О СХЕМЕ ВЗАИМОДЕЙСТВИЯ В КОМПЛЕКСЕ «АНАЛИЗ И СИНТЕЗ ЕСТЕСТВЕННОГО ЯЗЫКА И ИЗОБРАЖЕНИЙ»

С.С. Курбатов, *cul@nicevt.ru*

К.А. Найденова, *Naidenovaxen@gmail.com*

Г.К. Хахалин, *gkhakhalin@yandex.ru*

Комплекс: анализ/синтез ЕЯ и изображения + онтология

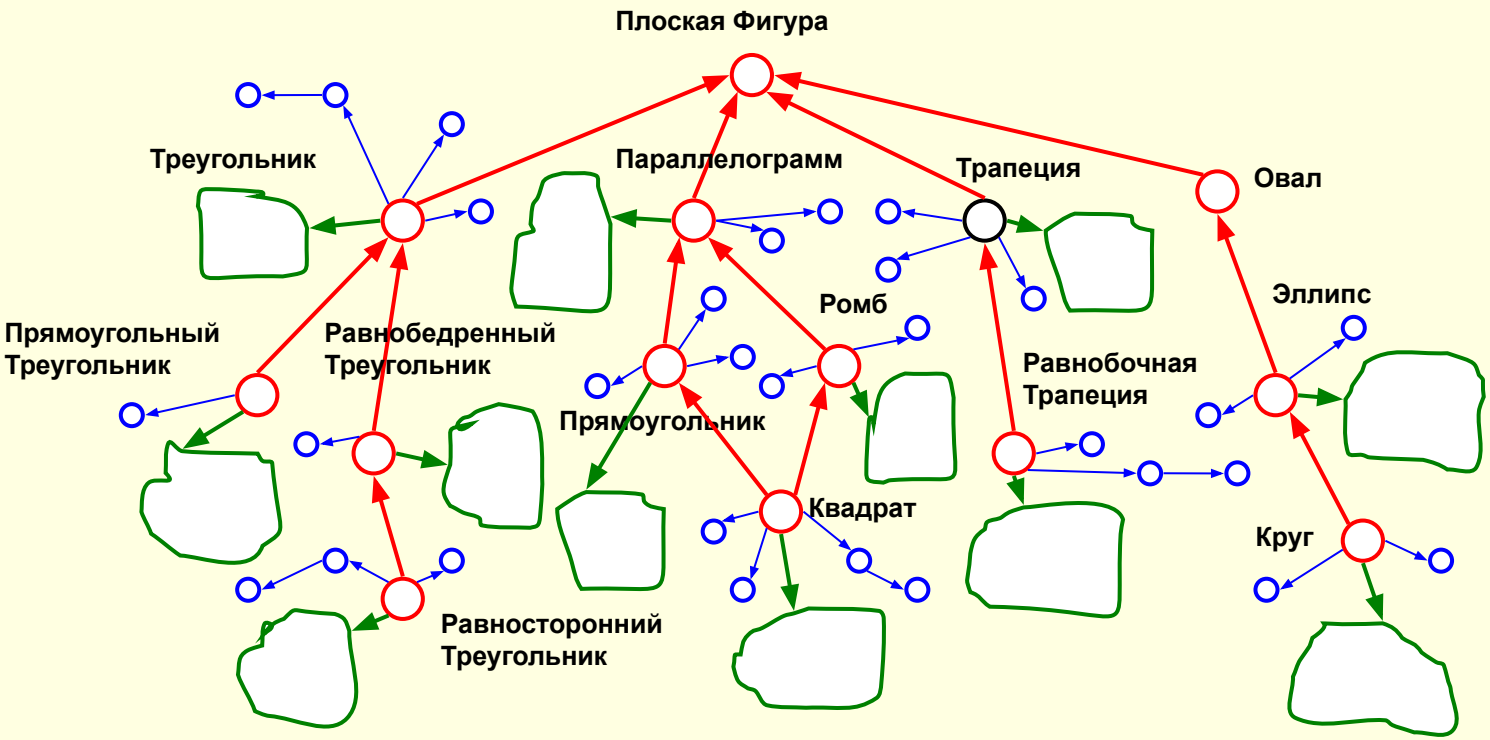
- Система анализа изображений.
- Система синтеза изображений.
- Система анализа ЕЯ-текста.
- Система синтеза ЕЯ-текста.
- Прикладная онтология - связующее звено между системами:
  - Прикладная область: плоские контурные планиметрические объекты.
  - Язык представления знаний в онтологии – семантический гиперграф.

