## Summer Internships 2011

<table>
<thead>
<tr>
<th>Company</th>
<th>Project</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP</td>
<td>Research and development of change correlation algorithms in a massively changing topological environment</td>
<td>3</td>
</tr>
<tr>
<td>Philips Medical</td>
<td>Software developers</td>
<td>4</td>
</tr>
<tr>
<td>SAP</td>
<td>Practical and scalable algorithms for repair and reconciliation in consistent mapping networks</td>
<td>5</td>
</tr>
<tr>
<td>Microsoft</td>
<td>Development engineer for advanced imaging technologies group (AIT) - XBox division</td>
<td>6</td>
</tr>
<tr>
<td>Microsoft</td>
<td>Bing mobile Israel group</td>
<td>6</td>
</tr>
<tr>
<td>Microsoft</td>
<td>identity and access in the cloud-anywhere access group</td>
<td>7</td>
</tr>
<tr>
<td>Microsoft</td>
<td>IEB knowledge-services Israel group – statistics proficiency</td>
<td>8</td>
</tr>
<tr>
<td>Microsoft</td>
<td>IEB knowledge-services Israel group – machine learning proficiency</td>
<td>8</td>
</tr>
<tr>
<td>Intel</td>
<td>A convertor from bit-vector logical model to OpenAccess model</td>
<td>9</td>
</tr>
<tr>
<td>Intel</td>
<td>Adding incremental capability to a memory save/restore engine</td>
<td>9</td>
</tr>
<tr>
<td>Intel</td>
<td>Conduct accuracy study for cell-based reliability checks</td>
<td>10</td>
</tr>
<tr>
<td>Intel</td>
<td>Conduct study on the most efficient stimuli for reliability verification</td>
<td>10</td>
</tr>
<tr>
<td>Intel</td>
<td>Integrating propriety core simulator into open source emulator</td>
<td>11</td>
</tr>
<tr>
<td>Intel</td>
<td>Research and development in Constraint Programming: improvement of arithmetical data domain representation and constraint propagation techniques</td>
<td>12</td>
</tr>
<tr>
<td>Intel</td>
<td>High-level parallelization of test-generation algorithms for multi-core validation</td>
<td>12</td>
</tr>
<tr>
<td>Intel</td>
<td>Develop OpenCL device agent for the graphic processor device</td>
<td>13</td>
</tr>
<tr>
<td>Intel</td>
<td>Develop prototype of OpenCL debugger agent on Linux</td>
<td>13</td>
</tr>
<tr>
<td>Intel</td>
<td>Turning core simulator to multithreading</td>
<td>14</td>
</tr>
<tr>
<td>IBM</td>
<td>Design and development of a ‘property-based DBs’</td>
<td>15</td>
</tr>
<tr>
<td>IBM</td>
<td>Revealing the true ‘visual anatomy’ of enterprise-models</td>
<td>16</td>
</tr>
<tr>
<td>IBM</td>
<td>Community analytics research in Web 2.0 and social media technologies group</td>
<td>17</td>
</tr>
<tr>
<td>IBM</td>
<td>Research and development of new methods for model matching and management using the IBM Jazz platform</td>
<td>18</td>
</tr>
<tr>
<td>IBM</td>
<td>Social stream analysis research in web 2.0 and social media technologies group</td>
<td>18</td>
</tr>
<tr>
<td>IBM</td>
<td>Visual analytics researcher for spatiotemporal data</td>
<td>19</td>
</tr>
<tr>
<td>IBM</td>
<td>Modeling maintenance process in software engineering</td>
<td>20</td>
</tr>
<tr>
<td>IBM</td>
<td>Software testing tools and methodology</td>
<td>20</td>
</tr>
<tr>
<td>IBM</td>
<td>Performance improvements</td>
<td>21</td>
</tr>
<tr>
<td>IBM</td>
<td>Advanced layout analysis of complex models for additional info contact</td>
<td>21</td>
</tr>
<tr>
<td>IBM</td>
<td>Develop a visual query language for enterprise information retrieval</td>
<td>22</td>
</tr>
<tr>
<td>IBM</td>
<td>Learning instruction stream behavior to improve random test program generators</td>
<td>23</td>
</tr>
<tr>
<td>IBM</td>
<td>Running a Floating-Point Test-Suite on Hardware</td>
<td>24</td>
</tr>
<tr>
<td>-----</td>
<td>---------------------------------------------</td>
<td>----</td>
</tr>
<tr>
<td>IBM</td>
<td>The relation between bugs and coverage</td>
<td>25</td>
</tr>
<tr>
<td>IBM</td>
<td>Exploring parallel techniques to solve the SAT problem</td>
<td>25</td>
</tr>
<tr>
<td>IBM</td>
<td>Combining many stream specifications to a single test</td>
<td>26</td>
</tr>
<tr>
<td>IBM</td>
<td>Optimized test suite generation for systems engineering models</td>
<td>27</td>
</tr>
<tr>
<td>Zoran</td>
<td>Linux SW developer</td>
<td>28</td>
</tr>
</tbody>
</table>
Research and development of Change correlation Algorithms in a massively changing topological environment

Company: HP Software, HP Labs

General Background: The HP Universal CMDB and Discovery system consistently detects and logs changes to the topological IT environment (changes in servers, network, applications etc...). HP UCMDB Configuration Manager is the application that monitors and visualizes these changes, matches it to compliance rules and controls the process of authorizing/validating it. This project will explore the benefit of data-mining for creating and maintaining smarter change validation policies. Change events will be analyzed to detect patterns of events that occur frequently. Unlike common data-mining tasks, such as clustering and classification, the patterns we seek do not divide the data to groups. Rather patterns may correspond to facets, or a limited view of a change process or a problem in the implementation of a change. The project will include implementing new data-mining algorithms, and applying these algorithms to real-world data provided by HP Software. Algorithm Implementation will be in Java. The project will be co-advised by HP Software and HP Labs

