

Опыт разработки и применения интеллектуально-графических компьютерных систем

Experience in developing and using intelligent graphical computer systems

Кузнецов В. В., Чижова И. А.

Kuznetsov V. V., Chizhova I. A.

В статье рассматриваются концептуальные, методологические и практические аспекты построения информационно-аналитических систем, в том числе систем прогноза и оценки, основывающихся на искусственном интеллекте и использовании баз знаний, а также реализация этих принципов в исследовательском варианте разработанной интеллектуально-графической системы на основе параметрических моделей месторождений и рудных полей Рудноалтайской минерагенической зоны, использованной для оценки перспективности рудных полей Змеиногорского рудного района.

Ключевые слова: интеллектуальная система, база знаний, банк данных, прогнозно-поисковые модели.

The article discusses the conceptual, methodological and practical aspects of building information and analytical systems, including forecasting and evaluation systems based on artificial intelligence and the use of knowledge bases, as well as the implementation of these principles in the research version of the developed intellectual and graphical system based on parametric models of deposits and ore fields of the Rudno-Altai mineragenic zone, used to assess the prospects of ore fields of the Zmeinogorsky ore district.

Keywords: intelligent system, knowledge base, data bank, predictive search models.

Для цитирования: Кузнецов В. В., Чижова И. А. Опыт разработки и применения интеллектуально-графических компьютерных систем. Руды и металлы 2021. № 1. С. 26–41. DOI: 10.47765/0869-5997-2021-10002.

For citation: Kuznetsov V. V., Chizhova I. A. Experience in developing and using intelligent graphical computer systems. Ores and metals, 2021, № 1, pp. 26–41. DOI: 10.47765/0869-5997-2021-10002.



История создания интеллектуальных систем берёт свое начало с 60-х годов прошлого столетия. Тогда Эдвард Фейгенбаум, исследователь в области искусственного интеллекта, как и многие учёные его времени, задумывался над тем, может ли машина думать и рассуждать подобно человеку и как много знаний в неё возможно вложить [5]. Собравшаяся команда единомышленников (Эдвард Фейгенбаум, Джошуа Ледерберг и примкнувший к ним Брюс Бученен) в стенах Стенфордского университета в 1965 г. начала работы по созданию первой экспертной системы. Бученен предложил хороший инструмент для создания экспертных систем, отделив механизм логического вывода от базы знаний. Он позволил разработать ряд прикладных экспертных систем. Первой системой в области геологии была система PROSPECTOR, которая разрабатывалась Центром искусственного интеллекта SRI International для Геологической службы США с 1974 по 1983 г. PROSPECTOR предназначена для поиска месторождений полезных ископаемых. В 1984 г. система точно предсказала существование месторождения молибдена, оценённого в многомиллионную сумму [2].

В дальнейшем при создании экспертных систем были учтены особенности и недостатки систем, разработанных ранее. В искусственном интеллекте обозначилось такое направление, как инженерия знаний, отвечающая за поиски передовых методов в сборе, представлении, хранении и преумножении информации.

Современные информационно-аналитические системы, в том числе системы прогноза и оценки, основывающиеся на искусственном интеллекте, базируются на совокупности постулатов, которую принято называть базой знаний интеллектуальной системы. База знаний черпает информацию из банка данных. Отсюда становится совершенно ясно, что основой технологии функционирования и успешного применения прогнозных комплексов компьютерных интеллектуальных систем является достоверная, определённым образом структурированная геоинформация из различных отраслей геологических знаний (собственно геологических, минералого-геохимических, геофизических и др.).

Интеллектуальная система – это программная система, имитирующая на компьютере мышление человека. Интеллектуально-графическая система (ИГС) – это система для решения прогнозных, поисковых и других задач с использованием техники рассуждений эксперта при анализе текстовой, числовой и графической информации об объекте исследования. Согласно современной типизации искусственного интеллекта по степени развития она относится к расширенному типу, поскольку направлена на поддержку человеческого мышления в процессе решения прогнозных и поисковых задач [1].

Методы искусственного интеллекта предполагают простой структурный подход к разработке сложных программных систем принятия решений. Кроме того, они предполагают высокую степень независимости отдельных частей программы, каждая из которых реализует определённый шаг решения одной или нескольких задач. При проектировании системы искусственного интеллекта прежде всего нужно определить цели, для достижения которых она предназначена. В первую очередь следует установить, к какому классу относится решаемая задача, и уметь описать её в нужных терминах.

Программная среда должна иметь все элементы, составляющие процесс принятия решения человеком: цели, факты, правила, механизмы вывода и упрощения.

Под технологией разработки ИГС понимается комплекс инструментальных средств и методик, обеспечивающих разработку надёжных систем для решения практических задач. Хотя ИГС и является программным продуктом, технология её разработки в силу специфики решаемых с её помощью задач (как правило, плохо формализованных) имеет свои особенности. Главное расхождение состоит в том, что обычные программы в начале технологического цикла разработки имеют чёткий жёсткий алгоритм получения конечного результата работы программы, который мало меняется на протяжении всего жизненного цикла системы.

Интеллектуальные системы, основанные на знаниях, напротив, такого алгоритма не имеют ни в начале, ни в конце жизненного цикла; при выбранной системе рассуждений резуль-

тат работы может и должен меняться на протяжении всего периода её существования.

Знания – это новый вид информации, который используют интеллектуальные системы, в отличие от традиционных. База знаний – организованная совокупность знаний, представленная в форме, которая допускает автоматическое или автоматизированное использование этих знаний с помощью ЭВМ. Выбирая нужную информацию, человеческий мозг подключает только относящиеся к делу факты, не перебирая все доступные ему знания. В системах искусственного интеллекта обязательно присутствует механизм упрощения для отбора знаний, необходимых в конкретной системе.

Часть интеллекта, которая помогает извлекать новые факты, называется механизмом вывода. Именно механизм вывода позволяет человеку учиться на опыте, так как он даёт возможность генерировать новые факты из уже существующих, применяя имеющиеся знания к новым ситуациям.

Знания предметной области – это совокупность специфических для данной системы фактов, целей и правил, необходимых системе знаний для решения поставленных задач. Знания включают в себя также понятия, свойства и отношения, составляющие эти факты, цели и правила.

Поскольку представление знаний является средством описания знаний человека, то желательно, чтобы его описательные возможности были как можно выше; с другой стороны, если форма представления знаний становится излишне сложной, то усложняются и механизмы упрощения выводов, при этом не только затрудняется проектирование интеллектуальной системы, но и возникает опасность потери достоверности выполнения её действий. В конечном итоге проектирование представления знаний предусматривает выработку всех этих условий, а затем и выбора решения на основе компромисса между ними.

