DAGSTUHL
REPORTS

Volume 2, Issue 6, May 2012

ISSN 2192-5283

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany.
Online available at http://www.dagstuhl.de/dagrep

Publication date
October, 2012

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Aims and Scope
The periodical *Dagstuhl Reports* documents the program and the results of Dagstuhl Seminars and Dagstuhl Perspectives Workshops.
In principal, for each Dagstuhl Seminar or Dagstuhl Perspectives Workshop a report is published that contains the following:

- an executive summary of the seminar program and the fundamental results,
- an overview of the talks given during the seminar (summarized as talk abstracts), and
- summaries from working groups (if applicable).

This basic framework can be extended by suitable contributions that are related to the program of the seminar, e.g. summaries from panel discussions or open problem sessions.

Contact
Schloss Dagstuhl – Leibniz-Zentrum für Informatik
Dagstuhl Reports, Editorial Office
Oktavie-Allee, 66687 Wadern, Germany
reports@dagstuhl.de

Report from Dagstuhl Seminar 12231

Future Internet for eHealth

Edited by

Katarzyna Wac[1], David Hausheer[2], Markus Fiedler[3], and Paolo Bonato[4]

1 **University of Geneva – Geneva, CH,** katarzyna.wac@unige.ch
2 **TU Darmstadt – Darmstadt, DE,** hausheer@ps.tu-darmstadt.de
3 **Blekinge Institute of Technology – Karlskrona, SE,** markus.fiedler@bth.se
4 **Harvard Medical School – Boston, US,** pbonato@partners.org

Abstract

From June 3-6, 2012, the Dagstuhl Seminar "Future Internet for eHealth" was held in Schloss Dagstuhl – Leibniz Center for Informatics. During this seminar, participants presented their current research, and ongoing work and open problems were discussed. The executive summary and abstracts of the talks given during the seminar are put together in this paper.

Seminar 03.–06. June, 2012 – www.dagstuhl.de/12231

1998 ACM Subject Classification C.2.1 Network Architecture and Design, H.1.2 User/Machine Systems, I.2.1 Applications and Expert Systems

Keywords and phrases Future Internet, Wearable Systems, eHealth, Aging, Chronic Care, Acute Care, Rehabilitation

Digital Object Identifier 10.4230/DagRep.2.6.1

1 Executive Summary

Katarzyna Wac
David Hausheer
Markus Fiedler
and Paolo Bonato

The paradox of life in the 21st century is that while advancements in technology and medicine enable us to live longer, our lifestyle choices increase the probability of becoming chronically ill earlier in our life and experience long-term limitations, requiring long-term social support. In 2005, 78% of European medical care spending was on chronic disease management, while 86% of deaths were due to such a disease. Yet, current health systems are designed for an acute cure rather than for a chronic care, leading to a continuous increase in healthcare costs. To achieve economically sustainable and affordable healthcare system, efficient and effective solutions are needed integrating technological advancements, and empowering the patients for better self-management, as well as healthcare teams for better decisions.

Recently, multiple initiatives have been established to shape the Internet of the future, supporting key application sectors such as healthcare, transportation, and energy, amongst others. At the same time, the emergence of next generation high bandwidth public wireless networks and miniaturized personal mobile devices have given rise to new mobile healthcare (mHealth) services. For example, highly customizable vital sign tele-monitoring of chronically ill patients can be provided based on body area networks (BAN) and mHealth applications. Such applications enable live-transmission of the data to healthcare providers, and real-time feedback to the patient, enabling her to self-manage her disease and health, respectively. Additionally, elderly people can benefit from applications that help them to stay in contact with their care teams, which are provided with valuable hints on the state of the elderly, thus in the long run facilitating economically sustainable care combined with an improved quality of life.

However, such applications do not emerge by themselves, but need to be carefully designed to support in an evolutionary way the existing healthcare workflows, fulfilling their duties at the given quality level and cost. Such a task can only be tackled in a multi-disciplinary way as it was a goal of this seminar; experts from healthcare, elderly care, insurance experts, together with experts from domains such as human-computer interaction, interactive application design, telecommunications, networking and economy teamed up to understand and support each other in designing and deploying future-proof eHealth services and applications based on Future Internet technology.

At large, the seminar addressed the following questions:

1. Which will be the key eHealth applications and services in the Future Internet?
2. Which are current and future quality requirements of eHealth applications and services?
3. Which business models are viable for future eHealth applications?
4. Which methodological support is required to design economically sustainable network-supported eHealth services?

Question 1 teamed up the participants around relevant use cases and facilitated discussions on the technical question 2 and the economical question 3, respectively. Question 4 addressed research needs from different domains and fertilized corresponding activities for advancing the topic of Future Internet for eHealth.

2 Table of Contents

3 Overview of Talks

3.1 Do We Have Smart Answers to Mary's Conundrum?

Albert Alonso (Hospital Clinic de Barcelona – Barcelona, ES)

Mary is a 79 years old woman that lives alone in her family home. She suffers from arthritis (legs and hands), has a long-standing digestive problem, a mitral valve leakage, and small mental lapses. She also suffers from technophobe ((*e.g.*, she hates ATM and care alarms). She is independent but sociable. Mary needs nutrition and mobility support, some help with shopping, bathing support and, above all, socialisation opportunities. She has no children, only a stepson and a niece – each 90 minutes away, being busy working. She's got 2 sisters, equally dependent, that do not drive. Her neighbours are elderly couples 70+, each husband with history of cancer.

The healthcare system approaches her as follows: her general practitioner seems to be uninterested, he just repeat medications and sees nothing curable; her cardiologist is only concerned with her medication compliance; her orthopedist seems to be undecided about how to help her; her neurologist is similarly unsure; her geriatrician – considers a possibility of a peace maker; her home nurse just visits her monthly and finally, her respite care is just checked annually.

So the question arises on who manages her health. Mary's health is seriously compromised if not helped holistically. The goal shall be to harmonize health care and social care and take into consideration the family and carers' own needs as factors.

After two decades of efforts in revising care provision, and piloting new modalities of care services, the number of Marys is on the rise. In spite the solutions that information and communication technologies have proposed, there are many that still see them as a doubtful joker in the pack. Does anyone have an answer to Mary? A smart one?

3.2 Adiposity, Obese Children and Adolescents and New Technology of Electronic Support

Gerald Bieber (Fraunhofer IGD – Rostock, DE)

Currently there is a trend in our society of changing behaviors, influenced by ubiquitous availability of computer workstations, television, computer games and Internet, as well as convenience of 'fast food' diet and high carbohydrate foods. Additionally, the physical activity of the humans is not sufficient, and as a result, the human weigh is too high on average. Today, only 1/3 of the population has normal weight. These all factors increase the risk of suffering from a disease; in the future a high increase of type 2 diabetes is expected. In Germany alone, already over 6 million of people are diabetic. In comparison, in 1960 only 0.6% of the population got the diabetes disease, yet currently, almost 10% of the Germans diabetic. The disease rate is expected to double again for the next 10 years. Diabetes is associated with dialysis, blindness or amputations; it is one of the most expensive and cruel diseases. The trend of the increasing disease rate can be reversed only by a change of lifestyle. Overweight children and adolescents usually retain their high weight when they become adults, but children can be educated to change their unhealthy behaviour.

In order to achieve a sustained motivation for lifestyle change, it is necessary to promote and practice health physical activity. Therefore, it is useful to assess the physical activity objectively and unobtrusively and use it in an individual treatment plan. For the treatment, it is necessary not only to detect whether the patient is active or passive, but what activity was performed over a period of time in chronological order. The set of identified activities allows an estimation of the energy expenditure and needed therapeutic support. Moreover, a motion-dependent treatments can be performed. The use of accelerometers to monitor physical activity by new algorithms for pattern recognition allows the identification of various physical activities. Jumping, running, walking, cycling or driving a car can be detected, additionally to sitting and resting. This can be done simply by wearing a standard smart phone, or a smart watch in combination with a special application deployed on it.

In our research we have developed the DiaTrace-technology, which transforms a smart phone to a mobile assistant for physical activity and food assessment. The mobile assistant rewards the user with electronic medals when he/she performs sufficient physical activity. Together with MEDIGREIF-Inselklinik, the DiaTrace technology was used for the treatment of obese children and adolescents and provides new insights in term of objective and subjective perception of food intake and physical activity. Furthermore, this technology can be used to design new treatment concepts, *e.g.*, for physical activity dependent medication.

3.3 Applied Health Technology and Old Age (in a Global Context)

Doris M. Bohman (Blekinge Institute of Technology – Karlskrona, SE)

My main interest is grounded in my ethnographic research in South Africa focusing on older people, where the use of mobile phones to a large extent is a natural part of their daily life. From this perspective, my interest focuses on mobile technology and health (*i.e.*, applied health technology) and on strategies how to increase the involvement of the potential users in design innovative solutions – the quadruple helix model. The main question arise – what can we learn from Africa when introducing and developing mobile technology based solutions in health care for older individuals in an European context.

The first objective to understand is what are the direct and indirect health benefits for older individuals and their families of using mobile technology/mobile phones in a context as Sweden/Europe and in a context as South Africa/Africa. Secondly, how does mobile technology/ mobile phones influence the intergenerational relations, mainly the relations between the older and the younger generations in relation to health services. Furthermore, which are the patterns of older individuals' use of mobile phones in health service and health promotion in Europe versus Africa and finally, which are the potential factors for formal health sector integration of mobile technology/mobile phones.

3.4 A Simplified Approach to the Design of Wearable Systems

Paolo Bonato (Harvard Medical School – Boston, US)

The development of wearable systems for mobile health (i.e. monitoring patients in the home and community settings) has mostly occurred in a vacuum, namely in engineering schools and companies without proper interaction between engineers and clinical personnel. Consequently, the field has witnessed a great deal of emphasis on the development of new technologies rather than on the design of solutions to well-defined clinical problems. This bias toward the development of new technologies has led to the incorrect assumption that all wearable systems for mobile health should allow clinicians to monitor patients continuously and ubiquitously. This is not correct. In fact, the majority of the clinical applications that fall within the umbrella of mobile health do not require continuous and ubiquitous patient monitoring, but rather often they 'only' require monitoring individuals for a few days at certain points in time (e.g. every few months) while they undergo a clinical intervention. To achieve the goal of delivering wearable systems that address real-world problems, it is essential that further developments in the field of wearable technology be put in the context of the specific clinical problems that one intends to address. Herein, we propose an 'oversimplified' approach to the design of wearable systems for mobile health that is based on focusing on three questions. (1) What would clinicians do with the data gathered using the wearable system? (2) How quickly would clinicians have to take action if wearable sensor data indicate that a clinical intervention is needed? (3) How long would clinicians need to monitor each individual to gather the clinical information of interest? Although these questions might come across as trivial, we argue that they capture fundamental aspects of the design of wearable systems that are critical to assess what technological solutions would be ideal on a case-by-case basis. For instance, proper understanding of the objectives of the data analysis procedures to be set in place would indicate whether real-time processing is required with significant implications from an engineering standpoint. If one intends to monitor patients at risk for severe arrhythmias, then the system to be developed would need to provide the ability of processing electrocardiographic data in real time and of relaying an alarm message to a caregiver when needed. If instead one intends to monitor patients with Parkinson's disease to titrate their medications, the data could be stored in data logging unit potentially worn by the patient and the data could be analyzed off line. It is clear that applications with such different requirements would lead to very different hardware and software implementations.

3.5 User Driven Service Design and the Future Internet for eHealth

Sara Eriksen (Blekinge Institute of Technology – Karlskrona, SE)

With an interdisciplinary background in Informatics and Human Work Science, I have been inspired by the Scandinavian tradition of Participatory Design (PD) as well as a near-sighted focus on everyday work practice.

The questions arise on how do people actually go about getting their work done, how does technology support everyday work, and when it does not, and furthermore, how can users' complaints be fed back in to the further design and development of technology to reduce frustration and constant work around.

Participatory design, I would claim, is more radical than User Centred Design. It lets the users have a say and I think this would be especially relevant in eHealth, where many of the users may be sick, elderly or have functional disabilities. However, I have found that PD is not all that radical anymore. In the most unexpected situations, I keep finding myself eye to eye with innovative users with great ideas. Not only do users have great ideas, they have attitudes. They expect to be able to do something useful with technology, and keep up with their friends and families on Facebook while they are at it. The question arises why not helping them to do so.

Maybe a large part of the problem when approaching eHealth is that we are not expecting this kind of initiative from what we perceive as marginalized user groups such as elderly, sick people and people with functional disabilities, who seem helpless. The question arises how can we more deliberately cultivate multiple methodological perspectives in technology design and development, and allow for a wider range of different user roles in our R&D projects with the aim of co-designing the Future Internet for eHealth.

3.6 FIT for eHealth – Show-maker or Show-stopper?

Markus Fiedler (Blekinge Institute of Technology – Karlskrona, SE)

Emerging Future Internet Technology (FIT) appears to be enabler and orchestrator, *i.e.*, some kind of show-maker, for a plethora of new apps and services in the eHealth domain. For instance, virtualisation/clouds and cooperative communications are believed to support the deployment of eye-catching eHealth apps and services in cost- and energy-efficient ways. However, user adoption is at stake if newly introduced apps and services do not (over-)fulfil requirements and expectations put by their (typically rather picky) users. Thus, it is time to identify potential show-stoppers that might hamper the successful dissemination of new eHealth apps and services to be built on top of FIT. Following up on this inventory, potential consequences for the design choices for FIT have to be addressed in order to yield an optimal user experience of eHealth services and thus success in the market.

3.7 eHealth to Fit

Geraldine Fitzpatrick (TU Wien – Vienna, AT)

Designing effective health IT has always been challenging as our own and others' experiences with electronic health records and related technologies show. New healthcare models around self-care, aged care and chronic disease management make this even more challenging as care moves out of clinical settings into patients' hands and people's homes. eHealth technology is seen as the critical enabler for this move. However the challenges of moving care into the home however should not be under-estimated. It is not just a case of moving technology from one setting to another, or creating more mobile apps. Rather, health and care embedded into everyday life (rather than in a discrete clinical encounter) presents a radically different space for health IT and for health practice. It fundamentally challenges our conceptualisations of patient, clinician, care and home and sets up new challenges for design and evaluation.

In previous research we have used mobile, wireless and sensor-based technologies, also tangible devices, to design for self-motivated care at home; we have also conducted various studies with older people and around AAL-home monitoring. From these experiences, plus the work others, we can identify a number of design challenges driven by the following issues: (1) whose voices are heard; (2) what implicit values or norms are we inscribing by design; (3) how do we design technologies to fit into lives and homes in a way that can both meet the needs of the health care system while also delivering value to the people who need to use the technologies; (4) how do we design for integration into the diverse spaces, routines and social contexts of everyday life; and (5) how do we design for the consequences of patient empowerment, shifting to a model of collaborative control and interpretation.

There are also complex evaluation challenges. In the health domain, clinical, randomized control trials are the standard for 'evidence' and they rely on relatively stable and controllable environments. Yet the implementations of many eHealth technologies are being appropriated in continually on-going ways and often specific to situated contexts that make it difficult to run such trials. Yet 'evidence' is still needed for large-scale systemic changes in healthcare delivery. We are still researching, for example, what new evaluation methods will allow us to orient to both process and outcomes equally, and how do we understand dynamic situated appropriation that will enable us to more systematically evolve new models of care and care processes. We need to understand and have a consent on what are measuring, what constitutes 'evidence' and for whom.

Who we think eHealth technologies are for, how we design for them, and what we evaluate, matters if we are to design technologies that really fit into people's lives and deliver value to both patients, their care providers and the healthcare system.

3.8 eHealth will Revolutionize Health Systems in Africa!

Caroline Franck (Hôpitaux Universitaires – Geneva, CH)

In 2000, Member States of the United Nations committed to the attainment of a set of quantifiable Millennium Development Goals for redressing poverty and improving lives by the year 2015. Even though the situation has improved, it is very unlikely in sub-Saharan Africa that the health – related goals will be achieved. Factors contributing to the failure of these goals are a shortage of trained health care professionals, especially in rural areas, and a lack of education. Both of these aspects are correlated with the following factors: limited resources, inequity and inefficiency.

The Internet and multimedia technologies are reshaping the way knowledge and services are delivered, due to the fact that it can be offered at relatively low cost and provide real-time distribution. Additionally e-learning creates a learner-centred, self-paced learning environment, which makes e-learning a real alternative to traditional classroom learning. Information Technology enables us to provide access to education, knowledge and research to the poor and remote villages on a large scale. Decreases in technology and infrastructure costs facilitate the use of these applications in low-resource settings like sub-Saharan Africa. For example, mini-VSAT connections and portable ultrasonography devices permit tele-medical applications in rural sub-Saharan regions. Centres of excellence and tele-consultations may decrease shortage of expertise in rural areas through providing remote access to specialists and empowering health care professionals, like for example midwives, to perform obstetrical evaluations.

Further on, the introduction of Health Information Systems will facilitate transparency and evidence base, which will largely contribute to enhancing accountability at all levels. They will also empower lower levels in the care or management pyramids, whose contribution will become increasingly visible. Such major developments in a national health policy constitute per se a major reform for health sectors that would take this direction.

3.9 Serious Games for Health

Stefan Goebel (TU Darmstadt – Darmstadt, DE)

"We do not stop playing when we get old, we just get old when we stop playing" (George Bernard Shaw)

In the health(care) application domain, a shift to prevention should take place. This covers both primary prevention – in the responsibility of individuals/private (and healthy) person – and secondary prevention in charge of the healthcare system (practically: healthcare insurance companies). For the private sector, motivating factors and incentives are necessary to convince individuals to do sports, exercise, change nutrition type or composition and have a healthy lifestyle at large, in order to avoid obesity, cardio problems or other diseases.

Games co-called respectively Serious Games – games being more than pure entertainment, with an additional overall purpose (*e.g.*, to improve the vital status) – provide an excellent opportunity. With respect to secondary prevention, healthcare insurance companies need some 'proof of concept' that game-based approaches do create the proposed (medical) effects. For that, comprehensive evaluation studies are necessary. Our research tackles these issues in a number of 'Games for Health' project in cooperation with technology providers, and stakeholders such as hospitals.

3.10 How To Get Technology, Healthcare & Policy Integrated?

Nick Guldemond (TU Delft – Delft, NL)

Worldwide, healthcare systems are considered unsustainable due an increase demand of care and associated rise of healthcare costs. Aging of societies and the growth of populations with chronic conditions makes a paradigm shift of western healthcare systems necessarily. The Chronic Care Model provides a framework for healthcare change, including a prominent role of the community and patients' self-management. Technology holds the promise to support the premise of the Chronic Care Model. How can technology be linked to the Chronic Care Model and what are the implications for policy, remains a question to be tackled.

3.11 People-centric eHealth

Mattia Gustarini (University of Geneva – Geneva, CH)

I research in the field of ubiquitous computing. So far, along my research I had the opportunity to learn upon the concept of eHealth, however mostly from the point of view of quality of service (QoS). For advanced solutions for QoS in eHealth, I have decided to put the user in the centre of my reasoning. Even before service provision, we as designers need to be clear what should happen around the user, and what are the devices at our disposal to implement eHealth keeping the user truly at the centre of our reasoning. By answering these questions, I have decided to use a smartphone as an eHealth platform, and a lot of research questions popped up. Firstly, we need to understand how to transform the smartphone in a credible, reliable sensor platform for eHealth applications, employing its internal, as well as external sensors. Secondly, how we can reach this goal without interfering with the 'normal' use of smartphone in terms of its efficiency, performance and resources management. Additionally, we need to understand how to involve the user to incentivize the use of eHealth application on smartphone, how to create a collaboration between all users and how to involve them in the collection of data that can be useful and usable for everyone. We shall research issues related to sharing and privacy management, as well as understand what to do with the data collected.

