Understanding the Unforeseen via the Intentional Stance

Stephanie Stacy, Alfredo Gabaldon, John Karigiannis, James Kubricht, Peter Tu
GE Research
{firstname.lastname, tu}@ge.com

Abstract

We present an architecture and system for understanding novel behaviors of an observed agent. Our approach uses analogy with past experiences to construct hypothetical rationales that attempt to explain the behavior of an observed agent. Moreover, we view analogies as partial; thus multiple past experiences can be blended to analogically explain an unforeseen event, leading to greater inferential flexibility. We argue that this approach results in more meaningful explanations of observed behavior than approaches based on surface-level comparisons. A key advantage of behavior explanation over classification is the ability to i) take appropriate responses based on reasoning and ii) make non-trivial predictions that allow for the verification of the hypothesized explanation. We provide a simple use case to demonstrate novel experience understanding through analogies in a gas station environment.

1. Introduction

Over the past decade, significant progress has been made in the development of computer vision systems which can detect and identify objects and events in real world video feeds. However, a computer’s understanding of the environment does not exhibit the depth and nuance of that formed by even a naïve human observer. Consider, for example, a common environment encountered often in daily life: the gas station. Video analytics systems can be installed in this environment to identify customers in the scene, critical landmarks, and even events that are relevant to the station’s operations (e.g., customers departing without removing fuel nozzles). However, if the station were to function without a human attendant, computer vision systems would also need to accurately detect and interpret events that are surprising or unanticipated. These situations generally arise due to a wide variety of opportunistic and intentional human behavior. The autonomous attendant of the future – when exposed to novel and unanticipated events – must be able to “get the gist of things” and respond appropriately to ensure that customers are safe and operations proceed in a desired manner. To establish such a capability, we argue that the autonomous attendant must understand the possibly free floating rationale behind such intentional behaviors.

1.1. The Intentional Stance

In [1] Dennett makes the argument for the intentional stance. Throughout our evolutionary history agents have developed various strategies and tactics for survival. Such methods were often the product of environmental reinforcement and as such may be viewed as some sort of hard-wired policy. In some cases such policies were handed down from generation to generation as a form of cultural or genetic inheritance. Thus in many cases the agents themselves may have had little to no understanding of why they did what they did, they just did it.

In a competitive environment, there is great advantage in having the capacity to predict what a rival agent might do. One may attempt to reverse engineer the hardwired policies embedded in the neural networks of a competitor, but such efforts are time consuming and may require extensive observation and experimentation. An alternative approach is to take the intentional stance. In this paradigm one assumes that the observed agent is a rational entity with desires, beliefs and the capacity for reason. Understanding the behavior of an agent then reduces to inferring the rationale behind the actions taken by the agent. Even if the agent is unaware of the reasons behind its actions, the rationale can still be thought of as existing, it is simply a property of the environment from which the agent evolved. In such cases the rationale is said to be free floating.

In this paper we consider the means by which the intentional stance can be taken. We propose that a form of analogical reasoning be used to discern the rationale behind an agent’s actions with as few as a single observation of such behaviors.
new novel experiences are obtained by observing an actor agent whose behavior – in accordance with the intentional stance – is assumed to have been motivated by some unobservable rationale. When constructing a rationale for novel observations, the observer does not know the underlying policy of the actor nor its state of mind; they can only observe the performed sequence of actions. The observed sequence of actions, alongside relevant background facts about the environment and the agent, are represented as a set of propositions encoding the target experience. We then compare the target against a library of base experiences for which the rationales are known. Structural Mapping Theory (SMT) \cite{5,6} offers a theoretical account of analogical mapping positing that psychological concepts are structured, and that analogies are made by comparing the structural similarity of relations and systems of relations contained in concepts that are known (base) and new (target). This theory has been supported by a body of empirical work (see \cite{7} for review) and developed into a computational model called the Structural Mapping Engine (SME) \cite{3,4}. Here, we leverage the SME framework to go beyond its traditional use of mapping relations to extrapolating rationales in novel scenarios.

2. Related Work

In AI, the field of plan, activity, and intention recognition aims at developing techniques for understanding the actions of an observed agent. Some approaches rely on pre-defined plan libraries as the hypotheses space, while others recast the recognition problem as a planning problem and use planning systems to solve it. Sukthankar et al.’s book \cite{12} is an excellent overview of the area.

Another related area is known as Inverse Reinforcement Learning \cite{11,10}, where the aim is to infer the observed agent’s reward function from its behavior, its sensory input, and a world model typically defined as a Markov Decision Process.

