

ИСПОЛЬЗОВАНИЕ ПАДЕЖНОЙ ГРАММАТИКИ ПРИ ИНФОРМАЦИОННОМ ПОИСКЕ В БАЗЕ ЗНАНИЙ ЭКСПЕРТНОЙ СИСТЕМЫ О КОНСТРУКЦИЯХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Н. И. Сидняев¹, Ю. И. Бутенко², Е. Е. Синева³

Аннотация: Раскрыты основные принципы построения экспертных систем в области авиационной промышленности. Предложен логико-лингвистический подход к информационному моделированию предметной области при проектировании базы знаний экспертной системы. Разработана модель представления знаний для описания конструкций летательных аппаратов, обеспечивающая информационно-структурную надежность базы знаний экспертной системы. Проведен анализ парадигматических особенностей авиакосмической терминологии с целью улучшения эффективности информационного поиска. Предложен метод разрешения лексической многозначности поискового запроса в базе знаний. Описан механизм вывода информации из базы знаний на основе падежной грамматики Ч. Филлмора, позволяющей определить семантико-синтаксическую структуру выводимого предложения. Показано, что посредством расстановки ограничений на число участников ситуации, которые может присоединять глагол, можно обеспечить эффективный вывод информации из базы знаний.

Ключевые слова: экспертная система; база знаний; летательный аппарат; терминология; знания

DOI: 10.14357/08696527210307

1 Введение

Поддержка принятия решений в сложных ситуациях необходима в различных областях деятельности, где требуются обоснованные, логически доказуемые аргументы, в том числе в оперативном и стратегическом управлении технологиями на авиационных предприятиях. При принятии решений необходимо проведение детального моделирования последствий предполагаемого решения, поиск оптимального пути достижения заданного результата при помощи технологий имитационного моделирования. С подобными задачами успешно справляются

¹Московский государственный технический университет им. Н. Э. Баумана, sidnyaev@yandex.ru

²Московский государственный технический университет им. Н. Э. Баумана, iuliiabutenko2015@yandex.ru

³Московский государственный технический университет им. Н. Э. Баумана, bolotovaee@mail.ru

экспертные системы — комплексы программных средств, способные заменять специалиста при решении сложных задач, возникающих в процессе диагностики, проектирования или эксплуатации аэрокосмической и другой техники [1]. Такие системы основаны на знаниях, полученных в процессе взаимодействия с экспертами в конкретной предметной области.

Параллельно с развитием различных направлений исследований в области искусственного интеллекта идет развитие информационных структур для представления знаний [2, 3]. Появились новые способы описания и представления данных, возникли фреймовые, списочные и иерархические структуры. Процесс построения баз знаний достаточно сложен и носит итеративный характер, заключающийся в циклической модификации баз знаний на основе результатов ее тестирования [4].

2 Абстрактная иерархическая структура как способ описания конструкций летательных аппаратов

Основной задачей проектирования базы знаний в области авиакосмического приборостроения является описание структуры летательного аппарата таким образом, чтобы оно предоставляло наиболее полную и непротиворечивую информацию об описываемом объекте. Если применить лингвистические модели в совокупности с математическим аппаратом, можно достичь более ясного и результативного описания и решения сложной технологической проблемы. Информация воспринимается специалистом на когнитивном уровне посредством основных единиц этого уровня — понятий, идей, концептов, поэтому для представления знаний часто применяется аппарат семантических сетей, который позволяет воссоздать модель, схожую с процессами мышления конструктора, технолога [5, 6]. В инженерии знаний под семантической сетью подразумевается граф с узлами, отображающими объекты предметной области, и дугами, обозначающими отношения между данными понятиями.

Для описания конструкций летательных аппаратов разработана интегрированная модель представления знаний — абстрактная иерархическая структура, объединяющая графовую структуру, фреймы и предикаты. Все знания о предметной области представлены в виде графа, в узлах которого содержатся фреймы — минимально возможные описания сущностей, записанные с помощью предикатов и предикатных символов [1]. Подобная структура гарантирует непротиворечивость описаний объектов и их отношений в реальном производстве.

