

## **Раздел. Общие правила разработки унифицированных семантических моделей баз знаний**

В рамках *Технологии OSTIS* выделяют три основных метода разработки баз знаний.

**1. Тест-ориентированный метод разработки баз знаний.** Данный метод основывается на реализации стартового диалога с конечным пользователем системы. На первом этапе проектирования базы знаний исследуются запросы пользователей к проектируемой базе знаний и составляется тестовый сборник вопросов, охватывающий вопросы пользователей по рассматриваемой предметной области, что предполагает выделение семантически полного набора вопросов, ответы на которые должны содержаться в первой версии базы знаний. На следующем этапе все вопросы систематизируются и объединяются в классы. Далее на все вопросы, входящие в тестовый сборник, записываются ответы, затем они переводятся на формальный язык. В процессе формализации ответов на вопросы выделяются ключевые узлы описываемой предметной области. Таким образом формируется первая версия базы знаний. Затем сборник вопросов расширяется и база знаний развивается эволюционно.

Такой метод проектирования баз знаний достаточно удобен при разработке баз знаний систем справочного назначения. К достоинствам данного подхода можно отнести эволюционное развитие базы знаний, которое позволяет быстро получить первый прототип.

**2. Метод поэтапной формализации, основанный на семантической структуризации исходных документов.** Данный подход к разработке баз знаний основывается на использовании в качестве источника информации для формализации некоторого документа (книги, справочника, учебника, энциклопедии и т.п.). На первом этапе разрабатывается структура базы знаний, далее каждому атомарному разделу в соответствие ставится файл, являющийся результатом трансляции данного раздела на какой-либо внешний язык. Вслед за этим начинается поэтапная формализация содержимого всех указанных файлов.

Процесс проектирования sc-модели базы знаний может быть декомпозирован следующим образом:

- выделение ветвей проектирования, соответствующих содержательной декомпозиции материала исходного документа (разделение на главы, параграфы, части и т.п.) вплоть до атомарных разделов. В данном случае каждый атомарный раздел документа соответствует атомарному разделу проектируемой базы знаний.
- проектирование семантической модели базы знаний заключается в детальной структуризации базы знаний, т.е. рассматривать структуру базы знаний, как иерархическую систему взаимосвязанных друг с другом предметных областей, представляемых в базе знаний. При таком рассмотрении процесса проектирования модели базы знаний можно выделить отдельные направления работ, соответствующие различным видам структуризации базы знаний, таких как построение теоретико-множественной онтологии описываемой предметной области, построение логической онтологии предметной области, построение терминологической онтологии предметной области, и т.д.

**3. Метод компонентного проектирования баз знаний, основанный на модели глобального семантического пространства человеческих знаний.**

Суть данного подхода заключается в использовании библиотеки многократно используемых компонентов баз знаний, пополняемой авторами баз знаний. Основой интеграции для всех компонентов служит унифицированная модель представления знаний с использованием SC-кода. Достоинством данного подхода является максимальное сокращение сроков разработки базы знаний. Фактически процесс разработки базы знаний сводится к выбору инженером необходимых компонентов баз знаний из библиотеки, а затем

их интеграции в единое семантическое пространство. Недостающие фрагменты базы знаний разрабатываются одним из приведенных выше методов.

Согласно данному методу разработка *sc-модели базы знаний* начинается с установки *Ядра sc-моделей баз знаний*, то есть набора базовых многократно используемых компонентов *sc-моделей баз знаний*, необходимых для работы даже первого прототипа *ostis-системы*.

Далее помимо базовых многократно используемых компонентов баз знаний, входящих в состав ядра базы знаний, каждая *дочерняя система* может быть дополнена другими компонентами из *Библиотеки многократно используемых компонентов sc-моделей баз знаний*.

Более конкретно процесс проектирования *sc-модели базы знаний* сводится к рассмотренным ниже этапам на примере разработки базы знаний *Интеллектуальной справочной системы по геометрии*.

## **1. Построение системы предметных областей и их онтологий, входящих в состав базы знаний проектируемой ostis-системы**

Первым и важнейшим этапом проектирования базы знаний является уточнение структуры описываемой предметной области или нескольких взаимосвязанных предметных областей. Уточнение такой структуры – это, прежде всего, уточнение класса исследуемых объектов, уточнение предмета исследования, уточнение всего семейства ключевых узлов семантической сети, представляющей предметную область. В рамках предметной области возможно выделение *частных предметных областей* на основе выделения подмножества из семейства классов исследуемых объектов или изменения набора исследуемых отношений.

