

Dansk Selskab for
DATA
logi

Maj 2016

– fra bestyrelsen

Foredrags- og aktivitetsoversigt:

Fredag den 10. juni 2016:

Kl. 16:00: Generalforsamling

Se indkaldelse inde i bladet

Ca. 16:30: Foredrag

Erik Frøkjær, DIKU: **Peter Naurs forståelse af datalogi som menneskelig aktivitet**

I faget datalogi beskæftiger vi os med data og dataprocesser, sådan som Peter Naur beskrev faget tilbage i 1966. Gennem alle årene var det samlende perspektiv for Peter Naur hans bestræbelser på at forstå fagets aspekter som menneskelig aktivitet. Med det som afsæt lykkedes det gang på gang for Peter Naur at udpege problemstillinger og angrebsvinkler, som på vældig mange områder var banebrydende - eller forud for sin tid.

Ca. 17:30: Specialeprisoverrækkelsen

for bedste datalogiske speciale 2013

Der er i år nomineret tre specialer i konkurrencen om specialeprisen.

Anbefalingerne fra specialevejlederne kan læses i dette blad.

Ca. 18:00 – 19:00: Reception

Ca. 19:00 – 22:00: Jubilæumsmiddag

I bladet findes endvidere oversigt over afholdte arrangementer i efteråret 2015.

På bestyrelsens vegne
Torben Mogensen, redaktør

Oplysninger om Dansk Selskab for Datalogi (afholdte medlemsmøde, adresser, m.v.) kan findes på selskabets hjemmeside <http://www.datalogi.dk>. Forespørgsler bedes sendt til datalogi@datalogi.dk.

Generalforsamling

Hermed indkaldes til ordinær generalforsamling

Fredag den 10. juni 2016 kl 16.00 - ca 16.30

Generalforsamlingen afholdes på DIKU, Universitetsparken 1, 2100 København Ø.

Dagsorden:

1. Valg af dirigent.

2. Formanden aflægger beretning.

3. Fremlæggelse af regnskab.

4. Fastsættelse af medlemsbidrag. Bestyrelsen foreslår uændrede satser:

Studerende: kr. 60

Pensionister: kr. 30

Andre: kr. 150

5. Valg af bestyrelsesmedlemmer

På valg er følgende:

Jacob Nørbjerg, CBS

Birgitte Taagholt, Oracle A/S

Bo Hansen Hansen, IT-Kartellel

Jacob Nørbjerg og Birgitte Taagholt modtager genvalg.

Bestyrelsen skal bestå af 6 medlemmer ifølge vedtægternes §4, men har i det sidste år kun haft 4 medlemmer. Det har ikke været muligt for bestyrelsen at rekruttere nye medlemmer. Det er derfor vigtigt, at interesserede blandt foreningens medlemmer tilmelder sig.

Kandidater kan melde sig til bestyrelsen eller opstille på selve generalforsamlingen.

Alle medlemmer af foreningen kan stille op.

6. Valg af revisorer

På valg er:

Kim Ørtved

Nils Andersen

Nils Andersen modtager genvalg.

Kim Ørtved har meddelt, at han ikke ønsker genvalg.

7. Indkomne forslag

Bestyrelsen foreslår ændringer til vedtægterne

§4 ændres til

"Bestyrelsen består af en formand, en næstformand, en sekretær, en kasserer og op til yderligere to medlemmer. Bestyrelsen konstituerer sig selv."

Begrundelse: Det er vanskeligt at finde kandidater til bestyrelsen og vi skønner ikke, at arbejdsbyrden kræver mere end 4 medlemmer. Banken kræver desuden, at bestyrelsens sammensætning lever op til vedtægterne.

Ny §8

"Foreningen tegnes af formand og kasserer i fællesskab."

Begrundelse: Ændringen skyldes lov om hvidvaskning af penge.

8. Eventuelt

Umiddelbart efter generalforsamlingen starter aftenens foredrag, hvorefter der er prisuddeling, reception og middag.

