

ПРОБЛЕМЫ РАЗРАБОТКИ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Мейтус В. Ю.

*Международный научно-учебный центр информационных технологий и систем НАН и МОН Украины
г. Киев, 03680 ГСП, Украина
тел. 044-5261319; e-mail: vmeitus@gmail.com*

Аннотация — Рассмотрены проблемы, связанные с созданием систем, наделенных интеллектом в том плане, что их поведение будет адекватным поведению человека, перед которым поставлены аналогичные задачи. Такими проблемами являются проблемы представления информации, которую использует интеллектуальная система, преобразования этой информации в процессе поиска решения задачи и проблему неопределенности, связанную с неоднозначностью, неточностью задания информации. В докладе предлагается рассматривать онтологический подход, основанный на использовании понятий и отношений, рассматриваемых в естественном языке, и построенных из них конструкций для формализации процесса решения задач интеллектуальной системой.

I. Введение

Интеллект – это способность системы создавать, разрабатывать, применять методы и процессы согласования наблюдений и фактов, связанных с внешней средой, с знаниями системы о среде, и на основании этого строить модели поведения системы, которые адекватны этой среде.

Под **адекватностью** поведения понимается способность системы правильно решать задачи и проблемы, возникающие в среде и предлагаемые системе как реакции среды на действия системы, что, в свою очередь, предполагает, что система адаптирует свое поведение к среде, находя существующие закономерности среды, если такая адаптация принципиально возможна, исходя из потенциала системы.

Наблюдения поступают из внешней среды по различным каналам – аудио, визуальным, тактильным. Можно рассматривать виртуальные наблюдения, возникающие в виде рассуждений, гипотез и предположений о характере зависимостей, существующих во внешней среде. Еще большим оказывается многообразие фактов и связей между ними, относящихся к разнообразным направлениям социальной и культурной деятельности человека. Этому многообразию соответствуют различные проявления интеллекта, отличающиеся по своему направлению и применению, но совпадающие по характеру процессов, определяющих эти направления.

Процессы согласования задаются последовательностью моделей, каждая из которых рассматривается как относительно мгновенное состояние отдельной области внешней среды, рассматриваемой в этот момент и представляемой интеллектуальной системой, т.е. сформированной с учетом действий среды и базы знаний системы на момент формирования модели. Сами моменты распределены во времени, поскольку формирование последующих моделей связывается с возникновением дополнительных требований или появлением неадекватности результатов, получаемых с помощью модели, и ожидаемых состояний внешней среды.

Интеллектуальная система – это система, в которую заложен механизм формирования интеллектуального поведения при решении задач, связанных с определенной предметной областью. Система

предназначена для достижения заданных целей и результатов в предметной области, которая представляет собой часть внешней среды и характеризуется некоторым набором условий, знаний, сценариев деятельности.

В последние годы создание систем, которые могут быть отнесены к интеллектуальным, получило новый стимул. Особенно это видно на примере систем, понимающих вопросы, заданные на обычном или близком к нему языке (Wolfram Alpha, IBM Watson, True Knowledge), создании систем управления роботами (Япония, Франция, США), ситуационный анализ (проект СУС), новое поколение медицинских экспертных систем. Можно в ближайшем будущем ожидать создания новых интеллектуальных систем для медицинского диагностирования, управления предприятиями, обучения, в основе которых будет положена большая база знаний, представляющая соответствующую предметную область, совместно с пониманием системой утверждений, относящихся к этой области, и формирование реакций, поведения системы на основе этого понимания.

Но это только небольшая часть предполагаемых разработок в этой области информатики. В условиях создания больших универсальных информационных систем типа Wikipedia, развития глобальных и локальных компьютерных сетей, возникновения мирового информационного пространства, интеллектуальные системы явятся безусловным шагом развития современных систем. Практически можно ожидать, что понятие «интеллектуальности» в приложении к компьютерным системам станет их естественным качеством, указание на которое будет опускаться, – других систем не будет. Поэтому сейчас желательно исследовать пути создания таких систем, начиная от согласования основных определений и до архитектуры и технологий их создания.