Опыт разработки экспертных систем показал, что наилучшим образом поставленным целям отвечают модели решающих правил. Они могут быть построены как на основе знаний экспертов, так и в результате математической об-

работки имеющейся базы данных по эталонным объектам (эвристические и расчётно-логические правила).

Технологии проектирования различаются по типу инструментальных средств, которые в них используются, а именно: типы представления знаний, формирования решений, технических средств, средств обучения, редактирования (базы знаний и базы данных), представления результатов и т. д.

Исследовательский вариант интеллектуальной системы может содержать большое количество различных интересных для пользователя средств, которые увеличивают стоимость системы; допускается неполный перечень средств, необходимых для целой технологии и системы.

Для описания выбранной технологии следует перечислить все компоненты системы и описать способы их построения и взаимосвязи. При построении интеллектуальных систем необходимо иметь структурную схему системы, чёткую процедуру формирования базы знаний, механизм принятия решений, обеспечивающий получение результата за приемлемое для пользователя время, удобный режим пользования системой, результат, представленный в привычной для пользователя форме.

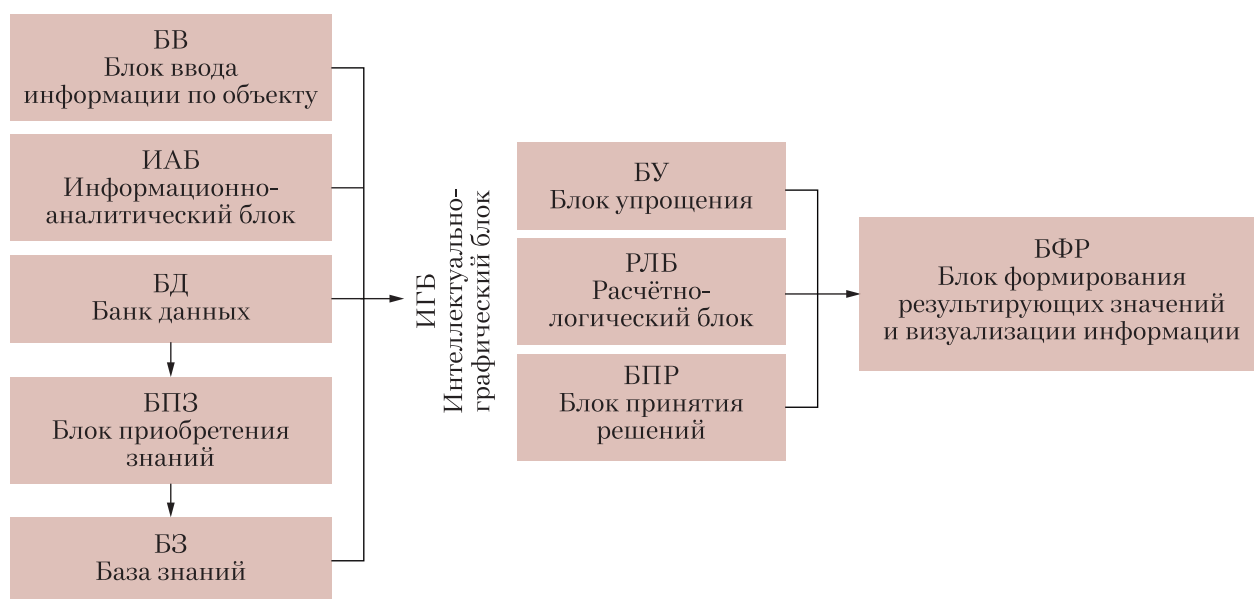
Структурная схема разработанной интеллектуально-графической системы представлена на рисунке. Система состоит из трёх основных взаимосвязанных блоков:

Блок 1 – информационно-аналитический (ИАБ) – на основе ГИС-технологии (в графической форме) обеспечивает визуализацию исходной информации по целевым параметрам из базы данных;

Блок 2 – интеллектуально-графический (ИГБ) – на основе технологии искусственного интеллекта обеспечивает построение базы знаний и работу системы принятия решений;

Блок 3 – является блоком формирования результатов (БФР) и оценки их достоверности и обеспечивает на основе ГИС-технологии визуализацию полученных заключений.

Сердцем всей системы является база знаний (БЗ), разработанная на картографической и фактографической информации по эталонным объектам различных металлогенических ран-



Структурная схема интеллектуально-графической системы

Structural scheme of intelligent graphical system (IGS)

гов (в зависимости от поставленной перед системой задачи). Основная компонента базы знаний – банк данных (БД), сформированный по данным параметрических моделей месторождений. База знаний строится не только на основе знаний экспертов, но и включает закономерности, получаемые на основе математического анализа имеющегося банка данных. Для этого имеется блок приобретения знаний (БПЗ), который содержит программные модули по статистической и логико-информационной обработке информации и классификации объектов.

Интеллектуально-графический блок включает в себя три подблока: блок упрощения (БУ) – для отбора тех знаний из БЗ, которые необходимы для решения поставленной задачи; расчётно-логический блок (РЛБ) – содержит программные модули для расчёта весовых функций объектов и вспомогательных параметров; блок принятия решений (БПР) – содержит программные модули многовариантного обчёта по заданным правилам многокритериального выбора.

Работа системы начинается с представления пользователю информации по распределению

прогнозных ресурсов, запасов, добычи и т. д. в изучаемом районе в пределах рудных районов, полей, месторождений, рудопоявлений и т. п., что позволяет наметить оптимальные площади для прогноза, поисков и оценки.

После их выбора пользователь и система вступают в диалог, причём этот режим предусматривает анализ и оценку как графического, так и фактографического материала. Важнейшими здесь являются интеллектуальные возможности системы, заключающиеся в ведении гибкого диалога, то есть система выбирает следующий вопрос на основании анализа предыдущих ответов; в случае достаточности информации по объекту для принятия решения опрос прекращается.

Характер и форма диалога интеллектуальной системы с пользователем определяется той базой знаний, которая была в неё заложена при проектировании. После каждого ответа проводится анализ полученных данных, и интеллектуальный блок формирует образ объекта. Каждый последующий вопрос следует только после ответа на предыдущий. Предусмотрена возможность возврата к ранее данным ответам с

их корректировкой, при этом машина выносит окончательные решения только после полного окончания сеанса.

Работа системы ориентирована на три типа заключений:

1. Заключение об ожидаемом формационном (минеральном) типе оруденения;
2. Опознание элементов рудного или околорудного пространства, к которому следует отнести объект оценки;
3. Заключение о масштабе прогнозируемого оруденения с приведением конкретных цифр прогнозных ресурсов.

