Integrating Planning and Recognition to Close the Interaction Loop

Richard G. Freedman
College of Information and Computer Sciences
University of Massachusetts Amherst
freedman@cs.umass.edu

Introduction

In many real-world domains, the presence of machines is becoming more ubiquitous to the point that they are usually more than simple automation tools for one-way interaction. As part of the environment amongst human users, it is necessary for these computers and robots to be able to interact back reasonably by either working independently around them or participating in a task, especially one with which a person needs help. Such interactions are now everywhere ranging from robots around homes and factories to virtual agents in mobile devices, video games, and automated dialogue systems. While interactive robots and computer systems have been implemented for various domains, most are specifically designed for a given domain such as industrial robotics (Levine and Williams 2014; Wurman, D'Andrea, and Mountz 2007), elderly care (Schwenk, Vaquero, and Nejat 2014; Fasola and Matarić 2013), etc. Just as there are domain-independent heuristics that can effectively find optimal solutions for any classical planning problem (Hoffmann and Nebel 2001; Helmert 2006), I introduce a domain-independent approach to performing interaction based on the integration of several research areas in artificial intelligence, particularly planning and plan recognition.

This interactive procedure requires several steps performed indefinitely as a loop: recognizing the user and environment from sensor data, interpreting the user's activity and motives, determining a responsive behavior, beginning to perform the behavior, and then recognizing everything again to confirm the behavior choice and replan if necessary. At the moment, the research areas addressing these steps, activity recognition, plan recognition, intent recognition, and planning, have all been primarily studied independently. However, pipelining each independent process can be risky in real-time situations where there may be enough time to only run a few steps. This leads to a critical question: *how do we perform everything under time constraints*? In this thesis summary, I propose a framework that *integrates these processes* by taking advantage of features shared between them. This includes my current work towards this preliminary system and outlines how I plan to complete the integration for a time-constrained interaction loop.

Background

One of the earliest areas of artificial intelligence, *planning* is the study of automated action selection. Early approaches usually involved representing the world as a list of logic statements and searching for a sequence of actions which would modify the list until it contained the set of goal conditions; the notation used for this is called *STRIPS*. Modern approaches range from improving search over STRIPS to decision theoretic planning with MDPs and its variants to approximation methods to handle uncertainties in the world.

As its inverse problem, *plan recognition* (PR) tries to identify the problem an agent is solving given its observed actions. The actions and problems are usually represented at a higher level such as STRIPS. *Activity recognition* (AR) works at the lower level by interpreting sensor data as higher-level actions. In addition to predicting current activity, *intent recognition* (IR) tries to predict the agent's specific goal or upcoming actions which allows some degree of foresight into the observed agent's behavior. Collectively, these fields of recognition are referred to as *PAIR* and have become a more popular area of research recently, including the topic of a Dagstuhl Seminar (Goldman et al. 2011). Various problems in PAIR are studied in other areas, sometimes under different names, making the literature vast, but they are still studied largely independently or pipelined in most these works.

One notable work which integrated plan and activity recognition was by Levine and Williams (2014) where, for a given plan with branching points based on a human's choice of actions, a robot would select actions to resolve broken causal links resulting from the human's action choice(s). Our approach differs from this work because the given plan provides instructions for the human to follow, but we do not restrict the human with directions. This assumed plan is fine for their intended factory domain, but insufficient for general interaction in any domain. Daily tasks may be intertwined over time or contain noise such as answering a ringing telephone while cooking or cleaning; such domains cannot assume that a human will follow a protocol.

Simultaneous Plan and Activity Recognition

The formulation of a typical recognition problem is as follows: given a sequence $\mathcal{O}$ of observations $o_1, o_2, \ldots, o_n$, determine which task(s) in library L the agent is performing. In AR, each o_i is a sensor reading and L is a set of actions or activities. Supervised machine learning and graphical models are usually used to infer the label in L which best describes $\mathcal{O}$. For PR, each o_i is a STRIPS action and L contains the

