Shubham **Agarwal**

Software Engineer · AI Dialogue Systems · AI Planning

38 Linnaean St., Apt. 1, Cambridge, MA, USA - 02138

(+1) 480-359-8381 | agarwalshubham2007@gmail.com | agarwalshubham2007.github.io | agarwalshubham2007 | agarwalshubham2007

Work Experience

IBM Research, MIT-IBM Watson AI Lab

Cambridge, MA, U.S.A

Research Software Engineer

Aug. 2017 - PRESENT

- **Multi-Agent Orchestration in a Conversational System for Digital Business Automation**
  - *Built a domain independent conversation system that automates a multi-step business process, such as Travel Pre-Approval.*
  - *The system orchestrates multiple business process agents potentially responding to a given user query. It also handles digression and disambiguation between user intention gracefully.*
  - *Responsibility: System Design, Development and Deployment of primary Orchestration modules - scoring, selecting and sequencing a set of agents*
  - *Led client facing implementation and deployment as rapid sprints and identified required patches*
  - *Shipped the system as a Python Package deployed on RedHat OpenShift cluster for our clients in Banking Industry, end of 2020.*
  - *Lead engineer to ship the core algorithm of orchestration to the product team of Watson Assistant in a beta release*
  - *Lead engineer of 2021 theme on Automatic Agent Composition of available agents and skills using AI Planning*
- **TraceHub : platform to generate automated time-series analytics using AI Planning and abstract data modeling**
  - *Awarded best technical demo at AAAI 2020, NYC*
  - *Led a team of 7, conducted scrums and sprints in an agile environment with multiple first-author publications and presentations in AAAI 2020, O'Reilly AI Conference*
  - *The platform generates data transformer pipelines using a planner*
  - *Increased - analytics search and generation, time efficiency by implementing a data modeling layer that abstracts away computation on underlying data*
  - *TraceHub presents a plug-n-play environment for SciPy and custom cloud hosted transformers and analytics*
- **MAi : Model Acquisition Interface for Declarative Dialogue Design of Goal Oriented Multi-Turn Conversation**
  - *Declarative specification of dialogue agent's behavior*
  - *Compile the complete implicit graphs from this compact specification using a non-deterministic AI planner*
  - *Frontend in Bootstrap and backend as a Flask server*
  - *Deployed overall service infrastructure using Docker containers*
  - *Hosted the containerized platform on Kubernetes clusters enabling high-availability and auto-scaling. CI/CD using Travis*
  - *Publication in AAAI Proceedings 2019*
- **HOVOR : Goal Oriented Dialogue Agent Executor for open domain models**
  - *Runtime execution of a deployed conversation agent's contingent plan*
  - *Action execution in the declarative environment based on its precondition & outcome determination best describing the real world change*
  - *A general purpose executor for open domain models including dialogue systems*
  - *AAAI Symposium, Intex, DEEP-DIAL 2019 publications*
- **Giving back to the community**
  - *Technical mentor for multiple interns*
  - *Delivered technical tutorials on Container Technology, Kubernetes and Dialogue Systems*
  - *Led a team of 6 and conducted technical interviews in a major recruiting event of MIT students*
  - *Lead and coordinator of monthly social and team building activities*

Arizona State University

Arizona, U.S.A

Instructor

Fall 2016

- Taught CSE 205 - Object Oriented Programming and Data Structures to 114 students
- Responsibilities included teaching course material through lectures, setting assignments, tests and term examinations for the students

EMC Corporation

Bangalore, India

Software Engineer Intern

July 2014 - Dec. 2014

- Developed RESTful web service for data copy management and security testing in AppSync product of EMC
- Increased developer's efficiency and decreased cognitive load through in-house web service usage
- Developed backend in Java and frontend in Flex. Code successfully merged into production

Honors & Awards

2020	**Best Technical Demo,** AAAI Conference	*New York, USA*
2019	**Reviewer,** AAAI Conference Demo Track	
2018	**Certificate of Appreciation,** First Patent Filed, IBM	*Cambridge, U.S.A*
2016	**Winner,** Annual Code Challenge, ASU	*Arizona, U.S.A*
2012	**Runner-Up,** IBM Technical Contest	*Pilani, India*

Skills

Deep Learning	Certificate of Completion - Coursera Deep Learning 5 course specialization, by Andrew Ng.
Cloud Services	Cloud Functions, IBM Cloud, Watson Assistant, Dialog Flow, Watson NLU
Containers	Docker, Kubernetes
Back-end	Flask, Flask-restx, Travis CI/CD, SQL and NoSQL Databases
Front-end	Django, Hugo, HTML5, CSS, JavaScript(jQuery, Ajax)
Programming	Python, JAVA, JavaScript, LaTeX, PDDL, C, C++

