Affordance-Based Similarity Measurement for Entity Types

Krzysztof Janowicz[1] and Martin Raubal[2]

[1]Institute for Geoinformatics
University of Muenster, Germany
janowicz@uni-muenster.de

[2]Department of Geography
University of California at Santa Barbara
raubal@geog.ucsb.edu

Abstract. When interacting with the environment subjects tend to classify entities with respect to the functionalities they offer for solving specific tasks. The theory of affordances accounts for this agent-environment interaction, while similarity allows for measuring resemblances among entities and entity types. Most similarity measures separate the similarity estimations from the context—the agents, their tasks and environment—and focus on structural and static descriptions of the compared entities and types. This paper argues that an affordance-based representation of the context in which similarity is measured, makes the estimations situation-aware and therefore improves their quality. It also leads to a better understanding of how unfamiliar entities are grouped together to ad-hoc categories, which has not been explained in terms of similarity yet. We propose that types of entities are the more similar the more common functionalities their instances afford an agent. This paper presents a framework for representing affordances, which allows determining similarity between them. The approach is demonstrated through a planning task.

1 Introduction

Understanding the interaction between agents and their environment is a fundamental research goal within cognitive science. The theory of affordances [1] describes how agents perceive action possibilities of entities within their environment, arising from both the physical structures of the entities and the agent. A major problem with this theory is that it does not account for cognitive and social processes. As argued by Chaigneau and Barsalou [2], function plays a prominent role in categorization, which also emphasizes the importance of affordances as part of human perception and cognition. The process of categorization itself can be explained in terms of similarity. With the exception of alignment models such as SIAM [3] most similarity theories assume that similarity is a static and decontextualized process. This contradicts the definition of affordances as inseparable constructs of agent and environment where entities are grouped around functionality. Similarity measures, such as MDSM [4] and SIM-DL [5], are context-aware, but at the same time reduce the notion of context to the domain of application, i.e., an unstructured set of entities or entity classes.

Measuring entity (type) similarity with respect to affordances requires their representation. The theory presented in this paper specifies such representation based on the conceptual design depicted in [6]. It utilizes an extended affordance theory [7], thus incorporating social-institutional constraints and goal definitions. The paper provides a context-aware similarity measure based on the hypothesis that entity types are the more similar the more common affordances their instances offer a specific user for solving a particular task. Hence the presented measurement theory offers a computational approach towards understanding how cognitive processes and social-institutional aspects interact in categorization. This view strongly correlates with the three main components of geographic information science, i.e., cognitive, computational, and social [8]. The presented framework provides additional insights into the grouping of unfamiliar entities to ad-hoc categories [9].

Starting with a review of related work on affordances and similarity measurement, the paper then introduces a formal representation of the extended affordance theory, which supports the separation of perceiving affordances from their execution [6]. Based on this representation similarity measures are developed that determine the similarity between entity types by comparing affordances. For that reason we decompose the language describing the affordances and transform it to conceptual spaces that support similarity measurement by providing a metric [10]. This leads to a representation of functions and actions in conceptual spaces [11]. The approach is demonstrated using a scenario from psychology and AI based planning, where an agent needs to change a light bulb, involving reasoning about what entities offer support for reaching the ceiling. The presented theory focuses on entity types but allows for modification to work on the level of individual entities as well.

2 Related Work

This section introduces the notion of affordance, its extended theory, and a functional representation framework. We then provide an overview of semantic similarity theories—focusing on those related to GIScience—and AI planning.

2.1 Gibson's theory of affordances

The term *affordance* was originally introduced by James J. Gibson who investigated how people visually perceive their environment [1]. His theory is based on ecological psychology, which advocates that knowing is a direct process: The perceptual system extracts invariants embodying the ecologically significant properties of the perceiver's world. Animal and environment form an inseparable pair and this complementarity is implied by Gibson's use of ecological physics.

Affordances must be described relative to the agent. For example, a chair's affordance "to sit" results from a bundle of attributes, such as "flat and hard surface" and "height", many of which are relative to the size of an individual. Later work with affordances builds on this *agent-environment mutuality* [12]. According to Zaff [13], affordances are measurable aspects of the environment, but only to be measured in relation to the individual. It is particularly important to understand the action relevant properties of the environment in terms of values intrinsic to the agent. Warren [14] demonstrates that the "climbability" affordance of stairs is more effectively specified as a ratio of riser height to leg length.

Several researchers have believed that Gibson's theory is insufficient to explain perception because it neglects processes of cognition. His account deals only with individual phenomena, but ignores categories of phenomena [15]. According to Eco [16], Gibson's theory of perception needs to be supplemented by the notion of *perceptual judgments*, i.e., by applying a cognitive type and integrating stimuli with previous knowledge.

Norman [17] investigated affordances of everyday things, such as doors, telephones, and radios, and argued that they provide strong clues to their operation. He recast affordances as the results from the mental interpretation of things, based on people's past knowledge and experiences, which are applied to the perception of these things. Gaver [18] stated that a person's culture, social setting, experience, and intentions also determine her perception of affordances. Affordances, therefore, play a key role in an experiential view of space [19, 20], because they offer a user centered perspective. Similarly, Rasmussen and Pejtersen [21] pointed out that modeling the physical aspects of the environment provides only a part of the picture. The overall framework must represent the mental strategies and capabilities of the agents, the tasks involved, and the material properties of the environment.

2.2 Extended theory of affordances

In this work we use an extended theory of affordances within a functional model, which supplements Gibson's theory of perception with elements of cognition, situational aspects, and social constraints. This extended theory suggests that affordances belong to three different realms: physical, social-institutional, and mental [7]. In a similar and recent effort, the framework of distributed cognition was used to describe and explain the concept of affordance [22].

Physical affordances require bundles of physical substance properties that match the agent's capabilities and properties—and therefore its interaction possibilities. One can only place objects on stable and horizontal surfaces, one can only drink from objects that have a brim or orifice of an appropriate size, and can be manipulated, etc. Common interaction possibilities are grasping things of a certain size with one's hands or walking on different surfaces. Physical affordances such as the sitability affordance of a chair depend on body-scaled ratios, e.g., doorways afford going through if the agent fits through the opening.

