AFRL-AFOSR-VA-TR-2023-0219

The Marvin Minsky Institute for Society of Mind Theory

Davis, Randall
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
77 MASSACHUSETTS AVE
CAMBRIDGE, MA, 02139
USA

12/01/2022
Final Technical Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE 20221201	**2. REPORT TYPE** Final	**3. DATES COVERED** **START DATE** 20161215	**END DATE** 20211014

4. TITLE AND SUBTITLE
The Marvin Minsky Institute for Society of Mind Theory

5a. CONTRACT NUMBER	**5b. GRANT NUMBER** FA9550-17-1-0081	**5c. PROGRAM ELEMENT NUMBER** 61102F
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
Randall Davis

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) MASSACHUSETTS INSTITUTE OF TECHNOLOGY 77 MASSACHUSETTS AVE CAMBRIDGE, MA 02139 USA	**8. PERFORMING ORGANIZATION REPORT NUMBER**

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington, VA 22203	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFRL/AFOSR RTA2	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)** AFRL-AFOSR-VA-TR-2023-0219

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Building on Genesis's self-reflection capability we will perform experiments where agents learn how solve problems in new and creative ways, learning by example and instruction, and distilling general procedures and recipes. Such systems will require reflective capabilities of various kinds in order to recognize, abstract, and align what they learn. To better understand the interface between language and other senses, we will experiment with biologicallyplausible models of vision, including architectures that can integrate top-down and bottom-up information. With a similar integrative approach, we will pursue human-scale models of language that can use knowledge and multi-modal senses to constrain meaning. We will integrate a new level of hypothetical reasoning capabilities into our cognitive models, enabling a range of more complex behaviors including forms of imagination, moral reasoning, evaluating, and planning. We believe all of this will be an important step toward understanding the competences necessary for human intelligence.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT** UU	**18. NUMBER OF PAGES** 16
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U		

19a. NAME OF RESPONSIBLE PERSON RICHARD RIECKEN	**19b. PHONE NUMBER** *(Include area code)* 696-9736

Accomplishments

Please list the main research objectives of this project

The main objective of the project is to advance the approach to Artificial Intelligence pioneered by Marvin Minsky, one of the founders of the field of Artificial Intelligence, and one of the founders of our lab, the MIT Computer Science and Artificial Intelligence Lab (then the MIT AI Lab). These theories are best expressed in his books, *The Society of Mind* and *The Emotion Machine*. Society of Mind theory is multifaceted and multidimensional, using a methodology of "thinking about thinking". It envisions the mind as a collection of communicating agents, each with perhaps differing viewpoints and problem solving approaches, that cooperate and compete in patterns that give rise to what we call intelligence. It emphasizes symbolic representations, commonsense knowledge, heuristic procedures, and reasoning by *panalogy* (parallel analogy). It proposes active memory architectures such as frames and K-lines ("knowledge lines"), used for what Minsky called *re-membering* and *re-minding*.

Please provide details of accomplishments during this reporting period

• A major direction of our work has been experimenting with several ways of **representing commonsense knowledge**, and integrating it with other approaches, particularly the modern machine learning approaches of word embeddings, transformers, and other language models. Our goal is to support **analogical reasoning with commonsense knowledge** and get the "**best of both worlds**" – the *confidence* of human-curated commonsense knowledge, and the *coverage* of large-scale language models trained on enormous corpora of text.

We developed *RetroGAN* [Colon-Hernandez, Xin, Lieberman, Havasi, Breazeal 21], which **integrates our commonsense knowledge base ConceptNet, with word embeddings**. RetroGAN uses the technique of *retrofitting*, which distorts a vector space of word embeddings, to reflect a corpus of symbolic assertions, in our case ConceptNet. RetroGAN then uses the popular generalized adversarial network (GAN) inference technique to learn a one-to-one mapping between concepts and their retrofitted counterparts.

That becomes especially useful in dealing with rare words or out-of-knowledge-base concepts. The mapping can be extended to "guess" what the meaning of novel concepts are, based on making analogies to already-understood concepts. This is inspired in part by the way some natural language understanding systems treat out-of-vocabulary words. Tests show that our techniques have better performance on rare and unusual concepts than conventional approaches.

• More recently, we have been experimenting with **commonsense inference in *transformer* models**. We introduced a technique called *hinting*, a kind of *prompting* that supplies additional input to a transformer model to constrain its inference [Colon-Hernandez, et al 22a]. We use prompts that are a hybrid of fully-specified assertions and partially-specified assertions that act as templates to constrain inference, taken from ConceptNet commonsense assertions.

Hinting also enables ***contextualized* inference**, in that it can perform inferences based on the context of a story, similar those presented by Winston's *Genesis* system. Genesis produces an *elaboration graph* (set of causally-connected events and descriptions). A sentence of the story is identified to give the system "something to talk about" and assure relevance. Since the generalization process is underconstrained, hinting assures that generalization takes place in a way that is relevant to the story context under consideration.

• We developed techniques for **commonsense-enabled parsing** [Xin, Lieberman, Chin 21], showing how to use commonsense knowledge to resolve ambiguities in natural language understanding, particularly prepositional phrase attachment and pronoun coreference. This is aimed at the context of enabling programming in natural language and conversational Siri-like agents, as explained in our report on our companion AFOSR project on Human-Machine Teaming. Our system, PatchComm, takes the result of a statistical parser, Spacy, and applies corrections if Spacy's choices are not consistent with the semantics of commonsense knowledge from ConceptNet. We are also exploring use of RetroGAN and our hinting techniques in this context.

