REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
12-09-2019	Final Report	15-Dec-2017 - 15-Mar-2019

4. TITLE AND SUBTITLE
Final Report: Towards an Open CommonSense Knowledge Base

5a. CONTRACT NUMBER
W911NF-18-1-0019

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHORS

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES

Carnegie Mellon University
5000 Forbes Avenue

Pittsburgh, PA 15213 -3589

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
72648-NS-DRP.5

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Abhinav Gupta
UU	UU	UU	UU		19b. TELEPHONE NUMBER 412-268-2067

RPPR Final Report
as of 06-Nov-2019

Agency Code:

Proposal Number: 72648NSDRP **Agreement Number: W911NF-18-1-0019**
INVESTIGATOR(S):

Name: Abhinav Gupta
Email: abhinavg@cs.cmu.edu
Phone Number: 4122682067
Principal: Y

Organization: **Carnegie Mellon University**
Address: 5000 Forbes Avenue, Pittsburgh, PA 152133589
Country: USA
DUNS Number: 052184116 EIN: 250969449

Report Date: 15-Jun-2019 Date Received: 12-Sep-2019
Final Report for Period Beginning 15-Dec-2017 and Ending 15-Mar-2019
Title: Towards an Open CommonSense Knowledge Base
Begin Performance Period: 15-Dec-2017 **End Performance Period:** 15-Mar-2019
Report Term: 0-Other

Submitted By: Abhinav Gupta Email: abhinavg@cs.cmu.edu
Phone: (412) 268-2067

STEM Degrees: **STEM Participants:**

Major Goals: The goal of this project was to
(a) Organize workshop for acquiring commonsense and reasoning with it. The workshop was supposed to investigate big issues related to commonsense learning and reasoning with goals to start a program.
(b) Inspire commonsense acquisition using human cognitive development.
(c) Introduce commonsense acquisition via robotics
(d) explore web-driven commonsense resources.
(e) Develop few benchmarks for commonsense reasoning via VQA

Accomplishments: Accomplishments:
(a) organized the commonsense workshop which included leading researchers in different fields of AI. We had two-days of open discussion on what topics are of importance to make progress in the field of commonsense learning. This led to a DARPA program on Machine Commonsense.

(b) For research, we focused on learning interpretable models of intuitive physics and using robots to learn commonsense knowledge. These resulted in publication at ECCV, NeruIPS and ICCV.

(c) We explored use of graphs for prediction models and also worked on benchmarks for commonsense reasoning via VQA (with AI2). AI2 is now part of MCS program for evaluations.

Training Opportunities: Nothing to Report

Results Dissemination: Apart from the accepted papers, the PI gave several talks in CVPR workshops including workshop on Cognitive Psychology and AI, Commonsense Reasoning (in conjunction with CVPR).

Honors and Awards: Abhinav Gupta received ONR Young Investigator Award.
Tom Mitchell received President's Medal, Stevens Institute of Technology

Protocol Activity Status:

Technology Transfer: Nothing to Report

PARTICIPANTS:

RPPR Final Report
as of 06-Nov-2019

Participant Type: PD/PI
Participant: Abhinav Gupta
Person Months Worked: 1.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Co PD/PI
Participant: Tom Mitchell
Person Months Worked: 1.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: Y
Other Collaborators:

Participant Type: Co-Investigator
Participant: Josh Tenenbaum
Person Months Worked: 1.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Other (specify)
Participant: Melissa Struhl
Person Months Worked: 9.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Technician
Participant: Mario Belledonne
Person Months Worked: 3.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Staff Scientist (doctoral level)
Participant: Ilker Yildirim
Person Months Worked: 3.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

RPPR Final Report
as of 06-Nov-2019

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: FOROUGH ARABSHAHI
Person Months Worked: 5.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Tao Chen
Person Months Worked: 7.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Victoria Dean
Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Dhiraj Gandhi
Person Months Worked: 7.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Technician
Participant: Deepthi Hegde
Person Months Worked: 2.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Wenxuan Zhou
Person Months Worked: 7.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Technician
Participant: Gaurav Pathak
Person Months Worked: 11.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Lerrel Pinto
Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Adithya Murali
Person Months Worked: 3.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Graduate Student (research assistant)
Participant: Samantha Powers
Person Months Worked: 2.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Other Professional
Participant: Bryan Kiesel
Person Months Worked: 3.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Other Professional
Participant: KATHRYN MAZAITIS
Person Months Worked: 7.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N

