**Semantic Parsing for Workflow Automation** **- Flight Reservation System**

**Shreejit Nair and Sowmya Jandhyala**

**Introduction**

This project is an attempt to apply the natural language into the machine (Machine Translation), so that it can be processed and interpreted just like human. This is an attempt to look forward to implementing natural language on Flight Reservation and Query System in order to increase the capability of Natural Language Processing.

So far in the current system, search techniques are seeking results through the keywords entered by the user with respect to those in the related databases or specific keyword based instructions. The Natural Language Processing engine would match all the keywords that have been collected before and then process it one at a time, continuously until the goal was discovered. In simple terms, NLP environments and tools can be classified into several categories such as:

* Translating Machine System
* Natural Language Interface and User Interface Management System
* Text Processing/Understanding System

**Problem Statement Description**

This mini project attempts at building meaningful workflow steps on an application using semantic parsing techniques and aims to translate natural language text into building steps that an underlying engine can understand and simulate based on previously provided inputs from an end user. The project takes a sample demo website for flight reservation systems and attempts to perform some key tasks such as:

1. Search for flight
2. Book a flight
3. Register a new user to the system

Whenever a user input is provided for which the agent is unable to perform all the necessary steps, a dialog prompt is provided to the end user asking for further instructions.

The application of the program can be extended beyond flight reservation systems to any external real world system for workflow automation or testing an application.

**Design and Structure of the system**

Natural Language Processing consists of 3 major components that are knowledge base, inference engine and user interface.

1. **Design of Knowledge Base Basically**

Knowledge base includes huge domain and heuristic facts. All of the gathered knowledge will be stored in the knowledge base in form of pattern matching. The system would need to have 2 kind of knowledge base. Both of which are always correlated to each other and accurate development architecture is necessary because it effects the final result.

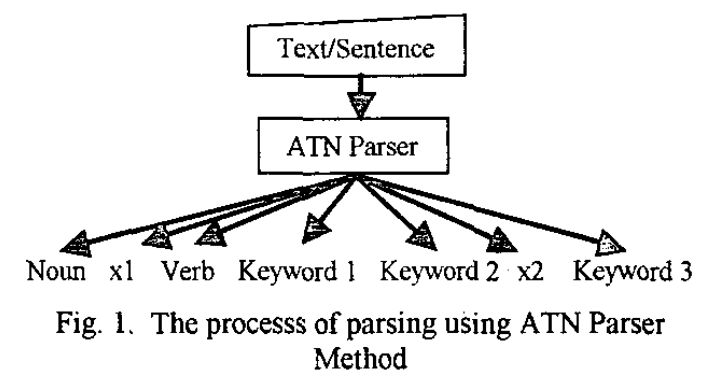
1. **Design of Inference Engine**

This application is mainly concerned with the semantics level,that is, understanding the meaning of a text entered by the user. There are 3 major techniques that had been used,

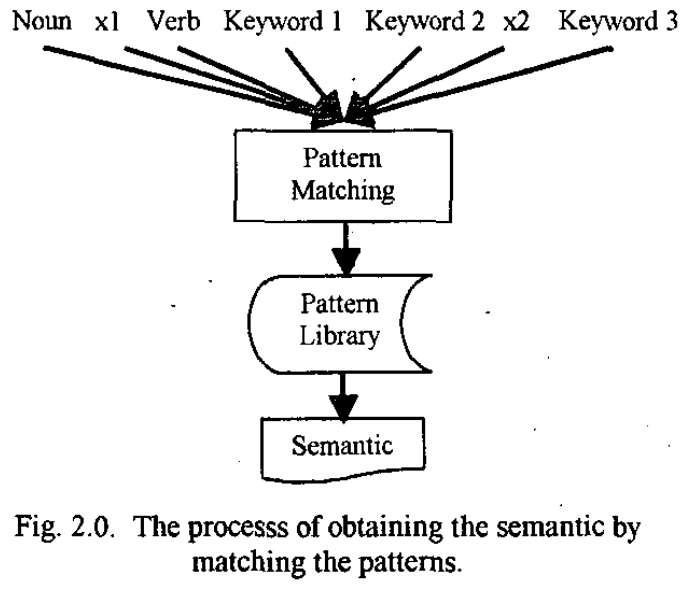
Augmented Transition Network Parser Technique to parsing the string of text (sentence) into the several parts (subjects, predicates, Nouns, Verbs, ... ) (look at Fig I),

Pattern-Matching Technique to match the parts to the Pattern Library in obtaining the keywords, then the semantic will be understandable (look at Fig 2)

And finally, is the Pattern-Matching Searching Technique that is the searching technique based on the keywords in the Knowledge Base (look at Fig 3).