• We demonstrated a new technique for **joint inference with multiple commonsense knowledge bases** [Colon-Hernandez, et al 22b]. While joint inference with fully consistent logical statements is simply a matter of conjoining them, inference with multiple commonsense assertions is trickier because of their imprecision and varying levels of generalization.

• We explored the use of **commonsense knowledge with transformer models** in understanding step-by-step procedures such as kitchen recipes, in collaboration with our colleague Dr. Howard Shrobe, who is working in a DARPA project to understand descriptions of procedures [Pei 21]. Pei augmented the GPT-2 model with assertions from ConceptNet and used it to try to predict next steps of procedures.

• We introduced the technique of *lensing* for **representing perspective or point of view in machine learning** [Dinakar, Lieberman 21]. The importance of handling multiple perspectives on problem solving and being able to shift perspective is a keystone of *Society of Mind*, best expressed in Minsky's dictum, "If you understand something in only one way, you don't understand it at all". It allows the user to criticize a machine learning model by example, following Minsky's notion of *critics*. Then it provides feedback to the machine learning algorithm to improve its behavior. Lensing can be applied to many different kinds of machine learning. You can "mix and match" lenses to view datasets through different perspectives. Lensing is used in a mixed-initiative manner described in the report on our companion AFOSR project on Human-Machine Teaming.

• We developed ***Astroparse***, a system to increase the abilities of reasoning systems, question answering systems, and related types of systems by enhancing the output of available statistical language processing systems and convert them to ternary expressions. This semantic representation, the ternary expression representation for knowledge and language, consists of linguistic triples (subject, relation, and object). A sentence or thought can be represented by a small number of structural ternary expressions—which highlight major syntacto-semantic relations—in conjunction with an extended collection of additional ternary expressions allowing the original sentence to be reconstructed exactly. Specifically, ternary expressions may capture structural relations (verb and arguments, prepositional attachment, etc.), syntactic features (tense, aspect, determiners, etc.), or lexical features (number, person, etc.). By prioritizing the most important relations—namely, the structural relations—the ternary expression representation allows for efficient indexing and matching. Furthermore, unlike typical knowledge “triples”, e.g., RDF, our ternary expressions support embedding, preserve instance identity across multiple uses, and distinguish instance identity across unrelated uses of the same word. By expanding the ability to produce ternary expressions, we can greatly increase the abilities of downstream reasoning systems.

How were the results disseminated to communities of interest?

Results were disseminated to the scientific community in Artificial Intelligence and Human-Computer Interaction through our publications, talks, teaching and mentoring students. Please see our Publications list.

We also produced several publications that communicated important topics relevant to the grant for the more general computer science community and for the general public.

• We contributed to a major article on the history of Logo [Solomon et al 20], the computer language for education designed by Seymour Papert and colleagues. Minsky made major contributions to Logo development, and Papert’s educational philosophy was highly influenced by Minsky’s “thinking about thinking”. Lieberman was a member of the original Logo research team at MIT. This was for the ACM History of Programming Languages conference/journal (HOPL), an event that occurs only once every ten years. HOPL is the archival record of the definitive history of almost every major programming language.

• Lieberman is editor of the blog for the journal ACM Transactions on Intelligent Interactive Systems (ACM TiiS). “Transactions” are the top-tier journal series of the main professional society for computer science. [Lieberman 21] discusses the important topic of ethics in Human Centered AI. It reviews an article and book by Ben Shneiderman, one of the founders of Human-Computer Interaction, that treats how companies and governments can ethically incorporate AI.

• Our work on commonsense reasoning was **reported in Scientific American**,
in an article by César Hidalgo, entitled “Why We Forgive Humans More Readily Than Machines”.

https://www.scientificamerican.com/article/why-we-forgive-humans-more-readily-than-machines/

• Ternary expressions are the native representation of the START Natural Language Processing System [Katz and Levin 88, Katz 97, Katz et al. 07]. START has been and is in use for natural language processing in multiple research systems and projects including the Genesis Project, the RITA system as part of the ASIST project, the MIT–Air Force Artificial Intelligence Accelerator, and research with the Office of Naval Research. Our recent work in increasing coverage by enhancing dependency parses is experimentally deployed in the START system and is in testing in some of these client systems.

• Our work on understanding social interactions ([Tejwani et al. 21] and [Netanyahu et al. 21]) was reported by MIT News in an articled entitled “Giving robots social skills”:

https://news.mit.edu/2021/robots-social-skills-1105

and in numerous other publications, such as:

https://techcrunch.com/2021/11/05/teaching-robots-to-socialize/

https://www.lifewire.com/robots-are-getting-more-social-to-better-understand-you-5209295

https://www.unite.ai/researchers-develop-framework-to-give-robots-social-skills/

What do you plan to do during the next reporting period to accomplish the goals and objectives?

• We will continue the development of RetroGAN, and assess its utility in some of the other components of our work: to assist in understanding the semantics of natural language, especially parsing; to apply to discourse with conversational agents and natural language programming; to understand descriptions of procedures, and other applications.

• We are developing a downstream application Deep Relationship Discovery (DRD) to perform search for candidates for analogies. Given two concepts, DRD will search for the probability of various relations occurring between them. These guide generalization in order to make analogies. For example, to make an analogy between birds and airplanes, it's important to know that they share the relation (Has-Part Wing).

• We continue to develop our promising *lensing* technique for representing point of view, especially for mental and physical health applications, following the successes of our cardiology, cyberbullying, and crisis counseling projects.