Other Collaborators:

Participant Type: Other Professional
Participant: Wei Yang
Person Months Worked: 2.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Technician
Participant: Pratyusha Sharma
Person Months Worked: 4.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

CONFERENCE PAPERS:

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: ECCV 2018
Date Received: 15-Nov-2018 Conference Date: 09-Sep-2018 Date Published:
Conference Location: Munich
Paper Title: Interpretable Intuitive Physics Model
Authors: Tian Ye, Xiaolong Wang, James Davidson, Abhinav Gupta
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: NIPS 2018
Date Received: 04-Sep-2019 Conference Date: 03-Dec-2018 Date Published: 03-Dec-2018
Conference Location: MONTREAL
Paper Title: Hardware Conditioned Policies for Multi-Robot Transfer Learning
Authors: Tao Chen, Adithya Murali, Abhinav Gupta
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 3-Accepted
Conference Name: International Conference on Computer Vision
Date Received: Conference Date: 27-Oct-2019 Date Published: 27-Oct-2019
Conference Location: Seoul
Paper Title: Compositional Video Prediction
Authors: Yufei Ye, Maneesh Singh, Shubham Tulsiani, Abhinav Gupta
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 06-Nov-2019

Publication Type: Conference Paper or Presentation **Publication Status:** 3-Accepted

Conference Name: NeurIPS

Date Received: 04-Sep-2019 Conference Date: 09-Dec-2019 Date Published: 10-Dec-2019

Conference Location: Vancouver

Paper Title: Modeling Expectation Violation in Intuitive Physics with Coarse Probabilistic Object Representations

Authors: Kevin A Smith, Lingjie Mei, Shunyu Yao, Jiajun Wu, Elizabeth Spelke, Joshua B Tenenbaum, Tomer D U

Acknowledged Federal Support: **Y**

Project Description

Research Area 1:
1.1 Disentangled Representations of Intuitive Physics

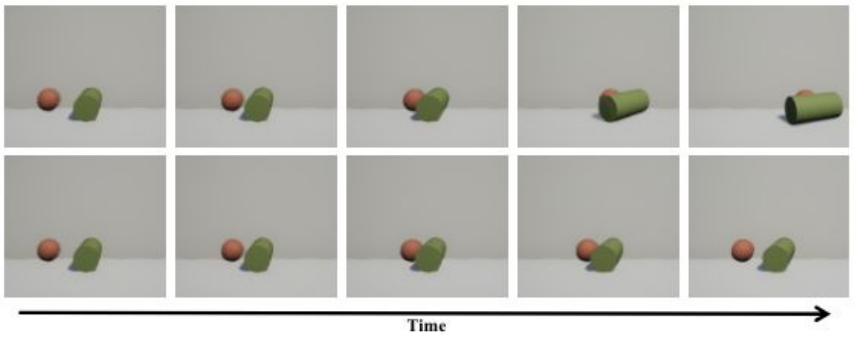

Figure 1. Interpretable Physics Models. Consider the sequences shown above. Not only we can predict the future frames of collisions but we can also predict the underlying factors that lead to such an inference. For example, we can infer the mass of cylinder is much higher in second sequence and therefore it hardly moves in the image. Our ability to infer meaningful underlying latent factors inspire us in this paper to learn an interpretable intuitive physics model.

Consider the collision image sequences shown in Figure 1. When people see these images, they not only recognize the shapes and color of objects but also predict what is going to happen next. For example, in the first sequence people can predict that the cylinder is going to rotate while in the second sequence the ball will bounce with no motion on cylinder. But beyond visual prediction, we can even infer the underlying latent factors which can help us explain the difference in visual predictions. For example, a possible explanation of the behavior between the two sequences, if we knew the ball's mass didn't change, is that the first sequence's cylinder was lighter than the ball whereas in the second sequence the cylinder was heavier than the ball. Beyond this we can deduce that that the cylinder in the first sequence was much lighter than the one in the second.