При описании структуры предметной области используются ключевые узлы, входящие в состав *Предметной области предметных областей*, входящей в *Библиотеку многократно используемых компонентов баз знаний* в виде компонента, следовательно, данный компонент необходимо добавить в состав базы знаний разрабатываемой системы.

Пример структуры *Предметной области Геометрии Евклида*:

### ***Предметная область Геометрии Евклида***

∈ предметная область

=> частная предметная область\*:

- *Предметная область геометрических точек*
- *Предметная область линий*
- *Предметная область планарных фигур*

=> частная предметная область\*:

- *Предметная область прямолинейных фигур*
- *Предметная область планарных углов*
- *Предметная область многоугольников*

=> частная предметная область\*:

- *Предметная область треугольников*
- *Предметная область четырехугольников*

- *Предметная область вписанных планарных фигур*
- *Предметная область кругов и окружностей*

- *Предметная область геометрических поверхностей*
- *Предметная область геометрических тел*

=> частная предметная область\*:

- *Предметная область многогранников и их поверхностей*
- *Предметная область непланарных углов*
- *Предметная область тел вращения и их поверхностей*
- *Предметная область конгруэнтности геометрических фигур*

Рассмотрение базы знаний с позиции ее соотношения с предметной областью позволяет рассматривать исследуемые объекты на различных уровнях детализации, которые отражаются в различных типах онтологий, описывающих определенное направление описания свойств объекта в рамках рассматриваемой предметной области. К таким типам таких онтологий относятся:

- *структурная спецификация предметной области* – описание связей рассматриваемой предметной области с другими предметными областями и ролей всех понятий, входящих в состав данной предметной области;
- *терминологическая онтология* – описание терминов и их синонимов ключевых понятий рассматриваемой предметной области, близких между собой терминов, этимологии терминов и правила построения идентификаторов экземпляров понятий;
- *теоретико-множественная онтология* – описание теоретико-множественных связей между понятиями рассматриваемой предметной области;
- *логическая онтология* – описание всех высказываний рассматриваемой предметной области;
- *логическая система понятий и их определений* – это структура, являющаяся надстройкой над логической онтологией, включающая описание системы определений понятий заданной предметной области с указанием набора понятий, через которые определяется каждое определяемое понятие рассматриваемой предметной области;
- *логическая система утверждений и их доказательств* – это структура, являющаяся надстройкой над логической онтологией, включающая описание системы утверждений рассматриваемой предметной области с указанием набора утверждений, через которые доказывается каждое утверждение;
- *онтология задач и решений задач* – описание конкретных задач, рассматриваемых в заданной предметной области, и их решений;
- *онтология классов задач и способов решения задач* – описание классов задач, рассматриваемых в заданной предметной области, и способов их решений. Данная структура является надстройкой над онтологией задач и решений задач.

Для описания указанной структуры предметной области используются ключевые понятия, входящие в состав многократно используемого компонента *sc-моделей баз знаний Предметная область онтологий*.

Таким образом, для всех разделов базы знаний, описывающих предметную область и ее спецификацию, имеет смысл задать типовую структуру. В качестве примера рассмотрим *Предметную область четырехугольников* и ее онтологии.

### ***Раздел. Предметная область четырехугольников***

*∈ предметная область и ее онтологии*

*<= декомпозиция раздела\*:*

- {
- *Предметная область четырехугольников*
- *Структурная спецификация предметной области четырехугольников*
- *Терминологическая онтология предметной области четырехугольников*
- *Теоретико-множественная онтология предметной области четырехугольников*
- *Логическая онтология предметной области четырехугольников*
- *Логическая иерархия понятий предметной области четырехугольников*
- *Логическая иерархия высказываний о предметной области четырехугольников*
- *Онтология задач и решений задач предметной области четырехугольников*
- *Онтология классов задач и способов решения задач предметной области четырехугольников*
- }

Далее рассмотрим подробнее фрагменты указанных выше разделов спецификации предметной области.

## 2. Разработка фрагментов используемых предметных областей

Любая предметная область также может содержать примеры конкретных объектов исследования, т.е. экземпляры классов исследования.