• We will continue to expand the coverage of our Astroparse system by training on additional language constructions. We will also explore expanding the capabilities of our system to produce ternary expressions using multiple dependency parsers. Our tree-matching algorithm is not specific to any one dependency parse schema and can be applied to the results of any dependency parse. Since various parsers have their own advantages and drawbacks, we may be able to achieve better results by integrating the ternary expressions created using each one.

Impacts

Development of the principal discipline(s) of the project

The principal discipline of our work is Artificial Intelligence. Our work provides fundamental contributions to the science of Artificial Intelligence in the following ways:

• **Integrating symbolic and subsymbolic methods in AI.** There is a longstanding debate between the value of symbolic and neural methods in AI. We don't view them as in opposition, but different views on the same problem. Roughly, symbolic views proceed "top-down" and subsymbolic "bottom-up", and there's no one right answer. Future progress depends upon incorporating both. During his lifetime, Minsky made important contributions to almost every aspect of AI, from his early work on neural nets in his thesis at Princeton, and the *Perceptrons* book with Papert (sometimes characterized, we believe incorrectly, as having "killed off" neural net research), to his better-known contributions to symbolic AI.

• **Understanding the role of commonsense knowledge in AI.** Commonsense knowledge is not straightforward to learn from text; it is rarely voiced explicitly because it is the stuff that "everybody knows". Reasoning with commonsense knowledge is not the same as reasoning with logical or factual knowledge, since it is often vague, context-dependent, and admits of many exceptions and caveats. Neither is it probabilistic reasoning, as people come to commonsense judgments without counting and classifying specific data instances. Our work contributes to understanding the salient and unique properties of commonsense reasoning.

• Our work contributes to **explainability, controllability, and transparency in AI.** Commonsense explanations are more readily understood and accepted by people than statistical or numeric analyses. People often express their desires for AI performance in commonsense terms, and AI systems must operationalize these desires. People need to "sanity check" AI, and understand the range of circumstances under which they should have confidence in AI systems.

Impact on society beyond science and technology

We had a major success in an application of lensing to Cardiology. Along with our medical colleagues, we participated in a study of how male and female patients report cardiac symptoms [Kreatsoulas et al 19]. A machine learning study used our lensing technique to analyze doctor-patient interviews, and extracted male-associated and female-associated viewpoints on angina (chest pain). We also mentioned this in our report on our companion AFOSR project on Human-Machine Teaming .

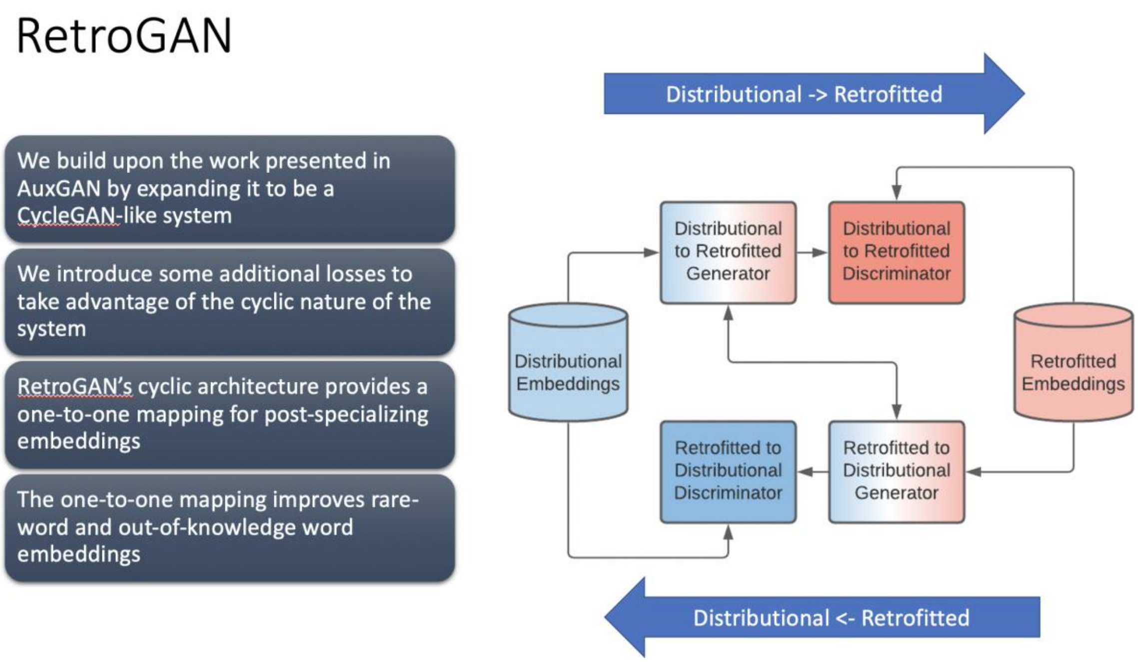
We were able to show that, although men and women express angina differently, the underlying conditions were the same. As a result, recently, the American Heart Association and the American College of Cardiology changed their diagnosis guidelines to eliminate a spurious distinction between "typical" and "atypical" angina [Gulati et al 21]. They explicitly cited [Kreatsoulas, Dinakar, et al 19] as strong evidence, and took wording for the guidelines, verbatim, from our paper. We expect this change to result in a significant reduction in the misdiagnosis of cardiac problems, the leading cause of death, in women.

Changes

No changes.