Исследовательский вариант ИГС построен на основе параметрических моделей месторождений и рудных полей Рудноалтайской минералогической зоны. В серии различных видов моделей рудных месторождений (геолого-поисковых, морфометрических, градиентно-векторных и др.), нацеленных главным образом на решение прикладных задач, геолого-поисковые модели занимают определённое место: они формируются в целях оптимизации поисковых и поисково-оценочных работ, в том числе на ранних стадиях их проведения, когда опойсковываемое геологическое пространство вскрывается, изучается и оценивается по единичным пересечениям. Кроме того, параметрические модели могут быть использованы при формировании геолого-генетических количественных моделей месторождений.

Параметрические модели представляют собой геолого-поисковые модели, в которых геологическое пространство, заключающее месторождение, охарактеризовано системой соподчинённых элементов, признаков и критериев, получивших применительно к различным частям пространства количественные (параметрические) оценки.

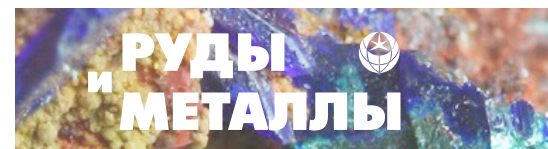
При построении параметрических моделей принято деление околорудного и рудного пространств на следующие части или зоны: собственно рудоносную (РМ), рудоносную фланговую (РФ), собственно надрудную (НРМ), надрудную фланговую (НРФ), надрудную периферийную (НРП), собственно подрудную (ПРМ), подрудную фланговую (ПРФ) и подрудную периферийную (ПРП). В целях формирования ба-

зы знаний интеллектуальной системы данная схема была несколько упрощена. Надрудные и подрудные пространства рассматривались как единое целое без расчленения на собственно над- и подрудную и фланговые части.

Зоны геологического пространства описаны системой факторов и признаков, в число которых, как следует из генетических построений и опыта прогноза и поисков месторождений колчеданного семейства, включены стратиграфолитолого-фациальные, магматические, структурные (в первую очередь, конседиментационные и синвулканические), а также показатели, характеризующие состав, морфологию, зональность и интенсивность гидротермально-метасоматических изменений; морфологию, состав и положение рудных тел в структурах месторождений и разрезах рудовмещающих толщ; геохимические и геофизические аномалии, сопровождающие месторождения.

Каждая из зон пространства получила количественные оценки следующих признаков и критериев или отдельных элементов, составляющих их: мощности рудоносных, подрудных и надрудных частей рудовмещающего разреза, отдельных слагающих их горизонтов и разновидностей пород; соотношения литологических разностей пород; размеры рудоносных структур, выраженные через отношения их длин по простиранию (L), вкrest простирания (H) и по вертикали (M), через мощности выполняющих структуры пород; на некоторых месторождениях – углы падения бортов структур, полученные при расчёте углов выклинивания рудных залежей и рудоносных горизонтов вкrest простирания (α) и по простиранию (β); параметры зон развития гидротермально-метасоматических пород и рудных тел, выраженные через величины отношений $L : H : M$; интенсивность проявления рудной минерализации через процентное содержание сульфидов к общему объёму вмещающих пород и величины соотношения главных рудообразующих компонентов; интенсивность первичных и вторичных геохимических ореолов и их количественные значения зональности ореолов.

Геофизические параметры моделей базируются на данных физических свойств горных по-



род и руд, а также на показателях, характеризующих локальные аномалии различных геофизических методов.

В качестве эталонных объектов отображены преимущественно детально разведанные месторождения, что обеспечило получение достаточно полного объёма информации, необходимой для определения границ различных зон околорудного пространства и достоверности расчёта количественных признаков этих частей.

Модель каждого типового месторождения включает блоки специализированных чертежей, отражающих геологические, минералогические, геохимические и геофизические признаки, таблицы их характеристик и параметрических показателей и геометризованную геолого-поисковую модель месторождения, выполненную, как правило, на основе реконструированных палеофациальных продольных или поперечных профилей. Модели представляют собой информационную основу для разработки новых технологий прогнозирования, поисков и оценки месторождений, параметрического содержания легенд к нагрузкам прогнозных карт различного масштаба, а также определения плотности поисковых и поисково-оценочных сетей.

Из разработанных моделей [3] следуют основные выводы:

Районы рудноалтайского типа эквивалентны структурно-формационным блокам в пределах металлогенических зон и представляют собой крупные вулканотектонические депрессии, выполненные вулканогенно-осадочными отложениями одной (или двух) рудоносных формаций. Палеоструктурный анализ рудных районов Алтая показал, что в их пределах выделяются краевые и внутренние депрессии. В пределах рудных районов выделяются рудные поля. Палеоструктуры рудных полей, как правило, представлены прогибами (или депрессиями второго порядка). В сложнопостроенных вулканотектонических депрессиях, контролирующих рудные районы, они занимают различное положение по отношению к поднятиям, имеющим вид вулканических построек, сложенных кремнекислыми вулканиками. Выделяются три типа прогибов: межвулканические, надвулканические и склоновые, отличающиеся друг от дру-

га морфологией, размерами и составом выполняющих их отложений.

Внутренняя их структура осложнена небольшими экстрезивными и лавовыми куполами, потоками и сопряжёнными с ними мелкими некомпенсированными впадинами. Кроме того, рудные поля часто приурочены к бортам вулканотектонических депрессий первого порядка, в пределах которых развиты экстрезивные купола и сопряжённые с ними впадины.

Колчеданно-полиметаллические месторождения Рудного Алтая по комплексу характеристик относятся к колчеданным месторождениям в вулканогенных комплексах (VMS) (В. В. Кузнецов, 2016) и ассоциируют с двумя рудоносными формациями (под термином «рудоносная формация» понимается геологическая формация, которая является и рудовмещающей, и рудогенерирующей [4]) – нижней (эмс – эйфель – ранний живёт), последовательно дифференцированной базальтсодержащей андезит-дацит риолитовой известково-кремнисто-терригенной, и верхней (поздний живёт – ранний фран), контрастной базальт-риолитовой кремнисто-терригенной. Эти формации обладают рядом характерных черт: принадлежностью вулканитов к калий-натриевой серии, существенным преобладанием кислых пород над основными, антидромным характером вулканизма и присутствием значительного количества осадочных и вулканогенно-осадочных пород (карбонатно-кремнисто-терригенных, углеродистоглинисто-кремнисто-терригенных, кремнистых туффилов и рифогенных известняков).

Средние соотношения Pb : Zn : Cu в рудах месторождений составляют 1 : 3 : 1 при широкой вариации этих соотношений. Главные минералы – пирит, галенит, сфалерит и халькопирит. Основные элементы-примеси представлены Ag, Au, Cd, Se, Te.

