

# ROY BAR-HAIM, PH.D

6 Sederot Kam., Ramat-Gan 52295, Israel ♦ Mobile: +972 506918281; Home: +972 36769345  
<http://www.roybarhaim.com> ♦ [barhair@gmail.com](mailto:barhair@gmail.com) ♦ Born 1973, Married+2

## Summary

- ♦ Expert in natural language processing (NLP), with many years of research experience in academia and industry.
- ♦ Many years of experience in managing R&D teams.
- ♦ Have worked on textual inference, Hebrew processing, information extraction, text mining, sentiment analysis, text classification, and more.
- ♦ Designed and created NLP tools and datasets that are used by dozens of researchers worldwide.

## Education

### 2005-2010 **Ph.D in Computer Science, Bar-Ilan University, Israel**

Advisor: Prof. Ido Dagan

- ♦ Thesis title: "*Semantic Inference at the Lexical-Syntactic Level*"
- ♦ Developed a robust and scalable semantic inference engine over natural language texts. My inference engine was the core of the Bar-Ilan textual entailment system, which participated in the Recognizing Textual Entailment (RTE) challenges in 2007-2009. Our team won the RTE search task in 2009.
- ♦ Co-organized the Second Recognising Textual Entailment challenge (RTE2, 2006). Activities included coordinating and supervising dataset collection and annotation, website management, submissions processing etc. These datasets have been used in dozens of research papers.

### 2001-2005 **M.Sc in Computer Science, *Cum Laude*, the Technion, Israel**

(Thesis: 94, courses: 91.1)

Advisor: Prof. Yoad Winter. Co-supervised by Prof. Khalil Sima'an, University of Amsterdam.

- ♦ Thesis title: "*Part-of-speech tagging of Hebrew and Other Semitic Languages*"
- ♦ Main results published in *Journal of Natural Language Engineering*.
- ♦ Research resulted in a state-of-the-art part-of-speech tagger and word segmenter for Hebrew. The tagger is maintained by the Knowledge Center for Processing Hebrew in the Technion, and is being used by researchers worldwide, as well as by commercial organizations. It has also been successfully ported to Arabic.

### 1992-1996 **B.Sc in Computer Science, *Summa Cum Laude*, the Technion, Israel**

(Average: 90.9)

## Research Internships

- ♦ August-September 2007: **Proteus Project, Department of Computer Science, New York University**. Worked on applying lexical entailment to event detection.  
Host: Prof. Satoshi Sekine.
- ♦ September 2002: **Institute for Logic, Language and Computation, University of Amsterdam**. Worked on syntactic and morphological disambiguation of Hebrew.  
Host: Prof. Khalil Sima'an.

# ROY BAR-HAIM, PH.D

6 Sederot Kam., Ramat-Gan 52295, Israel ♦ Mobile: +972 506918281; Home: +972 36769345  
<http://www.roybarhaim.com> ♦ [barhair@gmail.com](mailto:barhair@gmail.com) ♦ Born 1973, Married+2

## *Professional Experience*

- 2013- **Research Staff Member, IBM Research**  
Member of the *Machine Learning for Text Analytics* group at the IBM Haifa Research Lab.
- 2011-2013 **Manager, Algorithms Team, Leadspace**  
Headed the company's research and development in the areas of data and text mining, machine learning, information retrieval, text categorization, information extraction and semantic clustering.
- 2009-2011 **Manager, Text Mining Group, Digital Trowel**
- ♦ Built and managed a large R&D team consisting of five research engineers and three domain experts.
  - ♦ My team developed *The Stock Sonar*, a state of the art stock sentiment analysis engine based on precise information extraction. The Stock Sonar was chosen by *Dow Jones* to provide sentiment analysis and event extraction for their content. Research results were published in the IAAI-11 conference.
  - ♦ Developed novel methods for finding expert users in stock microblogs (such as StockTwits). Research results were published in the EMNLP 2011 conference.
  - ♦ Worked on several other projects in text mining, information extraction and sentiment analysis.
  - ♦ Coordinated Digital Trowel's participation in the [SCMS project](#), a joint project with four European partners that is funded in the EUREKA Eurostars! Programme.
- 2002-2009 **Researcher in Natural Language Processing, Bar-Ilan University & Technion**  
(Research activities are described under *Education*)
- 2002-2007 **Recitation instructor, Bar-Ilan University and the Technion**
- ♦ Bar-Ilan: Operating Systems, Automata and Formal Languages
  - ♦ Technion: Introduction to Natural Language Processing, File Systems
- 1996-2002 **Project leader, Israel Defense Forces, C4I Division**  
(rank: captain)
- ♦ Successfully managed large-scale software development projects through their entire life cycle, including requirement analysis, design and implementation, deployment, support and maintenance.
  - ♦ Technology included multithreading, client/server, messaging protocols and socket programming, databases, MS-Office automation, and GUI.
  - ♦ Responsibility included evaluation and recommendation of new technologies.

