

# Human Computation

# Synthesis Lectures on Artificial Intelligence and Machine Learning

## Editors

**Ronald J. Brachman**, *Yahoo Research*

**William W. Cohen**, *Carnegie Mellon University*

**Thomas Dietterich**, *Oregon State University*

## Human Computation

Edith Law and Luis von Ahn

2011

## Trading Agents

Michael P. Wellman

2011

## Visual Object Recognition

Kristen Grauman and Bastian Leibe

2011

## Learning with Support Vector Machines

Colin Campbell and Yiming Ying

2011

## Algorithms for Reinforcement Learning

Csaba Szepesvári

2010

## Data Integration: The Relational Logic Approach

Michael Genesereth

2010

## Markov Logic: An Interface Layer for Artificial Intelligence

Pedro Domingos and Daniel Lowd

2009

### Introduction to Semi-Supervised Learning

Xiaojin Zhu and Andrew B. Goldberg

2009

### Action Programming Languages

Michael Thielscher

2008

### Representation Discovery using Harmonic Analysis

Sridhar Mahadevan

2008

### Essentials of Game Theory: A Concise Multidisciplinary Introduction

Kevin Leyton-Brown and Yoav Shoham

2008

### A Concise Introduction to Multiagent Systems and Distributed Artificial Intelligence

Nikos Vlassis

2007

### Intelligent Autonomous Robotics: A Robot Soccer Case Study

Peter Stone

2007

© Springer Nature Switzerland AG 2022

Reprint of original edition © Morgan & Claypool 2011

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopy, recording, or any other except for brief quotations in printed reviews, without the prior permission of the publisher.

Human Computation

Edith Law and Luis von Ahn

ISBN: 978-3-031-00427-8      paperback

ISBN: 978-3-031-01555-7      ebook

DOI 10.1007/978-3-031-01555-7

A Publication in the Springer series

*SYNTHESIS LECTURES ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING*

Lecture #13

Series Editors: Ronald J. Brachman, *Yahoo Research*

William W. Cohen, *Carnegie Mellon University*

Thomas Dietterich, *Oregon State University*

Series ISSN

Synthesis Lectures on Artificial Intelligence and Machine Learning

Print 1939-4608    Electronic 1939-4616

# Human Computation

Edith Law and Luis von Ahn  
Carnegie Mellon University

*SYNTHESIS LECTURES ON ARTIFICIAL INTELLIGENCE AND MACHINE  
LEARNING #13*

## ABSTRACT

Human computation is a new and evolving research area that centers around harnessing human intelligence to solve computational problems that are beyond the scope of existing Artificial Intelligence (AI) algorithms. With the growth of the Web, human computation systems can now leverage the abilities of an unprecedented number of people via the Web to perform complex computation. There are various genres of human computation applications that exist today. Games with a purpose (e.g., the ESP Game) specifically target online gamers who generate useful data (e.g., image tags) while playing an enjoyable game. Crowdsourcing marketplaces (e.g., Amazon Mechanical Turk) are human computation systems that coordinate workers to perform tasks in exchange for monetary rewards. In identity verification tasks, users perform computation in order to gain access to some online content; an example is reCAPTCHA, which leverages millions of users who solve CAPTCHAs every day to correct words in books that optical character recognition (OCR) programs fail to recognize with certainty.

This book is aimed at achieving four goals: (1) defining human computation as a research area; (2) providing a comprehensive review of existing work; (3) drawing connections to a wide variety of disciplines, including AI, Machine Learning, HCI, Mechanism/Market Design and Psychology, and capturing their unique perspectives on the core research questions in human computation; and (4) suggesting promising research directions for the future.

## KEYWORDS

human computation, human-in-the-loop algorithms, output aggregation, active learning, latent class models, task routing, labor markets, games with a purpose, task design, crowdsourcing, mechanism design, incentives

*To Eric, Jacob and my parents.*

*– Edith*

*To Laura.*

*– Luis*

# Contents

	<b>Preface</b> .....	<b>xiii</b>
	<b>Acknowledgments</b> .....	<b>xv</b>
<b>1</b>	<b>Introduction</b> .....	<b>1</b>
1.1	Computation: Now and Then .....	1
1.2	What is Human Computation? .....	3
1.2.1	Explicit Control .....	4
1.3	Tackling AI Problems: From Vision to Biology .....	8
1.3.1	A Matter of Perception .....	8
1.3.2	The Language Barrier .....	9
1.3.3	Intuition into Computationally Intensive Problems .....	10
1.4	Overview .....	11
	 <b>PART I Solving Computational Problems</b> .....	 <b>13</b>
<b>2</b>	<b>Human Computation Algorithms</b> .....	<b>15</b>
2.1	A Definition of Algorithms .....	16
2.2	Building Blocks of Algorithms .....	17
2.2.1	Operations, Controls and Program Synthesis .....	17
2.3	Programming Frameworks .....	19
2.4	Evaluating Human Computation Algorithms .....	21
2.4.1	Correctness .....	21
2.4.2	Efficiency .....	21
2.5	Summary .....	23
<b>3</b>	<b>Aggregating Outputs</b> .....	<b>25</b>
3.1	Objective versus Cultural Truth .....	26
3.2	Classification .....	26
3.2.1	Latent Class Models .....	26