На рудноалтайских месторождениях ярко выражена взаимозависимость состава руд с составом и петрохимическими особенностями рудовмещающих толщ. С рудоносными геологическими формациями ассоциируют соответствующие рудные субформации:

- Свинцово-цинковая колчедансодержащая рудная субформация соответствует геоло-

гической последовательно дифференцированной базальтсодержащей андезит-дацит-риолитовой известково-кремнисто-терригенной формации, в составе которой кислые вулканы преобладают над основными, а осадочная составляющая – над вулканогенной. Для месторождений характерно примерно равное соотношение свинца и меди в рудах (при некотором преобладании свинца), а также присутствие повышенных содержаний цинка. Соотношение $Pb : Zn : Cu$ составляет $(1-2) : (6-0,2) : 1$. Коэффициент колчеданности $S : (Pb + Zn + Cu)$ варьирует от 0,5 до 1,5;

- Медно-свинцово-цинковая колчеданная рудная субформация соответствует геологической контрастной базальт-риолитовой кремнисто-терригенной формации, в которой отмечаются примерно равные соотношения кислых и основных вулканических пород, а также осадочной и вулканогенной составляющих. Для месторождений этой рудной субформации характерна незначительная роль свинца при близких значениях цинка и меди и высокой сернистости руд. Соотношение металлов $Pb : Zn : Cu$ составляет $(1-4) : (6-2) : 3$; коэффициент колчеданности – более 1,5.

Банк данных, созданный на основе параметрических моделей месторождений рудноалтайского типа является информационной основой её базы знаний. Структура, информационное наполнение БД и БЗ идентичны.

Банк данных и база знаний состоят из ряда разделов, включающих комплекс геологических, минералого-геохимических и геофизических сведений, имеющихся об эталонных месторождениях, что позволяет сформировать БЗ системы, которая даже при ограниченном количестве данных об объекте оценки (степень изученности которого всегда значительно ниже, чем у эталонных месторождений) принимает итоговое заключение о перспективности участка.

В структуру входят данные по следующим направлениям:

1. Породы, слагающие рудные поля и месторождения, и их формационная принадлежность;
2. Рудоконтролирующие и рудовмещающие структуры;

3. Метаморфические и гидротермально-метасоматические изменения;

4. Морфология и условия залегания рудных тел;

5. Минеральные типы руд (для рудовмещающей части разреза);

6. Зоны проявления сульфидной минерализации (для надрудной, фланговой и подрудной частей разреза);

7. Главные элементы геохимических ореолов;

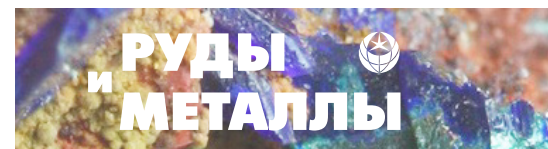
8. Геофизические поля.

Для каждого раздела приводятся качественные признаки, например, для первого раздела (типы осадочных горных пород и их количественное выражение) – процентное содержание породы в разрезе, её мощность, соотношение мощностей и т. д.).

В большинстве случаев при количественном описании того или иного признака используется не фиксированная жёсткая цифра содержания или длины и т. д., а диапазон значений. Это связано, с одной стороны, в специфике обработки и использования информации современными ПЭВМ, а с другой, – в точности самой информации, получаемой при геологических исследованиях.

В современной практике для нахождения границ диапазонов значений количественных характеристик признаков применяются два метода – чисто машинный, оперирующий приёмами математической статистики, и экспертный. Наиболее хороший результат получается при комбинированном использовании обоих – это метод выявления на основе методов математической статистики наиболее характерных, статистически выдержанных диапазонов с последующей экспертной их оценкой и корректировкой. Данная комбинированная методика была использована при формировании БЗ настоящей системы.

Банк данных содержит набор показателей каждой группы признаков, один или несколько из которых характерны для каждого из эталонных месторождений. Совокупность признаков, описывающих каждое месторождение, даёт общий список показателей БД. Типы полей БД сводятся к трём наиболее распространённым случаям:



1. Логические поля – на месторождении-эталоне или объекте оценки может быть строено только один из входящих в группу признаков, например, рудоносная формация, возраст отложений, их мощность, процентное количество пород в разрезе и т. д.

2. Комбинированные поля – из группы признаков значимыми могут являться от 1 до n (где n стремится к количеству признаков в группе). К этому типу относятся литолого-фациальные типы пород разреза, интрузивные образования, формы рудных тел, типы руд и т. п.

3. Цифровые поля – вводятся в БД только в случаях, когда нельзя по каким-либо причинам использовать диапазоны содержаний с логическим вводом данных.

Количество полей БД (БЗ) в целом не фиксировано и может быть как угодно велико. В то же время следует учитывать ряд обстоятельств.

С одной стороны, как уже отмечалось выше, необходима максимальная полнота описания, а с другой, в значительных по размерам структурах признаков пространств с наличием большого количества косвенных признаков на стадии выработки решающих правил и определения информационных весов возникают информационные шумы, которые затрудняют процедуру принятия решений, а в ряде случаев приводят к ложным заключениям. Начиная с первой в области геологии экспертной системы PROSPECTOR, проблема выбора оптимального размера признаков пространства стоит перед исследователями, занимающимися данным разделом информатики, и до настоящего момента нет единой точки зрения или каких-то строгих рекомендаций. При применении подобных систем исходя из опыта максимальную информационную нагрузку несут от 15–20 до 40–45 признаков, на которые обычно приходится до 90 % суммы информационных весов. Практика показывает, что 5–10-кратное превышение в исходном признаковом пространстве количества информативных признаков не приводит к каким-либо существенным информационным шумам и искажениям.

Специфика БД и БЗ системы, построенной на основе параметрических моделей месторождений, заключается в том, что для каждого из

объектов имеется не одно, а несколько описаний, совокупность которых характеризует эталон в целом. Так как параметрические модели строятся на основе расчленения объёма, занимаемого конкретным месторождением, на ряд пространств – рудоносное, фланговое, над- и подрудное, то и БД системы формируется по аналогии с моделями для рудоносного, флангового, над- и подрудного пространств. Совмещение шкалы признаков с матрицей пространств на каждом из объектов-эталонов даёт полную матрицу его описания. Все вместе эталонные месторождения формируют полный (исходный) БД системы.

База знаний использует всю совокупность признаков БД, но одновременно с этим может получать и дополнительные признаки путём их возможных (необходимых) сочетаний; необходимость формирования дополнительных исходных данных БЗ определяется экспертом, формирующим её, и закрепляется в виде правила, которое является обязательным для системы. Для исследовательского варианта на базе рудноалтайских месторождений наряду с конкретными породами, слагающими то или иное пространство разреза эталонного месторождения, сформированы их сочетания, характеризующие группы по масштабам и минеральному типу эталонных объектов. В обычном варианте при работе с системой эти дополнительные признаки БЗ формируются автоматически из ответов оператора при диалоге и широко используются для принятия окончательного решения. В других вариантах систем, использующих данную технологию, могут быть заданы, сформированы и оценены с точки зрения информативности любые сочетания из исходных признаков БД (конечно, они должны нести существенную геологическую нагрузку). Формирование дополнительных признаков БЗ – задача высококвалифицированного специалиста-геолога.