Vedtægter

- §1 Selskabets navn er Dansk Selskab for Datalogi. Selskabet har til opgave at arbejde for kendskabet til datalogi.
- §2 I selskabet kan optages enhver, der arbejder med praktisk eller teoretisk datalogi. Anmodning om optagelse rettes til et bestyrelsesmedlem.
- §3 Gæster kan overvære selskabets møder efter tilladelse af bestyrelsen.
- §4 Bestyrelsen består af en formand, en næstformand, en sekretær, en kasserer og yderligere to medlemmer. Bestyrelsen konstituerer sig selv.
- §5 Valg af bestyrelsesmedlemmer og revisorer sker på en generalforsamling for to år.
- §6 Selskabets møder ledes af et bestyrelsesmedlem. Der føres en protokol over møder og generalforsamlinger.
- §7 Selskabets regnskabsår er fra 1. maj til 30. april. Den årlige generalforsamling holdes i juni måned. Ekstraordinær generalforsamling indkaldes af bestyrelsen og kan forlanges af 1/5 af medlemmerne. Generalforsamling indkaldes med mindst 14 dages varsel. På den ordinære generalforsamling skal følgende punkter være på dagsordenen:
1. Valg af dirigent.
 2. Formanden aflægger beretning.
 3. Fremlæggelse af det af to revisorer reviderede regnskab for det forløbne regnskabsår.
 4. Fastsættelse af medlemsbidrag.
 5. Eventuelt

Beslutning træffes ved simpel majoritet.

Selskabets vedtægter kan ændres ved en ordinær generalforsamling, hvor 2/3 af de fremmødte medlemmer stemmer derfor, og når medlemmerne ved indkaldelsen er gjort bekendt med forslag til ændringer.

Specialeprisen 2015

Der er i år nomineret 3 specialer til konkurrencen om specialeprisen. Disse er listet nedenfor med anbefalingerne fra specialevejlederne.

Andreas Mathisen (AM) og Søren Krogh Sørensen (SKS), Institut for Datalogisk, Aarhus Universitet

The Effect of Indoor Positioning Performance on Human Activity Recognition

Vejleder: Kaj Grønbæk AU

Resumé af specialet:

I modsætning til udendørs positionering (GPS), der er meget udbredt i mobile Apps, så har indendørs positionering baseret på f.eks. WiFi signalstyrker endnu ikke opnået samme udbredelse. Dette speciales resultater giver et bedre grundlag for at udnytte indendørs positionering i praksis. Der gøres således op med det snævre fokus på positioneringspræcision, som dominerer den nuværende forskning. Det vises gennem omfattende data-analyser, at mere grovkornede indendørs positioneringsmetoder kan anvendes med gode resultater indenfor Human Activity Recognition (HAR). I specialet har AM og SKS sammen med en ph.d.-studerende udviklet en ny metode til automatisk detektering af faser i portørers arbejdsopgaver i det 300.000 kvm store Skejby Hospital. Denne metode er allerede publiceret i IEEE PerCom 2016 under titlen "Task Phase Recognition of Highly Mobile Workers in Large Building Complexes". Dernæst har AM og SKS undersøgt hvilken performance, der kan forventes af forskellige eksisterende positioneringsmetoder implementeret i et stort og dynamisk indendørs miljø. På baggrund af denne komparative analyse har de skrevet artikelmanuskript. Til sidst har de undersøgt i hvilken grad præcis positionering er nødvendig for at kunne løse praktiske HAR problemer i forbindelse med arbejdsopgaveanalyser, ligeledes giver de anvisninger på hvilke parametre man skal undersøge i forbindelse med praktisk anvendelse af indendørs positionering. Dette arbejde er også dokumenteret i et artikelmanuskript.

Begrundelse for indstillingen:

Specialet er et usædvanligt selvstændigt projekt, hvor de studerende har taget fat i flere teknisk vanskelige områder: smartphone sensing, WiFi baseret positionering, machine learning og genkendelse af menneskers aktiviteter (HAR) i indendørs miljøer. De har kombineret komplekse matematiske formalismer med omfattende programmering for at kunne gennemføre analyserne. De har således opnået en dyb forståelse af problemstillingerne og de har anvendt avancerede machine learning metoder til at implementere såvel positionerings- som HAR-løsninger. Specialets to sidste artikelmanuskripter er under indsendelse til publicering. AM og SKS er således både selvstændige, ambitiøse og utroligt arbejdsomme. Deres specialearbejdes kvalitet og omfang ligger langt over det niveau, som blot udtrykkes i et 12-tal. Censor og jeg var helt enige om, at i gamle dage ville dette klart være et 13-tal for et omfang og en kvalitet, der ligger langt ud over det sædvanlige.

Tony Beltramelli, ITU

Deep-Spying: Spying using Smartwatch and Deep Learning

Vejleder: Sebastian Risi, ITU

Resumé af specialet:

Wearable technologies are today on the rise, becoming more common and broadly available to mainstream users. In fact, wristband and armband devices such as smartwatches and fitness trackers already took an important place in the consumer electronics market and are becoming ubiquitous. By their very nature of being wearable, these devices, however, provide a new pervasive attack surface threatening users privacy, among others. In the meantime, advances in machine learning are providing unprecedented possibilities to process complex data efficiently. Allowing patterns to emerge from high dimensional unavoidably noisy data. The goal of this work is to raise awareness about the potential risks related to motion sensors built-in wearable devices and to demonstrate abuse opportunities leveraged by advanced neural network architectures. The LSTM-based implementation presented in this research can perform touchlogging and keylogging on 12-keys keypads with above-average accuracy even when confronted with raw unprocessed data. Thus demonstrating that deep neural networks are capable of making keystroke inference attacks based on motion sensors easier to achieve by removing the need for non-trivial pre-processing pipelines and carefully engineered feature extraction strategies. Our results suggest that the complete technological ecosystem of a user can be compromised when a wearable wristband device is worn.