## ROY BAR-HAIM, PH.D

6 Sederot Kam., Ramat-Gan 52295, Israel ♦ Mobile: +972 506918281; Home: +972 36769345  
<http://www.roybarhaim.com> ♦ [barhair@gmail.com](mailto:barhair@gmail.com) ♦ Born 1973, Married+2

- 1994-1996 **Developer (student position), IBM Haifa Research Lab**
- ♦ Design and implementation of algorithms for compiler optimization.
  - ♦ Co-author of four U.S. patents registered by IBM.
- 1987-1991 **Freelance developer**
- ♦ (Age 18) Developed an information system for a book store.
  - ♦ (Age 14) Developed educational software for learning music, published by Ramot press, Tel-Aviv University.

### *Publications*

1. Roy Bar-Haim, Ido Dagan, and Idan Szpektor. Benchmarking Applied Semantic Inference: The PASCAL Recognising Textual Entailment Challenges. Nachum Dershowitz and Ephraim Nissan, editors. *Language, Culture, Computation: Essays Dedicated to Yaacov Choueka*, Lecture Notes in Computer Science, Springer-Verlag, Berlin (3 volumes). To appear.
2. Roy Bar-Haim, Elad Dinur, Ronen Feldman, Moshe Fresko and Guy Goldstein. 2011. Identifying and Following Expert Investors in Stock Microblogs. In *Proceedings of EMNLP 2011*.
3. Ronen Feldman, Benjamin Rosenfeld, Roy Bar-Haim and Moshe Fresko. 2011. The Stock Sonar - Sentiment Analysis of Stocks Based on a Hybrid Approach. In *Proceedings of the Twenty-Third Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-11)*.
4. Shachar Mirkin, Roy Bar-Haim, Jonathan Berant, Ido Dagan, Eyal Shnarch, Asher Stern and Idan Szpektor. 2009. Addressing Discourse and Document Structure in the RTE Search Task. In *Proceedings of the Second Text Analysis Conference (TAC 2009)*.
5. Roy Bar-Haim, Jonathan Berant and Ido Dagan. 2009. A Compact Forest for Scalable Inference over Entailment and Paraphrase Rules. In *Proceedings of EMNLP 2009*.
6. Roy Bar-Haim, Khalil Sima'an and Yoad Winter. Part-of-Speech Tagging of Modern Hebrew Text. 2008. *Natural Language Engineering* 14(2): 223-251.
7. Idan Szpektor, Ido Dagan, Roy Bar-Haim and Jacob Goldberger. 2008. Contextual Preferences. In *Proceedings of ACL 2008*.
8. Ido Dagan, Roy Bar-Haim, Idan Szpektor, Iddo Greental, and Eyal Shnarch. 2008. Natural Language as the Basis for Meaning Representation and Inference. In *Proceedings of CICLING 2008*.
9. Roy Bar-Haim, Jonathan Berant, Ido Dagan, Iddo Greental, Shachar Mirkin, Eyal Shnarch, and Idan Szpektor. 2009. Efficient Semantic Deduction and Approximate

## ROY BAR-HAIM, PH.D

6 Sederot Kam., Ramat-Gan 52295, Israel ♦ Mobile: +972 506918281; Home: +972 36769345  
<http://www.roybarhaim.com> ♦ [barhair@gmail.com](mailto:barhair@gmail.com) ♦ Born 1973, Married+2

- Matching over Compact Parse Forests. In *Proceedings of the First Text Analysis Conference (TAC 2008)*.
10. Roy Bar-Haim, Ido Dagan, Iddo Greental and Eyal Shnarch. Semantic Inference at the Lexical-Syntactic Level. 2007. In *Proceedings of the Twenty-Second Conference on Artificial Intelligence (AAAI-07)*.
  11. Roy Bar-Haim, Ido Dagan, Iddo Greental, Idan Szpektor and Moshe Friedman. 2007. Semantic Inference at the Lexical-Syntactic Level for Textual Entailment Recognition. In *Proceedings of ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (WTEP-07)*.
  12. Roy Bar-Haim, Ido Dagan, Bill Dolan, Lisa Ferro, Danilo Giampiccolo, Bernardo Magnini and Idan Szpektor. 2006. The Second PASCAL Recognising Textual Entailment Challenge. In *Proceedings of the Second PASCAL Recognising Textual Entailment Challenge*, Venice, Italy.
  13. Roy Bar-Haim, Idan Szpektor and Oren Glickman. 2005. Definition and Analysis of Intermediate Entailment Levels. *ACL 2005 Workshop on Empirical Modeling of Semantic Equivalence and Entailment*.
  14. Roy Bar-Haim, Khalil Sima'an and Yoad Winter. 2005. Choosing an Optimal Architecture for Segmentation and POS-Tagging of Modern Hebrew. *ACL 2005 Workshop on Computational Approaches to Semitic Languages*.

