



УДК 004.822:514

## ПРОБЛЕМЫ СИНОНИМОВ В ТЕЗАУРУСЕ ИНТЕЛЛЕКТУАЛЬНОГО ПОМОЩНИКА ПРОЕКТАНТА

Боргест Н.М., Шустова Д.В., Одинцова С.А., Князихина Ю.Е.

Самарский государственный аэрокосмический университет имени академика С.П. Королева (национальный исследовательский университет), г. Самара, Россия

[borgest@yandex.ru](mailto:borgest@yandex.ru)

[ShustovaDV@yandex.ru](mailto:ShustovaDV@yandex.ru)

[sv-odincova@mail.ru](mailto:sv-odincova@mail.ru)

[jul4432@yandex.ru](mailto:jul4432@yandex.ru)

В работе рассматриваются проблемы, возникающие при создании тезауруса предварительного проектирования самолета в рамках проекта создания интеллектуального помощника проектанта. Приведены результаты разработки пилотской версии тезауруса с использованием предложенных решений.

**Ключевые слова:** информационно-поисковый тезаурус, эквивалентные классы, синонимы, многозначность.

### ВВЕДЕНИЕ

В русском языке, как и во многих других, есть большое количество многозначных слов, с которыми человек сталкивается в повседневной жизни. Вследствие различной формулировки и

длительный промежуток времени в различных предметных областях. Для описания терминов используется разрабатываемая онтология – система, состоящая из наборов понятий и

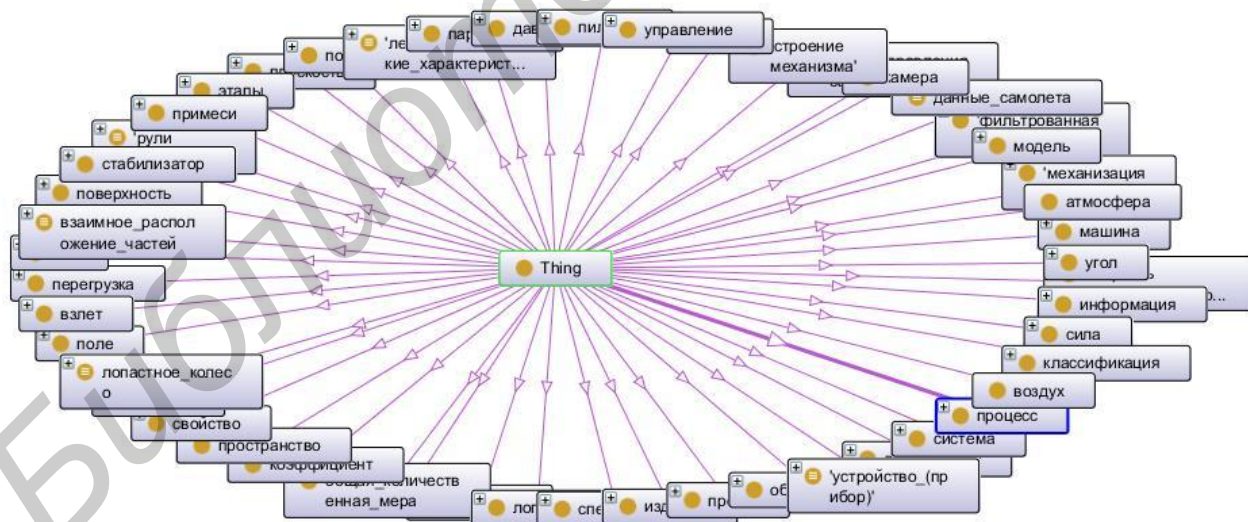


Рисунок 1- Основные классы тезауруса ПрО «Предварительное проектирование самолета»

интерпретации одних и тех же терминов, существует потребность в формировании словника, который даст единые и исчерпывающие знания о конкретном объекте. Разработка информационно-поискового тезауруса позволяет структурировать и унифицировать знания, накопленные за довольно

утверждений об этих понятиях, на основе которых можно описывать классы, отношения, функции и индивиды [Добров и др., 2009]. В данном случае, в качестве предметной области рассматривается этап предварительного проектирования самолета.

