

Newsletter

Contents

COVER STORY

DIRAC First Summer Workshop 1

FOCUS

- The IDIAP Research Institute 2
- Identifying unexpected 3

INSIDE DIRAC

- The IDIAP Research Institute 2
- Identifying unexpected 3

News

Happy birthday !

DIRAC is already one year old.



DIRAC First Summer Workshop

This year DIRAC will have its very first summer workshop. To organize this event Dirac has joined forces with IST-project CoSy (www.cognitivesystems.org) because of its similarity in goals :

- to investigate the cognitive processes of learning
- and understanding environments using data as retrieved by different sensor inputs.

The summer workshop is sponsored by the PASCAL Network of Excellence (www.pascal-network.org) and will carry the name "Summer Workshop on Multi-Sensory Modalities in Cognitive Science". It is scheduled to take place August 25-29 in Studienzentrum Gerzensee, located in an beautiful

- Visual Object Recognition Algorithms
- Image Matching and Camera Tracking
- Spatial Sound Processing
- Speech Communication by Humans and Machine
- Autonomous Robot Learning of Foundational Representations
- Developmental Algorithms
- Cognitive Architectures
- Markov Decision Processes

These tutorials will be brought by well-known lecturers in the field : Ales Leonardis, Bastian Leibe, Benjamin Kuipers, Bernt Schiele, Craig Boutilier, Frederic Kaplan, Hynek Hermansky, Joern Anemueeller, Matthias Scheutz, Rufin Vogels, Tinne Tuytelaars and Tomas Pajdla.



area in Switzerland which, with its stunning view, unspoiled nature and mild climate, beckons young and old. This quiet location, away from the hustle and bustle of our daily activities, offers an ideal place to contemplate the state-of-the-art and recent developments in the field of cognitive sciences. It will foster discussions between the participants from different scientific backgrounds, to see how different branches in cognitive sciences can be brought closer together in order to strive towards a common goal.

The program of the summer workshop will be spread over five days during which the participants may follow tutorials on various topics such as :

- Neural mechanisms of recognition and categorization

Each day will be closed by group discussions during which we will try to identify where these different topics may lead us in the future and to see how they can be geared towards a common goal, encouraging integration between the different disciplines in cognitive sciences.

To nourish the cultural appetite of our participants as well, a visit to the old town of Bern has been planned, which has been acknowledged as a UNESCO World Heritage site.

The summer workshop is open to 25 motivated students, researchers and industrial professionals. Online registration is possible until the 20th of May. For more information, see www.diracproject.org/workshop-2007

Cover Story

Newsletter

The IDIAP Research Institute



The IDIAP Research Institute (www.idiap.ch) is an independent, not-for-profit, research institute located in Martigny (Switzerland), and affiliated with the Swiss Federal Institute of Technology at Lausanne (EPFL), and the University of Geneva. Primarily funded by the Federal Government, the State of Valais, and the City of Martigny, IDIAP is involved in numerous national and international (EU and US) projects, as well as in multiple collaborative projects with the industry.

At the national level, IDIAP is also the "Leading House" of the National Center of Competence in Research (NCCR) on "Interactive Multimodal Information Management" (IM2, www.im2.ch). At the EU level, IDIAP is involved in numerous projects and Networks of Excellence, and is currently managing 2 Integrated Projects (AMIDA, www.amidaproject.org on the modeling of human-human communication in meetings, managed in collaboration with the University of Edinburgh, and DIRAC, www.diracproject.org on bio-inspired detection and identification of rare audio-visual events), and one STREP (MAIA, www.maia-project.org, on non invasive brain interaction with robots through Determination of intended action inferred from EEG signals). In the US, IDIAP is also involved in numerous collaboration with ICSI Berkeley (www.icsi.berkeley.edu), partner of a large DARPA project (GALE) and coordinator of a DTO project (Roadmap, as part of the VACE program).

IDIAP currently numbers more than 50 scientists including (including EPFL professors, seniors, postdoctoral researchers, PhD students and developers).

