Narrative Objects

Mihai Pomarlan[1] , **Robert Porzel**[2]
[1]Applied Linguistics, University of Bremen, Bremen, Germany
[2]Digital Media Lab, University of Bremen, Bremen, Germany
{pomarlan, porzel}@uni-bremen.de,

Abstract

In this work we operate on the view that the pragmatic employment of objects by an agent is manifested via their affordances for that specific agent. What affordances can manifest themselves in a particular situation depends, in part, on the dispositions offered by the particular objects involved. The contribution of this work is to construct and deploy a large scale formal model of dispositions for physical objects that can be employed to constrain and describe the roles they can play in narratives.

1 Introduction

A simple question such as "what *is* a particular object" turns out to have some complications if its answer is to be relevant for human beings. Hobbs remarks, by way of illustration, that a road is a line when we plan a trip, a surface when we drive on it, and a volume when we hit a pothole ([Hobbs, 1995], pg. 820). A road is always a road, of course, but in our common way to talk we confuse between what something *is*, and what it *is used as*. This confusion suggests that the use is more immediately relevant than the ontological classification, at least where agents pursuing pragmatic goals are concerned.

The relevance of "use as" requires that ontological modelling be able to describe both levels: of being, and of use – and of other ways to interact with the pragmatic goals of an agent. One approach to achieve this works by also modelling the roles that (world) entities play in a narrative an agent constructs about its interactions with the world [Porzel, 2021]. The entities that appear in the agent's narrative are classifications of world entities or ground objects based on the roles they play, rather than the ground objects themselves. A ground object such as a hammer can play various roles, e.g., a murder weapon, a paper weight or a door stopper. It can even be a tool to drive nails. The potential classifications of a physical object – the ways in which it can be *narrativized* – depend on its physical dispositions.

A disposition is a quality an object may have, relevant for questions such as "what can this do?" and "what can be done to it?" A knife can cut – it has a disposition allowing it to be a cutting instrument. A ball of dough can be cut – it has a disposition to be a patient of cutting. While the concept of disposition is complicated to model [Toyoshima *et al.*, 2022], for our purposes here this basic understanding will suffice: a disposition is a quality an object needs so that it can play a particular role in a particular action.

Dispositions tend to operate in, at least, pairs. Where the dispositions of knife and dough meet, there can be cutting. A block of wood can be cut too, though it would be hard to do so with a knife. There are more kinds of dispositions to being a cutting instrument, and not all of them are good fits for all the dispositions to being a cutting patient. Therefore, also needed is knowledge about what dispositions go together, to answer questions such as "what can I cut apples with?", "what can I contain boiling water with?" and so on.

If equipped with knowledge of object dispositions, and how these dispositions "match" and allow various affordances to manifest [Beßler *et al.*, 2020], an agent can tackle several problems it may encounter in its coping with the environment. It can select appropriate tools for tasks it needs to perform, or seek passable substitutes. It can also look at a scene and form an idea of what can happen, and how to work with, or against, such possibilities. If there is a puddle of oil lying around, people tend to be more careful with lit matches.

In our everyday coping with the world, we don't seem to usually tell ourselves stories about activities we master and render routine ([Agre, 1997], chapter 1; [Kahneman, 2011]). We just follow our goals, mostly avoid dangerous configurations of the world, without thinking about our decisions for too long. If however we had to narrativize – perhaps to teach someone else, perhaps to correct a perceived flaw in our action or learn something new – we would use knowledge of dispositions, and dispositional matching. We would assert that some tool is appropriate for a task, or explain that we did something to prevent something else from happening. Presumably, a similar capability is useful also for robotic agents for doing household chores; if nothing else, they should be able to explain themselves to human users/observers.

Therefore, a large scale formal model of dispositions of physical objects will be useful to the domestic service robots of the future. The contribution of this work is to construct and deploy such a model[1] and show how it can be employed to constrain and describe the roles objects can play in an agent's narratives.