При формировании БД использованы средние значения каждой параметрической составляющей, её максимальные и минимальные значения, а также коэффициенты их вариации.

Кроме фактографического БД сформирован графический банк и, соответственно, графическая БЗ. Его целевое назначение – использова-

ние в диалоговом режиме изображений (геологических разрезов и карт). Представление об объекте у специалиста-геолога строится на образе, представляющем собой совокупность ряда геологических признаков, воплощённых в геологические карты и разрезы, в связи с чем часто происходят затруднения и непонимание принципов работы с системами, основывающимися лишь на фактографической информации. Формирование банков данных изображений пока ещё малоиспользуемый путь при работе экспертных или интеллектуальных систем в виду значительных технологических затруднений ввода и анализа геологической графики. В то же время, это направление наиболее эффективно для совершенствования информационных систем. В настоящей работе сделан лишь первый шаг в этом направлении, позволяющий эксперту на основе растровых изображений геологических объектов (различных геологических карт, характеризующих рудные и нерудные обстановки) дать заключение о мере сходства объекта оценки с предложенным изображением. Это заключение наряду с фактографическими данными используется для принятия окончательного решения.

Выбор месторождений-эталонов БД и БЗ определяется той целевой задачей, которую должна решать интеллектуальная система, и типами объектов-эталонов (минеральными, формационными и т. д.), закономерности размещения и условия локализации которых существенно отличаются друг от друга. Цель выбора эталонных месторождений – сформировать по возможности устойчивые однородные группы при обучении системы, признаковые пространства которых имеют максимально возможное сходство. Количество эталонов в каждой группе должно быть таким, чтобы обеспечить нахождение устойчивых сочетаний информативных признаков. Исходя из опыта минимальное количество эталонов в группе – не менее трёх. При подборе эталонов по масштабам желательно, чтобы в каждой из групп (крупные, средние, мелкие) соблюдался непрерывный ряд изменения запасов. В настоящей работе использованы объекты-эталоны рудноалтайского типа месторождений.

Интеллектуальная система позволяет на основе использования разнообразных методов и методик решать ряд различных по виду задач в процессе реализации общей технологической схемы. В формализованном виде эти задачи можно свести к следующим типам:

1. Задачи распознавания образов. Сюда относятся все виды распознавания объектов от разделения их по минеральным типам до выявления части околорудного пространства и масштаба объектов.

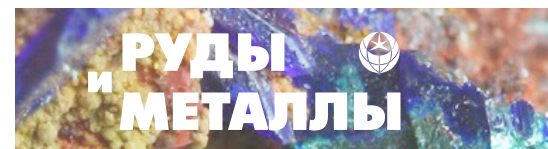
Для создания системы в первую очередь необходимо сформировать систему критериев для выделения объекта, определения его формационного типа и оценки масштабности. Для повышения надёжности получаемых критериев используются математические методы. Необходимым условием применимости методов является наличие однородных (по некоторому критерию) групп эталонных объектов, в результате сравнения которых представляется возможность формирования систем критериев с оценкой их значимости. При этом эталонная выборка объектов описывается в системе признаков, которые, по мнению эксперта, могут быть выбраны в качестве критериев.

Качество получаемой системы критериев (признаков) относительно решаемой задачи определяется на основе анализа эталонной и контрольной выборок объектов. При идентификации этих объектов на основе полученных критериев допускается не более 20 % ошибок.

Изучаемые признаки могут быть различной природы: качественные, балльные, количественные.

В формализованном виде в данном типе задачи можно выделить основные подтипы:

- Задача обучения распознаванию образов (распознавание по эталонной выборке объектов). Цель – отнесение объекта к одному из выделенных классов.
- Задача сокращения (минимизации) описания с оценкой значимости признаков для разделения объектов на классы. Цель – выделение системы информативных признаков.
- Задача таксономии (самообучение, классификации). Цель – формирование однородных групп объектов (признаков).



- Задача выделения типичных представителей классов объектов. Цель – формирование модели типичного представителя исследуемых классов объектов.

Для исследователя важно не только получение правильного результата идентификации объектов, но и выявление признаков, за счёт которых достигается искомый результат, выявление объектов, являющихся типичными представителями априорно или автоматически выделенных классов.

Существует ряд подходов к решению *задач распознавания*. Среди наиболее значимых могут быть выделены три подхода, основанные на идеях метрического, теоретико-вероятностного и комбинаторно-логического характеров соответственно. В зависимости от свойств объекта распознавания используется тот или иной подход.

Если в первом метрическом подходе предполагается, например, что объект изучения расположен в некотором метрическом пространстве и обладает некоторыми свойствами типа компактности, то во втором случае (теоретико-вероятностный подход), как правило, необходимо подчинение свойств объекта некоторым априорным вероятностным закономерностям. Оба эти подхода предполагают наличие большой обучающей выборки. Существенной чертой третьего подхода, называемого комбинаторно-логическим или логико-информационным, является отказ от предположений относительно свойств объекта (метризуемости, подчинённости вероятностным законам и др.), а также возможность вести распознавание при наличии довольно малой обучающей выборки. Наиболее общее свойство алгоритмов комбинаторно-логического типа состоит в том, что наряду с учётом информации по отдельным признакам производится целенаправленный поиск и использование сочетаний признаков, содержащих полную информацию о разделении классов эталонных объектов.

Задача обучения распознаванию образов формируется в следующем виде. Имеется некоторое множество наблюдений, которые относятся к M -классам. Требуется, используя информацию об этих наблюдениях и их принадлеж-

ности к классам, найти такое правило, с помощью которого можно было бы с минимальным числом ошибок классифицировать вновь появляющиеся наблюдения. Процесс поиска такого решающего правила обычно базируется либо на принципе перечисления объектов, либо на принципе общности свойств объектов. Если в первом случае требуется запоминание всей совокупности эталонных объектов для процесса распознавания, то во втором случае таблица эталонов требуется лишь на стадии обучения. Именно поэтому большинство авторов высказываются в пользу принципа общности свойств, который основан на предположении, что объекты, образующие один класс, обладают свойством подобия, отражённого в их характеристиках.

