OASIS[1] Lab Active Projects

Last updated: 9/8/23

This document exists to advertise potential opportunities for getting involved with research in the lab. If you are interested in working with us, please peruse these projects and identify which ones have the best fit with your interests and skillset, then email joelchan@umd.edu with an explanation of your selection, along with relevant information about your qualifications.

Synthesis Infrastructures

Project leads: Jay Patel, Siyi Zhu

Synthesizing prior knowledge of an area of study is a crucial step for scholarly and scientific creativity. However, doing this well is extremely costly, taking much more time and effort than many scholars are able to afford. As a result, many projects are undertaken without a sufficiently strong foundation of synthesis, especially interdisciplinary projects, which require synthesis across many different areas of study! This is a serious barrier to scientific progress. As Allen Newell famously said, "You can't play 20 questions with nature and win!"

We believe we can drastically reduce the overhead of synthesis if scholars have ready access to "semantic publications": scholarly assertions that are directly linked to underlying context and evidence (instead of whole papers). The core issue this project addresses is that creating semantic publications is currently prohibitively tedious and disconnected from normal scholarly workflows. In this project, we ask: How might we make small yet powerful changes to our existing tools for reading, annotating, and notetaking to enable scholars to create these semantic publications as part of their natural workflows, rather than a tedious separate task?

Here are some concrete tasks/ideas we could use help with (there may be more, but this is a good start):

1. Building on some think-aloud protocol observations, design and conduct a large-scale survey of scholars to discover how often and under what conditions scholars are *already* doing work that is necessary for creating a semantic publication. This will help validate our hypothesis that there is enough critical

[1] Open and Sustainable Innovation Systems

mass of interest and practices that we could build on.
Helpful skills: *survey design, basic descriptive statistical analysis*

2. Analyze think-aloud protocols of scholars doing synthesis, to identify events where scholars are attempting to reuse information from papers they've read and try to predict success/failure and time cost based on the ways in which that information was represented (e.g., annotations, notes)
Helpful skills: *basic statistical analysis (e.g., regression), willingness to learn content coding*
3. Design interactions and/or contribute to an open-source software extension for RoamResearch that enables people to write notes in close to prose, and author a formal, shareable discourse graph as a byproduct (read more here).
Helpful skills: *UX design/prototyping/mockups, Datalog/Datomic, Typescript, familiarity with RoamResearch and hypertext*
4. Help create proof of concept classifier that takes a PDF or article as input, and identifies snippets or sections that are likely to be needed later as context for understanding/evaluating/reusing a scholarly claim from the article.
Helpful skills: *programming (esp. in Python), machine learning (e.g., LSTM classifiers), PDF parsing (e.g., GROBID)*
5. Help with a lit review: read and take structured notes on relevant papers on this topic
Helpful skills: *comfort reading scholarly papers, background/familiarity with design concepts and ideas around scholarly communication, metascience, tools for thought, and sensemaking*, *comfort with hypertext notebooks like RoamResearch (or willingness to learn)*

This project has some limited funds for hourly research assistants (although we are actively seeking funding), but we welcome volunteers (if willing) and independent / directed research students.

Background material:

- CSCW poster on beyond itunes for papers
- CommonPlace Essay on Sustainable Authorship Models for Synthesis Infrastructures
- Description of discourse graphs publication model
- Talk at CMU HCII Seminar

Computational Analogy

Project lead: Jason Ding, Yow-Ting Shiue

Drawing analogies to other domains during problem solving is often a powerful source of new ideas. For example, the Wright brothers developed their core insight for their airplane design (the wing warping mechanism that enabled stable steering) based on an analogy to the twisting of a cardboard box!

The problem is that finding analogies "by hand" or "by memory" often fails, due to lack of knowledge or effort. With recent advances in machine learning and natural language processing, and the availability of open repositories of ideas (e.g., patent databases, research paper databases), there is immense potential to develop systems that can automatically find analogies from a problem you're working on to descriptions of ideas in other domains.

We have some working prototype analogy-finding systems that we're refining, and we are now focusing on prototyping systems that real innovators can actually use to help them come up with powerful new ideas.

Here are some concrete tasks/ideas we could use help with (there may be more, but this is a good start):

1. One persistent issue we have is the lack of a quantitative "gold standard" dataset of analogical matches we can use to tune our algorithms. We want to construct this dataset! For example, could we build this dataset by repurposing citation links between papers, crawling stories of analogy-driven innovation, and so on? **Helpful skills**: *text mining, comfort with APIs and web scraping/parsing, willingness to get nitpicky about the definition of analogy*
2. Frequently people miss the value of an analogical match because they can't make sense of it and how it might connect to their problem. But they probably know someone who might be closer to that analogical domain who might be able to help them make sense of it. They might even start a collaboration! Could we prototype a simpler people-matching recommender system for an organization that has people from different disciplines? A person could input a search query, and then the system could find not just matches, but also who in the organization has authored or cited papers that are related to those matches, and then recommend that the solution-seeker talk to them. Would this increase the "yield rate" of analogical matches? Would there be any downstream benefits for the collective intelligence of the organization? To maximize the value/impact of this

work, we'd like to figure out how to integrate AI models that can help with problem reformulation and abstraction, as well as analogical matching, into tools/workflows that innovators already use, such as social media, or chat systems like Slack or Discord.
Helpful skills: *comfort using machine learning models in a program; web APIs (maybe for sending/receiving emails); web development, writing bots/add-ons for Discord/Slack, comfort with applied machine learning (e.g., thoughtfully using transformer language models from huggingface)*