It is often not sufficient to derive affordances from physical properties alone because people act in environments and contexts with social and institutional rules [23]. The utilization of perceived affordances, although physically possible, is frequently socially unacceptable or even illegal. The physical properties of an open entrance to a subway station afford for a person to move through. In the context of public transportation regulations it affords moving through only when the person has a valid ticket. The physical properties of a highway afford for a person to drive her car as fast as possible. In the context of a specific traffic code it affords driving only as fast as allowed by the speed limit. Situations such as these include both physical constraints and social forces. Furthermore, the whole realm of social interaction between people is based on *social-institutional affordances*: Other people afford talking to, asking, and behaving in a certain way.

Physical and social-institutional affordances are the sources of *mental affordances*. During the performance of a task a person finds herself in different situations, where she perceives various physical and social-institutional affordances. For example, a public transportation terminal affords for a person to enter different buses and trains. It also affords to buy tickets or make a phone call. A path affords remembering and selecting, a decision point affords orienting and deciding, etc. In general, such situations offer for the person the mental affordance of deciding which of the perceived affordances to utilize according to her goal.

2.3 Functional representation of affordances

Our conceptual framework of affordances uses an adjusted version of the *HIPE theory of function*, which explains how functional knowledge is represented and processed [24]. This theory explains people's knowledge about function by integrating four types of conceptual knowledge: History, Intentional perspective, Physical environment, and Event sequences. Functional knowledge emerges during mental simulations of events based on these domains. The HIPE theory is well suited to the formalization of affordances because of their functional character [6]. Similar to functions, affordances are complex relational constructs, which depend on the agent, its goal and personal history, and the setting. The HIPE theory allows for representing what causes an affordance and therefore supports reasoning about affordances. More specifically, it is possible to specify which components are necessary to produce a specific affordance for a particular agent.

Figure 1 demonstrates the conceptual framework of the relation between the three affordance categories presented in section 2.2 during the process of an agent performing a task. The agent is represented through its physical structure (*PS*), spatial and cognitive capabilities (*Cap*), and a goal (*G*). Physical affordances (*Paff*) for the agent result from invariant compounds (*Comp*)—unique combinations of physical, chemical, and geometrical properties, which together form a physical structure—and the physical structure of the agent[1]. This corresponds to Gibson's original concept of affordance: a specific combination of (physical) properties of an environment taken with reference to an observer.

Social-institutional affordances (*SIaff*) are created through the imposition of social and institutional constraints on physical affordances—when physical affordances are perceived in a social-institutional context *Cont (SI)*. While performing a task the agent perceives various physical and social-institutional affordances within a spatio-temporal environment represented through *Env (S,T)*. This corresponds to HIPE's

[1] The arrows in Figure 1 represent a function that maps *Comp* and *Agent* to *Paff*.

notion of a physical system and allows for localizing the perception of affordances in space and time.

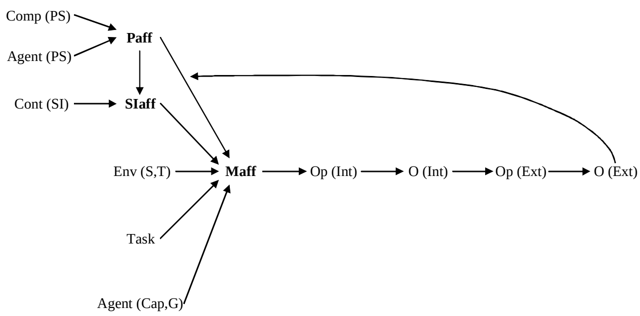

Figure 1: Functional representation of affordances for an agent.

Mental affordances (*Maff*) arise for the agent when perceiving a set of physical and social-institutional affordances in an environment at a specific location and time. Affordances offer possibilities for action as well as possibilities for the agent to reason about them and decide whether to utilize them or not, i.e., mental affordances. The agent needs to perform an internal operation *Op (Int)* to utilize a mental affordance. Internal operations are carried out on the agent's beliefs (including its history and experiences) and lead to an internal outcome *O (Int)*. In order to transfer such outcome to the world, the agent has to perform an external operation *Op (Ext)*, which then leads to an external outcome *O (Ext)*, i.e., some change of the external world. This external change, in turn, leads to new physical affordances, situated in social-institutional and spatio-temporal contexts.

2.4 Semantic similarity measurement

The notion of similarity originated in psychology and was established to determine why and how entities are grouped to categories, and why some categories are comparable to each other while others are not [25, 26]. The main challenge with respect to semantic similarity measurement is the comparison of meanings. A language has to be specified to express the nature of entities and metrics are needed to determine how (conceptually) close the compared entities are. While entities can be expressed in terms of attributes, the representation of entity types is more complex. Depending on the expressivity of the representation, language types are specified as sets of features, dimensions in a multidimensional space, or formal restrictions specified on sets using various kinds of description logics. Whereas some representation languages have an underlying formal semantics (e.g., model theory), the grounding of several representation languages remains on the level of an informal description. Because similarity is measured between entity types which are representations of concepts in human minds, it depends on what is said (in terms of computational representation) about entity types. This again is connected to the chosen representation language, leading to the fact that most similarity measures cannot be compared. Beside the question of representation, context is another major challenge for similarity assessments. In many cases meaningful notions of similarity cannot be determined without defining in respect to what similarity is measured [26, 27].

Similarity has been widely applied within GIScience. Based on Tversky's feature model [28], Rodríguez and Egenhofer [4] developed an extended model called Matching Distance Similarity Measure (MDSM) that supports a basic context theory, automatically determined weights, and asymmetry. Raubal and Schwering [29, 30] used conceptual spaces [10] to implement models based on distance measures within geometric space. The SIM-DL measure [5] was developed to close the gap between geo-ontologies described through various kinds of description logics, and similarity measures that had not been able to handle the expressivity of such languages. Various similarity theories [31, 32] have been developed to determine the similarity of spatial scenes.

