

6th International Conference on Learning Representations (ICLR 2018)

Vancouver, Canada
30 April - 3 May 2018

Volume 1 of 7

ISBN: 978-1-7138-7272-6

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2018) by International Conference on Learning Representations
All rights reserved.

Printed with permission by Curran Associates, Inc. (2023)

For permission requests, please contact International Conference on Learning Representations
at the address below.

International Conference on Learning Representations
2710 E Corridor Drive
Appleton, WI 54913

www.iclr.cc

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

VOLUME 1

ORAL PAPERS

SPHERICAL CNNs.....	1
<i>Taco S. Cohen, Mario Geiger, Jonas Köhler, Max Welling</i>	
ON THE INSUFFICIENCY OF EXISTING MOMENTUM SCHEMES FOR STOCHASTIC OPTIMIZATION.....	16
<i>Rahul Kidambi, Praneeth Netrapalli, Prateek Jain, Sham M. Kakade</i>	
CERTIFYING SOME DISTRIBUTIONAL ROBUSTNESS WITH PRINCIPLED ADVERSARIAL TRAINING.....	39
<i>Aman Sinha, Hongseok Namkoong, John Duchi</i>	
BREAKING THE SOFTMAX BOTTLENECK: A HIGH-RANK RNN LANGUAGE MODEL	73
<i>Zhilin Yang, Zihang Dai, Ruslan Salakhutdinov, William W. Cohen</i>	
BOOSTING DILATED CONVOLUTIONAL NETWORKS WITH MIXED TENSOR DECOMPOSITIONS	91
<i>Nadav Cohen, Ronen Tamari, Amnon Shashua</i>	
AMBIENTGAN: GENERATIVE MODELS FROM LOSSY MEASUREMENTS.....	115
<i>Ashish Bora, Eric Price, Alexandros G. Dimakis</i>	
BEYOND WORD IMPORTANCE: CONTEXTUAL DECOMPOSITION TO EXTRACT INTERACTIONS FROM LSTMS.....	137
<i>W. James Murdoch, Peter J. Liu, Bin Yu</i>	
ASK THE RIGHT QUESTIONS: ACTIVE QUESTION REFORMULATION WITH REINFORCEMENT LEARNING	152
<i>Christian Buck, Jannis Bulian, Massimiliano Ciaramita, Wojciech Gajewski, Andrea Gesmundo, Neil Houlsby, Wei Wang</i>	
VARIANCE REDUCTION FOR POLICY GRADIENT WITH ACTION-DEPENDENT FACTORIZED BASELINES	167
<i>Cathy Wu, Aravind Rajeswaran, Yan Duan, Vikash Kumar, Alexandre M Bayen, Sham Kakade, Igor Mordatch, Pieter Abbeel</i>	
SYNTHETIC AND NATURAL NOISE BOTH BREAK NEURAL MACHINE TRANSLATION	184
<i>Yonatan Belinkov, Yonatan Bisk</i>	
WASSERSTEIN AUTO-ENCODERS.....	197
<i>Ilya Tolstikhin, Olivier Bousquet, Sylvain Gelly, Bernhard Schoelkopf</i>	
PROGRESSIVE GROWING OF GANS FOR IMPROVED QUALITY, STABILITY, AND VARIATION.....	213
<i>Tero Karras, Timo Aila, Samuli Laine, Jaakko Lehtinen</i>	
ON THE CONVERGENCE OF ADAM AND BEYOND	239
<i>Sashank J. Reddi, Satyen Kale, Sanjiv Kumar</i>	
TRAINING AND INFERENCE WITH INTEGERS IN DEEP NEURAL NETWORKS	262
<i>Shuang Wu, Guoqi Li, Feng Chen, Luping Shi</i>	
ZERO-SHOT VISUAL IMITATION	276
<i>Deepak Pathak, Parsa Mahmoudieh, Guanghao Luo, Pulkit Agrawal, Dian Chen, Yide Shentu, Evan Shelhamer, Jitendra Malik, Alexei A. Efros, Trevor Darrell</i>	

LEARNING TO REPRESENT PROGRAMS WITH GRAPHS	292
<i>Miltiadis Allamanis, Marc Brockschmidt, Mahmoud Khademi</i>	
MULTI-SCALE DENSE NETWORKS FOR RESOURCE EFFICIENT IMAGE CLASSIFICATION	309
<i>Gao Huang, Danlu Chen, Tianhong Li, Felix Wu, Laurens Van Der Maaten, Kilian Weinberger</i>	
NEURAL SKETCH LEARNING FOR CONDITIONAL PROGRAM GENERATION.....	323
<i>Vijayaraghavan Murali, Letao Qi, Swarat Chaudhuri, Chris Jermaine</i>	
CONTINUOUS ADAPTATION VIA META-LEARNING IN NONSTATIONARY AND COMPETITIVE ENVIRONMENTS	340
<i>Maruan Al-Shedivat, Trapit Bansal, Yura Burda, Ilya Sutskever, Igor Mordatch, Pieter Abbeel</i>	
LEARNING DEEP MEAN FIELD GAMES FOR MODELING LARGE POPULATION BEHAVIOR.....	361
<i>Jiachen Yang, Xiaojing Ye, Rakshit Trivedi, Huan Xu, Hongyuan Zha</i>	
SPECTRAL NORMALIZATION FOR GENERATIVE ADVERSARIAL NETWORKS	376
<i>Takeru Miyato, Toshiki Kataoka, Masanori Koyama, Yuichi Yoshida</i>	
CHARACTERIZING ADVERSARIAL SUBSPACES USING LOCAL INTRINSIC DIMENSIONALITY	402
<i>Xingjun Ma, Bo Li, Yisen Wang, Sarah M. Erfani, Sudanthi Wijewickrema, Grant Schoenebeck, Dawn Song, Michael E. Houle, James Bailey</i>	
EMERGENCE OF LINGUISTIC COMMUNICATION FROM REFERENTIAL GAMES WITH SYMBOLIC AND PIXEL INPUT	417
<i>Angeliki Lazaridou, Karl Moritz Hermann, Karl Tuyls, Stephen Clark</i>	

POSTER PRESENTATIONS

LEARNING GENERAL PURPOSE DISTRIBUTED SENTENCE REPRESENTATIONS VIA LARGE SCALE MULTI-TASK LEARNING	430
<i>Sandeep Subramanian, Adam Trischler, Yoshua Bengio, Christopher J Pal</i>	
RECASTING GRADIENT-BASED META-LEARNING AS HIERARCHICAL BAYES	446
<i>Erin Grant, Chelsea Finn, Sergey Levine, Trevor Darrell, Thomas Griffiths</i>	
FLIPOUT: EFFICIENT PSEUDO-INDEPENDENT WEIGHT PERTURBATIONS ON MINI-BATCHES	459
<i>Yeming Wen, Paul Vicol, Jimmy Ba, Dustin Tran, Roger Grosse</i>	
DORA THE EXPLORER: DIRECTED OUTREACHING REINFORCEMENT ACTION-SELECTION	475
<i>Lior Fox, Leshem Choshen, Yonatan Loewenstein</i>	
IMPLICIT CAUSAL MODELS FOR GENOME-WIDE ASSOCIATION STUDIES	492
<i>Dustin Tran, David M. Blei</i>	
DEEP VOICE 3: SCALING TEXT-TO-SPEECH WITH CONVOLUTIONAL SEQUENCE LEARNING.....	510
<i>Wei Ping, Kainan Peng, Andrew Gibiansky, Sercan O. Arik, Ajay Kannan, Sharan Narang, Jonathan Raiman, John Miller</i>	
PREDICTING FLOOR-LEVEL FOR 911 CALLS WITH NEURAL NETWORKS AND SMARTPHONE SENSOR DATA.....	526
<i>William Falcon, Henning Schulzrinne</i>	
DIVIDE AND CONQUER NETWORKS.....	542
<i>Alex Nowak, David Folqué, Joan Bruna</i>	