Основой решения всех задач распознавания являются обнаружение и использование некоторых закономерностей, содержащихся в массиве исходных данных. В зависимости от характера исходных данных выбираются соответствующие алгоритмы распознавания, большинство из которых работает с непересекающимися классами (например, тестовые алгоритмы). Но на практике встречаются задачи, в которых исследуются «размытые» (пересекающиеся) классы. Это направление заслуживает особого внимания. Теория размытых множеств [6] достигла определённого успеха при решении подобных задач.

В качестве решающих правил обычно выбираются наиболее простые, поддающиеся логическому объяснению (интерпретации), позволяющие реализовывать простые оптимизационные процедуры его поиска. Именно поэтому наиболее широкое применение получили линейные решающие правила (дискриминантный анализ, вычисление сумм информационных весов объектов), голосование за отнесение объекта к классам. Используются два типа правил. При решении задач распознавания объектов по классам, выделенным по монотонно-изменяющемуся целевому свойству (например, масштабности), используется правило 1: объект относится к тому классу, в чьи пределы изменения попадает вес объекта, вычисленный на основе информативности признаков. В случае,

когда классы выделены без учёта изменения целевого свойства, используется правило 2: объект относится к тому классу, для которого вес объекта (функция принадлежности), вычисленный на основе информативности признаков, характерных для классов, принимает максимальное значение.

С проблемой распознавания тесно связаны задачи выделения существенных (информативных) признаков и минимизации признакового пространства.

Важность правильного выбора наиболее информативной системы признаков из некоторой исходной системы определяется также тем, что:

- указание наиболее информативных признаков оказывается достаточно интересным специалисту, изучающему внутренние связи сложных объектов исследования;
- сокращение числа признаков исходной системы, как правило, улучшает качество решения задачи распознавания.

Ясно, что отбирать существенные (информативные) признаки надо с учётом определённой цели, например, для выделения делящих свойств. Если признак характерен как для объектов класса A_i , так и для объектов класса A_j , то он будет малоинформативен при решении задачи распознавания объектов классов A_i и A_j . Существуют различные подходы к решению задачи минимизации системы признаков:

- без преобразования исходной системы признаков (выбор NM первых по информативности признаков, выброс наименее полезного признака, подсоединение наиболее полезного признака из оставшихся и т. д.);
- с переходом в новое признаковое пространство (факторный анализ, метод главных компонент).

В качестве меры важности признаков для разделения объектов на классы использовались веса, вычисленные на основе методов логико-информационного анализа.

Перейдём теперь к анализу возможных путей решения задачи таксономии. Напомним, что задачей таксономии (классификации) является получение разбиения объектов на группы на основании значений их признаков. При этом возможны два пути: разбиение проводит-

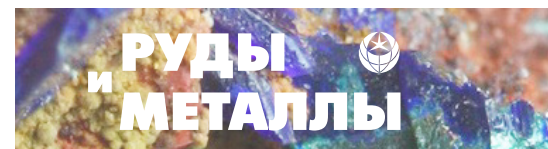
ся на фиксированное число классов (кластерный анализ); число получаемых классов заранее не фиксируется (алгоритм автоматической классификации). Построение классификации в подавляющем числе методов основывается на гипотезе компактности, которая предполагает выполнение двух условий: всегда возможен плавный переход от одного объекта к другому внутри класса так, что все промежуточные объекты будут восприниматься как объекты этого же класса, и нельзя плавно перейти к объектам другого класса; при небольших изменениях объекты не выходят за пределы данного класса.

Все методы основаны на некотором способе определения сходства объектов. Существует ряд способов количественного измерения меры сходства двух объектов: коэффициенты подобия; коэффициенты корреляции; показатели расстояния. Выбор способа вычисления меры сходства зависит от вида используемой информации, которая может носить как качественный, так и количественный характер. Коэффициенты корреляции и подобия указывают на наибольшее сходство в тех случаях, когда они имеют высокие положительные значения, в то время как показатели расстояния указывают на наибольшее сходство в тех случаях, когда они принимают наименьшие значения. Предпочтительней использовать те коэффициенты сходства, которые сохраняют своё значение при изменении масштабности измерения признаков (чем не обладает, к примеру, евклидово расстояние).

Основным требованием, предъявляемым к классификации, является следующее: результат классификации не зависит от порядка рассмотрения объектов.

После того, как получено некоторое разбиение множества объектов на непересекающиеся классы, встаёт интересная задача о выборе типичных представителей классов, характеризующих основные свойства сформированных классов. Такие центры выбираются, например, как точки сгущения либо как объекты, обладающие только признаками, имеющимися на всех объектах класса.

Таким образом, охарактеризованы возможные способы решения поставленных выше задач.



2. Задачи выявления связей параметрических характеристик между собой и с запасами металла. Решение такой группы задач осуществляется на основе применения методов математической статистики. Наиболее важными при этом являются методы общей статистики, анализа корреляционных связей и регрессионный анализ.

Общая статистика позволяет провести оценку главных элементов закономерностей изменчивости параметрических характеристик различных частей рудного и околорудного пространств, используя общепринятые параметры, такие как коэффициент вариации, дисперсию, среднеквадратичное отклонение, моду, медиану и др. Корреляционный анализ позволяет через коэффициент корреляции оценить значимость связей пар параметров между собой. Здесь не имеет существенного значения, положительная или отрицательная корреляционная связь установлена. Главное требование – устойчивость и высокий уровень значимости коэффициента, так как по нему выбираются те параметры, для которых могут быть установлены и рассчитаны связи через уравнения регрессии. Регрессионный анализ позволяет численно и графически представить выявленные закономерные связи между параметрическими характеристиками. Уравнения регрессии могут носить линейный и нелинейный вид в зависимости от сложности формы установленных связей.

3. Задачи оценки удалённости решаются на основе элементов математической статистики совместно с распознаванием образов. На основе теории распознавания образов выявляются элементы околорудного пространства, которые с помощью численных методов преобразуются в геометризованные модели (с использованием статистических оценок параметрических характеристик).

4. Задачи информационно-аналитические. Решаются с помощью ГИС-технологии (географической информационной системы). ГИС-технология обеспечивает выполнение работ по установке связи между огромными массивами числовых параметров объектов и их пространственным расположением. Главная цель – визуализация информации для более наглядно-

го представления изменчивости параметров в пространстве. Технические возможности современной вычислительной техники позволяют в виде карт и различных форм графиков и диаграмм отображать результаты на экране монитора и, в случае необходимости, на принтере. Применение ГИС-технологии является одним из принципиальных моментов в обеспечении достоверности информации.

Каждая из задач, на которые разделяется процесс распознавания объекта оценки, решается на основе выделения класса *информативных признаков*. Причём каждый информативный признак имеет свой информационный вес, определяющий значимость (важность) признака при решении данной конкретной задачи. Вес признака в данном случае рассматривается по абсолютной величине. Знак у величины признака обозначает, к какому из классов (по целевому свойству) он относится.