2.5 AI planning

Planning is the development of a strategy for solving a certain task and therefore a precondition for intelligent behavior. In terms of artificial agents, a plan is a chain of actions, or action sequence, where each action to be performed depends on some pre-conditions, i.e., a certain state of the world. Each action potentially causes effects or post-conditions that affect or trigger subsequent actions in the chain. The plan terminates when the intended goal is reached. A planner in Artificial Intelligence (AI) takes therefore three input variables: a representation of the initial state of the world, a representation of the intended outcome (goal), and a set of possible actions to be performed to reach the goal. Formally, a plan can be regarded as a triple $\langle O, I, A \langle p, q \rangle \rangle$ [33, 34] where O is the intended outcome, I the initial state of the world, and A a set of actions—each defined via pre- and post-conditions p, q. However, after executing actions the state of the world is changed, which impacts the future plan, therefore making planning a non-linear process. One distinguishes between offline and online planning. Offline planning separates the creation of the strategy and its execution into two distinct phases; this requires a stable and known environment. In contrast, online planning is suitable for unknown and dynamic environments where a pre-given set of behavioral rules and models cannot be determined. One of the main challenges within dynamic environments is that one can neither assume complete knowledge of the environment nor the availability of entities (as part of such environment) supporting certain actions.

3 Use Case

Contrary to classical planning, our vision of an affordance-based and similarity-driven planning service executes as follows: The agent determines an intended outcome (goal). Next, the agent selects a possible affordance descriptor (see section 4) from its internal knowledge base that either leads to the intended goal or is part of the chain towards it. The agent then needs to verify whether an entity of the type specified in the affordance descriptor is available within its immediate environment and if not, whether it can be substituted by an entity of a similar type. After that the agent can execute the (similar) action specified by the selected affordance. Thus, reaching a new state, the agent selects the next outcome to be reached towards the final goal and again chooses an appropriate affordance descriptor. The process terminates when the final outcome is reached or no supporting entity can be detected within the current environment.

To illustrate this approach, we introduce a use case, which is derived from the literature on ad-hoc categories [9, 35]. We assume that an agent needs to change a light bulb in an office room. Before doing so, the agent has to fetch an additional entity that raises it up to a certain level in order to reach the ceiling and change the bulb. In terms of affordances, the task is two-fold: the agent must find an entity that affords standing on and has to be movable to be carried or pushed to the required position. If a single entity lacks sufficient height, an additional affordance will come into play, namely that of being stackable. As illustrated in Figure 2 the office room contains several candidate entities, such as a desk, a chair, and books, which could be utilized to fulfill the task. Some entities are movable, stackable, and offer support for standing on them at the same time, while others fulfill these requirements only partially. We assume that the agent has the necessary capabilities to categorize entities accordingly. If an entity is of a certain type, it can be manipulated as specified in the affordance descriptor (section 4) stored in the agent's knowledge base.

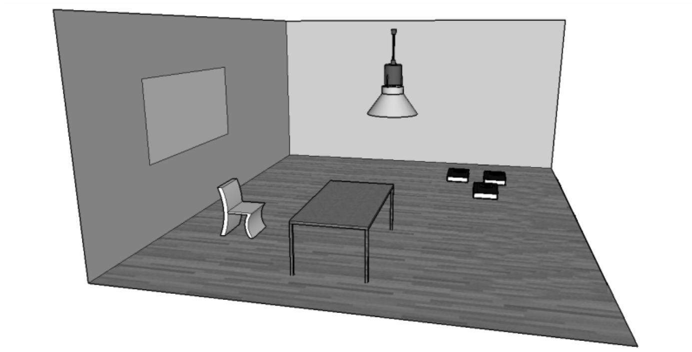

Figure 2: Candidate entities needed to change a light bulb.

Figure 3 shows a simplified representation of the 'changing a light bulb' scenario using the framework discussed in section 2.3. The agent perceives affordances involving the entities *desk*, *chair*, and *book* in the office environment, where the agent is spatio-temporally located. The task is changing the light bulb, which involves a series of sub goals. The physical structure of the desk affords the agent to move it, stand on it, stack it, and to climb it. The *Paff* of moving the desk is constrained through the following social context (or rule): Moving the desk will lead to scratches on the floor, therefore, one should not slide the desk across the floor, resulting in the *SIaff*(not move).

The chair affords the agent to move, stack, climb, and sit on it. Books in the office afford the agent to move and stack them. Notice that all of this knowledge, which had previously been acquired by the agent, is represented for the entity types desk, chair, and book. Perceived instances in the agent's environment are categorized with respect to these types and therefore the agent can utilize knowledge associated with them. This process is similar to a perceptual cycle [36] where a schema directs exploration and sampled objects modify the schema.

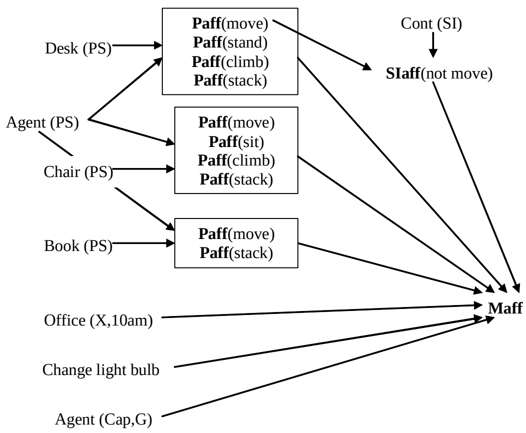

Figure 3: Functional affordance-based representation of use case.

All of these functions result in the top-level *Maff* for the agent, namely to evaluate whether the task of changing the light bulb can be fulfilled with the given constraints represented through the functions. More formally, the (interconnected) sets of physical and social-institutional affordances at a given point in space and time result in a set of mental affordances for the agent: $\{Paff, SIaff\}_{Env(S,T)} \Rightarrow \{Maff\}$. *Maffs* are therefore higher-order functions because *Paffs* and *SIaffs* are functions themselves.

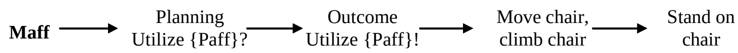

Figure 4: Functional activity process for the agent.

The second part of the process is represented in Figure 4. The agent performs internal operations (within the planning process), deciding whether the task can be performed based on the given functions. This is also where the agent performs affordance-based similarity measurement to find out which entities can be used for the task of changing the light bulb. The outcome of this operation is the decision to utilize a set of affordances. Moving and climbing the chair are external operations, which subsequently lead to the external outcome that the agent stands on the chair and can change the light bulb. This is an abstraction from more complex online planning tasks, as we assume that the same entity can be used for each step (moving, climbing, and standing).

4 Representation of Affordances

This section introduces a representation of affordances built upon the extended theory described in section 2.2, and provides the groundwork for the similarity measurements established in section 5.