Схема комплекса

Семантический гиперграф

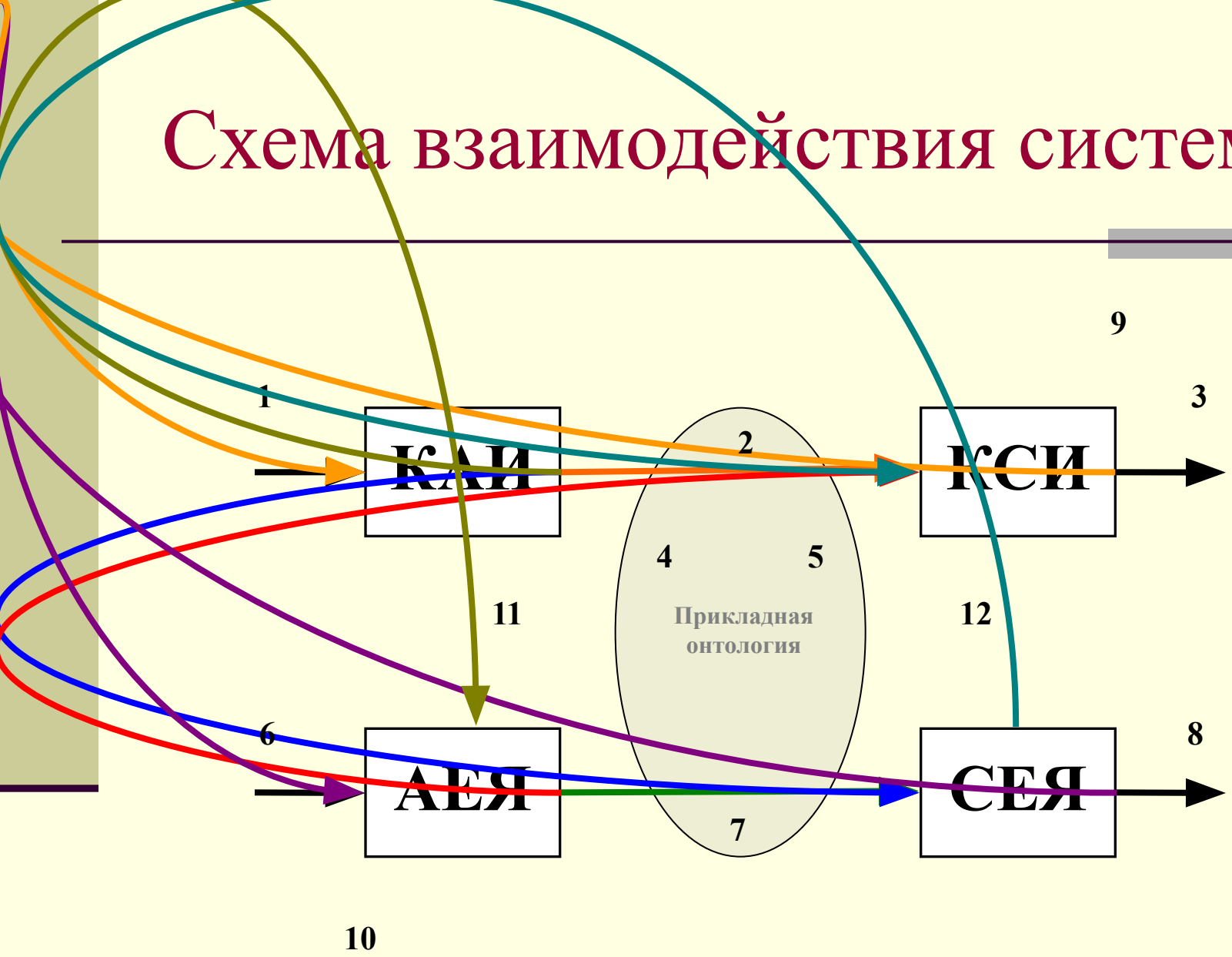
- Является расширением семантических сетей, где естественным образом представляются *n*-арные отношения.
- Предоставляет возможность задавать атрибутивную и структурную информацию об объектах и ситуациях.
- В зависимости от типов отношений можно реализовывать классифицирующие, функциональные, ситуационные, структурные сети и сценарии.