3 Отражение парадигматических особенностей авиакосмической терминологии в базе знаний

Неотъемлемой частью любого управления является его документационное обеспечение, основа которого — однозначно понимаемая и непротиворечивая терминология. Единое терминологическое обеспечение позволяет избегать недо-

понимания между специалистами как в рамках одной предметной области, так и за ее пределами.

Терминосистема — это некоторая лингвистическая модель, представляющая конкретную предметную область [7]. Аэрокосмическая терминология русского языка представляет собой сложную систему, включающую термины из различных тематических пространств: обозначения субъектов аэрокосмической сферы (*космонавт, летчик, пилот*), наименования организаций и структур области авиации и космонавтики (*НАСА, ИАТА, ИФАТКА*), обозначения действий аэрокосмической деятельности (*взлет, посадка, пилотирование, пристыковка*), наименования деталей и частей летательных аппаратов (*фюзеляж, крыло, подкрылки, двигатель*), обозначение характеристик и состояний летательных аппаратов (*работоспособность, надежность, безопасность*). Стандартизация терминологии данной предметной области — процесс трудоемкий, а отсутствие стандартизации приводит к появлению разных терминов для обозначения одних и тех же объектов и явлений, в связи с чем возникают случаи синонимии и лексической многозначности.

При построении баз знаний интеллектуальных систем синонимия считается крайне неблагоприятным явлением. Различия в номинации объектов приводят к трудностям в сфере профессиональной коммуникации, поэтому базу знаний экспертной системы предлагается пополнять всеми возможными вариантами синонимов тех или иных узкоспециальных и общенаучных терминов.

В качестве иллюстрации приведем несколько примеров синонимичных терминов: *интерцептор–спойлер; пусковое устройство–пусковая кнопка–стартер; шкала дальности–дистанционная шкала; спуск–снижение; приборная доска–приборная панель; маслоуловитель–маслосборник.*

В аэрокосмической терминологии также наблюдается семантическая оппозиция номинаций «свой–чужой», в результате чего возникают синонимичные пары, например: *космонавт–астронавт* (англ. astronaut); *космонавтика–астронавтика* (англ. astronautica); *высотомер–альтиметр* (франц. altimètre).

Равнозначность подобных синонимов необходимо отразить в проектируемой базе знаний во избежание недопонимания между экспертной системой и ее пользователями. Стоит отметить, что редко используемые аббревиатуры должны быть расшифрованы и понятны пользователям экспертной системы (например, *ШВП* — это *шасси на воздушной подушке* или *шаровой вытяжной парашют*).

Значительные трудности в процессе поиска информации в базе знаний вызывают явления лексической многозначности, т. е. сосуществование различных значений одного слова или фразы, например:

- Свеча* — 1. Приспособление для воспламенения горючей смеси (в двигателях внутреннего сгорания); 2. Крутой взлет, подъем (фигура высшего пилотажа).
Хвост — 1. Хвостовое оперение летательных аппаратов; 2. Вытянутое из пыли и газа кометное вещество, образующееся при приближении кометы к Солнцу.
Атмосфера — 1. Газообразная оболочка, окружающая Землю и некоторые другие планеты; 2. Единица измерения давления.

4 Применение падежной грамматики Ч. Филлмора для вывода знаний из экспертной системы

Один из необходимых компонентов базы знаний экспертной системы — механизм вывода, который содержит правила для решения конкретной задачи. Механизм вывода ссылается на информацию из базы знаний и выбирает факты и правила, которые будут применяться при попытке ответить на запрос пользователя. Данный механизм обеспечивает аргументацию информации в базе знаний и позволяет человеку получать информацию из базы знаний в виде предложений, построенных на естественном языке. Одним из наиболее удачных механизмов, позволяющих моделировать смысл высказываний на естественном языке, считается механизм, основанный на теории падежных грамматик Ч. Филлмора [8]. В падежной грамматике глагол является центром предложения и диктует в силу своего значения набор ролей (глубинных падежей), реализующихся в предложении посредством именных форм. Иными словами, падежная грамматика показывает связь существительного или местоимения с другими словами в предложении. Обрамление падежами подвержено определенным ограничениям, например каждый семантический падеж может встречаться в предложении только один раз [9]. Некоторые падежи являются обязательными, другие — необязательными. Обязательные падежи нельзя удалять, рискуя получить грамматически неправильные предложения.