3. Help with a lit review: read and take structured notes on relevant papers on this topic
Helpful skills: *comfort reading scholarly papers, background/familiarity with design concepts and ideas around NLP / language models, analogy, and creative cognition*, *comfort with hypertext notebooks like RoamResearch (or willingness to learn)*

This project has some limited funds for hourly research assistants, but we welcome volunteers (if willing) and independent / directed research students.

Background material:

- PNAS paper on computational analogy
- Talk at Protocol Labs

Democratizing Design

Project lead: Salma Elsayed-Ali

Participatory design has great promise for overcoming serious deficiencies in existing design processes that exclude user voices from the design process. For example, many of the harms of AI/ML systems could be addressed much earlier in the design process if the relevant stakeholders were deeply involved in the design process as equal partners, not just "data".

Unfortunately, many participatory design methods are very costly for both design teams and users, leading to serious limitations in who can participate. This hampers the vision of democratizing design. We want to understand how we might better balance breadth (how many users can get involved) and depth (how much they can participate) in participatory design, and what innovations to methods (e.g., VR/AR techniques, design protocols) might help with this.

Here are some concrete tasks/ideas we could use help with (there may be more, but this is a good start):

1. Help with a lit review: read and take structured notes on relevant papers on this topic
   Helpful skills: *comfort reading scholarly papers, background/familiarity with design concepts and ideas around diversity/inclusion, comfort with hypertext notebooks like RoamResearch (or willingness to learn)*
2. Help with analyzing data from interviews with participatory design experts.
   Helpful skills: *background in qualitative data analysis and open coding*
3. Help with prototyping interactions for an augmented reality remote design tool that allows people to take snapshots of sights/sounds in their homeworld and bring into a shared design space
   Helpful skills: *UX design/prototyping, Unity programming*

This project does not have funds for paid positions right now, but we welcome volunteers (if willing) and independent / directed research students.

Archive

Material Knowledge of ML for Design

Project lead/contact: none at the moment

As machine learning is becoming more ubiquitious, HCI researchers and practitioners are starting to investigate best practices around design methods for building systems with machine learning in them.

One important challenge is understanding the capabilities and limitations of machine learning as a "design material". This goes beyond reducing harm: we also want to connect the possibilities of machine intelligence with novel designs for human-computer/information interfaces!

There is a concept of "material knowledge" from design theory that might prove useful: for example, a furniture designer has a felt, tacit sense of the material properties of the wood she is working with, and can sense what design ideas are better or worse, and draw on this felt sense to think of new design ideas. We're curious if this concept might be useful for understanding what kind of knowledge is required to do good design with ML.

Here are some concrete tasks/ideas we could use help with (there may be more, but this is a good start):

1. Design, conduct, and analyze think-aloud observations and semi-structured interviews with expert ML practitioners to probe whether they have this "material knowledge" of the machine learning models they work with.
   Helpful skills: *background/familiarity with design concepts and ML, think-aloud and interview methods*

This project does not have funds for paid positions right now, but we welcome volunteers (if willing) and independent / directed research students.

Solution Diversity

Project lead: Joel Chan and Jason Ding

We often hear that being exposed to diverse solution examples can be helpful for creativity during design. But empirical studies of this have been quite mixed.

We're investigating how we might formally model the conditions under which diverse initial examples leads to (how much) creativity gains: for example, how might the type of problem, or problem-solving habits of the designer matter? We also want to make sure this model describes reality, so we want to compare the model's predictions to human data under the same conditions (e.g., solving similar problems).

Here are some concrete tasks/ideas we could use help with (there may be more, but this is a good start):

1. Help run some online experiments (e.g., on Amazon Mechanical Turk or similar) to get benchmark human performance data to compare with our model's predictions.
   Helpful skills: *web programming (Python/React stack), familiarity with crowdsourcing platforms, experiment design, advanced statistical analysis (e.g., mixed effects models)*
2. Help with a lit review: read and take structured notes on relevant papers on this topic
   Helpful skills: *comfort reading scholarly papers, background/familiarity with design concepts and ideas around creativity support tools and creative cognition*, *comfort with hypertext notebooks like RoamResearch (or willingness to learn)*

This project has some funds for a graduate assistantship position if there is an exceptional fit. We also welcome volunteers (if willing) and independent / directed research students.