3.12 Promoting Physical Activity with Mobile Devices

Jody Hausmann (University of Geneva – Geneva, CH)

My main research question is how can we deepen our understanding of own physical activity status – understand if we are sedentary or active, and if we are sedentary, what are the ways of getting enough motivation to stay active. Individual factors influence the modality over which the motivation shall be sparkled – being a textual/visual feedback or a meeting with a coach. A user shall have an easy, affordable and convenient way to track every day performance. Is it possible to provide it with a device that is non obtrusive, reliable and does not require many many manipulations – *e.g.*, a smartphone.

Our current research aims to answer these kind of questions. We assess the smartphone's feasibility to unobtrusively, continuously and in real-time track its user's physical activity intensity and quantity per day based on the World Health Organization (WHO) recommendations. Activity Level Estimator (ALE) is an Android OS application, which analyses how much time the user spends on walking and classifies it alone its intensity represented by a level (*i.e.*, low, moderate and vigorous). We conducted a study with 12 participants and we have calibrated and assessed the accuracy of ALE against a Gas Analysis System, a "gold standard". Our results showed an overall average accuracy of 90.06% per minute for walking activity. The next steps in this research will be to help the user to stay motivated.

3.13 eHealth Means Disruption and Change!

Rainer Herzog (Siemens Enterprise Communications – München, DE)

This talk addresses the disruption and change needed to be embraced to support eHealth in the Future Internet. The actors driving the disruption and change are healthcare providers, payers and patients. The issues to be considered are (1) mobility of patients and practitioners; (2) scarcity of care resources and competences; (3) need for intuitive user interfaces; (4) need for a team-based future care; (5) need for connected service platforms for healthcare provision, as well as (6) resources needed to educate an informed and empowered patient.

3.14 Applications of Virtual and Augmented Reality in Healthcare

Hannes Kaufmann (TU Wien – Vienna, AT)

Augmented Reality (AR) combines the real and the virtual and is interactive in real-time. In addition, virtual and real objects are registered in 3D space to pinpoint their positions for precise overlays. AR has been in the news a lot in recent years, mainly because of AR Apps on smartphones. Augmented Reality has a lot more to offer though. Hardware setups are versatile and range from mobile devices to immersive lab installations. Just as versatile are the application areas ranging from industrial uses (*e.g.*, automotive, manufacturing), training and education, modelling (architectural planning), design, visualization (*e.g.*, scientific, medical, and information), entertainment and more recently, the widening spectrum of possibilities in the medical domain, rehabilitation and therapy. An example of the latter is an EU FP7 ICT project on virtual rehabilitation – PLAYMANCER. It focused on developing serious games for cognitive behavioural therapy – specifically for patients with eating disorders and pathological gambling, and on serious games for the rehabilitation of chronic back pain patients. Rehabilitation for chronic pain follows a multidisciplinary approach, which despite the effort, often lacks long term success. Patients fail to translate skills learned in therapy to everyday life. In order to encourage continuous training and ensure impact at a wider scale when it comes to "Active Ageing", technology can and should be used to motivate people to exercise at home.

Movement data can provide medical experts with useful information regarding patients' home training, *e.g.*, duration, intensity and correctness. As an alternative home motion capture system, the low cost Microsoft Kinect was compared to an 8-camera precise optical motion capture system that we research. The results show that overall, the Kinect performs surprisingly well. It correctly captured some of the exercises used within the serious game. It cannot measure all required parameters (*e.g.*, head rotations cannot be detected) and lacks accuracy required for others (*e.g.*, velocities of hand/arm movement). For clinical evaluation such a device cannot be recommended due to large errors. However, for health related home use, a Kinect like depth camera can be used as a full body input device for serious games and other rehab or health programs that do not need supervision.

To reach the general public, using mobile technologies for health care purposes, seems to have the broadest impact. MIT's CATRA and EyeNETRA projects are suitable examples of these. Augmented Reality is predestined to contribute in these areas as well. Qualcomm recently announced the Tricorder X challenge to develop a mobile diagnostic tool based on a smartphone, making reliable health diagnoses available directly to "health consumers" in their homes.

My main interest is to find and develop applications of VR and AR that benefit people. What I have to offer is an extensive experience in tracking technologies, sensor fusion, mobile AR, immersive VR, motion capture, 3D displays and more. Currently, health related research proposals are under review for patients with mild dementia and Alzheimer, Parkinson patients and precise localization of blind people outdoors.

Regarding mobile and upcoming technologies for elderly people, excellent usability must be of uttermost importance. We will see a variety of new approaches and health technologies emerging until the end of this decade. With extensive funding opportunities on the horizon (*e.g.*, EU's HORIZON 2020), we should take our chances to develop and spread successful medical applications based on virtual, augmented and mixed reality.

3.15 Paradigm Shift in Patient Care Approach Could be Promoted by mHealth Usage

Zviad Kirtava (Tbilisi State Medical University – Tibilisi, GE)

In this talk we firstly discuss if eHealth will be a unifying tool between developed and developing countries or new frontier of digital divide. Secondly, we discuss if the mobile telemedicine and mobile tele-monitoring will be providing new path from inpatient-oriented care-centred model to outpatient-oriented prevention-centred one. Additionally, we address the issue of mobile learning, driven by cost of module developments and potentially propelled by an inter-academia to business cooperation, which would enable international students access high-quality training materials. Furthermore, we consider the possible roles of physicians and healthcare institutions, being supporters, observers or opponents of eHealth. One of the important questions is, if the business of healthcare would be threatened by universal access to patient data and competition.

3.16 eHealth Bike Solution for Cardiovascular and Pulmonary Rehabilitation

Willy Kostucki (Clinique Antoine Depage – Brussels, BE)

Physical inactivity is an independent risk factor for cardiovascular diseases, and sport is among the few non-controversial treatments to reduce the incidence of cardiovascular diseases. Improving the physical condition is as effective as the control of other risk factors like smoking, hypertension, diabetes, obesity, stress, and it competes with invasive treatment (stent) in several subgroups and this regardless to sex or age of the patient.

The opportunity resides on a paradox, *i.e.*, the more patients are at risk, the more intensive and well managed must be cardiac rehabilitation processes. Nevertheless, this rehabilitation is difficult to achieve, especially for patients that are physically limited due to their illness and eventual co-morbidities, as well as psychologically weakened, and for whose who are naturally afraid of physical exercise, which they tend to find boring. The other major hurdle to a "winning" rehabilitation is the "compliance" by the patient to the prescribed exercises.

The uniqueness of the proposed eHealth Bike Solution is the loop between the patient and the cardiologist that will continuously improve the acuity of the prescription and the efficiency of the training and thus the patient's health. Loop after loop. The eHealth Bike has also so many features never to be found on any even top market available fitness bikes today: Pad-driven application, blood-pressure monitoring, heart rate monitoring, unique ergonomic design, seat motorization (for the elderly), weight sensor, and so on. It tends to be more a health station rather than a simple bike. As for an accessibility, everything is web-based.

Searching investors is a bottleneck for those kinds of projects. Thus the question remains which business models are viable for future eHealth applications. Since cardiology is also a highly technical medical specialization with already many applications available, additional issue relates to the key eHealth applications and services in the Future Internet. I personally believe in the integration of the specializations instead of the present splitting in today's techno-medicine.

3.17 Assisted Living in a Smart City

Ernoe Kovacs (NEC Laboratories Europe – Heidelberg, DE)

The current Internet is a very successful world-wide infrastructure which enables different parties to communicate and offer and consume services. The success of the Internet is largely based on the adoption of two "hourglass patterns".

The first hourglass pattern is around the TCP/IP protocol stack in which the many applications are using the TCP/IP network layer (the waist of the hourglass) and in which that layer also hides the complexity and evolution on the lower networking layer. The big advantage of this is that the networking and application layer could evolve independently resulting in very dynamic progress in both areas. The second hourglass pattern is around Web technologies. Web technologies are enabling an explosion in services. They are (again) forming a commonly used waist enabling the technologies below it (*e.g.*, database, OS, scripting languages, middleware system, network) to evolve independently of the applications that are using it. The results are a worldwide market for services such as eCommerce, social networks, search, video services, information access, and many more.

Research in the Future Internet is driven by two aspects. Firstly, we are experiencing that the current Internet is reaching its limits. This is very obvious and observed with respect to the availability of addresses, but it can be also easily seen in many other areas, being it network management, content distribution, mobility management, and so on. Secondly, drive arises from the demand for better features in the network itself. With advanced features, better services can be build. For example, network virtualization and isolation will enable the creation of world-wide services which do not have to be aware, hardened and adapted to the fluctuating and harsh network conditions of today's Internet. Problems can be managed and solved in the virtualization layer, while the services can focus on providing their services. Clouds are promising endless computing resources, but need to be able to provide their data processing for each user as needed. The Internet-of-Things promises to deliver real-world information in aggregated and well-prepared forms to users and services – from around the world and ultimately in an open system. The search service is on for a new hour glass pattern that will enable to create the promising services of the future and shield them from the complexity of network, virtualisation, cloud, service orchestration, Internet-of-Things, and other technology evolution.

Evolution and progress in the eHealth sector is mirroring these developments and is also driving some advancement. First successes in eHealth were in the management of the overall medical and non-medical processes in medical facilities or across organisations. This goes hand-in-hand with the early phases of the Internet enabling world-wide and easy communication. Nowadays access to medical information, as well as to eHealth services is common practice. Serious work is now conducted with a goal of embedding sensor information into eHealth. The need for security and protection of privacy is driving respective research and would benefit from, *e.g.*, virtualization of networks.

The example of Ambient Assisted Living (AAL) in Smart Cities is illustrating this. Initially, AAL was applied in the area of smart homes or as a mobile services. These are isolated systems, usually under the control of a single authority. Technology is now equipping buildings, streets or places to become smart environments, and very soon complete cities will become a platform of a Smart Cities. In such environments, many providers of data, computing resources, knowledge processes, and useful services need to jointly work together for providing a open AAL environment.

This clearly calls for a Future Internet in which many players can be free and open, but also controlled and if needed, secured and isolated work together. Research has started here. The FI PPP project FI-Ware is searching for a configurable service platform for the Future Internet. It can serve as an example for first steps towards the new hourglass patter for the Future Internet, also benefiting advanced eHealth systems.

3.18 Data and Process Interoperability in eHealth

Lenka Lhotska (Czech Technical University – Prague, CZ)

Health records have been usually kept in the paper form. Doctors are still used to use it and consider it a good practice, as they are versatile and complete. Usual practice is to make human operators (usually nurses) enter the selected information from paper documentation into a Hospital Information System (HIS) during a night shift, which also raises the error ratio of the operator. The paper documentation usually contains free text, printed documents, filled forms and paper reports (*e.g.*, signals, results of diagnostics, analyses). Usually there is no backup of the paper documentation and the task of digitalization into Electronic Health Record (EHR) is inconceivable in such extent. Legislative defines the extent of a health record, therefore it is rather clear what belongs to the documentation. Patient has the right to look in the documentation and request a copy.

However, there is no usable standard for fully compliant EHR, although the health information is often saved in some form of a HIS (that are usually mutually incompatible).

There is no standard for signal data (*e.g.*, ECG, EEG) and the data are often stored in proprietary formats. The HL7 has been adopted in several countries as standard. However, the standard content of EHR is not yet well defined.

Although the legislation defines a mandatory back-up period for health records, there is no defined need on the readability of the medium. It is not clear whether a CD/DVD (or even diskette) is readable after 50 years, or whether devices for reading the media would be available. The question is also whether we are able to read 20 years old data/compression format. It is undefined whether the data should be archived as-is or stored with a transaction log (describing who/when/how changed any information, who approved the change) that cannot be modified (electronically signed). Another question is whether the data are stored securely (no stealth alternation can be made, consistency checks are performed). An important question is whether there exists an easy way to make the data anonymous for medical research, statistical purposes, business decisions, and so on. The answer for the latter question is that obviously no, as even the electronic documentation contains copied/-pasted text documents containing unstructured sensitive data. It is also important to ensure availability of data codebooks (such as ICD) that were up-to-date for each data in data-life timeline as the recent codebooks might not contain information referenced by archived data.

Actually there are insufficient legislative regulations on the health documentation with subsequent electronic health documentation legislative. Doctors are obliged to let the concerned party consult the documentation and must provide a copy on any data media. The problem is that the doctor usually inserts internal notes in the documentation which he does not want to reveal to public for multiple reasons – they entered subjective notes, the documentation contains information from a 3rd party as the patient was lying, due to the fear of prosecution, and so on. A question is whether the data from, *i.e.*, patient monitors (signals, trends, curves, etc.) belong to the hospital or to the patient, *i.e.*, whether the medical facility has the right to use data for any form of research. It is unclear whether the medical facility has even the right to obtain the raw data (without any pre-processing) from the appliance. It is not always easy to get the data in any readable format from the device manufacturer. And even there is a possibility to retrieve the data, it is a question whether it can be used in a clinical trial together with data from another appliance, as usually a filtering and pre-processing of data is performed that might not be comparable. It is crucial to have (also) access to the raw data.

There exits an unordered list of usual reasons that block the adoption and implementation of interoperable EHR. These reasons are not generally applicable to all the doctors nor IT professionals and should not be taken literally: (1) Doctors have been healing patients without any PC. They do not need any PC to perform medicine; (2) IT professionals ask the PC (EHR) to do things that are not present in the paper form (mainly the strutcuralization). Doctors often claim that about 40% of physicians do not have a PC. This is however questionable due to need of sending electronic reports to the health insurance companies; (3) Doctors do not want to be restricted by any structuralization or ontology when taking quick notes. They prefer to use free text. They also need to enter subjective and personal notes into the documentation that should not be seen by anybody outside the medical professionals; (4) People are in general unwillingly accept changes – the changes usually bring higher administrative load and the outcome is unclear; (5) Doctors are afraid of that with complete documentation it will be tougher to defend against liability claims; (6) In the discussions of IT professionals vs. doctors there is usually nobody accepted as a respected moderator (person) so the discussion goes deep into unimportant details and gets usually stuck for long time. It is extremely complicated to find any consensus in a reasonable time.

There is currently no fully implemented interoperable EHR standard. The E-prescription is running in a pilot stage now. Integration of the medical information is a problem: many standards, medical facilities are forced to buy cheap appliances with limited support. Medical facilities do not demand data in readable format for further processing. We are losing important data for long-term studies; also the sharing of chronically ill patient information is limited. There is no successful intent to adopt any working solutions available. Usually, there are attempts to find single (incompatible) solution (and reinvent a wheel). Integration is possible, however no one is forcing the medical facilities, HW/SW manufacturers to use interoperable standards. There must exist joint coordinated effort to unify the standards and to use the interoperable standards.

3.19 Wireless Multimedia Support for eHealth: The Status and the Way Forward

Maria Martini (Kingston University – Kingston upon Thames, GB)

Recent pervasive multimedia healthcare applications include hospital tele-consultation, health monitoring, intelligent emergency management systems, pervasive healthcare data access, ubiquitous mobile telemedicine, assisted living and remote surgery. Research projects are addressing such applications and the suitability of the current Internet to support these. An example at European level is the CONCERTO EU FP7 project, addressing content and context aware delivery for interactive multimedia healthcare applications. From a more strategic and interdisciplinary point of view, the eMobility European Technology Platform recently delivered a Strategic Applications Agenda, discussing the future key applications and challenges in three main domains, including Health and eInclusion. Recent challenges and future research directions of comprehensive wireless multimedia health monitoring and management will be discussed, including the concepts of context awareness, context quality issues, security and reliability, and autonomous and adaptable operation.

3.20 eHealth System for Diabetes Prevention, Monitoring and Treatment

Goran Martinovic (University of Osijek – Osijek, HR)

Diabetes, a chronic, incurable systemic metabolic disorder characterized by hyperglycemia, interferes with the exchange of carbohydrates, fats, and proteins in the body, which results in typical symptoms, and after a long period of time, it affects the structure of most organs and organ systems – the cardiovascular and the nervous systems. The number of people with diabetes is increasing steadily, even the number of people in their adolescence. This is a consequence of modern lifestyle, *i.e.*, an increase in the number of etiological factors of diabetes among which obesity is emphasized in particular. In the last thirty years, the number of diabetes sufferers has risen from about 160 to about 350 million people. Every year, about 3 million people die from diabetes-related causes worldwide. In Croatia, type 2 diabetes affects about 320,000 people (about 7% of the total population). The actual situation is probably much worse, since many do not feel sick and they do not have symptoms, hence get undiagnosed. The disease has a long asymptomatic phase, and at the time of clinical diagnosis, the disease is usually present for 7-12 years. About 30% of microvascular complications (mostly retinopathy, nephropathy and neuropathy) have already been developed by then. A patient with diabetes has also developed other risks for the development of microvascular complications, *i.e.*, obesity, hyperlipidemia, hypertension, smoking, sedentary lifestyle and the like. According to this, early disease detection by determining the blood glucose level is difficult, and early diagnosis through symptoms can significantly aid healing and reduce its complications. Also, early detection, monitoring and prevention of behavioural patterns that can cause a disease or accelerate the development of the disease, can alleviate or even prevent the disease itself.

Pervasive and ubiquitous information and communication technologies have been offering considerable support in the process of detection, treatment as well as monitoring and facilitation of diseases in general, including diabetes. However, in the ICT sector, patients and physicians long for an integrated and unified cooperative eHealth environment that would provide a lot more options to everyone. Wired and wireless networking at a relatively high speed, wireless sensor networks, body networks, web, mobile technologies and service-oriented computing environment, can increase support to measurement, collection, storage and analysis of numerous data. Some of these data are collected and stored off-line during medical check-ups in patients, by means of various questionnaires and from databases (*e.g.*, clinical data, symptomatic information, information about patient's lifestyles, certain laboratory parameters), while some are directly or indirectly measured, assessed, stored and analysed in real time (*e.g.*, glucose level, level of ketoacidosis, blood fat level, blood pressure values). Typically, on the basis of available data values, computer systems enable symptomatic and individual reaction often neglecting their mutual connection and interdependence. This connection and integration of knowledge is possible either off-line by using a PC, or it is done by a physician, which is not possible for many patients/data. In contrast to this, intelligent computer analysis of a large number of data collected for a great number of people/potential patients who live in different living conditions and behave differently, can provide the following: (1) multi-criteria risk assessment of developing diabetes based upon off-line (archived) data and data collected in real time; (2) learning from data and decision making, as well as forming and pinpointing habits and patterns of risky behaviour (*e.g.*, on the basis of glucose value, start an insulin pump or remind of the therapy; on the basis of the step counter, remind the patient (or a potential patient) of a physical activity, and related regular energy intake via an adequate composition of meals); (3) making a decision how to respond in case of life-threatening conditions of the patient, *e.g.*, a decreased blood glucose levels below permissible levels, a high level of ketons in the blood, deterioration of the condition due to diabetes-related deterioration and vice versa.

