

## **Автоматизированное извлечение конструкторско– технологических знаний**

*Национальный аэрокосмический университет им.Н.Е.Жуковского «ХАИ»*

Приобретение знаний (knowledge acquisition) в какой-либо предметной области есть передача потенциального опыта решения проблемы от некоторого источника знаний и преобразование его в вид, который позволяет использовать эти знания в программе [1,2]. Передача знаний выполняется в процессе достаточно длительных и пространных собеседований между инженером по знаниям и экспертом в предметной области (экспертом–технологом), способным достаточно четко сформулировать имеющийся у него опыт. Производительность такого процесса довольно низкая, поэтому функция приобретения конструкторско–технологических знаний является одним из главных «узких мест» при разработке и использовании систем гибкого автоматизированного проектирования технологических процессов (ТП) и создании конструкторско–технологической базы знаний предприятия с использованием нечетких экспертных правил [3]. Неудовлетворительные результаты такого подхода во многих предметных областях привели к необходимости автоматизации процесса передачи знаний специалистом машине и появлению автоматизированного извлечения знаний (automated knowledge elicitation).

### **Цель статьи**

Статья посвящена разработке методов и алгоритмов автоматизированного извлечения конструкторско–технологических знаний с использованием справочных баз данных предприятия, архива техпроцессов.

База знаний предприятия представляет собой набор модулей знаний, который пополняется в процессе эксплуатации системы.

Одним из простейших способов извлечения знаний у экспертов–технологов с использованием программных средств является технология извлечения знаний о предметной области посредством опроса через терминал в диалоговом режиме. Эксперту–технологу предлагается заполнить экранные формуляры, информация о которых затем считывается в структурированные объекты. Процедура извлечения знаний имеет циклический характер, алгоритм ее представлен на рис.1.

Данная методика ориентирована на частичную автоматизацию процесса извлечения знаний в ходе активного диалога интервьюируемого эксперта с программой и предполагает использование стратегии, направляемой знаниями о предметной области.

Одной из проблем является сложность использования баз конструкторско–технологических данных, хранилищ данных, накопленных на предприятии в результате использования различных диалоговых подсистем проектирования ТП для создания базы знаний. Решение этой проблемы возможно следующим

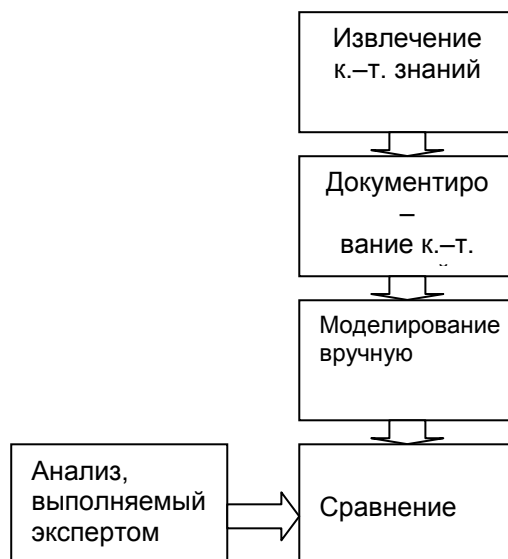


Рис. 1. Циклическая процедура приобретения конструкторско–технологических знаний

образом: 1 – использование различных справочников, хранящихся в базах данных предприятия, 2 – использование архива техпроцессов для автоматизированного извлечения знаний. В [4] рассмотрен алгоритм формирования нечетких экспертных правил выбора методов обработки поверхностей деталей с использованием имеющейся на предприятии справочной базы данных. Предложенный в [4] метод формирования исходного множества технологических альтернатив (ИМТА) для выбора методов обработки поверхностей позволяет использовать имеющуюся на предприятии справочную базу данных для генерирования нечетких экспертных правил базы конструкторско–технологических знаний (БКТЗ) предприятия. Недостатком известных в data–mining методов является то, что при их использовании генерируется большое количество правил, предоставляемых эксперту в проблемной области для оценки их «интересности» с дальнейшим использованием алгоритмов фильтрации [5]. Сложность представляет также формулировка посылок правил ЕСЛИ–ТО с привлечением экспертов проблемной области.

Существует два подхода при построении алгоритмов, генерирующих правила – использование логических и вероятностных методов [6]. В предыдущих статьях [4,7] описан алгоритм Reference–Knowledge\_module\_discovery (RKMD) автоматизированного извлечения конструкторско–технологических знаний, который позволяет генерировать минимальное множество нечетких экспертных правил (для выбора методов обработки поверхности) структуры, представленной в [4]. Алгоритм Reference–Knowledge\_module\_discovery построен на основе логических методов.