Closely related and with a cognitive science perspective, in this paper, we consider essentially the same problem while focusing on a version of it where the observed behavior is new to the observer and hence cannot be recognized by mapping it to one of the previously seen plans or by constructing a reward function in terms of known state and action features. As far as we know, unforeseen behavior recognition has received limited attention in the past, and our approach based on the intentional stance and inferring the rationale behind novel behavior through analogical reasoning is new.

3. Technical Description

Our work on behavior understanding under the intentional stance begins with the use of a Reinforcement Learning (RL) simulation system we call the Artificial Intentional-
ity Engine (AIE). The purpose of the AIE is to construct a set of test policies from which we hope to then construct various rationale that can be used to explain such polices. A description of the AIE is given in Section 3.1. In section 3.2 we describe how an SME framework can be used to construct an analogy between an observed behavior and a single prior experience where the rationale is already understood. We show that via such an analogy a rationale can then be constructed for the observed behavior. Finally, a methodology for considering multiple analogies between a variety of prior experiences can be used to synthesize a composite rationale for the observed behavior. In Section 4 experimental results are presented and comparisons are made with an approach based on Large Language models.

3.1. Artificial Intentionality Engine

For the purposes of generating behavior based policies, a simulation system known as the Artificial Intentionality Engine (AIE) has been developed. Within the AIE are a set of objects with various properties. The Agent has access to a set of actions which allows for travel between objects, transportation of objects and transformation of objects. Note that transformations may require the use of other collocated objects which operate as tools. Using reinforcement learning, an agent can develop policies for accomplishing various tasks.

A synopsis of five policies that have been constructed for experimental purposes is now given. Slumber involves an agent transitioning from a walking to a resting state due to fatigue. This transformation requires a bed to occur since it affords a place of rest. Dinner involves an agent consuming a chicken due to hunger. The concept of tools/utility is critical here since the agent must find a knife, transport the knife to the chicken and then dispatch it prior to consumption. Chopping involves an agent utilizing either a knife or an axe to transform a tree into lumber (two distinct patterns are observed, one for each tool). The concept of danger is relevant since transporting knives often leads to injury. Competition, like consumption, requires an agent to use a knife to consume a chicken, however there exists an animal in the environment which will consume the chicken immediately if it observes the agent picking up the knife. The agent must therefore dispatch the animal prior to attempting to consume the chicken. The rationale behind this behavior is similar to that of a cuckoo bird which must dispatch other entities in its nest to ensure its own survival. Weather requires the agent to discover the effects of the environment on its behavior. If the weather is good the agent may seek leisure, but if the weather is bad the agent must seek shelter.

In the following example, we show how multiple observations of a given behavior (in this case slumber) can be produced by exercising a learned policy. Initial observations are a series of state specifications represented as a set of predicates. These predicates define i) the state of each object, ii) the proximity of each object to other objects, and iii) which objects are currently held by the agent. An initial state specification is randomly generated, and the state of the environment at the conclusion of each action is also made available to the agent (see Figure 1). Three chronologies were constructed for this example, where a chronology embodies predicates that have changed from false to true or from true to false. Note that all but the first chronology contains spurious events (see Figure 2). Such chronologies represent the observation of a target experience. In the next section we consider methods for drawing an analogy between a target and a base experience.

3.2. Analogy Building under the Intentional Stance

SME aims to find the largest structurally consistent mapping between a chosen base experience in a library of understood experiences \( b \in B \) and novel target experience \( t \). Experiences are represented as a set of observations (expressions) over items (entities) that can include attributes, functions, and relations in predicate calculus form. Critically, SME focuses on structural analogies concerning relations (e.g. blood vessels are aqueducts) rather than surface level attributes (e.g. his face was red as a beet). As a result, analogies can be drawn between concepts that have different attributes but share structural similarities. More precisely, attributes are truth valued predicates with one input object, e.g. \( (\text{redAttr Beet}) \); functions also take one argument, but instead of being truth valued, they range over symbols representing objects or quantities, e.g. \( (\text{size Sun}) \); and relations are represented by predicates that take in multiple arguments, e.g. \( (\text{contains Blood, Veins}) \) which may...
themselves be functions, attributes, or even other relations, e.g. \((\text{greaterThan} \ (\text{mass} \ \text{Sun}) \ (\text{mass} \ \text{Planet}))\).