Data storage (preferably in the context of events, *i.e.*, respecting the domain, semantic and ontological principles), analysis, learning and decision making (based upon multi-criteria intelligent procedures, pattern recognition, self-study) require significant computing power and storage capacities (mostly on demand), inclusion of various sensor, mobile and wireless components, user interfaces and applications available and affordable to a wide range of average users, support at the level of social networks, support to cooperation, availability to a patient, a physician, but also as part of the public health and emergency and rescue system. With all these technologies, cloud computing seems to be the best integration and service-oriented solution. It assumes a high level of access security and data confidentiality at the level of transactions, a high level of environment availability, application of the principles of autonomic computing, as well as an acceptable business and financial model that will not jeopardize the operation and accessibility of such future eHealth systems.

3.21 Context is Everything: Using Sensors to Support Self and Collective Understanding

Dave Marvit (Fujitsu Labs of America Inc. – Sunnyvale, US)

Technology, and specifically sensing technology, is often viewed as supportive of reductionist thinking, after the statement "if we have all of the details at the lowest level then we will understand everything." While the reductionist fallacy is well understood, it is not as widely recognized that sensing technology may be the key to enabling systems thinking. Because all data is ultimately health data (since everything effects our mental and physical states), the use of sensing technology in the name of health offers a great opportunity to build systems that support contextual understanding. Work born of this philosophy has begun and will be (briefly) described in the examples including a generalized platform for mobile, continuous, multi-sensor monitoring, analysis, and service delivery – and some initial services based upon this platform.

3.22 Integrated Systems for Measuring and Intervening on the Exposome

Kevin Patrick (University of California – San Diego, US)

Advances in science and engineering have produced a world with increasingly ubiquitous information and communication technologies. Mobile, sensing, networking and computing technologies support the collection, analysis and use of increasing amounts of health-related data at every level: the genome, microbiome, biological and physiological parameters, and the exposome, the sum total of life-course health behaviours, environmental exposures (both healthy and unhealthy), stress, and social and economic influences on health. Traditional methods of addressing environmental causes of disease usually focus on a single element (*e.g.*, asbestos and mesothelioma) or risk factor (*e.g.*, alcohol and esophageal cancer; tobacco and lung cancer). The exposome is much more complex, multi-layered and multidimensional than this but we have been limited in our ability to measure it because we have not had sufficient technical infrastructure to do so. Mobile sensing, networking and computing technologies allow us to better understand the multi-layered and interconnected systems important to human health. Importantly, the same mobile and computer technologies that support the collection and analysis of these data also support increasingly "intelligent" systems-based interventions to improve both individual and population health.

The key eHealth applications of the future will leverage near real-time (if not real time) assessment and analysis of the exposome to provide individuals, health professionals (both medical and public health) and policy makers with actionable information to improve health. Quality needs for these applications and services will relate to fault tolerance and self-healing performance at every level: the device/sensor, the mobile phone or other personal computing device, networks, the cloud and other connected systems important to health.

Business models will vary from setting to setting depending on reward structures for health and illness-based care. In the US we have unique problems in this regard with the challenge being to move from a medical-industrial complex based on monetizing illness to a new model of monetizing wellness. Disrupting this process may well come from outside the traditional health care industry.

Methodological support is needed in all aspects of the health supply chain: improved design of sensors, mobile operating systems and networks; improved approaches to power management; human-computer interface issues; persuasive design; newer methods of data-fusion and analysis; and work-flow issues for those involved in health, to name a few.

3.23 Redesigning health in Europe for 2020

Terje Peetso (ICT for Health Unit, DG INFSO, European Commission – Brussels, BE)

This talk addresses the following issues to be addressed to support eHealth in the Future Internet. On one hand these issues include ownership, accessibility, and safety, as well as security related to health related information. On the other hand, health literacy, interoperability, and evidence are to be considered. Finally, the legislative framework and need for common standards need to be taken into account as well.

3.24 Fully Personalized, Content & Context Aware eHealth Services: Present, Future or Fiction?

Pawel Swiatek (Wroclaw University of Technology – Wroclaw, PL)

One of the main motivations for designing new architectures for the Future Internet is to meet challenges imposed on the ICT infrastructure by new applications. One of the most demanding application types are eHealth ones, where quality of the delivered services is often critical to the life of their users. The ICT challenges include among others:

1. Content awareness, meaning the sensitivity of data processing and transmission methods to the content being delivered to the end-user. Content awareness may emerge in: different processing of various data streams (*i.e.*, video encoding or sensor data encryption) and different forwarding methods (*e.g.*, routing) for various streams.

2. Context awareness consisting of different treatment (in terms of forwarding and processing methods) of traffic depending on the particular use-case scenario of application generating this traffic. Context may be connected for example with the type of networking device used by a user or users geographical localization.

3. User awareness understood as personalization of services delivered to end-user. Personalization is achieved by means of proper choice of data processing and transmission

methods according to functional and non-functional requirements stated by the user. Users requirements may be formulated explicitly or be a result of automatic recommendation which is based on the history of the application usage.

4. Sensor networks and applications covering not only eHealth and telemedicine scenarios but also such applications as: smart energy metering, vehicle networks, intelligent building infrastructure, etc. Each particular telemetry application involves specific types of data processing methods and transmission of large number of small portions of data often requiring real-time or near real-time end-to-end performance.

Augmentation of the current Internet architecture with the abovementioned functionalities will fulfil the assumptions of the pervasive computing paradigm where end-to-end delivery of eHealth services on the "anywhere & anytime" basis is facilitated by a cloud of distributed networking devices and loosely coupled application modules. The key feature of such an approach is the user-centricity where the user does not invoke any particular applications or service nor even specifies where the application should be executed.

Here we propose a general idea of delivering to complex eHealth services in a distributed networking environment. The main feature of the proposed idea is that the process of complex services delivery is aware of the content being delivered, the context of the services delivery and that the delivered services are personalized for each separate end-user. In order to achieve the content, context and user awareness, we propose a general scheme for signalling system, which task is to configure distributed application modules and network resources with respect to the requirements imposed by the content being delivered, context of services delivery, and the specific user's needs.

3.25 Integration of Test eHealth Applications with the Virtualized IPv6 QoS Network

Halina Tarasiuk (Warsaw Univ. of Technology – Warsaw, PL)

For now, International Telecommunication Union – Telecommunications Standardization Sector (ITU-T) is being defining the following service components of service stratum in NGN architecture: IPTV, PSTN/ISDN emulation, IP multimedia (IMS). The aim of these components is mainly to allow a communication for typical home/business users, *e.g.*, for IPTV applications or voice connections. However, nowadays we observe an intensive research on enablers for delivering eHealth services. These services usually would base on typical home users' services suite – like voice communication or video-conferencing, however, eHealth also requires to define new type of services, which since now have not been considered, *e.g.*, in NGN. Such services should support, *e.g.*, transmission of data from sensors monitoring personal health parameters such as: ECG, heart rate, blood pressure, wheezes detection for asthma patients, and so on. Since the above services are related to the human vital signs parameters, we need to consider, which of them have to be transmitted to a doctor or medical centre using which Quality of Service (QoS) levels of the underlying network, and as a consequence, which of them require utilization of the signalling system of the NGN architecture to establish a new connection.

The IPv6 QoS system is an approach researched in scope of the Future Internet research designed by polish national project "Future Internet Engineering" being a part of the larger

Polish IIP System. In scope of the IIP, we integrated the following test eHealth applications with the virtualized IPv6 QoS network: SmartFit, eDiab and eAsthma. THE IIP is now in the prototyping phase and will be tested in PL-LAB environment, which is established as a Polish national pilot network. We expect that PL-LAB will establish some co-operation with similar pilot networks.

The questions remains (1) what are further requirements posed on Future Internet for handling eHealth; (2) are the existing TCP/IP networks appropriate for Future Internet from the point of view of the eHealth requirements; (3) is a dedicated Service Control Function (SCF) for eHealth in NGN needed; and (4) what signalling protocols are applicable in the dedicated SCF for eHealth, choice being SIP in IMS or XMPP as for now.

3.26 Why is Health still Focused on a Hospital Based Regulated Market: The Concept of the Citizen as Health Coproducer

Vicente Traver Salcedo (Universidad Politécnica – Valencia, ES)

Most of the ICT investments in the field of eHealth are oriented to health care practices instead of being focused on primary prevention, especially taking into account how chronic conditions treatments are consuming most of the health care budget. As every decision in our life is affecting our health state and not only the decisions we take when we are in the health care service provision loop (which is less than 1% of our life), I am proposing to "break the wall".

The question how to do that remains unresolved. The focus of putting individuals in charge of managing their health and lifestyle and the fact that health is co-produced in their everyday life, gives rise to the question of what should be the new Rules of the Game (ROG) that need to be followed. There exists a traditional regulated healthcare system, where the ROG's have been defined and refined over the years with doubtful success. Based on PREVE EU project results, what we propose is the co-production of health from a citizen-based perspective, where so far no ROG's have been defined and new stakeholders and business opportunities are emerging.

Therefore, the future of a sustainable health system implies to extend the healthcare system. To achieve that, there there are a lot of questions about behaviour modelling, value creation, value networks and business models that need to be answered as technology is already available, yet technology itself has no value. Instead of that, value is created by the individual by using technology enabled services. This also reinforces the concept that the support provided from the outside to individuals to manage their lifestyle must be personalized and based on their free will. Therefore, ICT must fill the existing communication gap among all the stakeholders involved and the new ROG still needed to be defined, in a more complex scenario, involving a wide range of opportunities to improve our health and lifestyle.

3.27 QoE and QoS as Ingredients for eHealth Service Acceptability

Muhammad Ullah (Blekinge Institute of Technology – Karlskrona, SE)

The provision of eHealth services that have clinical efficacy, diagnostic accuracy and hold capabilities to fulfil end user expectations is still challenging, especially when service acceptance is the criterion for success. The contributing factors that may affect user acceptance include, on one side, the uncertainty about the capabilities of the deployed technologies for the desired eHealth service, and, on the other side, the resulting end-user perception or experiences of the delivered service quality. For user experiences (in turn, depending on their expectations and previous experiences), "quality of the delivered service" is considered as a criterion having a significant influence on service acceptability in most of the circumstances. To deal with acceptance of service quality by their end users, a notion of Quality of Experience (QoE) has emerged. QoE as by definition represents user experiences of the delivered service, stands for the user perception and level of satisfaction, motivates user inclination toward service acceptance or rejection. In other words, the higher is the QoE of the eHealth service the higher will be the acceptance ratio and vice versa. One of the major contributors to QoE is the Quality of Service (QoS) level of the end-to-end communication channel between the service provider and the service user. To avoid the mismatch between the capabilities of deployed technologies and end-user expectation, the QoE metrics of the highest importance in healthcare practices needs to be identified, especially those influencing the eHealth service efficacy, efficiency and effectivenss. The deployed technologies must assure at least the required levels of QoE.

3.28 Future Internet for eHealth: From QoS via QoE to QoL (and Back)

Katarzyna Wac (University of Geneva – Geneva, CH)

The key aim of the Future Internet for eHealth shall be to establish quality research on computing and communications systems deployed in the healthcare domain and delivering clinically validated eHealth services to patients, their professional and voluntary caregivers "anywhere-anytime-anyhow". The general approach shall be necessarily a transdisciplinary one, as many of the challenges in development and deployment phases of these systems need to be addressed considering not only the system, its services and their Quality of Service (QoS) in an operational environment, but, as sometimes missing in the current research – the actual system end-users, approaching eHealth not from the perspective of new hype technology, but from the perspective of new innovate ways to achieve their pre-defined goal in the healthcare domain. These end-users have thus different eHealth services' requirements, expectations and perceptions of the Quality of Experience (QoE), including system's usability, efficacy, efficiency and effectiveness. In transdisciplinary research projects, we shall aim to establish innovative methodologies to demonstrate the value of Future Internet for eHealth improving the QoS, the QoE and ultimately the health outcomes and QoL of its end-users, and thus the social welfare.

3.29 Free Access to All Medical Data, Why Not?

Marc van Anderlecht (Uni-Com Medical – Brussels, BE)

In my talk I present an idea of having a free access to all medical data. I do not expect positive responses based on the fact that people will decrease their expectations for medical data privacy. Even if the sharing of personal data is in increasing, some "cultural" or legal heritage will impeach this. The question is driven by the fact, that is that it is now possible to use dictionary for pathologies, use ISO norms for other data, and use "open source" or recognized files format (HL7, Dicom, etc.) for data. The interest for medical research should be more important than particular interest of some states, so the standardization coordination will be boosted in that direction and states will have an excellent incentive thanks to reimbursement policy. Standardization has already proven in many cases being very important for market competition, as well as for the community, that has a direct advantage in terms of decrease of costs for the use of medical devices.

By addressing those necessary points with respect to data format and its standardization, we can put forward the architecture of a global medical records' database. The idea is that instead of protecting all and each patient medical data, which is very difficult and expensive; we put all the medical records with a free access and protect only the identity of the patient. A patient key code, a physician key code (or any other protection used for example for banks) is enough to protect the identification of the patient related to his/her personal record. We suggest to increase the sharing of medical records and decrease the IT security costs.

Participants

- Albert Alonso
Hospital Clinic de Barcelona – Barcelona, ES
- Gerald Bieber
Fraunhofer IGD – Rostock, DE
- Doris M. Bohman
Blekinge Institute of Technology – Karlskrona, SE
- Paolo Bonato
Harvard Medical School – Boston, US
- Sara Eriksén
Blekinge Institute of Technology – Karlskrona, SE
- Markus Fiedler
Blekinge Institute of Technology – Karlskrona, SE
- Geraldine Fitzpatrick
TU Wien – Vienna, AT
- Caroline Franck
Hôpitaux Universitaires – Geneva, CH
- Stefan Göbel
TU Darmstadt – Darmstadt, DE
- Nick Guldemond
TU Delft – Delft, NL
- Mattia Gustarini
Univ. of Geneva – Geneva, CH
- Jody Hausmann
Univ. of Geneva – Geneva, CH
- Rainer Herzog
Siemens Enterprise Communications – München, DE
- Hannes Kaufmann
TU Wien – Vienna, AT
- Zviad Kirtava
Tbilisi State Medical Univ. – Tbilisi, GE
- Willy Kostucki
Clinique Antoine Depage – Brussels, BE
- Ernoe Kovacs
NEC Laboratories Europe – Heidelberg, DE
- Lenka Lhotská
Czech Technical University – Prague, CZ
- Maria Martini
Kingston University – Kingston upon Thames, GB
- Goran Martinovic
Univ. of Osijek – Osijek, HR
- Dave Marvit
Fujitsu Labs of America Inc. – Sunnyvale, US
- Kevin Patrick
University of California – San Diego, US
- Terje Peetso
European Coomission – Brussels, BE
- Pawel Swiatek
Wroclaw University of Technology – Wroclaw, PL
- Halina Tarasiuk
Warsaw University of Technology – Warsaw, PL
- Vicente Traver Salcedo
Universidad Politécnica – Valencia, ES
- Muhammad Ullah
Blekinge Institute of Technology – Karlskrona, SE
- Marc van Anderlecht
Uni-Com Medical – Brussels, BE
- Katarzyna Wac
Univ. of Geneva – Geneva, CH

Report from Dagstuhl Seminar 12241

Data Reduction and Problem Kernels

Edited by

Michael R. Fellows[1], Jiong Guo[2], Dániel Marx[3], and Saket Saurabh[4]

1 **Charles Darwin University, AU,** michael.fellows@cdu.edu.au
2 **Universität des Saarlandes, DE,** j.guo@mmci.uni-saarland.de
3 **Tel Aviv University, IL,** dmarx@cs.bme.hu
4 **The Institute of Mathematical Sciences – Chennai, IN,** saket@imsc.res.in

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 12241 "Data Reduction and Problem Kernels". During the seminar, several participants presented their current research, and ongoing work and open problems were discussed. Abstracts of the presentations given during the seminar as well as abstracts of seminar results and ideas are put together in this paper. The first section describes the seminar topics and goals in general. Links to extended abstracts or full papers are provided, if available.

Seminar 10.–15. June, 2012 – www.dagstuhl.de/12241
1998 ACM Subject Classification F.2 Analysis of Algorithms and Problem Complexity, G.2 2 Graph Theory, F.1.3 Complexity Measures and Classes
Keywords and phrases Preprocessing, Fixed-parameter tractability, Parameterized algorithmics
Digital Object Identifier 10.4230/DagRep.2.6.26
Edited in cooperation with Neeldhara Misra (Institute of Mathematical Sciences, Chennai, India, mail@neeldhara.com)

1 Executive Summary

Michael R. Fellows
Jiong Guo
Dániel Marx
Saket Saurabh

Preprocessing (data reduction or kernelization) is used universally in almost every practical computer implementation that aims to deal with an NP-hard problem. The history of preprocessing, such as applying reduction rules to simplify truth functions, can be traced back to the origins of Computer Science — the 1950's work of Quine, and much more. A modern example showing the striking power of efficient preprocessing is the commercial integer linear program solver CPLEX. The goal of a preprocessing subroutine is to solve efficiently the "easy parts" of a problem instance and reduce it (shrinking it) to its computationally difficult "core" structure (the *problem kernel* of the instance).

How can we measure the efficiency of such a kernelization subroutine? For a long time, the mathematical analysis of polynomial time preprocessing algorithms was neglected. The basic reason for this anomalous development of theoretical computer science, was that if we seek to start with an instance I of an NP-hard problem and try to find an efficient (P-time) subroutine to replace I with an equivalent instance I' with $|I'| < |I|$ then success would imply P=NP — discouraging efforts in this research direction, from a mathematically-powered point of view.

The situation in regards the systematic, mathematically sophisticated investigation of preprocessing subroutines has changed drastically with advent of parameterized complexity, where the issues are naturally framed. More specifically, we ask for upper bounds on the reduced instance sizes as a function of a parameter of the input, assuming a polynomial time reduction/preprocessing algorithm.

A typical example is the famous Nemhauser-Trotter kernel for the Vertex Cover problem, showing that a "kernel" of at most $2k$ vertices can be obtained, with k the requested maximum size of a solution. A large number of results have been obtained in the past years, and the research in this area shows a rapid growth, not only in terms of number of papers appearing in top Theoretical Computer Science and Algorithms conferences and journals, but also in terms of techniques. Importantly, very recent developments were the introduction of new lower bound techniques, showing (under complexity theoretic assumptions) that certain problems must have kernels of at least certain sizes, meta-results that show that large classes of problems all have small (e.g., linear) kernels — these include a large collection of problems on planar graphs and matroid based techniques to obtain randomized kernels.

Kernelization is a vibrant and rapidly developing area. This meeting on kernelization consolidated the results achieved in the recent years, discussed future research directions, and exploreed further the applications potential of kernelization algorithms, and gave excellent opportunities for the participants to engage in joint research and discussions on open problems and future directions. This workshop was also special as we celebrated the 60th birthday of one of the founder of parameterized complexity, Prof. Michael R. Fellows. We organised a special day in which we remembered his contributions to parameterized complexity, science in general and mathematics for children.

The main highlights of the workshop were talks on the solution to two main open problems in the area of kernelization. We give a brief overview of these new developments below.

The AND Conjecture

The OR-SAT problem asks if, given m formulas each of size n, at least one of them is satisfiable. In 2008, Fortnow and Santhanam showed that if there is a reduction from OR-SAT to any language L with the property that the reduction reduces to instances of size polynomial in n (independent of m) then the polynomial-time hierarchy collapses. Such a reduction is called an OR-distillation, and this work motivated the notion of an OR-composition, which produces a boolean OR of parameterized instances of a given problem, without any restriction on the size. It was then established that an OR-composition and a polynomial kernel cannot co-exist, because these ingredients can be combined to lead to an OR-distillation. Thus, an OR-composition counts as evidence against the existence of a polynomial kernel, and it has turned into a very successful framework for establishing kernel lower bounds.