Responsibilities: Research and development of a data mining algorithm as well as sample data generation engine

Required Skills: Knowledge in Java

Full/Part time position: Full 3 month position. Involves part work in HP Yehud office and part in HP Labs office in Haifa

Contact Details: yariv.snapir@hp.com, (HP Software R&D), ruth.bergman@hp.com (HP Labs)
Company: Philips Medical Systems Technologies Ltd

General Background: Participation in the development of medical imaging analysis applications, including clinical algorithms and advanced visualization solutions.

Responsibilities: Developing and optimizing 3D medical imaging algorithms with focus on segmentation and 3D editing tools.

Required Skills: M.Sc in computer science with background in developing high-performance algorithms in C/C++ and image processing or computer graphics experience.

Contact Details: Mark Rabotnikov, Mark.Rabotnikov@philips.com
Practical and Scalable Algorithms for Repair and Reconciliation in Consistent Mapping Networks

Company: SAP Research

General Background: SAP Research Israel is engaged in a Research Project about Consistent Mapping Network called NisB [http://www.nisb-project.eu/]. EPFL (the Group of Prof. Karl Aberer and Zoltan Miklos) and the Technion (the group of Prof. Avi Gal from IEM) participate in this project. The project deals with very large graphs of semantic entities and their relationships. The network contains uncertainty and inconsistencies. We developed novel algorithms, based on optimization and constraint satisfaction computational models that repair the networks thereby making them more consistent and of higher quality.

Responsibilities: We have a naive implementation of several algorithms in the domain that were developed by a research group in EPFL. We would like to re-implment these algorithms, thereby improving them and making them scalable for very large graphs.

Required Skills:
- Advanced programming skills (probably Java)
- Algorithmic approach
- Liking of the ideas of the “Semantic Web”

Full/Part time Position: Full time position for 3 months, in the summer: June – September 2011

Contact Details:
Eliezer Levy, Director, SAP Research Israel, eliezer.levy@sap.com, 054-2277128
Victor Shafran, Senior Researcher, SAP Research Israel, victor.shafran@sap.com, 052-3854883
Microsoft

Development Engineer for Advanced Imaging Technologies group (AIT) - XBox Division

Company: Microsoft (Location: Haifa)

General Background: Advanced Imaging Technologies group (AIT) is part of the Xbox division in Microsoft. It is focused on hardware and software technologies in the Kinect project.

A few words on Kinect: Microsoft’s Kinect for Xbox 360 is the latest motion sensing technology for gaming and entertainment. Instead of a controller, your body becomes the input device. It uses a full body tracking system which is based on 3D sensors and highly sophisticated computer-vision and machine-learning algorithms. Kinect’s aim is to revolutionize the entertainment in the living room. It sold 8 million units in the first 60 days after launch, on its way to become “the fastest-adopted consumer electronic device” (as reported by ReadWriteWeb).

Responsibilities: Prototype development.

Required Skills: The candidate’s mission involves planning and developing new prototypes using cutting-edge computer vision technologies. We are looking for highly motivated, self learning, independent and creative engineers with the following skills:

- The ability to plan thoroughly and implement new ideas.
- Knowledge in image-processing and computer-vision.
- Very good programming skills in C++, Win32 (Matlab – an advantage).

Contact Details: Please send your resume and grade sheet to: ildcjobs@microsoft.com
*Please write in the title of your mail: Summer Internship*

---

Bing Mobile Israel Group

Company: Microsoft (Location: Herzeliya)

General Background: What if people trusted you and let you know their whereabouts in return for interesting insights about themselves? Would you be able to tell them how much time they spend on commute, hiking, biking, and shopping? Would you be able to figure out what is their favorite food? Who are their favorite friends? Would you be able to suggest them places they would find interesting when they visit a town/country for the first time?

Would you be able to know how trendy they are in relation to the rest of the world?
We, at Bing Mobile IL team, believe that we can learn a lot about people by exploring his whereabouts, and provide them with useful information.

Project Description: We invite you to participate in a project that combines exploration of location data, inference and collaborative filtering, to provide a personalized view of users. The project involves experiencing with .Net and Windows Azure technologies in a creative and dynamic environment. New ideas are always welcome in our team.

Required Skills: C# is a big advantage and also experience with .Net technologies

Contact Details: Please send your resume and grade sheet to: ildcjobs@microsoft.com
*Please write in the title of your mail: Summer Internship*

Identity and Access in the Cloud-Anywhere Access Group

Company: Microsoft (Location: Herzeliya)

General Background: Our team develops the identity and access components for Microsoft cloud environments (Windows Azure and Office 365). The identity and access platform built by the team will enable cloud applications to securely identify the client (user & device) and its state (such as location and network conditions) as part of the authentication and authorization processes. The team is part of the Anywhere Access Group (AAG), which has a Microsoft-wide charter to enable users to connect from any device to any application securely.

Project Description: The project will explore new ways in which users can access cloud resources securely and from any location. Access grant will be granular, and made based on the user identity (including the mechanism used to authenticate the user), device identity, and device state.

