

УДК 519.7

Т. М. НЕОФИТНАЯ

ПРЕДСТАВЛЕНИЕ ПОНЯТИЙНЫХ ЗНАНИЙ В ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЕ ЭКОЛОГИЧЕСКОЙ ЭКСПЕРТИЗЫ СОСТОЯНИЯ ОКРУЖАЮЩЕЙ СРЕДЫ

Человек, являясь неотъемлемой частью биосферы, находится в постоянном взаимодействии с окружающей средой. Научно-технический прогресс, а также постоянное увеличение общей численности населения планеты, неуклонно влекут за собой расширение промышленного производства и энергетики. Негативное воздействие человека на окружающую среду непрерывно нарастает и оно в основном заключается в обусловленном его жизнедеятельностью антропогенном загрязнении биосферы, особенно сильно проявляющемся во время аварий и при других чрезвычайных ситуациях.

Защитные реакции окружающей среды не в состоянии своевременно нейтрализовать все последствия химических и энергетических выбросов и в результате последние приводят к нежелательным изменениям среды обитания. В то же время естественные процессы в природе часто оказывают большое негативное влияние на человечество, которое выражается в разнообразных стихийных бедствиях, таких как землетрясения, лавины, эпидемии, лесные пожары, наводнения и т. д.

Поэтому актуальной задачей является разработка системы экологической экспертизы состояния окружающей среды, предназначенной для проведения анализа текущего и прогнозирования будущего состояния окружающей среды с целью предотвращения чрезвычайных ситуаций, как антропогенного, так и природного характера, и оперативной ликвидации их последствий.

В настоящее время для решения таких задач привлекаются все доступные возможности вычислительной техники, т. к. без их помощи трудно себе представить обеспечение постоянного анализа безопасности и незамедлительной реакции на чрезвычайные ситуации, в том числе своевременную разработку стратегии действий по её ликвидации. Доказательством эффективности применения интеллектуальных программных систем для решения задач такого рода вместо традиционных процедурно-ориентированных систем, основанных на вычислительных алгоритмах, являются следующие особенности экологических проблем:

- необходимость оперирования не только числовыми данными, но и понятийными знаниями, которые выражаются только в символьной форме;
- отсутствие четкого алгоритма для решения такого рода проблем;
- не все знания определены строго, есть неточности и противоречия;
- использование не только теоретических, но и эмпирических знаний.

Известно, что основу любой интеллектуальной системы составляет база знаний, и успешность ее разработки влияет на гибкость и адаптивность системы, ее способность к решению различного рода интеллектуальных задач и тем самым определяет качество системы в целом [1]. Поэтому интеллектуальная система экологической экспертизы состояния окружающей среды должна иметь базу знаний (БЗ), в которой бы находились специальным образом организованные сведения о состоянии природной среды и антропогенных воздействиях. В связи с этим была поставлена задача – разработать систему поддержки базы знаний, отвечающую следующим требованиям:

- моделирование не только поверхностных, но и глубинных знаний;
- поддержка сложных иерархически организованных структур знаний;
- поддержка открытых баз знаний, в которых знания могут изменяться в процессе вывода.

Основой для построения базы знаний о проблемной области является концептуальная классификационная модель знаний. В концептуальной модели объекты (явления) реального мира, их свойства и связи между ними отображаются в виде понятий (концептов) и совокупности отношений на множестве понятий [2]. Это связано с тем, что “ни предметная область, ни терминосистема, не даны сознанию (интеллекту, абстрактному отражению, словесно-логическому мышлению) непосредственно, а только в виде соответствующих понятий, отражающих предметную область и выступающих в качестве содержания терминов” [3]. Поэтому необходимо, чтобы переход от концептуальной модели проблемной области к ее компьютерному представлению осуществлялся как можно более простым и естественным путем.