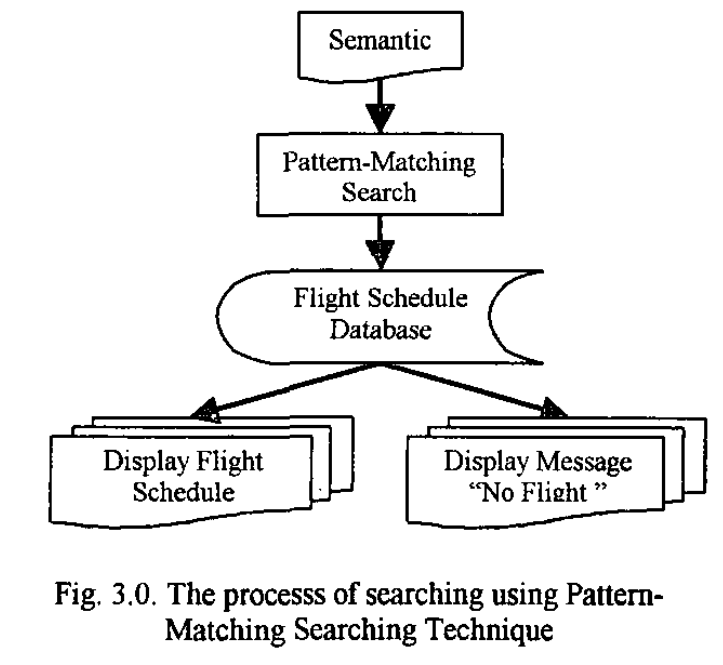


In this process, the system will store the entered sentence (input) temporary in the ‘virtual database’. Then, it’ll need to execute the ATN Parser Technique onto that Sentence and separate it into the several words.



On the other hand, the system would have to interpret the input and get the meaning from it. For that purpose, we would need to attempt to apply the Pattern Matching Technique. Technically, it’s involves 2 objects, namely, the input and pattern library.

Actually, the pattern library is one of the databases of this system, which contains a lot of types of patterns. And, it’s useful to identify the input (sentence) either it’s verb, nouns ... etc. So that, eventually we may decide whether the input was validate or not to be a complete sentence based on Malay Language Grammar Structure. As a final result, the system will translates to get the semantic before it is used in the next process.

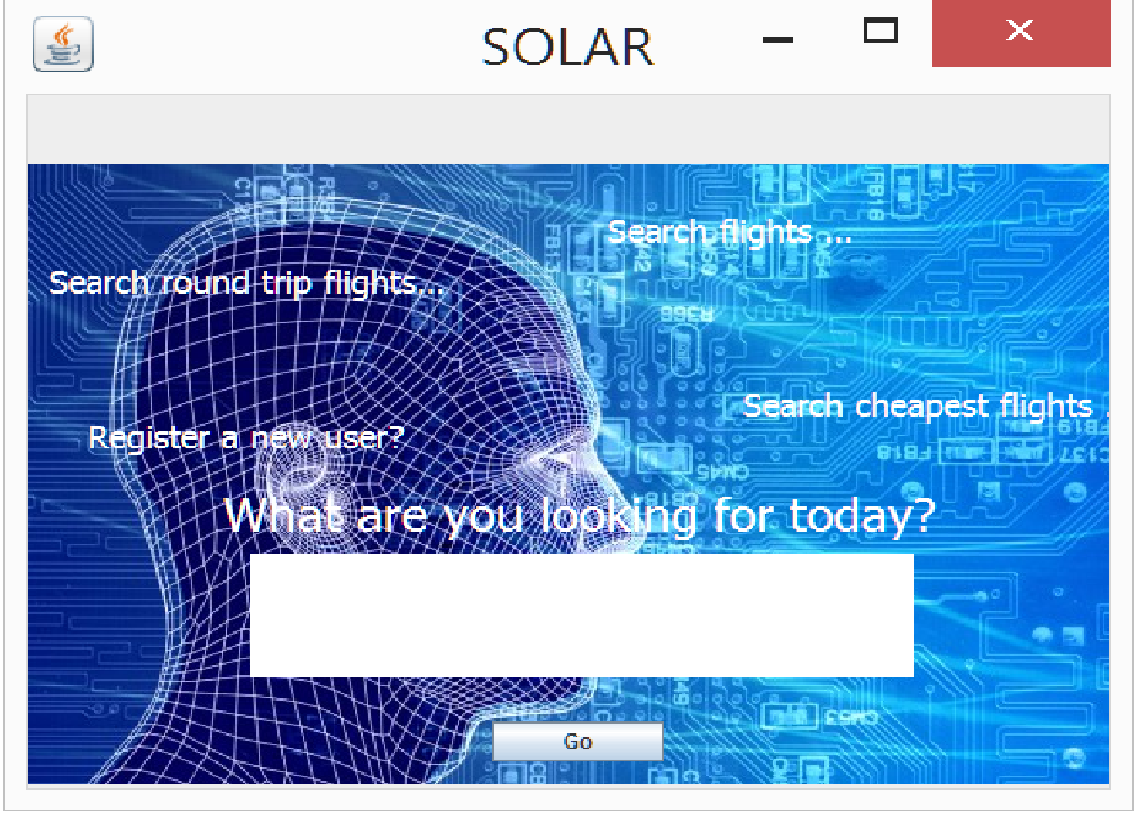


After we go through the first 2 processes, the next process is to focus on the searching technique. And, as we noticed before, the system will be using Pattern Matching and Searching Technique by matching against the semantic and the data in the flight reservation database.