Ниже приведен пример отображения в системе семантической окрестности объекта исследования Параллелограмм ABCD.



Рисунок 1 – Семантическая окрестность объекта исследования Параллелограмм ABCD

## 3. Разработка структурных спецификаций для каждой используемой предметной области

Рассмотрим фрагмент Структурной спецификации предметной области четырехугольников:

### **Предметная область четырехугольников**

=> онтология\*:

Структурная спецификация предметной области четырехугольников

=[

#### **Предметная область четырехугольников**

<= частная предметная область\*:

Предметная область многоугольников

≡ максимальный класс объектов исследования':

четырехугольник

≡ не максимальный класс объектов исследования':

- параллелограмм
- прямоугольник
- ромб
- квадрат
- трапеция
- равнобедренная трапеция
- прямоугольная трапеция

≡ элемент структуры':

- Квадрат ABCD
- Четырехугольник KLMN
- Трапеция BMNO

≡ исследуемое отношение':

- внутренний угол\*
- сторона\*
- высота\*
- площадь\*
- периметр\*
- средняя линия\*

]

Наличие указанного фрагмента в базе знаний позволяет задавать следующие вопросы системе:

- Что является объектом исследования данной предметной области?
- Какие отношения исследуются в данной предметной области?

### **4. Разработка терминологических онтологий для каждой используемой предметной области**

Рассмотрим фрагмент Терминологической онтологии предметной области четырехугольников:

#### **Предметная область четырехугольников**

=> онтология\*:

Терминологическая онтология предметной области четырехугольников

=[

#### **четырехугольник**

= многоугольник с четырьмя сторонами

= многоугольник с четырьмя углами

<= правила идентификации элементов\*:

{

- [Четырехугольник <суффикс >]
- [Чет-к <суффикс >]
- [Четырехугольник (<идентификатор точки>; <идентификатор точки>; <идентификатор точки>, <идентификатор точки>)]

}

#### **прямоугольник**

= четырехугольник с прямыми внутренними углами

<= правила идентификации элементов\*:

{

- [Прямоугольник <суффикс >]
- [Прям-к <суффикс >]
- [Прямоугольник (<идентификатор точки>; <идентификатор точки>; <идентификатор точки>, <идентификатор точки>)]

}

#### **ромб**

= четырехугольник с равными сторонами

= четырехугольник с конгруэнтными сторонами

<= правила идентификации элементов\*:

{

- [Ромб <суффикс >]
- [Ромб (<идентификатор точки>; <идентификатор точки>; <идентификатор точки>, <идентификатор точки>)]

}

#### **квадрат**

= ромб с прямыми углами

= прямоугольник с равными сторонами

= прямоугольник  $\cap$  ромб

<= правила идентификации элементов\*:

{

- [Квадрат <суффикс >]
- [Квадрат (<идентификатор точки>; <идентификатор точки>; <идентификатор точки>, <идентификатор точки>)]

}

#### **трапеция**

<= правила идентификации элементов\*:

{

- [Трапеция <суффикс >]
- [Трапеция (<идентификатор точки>; <идентификатор точки>; <идентификатор точки>, <идентификатор точки>)]

}

#### **равнобедренная трапеция**

= трапеция с равными сторонами

#### **прямоугольная трапеция**

= трапеция с прямыми углами при боковой стороне

]

Наличие указанного фрагмента в базе знаний позволяет задавать следующие вопросы *ostis-системе*:

- Какие синонимичные термины известны для заданного термина?
- По какому правилу идентифицируются элементы заданного множества?

## 5. Разработка теоретико-множественных онтологий для каждой используемой предметной области

Рассмотрим Фрагмент Теоретико-множественной онтологии предметной области четырехугольников:

### Предметная область четырехугольников

=> онтология\*:

Теоретико-множественная онтология предметной области четырехугольников

=[

**четыреугольник**

⊃ трапеция

**трапеция**

⊃ параллелограмм

⊃ равнобедренная трапеция

⊃ прямоугольная трапеция

**параллелограмм**

⊃ прямоугольник

⊃ ромб

**квадрат**

⊂ прямоугольник

⊂ ромб

**средняя линия\***

∈ бинарное отношение

=> область определения\*:

трапеция ∪ отрезок

=> первый домен\*:

трапеция

=> второй домен\*:

отрезок

]

Наличие указанного фрагмента в базе знаний позволяет задавать следующие вопросы *ostis-системе*:

- Как классифицируется заданное понятие?
- Какие надклассы известны для заданного понятия?
- Являются ли два заданных понятия пересекающимися?
- Какова область определения заданного отношения?