Для описания события в каждом фрейме иерархической структуры в первую очередь выделяется действие, которое обычно описывается глаголом. Далее определяются: объект, который действует — агенс; объект, над которым это действие выполняется — пациенс [10].

На каждое отношение накладывается множество ограничений, например структура глагола «пилотировать» может включать следующие семантические падежи: (1) агентив — *летчик пилотировал самолет*; (2) пациентив — *летчик пилотировал **самолет***; (3) инструменталис — *летчик пилотировал самолет **боковой ручкой управления***; (4) локатив — *летчик впервые пилотировал самолет **на аэродроме с плохим покрытием***. Возможно включение в структуру данного глагола и других семантических падежей, а также атрибутивных отношений, например: «*пилотировать каким образом?*» — «*умело*», «*профессионально*».

Такие ограничения необходимо накладывать для того, чтобы система могла строить правильные семантико-синтаксические структуры, что обеспечит связность и внутреннюю интерпретируемость знаний в экспертной системе.

Падежные отношения, а также соответствие между синтагматикой и линейной упорядоченностью слов в предложении можно представить в виде дерева непосредственно составляющих. Глубинная структура пропозиционального компонента любого простого предложения представляет собой конструкцию, состоящую из глагольной и именной группы, которые находятся в специальных помеченных отношениях (семантических падежах) ко всему предложению. Дере-

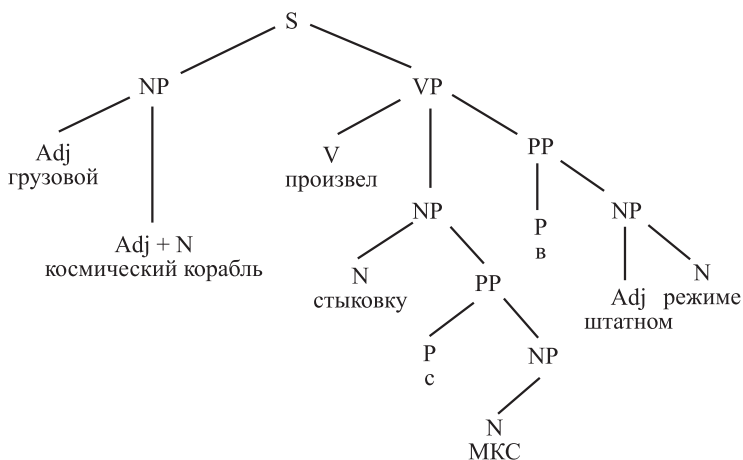


Рис. 1 Представление предложения в виде дерева составляющих

во составляющих позволяет представить синтаксическую структуру предложения в виде иерархии, составленной из связанных друг с другом непересекающихся единиц. Пример описания предложения «Грузовой космический корабль произвел стыковку с МКС в штатном режиме» с помощью дерева составляющих показан на рис. 1.

Исходя из этого, модель семантической целостности предложения (C) можно представить следующим образом:

$$C = \langle V^I, N^I \rangle,$$

где V — глагольная составляющая; N — именная составляющая; I — семантический падеж. Для вывода информации из базы знаний требуются также атрибутивные отношения для описания формы, цвета, размера и других характеристик объектов, а также дополнительные падежи цели, условия, времени и др., т. е. в зависимости от целей экспертной системы перечень, содержащий 6 основных падежей, может быть расширен.

Добавим к структуре рассматриваемого предложения дополнительные семантические падежи и применим теорию семантических падежей для описания семантической структуры знания о событии. Допустим, системе нужно вывести сообщение «Грузовой космический корабль пристыковался к МКС в 15:00 мск в штатном режиме. Груз доставлен на станцию». Семантическая сеть, описывающая событие доставки груза на МКС, представлена на рис. 2.