Mapping occurs in four steps: Local matching, Global matching, Candidate inference generation, and Structural evaluation, (see [4] for more algorithmic details). 1.) Local matching: Mapping first occurs locally over pairwise expressions \((b_j, t_k)\) contained in \(b\) and \(t\) according to a set of allowable match construction rules, \(M\text{Crule}(b_j, t_k) \in \{\text{True}, \text{False}\}\), in order to construct a set of match hypotheses: \(MH = \{b_j \in b, t_k \in t | M\text{Crule}(b_j, t_k)\}\). Here the traditional SME match rules are used (see [4]. Appendix A) which ignore properties or attributes of objects, but identically match relations such as \(\text{greaterThan}\) or \(\text{cause}\) and extend this to flexibly match predicates within a category (e.g. predicates labeled as “affordances”). This extension is well suited to mapping rationales because mapped analogies with the same structure may be indicative of a generalizable narrative of observed events, which can be understood by looking at category level structure. 2.) Global matching: These local matches are then combined using a greedy merge algorithm into maximally consistent global mappings (Gmaps) subject to two structural constraints. First, mapping must be one-to-one: an entity in the base and target cannot correspond with at most a single entity in the target and vice-versa. Second, supports must map: for \(MH(b_j, t_k)\) \(\subseteq MH\), any nested expressions within \(b_j\) and \(t_k\) must also map. It is possible for a base and target to have multiple Gmaps. 3.) Candidate inference generation: From each Gmap, expressions connected to the Gmap in the base that do not have matching expression in the target but are structurally consistent with the rest of the relations in the current Gmap are added as candidate inferences about the target. 4.) Structural evaluation: In the last step Gmaps are scored depending in both the number of expressions matched and the depth of the relations mapped. Each type of mapping (e.g. function) is given a weight and these scores are summed across all matches within the Gmap. Better mappings, and thus better analogies, exhibit larger degrees of higher order relations. The best Gmap is then presented as the proposed analogy map.

While inferences made using SME traditionally concern the underlying causal mechanism of phenomena, we are interested in the rationale behind a sequence of observed events produced by one or more agents. Thus, we have repurposed the SME algorithm to take in scenarios involving action and event trajectories generated by observing an agent(s) in a simulated or real world as opposed to solely mapping items and relations. This lets us integrate analogical reasoning with the AIE. Starting from a library of base experiences for which we understand the rationales, we can use SME to hypothesize rationales and understand novel experiences.

Here we demonstrate SME-based rationale mapping for two policies generated by the AIE, slumber and wood chopping. For Slumber, we assume that the agent has an existing base experience for the rationale behind traveling to a bonfire when cold: when the agent travels to the fire she then becomes comfortable and knows the reason is because the fire affords properties such as warmth and brightness and the agent is cold (an unfulfilled desire). When this prior experience maps to the new chronology of slumber generated by the AIE, the Bed maps to Fire; asleepTf, a transformation, maps to the predicate of comfortableTf; tiredDes, an unfulfilled desire, maps to coldDes; and affordances such as flatAff and softAff map to warmAff and brightAff. From this, new inferences can be made connecting the unfulfilled desires to affordances as well as rationale behind observed behaviors (see Figure 1).

Similarly, for the Chopping example, given known information in our base about the ability of a hammer and rock to both pound in a nail, we can make a full analogy to chopping wood using an axe versus a knife. Here, Axe is mapped to Hammer and Knife is mapped to Rock. The key aspect of this analogy is that, while both tools achieve the desired effect – transforming (poundedTf Nail) and (choppedTf Wood), there is an advantage (represented as the relation advantage to one choice over the other in terms of a property of the items, forceFn for the hammer and safetyFn for the axe). We are also able to infer the causal nature of the relationships and elements of the rationale, from the mapping (see Figure 4). Rationales for dinner, weather, and competition were similarly successfully mapped to base experiences of unlocking a door, day/night, and a pathogen in a host respectively but not shown here for brevity.

3.3. Iterative Rationale Construction

Except for the simplest cases, novel experiences often require multiple analogies with past experiences to construct a rationale that fully explains the observations. The objective is then to hypothesize a rationale for a new experience by collecting hypotheses drawn by analogy with
multiple past experiences which together form a rationale for the novel experience. This objective drives the idea of analogy synthesis by combining partial hypothesized rationales distributed over different experiences via sequential pairwise analogy evaluation. As described in Section 3.2, the SME constructs a mapping between a base $b$—a prior experience—and the target $t$—the novel situation. Given a set of bases $B = \{b_1, b_2, \ldots b_n\}$, corresponding to $n$ past experiences, the idea is to attempt a pairwise mapping $b_i \rightarrow t$ for each $b_i \in B$. From each pairwise mapping, a set of hypotheses $h$ is derived that represent candidate partial rationales. When $h \neq \emptyset$, the proposed algorithm augments the target $t$ with $h$. If $h$ is empty for all $b_i$, i.e., no new hypotheses are generated, the algorithm terminates.