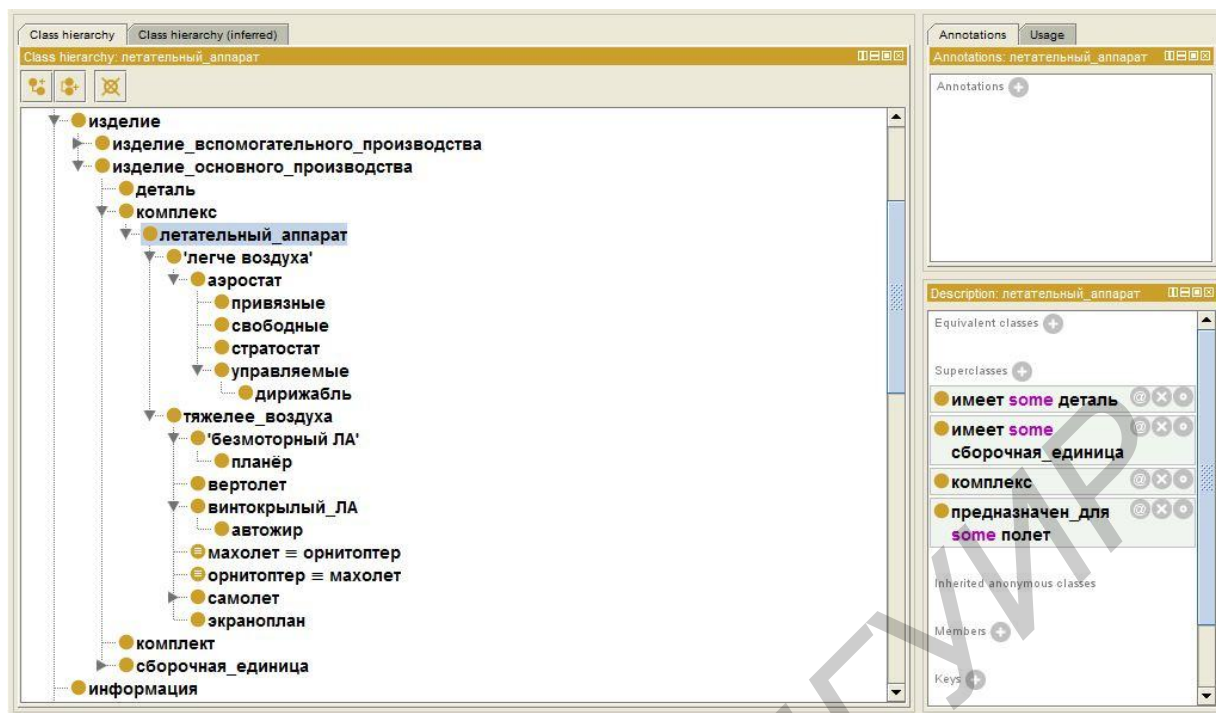


Рисунок 2 – Ассоциативные связи класса «Летательный аппарат»

Тезаурус представляет собой систематизированную совокупность понятий определенной отрасли науки, отражающих логические связи между терминами. Эти связи основываются на классовой иерархии, родовидовых и ассоциативных связях. Процесс визуализации такой системы упрощает понимание структуры рассматриваемой области.

На рисунке 1 показаны основные классы, входящие в состав тезауруса предварительного проектирования самолета.

Разработка тезауруса осуществляется на базе онтологического редактора Protégé, созданного в Стэнфордском университете. Несмотря на то, что Protégé, как и большинство конструкторов онтологий, представлен в англоязычной версии, он имеет интуитивный интерфейс, доступный подготовленному пользователю. Однако это не защищает последних от ряда проблем и ошибок при его использовании.

