

От чат-ботов и цифровых сотрудников

Колледж Семанта

От чат-ботов и цифровых сотрудников

Нидерланды Тел: +31-162-313323

Канада Тел: +1-514-331-0572



От чат-ботов и цифровых сотрудников



- Versiebeheer: 23.01.2021
- Auteur: Edward A .kool
- Datum: 9 maart 2021
- ©Lingvistica bv 2001-2021



SEMANTA

- Verwerkingspatroon
 - Uitgangspunt
 - Corpus
 - Internet of Things
 - Spreken met een virtuele gesprekspartner



EDDY

- Language teacher
- Internet helper
- Hoe het werkt



SEMANTUS

Контроль версий: 25. 08.2021

Контроль версий: 25. 08.2021	2
Знакомый.....	4
Абстрактный	4
Знакомство	5
Что такое чат-бот или собеседник?	7
Чат-бот.....	7
Языковойбот	7
Digital gesprekspartner	7
Human gesprekspartner.....	7
Xbot.....	8
Сценарист.....	8
Искусственный интеллект	9
Опыт работы с Xbot	9
Заключение	10
Семанта учитель	10
Шаблон обработки	10
Отправная точка	11
Корпус	11
Интернет вещей	11
Общение с виртуальным собеседником.....	12
Преподаватель языка	12
Интернет-помощник	12
Как это работает?	13

Знакомый

В 1970 году я был начинающим программистом. Cobol, Fortran, Assembler и RPG стали куском пирога. Позже дополнен Pascal, PL/1, Basic, C++, Java. Летним днем 1970 года, потеряв раздумья, я получил представление о будущем языков программирования и компьютеров в целом. Моя мысль остановилась на компьютерной программе, которая могла бы общаться более или менее по-человечески на голландском языке.

Мысль осталась со мной и да в 1993 году я познакомился с первым экземпляром ПАРС машинного перевода програама из Харькова в Украине. В 1994 году я продемонстрировал это на совещании surfnet об искусственном интеллекте.

В 1995 году меня познакомили с Интернетом, и я знал, чего хочу: переводить с помощью компьютера через Интернет. Это заняло до 2001 года, прежде чем появились голландская и английская версии. Позже расширился до немецкого и польского. До 2006 года это был прибыльный источник дохода.

В 2006 году мне исполнилось 58 лет, и это стало моим пенсионным возрастом, с тех пор сочетание языка и компьютеров стало моей повседневной деятельностью. Первоначально коммерчески обогатилась лексикона для 40 языков. С 2010 года добавлен разговорный чат-бот. До 2015 года с AIML и Pandorabots, а затем владеть программным обеспечением на HTML, PHP, Javascript и MySql.

В 2020 году архитектура для Semanta была готова, и теперь я готов позволить чат-ботам Semanta, в дополнение к автоматической и полуавтоматической генерации вопросов и ответов, сделать ее целостным и понятным.

Чат-бот способен выступать в роли цифрового собеседника. Разговор между человеком и гуманоидами или гуманоидами реконструируется из нескольких источников в понятную историю.

В отличие от Google, Microsoft и других технологических гигантов, основа для Semanta была заложена лексикой, грамматикой и семантическими свойствами языка.

Семанта превращает текст в понятное знание

Абстрактный

Semanta – это система, с помощью которой Lingvistica разрабатывает продукты и сервисы для процесса обучения для чат-ботов и цифровых представителей.

Этот документ предназначен для всех, кто хочет встретиться с посетителями сайта с виртуальным собеседником. Знания, необходимые для этого, содержатся на собственном сайте и вне его в «Интернете вещей». Semanta позволяет веб-мастеру анализировать источники в текстовом виде и, таким образом, позволяет собеседнику быть неотъемлемой частью веб-сайта. В данной статье обсуждается вопрос «[Семанта или как компьютерная программа могла научиться говорить](#)» и ответ на него Эда Кула.

Знакомство

Semanta – это система, с помощью которой Lingvistica разрабатывает продукты и сервисы для процесса обучения для чат-ботов и цифровых представителей.

Этот документ предназначен для всех, кто хочет встретиться с посетителями сайта с виртуальным собеседником. Знания, необходимые для этого, содержатся на собственном сайте и вне его в «Интернете вещей». Semanta позволяет веб-мастеру анализировать источники в текстовом виде и, таким образом, позволяет собеседнику быть неотъемлемой частью веб-сайта. В данной статье обсуждается вопрос «[Семанта или как компьютерная программа могла научиться говорить](#)» и ответ на него Lingvistica.

Альфа-версия 15.04.02 стала первой версией, с которой обратились заинтересованные стороны. Альфа-версия в основном использовалась для создания интереса и проверки того, насколько существует интерес к услугам Semanta.

От чат-ботов и цифровых сотрудников



Рисунок 1 Цифровой процесс Semanta

Что такое чат-бот или собеседник?

Чат-бот

Под чат-ботом Lingvistica понимается система компьютерных программ, баз данных и процедур, с которыми более или менее человеческое общение может осуществляться через Интернет.