Research at IDIAP pivots around six major themes, which guarantees strong synergies among the various research areas as well as a better circulation of information among researchers. The research themes span from machine learning and applications like data mining, recognition of models and samples, to speech processing (audio signal processing, sound segmentation, signal enhancement, real time and noise-robust speech recognition, speaker localization, automatic dictation systems, voice command systems, automatic retrieval and transcription of audio documents, low bit rate speech compression and transmission); methods and algorithms for image and multimedia data processing, including detection, tracking and recognition of people, objects, movements and expressions, as well as differentiation of various image sequences.

In DIRAC, IDIAP leads WP-6 and is the project coordinator. Integration between partners is a vital component for the success of the project; IDIAP has so far promoted and hosted several exchange visits between researchers within the consortium. IDIAP's main contribution in DIRAC are into:

WP1, (the Signal and Feature workpackage), where we are developing basic sound functionalities for the project. Here we have focused mainly on discrimination between speech and non-speech signals, with the goal of introducing a speech/non-speech detection (SND) module in the context of integrated systems for multi-person tracking and speech processing. We also carried out fundamental research in signal processing, particularly on acoustic and biophysical systems.

WP2, (Auditory Representations workpackage), where we aim at obtaining acoustic "objects" with different modeling and experimental techniques. Research work focused on a novel technique for identifying low probability words based on comparison of in-context and out-of-context derived frame-level posterior probabilities of phonemes. We also carried out fundamental research in non-linear signal processing and supervised learning

WP4, (Learning and Categorization workpackage) where we work on algorithms able to learn incrementally models of visual categories, and able to recognize their own level of competence (detection of ignorance). We focused on discriminative algorithms able to learn incrementally visual classes, and able to transfer knowledge across different acquisition devices.

In WP5 IDIAP contributed to development of a suitable strategy for identifying unexpected events based on fusion of top-down and bottom-up processes, which is the strategy being followed in WP2 for identifying unexpected out-of-vocabulary words. Further, IDIAP works on information fusion from several classifiers.

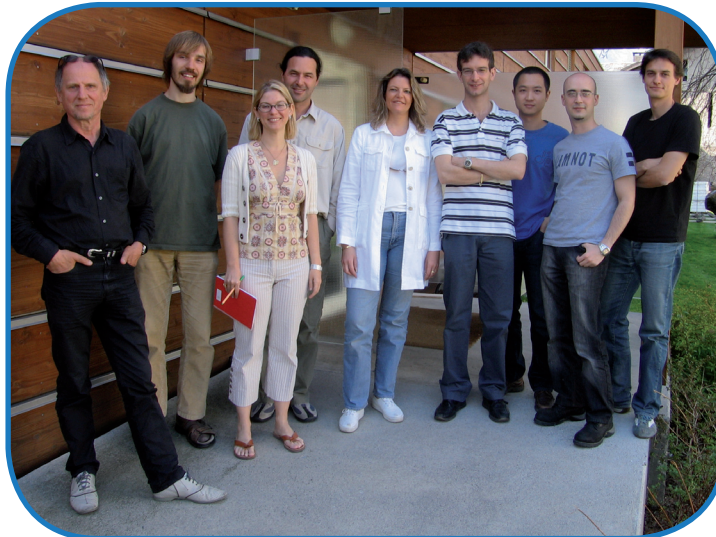
The present study uses auditory information streams where the large amounts of phoneme-labeled data needed for the training of the posterior

probability estimators are readily available. However, the studied techniques should be also applicable to fusion of audio-visual information.

The integration work package WP6 is headed by IDIAP with the objectives such as maintaining and reinforce collaborations among all partners by defining and developing appropriate scenarios for application of results of DIRAC research, identifying existing available audio, visual, and audio-visual databases that would support DIRAC research thrusts, and developing evaluation standards and protocols for DIRAC research.

The IDIAP team is currently composed of Barbara Caputo, Nicolas Fremaux, Mirko Hannemann, Hynek Hermansky, Andrew William Lovit, Jie Luo, Francesco Orabona, Nancy-Lara Robyr, Vincent Spano, Tamara Tasic, Fabio Valente.