Required Skills:
- Familiarity with development in managed environments (C# or Java)
- Creative thinking
- Ability to learn quickly new and complex areas
- Good knowledge of networking and security principles

Contact Details: Please send your resume and grade sheet to: ildcjobs@microsoft.com
*Please write in the title of your mail: Summer Internship*
IEB Knowledge-Services Israel Group – Statistics proficiency

Company: Microsoft (Location: Herzeliya)

General Background: If you are passionate about the cloud computing and large scale data systems which will serve millions throughout the world, we are offering a unique and great opportunity for talented students, to help us build and ensure our products meets the highest standards.

We invite you to be part of a very strong team responsible for developing the next generation recommendations and personalization systems for the cloud.

Required Skills:
- M.Sc or Ph.D in Computer Science
- Strong statistical abilities
- Experience in data mining or statistical work with large datasets
- Good coding skills
- Ability to work in a team

Contact Details: Please send your resume and grade sheet to: ildcjobs@microsoft.com
*Please write in the title of your mail: Summer Internship*

---

IEB Knowledge-Services Israel Group – Machine learning proficiency

Company: Microsoft (Location: Herzeliya)

General Background: If you are passionate about the cloud computing and large scale data systems which will serve millions throughout the world, we are offering a unique and great opportunity for talented students, to help us build and ensure our products meets the highest standards.

We invite you to be part of a very strong team responsible for developing the next generation recommendations and personalization systems for the cloud.

Required Skills:
- M.Sc or Ph.D in Computer Science
- Strong Machine Learning knowledge
- Good coding skills
- Ability to work in a team

Contact Details: Please send your resume and grade sheet to: ildcjobs@microsoft.com
*Please write in the title of your mail: Summer Internship*
A convertor from Bit-Vector logical model to OpenAccess model

Company: Intel

General Background: Formal verification is a very important process at Intel aiming to guarantee 100% of the chip functionality. Using world-class engines and solvers, we are building tools and methodologies which guarantee at Intel chips are sold in their best quality. To achieve that, Formal Property Verification (FPV) tools are aiming to check whether a given hardware design functionally satisfies a pre-defined property. Formal Equivalence Verification (FEV) tools are checking the functional equivalence between two given hardware designs. Both tools model hardware design in a logical presentation known as BitVector: this representation gives a flexible representation of both bit level as more abstract representation of vectors. OpenAccess is a public data model for VLSI design development. It is used at Intel to model circuit designs.

Responsibilities: This project is about developing a convertor from a BitVector data structure to OpenAccess data model. This utility involves deep knowledge in data structures. It is very important utility for the above tools as building OpenAccess model is done today in a non-efficient manner.

Required Skills: Good knowledge in data structures. Better to be familiar with performance and memory profiling tools

Full/Part time position: Full

Contact Details: Einat Keren, Einat.keren@intel.com

Adding incremental capability to a memory save/restore engine

Company: Intel

General Background: In order to design chips with billions of devices, Intel designers often use high-capacity tools that run for a long time. Some of these tools implement save/restore capabilities to allow the user to continue running on a separate machine or at a later time, and also to distribute the run across several machines. One way to implement save-restore is by using checkpointing (see here: http://checkpointing.org/).

Responsibilities: In this internship proposal, you will work on improving an internal checkpointing solution to work incrementally. This means that instead of generating a full dump of the memory image of the application, you will only generate a dump for pages of memory that have changed since the last checkpoint has taken place.
**Required Skills:** High motivation, ability to learn quickly, ability to tackle low-level issues related to Linux internals

**Full/Part time position:** Full-time

**Contact Details:** Einat Keren, Einat.keren@intel.com

---

**Conduct accuracy study for cell-based reliability checks**

**Company:** Intel

**General Background:** Tavor is the Reliability Verification (RV) tool developed in our group and used to check chip design for reliability issues such as electro-migration. Recently, we have developed a cell-based design (CBD) mode for this tool, than uses cell characterization to improve performance while maintaining acceptable accuracy. We would like to perform a study of the accuracy of the CBD mode in the tool for two of the stages in the tool – thermal simulation and checking the magnitude of the currents. This follows a successful study of the current calculation on the power nets in the last summer.

**Responsibilities:** Perform an accuracy study of Tavor Thermal and Tavor Checks CBD mode. Provide data that can be used for deciding on the modeling used in this mode to best balance accuracy vs. performance needs. Test suggestions for modeling improvements.

**Required Skills:** Programming (mostly scripts and/or tool such as Matlab) to analyze data and manipulate tool inputs/outputs. Electrical engineering background is applicable.

**Full/Part time position:** full time

**Contact Details:** Einat Keren, Einat.keren@intel.com

---

**Conduct study on the most efficient stimuli for Reliability Verification**

**Company:** Intel

**General Background:** Reliability Verification (RV) deals with several failure phenomena on VLSI products, such as Electro-migration (EM).

To assure that the design meets Intel RV targets we provide our RV tool, which simulates the electrical circuit to find the currents that its resistors bare.
As we are starting a new advanced implementation of this module, we are looking for a definition for the minimal stimuli required to apply on the circuit to find the worst case current on each of its resistor circuits.

**Responsibilities:** Use existing related material and articles, as well as in-house knowledge, to conduct a study that looks for this generation function. The work would have an iterative nature, of choosing a function, checking its results on a simulative environment, and improving according to results.

**Required Skills:** Good programming skills - programming language doesn’t matter, as the output of this work would be an external “off-session” output to the tool, which the tool can read.

Electrical Engineering background is highly applicable.