**Algorithm 1** Analogy Synthesis Algorithm

**Require:** $B \neq \emptyset$

**Require:** $t \neq \emptyset$

1. $i \leftarrow 1$
2. $b \leftarrow b_i$
3. repeat
   1. $newHyp \leftarrow False$
   2. for $b_i \in B$ do
      1. $h \leftarrow SME_{mapping}(b_i \rightarrow t)$
      2. if $h \neq \emptyset$ then
         1. $t \leftarrow t \cup h$
         2. $newHyp \leftarrow True$
      3. end if
   3. end for
4. until $newHyp = False$

The process of pairwise analogical mapping results in incremental augmentation of the target $t$ with hypothesized rationales that partially explain the observations, enabling an iterative full rationale construction.

Bases and target consist of a set of predicates of different categories. Binary predicates such as $(travelTo \ Door \ Customer)$ are relations. Unary predicates that end in $Aff$, $Fn$, and $Des$ are of type $affordance$, $function$, and $desire$, respectively. The SME treats each type of predicate differently. Relations match only if both the predicate name and the arguments match, while affordances, functions, and desires, match if they are of the same type, even if the predicate names are different e.g. $heightFn$ and $temperatureFn$, and the arguments match.

We have observed that the order in which past experiences $b_i$ are considered matters. With some orderings, new symbols representing conjectured objects are introduced while a different ordering would result in a matching with a current target symbol. The order also matters in terms of computation time, affecting the number of iterations through the set $B$ of bases before it terminates. Based on the intuition that discovering matches earlier is better, we introduce a sort of greedy heuristic method that considers the bases that are more likely to result in a higher number of matches first. The heuristic prioritizes bases with higher structural complexity, measured in terms of the number of edges, and higher degree of similarity with the predicates in $t$. Algorithm 2 shows this heuristic method.

**Algorithm 2** Heuristic on Sequencing Partial Analogies

1. $S_{b_i} \leftarrow Compute\ predicate\ similarity\ to\ t, \forall b_i \in B$
2. $Edges_{b_i} \leftarrow Compute\ number\ of\ edges, \forall b_i \in B$
3. $W_{b_i} \leftarrow S_{b_i} \cdot Edges_{b_i}$
4. Sort $b_i$ based on $W_{b_i}$

An alternative to the greedy heuristic would be to explore the hypothesis space exhaustively and then apply preference criteria to choose one set of rationales among those found. The preference can be based on a combination of minimal- ity (the smaller number of hypotheses the better) and coherence (the more facts and observations explained by a hypothesis the better). We leave exploring this to future work.

**4. Experiments**

To explore the plausibility of the approach, we have implemented Algorithm 1 and tested it on a representation of the gas pump experiences example described in the introduction. The corresponding rationale predictions are further compared with output from an open source large language model (LLM).\(^3\) Our experiments involve four base experiences:

1. A normal gas station visit: a customer travels to the gas stations, pumps gas, pays for the gas and leaves;
2. Dog chases a person: an aggressive dog chases a person and the person runs away;
3. Dark alley: a person walks into a dark alley, another person (a stranger) walks toward the first person and

---

\(^3\)https://huggingface.co/togethercomputer/GPT-NeoXT-Chat-Base-20B
The representation of the normal gas station visit base experience is shown in Figures 5. The representation of the other base experiences is similar. In general, the encoding of an experience consists of statements about object attributes and relationships, e.g., that the gas station is not a socializing area or that the customer wants gas. These facts are followed by a sequence of events similar to the chronologies generated by the AIE. The third element of the base is a rationale, represented by a proposition using the why relation. In each base, object names include a postfix such as .gsMt to make sure they are unique across bases.

In the novel experience, represented as in Figure 6, a customer visits the gas station to pump gas, then a stranger walks toward the customer, and the customer flees the scene. This being the target, it does not include a rationale in the form of a why-statement. The task is to use our iterative analogy approach to hypothesize what might have been the rationale(s) behind the observed behavior.