## 6. Разработка логических онтологий для каждой используемой предметной области

Рассмотрим фрагмент Логической онтологии предметной области четырехугольников

### Предметная область четырехугольников

=> онтология\*:

Логическая онтология предметной области четырехугольников

=[

**четыреугольник**

∈ ключевой *sc*-элемент\*:

Опр. (четыреугольник)

∈ определение

<= трансляция *sc*-текста\*:

...

∃ пример':  
 [**четырёхугольник** – это многоугольник с четырьмя сторонами.]  
 ∈ Русский язык  
 ...  
 ∈ утверждение  
 <= трансляция sc-текста\*:  
 ...  
 ∃ пример':  
 [Сумма мер внутренних углов четырёхугольника равна 360 угловых градусов.]  
 ∈ Русский язык  
 ...  
 ∈ утверждение  
 <= трансляция sc-текста\*:  
 ...  
 ∃ пример':  
 [Сумма квадратов длин диагоналей четырёхугольника равна удвоенной сумме  
 длин квадратов отрезков, соединяющих середины его противоположных  
 сторон.]  
 ∈ Русский язык  
 ...  
 ∈ утверждение  
 <= трансляция sc-текста\*:  
 ...  
 ∃ пример':  
 [Если около четырёхугольника можно описать окружность, то сумма мер его  
 любых противоположных углов равна 180 угловых градусов.]  
 ∈ Русский язык  
 ]

Наличие указанного фрагмента в базе знаний позволяет задавать следующие вопросы *ostis-системе*:

- Какие высказывания известны в рамках заданной предметной области?
- Какие высказывания являются аксиомами в рамках заданной предметной области?
- Как определяется/поясняется то или иное понятие?

## 7. Разработка логической системы понятий и их определений для каждой используемой предметной области

Рассмотрим фрагмент *Логической системы понятий и их определений предметной области четырёхугольников*

### *Логическая онтология предметной области четырёхугольников*

=> метазнание\*:

*Логическая система понятий и их определений предметной области четырёхугольников*

=

**четырёхугольник**

∈ ключевой sc-элемент':

...

∈ определение

<= трансляция sc-текста\*:

...

∃ пример':



[**четырёхугольник** – это многоугольник с четырьмя сторонами.]  
 ∈ Русский язык  
 <= используемые константы\*:  
 {  
 • многоугольник  
 • сторона\*  
 }  
  
**трапеция**  
 ∈ ключевой sc-элемент':  
 ...  
 ∈ определение  
 <= трансляция sc-текста\*:  
 ...  
 ∃ пример':  
 [**трапеция** – это четырёхугольник, у которого две противоположные стороны параллельны.]  
 ∈ Русский язык  
 <= используемые константы\*:  
 {  
 • четырёхугольник  
 • сторона\*  
 • параллельность\*  
 • противоположный\*  
 }  
 ]

Наличие указанного фрагмента в базе знаний позволяет задавать следующие вопросы *ostis-системе*:

- Через какие понятия определяется заданное понятие?
- Какие понятия определяются на основе заданного понятия?

## 8. Разработка логической системы утверждений для каждой используемой предметной области

В разрабатываемой системе по геометрии уместно привести типовые доказательства для некоторых теорем и типовые решения для некоторых примеров задач.

Для обеспечения такой возможности необходимо добавить в систему описание *Предметной области действий и задач*, содержащей все необходимые ключевые узлы, то есть *Раздел. Предметная область действий и задач*.

Рассмотрим фрагмент *Логической системы утверждений и их доказательств предметной области четырёхугольников*

### *Логическая онтология предметной области четырёхугольников*

=> метазнание\*:

*Логическая система высказываний и их доказательств предметной области четырёхугольников*

=[

#### **параллелограмм**

∈ ключевой sc-элемент':

Утв.(параллелограмм; диагональ\*; четырёхугольник)

∈ утверждение

<= трансляция sc-текста\*:

...