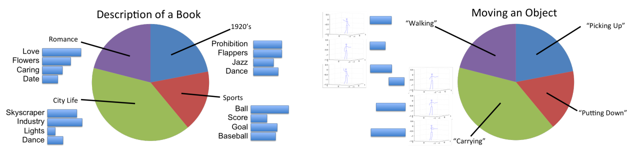

Figure 1: Analogy of Distributions for Topics and Actions

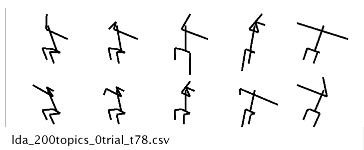

Figure 2: Learned Cluster of Postures Resembling Sitting

set of all tasks. A weighted matching method is often used to compare $\mathcal{O}$ to some of each task's solutions called *plans*, a sequence of STRIPS actions that satisfies the goal conditions. A recent method used to perform this matching is parsing hierarchical task networks, revealing a *parallel between PR and natural language processing* (NLP) (Geib and Steedman 2007). In particular, $\mathcal{O}$ is like a sentence and the breakdown of a complex task into subtasks is like a grammar of valid derivations.

I began the extension of this analogy by considering another problem in NLP: topic modeling. Unlike parsing which determines the underlying structure of a sentence, topic modeling investigates the concepts discussed in a collection of documents by finding clusters of related words called *topics*. For the popular Latent Dirichlet Allocation (LDA) topic model (Blei, Ng, and Jordan 2003), which was previously used for AR by Huỳnh, Fritz, and Schiele (2008), each topic $t \in T$ is simply a distribution over the set of words V and each document $d \in D$ is a distribution over the set of topics T; the respective distributions $\phi_{t \in T} : V \rightarrow [0, 1]$ and $\theta_{d \in D} : T \rightarrow [0, 1]$ are learned using unsupervised learning to model the training data. For interaction in a variety of domains, an unsupervised approach is more appealing because it is possible to learn a large number of activities in each domain as they are added to the system. Furthermore, LDA's *bag-of-words assumption* where each o_i is independent of the rest of $\mathcal{O}$ was appealing to begin integrating AR and PR due to the mismatch of the two sequence formations. A sensor records over time so that a single action has multiple consecutive o_i for AR, but a single STRIPS action is only one o_i for PR. Figure 1 illustrates how this analogously treats actions like topics. Each word is a sensor reading and the distribution over these readings describes an action while the task of the recording session is represented by the distribution of actions. *Hence inferring a topic with LDA performs AR and the distribution of the collection of inferences enables us to approximate PR simultaneously*. The results from running LDA on a small dataset provided evidence supporting my hypothesis since each topic contained postures resembling simple actions as in Figure 2 (Freedman, Jung, and Zilberstein 2014).

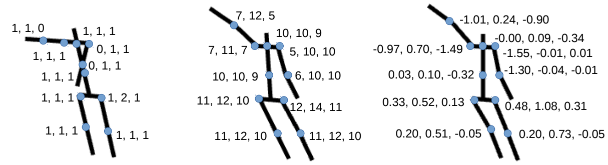

Figure 3: A posture reaching up towards a shelf with granularity 3 (left) and 21 (middle) from the original angles (right).

Representation of RGB-D Sensor Readings For robotics applications, Red, Green, Blue - Depth (RGB-D) sensors have become one of the most common sensors to use due to their present availability, affordability, and growing popularity. They provide a camera image with an overlaid infra-red scan to obtain three-dimensional point clouds of their surroundings, which aids segmentation and associating individual components' directions of motion (Herbst, Ren, and Fox 2013; Zhang, Nakamura, and Kaneko 2015) including a person's entire body as a stick figure rendering (Shotton et al. 2011). Because LDA is designed for inputs whose frequencies resemble words in a natural language, it is important to treat these stick figures in a similar manner. The process of converting data into a form for natural language methods is wordification (Perovšek et al. 2013).

The human posture is composed of fifteen joints via ten rotation matrices; though it is sufficient to create a "vocabulary" of postures as the concatenation of the rotation matrices' Eulerian transformations (roll, pitch, and yaw), the frequency of postures is not desirable. In particular, each wordified posture is in $[-\pi, \pi]^{30}$, which is an uncountably infinite vocabulary with a very small likelihood of duplicates. Zipf's Law states that not only should there be duplicate words in any natural language, but that the frequency of the i^{th} most frequent word is about twice that of the $(i+1)^{\text{th}}$ most frequent word. Hence we make the vocabulary finite and increase the likelihood of duplicates by *discretizing the space* with respect to a *granularity* parameter $g \in \mathbb{N}$, mapping each angle to $[0, g) \cap \mathbb{Z}$ as illustrated in Figure 3. The reduced vocabulary $\{0, 1, \ldots, g-1\}^{30}$ is still large in cardinality for small g, but many poses do not represent feasible body structures. For example, the limitations of each joint's range of motion will not form postures with hyperextended limbs, much the same as many combinations of orthographic letters do not form actual words in a language. Our initial investigation of g showed that increasing it reduces the number of duplicate words and that odd g have more duplicates than even g due to small body motions about the origin.

