Andrei Barbu

(315) 333-0621

andrei@0xab.com
abarbu@csail.mit.edu

May 28, 2024 http://0xab.com

Skills & interests

integration of language with robotics and vision, 3D vision, cognitive science, computer vision, functional programming, logic and constraint programming, stochastic and nondeterministic programming languages, linguistics, machine learning, operating systems, neuroimaging to understand the representation of knowledge and language in the brain, video event recognition, machine learning for accessibility and fairness

Languages

Haskell, Scheme, C, C++, Lisp, Elm, MATLAB, Java, Python, Clojure, JavaScript native English and Romanian, basic German and French

Education

Postdoctoral Associate

Massachusetts Institute of Technology, CSAIL

Part of the MIT/Harvard Center for Brains, Minds, and Machines Toyota Research Institute grant *Using vision and language to read minds* IBM-MIT grant *Laboratory for Brain-inspired Multimedia Machine Comprehension* Working on cross-modal reasoning and using vision to solve linguistic tasks.



2013 - 2016

PhD, Artificial Intelligence

Purdue University, School of Electrical and Computer Engineering

Title: Reasoning across language and vision in machines and humans Advisor: Jeffrey Mark Siskind PURDUE

2008 - 2013

language and high-level reasoning for solving inverse vision and manipulation problems grounding language in vision and robotic manipulation stochastic modeling via probabilistic programming optimization of large stochastic cognitive models segmentation and tracking of objects in videos with high-level semantic priors large-scale video search using sentential queries the neural representation of verbs, events, and compositionality TA for ECE473 and ECE570, Artificial Intelligence participated in the DARPA *Mind's Eye* program which investigated event recognition won the yearly evaluations against 11 other teams in both years of the program processed tens of millions of frames of video on a top-500 supercomputer.

participated in the DARPA *Mind's Eye* program which investigated event recognition won the yearly evaluations against 11 other teams in both years of the program processed tens of millions of frames of video on a top-500 supercomputer solved computer-vision and manipulation problems using AD-based optimization http://oxab.com/research/

2004 - 2008

Bachelor of Computer Science

University of Waterloo

complexity theory, real-time operating systems, graph theory, algorithm design, programming languages, medical imaging, linear programming, coding theory



Employment History

Research Scientist

Massachusetts Institute of Technology

Part of the MIT/Stanford Toyota Research Institute (TRI)

Part of the MIT/Harvard Center for Brains, Minds, and Machines (CBMM)

Part of the MIT/IBM Laboratory for Brain-inspired Multimedia Machine Comprehension (BM³C)

Part of the Air Force-MIT AI Accelerator

Developing robots that learn, understand, and follow commands

Working on deep video understanding grounding language in perception

Teaching and co-organizing the CBMM summer school (35 graduate students, 40 speakers)

ONR award Developing next generation AI vision systems by characterizing and exploiting untapped primate visual processing circuit motifs

NSF award NCS-FO: Studying language in the brain in the modern machine learning era

DARPA award Grounded Artificial Intelligence Language Acquisition (GAILA)

DARPA award Knowledge Management at Scale and Speed (KMASS)

Software Development Engineer

Amazon.com

migrated from a monolithic database design to a distributed service-oriented system designed & developed distributed services to collect metrics about internal systems coordinated with other teams to develop internal tools and alarms

May 2010 – *Aug.* 2010 Seattle, USA

Oct. 2016 - present

Cambridge, USA

amazon

Kernel Developer

Google Summer of Code 2008 - the GNU Hurd

designed and implemented an instrumentation framework for GNU Mach implemented the DWARF2 debugging standard debugger improvements bug fixes and other new features May 2008 – Aug. 2008 Purdue University, USA



Undergraduate Research Assistant

Advisor: Professor Jeffrey Mark Siskind

numerous improvements to a C frontend, including updates for C99 and GNUC improvements to the existing Scheme compiler ported the Scheme→C compiler to x86-64 ported various other applications to x86-64 significant performance improvements