Языковойбот

Под языковым ботом Lingvistica понимается система компьютерных программ, баз данных и процедур, с которыми можно общаться через Интернет. Архитектура Semanta позволяет веб-мастеру или веб-редактору использовать языкового бота для поддерживаемого языка. На данный момент существуют языковые боты для голландского, английского и русского языков.

Digital gesprekspartner

Lingvistica понимает цифрового собеседника как скрипт, который способен найти знания в одной или нескольких областях знаний и преобразовать их в полезный ответ. Каждая профессия а их более 1250 может быть представлена в цифровом виде собеседником.

Human gesprekspartner

Человек-собеседник, распознавание лиц или температура окружающей среды, и может двигаться на основе внешних или внутренних команд. Это не относится к фиксированным процессорам, таким как чат-боты.

Xbot

Xbot знает, где он находится, и может переходить из одной комнаты в другую по команде или по собственной инициативе и подниматься или спускаться по лестнице. Робот может приветствовать кого-то, кроме Эда Кула, зная, где находится **Эд Кул** и что он делает в любой момент. Выполнять действия, продиктованные временем. **Приготовьте** завтрак или чашку чая.

Начните использовать лекарство, предлагая его в определенное время.

Пожалуй, одной из самых важных функций является наблюдение за средой обитания, **Эд просыпается** вовремя, газ включен, выключен свет, низкий **нагрев**, закрыты шторы, заперта дверь. При обнаружении аварийной ситуации или отклонения вызовите автоматически. Все сообщения о поведении Эда Кула между Xbot можно интерпретировать текстуально.

Сценарист

Для Lingvistica сценарист VGPT является первым собеседником, обученным Semantus, чтобы позволить веб-редактору визуализировать сценарии для своего конкретного Xbot и настроить сценарий и разговорное поведение. Основываясь на палитре Xbot, см. выше, вы можете воплотить в жизнь своего собственного сценариста с помощью Semantus.

Искусственный интеллект

Фактический интеллект заключается в обработке языка и особенно семантических аспектов. Семантические аспекты словоформ определяются контекстом, в котором используются словоформы. По сути, робот – это механический объект с бионическими свойствами, дополненным процессором, в котором все аспекты робота объединяются. На данный момент производственные роботы вряд ли оснащены интеллектуальными речевыми технологиями. Только ограниченное количество роботов способно перевести общее количество событий вокруг робота в текст. Это касается всех сигналов от робота от датчиков, распознавания речи и распознавания лиц, температуры окружающей среды, влажности, качества воздуха, размеров помещения и т.д. Сигналы, которые, как и у человека, влияют на «мыслительный процесс» робота, и могут быть предложены в виде текста процессу обработки языка роботом.

Опыт работы с Xbot

Робот способен измерить рукопожатие человека и определить, было ли рукопожатие твердым, нормальным или легким. Интегрируя это в текущий разговор, Xbot гуманизируется.

То же самое относится и к окружающим звукам, из которых робот может выразить запрос выше ряда децибел в разговоре, чтобы снизить уровень шума.

В случае с роботами в здравоохранении возникает вопрос о том, насколько эффективно такие роботы по уходу могут общаться с пользователями.

Разработка носимой робототехники, экстремальным примером которой являются экзоскелеты, может помочь облегчить физически сложные профессии.

Заключение

Таким образом, вывод заключается в том, что все электромеханические сигналы робота, если их перевести в текст, интегрируются в процессе обработки языка.

Семанта учитель

Цифровой учитель – это особый виртуальный собеседник. Помимо ведения беседы, учитель или учитель должен уметь преобразовывать предложенную информацию в знания. Это увеличение знаний является непрерывным процессом, который осуществляется частично автоматически, а частично человеческими действиями. Lingvistica может обучить ваших сотрудников обучать ваших цифровых учителей. Семанта — это цифровой учитель, который может обучать других цифровых учителей. Версия Пандоры рассматривается Семантой как собеседник, с которым посетитель вступает в разговор и обучается через

Анализ выражения и разговорных элементов следует фиксированной схеме.

Шаблон обработки

- **Отбор.** Свободное высказывание, текстовый файл, веб-страницы веб-сайта.
- **Предварительная обработка.** Разделенные высказывания, словоформы и сегменты
- **Грамматический анализ.** Характер, чувство и практичность
- **Семантический анализ.** Отношения в контексте
- **Обогащение.** Дополнения к характеру, чувствам и языковым характеристикам
- **Захват и преобразование.** В зависимости от принимающего цифрового партнера, этот процесс является автоматическим или ручным.

Программа O

Версия доступна веб-мастеру. Роль Lingvistica заключается в том, чтобы направлять веб-мастера в применении функциональности semanta для своего веб-сайта. Поскольку виртуальные собеседники могут быть представлены как обучающиеся виртуальные роботы, существует также потребность в учителях для этих роботов. Lingvistica отвечает на это продуктами и услугами, которые упрощают построение, обслуживание и эксплуатацию чат-ботов и обучают цифровых учителей, которых можно использовать для обучения чат-ботов.