В процессе поиска информации в базе знаний, основанной на семантической сети, происходит сопоставление общей сети с сетью запроса для поиска нужного фрагмента информации. При выводе в семантических сетях также применяется метод перекрестного поиска, в процессе которого сопоставляются не узлы семантической сети, а дуги — отношения между объектами.



Рис. 2 Пример семантической сети, описывающей событие доставки груза на МКС

5 Заключение

В работе раскрыты основные принципы построения интеллектуальных систем в авиакосмической промышленности. Поднимается проблема стандартизации терминологии как одного из ключевых этапов разработки баз знаний экспертных систем с целью сведения разнообразно представленной информации к общепонятному виду, что позволит организовать ее совместное и многократное использование специалистами. Для выявления парадигматических особенностей терминологии авиакосмической области проведен анализ лексических единиц, входящих в тематическую группу «Авиакосмическое приборостроение». В связи с тем что стандартизация терминологии является трудоемким процессом, предложено пополнять проектируемую базу знаний всеми возможными вариантами синонимичных терминов. Рассмотрен механизм вывода информации из базы знаний экспертной системы на основе падежной грамматики Ч. Филлмора, которая позволяет определить семантико-синтаксическую структуру выводимого предложения. Показано, что посредством расстановки ограничений на число актантов (участников ситуации), которые может присоединять глагол, можно обеспечить эффективный вывод предложений на естественном языке.

Литература

1. Сидняев Н. И., Бутенко Ю. И., Болотова Е. Е. Экспертная система производственного типа для создания базы знаний о конструкциях летательных аппаратов // Авиакосмическое приборостроение, 2019. № 6. С. 38–52. doi: 10.25791/aviakosmos.06.2019.676.
2. Сигов А. С., Нечаев В. В., Кошкарёв М. И. Архитектура предметно-ориентированной базы знаний интеллектуальной системы // Int. J. Open Information Technologies, 2014. № 12. Р. 1–6.
3. Сидняев Н. И., Бутенко Ю. И., Болотова Е. Е. Логическая модель требований информационно-системной надежности для баз знаний интеллектуальных систем // Программная инженерия, 2020. № 4. С. 195–204. doi: 10.17587/prin.11.195-204.

4. *Helbig H.* Knowledge representation and the semantics of natural language. — Berlin/Heidelberg/New York: Springer, 2006. 655 p.
5. *Woods D. D.* Cognitive technologies: The design of joint human-machine cognitive systems // *AI Mag.*, 1985. Vol. 6. Iss. 4. P. 86–92.
6. *Ломов П. А., Шишаев М. Г.* Формирование когнитивных фреймов на основе онтологических паттернов для визуализации онтологий // *Информационные системы и технологии*, 2015. № 6(92). С. 12–22.
7. *Даниленко В. П.* Русская терминология: опыт лингвистического описания. — М.: Наука, 1977. 246 с.
8. *Fillmore Ch.* The case for case // *Universals in linguistic theory* / Eds. E. Bach, R. T. Harms. — New York, NY, USA: Holt, Rinehart, and Winston, 1968. P. 1–88.
9. *Богданов В. В.* Структурно-семантическая организация предложения. — Л.: ЛГУ, 1977. 205 с.
10. *Jackendoff R. S.* The status of thematic relations in linguistic theory // *Linguistic Inq.*, 1987. Vol. 16. P. 369–411.

Поступила в редакцию 18.08.21

USE OF CASE GRAMMAR IN INFORMATION SEARCH IN THE EXPERT SYSTEM KNOWLEDGE BASE ON AIRCRAFT STRUCTURES

N. I. Sidnyaev, Yu. I. Butenko, and E. E. Sineva

Bauman Moscow State Technical University, 5-1 2nd Baumanskaya Str., Moscow 105005, Russian Federation

Abstract: The paper reveals the basic principles of building expert systems in the aviation industry. A model of knowledge representation that provides information and structural reliability of the expert system's knowledge base for describing aircraft structures has been developed. The article analyzes the paradigmatic features of aerospace terminology in order to improve the efficiency of information search. A method for resolving lexical ambiguity of a search query in the knowledge base is proposed. Inference engine and explanation mechanism of the knowledge base premised on the Charles Fillmore's case grammar is described. The grammatical cases make it possible to determine the semantic and syntactic structure of the output sentence. Restrictions show the number of participants in the situation that a verb can attach, it is possible to provide effective output of information from the knowledge base.

