

# Journée du Département 4 // Department 4 Day

Venez découvrir le  
département 4 du Loria !

On y étudie le traitement  
des langues et des  
connaissances.

20 juin 2018

8 h 30 - 12 h 30

Amphi Gilles Kahn / Loria

Come and discover  
Loria's department 4!

We work on knowledge  
and language  
management.

8 h 30 : *Accueil // Welcoming*

8 h 35 : Tien Hoa Le (Synalp)  
*Variational sequence to sequence learning*

8 h 55 : Tatiana Makhalova (Orpailleur)  
*A study on what MDL can do for FCA*

9 h 05 : Clément Beysson (Sémagramme)  
*Quantificateurs généralisés dynamiques pour l'analyse  
discursive*

9 h 25 : Marion Barré (Synalp)  
*Modelisation of the impact of the time changing risk  
aversion of investors on prices by deep neural networks*

9 h 35 : Pierre Monnin (Orpailleur)  
*PGxO: an ontology to represent pharmacogenomic  
relationships of various provenances, enabling further  
comparison*

10 h 55 : Pierre Ludmann (Sémagramme)  
*Dynamic Construction of Discourse Relations*

10 h 05 : Théo Biasutto--Lervat (Multispeech)  
*Multimodal Coarticulation Modeling, towards the  
animation of an intelligible talking head*

10 h 25 : *Pause // Break*

10 h 45 : Anastasia Shimorina (Synalp)  
*Natural Language Generation from Semantic Web Data*

10 h 55 : Kevin Dalleau (Orpailleur)

11 h 15 : Sunit Sivasankaran (Multispeech)  
*Keyword based speaker localization: Localizing a target  
speaker in a multi-speaker environment*

11 h 25 : Émilie Colin (Synalp)  
*Natural Language Processing and Automatic Generation  
of Grammar Exercises*

11 h 45 : Nicolas Turpault (Multispeech)  
*Semi supervised learning for sound scene analysis in real  
environments*

11 h 55 : Nyoman Juniarta (Orpailleur)  
*Sequential Pattern Mining using FCA and Pattern  
Structures for Analyzing Visitor Trajectories in a Museum*

12 h 15 : Maria Boritchev (Sémagramme)  
*Modeling dialogues in a dynamic theory of types*

12 h 25 : *Clôture // Closing*

### **Tien Hoa Le, *Variational sequence to sequence learning***

Recently, variational latent variable is gradually considered as a natural candidate to incorporate rich nuances of context to generate fluent and human-readable text. In this work, we explore various ways to incorporate it to sequence-to-sequence (seq2seq) framework. It includes injecting latent variable between encoder-decoder, feeding per decoder timestep independently and dependently with a simple normal distribution prior or a normal prior evolving through time. We're interested to see if the position, the plurality, the recurrent connection and the form prior distribution could be a decisive factor in employing VAE in seq2seq. The models are designed to be simple while on purpose, we hope that it still can incorporate the most variations into seq2seq model. We provide experiments on summarization datasets to show the effectiveness of each model.

### **Tatiana Makhalova, *A study on what MDL can do for FCA***

Formal Concept Analysis can be considered as a classification engine able to build classes of objects with a description or concepts and to organize these concepts within a concept lattice. The concept lattice can be navigated for selecting significant concepts. Then the problem of finding significant concepts among the potential exponential number of concepts arises. Some measures exist that can be used for focusing on interesting concepts such as support, stability and others. MDL (minimum description length) is also a good candidate that was rarely used until now in FCA for such objective. This is the purpose of this paper to present, explain and evaluate what MDL can bring to FCA practitioners as a measure for selecting significant and representative concepts. First experiments are also presented for evaluating the current approach.

### **Clément Beysson, *Quantificateurs généralisés dynamiques pour l'analyse discursive***

Le discours est par essence quelque chose de dynamique. Au fil des mots que l'on ajoute, nous modifions le contexte ainsi que les entités auxquelles il sera possible de faire référence par la suite. En particulier, les déterminants nous permettent d'introduire de nouveaux référents et interagissent avec ceux déjà introduits.

L'objectif de mes travaux est, en partant de la formalisation en quantificateurs généralisé des déterminants, de modéliser ces aspects dynamiques. De plus, comme il n'est pas toujours possible de réduire la sémantique à des propriétés sur les individus, il est nécessaire de disposer d'une théorie gérant aussi bien les référents singulier que pluriel.

### **Marion Barré, *Modelisation of the impact of the time changing risk aversion of investors on prices by deep neural networks***

Our main purpose is to be able to predict the dynamics of some given asset prices and then to build an investment strategy based on this prediction. Many price prediction models already exist in finance. They nevertheless present two main disadvantages. Firstly, most of them rely on unproven or questionable hypotheses such as market efficiency or non-arbitrage. Secondly, they fail to take human biases into account.