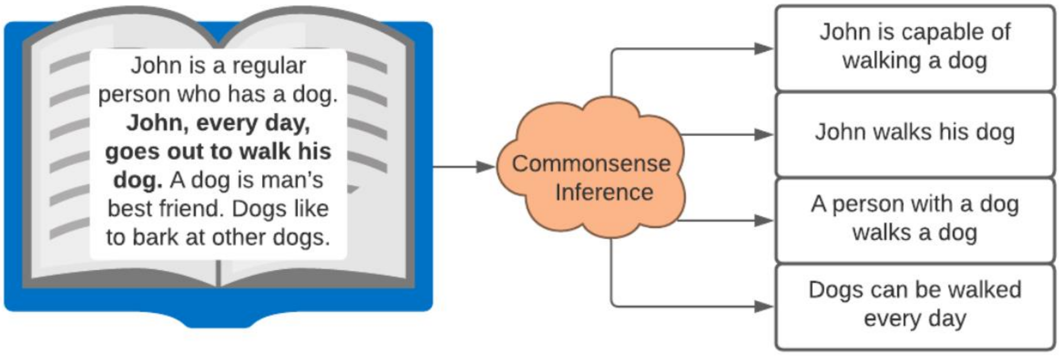
Technical Updates

• RetroGAN: A Cyclic Post-Specialization System for Improving Out-of-Knowledge and Rare Word Representations



Retrofitting is a technique used to move word vectors closer together or further apart in their space to reflect their relationships in a Knowledge Base (KB). However, retrofitting only works on concepts that are present in that KB. RetroGAN uses a pair of Generative Adversarial Networks (GANs) to learn a one- to-one mapping between concepts and their retrofitted counterparts. It applies that mapping (post-specializes) to handle concepts that do not appear in the original KB in a manner similar to how some natural language systems handle out-of-vocabulary entries. We test our system on three word-similarity benchmarks and a downstream sentence simplification task, and achieve the state of the art (CARD-660). Altogether, our results demonstrate our system's effectiveness for out-of-knowledge and rare word generalization.

• **Adversarial Transformer Language Models for Contextual Commonsense Inference**



Contextualized commonsense inference is the task of generating commonsense assertions or facts from a given story and a sentence from the story. This task is hard even for modern contextual language models. Some of the problems with the task are lack of controllability for topics of the inferred assertions, lack of commonsense knowledge during pre-training, and possibly hallucinated or false assertions. In this work we tackle these three challenges by first utilizing a technique called ``hinting" which is a hybrid prompting technique that augments training data by adding ``hints" that are part of assertions. This serves as a control signal for the language model to ``talk" or generate assertions about whatever is in the hint. Secondly, we combine three knowledge graphs in a textual manner by aligning the assertions in the knowledge graphs with a story and a target sentence, and replace their symbolic assertions with textual versions of them. This combination allows us to train a single model to perform joint inference with these knowledge graphs. Thirdly, we train a combination of two language models in an adversarial manner such that one model generates plausible assertions and another scores them on the factuality of them. Altogether we present a system that can controllably generate assertions using joint inference, and can score these assertions.

• PatchComm natural language programming

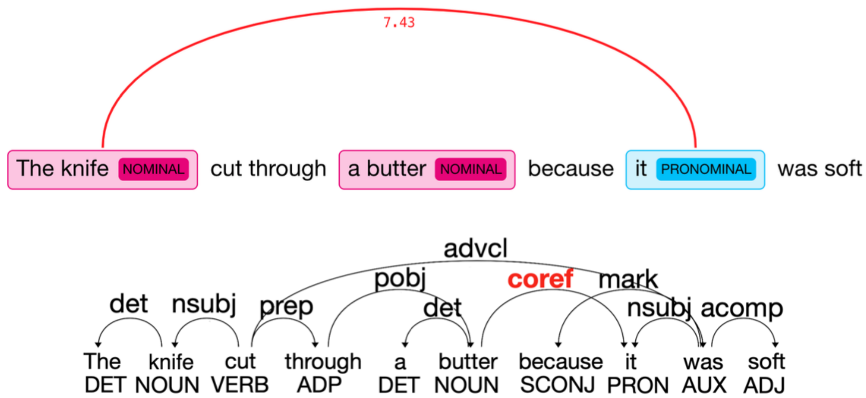
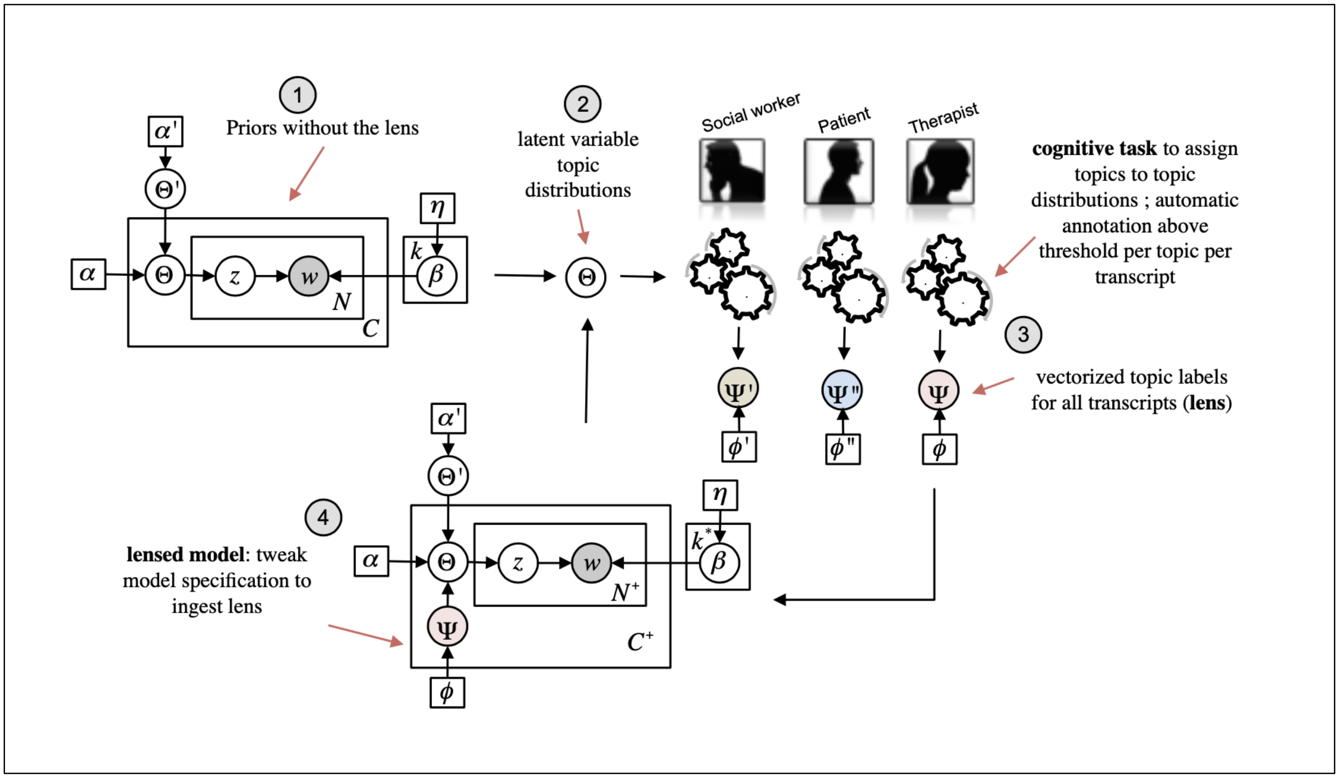