MONOTONIC CHUNKWISE ATTENTION	557
<i>Chung-Cheng Chiu, Colin Raffel</i>	
VARIATIONAL MESSAGE PASSING WITH STRUCTURED INFERENCE NETWORKS.....	573
<i>Wu Lin, Nicolas Hubacher, Mohammad Emtiyaz Khan</i>	
ALL-BUT-THE-TOP: SIMPLE AND EFFECTIVE POSTPROCESSING FOR WORD REPRESENTATIONS.....	586
<i>Jiaqi Mu, Pramod Viswanath</i>	
SGD LEARNS OVER-PARAMETERIZED NETWORKS THAT PROVABLY GENERALIZE ON LINEARLY SEPARABLE DATA.....	611
<i>Alon Brutzkus, Amir Globerson, Eran Malach, Shai Shalev-Shwartz</i>	
UNIVERSAL AGENT FOR DISENTANGLING ENVIRONMENTS AND TASKS.....	628
<i>Jiayuan Mao, Honghua Dong, Joseph J. Lim</i>	
POLICY OPTIMIZATION BY GENETIC DISTILLATION	648
<i>Tanmay Gangwani, Jian Peng</i>	
LATENT SPACE ODDITY: ON THE CURVATURE OF DEEP GENERATIVE MODELS.....	664
<i>Georgios Arvanitidis, Lars Kai Hansen, Søren Hauberg</i>	
DEPTHWISE SEPARABLE CONVOLUTIONS FOR NEURAL MACHINE TRANSLATION.....	679
<i>Lukasz Kaiser, Aidan N. Gomez, Francois Chollet</i>	
GENERATING NATURAL ADVERSARIAL EXAMPLES.....	689
<i>Zhengli Zhao, Dheeru Dua, Sameer Singh</i>	
ROBUSTNESS OF CLASSIFIERS TO UNIVERSAL PERTURBATIONS: A GEOMETRIC PERSPECTIVE	704
<i>Seyed-Mohsen Moosavi-Dezfooli, Alhussein Fawzi, Omar Fawzi, Pascal Frossard, Stefano Soatto</i>	
LEARNING APPROXIMATE INFERENCE NETWORKS FOR STRUCTURED PREDICTION.....	719
<i>Lifu Tu, Kevin Gimpel</i>	
PARAMETRIZED HIERARCHICAL PROCEDURES FOR NEURAL PROGRAMMING	734
<i>Roy Fox, Richard Shin, Sanjay Krishnan, Ken Goldberg, Dawn Song, Ion Stoica</i>	
NEURAL SPEED READING VIA SKIM-RNN.....	749
<i>Minjoon Seo, Sewon Min, Ali Farhadi, Hannaneh Hajishirzi</i>	
ROUTING NETWORKS: ADAPTIVE SELECTION OF NON-LINEAR FUNCTIONS FOR MULTI-TASK LEARNING	763
<i>Clemens Rosenbaum, Tim Klinger, Matthew Riemer</i>	
BACKPROPAGATION THROUGH THE VOID: OPTIMIZING CONTROL VARIATES FOR BLACK-BOX GRADIENT ESTIMATION.....	779
<i>Will Grathwohl, Dami Choi, Yuhuai Wu, Geoff Roeder, David Duvenaud</i>	
DEEP REWIRING: TRAINING VERY SPARSE DEEP NETWORKS.....	796
<i>Guillaume Bellec, David Kappel, Wolfgang Maass, Robert Legenstein</i>	

VOLUME 2

GAUSSIAN PROCESS BEHAVIOUR IN WIDE DEEP NEURAL NETWORKS	820
<i>Alexander G. De G. Matthews, Jiri Hron, Mark Rowland, Richard E. Turner, Zoubin Ghahramani</i>	
A COMPRESSED SENSING VIEW OF UNSUPERVISED TEXT EMBEDDINGS, BAG-OF-N-GRAMS, AND LSTMS.....	835
<i>Sanjeev Arora, Mikhail Khodak, Nikunj Saunshi, Kiran Vodrahalli</i>	

INTERPRETABLE COUNTING FOR VISUAL QUESTION ANSWERING	856
<i>Alexander Trott, Caiming Xiong, Richard Socher</i>	
ON UNIFYING DEEP GENERATIVE MODELS	874
<i>Zhiting Hu, Zichao Yang, Ruslan Salakhutdinov, Eric P. Xing</i>	
OVERCOMING CATASTROPHIC INTERFERENCE USING CONCEPTOR-AIDED BACKPROPAGATION.....	893
<i>Xu He, Herbert Jaeger</i>	
STOCHASTIC GRADIENT DESCENT PERFORMS VARIATIONAL INFERENCE, CONVERGES TO LIMIT CYCLES FOR DEEP NETWORKS	904
<i>Pratik Chaudhari, Stefano Soatto</i>	
PARALLELIZING LINEAR RECURRENT NEURAL NETS OVER SEQUENCE LENGTH.....	924
<i>Eric Martin, Chris Cundy</i>	
DISTRIBUTED FINE-TUNING OF LANGUAGE MODELS ON PRIVATE DATA	933
<i>Vadim Popov, Mikhail Kudinov, Irina Piontkovskaya, Petr Vytovtov, Alex Nevidomsky</i>	
STOCHASTIC VARIATIONAL VIDEO PREDICTION	945
<i>Mohammad Babaeizadeh, Chelsea Finn, Dumitru Erhan, Roy H. Campbell, Sergey Levine</i>	
MITIGATING ADVERSARIAL EFFECTS THROUGH RANDOMIZATION.....	960
<i>Cihang Xie, Jianyu Wang, Zhishuai Zhang, Zhou Ren, Alan Yuille</i>	
LEARNING TO COUNT OBJECTS IN NATURAL IMAGES FOR VISUAL QUESTION ANSWERING	976
<i>Yan Zhang, Jonathon Hare, Adam Prügel-Bennett</i>	
REGULARIZING AND OPTIMIZING LSTM LANGUAGE MODELS	993
<i>Stephen Merity, Nitish Shirish Keskar, Richard Socher</i>	
LARGE SCALE OPTIMAL TRANSPORT AND MAPPING ESTIMATION.....	1006
<i>Vivien Seguy, Bharath Bhushan Damodaran, Remi Flamary, Nicolas Courty, Antoine Rolet, Mathieu Blondel</i>	
SEMANTICALLY DECOMPOSING THE LATENT SPACES OF GENERATIVE ADVERSARIAL NETWORKS	1021
<i>Chris Donahue, Zachary C. Lipton, Akshay Balsubramani, Julian McAuley</i>	
A NEURAL REPRESENTATION OF SKETCH DRAWINGS	1040
<i>David Ha, Douglas Eck</i>	
DEEP CONTEXTUALIZED WORD REPRESENTATIONS	1056
<i>Authors Anonymous</i>	
ADAPTIVE QUANTIZATION OF NEURAL NETWORKS	1073
<i>Soroosh Khoram, Jing Li</i>	
MIXED PRECISION TRAINING	1086
<i>Paulius Micikevicius, Sharan Narang, Jonah Alben, Gregory Diamos, Erich Elsen, David Garcia, Boris Ginsburg, Michael Houston, Oleksii Kuchaiev, Ganesh Venkatesh, Hao Wu</i>	
DEEP NEURAL NETWORKS AS GAUSSIAN PROCESSES	1098
<i>Jaehoon Lee, Yasaman Bahri, Roman Novak, Samuel S. Schoenholz, Jeffrey Pennington, Jascha Sohl-Dickstein</i>	
ALTERNATING MULTI-BIT QUANTIZATION FOR RECURRENT NEURAL NETWORKS	1115
<i>Chen Xu, Jianqiang Yao, Zhouchen Lin, Wenwu Ou, Yuanbin Cao, Zhirong Wang, Hongbin Zha</i>	
GLOBAL OPTIMALITY CONDITIONS FOR DEEP NEURAL NETWORKS	1128
<i>Chulhee Yun, Suvrit Sra, Ali Jadbabaie</i>	

DETECTING STATISTICAL INTERACTIONS FROM NEURAL NETWORK WEIGHTS.....	1142
<i>Michael Tsang, Dehua Cheng, Yan Liu</i>	
TOWARDS BETTER UNDERSTANDING OF GRADIENT-BASED ATTRIBUTION METHODS FOR DEEP NEURAL NETWORKS	1163
<i>Marco Ancona, Enea Ceolini, Cengiz Öztireli, Markus Gross</i>	
TRAINING GENERATIVE ADVERSARIAL NETWORKS VIA PRIMAL-DUAL SUBGRADIENT METHODS: A LAGRANGIAN PERSPECTIVE ON GAN.....	1179
<i>Xu Chen, Jiang Wang, Hao Ge</i>	
MEMORY-BASED PARAMETER ADAPTATION	1193
<i>Pablo Sprechmann, Siddhant M. Jayakumar, Jack W. Rae, Alexander Pritzel, Adria Puigdomenech Badia, Benigno Uria, Oriol Vinyals, Demis Hassabis, Razvan Pascanu, Charles Blundell</i>	
LEARN TO PAY ATTENTION.....	1209
<i>Saumya Jetley, Nicholas A. Lord, Namhoon Lee, Philip H. S. Torr</i>	
GRADIENT ESTIMATORS FOR IMPLICIT MODELS	1223
<i>Yingzhen Li, Richard E. Turner</i>	
LEARNING ONE-HIDDEN-LAYER NEURAL NETWORKS WITH LANDSCAPE DESIGN.....	1242
<i>Rong Ge, Jason D. Lee, Tengyu Ma</i>	
ACTIVE NEURAL LOCALIZATION	1279
<i>Devendra Singh Chaplot, Emilio Parisotto, Ruslan Salakhutdinov</i>	
TEMPORALLY EFFICIENT DEEP LEARNING WITH SPIKES	1294
<i>Peter O'Connor, Efstratios Gavves, Matthias Reisser, Max Welling</i>	
THE KANERVA MACHINE: A GENERATIVE DISTRIBUTED MEMORY	1313
<i>Yan Wu, Greg Wayne, Alex Graves, Timothy Lillicrap</i>	
CAN NEURAL NETWORKS UNDERSTAND LOGICAL ENTAILMENT?.....	1329
<i>Richard Evans, David Saxton, David Amos, Pushmeet Kohli, Edward Grefenstette</i>	
UNDERSTANDING IMAGE MOTION WITH GROUP REPRESENTATIONS.....	1344
<i>Andrew Jaegle, Stephen Phillips, Daphne Ippolito, Kostas Daniilidis</i>	
VITERBI-BASED PRUNING FOR SPARSE MATRIX WITH FIXED AND HIGH INDEX COMPRESSION RATIO	1358
<i>Dongsoo Lee, Daehyun Ahn, Taesu Kim, Pierce I. Chuang, Jae-Joon Kim</i>	
VOICELOOP: VOICE FITTING AND SYNTHESIS VIA A PHONOLOGICAL LOOP.....	1374
<i>Yaniv Taigman, Lior Wolf, Adam Polyak, Eliya Nachmani</i>	
FEW-SHOT LEARNING WITH GRAPH NEURAL NETWORKS	1388
<i>Victor Garcia Satorras, Joan Bruna Estrach</i>	
BEYOND SHARED HIERARCHIES: DEEP MULTITASK LEARNING THROUGH SOFT LAYER ORDERING	1401
<i>Elliot Meyerson, Risto Miikkulainen</i>	
CAUSALGAN: LEARNING CAUSAL IMPLICIT GENERATIVE MODELS WITH ADVERSARIAL TRAINING	1415
<i>Murat Kocaoglu, Christopher Snyder, Alexandros G. Dimakis, Sriram Vishwanath</i>	
TRUNCATED HORIZON POLICY SEARCH: COMBINING REINFORCEMENT LEARNING & IMITATION LEARNING	1448
<i>Wen Sun, J. Andrew Bagnell, Byron Boots</i>	
TOWARDS IMAGE UNDERSTANDING FROM DEEP COMPRESSION WITHOUT DECODING	1462
<i>Robert Torfason, Fabian Mentzer, Eirikur Agustsson, Michael Tschannen, Radu Timofte, Luc Van Gool</i>	

