Sri Lanka Association for Artificial Intelligence
First Sinhala Chatbot in action

Budditha Hettige
Department of Statistics and Computer Science,
Faculty of Applied Science,
University of Sri Jayewardenepura, Sri Lanka.

&

Asoka S. Karunananda
Faculty of Information Technology, University of Moratuwa,
Sri Lanka.
Introduction

- Introduction: Chat systems
- Structure of a Chatbot System
- Some Popular Chatbot Systems
- Design - Sinhala Chatbot System
- Implementation
- Chatbot in action
- Further work
Chat systems

- Computer-based chat system is one of the most popular commutation methods used in the modern world
- Enable communication using natural languages such as English
- Types of Chat Systems
  - Human-Human dialog systems
  - Human-Computer dialog systems (Chatbot)
Human-Human dialog systems

- Development is easy
- Work only as a mediator between two humans who actually manipulates the respective natural language
- Do not need machine level natural language processing abilities
- Example: Yahoo Messenger, MSN Messenger
Human-Computer dialog systems (Chatbot)

Is a more challenging task

All these chat systems are available in English language
Overview: Chatbot System

Human input

Analyzer

Knowledge Identification Engine

Knowledge Base

Computer Output

Generator
Analyzer reads input sentence from user and analyze Syntax of the given sentence

- Identify appropriate Tags and Patterns
Knowledge Identification Engine

- Reads Tags and Patterns from analyzer and find the suitable answer from Knowledge base
- Use some Search mechanism to identify the correct answer
Knowledge Base

- Brain of the chatbot system
- Is a database
- Contains knowledge of the chatbot system
Generate appropriate sentence
ELIZA is an early Artificial Intelligent program that was written in the mid 1960s by Joseph Weizenbaum to simulate a non-directive psychotherapist.

- Input sentences are analyzed on the basis of decomposition rules
- ELIZA had very limited natural language processing capabilities
Elizabeth is another Chatbot system that is an adaptation of the Eliza system. Elizabeth uses to store knowledge as a script in a text file, with each line started with a script command notation. Elizabeth has the ability to produce a grammar structure analysis of a sentence using a set of input transformation rules to represent grammar rules.
A.L.I.C.E

Artificial Linguistic Internet Computer Entity

- is a software robot or program that you can chat with using natural language
- ALLICE uses AIML files to implement its knowledge

It was developed by the Alicebot free software community during 1995-2000 to enable people to input dialogue pattern knowledge into Chatbots based on the ALICE free software technology.
ALLICE uses pattern-matching algorithm to identify user input and this algorithm uses depth-first search techniques.

ALICE has passed the Turing test in two consecutive years.
Sinhala Chatbot System

- Can Communicate through Sinhala Natural Language
- Can answer simple questions
- Can do the small operations
Design: Client-Server model

- Client-server model
- Client can use chatbot through client server network
Server side Design

Network

Server Socket

Morphological analyzer

Lexical Dictionaries

Sinhala Parser

Knowledge Identification Engine

Knowledge base

Application Module

Sinhala Composer

Morphological Generator
Server Socket

- Reads data string from clients
- Pass it into Sinhala Language passing system
Sinhala language passing system contains

- Morphological Analyzer
- Sinhala Parser
- Three Dictionaries
- Morphological Generator
- Sinhala Composer
Morphological Analyzer

- Reads a data string (sentence) word by word
- Identifies grammatical information
- Send each information into Sinhala parser
Sinhala Parser

- Analyze syntax of the Sinhala Sentence
- Identify the sentence patterns
Three Dictionaries

- Base dictionary
  - Stores base words
- Rule dictionary
  - Stores Grammatical rules
- Concept dictionary
  - Stores synonyms and antonyms
Morphological Generator and Sinhala composer

- Morphological generator generates appropriate words
- Sinhala Composer compose appropriate Sinhala sentence
Knowledge Identification Engine

- Reads all the information given from Sinhala Language passing System
- It work as inference engine
- Uses pattern machine algorithms to identify user input
Some question patterns

- **msg** – Message
- **qyn** – Question with yes/no answer form
- **qni** – Question with more answers
- **qwc** – Question with command
- **qun** – Unknown question
- **qda** – Question with direct answer
Knowledge base

- Stores all the requires knowledge in a chatbot system
- Implemented using SWI-Prolog database
Application module

- Can run appropriate commands and read the results

Example
- Can run System Command
- Open / Close applications etc.
Software requirements

- Software
  - SWI- Prolog 1.4.7
  - JDK 1.4
Sinhala Chatbot in action

Chat  Settings  Help

��: කිසේ නමුත් විශේෂකරොට?
.respond: කිසිමදී පැරණිතර මහවිඳුර පැතිලිකුවට පෙන්සියේ කිසිමදී මල්ලමක් සිදුකල මේවා.

��: නිල් බුද්ධම පැවැති?
.respond: පැවැතිය මේම 200 මෙල් නිකන්ත්තා 10 ම්‍රාලයේ.

��: විශේෂකරු 45 විශේෂකරු?
.respond: 0.7071067811865475

��: නිකන්ත්තා පැවැති?
.respond: 22 39 මේම නැසේය 42 මේම.

��: මෙමත්?
.respond: පැමිණිදී පැමිණිදී පැමිණිදී මෙමත්?

your Message  Send
How Sinhala chatbot Works

- Input Sentence

- Morphological Analyzer identified
  - adjective (හොඳ).
  - noun (මුස්ක).
  - verb (ලැදු).
How Sinhala chatbot Works

- Sinhala Parser

- Knowledge Identification Engine Pattern \( \text{today, date, qu} \).

- Knowledge Base

\[
\text{do\_action(\text{today, date, qu})}.
\]
How Sinhala chatbot Works

- Application Module
  \[\text{printtoday(PD):- date(A), assert(A), date(Y,M,D), retract(A), mounth(M,Mo), string_concat(Y, ', ', Year), string_concat(Mo, ', ', Month), string_concat(Year, Month, YM), string_concat(YM, ', ', PYM), string_concat(PYM, D, P), string_concat(P, ' fjksod', PD).}\]

- Morphological Generator and Composer generate appropriate Sinhala sentence
How Sinhala chatbot Works

Chatbot System Output

[Sinhala Chatbot window with text in Sinhala]

- : දෙපු මිදුරු මවැලේ
- : 2006 වර්ෂයට ඇති 11 යොදා
Further work

Extending the chatbot to operate on a more specific domain
Thank you!