Fig. 4. PATCHCOMM makes a more commonsensical coreference resolution than NeuralCoref alone does. Note that in the bottom figure, PATCHCOMM automatically embeds its coreference resolution result into its spaCy-style parsing output.

Natural language understanding (NLU) has improved so much in recent years that it may soon be feasible to replace conventional programming languages with user interactions in natural languages. To accomplish this, it will be necessary to use NLU modules specifically adapted for the natural language programming task. We present a natural language programming system, PATCHCOMM PRO, and its associated NLU module, PATCHCOMM. In its current version, PATCHCOMM uses large-scale commonsense knowledge to guide a state-of-the-art statistical parser to resolve linguistic ambiguities, notably prepositional phrase attachment and pronoun reference. Rather than simply translating the user's utterances into code, PATCHCOMM PRO treats the natural language programming task as a discourse between the user and the machine, where the user talks about how the program should work and the machine tries to realize it. PATCHCOMM PRO's advantages are in explainability and usability. It first produces transparent, symbolic rules that turn discourses into frame semantic representations, and then renders these representations.

• **Lensing technique for representing point of view in machine learning**



Many datasets represent a combination of several viewpoints – different ways of looking at the same data that lead to different generalizations. For example, a corpus with examples generated by different people may be mixtures of many perspectives and can be viewed with different perspectives by others. It isn't always possible to represent the viewpoints by a clean separation, in advance, of examples representing each viewpoint and train a separate model for each viewpoint. We introduce lensing, a mixed-initiative technique to (1) extract 'lenses' or mappings between machine-learned representations and perspectives of human experts, and to (2) generate 'lensed' models that afford multiple perspectives of the same dataset.

• **Astroparse**

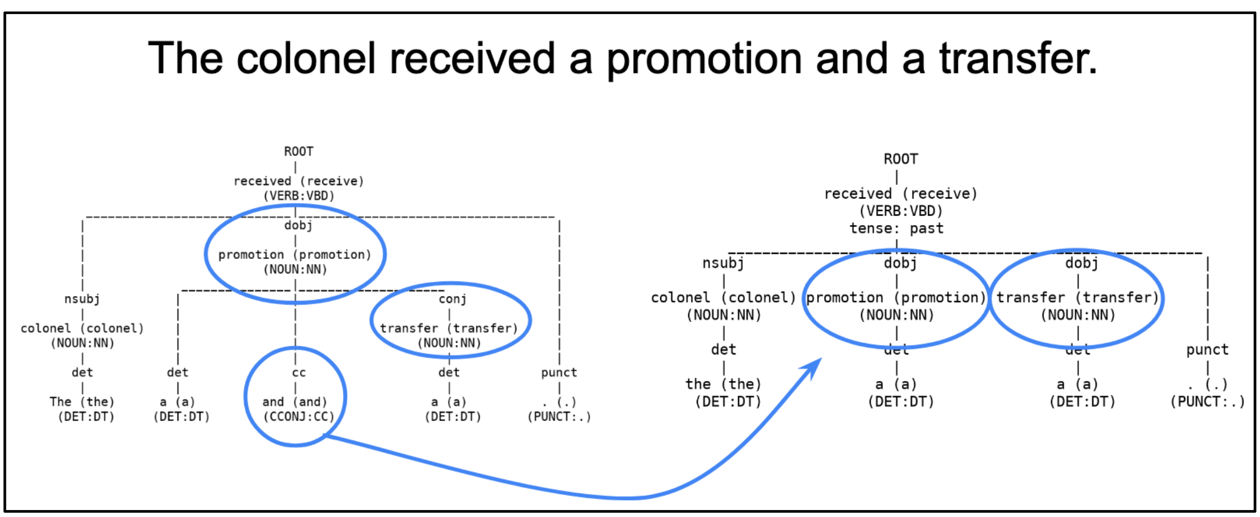
In recent years, advancements have been made in statistical parsing, or techniques which use a large amount of corpus data to build models for analyzing the roles of words in a sentence. State-of-the-art tools such as spaCy can be used to create dependency parses which can then be used for downstream tasks. Dependency parses can be computed quickly and robustly with fairly high accuracy. However, dependency parses can be difficult to work with directly, because they provide a superficial analysis by nature.

