

КОНЦЕПЦИЯ МОДЕЛИРОВАНИЯ ЗНАНИЙ ДЛЯ АВТОМАТИЗИРОВАННОЙ ТЕХНИКО-ЭКОНОМИЧЕСКОЙ ОЦЕНКИ МЕСТОРОЖДЕНИЙ НЕФТИ И ГАЗА

Ю.Г. Богаткина, Н.А. Еремин, О.Н. Сарданашвили
ИПНГ РАН, e-mail: ubgt@mail.ru

Из истории развития вычислительной техники известно, что путем машинной обработки данных, автоматизированного контроля и управления различными объектами и процессов стало возможно значительно сократить время выполнения рутинных работ по сбору, систематизации и обработке накопленной информации в проблемно-ориентированных компьютерных системах по технико-экономической оценке месторождений нефти и газа. Однако, как показала практика применения таких систем, долгое время существенным недостатком являлась их недостаточная адаптируемость к растущим потребностям пользователей [1, 2]. В первую очередь это было связано с технологиями программирования, не дающими возможности изменять алгоритмы обработки информации по ходу решения задачи. Слабо развитые средства компьютерного диалога требовали привлечения высококвалифицированных программистов. Во многих случаях полученные результаты подлежали долгому кропотливому анализу экспертами. Таким образом, недостаточная гибкость и модифицируемость этих систем потребовала создания совершенно новых подходов обработки информации с применением средств вычислительной техники.

Известно, что большой вклад в теорию построения систем, ориентированных на обработку больших объемов исходных данных и способных гибко адаптироваться к изменяющимся условиям обработки информации, внесло новое направление, базирующееся на теории искусственного интеллекта [3–7].

Благодаря этому направлению был возможен переход от машинного представления процедур к машинному представлению экспертных знаний. Работы, связанные с этой проблемой, развивались по двум направлениям: бионическому и программно-прагматическому [3].

Наиболее ярко свое применение нашло программно-прагматическое направление. Так, теория логико-лингвистического моделирования привела к появлению такого понятия как «базы знаний» (БЗ). С помощью логико-лингвистических формализмов стало возможным представление в машинной форме декларативных и процедурных знаний. При этом особый интерес стали представлять задачи:

- классификации информации;
- логического вывода;
- машинной семантики.

Наглядно можно показать, что модели знаний подразделяются на два основных типа (рис. 1):

- **логические**, в основу которых положено понятие формальной системы;
- **эвристические**, представляемые набором средств, передающих специфические особенности той или иной предметной области.