The question of whether there is similar evidence against the existence of an AND-distillation (defined analogously) has since been open. Such a result would imply that problems that have AND-compositions are also unlikely to admit polynomial kernels, and would therefore be a significant addition to the kernel lower bound toolkit. The question has been a central open problem for the kernelization community and was settled by Drucker in his work on classical and quantum instance compression. The route to the result is quite involved, and forges new connections between classical and parameterized complexity.

Tools from Matroid and Odd Cycle Traversal

The Odd Cycle Traversal problem asks if, given a graph G, there is a subset S of size at most k whose removal makes the graph bipartite. Equivalently, the question is if there is a subset S of size at most k that intersects every odd cycle in G. The problem was first shown to be FPT by Reed, Smith, and Vetta in 2004, and this was also the first illustration of the technique of iterative compression. However, the question of whether the problem admits a polynomial kernel was among the main open questions in the study of kernelization.

A breakthrough was recently made in work by Kratsch and Wahlström, providing the first (randomized) polynomial kernelization for the problem. It is a novel approach based on matroid theory, where all relevant information about a problem instance is encoded into a matroid with a representation of size polynomial in k.

Organization of the seminar and activities

The seminar consisted of twenty two talks, a session on open questions, and informal discussions among the participants. The organizers selected the talks in order to have comprehensive lectures giving overview of main topics and communications of new research results. Each day consisted of talks and free time for informal gatherings among participants. On the fourth day of the seminar we celebrated the 60th birthday of Mike Fellows, one of the founder of parameterized complexity. On this day we had several talks on the origin, history and the current developments in the field of parameterized complexity.

2 Table of Contents

Open Problems

3 Overview of Talks

3.1 Graph decompositions for algorithms and graph structure

Bruno Courcelle (Université Bordeaux, FR)

Several graph decompositions are important for algorithmic purposes, and not only tree-decompositions, rank-decompositions and those for clique-width. Many of them lead to "multi-kernelization" as they reduce a problem to several related problems for "prime" or "indecomposable" subgraphs.

I will review the algorithmic properties and uses of several known *canonical* decompositions: Tutte decomposition in 3-connected components, modular decomposition and split decomposition.

I will introduce a new one for strongly connected graphs, linked to Tutte decomposition that I call the **atomic decomposition**. The initial motivation is the study of Gauss words (curves in the plane) but there are other applications in view. It is related but different to a noncanonical decomposition of the same graphs by Knuth (1974)

3.2 (Non)constructive advances

Hans L. Bodlaender (Utrecht University, NL)

The talk surveys early results of Fellows and Langston and memorates Mike Fellows contributions to the field.

3.3 Tight Compression Bounds for Problems in Graphs with Small Degeneracy

Marek Cygan (University of Warsaw, PL)

We study kernelization in d-degenerate graphs. It is known that a few problems admit $k^{O(d)}$ kernels in d-degenerate graphs, including Induced Matching, Independent Dominating Set, Capacitated Vertex Cover, Connected Vertex Cover. Moreover a $k^{O(d^2)}$ kernel is known for Dominating Set. Simple reductions show that for Capacitated Vertex Cover and Connected Vertex Cover $k^{\Omega(d)}$ lower bounds exist. We show $k^{\Omega(d)}$ lower bounds for Induced Matching and Independent Dominating Set.

Furthermore, most interestingly, we also prove $k^{\Omega(d^2)}$lower bound for Dominating Set, which matches the known upper bound by Philip et al. [TALG] for this problem as well.

3.4 New Evidence for the AND- and OR-Conjectures

Andrew Drucker (MIT – Cambridge, US)

In the OR(SAT) problem, one is given a collection of Boolean formulas, each of length at most k, and wants to know whether at least one is satisfiable. Similarly, in the AND(SAT) problem, one wants to know whether all the formulas are individually satisfiable.

These problems are not known to have polynomial kernels. Work beginning with [Harnik and Naor '06; Bodlaender, Downey, Fellows, and Hermelin '08] has established that, if OR(SAT) is not polynomially kernelizable, then many other natural problems fail to have polynomial kernels. Bodlaender et al. also showed that the "kernelization-hardness" of AND(SAT) would imply a number of other hardness results. Thus, these two hypotheses, the "OR-" and "AND- conjectures," have a great deal of explanatory power. But should we believe them? In support of the OR-conjecture, [Fortnow and Santhanam '08] showed that OR(SAT) does not have polynomial kernels unless NP is in coNP/poly.

In this work we provide equally strong evidence for the AND-conjecture: if AND(SAT) has poly kernels then NP is in coNP/poly, and even in SZK/poly. We also extend the hardness evidence for OR(SAT) in several ways; for instance, we give the first strong evidence against probabilistic kernelizations for OR(SAT) with two-sided bounded error. To prove our results, we exploit the information bottleneck of a kernelization reduction, using a new, general method to "disguise" information being fed into a compressive mapping.

3.5 Train marshaling is fixed parameter tractable

Rudolf Fleischer (German University of Technology – Oman, OM)

The train marshalling problem is about reordering the cars of a train using as few auxiliary rails as possible. The problem is known to be NP-complete. We show that it is fixed parameter tractable (FPT) with the number of auxiliary rails as parameter.

3.6 Parameterized Complexity of the Workflow Satisfiability Problem

Gregory Z. Gutin (RHUL – London, GB)

Joint work of Jason Crampton, Gregory Z. Gutin and Anders Yeo.

The Workflow Satisfiability Problem (WSP) defined below arises in Access Control in Information Security.

In WSP, we are given a set S of steps and a set U of users and asked to decide whether there is a function $\pi : S \to U$ that satisfies some constraints. Firstly, each step can be assigned (mapped to) some subset of U. Secondly, there are some relations ρ on U (ie., $\rho \subseteq U \times U$) such that all constraints of the type $(\rho, S', S")$, where S', S'' are subsets of S, must be satisfied meaning that there exist $s' \in S'$ and $s'' \in S''$ such that $(\pi(s'), \pi(s'')) \in \rho$. Examples of ρ include $=$ and $\neq$.

Wang and Li (ACM Trans. Inf. Syst. Secur., 2010) proved that WSP is NP-hard. They also observed that $k = |S|$ is relatively small (with respect to $n = |U|$) and proved that k-WSP is W[1]-hard. They obtained a fixed-parameter algorithm for special cases of k-WSP when only relations $=$ and $\neq$ are allowed.

Using a result of Bjorklund, Husfeldt and Koivisto (SIAM J. Comput., 2009) we obtain a new fixed-parameter algorithm that significantly improves the runtime of Wang and Li and widen the special case for which k-WSP is fpt (including there organizations with hierarchical structures). In particular, we improve a result of Fellows, Friedrich, Hermelin, Narodytska, and Rosamond (IJCAI 2011). We also investigate the existence of polynomial-size kernels and obtain both positive and negative results using, in particular, a result of Dom, Lokshtanov and Saurabh (ICALP 2009).

3.7 Faster than Courcelle's Theorem on Shrubs

Petr Hlineny (Masaryk University, CZ)

URL http://arxiv.org/abs/1204.5194

Famous Courcelle's theorem claims FPT solvability of any MSO2-definable property in linear FPT time on the graphs of bounded tree-width (alternatively, of MSO1 on clique-width by Courcelle-Makowsky-Rotics). A drawback of this powerful algorithmic metatheorem is that its runtime has a nonelementary dependence on the quantifier alternation depth of the defining formula. This is indeed unavoidable in full generality (even on trees) as shown by Frick and Grohe.

We show a new kernelization approach to this problem, giving an MSO model checking algorithm on trees of bounded height in FPT with elementary dependence on the formula; actually, we "trade" a nonelementary runtime dependence on the formula for a nonelementary dependence of our kernel on the tree height. This implies a faster (than Courcelle's) new algorithm for all MSO2-definable properties on the graphs of bounded tree-depth, and similarly a faster algorithm for all MSO1-definable properties on the classes of bounded shrub-depth.

3.8 Preprocessing Subgraph and Minor Problems: When Does a Small Vertex Cover Help?

Bart Jansen (Utrecht University, NL)

We prove a number of results around kernelization of problems parameterized by the vertex cover of a graph. We provide two simple general conditions characterizing problems admitting kernels of polynomial size. Our characterizations not only give generic explanations for the existence of many known polynomial kernels for problems like Odd Cycle Transversal, Chordal Deletion, Planarization, η-Transversal, Long Path, Long Cycle, or H-packing, they also imply new polynomial kernels for problems like $\mathcal{F}$-Minor-Free Deletion, which is to delete at most k vertices to obtain a graph with no minor from a fixed finite set $\mathcal{F}$.

While our characterization captures many interesting problems, the kernelization complexity landscape of problems parameterized by vertex cover is much more involved. We demonstrate this by several results about induced subgraph and minor containment, which we find surprising. While it was known that testing for an induced complete subgraph has no polynomial kernel unless NP is in coNP/poly, we show that the problem of testing if a graph contains a given complete graph on t vertices as a minor admits a polynomial kernel. On the other hand, it was known that testing for a path on t vertices as a minor admits a polynomial kernel, but we show that testing for containment of an induced path on t vertices is unlikely to admit a polynomial kernel.

3.9 Max-Cut Parameterized Above the Edwards-Erdos Bound

Mark Jones (RHUL – London, GB)

We study the problem Max Cut: Given a graph, find a bipartite subgraph with the most edges. The Edwards-Erdos bound states that for any connected graph with n vertices, m edges, there is a bipartite subgraph with at least $m/2 + (n-1)/4$ edges.

We study Max Cut parameterized above this bound: Given a connected graph with n vertices, m edges, decide whether there is a bipartite subgraph with at least $m/2+(n-1)/4+k$ edges. We show that the problem is fixed-parameter tractable with running time $2^{(3k)}n^{O(1)}$, and has a kernel of size $O(k^5)$.

3.10 Data Reduction for Finding Diameter-Two Subgraphs

Christian Komusiewicz (TU Berlin, DE)

Given an undirected graph $G = (V, E)$ and an integer $l > 1$, the NP-hard 2-club problem asks for a vertex set $S \subseteq V$ of size at least l such that $G[S]$ has diameter at most 2.

We study the 2-club problem with respect to many-to-one- and Turing-kernelizability for a variety of parameters such as bandwidth of G, vertex cover size of G, the dual parameter $|V| - l$, and the feedback edge set number of G.

3.11 Kernel lower bounds using co-nondeterminism: Finding induced hereditary subgraphs

Stefan Kratsch (Utrecht University, NL)

This work further explores the applications of co-nondeterminism for showing kernelization lower bounds. The only known example excludes polynomial kernelizations for the RAMSEY(k) problem of finding an independent set or a clique of at least k vertices in a given graph (Kratsch, SODA 2012). We study the more general problem of finding induced subgraphs on k vertices fulfilling some hereditary property Π, called Π-INDUCED SUBGRAPH(k). The problem is NP-hard for all non-trivial choices of Π by a classic result of Lewis and Yannakakis (JCSS 1980). The parameterized complexity of this problem was classified by Khot and Raman (TCS 2002) depending on the choice of Π. The interesting cases for kernelization are for Π containing all independent sets and all cliques, since the problem is trivial or W[1]-hard otherwise.

Our results are twofold. Regarding Π-INDUCED SUBGRAPH(k), we show that for a large choice of natural graph properties Π, including chordal, perfect, cluster, and cograph, there is no polynomial kernel with respect to k. This is established by two theorems: one using a co-nondeterministic variant of cross-composition and one by a polynomial parameter transformation from RAMSEY(k).

Additionally, we show how to use improvement versions of NP-hard problems as source problems for lower bounds, without requiring their NP-hardness. E.g., for Π-INDUCED SUBGRAPH(k) our compositions may assume existing solutions of size $k-1$. We believe this to be useful for further lower bound proofs, since improvement versions simplify the construction of a disjunction (OR) of instances required in compositions. This adds a second way of using co-nondeterminism for lower bounds

3.12 Planar $\mathcal{F}$-Deletion: Kernelization, Approximation and FPT Algorithms (I)

Daniel Lokshtanov (University of California – San Diego, US)

In the $\mathcal{F}$-Deletion problem you are given a graph G and integer k and asked whether there is a set S on at most k vertices such that G does not contain any minors from $\mathcal{F}$, where $\mathcal{F}$ is a finite list of graphs. We show that if $\mathcal{F}$ contains at least one planar graph, then the $\mathcal{F}$-Deletion problem admits polynomial kernels, constant factor approximation algorithms. If additionally all graphs in $\mathcal{F}$ are connected the $\mathcal{F}$-Deletion problem admits $c^k \cdot n$ time FPT algorithms. On the way we develop some new and interesting tools. Our results are stringed together by a common theme of polynomial time preprocessing

3.13 FPT suspects and tough customers: Open problems of Downey and Fellows

Dániel Marx (MTA – Budapest, HU)

We give an update on the status of open problems from the book "Parameterized Complexity" by Downey and Fellows.

3.14 Planar-$\mathcal{F}$ Deletion: Approximation, Kernelization and Optimal FPT Algorithms (II)

Neeldhara Misra (The Institute of Mathematical Sciences – Chennai, IN)

The notion of protrusions – constant treewidth subgraphs that can be separated from the instance by constant-sized separators – has been very useful in the context of kernelization algorithms on sparse graphs. When the optimization problem in question has certain properties, protrusions lend themselves to vastly general reduction rules, leading to a number of interesting meta theorems on sparse graphs. Unfortunately, however, the technique is not easily amenable to work the same way on general graphs.

In particular, for the Planar $\mathcal{F}$-deletion problem on general graphs, it turns out that even for apparently simply cases, non-trivial degree reduction rules crafted "by hand" have to come into play before protrusion-based reductions can be applied. It is not clear that this approach is amenable to generalization for more complex cases.

We therefore revisit the notion of a protrusion and introduce a more flexible variant, namely a near-protrusion. Informally, a near-protrusion is a subgraph which can become a protrusion in the future, after removing some vertices of some optimal solution. The usefulness of near-protrusions is that they allow us to find an irrelevant edge, i.e., an edge which removal does not change the problem.

We give a brief overview of the ideas involved in making protrusion-based reductions work in more general situations.

3.15 Planar-$\mathcal{F}$ deletion in parameterized single exponential time

Christophe Paul (CNRS, Université Montpellier II, FR)

Let $\mathcal{F}$ be a finite family of graphs containing at least one planar graph. In the parameterized PLANAR-$\mathcal{F}$ DELETION problem, we are given an n-vertex graph G and a non-negative integer k (the parameter), and the question is whether G has a set X of vertices of size at most k such that $G - X$ is H-minor-free for every H in $\mathcal{F}$. This problem encompasses a number of well-studied parameterized problems such as Vertex Cover, Feedback Vertex Set, or Treewidth-t Vertex Deletion for every value of $t \geq 0$. We present a algorithm for the parameterized PLANAR-$\mathcal{F}$ DELETION problem running in parameterized single-exponential time. Our approach significantly deviates from previous work as we do not use any reduction rule, but instead we apply a series of branching steps. This allows us to deal, in particular, with the case where the graphs in $\mathcal{F}$ are not necessarily connected, which was not known to admit a single-exponential algorithm

3.16 Graph separation: New incompressibility results

Marcin Pilipczuk (University of Warsaw, PL)

In the talk we plan to present the recent developments on the kernelization hardness of graph separation problems. We show that, unless NP is contained in coNP/poly, the following parameterized problems do not admit a polynomial kernel:

- Directed Edge/Vertex Multiway Cut, parameterized by the size of the cutset, even in the case of two terminals,
- Edge/Vertex Multicut, parameterized by the size of the cutset,
- and k-Way Cut, parameterized by the size of the cutset.

Our results complement very recent developments in designing parameterized algorithms for cut problems by Marx and Razgon [STOC'11], Bousquet et al. [STOC'11], Kawarabayashi and Thorup [FOCS'11] and Chitnis et al. [SODA'12].

The presented results are included in the ICALP'12 paper "Clique cover and graph separation: New incompressibility results" (joint work with Marek Cygan, Stefan Kratsch, Michal Pilipczuk and Magnus Wahlstrom).

3.17 Tight bounds for Edge Clique Cover

Michal Pilipczuk (University of Bergen, NO)

In the EDGE CLIQUE COVER problem, given a graph G and an integer k, we ask whether the edges of G can be covered with k complete subgraphs of G or, equivalently, whether G admits an intersection model on k-element universe. Gramm et al. [JEA 2008] have shown a set of simple rules that reduce the number of vertices of G to 2^k, and no algorithm is known with significantly better running time bound than a brute-force search on this reduced instance. In this work we show that the approach of Gramm et al. is essentially optimal: we present a polynomial time algorithm that reduces an arbitrary 3-CNF-SAT formula with n variables and m clauses to an equivalent EDGE CLIQUE COVER instance (G, k) with $k = O(\log n)$ and $|V(G)| = O(n + m)$. This implies that EDGE CLIQUE COVER does not admit an FPT algorithm that has better than doubly-exponential running time dependency on k, unless ETH fails. Moreover, we exclude subexponential kernels for the problem under ETH and under NP not contained in coNP/poly. This refines previous work together with Stefan Kratsch and Magnus Wahlstroem [ICALP 2012], in which we proved that polynomial kernelization would contradict the second complexity assumption.

3.18 Linear Kernels on Graphs Excluding a Topological Minor

Somnath Sikdar (RWTH Aachen, DE)

In this talk, we will sketch a proof of the following result: a parameterized graph problem that has finite integer index and satisfies a property that we call "treewidth-bounding" admits a linear kernel on the class of H-topological-minor free graphs, where H is an arbitrary but fixed graph. This builds on earlier work on the existence of linear kernels by Bodlaender et al. on graphs of bounded genus and by Fomin et al. on H-minor-free graphs. This result implies that several problems, including Chordal Vertex Deletion, Feedback Vertex Set and Edge Dominating Set, admit linear kernels on H-topological-minor-free graphs.

3.19 A Polynomial kernel for Proper Interval Vertex Deletion

Yngve Villanger (University of Bergen, NO)

It is known that the problem of deleting at most k vertices to obtain a proper interval graph (Proper Interval Vertex Deletion) is fixed parameter tractable. However, whether the problem admits a polynomial kernel or not was open. Here, we answers this question in affirmative by obtaining a polynomial kernel for Proper Interval Vertex Deletion. This resolves an open question of van Bevern, Komusiewicz, Moser, and Niedermeier

3.20 Uses of Matroids in Kernelization

Magnus Wahlström (MPI für Informatik – Saarbrücken, DE)

In some recent results (Kratsch and Wahlström, SODA 2012; Kratsch and Wahlström, preprint, 2012), tools from matroid theory have shown themselves to have powerful applications in polynomial kernelization; in particular, a tool known as representative sets (Marx, 2006; Lovász, 1980) has proved itself very useful.

In this talk, I will give an overview of the use of these tools, illustrating with applications to kernels for Almost 2-SAT and for graph cut problems.

3.21 Different parameterizations of the Test Cover problem

Anders Yeo (RHUL – London, GB)

In the Test Cover problem we are given a set $\{1, \ldots, n\}$ of items together with a collection, T, of distinct subsets of these items called tests. We assume that T is a test cover, i.e., for each pair of items there is a test in T containing exactly one of the items. The objective is to find a minimum size subcollection of T which is still a test cover.