Процесс работы интеллектуальной системы определяется информацией, полученной из предметной области, и когнитивным опытом системы, отраженном в ее базе знаний. На основании этой информации строится модель предметной области (ее части) и анализируются возможные действия системы, из которых выбирается оптимальное действие. Его результаты сравниваются с ожидаемыми результатами, построенными системой, исходя из модели. В модель вносятся соответствующие коррективы, которые должны приблизить системный результат к реальности. И все повторяется заново, пока не будет получен удовлетворительный результат адаптации.

К сожалению, это удается не всегда. Модель может оказаться очень сложной, а ее упрощения будут не достаточно адекватными среде, в результате чего возможные корректировки не улучшают адаптацию. В самом знании о предметной области имеются существенные пробелы, а действия системы не соответствуют исходной модели. «Тому в истории мы тьму примеров видим». Но эти проблемы, характерные для всей человеческой цивилизации, не должны явиться непреодолимым препятствием на пути ис-

следования. Само развитие и постоянное совершенствование технологий, развитие науки и процессов познания, характерных для человечества, является демонстрацией того, что создание интеллектуальных систем, прототипом которых является человек, возможно. Как говорит латинская поговорка: «Тот сделал уже полдела, кто уже начал».

II. Проблемы создания интеллектуальных систем

Общая технология построения интеллектуальной системы (ИС) – системы, обладающей интеллектом, – основана на разработке следующих направлений создания компонент этой системы, реализовывающих отдельные процессы, обеспечивающие функционирование системы в целом:

1. Создание подсистемы, которая пополняет, сберегает и согласовывает знания ИС, осуществляет их эффективный поиск и выдачу по внешним запросам. Это формирование и работа с базой знаний ИС.

2. На основе знаний, хранящихся в базе знаний, построение моделей внешней среды. Эти модели лежат в основе формирования поведения ИС во внешней среде.

3. По информации, получаемой из среды, выделение знаний, отражающих динамику изменения среды и используемых для развития и корректировки базы знаний.

4. Накопление информации о способах формирования знаний и методах построения моделей внешней среды. Эта информация является основой для построения системой различных процессов адаптации к внешней среде.

Такую технологию будем называть интеллектуальной технологией, хотя иногда к интеллектуальным относят технологии, которые включают элементы, характерные для ИС.

Создание интеллектуальной технологии предполагает построение в отдельности каждой технологии, обеспечивающей выполнение действий, определяемых этими четырьмя пунктами. В результате получим четыре подтехнологии, которые нужно объединить в единую технологию для создания ИС. В зависимости от методов и способов построения отдельных компонент, их полной или частичной интеграции в единую методику получаем разные интеллектуальные технологии.

Соответственно и общая линия разработки и создания ИС опирается в решение целого ряда проблем, связанных с сформулированными направлениями. Сама схема разработки предполагает возможность использования итеративной методики создания ИС. Причем сначала решаются проблемы, определяемые первыми двумя направлениями, что позволяет построить базовую ИС, ориентированную на некоторую статическую предметную область. А затем уже реализовывать программу создания более сложных ИС, развивая и модифицируя это базовое представление.

Сформулируем проблемы, связанные с разработкой интеллектуальных технологий и требующие первоочередного решения.

Первая группа проблем определяется как **выбор формы представления информации**, которую использует ИС. В этой форме задаются знания, связанные с рассматриваемой предметной областью, на ней основаны преобразования и выводы новых знаний, утверждений, следствий и базируются

механизмы таких преобразований и выводов. Форма представления должна отражать свойства предметной области и максимально упрощать работу с составляющими области, чтобы обеспечить решение задач, возникающих в этой области.

В качестве такой формы обычно рассматривается некоторый язык, на котором представляется информация, или способ, устройство, механизм, с помощью которого реализуются отдельные элементы языка представления информации. Примерами таких языков могут служить использования:

- логические исчисления, например, в виде языка узкого исчисления предикатов и классической логики;
- различные варианты нейронных сетей и связанных с ними процедур;
- системы агентов как способ задания и накопления информации о среде, направленные на совместное решение задач, поставленных перед агентами;
- семантические сети;
- представления, основанные на использовании онтологий и онтологических пространств;
- представление на основе естественного языка, допускающее удобную компьютерную обработку, и др.

