**Язык AIML**

Для работы AIML-бота необходимо 2 компонента: файл AIML и программа (интерпретатор), которая обрабатывает входящие реплики собеседника и формирует ответную реплику в соответствии с файлом AIML. Файл AIML - содержит набор категорий (category). Категория открывается тегом <category> и закрывается тегом </category>. Категория включает паттерны (pattern) и темплейты (template). Паттерн открывается тегом <pattern> и закрывается тегом </pattern>. Темплейт открывается тегом <template> и закрывается тегом </template>. Категория может включать только один паттерн и один темплейт. Паттерн и темплейт должны находиться внутри категории.

Интерпретатор AIML работает очень просто. Он берет входящую фразу собеседника и сравнивает ее со всеми имеющимися паттернами в файле AIML. В случае совпадения входящей фразы с паттерном в какой-либо категории интерпретатор в качестве ответной реплики подставляет темплейт из этой категории.

Паттерн пишется большими буквами. Темплейт может включать одну реплику Или несколько реплик. В этом случае необходимо использовать дополнительный тег <random>. В этом случае в качестве ответной реплики будет произвольно выбрана одна из реплик, находящаяся между тегами <li></li>.

В темплейте также возможно создать ссылку на другой паттерн. Делается это с помощью тега <srai>.

Так как невозможно внести информацию обо всём в бота, то можно сделать вид, что он знает тему разговора Создать такую ответную реплику нам удается с помощью тега <star/>.

Разговорная речь, как правило, содержит некоторое количество т.н. слов-паразитов, которые человеческий мозг воспринимает без проблем, в то время как для мозга бота, написанного на AIML, такие слова могут стать серьезной проблемой.

Для бота “Сколько тебе лет?” и “И сколько тебе лет?” - разные вопросы, поэтому слова паразиты в начале обрезаем.

Для того, чтобы бот не только разговаривал, но и выражал эмоции в ответ на реплики собеседника стандарт AIML дополнен тегом <action>

**Оcновы языка AIML**

AIML (Artificial Intelligence Markup Language) — язык разметки для

искусственного интеллекта. AIML — подмножество (диалект) XML, он

предназначается для создания виртуальных собеседников.

Это один из самых первых и самых простых языков которые используется для создания ботов.

Первым виртуальным собеседником была программа Элиза, разработанная в 1966 году. Её создатель, Джозеф Вейценбаум, построил её базу знаний по принципу шаблонных ответов. Подобный метод был использован в программе A.L.I.C.E. в 1995 году, но теперь база знаний стала более структурированной. Для построения такой базы как раз и был создан AIML. Этот язык разрабатывался Ричардом Уолессом — инициатором проекта ALICE — и всемирным сообществом свободного ПО с 1995 по 2002 год.

Версия AIML, используемая в ALICE, распространяется под лицензией GNU GPL, поэтому существует много свободных реализаций этого языка, а также программ, использующих этот язык (интерпретаторов). Наиболее активно используемая версия интерпретатора — Программа D, написанная на языке Java. Существуют также формальная спецификация языка и XML-схема, данная консорциумом W3C.

<sr/> это сокращенная запись <srai><star/></srai>, то есть ссылка на

звездочку.

Ключевой символ «\_» Этот символ отбрасывает все слова, которые были написаны до искомого слова и обрабатывает только слово, которое должно быть после данного символа. В случае если искомое слово было написано слитно с «\_», то искомое слово как отдельное не найдётся.

<Set> тег используется для установки значения в переменной

Тег <Get> используется для получения значения из переменной.

Тег <that> используется в AIML для реагирования по контексту.

Тег <topic> используется в AIML для хранения контекста, чтобы позже разговор мог быть продолжен на основе этого контекста. Как правило, тег <topic> используется в Да / Нет типе разговора.

Тег < think > позволяет задавать значения переменным, сохраняя их в памяти.

Тег < condition > похож на переключение высказывания в языке программирования. Это помогает реагировать на соответствующие входные данные.

Этот тег (<learnf>) позволяет в автоматическом режиме создавать шаблон. То есть он создаёт отдельный файл с расширением .aiml, куда записывает новый шаблон и ответ на него.

Чтобы сделать полноценного бота на русском языке - нужно обязательно

выделять "семантическое ядро" фразы.

В AIML средства, позволяющие обрабатывать такие конструкции, очень

неудобны и делать бота на русском языке трудоемко и сложно.

Кроме того, язык AIML неочевиден, перегружен синтаксисом тегов XML и является источником трудностей для лингвистов и генератором трудно уловимых ошибок.

Поэтому, для написания русскоязычных ботов используется диалект с

трансляцией в AIML или другие языки диалогов без наследия XML и AIML.