

УДК 004.89

И.В. ШОСТАК, М.А. ДАНОВА

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Украина

ТЕХНОЛОГИЯ СИНТЕЗА КОНЦЕПТУАЛЬНЫХ МОДЕЛЕЙ В РАМКАХ КОМПЬЮТЕРИЗАЦИИ ПРОЦЕССА ПРОГНОЗИРОВАНИЯ НАУЧНО-ТЕХНИЧЕСКОГО РАЗВИТИЯ РЕГИОНА

Проведен обзор публикаций, посвященных особенностям применения технологии Форсайт для прогнозирования научно-технического развития региона. Приведена постановка задачи синтеза моделей объектов, задействованных в процессе прогнозирования научно-технического развития. Сформулирован комплекс требований, которым должны удовлетворять модели, синтезированные в процессе прогнозирования. Описан онтологический подход к созданию компьютерной системы прогнозирования научно-технического развития. Показана адекватность онтологических моделей объектов, задействованных в процессе прогнозирования, целям моделирования.

Ключевые слова: онтологическая система, концептуальные модели, научно-техническое развитие, Форсайт, компьютерная система прогнозирования, онтологические модели знаний.

Введение

В современных условиях ведущие мировые державы стремятся занять лидирующие позиции в сфере высокотехнологичных производств и инновационных технологий.

Главной целью их научно-технической политики является повышение эффективности исследований в наиболее перспективных областях знаний [1, 2]. При этом ни одна из развитых, а тем более развивающихся, стран не может себе позволить вести полномасштабные исследования на современном уровне по всем научным направлениям, поскольку получение новых знаний требует больших затрат на оборудование и подготовку специалистов. В связи с этим возникла необходимость выбора приоритетных направлений, на которых должны быть сконцентрированы основные усилия правительства и инвестиции бюджетных средств.

Указанная задача, являясь важнейшей при формировании научно-технической политики, предполагает разработку инструментов определения научно-технологических приоритетов, а также механизмов их реализации [3, 4]. В большинстве государств периодически формируются специальные программы, определяющие приоритетные области развития науки и техники [5 – 8]. Методы, используемые в процессе разработки этих программ, получили обобщающее название Форсайт, от английского Foresight – «предвидение», и к настоящему моменту зарекомендовали себя как наиболее эффективный инструмент выбора приоритетов в сфере науки и технологий [9].

На сегодня отсутствует единая модель Форсайта [10], каждая страна приспособливает данную методологию под свои конкретные цели и потребности [11, 12]. Как в развитых странах мира, так и в нашей стране [3], одним из наиболее распространенных методов программы Форсайта стали перечни критических технологий. Этот подход применяется не только на уровне страны в целом, но и при прогнозировании развития отдельных отраслей экономики, тематических областей, регионов и других объектов [11 – 13]. Определение понятия регион и специфика процесса прогнозирования его научно-технического развития (НТР) была рассмотрена авторами ранее, в работе [14]. Результатом становятся, как правило, перечни технологий, либо направления исследований и разработок, которые требуют первоочередного внимания.

Реализация данного подхода осуществляется с привлечением большого числа экспертов из разных областей, а существующие на данный момент инструменты прогнозирования [15, 16] направлены, в основном, на автоматизацию конкретного (иногда нескольких) этапа процесса прогнозирования. Это, в свою очередь, порождает ряд проблем, связанных с субъективностью и трудоемкостью процесса прогнозирования, выбором методологии и инструментов для построения прогнозов, а также отсутствием комплексной его компьютеризации.

Цель данной статьи состоит в изложении онтологического подхода к созданию технологии синтеза концептуальных моделей в рамках компьютеризации процесса прогнозирования НТР региона.

1. Постановка задачи

В качестве исходных данных для прогнозирования НТР региона рассмотрим обязательные этапы метода критических технологий, как наиболее популярного, на сегодняшний день, варианта реализации методологии Форсайт. К этим этапам относятся: формирование группы экспертов; составление первоначального списка технологий; выбор системы критериев для оценки технологий; оценка первоначального списка технологий экспертами по выбранной системе критериев; формирование перечня критических технологий с учетом результатов оценки.