4.1. Experiment 1A: Heuristic Method

Applying the heuristic method outlined in the previous section to our set of bases and target suggests the following ordering of the bases:
1. Normal gas station visit
2. Dark alley
3. Dog chases a person
4. Car fire

Next we apply Algorithm 1 modified to consider the bases in the order suggested by the heuristics. The results from the first base include four new hypotheses, one of them being a partial rationale, shown as a graph in Figure 7, for the target behavior. This partial analogy gives us an explanation for why the customer travelled to the gas station. Other behavior remains to be explained.

From the dark alley base, we obtain seven new hypotheses, three of which are discarded because they involve unobserved events. This yields no rationale, but does yield an important piece of information: that the stranger in the non-social gas station is dangerous. These hypotheses are depicted in Figure 8.

The next base, the dog chase experience, results in five new hypotheses, including a rationale for the fleeing behavior and another hypothesis around the danger affordance of the person—that the person is aggressive. The hypotheses are depicted in Figure 9.

At this point, we have collected rationales for all the events in the target and two hypotheses that explain why the person that walked to the customer has the danger affordance: because he is a stranger or because he is aggressive. The algorithm will next consider the last base, the car fire. As expected, given all the hypotheses we have collected so far this base does not yield any further hypotheses. The algorithm will then do another iteration over the set of bases, which will generate no further hypotheses, and terminate.

---

4We assume that all actions are observable and hence hypotheses including unobserved events can be discarded.
4.2. Experiment 1B: Comparison With Large Language Model

The four scenarios from Experiment 1A were presented to the LLM in textual form as few-shot examples in the prompt, followed by the novel scenario and a question asking for the rationale behind the person’s behavior. Since the LLM we used is fine-tuned for dialogue-type of interactions, we pose the problem as a dialogue between a human and a chat-bot where the human describes a scenario and asks for a rationale, and then the bot responds with one. In the following sample run, all of the text is part of the prompt except for the LLM’s output shown in italics at the end:

**human**: There is a gas station where people do not typically socialize. A customer wants gas, and the gas station sells gas. The customer travels to the gas station, pumps some gas, pays for the gas, and leaves the gas station. What is the customer’s rationale?

**bot**: The customer’s rationale is that the customer travels to a gas station and pumps gas because the customer wants gas and the gas station sells gas.

**human**: There is a dark alley which is not a social area. A stranger is in the dark alley. The stranger desires criminal behavior. Strangers in dark alleys have the danger affordance. A second person walks into the dark alley. The stranger then approaches the second person. Then the stranger attacks the second person. What is the stranger’s rationale?

**bot**: The stranger’s rationale is that the stranger attacks the second person because the stranger desires criminal behavior.

**human**: A person desires safety and there is a dog nearby that is aggressive. Aggressive dogs have the danger affordance. The dog start running towards the person, and then the person flees. What is the person’s rationale?

**bot**: The person’s rationale is that the person flees because the person desires safety and the dog affords danger.

**human**: There is a gas station which is not a social area, and there is a customer that desires safety. The customer wants gas and the gas station sells gas. The customer travels to the gas station and pumps gas. Then the car catches on fire, and the customer flees. The car catching fire causes the car to afford danger. What is the customer’s rationale?

**bot**: The customer’s rationale is that the customer flees because the customer desires safety and the car affords danger.

**human**: There is a gas station that sells gas where people do not typically socialize. Another person, a stranger, is nearby. The stranger desires safety. The customer drives to the gas station and pumps gas. Then the car catches on fire, and the customer flees. The car catching fire causes the car to afford danger. What is the customer’s rationale?

**bot**: The customer’s rationale is that the customer flees because the customer desires safety and the car affords danger.

**human**: There is a gas station that sells gas where people do not typically socialize. Another person, a stranger, is nearby. The stranger desires safety. The customer drives to the gas station and pumps gas. Then the car catches on fire, and the customer flees. The car catching fire causes the car to afford danger. What is the customer’s rationale?

**bot**: The customer’s rationale is that the customer flees because the customer desires safety and the car affords danger.

The answer from the LLM is close but does not quite get to the main point, that based on the past experiences, the stranger has the danger affordance. Another potential issue with LLMs is that different runs with the same prompt often result in different answers. Another run with the same prompt shown above generated the answer:

**bot**:

The customer desires safety, so the customer runs away because the customer wants to avoid the person.

This time the answer includes an unexpected assertion (the customer is a stranger) that does not follow from the information in the prompt.