Фрагмент прикладной онтологии на языке семантического гиперграфа



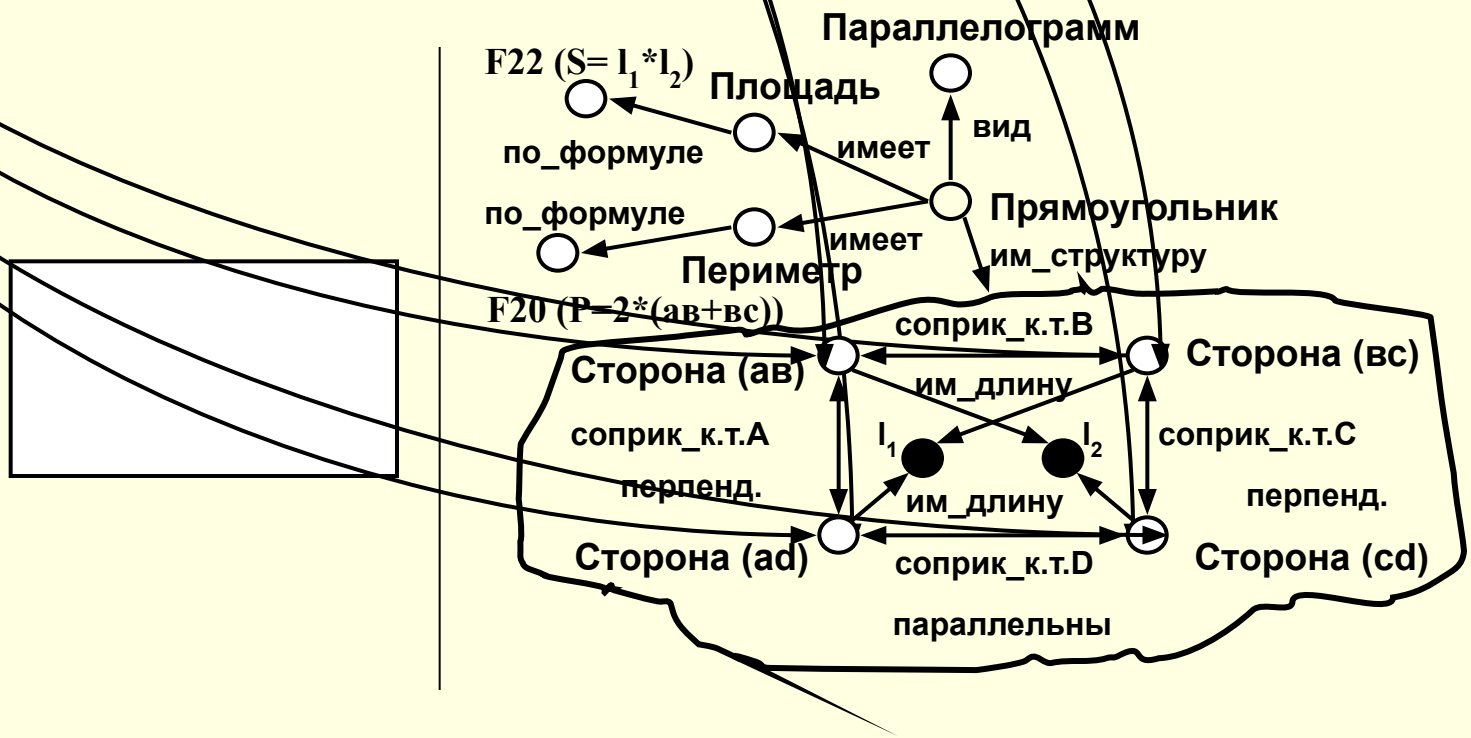
Равнобочная трапеция

Схема взаимодействия систем



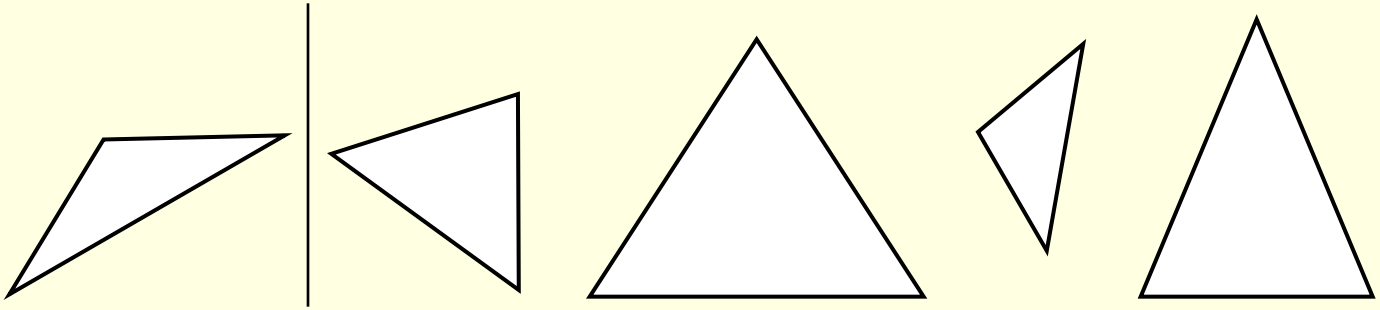
Анализ/понимание изображения

- **{1, 2} – вход: изображение; в онтологии: описание зрительной ситуации (всей или части) на ЯПЗ**



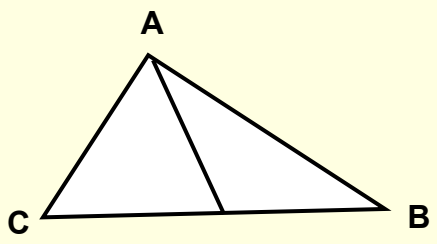
Анализ изображения и проверка результатов анализа по синтезу объектов

- {1, 2, 3} – вход: изображение; в онтологии: описание зрительной ситуации на ЯПЗ; выход: множество примеров изображений класса объектов.



Генерирование текста по рисунку

- {1, 4, 8} – вход: изображение; в онтологии: описание геометрической ситуации на ЯПЗ; выход: ЕЯ-тексты, описывающие геометрическую ситуацию.



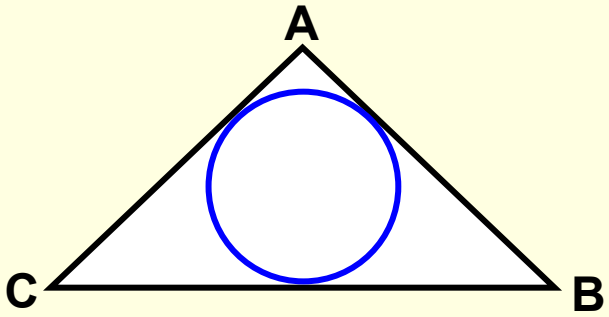
ЕЯ-тексты:

- *Треугольник ABC, две стороны которого равны 6 и 8, а медиана, заключенная между ними, равна 5.*
- *Треугольник со сторонами AB=8, AC=6. Медиана, проведенная из вершины A, равна 5.*
- *Задан треугольник ABC. Сторона AC (равна 6) меньше стороны AB (равна 8). Из вершины A к основанию CB опущена медиана AD (равна 5).*

Создание рисунка по тексту

- {6, 5, 3} – вход: ЕЯ-текст; в онтологии: семантическое описание текстовой ситуации на ЯПЗ; выход: изображение (множество изображений), описывающее текстовую ситуацию.

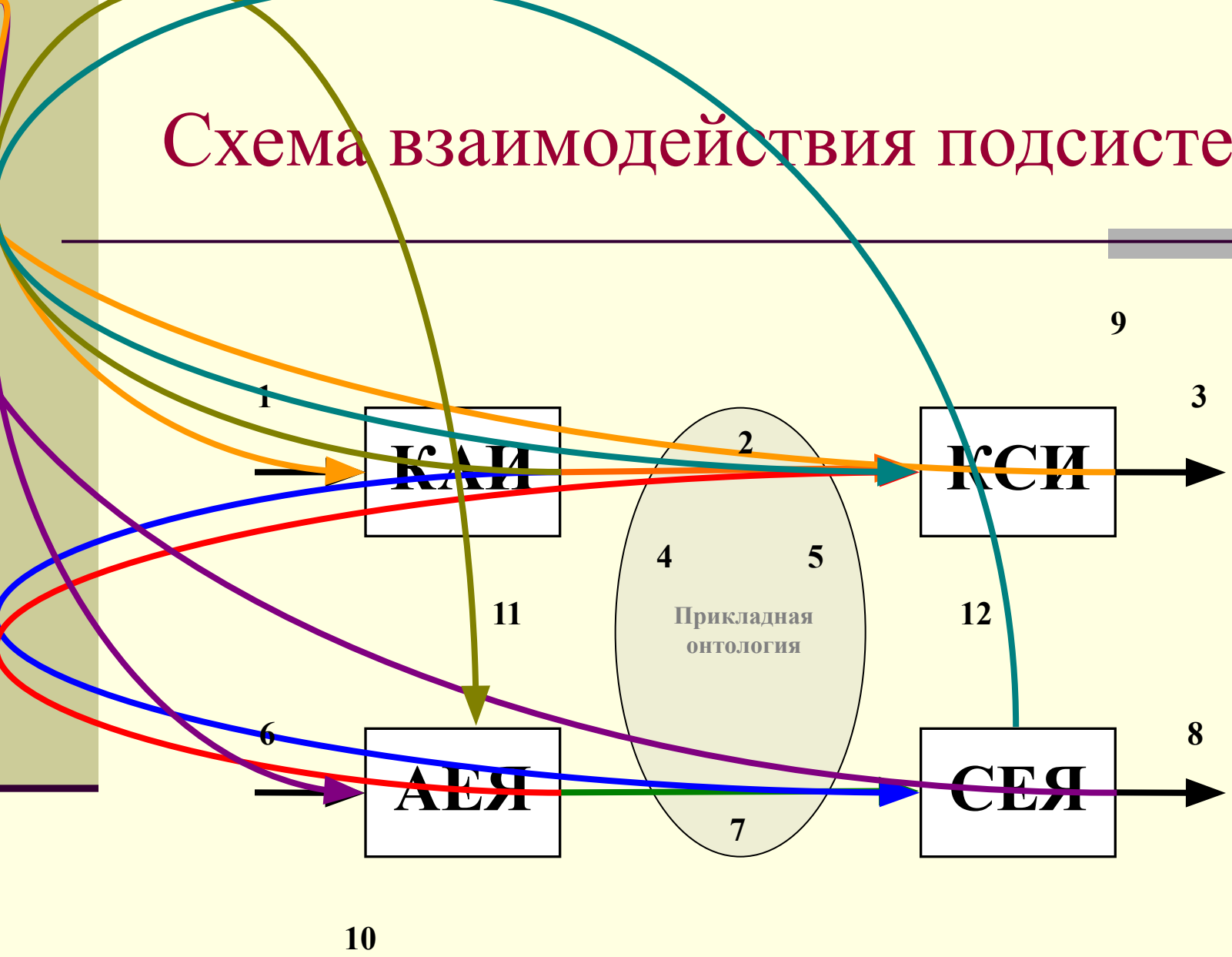
<u>ЕЯ-текст</u>: *В треугольник АВС вписана окружность*.



