Marco Grazioso

Computational Linguistics Technique for the development of ChatBot architectures for commercial applications

Tutor: Francesco Cutugno
co-Tutor: Valentina Russo

Cycle:XXXVI Year: I

My background

- MSc degree in Computer Science with a thesis in Human Computer Interaction
- URBAN/ECO Research Center
- PhD started 1 November 2020
- Scholarship funded by Logogramma s.r.l.

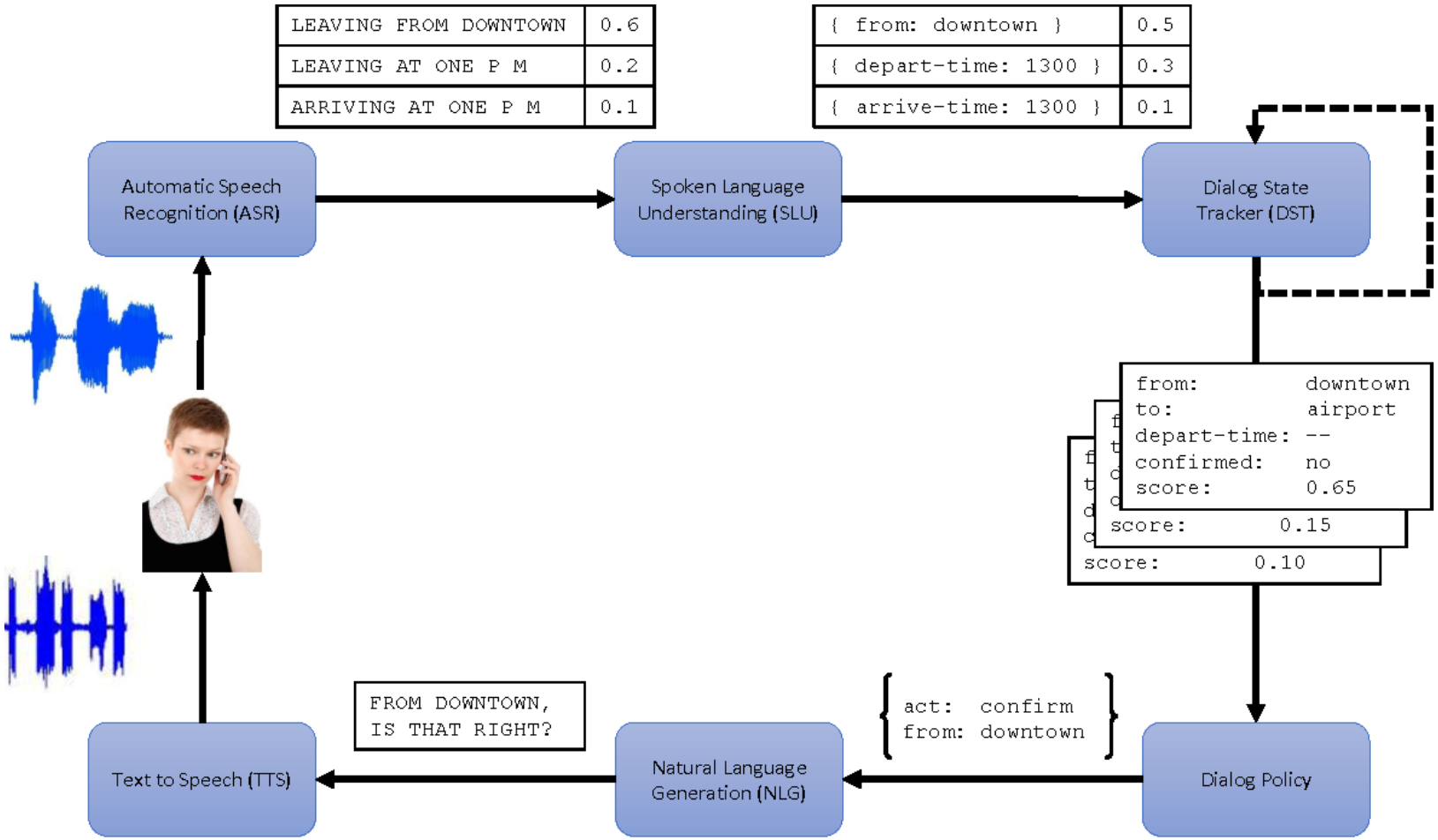
Research field of interest

- In the field of Artificial Intelligence, the **Dialogue Systems (DS)** are becoming more and more attractive for the scientific community
- Intelligent DSs need to be capable of interacting with users in a natural and intelligent way
- Several Natural Language Processing tasks contribute to the developing of a DS:
  - Recognition of user communicative intention (E.g. a booking a flight)
  - Recognition of named entity (E.g. departure/landing airport)
  - Slot filling: associates the entities with the parameter of a specific query
  - Dialogue state tracking: track the dialogue history (utterances, slots and more) to select the next system action.