1. **Design of User Interface**

We have designed a Java Swing interface which responds to queries entered by user in Natural text, tries to interpret the commands, resolve the path of action and perform the necessary steps for the user on the screen.

Nice to have: Additionally, an attempt would be made to build an interactive layer for incomplete instructions. For eg. If the user inputs “Search for flights”, the system should ask for destination and take the destination as input from the user.

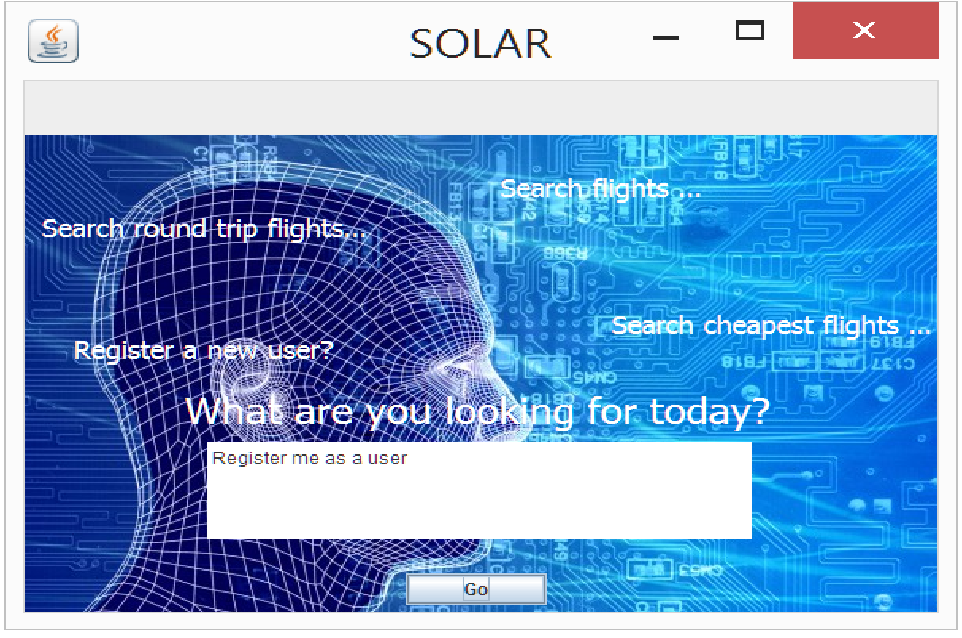
 The user interface is a Java Swing application that accepts queries from the user. The system or agent called SOLAR tries to parse / interpret the user query provided in natural language interprets the commands and attempts to reproduce the actions that the user has provided. On finding the appropriate path it performs the necessary steps, if not it invites the user to provide more information to proceed with the instructions.

Data Set: We would be creating custom annotated data simulating various permutations and combinations of how a real world user would attempt at searching for flights on the system.

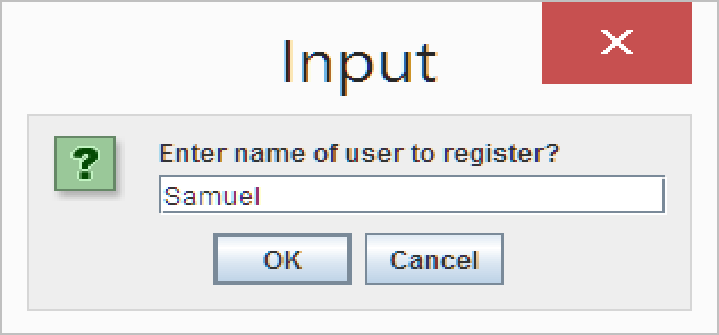
Sample website being used for automation purposes:

<http://newtours.demoaut.com/>

If the agent is unable to parse instructions the user is presented with a dialog interface asking for the required inputs to complete the set of instructions.



SOLAR is unable to register a new user with limited information



User is prompted with an input box to complete further instructions to proceed so that the agent has all the required information available to complete registration of a new user.

**References**

1. http://researcher.watson.ibm.com/researcher/files/in-saurabhsinha/ata-techrep-RI11014.pdf