It is difficult to compare output from the LLM with output from the SME since concepts which are explicit in predicate-based representations may be implied in natural language. Nevertheless, these results demonstrate that
while LLMs are capable of generating rationale inferences that are interpretable and rich with background knowledge, they are not sensitive to the underlying conceptual structures described in natural language. Specifically, the scenario descriptions were manufactured based on explicit, structured, predicate-based representations which the SME is guaranteed to consider but may be disregarded by the LLM. Although rationale inferences from LLMs might cohere with structural analogies formed between commonplace situations, it appears that this capability breaks down when objects, attributes, and affordances deviate from what is expected in daily life. This poses a particular challenge when searching for distant analogies between abnormal world states—perhaps a fundamental aspect of human creativity [8].

5. Conclusion

In order to understand and formalize rationale, we have introduced the components of a novel modeling pipeline. First, the AIE can be leveraged to generate behaviors and construct chronologies of events, then SME provides a mechanism to map analogies between previous experiences and new observations, and finally novel synthesis techniques allow this mapping to form flexible, partial analogies.

We argue that a given analogy may produce various hypotheses which may all potentially be valid. For example, when considering the prior experience “man runs away from fire,” one might produce the hypothesis that people run because they are afraid of a dangerous thing. Another prior experience might be “man starts running when the pistol of the race fires.” In this case it is not fear that causes the person to run, it is the desire to win the race or even to increase one’s fitness. We therefore argue that both hypotheses could potentially be correct. Our algorithm for analogy synthesis provides a partial mapping strategy designed to generate these plausible explanations through comparison to existing experiences. Having said that, a critical future step in this research effort is to study the mechanism used to validate generated hypotheses and autonomously reason over the most appropriate hypothesis that best explains the observation of the current novel situation.

This is particularly relevant when considering analogical inferences produced by LLMs. Given a particular set of contextual features in a novel situation, there may exist a subset of base (source) analogs which should be used to generate relevant rationales. However, it appears difficult to constrain the LLM dialogue process to focus entirely on these relevant examples. Instead, background knowledge (which is assumed to be rigid across experiences) inevitably seeps in and potentially disrupts the analogical inference and mapping processes. For example, with the Normal Gas Station Visit scenario, the LLM reports that gas stations generally afford safety which is not relevant since the customer explicitly traveled to the station solely to purchase gas. Moreover, this specific gas station likely does not afford much safety since this customer is clearly being attacked. Future work should therefore aim to determine how LLMs might be prompted to consider abnormal factors in their environment and adjust their language accordingly. That being said, the LLM results point towards many potential applications in AI reasoning systems.

References

[1] D. C. Dennett. Kinds of minds – toward an understanding of consciousness. 1996. 1
[2] H. Douglas and E. Sander. Surfaces and essences: Analogy as the fuel and fire of thinking. 2013. 2
[3] B. Falkenhainer, K. D. Forbus, and D. Gentner. The structure-mapping engine. In Proceedings of the Fifth AAAI National Conference on Artificial Intelligence, pages 272–277, 1986. 2
[4] B. Falkenhainer, K. D. Forbus, and D. Gentner. The structure-mapping engine: Algorithm and examples. Artificial Intelligence, 41(1):1–63, 1989. 2, 4
[5] D. Gentner. Structure-mapping: A theoretical framework for analogy. Cognitive Science, 7(2):155–170, 1983. 2
[6] D. Gentner. The mechanisms of analogical learning. In S. Vosniadou and A. Ortony, editors, Similarity and analogical reasoning, pages 199–241. Cambridge University Press, Cambridge: MA, 1989. 2
[7] D. Gentner and A. B. Markman. Structure mapping in analogy and similarity. American psychologist, 52(1):45, 1997. 2
[8] A. E. Green. Creativity, within reason: Semantic distance and dynamic state creativity in relational thinking and reasoning. Current Directions in Psychological Science, 25(1):28–35, 2016. 8
[9] K. J. Holyoak and P. Thagard. Mental leaps: Analogy in creative thought. MIT press, 1996. 2
[10] A. Y. Ng and S. Russell. Algorithms for inverse reinforcement learning. In ICML, 2000. 2
[11] S. Russell. Learning agents for uncertain environments. In COLT, pages 101–103, 1998. 2
[12] G. Sukthankar, C. Geib, H. Bui, D. Pynadath, and R. P. Goldman. Plan, activity, and intent recognition: Theory and practice. Newnes, 2014. 2