Extensions for Additional Features I extended LDA to a generative model that also considers the presence of nearby objects and/or global temporal patterns. The set of objects with which the user can interact are represented using a second vocabulary in a separate LDA model, but it shares the same action/topic as the posture at each respective timestep. The temporal relation is captured in the composite topic model (Griffiths et al. 2004) by embedding LDA as a state within a hidden Markov model (HMM), enforcing a syntactical structure with the HMM where *one state is a 'blank' for semantic words/postures/objects derived by LDA*. The improved log-likelihoods of observed task executions with our new topic model variations serve as evidence that the information provided by these two factors are not only independent, but assist disambiguating actions that contain common postures (Freedman, Jung, and Zilberstein 2015). In future work, I will investigate additional variations and their insights for PAIR.

Integration of Planning with Plan Recognition

Ramírez and Geffner (2010) introduced a compilation of PR problems into classical planning problems. It assigned a distribution over sets of goal conditions $\mathcal{G}$ instead of over precomputed plans; they refer to this generalization as a domain instead of a library. Bayes's Rule compares the compiled classical planning problems for each entry of the domain against each other using the most optimal plans with and without $\mathcal{O}$ as a subsequence. This accounts for the probability of the agent solving each task conditioned on its observed actions, considering optimal (shorter) plans to be more likely. While the accuracy for the method is very strong, a temporal plot of the probabilities showed that it only achieved this accuracy towards the *completion of the plan* when the final actions were observed.

While their compilation is excellent when the plan's completing actions are observed, this is not as practical in interactions because the observed agent will likely be almost finished executing a plan by the time the machine can respond. In collaboration with Fukunaga (University of Tokyo), I have proposed two approaches to address this (Freedman and Fukunaga 2015). The first one generates a *dynamic prior* for Bayes's Rule that removes the bias for shorter plans and converges to the true prior over time. The second one counts the number of linearizations of a partially-ordered plan in order to account for the number of optimal plans rather than their length alone.

The updated distribution over $\mathcal{G}$ can be used to aggregate the lists of logic statements which must be true for each goal and *identify the most necessary conditions*. If a set of conditions is shared between the most likely tasks, then satisfying them should be required regardless of the task. Thus a second pass of the planner on a variation of the original classical planning problem should yield a plan that the machine may execute to interact properly even if the recognized task is still ambiguous. In order to consider potential coordination between the observed agent and the machine, we assume that the response problem is a centralized multi-agent planning problem and that the observed agent will perform its actions assigned from that plan. In reality, the agents are decentralized and this synchronization will likely not hold; thus the interaction loop begins again with the recognition steps to determine how the observed agent reacts.

Status of Thesis

I plan to close the interaction loop by completing the integration of these processes. Besides continuing the works above, there are several key remaining tasks. The most important one is bridging the gap between simultaneous PR and AR and the integration of PR with planning. Although it seems trivial because both contain PR, they do not align due to the unsupervised nature of topic models. Recognized actions are clusters of postures (or other forms of sensor data) without annotation while actions in the newer research are assumed to be in STRIPS which is designed by humans. I have begun to identify methods for autonomously extracting features of the postures with the greatest probability mass in each cluster and using them to describe the respective action (Freedman and Zilberstein 2015). In addition to applying this automated feature extraction process to sensor data, I am exploring analogies in topic modeling for natural language data with Wallach (Microsoft Research).