LEVERAGING GRAMMAR AND REINFORCEMENT LEARNING FOR NEURAL PROGRAM SYNTHESIS.....	1479
<i>Rudy Bunel, Matthew Hausknecht, Jacob Devlin, Rishabh Singh, Pushmeet Kohli</i>	
DEEP LEARNING FOR PHYSICAL PROCESSES: INCORPORATING PRIOR SCIENTIFIC KNOWLEDGE	1494
<i>Emmanuel De Bezenac, Arthur Pajot, Patrick Gallinari</i>	
EMERGENCE OF GRID-LIKE REPRESENTATIONS BY TRAINING RECURRENT NEURAL NETWORKS TO PERFORM SPATIAL LOCALIZATION.....	1512
<i>Christopher J. Cueva, Xue-Xin Wei</i>	
LEARNING LATENT PERMUTATIONS WITH GUMBEL-SINKHORN NETWORKS.....	1531
<i>Gonzalo Mena, David Belanger, Scott Linderman, Jasper Snoek</i>	
EVIDENCE AGGREGATION FOR ANSWER RE-RANKING IN OPEN-DOMAIN QUESTION ANSWERING.....	1553
<i>Shuohang Wang, Mo Yu, Jing Jiang, Wei Zhang, Xiaoxiao Guo, Shiyu Chang, Zhiguo Wang, Tim Klinger, Gerald Tesauro, Murray Campbell</i>	
FEARNET: BRAIN-INSPIRED MODEL FOR INCREMENTAL LEARNING.....	1567
<i>Ronald Kemker, Christopher Kanan</i>	
SEMANTIC INTERPOLATION IN IMPLICIT MODELS.....	1583
<i>Yannic Kilcher, Aurelien Lucchi, Thomas Hofmann</i>	
THE REACTOR: A FAST AND SAMPLE-EFFICIENT ACTOR-CRITIC AGENT FOR REINFORCEMENT LEARNING	1611
<i>Audrunas Gruslys, Will Dabney, Mohammad Gheshlaghi Azar, Bilal Piot, Marc Bellemare, Remi Munos</i>	

VOLUME 3

MIXED PRECISION TRAINING OF CONVOLUTIONAL NEURAL NETWORKS USING INTEGER OPERATIONS.....	1630
<i>Dipankar Das, Naveen Mellempudi, Dheevatsa Mudigere, Dhiraj Kalamkar, Sasikanth Avancha, Kunal Banerjee, Srinivas Sridharan, Karthik Vaidyanathan, Bharat Kaul, Evangelos Georganas, Alexander Heinecke, Pradeep Dubey, Jesus Corbal, Nikita Shustrov, Roma Dubtsov, Evarist Fomenko, Vadim Pirogov</i>	
AN EFFICIENT FRAMEWORK FOR LEARNING SENTENCE REPRESENTATIONS	1641
<i>Lajanugen Logeswaran, Honglak Lee</i>	
HIERARCHICAL SUBTASK DISCOVERY WITH NON-NEGATIVE MATRIX FACTORIZATION.....	1657
<i>Adam C. Earle, Andrew M. Saxe, Benjamin Rosman</i>	
SPECTRALNET: SPECTRAL CLUSTERING USING DEEP NEURAL NETWORKS	1668
<i>Uri Shaham, Kelly Stanton, Henry Li, Ronen Basri, Boaz Nadler, Yuval Kluger</i>	
LEARNING FROM NOISY SINGLY-LABELED DATA.....	1688
<i>Ashish Khetan, Zachary C. Lipton, Animashree Anandkumar</i>	
AN IMAGE REPRESENTATION BASED CONVOLUTIONAL NETWORK FOR DNA CLASSIFICATION	1703
<i>Bojian Yin, Marleen Balvert, Davide Zambrano, Alexander Schoenhuth, Sander Bohte</i>	
ADAPTIVE DROPOUT WITH RADEMACHER COMPLEXITY REGULARIZATION	1717
<i>Ke Zhai, Huan Wang</i>	
QUANTITATIVELY EVALUATING GANS WITH DIVERGENCES PROPOSED FOR TRAINING.....	1734
<i>Daniel Jiwoong Im, He Ma, Graham W. Taylor, Kristin Branson</i>	

MEASURING THE INTRINSIC DIMENSION OF OBJECTIVE LANDSCAPES	1764
<i>Chunyuan Li, Heerad Farkhoor, Rosanne Liu, Jason Yosinski</i>	
LIFELONG LEARNING WITH DYNAMICALLY EXPANDABLE NETWORKS	1788
<i>Jaehong Yoon, Eunho Yang, Jeongtae Lee, Sung Ju Hwang</i>	
LEAVE NO TRACE: LEARNING TO RESET FOR SAFE AND AUTONOMOUS REINFORCEMENT LEARNING	1799
<i>Benjamin Eysenbach, Shixiang Gu, Julian Ibarz, Sergey Levine</i>	
LATENT CONSTRAINTS: LEARNING TO GENERATE CONDITIONALLY FROM UNCONDITIONAL GENERATIVE MODELS	1817
<i>Jesse Engel, Matthew Hoffman, Adam Roberts</i>	
ACTION-DEPENDENT CONTROL VARIATES FOR POLICY OPTIMIZATION VIA STEIN IDENTITY	1839
<i>Hao Liu, Yihao Feng, Yi Mao, Dengyong Zhou, Jian Peng, Qiang Liu</i>	
UNDERSTANDING SHORT-HORIZON BIAS IN STOCHASTIC META-OPTIMIZATION	1855
<i>Yuhuai Wu, Mengye Ren, Renjie Liao, Roger Grosse</i>	
IMPROVING GAN TRAINING VIA BINARIZED REPRESENTATION ENTROPY (BRE) REGULARIZATION	1872
<i>Yanshuai Cao, Gavin Weiguang Ding, Kry Yik-Chau Lui, Ruitong Huang</i>	
ENSEMBLE ADVERSARIAL TRAINING: ATTACKS AND DEFENSES	1894
<i>Florian Tramèr, Alexey Kurakin, Nicolas Papernot, Ian Goodfellow, Dan Boneh, Patrick McDaniel</i>	
PIXELNN: EXAMPLE-BASED IMAGE SYNTHESIS	1914
<i>Aayush Bansal, Yaser Sheikh, Deva Ramanan</i>	
LEARNING DIFFERENTIALLY PRIVATE RECURRENT LANGUAGE MODELS	1926
<i>H. Brendan McMahan, Daniel Ramage, Kunal Talwar, Li Zhang</i>	
DEEP LEARNING AS A MIXED CONVEX-COMBINATORIAL OPTIMIZATION PROBLEM	1940
<i>Abram L. Friesen, Pedro Domingos</i>	
LEARNING HOW TO EXPLAIN NEURAL NETWORKS: PATTERNNET AND PATTERNATTRIBUTION	1954
<i>Pieter-Jan Kindermans, Kristof T. Schütt, Maximilian Alber, Klaus-Robert Müller, Dumitru Erhan, Been Kim, Sven Dähne</i>	
IMPROVING GANS USING OPTIMAL TRANSPORT	1970
<i>Tim Salimans, Han Zhang, Alec Radford, Dimitris Metaxas</i>	
DECISION-BASED ADVERSARIAL ATTACKS: RELIABLE ATTACKS AGAINST BLACK- BOX MACHINE LEARNING MODELS	1983
<i>Wieland Brendel, Jonas Rauber, Matthias Bethge</i>	
LEARNING A NEURAL RESPONSE METRIC FOR RETINAL PROSTHESIS	1995
<i>Nishal P Shah, Sasidhar Madugula, Ej Chichilnisky, Yoram Singer, Jonathon Shlens</i>	
INTRINSIC MOTIVATION AND AUTOMATIC CURRICULA VIA ASYMMETRIC SELF- PLAY	2010
<i>Sainbayar Sukhbaatar, Zeming Lin, Ilya Kostrikov, Gabriel Synnaeve, Arthur Szlam, Rob Fergus</i>	
SMASH: ONE-SHOT MODEL ARCHITECTURE SEARCH THROUGH HYPERNETWORKS	2026
<i>Andrew Brock, Theo Lim, J. M. Ritchie, Nick Weston</i>	
MATRIX CAPSULES WITH EM ROUTING	2048
<i>Geoffrey E Hinton, Sara Sabour, Nicholas Frosst</i>	