Вторая группа проблем – это проблемы **выбора адекватных механизмов преобразований информации**, соответствующих понятиям, отношениям и свойствам предметной области. Примером таких механизмов в зависимости от формы задания знаний могут служить известные в психологии интеллектуальные функции или методы логического (индуктивного, абдуктивного) вывода для логической формы задания информации.

Существование этих проблем определяется необходимостью перехода от известной информации, заданных знаний, условий, утверждений, к новым утверждениям о среде. Здесь могут быть использованы существующие математические модели и методы, как это сделано, например, в системе Wolfram Alpha.

Сюда же относятся и проблемы, связанные с **планированием порядка действий** по преобразованию информации. Идет речь о том, что, если это возможно, осуществлять не случайный перебор, а направленные действия по получению необходимого результата. Или, если направление таких действий отсутствует, то сформулировать критерии, по которым выполняется случайный перебор (например, принципы селекции в генетических алгоритмах).

Третья группа проблем – это проблемы **неопределенности информации**. Этот класс проблем связан с неполнотой и неточностью языка, в котором человек представляет знания и факты, связанные со знаниями. Информация, выраженная в языковой форме, тесно связана с тем объектом, который эту информацию представляет. Поэтому разрабатываются специальные языки, построенные по аксиоматическому принципу, например, язык математики, в котором выражения имеют точное значение. Но в этом, вероятно, и заключается сложность освоения математики для человека, ориентированного на разнообразные ассоциации и образы, соотносимые с одним и тем же объектом.

Чем формальнее язык представления знаний, тем меньше неопределенность этих представлений, но сложнее механизмы вывода в этом языке. Поэто-

му очень часто математик исходя из некоторых общих ассоциаций и аналогий сначала формулирует в виде гипотезы некоторый результат, а уже потом ищет формальное обоснование этого результата, его вывод, доказательство. И не всегда находит этот вывод самостоятельно, формулируя предполагаемый результат как открытую проблему и привлекая к ней внимание целого сообщества исследователей.

Проблемы неопределенности как раз связаны с поиском методов, позволяющих находить неформальные результаты, опираясь на механизмы, участвующие в связях между близкими по характеру знаниями, например, механизм индуктивного вывода, или существование выводов, основанных на отношениях подобия, ассоциации, аналогии. Или, наконец, используя в одном выводе различные формы представления знаний и описания задач со своими способами обработки этих знаний.

Сама методика разработки технологии построения ИС (ИНТ) может основываться на выборе некоторых базовых предпосылок как основы, на которой решаются рассматриваемые проблемы. В докладе предлагается для формализации процесса решения задач ИС рассматривать онтологический подход, основанный на использовании понятий и отношений, рассматриваемых в естественном языке. Перспективность этого подхода определяется тем, что на языковом уровне решает задачи и обменивается информацией человек.

Формально **модель онтологии** Ω представляется как семерка символов $\Omega = (E, N, R, Pr, \Lambda, I, \lambda)$, где E – представление предметной области, для которой строится онтология, в пространстве признаков V , N – спецификация понятий области E , их онтологических свойств и признаков, R – спецификация отношений между понятиями области E , Pr – правила порождения в онтологии новых понятий из исходных понятий и отношений, Λ – логика, выбранная для оценки характера соответствия между онтологическими спецификациями и предметной областью, I – оператор интерпретации, $I: N \cup R \rightarrow E$, λ – логическая функция. Если понятию о чем идет речь, то можем употребить выражение «онтология Ω » имея в виду модель онтологии Ω .

Расширяя онтологию за счет добавления к ней экстенционалов и интенционалов понятий и отношений онтологии и вводя операции комбинации элементов пополненного множества получаем онтологическое пространство (О-пространство), в котором можно непосредственно решать задачи, представленные с помощью понятий и отношений онтологии. С одной стороны, решение представляется как последовательность элементов О-пространства, а, с другой, это решение может быть проверено путем перехода в пространство признаков непосредственным вычислением всех функций, входящих в онтологическое представление решения задачи.

Формально для задания онтологического пространства вводится алгебра конструкций K_Ω , определяемая следующим образом.

а) Каждое понятие $A \in N$ (концепт A из N) есть конструкция.

б) Если заданы конструкции A, B , то новая конструкция AB , формируется как последовательная композиция A и B , является конструкцией, определяемой вырожденным отношением.

