10:00 Coffee & Tagging
10:30 Opening Remarks and Greeting
Shlomi Dolev and Ehud Gudes

10:40 A Pattern Based Approach for Secure Database Design
Jenny Abramov, Ben-Gurion University, Israel

Abstract: The perception of threats in the domain of database systems, and provability information was shown to be extremely useful for different purposes, such as threat management, cost, security, compliance and probabilistic databases. The advances in this area are generally achieved separately, not concurrently, for query languages of various expressivity. What was missing was a comprehensive approach, and this is what provability information can provide.

We introduce in this work a novel approach for pattern matching that aims at efficiently processing patterns which comprise all levels of complexity. We present a formal model for pattern matching and demonstrate its usage in various applications and compare techniques for complex pattern types, taking various semantic interpretations into account. A cost model is presented, balancing processing history and event outcomes according to user preference. Pattern cost is then generated using simulation-based techniques. This work advances the state-of-the-art by analyzing complex event processing logics and by using explicit means to optimize elements that were considered ‘black box’. One empirical study employs encouraging results, with improvement gaps of up to 8.5% relative to the optimal solutions that are used in the current state-of-the-art systems.

11:15 Private Data Analysis
Kobi Nissim, Ben-Gurion University and Microsoft, Israel

Abstract: This paper presents a novel approach to the problem of securing data analysis. A notion that has entered in this line of research is Differential Privacy. I will motivate differential privacy, show basic techniques for constructing differentially private analyses, and relate differential privacy with other notions of theoretical computer science. No prior background will be assumed.

11:55 Coffee Break

12:10 Towards Provenance for SQL (and beyond!)
Daniel Deutch, Ben-Gurion University, Israel

Abstract: The processing of events allows enterprises and individuals to move from reactive mode, where the computerized system reacts to an event or a detected situation based on pre-defined policies, to general reactive mode, where the computerized system reacts to a user-defined policy. The challenge here is to provide a cost-effective way to generate such policies, while also enabling effective monitoring of the policy execution.

12:50 Towards Proactive Event-driven Computing
Ofer Ezion, IBM Haifa Research, Israel

Abstract: The processing of events allows enterprises and individuals to move from reactive mode, where the computerized system reacts to the user's request to reactive mode, where the computerized system reacts to an event or a detected situation based on pre-defined policies. The challenge here is to provide a cost-effective way to generate such policies, while also enabling effective monitoring of the policy execution.

13:30 Lunch Break

14:15 Keynote lecture – Map Reduce And Its Children
Jeff Ullman, Stanford University, USA

Abstract: MapReduce is a programming system that is widely used for running distributed data analysis. It is a good choice for large-scale data analysis because it provides a simple, high-level interface for programmers, and it is easy to build on top of existing software. It is also a good choice for large-scale data analysis because it provides a simple, high-level interface for programmers, and it is easy to build on top of existing software.

15:15 Parallel XML Query Algorithm
Lila Shnaderman, Technion, Israel

Abstract: We present the Parallel Path stack algorithm (PPS) and the Parallel Twig stack algorithm (PTS) that are novel and efficient algorithms for matching XML query twig patterns in a parallel multi-threaded computing platform. PPS and PTS are based on the PathStack and TwigStack Algorithms of [11][12][13][14]. These algorithms employ a sophisticated technique for filtering string matching to reduce the size of the search tree. We provide an efficient algorithm for matching XML twig patterns in a parallel multi-threaded computing platform.

16:05 On Provenance Minimization
Yael Amsterdamer, TAU, Israel

Abstract: Provenance information has been proved to be very effective in capturing the computational process performed by queries, and has been used extensively in a number of different applications. In this work, we study the problem of minimizing the size of provenance information, namely the part of provenance that appears independently of the query plan that is run. This provenance can be informative as it describes the part of the computation that is not captured by the query plan. It is useful for capturing the relationship between the query and the computation that the user requested.

16:45 Efficient Entity resolution with MFBlocks
Batyka Kenig, Technion, Israel

Abstract: Entity resolution is the process of discovering groups of tuples that correspond to the same real-world entity. In order to achieve high-quality results, it is essential to effectively compute and analyze the provenance of the results. We are interested in understanding the value of provenance in this context and how it can be used to improve the quality of the results. Our work demonstrates that considering the provenance of the results can greatly improve the quality of the entity resolution process.

17:20 Diversified Recommendations for Semantic Data
Rubi Boim, TAU, Israel

Abstract: The research considers a popular class of recommender systems that are based on Collaborative Filtering (CF) and proposes a novel technique for diversifying the recommendations that they give to users. Our method is based on a novel approach to diversification that is inspired by the problem of generating diverse sets of recommendations and by the idea of using randomness to improve the quality of the recommendations. We demonstrate the effectiveness of our approach on a real-world dataset and show that it can improve the quality of the recommendations significantly.

17:40 End of Database Day

We present the Parallel Path stack algorithm (PPS) and the Parallel Twig stack algorithm (PTS) that are novel and efficient algorithms for matching XML query twig patterns in a parallel multi-threaded computing platform. PPS and PTS are based on the PathStack and TwigStack Algorithms of [11][12][13][14]. These algorithms employ a sophisticated technique for filtering string matching to reduce the size of the search tree. We provide an efficient algorithm for matching XML twig patterns in a parallel multi-threaded computing platform.