I am also considering the application of constraint optimization to align LDA clusters (the recognized actions) with STRIPS operators using ordering of each $\mathcal{O}$ and these extracted features. Due to the frequency of observations for sensors compared to higher-level actions, $|\mathcal{O}_{\text{sensor}}| \geq |\mathcal{O}_{\text{STRIPS}}|$. So there is not necessarily a bijective mapping between the two sequences; however, a single STRIPS action should be associated with a particular subset of LDA clusters for its respective postures. This means we can evaluate a constraint optimization problem of the following form:

- Assign each variable $s_i \in \mathcal{O}_{\text{sensor}}$ a value $p_j \in \mathcal{O}_{\text{STRIPS}}$
- Preserve sequence ordering with constraints $s_1 = p_1$, $s_{|\mathcal{O}_{\text{sensor}}|} = p_{|\mathcal{O}_{\text{STRIPS}}|}$, and $s_i = p_j \Rightarrow s_{i+1} \in \{p_j, p_{j+1}\}$
- Ensure that each STRIPS action is associated with a limited subset of LDA clusters with minimization constraints $\min \text{Var}(\tau_a)$ for each STRIPS action a where $\tau_a : T \rightarrow [0, 1]$ is a probability distribution over the LDA clusters to which observations of a have been assigned.

Due to its higher-level representation, breaking a single STRIPS action into smaller subactions like a hierarchical task network (HTN) (Erol, Hendler, and Nau 1994) may facilitate the alignment process with additional constraints for each subaction. However, it would also be necessary to perform additional search to find the correct HTN breakdown for the alignment. This introduces new challenges of identifying *heuristics to evaluate snapshots of optimality*. A visualization of our extended generative model for recognition and this HTN alignment is shown in Figure 4.

The integration of other components such as planning and execution have previously been studied in such areas as metareasoning (Russell and Wefald 1991; Zilberstein 2011). After implementing and testing the proposed integrations of recognition and planning, it will be ideal to integrate IR to better predict upcoming actions so that the machine does not interfere with the observed user. For this, I intend to investigate the planning graph (Blum and Furst 1997) and determine how to probabilistically select action nodes which are more likely to be executed. Once these are all in tact, the preliminary interaction loop will be complete and optimization will be necessary to make it usable under realistic time constraints. For example, the work that currently integrates planning and plan recognition runs a classical planner $2|\mathcal{G}|$ times from the same initial state to identify all the plans for

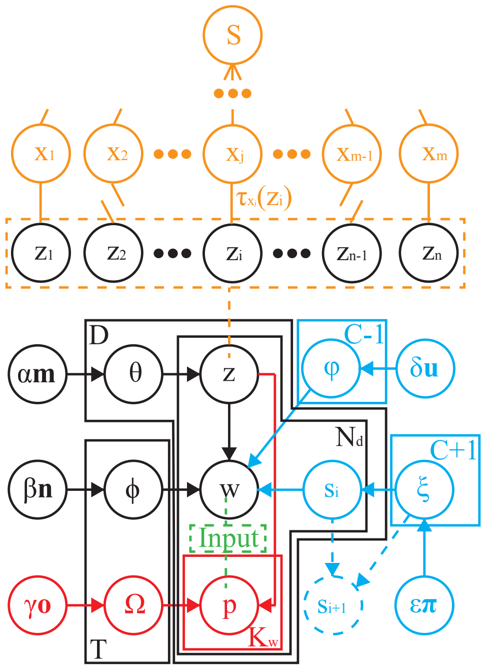

Figure 4: Layout of integrated plan and activity recognition generative model with wordified sensor inputs and objects (green), LDA (black), extension for objects (red), HMM for the temporal extension (blue), and aligned HTN (orange) where S is the sequence of STRIPS actions to break down.

the Bayes's Rule computation, but there is research on finding multiple goals in a single heuristic search so that the classical planner only needs to run one time (Davidov and Markovitch 2006). This avoids redundant expansion of the state space.

References

Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent Dirichlet allocation. *Journal of Machine Learning Research* 3:993–1022.

Blum, A. L., and Furst, M. L. 1997. Fast planning through planning graph analysis. *Artificial Intelligence* 90(1-2):281–300.

Davidov, D., and Markovitch, S. 2006. Multiple-goal heuristic search. *Journal of Artificial Intelligence Research* 26:417–451.

Erol, K.; Hendler, J.; and Nau, D. S. 1994. HTN planning: Complexity and expressivity. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, 1123–1128.

Fasola, J., and Matarić, M. J. 2013. Socially assistive robot exercise coach: Motivating older adults to engage in physical exercise. In Desai, J. P.; Dudek, G.; Khatib, O.; and Kumar, V., eds., *Experimental Robotics*, volume 88 of *Springer Tracts in Advanced Robotics*. Springer International Publishing. 463–479.