IDIAP is also the site coordinating the project. Nancy-Lara Robyr is the program manager, and Vincent Spano is the responsible for the IDIAP web-page.



From left to right : Hynek Hermansky, Mirko Hanneman, Barbara Caputo, Vincent Spano, Nancy-Lara Robyr, Fabio Valente, Jie Luo, Francesco Orabona, Nicolas Fremaux

Newsletter

Identifying unexpected

Inesperata accidunt magis saepe quam quae speres, i.e. things you do not expect happen more often than things you do expect, warns Plautus (cca 200 B.C.). This Plautus' motto has been adopted by DIRAC to emphasise the consortium strive for handling surprising sensory inputs data that could represent a new danger or a new opportunity. A hypothesized cognitive process in processing of sensory inputs is illustrated in Fig. 1.

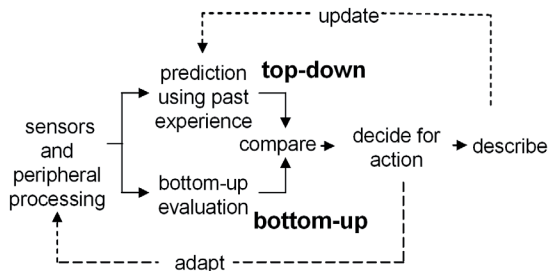


Fig. 1 Cognitive processing of sensory inputs

The sensory input triggers a predictive process in the upper path that uses top-down knowledge from the past experience and generates predicted components of the scene. In parallel, the scene components are also estimated directly (i.e. without the use of the top-down knowledge) from the input. A comparison between the two sets of components may indicate an unexpected item and an appropriate action can be undertaken. This action may involve further investigation of the item, and its description and adoption to the set of prior beliefs.

Detection of unexpected (out-of-vocabulary) words in machine recognition of speech

Our laboratory at IDIAP Research Institute is working on identification and description of low-probability words. The technique is shown in Fig. 2. First, the context-unconstrained phoneme probabilities are estimated. These are subsequently being used in the search for the most likely stochastic model of the input utterance. A side product of this search are context-constrained phoneme probabilities.

An example of a typical result is shown in the lower part of the Figure. An inconsistency between the two information streams could indicate unexpected out-of-vocabulary word. In this example, the HMM model inconsistency was introduced by removing the word 'three' from the recognizer vocabulary. The correct phoneme sequence for the word 'three' is represented in the HMM-derived posterigram (replaced by a sequence /z/iy//r//oh/ of the in-vocabulary word 'zero'). The ANN derived probabilities indicate in this case the correct sequence /th//r//iy/ for the out-of-vocabulary word 'three'. Comparison of the respective posterior probability density functions by evaluating their relative entropy (KL divergence), shown smoothed by 100 ms square time window as a function of time in the lower part of the figure, indicates HMM model inconsistency in the neighbourhood of the out-of-vocabulary word 'there' (The figure is adopted from1 that also gives more details).

Several additional thoughts

- Being able to identify which words are not in the lexicon of the recognizer, and being able to provide an estimate of their pronunciation, may allow for inclusion of these new words in the pronunciation dictionary, thus leading to ASR system that would be able to improve its performance as being used over time, i.e. to learn.
- The inconsistency between in-context and out-of-context probability streams does not have to indicate only the presence of unexpected lexical item but could also indicate any other inadequacy of the model.
- Further, it may also indicate corrupted input data when the in-context probability estimation could yield more reliable estimate than the unconstrained out-of-context stream.