На этом пути важной задачей является выбор соответствующего формализма представления знаний. В работах по искусственному интеллекту накоплен широкий аппарат моделей и методов представления знаний. Для решения реальных и сложных задач в слабоформализованных предметных областях, к которым в том числе относится область экологии, необходимо использовать модели и методы когнитивного подхода к представлению знаний, который основывается на изучении особенностей естественного интеллекта и, в частности, понятийной системы человека [4]. В связи с этим предпочтение было отдано эвристическим моделям (семантические сети, фреймы и продукции), которые позволяют адекватно отразить семантику проблемной области, и обладают большой выразительной силой и эвристической мощностью. Однако каждая из этих моделей представления знаний имеет определенные достоинства и недостатки [5] и характеризуется своей эффективной областью применения [6].

Поэтому для реализации мощной системы, основанной на знаниях, целесообразно провести интеграцию всех перечисленных выше моделей представления знаний с целью использования преимуществ каждой модели и представления знаний различной природы. В связи с этим в данной системе используется смешанная модель, включающая фреймы, семантические сети и продукции. В этой модели для описания понятий и отношений между ними предлагается сеть фреймов, а для описания операционной семантики проблемной области задействованы продукции.

Выбор в качестве модели представления понятийных знаний сети фреймов обусловлен рядом причин. Во-первых, с помощью фреймовых структур данных могут быть эффективно представлены как абстрактные знания (понятия о классах объектов), так и конкретные знания (единичные понятия) [7, 4]. Во-вторых, в рамках фреймового формализма возможно описание различных типов отношений между понятиями с помощью введения специальных слотов связей [8, 9]. В-третьих, существует возможность явного определения родовидовых иерархий понятий [4, 6]. Благодаря этому становится возможной работа с большими объемами знаний, т. к. существенно облегчается поиск необходимых знаний. Таким образом, во фреймовой модели учитывается иерархичность, присущая человеческим знаниям. В-четвертых, отличительной особенностью фреймовых сетей является наличие механизма наследования свойств [4, 6], позволяющего устранить излишнюю избыточность информации, т. к. вся общая информация для некоторой совокупности фреймов хранится во фрейме-родителе, а в остальных фреймах указывается лишь ссылка на эту информацию. Кроме того, благодаря использованию механизма наследования свойств в процессе приобретения знаний упрощается ввод новых знаний в систему, т. к. для каждого нового понятия необходимо будет только указать его отличия от родового понятия.

В настоящее время существует много интерпретаций фреймового формализма представления знаний. Прежде всего, это обусловлено чрезмерной абстрактностью теории фреймов, предложенной Минским М. [10] и использованием фреймового формализма представления знаний для различных приложений (распознавание образов, построение экспертных

систем, понимание естественного языка). Поэтому в данной работе была проведена адаптация данного формализма с целью реализации средств поддержки базы понятий о проблемной области. Далее следует описание характерных особенностей фреймовой сети, разработанной для представления понятийных знаний в данной системе, и обосновываются ее основные отличия от распространенных систем фреймового типа.

В основу фреймовой сети положена иерархия фреймов, в которую фреймы соединяются с помощью родовидовых связей. В качестве идентификатора фрейму присваивается имя. Это имя должно быть единственным во всей фреймовой системе. Любой фрейм состоит из набора слотов – признаков описываемого им понятия. Слот – это структура, состоящая из имени слота, указателя атрибута (типа данных), указателя наследования и значения слота. Имя слота должно быть уникальным в рамках фреймов, компонентом которых является данный слот. Кроме того, значение слота должно совпадать с указанным типом данных для этого слота и должно выполняться условие наследования.

Каждый фрейм решено оснащать рядом стандартных слотов, которые будут зарезервированы системой для выполнения специальных функций и недоступны в явном виде для конечного пользователя. Среди них, с целью реализации фреймовых сетей, предлагаются следующие:

- PARENT – слот, указывающий фрейм-родитель;
- CHILDS – слот, указывающий дочерние фреймы;
- USEINSLOT – слот, указывающий на слот во фрейме, в котором указатель на текущий фрейм является значением слота.