CASCADE ADVERSARIAL MACHINE LEARNING REGULARIZED WITH A UNIFIED EMBEDDING	2063
<i>Taesik Na, Jong Hwan Ko, Saibal Mukhopadhyay</i>	
DIFFUSION CONVOLUTIONAL RECURRENT NEURAL NETWORK: DATA-DRIVEN TRAFFIC FORECASTING	2079
<i>Yaguang Li, Rose Yu, Cyrus Shahabi, Yan Liu</i>	
NEURAL LANGUAGE MODELING BY JOINTLY LEARNING SYNTAX AND LEXICON.....	2095
<i>Yikang Shen, Zhouhan Lin, Chin-Wei Huang, Aaron Courville</i>	
WRPN: WIDE REDUCED-PRECISION NETWORKS	2113
<i>Asit Mishra, Eriko Nurvitadhi, Jeffrey J Cook, Debbie Marr</i>	
LEARNING LATENT REPRESENTATIONS IN NEURAL NETWORKS FOR CLUSTERING THROUGH PSEUDO SUPERVISION AND GRAPH-BASED ACTIVITY REGULARIZATION.....	2124
<i>Ozsel Kilinc, Ismail Uysal</i>	
LEARNING WASSERSTEIN EMBEDDINGS.....	2140
<i>Nicolas Courty, Rémi Flamary, Mélanie Ducoffe</i>	
REINFORCEMENT LEARNING ON WEB INTERFACES USING WORKFLOW-GUIDED EXPLORATION	2153
<i>Evan Zheran Liu, Kelvin Guu, Panupong Pasupat, Tianlin Shi, Percy Liang</i>	
A DEEP REINFORCED MODEL FOR ABSTRACTIVE SUMMARIZATION	2168
<i>Romain Paulus, Caiming Xiong, Richard Socher</i>	
A FRAMEWORK FOR THE QUANTITATIVE EVALUATION OF DISENTANGLED REPRESENTATIONS.....	2181
<i>Cian Eastwood, Christopher K. I. Williams</i>	
GANITE: ESTIMATION OF INDIVIDUALIZED TREATMENT EFFECTS USING GENERATIVE ADVERSARIAL NETS.....	2196
<i>Jinsung Yoon, James Jordon, Mihaela Van Der Schaar</i>	
MULTI-TASK LEARNING FOR DOCUMENT RANKING AND QUERY SUGGESTION	2218
<i>Wasi Uddin Ahmad, Kai-Wei Chang, Hongning Wang</i>	
ACTIVE LEARNING FOR CONVOLUTIONAL NEURAL NETWORKS: A CORE-SET APPROACH.....	2232
<i>Ozan Sener, Silvio Savarese</i>	
UNSUPERVISED REPRESENTATION LEARNING BY PREDICTING IMAGE ROTATIONS.....	2245
<i>Spyros Gidaris, Praveer Singh, Nikos Komodakis</i>	
A HIERARCHICAL MODEL FOR DEVICE PLACEMENT.....	2261
<i>Azalia Mirhoseini, Anna Goldie, Hieu Pham, Benoit Steiner, Quoc V. Le, Jeff Dean</i>	
MULTI-LEVEL RESIDUAL NETWORKS FROM DYNAMICAL SYSTEMS VIEW	2272
<i>Bo Chang, Lili Meng, Eldad Haber, Frederick Tung, David Begert</i>	
DEEP GAUSSIAN EMBEDDING OF GRAPHS: UNSUPERVISED INDUCTIVE LEARNING VIA RANKING.....	2286
<i>Aleksandar Bojchevski, Stephan Günnemann</i>	
NATURAL LANGUAGE INFERENCE OVER INTERACTION SPACE	2299
<i>Yichen Gong, Heng Luo, Jian Zhang</i>	
SPATIALLY TRANSFORMED ADVERSARIAL EXAMPLES.....	2314
<i>Chaowei Xiao, Jun-Yan Zhu, Bo Li, Warren He, Mingyan Liu, Dawn Song</i>	
THERMOMETER ENCODING: ONE HOT WAY TO RESIST ADVERSARIAL EXAMPLES	2344
<i>Jacob Buckman, Aurko Roy, Colin Raffel, Ian Goodfellow</i>	

DEEP GRADIENT COMPRESSION: REDUCING THE COMMUNICATION BANDWIDTH FOR DISTRIBUTED TRAINING	2366
<i>Yujun Lin, Song Han, Huizi Mao, Yu Wang, Bill Dally</i>	
IDENTIFYING ANALOGIES ACROSS DOMAINS	2380
<i>Yedid Hoshen, Lior Wolf</i>	
HIERARCHICAL AND INTERPRETABLE SKILL ACQUISITION IN MULTI-TASK REINFORCEMENT LEARNING	2392
<i>Tianmin Shu, Caiming Xiong, Richard Socher</i>	
LEARNING PARAMETRIC CLOSED-LOOP POLICIES FOR MARKOV POTENTIAL GAMES	2407
<i>Sergio Valcarcel Macua, Javier Zazo, Santiago Zazo</i>	
SELF-ENSEMBLING FOR VISUAL DOMAIN ADAPTATION	2423
<i>Geoff French, Michal Mackiewicz, Mark Fisher</i>	

VOLUME 4

NOISY NETWORKS FOR EXPLORATION	2438
<i>Meire Fortunato, Mohammad Gheshlaghi Azar, Bilal Piot, Jacob Menick, Matteo Hessel, Ian Osband, Alex Graves, Volodymyr Mnih, Remi Munos, Demis Hassabis, Olivier Pietquin, Charles Blundell, Shane Legg</i>	
MGAN: TRAINING GENERATIVE ADVERSARIAL NETS WITH MULTIPLE GENERATORS	2459
<i>Quan Hoang, Tu Dinh Nguyen, Trung Le, Dinh Phung</i>	
PROGRESSIVE REINFORCEMENT LEARNING WITH DISTILLATION FOR MULTI-SKILLED MOTION CONTROL	2483
<i>Glen Berseth, Cheng Xie, Paul Cernek, Michiel Van De Panne</i>	
MODEL COMPRESSION VIA DISTILLATION AND QUANTIZATION	2498
<i>Antonio Polino, Razvan Pascanu, Dan Alistarh</i>	
THE HIGH-DIMENSIONAL GEOMETRY OF BINARY NEURAL NETWORKS	2519
<i>Alexander G. Anderson, Cory P. Berg</i>	
DEEP LEARNING WITH LOGGED BANDIT FEEDBACK	2534
<i>Thorsten Joachims, Adith Swaminathan, Maarten De Rijke</i>	
TRAINING CONFIDENCE-CALIBRATED CLASSIFIERS FOR DETECTING OUT-OF-DISTRIBUTION SAMPLES	2546
<i>Kimin Lee, Honglak Lee, Kibok Lee, Jinwoo Shin</i>	
UNSUPERVISED CIPHER CRACKING USING DISCRETE GANS	2562
<i>Aidan N. Gomez, Sicong Huang, Ivan Zhang, Bryan M. Li, Muhammad Osama, Lukasz Kaiser</i>	
POLAR TRANSFORMER NETWORKS	2577
<i>Carlos Esteves, Christine Allen-Blanchette, Xiaowei Zhou, Kostas Daniilidis</i>	
FIDELITY-WEIGHTED LEARNING	2591
<i>Mostafa Dehghani, Arash Mehrjou, Stephan Gouws, Jaap Kamps, Bernhard Schölkopf</i>	
WHAI: WEIBULL HYBRID AUTOENCODING INFERENCE FOR DEEP TOPIC MODELING	2611
<i>Hao Zhang, Bo Chen, Dandan Guo, Mingyuan Zhou</i>	
APPRENTICE: USING KNOWLEDGE DISTILLATION TECHNIQUES TO IMPROVE LOW-PRECISION NETWORK ACCURACY	2626
<i>Asit Mishra, Debbie Marr</i>	
UNDERSTANDING DEEP NEURAL NETWORKS WITH RECTIFIED LINEAR UNITS	2643
<i>Raman Arora, Amitabh Basu, Poorya Mianjy, Anirbit Mukherjee</i>	
HEXA CONV	2660
<i>Emiel Hoogeboom, Jorn W. T. Peters, Taco S. Cohen, Max Welling</i>	