**Full/Part time position:** Full time.

**Contact Details:** Einat Keren, Einat.Keren@intel.com

---

**Integrating propriety core simulator into open source emulator**

**Company:** Intel

**General Background:** Intel develops and employs a wide range of simulation tools, ranging from detailed physical models describing the manufacturing process to functional models simulating the x86 architecture. This internship will involve work on an internal, high-speed x86 core simulator.

**Responsibilities:** In this internship proposal, you will work on integrating an internal CPU simulator into qemu – a generic and open-source machine emulator. Integration itself will be done in a shared-library manner, so that qemu is not contaminated with Intel-specific IP; the challenge will be to enable the integration while still maintaining good performance for the simulator.

**Required Skills:** Good knowledge of data structures, high motivation, ability to learn foreign code quickly, ability to tackle hard/non-textbook problems

**Full/Part time position:** Full-time

**Contact Details:** Einat Keren, Einat.keren@intel.com
Research and development in Constraint Programming: improvement of arithmetical data domain representation and constraint propagation techniques

**Company:** Intel

**General Background:** Constraint Programming is the core engine of test generation solution in the industry (e.g. Intel, IBM, Cadence, etc...). It is a very popular approach of general problem representation and solving, successfully used for a vast number of hard problems. The research will focus on cutting-edge methods for efficient solving and modeling of numeric variables in CP. This development is done by senior developers who are recognized worldwide as leading in CSP technology and with tight interaction with customers.

**Responsibilities:** The development will be part of Test Generation tool that is being used worldwide by multiple Intel’s leading projects. Developing an algorithm and implementing a working prototype. Optional: Publish a paper describing the research and submitting it to a CP conference.

**Required Skills:** Motivation for leading a research; knowledge in algorithms and complexity; experience in C++; good mathematical background. Advantage for students with background in AI and CP or SAT/SMT.

**Full/Part time position:** Full time for summer.

**Contact Details:** Einat Keren, Einat.keren@intel.com

---

High-level parallelization of test-generation algorithms for multi-core validation

**Company:** Intel

**General Background:** The development will be part of Test Generation tool that is being used worldwide by multiple Intel’s leading projects.

Our group develops Test Generation solutions for functional validation of Intel processors. These tools are currently single-thread based. The goal in this research & development prototype is to develop concurrent algorithms that can bring performance breakthrough in the tools. SW environment for development: Linux, C++, Intel TBB library, Intel Parallel Studio.

**Responsibilities:** Development of an algorithm and implementing a working prototype; evaluation of the quality and effectiveness. Optional: writing a paper and submitting to a conference.
**Required Skills:** Motivation for leading a research; must: strong knowledge in C++; advantage: experience in the concurrent programming

**Full/Part time position:** Full time for summer

**Contact Details:** Einat Keren, Einat.keren@intel.com

---

**Develop OpenCL device agent for the graphic processor device**

**Company:** Intel

**General Background:** OpenCL is a new emerging programming open standard that for parallel programming of heterogeneous systems. (http://www.khronos.org/opencl/)

Project Description: Develop OpenCL device agent for the graphic processor device. The Graphic agent will connect the CPU runtime to the processor graphic via OpenCL API's. After developing the prototype the student will measure performance of the processor graphic with and without his agent. The Student will analyze different scenarios of load balancing of the application between GPU and CPU devices by using OpenCL

In this research the student will need to ramp on OpenCL standard and on Intel OpenCL product. The Student will perform performance study on the prototype on 10 workloads using advanced performance analysis tools.

**Full/Part time Position:** Prefer Full time, optional part time.

**Contact Details:** Einat Keren, Einat.keren@intel.com

---

**Develop prototype of OpenCL debugger agent on Linux**

**Company:** Intel

**General Background:** OpenCL is a new emerging programming open standard that for parallel programming of heterogeneous systems. (http://www.khronos.org/opencl/)

Project Description: The student will develop prototype of OpenCL debugger agent on Linux. The debugger agent will be a plug in of GDB debugger and connected to existing debugger backend in the Intel OpenCL product. The student will need to ramp on OpenCL,Intel OpenCL product and GDB debugging interfaces.

**Responsibilities:** Interesting projects that will expose student to the new OpenCL standard as well as to writing efficient code on Intel future processors.
**Required Skills:** The job is suitable for a software engineer with reach background in C\C++. Knowledge in compilers is advantage. Knowledge in optimized code written on the CPU is advantage.

**Full/Part time Position:** Prefer Full time, optional part time.

**Contact Details:** Einat Keren, Einat.keren@intel.com

---

**Turning core simulator to multithreading**

**Company:** Intel

**General Background:** Intel develops and employs a wide range of simulation tools, ranging from detailed physical models describing the manufacturing process to functional models simulating the x86 architecture. This internship will involve work on an internal, high-speed x86 core simulator.

**Responsibilities:** In this internship proposal, you will first change the simulator to work in a multi-threaded fashion, and then work on studying and improving the multi-threaded performance of the simulator, with the goal of reaching perfect scaling (N simulated cores running on N real cores as fast as 1 simulated core running on 1 real core).

**Required Skills:** Good knowledge of data structures, high motivation, ability to learn quickly, ability to tackle hard/non-textbook problems

**Full/Part time position:** Full-time

**Contact Details:** Einat Keren, Einat.keren@intel.com
Design and Development of a ‘Property-based DBs’

Company: IBM, Haifa Research Lab

General Background: The combination of globalization in the internet era and the well-observed exponential improvement in technological means, processing power, hardware, and software, could all be conceptualized as one major force – “collapsibility” - which is constantly affecting and re-shaping organizational IT infrastructures. This force is translated into various organizational IT related strategies including: outsourcing, off-shoring, consolidation, and virtualization. Such strategic efforts impose new challenges to the realization of supporting IT infrastructures. One such realization which is provisioned to become maybe the most prominent is the one of “cloud computing” – an architectural configuration which is aimed to enable the consumption of IT resources in the same way and ease that other resources such as ‘gas’ and ‘telecommunication’ are being consumed. This technology is expected to allow end users the utilization of core IT resources such as computing power, and storage, yet being completely detached from the burden associated with the actual management and maintenance of these resources.