В первую очередь, логическое построение структур влияет на восприятие пользователем информационно-поискового тезауруса. Это является следствием того, что у каждого человека существует свое собственное понимание окружающего мира и предметной области.

Таким образом, несмотря на то, что разрабатываемый тезаурус ориентирован на использование различными пользователями, и, казалось бы, должен быть универсальным и контекстно адаптирован к пользователям разной профессиональной культуры, но в конечном итоге он является весьма субъективным проектом. Существующие терминологические ГОСТы, ОСТы, профессиональные энциклопедии и словари

лишь частично снимают вопрос о субъективности «готового продукта».

Во-вторых, сама по себе классовая иерархия не является универсальной, так как далеко не все термины можно описать с помощью отношений типа род-вид. Зачастую приходится использовать дополнительные ассоциативные связи (см. рисунок 2) для доопределения слов. Но и это не является гарантом того, что можно будет увидеть полную картину создаваемой предметной области.

Ранее, в публикации [Боргест и др., 2012] рассматривалась проблема задания иерархических и ассоциативных связей в тезаурусе. Было принято решение включения вспомогательных слов для установления связей ассоциации. Тезаурус концептуального проектирования самолета является инновационным проектом, поэтому процесс его создания сопровождается выявлением проблем, не рассматриваемых ранее. В ходе разработки проекта было выявлено несколько важных вопросов.

## ДОБАВЛЕНИЕ СИНОНИМОВ В ТЕЗАУРУС

Ограниченные возможности самого онтологического редактора Protégé устанавливают определенные рамки при его использовании. Например, в редакторе исключена возможность создания одинаковых терминов. С одной стороны, это помогает избежать повторного добавления слова. Этот фактор имеет большое значение в том случае, когда в тезаурус входит большое число терминов, так как не нужно проверять каждый добавляемый термин с помощью функции поиска, и таким образом, ускорить процесс работы. С другой стороны, отсутствие возможности создания

одинаковых терминов выявляет определенные трудности в случае добавления синонимов и омонимов.

В случае однозначности рассматриваемых терминов, в Protégé предусмотрена возможность создания эквивалентных классов, то есть значения этих слов приравниваются друг к другу. Но более

В случае задания термина «строение механизма» как синонима термина «устройство» получаем эквивалентный ряд: «прибор = устройство = строение механизма». Но при этом образуется нарушение логической цепочки, так как приравнивание терминов «прибор» и «строение механизма» не является верным. При поиске решения этой проблемы просматривались сайты и