TRUST-PCL: AN OFF-POLICY TRUST REGION METHOD FOR CONTINUOUS CONTROL	2671
<i>Ofir Nachum, Mohammad Norouzi, Kelvin Xu, Dale Schuurmans</i>	
GENERATIVE NETWORKS AS INVERSE PROBLEMS WITH SCATTERING TRANSFORMS.....	2685
<i>Tomás Angles, Stéphane Mallat</i>	
I-REVNENET: DEEP INVERTIBLE NETWORKS.....	2695
<i>Jörn-Henrik Jacobsen, Arnold W. M. Smeulders, Edouard Oyallon</i>	
TRAINING GANS WITH OPTIMISM	2706
<i>Constantinos Daskalakis, Andrew Ilyas, Vasilis Syrgkanis, Haoyang Zeng</i>	
DON'T DECAY THE LEARNING RATE, INCREASE THE BATCH SIZE.....	2736
<i>Samuel L. Smith, Pieter-Jan Kindermans, Chris Ying, Quoc V. Le</i>	
RELATIONAL NEURAL EXPECTATION MAXIMIZATION: UNSUPERVISED DISCOVERY OF OBJECTS AND THEIR INTERACTIONS	2747
<i>Sjoerd Van Steenkiste, Michael Chang, Klaus Greff, Jürgen Schmidhuber</i>	
A SIMPLE NEURAL ATTENTIVE META-LEARNER	2762
<i>Nikhil Mishra, Mostafa Rohaninejad, Xi Chen, Pieter Abbeel</i>	
CRITICAL PERCOLATION AS A FRAMEWORK TO ANALYZE THE TRAINING OF DEEP NETWORKS.....	2779
<i>Zohar Ringel, Rodrigo Andrade De Bem</i>	
ATTACKING BINARIZED NEURAL NETWORKS.....	2796
<i>Angus Galloway, Graham W. Taylor, Medhat Moussa</i>	
SIMULATED+UNSUPERVISED LEARNING WITH ADAPTIVE DATA GENERATION AND BIDIRECTIONAL MAPPINGS	2810
<i>Kangwook Lee, Hoon Kim, Changho Suh</i>	
ESPRESSO: EFFICIENT FORWARD PROPAGATION FOR BINARY DEEP NEURAL NETWORKS.....	2825
<i>Fabrizio Pedersoli, George Tzanetakis, Andrea Tagliasacchi</i>	
EXPRESSIVE POWER OF RECURRENT NEURAL NETWORKS.....	2835
<i>Valentin Khrulkov, Alexander Novikov, Ivan Oseledets</i>	
MEMORY AUGMENTED CONTROL NETWORKS.....	2847
<i>Arbaaz Khan, Clark Zhang, Nikolay Atanasov, Konstantinos Karydis, Vijay Kumar, Daniel D. Lee</i>	
MIXUP: BEYOND EMPIRICAL RISK MINIMIZATION.....	2866
<i>Hongyi Zhang, Moustapha Cisse, Yann N. Dauphin, David Lopez-Paz</i>	
MODULAR CONTINUAL LEARNING IN A UNIFIED VISUAL ENVIRONMENT.....	2879
<i>Kevin T. Feigelis, Blue Sheffer, Daniel L. K. Yamins</i>	
SPARSE PERSISTENT RNNs: SQUEEZING LARGE RECURRENT NETWORKS ON-CHIP.....	2905
<i>Feiwen Zhu, Jeff Pool, Michael Andersch, Jeremy Appleyard, Fung Xie</i>	
GRAPH ATTENTION NETWORKS	2920
<i>Petar Velićković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, Yoshua Bengio</i>	
N2N LEARNING: NETWORK TO NETWORK COMPRESSION VIA POLICY GRADIENT REINFORCEMENT LEARNING	2932
<i>Anubhav Ashok, Nicholas Rhinehart, Fares Beainy, Kris M. Kitani</i>	
NEUMANN OPTIMIZER: A PRACTICAL OPTIMIZATION ALGORITHM FOR DEEP NEURAL NETWORKS.....	2953
<i>Shankar Krishnan, Ying Xiao, Rif. A. Saurous</i>	

PARAMETER SPACE NOISE FOR EXPLORATION	2968
<i>Matthias Plappert, Rein Houthoofd, Prafulla Dhariwal, Szymon Sidor, Richard Y. Chen, Xi Chen, Tamim Asfour, Pieter Abbeel, Marcin Andrychowicz</i>	
SKIP CONNECTIONS ELIMINATE SINGULARITIES	2986
<i>Emin Orhan, Xaq Pitkow</i>	
WHEN IS A CONVOLUTIONAL FILTER EASY TO LEARN?.....	3008
<i>Simon S. Du, Jason D. Lee, Yuandong Tian</i>	
AUTO-ENCODING SEQUENTIAL MONTE CARLO.....	3028
<i>Tuan Anh Le, Maximilian Igl, Tom Rainforth, Tom Jin, Frank Wood</i>	
SOBOLEV GAN.....	3045
<i>Youssef Mroueh, Chun-Liang Li, Tom Sercu, Anant Raj, Yu Cheng</i>	
DEEP AUTOENCODING GAUSSIAN MIXTURE MODEL FOR UNSUPERVISED ANOMALY DETECTION.....	3072
<i>Bo Zong, Qi Song, Martin Renqiang Min, Wei Cheng, Cristian Lumezanu, Daeki Cho, Haifeng Chen</i>	
CAN RECURRENT NEURAL NETWORKS WARP TIME?.....	3091
<i>Corentin Tallec, Yann Ollivier</i>	
FASTGCN: FAST LEARNING WITH GRAPH CONVOLUTIONAL NETWORKS VIA IMPORTANCE SAMPLING	3104
<i>Jie Chen, Tengfei Ma, Cao Xiao</i>	
A PAC-BAYESIAN APPROACH TO SPECTRALLY-NORMALIZED MARGIN BOUNDS FOR NEURAL NETWORKS.....	3119
<i>Behnam Neyshabur, Srinadh Bhojanapalli, Nathan Srebro</i>	
EMERGENT COMMUNICATION THROUGH NEGOTIATION	3128
<i>Kris Cao, Angeliki Lazaridou, Marc Lanctot, Joel Z Leibo, Karl Tuyls, Stephen Clark</i>	
DO GANS LEARN THE DISTRIBUTION? SOME THEORY AND EMPIRICS.....	3143
<i>Sanjeev Arora, Andrej Risteski, Yi Zhang</i>	
AUTO-CONDITIONED RECURRENT NETWORKS FOR EXTENDED COMPLEX HUMAN MOTION SYNTHESIS.....	3159
<i>Yi Zhou, Zimo Li, Shuangjiu Xiao, Chong He, Zeng Huang, Hao Li</i>	
GUIDE ACTOR-CRITIC FOR CONTINUOUS CONTROL.....	3172
<i>Voot Tangkaratt, Abbas Abdolmaleki, Masashi Sugiyama</i>	
KERNEL IMPLICIT VARIATIONAL INFERENCE	3196
<i>Jiaxin Shi, Shengyang Sun, Jun Zhu</i>	
MODEL-ENSEMBLE TRUST-REGION POLICY OPTIMIZATION.....	3219
<i>Thanard Kurutach, Ignasi Clavera, Yan Duan, Aviv Tamar, Pieter Abbeel</i>	

VOLUME 5

UNSUPERVISED LEARNING OF GOAL SPACES FOR INTRINSICALLY MOTIVATED GOAL EXPLORATION.....	3234
<i>Alexandre Péré, Sébastien Forestier, Olivier Sigaud, Pierre-Yves Oudeyer</i>	
PIXELDEFEND: LEVERAGING GENERATIVE MODELS TO UNDERSTAND AND DEFEND AGAINST ADVERSARIAL EXAMPLES.....	3260
<i>Yang Song, Taesup Kim, Sebastian Nowozin, Stefano Ermon, Nate Kushman</i>	
SIMULATING ACTION DYNAMICS WITH NEURAL PROCESS NETWORKS	3280
<i>Antoine Bosselut, Omer Levy, Ari Holtzman, Corin Ennis, Dieter Fox, Yejin Choi</i>	