This problem is NP-hard, so we consider the following parameterizations of the problem, where k is the parameter and m is the number of tests available.

1. Is there a solution with at most k tests?
2. Is there a solution with at most $n - k$ tests?
3. . Is there a solution with at most $m - k$ tests, where m is the size of T?
4. Is there a solution with at most $(\log n) + k$ tests?

The above is of interest as n and m are upper bounds for the size of an optimal solution and $\log n$ is a lower bounds. We state the FPT-complexities of the above parameterizations and focus on (non-)polynomial kernel results. In particular we will illustrate why parameterization 1 has no polynomial kernel (unless NP is a subset of coNP/poly).

4 Open Problems

4.1 Above Guarantee Independent Set on Planar Graphs

Venkatesh Raman (The Institute of Mathematical Sciences – Chennai, IN)

It is well known that every planar graph admits an independent set on at least $n/4$ vertices, as an easy consequence of the Four Color Theorem. The above guarantee version of the question involve asking for an independent set of size at least $\frac{n}{(4+k)}$. The parameterized complexity of this question, parameterized by k, is open. As an aside, we note that the question is non-trivial even when $k = 1$.

4.2 Biclique

Mike Fellows (Charles Darwin University – Darwin, AU)

The parameterized complexity of finding $K_{k,k}$ as a subgraph when parmeterized by k is a long-standing and notorious open problem.

The multicolored variant (where the vertex set is partitioned into $2k$ parts and we would like to find a subgraph that involves exactly one vertex from each part) is known to be W[1]-hard (see Appendix, [10]). It is also known that counting bicliques is #W[1]-hard parameterized by k. If the input graph has no induced paths of length s, then the problem is fixed-parameter linear parameterized by k and s [1].

4.3 Chromatic Number of P_5-free graphs

Fedor Fomin (University of Bergen, NO)

The chromatic number of a P_4-free graph can be computed in polynomial time and it is NP-hard to compute the chromatic number of a P_5-free graph. However, the question of whether the chromatic number of a P_5-free graph is at most k admits a XP algorithm [21]. It is open as to whether there the problem is FPT.

4.4 Clique for Line Segments

Dániel Marx (MTA – Budapest, HU)

The problem of finding a maximum sized clique on the intersection graph of line segments in the plane is known to be NP-hard [4]. However, the parameterized complexity remains open.

4.5 Cliquewidth

Daniel Lokshtanov (University of California – San Diego, US)

While problems parameterized by cliquewidth are studied extensively, computing cliquewidth remains elusive.

A well-known fact is that if the tree-width of a graph is t then its clique-width is bounded by $3 \cdot 2^{t-1}$ [7]. On the other hand, complete graphs have clique-width 2 and unbounded tree-width. However, for sparse graphs the treewidth and cliquewidth are linearly related.

Hlineny and Oum obtained an algorithm running in polynomial time and computing $(2^{k+1} - 1)$-expressions for a graph G of clique-width at most k [20].

An FPT algorithm for computing cliquewidth remains an open problem.

4.6 Contraction Decomposition Beyond H-minor Free Graphs

MohammadTaghi Hajiaghayi (University of Maryland – College Park, US)

A contraction decomposition is a partition of the edges of a graph into a desired number k of color classes such that contracting the edges in any one color class results in a graph of treewidth linear in k.

A series of results on contraction decompositions finally culminated in this paper [11], which showed that such decompositions exist for H-minor free graphs and can be computed in polynomial time. This leads to a general framework for approximation and fixed-parameter algorithms for problems closed under contractions in graphs excluding a fixed minor. For example, one of the implications is (another) fixed-parameter algorithm for k-cut in H-minor-free graphs, which was an open problem of Downey et al. even for planar graphs.

Can such decompositions be constructed for other classes of graphs?

4.7 Directed Feedback Vertex Set

Saket Saurabh (The Institute of Mathematical Sciences – Chennai, IN), Stefan Kratsch (Utrecht University, NL), and Magnus Wahlström (MPI für Informatik – Saarbrücken, DE)

Directed Feedback Vertex Set is known to be FPT parameterized by solution size [5], however, the question of whether it admits a polynomial kernel remains open. Little progress has been made on this question, even on special graph classes.

4.8 Disjoint Paths

Saket Saurabh (The Institute of Mathematical Sciences – Chennai, IN)

The Disjoint Paths problem asks for vertex disjoint paths connecting k terminal pairs. It is known to be NP-complete even on planar graphs [25], and is FPT parameterized by k. However, the FPT result relies on graph minor theory. An explicit FPT algorithm that avoids this route – even an explicit FPT approximation – remains open.

4.9 Even Set

Mike Fellows (Charles Darwin University – Darwin, AU)

The Even Set problem takes a red/blue bipartite graph as input and asks for a non-empty set of at most k red vertices R such that each blue vertex has an even number of neighbors in R. The problem is known to be NP-complete [29] and the exact version of the problem, where $|R| = k$, is W[1]-hard [13]. The problem can be reformulated in a number ways in a variety of contexts, and its parameterized complexity is an important unsolved problem.

4.10 Group Feedback Edge/Vertex Set

Stefan Kratsch (Utrecht University, NL) and Magnus Wahlström (MPI für Informatik – Saarbrücken, DE)

Group Feedback Vertex Set (GFVS) is a broad generalization of the Odd Cycle Transversal problem (OCT). In this problem, the input is a graph G with edges labeled by elements from a group Γ, and the task is to delete vertices (or edges) to remove any cycles whose labels do not sum up to zero. OCT corresponds to the case where Γ =GF(2); see the literature for more precise definitions. The problem is known to be FPT, even in quite general variants; see [19, 9]. When Γ is fixed, the problem has a polynomial kernel [23], but the case of a non-fixed group is open and interesting. Depending on the choice of group representation and parameter, the strength of the problem seems to vary, but it is known to subsume both Multiway Cut, for an arbitrary large-enough group [23], and Subset Feedback Vertex Set, for a group with $2^{O(n)}$ elements given via oracle access [9]. A polynomial kernel for this most-general setting would be surprising, but even this is not excluded; failing this, though, the question is for what settings and parameter combinations the problem allows a polynomial kernel.
Open: GFVS(k+$|\Gamma|$) and GFVS(k). (In explicit or oracle representation of the group)

4.11 Knapsack Parameterized by Items

Stefan Kratsch (Utrecht University, NL) and Magnus Wahlström (MPI für Informatik – Saarbrücken, DE)

We know that subset sum is W[1]-hard if parameterized by the size of the subset of numbers sought. On the other hand, it is also admits a randomized polynomial kernel parameterized by n. Does an analogous result hold (in the kernelization context) for knapsack parameterized by the number of items?

4.12 Lower Bounds for Turing Kernels

Mike Fellos (Charles Darwin University – Darwin, AU)

Turing kernels were introduced as a potential coping strategy for problems that are not expected to have polynomial kernels under standard complexity-theoretic assumptions [16]. Informally speaking, these are "many polynomial kernels" – that are independent of each other and can therefore be processed in parallel. The number of kernels is allowed to be a function of n. Formulating a lower bound framework for Turing kernels remains an open problem.

4.13 Multiway Cut

Stefan Kratsch (Utrecht University, NL) and Magnus Wahlström (MPI für Informatik – Saarbrücken, DE)

Arguably a polynomial kernelization for MULTIWAY CUT parameterized by the size k of the requested cutset is one of the biggest open questions in kernelization of cut problems (along with DFVS). It is known that there are randomized polynomial kernels with $O(k^{s+1})$ vertices when the number of terminals is bounded by some constant s, and with $O(k^3)$ vertices when terminals are deletable (the latter is equivalent to having terminal-degrees equal to one) [23]. The main open question is what happens when s may be unbounded; note though, that known reduction rules give $s \leq 2k$ so parameterization by $k + s$ is just as hard. Furthermore, the edge deletion variant holds independent interest.
Open: Multiway Cut(k).

4.14 Multicut

Stefan Kratsch (Utrecht University, NL) and Magnus Wahlström (MPI für Informatik – Saarbrücken, DE)

The standard parameterization by the size of the requested cutset was recently showed not to admit a polynomial kernelization under the standard assumption [8], following the recent breakthrough results that show its fixed-parameter tractability [3, 27]. However, the used cross-composition from 3-Multiway Cut creates a large number of terminal pairs [8]. Hence, it is interesting to know whether parameterization by $k + s$ (here s is the number of terminal pairs) is helpful for getting a polynomial kernelization. Similarly to MULTIWAY CUT, there is a randomized polynomial kernelization when the number of terminal pairs is bounded by some constant s. Note that deleteable terminals do not help, since terminals can be easily copied without creating undesired requests (unlike for MULTIWAY CUT).
Open: Multicut($k + s$).

4.15 Multiway Cut and Multicut in directed graphs

Stefan Kratsch (Utrecht University, NL) and Magnus Wahlström (MPI für Informatik – Saarbrücken, DE)

It is known that the standard parameterizations of DIRECTED MULTIWAY CUT and DIRECTED MULTICUT do not admit polynomial kernelizations even when there are only two terminals respectively one terminal pair [8]. Note that DIRECTED MULTIWAY CUT is FPT [6] and DIRECTED MULTICUT is W[1]-hard [27]. For DIRECTED MULTICUT the restriction to directed acyclic graphs (DAGs) remains W[1]-hard when parameterized by the cutset only but it is

FPT when parameterized by the cutset k plus the number s of terminal pairs [22]; this leaves open whether it admits a polynomial kernelization parameterized by $k+s$ or parameterized by k and with s fixed (the restriction to DAGs prevents the lower bound construction used for general directed graphs).
Open: Multicut-in-DAGs($k+s$) and s-Multicut-in-DAGs(k).

4.16 Parameterized Approximation for Dominating Set

Mike Fellows (Charles Darwin University – Darwin, AU)

The following question is open: Is there an FPT algorithm that, given a graph G and parameter k, either determines that G has no k-Dominating Set, or or produces a dominating set of size at most $g(k)$ (where $g(k)$ is some fixed function of k?

It is known that there is no such FPT algorithm for $g(k)$ of the form $(k+c)$ (where c is a fixed constant), unless FPT = W[2]. Also, it is known that there is no such FPT algorithm for *any* $g(k)$ for the Independent Dominating Set problem unless FPT = W[2] [14]. The Threshold Set problem is also known to be FPT inapproximable for any function g unless FPT = W[1] [26].

4.17 Polynomial Kernels for $\mathcal{F}$-deletion

Daniel Lokshtanov (University of California – San Diego, US)

The $\mathcal{F}$ deletion problem asks for a subset of vertices of size at most k whose removal makes a graph H-minor free for allowed $H \in \mathcal{F}$. It is known that the problem admits a polynomial kernel (parameterized by k) if $\mathcal{F}$ contains at least one planar graph [17], but the kernelization complexity is open for the case when $\mathcal{F}$ contains only non-planar graphs.

4.18 Polynomial Kernel for Imbalance

Saket Saurabh (The Institute of Mathematical Sciences – Chennai, IN)

The problem of checking if a graph admits a layout with imbalance at most k is known to be FPT parameterized by k [24]. The question of whether the problem admits a polynomial kernel is open.

4.19 Quadratic Integer Programming

Daniel Lokshtanov (University of California – San Diego, US)

While integer linear programs are known to have FPT algorithms parameterized by the number of variables, the parameterized complexity is unsettled for quadratic integer programs parameterized by the number of variables.

A FPT algorithm would imply that Optimal Linear Arrangement is FPT when parameterized by Vertex Cover [15].

4.20 Treewidth

Hans L. Bodlaender (Utrecht University, NL)

Treewidth is a fundamental structural parameter that has played a central role in algorithmic graph theory in general and in obtaining FPT algorithms in particular. The complexity of computing treewidth is of great interest, and here are some of the challenging open problems.

1. What is the complexity of treewidth on planar graphs? A (3/2)-approximation is known since branchwidth can be computed in polynomial time [28], but even NP-hardness remains open. It is also known that the treewidth of a planar graph is linear in the tree-length of the graph [12].
2. Is there a constant-factor approximation for treewidth? The answer is in the negative assuming the Small Set Expansion conjecture [2].
3. Is $2^{(O(k^3))}$ optimal for tree width in terms of the function of k?

References

1 Aistis Atminas, Vadim V. Lozin, and Igor Razgon. Linear time algorithm for computing a small biclique in graphs without long induced paths. In *Algorithm Theory - SWAT, Proceedings*, volume 7357 of *Lecture Notes in Computer Science*, pages 142–152. Springer, 2012.
2 Per Austrin, Toniann Pitassi, and Yu Wu. Inapproximability of treewidth, one-shot pebbling, and related layout problems. volume 7408 of *Lecture Notes in Computer Science*, pages 13–24. Springer, 2012.
3 Nicolas Bousquet, Jean Daligault, and Stéphan Thomassé. Multicut is fpt. In Fortnow and Vadhan [18], pages 459–468.
4 Sergio Cabello, Jean Cardinal, and Stefan Langerman. The clique problem in ray intersection graphs. In *Algorithms - ESA - 20th Annual European Symposium*, volume 7501 of *Lecture Notes in Computer Science*, pages 241–252. Springer, 2012.
5 Jianer Chen, Yang Liu 0002, Songjian Lu, Barry O'Sullivan, and Igor Razgon. A fixed-parameter algorithm for the directed feedback vertex set problem. *J. ACM*, 55(5), 2008.
6 Rajesh Hemant Chitnis, MohammadTaghi Hajiaghayi, and Dániel Marx. Fixed-parameter tractability of directed multiway cut parameterized by the size of the cutset. In Yuval Rabani, editor, *SODA*, pages 1713–1725. SIAM, 2012.
7 Corneil and Rotics. On the relationship between clique-width and treewidth. *SICOMP: SIAM Journal on Computing*, 34, 2005.

8 Marek Cygan, Stefan Kratsch, Marcin Pilipczuk, Michal Pilipczuk, and Magnus Wahlström. Clique cover and graph separation: New incompressibility results. *CoRR*, abs/1111.0570, 2011. To appear in ICALP 2012.

9 Marek Cygan, Marcin Pilipczuk, and Michal Pilipczuk. On group feedback vertex set parameterized by the size of the cutset. *CoRR*, abs/1112.6255, 2011.

10 Holger Dell and Dániel Marx. Kernelization of packing problems. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 68–81. SIAM, 2012.

11 Erik D. Demaine, MohammadTaghi Hajiaghayi, and Ken ichi Kawarabayashi. Contraction decomposition in h-minor-free graphs and algorithmic applications. In *Proceedings of the 43rd ACM Symposium on Theory of Computing, STOC*, pages 441–450. ACM, 2011.

12 Youssou Dieng and Cyril Gavoille. On the tree-width of planar graphs. *Electronic Notes in Discrete Mathematics*, 34:593–596, 2009.

13 Downey and Fellows. Fixed-parameter tractability and completeness I: Basic results. *SICOMP: SIAM Journal on Computing*, 24, 1995.

14 R G. Downey, M. R. Fellows, C. McCartin, and F. A. Rosamond. Parameterized approximation of dominating set problems. *Information Processing Letters*, 109(1):68–70, 2008.

15 Michael R. Fellows, Daniel Lokshtanov, Neeldhara Misra, Frances A. Rosamond, and Saket Saurabh. Graph layout problems parameterized by vertex cover. In *Algorithms and Computation, 19th International Symposium, ISAAC*, volume 5369 of *Lecture Notes in Computer Science*, pages 294–305. Springer, 2008.

16 Henning Fernau, Fedor V. Fomin, Daniel Lokshtanov, Daniel Raible, Saket Saurabh, and Yngve Villanger. Kernel(s) for problems with no kernel: On out-trees with many leaves. In *26th International Symposium on Theoretical Aspects of Computer Science (STACS 2009)*, volume 3 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 421–432, Dagstuhl, Germany, 2009. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

17 Fedor V. Fomin, Daniel Lokshtanov, Neeldhara Misra, and Saket Saurabh. Planar f-deletion: Approximation, kernelization and optimal fpt algorithms. In *Foundations of Computer Science, FOCS*, 2012. (To Appear).

18 Lance Fortnow and Salil P. Vadhan, editors. *Proceedings of the 43rd ACM Symposium on Theory of Computing, STOC 2011, San Jose, CA, USA, 6-8 June 2011.* ACM, 2011.

19 Sylvain Guillemot. Fpt algorithms for path-transversal and cycle-transversal problems. *Discrete Optimization*, 8(1):61–71, 2011.

20 Petr Hlinený and Sang il Oum. Finding branch-decompositions and rank-decompositions. *SIAM J. Comput*, 38(3):1012–1032, 2008.

21 Chính T. Hoàng, Marcin Kaminski, Vadim V. Lozin, Joe Sawada, and Xiao Shu. Deciding k -colorability of P $_5$-free graphs in polynomial time. *Algorithmica*, 57(1):74–81, 2010.

22 Stefan Kratsch, Marcin Pilipczuk, Michal Pilipczuk, and Magnus Wahlström. Fixed-parameter tractability of multicut in directed acyclic graphs. *CoRR*, abs/1202.5749, 2012. To appear in ICALP 2012.

23 Stefan Kratsch and Magnus Wahlström. Representative sets and irrelevant vertices: New tools for kernelization. *CoRR*, abs/1111.2195, 2011.

24 Daniel Lokshtanov, Neeldhara Misra, and Saket Saurabh. Imbalance is fixed parameter tractable. In *Computing and Combinatorics, 16th Annual International Conference, COCOON*, volume 6196 of *Lecture Notes in Computer Science*, pages 199–208. Springer, 2010.

25 James F. Lynch. The equivalence of theorem proving and the interconnection problem. *SIGDA Newsl.*, 5(3):31–36, 1975.

26 Dániel Marx. Completely inapproximable monotone and antimonotone parameterized problems. In *IEEE Conference on Computational Complexity*, pages 181–187. IEEE Computer Society, 2010.

27 Dániel Marx and Igor Razgon. Fixed-parameter tractability of multicut parameterized by the size of the cutset. In Fortnow and Vadhan [18], pages 469–478.
28 P. D. Seymour and R. Thomas. Call routing and the ratcatcher. *Combinatorica*, 14(2):217–241, 1994.
29 Vardy. Algorithmic complexity in coding theory and the minimum distance problem. In *STOC: ACM Symposium on Theory of Computing (STOC)*, 1997.

