

ИИС

01

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

Подходы к созданию систем искусственного интеллекта

- * **1) Нейрокибернетический подход** основан на построении самоорганизующихся систем, состоящих из множества элементов, функционально подобных нейронам головного мозга.
- * **Основную идею этого подхода** можно сформулировать следующим образом:
- * Единственный объект, способный мыслить, — это человеческий мозг. Поэтому любое «мыслящее» устройство должно каким-то образом воспроизводить его структуру.
- * **2) Логический подход** основан на выявлении и применении в интеллектуальных системах различных логических и эмпирических приемов (эвристик), которые применяет человек для решения каких-либо задач. **Эвристика** – правило теоретически не обоснованное, но позволяющее сократить количество рассматриваемых вариантов в процессе поиска.
- * **В основу этого подхода был положен принцип**, противоположный нейрокибернетике.

Данные и знания

- * **Процедурная** информация о веществе («растворена») в программах (алгоритмах), которые выполняются в процессе решения задач
- * **декларативная** – в данных, с которыми эти программы работают.

особенности (свойства) декларативных знаний

- * 1. **Идентифицируемость.** Каждая информационная единица должна иметь уникальное имя, по которому информационная система выделяет ее среди других.
- * 2. **Внутренняя интерпретируемость.** Информационные единицы, как правило, содержат сведения о том, как их использовать.
- * 3. **Структурированность.** Информационные единицы должны иметь четкую, но в тоже время гибкую структуру, позволяющую задавать рекурсивные¹ определения информационных единиц или вложенность одних информационных единиц в другие.
- * 4. **Связность.** В информационной базе между информационными единицами должна быть предусмотрена возможность установления связей различного типа. Например, таких как функциональные отношения, причинно-следственные связи, структурные и семантические отношения, отношения несовместимости информационных единиц в едином контексте, отношение релевантности и т.д. Отношение релевантности (смысловой и ситуационной близости) информационных единиц позволяет находить знания, близкие к уже найденным.
- * 5. **Активность.** В традиционных информационных системах данные пассивны, а команды активны. В ИИС, как и у человека, актуализации тех или иных действий способствуют знания, имеющиеся в системе. Таким образом, выполнение программ в ИИС инициируется текущим состоянием информационной базы, т.е. декларативными знаниями.

ОСНОВНЫЕ ОТЛИЧИЯ ЗНАНИЙ ОТ ДАННЫХ

- * - знания более структурированы;
- * - знания самоинтерпретируемы;
- * - в знаниях существенное значение имеют не только атомарные элементы (как в данных), но и взаимосвязи между ними;
- * - знания активны в отличие от пассивных данных, т. е. знания могут порождать действия системы, использующей их.

- 
- * **Знания** — это закономерности предметной области (факты, принципы, связи, законы), полученные в результате практической деятельности и профессионального опыта, позволяющие специалистам ставить и решать задачи в этой области.
 - * **Знания** — это хорошо структурированные данные (данные о данных, метаданные).

База данных и база знаний

- * Для хранения данных используются базы данных (для них характерны большой объем и относительно небольшая удельная стоимость информации), для хранения знаний — базы знаний (небольшого объема, но исключительно дорогие информационные массивы). **База знаний** — это совокупность знаний, описанных с использованием выбранной формы (модели) их представления. База знаний — основа любой интеллектуальной системы

Классификация знаний по степени основательности

- * - **поверхностные (конкретные, экстенсиональные)** - конкретные факты, сведения о видимых взаимосвязях между отдельными событиями и фактами в предметной области, эвристики и закономерности, полученные опытным путем и используемые при отсутствии общих теорий.
- * Пример: «Если нажать на кнопку звонка, раздастся звук. Василий Петрович - инженер второй категории. Иванов имеет автомобиль»;
- * - **глубинные (абстрактные, интенциональные)** - совокупность основных закономерностей и аксиом в конкретной предметной области. В эту совокупность могут входить абстракции, аналогии, формулы, законы, правила, методики, схемы, отображающие структуру и природу процессов, и т.д. Эти знания объясняют явления и могут использоваться для прогнозирования поведения объектов.
- * Пример: «Электрическая схема звонка. Инструкция по определению наличной провозной и пропускной способности станций и перегонов. Закон Бойля-Мариотта».