Преимуществами предложенного алгоритма является возможность автоматизации процесса генерации предпосылок нечетких экспертных правил благодаря использованию справочной информации баз конструкторско–технологических данных, что позволяет сократить сроки создания базы знаний

предприятия. Формулировка посылок правил производится в соответствии с классификатором поверхностей [3]. Алгоритм Reference–Knowledge\_module\_discovery генерирует минимальное множество правил, предоставляемых эксперту в проблемной области (эксперту–технологу) благодаря использованию конструкторско–технологических справочников базы данных предприятия, что позволяет сократить затраты на создание базы знаний (рис.2).

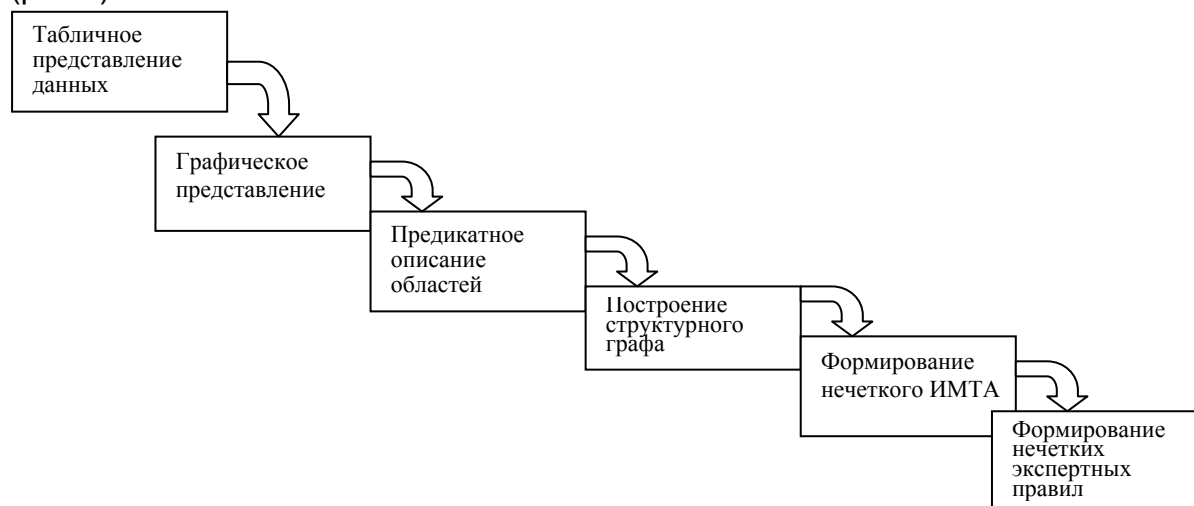


Рис. 2. Алгоритм RKMD формирования нечетких экспертных правил с использованием справочной базы данных предприятия

```

Колличество альтернативных методов обработки:
P1 100000000 1
P2 110000000 2
P3 010000000 1
P4 001000000 1
P5 001000001 2
P6 001010000 2
P7 001110000 3
P8 001110001 4
P9 000100001 2
P10 000000001 1
P11 000001000 1
P12 000101000 2
P13 000101100 3
P14 000101001 3
P15 000101101 4
P16 000000101 2
P17 000000100 1
P18 000000111 3
P19 000000010 1

Postroenie strukturnogo grafa
metodov obrabotki:
R1 <P1, P2>
R2 <P2, P3>
R3 <P4, P5, P6, P7, P8>
R4 <P7, P8, P9, P12, P13, P14, P15>
R5 <P6, P7, P8>
R6 <P11, P12, P13, P14, P15>
R7 <P13, P15, P16, P17, P18>
R8 <P18, P19>
R9 <P5, P8, P9, P10, P14, P15, P16, P18>

Podoblast zagotovki:
1
Podoblast detali:
6

Koli4estvo putej: 1
Put 1,2,3,
Koli4estvo putej: 2
Put 1,2,3,4,5,
Put 1,2,3,5,
  
```

Рис. 3. Результаты работы алгоритма Reference–Knowledge\_module\_discovery.

Сгенерированные правила в диалоговом режиме выводятся на экран эксперту–технологу для задания функций принадлежности каждой альтернативе и удаления «неинтересных» и некорректных с точки зрения технолога правил путем задания значения  $\mu_i=0$ . Данный алгоритм был апробирован на примере

справочной информации о методах обработки наружных цилиндрических поверхностей. промежуточные результаты работы программы представлены на рис. 3. Программа сгенерировала 51 нечеткое экспертное правило для эксперта–технолога.