Перефразирование входного текста (анализ/синтез текста)

- **Входной текст**: *Постройте круг данного радиуса r, который включает точку с координатами X1 и Y1.*
- **Выходные тексты**:
  - *Задана точка с координатами X1 и Y1. Она находится внутри круга с радиусом r. Построить такой круг.*
  - *Построить круг с радиусом r, включающий точку с координатами X1 и Y1.*
  - *Внутри круга находится точка с координатами X1 и Y1. Постройте такой круг, радиус которого равен r.*
- Выходной текст может быть представлен на языке, отличным от входного языка:
  - *To construct the circle with radius r which includes a points with co-ordinates X1 and Y1.*

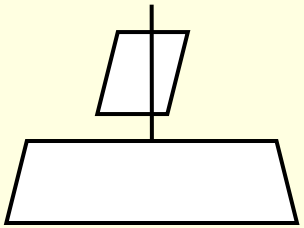
Схема взаимодействия подсистем



Пример наполнения онтологии

- *Кораблик – это равнобочная трапеция, к большему основанию которой перпендикулярно проведен отрезок прямой длиной не более 2/3 этого основания. Этот отрезок прямой касается большего основания трапеции в его середине. Отрезок прямой пересекает параллелограмм. Стороны параллелограмма не более 2/3 и не менее 1/4 длины отрезка.*

Структура понятия «Кораблик» и пример изображения



Претенденты на системы комплекса

- Система Spotlight (фирма CSOFT) – «нижние» уровни анализа объектов изображения.
  - Основные функции: сканирование, фильтрация, коррекция, распознавание геометрических примитивов:
    - отрезки, дуги, окружности, символы, произвольные кривые (полилинии)
- Система «ПСИ» (Власов А.В., Аредова И.И.) – синтез объектов изображения.
  - Основные функции: означивание параметров объектов изображения по структурам модели; расчет экземпляров объектов; выявление недостатка/противоречивости данных и т.д.

Претенденты на системы комплекса

- Лингвистический транслятор «Адалит» (Хахалин Г.К. и др.) – анализ/синтез ЕЯ-текста.
  - Основные функции: морфологический, синтаксический и семантический анализ элементов ЕЯ-текста, включающего простые, осложненные, сложные, эллиптические и анафорические предложения. Элементы синтеза ЕЯ-текста.
- Язык представления знаний для прикладной онтологии (Курбатов С.С., Хахалин Г.К.) – *семантический гиперграф*.
  - Основные функции: стандартные процедуры обработки гиперграфов (добавление/удаление вершин, ребер; стягивание, свертка и т.д.).

Возможное применение:

- Системы обучения (естественному языку с картинками, геометрии и т.д.).
- Робототехника.
- Система сурдоперевода («текст → аватар»).
- Система описания изображений на ЕЯ для поиска.

Спасибо за внимание!

- Дополнительную информацию можно получить, обратившись к докладчикам или по e-mail: *gkhakhalin@yandex.ru*

Студенты и аспиранты!

- Если Вы хотите работать над дипломом или над диссертацией по темам «Лингвистический анализ/синтез естественного языка» или «Концептуальный анализ/синтез изображений», обращайтесь к докладчику или по e-mail: *gkhakhalin@yandex.ru*
- Престиж и качество гарантируются!