по степени теоретической обоснованности

- * - **фактические** - хорошо известные (аксиомы) и теоретически обоснованные законы и правила, а также конкретные достоверные факты и сведения;
- * - **эвристические** - знания, теоретически не обоснованные, но подтвержденные в результате многолетней практики или экспериментов;

по способу представления и использования в интеллектуальных информационных системах

- * - **декларативные (факты);**
- * - **процедурные (правила);**
- * - **метазнания (знания о знаниях)** - знания, предписывающие способы и порядок использования знаний, а также их свойства

по степени детерминированности (определенности)

- * **детерминированные (точные)** - знания, воспринимаемые всеми однозначно (например, Путин президент России, $S_{\text{круга}} = \pi R^2$, в сутках 24 часа). Как правило, это аксиомы, теоретические обоснованные законы и количественные характеристики. В любом случае, эти знания могут быть интерпретированы либо как полностью истинные (true, 1) либо как полностью ложные (false, 0);
- * - **недетерминированные (нечеткие, неточные)** - знания, воспринимаемые неоднозначно или степень достоверности которых отличается от 0 или 1 (например, достоверность того, что Иванов умный - 0.7). Как правило, это качественные характеристики, такие как рост (высокий, средний, низкий, карликовый), возраст (старый, пожилой, в расцвете сил, молодой, юный)

Модели представления знаний

Наименование	Способ представления знаний	Примечания
Алгоритмические	Тексты программ. Блок-схемы.	
Продукционные	Правила «ЕСЛИ (условие) - ТО (результат или действие)».	Подвид алгоритмических моделей.
Семантические сети	Граф, вершины которого - понятия, а дуги - связи между ними	В вершинах находятся атомарные понятия.

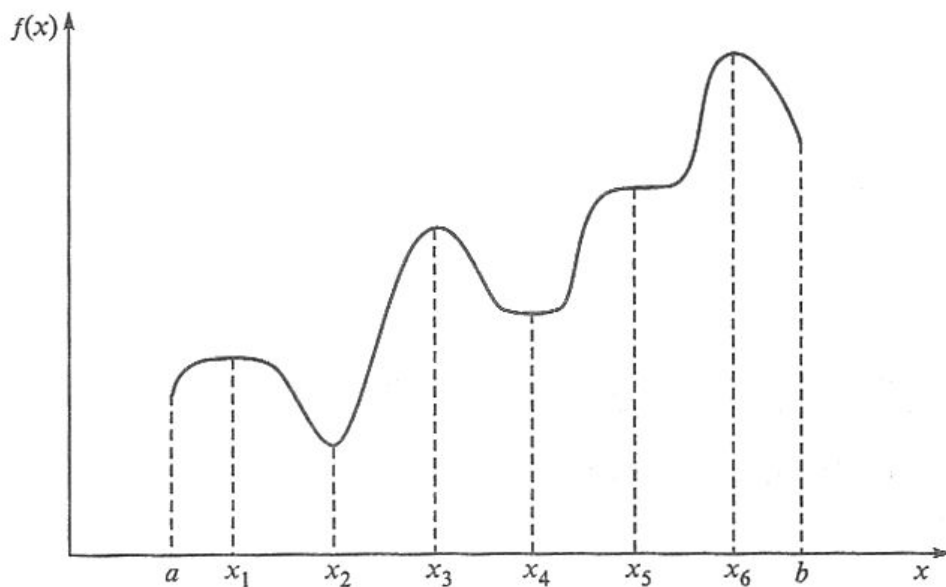
Наименование	Способ представления знаний	Примечания
Диаграммы «сущность-связь» (Entity-Relationship Diagrams, ERD)	Аналогично семантическим сетям.	Используются для отображения состава и структуры БД.
Фреймы	Структуры, описывающие взаимосвязь неких концептуальных понятий (объектов, ситуаций, законов, правил и т.д.).	
Объектно-ориентированные	Тексты программ на языках объектно-ориентированного программирования. Унифицированный язык моделирования (UML, Unified Modeling Language) - диаграммы, описывающие систему (объект).	Диаграммы отображают состав, структуру и поведение системы (объекта). Каждая диаграмма, как правило, представляется связным графом. Используются, в основном, при проектировании информационных систем с использованием объектно-ориентированного подхода.