Humans demonstrate the profound ability to understand the underlying physics of the world and use it to predict the future. And we use this physical commonsense for not only rich understanding but also for physical interactions. The question arises as to whether this physical commonsense is just an end-to-end model with intermediate representations being a black-box? Or does a physical commonsense model require explicit and meaningful intermediate representations? For humans, the answer appears to be the latter. We can predict the future if some underlying conditions are changed. For example, we can predict that if we increase the speed with which the ball is thrown 10x in the second sequence then the cylinder will rotate.

We focus on the task of learning an intuitive model of physics. Unlike some recent efforts, where the goal is to learn physics in an end-to-end manner with little-to-no constraints on intermediary

layers, we focus on learning an {\bf interpretable} model. More specifically, the bottleneck layers in our network model physical properties such as mass, friction, etc.

Learning an interpretable intuitive physics model is, however, quite a challenging task. For example, Wu et al. attempts to build a model but the inverse graphics engine infers physical properties such mass and friction. These properties are then used with neural physics engine or simulators for prediction. But can we really infer physical properties from the few frames of such collisions? Can we separate friction from mass, restitution by observing the frames? The fact is most of these physical factors are so dependent that it is infeasible to infer the exact values of physical properties. For example we can determine ratios between properties but not the precise value of both (e.g., we can determine the relative mass between two objects but not the exact value for both from a sequence). This is precisely the reason why in Wu et al. only one factor is inferred from motion and the other factor is directly correlated the appearance. Furthermore, the learned physics model is quite domain-specific and will not generalize--even across different shapes.

To tackle these challenges, we propose an interpretable intuitive physics model, where specific dimensions in the bottleneck layers correspond to different physical properties. The bottleneck layer models the distribution rather than infer precise values of mass, speed and friction. In order to demonstrate that our system models these underlying physical properties, we train our model on collision of different shapes (cube, cone, cylinder, spheres etc.) and test on collisions of unseen combination of shapes altogether. We also demonstrate the richness of our model by predicting the future states under different physical conditions. For example, how will the future frames look if the friction is doubled etc.

Our contributions include: (a) intuitive physics model that disentangles different physical properties in an interpretable way; (b) a staggered training algorithm designed to distinguish the subtleties between different physical quantities; (c) generalization to different shapes and physical quantity combinations. But most importantly, (d) the ability to adapt future predictions when physical environments change. Note (d) is different from generalization because the first four frames are observed for a completely different physical scene but the hallucination/prediction is done for a different scene.

1.2 Disentangled Representations of Intuitive Physics

People have a rich understanding of everyday physics that they use to predict how the future might unfold, plan actions, and manipulate tools. This commonsense reasoning includes a set of early developing or possibly innate expectations about the behavior of objects, which are part of ‘core knowledge’. For example, even very young infants generally expect objects to remain coherent wholes that follow spatially contiguous paths, and not wink in and out of existence. Intelligent machines that can interact with the physical world in a human-like way should hold human-like physical intuitions. We propose that human-like physical understanding is based on explicit object-centric representation and their associated dynamics, similar to the idea of a

Mental Game Engine. Such object representations have constant physical properties regardless of their perceptual appearance, an appearance that can differ greatly depending on factors such as viewpoint, distance, and occlusion. We assume these object-based representations as given, and propose that the main computational burden is learning how visual input is parsed into these representations.

We present a new model, “Approximate Derenderer, Extended Physics, and Tracking” (ADEPT), that closes the loop between cognitive models of intuitive physics which assume object parses are given, and computer vision models that parse images into physical object representations. Importantly, we suggest that perception does not have to be exact to capture basic physical expectations. Approximate perception allows the model to trivially extend to novel objects at a loss of object identity information; this is similar to the way young infants can reason about new objects but are insensitive to changes of object shape. Our model (i) learns to approximately parse novel arbitrarily shaped objects into approximate geometric forms, (ii) makes extended predictions about future world states, by using a robust dynamics model that combines a physics engine with qualitative state-changes, and (iii) uses a particle filter to tie together parsing and predicting, allowing it to track objects over occlusion. In the spirit of Riochet et al. [2018] and Piloto et al. [2018], we evaluate our model using the Violation of Expectations (VoE) paradigm from developmental psychology. In this paradigm, models are show scenes that are matched as closely as possible, except that one scene contains an event that violates intuitive physical expectations (e.g., an object disappearing; Fig 1). A model passes the test if its predictions diverge from observations more strongly in the violation video than the control.