Keywords: expert system; knowledge base; aircraft; terminology; knowledge

DOI: 10.14357/08696527210307

References

1. Sidnyaev, N. I., Yu. I. Butenko, and E. E. Bolotova. 2019. Ekspertnaya sistema produktsionnogo tipa dlya sozdaniya bazy znaniy o konstruktsiyakh letatel'nykh apparatov [Rule-based expert system for creating a knowledge base on aircraft structures]. *Aviakosmicheskoe priborostroenie* [Aerospace Instrument-Making] 6:38–52. doi: 10.25791/aviakosmos.06.2019.676.
2. Sigov, A. S., V. V. Nechaev, and M. I. Koshkaryov. 2014. Arkhitektura predmetno-orientirovannoy bazy znaniy intellektual'noy sistemy [Architecture of the subject-oriented knowledge base of the intellectual system]. *Int. J. Open Information Technologies* 12:1–6.
3. Sidnyaev, N. I., Yu. I. Butenko, and E. E. Bolotova. 2020. Logicheskaya model' trebovaniy informatsionno-sistemnoy nadezhnosti dlya baz znaniy intellektual'nykh sistem [Logical model of information and system reliability requirements for knowledge bases of intellectual systems]. *Programmnaya inzheneriya* [Software Engineering] 4:195–204. doi: 10.17587/prin.11.195-204.
4. Helbig, H. 2006. *Knowledge representation and the semantics of natural language*. Berlin/Heidelberg/New York: Springer. 655 p.
5. Woods, D. D. 1985. Cognitive technologies: The design of joint human-machine cognitive systems. *AI Mag.* 6(4):86–92.
6. Lomov, P. A., and M. G. Shishaev. 2015. Formirovanie kognitivnykh freymov na osnove ontologicheskikh patternov dlya vizualizatsii ontologiy [Approach to ontology visualization based cognitive frames]. *Informatsionnye sistemy i tekhnologii* [Information Systems and Technologies] 6(92):12–22.
7. Danilenko, V. P. 1977. *Russkaya terminologiya: opyt lingvisticheskogo opisaniya* [Russian terminology: Experience of linguistic description]. Moscow: Nauka. 246 p.
8. Fillmore, C. J. 1968. The case for case. *Universals in linguistic theory*. Eds. E. Bach and R. T. Harms. New York, NY: Holt, Rinehart, and Winston. 1–88.
9. Bogdanov, V. V. 1977. *Strukturno-semanticheskaya organizatsiya predlozheniya* [Structural and semantic organization of a sentence]. Leningrad: LSU. 205 p.
10. Jackendoff, R. S. 1987. The status of thematic relations in linguistic theory. *Linguistic Inq.* 16:369–411.

Received August 18, 2021

Contributors

Sidnyaev Nikolay I. (b. 1955) — Doctor of Science in technology, professor, Head of Department, Bauman Moscow State Technical University, 5-1 2nd Baumanskaya Str., Moscow 105005, Russian Federation; sid-nyaev@yandex.ru

Butenko Yuliya I. (b. 1987) — Candidate of Science (PhD) in technology, associate professor, Bauman Moscow State Technical University, 5-1 2nd Baumanskaya Str., Moscow 105005, Russian Federation; iuliiabutenko2015@yandex.ru

Sineva Elizaveta E. (b. 1998) — Master student, Faculty of Fundamental Sciences, Bauman Moscow State Technical University, 5-1 2nd Baumanskaya Str., Moscow 105005, Russian Federation; bolotovae@mail.ru