Э пример':

[Если диагонали четырехугольника пересекаются и точкой пересечения делятся пополам, то этот четырехугольник - параллелограмм]

∈ Русский язык

Э главный ключевой sc-элемент':

...

=> основное доказательство\*:

Док-во. Утв. (параллелограмм; диагональ\*; четырехугольник)

<= используемые утверждения\*:

{

• Утв. (вертикальные углы\*; конгруэнтность\*)

• Утв. (треугольник; конгруэнтность\*; сторона\*; внутренний угол\*)

• Опр. (внутренние накрест лежащие углы\*)

• Опр. (параллелограмм)

}

]

Теперь, используя упомянутые ключевые узлы, добавим в систему пример доказательства теоремы о равенстве вертикальных углов:

```
...
= основное доказательство*:
  stat_quadilateral_parallelogram_diagonal_point_intersect_middle
= декомпозиция sc-действия*:
  ...
  Э rrel_1:
    ...
    Э ...
    Э ...
    Э ...
    Э ...
    Э ...
    Э ...
    Э ...
    Э ...
  € successfully_completed_action
```

Рисунок 2 – Пример отображения sc-текста общей структуры доказательства, представленного в SCn-коде

```
...
= результат*:
  ...
  = трансляция sc-текста*:
    ...
    Э 

Прямая AD параллельна прямой BC. ...


    € Русский язык
  = последовательность действий*:
    ...
  = последовательность действий*:
    ...
    Э rrel_2:
      stat_parallelism_congruence_straight_line_internal_cross_lying_angles
    Э rrel_1:
      ...
    € ...
    ⇒ декомпозиция sc-действия*:
      ...
    € successfully_completed_action
    € sc-действие применения логического утверждения
```