Begrundelse for indstillingen:

In Tony Beltramelli's Master's thesis titled "Deep-Spying: Spying using Smartwatch and Deep Learning" he demonstrates the practicality of side-channel attacks using motion sensors built-in wearable devices. Tony employs deep recurrent neural networks to process the signal and infer keys typed by the user. The result in his thesis are significant because they show that this type of machine learning algorithm can be applied to perform touchlogging and keylogging attacks with above-average accuracy while removing the need for non-trivial pre-processing pipelines and carefully engineered feature extraction strategies.

His work has received much media attention, and was featured in both international media (Wired UK, Version2.dk, Forbes, Huffington Post, El Pais, Naked Security, Vice, geek.com, Gizmodo, Lifehacker Australia, XDA Developers, Softpedia) and on Danish national TV (TV2 NewScience, 22. January 2016). A paper based on Tony's thesis is currently under review at the International Conference on Machine Learning (ICML 2016).

A demo video of the project is available at <https://youtu.be/ZBwSfvnoq5U>. The public interest in his work is demonstrated by over 50,000 video views since it was uploaded in December 2015.

Kristoffer Aalund Søholm and Sebastian Paaske Tørholm, DIKU

Ordered Finite Action Transducers for High-Performance Stream Processing

Vejleder: Fritz Henglein, DIKU

Resumé af specialet og begrundelse for indstillingen:

Kristoffer Søholm’s and Sebastian Tørholm’s speciale is an impressive *full stack* piece of computer science ranging from developing a fundamental, general transducer model for compositional stream processing to its blazingly fast implementation that hands-down beats comparable tools in expressiveness or performance—especially performance—or *both*. This includes tools such as Google’s state-of-the-art regex processor RE2.

In the KMC project we¹ had just shown theoretical optimality results on streaming implementation and asymptotic performance of grammatically specified input processing by devising algorithms and compilation techniques incorporated in our domain-specific language *Kleenex* that emerged from new connections between type theory, language theory and, by then independently revitalized, transducer theory; and we had achieved practical performance that had already garnered positive attention.

As they compactly and correctly write themselves, Kristoffer and Sebastian have contributed to this effort

“the design and implementation of pipelines and actions in the language, an alternative formalization of Kleenex in terms of our OFAT [Ordered Finite Action Transducer] model, a time bound for the full Kleenex language, the exploration of use cases for Kleenex, and benchmarks against regular expression matching tools” and “optimizations targeting both the compile time and code generation, bug fixes, testing and tool-building for the project”.

Remarkably, during their 5-month project period, in which they started, conducted, wrote up and defended their thesis, they also worked on and contributed their findings to a joint submission to the 43rd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL) 2016, the longest running and one of the highly – if not the most – reputable and selective programming language research conferences. The paper was accepted and presented at POPL in January, 2016.

Some noteworthy contributions from Kristoffer and Sebastian’s thesis didn’t make it into the POPL paper—for specific reasons. In their thesis, they came up with and implemented a number of use cases illustrating the diversity of Kleenex’s applicability: large-scale data conversion, on-line log file rewriting, syntax highlighting, a full (yet only half-page long) IRC processor. And a *Man-in-the-Middle attack* operating on HTML-forms. They write that, a bit disturbingly, “several of Kleenex’s properties make it a good fit for this domain” since it is “easy to express the conditions to do specifically targeted attacks in Kleenex”. That particular use case didn’t make it into the POPL paper on advice by the senior co-author. . .

¹My Ph.D. students Bjørn Bugge Grathwohl, Ulrik Terp Rasmussen and myself.

Another remarkable part of Kristoffer and Sebastian’s thesis that didn’t make it into the POPL paper is their empirical comparison with DReX, a combinatory functional language for streaming string transducers developed by researchers at the University of Pennsylvania with publications at the premier conferences Logic in Computer Science (LICS) 2014 and POPL 2015. Kristoffer and Sebastian devised Kleenex specifications that are (objectively) much more compact than the corresponding DReX specifications and (arguably) considerably more natural. They found that the DReX implementation was limited to processing only small inputs (less than 2 MB), where DReX ran for a couple of minutes. Kleenex took 10-20 *milliseconds*. Since DReX was intended more as a proof of concept, it was not included in the empirical comparisons in the POPL 2016 paper.

