

КОНЦЕПЦИЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ В ОБУЧАЮЩИХ СИСТЕМАХ В ПРОМЫШЛЕННОСТИ

О.А. Козлов, В.И. Сердюков, Е.В. Садков

Россия, г. Москва

Задача расширения функциональных возможностей обучающей системы, обеспечивающих достижение целей обучения оптимальным образом, решается в настоящее время за счет ее интеллектуализации, на основе моделей предметной области, обучаемого и процесса обучения [1].

Технологические процессы, в которых осуществляется преобразование одних объектов в другие с целью создания материального или информационного продукта, – один из самых сложных для формализации разделов человеческих знаний.

Расширение функциональных возможностей средств электронного обучения, интеллектуализация обучающей системы возможны после построения ряда информационных моделей, основными из которых являются модели обучаемого, процесса обучения и предметной области.

Одной из основных характеристик обучаемого является показатель уровня усвоения учебного материала α , который может меняться от осмысленного понимания новой информации до уровня творческой деятельности и создания новых знаний [2]. На основе оценки уровня усвоения студентами учебных элементов выполняется коррекция учебного материала в процессе обучения. Модели процесса обучения и предметной области определяют набор логических связей и декомпозицию учебного материала. Его структуризация обычно выполняется на основе дидактических рекомендаций, зависит от педагогического мастерства и слабо поддерживается рабочими методиками.

Компонентами, которые необходимо формализовать при построении модели предметной области, описывающей технологические процессы и

конструкторские решения, являются вербальная модель (снабженная иллюстрациями, формулами, таблицами и т.п.), а также большое количество экспериментальных данных [3]. Целый ряд конструкторско-технологических задач в промышленности являются слабоформализованными. Для принятия конкретных решений в производственной сфере необходим значительный предшествующий опыт специалиста, снижающий неопределенность проектной задачи. В таких условиях обучающая система должна быть основана на знаниях.

Выделение категорий технологических объектов выполняется специалистом в предметной области по сочетаниям свойств всех или отобранных выходных признаков. Число атрибутов, идентифицирующих технологические объекты в базе знаний (БЗ) технологических процессов, исчисляется сотнями. Для представления знаний, эксплицированных специалистами, и извлечения знаний из массива данных, образованных триадами «объект–атрибуты–значение», необходимо использовать методологию обучения распознаванию образов для автоматического построения обобщающих правил, описывающих принадлежность ситуаций к классам. В случаях, когда объем обучающих данных ограничен, в особенности при наличии большого числа классов, определена необходимость рассмотрения прогностических свойств построенной системы решающих правил.

Решающие правила, построенные в процессе обучения распознаванию образов, используются для представления динамических знаний о технологическом процессе в виде набора продукций. Причинно-следственные связи между классами объектов и их признаками, отражающие закономерности в обучающих последовательностях, чаще всего устанавливаются с помощью логических решающих правил. Эти правила при малой мере сложности являются универсальными для эффективного решения прикладных задач. Они легко поддаются семантической интерпретации и позволяют автоматически формировать сообщения типа «на заданном оборудовании, в заданной

последовательности, с заданными режимами подвергнуть исходные элементы определенному воздействию”.

Формально технологический оператор можно описать математической моделью [4]:

$$Y(t) = W(t, U) X(t),$$

где $X(t)$ – свойства сырья или материалов, участвующих в преобразовании в выходной продукт; t – время; U – конструкционные и режимные параметры оборудования, а также параметры входных и выходных технологических потоков; $W(t, U)$ – оператор преобразования, отображающий пространства X и U в пространство значений выходных переменных Y ; $Y(t)$ – потребительские свойства готового продукта. Часто перечисление заданных свойств Y является вербально формулируемой целью ТП.

На практике любой технологический процесс представляет последовательность ряда частных технологических подпроцессов (этапов). Такая технологическая цепочка может реализовываться последовательно или параллельно. Параметрами, характеризующими все этапы технологической цепочки, являются [5]:

$X = X_C \cup X_M \cup X_D$ – характеристики исходных компонентов, где

$X_C = X_{C_1} \cup X_{C_2} \cup \dots$ – характеристики используемого сырья,

$X_M = X_{M_1} \cup X_{M_2} \cup \dots$ – характеристики материалов,

$X_D = X_{D_1} \cup X_{D_2} \cup \dots$ – параметры деталей;

$U = \{ U_{C_1}, U_{C_2}, \dots, U_{M_1}, \dots, U_{D_K} \}$ – параметры оборудования и

технологических потоков на этапах обработки сырья, материалов, деталей;

$Y = Y_{\Pi} \cup Y_{\exists}$ – потребительские свойства продукции, где $Y_{\Pi} = Y_{\Pi_1} \cup Y_{\Pi_2} \cup \dots$

– показатели качества произведенной продукции, Y_{\exists} – показатели работоспособности изделия после его эксплуатации.

Сложные многоконтурные технологические системы могут быть представлены эквивалентной разомкнутой схемой путем выделения отдельных подсистем [6].

Возможные наименования продукции, сортность, уровни потребительских и эксплуатационных показателей качества при одной и той же структуре технологической цепочки представляет собой множество классов Ω , содержащее M элементов. Классификация выполняется специалистом-экспертом или автоматически по сочетанию потребительских свойств продукции Y . В трудноформализуемых ситуациях – с использованием весовых коэффициентов и функций принадлежности нечетких переменных размытым множествам-классам.

Составление алгоритма управления, описывающего, при необходимости, применение оборудования в режимах пуска, останова и нормальной эксплуатации, выполняется отдельно для каждого класса продукции.

Выбор тот или иной метод конструирования набора решающих правил на начальном этапе формирования базы знаний позволяет анализ причин и видов априорной неопределенности. В случае недостаточности глобального множества признаков, характеризующих технологический процесс, при отсутствии возможности его пополнения для повышения точности решающего правила целесообразно использовать дополнительные источники информации: последовательное (неоднократное) измерение признака(ов) в процессе наблюдения за изготовлением готового продукта или учет опыта эксперта при построении решающего правила. Последовательный алгоритм выделения класса с новыми потребительскими свойствами применяется при статистической неопределенности задачи обучения распознаванию образов, по запросу на дополнительные измерения, до достижения заданной точности распознавания. В ситуации нестатистической неопределенности используется нечеткое решающее правило, позволяющее учесть субъективную информацию эксперта.

Литература

1. Башмаков А.И. Разработка компьютерных учебников и обучающих систем. М.: Информационно-издательский дом «Филин», 2003. 616 с.

2. Смирнов А.В. Многоагентный подход к построению систем интеграции знаний / А.В.Смирнов, Т.В. Левашова, М.П. Пашкин, Н.Г. Шилов //Известия вузов. Приборостроение. 2003. Т. 46. № 5. С. 13-19.

3. Баранов В.Ю. Формы ответов при автоматизации контроля знаний // Информатика и образование. 2004 №8. С. 17-22.

4. Дуплик С.В., Кюршонов А.С. Модель адаптивного тестирования на нечеткой математике // Информатика и образование. 2004. №11. С. 57-65.

5. Тарасов В.А. Генерация тестов с выборочными ответами на языке JavaScript // Информатика и образование. 2004. №7. С. 113-115.

6. Городецкий В., Самойлов В., Малов А. Технология обработки данных для извлечения знаний: обзор состояния исследований // Новости искусственного интеллекта. 2002. №№ 3-4.