[1]The knowledge graph is available at https://github.com/ease-crc/ease_lexical_resources

2 State of the Art

2.1 Dispositions and Affordances

Affordances were informally defined by Gibson as what an environment offers to an agent [Gibson, 1979]. Formal accounts have been proposed (affordances as qualities [Ortmann and Kuhn, 2010], as events [Moralez, 2016], as designs [Awaad *et al.*, 2014], as relations [Chemero, 2003], [Beßler *et al.*, 2020] etc.; overview in [Toyoshima *et al.*, 2022]). Turvey proposed a theory [Turvey, 1992] stating the disposition of a "bearer" can realize an event when it encounters a "trigger" disposition under some "background" conditions. An ontological model for dispositions based on Turvey's insights was recently proposed [Toyoshima, 2018], [Toyoshima *et al.*, 2022]. Learning affordances from interaction has been investigated [Moldovan *et al.*, 2012], [Detry *et al.*, 2009], [Ugur and Piater, 2017].

2.2 Commonsense Knowledge Graphs

Several benchmarks are available to test an agent's ability to answer commonsense queries, e.g. WinoGrande [Sakaguchi *et al.*, 2020]. State of the art AI either sits below human performance, or is not reliable in providing appropriate answers.

Embodying commonsense reasoning in a computational model is complex for many reasons [Davis, 2017], such as the need for commonsense knowledge. To address this, several commonsense knowledge graphs have been constructed: the CommonSense Knowledge Graph (CSKG; [Ilievski *et al.*, 2020]), itself a merge of among others ConceptNet [Speer *et al.*, 2016], ATOMIC [Sap *et al.*, 2019], ImageNet [Deng *et al.*, 2009]. The most famous commonsense knowledge graph was constructed by Cyc [Lenat, 1995], but unfortunately its open- and research access versions have been discontinued.

Linguistic resources have also been used for commonsense reasoning: WordNet [Fellbaum, 1998], FrameNet [Ruppenhofer *et al.*, 2010], VerbNet [Schuler, 2006]. These capture relations between word meanings, descriptions of scenarios, and thematic roles and their selectional restrictions.

Our work builds from existing knowledge graphs by selecting, correcting, and integrating knowledge items from them with new sources. We obtained a rich knowledge store for answering questions about object capabilities and uses.

3 Limits of Open Knowledge Sources

Given the abundance of large knowledge graphs that have some connection to commonsense, are more such graphs needed or useful? We argue here that they are.

We initially used OpenCyc as a foundation for representing knowledge for a robot doing household activity. DOLCE UltraLite [Mascardi *et al.*, 2007] provided a more cognitively motivated foundation, and we then developed an ontological model on top of it to distinguish between object class and object use [Beßler *et al.*, 2021; Beßler *et al.*, 2020], a distinction which OpenCyc does not make consistently. We will look again at the knowledge items in Open-/ResearchCyc in the future. In this paper we discuss the more recent, open-access knowledge graphs such as CSKG.

Modern knowledge graph projects emphasize their broad coverage, manifested in millions of triples. Indeed, CSKG contains plenty of *trivia* knowledge. Knowledge for everyday activities is sparser. E.g., there are no *UsedFor* triples for *wipe/v/wn/contact* – there is no knowledge about what tools would be appropriate for an everyday task such as wiping.

The large scale knowledge graphs of today would probably not exist without automatic techniques and cheap crowdsourcing. However, this is prone to introducing errors[2]. We do not intend to be critical of the CSKG project, which formed the basis of our own. We however became very aware during our work of how problematic it is for inference.

We position our work, relative to other knowledge graphs, as specialized on everyday knowledge and reasoning. More human attention was spent on vetting the items, with newer ontological modelling approaches [Beßler *et al.*, 2021] and new sources of knowledge about object use obtained from games with a purpose [Pfau and Malaka, 2020].

4 SOMA_DFL

We now describe what are the sources for our knowledge graph and how it is structured and organized into an ontology for the purposes of reasoning and answering queries.