We found that we are able to produce ternary expressions using a trainable tree-matching algorithm. During each training step, we present a sentence exhibiting linguistic phenomena of interest together with its corresponding ternary expressions, and the algorithm records properties of the dependency parse subtree where constituents of the ternary expressions occur. When parsing new sentences, our system detects subtree patterns it has seen, and it uses local features of the dependency parse to reproduce the corresponding ternary expressions. By taking a local approach to constructing ternary expressions from dependency parses, our system can parse sentences of arbitrary nesting and length, provided that the underlying dependency parse is accurate. We can thus produce ternary expressions for a larger class of sentences, taking advantage of dependency parsers' robustness. Downstream reasoning systems can use the ternary expressions for indexing and matching made more effective than using raw dependency parses by highlighting structural ternary expressions for the first pass of matching.

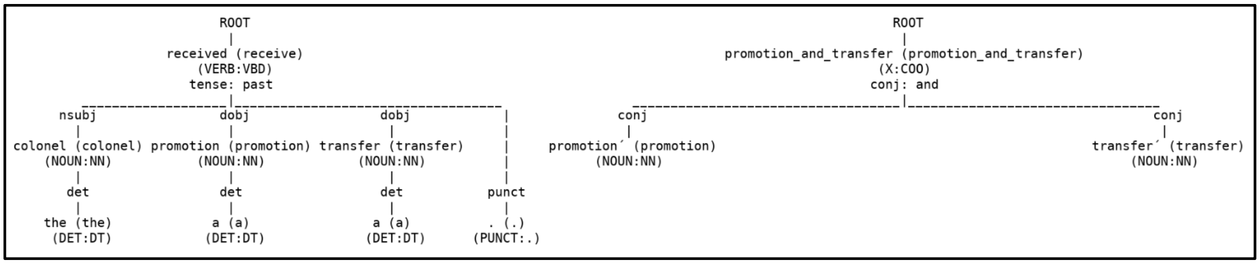
In addition to highlighting important relations, ternary expressions are more effective than dependency parses because they capture many relations that are altogether missed by the latter. We identified various language constructions where dependency parses by nature fail to recognize important relations, including but not limited to:

- Coordinate structures (conjunctions)
- Imperative constructions with covert subject
- Embedded clauses with covert subject
- Relative clauses with overt or covert relativizers

To produce correct ternary expressions for these constructions, our system, prior to detecting subtree patterns in the dependency parse, performs transformations which recover the missing relations. For a detailed example, in the sentence "The colonel received a promotion and a transfer", a dependency parser is inherently unable to relate the information that the colonel "received a transfer", because a superficial analysis only reveals that "transfer" is conjoined to "promotion". When faced with a sentence like this containing coordinate structures, our system traverses the parse to find conjuncts and attach them directly where they should relate to the rest of the parse. Only after deriving important dependency relations do we produce ternary expressions. Our ensuing ternary expression representation can then allow for answering questions such as "Who received a transfer?" (the colonel), or "What did the colonel receive?" (two things: a promotion and a transfer).



To allow the capture of important relations while including enough information to reconstruct a sentence completely, our system also keeps a record of the hierarchy within coordinate structures. We expand the dependency parse representation to allow for multiple roots, some of which connect conjuncts according to their original hierarchy in the dependency parse (for example, in the sentence "The colonel wanted a promotion, or a transfer and a raise", "a transfer and a raise" are grouped more closely than "a promotion or a transfer".)



To expand our system's coverage while maintaining the necessary relations, we also derive missing relationships for many other cases. In imperative constructions, we recover, e.g., the missing "you" in "(you) be careful!" In embedded clauses, we recover, e.g., the relation between "Air Force" and "to expand" in the sentence "The Air Force continues to expand its knowledge of sonic boom." In relative clauses, we recover, e.g., the relation between "doing" and "training" in the sentence "The exercise is validating all of the training the 34th SOS has been doing." In all such constructions, the missing relations are critical for downstream reasoning, as they allow us to answer questions such as "Who should be careful?", "Who is expanding their knowledge?", and "What has the 34th SOS been doing?", respectively.

Publications

Barbu et al. 19
Barbu A, Mayo D, Alverio J, Luo W, Wang C, Gutfreund D, Tenenbaum J, Katz B. “ObjectNet: A Large-scale Bias-controlled Dataset for Pushing the Limits of Object Recognition Models.” The Thirty-third Annual Conference on Neural Information Processing Systems (NeurIPS), 2019.

Berzak et al. 17
Berzak Y, Nakamura C, Flynn S, Katz B. “Predicting Native Language from Gaze.” Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL), 2017.

Berzak et al. 18
Berzak Y, Katz B, Levy R. “Assessing Language Proficiency from Eye Movements in Reading.” Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL), 2018.

Berzak et al. 22
Berzak Y, Nakamura C, Smith A, Weng E, Katz B, Flynn S, Levy R. “CELER: A 365-Participant Corpus of Eye Movements in L1 and L2 English Reading.” Open Mind: Discoveries in Cognitive Science, 2022.