Based on previous definitions [37, 38], we specify an affordance *A* as a triple *⟨O, E, {AC}⟩*. The outcome *O* is the change of world state after execution of the actions *AC* with respect to manipulated entities of type *E*. The same affordance can be realized by several actions—each described by physical *PH* and social-institutional *SI* constraints, i.e., pre-conditions. *AC* is therefore defined as a set of actions $\{ac_1(ph_1,si_1),..., ac_n(ph_n,si_n)\}$. Constraints are tied to a certain action with respect to an entity (of a given type), while the outcome is equal for all actions defined for the affordance *A*. Therefore the outcome can also be regarded as the post-condition of all actions of *A*. For the representation of outcome, physical, and social-institutional constraints we employ unary and binary predicates, and apply a restricted kind of predicate logic. The predicates (predicate variables) are part of the agent's internal knowledge base, i.e., no semantic problems arise from the question of what a predicate, such as hasPosition, means to the agent. This knowledge base also contains information about the inter-relation between predicates, such as hierarchies or (spatial and temporal) neighborhood models. The predicate logic used to specify actions and outcome can be regarded as a subset of first order logic. Valid operators and quantifiers are:

- Operators (constructors): logical *not* (¬), *and* (∧), and *implication* (→)
- Quantifiers: ∀, ∃, and ∃! (*exactly one*)
- Arithmetic operators: <, >, ≤, ≥, ≠ and = applied to $\mathbb{R}^+$

In addition, we use some abbreviations to improve readability. This way, some necessary assumptions are made about the later transformation to conceptual spaces without the need of going into detail about the relation between logics and conceptual spaces (see also [39]), and the problem of mapping.

- (∀e∃!r P(e,r) ∧ E(e) ∧ R(r) ∧ R(x) ∧ (r ≤ x)) ∧ …→ Q(e) is abbreviated by Q: P(e, ≤x) ∧ …; where P,Q are predicates, e is an instance of E, and r, x are real numbers. This allows for statements such as the one shown for carry(ability) below.
- The same way ∀e∃!f P(e,f) ∧ E(e) ∧ F(f) is abbreviated by P(e,F(f)); where f is an instance of F such as in hasPosition(e,Position(p)). This states that an entity needs to have a position from where it is moved to another one. The same kind of statement can be made by adding negation, such as on(e, ¬Parquet(p)).

The perception and execution of affordances is modeled in terms of mapping statements, i.e., predicates connected via logical *and*, to Boolean values. Physical constraints describe statements about the physical properties of entities that need to be *true* before the specified action can be performed with respect to entities of the type *E*. Social-institutional constraints specify statements about social aspects regarding the interaction with entities (again abstracted to type level) that need to be *true* before the specified action can be performed. Both types of constraints are specific with respect to the agent perceiving the affordance. We claim that entity types *E* are specified only via what their instances afford a given agent and are the more similar the more common or similar affordances they support.

Summarizing, a certain type of entities affords something to a specific agent if the agent can perform actions on such entities; i.e., *A* is *true* with respect to the agent if at least one of the actions of *AC* is performable (its *PH* and *SI* constraints are satisfied, i.e., *true*) and after realizing the affordance *A* the state of the world changes as specified in *O* (i.e., if the predicates specified for the outcome are *true*). Consequently, although an agent can perceive the affordance of something to be moveable, it could fail to move the entity because of external factors not explicitly stated in the action constraints. This reflects both the separation of internal operation and outcome, and external operation and outcome described in section 2.3. Note that because we assume the agent to be fixed it is not part of the affordance definition itself but its physical and social-institutional context is defined via constraints on the actions.

In terms of the light bulb scenario, an affordance descriptor for moveability of desks is specified as follows:

```
Move-ability (
 Outcome (O): hasPosition(e, Position(y)) ∧ y ≠ x
 Entity Type (E): Desk
 Actions (AC):
  carry(PH:hasPosition(e, Position(x))∧ WeightKg(e, ≤20)∧ LengthCm(e, ≤100)∧ …)
  push(PH:hasPosition(e, Position(x)))∧ WeightKg(e, ≤100)∧ …
         SI: on(e, ¬Parquet(p))
 …)
```

For our agent desks afford moving if they can be either pushed or carried from a position *x* to another position (specified as a position *not* being the start location *x*). Due to the agent's physical capabilities it is able to carry desks with a weight below 21kg and a length of up to 100cm. Pushing the desk is an alternative action and could be even performed with heavier desks (up to 100kg). Pushing though may damage floors. Therefore an entity of type desk is moveable if it either weighs less than 21kg and is not longer than 100cm, or weighs less than 100kg but the supporting floor is resistant to damage caused by sliding heavy entities across it (parquet is not). In this example the restrictions are mostly based on the physical capabilities of the agent and the structure of solid entities. In other ad-hoc categories such as 'things to extinguish a fire' the candidate entities (e.g., water, sand, or a blanket) also differ in their consistency. It is still possible that the agent perceives the affordance and tries to carry the desk but during execution recognizes that for external reasons the desk cannot be moved (e.g., because the desk is mounted to the floor). Depending on these restrictions and the abilities of the agent some desks may not be movable at all. The question whether such entities should still be categorized as desks is not discussed here but will be taken up again in the future work section.

This example also points to the connection between affordances. Depending on the granularity, one may argue that there are explicit pushability and carryability affordances that can be defined as sub-affordances of moveability. In addition, the social-institutional restriction introduced for pushing desks could also be perceived as an affordance (damageability). Damageability of floors is then defined via an outcome *O* specifying that the state (in this case, the surface) of the floors is changed by several actions *AC*.

5 Affordance-based Similarity Measurement

This section describes how similarity between entity types can be measured based on the assumption that types of entities are the more similar, the more common functionalities their instances afford to a given agent. We therefore introduce a framework that specifies what parts of the affordances are compared and how. After defining what makes affordances similar, we use the similarity values determined between affordances to develop a similarity measure for entity types. As we cannot directly handle the expressivity of the affordance representation with respect to similarity measurement, the affordance descriptors need to be transformed to regions within conceptual spaces. Similarity between these regions can be computed using existing measures. Overall similarity is then described as the weighted sum of the individual similarity values measured between affordance descriptors. This step is comparable to a weighted (ω) Tversky Ratio Measure such as used in MDSM. At last the same kind of measure is applied to determine entity type similarity.