Принципиально все признаки базы знаний имеют некоторый информационный вес, но для большинства признаков он по абсолютной величине небольшой. В связи с этим они не рассматриваются. Как правило, наибольшие веса имеет ограниченное количество признаков, которые определяют результат решения целевой задачи.

Коротко рассмотрим информативные признаки каждой из задач.

1. Разделение объектов на минеральные типы – полиметаллический и колчеданно-полиметаллический – опирается на 12 наиболее информативных признаков, среди которых максимальное значение имеют типы геологических формаций и субформаций, магматических пород, рудоконтролирующие и рудовмещающие структурные элементы.

2. Выявление пространства, к которому относится объект наблюдений, – для колчеданно-полиметаллического и полиметаллического минеральных типов. Наибольшее значение имеют признаки, характеризующие литолого-фацимальное строение разреза, состав и характер проявления околорудных метасоматитов, состав и морфологию рудной минерализации.

3. Оценка масштаба объекта – для колчеданно-полиметаллического типа; для всех про-

странств наиболее значимы формационные, литолого-фациальные и структурные признаки, для полиметаллических объектов кроме перечисленных выше информативными являются и геохимические признаки.

При работе в диалоге используется вся совокупность признаков БЗ, связанных с решением поставленной задачи, но при принятии решений основная роль отводится наиболее информативным признакам. Полный анализ последних показывает, что они отражают те основные представления об условиях локализации алтайских объектов различного типа и масштаба, которые развиваются большим кругом исследователей последние два десятилетия.

На основе систем признаков и их сочетаний решаются задачи распознавания эталонных месторождений, а в дальнейшем и объектов оценки.

1. Разделение эталонных объектов по минеральным типам. Целевое свойство – минеральный тип месторождений: колчеданно-полиметаллический (1-й класс) и полиметаллический (2-й класс).

Для 1-го класса использовались 36 объектов, из которых правильно распознаны 31 (доля распознавания 86,1 %), а для 2-го – 44 объекта, правильно из которых распознаны 37 (доля распознавания 84,1 %).

2. Распознавание частей околорудного пространства на эталонных объектах полиметаллического минерального типа. Использовались 11 эталонных месторождений различного масштаба. Правильно распознано от 72,7 до 81,8 % объектов. Максимальные значения качества распознавания характеризуют рудоносное пространство, а минимальные – над- и подрудное.

3. Распознавание частей околорудного пространства на эталонных объектах колчеданно-полиметаллического минерального типа. Использовались девять эталонных месторождений. Правильно распознано от 77,8 до 88,9 %. Так же, как и в предыдущем случае, максимальные значения характеризуют рудоносные и фланговые части пространства, а минимальные – над- и подрудные.

4. Распознавание объектов по масштабам проводилось отдельно по каждой из частей око-

лорудного пространства. В каждом случае на основании перебора нескольких вариантов информативных признаков достигалось максимально возможное различие эталонных групп (мелкие, средние, крупные объекты). Граница классов определялась через полусумму максимального и минимального значений информационного веса эталона в соседних классах. В трёх случаях эта граница не зафиксирована ввиду перекрытия информационных весов соседних классов для подрудного и надрудного пространств.

В качестве примера приведём информационные веса объектов при распознавании масштаба оруденения (табл. 1).

В целом при решении любой из приведённых выше задач доля принятия ошибочного решения не превышает 28 %, что следует признать удовлетворительным. Анализ неправильно распознанных геологических ситуаций показывает, что наибольший процент ошибок связан с над- и подрудными пространствами, где геоинформационная база наименее достоверна, или с объектами, имеющими переходные черты между минеральными типами или запасы, близкие к границам классов.

На основе созданной системы была оценена перспективность рудных полей Змеиногорского рудного района Рудного Алтая (табл. 2). Так как в ИГС в обучающем блоке использовались запасы хорошо разведанных и эксплуатируемых месторождений, то и результат оценки может быть приравнен к потенциальным запасам без разбивки на категории. Сравнительный анализ с известными балансовыми запасами месторождений и апробированными прогнозными ресурсами площадей в пределах рудных полей позволяет сделать следующие выводы.

Корбалихинское рудное поле. Ресурсы принадлежат выявленным в результате проведения поисковых работ Каменскому и Давыдовскому проявлениям. Увеличение их при переводе в запасы является маловероятным. В то же время рудные тела западного участка крупного Корбалихинского месторождения на глубинах 700–900 м не оконтурены по падению, в связи с чем прирост запасов на глубоких горизонтах вполне возможен. Выделение новых



1. Информационные веса эталонных объектов при распознавании масштаба оруденения

1. Information weights of standard deposits in mineralization scale recognition

Класс объектов по крупности (запасы, тыс. т металла)	Эталонное месторождение	Части рудоносного пространства			
		рудоносная	рудоносная фланговая	надрудная	подрудная
Колчеданно-полиметаллический минеральный тип					
Мелкие (<500)	Берёзовогорское	0,8	0,612	0,832	0,474
	Лазурское	0,154	0,238	0,309	0,654
	Ново-Золотушинское	0,656	0,293	0,424	0,647
Граничные значения		1,08	0,93	1,03	0,96
Средние (500–1000)	Белоусовское	1,354	1,254	1,315	1,260
	Иртышское	1,395	1,394	1,236	2,069
Граничные значения		1,74	1,53	1,75	?
Крупные (>1000)	Корбалихинское	2,688	2,534	2,532	2,985
	Орловское	2,091	1,662	2,179	2,013
	Малеевское	2,688	2,392	2,532	2,059
Полиметаллический минеральный тип					
Мелкие (<500)	Семёновское	0,9	0,892	0,875	0,886
	Змеиногорское	0,694	0,688	1,222	1,014
	Зареченское	0,594	1,016	1,115	0,577
	Захаровское	0,496	0,482	0,227	0,516
Граничные значения		1,01	1,06	?	?
Средние (500–1000)	Рубцовское	1,473	1,545	1,348	1,501
	Таловское	1,125	1,536	1,338	1,361
	Юбилейное	1,329	1,103	1,023	1,173
	Среднее	1,650	1,428	1,195	0,946
Граничные значения		1,86	1,62	1,58	1,55
Крупные (>1000)	Греховское	2,553	2,539	2,308	2,214
	Чекмарь	2,076	1,697	1,813	1,605
	Зыряновское	2,522	2,294	2,438	2,368

2. Сравнительная характеристика результатов оценки с использованием ИГС рудных полей Змеиногорского рудного района с известными в их пределах запасами и ресурсами

2. Comparative characterization of assessment results using Zmeinogorsky ore district fields IGS versus their known reserves and resources

Рудные поля	Оценка на основе ИГС Zn + Pb + Cu, тыс. т	Известные запасы и ресурсы Zn + Pb + Cu, тыс. т		
		ABC ₁ C ₂	P ₁	P ₂
Корбалихинское	4500	3261,1	311	316
Лазурское	1110	214,8	-	310
Берёзовогорское	1560	222	177	247
Змеиногорское	3800	958	899	1793
Комиссаровское	280	-	176,4	81,2
Зайцевское	1300	-	-	708
Черепановское	870	-	815	300
Кандидатское	1300	-	-	1555
Белоглинское	750	-	-	370

участков для постановки поисковых работ в пределах рудного поля маловероятно.