Cheng et al. 21
Cheng E, Kuo Y-L, Cases I, Katz B, Barbu A. “Toward Modeling the Emergence of Symbolic Communication.” ICRA Workshop on Social Intelligence in Humans and Robots (ICRA SIHR), 2021.

Cheng et al. 22
Cheng E, Kuo Y-L, Correa J, Katz B, Cases I, Barbu A. “Quantifying the Emergence of Symbolic Communication.” Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci), 2022.

Colon-Hernandez, Lieberman, Havasi 19
Does a Dog Desire Cake?: Expanding Knowledge Base Assertions Using Relationship Discovery, Pedro Colon-Hernandez, Henry Lieberman, Catherine Havasi, Conference on Neural Information Processing (Neurips), Workshop on Knowledge Representation and Machine Learning, Vancouver, December 2019.

Colon-Hernandez et al. 21
RetroGAN: A Cyclic Post-Specialization System for Improving Out-of-Knowledge and Rare Word Representations. Pedro Colon-Hernandez, Yida Xin, Henry Lieberman, Catherine Havasi, Cynthia Breazeal, Conference of the Association of Computational Linguistics / International Joint Conference on Natural Language Processing (ACL/IJNLP), August 2021.

Colon-Hernandez et al. 22a
Can Language Models Take A Hint? Prompting for Controllable, Contextualized Commonsense Inference, Pedro Colon-Hernandez, Yida Xin, Henry Lieberman, Cynthia Brezeal, submitted to Association of Computational Linguistics (ACL-2022).

Colon-Hernandez et al. 22b
Adversarial Transformer Language Models for Contextual Commonsense Inference, Pedro Colon-Hernandez, Henry Lieberman, Yida Xin, Claire Yin, Cynthia Breazeal and Peter Chin, submitted to the Semantic Web Journal, Special Issue on Commonsense Knowledge and Reasoning, IOS Press, 2022.

Conwell C, Mayo D, Buice M, Katz B, Alvarez G, Barbu A. “Neural Regression, Representational Similarity, Model Zoology & Neural Taskonomy at Scale in Rodent Visual Cortex.” The Thirty-fifth Annual Conference on Neural Information Processing Systems (NeurIPS), 2021.

Dinakar and Lieberman 21
Lensing Machines: Representing Perspective in Latent Variable Models, Karthik Dinakar and Henry Lieberman, Conference on Advances in Cognitive Systems (ACS-21), November 2021.

Fry and Lieberman 18
Why Can’t We All Just Get Along?: How Science Can Enable A More Cooperative Future, Christopher Fry and Henry Lieberman, self-published book, February 2018,

http://www.whycantwe.org/

Hu 22
Transforming dependency parses into ternary expressions for enhanced indexing and matching, Henry Hu, MIT MEng thesis, May 2022.

Idiart et al. 19
Idiart M, Villavicencio A, Katz B, Rennó-Costa C, Lisman J. “How the Brain Represents Language and Answers Questions? Using an AI System to Understand the Underlying Neurobiological Mechanisms.” Frontiers in Computational Neuroscience, Volume 13, 2019.

Katz et al. 18
Katz B, Borchardt G, Felshin S, Mora F. “A Natural Language Interface for Mobile Devices.” In Kent L. Norman and Jurek Kirakowski (Eds.), The Wiley Handbook of Human Computer Interaction, Volume 2, First Edition, 539–559, John Wiley & Sons, 2018.

Kuo 18a
Kuo Y-L, Barbu A, Katz B. “Deep Compositional Models for Robotic Planning and Language.” In Proceedings of 2018 IROS Workshop on Language and Robotics (LangRobo-2018), October 2018.

Kuo et al. 18b
Kuo Y-L, Barbu A, Katz B. “Deep Sequential Models for Sampling-based Planning.” International Conference on Intelligent Robots (IROS), 2018.

Kuo et al. 20
Kuo Y-L, Katz B, Barbu A. “Deep Compositional Robotic Planners that Follow Natural Language Commands.” Workshop on Visually Grounded Interaction and Language (ViGIL) at the Thirty-third Annual Conference on Neural Information Processing Systems (NeurIPS), 2019.

Kuo et al. 21a
Kuo Y-L, Katz B, Barbu A. “Compositional Networks Enable Systematic Generalization for Grounded Language Understanding.” 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2021.

Kuo et al. 21b
Kuo Y-L, Katz B, Barbu A. “Compositional RL Agents that Follow Language Commands in Temporal Logic.” Frontiers in Robotics and AI, section Robot and Machine Vision, 2021.

Kuo Y-L, Huang X, Barbu A, McGill SG, Katz B, Leonard JJ, Rosman G. “Trajectory Prediction with Linguistic Representations.” International Conference on Robotics and Automation (ICRA), 2022.

Lieberman 17
Why Can’t We All Just Get Along?, TED talk, Boston, November 2017.
https://www.ted.com/talks/henry_lieberman_why_can_t_we_all_just_get_along

Lieberman 18a
Wearing Our Hearts on Our Sleeves,
ACM Transactions on Intelligent Interactive Systems,
Blog Editor's post, November 2018.
https://medium.com/acm-tiis

Lieberman 18b
Pay Attention to the Man Behind the Curtain,
ACM Transactions on Intelligent Interactive Systems, Special Issue on Human-Centered Machine Learning, Blog Editor's post, August 2018.
https://medium.com/acm-tiis

Lieberman 18c
Owning the Wind, Invited presentation, Workshop on GDPR (EU regulations on data privacy), Computers and Human Interaction Conference (CHI-18) Montréal, April 2018.