Contrary to most previous pieces of research, the rationale is not to use deep neural networks directly for pure time series prediction. We rather aim for exploiting deep networks in subtasks of the market analysis processing chain, merging them with financial concepts and integrating them in structured models where they can bring novel insights.

### **Pierre Monnin, *PGxO: an ontology to represent pharmacogenomic relationships of various provenances, enabling further comparison***

Pharmacogenomics (PGx) aims at identifying relationships between genomic variations and variability in drug response, including adverse drug effects. Knowledge in pharmacogenomics is typically composed of units that have the form of ternary relationships gene variant – drug – adverse event. Such a relationship states that an adverse event may occur for patients having the specified gene variant and being exposed to the specified drug. A huge number of pharmacogenomic relationships are available in the literature and specialized knowledge bases. On the other hand, PGx knowledge can also be discovered from the mining of Electronic Health Records (EHRs) and, in this case, may either correspond to new knowledge or confirm state-of-the-art knowledge that lacks "clinical counterpart" or validation. Therefore, there is a need for knowledge units from distinct sources to be further compared. To this end, we used Semantic Web technologies and developed PGxO, a small ontology to represent PGx relationships and their components. We also defined and experimented a set of rules, as a first approach to compare knowledge units from various sources. The ontology, the rules and their experimentations constitute a first step towards a more complete framework and illustrate some of the remaining difficulties for knowledge comparison in PGx.

## **Pierre Ludmann, *Dynamic Construction of Discourse Relations***

To account discourse relations and cohesion, Asher and Lascarides developed the Segmented Discourse Representation Theory (SDRT) on the ground of Kamp's Discourse Representation Theory (DRT). Yet DRT's dynamic logic rely on a model of imperative computation, hardly compatible with the applicative model basing Montague semantics. In order to retrieve the compositionality, de Groote and Lebedeva propound a redevelopment of dynamic logic, based on Church's simple theory of types. Eventually, Asher and Pogodalla combined the two investigation lines laying the foundation for a system conciliating SDRT and the dynamic approach based on the simple theory of types. In the wake of these works, mine is to define a dynamic model of discourse, based on type theory, that accounts for discourse relations.

## **Théo Biasutto--Lervat, *Multimodal Coarticulation Modeling, towards the animation of an intelligible talking head***

Speech is the natural communication medium between human beings. Thus, an animated talking head with realistic speech movements should be an effective human-machine interface. Moreover, it's well established that viewing the speaker's face tends to increase speech intelligibility. Hence such technology could have several applications: communication aids for hard-of-hearing people, embodied conversational agents for a noisy place, and more. However, proper computation of speech gestures should take account of coarticulation phenomenons, i.e. the influence of a sound and its articulation on another phoneme's articulation. This influence can be both retentive and anticipative, e.g. the lip movement associated with /s/ is different in 'see' and 'sue' because of the anticipation of the protrusion needed by /u/. Motivated by the recent success of deep learning techniques in many language-related tasks, we try to model these coarticulation effects with a neural network approach. As coarticulation is highly temporal-dependent, recurrent neural networks seem to be a tool of choice by their capacities to deal with time series. In a supervised fashion, we learn to generate the articulatory trajectories given the phoneme sequence and some acoustic features.

## **Anastasia Shimorina, *Natural Language Generation from Semantic Web Data***

There are two traditional ways to evaluate natural language generation (NLG) systems: using automatic metrics and collecting human judgments. Automatic metrics are mostly calculated based on the n-gram overlap between references and a hypothesis produced by an NLG system. Human judgments are scores assigned to different linguistic aspects (fluency, grammaticality, semantic adequacy) of produced hypotheses. This talk will present two types of correlation analysis carried out between automatic metrics and human scores. Depending on the usage of a system-level or sentence-level correlation analysis, correlation results between automatic scores and human judgments were inconsistent. Results obtained on NLG data showed a similar trend that was already observed in Machine Translation: automatic metrics correlate poorly with human judgments at the sentence level. Future directions of the PhD research will be also outlined: testing a capability of NLG neural models to generalise; finding a way to automatically spot deficiencies in produced outputs; comparing performance of neural generation models on languages other than English.