Наименование	Способ представления знаний	Примечания
Логические	Формальный язык с ограниченным синтаксисом и семантикой.	Логика высказываний. Логика предикатов первого порядка. Нечеткая логика. И т.д.
Сценарии	Описание взаимосвязанных фактов или стандартной последовательности действий в определенной ситуации.	
Искусственные нейронные сети	Ориентированный граф, состоящий из множества нейроподобных элементов и связей между ними.	

КРАТКИЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ОПТИМИЗАЦИИ

- 
- * **Оптимизация** (в математике, информатике и исследовании операций) - задача нахождения экстремума (минимума или максимума) целевой функции в некоторой области конечномерного векторного пространства, ограниченной набором линейных и/или нелинейных равенств и/или неравенств.
 - * **Оптимизация** - процесс приведения системы в наилучшее (оптимальное) состояние.

- 
- * В задачах оптимизации **имеется целевая функция**, зависящая от одной или нескольких переменных. Требуется найти такие значения переменных, при которых функция примет оптимальное (наилучшее) значение. При этом на переменные могут накладываться различные ограничения.
 - * Наилучшее значение функции достигается в так называемых точках минимума или максимума (**экстремальных точках**)



- * точка x_0 является точкой максимума функции $f(x)$, если соблюдается неравенство
- * $f(x_0 - \Delta x) \leq f(x_0) \geq f(x_0 + \Delta x)$,
- * где Δx – достаточно малое приращение значения переменной x .
- * - точка x_0 является точкой минимума функции $f(x)$, если соблюдается неравенство
- * $f(x_0 - \Delta x) \geq f(x_0) \leq f(x_0 + \Delta x)$.
- * для интервала $[a, b]$ точки x_1 , x_2 , x_3 , x_4 и x_6 являются экстремальными точками.

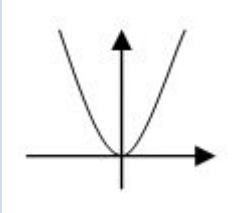
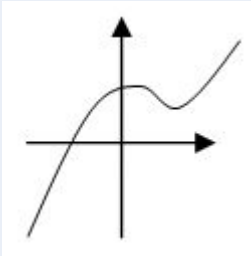
Значение $f(x_6)$ называется глобальным или абсолютным максимумом, а значения $f(x_1)$ и $f(x_3)$ – локальными или относительными максимумами. Аналогично, $f(x_2)$ – глобальный минимум и $f(x_4)$ – локальный минимум.

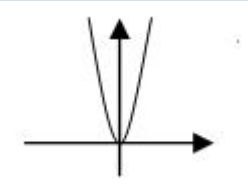
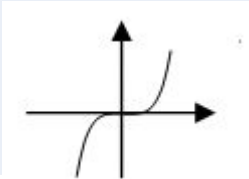
x_1 , являясь точкой максимума функции $f(x)$, отличается от остальных локальных максимумов $f(x)$ т.к. по крайней мере в одной точке ее окрестности значение функции $f(x)$ совпадает с $f(x_1)$. Точка x_1 по этой причине называется нестрогим (слабым) максимумом функции, а точки x_3 и x_6 – строгими максимумами. Нестрогий максимум, следовательно, подразумевает наличие (бесконечного количества) различных точек, которым соответствует одно и то же максимальное значение функции.