Необходимо получить процедуру синтеза моделей объектов, задействованных в процессе прогнозирования НТР. Построенная модель должна обеспечивать следующее:

- гомогенность среды моделирования (все модели ансамбля должны иметь одну и ту же математическую структуру);
- повторность использования;
- возможность верификации;
- адекватное представление трудно-формализуемых понятий ПрО;
- непосредственную трансляцию математической структуры в программный код.

2. Онтологический подход к созданию компьютерной системы прогнозирования НТР

Современные методы анализа документопотоков и баз метаданных, а также технологии управления информационно-аналитическими и интеллектуально информационными системами только начинают привлекаться для решения заданий научно-технологического прогнозирования. На сегодня в мировой науке сформировался ряд новых концепций относительно сохранения, обработки и анализа информации, которые позволяют за счет расширения использования формализованных методов анализа данных обеспечить высокую достоверность и оперативность научно-технологического прогнозирования. Среди них следует отметить:

- технологии интеллектуального анализа данных (data mining), выделения знаний, из баз данных (knowledge discovery);
- онтологические системы (ontology system);
- методы когнитивной графики и визуализации информации (knowledge visualizing, mapping scientific frontiers);
- концепцию информационных хранилищ (data warehouse), оперативной аналитической обработки данных (on-line analytical processing, OLAP, MOLAP);

- анализ веб-цитирования (mining a web citation databases);

- концепции открытых инноваций (open innovation) и „технологической разведки” (technology scanning/scan, technology sourcing, technology scouting).

Одним из важнейших и перспективных направлений в области формализации знаний, которое дает возможность использования накопленных знаний для компьютерной обработки, являются онтологии [17].

Онтология – целостная структурная спецификация некоторой предметной области, ее формализованное представление, которое включает словарь (или имена) указателей на термины предметной области и логические выражения, описывающие, как они соотносятся друг с другом [17].

Онтологический подход позволяет создавать системы, в которых накопленные знания становятся доступными для большинства пользователей. Основные преимущества этого подхода:

- онтология представляет пользователю целостный, системный взгляд на определенную предметную область (ПрО);
- знания о ПрО представлены единообразно, что упрощает их восприятие;
- построение онтологии позволяет восстановить недостающие логические связи ПрО.

Одна из наиболее важных функций онтологии – функция интеграции разнородных данных и знаний специалистов различных областей знаний. Онтологический инжиниринг подразумевает глубокий структурный анализ предметной области, а именно:

- выделение концептов – базовых понятий данной предметной области;
- определение «высоты дерева онтологий» – количества уровней абстракции;
- распределение концептов по уровням;
- построение связей между концептами (определение отношений и взаимодействий базовых понятий);
- консультации с различными специалистами для исключения противоречий и неточностей.

Методология построения онтологии предполагает рассмотрение следующих важных вопросов:

- обозначение целей и области применения создаваемой онтологии;
- построение онтологии, которое включает: фиксирование знаний о ПрО; кодирование; выбор или разработку специального языка для представления онтологии; непосредственно задание фиксированной концептуализации на выбранном языке представления знаний;
- совместное использование людьми или программными агентами общего понимания структуры информации;

- обеспечение возможности использования знаний ПрО;
- создание явных допущений в ПрО, лежащих в основе реализации;
- отделение знаний ПрО от оперативных знаний;
- анализ знаний в ПрО.

Современный уровень технологий построения сложных информационных систем характеризуется модульностью построения проекта, достигаемой за счет алгоритмической или объектно-ориентированной декомпозиции. Следствием этого стремления является иерархичность проекта, формирование которой представляет собой нетривиальную задачу, и только если уровень абстрагирования лежит в пределах $7+/-2$ уровня, то возможно конгруэнтное управление системой в целом. Ограничение в $7+/-2$ уровня абстрагирования, делает невозможным применение стандартных технологий для построения такой высокоразвитой экспертной системы (ЭС), как система прогнозирования НТР (СПНТР).

Для построения ЭС традиционно используют продукционные, семантические, фреймовые, нейронные и другие модели, которые позволяют обойти эти ограничения. Так, к примеру, база знаний (БЗ) ЭС может содержать многие тысячи продукционных правил, однако, это делает проблематичным использование объектно-ориентированных моделей и, в некоторой степени, алгоритмических [17].