Отправная точка

Для того, чтобы уметь «разговаривать», компьютерная программа нуждается в знаниях. Для *Lingvistica* это знание в виде текстов. Тексты, которые могут быть предложены во всех видах форм. Каждое выражение через Интернет содержит значимую информацию.

Для веб-сайта *Semanta* использует Интернет в качестве источника, чтобы превратить его в обрабатываемый текст. К каждой текстовой форме, корпусу, текстовому файлу, обычному тексту или URL-адресу добавляется язык, тема беседы и область знаний, с которой может начаться собеседник. Язык, на котором «разговаривает» собеседник, определяется местоположением посетителя в Интернете и/или содержанием предлагаемого текста.

Если размер текста превышает 30 КБ, текстовый файл будет характеризоваться как корпус. Перед обработкой текста файл должен быть загружен на сервер *Semanta*. Это относится к корпусу, текстовому файлу и обычному тексту. *Semanta* разработала скрипты для каждой формы. После того, как текст будет загружен, *Semanta* разделит контент на предложения и фразы, которые, в свою очередь, могут быть отредактированы индивидуально. Из всех предложений и фраз грамматические знания записываются на основе отдельных словоформ.

Корпус

Собеседник использует корпус для общения с посетителем веб-сайта. Корпус состоит из отношений между словоформами, текстовыми сегментами, веб-страницами. Корпус содержит «дословные» представления формы выражения по теме в Интернете. Корпус является частью области знаний, которая может быть использована одним или несколькими собеседниками. **Текстовый файл** Для файлов размером до 30 КБ, *Semanta* предлагает возможность загружать текстовые файлы с расширениями: **, ТХТДАТ, АИМЛ** и предлагать их для анализа. Время обработки составляет около 5 секунд на КБ текста.

Обычный текст Текст до 1024 символов рассматривается *Semanta* как «обычный» текст. Для текстов выше этот номер может быть предложен *Semanta* в виде текстового файла. *Semanta* использует «простой текст» через диалог с посетителем веб-сайта. Беседа ведется короткими вопросами и ответами. При обработке обычного текста также может быть указан URL-адрес, из которого извлекается текстовая информация и предлагается *Semanta*. Текст является «сырым» и требует от пользователя выбора из найденного текста.

Интернет вещей

В «интернете вещей» любой объект с IP-адресом и достаточным количеством программного обеспечения может быть представлен в качестве виртуального собеседника. Вы можете спросить свои часы, в какое время, и сообщить стиральной машине, в какой программе следует повернуть белье. В соответствии с этим, *Semanta* может, в частности, внести свой вклад в разработку текстовых элементов при создании карт знаний или *google Knowledge Graphs* для вашего сайта.

Общение с виртуальным собеседником

Для реализации речи в сервисах Semanta тестируется **Nuance** и **ReadSpeaker**, если отдельные сервисы могут быть связаны. Также Pandora.org предоставляет говорящего собеседника, для которого мы разработали версию Semanta. Растущий спрос на общение на разговорном языке был решен в архитектуре программного обеспечения Semanta и может быть легко реализован.

Преподаватель языка

Семантус – это первая реализация ее роли **УЧИТЕЛЯ ЯЗЫКА**, с помощью которой Семанта пытается найти ответ на вопрос «Может ли компьютер научиться говорить?» Lingvistica заложила основу для положительного ответа на этот вопрос. Основываясь на нашей технологии Semanta, можно предоставить посетителю из Семанты, **не говорящего на голландском языке**, инструменты для знакомства с голландским языком с его родного языка. Основано на одном слове из голландского словаря. То, что относится к голландскому языку, также относится ко всем другим языкам, для которых мы разработали услуги и продукты.

В консультации с различными агентствами и интернет-исследованиями Lingvistica создала в Семанте, чтобы начать с того, что Lingvistica считает прорывом в области языкового редактирования в Нидерландах.

Интернет-помощник

Виртуальная собеседник Сценарист является первой реализацией своей роли **ИНТЕЛЛЕКТУАЛЬНОГО ПОМОЩНИКА**, с помощью которого Семанта пытается найти ответ на вопрос «Может ли компьютер научиться разговаривать?» Lingvistica заложила основу для положительного ответа на этот вопрос.

Основываясь на нашей технологии Semanta, можно предоставить посетителю инструменты от Semanta для генерации вопросов и ответов с его родного языка с партнером по обсуждению из выбранного корпуса.

Как это работает?

Это делается двумя способами:

- Семантические характеристики отдельных словоформ
- Грамматические сегменты от 2 до 5 словоформ

В зависимости от разделения ролей в программе обучения, роль Семанты, нашего первого **цифрового учителя**, может быть заполнена для виртуального **собеседника**. Под руководством человека-веб-редактора инструменты Semanta могут быть использованы для обучения чат-бота, аватара или робота воспроизводимым знаниям. Информация может поступать из отдельных выражений, текстов, текстовых файлов, веб-сайтов, википедии, bol.com, Google и т. Д. Предполагается, что каждый текст состоит из неструктурированного числа выражений, в которых могут содержаться вопросы и ответы. Знания, которые в конечном итоге могут быть раскрыты через **интеллектуальные разговоры** между человеком и виртуальными собеседниками.