в) Пусть существует правило $\pi \in Pr$, которое для понятий A_1, A_2, \dots, A_n и отношения $\xi \in R$ определяет

новое понятие B . Тогда конструкция K_B , соответствующая понятию B , задается выражением $K_B = \pi(\xi, A_1, A_2, \dots, A_n)$ и называется **понятийной структурой** для понятия B .

д) Других конструкций в алгебре K_Ω нет.

Алгебра конструкций может быть частично упорядочена продолжением частичного порядка, введенного для концептов онтологии, на элементы алгебры K_Ω . Частично упорядоченная алгебра конструкций K_Ω называется **базовым онтологическим пространством**, определенным онтологией Ω .

Содержательно использование конструкции определяет представление любого понятия в виде некоторой структуры, а сами конструкции рассматриваются как предложения в языке, построенном из онтологических понятий и отношений. Эти предложения выражают само понятие и его связи с другими понятиями, характеризующими составляющие рассматриваемой предметной области. На конструкции может быть продолжен оператор интерпретации, заданный для онтологии Ω .

Алгебра конструкций является основой для построения **основного онтологического пространства** в виде некоторой категории, объектами которой являются, конструкции из алгебры K_Ω , вместе с их конечными произведениями и копроизведениями. Рассмотренное выше О-пространство является одним из примеров задания онтологического пространства, если для онтологии Ω отсутствует задание множества формирующих правил Pr .

Архитектура процесса разработки ИНТ предполагает последовательное выполнение следующих шагов создания ИНТ – **концептуального, логического и физического**. Каждый из этих шагов разделяется на несколько этапов создания технологии.

Концептуальный шаг включает в себя следующие этапы. Первый этап (K1) – формулировка принципов и гипотез, которые лежат в фундаменте ИНТ. Второй этап (K2) – разработка концептуальной модели базы знаний на основе принципов, сформулированных на предыдущем этапе K1. Третий этап (K3) – разработка моделей решения задач, опирающихся на информацию из базы знаний с учетом представления этой базы. Четвертый этап (K4) – создание моделей концептуального представления самих задач и концептуальных механизмов преобразования информации для решения задач.

Логический шаг проектирования ИНТ заключается, во-первых, в согласовании концептуальной модели базы знаний с механизмами преобразования информации. Во-вторых, с выбором модели устранения или, по крайней мере, упорядочивания методов элиминации неопределенностей при анализе знаний, утверждений, условий, связываемых с решаемыми задачами. В процессе логического проектирования конструируется логическое пространство решения задач при применении ИНТ, создается логическая модель ИНТ. Но о физической реализации самой технологии речь не идет.

Физический шаг разработки ИНТ предполагает реализацию логической модели технологии в виде базы знаний, ее системы управления, и программной системы, обеспечивающей выполнение основных процессов работы ИС, решающей поставленные задачи.

Поскольку ИНТ тесно связана с внешней средой, то процесс ее разработки может носить итеративный характер, где итерация определяется постепенным

расширением внешней среды, за счет присоединения новых предметных областей. Например, при разработке ИС, ориентированной на доказательство алгебраических утверждений, предметная область, определяемая конечными алгебраическими группами с их многообразными разветвлениями расширяется за счет добавления бесконечных, а затем и топологических групп. Или к теории колец добавляются модули, а затем и поля. Так строится тот раздел математики, который называется алгеброй.

В то же время можно создавать ИНТ ориентированные на ситуационный экономический или политический анализ. Здесь будут работать другие механизмы вывода, более близкие к тем, которые характерны для людей анализирующих аналогичные ситуации. Но методы построения архитектуры таких технологий и систем остаются одинаковыми, хотя концептуальные и логические шаги проектирования меняются. Отличается используемая информация и механизмы ее обработки. Вводятся вероятностные методы оценки ситуаций и рисков. Но то же самое находим и у людей – математики по технологиям своей работы отличаются от политологов, а языковеды – от экономистов.