Kristoffer and Sebastian’s thesis report is a 100 pages: self-contained, illustrative and kept in a pleasantly compact style of exposition deserving of the topic at hand. It speaks through its final results.

Some aspects of how the work evolved are worth hinting at, though. I had given a talk entitled “Towards regular expression parsing at 1 Gbps ” at a Dagstuhl seminar, where I speculated that various techniques such as multistriding had the potential of pushing Kleenex performance to 1 Gbps on a single core, at a time when Kleenex was operating at a sustained rate of about 200 Mbps.² Shortly after Kristoffer and Sebastian had started on their thesis project I had to change the title of my talk: They *doubled* the sustained performance of Kleenex in a jiffy by devising clever systems level I/O processing.³

Another hidden aspect of the thesis is the *large and significant amount* of experimental work that went into it. During a meeting I had conjectured that specializing the bit string operations to specific string lengths might yield a significant performance improvement, something that required developing a full data flow analysis for automatically generated streaming string transducers produced in the bowels of a both advanced and necessarily complicated implementation developed by Ulrik Terp Rasmussen and Bjørn Grathwohl; and devising and implementing bit string specialization for it, to boot. The *following* week Kristoffer and Sebastian were back and reported, without much ado, that “yes, data flow analysis was a useful optimization and no, bit string code specialization didn’t yield significant speed improvements” supported by design, implementation, evaluation and benchmarking of the corresponding optimizations and back-of-the-envelope theoretical arguments in their support. It was almost unsettling to see such an amount of work delivered with a simple “here is the result”.

²Scaling out in *parallel* is an additional, orthogonal path to performance optimization; it multiplies, rather than subsumes, sequential performance optimizations.

³That wasn’t the kind of techniques I had thought of, but it meant that there was a potential for improvements beyond what I had previously thought reasonably possible on mundane stock hardware: The talk’s title became “Towards regular expression parsing at 5 Gbps”.

Tidligere arrangementer

I efteråret 2015 er der afholdt følgende aktiviteter:

30. november 2015: Future of our Past: Dis-Membering Social Memory into Bits and Bytes

Assistant professor Attila Marton, Department of IT Management, CBS

Memory invokes the dynamic interplay between forgetting and remembering as the basis for the construction of a past aligned to what the future is expected to bring.

Propelled by the immense capacities to store data, the ideal of information technologies as a remedy against forgetting seems to be an overstatement considering the fact that digital media are based on the most radical classification possible – the classification of everything and anything into only two classes of 0 and 1. Digital memory is the dis-membrance of singular events into binary digits leading to the counter-intuitive conclusion that the storing of binary-based data itself is a way of forgetting.

Re-membrance, in turn, is the computational reconstruction of artefacts composed out of binary digits. Thus conceived, binary-based digital media are very ill-fitted for mnemonic purposes, since, for the first time, the conservation of a communication medium as such is not enough for societal remembering.

28. oktober 2015: Præsentation af tre af de nominerede specialer til specialeprisen 2014

Eva Rotenberg, DIKU (vindespeciale): Maintaining alterable planar embeddings of dynamic graphs.

Peter Urbak, DAIMI, AU: A Formal Study of Moessner's Sieve

Jacques Holst og Martin Kjeldsen, ITU: Identifying Payment Instrument Characteristics through the use of Repertory Grid

Adresser

Dansk Selskab for Datalogis hjemmeside er <http://www.datalogi.dk>. Bestyrelsen kan kontaktes på E-post datalogi@datalogi.dk.

Bestyrelsen

Formand

Lektor Jacob Nørbjerg
Institut for Datalogi,
Aalborg Universitet
Selma Lagerlöfs Vej 300
9220 Aalborg Ø
Tlf: 9940 3546
E-post: formand@datalogi.dk

Næstformand og Webmaster

Bo Hansen Hansen
C. F. Richs Vej 2, 4. sal
2000 Frederiksberg
Tlf: 60 68 47 27
E-post: bo@datalogi.dk

Kasserer

Cand. Scient. Birgitte Tågholt
Oracle Danmark
Metalbuen 66
2750 Ballerup
Tlf: 22 28 65 00
E-post: kasserer@datalogi.dk

Redaktør

Torben Ægidius Mogensen
Datalogisk Institut
Københavns Universitet
Universitetsparken 5
2100 København Ø
E-post: torben@datalogi.dk

**Dansk Selskab for
DATALOGI**