Рисунок 3 – Пример отображения sc-текста шага доказательства

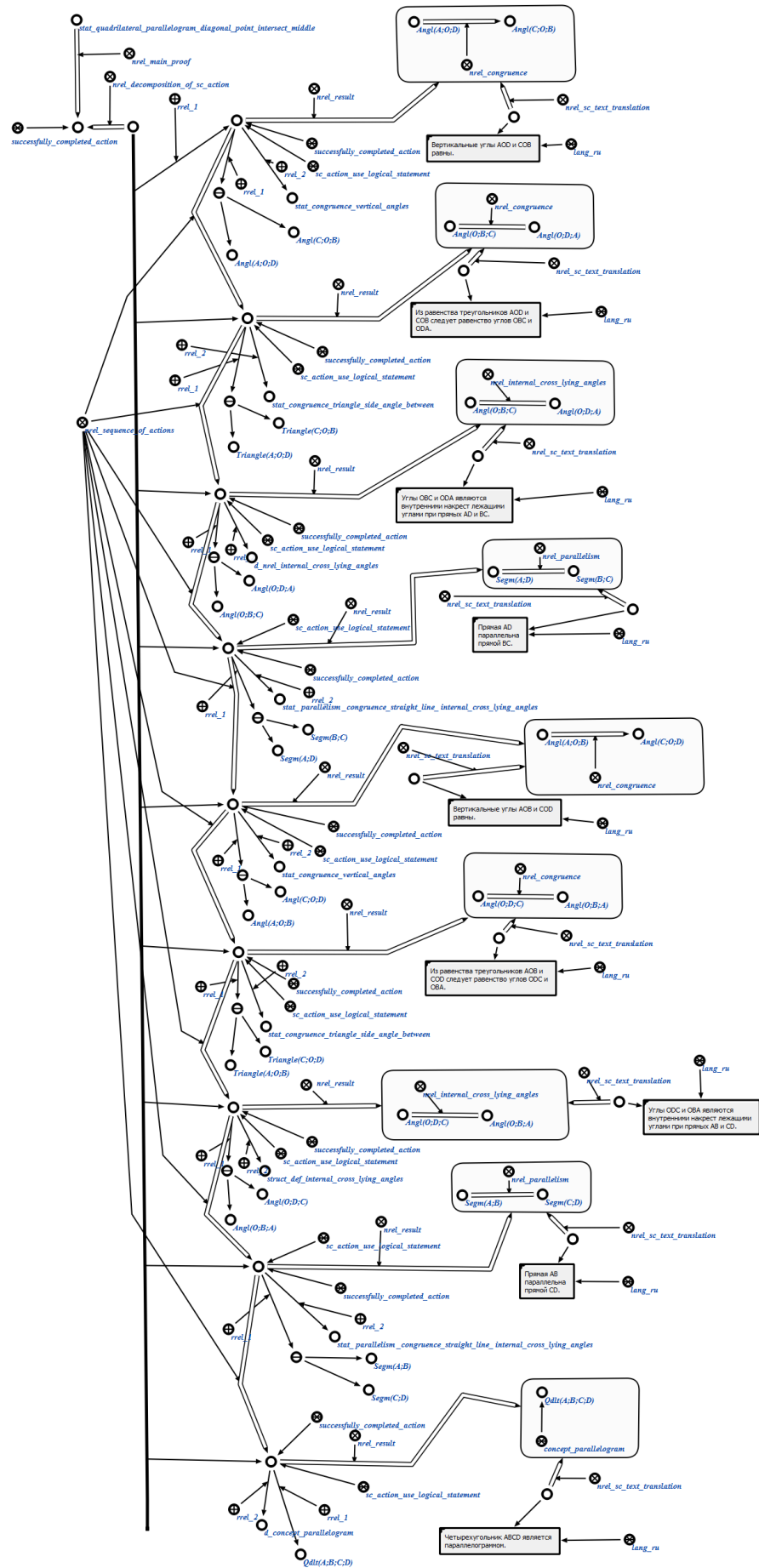


Рисунок 4 – Пример sc-текста доказательства, представленного в SCg-коде

Внесенное в систему доказательство можно просмотреть при помощи базовых средств навигации, как показано на рисунках выше, однако это не всегда удобно, поэтому имеет смысл использовать специальный *sc-агент поиска доказательства заданного утверждения*, о чем будет сказано ниже.

## 9. Разработка онтологий задач и решений для каждой используемой предметной области

Аналогичным образом можно внести в систему пример решения, например, следующей задачи: «В параллелограмме ABCD  $\sin(C) = 5/7$ .  $AD = 7$ . Найдите высоту, опущенную на сторону АВ».

Ниже приведен фрагмент *Онтологии задач и решений задач предметной области четырехугольников*.

### *Предметная область четырехугольников*

=> онтология\*:

*Онтология задач и решений задач предметной области четырехугольников*

= [

***параллелограмм***

∈ ключевой *sc-элемент*’:

*Задача. Нахождение высоты параллелограмма по длине стороны и синусу внутреннего угла*

∈ задача

<= трансляция *sc-текста*\*:

...

∃ пример’:

[В параллелограмме ABCD  $\sin(C) = 5/7$ .  $AD = 7$ . Найдите высоту, опущенную на сторону АВ]

∈ *Русский язык*

=> решение\*:

*Решение. Задача. Нахождение высоты параллелограмма по длине стороны и синусу внутреннего угла*

<= используемые утверждения\*:

{

• Утв.(параллелограмм; *противолежащий\**; *внутренний угол\**; *конгруэнтность\**)

• Опр.(синус\*)

}

]

Наличие указанного фрагмента в базе знаний позволяет задавать следующие вопросы *ostis-системе*:

- Примеры каких задач известны в рамках заданной предметной области?
- Как решается заданная задача?
- С помощью каких утверждений решается заданная задача?

Формальная запись условия этой задачи на языке SCg представлена на рисунке 8.

Предлагаемый вариант решения данной задачи состоит из следующих шагов:

- $\sin(A) = \sin(C) = \frac{5}{7}$ ;
- $\sin(A) = \frac{5}{7} = \frac{DH}{AD} = \frac{DH}{7}$ .  $DH = 5$ .

Решение по шагам в формальном представлено на рисунке 9.

В системе решение можно просмотреть по шагам при помощи базовых *sc-агентов информационного поиска*, однако, как и в случае с доказательствами, это не всегда удобно, поэтому имеет смысл добавить в систему *sc-агент поиска решения задачи*, позволяющий просмотреть все решение сразу.

## **10. Разработка онтологий классов задач и способов решения задач для каждой используемой предметной области**

Рассмотрим фрагмент *Онтологии задач и решений задач предметной области четырехугольников*.