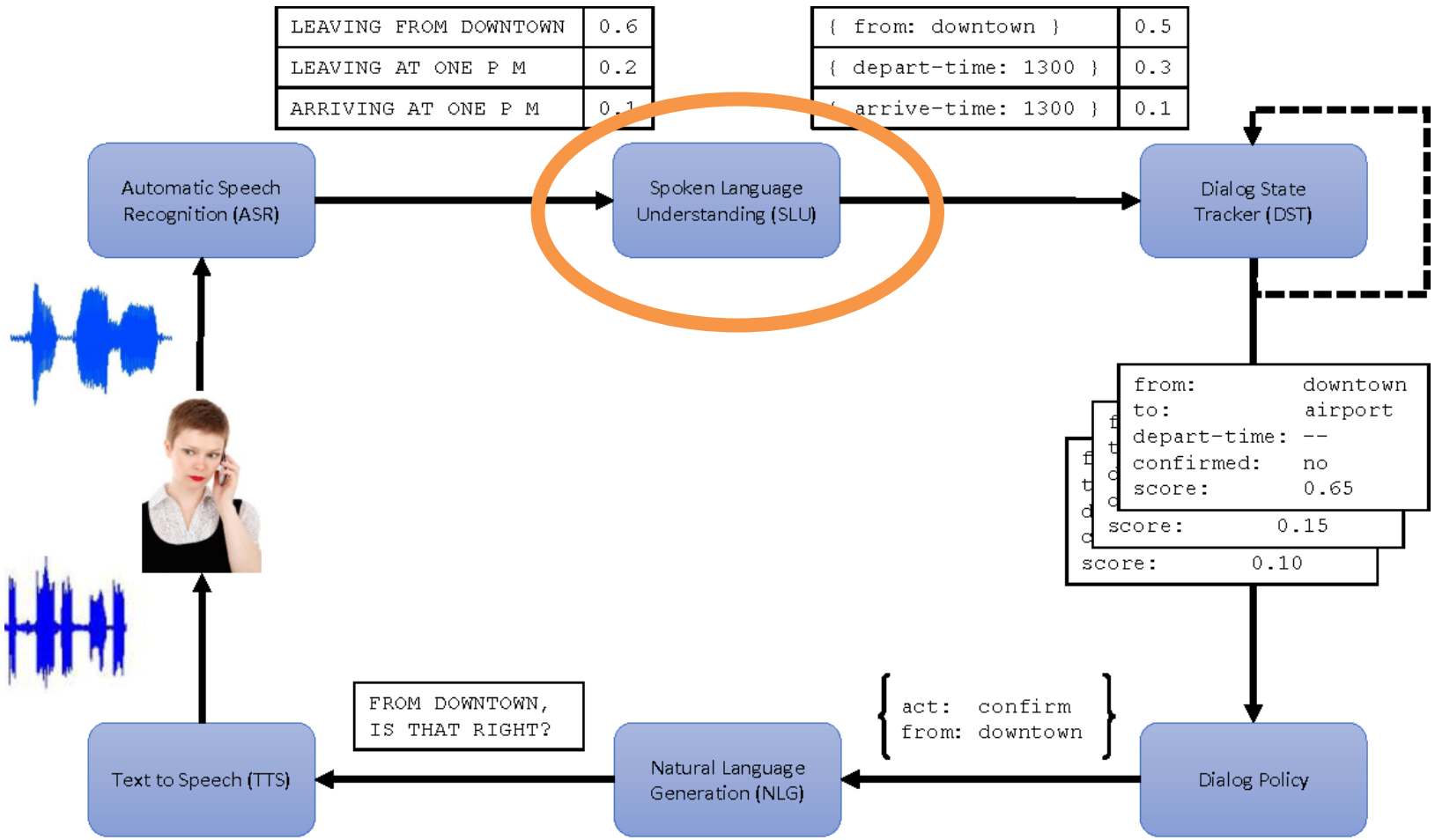
Summary of study activities

- Ad hoc PhD courses / schools
  - Digital Forensics' methods, practices and tools
  - AIRO PhD School 2021: Optimization and Data Science: Trends and Applications
- Courses borrowed from MSc curricula
  - Game Engine Architectures and Interactive Experiences
  - Human Language Technologies
- Conferences / events attended
  - CLiC-it: Seventh Italian Conference on Computational Linguistics
  - AI-HCI: 2ND INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE IN HCI
  - Lectures on Computational Linguistics 2021