В свою очередь, применение только объектно-ориентированного метода приводит к потере масштабируемости системы непрофессиональным программистом (пользователем системы).

Решением, позволившим получить преимущества использования объектно-ориентированных моделей и традиционных моделей для создания ЭС, является применение онтологических моделей знаний [18].

Для придания математической строгости представлению знаний в форме онтологической модели целесообразно вначале транслировать концепты описания предметной области в теоретико-множественную модель, на которой затем определяется модель формальной семантики. Это дает возможность строго описать механизм взаимодействия онтологической системы с другими системами такой же природы в распределенной информационной среде.

2.1. Особенности онтологической компоненты в составе СПНТР

Предметную область «Прогнозирование НТР» представим двойкой объектов

$$KWG = \langle T_{const}, T_{dyn} \rangle, \quad (1)$$

где $T_{const} = \{t_c\}$, $c \in \hat{\Omega}$ – неизменяющаяся составляющая, которая в онтологической системе представлена графом всех возможных состояний; $T_{dyn} = \{t_d\}$, $d \in \tilde{\Omega}$ – динамическая составляющая, задающая правила переходов по графу состояний.

При этом Ω – множество индексов объектов ПрО, $\Omega \supset \hat{\Omega} \cup \tilde{\Omega}$. В такой интерпретации, в онтологической компоненте составляющая T_{const} соответствует по форме БЗ продукционного типа, T_{dyn} – рабочей памяти, которая описывает текущую стадию решения задачи прогнозирования.

Для представления знаний в виде онтологической модели целесообразно рассматривать статическую составляющую, как множество концептов T , полученных в ходе таксономической декомпозиции. Эта одна из ранних стадий разработки онтологической модели. Начальным этапом является стадия выделения назначения концептуальной модели, области определения (границ ПрО) и области значений [19]. После этого осуществляется сбор информации о ПрО и выделение ее концептов – так называемый этап таксономической декомпозиции.

Связь любой пары концептов t_i, t_j ; $i \in \hat{\Omega}, j \in \tilde{\Omega}, i \neq j$ из T_{const} может быть одного из трех типов: «is_a»; «part_of»; «:» (наследование).

Определение 1. Словарь онтологической модели, расширенный типами и диапазонами значений из некоторой системы типов X является полной решеткой.

Словарь онтологии X может быть представлен линейно упорядоченными доменами, являющимися множеством потенциально допустимых значений данного типа.

Воспользуемся следующим формальным описанием онтологической структуры:

$$\Phi : I \rightarrow T, \quad (2)$$

где I – нерасширенный словарь онтологии; T – множество концептов; Φ – функция состояния онтологической структуры.

Причем отображение множества идентификаторов I во множество концептов T такое, что

$$\forall \tau \in T \exists I_s \mid \tau : I_s \rightarrow S, \quad (3)$$

где S – множество полей.

Отображение множества идентификаторов полей концепта во множество полей S , однозначно определяет сущность (концепт), что, в свою очередь, дает возможность введение операций над множеством $S \equiv \text{expr}(X)$, представляющие множество λ – выражений над системой типов X :

$$\text{expr}(X) \subset \Lambda(X), \quad (4)$$

включающих в себя отношения между концептами, ссылки на другие поля концептов, а так же ссылки на другие концепты.

При создании онтологической компоненты СПНТР необходимо введение отношения наследования « : » между концептами, определяющее иерархию понятий предметной области и соответственно используемое для заимствования свойств (полей) правил от родительских классов. Необходимо различать одиночное и множественное наследование: в первом случае отношение « : » является функцией по правому аргументу, т.е. одному дочернему концепту может соответствовать только один родитель, что трансформирует иерархию в дерево; во втором случае такое ограничение отсутствует. Кроме того, необходимо также различать статическое наследование, при котором отношение наследования остается постоянным в процессе логического вывода, и динамическое наследование, когда конфигурация иерархии наследования в процессе вывода изменяется.

Будем рассматривать отношение наследования, индуцированное значением полей “parent” дочерних концептов, т.е.

$$T : G \Leftrightarrow \|T.parent\| = G, \quad (5)$$

где $\|T.parent\|$ - операция вычисления значения поля, которая будет уточняться далее в соответствии с определяемой семантикой.