5.1 Similarity between Affordances

Each affordance is specified by the change in world state its execution causes and the actions performed on certain types of entities to achieve this outcome. As the entity types are only described in terms of what they afford, similarity between affordances depends on the action and outcome specification. To evaluate whether an affordance defined for a certain entity type is valid in the context of a specific agent and entity, predicates are resolved to Boolean values. Similarity though rests on the assumption that affordances are the more similar the more similar their descriptors are. As no metric can be defined to reason about the similarity of predicates in general, we define mappings for the predicates to quality dimensions within a conceptual space [10], hence being able to utilize a metric for comparison. Similarity measures are asymmetric, therefore the direction of the comparison must be taken into account. In the following, the index *s* is used for source while *t* determines the target, i.e., the compared-to predicate.

First we consider predicates that map entities to non-negative real numbers ($\mathbb{R}^+$). Such predicates can either describe facts about entities of the type specified in the affordance or external entity types. The predicates are transformed to dimensions and the numeric values to upper or lower bounds of the dimensions. If a predicate maps to a single value, lower and upper bounds are equal. If no lower bound is specified it is set to 0 or in the case of upper bound to infinity. Dimensions referring to the entity type *E* together form a conceptual space while dimensions referring to other entity types form conceptual spaces for those types[2]. This is also the reason why the action and outcome descriptors cannot be directly utilized to determine the similarity between entity types. Physical and social-institutional constraints as well as outcomes may directly refer to the specified entity type or to its environment, e.g., via external types.

[2] In such cases similarity is determined in a recursive way as entity types are again explained in terms of affordances.

In cases where predicates map between entities the transformation to dimensions is more complex. The predicates can still be represented as dimensions but only on a Boolean (i.e., nominal) scale[3]. This means that predicates p_s and q_t are equal (similarity is 1) if $p_s = q_t$, which also includes rewriting rules (such as De Morgan rules for $\wedge$ and $\vee$ as well as for $\exists$ and $\forall$) or if q_t can be inferred from p_s. These inference mechanisms include standard inference rules, such as elimination and introduction but also spatio-temporal reasoning, etc. A typical example can be constructed assuming that p_s: PO(X, Y) and q_t: ¬DC(X, Y) in terms of the Region Connection Calculus [40]. Note that the same example (and other inference rules) does not work in the opposite way, which is consistent with keeping similarity measurement asymmetric. In other cases similarity is 0.

The connection between single predicates using $\wedge$ is preserved within the structure of conceptual spaces by the amalgamation (+) of similarity values (Equation 2). The pre-processing step of turning predicate-based descriptions to conceptual spaces is applied to any physical and social-institutional constraints of all action descriptors $ac \in AC$ and the outcome descriptors O of source and target affordance A_s and A_t. This creates at least one conceptual space for each action and outcome of A_s and A_t. The process is computationally expensive, but it is static and therefore easy to cache offline. The process is depicted in Figure 5.

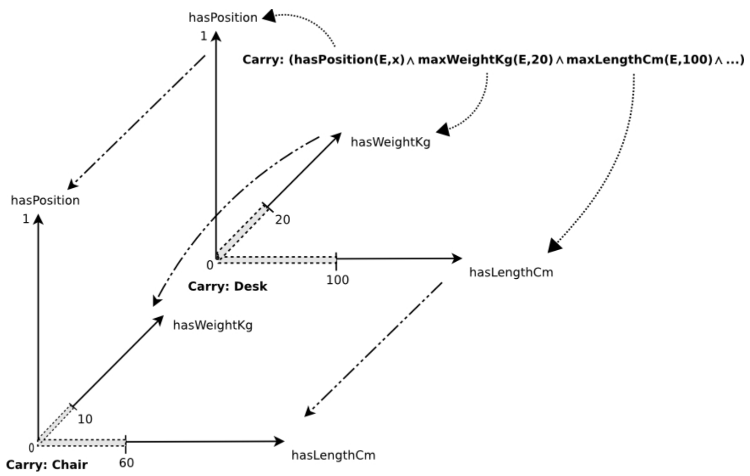

Figure 5: Creating and mapping dimensions from predicates.

After transformation to conceptual spaces an alignment procedure [5, 41] must be established to determine which conceptual spaces of the descriptors of A_s and A_t are mapped for similarity measurement. Each conceptual space describes either the capabilities (in terms of physical or social-institutional constraints) of the agent to perform an action with respect to an entity type or the desired outcome for all actions specified for the affordance. From this the following alignment rules can be derived.

[3] While this approach is used in most of the literature on similarity measures within conceptual spaces, one may argue that such mapping violates the notion of dimensions and regions.

- Spaces representing the outcome aspects of A_s are mapped to such from A_t.
- Spaces representing physical aspects or social-institutional aspects of an action from A_s are mapped to such from A_t.
- Space specifying dimensions for different types of entities cannot be mapped.
- Physical and social-institutional aspects of an action from A_s are jointly mapped to their counterparts of an action from A_t and not separately to different actions.
- Action names are unique within the agent's knowledge base, therefore action descriptors of A_s and A_t are mapped if both share the same name, or if both are situated within the same hierarchy or neighborhood model. In the latter case, the maximal possible similarity is decreased to the similarity within the hierarchy or neighborhood [5, 42].
- If no counterpart for a conceptual space representing constraints of an action can be found, the similarity value is 0.

Finally, after constructing conceptual spaces for the affordance descriptors, semantic similarity between conceptual regions can be measured. Semantic distances are calculated based on the standardized differences of the values for each quality dimension. The final values are normalized by the number of dimensions used in the calculation. This way, a semantic distance function between two conceptual regions can be established [43]. Here, quality dimensions are represented on either Boolean or interval scale. For Boolean dimensions, the values can take 1 or 0, therefore semantic similarity between two conceptual regions for each Boolean dimension is either 1 (completely similar) or 0 (completely dissimilar). In order to calculate asymmetric similarity for two intervals we consider a simplified version of the *line alongness ratio* for topological line-line relations [44]. Positive similarity values of intervals result only if there exists at least a partial overlap between intervals. If such overlap does not exist, the similarity evaluates to 0, i.e., complete dissimilarity. This makes sense for the described scenario, because minimum and maximum values for dimensions are hard constraints for the agent. The calculation of interval similarity is given in Equation 1.

$$Sim_{Int}(I_i,I_j) = length(I_i \cap I_j)/length(I_i) \text{ with } i,j \in \{1,2\},\ i \neq j \quad (\mathbf{1})$$