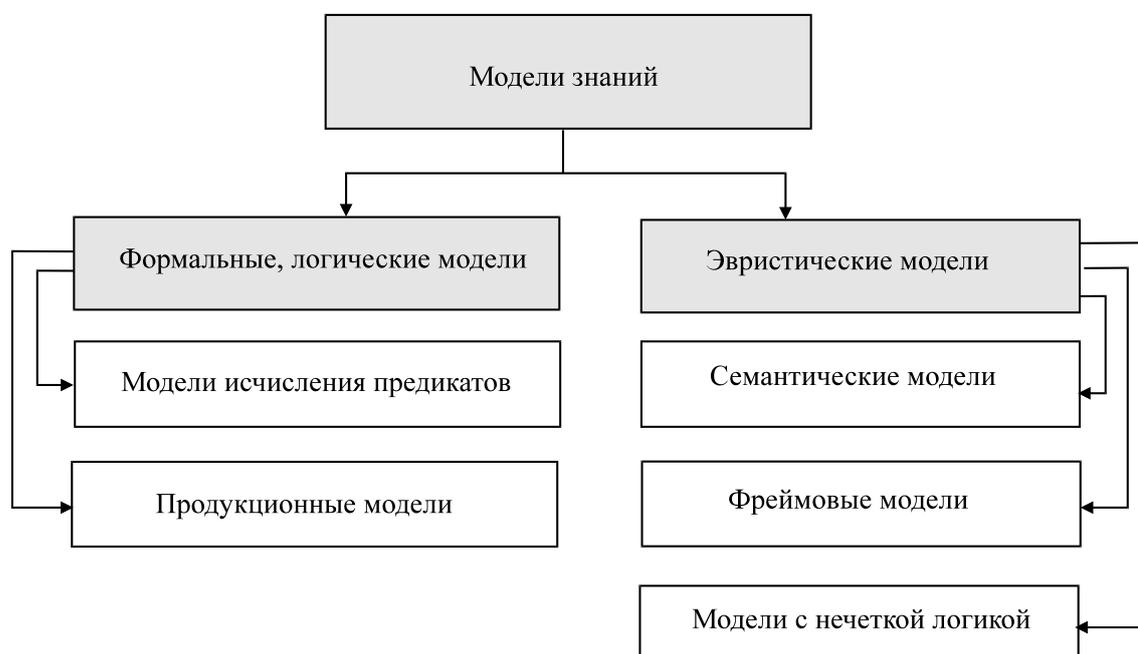


Рис. 1. Классификация моделей знаний

Отметим, что с помощью методов логического вывода и машинной семантики можно представлять формализованные знания (факты), истинность или ложность которых можно доказать. При этом предполагается анализ обрабатываемой технико-экономической информации по вариантам разработки месторождений с точки зрения построения сложно-структурированных иерархических систем с целью решения задачи синтеза расчетных аналитических алгоритмов. Эти методы предполагают изучение семантических понятий, а также методологическую и практическую целесообразность совместной человеко-машинной деятельности.

Применяемые с этой целью семантические сети, с помощью которых строятся различные фразы, позволяют вести осмысленный диалог человека с машиной, в

результате чего возможно обучение пользователей наиболее эффективному представлению декларативных и процедурных знаний эксперта о решении задачи определения основных технико-экономических показателей по заданному расчетному алгоритму. Вершины такой сети должны соответствовать сущностям (атрибутам моделируемых расчетных алгоритмов), а также отношениям и связям между ними [8, 9].

Моделируемые таким образом алгоритмы могут составлять поле возможных решений задачи и накапливаться в интеллектуальной системе в процессе ее функционирования. Алгоритмы могут быть заложены в систему априорно, не имея при этом статуса абсолютной истинности.

Семантические модели, послужившие основой для формирования проблемно-ориентированной БЗ, представленных в работах [3, 4]. Такие структуры включают в себя логико-атрибутивные отношения между исследуемыми объектами и предназначаются для решения задач планирования пути достижения нужной цели из некоторой фиксированной начальной ситуации. Результатом решения таких задач является план действий, представляющий собой упорядоченную совокупность операций над исследуемыми объектами. При этом план действий определяется путем, ведущим к целевой вершине функциональной семантической сети (семантического графа). В частности, для решения на функциональных сетях задач вычислительного характера формируются отношения типа: «часть–целое», «задача–подзадача», «общий случай–частный случай» [10–12].

Метод планирования в пространстве состояний был использован в первой российской программе ПРИЗ (пакет прикладных инженерных задач) [6]. В последних версиях этот пакет представлял неориентированную сеть с вершинами двух типов: дескрипторов, отражающих параметры математической модели, и спецификаторов, отражающих функциональные зависимости между параметрами. Отметим, что функциональная нагрузка дуг переносится на вторую долю (множество вершин спецификаторов) неориентированного двудольного графа, которым представлена семантическая сеть. Здесь спецификаторы являются функциональными выражениями вида $f(x_1, x_2, \dots, x_n)=0$, которые допускают явное разрешение, хотя бы относительно одного из аргументов x_i . С таким спецификатором связано n дескрипторов из множества $\{x_1, x_2, \dots, x_n\}$. Два спецификатора соединены друг с другом некоторым дескриптором, где последний является связующей вершиной, через которую может быть образован ориентированный путь, ведущий от исходных дескрипторов к целевым.

Такой подход становится наиболее удобным для представления неявных отношений между вершинами-параметрами функциональной семантической сети. Сеть содержит в себе возможные решения, которые выявляются в процессе ориентации сети в пространстве ее возможных состояний [3].