Создание ИС для научных исследований является частью является одним из направлений, связанных с разработкой новых средств, используемых в научной деятельности. В большинстве случаев это направление реализуется в виде систем, обеспечивающих эту деятельность подходящей научной информацией. Это очень важная проблема, поскольку общий рост объемов исследований и публикаций научных результатов, увеличивает затраты на предварительный поиск и анализ этой информации. Информационные поисковики выдают десятки тысяч сайтов, относящихся к запросу, но иногда пропускают существенную информацию. Поэтому решение проблемы отыскания в современных информационных сетях адекватных необходимости результатов и методов является одним из важных направлений обеспечения научных разработок.

Не менее важна и вторая проблема: поиска в научных книгах и статьях тех методов, которые могут пригодиться в конкретном научном исследовании. Существуют десятки работ, относящихся в одной и той же научной теме, написанных на разных языках и в разное время. Хотелось бы, чтобы ИС по заданной ситуации нашла и указала ту часть статьи или книги, которая относится к этой ситуации.

Решение обеих проблем так или иначе сводится к пониманию текстов, представленных на одном или нескольких языках. Понимание рассматривается как представление утверждений или текстов с помощью понятий и отношений, заранее определенных для заданной предметной области. Но это общая проблема любой ИС.

Кардинальная задача анализа текстов – перейти от понимания, основанного на сопоставлении наборов ключевых слов, к пониманию утверждений и результатов на основе онтологических представлений соответствующих предметной области, определяемой научным исследованием. Поэтому и направление работы лежит рамках разработки универсальных ИС, пригодных для использования в широком классе научных разработок в основном за счет изменения соответствующей базы знаний, которая настраивается на избранную предметную область.

PROBLEMS OF DEVELOPMENT OF INTELLECTUAL SYSTEMS

Mejtus V. Yu.

*International Research and Training Center for Information Technologies and Systems
NAS and MES Ukraine
Kiev, 03680 GSP, Ukraine
Ph. 044-5261319; e-mail: vmeitus@gmail.com*

Abstract — The problems connected to creation of intelligence systems are considered. There are problems of information representation, transformation of this information in the process of search of the solving a problem and the issue of uncertainty connected to ambiguity, an inaccuracy of the information definition. It is offered to consider the ontology approach based on usage of concepts and relations, relevant to a natural language, and the constructions constructed of them for formalization of process of the decision of tasks by intellectual system.

I. Introduction

Intelligence is ability of a system to create, develop, apply methods and processes of the coordination of observations and the facts connected to an external environment, with knowledge of system of this environment, and on this basis to build behaviour models which are adequate to this environment.

The intellectual system is a system in which the mechanism of formation of intellectual behaviour is put at the decision of the tasks connected to certain data domain. The system is intended for achievement of the given purposes and results in data domain which represents a part of an external environment and characterized by some set of conditions, knowledge, activity scenarios.

II. Problems of Creation of Intellectual Systems

The general technology of creation of an intellectual system (IS) - the system possessing intelligence, - is based on development of following directions of creation of components of this system, fulfilling the separate processes providing functioning of the system as an integral part.

As per the problems connected to development of intellectual technologies and demanding the prime solution is as follows:

- the first group of problems is defined as a choice of the form of information representation which uses the IS,
- the second group is problems of a choice of adequate mechanisms of the information transforms corresponding to concepts, relations and properties of data domain,
- the third group of problems is problems of uncertainty of the information. This class of problems is connected to incompleteness and inaccuracy of a language in which the person represents the knowledge and facts connected to knowledge.

For formalization of process of the decision of tasks of the IS it is offered to consider the ontology approach based on usage of concepts and relations, considered in a natural language. The **ontology model** Ω is represented as seven characters $\Omega = (E, N, R, Pr, \Lambda, I, \lambda)$, where E - data domain representation for which the ontology is under construction, in space of signs V, N - the specification of concepts of area E, their ontological properties and signs, R - the specification of relations between concepts of area E, Pr - result rules in ontology of new concepts from initial concepts and relations, Λ - the logic selected for estimation of a character of correspondence between ontological specifications and data domain, I - the operator of interpretation, $I: N \cup R \rightarrow E$, λ - a logic function. If it is clear about what there is a speech we can use expression "ontology Ω " meaning ontology model Ω .

Expanding ontology at the expense of adding to it extensional and intentional concepts and relations of ontology and entering operations of a combination of elements of the replenished set, we obtain the ontology space (O-space) in which it is possible to solve immediately the problems presented by means of concepts and relations of ontology.