The final measure for semantic similarity between two conceptual regions X and Y is depicted in Equation 2, where *Scale* refers to either Boolean or interval, S_i and S_j are the respective values of a quality dimension to be compared, and n refers to the number of dimensions.

$$Sim_{CS}(X,Y) = \Sigma(Sim_{Scale}(S_i,S_j)) / n \quad (\mathbf{2})$$

After being able to determine the similarity within conceptual spaces, the similarity between affordances is defined as depicted in Equation 3.

$$Sim_A(A_s,A_t) = \omega_{ac} * {}^1/_n \Sigma\, sim_{AC} + \omega_o * sim_O;\ \text{where}\ \ \Sigma\omega = 1 \quad (\mathbf{3})$$

$\mathrm{Sim_A}$ is specified as the weighted (ω) and normalized sum of similarities for compared actions ($\mathrm{sim_{AC}}$) and outcomes ($\mathrm{sim_O}$) expressed as similarity within conceptual spaces. While $\mathrm{sim_O}$ is directly determined from the outcome conceptual space of A_s and A_t, the similarity between actions is determined via the weighted sum of the similarities for the physical and social-institutional aspects (Equation 4). The number of compared actions *n* in Equation 3 represents alignable actions, not the total number of all available actions.

$$Sim_{AC} = \omega_{ph} * sim_{ph} + \omega_{si} * sim_{si};\ where\ \ \Sigma\omega = 1 \quad (4)$$

Summarizing, as depicted in Figure 6, affordances can be compared by mapping their descriptors to conceptual spaces and expressing predicates as dimensions of such spaces. The overall similarity is then defined as the weighted sum for the individual similarities computed for actions and outcomes, where the former depend on the similarity values computed for their physical and social-institutional constraints (regarding a certain type of entity and agent).

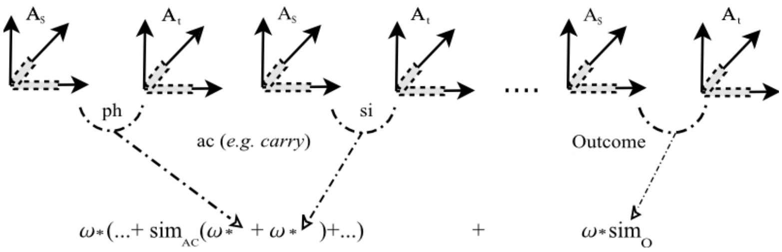

Figure 6: Comparing affordances via their descriptors.

5.2 Similarity between Entity Types

Based on our core assumption that entity types are the more similar the more common functionalities their instances afford an agent (in solving a given task), similarity between types can be determined as depicted in Equation 5.

$$Sim_E(E_s,E_t) = {}^1/_n * \Sigma Sim_A \quad (5)$$

Again, this raises the question which affordances should be compared. The theory presented in this paper is driven by the idea of solving tasks via the agent's interaction with its environment and hence focus on achieving goals. Therefore, affordances are selected for comparison with respect to their outcome, i.e., the affordances A_s and A_t with the highest outcome similarity (sim_O) are matched.

5.3 Weights

Weighting is a useful method to adjust the similarity measurement process to better fit a given task. However, skeptical readers may argue that weights can be used to manually tweak the numbers until they fit the expected results. This section briefly discusses the role of weights in the presented framework to counteract such argumentation. The weights introduced in section 5.1 are comparable to the commonality and variability weights introduced as part of the context theory of

MDSM [4] and can therefore be automatically determined without the need of manual adjustment. While MDSM distinguishes between attributes, functions, and parts to describe and weight entity types, our approach can be weighted with respect to physical and social-institutional constraints as well as outcomes. By using such weights the importance of these aspects can be automatically adjusted to improve the quality of the similarity estimations. Commonality increases the similarity of such kinds of descriptors that are common for most compared entity types and tends to increase overall similarity. In contrast, variability strengthens the weight of such aspects that are unique for certain types of entities and tends to decrease overall similarity. In terms of the light bulb use case, social-institutional aspects may become more relevant for the measured similarity if the physical constraints of the entities are more or less the same (or vice versa).

The proposed weights can be used to define aspects as irrelevant for a certain task (by setting the weight to 0). If a task is of fundamental importance—as can be imagined for the wheelchair example described in [6]—it may be reasonable to ignore social-institutional constraints and focus on the outcome aspects. In other words, the proposed weights can be either automatically determined or used as exclusion factors depending on the task. In both cases the weights do not require manual pre-settings.

6 Application to Use Case

This section applies the presented framework to the light bulb use case. After perceiving the available entities in its environment the agent recognizes that it is not allowed to slide the desk towards the light bulb (modeled as social-institutional constraint). The agent's physical capabilities prevent it from carrying the desk, therefore other entities must be used to perform the task. From previous interaction with its environment the agent knows several facts about certain types of entities. This information is stored as affordance descriptors in the agent's knowledge base. To find out whether other entities (E_t) can be utilized, the agent compares the affordance descriptors (relevant for performing the task) of desks (E_s) with those of books and chairs (figure 3).

As argued in section 4 we assume that the physical constraints on carry und push depend on the abilities of the agent and the structure of the moved entities, leaving other aspects, such as texture, shape, and minimal size (which affect graspability) aside. The following specification for moveability of books does not state that the agent is physically unable to carry heavier books, but represents its current knowledge about the interaction with books[4].

```
Move-ability (
 Outcome (O): hasPosition(e, Position(y)) ∧ y ≠ x
 Entity Type (E): Book
 Actions (AC):
  carry(PH:hasPosition(e, Position(x))∧ WeightKg(e, ≤ 3)∧ LengthCm(e, ≤ 30)∧ …)
  push(PH:hasPosition(e, Position(x)))∧ WeightKg(e, ≤ 3)∧ …
 …)
```

[4] Instead of assuming the previously acquired knowledge as an upper bound, one may also argue for positive infinity as an upper bound.

Contrary to desks, social-institutional constraints are not defined for moving books. The comparison between moveability of desks and books yields a similarity of 0.68. In the case of chairs, where maximal weight and length were set to 10kg respectively 60cm, the resulting similarity is 0.86. This reflects the fact that, with respect to weight and length, the experience of moving desks and chairs is more similar than between desks and books.