Более того, если допустить значение поля “parent” некоторого спискового типа, то возможно определить отношение множественного наследования как $T : G \Leftrightarrow \|T.parent\| \in G$. Определенное таким образом отношение наследования будет динамическим, так как значение поля будет вычисляться в соответствии с правилами определенной ниже семантики, т.е. будет возможно определение родителя посредством продукционных правил, либо операции спецификации.

Выводы

Построение концептуальных моделей ПрО «Прогнозирование НТР» в форме онтологической системы даст возможность решить ряд важных прикладных проблем, а именно:

- снизить сложность формализации неявных правил, отношений и обычного восприятия, а также формализовать правила, неподдающиеся описанию с помощью языка OWL (Web Ontology Language);

- обеспечить возможность в строго формализованной форме проводить операции верификации БЗ, а также обеспечивать внутреннюю и внешнюю согласованность знаний в ней;

- проводить непосредственную трансляцию на основе формального синтаксиса и семантики с языка представления знаний во внутреннее представление СПНТР. Формальная семантика может лечь в основу

реализации соответствующего программного прототипа на функциональном языке программирования;

- формальное описание позволяет однозначно и строго представить синтаксис и семантику языка, исключая этим различные толкования языковых конструкций;

- найдет свое решение задача управления, наполнения, осуществления поиска, предоставления для повторного использования данных и знаний и поддержания онтологий с ростом количества аксиом, правил, концептов и концептуальных отношений.

Литература

1. Денисов, Ю.Д. *Технологическое прогнозирование и научно-технические приоритеты в индустриально развитых странах [Текст]* / Ю.Д. Денисов, А.В. Соколов. – М.: ЦИСН, 1998. – 235 с.
2. Зинов, В.Г. *Прогнозирование научно-технического развития [Текст]: учеб. пособие* / В.Г. Зинов, Ю.Н. Кусакина. – М.: ГАУ, 2005. – 130 с.
3. *Про пріоритетні напрями розвитку науки і техніки [Текст]: закон України від 11 липня 2001р. № 2623-III // Відомості Верховної Ради України. 2001. – № 48. – С. 253.*
4. *Про пріоритетні напрями інноваційної діяльності в Україні [Текст]: закон України від 16 січня 2003р. № 433-IV // Відомості Верховної Ради. – 2003. – № 13. – С. 354.*
5. Loveridge, D. *United Kingdom Foresight Programme [Text]* /D. Loveridge, L. Georghiou, M. Neveda PREST. – University of Manchester, 2001 – 200 p.
6. Попович О.С. *Стан формування цілісної системи пріоритетних напрямів розвитку науки і техніки в Україні* / О.С. Попович // *Проблеми науки.* – 2002. – № 7. – С. 31 – 35
7. *Об утверждении плана мероприятий по реализации основ политики Российской Федерации в области развития науки и технологий на период до 2010 года и дальнейшую перспективу [Текст]: распоряжение Правительства РФ от 7 февраля 2006 г. № 156 // Собр. законодательства РФ. – 2003. – № 30. – С. 150*
8. Малицький, Б.А. *Методичні рекомендації щодо проведення прогностно-аналітичного дослідження в рамках Державної програми прогнозування науково-технологічного та інноваційного розвитку України [Текст]* / Б.А. Малицький, О.С. Попович В.П. Соловійов. – К.: Фенікс, 2004. – 52 с.
9. *Что такое Форсайт? Форсайт-центр: государственный комитет – Высшая школа экономики [Электронный ресурс]. – Режим доступа: <http://foresight.hse.ru>. – 10.05.2012 г.*
10. Шелюбская, Н.В. *Форсайт – механизм определения приоритетов формирования общества знаний стран Западной Европы [Текст]* / Н.В. Шелюбская. – К.: Фенікс, 2007. – 60 с.
11. Ben, R. Martin. *Technology foresight in a rapidly globalizing [Text]* / R. Martin. Ben // *International*

Practice in Technology Foresight. Vienna: UNIDO. – 2002. – 14 p.