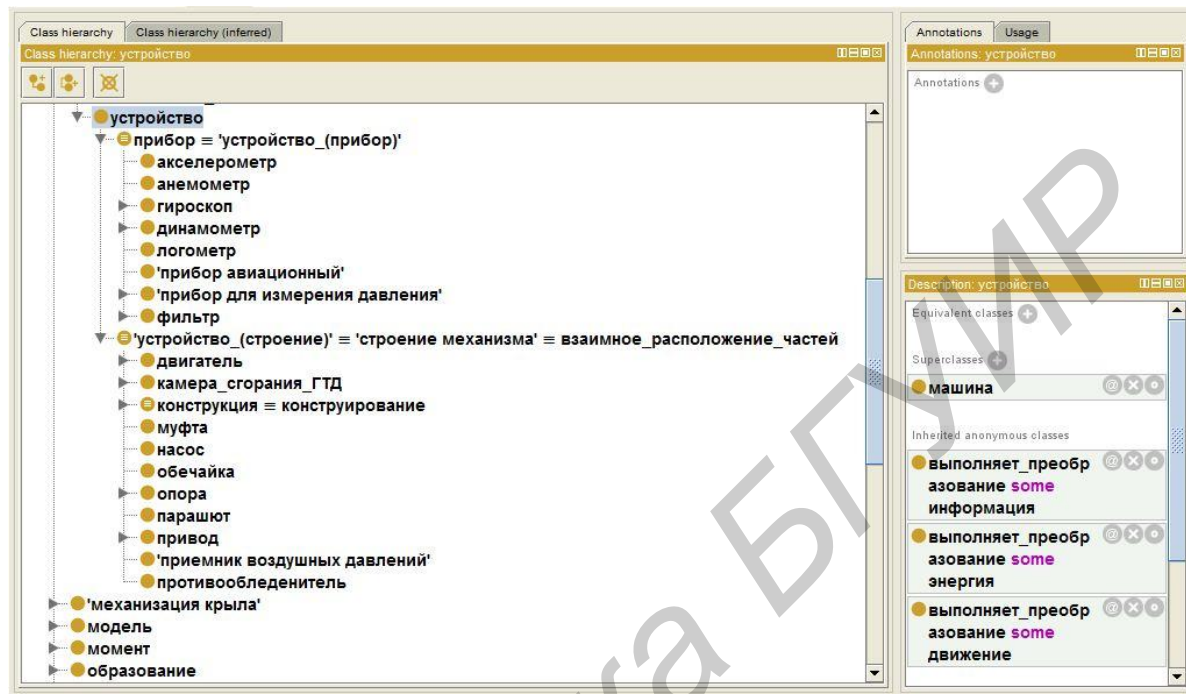


Рисунок 3 – Использование пояснительных слов в классе «Устройство»

сложная ситуация наблюдается в случае добавления многозначных, синонимичных терминов.

В качестве примера рассмотрим взаимосвязь слов «устройство», «строение механизма» и «прибор».

В процессе работы над тезаурусом в качестве главного, но не единственного источника используется словарь авиационных терминов [Боргест и др., 1992], который дает следующее определение прибора:

*«ПРИБОР АВИАЦИОННЫЙ – измерительное устройство, входящее в состав авиационной системы индикации экипажу самолета или вертолета и сочетающее в одном корпусе индикационную и приемно-измерительную часть или датчик».*

Таким образом, устройство и прибор являются синонимами, что задается в программе с помощью эквивалентных классов. В то же время, под устройством в авиационной отрасли зачастую понимается строение механизма.

форумы [Щербак, 2012], [Semantic, 2012], на которых ведется обсуждения вопросов, касающихся создания онтологий в Protégé. Однако эта тема пока не была затронута, поэтому было предложено следующее решение.

Для облегчения работы с синонимами и решения вопросов, касающихся эквивалентности классов в данном случае, было решено указывать значение, в котором употребляется данный термин, с помощью пояснительных слов. На рисунке 3 приведен пример использования пояснительных слов для класса «Устройство».

Таким образом, в графическом виде класс «Устройство» как «прибор» и как «строение механизма», будет иметь следующий вид (см. рисунок 4).

Подводя итог вышесказанному, можно сделать вывод, что использование пояснительных слов является рациональным способом выхода из ситуации, когда имеются 2 или более синонимичных слова.