Freedman, R. G., and Fukunaga, A. 2015. Integration of planning with plan recognition using classical planners (extended abstract). In *Proceedings of AAAI 2015 Fall Symposium on AI for Human-Robot Interaction*, 48–50.

Freedman, R. G., and Zilberstein, S. 2015. Automated interpretations of unsupervised learning-derived clusters for activity recognition. In *Workshop on Learning for Human-Robot Collaboration*.

Freedman, R. G.; Jung, H.-T.; and Zilberstein, S. 2014. Plan and activity recognition from a topic modeling perspective. In *Proceedings of the Twenty-Fourth International Conference on Automated Planning and Scheduling*, 360–364.

Freedman, R. G.; Jung, H.-T.; and Zilberstein, S. 2015. Temporal and object relations in unsupervised plan and activity recognition. In *Proceedings of AAAI 2015 Fall Symposium on AI for Human-Robot Interaction*, 51–59.

Geib, C. W., and Steedman, M. 2007. On natural language processing and plan recognition. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence*, 1612–1617.

Goldman, R. P.; Geib, C. W.; Kautz, H.; and Asfour, T. 2011. Plan Recognition – Dagstuhl Seminar 11141. *Dagstuhl Reports* 1(4):1–22.

Griffiths, T. L.; Steyvers, M.; Blei, D. M.; and Tenenbaum, J. B. 2004. Integrating topics and syntax. In *Proceedings of the Eighteenth Annual Conference on Neural Information Processing Systems*, 537–544.

Helmert, M. 2006. The fast downward planning system. *Journal of Artificial Intelligence Research* 26(1):191–246.

Herbst, E.; Ren, X.; and Fox, D. 2013. RGB-D flow: Dense 3-D motion estimation using color and depth. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 2276–2282.

Hoffmann, J., and Nebel, B. 2001. The FF planning system: Fast plan generation through heuristic search. *Journal of Artificial Intelligence Research* 14:253–302.

Huỳnh, T.; Fritz, M.; and Schiele, B. 2008. Discovery of activity patterns using topic models. In *Proceedings of the Tenth International Conference on Ubiquitous Computing*, 10–19.

Levine, S., and Williams, B. 2014. Concurrent plan recognition and execution for human-robot teams. In *Proceedings of the Twenty-Fourth International Conference on Automated Planning and Scheduling*, 490–498.

Perovšek, M.; Vavpetič, A.; Cestnik, B.; and Lavrač, N. 2013. A wordification approach to relational data mining. In Fürnkranz, J.; Hüllermeier, E.; and Higuchi, T., eds., *Discovery Science*, volume 8140 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg. 141–154.

Ramírez, M., and Geffner, H. 2010. Probabilistic plan recognition using off-the-shelf classical planners. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, 1121–1126.

Russell, S. J., and Wefald, E. 1991. Principles of metareasoning. *Artificial Intelligence* 49(1-3):361–395.

Schwenk, M.; Vaquero, T.; and Nejat, G. 2014. Schedule-based robotic search for multiple residents in a retirement home environment. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2571–2577.

Shotton, J.; Fitzgibbon, A.; Cook, M.; Sharp, T.; Finocchio, M.; Moore, R.; Kipman, A.; and Blake, A. 2011. Real-time human pose recognition in parts from a single depth image. In *Proceedings of the Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition*, 1297–1304.

Wurman, P. R.; D'Andrea, R.; and Mountz, M. 2007. Coordinating hundreds of cooperative, autonomous vehicles in warehouses. In *Proceedings of the Nineteenth National Conference on Innovative Applications of Artificial Intelligence*, volume 2 of *IAAI'07*, 1752–1759. Vancouver, British Columbia, Canada: AAAI Press.

Zhang, B.; Nakamura, T.; and Kaneko, M. 2015. Adaptive fusion of multi-information based human identification for autonomous mobile robot. In *Proceedings of the Twenty-fourth IEEE International Symposium on Robot and Human Interactive Communication*.

Zilberstein, S. 2011. Metareasoning and bounded rationality. In Cox, M., and Raja, A., eds., *Metareasoning: Thinking about Thinking*. Cambridge, MA, USA: MIT Press. 27–40.