	3.2.2 Learning from Imperfect Data .....	29
3.3	Beyond Classification .....	30
	3.3.1 Ranking and Voting .....	31
	3.3.2 Clustering .....	31
	3.3.3 Structured Outputs .....	32
	3.3.4 Beliefs .....	32
3.4	Summary .....	32
<b>4</b>	<b>Task Routing .....</b>	<b>35</b>
4.1	Push versus Pull Approaches .....	36
4.2	Push Approach .....	37
	4.2.1 Allocation .....	37
	4.2.2 Matching .....	38
	4.2.3 Inference .....	38
4.3	Pull Approach .....	40
	4.3.1 Search and Visualization .....	40
	4.3.2 Task Recommendation .....	40
	4.3.3 Peer Routing .....	41
4.4	Evaluation Criteria .....	41
4.5	Summary .....	42
	 <b>PART II Design .....</b>	 <b>43</b>
<b>5</b>	<b>Understanding Workers and Requesters .....</b>	<b>45</b>
5.1	Markets .....	46
	5.1.1 Mechanical Turk and Paid CrowdSourcing .....	46
	5.1.2 Security and Access .....	48
	5.1.3 Gamers .....	49
	5.1.4 Citizen Science .....	51
	5.1.5 Learners .....	53
	5.1.6 Temporary Markets .....	54
5.2	Supporting End Users .....	54
	5.2.1 Workers .....	54
	5.2.2 Requesters .....	55
5.3	Summary .....	56

<b>6</b>	<b>The Art of Asking Questions</b>	<b>57</b>
6.1	Designing Tasks	57
6.1.1	Information	58
6.1.2	Granularity	60
6.1.3	Independence	60
6.1.4	Incentives	61
6.1.5	Quality Control	62
6.2	Eliciting Truthful Responses	62
6.2.1	Human Computation Games	63
6.2.2	Leveraging Communication	64
6.2.3	Explicitly Preventing Bad Outputs	67
6.2.4	A Brief Survey of Games and Mechanisms	69
6.3	Summary	70
	 <b>PART III Conclusion</b>	 <b>71</b>
<b>7</b>	<b>The Future of Human Computation</b>	<b>73</b>
7.1	Research Directions	73
7.1.1	Interweaving Human and Machine Intelligence	73
7.1.2	Fostering Long-Term Relationships	74
7.1.3	Designing Organizations and Task Markets	74
7.2	Conclusion	75
	 <b>Bibliography</b>	 <b>77</b>
	 <b>Authors' Biographies</b>	 <b>105</b>

# Preface

Human computation is a new and evolving research area that centers around harnessing human intelligence to solve computational problems (e.g., image classification, language translation, protein folding) that are beyond the scope of existing Artificial Intelligence (AI) algorithms. With the growth of the Web, human computation systems can now leverage the abilities of an unprecedented number of people via the Web to perform complex computation. There are various genres of human computation applications that exist today. Games with a purpose (e.g., the ESP Game) specifically target online gamers who generate useful data (e.g., image tags) while playing an enjoyable game. Crowdsourcing marketplaces (e.g., Amazon Mechanical Turk) are human computation systems that coordinate workers to perform tasks in exchange for monetary rewards. In identity verification tasks, users perform computation in order to gain access to some online content; an example is reCAPTCHA, which leverages millions of users who solve CAPTCHAs every day to correct words in books that optical character recognition (OCR) programs fail to recognize with certainty.

Despite the variety of human computation applications, there exist many common core research issues. How can we design mechanisms for querying human computers such that they are incentivized to generate truthful outputs? What are some techniques for aggregating noisy or complex outputs from multiple human computers in the absence of ground truth? How do we effectively assign tasks to human computers in order to satisfy the objectives of both the system (e.g., quality, budget and time constraints) and the workers (e.g., desire to succeed, to learn, to be entertained)? What classes of computational problems can be efficiently answered using human computation? What are some programming paradigms for designing human computation algorithms? How can human computation systems leverage the joint efforts of both machines and humans?

This book is aimed at achieving four goals: (1) defining human computation as a research area; (2) providing a comprehensive review of existing work; (3) drawing connections to a wide variety of disciplines, including AI, Machine Learning, HCI, Mechanism/Market Design and Psychology, and capturing their unique perspectives on the core research questions in human computation; and (4) suggesting promising research directions in the field. Additional supplementary materials for this book (including the slide for the AAAI 2011 tutorial on “Human Computation: Core Research Questions and State of the Art”) can be found at <http://humancomputation.com/book>. We hope that this book will be a useful resource in the years to come for both newcomers and seasoned researchers who are interested in human computation, or more generally, the study of computational systems with humans in the loop.

Edith Law and Luis von Ahn  
July 2011

# Acknowledgments

As a relatively young field, the idea of human computation is not yet well defined. This book is in part an accumulation of ideas from people in the field with whom we had discussions and debates about the definitions, scope and future directions of human computation. These people include Paul Bennett (Microsoft Research), Eric Blais (Carnegie Mellon University), Yiling Chen (Harvard), David A. Grier (George Washington University), Eric Horvitz (Microsoft Research), Panos Ipeirotis (New York University), Winter Mason (Yahoo! Research), Rob Miller (MIT), Tom Mitchell (Carnegie Mellon University), David Parkes (Harvard University), Paul Resnick (University of Michigan), and Haoqi Zhang (Harvard University). Last, but not least, we thank our families for their unfailing support throughout this time-consuming, but invaluable, project.

The writing of this book is supported by the National Science Foundation Social-Computational Systems (SoCS) program under Grant No. IIIS-0968487. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Edith Law and Luis von Ahn  
July 2011