12. *Winning Through Foresight: A Strategy Taking the Foresight Programme to the Millennium [Text]. // Office of Science and Technology. – London, 1996. – 117 p.*

13. Дуб, А.В. *Инновационные приоритеты для энергетического машиностроения: опыт отраслевого Форсайта [Текст] / А.В. Дуб, С.А. Шаинов // Форсайт. – 2007. – №3. – С. 4 – 11.*

14. Шостак, И.В. *Проблемы комплексной компьютеризации процесса прогнозирования научно-технического развития региона [Текст] / И.В. Шостак, М.А. Данова // Радиоэлектронні і комп'ютерні системи. – 2012. – № 7 (59). – С. 236 – 240.*

15. *Программный продукт FORECAST [Электронный ресурс]. – Режим доступа к программе: <http://www.forecastpro.com>. – 10.05.2012 г.*

16. *Программный продукт STATISTICA [Электронный ресурс]. – Режим доступа к программе: <http://www.statsoft.ru>. – 10.05.2012 г.*

17. Буч, Г. *Объектно-ориентированный анализ и проектирование с примерами на C++ [Текст] / Г.Буч. – М.: Бином, 2001. – 558 с.*

18. Лапишин, В.А. *Онтологии в компьютерных системах [Текст] / В.А. Лапишин. – М.: Научный мир, 2010. – 224 с.*

19. Загоруйко, Н.Г. *Прикладные методы анализа данных и знаний [Текст] / Н.Г. Загоруйко. – Новосибирск, 1999. – 215 с.*

Поступила в редакцию 4.06.2012

Рецензент: д-р техн. наук, проф., проф. кафедры ПО ЭВМ С.Ю. Шабанов-Кушнаренко, Харьковский национальный университет радиоэлектроники, Харьков, Украина.

ТЕХНОЛОГІЯ СИНТЕЗУ КОНЦЕПТУАЛЬНИХ МОДЕЛЕЙ В РАМКАХ КОМП'ЮТЕРИЗАЦІЇ ПРОЦЕСУ ПРОГНОЗУВАННЯ НАУКОВО-ТЕХНІЧНОГО РОЗВИТКУ РЕГІОНУ

I.V. Shostak, M.A. Danova

Проведений огляд публікацій, присвячених особливостям вживання технології Форсайт для прогнозування науково-технічного розвитку регіону. Приведена постановка завдання синтезу моделей об'єктів, задіяних в процесі прогнозування науково-технічного розвитку. Сформульований комплекс вимог, яким повинні задовольняти моделі, синтезовані в процесі прогнозування. Описаний онтологічний підхід до створення комп'ютерної системи прогнозування науково-технічного розвитку. Показана адекватність онтологічних моделей об'єктів, задіяних в процесі прогнозування, цілям моделювання.

Ключові слова: онтологічна система, концептуальні моделі, науково-технічний розвиток, форсайт, комп'ютерна система прогнозування, онтологічні моделі знань.

TECHNOLOGY OF SYNTHESIS OF CONCEPTUAL MODELS IN THE FORECASTING PROCESS OF COMPUTERIZATION OF FORECASTING SCIENTIFIC AND TECHNICAL DEVELOPMENT OF REGION

I.V. Shostak, M.A. Danova

The review of the publications devoted to features of application of technology Foresight for forecasting of scientific and technical development of the region is carried out. Statement of a problem of synthesis of models of the objects involved in process of forecasting of scientific and technical development is given. The complex of requirements with which should satisfy the models synthesized in the course of forecasting is formulated. The ontologic approach to creation of computer system of forecasting of scientific and technical development is described. Adequacy of ontologic models of the objects involved in process of forecasting is shown, to the modeling purposes.

Key words: ontological system, the conceptual model, scientific and technical development, Foresight, computer forecasting system, ontological knowledge model.

Шостак Игорь Владимирович – д-р техн. наук, профессор, профессор каф. инженерии ПО Национального аэрокосмического университета им. Н.Е. Жуковского «ХАИ», Харьков, Украина.

Данова Мария Александровна – аспирант каф. инженерии ПО Национального аэрокосмического университета им. Н.Е. Жуковского «ХАИ», Харьков, Украина, e-mail: danovamariya@gmail.com.