Participants

- Faisal Abu-Khzam
Lebanese American Univ., LB
- Cristina Bazgan
Université Paris-Dauphine, FR
- Hans L. Bodlaender
Utrecht University, NL
- Bruno Courcelle
Université Bordeaux, FR
- Marek Cygan
University of Warsaw, PL
- Rodney Downey
Victoria Univ. – Wellington, NZ
- Andrew Drucker
MIT – Cambridge, US
- Michael R. Fellows
Charles Darwin University – Darwin, AU
- Henning Fernau
Universität Trier, DE
- Rudolf Fleischer
German University of Technology – Oman, SAS
- Fedor V. Fomin
University of Bergen, NO
- Robert Ganian
Goethe-Universität Frankfurt am Main, DE
- Serge Gaspers
UNSW – Sydney, AU
- Archontia Giannopoulou
National and Kapodistrian University of Athens, GR
- Jiong Guo
Universität des Saarlandes, DE
- Gregory Z. Gutin
RHUL – London, GB
- MohammadTaghi Hajiaghayi
University of Maryland – College Park, US
- Pinar Heggernes
University of Bergen, NO
- Danny Hermelin
MPI für Informatik – Saarbrücken, DE
- Petr Hlineny
Masaryk University, CZ
- Juraj Hromkovic
ETH Zürich, CH
- Falk Hüffner
TU Berlin, DE
- Bart Jansen
Utrecht University, NL
- Mark Jones
RHUL – London, GB
- Iyad A. Kanj
DePaul University – Chicago, US
- Eun Jung Kim
Université Paris-Dauphine, FR
- Christian Komusiewicz
TU Berlin, DE
- Stefan Kratsch
Utrecht University, NL
- Michael A. Langston
University of Tennessee, US
- Daniel Lokshtanov
University of California – San Diego, US
- Dániel Marx
MTA – Budapest, HU
- Kurt Mehlhorn
MPI für Informatik – Saarbrücken, DE
- Daniel Meister
Universität Trier, DE
- Neeldhara Misra
The Institute of Mathematical Sciences – Chennai, IN
- Matthias Mnich
Universität des Saarlandes, DE
- Anil Nerode
Cornell University, US
- Rolf Niedermeier
TU Berlin, DE
- Christophe Paul
CNRS, Univ. Montpellier II, FR
- Geevarghese Philip
MPI für Informatik – Saarbrücken, DE
- Marcin Pilipczuk
University of Warsaw, PL
- Michal Pilipczuk
University of Bergen, NO
- Venkatesh Raman
The Institute of Mathematical Sciences – Chennai, IN
- Fahimeh Ramezani
MPI für Informatik – Saarbrücken, DE
- Felix Reidl
RWTH Aachen, DE
- Frances A. Rosamond
Charles Darwin University – Darwin, AU
- Peter Rossmanith
RWTH Aachen, DE
- Noy Rotbart
University of Copenhagen, DK
- Ignasi Sau Valls
CNRS, Univ. Montpellier II, FR
- Saket Saurabh
The Institute of Mathematical Sciences – Chennai, IN
- Ildiko Schlotter
Budapest Univ. of Technology & Economics, HU
- Hadas Shachnai
Technion – Haifa, IL
- Somnath Sikdar
RWTH Aachen, DE
- Karolina Soltys
MPI für Informatik – Saarbrücken, DE
- Ulrike Stege
University of Victoria, CA
- Ondrej Suchy
Universität des Saarlandes, DE
- Stefan Szeider
Vienna Univ. of Technology, AT
- Jan Arne Telle
University of Bergen, NO
- Dimitrios M. Thilikos
National and Kapodistrian University of Athens, GR
- Yngve Villanger
University of Bergen, NO
- Magnus Wahlström
MPI für Informatik – Saarbrücken, DE
- Gerhard J. Woeginger
TU Eindhoven, NL
- Anders Yeo
RHUL – London, GB

Report from Dagstuhl Seminar 12261

Putting Data on the Map

Edited by
Stephen Kobourov[1], Alexander Wolff[2], and Frank van Ham[3]

1 **University of Arizona, Tucson, US,** kobourov@cs.arizona.edu
2 **Universität Würzburg, DE,**
http://www1.informatik.uni-wuerzburg.de/en/staff/wolff_alexander
3 **IBM Software Group, NL,** frankvanham@nl.ibm.com

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 12261 "Putting Data on the Map".

Seminar June 24–29, 2012 – www.dagstuhl.de/12261
1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems
Keywords and phrases Information Visualization, Cartography, GIScience, Computational Geometry, Graph Drawing, Cognition
Digital Object Identifier 10.4230/DagRep.2.6.51
Edited in cooperation with Philipp Kindermann

1 Executive Summary

Stephen Kobourov
Alexander Wolff
Frank van Ham

Visualization allows us to perceive relationships in large sets of interconnected data. While statistical techniques may determine correlations among the data, visualization helps us frame what questions to ask about the data. The design and implementation of algorithms for modeling, visualizing and interacting with large relational data is an active research area in data mining, information visualization, human-computer interaction, and graph drawing.

Map representations provide a way to visualize relational data with the help of conceptual maps as a data representation metaphor. In a narrow sense, a map representation of a graph is a contact graph representation where the adjacency of vertices is expressed by regions that share borders. Such representations are, however, limited to planar graphs by definition. We can extend the notion of a map representation to non-planar graphs by generalizing the idea as follows: clusters of well-connected vertices form countries, and countries share borders when neighboring clusters are tightly interconnected.

Information spatialization and cartograms also connect the notions of data with those of maps. Cartograms redraw an existing geographic map such that the country areas are proportional to some metric (e.g., population), an idea that dates back to a paper by Raisz in 1934 and is still popular today. *Spatialization* is the process of assigning two- or three-dimensional coordinates to abstract data points, ideally such that the spatial mapping has much of the characteristics of the original high-dimensional space. Multi-dimensional scaling or principal component analysis are techniques that allow us to spatialize high-dimensional data. Techniques like information landscapes can then be used to convert the resulting two-dimensional coordinates into meaningful three-dimensional landscapes.

Providing efficient and effective data visualization is a difficult challenge in many real-world software systems. One challenge lies in developing algorithmically efficient methods to visualize large and complex data sets. Another challenge is to develop effective visualizations that make the underlying patterns and trends easy to see. And finally, we need to allow users to interactively access, analyze, and filter these patterns in an intuitive manner. All of these tasks are becoming increasingly more difficult due to the growth of the data sets arising in modern applications, as well as due to their highly dynamic nature.

Topics of the Seminar

Graph representations of side-to-side touching regions tend to be visually appealing and have the added advantage that they suggest the familiar metaphor of a geographical map. Traditional maps offer a natural way to present geographical data (continents, countries, states) and additional properties defined with the help of contours (topography, geology, rainfall).

An important difference between drawings of graphs and maps is the following: graphs are usually *drawn on* the plane (using small placeholder symbols for vertices and curves for edges), whereas maps *fill* the plane (or a sufficiently large area). We want to explore this new paradigm.

In the process of data mining and data analysis, clustering is an extremely important step. It turns out that maps are very helpful in dealing with clustered data. There are several reasons why a map representation of clusters can be helpful. First, by explicitly defining the boundary of the clusters and coloring the regions, we make the clustering information clear. Second, as most dimensionality-reduction techniques lead to a two-dimensional positioning of the data points, a map is a natural generalization. Finally, while it often takes us considerable effort to understand graphs, charts, and tables, a map representation is intuitive, as most people are very familiar with maps and even enjoy carefully examining maps.

When designing algorithms to produce maps for abstract data, we can leverage cartography and GIS expertise in order to answer critical questions such as how regions and geographic networks (such as street or river networks) are represented on traditional geographic maps, how they are labeled (an interesting problem in its own right) and how (boundary) lines are simplified (through a process called *cartographic generalization*), or even schematized, in order to focus on important features. Therefore, participation of people from several diverse areas is essential for the success of our seminar.

Aims of the Seminar

The main goal of this seminar was to foster co-operation between researchers with interests in data visualization coming from the information visualization, human-computer interaction, data mining, graph drawing, and GIS communities. The specific aims of the Dagstuhl seminar were:

1. To bring together researchers working on visualization from a theoretical point of view (graph theory, computational geometry), from a practical point of view (information visualization, HCI), and from a map point of view (cartography, GIS).

2. To identify specific theoretical and practical problems that need to be solved in order to make it possible to create full-fledged conceptual maps as an interactive and scalable data-representation metaphor and to begin working on these problems at the seminar.
3. To formulate the findings as a first step to the solutions of the problems under consideration and to define future research directions.

In order to promote the communication and cooperation between the diverse set of participants, we used a non-traditional format, which included survey presentations, open problem sessions, demo sessions, open mic sessions, problem solving sessions, as well as an exhibition of map-based visualizations. The exhibition entitled "Beyond the Landscape" was organized by seminar participant Maxwell Roberts and by seminar co-organizer Alexander Wolff. It was opened on June 26 by the scientific director of Schloss Dagstuhl, Prof. Reinhard Wilhelm.

Achievements of the Seminar

The achievements in the seminar were numerous and varied. Some of the more important ones can be summarized as follows:

1. On Monday and Tuesday, we enjoyed five survey lectures; see Section 3. Jason Dykes discussed geographic data visualization. Sara Fabrikant presented the cartographic and geovisual perspective. Stephen Kobourov talked about visualizing relational data with the help of the map metaphor. Stefan Felsner illustrated connections with geometry and graph theory. Falko Schmid discussed maps and the interaction with geographic data on small mobile devices. Beyond the survey lectures, a highlight of the seminar was the Friday morning lecture by psychology and perception expert Barbara Tversky; see Section 3.6.
2. We also had a number of stimulating presentations and demos of new software. In particular, new approaches to the layout of large and/or dynamic graphs as well as new visualization paradigms were presented.
3. A number of relevant open problems were formulated early in the seminar and working groups formed around related open problems. The groups then worked by themselves; formalizing and solving their specific theoretical and practical challenges. Below is a list of the working group topics.

   a. Geometric properties of cartograms; convex cartograms
   b. Evaluation of maps and graphs
   c. Metro map visualization
   d. Semantic word cloud visualization
   e. Edge bundling problems
   f. Multi-dimensional temporal data on maps
   g. Map distortion based on (dis)similarity
   h. Work flow for creating maps out of relational data
   i. Maps based on space-filling curve ordering
   j. Multi-scale map generalizations

The last three days of the seminar were dedicated to working group effort. Several of the groups kept their focus on the original problems as stated in the open problem session, while other groups modified and expanded the problems; see Section 4. On the last day of the seminar we heard progress reports from all but two of the groups. We are expecting several research publications to result directly from the Seminar.

Arguably the best, and most-appreciated, feature of the seminar was the opportunity to engage in discussion and interactions with experts in various fields with shared passion about maps. The aforementioned exhibition “Beyond the Landscape” made topics of the seminar visible and raised new questions.

In summary, it is our impression that the (56!) participants enjoyed the great scientific atmosphere offered by Schloss Dagstuhl and profited from the scientific program. We are grateful for having had the opportunity to organize this seminar. We thank Philipp Kindermann for helping us to put this report together.

2 Table of Contents

3 Overview of Talks

3.1 (Geo)Visualization at the giCentre

Jason Dykes (City University London, GB)

Joint work of Wood, Jo; Slingsby, Aidan

In my survey talk I drew upon a number of examples of geo and information visualization to explore. The objective was to demonstrate and situate the giCentre approach to putting data on various kinds of maps and map like graphics.

I began by showing how the cdv [1] and panorama [2] applications used linked graphics to relate spatial, semi-spatial and aspatial views of data.

I then drew attention to the different emphases in 'mapping' as we try to combine aspects of the geographic and aspects of the statistical or structural information in our graphics. Examples included aspatial bar charts, 1-dimensional spatial bar charts, choropleth alternatives, etc.

This can be depicted in a 'back of the envelope' sketch showing the trade-off between positional geography and statistical / structural information; see Figure 1. The objective of much of the design activity involved in 'putting data on the map' is to identify viable positions above the line depicting the trade-off. How can we show the geography that we need to achieve the tasks for which our 'maps' are designed whilst providing adequate statistical and / or structural information to support this activity?

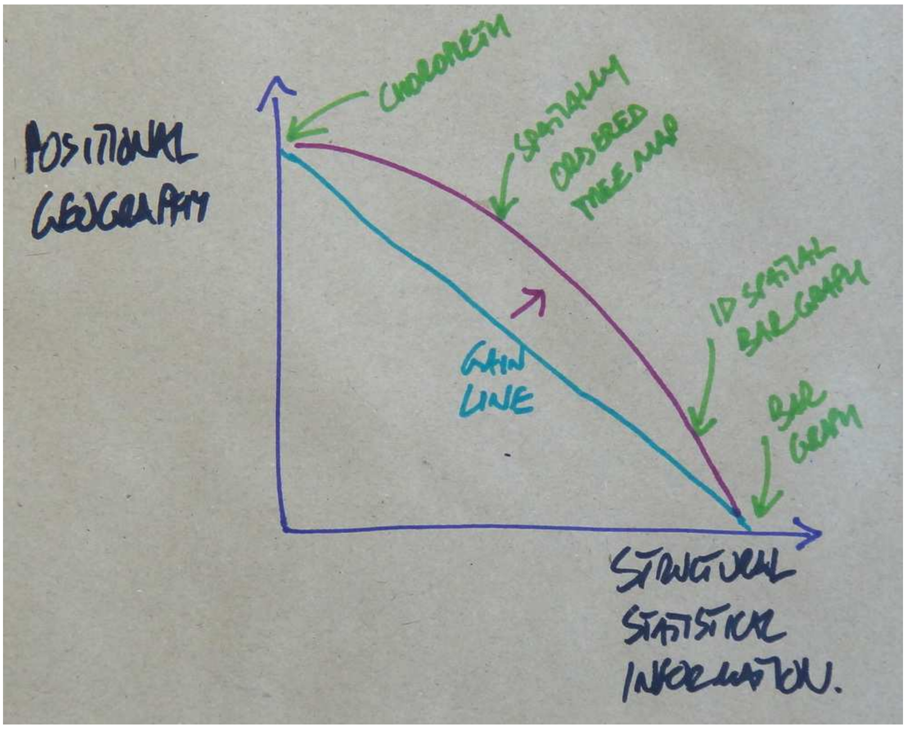

Figure 1 Positional Geography vs. Structural / Statistical Information. Ways in which maps and statistical graphics address this trade-off and ideas for crossing the gain line to show both effectively.

Figure 2 A spatially ordered TreeMap of postcodes: London postal areas, districts, sectors and units are sized by the number of postal addresses.

Spatially ordered TreeMaps [11] were introduced as a means of adding geographic information to hierarchical representations of data and establishing a potentially useful position along or beyond this line; see Figure 2.

Distortions in the mapping between geographic space graphic space were discussed and ways of visualizing and addressing them introduced with a focus on Bernhard Jenny's work with MapAnalyst [5, 6, 7].

Difficulties in associating places with geographic spaces and the personal, emotional and uncertain nature of place were introduced [8] along with some ways of using community contributed information to gain some insights into these [4, 9].

These perspectives lead to questions about what we are mapping. Many of these are open:

- How much geography do we need in our maps?
  - Which maps? Which spaces / places?
- Can we learn new geographies?
  - Who? Which geographies? How personalised are these?
- Can effective symbolism and interactive cartography help us with this learning process?
  - How? Who? When?!

I showed how at the giCentre we have been using hierarchical grids to show information about the London Cycle Hire scheme and bias in the London elections [13].

BikeGrid shows current and historical data relating to the stations comprising the London Cycle Hire Scheme in a semi-spatial grid view with geographic clues and animated transitions [3] to help orientate the map reader; see Figure 3.

BallotMaps show voting patterns in semi-spatial and aspatial views with animated transitions [3] to help orientate the map reader; see Figure 4.

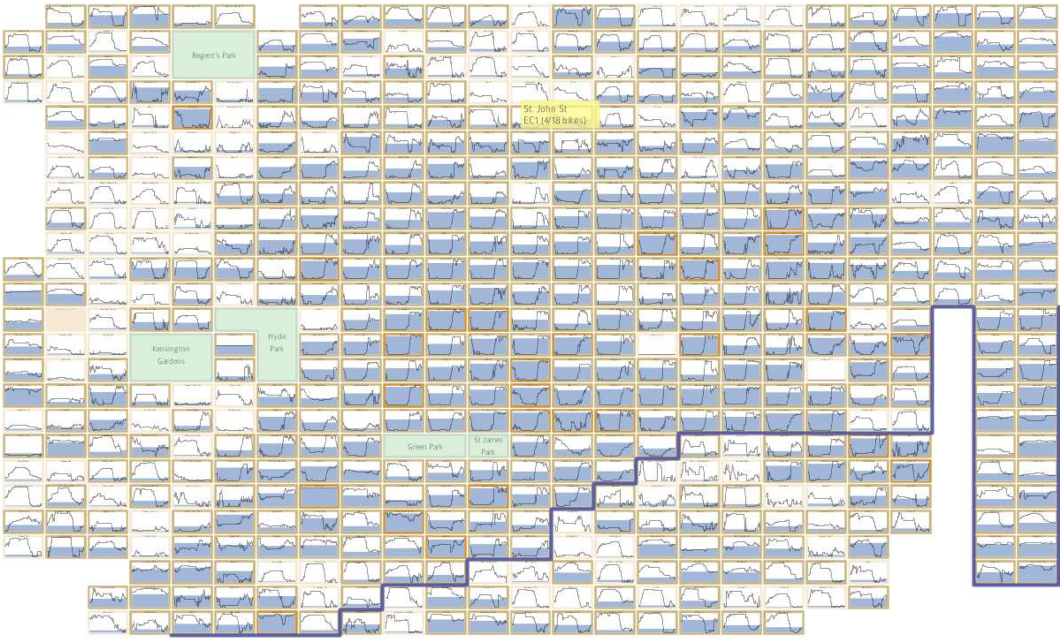

Figure 3 BikeGrid uses an semi-spatial view to aid geographic comparisons of bike station capacity over time.

Figure 4 A BallotMap showing the boroughs of London arranged in a semi-spatial configuration. Each borough is equally sized and split and ordered by party, position of candidate on the ballot paper within party and electoral success and coloured by party and electoral success.

Using these semi-spatial geographies at multiple levels of the graphic hierarchy enables us to create maps within maps. This enables us to see the flows between origins and destinations through OD Maps [12].

Each of these views can be described using the hierarchical visualization expression language HiVE [10].

I concluded by contending that crossing the 'gain line' (see Figure 1) involves:

- Cognitively informed symbolism
- Task, user and data dependent solutions
- Effective (novel?) interaction
- Learning
- Ideation with data and technology
- Evaluation with informed users

To put data more usefully on more effective maps it seems to me that we need to know more about crossing the gain line:

- When and where this is achieved
- Computational and numeric approaches to support this
- Effects of learning
- Means of supporting learning
- With a variety of: phenomena; people; data sets; tasks
- In a range of application domains

References

1 J. A. Dykes. Cartographic visualization: Exploratory spatial data analysis with local indicators of spatial association using Tcl/Tk and cdv. *The Statistician*, 47(3):485–497, 1998.

2 J. A. Dykes. An approach to virtual environments for fieldwork using linked geo-referenced panoramic imagery. *Computers, Environment and Urban Systems*, 24(2):127–152, 2000.

3 J. Heer and G. Robertson. Animated transitions in statistical data graphics. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1240–1247, 2007.

4 L. Hollenstein and R. Purves. Exploring place through user-generated content: Using flickr tags to describe city cores. *Journal of Spatial Information Science*, 1:21–48, 2010.

5 B. Jenny. Geometric distortion of schematic network maps. *Bulletin of the Society of Cartographers*, 40:15–18, 2006.

6 B. Jenny, T. Patterson, and L. Humi. Graphical design of world map projections. *International Journal of Geographical Information Science*, 24(11):1687–1702, 2010.

7 B. Jenny, A. Weber, and L. Hurni. Visualising the planimetric accuracy of historical maps with mapanalyst. *Cartographica*, 42(1):89–94, 2007.

8 D. R. Montello, M. F. Goodchild, J. Gottsegen, and P. Fohl. Where's downtown?: Behavioural methods for determining referents of vague spatial queries. *Spatial Cognition and Computation*, 3(2–3):185–204, 2003.