ON THE INFORMATION BOTTLENECK THEORY OF DEEP LEARNING	3295
<i>Andrew Michael Saxe, Yamini Bansal, Joel Dapello, Madhu Advani, Artemy Kolchinsky, Brendan Daniel Tracey, David Daniel Cox</i>	
DEEP SENSING: ACTIVE SENSING USING MULTI-DIRECTIONAL RECURRENT NEURAL NETWORKS.....	3322
<i>Jinsung Yoon, William R. Zame, Mihaela Van Der Schaar</i>	
LEARNING FROM BETWEEN-CLASS EXAMPLES FOR DEEP SOUND RECOGNITION	3341
<i>Yuji Tokozume, Yoshitaka Ushiku, Tatsuya Harada</i>	
EMERGENT COMMUNICATION IN A MULTI-MODAL, MULTI-STEP REFERENTIAL GAME	3354
<i>Katrina Evtimova, Andrew Drozdov, Douwe Kiela, Kyunghyun Cho</i>	
DEBIASING EVIDENCE APPROXIMATIONS: ON IMPORTANCE-WEIGHTED AUTOENCODERS AND JACKKNIFE VARIATIONAL INFERENCE	3367
<i>Sebastian Nowozin</i>	
DECISION BOUNDARY ANALYSIS OF ADVERSARIAL EXAMPLES.....	3383
<i>Warren He, Bo Li, Dawn Song</i>	
DIVIDE-AND-CONQUER REINFORCEMENT LEARNING	3398
<i>Dibya Ghosh, Avi Singh, Aravind Rajeswaran, Vikash Kumar, Sergey Levine</i>	
STABILIZING ADVERSARIAL NETS WITH PREDICTION METHODS	3412
<i>Abhay Yadav, Sohil Shah, Zheng Xu, David Jacobs, Tom Goldstein</i>	
EVALUATING THE ROBUSTNESS OF NEURAL NETWORKS: AN EXTREME VALUE THEORY APPROACH	3433
<i>Tsui-Wei Weng, Huan Zhang, Pin-Yu Chen, Jinfeng Yi, Dong Su, Yupeng Gao, Cho-Jui Hsieh, Luca Daniel</i>	
NOT-SO-RANDOM FEATURES	3451
<i>Brian Bullins, Cyril Zhang, Yi Zhang</i>	
MEMORIZATION PRECEDES GENERATION: LEARNING UNSUPERVISED GANS WITH MEMORY NETWORKS	3469
<i>Youngjin Kim, Minjung Kim, Gunhee Kim</i>	
INITIALIZATION MATTERS: ORTHOGONAL PREDICTIVE STATE RECURRENT NEURAL NETWORKS.....	3484
<i>Krzysztof Choromanski, Carlton Downey, Byron Boots</i>	
REINFORCEMENT LEARNING ALGORITHM SELECTION	3501
<i>Romain Laroche, Raphael Feraud</i>	
TREEQN AND ATREEC: DIFFERENTIABLE TREE-STRUCTURED MODELS FOR DEEP REINFORCEMENT LEARNING	3531
<i>Gregory Farquhar, Tim Rocktäschel, Maximilian Igl, Shimon Whiteson</i>	
SYNTHESIZING REALISTIC NEURAL POPULATION ACTIVITY PATTERNS USING GENERATIVE ADVERSARIAL NETWORKS	3546
<i>Manuel Molano-Mazon, Arno Onken, Eugenio Piasini, Stefano Panzeri</i>	
DEEP COMPLEX NETWORKS	3570
<i>Chiheb Trabelsi, Olexa Bilaniuk, Ying Zhang, Dmitriy Serdyuk, Sandeep Subramanian, Joao Felipe Santos, Soroush Mehri, Negar Rostamzadeh, Yoshua Bengio, Christopher J Pal</i>	
COMPRESSING WORD EMBEDDINGS VIA DEEP COMPOSITIONAL CODE LEARNING	3589
<i>Raphael Shu, Hideki Nakayama</i>	
A DIRT-T APPROACH TO UNSUPERVISED DOMAIN ADAPTATION	3602
<i>Rui Shu, Hung Bui, Hirokazu Narui, Stefano Ermon</i>	

NEURAL-GUIDED DEDUCTIVE SEARCH FOR REAL-TIME PROGRAM SYNTHESIS FROM EXAMPLES.....	3621
<i>Ashwin Kalyan, Abhishek Mohta, Oleksandr Polozov, Dhruv Batra, Prateek Jain, Sumit Gulwani</i>	
UNSUPERVISED NEURAL MACHINE TRANSLATION.....	3636
<i>Mikel Artetxe, Gorka Labaka, Eneko Agirre, Kyunghyun Cho</i>	
VARIATIONAL CONTINUAL LEARNING	3648
<i>Cuong V. Nguyen, Yingzhen Li, Thang D. Bui, Richard E. Turner</i>	
STOCHASTIC ACTIVATION PRUNING FOR ROBUST ADVERSARIAL DEFENSE	3666
<i>Guneet S. Dhillon, Kamyar Azizzadenesheli, Zachary C. Lipton, Jeremy D. Bernstein, Jean Kossaifi, Aran Khanna, Animashree Anandkumar</i>	
MASTERING THE DUNGEON: GROUNDED LANGUAGE LEARNING BY MECHANICAL TURKER DESCENT	3679
<i>Zhilin Yang, Saizheng Zhang, Jack Urbanek, Will Feng, Alexander Miller, Arthur Szlam, Douwe Kiela, Jason Weston</i>	
COULOMB GANS: PROVABLY OPTIMAL NASH EQUILIBRIA VIA POTENTIAL FIELDS	3695
<i>Thomas Unterthiner, Bernhard Nessler, Calvin Seward, Günter Klambauer, Martin Heusel, Hubert Ramsauer, Sepp Hochreiter</i>	
RESIDUAL LOSS PREDICTION: REINFORCEMENT LEARNING WITH NO INCREMENTAL FEEDBACK.....	3716
<i>Hal Daumé III, John Langford, Amr Sharaf</i>	
THE ROLE OF MINIMAL COMPLEXITY FUNCTIONS IN UNSUPERVISED LEARNING OF SEMANTIC MAPPINGS	3739
<i>Tomer Galanti, Lior Wolf, Sagie Benaim</i>	
BOUNDARY SEEKING GANS.....	3807
<i>R Devon Hjelm, Athul Paul Jacob, Adam Trischler, Gerry Che, Kyunghyun Cho, Yoshua Bengio</i>	
TWIN NETWORKS: MATCHING THE FUTURE FOR SEQUENCE GENERATION.....	3824
<i>Dmitriy Serdyuk, Nan Rosemary Ke, Alessandro Sordoni, Adam Trischler, Chris Pal, Yoshua Bengio</i>	
LEARNING INTRINSIC SPARSE STRUCTURES WITHIN LONG SHORT-TERM MEMORY.....	3836
<i>Wei Wen, Yuxiong He, Samyam Rajbhandari, Minjia Zhang, Wenhan Wang, Fang Liu, Bin Hu, Yiran Chen, Hai Li</i>	
TD OR NOT TD: ANALYZING THE ROLE OF TEMPORAL DIFFERENCING IN DEEP REINFORCEMENT LEARNING	3850
<i>Artemij Amiranashvili, Alexey Dosovitskiy, Vladlen Koltun, Thomas Brox</i>	
DISTRIBUTED PRIORITIZED EXPERIENCE REPLAY	3864
<i>Dan Horgan, John Quan, David Budden, Gabriel Barth-Maron, Matteo Hessel, Hado Van Hasselt, David Silver</i>	
BI-DIRECTIONAL BLOCK SELF-ATTENTION FOR FAST AND MEMORY-EFFICIENT SEQUENCE MODELING	3883
<i>Tao Shen, Tianyi Zhou, Guodong Long, Jing Jiang, Chengqi Zhang</i>	
CRITICAL POINTS OF LINEAR NEURAL NETWORKS: ANALYTICAL FORMS AND LANDSCAPE PROPERTIES	3901
<i>Yi Zhou, Yingbin Liang</i>	
LEARNING A GENERATIVE MODEL FOR VALIDITY IN COMPLEX DISCRETE STRUCTURES	3920
<i>Dave Janz, Jos Van Der Westhuizen, Brooks Paige, Matt Kusner, José Miguel Hernández-Lobato</i>	

HIERARCHICAL DENSITY ORDER EMBEDDINGS	3932
<i>Ben Athiwaratkun, Andrew Gordon Wilson</i>	
A NEW METHOD OF REGION EMBEDDING FOR TEXT CLASSIFICATION	3947
<i>Chao Qiao, Bo Huang, Guocheng Niu, Daren Li, Daxiang Dong, Wei He, Dianhai Yu, Hua Wu</i>	
FIX YOUR CLASSIFIER: THE MARGINAL VALUE OF TRAINING THE LAST WEIGHT LAYER	3959
<i>Elad Hoffer, Itay Hubara, Daniel Soudry</i>	
ACTIVATION MAXIMIZATION GENERATIVE ADVERSARIAL NETS	3970
<i>Zhiming Zhou, Han Cai, Shu Rong, Yuxuan Song, Kan Ren, Weinan Zhang, Jun Wang, Yong Yu</i>	
SYNTAX-DIRECTED VARIATIONAL AUTOENCODER FOR STRUCTURED DATA.....	3994
<i>Hanjun Dai, Yingtao Tian, Bo Dai, Steven Skiena, Le Song</i>	
MANY PATHS TO EQUILIBRIUM: GANS DO NOT NEED TO DECREASE A DIVERGENCE AT EVERY STEP	4011
<i>William Fedus, Mihaela Rosca, Balaji Lakshminarayanan, Andrew M. Dai, Shakir Mohamed, Ian Goodfellow</i>	

VOLUME 6

EMPIRICAL RISK LANDSCAPE ANALYSIS FOR UNDERSTANDING DEEP NEURAL NETWORKS.....	4032
<i>Pan Zhou, Jiashi Feng</i>	
ON THE REGULARIZATION OF WASSERSTEIN GANS	4092
<i>Henning Petzka, Asja Fischer, Denis Lukovnikov</i>	
MAXIMUM A POSTERIORI POLICY OPTIMISATION.....	4116
<i>Abbas Abdolmaleki, Jost Tobias Springenberg, Yuval Tassa, Remi Munos, Nicolas Heess, Martin Riedmiller</i>	
TOWARDS DEEP LEARNING MODELS RESISTANT TO ADVERSARIAL ATTACKS	4138
<i>Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, Adrian Vladu</i>	
LEARNING TO TEACH	4161
<i>Yang Fan, Fei Tian, Tao Qin, Xiang-Yang Li, Tie-Yan Liu</i>	
THE POWER OF DEEPER NETWORKS FOR EXPRESSING NATURAL FUNCTIONS.....	4177
<i>David Rolnick, Max Tegmark</i>	
ONLINE LEARNING RATE ADAPTATION WITH HYPERGRADIENT DESCENT	4191
<i>Atilim Gunes Baydin, Robert Cornish, David Martinez Rubio, Mark Schmidt, Frank Wood</i>	
BOOSTING THE ACTOR WITH DUAL CRITIC.....	4202
<i>Bo Dai, Albert Shaw, Niao He, Lihong Li, Le Song</i>	
SCAN: LEARNING HIERARCHICAL COMPOSITIONAL VISUAL CONCEPTS.....	4222
<i>Irina Higgins, Nicolas Sonnerat, Loic Matthey, Arka Pal, Christopher P Burgess, Matko Bošnjak, Murray Shanahan, Matthew Botvinick, Demis Hassabis, Alexander Lerchner</i>	
TEMPORAL DIFFERENCE MODELS: MODEL-FREE DEEP RL FOR MODEL-BASED CONTROL	4246
<i>Vitchyr Pong, Shixiang Gu, Murtaza Dalal, Sergey Levine</i>	
DECOUPLING THE LAYERS IN RESIDUAL NETWORKS	4260
<i>Ricky Fok, Aijun An, Zana Rashidi, Xiaogang Wang</i>	
CGANS WITH PROJECTION DISCRIMINATOR	4271
<i>Takeru Miyato, Masanori Koyama</i>	