Функциональные семантические сети могут являться формой представления экспертных знаний для технико-экономической оценки нефтегазовых инвестиционных проектов. Их можно представить в виде взаимосвязанных семантических графов (подсетей), которые обладают большой выразительностью, и кроме того, с их помощью можно сформировать некоторую исходную базу для дальнейшего накопления знаний эксперта о решаемой задаче.

Надо отметить, что разработанная модель функциональной семантической сети для расчета основных технико-экономических показателей состоит из множества взаимосвязанных компонентов в виде отдельных блоков. Каждый блок соответствует определенному классу, включающему множества семантических подсетей. При этом подсети содержат возможные варианты расчета технико-экономического показателя, указанного в блоке. На рис. 2 показана укрупненная модель расчета основных технико-экономических показателей предметной области.

Объединение подсетей в класс осуществляется по принципу структурного описания сетей Хейса [4]. Подобно сетям Хейса моделируемые подсети, принадлежащие одному классу, объединены общей структурой, позволяющей выбирать необходимую подсеть из рассматриваемого класса. Характерной особенностью такого графа является правило связи вершин друг с другом, где каждая вершина-параметр, соединяющая функциональные зависимости, может иметь на входе не более одной дуги.

Подсети, входящие в отдельные блоки, изначально имеют неориентированный вид, так как могут содержать альтернативные варианты решений искомым параметров. На рис. 3 показан пример таких подсетей. В одном случае представлены варианты расчета искомого показателя для нефтяного месторождения (А), в другом случае те же условия выбора соответствуют газовому месторождению (В).

Каждая подсеть может быть разбита на элементарные компоненты, являющиеся элементарным базисом. Элементарный базис рассмотрим как сеть, содержащую не более одного формального отношения. Будем называть такую сеть структурной связкой.