### ***Онтология задач и решений задач предметной области четырехугольников***

=> *метазнание\**:

*Онтология классов задач и способов решения задач предметной области четырехугольников*

=

*Задача. Нахождение высоты параллелограмма по длине стороны и синусу внутреннего угла*

*∈ задача поиска заданной величины*

*∈ задача на параллелограммы*

*∈ задача на четырехугольники*

*∈ задача планиметрии*

]

Наличие указанного фрагмента в базе знаний позволяет задавать следующие вопросы *ostis-системе*:

- Какие классы задач известны для заданной предметной области?
- Какие способы решения класса задач известны для заданной предметной области?

Таким образом, база знаний *дочерней sc-системы* может быть пополнена как за счет использования компонентов из *Библиотеки многократно используемых компонентов sc-моделей баз знаний*, так и за счет добавления разработчиком в систему предметных знаний из рассматриваемой области.

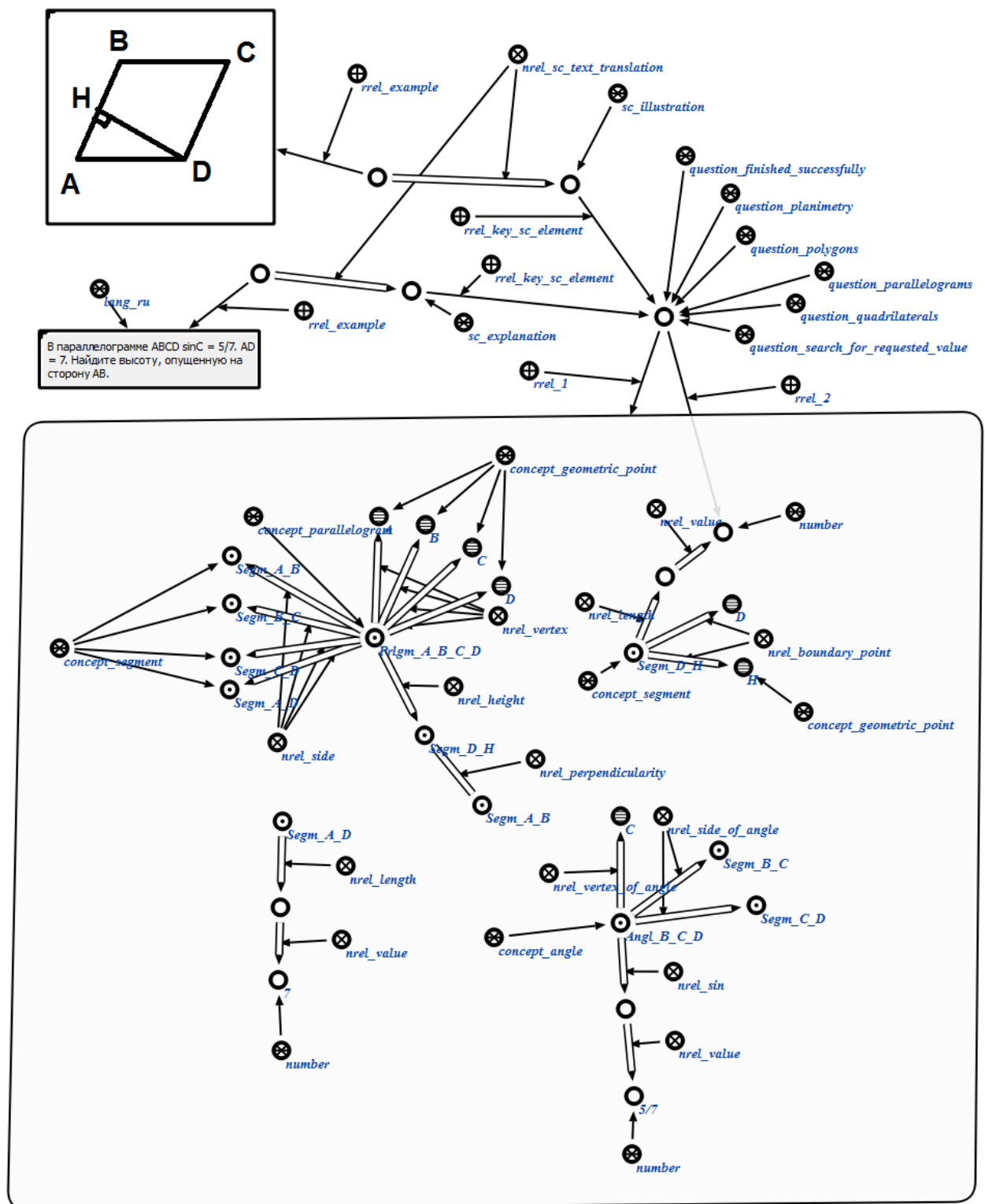


Рисунок 5 – Условие задачи, представленное в SCg-коде

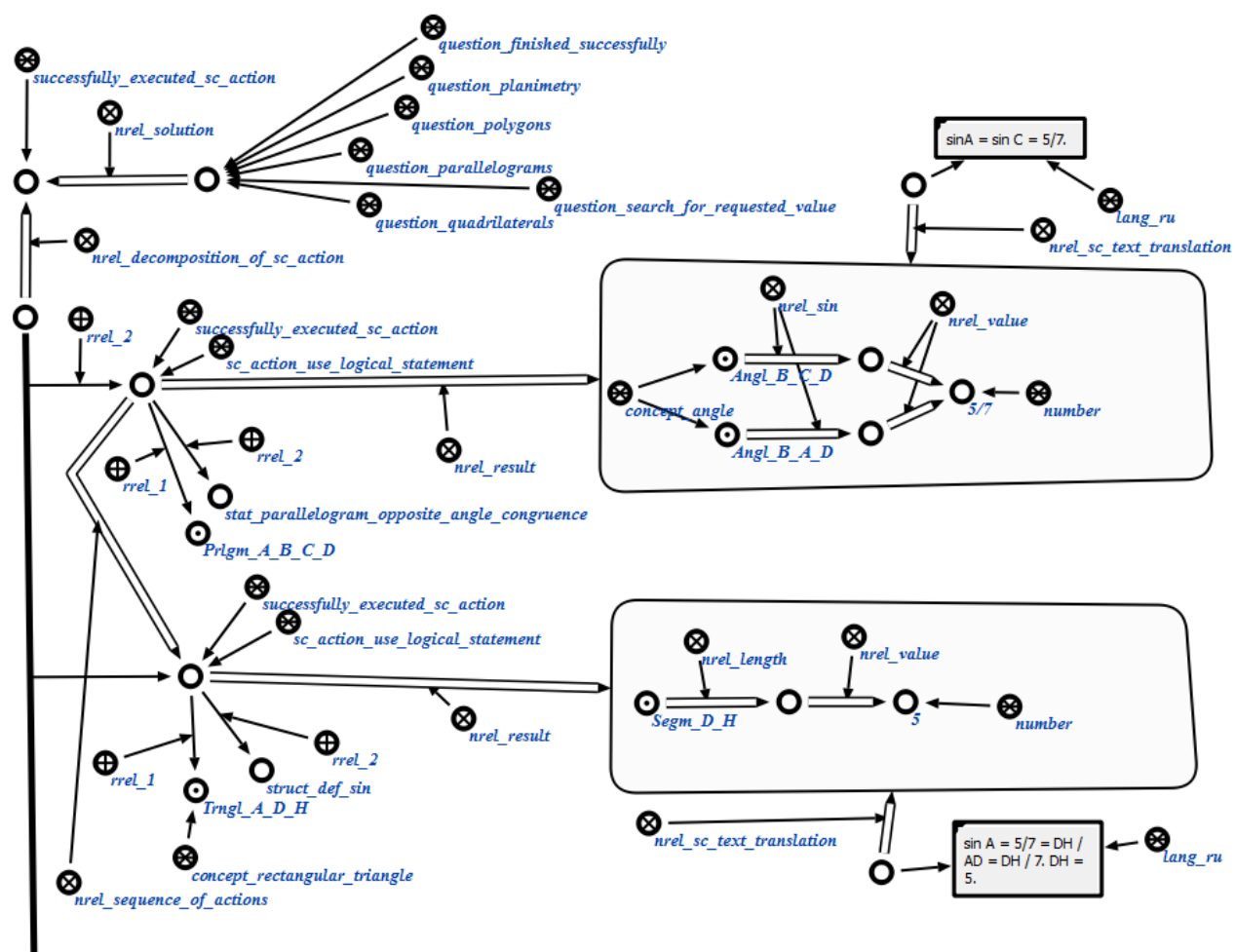


Рисунок 6 – Решение задачи, представленное в SCg-коде