Project Description: This project will focus on the augmentation of the traditional ‘relational’ database model to better align it with the challenges imposed by cloud-computing. Specifically, an innovative model - namely, the ‘property-based’ model (Parsons & Wand, 2000) - is proposed as an alternative architectural base-line for designing new type of transactional data-management repositories for organizations. The project task will include the following activities:

- Survey of existing literature to identify potential theoretical frameworks to support the notion of a single-enterprise database for all applications (namely, a ‘Cloud DBs’).
- Requirement and design specifications for a data repository framework that relies on the ‘property-based’ model to enable the vision of Cloud DBs.
- Implementation of a prototype to evaluate the theorized framework.

Responsibilities: We offer an exploratory research project, possibly leading to a scientific publication, in addition to participating in developing a novel solution with a clear productization roadmap.

Required Skills: Databases, Knowledge Representation, and Software Engineering.

Full/Part time Position: Full time student position for the summer period.

Contact Details: liorli@il.ibm.com
Revealing the true ‘visual anatomy’ of Enterprise-Models

**Company:** IBM, Haifa Research Lab

**General Background:** Conceptual modeling is the activity of formally describing some aspects of the physical and social world around us for purposes of understanding and communication. In most actual settings, the elements captured in such models are expressed using visual diagrams, formally referred to as the ‘visual notation’ or the ‘concrete syntax’ of a specific conceptual modeling language. Such notations are expected to adhere to some predefined restrictions and rules, to ensure a faithful match between the diagram and the actual behavior or structure of the domain being illustrated. Specifically, such diagrams play a significant role when the domain of interest is the one of Enterprise Architectures. Yet, there is very little work in the literature how to create “good” diagrams, ones that posses visual properties that most effectively communicate their referenced domain.

Project Description: This project is aimed at evaluating a recent research work (Daniel L. Moody, 1999) which suggests several principles about the “physics” of notations. Specifically it will investigate how the expected visual properties of diagrams that are used in actual Enterprise Conceptual Models. To accomplish this objective, a pre-collected sample of enterprise architecture diagrams will be used as a dataset to analyze the actual degree to which Moody’s principles are implemented in real organizational settings. The project task will include two major activities: (1) testing Moody’s principles with respect to the dataset (may require the development of a tool to assist with the automatic analysis of visual features), and (2) identification of new visual characteristics that may have not been identified yet, but ones that may be apparent in the domain of enterprise conceptual models.

**Responsibilities:** We offer a confirmatory research project, possibly leading to a scientific publication, with also the potential to further contribute to the scarce theoretical principles that underlie the creation of effective enterprise models.

**Required Skills:** Conceptual Modeling, Systems Analysis and Design, Programming with MS Office, Background in Statistics and advantage.

**Full/Part time Position:** Full time student position for the summer period.

**Contact Details:** liorli@il.ibm.com
Community analytics research in Web 2.0 and Social Media Technologies group

Company: IBM, Haifa Research Lab

General Background: Online communities are becoming increasingly popular on the Web and within social network sites (e.g., Facebook and LinkedIn). The abundance of communities makes a challenge both for the user to discover and choose which communities to join and for the community owners to attract new audience. In this project, the goal would be to improve our existing community recommendation algorithm, which is based on related people and related tags [1][2]. The current algorithm refers to a community as a general resource (similarly to a bookmark, a file, or a blog) and does not take advantage of particular community features, such as common membership, the overall level of activity within the community, or the profile of an average member in the community.

The goal of the project is to implement new enhancements to the community recommendation algorithm and conduct a large-scale experiment to evaluate the effectiveness of the different enhancements. The outcomes of this project are twofold: (1) publishing a research paper that compares the performance of different community recommendation algorithms and shows improvement over state-of-the-art, and (2) the results of the project will serve to enhance the community recommendation feature in IBM's leading-market enterprise social software product, Lotus Connections.


Requirements: Graduate students with strong engineering skills and excellent research skills who can work as part of a team. Java programming knowledge is a must, due to the short term of the project. Web development skills are an advantage and so is existing publication experience.

Full/Part time Position: The project fits a 3-month internship on a full position basis.

Contact Details: Ido Guy, ido@il.ibm.com, 04-8296345, 054-6976345
Research and development of new methods for model matching and management using the IBM Jazz platform

**Company:** IBM, Haifa Research Lab

**General Background:** Jazz ([http://www.jazz.net](http://www.jazz.net)) is an open platform developed by IBM Rational and serves the new generation of IBM model-based software development tools. The platform integrates some 20 and more open source projects and standards for tools interoperability and model-based design. Our team takes Jazz into the system engineering domain via EU projects in cooperation with leading industrial and research partners in Europe and Israel.

**Responsibilities:** Participate in a European project ([http://www.sprint-iot.eu/](http://www.sprint-iot.eu/)) in which multiple modeling tools are integrated into a Jazz-based environment where all models share a common representation and linked to each other. You will be responsible to investigate methods to discover and apply matching among the models, and implement such algorithm into the common environment which is represented as RDF triples within Jazz. This work may result in a journal or a conference paper.