### *Patents*

1. U.S. Patent Application 13/367562: System and Method for Identifying Potential Customers. Amnon Mishor, Roy Bar-Haim, and Eyal Shpin .Filed February 7, 2012
2. U.S. Patent 5,774,730: Method and apparatus for improving colorability of constrained nodes in an interference graph within a computer system. Nava Arela Aizikowitz, Liviu Asnash, Roy Bar-Haim, Orit Edelstein, Mircea Namolaru, Edward Curtis Prosser, Robert Ralph Roediger, and William Jon Schmidt. Filed July 31, 1995.
3. U.S. Patent 5,761,514: Register allocation method and apparatus for truncating runaway lifetimes of program variables in a computer system. Nava Arela Aizikowitz, Roy Bar-Haim , Edward Curtis Prosser, Robert Ralph Roediger, and William Jon Schmidt. Filed August 31, 1995
4. U.S. Patent 5,946,491: Register allocation method and apparatus for generating spill code as a function of register pressure compared to dual thresholds. Nava Arela Aizikowitz, Liviu Asnash, Roy Bar-Haim, Edward Curtis Prosser, Robert Ralph Roediger, and William Jon Schmidt. Filed June 6,1996.

## ROY BAR-HAIM, PH.D

6 Sederot Kam., Ramat-Gan 52295, Israel ♦ Mobile: +972 506918281; Home: +972 36769345  
<http://www.roybarhaim.com> ♦ [barhair@gmail.com](mailto:barhair@gmail.com) ♦ Born 1973, Married+2

5. U.S. Patent 5,890,000: Cooperation of global and local register allocators for better handling of procedures. Nava E. Aizikowitz, Roy Bar-Haim, Orit Edelstein, Edward Curtis Prosser, and Robert Ralph Roediger. Filed December 4, 1996.

### *Academic Service*

#### **Journal Reviewer:**

- Transactions of the Association for Computational Linguistics (TACL)
- Natural Language Engineering (NLE)
- ACM Transactions on Intelligent Systems and Technology (ACM TIST)

#### **Conference PC Member/Reviewer:**

- ♦ COLING 2008, 2014
- ♦ EMNLP 2008, 2009, 2012
- ♦ ACL 2008

#### **Organization:**

- ♦ Co-organizer of the Second PASCAL Recognising Textual Entailment Challenge (RTE-2), 2006.

### *Talks and Presentations*

1. *Mining the Web for Stock Sentiment*. January 22, 2012. IBM Machine Learning Seminar, Tel-Aviv.
2. *Identifying and Following Expert Investors on Twitter*. In ISCOL 2011, Israeli Seminar on Computational Linguistics. June 2, 2011. Bar-Ilan University, Ramat-Gan.
3. *SSA - A Hybrid Approach to Sentiment Analysis of Stocks*. In ISCOL 2010, Israeli Seminar on Computational Linguistics. June 16, 2010, Tel-Aviv University.
4. *Applied Semantic Inference over Natural Language Expressions*. September 2009. IBM Haifa Research Lab.
5. *Efficient Entailment Inference over Compact Parse Forests*. In BISFAI 2009, The 10th Bar-Ilan Symposium on the Foundations of Artificial Intelligence, June 14-15, 2009, Bar-Ilan University, Ramat-Gan, Israel.
6. Roy Bar-Haim and Ido Dagan. 2008. Efficient Semantic Inference over Language Expressions. *NSF Symposium on Semantic Knowledge Discovery, Organization and Use*, New York University.
7. *Efficient Semantic Deduction and Approximate Matching over Compact Parse Forests*. In IAAI Symposium, The Israeli Association for Artificial Intelligence, Ashkelon Academic College. November 12, 2008.
8. *Semantic Inference at the Lexical-Syntactic Level* (extended talk). Microsoft Research and New York University. 2007.
9. *Semantic Inference at the Lexical-Syntactic Level*. In BISFAI 2007, The 9th Bar-Ilan Symposium on the Foundations of Artificial Intelligence, June 20-22, 2007, Bar-Ilan University, Ramat-Gan, Israel.

## ROY BAR-HAIM, PH.D

6 Sederot Kam., Ramat-Gan 52295, Israel ♦ Mobile: +972 506918281; Home: +972 36769345  
<http://www.roybarhaim.com> ♦ [barhair@gmail.com](mailto:barhair@gmail.com) ♦ Born 1973, Married+2

10. *Part-of-Speech Tagging for Hebrew and Other Semitic Languages*. In ISCOL 2005 - Israeli Seminar on Computational Linguistics. June 22, 2005, the Technion, Haifa.
11. *Corpus-Based Morphological and Syntactic Analysis of Hebrew*. September 2002. Institute for Logic, Language and Computation, University of Amsterdam.

### *Honors and Awards*

- ♦ Doctoral Fellowship of Excellence, 2005-2008, Bar-Ilan University
- ♦ M.Sc *Cum Laude*, the Technion, 2005
- ♦ Distinguished officer award, IDF C4I Division, 2000
- ♦ B.Sc *Summa Cum Laude*, the Technion, 1996
- ♦ President's list, the Technion, 1994, 1996
- ♦ Dean's list, the Technion, 1993, 1995

### *Technical Skills*

- ♦ Programming Languages: Java, Perl, C#, Visual C++, Visual Basic
- ♦ Operating Systems: Windows, Linux/Unix.