Данный алгоритм позволяет генерировать минимальное количество правил

$$N = \sum_{i=1}^{mz} (np - i),$$

где  $np$ ,  $mz$  – количество подмножеств  $P_i$  значений  $IT$ ,  $R_a$  для поверхности детали и заготовки [4,7].

Количество правил, которые были бы сгенерированы путем полного перебора  $2^{np}=2^9=512$ , ( $np$  – количество методов обработки [4]), большинство из которых окажутся отвергнутыми при дальнейшем анализе экспертом–технологом.

Генерацию нечетких экспертных правил для создания базы конструкторско–технологических знаний можно также осуществлять с использованием метода, основанного на объединении ассоциативных и неожиданных (*unexpected*) правил [5], полученных путем извлечения знаний из архива техпроцессов, баз конструкторско–технологических данных предприятия. При таком подходе формулировка посылок правил также производится в соответствии с классификатором поверхностей [3]. Для рассмотренного примера построения нечетких экспертных правил выбора методов обработки наружных цилиндрических поверхностей множество технологических альтернатив  $\Omega$  в общем случае состоит из  $N$  элементов – альтернативных методов обработки. Генерируется нечеткое экспертное правило вида

**ЕСЛИ** *поверхность=(основная, наружная, круглая в поперечном сечении, цилиндрическая)*

**И** *заготовка*( $IT, R_a$ ) (<от...до..>, <от...до...>)

**И** *требуемые*( $IT, R_a$ ) (<от...до..>, <от...до...>)

**ТО** *методы обработки* =  $\{(\Omega_1, \mu_1), (\Omega_2, \mu_2), \dots, (\Omega_n, \mu_n)\}$ ,

содержащее некоторое количество альтернативных методов обработки поверхности. С помощью запроса извлекаются все записи из базы данных – архива техпроцессов, удовлетворяющие посылке правила ЕСЛИ. Пусть таких записей окажется  $k$ . Количество альтернативных методов обработки  $\Omega_i$  среди этих записей обозначим  $L$ . Тогда значение функции принадлежности  $\mu_i (i = \overline{1, L})$  для каждого альтернативного метода обработки  $\Omega_i$  в правиле вычисляется по формуле:

$$\mu_i = \frac{NM_L}{k},$$

где  $NM_L$  – количество записей, извлеченных из архива техпроцессов, для которых применялся метод обработки номер  $L$ . В дальнейшем полученные автоматизировано нечеткие экспертные правила предъявляются для анализа и фильтрации эксперту–технологу. Блок–схема алгоритма представлена на рис. 4.