В теории фреймов принято различать фреймы для описания абстрактных и конкретных знаний, их именуют прото-фреймами и экзо-фреймами соответственно [4]. Если воспользоваться языком формальной логики, прото-фрейм будет соответствовать понятию о некотором классе объектов предметной области, а экзо-фрейм – единичному понятию. Для облегчения поиска знаний во фреймовой сети целесообразно отдельно выделять фреймы-категории, содержащие наиболее общую информацию, истинную для всех остальных фреймов, порожденных от данного. Фрейм-категория всегда является вершиной иерархии фреймов и не имеет предков.

Таким образом, в системе используются следующие обозначения для различных типов фреймов:

- С (Category) – предлагается применять для обозначения фрейма, находящегося на самом верхнем уровне иерархии фреймов;
- Р (Prototype) – используется для обозначения фреймов, имеющих как предков, так и потомков;
- Е (Example) – используется для обозначения фреймов, не имеющих потомков (описывающих единичные понятия предметной области).

При определении фрейма-категории достаточно ввести имя фрейма и необходимую информацию о слотах. В предельном случае допускается существование вырожденного фрейма, т. е. фрейма который характеризуется только именем и не имеет никаких слотов, кроме системных. Все остальные фреймы наследуются от фреймов-категорий. Для их определения обязательно требуется указать имя фрейма-родителя и имя, присваиваемое данному фрейму. Фрейм-потомок автоматически наследует все слоты своего предка, а к ним уже могут быть добавлены слоты, специфичные для данного фрейма. Каждый фрейм в системе может иметь свои конкретные экземпляры. Экземпляр фрейма определяется только путем означивания его слотов, добавление новых слотов недопустимо.

На основе анализа экологической информации была выявлена необходимость в следующих типах данных:

- I (Integer) – целое число;

- R (Real) – действительное число;
- C (Char) – символ;
- B (Boolean) – логическое значение;
- T (Text) – текст;
- L (List) – список;
- D (Range) – диапазон;
- L (List of frames) – указатель на фрейм или группу фреймов.

Указатели I, R, C, B, T используются для обозначения простых и заранее определенных в системе типов данных Integer, Real, Char, Boolean и Text соответственно. Простые типы используются для конструирования сложных типов данных Range и List.

Введение типа данных Range позволяет использовать в качестве значения слота некоторый числовой диапазон. В этом случае значение слота задается приблизительно – в виде интервала возможных значений. Например, температура воздуха в данной местности не выходит за пределы $-35 - +40^{\circ}\text{C}$. При выборе типа Range необходимо указать, какой из простых числовых типов данных (Integer, Real, Char) используется. Таким образом, при определении диапазона значений слота, по существу, накладываются некоторые ограничения на область определения выбранного числового типа данных. При этом указывается верхняя и/или нижняя граница диапазона, и задаются ее свойства (открытая, замкнутая). Если были определены не все границы диапазона, то по умолчанию используется граница области определения выбранного простого типа данных.

Тип данных List – это список возможных значений из области определения некоторого простого типа данных. Список может быть задан аналитически – с помощью правила выборки значений из области определения простого числового типа данных или прямым перечислением всех элементов списка.

Интерпретация сложных типов данных Range и List происходит динамически, т. е. в ходе выполнения программы. Более того, для каждого слота возможно задание значений разных типов. Таким способом можно одновременно хранить субъективные и объективные данные для одной и той же характеристики фрейма. Например, субъективной оценкой температуры воздуха в данной местности может быть значение “жарко”, а объективной – “ 35°C ”

Тип данных List of frames занимает особое место в системе. С его помощью легко организовать отношения типа “часть-целое”, например, можно создать слот с именем “Состоит из” и присвоить ему в качестве значения несколько указателей на фреймы из других классификаций. В качестве указателя в системе используется имя фрейма. Вообще говоря, введение типа данных List of frames позволяет моделировать любые сложные сетевые структуры знаний. В этом случае имя слота необходимо рассматривать как название связи между фреймами. Преимуществом фреймовой сети является возможность хранить непосредственно в данном фрейме всю информацию о его связях с другими фреймами, благодаря чему значительно упрощается поиск релевантной информации о некотором фрейме, поскольку ведется просмотр не всей БЗ, а рассматриваются только фреймы, связанные с данным прямым или опосредованным путем.