IMPROVING THE UNIVERSALITY AND LEARNABILITY OF NEURAL PROGRAMMER-INTERPRETERS WITH COMBINATOR ABSTRACTION	4294
<i>Da Xiao, Jo-Yu Liao, Xingyuan Yuan</i>	
COMMUNICATION ALGORITHMS VIA DEEP LEARNING	4310
<i>Hyeji Kim, Yihan Jiang, Ranvir B. Rana, Sreeram Kannan, Sewoong Oh, Pramod Viswanath</i>	
META LEARNING SHARED HIERARCHIES.....	4327
<i>Kevin Frans, Jonathan Ho, Xi Chen, Pieter Abbeel, John Schulman</i>	
DEEP BAYESIAN BANDITS SHOWDOWN: AN EMPIRICAL COMPARISON OF BAYESIAN DEEP NETWORKS FOR THOMPSON SAMPLING	4338
<i>Carlos Riquelme, George Tucker, Jasper Snoek</i>	
COMPOSITIONAL ATTENTION NETWORKS FOR MACHINE REASONING.....	4365
<i>Drew A. Hudson, Christopher D. Manning</i>	
DEFENSE-GAN: PROTECTING CLASSIFIERS AGAINST ADVERSARIAL ATTACKS USING GENERATIVE MODELS.....	4385
<i>Pouya Samangouei, Maya Kabkab, Rama Chellappa</i>	
DEEP LEARNING AND QUANTUM ENTANGLEMENT: FUNDAMENTAL CONNECTIONS WITH IMPLICATIONS TO NETWORK DESIGN	4402
<i>Yoav Levine, David Yakira, Nadav Cohen, Amnon Shashua</i>	
LEARNING SPARSE NEURAL NETWORKS THROUGH L ₀ REGULARIZATION	4431
<i>Christos Louizos, Max Welling, Diederik P. Kingma</i>	
GENERALIZING HAMILTONIAN MONTE CARLO WITH NEURAL NETWORKS	4444
<i>Daniel Levy, Matt D. Hoffman, Jascha Sohl-Dickstein</i>	
META-LEARNING AND UNIVERSALITY: DEEP REPRESENTATIONS AND GRADIENT DESCENT CAN APPROXIMATE ANY LEARNING ALGORITHM	4459
<i>Chelsea Finn, Sergey Levine</i>	
UNBIASED ONLINE RECURRENT OPTIMIZATION	4479
<i>Corentin Tallec, Yann Ollivier</i>	
CONSEQUENTIALIST CONDITIONAL COOPERATION IN SOCIAL DILEMMAS WITH IMPERFECT INFORMATION.....	4491
<i>Alexander Peysakhovich, Adam Lerer</i>	
SKIP RNN: LEARNING TO SKIP STATE UPDATES IN RECURRENT NEURAL NETWORKS	4507
<i>Victor Campos, Brendan Jou, Xavier Giró-I-Nieto, Jordi Torres, Shih-Fu Chang</i>	
GENERATING WIKIPEDIA BY SUMMARIZING LONG SEQUENCES	4524
<i>Peter J. Liu, Mohammad Saleh, Etienne Pot, Ben Goodrich, Ryan Sepassi, Lukasz Kaiser, Noam Shazeer</i>	
FRATERNAL DROPOUT	4542
<i>Konrad Zolna, Devansh Arpit, Dendi Suhubdy, Yoshua Bengio</i>	
LEARNING DISCRETE WEIGHTS USING THE LOCAL REPARAMETERIZATION TRICK.....	4556
<i>Oran Shayer, Dan Levi, Ethan Fetaya</i>	
FEW-SHOT AUTOREGRESSIVE DENSITY ESTIMATION: TOWARDS LEARNING TO LEARN DISTRIBUTIONS.....	4568
<i>Scott Reed, Yutian Chen, Thomas Paine, Aäron Van Den Oord, S. M. Ali Eslami, Danilo Rezende, Oriol Vinyals, Nando De Freitas</i>	
EMERGENT COMPLEXITY VIA MULTI-AGENT COMPETITION	4579
<i>Trapit Bansal, Jakub Pachocki, Szymon Sidor, Ilya Sutskever, Igor Mordatch</i>	
MEMORY ARCHITECTURES IN RECURRENT NEURAL NETWORK LANGUAGE MODELS	4591
<i>Dani Yogatama, Yishu Miao, Gabor Melis, Wang Ling, Adhiguna Kuncoro, Chris Dyer, Phil Blunsom</i>	

ADVERSARIAL DROPOUT REGULARIZATION	4601
<i>Kuniaki Saito, Yoshitaka Ushiku, Tatsuya Harada, Kate Saenko</i>	
NEURAL MAP: STRUCTURED MEMORY FOR DEEP REINFORCEMENT LEARNING.....	4616
<i>Emilio Parisotto, Ruslan Salakhutdinov</i>	
MASKGAN: BETTER TEXT GENERATION VIA FILLING IN THE _____	4636
<i>William Fedus, Ian Goodfellow, Andrew M. Dai</i>	
ON THE DISCRIMINATION-GENERALIZATION TRADEOFF IN GANS	4653
<i>Pengchuan Zhang, Qiang Liu, Dengyong Zhou, Tao Xu, Xiaodong He</i>	
INTERACTIVE GROUNDED LANGUAGE ACQUISITION AND GENERALIZATION IN A 2D WORLD	4679
<i>Haonan Yu, Haichao Zhang, Wei Xu</i>	
GO FOR A WALK AND ARRIVE AT THE ANSWER: REASONING OVER PATHS IN KNOWLEDGE BASES USING REINFORCEMENT LEARNING.....	4708
<i>Rajarshi Das, Shehzaad Dhuliawala, Manzil Zaheer, Luke Vilnis, Ishan Durugkar, Akshay Krishnamurthy, Alex Smola, Andrew McCallum</i>	
TOWARDS SYNTHESIZING COMPLEX PROGRAMS FROM INPUT-OUTPUT EXAMPLES.....	4726
<i>Xinyun Chen, Chang Liu, Dawn Song</i>	
LOSS-AWARE WEIGHT QUANTIZATION OF DEEP NETWORKS	4757
<i>Lu Hou, James T. Kwok</i>	
EFFICIENT SPARSE-WINOGRAD CONVOLUTIONAL NEURAL NETWORKS.....	4773
<i>Xingyu Liu, Jeff Pool, Song Han, William J. Dally</i>	
DEMYSTIFYING MMD GANS	4783
<i>Mikolaj Binkowski, Danica J. Sutherland, Michael Arbel, Arthur Gretton</i>	
SENSITIVITY AND GENERALIZATION IN NEURAL NETWORKS: AN EMPIRICAL STUDY	4819
<i>Roman Novak, Yasaman Bahri, Daniel A. Abolafia, Jeffrey Pennington, Jascha Sohl-Dickstein</i>	
QANET: COMBINING LOCAL CONVOLUTION WITH GLOBAL SELF-ATTENTION FOR READING COMPREHENSION	4839
<i>Adams Wei Yu, David Dohan, Minh-Thang Luong, Rui Zhao, Kai Chen, Mohammad Norouzi, Quoc V. Le</i>	

VOLUME 7

DISTRIBUTED DISTRIBUTIONAL DETERMINISTIC POLICY GRADIENTS.....	4855
<i>Gabriel Barth-Maron, Matthew W. Hoffman, David Budden, Will Dabney, Dan Horgan, Dhruva Tb, Alistair Muldal, Nicolas Heess, Timothy Lillicrap</i>	
EIGENOPTION DISCOVERY THROUGH THE DEEP SUCCESSOR REPRESENTATION.....	4871
<i>Marlos C. Machado, Clemens Rosenbaum, Xiaoxiao Guo, Miao Liu, Gerald Tesauro, Murray Campbell</i>	
THE IMPLICIT BIAS OF GRADIENT DESCENT ON SEPARABLE DATA.....	4893
<i>Daniel Soudry, Elad Hoffer, Mor Shpigel Nacson, Nathan Srebro</i>	
COUNTERING ADVERSARIAL IMAGES USING INPUT TRANSFORMATIONS	4914
<i>Chuan Guo, Mayank Rana, Moustapha Cisse, Laurens Van Der Maaten</i>	
LEARNING TO CLUSTER IN ORDER TO TRANSFER ACROSS DOMAINS AND TASKS.....	4926
<i>Yen-Chang Hsu, Zhaoyang Lv, Zsolt Kira</i>	
CERTIFIED DEFENSES AGAINST ADVERSARIAL EXAMPLES	4946
<i>Aditi Raghunathan, Jacob Steinhardt, Percy Liang</i>	
VARIATIONAL IMAGE COMPRESSION WITH A SCALE HYPERPRIOR.....	4961
<i>Johannes Ballé, David Minnen, Saurabh Singh, Sung Jin Hwang, Nick Johnston</i>	