References

1. H. Ketabdard and H. Hermansky: Identifying and dealing with unexpected words using in-context and out-of-context posterior phoneme probabilities, **IDIAP Research Report 06-68**, 2006

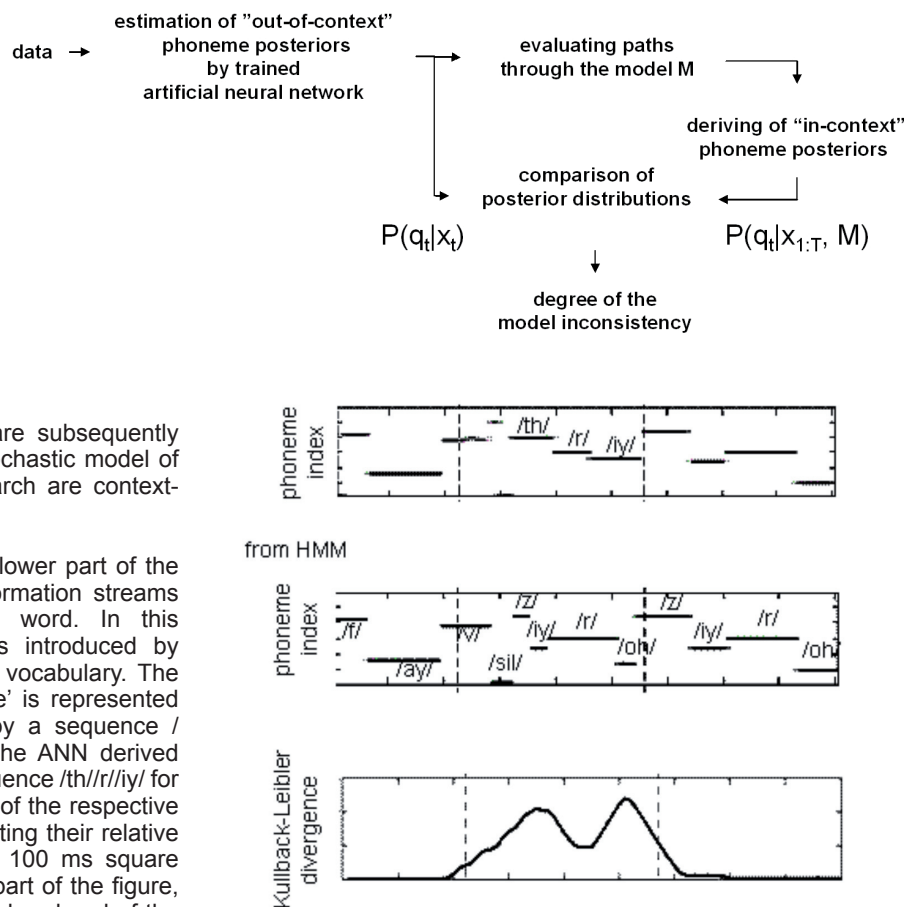


Fig. 2 Discovery of out-of-vocabulary words

Newsletter

News & Events

DIRAC Undergraduate Internship Program

DIRAC offers undergraduates a chance to perform a thesis at one of its institutions. It provides for an effective preparation for a range of research in cognitive engineering which undergraduates might attempt during their future careers. To achieve this, a thesis proposal will be carried out under supervision at the DIRAC institute which offers the proposal, and the undergraduate will be funded to cover for travel and living expenses.

Currently the DIRAC project offers the following thesis and internship positions:

- Feature Selection driven by Visual Categories
- Fast Matching for Object Recognition
- Pose estimation for mobile platforms
- Simultaneous Mapping and Localization
- Characterization of Auditory Stimuli
- Ultra fast image processing on heterogeneous multicore processors
- Robot self-localization through matching between an incoming image and a textured 3D model
- Internship in Oldenburg

Please consult the following webpage for detail on www.diracproject.org/trainingprogram/Undergraduate_Program and more information on the thesis www.diracproject.org/trainingprogram/Internships_Proposals

Telluride 2007



The three-week workshop focused on neuromorphic engineering will be held from July 1st till Saturday July 21st 2007 in Telluride, Colorado.

The goal of the workshop is to bring together young investigators and more established researcher's from academia with their counterparts in industry and national laboratories, working on both neurobiological as well as engineering aspects of sensory systems and sensory-motor integration.

The workshop is intense and active, with demonstration systems and hands-on-experience for all participants. These demonstrations are supplemented with lectures, but the focus is on the projects and interactions between participants rather than on passive exposure to knowledge.