July 2006 – *Sept.* 2006 Purdue University, USA

PURDUE UNIVERSITY

Software Developer and Knowledge Systems Architect

Dalin Software SA

ontology, temporal, and spatial logic frameworks designed extensions to OWL, the Ontology Web Language designed and developed tools for querying and updating OWL+SWRL ontologies

July 2005 – May 2006 Laussane, Switzerland

Publications

Revealing Vision-Language Integration in the Brain with Multimodal Networks V. Subramaniam, C. Conwell, C. Wang, G. Kreiman, B. Katz, I. Cases, A. Barbu International Conference on Machine Learning (ICML)	July 2024 P1
How hard are computer vision datasets? Calibrating dataset difficulty to viewing time D. Mayo, J. Cummings, X. Lin, D. Gutfreund, B. Katz, A. Barbu Neural Information Processing Systems (NeurIPS)	December 2023 P2
BrainBERT: Self-supervised representation learning for intracranial recordings C. Wang, V. Subramaniam, A. Yaari, G. Kreiman, B. Katz, I. Cases, A. Barbu International Conference on Learning Representations (ICLR)	<i>March</i> 2023
Zero-shot linear combinations of grounded social interactions with Linear Social MDPs R. Tejwani, Y. Kuo, T. Shu, B. Stankovits, D. Gutfreund, J. Tenenbaum, B. Katz, A. Barbu Conference on Artificial Intelligence (AAAI)	January 2023 P4
The Aligned Multimodal Movie Treebank: An audio, video, dependency-parse treebank A. Yaari, J. DeWitt, H. Hu, B. Stankovits, S. Felshin, Y. Berzak, H. Aparicio, B. Katz, I. Cases, A. Barbu Empirical Methods in Natural Language Processing (EMNLP)	September 2022 P5
Developing a Series of AI Challenges for the United States Department of the Air Force V. Gadepally, G. Angelides, A. Barbu, A. Bowne, L. Brattain, T. Broderick, A. Cabrera, G. Carl, R. Carter, M. Cha, E. Cowen, J. Cummings, B. Freeman, J. Glass, S. Goldberg, M. Hamilton, T. Heldt, K. W. Huang, P. Isola, B. Katz, J. Koerner, Y. Lin, D. Mayo, K. McAlpin, T. Perron, J. Piou, H. M. Rao, H. Reynolds, K. Samuel, S. Samsi, M. Schmidt, L. Shing, O. Simek, B. Swenson, V. Sze, J. Taylor, P. Tylkin, M. Veillette, M. L Weiss, A. Wollaber, S. Yuditskaya, J. Kepner IEEE Conference on High Performance Extreme Computing (HPEC)	September 2022 P6
Quantifying the Emergence of Symbolic Communication E. Cheng, Y. Kuo, I Cases, B. Katz, A. Barbu CogSci	August 2022 P7
Incorporating rich social interactions into MDPs R. Tejwani, Y. Kuo, T. Shu, B. Stankovits, D. Gutfreund, J. Tenenbaum, B. Katz, A. Barbu International Conference on Robotics and Automation (ICRA) Excellent Paper award at IROS Cognitive and Social Aspects of Human Multi-Robot Interaction Workshop	<i>March</i> 2022 P8
Trajectory prediction with linguistic representations Y. Kuo, X. Huang, A. Barbu, S. McGill, B. Katz, J. Leonard, G. Rosman International Conference on Robotics and Automation (ICRA)	March 2022 P9
Neural Regression, Representational Similarity, Model Zoology & Neural Taskonomy at Scale in Rodent Visual Cortex	December 2021
C. Conwell, D. Mayo, A. Barbu, M. Buice, G. Alvarez, B. Katz Neural Information Processing Systems (NeurIPS)	P10
Perceiving social events in a physical world T. Shu, A. Netanyahu, M. Kryven, J. Muchovej, N. Shenoy, B. Katz, A. Barbu, T. Ullman, J. Tenenbaum Vision Sciences Society Annual Meeting	September 2021
Social Interactions as Recursive MDPs R. Tejwani, Y. Kuo, B. Katz, and A. Barbu Conference on Robot Learning (CoRL)	August 2021 P12
Compositional Networks Enable Systematic Generalization for Grounded Language Understanding Y. Kuo, B. Katz, and A. Barbu Findings of Empirical Methods in Natural Language Processing (EMNLP)	August 2021 P13