Классическая теория оптимизации

- * первая производная функции $f(x)$ (тангенс угла наклона касательной к графику функции) равна 0 во всех ее экстремальных точках. Однако это условие выполняется и в точках перегиба и седловых точках (x_5)

- * x_i , являющиеся решениями уравнения $f'(x) = 0$, – это либо точки экстремума либо точки перегиба. Условие $f'(x) = 0$ **называют необходимым условием наличия экстремума.**
- * Исследование производных высших порядков позволяет убедиться, что точка x_i – экстремум и более того, является она точкой максимума или минимума. Для этого необходимо найти вторую производную $f''(x)$ и в нее подставить значения x_i , полученные при решении уравнения $f'(x) = 0$:
 - * - если $f''(x_i) > 0$ – то в точке x_i минимум функции;
 - * - если $f''(x_i) < 0$ – то в точке x_i максимум функции;
 - * - если $f''(x_i) = 0$ – то необходимо исследовать следующие производные. В этом случае, если первые $(n-1)$ производных равны 0 и $f^{(n)}(x_i) \neq 0$, то в точке x_i функция имеет:
 - * - точку перегиба, если n – нечетное;
 - * - минимум, если n – четное и $f^{(n)}(x_i) > 0$;
 - * - максимум, если n – четное и $f^{(n)}(x_i) < 0$.
- * **Эти условия называются достаточными**

№ п/п	Функция	Экстремумы	График функции
1	$y = f(x) = x^2$	$f'(x) = 2x$ $2x = 0$ $x_0 = 0$ – точка экстремума или перегиба. $f''(x) = 2 \rightarrow 2 > 0$ $x_0 = 0$ – точка минимума.	
2	$y = f(x) = x^3 - 2x^2 + x + 1$	$f'(x) = 3x^2 - 4x + 1$ $3x^2 - 4x + 1 = 0$ $x_1 = 1/3$ и $x_2 = 1$ – точки экстремума или перегиба. $f''(x) = 6x - 4 \rightarrow$ $f''(x_1) = 6 * 1/3 - 4 = -2$ $f''(x_1) < 0$ $x_1 = 0$ – точка максимума; $f''(x_2) = 6 * 1 - 4 = 2$ $f''(x_2) > 0$ $x_2 = 0$ – точка минимума.	

№ п/п	Функция	Экстремумы	График функции
3	$y = f(x) = 2x^4$	$f'(x) = 8x^3$ $8x^3 = 0$ $x_0 = 0$ – точка экстремума или перегиба. $f''(x) = 24x^2$ $f''(x_0) = 24 * 0^2 = 0.$ $f^{(3)}(x) = 48x$ $f^{(3)}(x_0) = 48 * 0 = 0.$ $f^{(4)}(x) = 48$ $48 > 0$ и $n = 4$ – четное $x_0 = 0$ – точка минимума.	
4	$y = f(x) = 2x^3$	$f'(x) = 6x^2$ $6x^2 = 0$ $x_0 = 0$ – точка экстремума или перегиба. $f''(x) = 12x$ $f''(x_0) = 12 * 0^2 = 0.$ $f^{(3)}(x) = 12$ $n = 3$ – нечетное $x_0 = 0$ – точка перегиба.	

НЕЧЕТКИЕ ЗНАНИЯ И СПОСОБЫ ИХ ОБРАБОТКИ

Недетерминированность выводов

- * **Недетерминированность выводов.** Это характерная черта большинства интеллектуальных информационных систем. Недетерминированность означает, что заранее путь решения конкретной задачи в пространстве ее состояний определить невозможно. Поэтому в большинстве случаев методом проб и ошибок выбирается некоторая цепочка логических заключений, согласующихся с имеющимися знаниями, а в случае если она не приводит к успеху, организуется перебор с возвратом для поиска другой цепочки и т.д.
- * Например, выезжая на автомобиле, следует учитывать состояние дорог, транспорта, погодные условия и т.д. При нарушении одного из предположений, например, из-за пробки на обычном маршруте, планы меняются и выбирается альтернативный маршрут.

Многозначность

- * **Многозначность.** Многозначность интерпретации — обычное явление в задачах распознавания. При понимании естественного языка серьезными проблемами становятся многозначность смысла слов, их подчиненности, порядка слов в предложении и т.п. Проблемы понимания смысла возникают в любой системе, взаимодействующей с пользователем на естественном языке. Распознавание графических образов также связано с решением проблемы многозначной интерпретации.