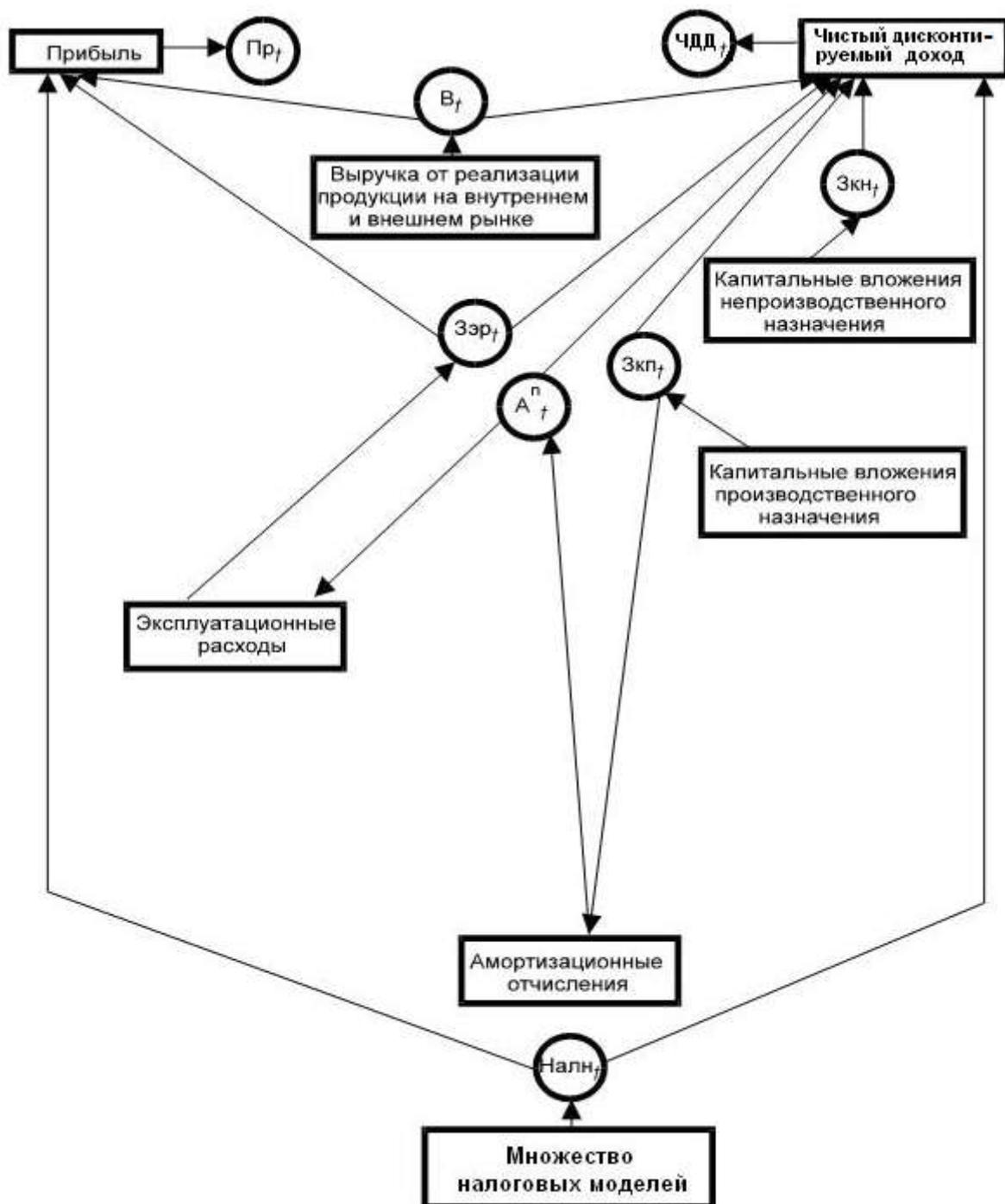


Рис. 2. Укрупненная модель расчета основных технико-экономических показателей предметной области

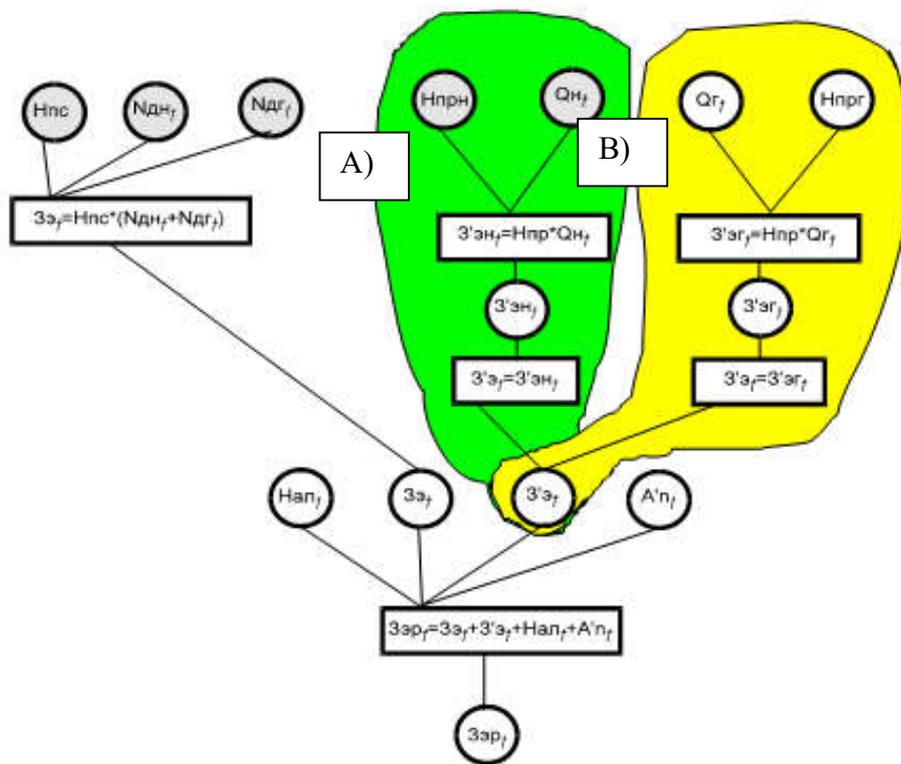


Рис. 3. Пример альтернативных решений на семантических подсетях

Покажем теперь, что множество элементарных сетей (структурных связей) можно классифицировать и на этой основе получать семантическую структуру произвольной сложности.