We tested the ADEPT model on eight different scenarios, which replicate tests from developmental psychology and capture different aspects of early core object knowledge. These scenarios examine concepts such as permanence (objects do not appear or disappear for no reason), continuity (objects move along connected trajectories), and solidity (objects cannot move through one another). We compared ADEPT to models that learn physics without explicit object representations, and found that only it discriminates violations from control stimuli at above chance rates in all eight scenarios, and did so at similar rates as humans.

Figure 1: Frames taken from a physically implausible video, in which a yellow cube seems to disappear behind the occluder.

Research Area 2: Graphs for Visual Prediction

A single image of a scene allows us humans to make a remarkable number of judgments about the underlying world. For example, consider the two images on the left in Fig 2. We can easily infer that the top image depicts some stacked blocks, and the bottom shows a human with his arms raised. While these inferences showcase our ability to understand what is, even more remarkably, we are capable of predicting what will happen next. For example, not only do we know that there are stacked blocks in the top image, we understand that the blue and yellow ones will topple and fall to the left. Similarly, we know that the person in the bottom image will lift his torso while keeping his hands in place. In this work, we aim to build a model that can do the same -- from a {single} (annotated) image of a scene, predict at a pixel level, what the future will be.

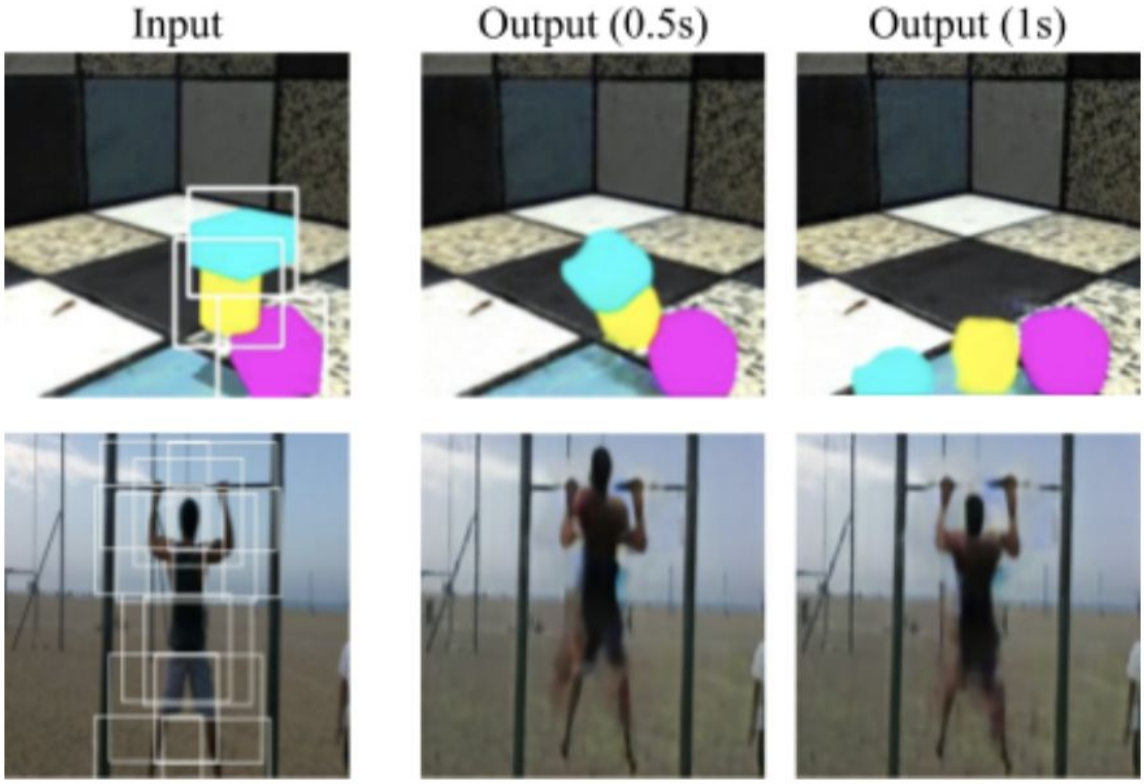

Figure 2: Given a still image with locations of entities (objects or joints), we predict a sequence of future frames. We visualize two frames from the predicted sequence for the given inputs.