Неточность и ненадежность знаний и выводов

- * **Неточность и ненадежность знаний и выводов.** Как было отмечено выше, количественные данные (знания) могут быть неточными. Неточность в основном связана с объективными причинами: несовершенство измерительных приборов (школьной линейкой нельзя измерять объекты меньше миллиметра или больше километра), несоблюдения условий проведения замеров (повышенная или пониженная температура, влажность и т.п.) и т.д. При этом существуют различные способы оценки такой неточности, разрабатываемые в рамках теории измерений.

- * Ненадежность знаний в большей степени связана с субъективными причинами: **отсутствием формальных процедур получения точных данных**, вероятностной природой поступающих данных, недостаточной математической (логической) обоснованностью используемых правил.
- * Ненадежность означает, что для оценки достоверности знаний нельзя применить двухбалльную шкалу (1 – абсолютно надежные, 0 – недостоверные).
- * Применяют вероятностные оценки тех или иных знаний, как в части фактов, так и правил вывода. Так, утверждение $p(\text{высокий(вася)}) = 0.75$ можно интерпретировать как вероятность того, что Вася высокий на три четверти истинна. Утверждение $p(\text{ЕСЛИ настроение_преподавателя} = \text{«хорошее» И знание_ответа_на_билет} = \text{«нулевые» ТО оценка_за_экзамен} = \text{«не меньше 3»}) = 0.8$ определяет вероятность истинности правила.

Неполнота знаний и немонотонная логика

- * Абсолютно полных знаний не бывает, поскольку процесс познания бесконечен.
- * Состояние базы знаний должно изменяться с течением времени.
- * При добавлении новых знаний возникает опасность получения противоречивых выводов, т.е. выводы, полученные с использованием новых знаний, могут опровергать те, что были получены ранее.

Модель закрытого мира

- * **Модель закрытого мира** предполагает жесткий отбор знаний, включаемых в базу
- * база знаний заполняется исключительно верными понятиями, все, что ненадежно или неопределенно, заведомо считается **ЛОЖНЫМ**
- * Все, что известно базе знаний, является истиной, а остальное - ложью.
- * Модель имеет ограниченные возможности представления знаний и таит в себе опасность получения противоречий при добавлении новой информации.

Монотонность

- * В системах, построенных по принципу модели закрытого мира, добавление новых фактов **не нарушает справедливость ранее полученных выводов**. Это свойство логических выводов называется **МОНОТОННОСТЬЮ**.
- * Ральные знания, закладываемые в интеллектуальных информационных системах, крайне редко бывают полными.

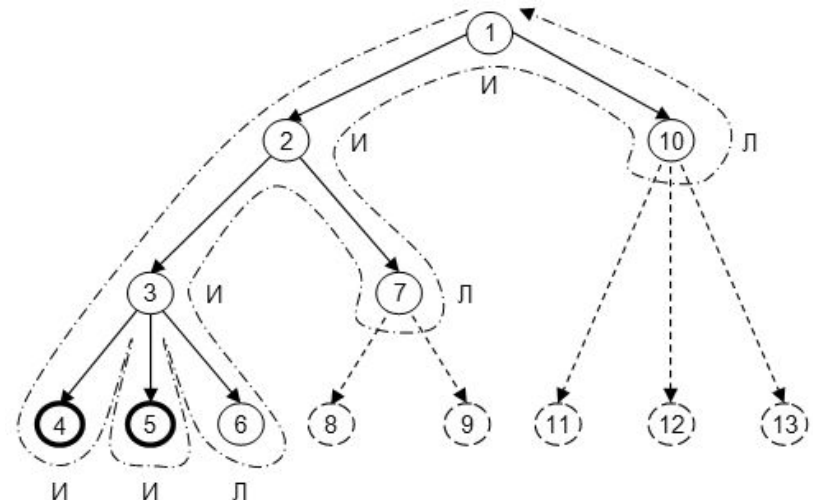
- 
- * «Птицы летают».
 - * «Пингвин не летает».
 - * «Пикколо — птица».