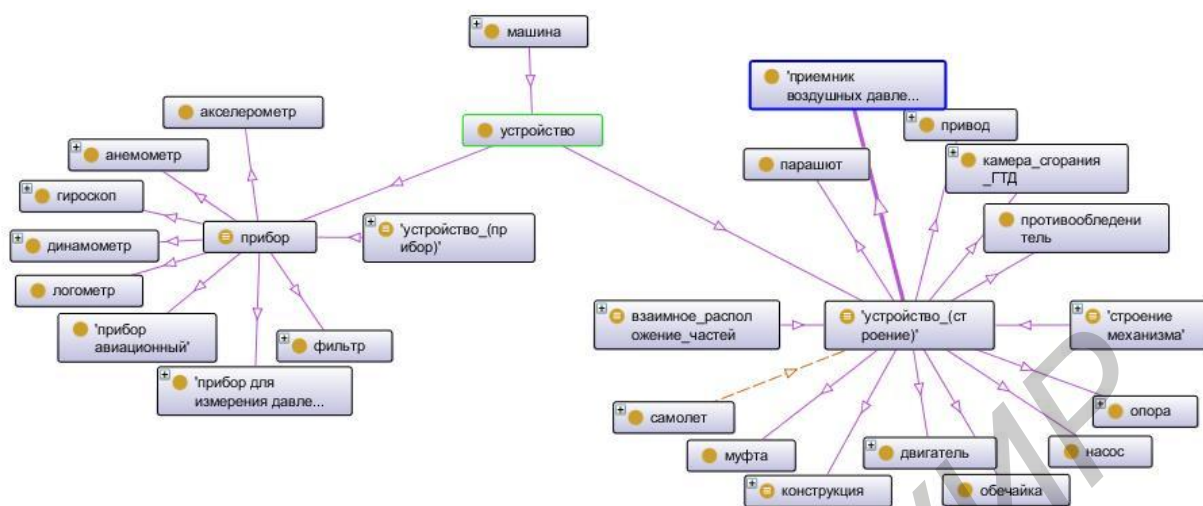


Рисунок 4 – Классификация термина «Устройство»

Одним из наиболее обсуждаемых в последнее время вопросов является объединение онтологий. С этой целью были созданы такие инструментальные средства, как PROMPT, Chimaera, OntoMerge и прочие [Овдей и др., 2004]. В интеллектуальном помощнике проектанта предметной области предварительного проектирования самолета используется онтологический редактор Protégé версии 4.1.beta.

## ЗАКЛЮЧЕНИЕ

В данной статье описываются возможные варианты решения синонимических проблем, возникающих в процессе создания тезауруса предварительного проектирования самолета для интеллектуального помощника проектанта.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- [Боргест и др., 1992] Боргест Н.М., Данилин А.И., Комаров В.А. Краткий словарь авиационных терминов/ Под ред. В.А. Комарова. – М.: Изд-во МАИ, 1992. – 224 с.
- [Боргест и др., 2012] Боргест Н.М., Гиматинова С.Р., Шустова Д.В., Иерархические и ассоциативные связи между терминами в тезаурусе на примере словаря проектанта, Вестник СГАУ, №3, 2012
- [Добров и др., 2009] Добров Б.В., Иванов В.В., Лукашевич Н.В., Соловьев В.Д.. Онтологии и тезаурусы: модели, инструменты, приложения. Изд-во: ИНТУИТру. 2009 – 176 с. <http://www.intuit.ru/department/expert/ontoth/1/>.
- [Овдей и др., 2004] Овдей О.М., Проскудина Г.Ю. Обзор инструментов инженерии онтологий. Электронные библиотеки. РНЭЖ. - 2004 - Том 7 - Выпуск 4.
- [Щербак, 2012] SHCHERBAK.NET - блог PhD Щербак Сергея - о том новом, что будет, когда мы забудем об HTML и вспомним об RDF. <http://shcherbak.net/>
- [Semantic, 2012] Semantic Future by SWUG. Форум Semantic Web User Group. <http://forum.semanticfuture.net/>

## PROBLEMS OF SYNONYMS IN THE THESAURUS OF INTELLIGENT ASSISTANT OF DESIGNER

Borgest N.M., Shustova D.V., Odintsova S.A., Knyazihina Y.E.

Samara State Aerospace University named after academician Korolev S.P. (National Research University), Samara, Russia

borgest@yandex.ru  
ShustovaDV@yandex.ru  
sv-odincova@mail.ru  
jul4432@yandex.ru

## INTRODUCTION

Thesaurus allows you to structure the knowledge gained in different fields. Special systems, called ontologies, are used for these purposes. Process of visualization of such a system simplifies the understanding of the structure of a considered area.

## MAIN PART

Ontological editor Protégé is used to develop a thesaurus. Its capacity is limited, and it can cause some problems. One of these problems is the difficulty of adding the same terms.

## CONCLUSION

In work the solution of problem, arising during creation of thesaurus of preliminary aircraft design area are proposed. Some issues to be addressed in the future research are also named. Intermediate results of thesaurus developing are presented.