A key factor in the ability to make these predictions is that we understand scenes in terms of `entities', that can move and interact e.g.\ the blocks are separate objects that move; the human body's motion can similarly be understood in terms of the correlated motion of the limbs. We operationalize this ideology and present an approach that instead of directly predicting future frames, learns to predict the future locations and appearance of the entities in the scene, and via these composes a prediction of the future frame. The modeling of appearance and the learned composition allows our method to leverage the benefits of independent per-entity representations while allowing for reasoning in pose changes or overlap/occlusions in pixel space.

Although our proposed factorization allows learning models capable of predicting the future frames via entity-based reasoning, this task of inferring future frames from a single input image is fundamentally ill-posed. To allow for the inherent multi-modality of the prediction space, we propose to use a trajectory-level latent random variable that implicitly captures the ambiguities over the whole video and train a future predictor conditioned of this latent variable. We demonstrate that modeling the ambiguities using this single latent variable instead of per-timestep random variables allows us to make more realistic predictions as well as sample diverse plausible futures.

We validate our approach using two datasets where the `entities' either represent distinct objects, or human body joints, and demonstrate that the same method allows for predicting future frames across these diverse settings. We demonstrate: (a) the benefits of our proposed entity-level factorization; (b) ability of the corresponding learned decoder to generate future frames; (c) capability to sample different futures.

Research Area 3: VQA Benchmark for Commonsense

Visual Question Answering (VQA) has recently become a prominent topic in computer vision. It is not only a fertile ground for vision and language research but it also provides a proxy to evaluate AI models for the task of scene understanding. However, recent research shows that current VQA tasks suffer from several deficiencies: first, the biases of language are strong and can lead to the solutions without looking at visual data. Even if the biases are corrected, most VQA questions are in form of simple counting, yes/no and visual detection. Therefore, most of the existing datasets do not require reasoning or association with external knowledge. However, the most difficult and interesting questions ideally require knowing more than what the question entails or what information is contained in the images. For example, consider the question in Figure 2, which asks about the relation between the teddy bear and an American president. The information in the image is not complete for answering the question. We need to link the image content to the external knowledge sources, such as the sentences on the right side of the figure taken from Wikipedia. Given the question, image, and Wikipedia sentences, there is now enough information to answer the questions: Teddy Roosevelt!

We introduce the OK-VQA, which only includes questions that require external resources for answering them. The current VQA datasets, such as \cite{antol15} or \cite{zhu16}, are not suitable for this task since they mainly focus on questions that can be answered given the image (for instance, counting questions or question about the color, etc). OK-VQA includes more than 10,000 questions for training and evaluation. Our experimental evaluations show that the performance of the state-of-the-art VQA models drastically degrades on our new dataset.

Traditional approaches to VQA, which are typically based on a combination of CNNs and RNNs or more advanced attention-based approaches, are not suitable for these types of questions that require external knowledge. The main reason is that they are limited to the information provided in the image and the question and they mainly learn patterns that appear in the training data. Fact-based VQA systems have also been proposed, where a set of supporting facts is provided for each question-answer pair. In this work, we take a step further and introduce the task of knowledge-based Visual Question Answering, where answering a question requires a source of external knowledge, either in structured or unstructured form. We explore the latter: our goal is to explore free-from knowledge from the web (Wikipedia in our case) to answer questions that are related to the image.

Using web data for VQA tasks poses various key challenges: (1) What queries should be generated given the question and the image content? (2) How to find the relevant articles (knowledge sources) among the retrieved results? (3) How to find the correct answer in these unstructured knowledge bases? (4) How to integrate the knowledge retrieval module with the rest of the pipeline for comprehending the question and the image?