- * можно получить заключение «Пикколо летает»
- * сделать вывод о том, что «Пингвин не является птицей».
- * Если в базу знаний добавить факт «Пикколо — пингвин», то получим противоречащие предыдущим заключения: «Пикколо не летает» и «Пикколо не является птицей».

Способы устранения и/или учета нечеткости знаний

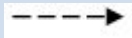
Учет недетерминированности вывода. метод перебора с возвратами

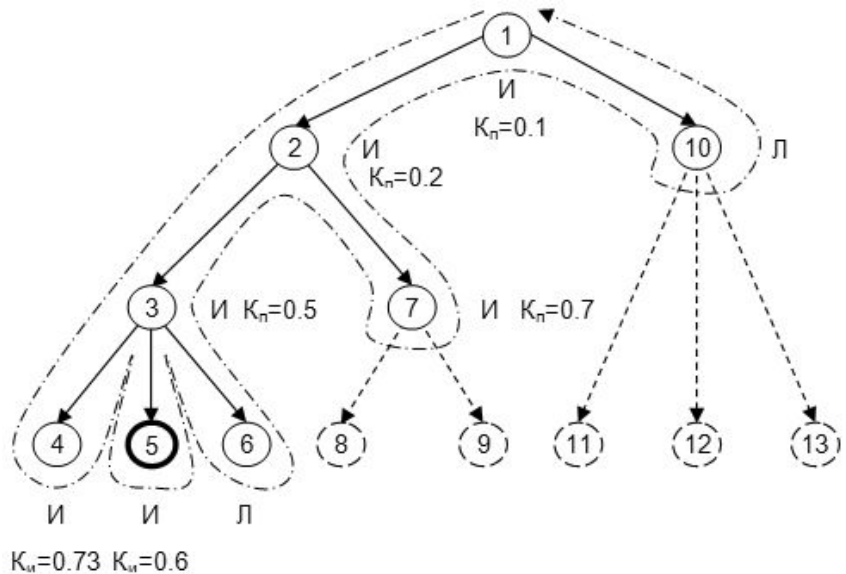
- * Все пространство поиска можно представить в виде дерева, узлами которого являются частичные или итоговые решения задачи (необязательно верные).
- * По мере поиска решения, удовлетворяющего условиям (требованиям, ограничениям), постепенно строится это дерево (выполняется обход узлов).
- * Если в листьях (узлах последнего уровня) решение удовлетворяет требуемым условиям, то оно и есть результат поиска. В общем случае, решений может быть несколько (узлы 4 и 5).
- * Если при обходе дерева система попадает в узел, решение в котором не удовлетворяет (противоречит) условиям задачи, тогда система возвращается к предыдущему узлу и продолжает поиск в альтернативном направлении



Условные обозначения

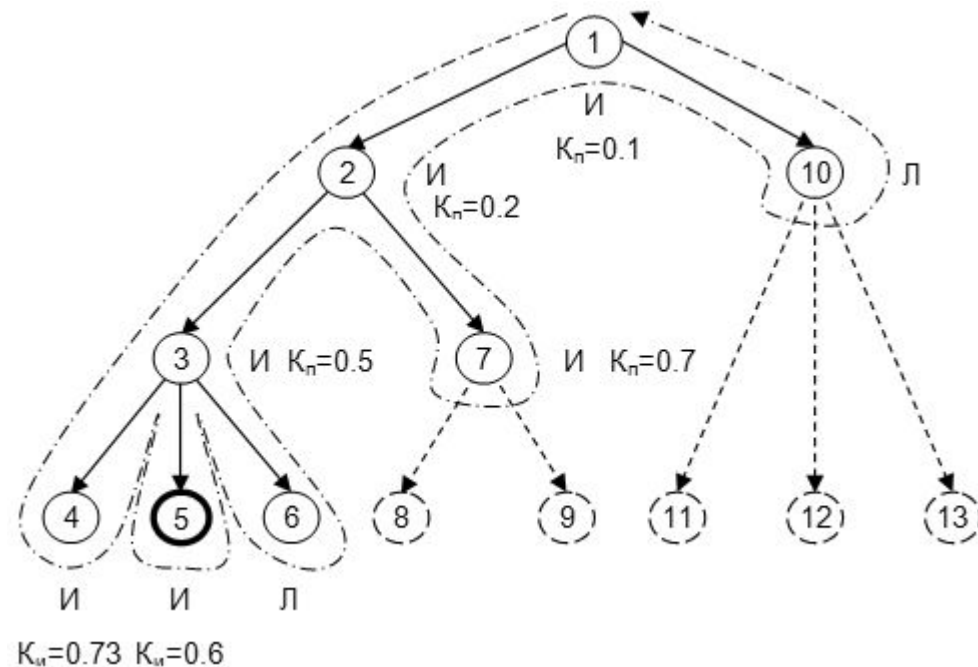
 направление поиска из узлов,
решение в которых удовлетворяет