## **Kevin Dalleau**

Many unsupervised learning algorithms rely on a (dis)similarity measure to evaluate the pairwise distance between samples. Despite the large number of measures already described in the literature, in many applications, the set of available metrics is reduced by intrinsic characteristics of the data and of the chosen algorithm. In this talk, I will present a novel method based on extremely randomized trees. This method, UET (Unsupervised Extremely Randomized Trees), is based on the work of T.Shi and S.Horvath and can be generalized to heterogeneous data, enabling the clustering of datasets with mixed-types variables. The evaluation of the methods shows promising results, both in terms of discrimination of instances and in terms of running time.

## **Sunit Sivasankaran, *Keyword based speaker localization: Localizing a target speaker in a multi-speaker environment***

Speaker localization is a hard task, especially in adverse environmental conditions involving reverberation and noise. In this work we introduce the new task of localizing the speaker who uttered a given keyword, e.g., the wake-up word of a distant-microphone voice command system, in the presence of overlapping speech. We employ a convolutive neural network based localization system and investigate multiple identifiers as additional inputs to the system in order to characterize this speaker. We conduct experiments using ground truth identifiers which are obtained assuming the availability of clean speech and also in realistic conditions where the identifiers are computed from the corrupted speech.

We find that the ground truth time-frequency mask corresponding to the target speaker provides the best localization performance and we propose methods to estimate such a mask in adverse reverberant and noisy conditions using the considered keyword.

### **Émilie Colin, *Natural Language Processing and Automatic Generation of Grammar Exercises***

This thesis aims to explore deep learning methods to generate correct and fluent sentences. It is part of the METAL collaborative project on e-learning. The current research goal is to generate under constraints. For instance, given the input *mouse eat cheese/cat eat mouse* and a subject relative clause constraint, the generated sentence might be: *The cat eats the mouse that eats cheese*. The motivation for generating under constraints is that it can be used to generate sentences that are lexically and syntactically appropriate to automatically generate grammar exercises.

As a first step towards that goal, we have built a training corpus using a dataset called WebNLG (semantic web data associated with lexicalizations) and enriched the input data with syntactic labels indicating the syntactic constructs occurring in the corresponding sentence. We then used a bidirectional LSTM (with attention mechanism) associated with a decoder for the generation and have demonstrated that syntactically constrained generation increases diversity while maintaining a good level of linguistic (BLEU : 65.11) and syntactic adequacy (Synt : 91% sentences).

### **Nicolas Turpault, *Semi supervised learning for sound scene analysis in real environments***

Audio scene analysis encompasses of audio scene classification, audio tagging and audio event detection. The aim first two tasks is to assign one or several labels to an audio segment regardless of time localization. On the other hand, audio event detection, which is our field of research, aims to detect when an event happens and the type of event. This problem is far from being solved because supervised approaches proposed until now require audio data and annotations indicating the time boundaries of the events. While it is rather simple to get this kind of annotations for synthetic sounds, it is hardly feasible to annotate a large amount of real world audio data. Therefore, real audio data is available (Audioset, Freesound,...), but annotations are generally limited to segment level annotations (without time boundaries). Besides, the annotations are often obtained automatically and their quality can be rather low. In a first step, to overcome this label quality problem, we consider only a subset of the dataset that has accurate segment-level labels (usually verified by humans). The rest of the dataset is then considered as unlabeled. We then propose to develop a semi supervised approach based on this dataset partition. Our model combine the triplet loss used by Google in a fully unsupervised manner and extend it to a semi supervised setting.

### **Nyoman Juniarta, *Sequential Pattern Mining using FCA and Pattern Structures for Analyzing Visitor Trajectories in a Museum***

We are working on mining visitor trajectories in Hecht Museum (Haifa, Israel), within the framework of CrossCult European Project about cultural heritage. We present a theoretical and practical research work about the characterization of visitor trajectories and the mining of these trajectories as sequences. The mining process is based on two approaches in the framework of FCA, namely the mining of subsequences without any constraint and the mining of frequent contiguous subsequences. Both approaches are based on pattern structures. In parallel, a similarity measure allows us to build a hierarchical classification which is used for interpretation and characterization of the trajectories w.r.t. four well-known visiting styles.

### **Maria Boritchev, *Modeling dialogues in a dynamic theory of types***

Studying dialogues is a major subject in natural language processing. It is a complex task which links approaches from fields such as semantics, pragmatics and more generally logic and cognition. My work focuses on a compositional approach to dynamic models of questions and answers mechanisms in a dialogue setting. We want to control the storage of the data that is considered as available to all participants of a conversation. We work towards applying our model to real-life examples of dialogue interactions provided by three different corpora. Machine Learning is now the most popular way to solve NLP problems, however here, we focus on developing formal and computational tools to handle specifics of dialogue interactions on top of already existing methods for general discourse. In this talk, I present an approach to dialogue modeling in a dynamic framework.