once more for their generous support of the event.

Jan Sprenger
Tilburg Center for Logic and Philosophy of Science

Computational Linguistic Aspects of Grammatical Inference, 30–31 March

The International Community of Grammatical Inference organizes biennial conferences, called ICGI, in the even years and in the odd years it organizes co-located workshops or tutorials. These events are typically co-located with conferences that fall within the application areas of grammatical inference: in previous occasions machine learning or artificial intelligence conferences were chosen. This year, a workshop, called Computational Linguistic Aspects of Grammatical Inference (CLAGI) was organized by Menno van Zaanen and Colin de la Higuera. This event was co-located with the triennial conference of the European Chapter of the Association Computational linguistics in Athens, Greece.

The CLAGI workshop consisted of eight talks, an invited talk and a panel session. The talks were divided into three sessions. The invited talk was given by Damir Cavar. He gave an overview of work in the field of grammatical inference applied to linguistics. This illustrates the search for the holy grail: a general grammar induction model that learns all aspects of natural languages.

The first session contained three talks on transduction. Jeroen Geertzen presented a novel grammatical inference system that finds regularities in (human) dialogs and uses these regularities to predict future dialog acts. Dana Angluin and Leonor Becerra-Bonache concentrated on a formal description of the language learning task where the focus lies on incorporating semantics. Finally, Jorge Gonzalez and Francisco Casacuberta described their machine translation toolkit that uses a transducer inference system.

The second session was on language models and parsing. Alexander Clark, Remi Eyraud and Amaury Habrard discussed properties of the class of contextual binary feature grammars (which are known to be efficiently learnable) in comparison to the class of natural languages. Next, Herman Stehouwer and Menno van Zaanen illustrated an application of parallel n-gram language models in the context of typographical error correction. This was followed by a presentation by Marie-Hélène Candito, Benoit Crabbé and Djamé Seddah describing experiences with statistical parsing of French. Finally, Franco M. Luque and Gabriel Infante-Lopez investigated the performance of unambiguous non-terminally separated grammars in the context of natural language learning.

The last session on morphology contained one talk. Katya Pertsova analyzed a collection of learners that analyze boolean partitions. These learners are applied to the task of learning morphological paradigms.

Following the talks, Jeroen Geertzen, Alexander Clark, Colin de la Higuera and Menno van Zaanen briefly presented their experiences and ideas on competitions in the areas of computational linguistics and grammatical inference. This lead to a panel discussion with much interaction from the audience on possibilities of future competitions and their impact.

Overall, the workshop contained a wide variety of talks, ranging from technical and theoretical research on learnability of language to descriptions of practical natural language learning experiments. This variety illustrated the wide applicability of grammatical inference, but at the same time the workshop showed that people in the field of computational linguistics are unfamiliar with the field of grammatical inference and vice versa, even though their research is highly relevant to both fields. This workshop should be seen as an initial attempt at bringing these fields closer together.

Menno van Zaanen
ILK, Tilburg

Colin de la Higuera
Laboratoire Hubert Curien, Sait-Etienne (France)

Sparsity in Machine Learning and Statistics, 1–3 April

Sparsity has emerged as one of the most important modelling tools of the past decade. Its popularity is easy to grasp; it is both conceptually simple, and the estimation methods that follow as its consequence have already been implemented for several decades in geophysics, as well as in image estimation, signal processing, statistics and machine learning.

The aim of the workshop on Sparsity in Machine Learning and Statistics, organized by the UCL Centre for Computational Statistics and Machine Learning (CSML) and sponsored by the PASCAL Network of Excellence and the Royal Statistical Society, was not only to highlight research questions that are currently receiving particular attention, but also to draw the common threads from different fields, in particular statistics, signal processing and machine learning. This second aim of the workshop aligns closely with the CSML Centre that is working to promote cross-fertilisation between these fields as well as application of relevant techniques across the sciences.

The concept of sparsity is simple; in whatever family of models we may assume our data was generated