Лазурское рудное поле. Известное Лазурское месторождение разведано и хорошо изучено. Проведённые в пределах рудного поля поисковые работы позволяют оценить Мосляноско-Пихтовский участок по категории прогнозных ресурсов $P_1 + P_2$ примерно в 540 тыс. т суммы металлов. Исходя из проведённой нами оценки на основе ИГС в пределах рудного поля можно ожидать выявление ещё одного объекта с ресурсами 500–600 тыс. т, скорее всего, в пределах северо-западной части Лазурской вулканотектонической депрессии.

Змеиногорское рудное поле. Полученный результат оценки показывает высокую степень сходимости с имеющимися запасами и ресурсами Зареченского, Среднего месторождений и вновь выявленного Петровского проявления.

Зайцевское, Белоглинское, Кандидатское рудные поля характеризуются ресурсами категории P_2 . Проведённая оценка на основе применения ИГС позволяет рассчитывать на зна-

чительный прирост ресурсов в результате проведения поисковых работ, а в дальнейшем – на выявление в каждом из рудных полей средних по масштабам месторождений.

Комиссаровское и Черепановское рудные поля. Проведённая оценка показала низкий потенциал указанных рудных полей, полностью исчерпываемый имеющимися в их пределах рудопрооявлениями, не имеющими промышленного значения.

Таким образом, использование ИГС показывает свою высокую эффективность при прогнозе и поисках колчеданно-полиметаллических месторождений, её применение позволило заново переоценить известные рудные поля и выявить приоритетные направления для постановки поисковых работ в рудном районе.

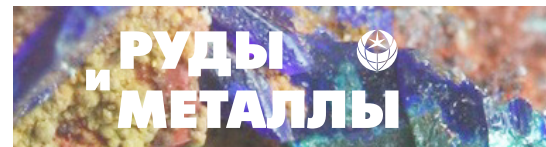
В заключение отметим, что повышение эффективности интеллектуальных систем заключается в совершенствовании их геоинформационной базы – базы знаний – и в возрастании интеллектуальных возможностей собственно компьютерной технологии.

Список литературы

1. *Виды искусственного интеллекта.* – URL: https://spravochnick.ru/informatika/ponyatie_iskusstvennogo_intellekta/vidy_iskusstvennogo_intellekta_1/ (дата обращения: 12.04.2021).
2. *Красильников Виталий. Эволюция экспертных систем.* – URL: <https://www.kv.by/archive/index-2005401103.htm> (дата обращения: 12.04.2021).
3. *Кузнецов В. В., Кудрявцева Н. Г., Серавина Т. В., Мурзин О. В., Корчагина Д. А., Кузнецова С. В., Миляев С. А. Основы прогноза и поисков колчеданно-полиметаллических месторождений Рудного Алтая.* – М. : ЦНИГРИ, 2019. – 205 с.
4. *Металлогенический кодекс России.* – М. : Геокарт-ГЕОС, 2012. – 126 с.
5. *Rheingold H. Expert systems. Computers as sages.* – URL: www.atariarchives.org/deli/expert_systems.php (дата обращения: 12.04.2021).
6. *Zadey L. A. Lotfi A. Zadeh, King-Sun Fu, Kokichi Tanaka. Fuzzy sets and their applications to cognitive and decision processes. Proceedings of the U.S. – Japan Seminar on Fuzzy Sets and Their Applications, Held at the University of California, Berkeley, California, July 1–4, 1974.* – USA : Academic press, 1975. – 506 p.

References

1. *Vidy iskusstvennogo intellekta [Types of artificial intelligence], available at: https://spravochnick.ru/informatika/ponyatie_iskusstvennogo_intellekta/vidy_iskusstvennogo_intellekta_1/ (Accessed: 12.04.2021).*
2. *Krasil'nikov Vitalii. Evolyutsiya ekspertnykh system [Evolution of expert systems], available at: <https://www.kv.by/archive/index-2005401103.htm> (Accessed: 12.04.2021).*



3. Kuznetsov V. V., Kudryavtseva N. G., Seravina T. V., Murzin O. V., Korchagina D. A., Kuznetsova S. V., Milyaev S. A. Osnovy prognoza i poiskov kolchedanno-polimetallicheskih mestorozhdenii Rudnogo Altaya [Fundamentals of forecasting and prospecting for pyrit-polymetallic deposits in the Rudny Altai], Moscow, TsNIGRI Publ., 2019, 205 p.
4. Metallogenicheskii kodeks Rossii [Metallogenic Code of Russia], Moscow, Geokart-GEOS Publ., 2012, 126 p.
5. Rheingold H. Expert systems. Computers as sages, available at: www.atariarchives.org/deli/expert_systems.php (Accessed: 12.04.2021).
6. Zadey L. A. Lotfi A. Zadeh, King-Sun Fu, Kokichi Tanaka. Fuzzy sets and their applications to cognitive and decision processes. Proceedings of the U.S. – Japan Seminar on Fuzzy Sets and Their Applications, Held at the University of California, Berkeley, California, July 1–4, 1974, USA, Academic press, 1975, 506 p.

Авторы

Кузнецов Владимир Вениаминович

кандидат геолого-минералогических наук
заведующий отделом цветных металлов¹
okt@tsnigri.ru

Чижова Ирина Александровна

доктор геолого-минералогических наук
ведущий научный сотрудник²
I.Chizhova@yandex.ru

¹ ФГБУ «Центральный научно-исследовательский геологоразведочный институт цветных и благородных металлов», г. Москва

² ФГБУН Институт геологии рудных месторождений, петрографии, минералогии и геохимии РАН, г. Москва

Authors

Kuznetsov Vladimir Veniaminovich

PhD
Head of base metals department¹
okt@tsnigri.ru

Chizhova Irina Alexandrovna

PhD
leading researcher²
I.Chizhova@yandex.ru

¹ FSBI Central Research Institute of Geological Prospecting for Base and Precious Metals, Moscow, Russia

² FSBIS Institute of Ore Deposit Geology, Petrography, Mineralogy and Geochemistry, Moscow, Russia