

Statement of Purpose

Ashutosh Kumar Nirala

About me: I want to work on household robotics [object manipulation and scene understanding], which I believe is the right step towards our ultimate innovation of Artificial General Intelligence (AGI). *Making a humble contribution to this unprecedented opportunity, creation of strong AI, is my mission.*

Since my undergrad days, I have been working on building toy robots. However, my fascination with AI really started after my graduation. During that time I naively thought of parsing natural language sentences, as per the rules of English grammar, and converting them into computer code. After a little experimentation, I realized it to be a very challenging task and decided to do masters to pursue this further.

Masters at IIT Bombay: I joined the M.Tech. program for Computer Science at Indian Institute of Technology, Bombay, India in 2010.

In the first year, I studied *NLP* and *AI*, taught by Prof. Pushpak Bhattacharya, *ML* and *Statistical Relational Learning*, taught by Prof. Ganesh Ramakrishnan and *Advanced ML*, taught by Prof. Sunita Sarawagi. I soon realized that natural language is a medium of communication between intelligent entities (humans) and, thus in some sense, it refers to states which the entities have a shared understanding of. Thus to fully interpret the NL, in the way we humans do, a NLP system must have an internal representation of the states being referred to. This indeed is a tough nut to crack. To the best of my knowledge there are no NLP system which are doing this, at least in a general way.

In my next year, I shifted my focus towards computer vision (CV), as it deals directly with the physical world. I studied *Digital Image Processing* and *Computer Vision*, taught by Prof. Sharat Chandran. I realized that while NLP works at very abstract level, CV algorithms operate at a very granular level. For example most of the vision systems simply classify an image in the image space, without actually reasoning about the content, unlike humans.

Around the same time (2011) few open source communities like **PCL** and **ROS** developed and released state of the art algorithms for Vision and robotics planning, respectively. *The ROS community at Willow Garage made it possible for anyone to do realistic robotics experiments in simulation (like Gazebo simulator), and it was very useful for me.* For my CV course project, I tried to describe spatial relation of a person's movement with respect to objects in the environment, using Kinect. Point clouds were segmented using PCL, and were tagged manually as chair, table or shelf. This scene was then recreated in the virtual environment with tagged objects. The system was able to identify motion of a person like 'towards the chair', 'along the table' etc. This could be used to interpret such commands given in NL, by choosing a path (among a set of paths in consideration) satisfying the given spatial relation.

Masters project: I did my masters project '*Interpreting Natural Language for Household Robotics*' under the guidance of Prof. Pushpak Bhattacharya. The goal of the project was to simulate the execution of commands given in NL.

ROS was used as middleware and an indoor studio apartment was simulated/created in Gazebo with a robot (PR2). A rule based system with blackboard architecture was developed for the same. A LISP like language was developed to interpret the rules. The system was capable of handling navigational commands like 'move in front of the yellow chair'. The movement was shown in a simulator, and was also verified with the help of a physics engine, during execution. While rules could be easily added for object manipulation, like pick and place, the robot was not able to perform the grasp action (object detection, and grasping pose) reliably, and thus it was not integrated with the final system¹.

Work at Microsoft: After completing my M.Tech., I got a job offer from Microsoft. A part of my job included data analysis of a product, which included analysis of customer feedback, competitors, emerging

¹For short summary of my MTP please see <https://drive.google.com/open?id=0B2aSa7qvjD2JWm1aUjZ5NjJLaWs>. Complete report at: <https://drive.google.com/open?id=0B2aSa7qvjD2JNXIKcko4Z3RMMUE>

trends, etc. However, with agile development, it is difficult to do it exhaustively in a short amount of time. *I worked independently to automate this procedure and created a pipeline to automatically generate reports using the customer data and feedback for all the apps owned by my team. I was awarded, a star award, for this effort.*

When I joined Microsoft in 2010, there was no role for data scientists in the company (outside Microsoft Research). Later, with company reorg, such a role was created and I served as a data scientist for my last 3 months at Microsoft, under *Dr Sivasubramanian Narayanan*. My work was to gauge the trend and usage of Microsoft's Cloud Computing Platform: Azure.

Looking ahead: My experience from industry as well as academia suggests that, research is an ideal way, for me, to serve our society. I would like to devote myself towards research and doing so, contribute to the development of AGI. The first step towards that would be a PhD in the relevant area.

I strongly believe that some major breakthrough in AGI is bound to happen in our lifetime. For example, in a decade or so, many homes would have a home butler to do household chores, and people would be able to converse with it in NL. NLP systems would be much more sophisticated and we would be able to converse with them like we would converse with a savant. In doing so, I am confident, that we would get closer to the ultimate goal of creating an AGI.

Household Robotics: A mobile system accoutred with dexterous arms and vision system, could potentially execute many household tasks like cleaning, arranging stuff, operating household equipments and even cooking. However, household environment is quite challenging for object manipulation, as objects and their poses are hard to detect in this setting. Objects are constantly being moved around, some gets deformed, novel objects are introduced, objects are often occluded, lighting conditions changes etc. Moreover there are safety issues involved as well in introducing any mobile system in household.

I believe that a ceiling mounted robot, with limited movement (say in a kitchen) could be designed to overcome some of the issues, simply by it's design. Being ceiling mounted, and confined to a limited rail, it would greatly solve the navigation and safety concerns. A separate system, having multiple cameras placed in the environment could track objects continuously, thereby making object detection quick when needed. I believe such a system could immediately find use in a household.

Why PhD at Iowa State University (CSE): During my M.Tech project, I worked on the NLP and robotics in simulation. I was the only one in my research group (under my guide) working on this area. I believe a collaborative environment is of utmost importance in doing research, and Iowa State University provides a rich and wide area of research including robotics and vision. I am especially interested in working with professors Yan-Bin Jia. I believe the research being done by his group matches very well with my research interests.

If selected at Iowa State University, I would get a wonderful opportunity to explore field of robotics in depth. Working with the AI and robotics group will be a privilege for me and will help me in contribute something original to the field.