**Requirements:** Modeling, Java; familiarity with Eclipse; useful to know: Restful programming, RDF, Java script and dojo.

**Full/Part time position:** Full

**Contact Details:** Uri Shani, [shani@il.ibm.com](mailto:shani@il.ibm.com), 054-697-6282, 04-8296282

---

Social stream analysis research in Web 2.0 and Social Media Technologies group

**Company:** IBM, Haifa Research Lab

**General Background:** Social stream applications (activity streams), such as Twitter or the Facebook newsfeed, are becoming more and more dominant on the web. Through such streams, users share their activities, opinions, ideas, favorite links, photos, and videos, and in parallel get similar updates from their social environment. Similarly to their counterparts on the Web, social stream applications have also emerged inside organizations. Yammer, Chatter, Jive What Matters, and Lotus Connections Activity Stream are a few prominent examples. In this project, the goal is to analyze the stream on both the personal and organizational level. On the personal level, challenges such as attention management and event aggregation are to be addressed. On the organizational level, trend detection and sentiment analysis may aid unveiling organizational mood, reaction to processes and IT tools, and flow of information between business units (or lack thereof). We will narrow our research questions to one or more of the
above and focus on researching it. Our goal would be to publish a scientific paper on the on the
hand, and contribute to the IBM newly formed vision around unified enterprise streams on the
other hand.

Requirements: Graduate students with strong engineering skills and excellent research skills
who can work as part of a team. Java programming knowledge is a must, due to the short term
of the project. Web development skills are an advantage and so is existing publication
experience.

Full/Part time position: The project fits a 3-month internship on a full position basis.

Contact Details: Ido Guy, ido@il.ibm.com, 04-8296345, 054-6976345

Visual Analytics Researcher for Spatiotemporal Data

Company: IBM, Haifa Research Lab

General Background: Visual Analytics is an emerging area of research and practice – taking
visualization one step further - that aims at integrating the human perceptual capacity for visual
information exploration and the processing power of computers to form knowledge discovery
environment, or make well-informed decisions in complex situations. The great achievement of
Visual Analytics is of finding ways to involve and integrate the analysts in the discovery pipeline
as an integer and active part thereof. Visual Analytics is an inherently multi-disciplinary field that
aims to combine the findings of various research areas as Human-Computer Interaction (HCI),
Usability Engineering, Cognitive and Perceptual Science, Decision Science, Information
Visualization, Scientific Visualization, Databases, Data Mining, Statistics, Knowledge Discovery,
Data Management & Knowledge Representation, Statistics, Interaction, Geography and
Geospatial Analytics, Graphics and Rendering.

Responsibilities: Independent research in the field of Visual Analytics, by defining an application
area and prototypically showing how visual analytics can be applied in the defined field yielding
more efficient performance of the analyst. Suggested application areas: traffic, travel and
logistics in which the optimization of routes, and the extraction of factors affecting logistic and
travel accidents and general performance are required.

Requirements:
1. Independent research work
2. Publication in an international conference / journal
3. Implementation of a prototype

Full/Part time position: Full time position

Contact Details: Peter Bak, peterba@il.ibm.com, 04-8296537, 052-6677988
Modeling maintenance process in software engineering

Company: IBM, Haifa Research Lab

General Background: A typical software engineering company invests a significant amount of resources in Sustaining Engineering and Maintenance (SEM) - continuing engineering and technical support that follows release of a new product. A major part of this effort is dedicated to issues that arise when a customer reports a problem or a defect in a software product, or a product defect is discovered internally. Therefore, it is of major importance to understand the SEM process, model it and detect significant trends. It was empirically discovered that the defect handling time typically has a heavy-tailed distribution: defects with very large handling time arise occasionally. This phenomenon has important implications in Statistical Process Control for SEM. However, its roots are still not understood well. In this research, we develop a model of the SEM process and fit SEM empirical data to this model using combination of theoretical and simulation methods.

Responsibilities: Jointly with other IBM researchers, design models of the SEM process. Work with the queueing simulator developed in HRL and enhance it for the needs of our project. Develop simulation and theoretical methods that fit the SEM model to empirical data. This work may result in a journal or a conference publication.

Requirements: knowledge in probability theory, stochastic models and statistics; Java; familiarity with Eclipse; useful to know: Matlab, queueing theory.

Full/Part time position: Full

Contact Details: Sergey Zeltyn, sergeyz@il.ibm.com, 052-6658807, 04-8296144

Software Testing Tools and Methodology

Company: IBM, Haifa Research Lab

Responsibilities: Working with a team of IBM researchers in the area of advanced software testing technology. With the introduction of many core and distributed hybrid large scale commercial systems the challenge of testing and analyzing these systems is constantly increasing. The candidate will be part of a team that developed advanced tools in this domain.

Requirements:
- CS courses
- Good Software engineering
- Creative ideas
• Good background in computer science or related areas
• Open minded

**Full/Part time position:** Both

**Contact Details:** Bilha Mendelson: [Bilha@il.ibm.com](mailto:Bilha@il.ibm.com), Eitan Farchi: [Farchi@il.ibm.com](mailto:Farchi@il.ibm.com)

---

**Performance Improvements**

**Company:** IBM, Haifa Research Lab

**Responsibilities:** Working with a team of IBM researchers in the area of advanced compiler and/or post-link optimization and/or advanced performance tools. The focus is to better utilize modern architectures and systems. The candidate will be part of a team that developed advanced tools in this domain.

