begin{vmatrix} 0 & 0.4 & 1 \\ 0.6 & 0.6 & 0.6 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 0.4 & 0.4 & 1 \end{vmatrix}$$

т. е. ***Qi* =0,4/100+0,4/200+1/300**. Условно полученный

результат может быть интерпретирован как **«почти большой».**