9 T. Rattenbury and M. Naaman. Methods for extracting place semantics from flickr tags. *ACM Transactions on the Web*, 3(1):1, 2009.

10 A. Slingsby, J. Dykes, and J. Wood. Configuring hierarchical layouts to address research questions. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):977–984, 2009.

11 J. Wood and J. A. Dykes. Spatially ordered treemaps. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1348–1355, 2008.

12 J. Woods, J. A. Dykes, and A. D. Slingsby. Visualization of origins, destinations and flows with od maps. *The Cartographic Journal*, 47(2):117–129, 2010.

13 J. Woods, A. D. Slingsby, and J. A. Dykes. Visualizing the dynamics of london's bicycle hire scheme. *Cartographica,*, 46(4):239–251, 2011.

3.2 Connecting the Dots: A Cartographic Perspective

Sara Fabrikant (Universität Zürich, CH)

For more than 5000 years cartographers have systematically transformed collected (commonly multivariate) spatial data into a two-, three- or four-dimensional visuo-spatial displays. This process is typically performed by applying scientific (i.e., systematic, transparent, and reproducible) cartographic design methods, as well as aesthetic expressivity. In recent years, cartographers and GIScientists have become involved in extending geographic concepts and cartographic design approaches to the depiction of massive, non-geographic data archives. These so-called information spaces also incorporate explicit geographic metaphors with the intention to create a graphic representation that is easier to comprehend for information seekers.

In this presentation I propose an empirically validated design framework for the construction of cartographically sound spatialized network displays based on spatial metaphors. As empirical studies on spatialized networks suggest, basic geographic principles and cartographically informed design guidelines enable information designers to not only construct conceptually robust and usable semantic network spaces, but also allow information seekers to more efficiently extract knowledge buried in large digital data archives.

3.3 From Data to Maps

Stephen Kobourov (University of Arizona, Tucson, US)

Main reference E. Gansner, Y. Hu, S. G. Kobourov, "Visualizing Graphs and Clusters as Maps," IEEE Computer Graphics and Applications, vol. 30, no. 6, p. 54–66, 2010.
URL http://www.cs.arizona.edu/~kobourov/PROJECTS/maps.html

Information visualization can be invaluable in making sense out of large data sets. However, traditional visualization methods often fail to capture the underlying structural information,clustering, and neighborhoods. Our approach for visualizing relational data as a map provides a way to overcome some of the shortcomings with the help of the geographic map metaphor. While graphs, charts, and tables often require considerable effort to comprehend, a map representation is more intuitive, as most people are very familiar with maps and even enjoy carefully examining maps. The effectiveness of the map representation algorithm is illustrated with applications in recommendation systems for TV shows, movies, books,and music. Several interesting and challenging geometric and graph theoretic problems underlie this approach of creating maps from graphs. Specifically, we review recent progress on contact representations,rectilinear cartograms, and maximum differential coloring.

3.4 Graph Representations: Rectangles, Squares and Prescribed Area

Stefan Felsner (TU Berlin, DE)

Main reference S. Felsner, "Rectangle and Square Representations of Planar Graphs," to appear in "Thirty Essays in Geometric Graph Theory" edited by J. Pach.
URL http://page.math.tu-berlin.de/~felsner/Paper/geom-rep.pdf

In the first part of this survey we consider planar graphs that can be represented by a dissections of a rectangle into rectangles. In rectangular drawings the corners of the rectangles represent the vertices. The graph obtained by taking the rectangles as vertices and contacts as edges is the rectangular dual. In visibility graphs and segment contact graphs the vertices correspond to horizontal or to horizontal and vertical segments of the dissection. Special orientations of graphs turn out to be helpful when dealing with characterization and representation questions. Therefore, we look at orientations with prescribed degrees, bipolar orientations,separating decompositions, and transversal structures.

In the second part we ask for representations by a dissection of a rectangle into squares. We review results by Brooks et al. (1940), Kenyon (1998) and Schramm (1993), and discuss a technique of computing squarings via solutions of systems of linear equations.

In the third part we report on contact representations of planar graphs with regions of prescribed area (cartogram representations). In joint work with Alam, Biedl, Kaufmann, Kobourov and Ueckerdt, we have recently shown that planar triangulations admit such a representation with rectilinear polygons with 8 corners. The proof is based on Schnyder woods and the notion of area universal layouts.

3.5 Removing Data From The Map – How Information Reduction and Tailored Interaction Makes Maps Usable

Falko Schmid (Universität Bremen, DE)

Joint work of Schmid, Falko; Kuntzsch, Colin; Winter, Stephan; Kazerani, Aisan; Preissig, Benjamin;
Main reference F. Schmid, C. Kuntzsch, S. Winter, A. Kazerani, B. Preissig, "Situated Local and Global Orientation in Mobile You-Are-Here Maps," in *Proceedings of the Int'l Conf. on Mobile Human-Computer Interaction (MobileHCI'10)*, pp. 83–92. ACM, 2010.
URL http://dx.doi.org/10.1145/1851600.1851617

In my talk I gave a brief survey of topics related to Small Display Cartography. I covered topics such as transformations, off-screen visualization, map-based interaction, and schematization.

I showed that instead of adding data on the map, one has to consider the TEAR model of map creation: by considering the task T, the environment E, and the agent or user A, it is possible to create a tailored representation R.

3.6 Cognitive Tools (brief talk)

Barbara Tversky (Stanford University, US)

Main reference B. Tversky, "Visualizations of thought," *Topics in Cognitive Science*, 3 (2011) 499–535.
URL http://dx.doi.org/10.1111/j.1756-8765.2010.01113.x

Thought quickly overwhelms the mind and uses the world, as words, gestures,diagrams. Diagrams and sketches map elements and relations of ideas that are spatial or metaphorically spatial to elements and relations in space. Space carries meaning, e.g., vertical is more, good, power, health, strength. Horizontal is more neutral, but reading order provides direction. Elements carry meaning, concrete and abstract. Points are places, ideas; lines connect/relate them in road, social, brain, etc. networks. Gestures use analogous ways of expressing meanings. Both diagrams and gestures (actions) are instrumental to thought. The designed world—on shelves, buildings, streets, etc.—is a diagram created by actions and carries abstract meanings, 1–1, hierarchies, symmetries, embeddings.

4 Working Groups

4.1 A Note on Representing Data on Maps

Mohammed Jawaherul Alam, Walter Didimo, Stefan Felsner, Ferran Hurtado, Marc van Kreveld, Giuseppe Liotta, Pavel Valtr, and Kevin Verbeek

Here we address problems related to cartography, where we are given a planar map and we want to redraw the map so that the face-areas in the map represent some prescribed data. Formally, let M be a planar map and F be the set of faces in M. A weight function $w\colon F \to \mathbb{R}^+$ on the faces assigns a positive weight to the faces of M. We want to redraw M so that each face $f \in F$ has area proportional to $w(f)$. We discussed two approaches for these problems.

In the first approach we first fix a "centroid" point inside each face of M and then we compute a weighted voronoi diagram or power diagram with respect to these centroid points in order to obtain a final map M'.

In the second approach, we consider the following problem: we are given a planar map M drawn inside a fixed outer boundary B and a weight function w on the faces of M. We want to redraw M in such a way the topology and the outer boundary remain fixed (vertices may move around the boundary) and the area of each face f is proportional to $w(f)$. In a restricted version of the problem the given map is convex and we also want to maintain the convexity of the regions.

In both the problems we have some positive as well as negative results. Furthermore, we pointed out open questions and future direction in both these problems. The details can be found at http://www.cs.arizona.edu/~mjalam/dagstuhl_12261/group_report.pdf.

4.2 Computation of Wordles with Semantic Constraints

Sara Irina Fabrikant, Stephen Kobourov, Anna Lubiw, Martin Nöllenburg, Yoshio Okamoto, Claudio Squarcella, Torsten Ueckerdt, and Alexander Wolff

A "wordle" is a visual representation for text data where prominent features of words (e.g., their relative frequency or importance) are shown with font size or color [5]. Wordles are generally regarded as beautiful visualizations for websites and blogs, and are usually employed to show the main contents or topics of a web page.

During the Dagstuhl Seminar "Putting Data on the Map" our working group focused on algorithms and methods to compute wordles with specific semantic and geometric constraints, such as the proximity between words that are related to each other. We strive to create wordles that transfer more information, such as a chronological order of the words or close semantic relations between certain words.

Wordles can be created with a number of online and offline services (see, e.g., [1]). An algorithm for the computation of appealing wordles is described in [4]: it features a heuristic for the computation of wordles and focuses on aesthetic criteria.

We proposed and studied various constraints for the computation of wordles. In all of our settings we apply some simplifications to the underlying problem. First of all, each word (or group of words) is considered as a rectangle. In a wordle no two rectangles overlap. Height and width are given, but sometimes we allow 90-degree rotations.

As a first step we imagined to draw wordles on a time axis, fixing each rectangle by time interval and packing all the rectangles using the minimum height ("time flow wordle"). The formulation is equivalent to a scheduling problem [2, 3], which is NP-hard. We independently found counterexamples for greedy algorithms, sketched a proof for weak NP-hardness in the general case and found algorithms for more constrained cases.

We also tackled the problem of realizing a wordle with semantic proximity between selected pairs of words, i.e., where two words must touch in the wordle if they are given as semantically close to each other. We compared this problem with known problems (e.g., rectangular dual of a planar graph) and focused on specific aspects like the presence of holes between words.

We restricted the problem to proximity relations that are represented by trees and sketched NP-hardness proofs for fixed, as well as, free embedding:

- We sketched an NP-hardness proof reducing from 3-Partition for the case when the embedding of the tree is not given.
- We sketched a weak NP-hardness proof reducing from SubsetSum for the case when the embedding of the tree is given.

We further restricted ourselves to hierarchical trees (directed trees with unique source) and required that every contact between two words must be horizontal with the hierarchically superordinated word being on top.

- We sketched an NP-hardness proof for the case when the embedding of the hierarchical tree is not given.
- We developed a polynomial time algorithm for the case when the embedding of the hierarchical tree is given.
- We extended the polynomial time algorithm to four hierarchies (corresponding to the four sides top, bottom, left and right) with a common root.

We thank Michael Kaufmann for discussions.

References

1 Wordle – beautiful word clouds. http://www.wordle.net.
2 A. L. Buchsbaum, H. Karloff, C. Kenyon, N. Reingold, and M. Thorup. *OPT* versus *LOAD* in dynamic storage allocation. *SIAM Journal on Computing*, 33(3):632–646, 2004.
3 H. A. Kierstead. A polynomial time approximation algorithm for dynamic storage allocation. *Discrete Mathematics*, 87(2-3):231–237, 1991.
4 F. B. Viegas, M. Wattenberg, and J. Feinberg. Participatory visualization with Wordle. *IEEE Transactions on Visualization and Computer Graphics*, 15:1137–1144, 2009.
5 Wikipedia. Tag cloud. http://en.wikipedia.org/wiki/Tag_cloud.

4.3 The Role of Animation and Interaction for Exploring Geotemporal Data

Pierre Dragicevic, Sara Irina Fabrikant, Christophe Hurter, William Mackaness, and Barbara G. Tversky

Visualizing geotemporal data is challenging. Known approaches include
1. static views (e.g., small multiples),
2. animated views (e.g., animated maps) and
3. interactive views (e.g., dynamic queries).

The relative effectiveness of these approaches – and of animation in particular – is subject of debate. Numerous studies have been carried out in various fields but are difficult to compare and generalize. In order to better understand which approach works when and why, we propose to place animation and interaction on two ends of a continuum where the locus of control is shared between the system, the end user, and possibly third-party users. We expect that a study comparing various loci of control (i.e., automatic animation vs. manual exploration vs. showing explorations performed by another user) will help us understand if some animations are ineffective due to their poor design or due to a lack of user control.

4.4 A Taxonomy of Temporal Visualizations Based on Space-Time Cube Operations

Daniel Archambault, Sheelagh Carpendale, Pierre Dragicevic, Christophe Hurter, and Ying Yang

Although there are a number of surveys and textbooks on temporal data visualizations, there is still a need to clearly structure the large amount of previous work in the area. Taxonomies and typologies are essential for reflecting on existing techniques, designing new techniques, and teaching information visualization. A possibly useful classification is based on geometrical operations on a space-time cube. Although the space-time cube has been used as a visualization metaphor, it has never been used to classify visualization techniques. This framework can capture most known techniques for representing dynamic 2D data (e.g., geotemporal data), 2D and 3D visualization techniques alike. The goal of this transversal project is to write a taxonomy and/or a survey article on temporal visualizations based on space-time cube operations or a similar concept.

4.5 A Transdisciplinary Survey of Multiscale Rendering Techniques

Sheelagh Carpendale, Pierre Dragicevic, Christoph Hurter, William Mackaness, and Monika Sester

The idea of displaying data differently depending on scale is an important research topic in many disciplines, including in cartography, HCI and infovis. However, each discipline uses a different terminology and (mostly) keeps ignoring the work from other disciplines. The goal of this project is to connect concept and ideas for a better transdisciplinary awareness and for a faster progress on the problem of multi scale rendering in general. The goal will be to

1. identify disciplines that use multi scale rendering (possibly inviting experts from fields not represented in the Dagstuhl seminar),
2. identify the conceptual overlaps between the terms used in each field (e.g., semantic zooming vs. visual aggregation vs. map generalization),
3. identify the techniques used in each field (e.g., reduction, etc.),
4. connect these techniques conceptually and algorithmically and
5. discuss the possibility of re-using techniques and algorithms across disciplines.

4.6 Putting Maps on a Curve

David Auber, Sergi Cabello, Fabrizzio Frati, Herman Haverkort, Martin Gronemann, Michael Kaufmann, and Ignaz Rutter

We study several problems that arise in the context of drawing clustered graphs with a map metaphor. We prove that it is NP-hard to find representations where all vertices that are adjacent in the input graph touch in the map. Due to this inherent difficulty, we suggest a heuristic approach based on alternating between a 1-dimensional and a 2-dimensional layout problem and study the complexity of the individual steps in this heuristic.

One way to draw a clustered graph $G = (V, E)$ in a map metaphor, where the clustering is given as a tree T whose leaves are the vertices of G, is to choose an embedding of the tree, which determines a linear ordering $\mathcal{L}$ of the vertices of the graph. The map $M(G, \mathcal{L})$ is then generated by laying out the landmass for each leaf along a space-filling curve in 2-dimensional space according to the ordering $\mathcal{L}$. This has the advantage that the clusters given by the tree are automatically connected. However, it gives little guarantees on the relative positions of countries. In particular, adjacent vertices of the graph may be placed far apart in the map. We study the problem of finding linear orderings $\mathcal{L}$ represented by the cluster tree T such that adjacent vertices are close in $M(G, \mathcal{L})$.

4.6.1 Our Work

We study the complexity of finding a linear ordering $\mathcal{L}$ represented by T that optimizes the proximity of vertices that are adjacent in G. We showed that it is NP-complete to decide whether an ordering exists such that landmasses of adjacent vertices touch each other. Hence it is not even possible to efficiently approximate the distances.

On the theoretical side, we consider the more restricted case, where we just put the ordering $\mathcal{L}$ on the real line, trying to minimize optimize proximity of adjacent vertices. Two natural objective functions are to either measure the total length of the edges, when putting points at unit distance on the real line, or to measure the number of crossings when adding the edges of G to the embedding of T without crossing T. The first problem is called LinearArrangement and is well-studied. It is generally NP-hard, admits a $O(\log n)$-approximation and can be solved by dynamic programming for balanced bounded-degree trees [1]. We denote the second problem of minimizing the number of crossings by MinCrossing. It is at least as hard as the closely related tanglegram problem [2]. However, the machinery showing that CrossingNumber is FPT with respect to the number of crossings [3] can be applied to MinCrossing as well.

On the practical side, to circumvent the inherent difficulty to preserve distances on the 2D space-filling curve, we suggest a more heuristic approach. The idea is to alternate between the layout along a 2D space-filling curve (optimizing proximity, possibly violating representability by T) and a more tractable 1D ordering problem that enforces representability by T.

More precisely, we start out with an initial ordering $\mathcal{L}$ represented by T and consider the corresponding map $M(G, \mathcal{L})$ and iterate the following steps.

1. Optimize the 2D layout using a force-directed approach with attracting forces for landmasses that should be close. Call the resulting map M'.
2. Traverse the space-filling curve and collect the landmasses as they occur along the curve. This yields a linear ordering $\mathcal{L}'$ of the landmasses and hence of the vertices of G.
3. Compute an ordering $\mathcal{L}''$ that is represented by T and as similar as possible to $\mathcal{L}'$.
4. Start over with step (1) and map $M(G, \mathcal{L}'')$.

The main issue, aside from choosing suitable forces for step (1), is the implementation of step (3). Inspired by the linear ordering problems mention in the introduction, we suggest three ways to measure the similarity of two linear orderings $\mathcal{L}$ and $\mathcal{L}'$. Namely, either the sum of all displacements, the maximum displacement and the number of transpositions. We call the corresponding problems TotalDisplacement, MinMaxDisplacement and MinTranspositions. We showed that using approaches similar to the one for LinearArrangement [1], one can obtain efficient algorithms for all three problems on balanced bounded-degree trees.

References

1. R. Bar-Yehuda, G. Even, J. Feldman, and J. Naor. Computing an optimal orientation of a balanced decomposition tree for linear arrangement problems. *Journal of Graph Algorithms and Applications*, 5(4):133–160, 2001.
2. K. Buchin, M. Buchin, J. Byrka, M. Nöllenburg, Y. Okamoto, R. I. Silveira, and A. Wolff. Drawing (complete) binary tanglegrams — hardness, approximation, fixed-parameter tractability. *Algorithmica*, 62(1–2):309–332, 2012.
3. K. Kawarabayashi and B. Reed. Computing crossing number in linear time. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing (STOC'07)*, pages 382–390, 2007.

4.7 A Generic Work-Flow for Making Maps out of Graphs

Robert P. Biuk-Aghai, Joe Fowler, Jan-Hendrik Haunert, Petra Mutzel, Frank van Ham, and Marc van Krefeld

GMap [2] is a new approach to visualizing graphs and clusters as maps. Depending on the application, one may have different goals for an alternative map framework to GMap. No one framework is necessarily ideal, and often there will be trade offs to consider. Our proposed framework primarily tries to address the needs of the Wikipedia graph, since GMap (as currently implemented in GraphViz as `gvmap`) does not meet several aesthetic criteria including

- *fixed areas of each country*—based on the logarithm of the size of the topic,
- *logical adjacencies and proximities*—very similar topics should be adjacent or nearby, while unrelated topics need to be sufficiently separated, and
- *contiguous boundaries*—i.e., non-fragmented countries.

The well-defined real-world graph whose map representation (once fully realized to meet its constraints) should have broad appeal, validating our approach (if successful).

4.7.1 Approach and Aims

We focused on producing a recursive rectangular cartogram representation that could potentially meet some specific guarantees, such as area, width, and/or aspect-ratio requirements. We proposed a framework that addresses several of GMap's potential shortcomings (depending on the application), namely,

- *fragmentation*,
- *highly irregular boundaries*, and
- *no control over relative positioning.*

All three shortcomings are consequences of using force-directed/MDS placement with standard node overlap removal to create Voronoi-based maps.