  - ICMI: 23rd ACM International Conference on Multimodal Interaction

Dialogue System Architecture



Dialogue System Architecture



Research activity: Overview

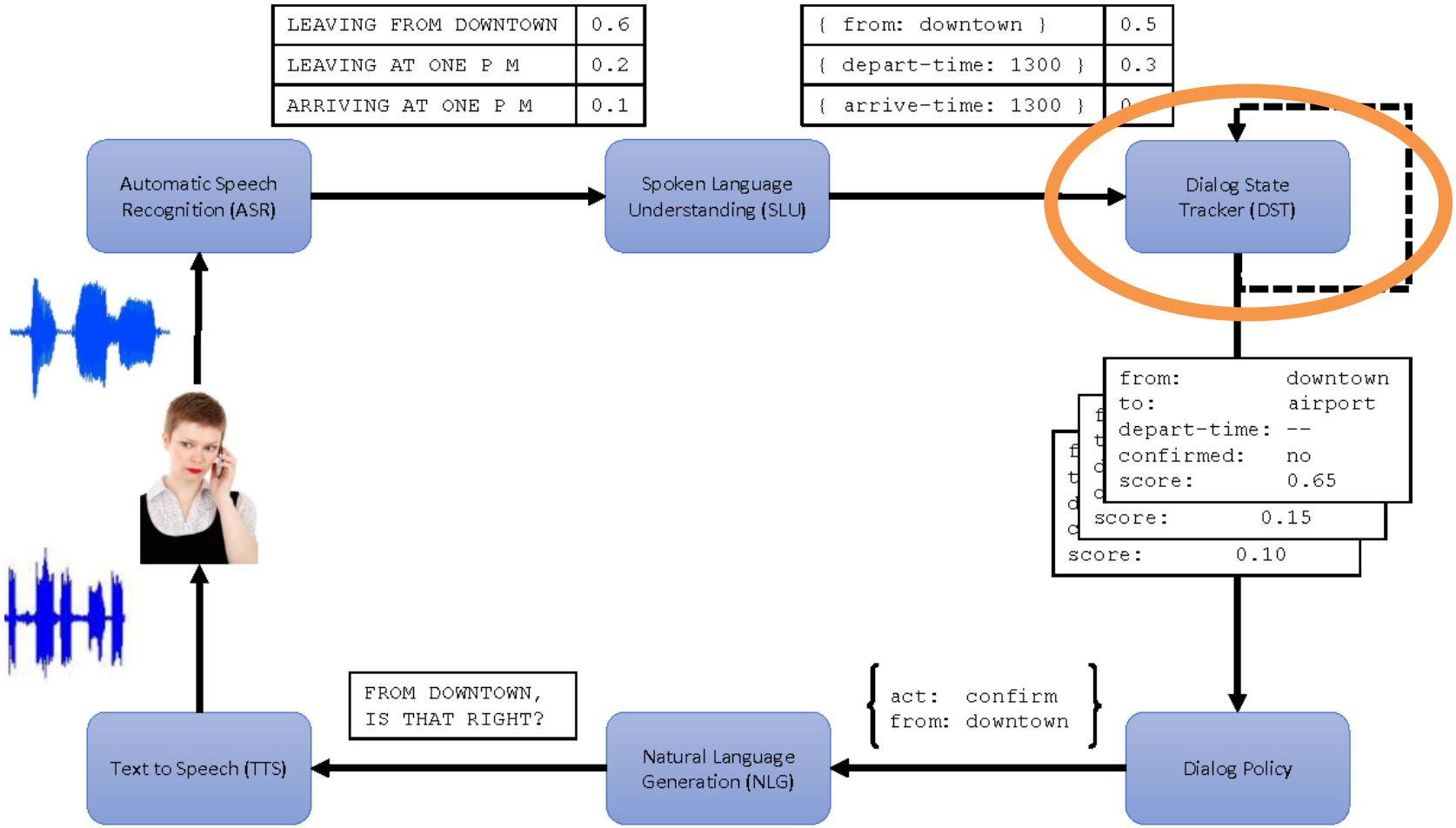
- Problems:
  - *How to recognize the user request ?*
  - *How to extract the entities from the request?*
  - *How to store the domain knowledge?*
- Objective
  - Design a scalable and domain adaptable infrastructure
  - *Train state of the art models for intent and entity recognition*
  - *Evaluate the models to analyze errors and make improvements*