**Requirements:**
- CS courses. In particular: Algorithms and Data Structures, Graph Theory, Computer architecture basics, Operating Systems
- Creative ideas
- Good background in computer science
- Basic understanding computer architectures
- Open minded

**Full/Part time position:** Both

**Contact details:** Bilha Mendelson: [Bilha@il.ibm.com](mailto:Bilha@il.ibm.com)

---

**Advanced layout analysis of complex models for additional info contact**

**Company:** IBM, Haifa Research Lab

**General Background:** As more and more models are created using free form drawing tools a need to understand the models and harvest them to rigid modeling tools increases. This is typically done in several phases: crawl the enterprise and collect the models, cleanse and index the models and finally understand the models and transform them to a standardize model.
Project Description: A crucial part of the effort to understand a diagram, drawn in a free form tool, is to conduct layout analysis of the diagram. To meet this end, one needs to research the existing state of the art layout analysis algorithms and implement them within an existing framework. Such a layout analysis should identify advanced relationship between objects and shapes which appear on the drawing pane. Examples to such patterns are containment relationship, connectivity of shapes, relative positioning and size.

**Responsibilities:** We offer exploratory Research project possibly leading to a scientific publication in addition to participating in developing a novel solution with a clear productization roadmap.

**Requirements:** Strong algorithmic background

**Full/Part time Position:** Full time student position for the summer period

**Contact details:** davida@il.ibm.com

---

Develop a Visual Query Language for Enterprise Information Retrieval

**Company:** IBM, Haifa Research Lab

**General Background:** As more business users are exposed to enterprise information, the need for an easy method to formulate queries is essential. Business users fall short when it comes to using SQL-like languages and prefer an ‘easy to use’ visual mean instead.

Project Description: Design and implement a visual query language and interface and integrate it with an existing web application. The work includes: researching existing prior art of visual queries languages and their supported environment, choosing the best of breed and implementing it using Web2.0 technologies (JavaScript, Dojo, Ajax).

**Responsibilities:** We offer exploratory Research project possibly leading to a scientific publication in addition to participating in developing a novel solution with a clear productization roadmap.

**Requirements:** Strong algorithmic background.

**Full/Part time Position:** Full time student position for the summer period.

**Contact details:** davida@il.ibm.com
Learning Instruction Stream Behavior to Improve Random Test Program Generators

Company: IBM, Haifa Research Lab

General background: In processor verification, random test program generators are used to generate stimuli whose goal is to trigger interesting events in the verified processor and help expose bugs in it. These generators receive as their input test specifications or test templates and generate, out of these specifications, many random test cases.

To generate tests out of the test specifications, the test generators need knowledge about the architecture of the verified processor. For example, the number of types of operands in each instruction, memory translations, etc. This information is usually encoded in the generators or attached to them in an external model of the verified processor architecture.

In a similar manner, information about the microarchitectural behavior of the processor can help test generators to better fill requests in the test specifications and generate higher quality tests. For example, knowing the execution time of instructions can be used to generate instructions with long latency when stalling the pipelines is needed and instructions with short latency when high throughput is needed. While such information on the microarchitectural behavior of the verified processor is readily available (e.g., in microarchitectural documentation), incorporating it in random test program generators is often a tedious, time consuming, and error prone task.

Last year we had an intern project on automatic learning of microarchitectural properties of single instructions and incorporation of this knowledge into GenesysPro, IBM’s processor-level random test generator. The work resulted in a tool prototype that demonstrated the benefits of the automatic learning and a paper that was accepted to the Design Automation Conference – DAC (this is a top ranking conference with a 20% acceptance rate).

Requirements: This year we would like to extend this work to learn properties of streams of instructions, or in other words, stream conditions that lead to specific microarchitectural events. For example, the conditions needed to activate a certain bypath between two pipelines. The work comprises two major steps. The first step is finding or developing machine learning algorithms that can learn rules about the behaviors of streams. The second step is incorporating the learnt rules into GenesysPro in a way that will allow it to use them to improve the quality of generated tests.

We see the work producing two types of outcomes. The first outcome is a prototype for a system that can extract rules about the behavior of instruction streams out of simulation traces. The second outcome is a paper describing the work and analyzing its benefits.

Contact details: Avi Ziv aziv@il.ibm.com
Running a Floating-Point Test-Suite on Hardware

**Company:** IBM, Haifa Research Lab

**General background:** Verification of floating-point data path is a well recognized challenge. A large variety of tools and techniques were developed to meet this challenge. One of the most popular techniques is to run a set of test-cases (a test suite) on the simulated design and compare the expected result provided by the test suite with actual results obtained by the design. Due to the performance limitations of existing simulation platforms, this approach is not applicable for very large tests suites, such as the one developed at IBM for the floating point IEEE standard.

We propose an alternative technique that enables running the test suites directly on the verified hardware (rather than on a simulation of the hardware design). This technique requires the generation of an assembly program from the test suite. The generated program initializes the floating point registers with desired inputs, runs the instructions, and compares result with the correct value.

This project involves the development of a tool that receives a test suite for floating point data path verification, and outputs an assembly program that will run the tests on the hardware. This program could later be used by, for example, logic designers to verify that their changes did not break the existing functionality (regression testing).

One of this internship's goals is to publish a paper with the project results in one of the relevant scientific/industrial conferences.