Указатель наследования используется для задания режима наследования текущим слотом текущего фрейма значения соответствующего слота родительского фрейма [4, 6]. Обычно во фреймовых системах используется predetermined набор указателей наследования. Это позволило сформировать список стандартных указателей наследования, используемых в тех или иных фреймовых системах: U (Unique) – слот наследуется, но данные в каждом фрейме могут принимать любые значения; S (Same) – наследование тех же значений данных; R (Range) – ограничение диапазона значений данных; M (Member) – ограничение выбора из элементов, перечисленных во фрейме верхнего уровня; I (Independence) – без наследования, O (Override) – наследование тех же значений данных, если не введено собственное значение [4 – 6, 7 – 9].

Анализ показал, что введение такого количества указателей наследования является избыточным. Указатели S, O, I касаются степени выраженности наследования текущим слотом текущего фрейма значения соответствующего слота родительского фрейма. Указатели R и M показывают, как будет работать механизм наследования, если типом данных слота являются Range (диапазон) и List (список) соответственно. Так, для типа данных Range наследование может осуществляться только в пределах диапазона, указанного во фрейме-родителе, а для типа данных List наследоваться могут только элементы из списка значений слота родительского фрейма. Естественно, что для других типов данных указатели наследования R и M не работают. Поэтому в разрабатываемой системе используются только три указателя наследования: а) I – значение родительского слота не наследуется; б) O – значение наследуется только, если не введено собственное значение; в) S – значение в родительском фрейме наследуется всегда, а интерпретация наследования будет зависеть от типа данных указанного для слота. В случае простых типов данных (I, R, B, C, T) она очевидна и перечисленные выше три типа наследования не нуждаются в уточнении. Для типов данных Range и List наследование осуществляется так, как описано выше.

Необходимо описать механизм наследования, когда типом данных является List of frames, т. е. в качестве значения слота используются указатели на группу фреймов, являющихся элементами другой родовидовой классификации. Для простоты рассмотрим случай, когда значением слота является указатель на один фрейм. Тогда при наследовании этому слоту могут быть присвоены только ссылки на те фреймы, которые находятся в одной иерархии с данным и при этом они должны быть порожденными от него. Для примера рассмотрим схему, изображенную на рисунке. Из этой схемы видно, что местоположением промышленной печи является цех, разновидностями промышленных печей являются вагранки, мес-

