

А.П. Гуменюк
Криворожский государственный
педагогический университет
IV курс, группа И-05
Научный руководитель:
к.пед.н., доц. С.А. Семериков

РАЗРАБОТКА ИНТЕРНЕТ-СОБЕСЕДНИКА

Виртуальный собеседник (англ. Chatterbot) – это компьютерная программа, которая создана для имитации речевого поведения человека при общении с одним или несколькими пользователями. По отношению к виртуальным собеседникам употребляется также название *программа-собеседник*.

Предполагается, что идеальная программа-собеседник должна пройти тест Тьюринга – тест, предложенный Аланом Тьюрингом в 1950 г. в статье «Вычислительные машины и разум» (Computing machinery and intelligence) для проверки, является ли компьютер разумным в человеческом смысле слова.

Тест был инспирирован салонной игрой, в ходе которой гости пытались угадать пол человека, находящегося в другой комнате, путем написания вопросов и чтения ответов. В оригинальной формулировке Тьюринга человек должен был притворяться человеком противоположного пола, а тест длился 5 минут. Сейчас эти правила не считаются необходимыми и не входят в спецификацию теста. Пока что ни одна программа и близко не подошла к прохождению теста.

Дело в том, что цели конкретных диалогов между людьми различаются. Можно просто «поболтать», а можно обсудить важную проблему. Реализация последнего типа диалога представляет дополни-

тельную проблему: научить программу мыслить. Поэтому функциональность большинства современных программ ограничивается возможностью ведения незатейливой беседы.

Создание виртуальных собеседников граничит с проблемой общего искусственного интеллекта, то есть единой системы (программы, машины), моделирующей интеллектуальную деятельность человека.

Виртуальные собеседники работают с «живым» языком. Обработка естественного языка, особенно разговорного стиля, – острая проблема искусственного интеллекта.

AIML (Artificial Intelligence Markup Language) – язык разметки для искусственного интеллекта. AIML – подмножество (диалект) XML, он предназначается для создания виртуальных собеседников.

Категории (category) в AIML – основная структурная единица. Категория состоит по крайней мере из двух элементов: образца (pattern) и шаблона (template). Простой пример:

```
<category>
  <pattern>Какое ваше имя</pattern>
  <template>Моё имя - Олег</template>
</category>
```

Если эта категория содержится в базе знаний, интерпретатор на вопрос «Какое ваше имя?» будет отвечать «Моё имя – Олег».

Образец – символьная строка, соответствующая одному или нескольким пользовательским запросам. Такие образцы, как

```
КАКОЕ ВАШЕ ИМЯ
```

будут соотноситься только с одним вопросом (причём без учёта регистра). Однако образцы могут содержать символы обобщения, заменяющие одно или несколько слов. Образец:

```
КАКОЕ ВАШЕ *
```

будет соответствовать бесчисленному множеству вопросов, например: «Какое ваше имя», «Какое ваше отчество», «Какое ваше мнение об этом» и т.д.

Подобный синтаксис намного проще, чем регулярные выражения. Он соответствует потребностям программ-собеседников, а если все же его не хватает, сами интерпретаторы могут предоставить дополнительные функции по обработке запросов.

Шаблоны обозначают ответы на распознанные образцы. Шаблон может быть и простым:

Моё имя - Олег.

а может и содержать переменные:

Моё имя - <bot name="имя"/>.

Вы сказали, что Вам <get name="возраст пользователя"/> лет.

В первом случае интерпретатор подставит своё «имя», а во втором – возраст пользователя (если возраст известен).

Шаблоны могут содержать базовое текстовое форматирование, условные и случайные части. Возможны также перенаправления (задаются элементом `srai`). Это может использоваться для обозначения синонимии и эквивалентности фраз:

```
<category>
  <pattern>КАКОЕ ВАШЕ ИМЯ</pattern>
  <template>Моё      имя      -      <bot
name="имя"/>.</template>
</category>
<category>
  <pattern>КАК ВАС ЗОВУТ</pattern>
  <template>
    <srai>какое ваше имя</srai>
  </template>
</category>
```

Шаблоны могут содержать форматирование, нацеленное на специфический интерфейс, и обрабатываемое конкретным интерпретатором.