- Methodology
  - Storage of the domain knowledge into a graph database in an domain-independent way
  - Definition of a training procedure retrieving training data from the graph db
  - Development of state of the art transformer-based models to perform intent and entity recognition
  - Realization of a web-based test application to evaluate intent recognition model

Products

[P1]	Grazioso M., Podda A.S., Barra S., Cutugno F. (2021) Natural Interaction with Traffic Control Cameras Through Multimodal Interfaces. In: Degen H., Ntoa S. (eds) Artificial Intelligence in HCI. HCII 2021. Lecture Notes in Computer Science, vol 12797. Springer, Cham. https://doi.org/10.1007/978-3-030-77772-2_33
[P2]	Russo V., Mancini A., Grazioso M., Di Bratto M., A dialogue-oriented approach to the design of a knowledge base supporting complex pragmatic scenarios" **submitted** at IJCoL(Italian Journal of Computational Linguistics)
[P3]	D'Asaro F. A., Raggioli, L., Malek S., Grazioso M., Rossi S., An Application of a Runtime Epistemic Probabilistic Event Calculus to Decision-making in e-Health Systems **accepted with revision** by the journal Theory and Practice of Logic Programming
[P4]	Developed a chatbot prototype performing intent recognition, named entity recognition and slot filling
[P5]	Developed a multimodal interaction system (voice and gestures) performing intent recognition and named entity recognition (see the publication Natural Interaction with Traffic Control Cameras Through Multimodal Interfaces)

Next year


LEAVING FROM DOWNTOWN 0.6
LEAVING AT ONE P M 0.2
ARRIVING AT ONE P M 0.1
{ from: downtown } 0.5
{ depart-time: 1300 } 0.3
{ arrive-time: 1300 }
Automatic Speech Recognition (ASR)
Spoken Language Understanding (SLU)
Dialog State Tracker (DST)
from: downtown
to: airport
depart-time: --
confirmed: no
score: 0.65
score: 0.15
score: 0.10
act: confirm
from: downtown
FROM DOWNTOWN, IS THAT RIGHT?
Text to Speech (TTS)
Natural Language Generation (NLG)
Dialog Policy

Dialogue State Tracking

- A dialogue state is composed of the current user utterance together with the dialogue history up to the current turn
- Representing the state, inferring the current state and selecting the next system action are challenging tasks
- Purposes:
  - Study state of the art approaches
  - Develop an innovative approach
  - Evaluate and compare it with other approaches on benchmark datasets: DSTC-1, DSTC-2, DSTC-3 and MultiWOZ
  - Develop a prororype and evaluate it in a real case study