To address these challenges, we propose our Wikipedia extraction method. Our method generates queries based on the question and objects and scene information. It then retrieves the queried articles from Wikipedia and ranks their content using our proposed ArticleNet. The most relevant content is then integrated with the image and question processing pipeline to compute the final answer. One of the key distinguishing factors of our method is the handling of unstructured knowledge (in contrast to methods based on structured and annotated knowledge bases or facts).

We show that our ArticleNet achieves remarkable results in retrieving the relevant sentences among the Wikipedia articles. Moreover, we show that our method outperforms the baseline VQA models when it relies on the external knowledge.

Our contributions in this project are three-fold: (a) we introduce OK-VQA dataset, which includes only questions that require external resources for answering them; (b) we show how existing VQA systems fail on the task of knowledge-based VQA; (c) finally, we propose a model which exploits the unstructured knowledge and show improvements over a state-of-the-art VQA model for the task of knowledge-based VQA.

Research Area 4: Generalization of action policies to different hardware

In recent years, we have seen remarkable success in the field of deep reinforcement learning (DRL). From learning policies for games to training robots in simulators \cite{ddpg}, neural network based policies have shown remarkable success. But will these successes translate to real world robots? Can we use DRL for learning policies of how to open a bottle, grasping or even simpler tasks like fixturing and peg-insertion? One major shortcoming of current approaches is that they are not sample-efficient. We need millions of training examples to learn policies for even simple actions. Another major bottleneck is that these policies are specific to the hardware on which training is performed. If we apply a policy trained on one robot to a different robot it will fail to generalize. Therefore, in this paradigm, one would need to collect millions of examples for each task and each robot.

But what makes this problem even more frustrating is that since there is no standardization in terms of hardware, different labs collect large-scale data using different hardware. These hardware vary in degrees of freedom (DOF), kinematic design and even dynamics. Because the learning process is so hardware-specific, there is no way to pool and use all the shared data collected across using different types of robots, especially when the robots are trained under torque control. There have been efforts to overcome dependence on hardware properties by learning invariance to robot dynamics using dynamic randomization. However, learning a policy invariant to other hardware properties such as degrees of freedom and kinematic structure is a challenging problem.

In this project, we propose an alternative solution: instead of trying to learn the invariance to hardware; we embrace these differences and propose to learn a policy conditioned on the hardware properties itself. Our core idea is to formulate the policy as a function of current state and the hardware properties. So, in our formulation, the policy decides the action based on current state and its own capabilities (as defined by hardware vector). But how do you represent the robot hardware as vector? In this project, we propose two different possibilities. First, we propose an explicit representation where the kinematic structure itself is fed as input to the policy function. But such an approach will not be able to encode robot dynamics which might be hard to measure. Therefore, our second solution is to learn an embedding space for robot hardware itself. Our results indicate that encoding the kinematic structures explicitly enables high success rate on zero-shot transfer to new kinematic structure. And learning the embedding vector for hardware implicitly without using kinematics and dynamics information is able to give comparable performance to the model where we use all of the kinematics and dynamics information. Finally, we also demonstrate that the learned policy can also adapt to new robots with much less data samples via finetuning.

Workshop

We organized a two-day workshop on commonsense and AI. This workshop was attended by:

1. Leslie Kaebling
2. Josh Tenenbaum
3. Yejin Choi
4. Ernst Davis
5. William Cohen
6. Oren Etzioni
7. Tom Mitchell
8. Jitendra Malik
9. Yann Lecun
10. Dieter Fox
11. Fernando Pereira
12. Elizabeth Spelke

The workshop focused on discussions regarding what are areas to focus in commonsense. How do we represent commonsense? How do we learn it? and What are the benchmark tasks in commonsense learning? Some key directions emerged from the workshop including a benchmark that AI2 is developing. There were also key discussions on how cognitive development and psychology can play a role in developing commonsense-based reasoning approaches. This workshop led to DARPA program on Machine CommonSense. It also led to key contributors which are now part of program including AI2.