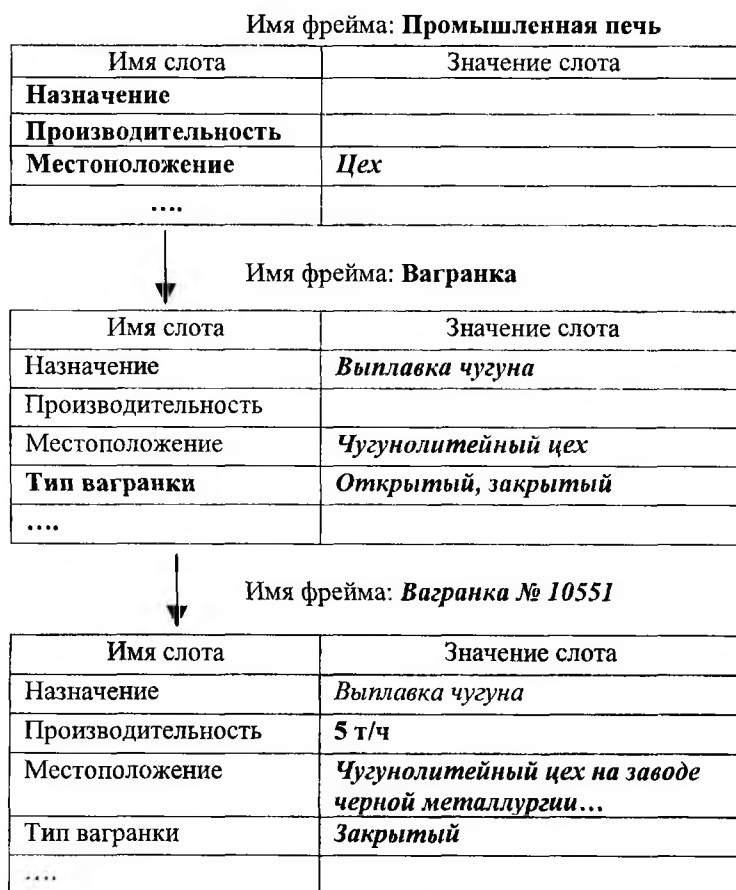


Рис. Фрагмент фреймовой сети.

тоположением вагранки обычно является цех определенного типа, а именно чугунолитейный цех. При определении значения слота “Местоположение” фрейма “Вагранка” (в случае указателя наследования S) будет производиться контроль ввода, ограничивающий выбор значений только именами фреймов-потомков фрейма “Цех”. Аналогично осуществляется наследование, когда значение слота – это целый список указателей на фреймы, только в этом случае переопределение значения производится для каждого фрейма отдельно.

Предложенный механизм интерпретации наследования для типа данных List of frames естественно называть родовидовым наследованием значений слотов фрейма. Он позволяет устранить избыточность информации и проводить контроль ввода знаний, когда в качестве значения слота используется имя фрейма, являющегося вершиной другой родовидовой иерархии фреймов. Следует отметить, что такой вид наследования по имеющимся данным еще не используется в существующих интеллектуальных системах, основанных на фреймовом формализме.

Таким образом, в статье с учетом особенностей когнитивных структур мышления человека обоснован выбор фреймовой сети в качестве модели представления понятийных знаний. Проведена адаптация фреймового формализма представления знаний и описаны характерные особенности предложенной модели, касающиеся типов фреймов, типов данных для значений слота, указателей наследования и механизма наследования. Предложенная модель использовалась при реализации программных средств поддержки БЗ в интеллектуальной системе экологической экспертизы состояния окружающей среды.

Список литературы: 1. *Гаврилова Т. А., Червинская К. Р.* Извлечение и структурирование знаний для экспертных систем. М.: Радио и связь, 1992. 200 с. 2. *Полищук Ю.М., Хон В.Б.* Теория автоматизированных банков информации. М.: Высш. шк., 1989. 184 с. 3. *Соловьева Е. А.* Естественная классификация. Харьков: ХТУРЭ, 1999. 222 с. 4. *Представление и использование знаний: Пер. с япон. / Под ред. Уэно Х., Исидзука М. М.: Мир, 1989. 220 с.* 5. *Георгиев В. О.* Модели представления знаний предметных областей диалоговых систем (обзор) // Известия АН СССР. Техническая кибернетика. 1993. N 5. С. 24-44. 6. *Диалоговые системы представления знаний / Кокорева Л. В. и др. К.: Наук. думка, 1992. 448 с.* 7. *Барсуков А.В., Гаврилов А.В., Олейник Е.В.* Представление знаний в системе распределенных баз знаний и данных в INTERNET / INTRANET // Труды международной научно-технической конференции “Научные основы высоких технологий”. Новосибирск: 1997. Т. 2. С. 212-217. 8. *Андриенко Г.Л., Андриенко В.Л.* Построение информационно-аналитических MULTIMEDIA-систем, основанных на знаниях // Теория и системы управления. 1995. № 5. С. 160-172. 9. *Искусственный интеллект: Справочник.* В 3-х кн. Кн. 2: Модели и методы / Под ред. Д.А. Поспелова. М.: Радио и связь, 1990. 304 с. 10. *Минский М.* Фреймы для представления знаний: Пер. с англ. М.: Энергия, 1979. 152 с.

Поступила в редколлегию 5.08.2000