**Responsibilities:** a student undertaking this internship will:
- Get familiarized with our tools for floating-point test generation
- Get familiarized with out hardware-based tools
- Practice both assembler and C++ programming
- Design and implement a solution for the described problem
- Participate in writing a paper to one of the relevant scientific/industrial conferences, reporting on the results of this experiment

**Requirements:**
- C++ programming skills are must
- Software engineering and Object Oriented Programming is an advantage

**Contact details:** Wisam Kadry wisamk@il.ibm.com
The Relation between Bugs and Coverage

**Company:** IBM, Haifa Research Lab

**General background:** Statistical information, such as the number of bugs found, and coverage data are two of the most useful measures on the status and progress of the verification process. Intuitively, it seems that coverage and bugs should be strongly related. Progress in coverage usually means reaching new areas or features in the design. This, in turn, increases the potential of finding new bugs in the design. However, this intuitive relation has not been confirmed. In fact, the actual relationship between coverage and bug measures is not fully understood and has not been thoroughly investigated so far.

**Responsibilities:** We propose to conduct a controlled experiment on a real-life design and verification environment that will investigate the relation between coverage and bugs both in the behavior of individual tests (e.g., does a test that detects a bug show a unique coverage profile?) and the behavior of the entire verification process (e.g., is there a relation between the progress of coverage and the number of accumulated bugs?). The goal of the research is to develop some understanding about the relations between coverage and design bugs, which can be used by verification teams to better manage the verification process. The results of the research will be documented in a paper that will be submitted to a top verification conference and / or a relevant journal.

**Contact details:** Avi Ziv aziv@il.ibm.com

---

Exploring Parallel Techniques to Solve the SAT Problem

**Company:** IBM, Haifa Research Lab

**General background:** The need to solve satisfaction and optimization problems is fundamental in many domains. These problems are NP-hard, and therefore attract much attention from various academic communities, ranging from pure-theory research to specific domains and specific types of problems. For example, the Formal Verification community has spent vast research resources into gaining better understanding and creating better solution techniques to the Boolean Satisfiability (SAT) problem. Constraint satisfaction problems (CSP), a generalization of SAT into non-Boolean discrete variables constrained by constraints of any form, is at the core of many other industrial applications. Efficiently solving such problems, either in their satisfaction or optimization versions, is the cornerstone of Operations Research. A large class of such problems lend themselves naturally to bit-wise logic manipulation just like in the signal processing and imaging domains. However, to date, all commercially available solutions rely on (serial) software. Taking the serial-software approach quickly brings about a limit to what these tools can be used for. That is, both industry and academia are restricted in the application of such problems because the solvers themselves are limited (in terms of
memory usage and runtime). Non-negligible research has been invested in parallelizing solvers, but these works remain at an academic level, and have not matured to commercial usage.

**Responsibilities:** a student undertaking this internship will be required to:
- Study existing literature on this topic
- Construct a software simulator which will measure the effectiveness of the proposed solution
- Design and implement several stochastic search engines, and evaluate their performance.

**Requirements:** We aim to explore different parallel techniques to solve SAT. We intend to combine parallel, configurable, hardware (FPGAs) together with stochastic search technique, to construct a first of a kind solver, which can solve hard SAT formula orders of magnitude faster than standard solvers.

**Contact details:** Amir Nahir nahir@il.ibm.com

---

**Combining Many Stream Specifications to a Single Test**

**Company:** IBM, Haifa Research Lab

**General background:** In processor verification, random test program generators are used to generate stimuli whose goal is to trigger interesting events in the verified processor and help expose bugs in it. These generators receive as their input test specifications or test templates and generate, out of these specifications, many random test cases.

**Requirements:** In this work, we begin with a test specification that includes many small scenarios. The scenarios comprise small sets of instructions and constraints that relate the instructions. For example, a load-hit-store scenario has a store instruction followed by a load instruction with constraints that the two instructions access the same memory address, the distance between the instructions in the program is less than 5 instructions, and no instruction in the middle accesses the memory. Out of these constraints, this project is mostly interested in scheduling-type constraints, such as the last two constraints in the example. The goal of the project is to define and implement an algorithm for scheduling the instructions in a ways that produce interesting test cases. There are aspects to our combination problem, which make it different from traditional scheduling problems. First, our main goal is not optimization (of time or resources), but increasing test quality. Second, we allow (and even encourage) instructions from different streams to merge into a single instruction.

We see the work producing two types of outcomes. The first outcome is a prototype for the scheduling algorithm that will be integrated into our stream-level stimuli generator. The second outcome is a paper describing the work and analyzing its benefits.

**Contact details:** Avi Ziv aziv@il.ibm.com
Optimized Test Suite Generation for Systems Engineering Models

Company: IBM, Haifa Research Lab

General background: The research is in the field of VVT (Verification, Validation and Testing) for Systems Engineering.
The role involves the research of UML based behavioral models in systems engineering with the purpose of generating optimized tests suites.

The following lists the details of the job description:
• Learn how to use IBM modeling tools (e.g. Rational Rhapsody)
• Design the modeling of optimization related information for the purpose of generating optimized test suites
• Design & implement an infrastructure for the mass generation of a variety of optimized tests suites for behavioral models (e.g. cost, time optimizations)
• Design & implement an infrastructure for the generation of field tests
• Participate in writing a paper on the topic at hand to one of the relevant scientific/industrial conferences

Requirements:
• Excellent technical skills in software engineering
• Excellent autodidactic abilities
• Excellent interpersonal skills
• Motivated and driven individual
• SysML / UML knowledge is an advantage

Contact details: Ronen Levy ronenl@il.ibm.com
Linux SW developer

Company: Zoran

General background: Software for set-top-box receivers and multi media players

Responsibilities: Development of unique services for Linux kernel

Requirements: Experience in Linux user and kernel space development

Full/Part time position: Preferably full time

Contact Details: Alex Matushevsky alex.matushevsky@zoran.com