The same comparison is also computed for standability, stackability, and climbability, and finally leads to a similarity of 0.36 between desks and books and 0.64 between desks and chairs. Due to the situated nature of similarity and categorization the results cannot be used to argue that chairs or books are similar to desks in general [27]. However, to fulfill the given task, i.e., to change the light bulb the agent can conclude from previously acquired knowledge that chairs are possible candidates (internal operation). While trying to utilize the individual chair (external operation) the agent may fail because—contrary to other chairs—the available chair might not have sufficient stability, and could then utilize the books. If the chair is suitable to solve the task, the agent adds a new affordance (restricted by the physical properties of the used chair) about chairs to its knowledge base. In case of the books, standability and climbability are added. This relates to the fact that humans cannot perceive all information about the physical properties of a certain entity and therefore reason on a category level (based on previous knowledge). This post-processing of entity types can be regarded as a learning process.

7 Conclusions and Future Work

The presented methodology provides a framework for the conceptual affordance representation discussed in [6] and specifies how to measure similarity between affordances and entity types. The formalization captures important aspects of the conceptual design, such as the distinction between physical and social-institutional constraints. Via the outcome specification our approach is able to distinguish between the perception of an affordance and its execution. Similarity measurement is not a static procedure but modeled as a situated process [26, 35] within a context formed by the actor, task, and environment. On the one hand this provides insight into similarity-based categorization of unfamiliar entities (entity types) such as for ad-hoc categories [9]. On the other hand it allows for similarity-based reasoning and planning.

Further work should focus on providing a detailed formal system underlying the presented affordance theory. The important question is not directed to whether a more expressive language can improve the computational representation of affordances but whether such representation still allows for similarity measurement. In this paper simple predicates were transformed to regions in conceptual spaces to determine their similarity. One obvious problem thereby is that all dimensions are regarded as independent from each other, which is normally not the case. For instance, restrictions such as weight and length defined for the actions of the moveability affordance influence each other. In addition, only such predicates were allowed that could be mapped to the conceptual space representation. To overcome these restrictions and allow for more complex logical statements within the affordance definition, one could consider the integration of approaches such as the conceptual spaces logic [39] or similarity measurements a la SIM-DL. Analogy may be an additional tool to compare logical statements that, due to the nature of predicate logic, cannot be compared in terms of similarity. The presented theory would also benefit from the integration of measures focusing on similarity between spatial scenes [31, 32] and case-based reasoning [45]. While we have introduced AI planning to argue for the representation of affordance as a triple $\langle O, E, \{AC\}\rangle$, further work is required to adopt the presented methodology to real planning and learning scenarios. This will involve the interaction with several types of entities to solve a certain task, instead of trying to find one entity (of a given type) that can be used to fulfill all subtasks. Such an extended approach can then be tested with real geographic data sets.

Additional research should investigate the relationship between affordances and how affordances can be combined. While our work focuses on entity types the model can be adapted to entities as well. This raises the question whether an entity is still of some sort independent of whether it offers certain affordances (e.g., a broken cup). Considering temporal aspects, one may argue that an occupied chair (therefore not supporting sitability) should not be categorized as chair anymore. The agent should not reason about entities in terms of categories, such as desks, books, and chairs, but as members of sets determined by affordances. Each 'light bulb changing support'-entity is then defined as a member of the ad-hoc formed set 'moveable *and* standable *and* climbable (*and* stackable)'. The task of similarity is therefore to distinguish between central and radial entities.

8 Acknowledgments

The comments from three anonymous reviewers provided useful suggestions to improve the content and clarity of the paper. We also thank Johannes Schöning for providing the graphical illustration of the use case. Partial funding for this work came from the SimCat project granted by the German Research Foundation (DFG Ra1062/2-1).

References

1. Gibson, J., *The Theory of Affordances*, in *Perceiving, Acting, and Knowing - Toward an Ecological Psychology*, R. Shaw and J. Bransford, Editors. 1977, Lawrence Erlbaum Ass.: Hillsdale, New Jersey. p. 67-82.
2. Chaigneau, S. and L. Barsalou, *The Role of Function in Categories.* Theoria et Historia Scientiarum, forthcoming.
3. Goldstone, R. and D. Medin, *Similarity, interactive activation, and mapping: An overview*, in *Analogical Connections: Advances in Connectionist and Neural Computation Theory*, K. Holyoak and J. Barnden, Editors. 1994, Ablex: Norwood, NJ. p. 321-362.
4. Rodríguez, A. and M. Egenhofer, *Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure.* International Journal of Geographical Information Science, 2004. **18**(3): p. 229-256.