	направление поиска из узлов, решение в которых удовлетворяет условиям
	направление поиска из узлов, решение в которых не удовлетворяет условиям
	маршрут обхода узлов
	узел, в котором проверялись условия. Рядом с узлом стоит «И», если решение в нем удовлетворяет условиям, в противном случае стоит «Л»
	узел, непопавший в маршрут обхода

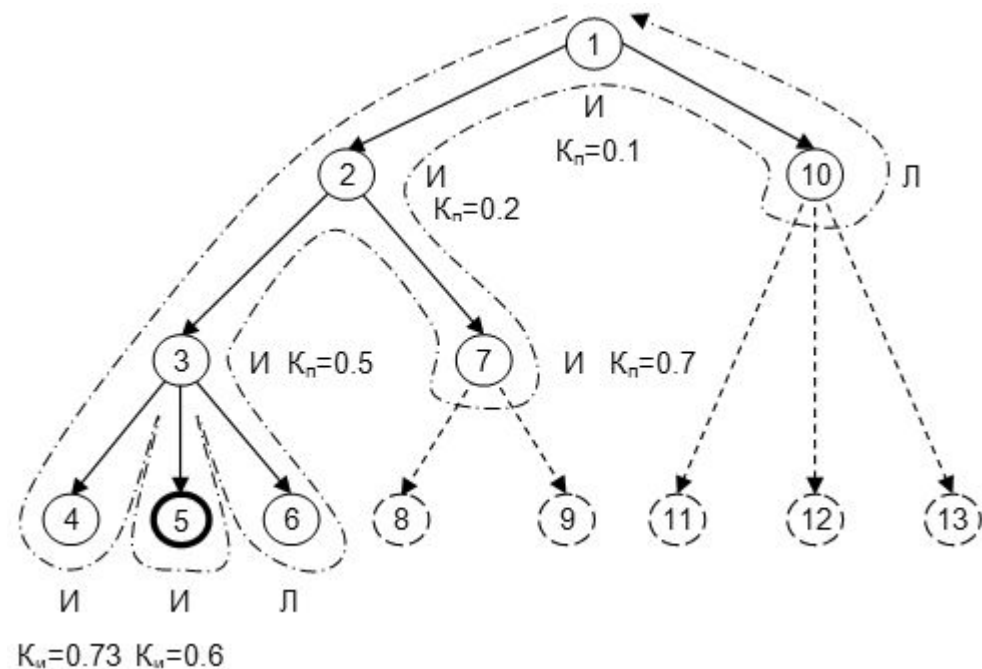


метод частичного (неявного) перебора

- * в узлах помимо проверки допустимости (соответствия ограничениям) считается и проверяется значение выбранного критерия.
- * Если значение этого критерия (при минимизации целевой функции) в некотором узле J больше, чем значение, полученное в другом итоговом узле I (представляющем допустимое решение задачи), то дальнейший поиск из узла J не ведется (т.к. значение критерия не станет лучше)