Education

Arizona State University(ASU)

Arizona, USA

Master of Science, Computer Science

Aug. 2015 - Aug. 2017

- Thesis: "Aligning English Sentences with Abstract Meaning Representation Graphs using Inductive Logic Programming"
- Advisor: Dr. Chitta Baral, Professor at Computer Science Department, ASU
- CGPA: 3.72/4

Birla Institute of Technology and Science, Pilani

Rajasthan, India

Bachelors, Information Systems

May 2011 - July 2015

- Semester long tech corporate internship experience at EMC Corp.
- CGPA: 7.24/10

Patents & Publications

D3BA: A Tool for Optimizing Business Processes Using Non-Deterministic Planning Tathagata, C.; Agarwal, S.; Khazaeni, Y.; Rizk, Y.; Ishakian, V. In *AI4BPM Workshop*, 2020.

U.S. Patent 20200074332, "REAL-WORLD EXECUTION OF CONTINGENT PLANS" Agarwal, S.; Muise, C.; Vodolan, M.; Bajgar, O.; Lastras, L., March 5, 2020.

TraceHub: A Platform to bridge gap between State-Of-The-Art Time-Series Analytics and Datasets Agarwal, S.; Muise, C.; Agarwal, M.; Upadhyay S.; Tang Z.; Zeng Z.; Khazaeni Y. In *AAAI Deomonstration Program*, 2020.

TraceHub: Bridging gap between Time-Series Analytics and Datasets Agarwal, S.; Muise, C.; Agarwal, M.; Upadhyay S.; Tang Z.; Zeng Z.; Khazaeni Y. In *O'Reilly AI Conference*, 2019.

MAI : An Intelligent Model Acquisition Interface For Interactive Specification Of Dialog Agents Tathagata, C.; Muise, C.; Agarwal, S.; Lastras, L. In *AAAI Demonstrations Program*, 2019.

Generating Dialogue Agents Via Automated Planning Botea, A.; Muise, C.; Agarwal, S.; Alkan, O.; Bajgar, O.; Daly, E.; Kishimoto, A.; Lastras, L.; Marinescu, R.; Ondrej, J.; Pedemonte, P.; Vodolan, M. In *The Second AAAI Workshop on Reasoning and Learning for Human-Machine Dialogues (DEEP-DIAL)*, 2019.

Executing Contingent Plans: Challenges in Deploying Artificial Agents. Muise, C.; Vodolan, M.; Agarwal, S.; Bajgar, O.; and Lastras, L. In *Fall Symposium on Integrating Planning, Diagnosis, and Causal Reasoning*, 2018.

Aligning English Sentences with Abstract Meaning Representation Graphs using Inductive Logic Programming. Agarwal, S. Diss. *Arizona State University*, 2017.

Automated software test data generation using improved search procedure Agarwal, S.; Bhatter, A. In *Lecture Notes on Software Engineering 3, no. 2, 152*, 2015.

Academic Projects

Aligning English Sentences With AMR graphs using ILP

Arizona, U.S.A

ASU, Masters Thesis

- AMR: Semantic formalism to English natural language encoding meaning of a sentence in a rooted graph
- Idea of approach was predicting concepts invoked by words in a sentence is same as aligning words to those concepts
- Extracted linguistic background knowledge from sentences like lemma, part of speech, modals, named entities, question tokens.
- Concepts in AMR split in nine categories. Learnt ILP rules for each category that invoke AMR concepts from sentence-AMR graph pairs in the training data
- Learnt ILP rules using open source system XHAIL deriving hypothesis in three steps : grounding, finding kernel, hypothesis generation
- Dataset consisted of 13050 AMR/English sentence pairs inclusive of 200 development and test pairs.
- Performance of the aligner was measured using precision, recall and f-score measures on test dataset [P=0.971 | R=0.858 | F=0.91]

Semantic Search on Movie Database *Arizona, U.S.A*

ASU, NLP Course Project

- Created a text based system that predicts movie names on input user query
- Dataset of movie summaries text crawled from IMDB
- Proposed and implemented a semantic approach to find similarity between query and movie summary texts. Created a movie graph of events with Characters(nodes) and Events(edges)
- Used a semantic K(knowledge)-parser to extract events from query and movie summaries
- Used multiple similarity scores to calculate similarity between input and movie graphs.
- Used NER similarity using Stanford CoreNLP, Term similarity using WS4J's PATH, LIN, LESK algorithms and Tf-IDF
- Used NLTK for NER detection and for name-co-reference unification of text
- Evaluated results using a hand prepared test dataset of 50 movies
- Led and coordinated team of 3