Measuring Social Biases in Grounded Vision and Language Embeddings C. Ross, B. Katz, and A. Barbu Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)	June 2021 P14
Multi-resolution modeling of a discrete stochastic process identifies causes of cancer A. Yaari, M. Sherman, O. Clarke Priebe, P. Loh, B. Katz, A. Barbu, B. Berger <i>International Conference on Learning Representations (ICLR)</i>	May 2021 P15
PHASE: PHysically-grounded Abstract Social Events for Machine Social Perception A. Netanyahu, T. Shu, B. Katz, A. Barbu, J. Tenenbaum Conference on Artificial Intelligence (AAAI) Best paper award at the NeurIPS SVRHM workshop	January 2021 P16
Genetic and epigenetic factors interact to explain somatic mutation rates at multiple length scales	October 2020
A. Yaari, M. Sherman, O. Clarke Priebe, P. Loh, A. Barbu, B. Katz, B. Berger American Society of Human Genetics	P17
Best paper award semi-finalist and reviewer's choice award https://www.abstractsonline.com/pp8/#!/9070/presentation/3409	
Learning a natural-language to LTL executable semantic parser for grounded robotics C. Wang, C. Ross, Y. Kuo, B. Katz, and A. Barbu Conference on Robot Learning (CoRL)	November 2020 P18
Encoding formulas as deep networks: Reinforcement learning for zero-shot execution of LTL formulas	October 2020
Y. Kuo, B. Katz, and A. Barbu International Conference on Intelligent Robots (IROS)	P19
Deep compositional robotic planners that follow natural language commands Y. Kuo, B. Katz, and A. Barbu International Conference on Robotics and Automation (ICRA)	May 2020 P20
ObjectNet: A large-scale bias-controlled dataset for pushing the limits of object recognition models	December 2019
A. Barbu, D. Mayo, J. Alverio, W. Luo, C. Wang, D. Gutfreund, J. Tenenbaum, and B. Katz Neural Information Processing Systems (NeurIPS) https://objectnet.dev	P21
Deep video-to-video transformations for accessibility with an application to photosensitiv-	June 2019
ity D. Banda, A. Barbu, and B. Katz Pattern Recognition Letters (PatRec)	P22
Partially Occluded Hands: A new dataset for single-image hand pose estimation B. Myanganbayar, C. Mata, G. Dekel, B. Katz, G. Ben-Yosef, A. Barbu Asian Conference on Computer Vision (ACCV)	December 2018 P23
First steps to understanding how symbolic communication arises J. Correa, Y. Kuo, B. Katz, and A. Barbu Workshop on Emergent Communication at NeurIPS	December 2018 P24
Grounding semantic parsing using caption videos C. Ross, A. Barbu, Yevgeni Berzak, Battushig Myanganbayar, and B. Katz Empirical Methods in Natural Language Processing (EMNLP) http://aclweb.org/anthology/D18-1285	November 2018 P25
Deep sequential models for sampling-based planning Y. Kuo, A. Barbu, and B. Katz International Conference on Intelligent Robots (IROS) http://0xab.com/papers/iros2018.pdf	October 2018 P26