Формально каждая структурная связка будет обозначаться буквой характеризующей тип сети и переменными в скобках, идентифицирующими входные и выходные вершины-параметры (табл. 1).

Надо отметить, что автоматизированная технико-экономическая оценка разработки месторождений природных углеводородов должна включать в себя комплекс методов представления, обработки и интерпретации экспертной информации, которые можно разделить на теоретические и прикладные. К первым относятся принципы построения моделей предметной области, алгоритмы обработки экспертной информации. В рамках вторых формируются данные для проведения расчетов, проводятся исследования на базе сформированных методик и моделей.

Таблица 1

Типы элементарных ориентированных связей, применяемых при моделировании функциональной семантической сети

Тип перехода	Графическое представление	Условное обозначение
Простое копирование информации		$T(X, Y)$
Объединение с использованием формулы		$J(X_1, X_2, \dots, X_n, Y)$
Объединение с использованием процедуры		$Y(X_1, X_2, \dots, X_n, Y')$

Каждый нефтегазовый проект может иметь индивидуальные модели расчетов, основывающиеся на больших объемах исходной технико-экономической информации и различных прогнозных вариантах разработки и обустройства месторождений нефти и газа. При этом возникает необходимость получения технико-экономических прогнозов по освоению залежей углеводородов, а также определения степени влияния регулирующих показателей (добыча, цены, капитальные и эксплуатационные затраты) на принятие решений с учетом неопределенности и случайности.

Поэтому в основу предлагаемой концепции предлагается положить идею, заключающуюся в выделении из экспертной информации важнейших компонент знаний и представлении их в виде, обеспечивающем готовность к оперативному использованию. Сформированные компоненты знаний могут быть применены непосредственно в процессе создания прикладных расчетных программ по оценке нефтегазовых проектов. Отметим, что Знание в теории искусственного интеллекта (ИИ) и экспертных систем – это совокупность информации и правил вывода (у индивидуума, общества или системы ИИ) о мире, свойствах объектов, закономерностях процессов и явлений, а также правилах использования их для принятия решений. Главное отличие знаний от данных состоит в их структурности и активности. Появление в БЗ новых фактов или установление новых связей может стать источником изменений в принятии решений.

В основе построения БЗ для технико-экономической оценки месторождений нефти и газа лежат принципы, ориентированные на создание интеллектуальных систем, разработанных, в частности, на базе семантических моделей [3, 4]. При этом основной целью является извлечение знаний, которыми располагают специалисты, и представление их в наглядной форме. Это должно давать возможность всестороннего использования технико-экономической информации по проекту в процессе получения экспертных заключений.

Необходимо сказать, что инженерия знаний, как наука, исторически развивалась на основе совершенно иных вычислительных моделей, нежели традиционная программная инженерия. Характерный перечень терминов, используемых в инженерии программ, представляется понятиями: «программа», «описание», «оператор», «переменная» и пр. Для инженерии знаний используются понятия и методы математической логики, теории продукций, фреймов, машинной семантики, аппарата теории нечетких множеств. Опыт в этой области накапливался многими специалистами, называемыми инженерами по знаниям. На наш взгляд, на практике для решения очень многих прикладных задач, в том числе и для оценки нефтегазовых проектов, должно использоваться именно сочетание методов программной (инструментальной) инженерии по обработке экспертной информации и методов инженерии знаний.

Отдельно отметим, что инженерия знаний в инвестиционном проектировании может включать в себя такие понятия как «добыча данных» и «добыча знаний», последнее является довольно новым понятием в системотехнике. Возникновение этого термина

связано с развитием средств и методов обработки данных, алгоритмов, заданий, естественного языка и аналитических выкладок. В рамках указанной проблемы под данными будем понимать всю совокупность фактографической геолого-технологической и нормативной информации, которая классифицируется и сохраняется, но не организуется для передачи какого-либо специального смысла. К знаниям будем относить расчетные и оценочные экономико-математические модели в инвестиционном нефтегазовом проектировании.