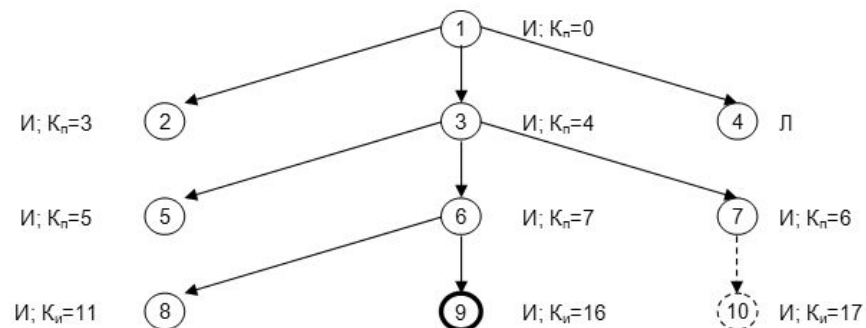


- * Оптимальным решением в рассмотренном примере является решение в узле 5.
- * Поиск решения из узла 7 не выполнялся, т.к. решения в узлах 8 и 9 дадут заведомо худшее решение по сравнению с узлом 5.
- * Ограничением использования этого метода является как монотонность выводов, так и монотонность критерия (целевой функции).
- * значение критерия в узле, из которого ведется поиск, не может быть больше значения критерия в ниже лежащих узлах. В случаях, когда надо максимизировать целевую функцию, считают величину обратную критерию $1/K$.

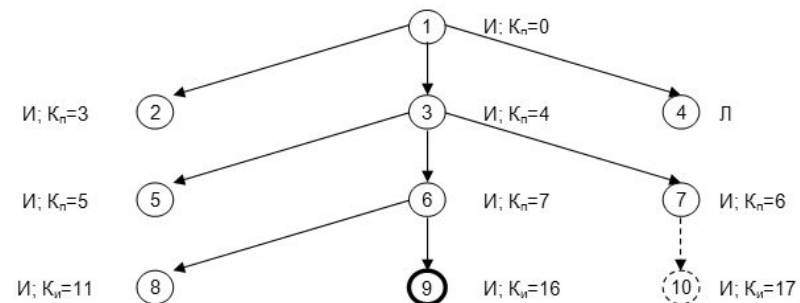


алгоритмом A*

- * Этот эвристический алгоритм используют в тех случаях, когда даже частичный обход полного дерева выполнить нереально.
- * Для всех узлов, удовлетворяющих ограничениям и в которые можно попасть из корневого узла, вычисляется значение критерия.
- * Дальнейший процесс поиска выполняется только из узла, в котором значение критерия максимально (минимально), остальные ветви поиска не рассматриваются.



- * Недостатком данного алгоритма является возможный пропуск оптимального решения (в примере узел 10), но найденное итоговое решение очень часто является оптимальным или, по крайней мере, эффективным.



Устранение многозначности

- * Вопросами устранения многозначности смысла слов, фраз и предложений занимается **теория формальных грамматик**.
- * Развитием классической логики является **многозначная логика**. Помимо значений истинности **true** и **false** в рассуждениях используется другие значения, например, **unknown**. Такой подход может обеспечить отделение ложных утверждений от утверждений, истинность которых просто неизвестна.

- 
- * - коэффициенты уверенности;
 - * - нечеткие множества и нечеткая логика;
 - * - вероятностный подход на основе теоремы Байеса;
 - * - модифицированный байесовский подход;
 - * - теория доказательства (обоснования) Демпстера-Шафера.
 - *

Коэффициенты уверенности

- * **Коэффициент уверенности (КУ)** – это неформальная оценка, которую эксперт добавляет к заключению. КУ вычисляется по формуле
- * $KU(H | E) = MD(H | E) - MND(H | E)$,
- * где $KU(H | E)$ - уверенность в гипотезе H с учетом свидетельств E ;
- * $MD(H | E)$ и $MND(H | E)$ — мера доверия и недоверия к гипотезе H при свидетельствах E .
- *
- * MD и MND измеряются от 0 (абсолютная ложь) до 1 (абсолютная истина), включая промежуточные значения. Соответственно, KU может изменяться от -1 (абсолютная ложь) до +1 (абсолютная истина), причем 0 означает полное незнание.
- * Формула не позволяет отличить случай противоречащих свидетельств (MD и MND обе велики) от случая недостаточной информации (MD и MND обе малы), что иногда бывает полезно.