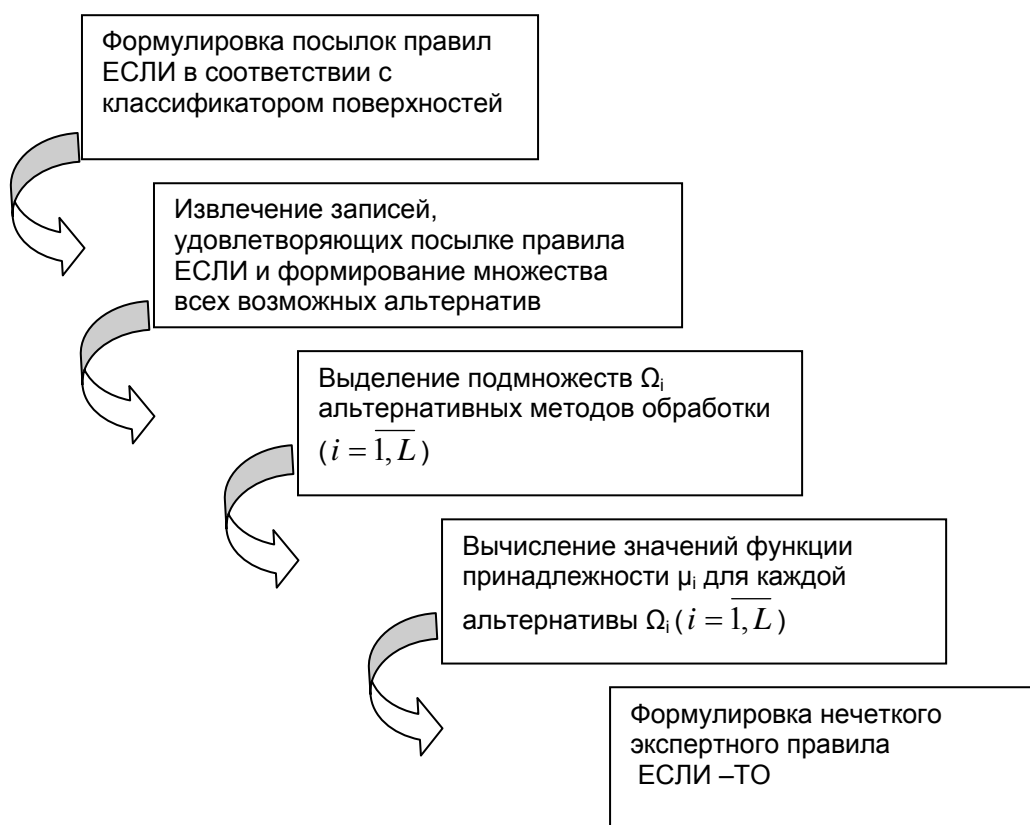


Рис.4 – Алгоритм автоматизированного формирования нечетких экспертных правил БКТЗ для выбора методов обработки поверхностей

Данный метод позволяет ускорить процесс создания базы знаний предприятия, использовать статистическую информацию, имеющуюся на предприятии для вычисления значений функций принадлежности  $\mu_i$  без проведения дорогостоящей процедуры экспертного оценивания альтернатив. Преимуществами данного алгоритма является то, что он генерирует множество альтернативных методов обработки из уже известных и применявшихся на данном предприятии, а не использует знания предметной области вообще.

### Выводы:

1. Построение базы знаний предприятия в виде набора модулей знаний позволяет автоматизировать процесс заполнения базы знаний, использовать справочные базы данных предприятия для формирования правил, наращивать базу знаний в процессе ее эксплуатации.
2. Преимуществами предложенных алгоритмов является возможность автоматизации процесса генерации предпосылок нечетких экспертных правил

благодаря использованию справочной информации баз конструкторско–технологических данных, что позволяет сократить сроки создания базы знаний предприятия.

3. Алгоритм Reference–Knowledge\_module\_discovery генерирует минимальное множество правил, предоставляемых эксперту в проблемной области (эксперту–технологу) благодаря использованию конструкторско–технологических справочников базы данных предприятия, что позволяет сократить затраты на создание базы знаний.

4. Рассмотренный алгоритм автоматизированной генерации нечетких экспертных правил с использованием архива техпроцессов позволяет использовать статистическую информацию, имеющуюся на предприятии для вычисления значений функций принадлежности.

5. В дальнейшем целесообразно применение алгоритмов автоматизированного извлечения конструкторско–технологических знаний для различных этапов проектирования ТП с использованием архива техпроцессов предприятия.

### Список литературы

1. Buchanan B.G., Barstow D., Bechtel R., Bennet J., Clancey W., Kulikowski C., Mitchell T.M. and Waterman D.A. Constructing an expert system. In Building Expert Systems (Haes–Roth F, Waterman D.A. and Levat D., eds.), Chapter 5. Reading, MA: Addison–Wesley.
2. Джексон П. Введение в экспертные системы. : Пер. с англ. : Уч. Пос.– М.:Издательский дом «Вильямс», 2001.–621с.
3. Лапа М.В. Интеллектуальный метод гибкого проектирования технологических процессов деталей приборов // Дис.канд.т.н. – Київ: НТУУ“КПІ”. – 2004. – 198с.
4. Лапа М.В. Формирование нечеткого исходного множества технологических альтернатив для выбора методов обработки поверхностей // Открытые информационные и компьютерные технологии: Сб. науч. тр. Вып. №31. – Харьков: Нац. аэрокосмич. ун–т.– 2006. – С.215–220.
5. B. Padmanabhan and A. Tuzhilin. On Characterization and Discovery of Minimal Unexpected Patterns in Rule Discovery//IEEE Transactions on knowledge and data engineering, № 2.– February 2006.–Pp. 202–216.
6. M.J. Druzdzel and L.C. van der Gaag. Building Probabilistic Networks: Where Do the Numbers Come From? //IEEE Trans. Knowledge and Data Eng., № 4, July/Aug. 2000.–Pp. 481–486.
7. Лапа М.В. Алгоритм формирования нечетких экспертных правил для выбора методов обработки поверхностей деталей// Вестник Харк. нац. ун–та. – №733–2006. Сер. "Математическое моделирование. Информационные технологии. Автоматизированные системы управления".