Deep compositional models for robotic planning and language Y. Kuo, A. Barbu, and B. Katz Workshop on Language and Robotics at IROS	October 2018 P27
Temporal Grounding Graphs for Language Understanding with Accrued Visual-Linguistic Context R. Paul, A. Barbu, S. Felshin, B. Katz, and N. Roy International Joint Conference on Artificial Intelligence (IJCAI) http://0xab.com/papers/ijcai17.pdf	August 2017 P28
Temporal Grounding Graphs for Language Understanding with Accrued Visual-Linguistic Context R. Paul, A. Barbu, S. Felshin, B. Katz and N. Roy Language Grounding for Robotics Workshop at ACL http://0xab.com/papers/acl17.pdf	July 2017 P29
Saying What You're Looking For: Linguistics Meets Video Search D. Barrett, A. Barbu, N. Siddharth, and J. M. Siskind IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) http://0xab.com/papers/pami.pdf	October 2016 P30
Anchoring and Agreement in Syntactic Annotations Y. Berzak, Y. Huang, A. Barbu, A. Korhonen, and B. Katz Empirical Methods in Natural Language Processing (EMNLP) https://arxiv.org/pdf/1605.04481v3	September 2016 P31
Do You See What I Mean? Visual Resolution of Linguistic Ambiguities Y. Berzak, A. Barbu, D. Harari, B. Katz, and S. Ullman <i>Empirical Methods in Natural Language Processing (EMNLP)</i> http://0xab.com/papers/emnlp2015.pdf	September 2015 P32
A Compositional Framework for Grounding Language Inference, Generation, and Acquisition in Video H. Yu, N. Siddharth, A. Barbu, and J. M. Siskind Journal of Artificial Intelligence Research (JAIR) http://0xab.com/papers/jair.pdf	April 2015 P33
Seeing is <i>Worse</i> than Believing: Reading People's Minds Better than Computer-Vision	September 2014
Methods Recognize Actions A. Barbu, D. P. Barrett, W. Chen, N. Siddharth, C. Xiong, J. J. Corso, C. D. Fellbaum, C. Hanson, S. J. Hanson, S. Hélie, E. Malaia, B. A. Pearlmutter, J. M. Siskind, T. M. Talavage, and R. B. Wilbur European Conference on Computer Vision (ECCV) http://0xab.com/papers/eccv2014.pdf	P34
Seeing What You're Told: Sentence-Guided Activity Recognition In Video N. Siddharth, A. Barbu, J. M. Siskind Human-Machine Communication for Visual Recognition and Search Workshop at ECCV ftp://ftp.ecn.purdue.edu/qobi/cvpr2014.pdf	September 2014 P35
Language-driven video retrieval A. Barbu, N. Siddharth, J. M. Siskind Vision Meets Cognition at CVPR http://www.visionmeetscognition.org/fpic2014/Camera_Ready/Paper%2033.pdf	June 2014 P36
Seeing what you're told: sentence-guided activity recognition in video N. Siddharth, A. Barbu, and J. M. Siskind IEEE Conference on Computer Vision and Pattern Recognition (CVPR) http://0xab.com/papers/cvpr2014.pdf	June 2014 P37

Recognizing human activities from partially observed videos *June* 2013 Y. Cao, D. Barrett, A. Barbu, N. Siddharth, H. Yu, A. Michaux, Y. Lin, S. Dickinson, J. M. P38 Siskind, and S. Wang IEEE Conference on Computer Vision and Pattern Recognition (CVPR) poster 472/1870 (25.2%) http://0xab.com/papers/cvpr2013.pdf December 2012 Seeing unseeability to see the unseeable N. Siddharth, A. Barbu, and J. M. Siskind P39 Advances in Cognitive Systems, 2:77-94 oral 14/38 (37%), conference presentation associated with http://0xab.com/papers/acs2012a.pdf Simultaneous object detection, tracking, and event recognition December 2012 A. Barbu, N. Siddharth, A. Michaux, and J. M. Siskind P40 Advances in Cognitive Systems, 2:203-20 oral 14/38 (37%), conference presentation associated with http://0xab.com/papers/acs2012b.pdf Video in sentences out August 2012 A. Barbu, A. Bridge, Z. Burchill, D. Coroian, S. Dickinson, S. Fidler, A. Michaux, S. P41 Mussman, N. Siddharth, D. Salvi, L. Schmidt, J. Shangguan, J. M. Siskind, J. Waggoner, S. Wang, J. Wei, Y. Yin, and Z. Zhang Conference on Uncertainty In Artificial Intelligence (UAI), pp. 102–12 oral 24/304 (8%) http://0xab.com/papers/uai2012.pdf A visual language model for estimating object pose and structure May 2011 in a generative visual domain N. Siddharth, A. Barbu, and J. M. Siskind P42 IEEE International Conference on Robotics and Automation (ICRA) oral 982/2004 (49%) http://0xab.com/papers/icra2011.pdf Learning physically-instantiated robotic game play through visual observation May 2010 A. Barbu, N. Siddharth, and J. M. Siskind P43 IEEE International Conference on Robotics and Automation (ICRA) oral 856/2062 (42%) http://0xab.com/papers/icra2010.pdf **Patents** Deep Compositional Robotic Planners That Follow Natural Language Commands in review Y. Kuo, Boris Katz, A. Barbu A1 Provisional 62/944,924 and 62/944,932 Computer method and apparatus making screens safe for those with photosensitivity July 2022 A. Barbu, D. Banda, B. Katz *US patent US11381715* Correlating videos and sentences November 2015 J. M. Siskind, A. Barbu, N. Siddharth, and H. Yu A3 US patent US9183466 **Technical Reports** Bias and Agreement in Syntactic Annotations May 2016 Yevgeni Berzak, Yan Huang, A. Barbu, Anna Korhonen, Boris Katz T1 arXiv:1605.04481