HIERARCHICAL REPRESENTATIONS FOR EFFICIENT ARCHITECTURE SEARCH	5008
<i>Hanxiao Liu, Karen Simonyan, Oriol Vinyals, Chrisantha Fernando, Koray Kavukcuoglu</i>	
TRAINING WIDE RESIDUAL NETWORKS FOR DEPLOYMENT USING A SINGLE BIT FOR EACH WEIGHT	5021
<i>Mark D. McDonnell</i>	
LEARNING SPARSE LATENT REPRESENTATIONS WITH THE DEEP COPULA INFORMATION BOTTLENECK	5036
<i>Aleksander Wiecek, Mario Wieser, Damian Murezzan, Volker Roth</i>	
ON THE IMPORTANCE OF SINGLE DIRECTIONS FOR GENERALIZATION.....	5051
<i>Ari S. Morcos, David G. T. Barrett, Neil C. Rabinowitz, Matthew Botvinick</i>	
META-LEARNING FOR SEMI-SUPERVISED FEW-SHOT CLASSIFICATION.....	5066
<i>Mengye Ren, Eleni Triantafyllou, Sachin Ravi, Jake Snell, Kevin Swersky, Joshua B. Tenenbaum, Hugo Larochelle, Richard S. Zemel</i>	
HYPERPARAMETER OPTIMIZATION: A SPECTRAL APPROACH.....	5081
<i>Elad Hazan, Adam Klivans, Yang Yuan</i>	
GENERALIZING ACROSS DOMAINS VIA CROSS-GRADIENT TRAINING.....	5099
<i>Shiv Shankar, Vihari Piratla, Soumen Chakrabarti, Siddhartha Chaudhuri, Preethi Jyothi, Sunita Sarawagi</i>	
ON THE STATE OF THE ART OF EVALUATION IN NEURAL LANGUAGE MODELS	5111
<i>Gábor Melis, Chris Dyer, Phil Blunsom</i>	
LEARNING TO SHARE: SIMULTANEOUS PARAMETER TYING AND SPARSIFICATION IN DEEP LEARNING.....	5121
<i>Dejiao Zhang, Haozhu Wang, Mario Figueiredo, Laura Balzano</i>	
EMERGENT TRANSLATION IN MULTI-AGENT COMMUNICATION.....	5135
<i>Jason Lee, Kyunghyun Cho, Jason Weston, Douwe Kiela</i>	
DEEP ACTIVE LEARNING FOR NAMED ENTITY RECOGNITION.....	5153
<i>Yanyao Shen, Hyokun Yun, Zachary C. Lipton, Yakov Kronrod, Animashree Anandkumar</i>	
NON-AUTOREGRESSIVE NEURAL MACHINE TRANSLATION	5168
<i>Jiatao Gu, James Bradbury, Caiming Xiong, Victor O. K. Li, Richard Socher</i>	
TOWARDS NEURAL PHRASE-BASED MACHINE TRANSLATION	5181
<i>Po-Sen Huang, Chong Wang, Sitao Huang, Dengyong Zhou, Li Deng</i>	
FUSIONNET: FUSING VIA FULLY-AWARE ATTENTION WITH APPLICATION TO MACHINE COMPREHENSION.....	5195
<i>Hsin-Yuan Huang, Chenguang Zhu, Yelong Shen, Weizhu Chen</i>	
UNSUPERVISED MACHINE TRANSLATION USING MONOLINGUAL CORPORA ONLY.....	5215
<i>Guillaume Lample, Alexis Conneau, Ludovic Denoyer, Marc'Aurelio Ranzato</i>	
AUTOMATICALLY INFERRING DATA QUALITY FOR SPATIOTEMPORAL FORECASTING	5229
<i>Sungyong Seo, Arash Mohegh, George Ban-Weiss, Yan Liu</i>	
COMBINING SYMBOLIC EXPRESSIONS AND BLACK-BOX FUNCTION EVALUATIONS IN NEURAL PROGRAMS.....	5240
<i>Forough Arabshahi, Sameer Singh, Animashree Anandkumar</i>	
PROXIMAL BACKPROPAGATION.....	5252
<i>Thomas Frerix, Thomas Möllenhoff, Michael Moeller, Daniel Cremers</i>	
DCN+: MIXED OBJECTIVE AND DEEP RESIDUAL COATTENTION FOR QUESTION ANSWERING.....	5266
<i>Caiming Xiong, Victor Zhong, Richard Socher</i>	

IMITATION LEARNING FROM VISUAL DATA WITH MULTIPLE INTENTIONS.....	5276
<i>Aviv Tamar, Khashayar Rohanimanesh, Yinlam Chow, Chris Vigorito, Ben Goodrich, Michael Kahane, Derik Pridmore</i>	
RETHINKING THE SMALLER-NORM-LESS-INFORMATIVE ASSUMPTION IN CHANNEL PRUNING OF CONVOLUTION LAYERS	5293
<i>Jianbo Ye, Xin Lu, Zhe Lin, James Z. Wang</i>	
AN ONLINE LEARNING APPROACH TO GENERATIVE ADVERSARIAL NETWORKS.....	5304
<i>Paulina Grnarova, Kfir Y Levy, Aurelien Lucchi, Thomas Hofmann, Andreas Krause</i>	
SCALABLE PRIVATE LEARNING WITH PATE.....	5328
<i>Nicolas Papernot, Shuang Song, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, Ulfar Erlingsson</i>	
SEMI-PARAMETRIC TOPOLOGICAL MEMORY FOR NAVIGATION	5362
<i>Nikolay Savinov, Alexey Dosovitskiy, Vladlen Koltun</i>	
MULTI-MENTION LEARNING FOR READING COMPREHENSION WITH NEURAL CASCADES	5378
<i>Swabha Swayamdipta, Ankur P. Parikh, Tom Kwiatkowski</i>	
DYNAMIC NEURAL PROGRAM EMBEDDINGS FOR PROGRAM REPAIR	5390
<i>Ke Wang, Rishabh Singh, Zhendong Su</i>	
MINIMAL-ENTROPY CORRELATION ALIGNMENT FOR UNSUPERVISED DEEP DOMAIN ADAPTATION	5402
<i>Pietro Moreerio, Jacopo Cavazza, Vittorio Murino</i>	
LARGE SCALE DISTRIBUTED NEURAL NETWORK TRAINING THROUGH ONLINE DISTILLATION	5417
<i>Rohan Anil, Gabriel Pereyra, Alexandre Passos, Robert Ormandi, George E. Dahl, Geoffrey E. Hinton</i>	
VARIATIONAL NETWORK QUANTIZATION	5429
<i>Jan Achterhold, Jan Mathias Koehler, Anke Schmeink, Tim Genewein</i>	
LEARNING AN EMBEDDING SPACE FOR TRANSFERABLE ROBOT SKILLS	5447
<i>Karol Hausman, Jost Tobias Springenberg, Ziyu Wang, Nicolas Heess, Martin Riedmiller</i>	
A SCALABLE LAPLACE APPROXIMATION FOR NEURAL NETWORKS	5463
<i>Hippolyt Ritter, Aleksandar Botev, David Barber</i>	
NERVENET: LEARNING STRUCTURED POLICY WITH GRAPH NEURAL NETWORKS	5478
<i>Tingwu Wang, Renjie Liao, Jimmy Ba, Sanja Fidler</i>	
WAVELET POOLING FOR CONVOLUTIONAL NEURAL NETWORKS.....	5504
<i>Travis Williams, Robert Li</i>	
GENERATIVE MODELS OF VISUALLY GROUNDED IMAGINATION.....	5516
<i>Ramakrishna Vedantam, Ian Fischer, Jonathan Huang, Kevin Murphy</i>	
COMPOSITIONAL OBVERTER COMMUNICATION LEARNING FROM RAW VISUAL INPUT	5538
<i>Edward Choi, Angeliki Lazaridou, Nando De Freitas</i>	
KRONECKER-FACTORED CURVATURE APPROXIMATIONS FOR RECURRENT NEURAL NETWORKS.....	5556
<i>James Martens, Jimmy Ba, Matt Johnson</i>	
SEARNN: TRAINING RNNs WITH GLOBAL-LOCAL LOSSES	5581
<i>Rémi Leblond, Jean-Baptiste Alayrac, Anton Osokin, Simon Lacoste-Julien</i>	
MULTI-VIEW DATA GENERATION WITHOUT VIEW SUPERVISION.....	5597
<i>Mickael Chen, Ludovic Denoyer, Thierry Artières</i>	

SMOOTH LOSS FUNCTIONS FOR DEEP TOP-K CLASSIFICATION	5620
<i>Leonard Berrada, Andrew Zisserman, M. Pawan Kumar</i>	
LEARNING ROBUST REWARDS WITH ADVERSERIAL INVERSE REINFORCEMENT LEARNING.....	5645
<i>Justin Fu, Katie Luo, Sergey Levine</i>	
A BAYESIAN PERSPECTIVE ON GENERALIZATION AND STOCHASTIC GRADIENT DESCENT	5660
<i>Samuel L. Smith, Quoc V. Le</i>	

Author Index