5. Janowicz, K., *Sim-DL: Towards a Semantic Similarity Measurement Theory for the Description Logic ALCNR in Geographic Information Retrieval*, in *SeBGIS 2006, OTM Workshops 2006*, R. Meersman, et al., Editors. 2006, Springer: Berlin. p. 1681 - 1692.
6. Raubal, M. and R. Moratz, *A functional model for affordance-based agents*, in *Affordance-Based Robot Control*, J. Hertzberg and E. Rome, Editors. forthcoming 2007, Springer: Berlin.
7. Raubal, M., *Ontology and epistemology for agent-based wayfinding simulation.* International Journal of Geographical Information Science, 2001. **15**(7): p. 653-665.
8. Goodchild, M., et al., *Introduction to the Varenius project.* International Journal of Geographical Information Science, 1999. **13**(8): p. 731-745.
9. Barsalou, L., *Ad hoc categories.* Memory & Cognition, 1983. **11**: p. 211-227.
10. Gärdenfors, P., *Conceptual Spaces - The Geometry of Thought*. 2000, Cambridge, MA: Bradford Books, MIT Press. 307.
11. Gärdenfors, P., *Representing actions and functional properties in conceptual spaces*, in *Body, Language and Mind*, J. Zlatev, et al., Editors. forthcoming, Mouton de Gruyter: Berlin.
12. Gibson, J., *The Ecological Approach to Visual Perception*. 1979, Boston: Houghton Mifflin Company. 332.
13. Zaff, B., *Designing with Affordances in Mind*, in *Global Perspectives on the Ecology of Human-Machine Systems*, J. Flack, et al., Editors. 1995, Lawrence Erlbaum Associates: Hillsdale, New Jersey. p. 238-272.
14. Warren, W., *Constructing an Econiche*, in *Global Perspectives on the Ecology of Human-Machine Systems*, J. Flack, et al., Editors. 1995, Lawrence Erlbaum Associates: Hillsdale, New Jersey. p. 210-237.
15. Lakoff, G., *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. 1987, Chicago: The University of Chicago Press. 614.
16. Eco, U., *Kant and the Platypus - Essays on Language and Cognition*. English translation ed. 1999, San Diego, New York: Harvest Book Harcourt. 464.
17. Norman, D., *The Design of Everyday Things*. 1988, New York: Doubleday. 257.
18. Gaver, W., *Technology Affordances*, in *Human Factors in Computing Systems, CHI'91 Conference Proceedings*. 1991, ACM Press: New York. p. 79-84.
19. Lakoff, G., *Cognitive Semantics*, in *Meaning and Mental Representations*, U. Eco, M. Santambrogio, and P. Violi, Editors. 1988, Indiana University Press: Bloomington. p. 119-154.
20. Kuhn, W., *Handling Data Spatially: Spatializing User Interfaces*, in *SDH'96, Advances in GIS Research II, Proceedings*, M. Kraak and M. Molenaar, Editors. 1996, International Geographical Union: Delft. p. 13B.1-13B.23.
21. Rasmussen, J. and A. Pejtersen, *Virtual Ecology of Work*, in *Global Perspectives on the Ecology of Human-Machine Systems*, J. Flack, et al., Editors. 1995, Lawrence Erlbaum Associates: Hillsdale, New Jersey. p. 121-156.
22. Zhang, J. and V. Patel, *Distributed cognition, representation, and affordance.* Pragmatics & Cognition, 2006. **14**(2): p. 333-341.
23. Smith, B., *Les objets sociaux.* Philosophiques, 1999. **26**(2): p. 315-347.
24. Barsalou, L., S. Sloman, and S. Chaigneau, *The HIPE Theory of Function*, in *Representing functional features for language and space: Insights from perception, categorization and development*, L. Carlson and E. van der Zee, Editors. 2005, Oxford University Press: New York. p. 131-147.
25. Goldstone, R. and J. Son, *Similarity*, in *Cambridge Handbook of Thinking and Reasoning*, K. Holyoak and R. Morrison, Editors. 2005, Cambridge University Press: Cambridge.
26. Medin, D., R. Goldstone, and D. Gentner, *Respects for Similarity.* Psychological Review, 1993. **100**(2): p. 254-278.

27. Goodman, N., *Seven strictures on similarity*, in *Problems and projects*, N. Goodman, Editor. 1972, Bobbs-Merrill: New York. p. 437-447.
28. Tversky, A., *Features of Similarity.* Psychological Review, 1977. **84**(4): p. 327-352.
29. Raubal, M., *Formalizing Conceptual Spaces*, in *Formal Ontology in Information Systems, Proceedings of the Third International Conference (FOIS 2004)*, A. Varzi and L. Vieu, Editors. 2004, IOS Press: Amsterdam, NL. p. 153-164.
30. Schwering, A. and M. Raubal, *Spatial Relations for Semantic Similarity Measurement*, in *Perspectives in Conceptual Modeling: ER 2005 Workshops CAOIS, BP-UML, CoMoGIS, eCOMO, and QoIS, Klagenfurt, Austria, October 24-28, 2005.*, J. Akoka, et al., Editors. 2005, Springer: Berlin. p. 259-269.
31. Li, B. and F. Fonseca, *TDD - A Comprehensive Model for Qualitative Spatial Similarity Assessment.* Spatial Cognition and Computation, 2006. **6**(1): p. 31-62.
32. Nedas, K. and M. Egenhofer, *Spatial Similarity Queries with Logical Operators*, in *SSTD '03 - Eighth International Symposium on Spatial and Temporal Databases, Santorini, Greece*, T. Hadzilacos, et al., Editors. 2003. p. 430-448.
33. Fikes, R. and N. Nilsson, *STRIPS: a new approach to the application of theorem proving to problem solving.* Artificial Intelligence, 1971. **2**: p. 189-208.
34. Erol, K., J. Hendler, and D. Nau, *Complexity results for HTN planning.* Annals of Mathematics and Artificial Intelligence, 1996. **18**(1): p. 69-93.
35. Barsalou, L., *Situated simulation in the human conceptual system.* Language and Cognitive Processes, 2003. **5**(6): p. 513-562.
36. Neisser, U., *Cognition and Reality - Principles and Implications of Cognitive Psychology*. Books in Psychology, ed. R. Atkinson, et al. 1976, New York: Freeman. 230.
37. Stoffregen, T., *Affordances and events.* Ecological Psychology, 2000. **12**: p. 1-28.
38. Rome, E., et al. *Towards Affordance-based Robot Control*. in *Dagstuhl Seminar 06231 "Affordance-based Robot Control", June 5-9, 2006*. 2006. Dagstuhl Castle, Germany.
39. Fischer Nilsson, J. *A Conceptual Space Logic*. in *9th European-Japanese Conference on Information Modelling and Knowledge Bases*. 1999. Hachimantai, Iwate, Japan.
40. Cohn, A. and S. Hazarika, *Qualitative Spatial Representation and Reasoning: An Overview.* Fundamenta Informaticae, 2001. **46**(1-2): p. 2-32.
41. Markman, A. and D. Gentner, *Structure mapping in the comparison process.* American Journal of Psychology, 2000. **113**: p. 501-538.
42. Rada, R., et al., *Development and Application of a Metric on Semantic Nets.* IEEE Transactions on Systems, Man and Cybernetics, 1989. **19**: p. 17-30.
43. Schwering, A. and M. Raubal, *Measuring Semantic Similarity between Geospatial Conceptual Regions*, in *GeoSpatial Semantics - First International Conference, GeoS 2005, Mexico City, Mexico, November 2005*, A. Rodriguez, et al., Editors. 2005, Springer: Berlin. p. 90-106.
44. Nedas, K., M. Egenhofer, and D. Wilmsen, *Metric Details of Topological Line-Line Relations.* International Journal of Geographical Information Science, 2007. **21**(1): p. 21-48.
45. Stahl, A. and T. Gabel. *Optimizing Similarity Assessment in Case-Based Reasoning*. in *21st National Conference on Artificial Intelligence (AAAI-06)*. 2006. Boston, USA: AAAI Press.