



Department of
Electrical Engineering

香港城市大學
City University of Hong Kong



Hong Kong Section 香港分會
CAS/COM Chapter

For Posting

CITY UNIVERSITY OF HONG KONG

DEPARTMENT OF ELECTRICAL ENGINEERING & IEEE HK SECTION CAS/COM
CHAPTER

Presents a seminar on

Coherent Gradients: An Approach to Understanding Generalization in Gradient Descent-Based Optimization

Dr. Sat Chatterjee
Senior Staff Engineering Manager
Google AI
Mountain View, CA, USA

Date : 18 Oct 2019 (Friday)

Time : 10:00 am - 11:00 am

Venue : G6302, 6/F, Green Zone, Yeung Kin Man
Academic Building, CityU

Abstract

An open question in the Deep Learning community is why neural networks trained with Gradient Descent generalize well on real datasets even though they are capable of fitting random data. We propose an approach to answering this question based on a hypothesis about the dynamics of gradient descent that we call Coherent Gradients: Gradients from similar examples are similar and so the overall gradient is stronger in certain directions where these reinforce each other. Thus changes to the network parameters during training are biased towards those that (locally) simultaneously benefit many examples when such similarity exists. We support this hypothesis with heuristic arguments and perturbative experiments and outline how this can explain several common empirical observations about Deep Learning. Furthermore, our analysis is not just descriptive, but prescriptive. It suggests a natural modification to gradient descent that can greatly reduce overfitting.

Biography

Sat Chatterjee is a Senior Staff Engineering Manager at Google AI in Mountain View, California. In addition to basic research in Machine Learning, he is interested in various applications of Machine Learning ranging from Electronic Design Automation to Algorithmic Trading. Previously, he was a Senior Vice President at Two Sigma Investments in New York City where most recently he founded the Deep Learning/Core AI organization and previously headed Engineering for the Market Making and Index Arbitrage business. Prior to that he worked in Formal Verification of Microarchitecture at Intel's Strategic CAD Labs in Hillsboro, Oregon. He has a PhD in Computer Science from the University of California at Berkeley, where his focus was on efficient algorithms for logic synthesis and formal verification.

~~~ All are welcome ~~~