Добыча знаний с целью проведения технико-экономической оценки по проекту, в первую очередь, должна сводиться к поиску информации, которую надо извлечь из литературы, экспериментальных действий, или знаний от специалистов в рассматриваемой области (консультантов). Работа по добыче знаний должна выполняться инженером по знаниям (инженером-системотехником), хорошо знающим предметную область, принципы системного подхода, и выражающим всю информацию о проекте в систематизированном виде. Добыча данных и знаний должна производиться с целью создания универсальных программных интерфейсов для хранения и управления потоками исходной геолого-технологической и экономической информации по проекту под контролем высококвалифицированных экспертов [10]. Экспертная информация может быть сохранена и организована различными путями до и после ее использования.

Для проведения технико-экономической оценки месторождений нефти и газа разработана специализированная интеллектуально-логическая система (ИЛС) «Граф», созданная в Институте проблем нефти и газа Российской академии наук (ИПНГ РАН).

Необходимо отметить, что функционирование интеллектуальной системы удовлетворяет различным по форме и содержанию информационным потребностям пользователей в прикладной области и обеспечивать выдачу:

- отчетных документов;
- информации, хранящейся в соответствующих подсистемах БЗ;
- информации, которая может быть получена в результате взаимодействия знаний, хранящихся в БЗ;
- специализированного программного обеспечения, полученного на основе использования БЗ;
- проектных оценочных решений, формируемых в результате проведения расчетов.

Основные методы разработки и специфика применения системы предусматривают различные математические модели и обоснованность выбора той или иной модели в зависимости от характера решаемой задачи. Так, с помощью опросных электронных шаблонов можно формировать структуру информационных баз данных, с помощью когнитивных карт (семантических подсетей) представлять аналитические выкладки технико-экономических моделей в виде базы знаний, с помощью графиков хранить и интерпретировать технико-экономическую информацию по проекту, а с помощью теории нечетких множеств определять степень риска инвестиционных прогнозов [12].

Дадим схематический вид диалогового интерфейса ИЛС «Граф» (рис. 4).

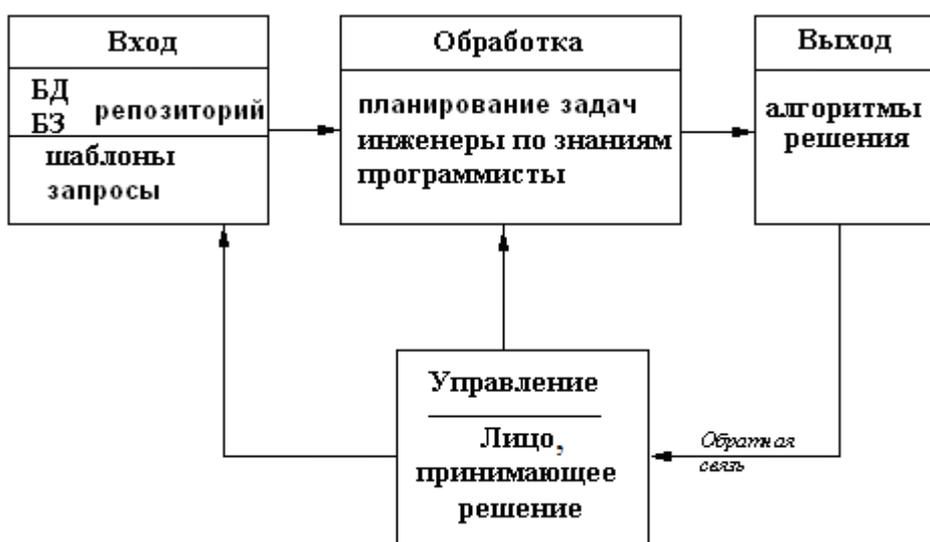


Рис. 4. Диалоговый интерфейс ИЛС «Граф»