4.7.2 Identified Issues with the GMap Framework

We considered a variety of alternatives including removing edges with high betweenness to unroll a graph once a force-directed algorithm is applied. However, unless we used some type of tree-map or cartogram approach to then take that embedding to produce a map, we were left with four problems inherent to the GMap framework:

1. *How does one modify the node-overlap removal algorithm to respect clustering and relative placement?*
   We want to eliminate overlapping labels that can produce useless fragmentation.
2. *How does one reposition or eliminate points so that the Voronoi diagram has less irregular boundaries?*
   Repositioning points may violate country adjacencies. Having too few points can produce blocky maps, while having too many can yield highly jagged boundaries.

3. *How does one achieve guarantees regarding area or aspect ratio?*
   Current method is to specify font size and shore line depth that has weak bounds.
4. *How does one best ensure having the desired adjacencies and non-adjacencies?*
   Resulting Voronoi cells are highly dependent on the outcome of node overlap removal algorithm where adjacencies can be drastically altered.

4.7.3 Proposed Generic Graph-to-Map Framework

Rather than directly addressing these questions, we opted to employ a non-force directed method using recursive rectangular cartograms, with SPQR-trees used to efficiently extract a suitable planar subgraph (with no separating triangles) and find an embedding to reduce overall edge lengths. This circumvents the four problems above, giving an alternative to the GMap force-directed/MDS paradigm. More specifically, our framework takes a weighted hierarchical clustered graph as input and has the following steps:

1. Extract a planar subgraph with no separating triangles prioritizing edges with low betweenness and greater weights.
2. Using SPQR-trees, greedily determine an embedding that reduces lengths of edges with higher priority.
3. Fully triangulate graph by inserting dummy nodes and edges without creating separating triangles.
4. Given that the graph is now 4-connected, obtain a rectangular dual.
5. Solve for a proportional representation using iterative linear programming for desired areas, aspect ratios and proximity.
6. Recurse on each rectangular region representing a non-trivial cluster.

4.7.4 Open Problems with Possible Approaches

In developing this framework, we were confronted with the following several open problems for which we formulated tentative high-level approaches.

1. *How can SPQR-trees be efficiently used to extract a maximal planar subgraph while avoiding separating triangles?*
   Use an SPQR-tree approach to find a maximal planar subgraph (in terms of $\frac{betweenness}{edge\ weight}$), and then attempt to swap edges and/or expand nodes locally. Given that separating triangles are uncommon, in practice one may always eliminate these.
   Ideally, one would wish to obtain an efficient method to guarantee their removal.
2. *How do you then use the same SPQR-tree data structure to search overall all possible combinatorial embeddings to reduce edge crossings of heavily weighted edges?*
   This problem is akin to crossing minimization, where one can also use various types of integer linear programming techniques to optimally solve. Also, there may be reasonable approximation algorithms that may be suitable.
3. *What is the best strategy of inserting dummy nodes and weighting their incident edges so as to not introduce separating triangles?*
   Insert chains of length 2 in lieu of edges to fully triangulate the graph, where then the dummy node would also be adjacent to the two neighbors of its endpoints.
   Weight the dummy node according to the dissimilarity between endpoints so as to produce a lake or a river whose width would be dependent on the similarity of opposing adjacent countries.

4. *How can we find a sufficiently proportional rectangular representation (permitting recursive subdivision)?*
   We can solve for a desirable proportional representation in terms of area, aspect ratios, and/or adjacencies/proximities with given error bounds as follows:

   a. Fix horizontal segments while allowing vertical segments to shift left or right.
   b. Solve associated linear program for optimal x-coordinates of vertical segments.
   c. Fix vertical segments while allowing the horizontal segments to shift up or down.
   d. Solve associated linear program for optimal y-coordinates of horizontal segments.
   e. Repeat steps 1 to 4 until solutions of steps 2 and 4 converge.

   If the solution space is convex, it may be possible to restate this problem in terms of quadratic equations with linear constraints permitting a polynomial-time solution (not requiring this iterative back and forth) for a given error bound. Typically, though, the solution space is not convex, or even connected.
   Furthermore, a 4-connected graph can have many alternate rectangular representations. For the best result, we may have to perform the iterative shift approach for each of them. Optimizing criteria over rectangular representations was considered by Buchin et al. [1].
5. *What guarantees can we ensure for the final map representation?*
   For this to be viable alternative, we should demonstrate that this method has advantages, preferably with guaranteed constraints in terms of area, aspect ratio, proximities, relative positioning, and/or adjacencies within some fixed error, over the current GMap framework. The advantages should be sufficiently compelling to adopt, compensating for having a more complex implementation than GMap.
   Our method (as proposed) requires using an involved SPQR-tree data structure in conjunction with linear programming techniques, which may not be efficient for large instances. This is in contrast to using a simple force-directed algorithm with Voronoi overlay in the case of GMap, which in its most efficient implementation runs in time proportional to $O(n \log n)$ using a force-directed algorithm using a Barnes-Hut approach with quadtrees to approximate distant force.

4.7.5 Conclusion

If we can obtain a reasonable, time-efficient implementation that can create reasonably proportional maps for most real-world weighted graphs, then we can clearly demonstrate the viability of this proposed framework. Moreover, in developing this framework, we have the opportunity to address several open theoretical problems, whose solutions have practical applications. We can further strengthen our results by comparing such an implementation against existing GMap implementations, and evaluate how well our proposed framework fares both in terms of visual aesthetics and desired constraints.

References

1 K. Buchin, B. Speckmann, and S. Verdonschot. Optimizing regular edge labelings. In U. Brandes and S. Cornelsen, editors, *Proceedings of the 18th International Conference on Graph Drawing (GD'10)*, volume 6502 of *Lecture Notes in Computer Science*, pages 117–128, Berlin, Heidelberg, 2011. Springer-Verlag.
2 E. R. Gansner, Y. Hu, and S. G. Kobourov. Gmap: Visualizing graphs and clusters as maps. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis'10)*, pages 201–208, 2010.

4.8 Area Labeling

Jan-Henrik Haunert, Bernhard Jenny, Philipp Kindermann, Sergey Pupyrev, and Falko Schmid

An instance to the area-labeling problem is given with a geographic region, its name, and a set of obstacles.

- A geographic region may be given as a polygon. Additionally, we consider regions with uncertain boundaries (e.g., as proposed by Jones et al. [5]). A region R with an uncertain boundary is represented with a function $f_R \colon \mathbb{R}^2 \to \mathbb{R}$ that assigns a membership value to every point in the plane; $f_R(p)$ measures to what degree point p belongs to R.
- In our basic problem, each name consists of a single word of multiple letters.
- An obstacle is an object (e.g., a road or a house) that already lies in the map.

A solution should satisfy the following criteria.

- The label should follow a support line ℓ of low curvature that approximates the given region R. More precisely, if R is a polygon, ℓ should approximate the medial axis of the region. If R is given with a membership function f_R, ℓ should cover a preferably large set of points of high membership values.
- The label should preferably be centered on the support line ℓ, that is, the distance between the center point c of ℓ and the center γ of the label should be small.
- Each two consecutive letters must have the same distance. This distance is a variable that we denote by δ.
- The distance δ should be close to d, which is the distance between two consecutive letters if they are evenly spread over the whole extent of ℓ.
- A letter should not be placed on an obstacle; this may be a hard constraint, but we also consider a variant where we aim to minimize a cost function that depends on how many (and which) obstacles become occluded by a letter.

4.8.1 Related Work

The problem of labeling a geographic region has been discussed by multiple authors [1, 2, 3, 4].

Barrault [1] has developed an automatic method that first selects a discrete set of candidates for the support line of the label. Then, for each candidate line ℓ, an optimal label position is searched by explicitly testing a large number of values for three variables, namely the offset of the first letter from the begin of ℓ, the distance between two consecutive letters, and the distance between two consecutive words. Each solution is assessed based on multiple criteria, also considering overlaps between letters and other map objects. The overall best solution is returned.

While the method of Barrault only tests a finite number of solutions, we propose an approach by continuous optimization that for a given support line optimizes over all possible solutions. Furthermore, Barrault has not considered geographic regions with uncertain boundaries and his method for selecting candidates for the support line of the label, which is based on a skeleton of the input polygon, cannot easily generalized to regions with uncertain boundaries. We therefore propose a (heuristic) force-directed method that tries to move a randomly sampled initial support line to favorable points while keeping its curvature low.

4.8.2 Outline of the Algorithm

1. For a given object and a label L of n letters, compute a good set of candidates for the support line, i.e., a set of lines on which L may be placed.
2. For each candidate line ℓ from step 1, find all $\mathcal{O}(n^2m^2)$ solutions (with m being the number of obstacles intersected by ℓ) in which at least two letters touch an obstacle.
3. For each solution from step 2, optimally adjust letter positions by quadratic programming.
4. Return the best solution found in step 3.

For the special case that the region is a polygon and the text is required to be horizontal, we have an efficient sweep-line algorithm that finds all interesting support lines. This replaces our heuristic force-directed method.

References

1 M. Barrault. A methodology for placement and evaluation of area map labels. *Computers, Environment and Urban Systems*, 25(1):33–52, 2001.

2 H. Freeman. Automated cartographic text placement. *Pattern Recognition Letters*, 26(3):287–297, 2005.

3 E. Imhof. Positioning names on maps. *Cartography and Geographic Information Science*, pages 128–144, Oct. 1975.

4 C. B. Jones, R. S. Purves, P. D. Clough, and H. Joho. Modelling vague places with knowledge from the web. *International Journal of Geographical Information Science*, 22(10):1045–1065, 2008.

5 I. Pinto and H. Freeman. The feedback approach to cartographic areal text placement. In *Proceedings of the 6th International Workshop on Advances in Structural and Syntactical Pattern Recognition*, SSPR '96, pages 341–350, London, UK, 1996. Springer-Verlag.

4.9 A Metro Map Problem

Maxwell J. Roberts, Andreas Reimer, Yoshio Okamoto, and Therese Biedl

This problem concerns optimizing a linear schematic of a rail network by identifying the angles most congruent with the structure of the network. By choosing the most appropriate angles, trajectory corrections to the lines can be minimized, and hence these can be presented more simply. However, the use of multiple angles is potentially a source of complexity in its own right. Hence the point of view of automation of schematic map design, is it possible to identify the minimum number of angles necessary in order to provide well-optimized line trajectories, maintaining network topology and without unacceptable levels of topographical distortion.

The group tackled this problem bottom-up, trying to identify the stages that a human designer would go through in schematizing a network and answering the question, trying to provide the straightest lines with the minimum angles. In doing so, a task analysis was created, identifying which specific stages can potentially be addresses using current, or slightly modified computer algorithms, versus stages for which new computer algorithms will need to be developed.

- Identify priority region of the network (typically the Central Business District).
- Apply suitable transformation (to enlarge the center in relation to suburbs).

- Identify difficult regions (dense stations), lines with complex versus simple trajectories.
- Identify local traditions (e.g., key landmarks) and focal points.
- Apply edge straightening routines to line trajectories.
- Straighten edges further in conjunction with adjusting trajectories in line with coherence criteria (parallel lines etc.).
- Identify a grid, rotate approximately to it.
- Snap remaining angles to grid.

In the process of outlining these stages, the following observations were made:

- Identifying and applying coherence principles almost makes the map self-organizing.
- Parallel lines may be more important than precisely intervalled angles.
- The Importance of parallel lines may proportional to edge length, and the number of edges may also be a factor (e.g., it is more important that four long lines are parallel than two short ones).
- Automation may be easier to implement if human design techniques are followed, perhaps permitting complex networks to be schematized for the first time.
- Vienna is an octilinear city, but not Barcelona.
- A very abstract distorting schematic might be acceptable if clear benefits of simplification.

Overall, the following 'next steps' were identified:

- Perform similar analyses for other cities of similar complexity
- If there is no clear solution, revisit edge simplification and coherence transformations
- Identify hierarchy of design criteria
- Develop algorithms to maximize parallel lines on graphs
- Develop algorithms to snap to grid, not a trivial adjustment even for a few degrees

4.10 Showing Dissimilarity Data on Cartographic Maps

Sheelagh Carpendale, Tim Dwyer, Jason Dykes, Nathalie Henry-Riche, Arlind Nocaj, and Bettina Speckmann

Given a dataset containing pairwise dissimilarities for points on a map, we wish to explore various ways to distort or overlay the map in order to best show those dissimilarities in the context of the geography. Our workshop discussions on this topic turned into three separate lines of inquiry that we will pursue further in the coming months:

1. Inspired by the “Wrap/Distort/Cut” project of Van Wijk [1], we take the original geographic placement and map to the surfaces of sphere, torus, etc. Using, for example, SOM, we distort on surface in order to minimize dissimilarity error. We find cut paths through the surface in order to:
   - minimize dissimilarity error in final projection.
   - split geographically uninteresting places (e.g., deserts, oceans)
2. A framework for combining overlays and simple global distortion for showing dissimilarity on maps. We want to experiment with trading-off distortion of the map with overlays (such as “ridges” or “bridges”) between points to show the degree of their similarity in the underlying dataset.

3. There is scope for the development of an entire taxonomy, mapping out the design space around this problem and hopefully leading to a reasoned exploration of possible alternatives for displaying dissimilarity data on the map.

References

1 J. J. van Wijk. Unfolding the Earth: Myriahedral projections. *The Cartographic Journal*, 45(1):32–42, 2008.

4.11 Evaluating Bundling Quality Using Image-Based Techniques

Stephan Diehl, Seok-Hee Hong, Quan Nguyen, Monika Sester, and Alexandru Telea

In the last few years, over a dozen algorithms were proposed to simplify complex graph drawings by *edge bundling*. Edge bundling, at a high level, can be seen as a technique that trades off overdraw for clutter, *i.e.*, generates graph drawings where, on the one hand, several edges share the same screen space (thus, the overdraw), but on the other hand one obtains more empty space between groups of edges, or bundles (thus, the clutter reduction). Such drawings thus can successfully convey the coarse structure of a graph, *e.g.*, the main communication paths in a network.

Although many techniques exist that produce bundled edge drawings with different looks, it is not clear what are the properties of a *good* bundled edge drawing. Moreover, although significant work exists on the aesthetics and desirable properties of straight-line graph drawings, such properties cannot be directly taken over to bundled edge drawings.

One promising new avenue would be to use *image-based techniques* to analyze the quality of a bundled edge drawing. For this, we analyze the final image (in which bundled edges have been drawn) rather than the geometric (polyline) representation of the bundled graph. One of the advantages of this approach is that we can directly use many existing image-processing techniques, such as filtering, segmentation, and edge detection.

Relevant questions identified in this working group are as follows:

1. *What is a bundled edge drawing?*
We need to answer this question before we are able to propose desirable properties thereof as well as quality criteria. We propose an analogy between a bundled graph and an image based on the following elements (in the table below, the left column shows data structures and algorithms for bundled graphs; the right column shows image-processing operations).

Edge bundling drawing	**Image properties**
an edge / a bundle	a pixel / an image segment
the bundling process	image segmentation
bundle smoothing	image blurring
edge curvatures	segment contour smoothness

2. *How to model the quality of an edge bundling?*
Following the above graph-image analogy, we can now encode various bundled-graph quality aspects into the corresponding image characteristics:

Edge bundling quality factors	**Image properties**
smoothness and continuity	texture richness
ink/whitespace ratio	image histogram sharpness
bundling strength	image-edge histogram sharpness
disorder/entropy	image histogram flatness
pattern segregation	image histogram inter-peak distance

3. *How to quantify the quality of an edge bundling?*
Besides modeling the quality of a bundled graph by image properties, we need to *quantify* it. For this, we propose a simple but flexible quality metric B. B should have two ingredients: a distance-metric term (encoding the domain knowledge and user requirements to what a good bundling is); and a distribution sharpness term (encoding how well the bundles are segregated in the image according to the eariler-mentioned distance metric). A simple instance of B would be the ratio between the *bundling strength*, *e.g.* the ink/non-ink ratio in the image, to the image-space difference between the original and bundled drawings.
4. *Optimizing an edge bundling*
Now that we have a model for bundling quality, we can propose techniques to exploit this metric to optimize an edge bundling. Two promising directions are: (a) showing the bundling error B of a given graph, to let users understand how much the drawing deviates from the "ideal" one; and (b) Computing an optimal graph bundling by using (non)parametric optimization techniques for increasing the value of B.

Participants

- Muhammad Jawaherul Alam
Univ. of Arizona – Tucson, US
- Daniel Archambault
University College Dublin, IE
- David Auber
Université Bordeaux, FR
- Peter Bak
IBM – Haifa, IL
- Therese Biedl
University of Waterloo, CA
- Robert P. Biuk-Aghai
University of Macau, CN
- Sergio Cabello
University of Ljubljana, SI
- Sheelagh Carpendale
University of Calgary, CA
- Walter Didimo
University of Perugia, IT
- Stephan Diehl
Universität Trier, DE
- Pierre Dragicevic
Université Paris Sud, FR
- Tim Dwyer
Microsoft – Redmond, US
- Jason Dykes
City University – London, GB
- Sara Irina Fabrikant
Universität Zürich, CH
- Stefan Felsner
TU Berlin, DE
- Martin Fink
Universität Würzburg, DE
- Joe Fowler
Univ. of Arizona – Tucson, US
- Fabrizio Frati
Università di Roma III, IT
- Martin Gronemann
Universität Köln, DE
- Jan-Henrik Haunert
Universität Würzburg, DE
- Herman J. Haverkort
TU Eindhoven, NL
- Nathalie Henry Riche
Microsoft – Redmond, US
- Seok-Hee Hong
The University of Sydney, AU
- Ferran Hurtado
UPC – Barcelona, ES
- Christophe Hurter
ENAC – Toulouse, FR
- Bernhard Jenny
Oregon State University – Corvallis, US
- Michael Kaufmann
Universität Tübingen, DE
- Philipp Kindermann
Universität Würzburg, DE
- Stephen Kobourov
Univ. of Arizona – Tucson, US
- Jan Kratochvil
Charles University – Prague, CZ
- Giuseppe Liotta
University of Perugia, IT
- Anna Lubiw
University of Waterloo, CA
- William A. Mackaness
University of Edinburgh, GB
- Petra Mutzel
TU Dortmund, DE
- Quan Nguyen
The University of Sydney, AU
- Arlind Nocaj
Universität Konstanz, DE
- Martin Nöllenburg
KIT – Karlsruhe Institute of Technology, DE
- Yoshio Okamoto
Univ. of Electro-Communications – Tokyo, JP
- Sergey Pupyrev
Ural State Univ. – Ekaterinburg, RU
- Andreas Reimer
GFZ Potsdam, DE
- Maxwell J. Roberts
University of Essex, GB
- Ignaz Rutter
KIT – Karlsruhe Institute of Technology, DE
- Falko Schmid
Universität Bremen, DE
- Monika Sester
Leibniz Univ. Hannover, DE
- Bettina Speckmann
TU Eindhoven, NL
- Claudio Squarcella
Università di Roma III, IT
- Alexandru C. Telea
University of Groningen, NL
- Barbara Tversky
Stanford University, US
- Torsten Ueckerdt
Charles University – Prague, CZ
- Pavel Valtr
Charles University – Prague, CZ
- Frank van Ham
IBM Software Group, NL
- Marc van Kreveld
Utrecht University, NL
- Kevin Verbeek
TU Eindhoven, NL
- Dorothea Wagner
KIT – Karlsruhe Institute of Technology, DE
- Alexander Wolff
Universität Würzburg, DE
- Jing Yang
University of North Carolina at Charlotte, US