http://arxiv.org/abs/1605.04481

The Compositional Nature of Event Representations in the Human Brain *June* 2013 A. Barbu, N. Siddharth, C. Xiong, J. J. Corso, C. D. Fellbaum, C. Hanson, S. J. Hanson, S. T2 Hélie, E. Malaia, B. A. Pearlmutter, J. M. Siskind, T. M. Talavage, and R. B. Wilbur arXiv:1306.2293 http://arxiv.org/abs/1306.2293 Large-scale automatic labeling of video events with verbs *April* 2012 based on event-participant interaction A. Barbu, A. Bridge, D. Coroian, S. Dickinson, S. Mussman, S. Narayanaswamy, D. Salvi, L. Т3 Schmidt, J. Shangguan, J. M. Siskind, J. Waggoner, S. Wang, J. Wei, Y. Yin, and Z. Zhang arXiv:1204.3616 http://arxiv.org/abs/1204.3616 Workshops organized Linguistics Meets Image and Video Retrieval at ICCV 2019 October 2019 Amrita Saha, Hui Wu, Adriana I. Kovashka, A. Barbu, Xiaoxiao Guo, Karthik Sankar-W1 anarayanan, Samarth Bharadwaj, and Yupeng Gao 3 invited speakers, hosted the Fashion IQ challenge https://sites.google.com/view/lingir June 2019 Vision and Language at CVPR 2019 A. Barbu, Siddharth N., D. Gutfreund 22 accepted papers, 5 invited speakers, over 200 attendees languageandvision.com June 2018 Combining Vision and Language ACCV 2018 Qi Wu, Peng Wang, Chuang Gan, Fumin Shen, Xiaodong He, A. Barbu, Anton van den W3 Hengel A workshop introducing new researchers to problems in vision and language, 7 invited speakers http://qi-wu.me/accv_v21/ Vision and Language at CVPR 2018 *June* 2018 A. Barbu, Siddharth N., Tao Mei, Y. Berzak, N. Shukla, J. Luo, R. Sukthankar, D. Gutfreund W4 16 accepted papers, 7 invited speakers, over 200 attendees languageandvision.com Vision and Language at CVPR 2017 July 2017 A. Barbu, Tao Mei, Siddharth Narayanaswamy, Puneet Kumar Dokania, Quanshi Zhang, W5 Nishant Shukla, Jiebo Luo, Rahul Sukthankar 17 accepted papers, 7 invited speakers, over 200 attendees languageandvision.com/2017.html June 2015 Vision and Language at CVPR 2015 A. Barbu, Georgios Evangelopoulus, Daniel Harari, Krystian Mikolajczyk, Siddharth W6 Narayanaswamy, Caiming Xiong, Yibiao Zhang 14 accepted papers, 9 invited speakers, over 100 attendees languageandvision.com/2015.html PhD students supervised Christopher Arnold 2023-present Multimodal story understanding Co-supervised with Boris Katz. 2022-2023 Ravi Tejwani Social perception Co-supervised with Boris Katz. Christopher Z Wang 2021-present Object recognition

Co-supervised with Boris Katz.