Lieberman 19
In Robots We Trust?, Henry Lieberman,
ACM Transactions on Intelligent Interactive Systems,
Blog Editor's post, February 2019.
https://medium.com/acm-tiis

Lieberman 20
Intrinsic and Extrinsic Motivation in Intelligent Systems, Henry Lieberman, First International Workshop on Self-Supervised Learning, Cambridge, MA, Proceedings of Machine Learning Research 131:62-71, 2020.

Lieberman 21
Mind the (Ethical) Gap, Henry Lieberman, ACM Transactions on Intelligent Interactive Systems, Blog editor's post, May 2021.
https://medium.com/acm-tiis/mind-the-ethical-gap-a85302c1efca

Lieberman, Fry 18
Hippies Values Really Did Build the Internet, Henry Lieberman and Christopher Fry, Communications of the ACM, September 2018. Includes commentary also by New York Times journalist John Markoff, and CACM editors Andrew Chien and Moshe Vardi.

Lieberman, Kuo, Staneva 18
Debugging Probabilistic Programming: Lessons from Debugging Research, Henry Lieberman, Yen-Ling Kuo, Valeria Staneva, 1st Conference on Probabilistic Programming, Cambridge, MA, October 2018.

Lieberman, Williams, Winston 19
Commonsense Reasoning for Story Understanding, Henry Lieberman, Bryan Williams, Patrick Winston, Advances in Cognitive Systems, Workshop on Story-Enabled Intelligence, Cambridge,

MA, August 2019.

Mayo et al. 22
Mayo D, Lu D, Zhang C, Cummings J, Lin X, Katz B, Glass J, Barbu A. “Growing ObjectNet: Adding speech, VQA, occlusion, and measuring dataset difficulty.” ICML Workshop on Shift happens: Crowdsourcing metrics and test datasets beyond ImageNet, Thirty-ninth International Conference on Machine Learning (ICML), 2022.

Myanganbayar et al. 18
Myanganbayar B, Mata C, Dekel G, Katz B, Ben-Yosef G, Barbu A. “Partially Occluded Hands: A Challenging New Dataset for Single-image Hand Pose Estimation.” Asian Conference on Computer Vision (ACCV), 2018.

Netanyahu et al. 21
Aviv Netanyahu, Tianmin Shu, Boris Katz, Andrei Barbu, Joshua B. Tenenbaum. “PHASE: PHysically-grounded Abstract Social Events for Machine Social Perception.” Proceedings of the AAAI Thirty-Fifth Conference on Artificial Intelligence (AAAI), 2021.

Palmer et al. 21
Palmer I, Rouditchenko A, Barbu A, Katz B, Glass J. “Spoken ObjectNet: A Bias-Controlled Spoken Caption Dataset.” The 22nd Annual Conference of the International Speech Communication Association (Interspeech), 2021.

Paul 17
Paul R, Barbu A, Felshin S, Katz B, Roy N. “Temporal Grounding Graphs for Language Understanding with Accrued Visual-linguistic Context.” International Joint Conferences on Artificial Intelligence (IJCAI), 2017.

Pei 21
Augmenting Transformers for Open Domain Procedural Text Comprehension, Yixuan Pei, MIT MEng thesis, September 2021.

Ross et al. 18
Ross C, Barbu A, Katz., B. “Grounding language acquisition by training semantic parsers using captioned videos.” 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2018.

Ross et al. 19
Ross C, Mao C, Katz B, Barbu A. “Learning Language from Vision.” Workshop on Visually Grounded Interaction and Language (ViGIL) at the Thirty-third Annual Conference on Neural Information Processing Systems (NeurIPS), 2019.

Ross et al. 21
Ross C, Katz B, Barbu A. “Measuring Social Biases in Grounded Vision and Language Embeddings.” 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics (HLT/NAACL), 2021.

Solomon, et al. 21
History of Logo, Cynthia Solomon, Brian Silverman, Henry Lieberman, Brian Harvey, Mark L. Miller, Ken Kahn, Margaret Minsky, Artemis Papert, Fourth International Conference on the History of Programming Languages (HOPL4), London, 2021.

Tejwani et al. 21a
Ravi Tejwani, Yen-Ling Kuo, Tianmin Shu, Boris Katz, Andrei Barbu. "Social Interactions as Recursive MDPs." Conference on Robot Learning (CoRL), 2021.

Tejwani et al. 21b
Tejwani R, Katz B, Breazeal C. "Migratable AI: Investigating Users' Affect on Identity and Information Migration of a Conversational AI Agent." International Conference on Social Robotics, 257-267, 2021.

Tejwani et al. 22
Tejwani R, Kuo Y-L, Shu T, Stankovits B, Gutfreund D, Tenenbaum J, Katz B, Barbu A. "Incorporating Rich Social Interactions into MDPs" International Conference on Robotics and Automation (ICRA), 2022.

Wang et al. 20
Wang C, Ross C, Kuo Y-L, Katz B, Barbu A. Learning a natural-language to LTL executable semantic parser for grounded robotics. 4th Conference on Robot Learning (CoRL 2020), Cambridge MA, USA.

Williams et al. 17
Understanding Stories with Large-Scale Commonsense, Bryan Williams, Henry Lieberman, and Patrick Winston, Thirteenth International Symposium on Commonsense Reasoning, London, November 2017.

Xin et al. 21
PatchComm: Using Commonsense Knowledge to Guide Syntactic Parsers, Yida Xin, Henry Lieberman, Peter Chin, International Conference on Principles of Knowledge Representation and Reasoning (KR 21), November 2021.