Диалоговый интерфейс позволяет собирать, обрабатывать, хранить, анализировать и интерпретировать информацию об инвестиционном проекте с целью его технико-экономической оценки. Через интерфейс осуществляется доступ к *входной информации* (БД, БЗ, шаблоны данных) и *выходной информации* в виде отчетных форм. Обработанная ИЛС информация в наглядной форме посылается пользователю-эксперту (лицу или лицам, принимающим решения) через посредничество инженера по знаниям. В системе существует механизм *обратной связи*, который контролируется экспертом, инженером по знаниям и программистом. Итерационный процесс принятия решений по инвестиционному проекту приводит пользователей системы к оптимальным и взаимосогласованным решениям на основании заложенных в систему технико-экономических моделей.

По способу формирования решения ИЛС относится к классу аналитических систем, предполагающих выбор решений из множества известных альтернатив. При этом предполагается, что в процессе поиска решения последовательность формируемых ситуаций (алгоритмов) не оборвется до получения решения. По способу учета временного признака система решает задачи при изменяемых в процессе решения данных и знаниях с целью пересмотра в процессе решения полученных ранее результатов, а по виду используемых данных обрабатывает прогнозную технико-экономическую информацию по вариантам разработки месторождений нефти и газа. По числу используемых источников знаний ИЛС сориентирована на множество источников, которые являются альтернативными (множество миров) или дополняющими друг друга (кооперирующими).

Таким образом, совокупность рассматриваемых выше характеристик определяют особенности прикладной ИЛС, а решаемые ею задачи можно характеризовать следующими аспектами: числом и сложностью прикладных алгоритмов (правил), их связностью, пространством поиска и количеством активных пользователей, формирующих предметную область.

Статья написана в рамках выполнения государственного задания (тема «Фундаментальный базис инновационных технологий нефтяной и газовой промышленности», № АААА-А16-116031750016-3).

ЛИТЕРАТУРА

1. *Еремин Н.А.* Моделирование месторождений углеводородов методами нечеткой логики. М.: Наука, 1994. 460 с.
2. *Волков С.И., Романов А.Н.* Организация машинной обработки экономической информации. М.: Финансы и статистика, 1988. 399 с.
3. *Поспелов Г.С.* Искусственный интеллект – основа новой информационной технологии. М.: Наука, 1988. 280 с.
4. *Вагин В.Н.* Дедукция и обобщение в системах принятия решений. М.: Наука, 1988. 384 с.
5. *Змитрович А.И.* Интеллектуальные информационные системы. Минск: ТетраСистемс, 1997. 367 с.
6. *Кахро М.И., Калья А.П., Тыгуу Э.Х.* Инструментальная система программирования на ЕС ЭВМ (ПРИЗ). М.: Финансы и статистика, 1988. 175 с.

7. *Абдикеев Н.М.* Проектирование интеллектуальных систем в экономике. М.: Российская экономическая академия им. Г.В.Плеханова, 2003. 312 с.

8. *Богаткина Ю.Г., Пономарева И.А.* О методах моделирования экономических показателей в автоматизированной системе технико-экономической оценки месторождений нефти и газа // Автоматизация, телемеханизация и связь в нефтяной промышленности. 2014. № 1. С. 10–13.

9. *Богаткина Ю.Г.* Методические предпосылки создания базы знаний в логической системе технико-экономической оценки месторождений нефти и газа // Автоматизация, телемеханизация и связь в нефтяной промышленности. 2014. № 2. С. 13–16.

10. *Богаткина Ю.Г., Степанкина О.А.* Структура интеллектуального интерфейса в логической системе «граф» // Автоматизация, телемеханизация и связь в нефтяной промышленности. 2015. № 1. С. 25–30.

11. *Богаткина Ю.Г.* Особенности программных продуктов для оценки экономической эффективности нефтегазовых инвестиционных проектов // Автоматизация, телемеханизация и связь в нефтяной промышленности. 2015. № 2. С. 21–26.

12. *Богаткина Ю.Г., Пономарева И.А., Еремин Н.А.* Применение информационных технологий для экономической оценки нефтегазовых инвестиционных проектов: Монография. М.: Макс Пресс, 2016. 148 с.