David Mayo Object recognition Co-supervised with Boris Katz.	2020-present S4
Aviv Netanyahu Social interactions Co-supervised with Boris Katz.	2019-2020 S5
Adam Yaari Language and neuroscience CEO and founder of Serinus Biosciences. Co-supervised with Boris Katz.	2018-present S6
Yen-Ling Kuo Grounding language in planning Faculty at University of Virginia. Co-supervised with Boris Katz.	2017-2022 S7
Candace Ross Grounded language acquisition Now a Postdoc at FAIR. Co-supervised with Boris Katz.	2016-2022 S8
Master's students supervised	
Abdulrahman Alabdulkareem Understanding modalities in detail Co-supervised with Boris Katz.	current M1
Anastasiia Uvarova Modeling the development of language Co-supervised with Boris Katz.	current M2
Laura Queipo Causal summarization Co-supervised with Boris Katz.	current M3
Vighnesh Subramaniam Understanding multimodal integration in the brain Co-supervised with Boris Katz.	current M4
Meagan R. Jens Machines that change the rules to win Co-supervised with Boris Katz.	2023 M5
Dana Rosenfarb Creating a large-scale benchmark for language in the brain Co-supervised with Boris Katz.	2022 M6
Felipe Monsalve Developing a secure ML platform for healthcare applications Co-supervised with Boris Katz.	2022 M7
Emily Cheng Understanding language evolution Now a PhD student. Co-supervised with Boris Katz.	2022 M8
Aaditya K. Singh Deep Attentional Modulation for Zero-shot Learning in Object Recognition Now at PhD student at UCL. Co-supervised with Boris Katz.	2021 M9
Christopher Z. Wang Modeling child language acquisition as grounded semantic parsing <i>Now at PhD student at MIT. Co-supervised with Boris Katz.</i>	2021 M10

Cognitively-plausible translation without aligned corpora Co-supervised with Boris Katz.	M11
Julian Alverio Robots that understand language Co-supervised with Boris Katz.	2021 M12
David Mayo Understanding human and machine object recognition at scale Co-supervised with Boris Katz.	2018 M13
Dalitso Banda Removing dangerous visual stimuli from videos Co-supervised with Boris Katz.	2018 M14
Battushig Myanganbayar Inverse graphics for hand-pose reconstruction Co-supervised with Boris Katz now at Apple.	2018 M15
Maria Ryskina Imaginative Paraphrase Recognition via Join Vision-Language Video Analysis Co-supervised with Boris Katz. Now a PhD student at CMU.	2016 M16
Sergey Voronov Common Sense Reasoning Through Imagination Co-supervised with Boris Katz. Now a PhD student at University of Ohio.	2016 M17
Nicolas Rakover A uniform Representation for Visual Concepts Co-supervised with Boris Katz. Now at Google.	2016 M18
Undergraduate students supervised	
Nazar Ilamanov Co-supervised with Boris Katz.	present U1
Julian Alverio Co-supervised with Boris Katz. Now a Masters student at MIT.	2018 U2
Gil Dekel Co-supervised with Boris Katz. Now at Google.	2016-2018 U3
Battushig Myanganbayar Co-supervised with Boris Katz. Now a Masters student at MIT.	201 <i>6</i> U4
David Mayo Co-supervised with Boris Katz. Now a Masters student at MIT.	2016 U5
Corinn Herrik Co-supervised with Boris Katz.	201 <i>6</i> U6
Sarah Coe Co-supervised with Boris Katz.	2015-2016 U7
Candace Ross Co-supervised with Boris Katz. Now a PhD student at MIT.	2015 U8
Victor Cabrera Co-supervised with Boris Katz.	2015 U9
Matthew Evanusa Co-supervised with Boris Katz.	2014 U10
Matthew Arbesfeld Co-supervised with Boris Katz.	2014-2015 U11

2021

M11

Allison Fu