See ine-web.org/telluride-conference-2007/telluride-overview/index.html for more information.

DIRAC's Publications

The following papers have been published since our last Newsletter. For more details, please visit our website www.diracproject.org/publications

Bayesian Linear Gaussian State Space Models for Biosignal Decomposition
Silvia Chiappa and David Barber
[Signal Processing Letters](#)

Speech Enhancement and Recognition in Meetings with an Audio-Visual Sensor Array
Hari Krishna Maganti, Daniel Gatica-Perez, and Iain McCowan
[IEEE Transactions on Audio, Speech and Language Processing](#)

A Novel Gaussian Sum Smoother for Approximate Inference in Switching Linear Dynamical Systems
David Barber
[Journal of Machine Learning Research](#), David Maxwell Chickering, MIT Press, 7, 2515-2540

Structure from Motion with Wide Circular Field of View Cameras
Branislav Micusik and Tomas Pajdla
[IEEE Transactions on Pattern Analysis and Machine Intelligence](#), IEEE, vol., pp. 1135-1149

Bayesian Linear Gaussian State Space Models for Biosignal Decomposition
Silvia Chiappa and David Barber
[Neural Information Processing Systems](#), Bernhard Schölkopf, John Platt, and Thomas Hofmann, MIT press

A Bayesian Treatment of Gain Adaptation in Switching AR-HMMs
Bertrand Mesot and David Barber
[ICASSP 2007](#)

Audiovisual Category Transfer - an electrophysiological study in rodents
Fillbrandt A, Deliano M, Ohl FW

Appetitive and Aversive Reinforcement and Their Interaction During Auditory Learning
Micheal AI, Wetzel W, Scheich H, Ohl FW

Structural Left-Right Asymmetries in Rodent Auditory Cortex
Laszcz A, Goldschmidt J, Ohl FW, Scheich H, Schildt M, Schulze H, Wetzel W, Zuschratter W, Buderger E

Disentangling the Contribution of Intracortical and Thalamo-Cortical Projections

to the Generation of Subthreshold Spectral Receptive Fields in The Auditory Cortex
Jeschke M, Deliano M, Ohl FW

Spatial and Temporal Activity Characteristics in Primary Auditory Cortex Investigated with Current Source Density Analysis under Pharmacological Manipulation
Happel MFK, Jeschke M, Deliano M, Ohl FW

Backward Masking Effects Produced by Intracortical Microstimulation in the Auditory Cortex
Engelhorn A, Dann B, Deliano M, Ohl FW

States in the Ongoing Cortical Activity Carrying Information in Discrimination Learning of Differential Electrical Stimulation
Deliano M, Scheich H, Ohl FW

Unsupervised Speech / Non-speech Detection for Automatic Speech Recognition in Meeting Rooms
Hari Krishna Maganti, Petr Motlicek, Daniel Gatica-Perez
[IEEE International Conference on Acoustics, Speech, and Signal Processing \(ICASSP\)](#)

SVM-based transfer of knowledge across robotic platforms
Jie Luo, Andrzej Pronobis, Barbara Caputo
[Proceedings of the International Conference of Vision Systems](#)

A Novel Gaussian Sum Smoother for Approximate Inference in Switching Linear Dynamical Systems
David Barber and Bertrand Mesot
[Neural Information Processing Systems](#).

Discriminant linear Processing of Time-Frequency Plane
Fabio Valente, Hynek Hermansky
[International conference on speech communication and technology](#), Interspeech, pp. 349-352

Segmentation Based Multi-Cue Integration for Object Detection
Bastian Leibe, Krystian Mikołajczyk, Bernt Schiele
[British Machine Vision Conference \(BMVC'06\)](#)

Speaker Localization for Microphone Array-Based ASR: The Effects of Accuracy on Overlapping Speech
Hari Krishna Maganti and Daniel Gatica-Perez
[Proceedings of International Conference on Multimodal Interfaces \(ICMI'06\)](#)