4.1 Knowledge Sources

Our primary source is CSKG [Ilievski *et al.*, 2020], which combines several knowledge resources. We have focused on the part of CSKG that is comprised of entities with an associated English WordNet synset[3]. We selected the part of the object taxonomy that refers to tools, buildings, and food. Some of these entities also have associated *UsedFor* and *CapableOf* triples, where the third member of a triple corresponds to an action. We have also selected *MannerOf* triples between actions in *UsedFor* and *CapableOf* triples. We added triples linking actions to VerbNet 3.2 classes, and so to thematic roles and selectional restrictions on role fillers. We have done extensive manual corrections on the collected triples. We added triples describing what items can be used together during an action, Some of this knowledge comes from WordNet synset definitions, some from the work of our colleagues on games with a purpose [Pfau and Malaka, 2020].

4.2 Structure

For reasoning and query answering, the triples from the knowledge graph are organized into an ontology. We now use OWL-DL, but plan to add support for non-monotonic inference as it resembles more the default-with-exceptions pattern that human rules often follow. All birds fly, except those that do not – and it is convenient both to keep the default rule as well as an open-ended list of exceptions.

The ontology we produce is built on top of DUL [Mascardi *et al.*, 2007] and the SOcio-physical Model of Activity (SOMA [Beßler *et al.*, 2021]). These provide higher-level knowledge, in particular that *Tasks* are classify events and define *Roles* which can be filled by appropriate *Objects*.

[2] In CSKG, we find that *amputate/v/wn/medicine* is a manner of *audiotape/n/wn/artifact*, *authorize/v/wn/communication*, *fructify/v/wn/body*, and several other such doubtful relations.

[3] In CSKG, the names of such entities can be identified by a "/c/en/" prefix and a "/wn/" infix. An example of such an entity is */c/en/cut/v/wn/contact*. In this paper we will omit the "/c/en/" prefix.

4.3 Axiomatization

Our dispositional theory asserts that to play a role, an object must have a disposition for it. From VerbNet knowledge, we produce axioms asserting that if an object fills a particular role in a particular task then it must also obey the appropriate semantic restrictions. E.g., the following axioms

$classifies^{-1}\ (Patient \sqcap \exists defines^{-1}\ chew) \sqsubseteq$
$\exists hasQuality\ chew.Patient$
$\exists hasQuality\ chew.Patient \sqsubseteq comestible \sqcap solid$

assert that what plays the patient role for chewing must have the chewable disposition, and therefore comestible and solid.

MannerOf triples are interpreted as describing a taxonomy, and the roles a manner defines are connected to the roles defined by the task it is a manner of. E.g., the following axioms

$chew \sqsubseteq grind$
$chew.Patient \sqsubseteq grind.Patient$

assert that chewing is a kind of grinding and a disposition to be chewable is a disposition to be grindable.

UsedFor triples are interpreted as asserting that an object can play an instrumental role in a particular task. *CapableOf* triples are interpreted as asserting that an object can play a passive role, usually *Patient*. Triples describing combinations of items are interpreted as general assertions about all items of the categories mentioned in the triple.

5 Queries

Reasoning with SOMA_DFL consists in performing subsumption inference tasks between concepts in the ontology and query concepts defined by the user. We query SOMA_DFL with the reasoner Konclude. The ontology is somewhat large, with 22527 object classes coming from CSKG and related resources, and 45538 subclassof/equivalentclasses axioms. Further, the ontology uses the full expressivity of OWL-2. Nonetheless, performing disposition queries is fast – less than a milisecond – as long as a cache is constructed first, which takes about 10 seconds using Konclude 0.7.0 on an Intel®Core™i5-7500 CPU @ 3.40GHz with 8GB RAM. Table 1 shows some example queries.

What can you use (as instrument) to –	
bake	broiler, dutch oven, gas oven, microwave, oven, rotisserie, tandoor
cut	adz, auger, ax, axe, backsaw, (100 more)
wipe	cleaning pad, handkerchief, paper towel, piece of cloth, sponge, swab, towel
What can you use (as object acted on) to –	
burn	album, anthracite, art paper, autograph, card, (204 more)
eat	acerola, ackee, acorn squash, adobo, aitchbone, (1491 more)
spill	absynth, acetone, acidophilus milk, aioli, alcohol, (709 more)
What can you do with –	
bowl	break, break into, change, contain, cover, (25 more)
knife	carve, change, cube cut, dice, (14 more)
paper	burn, change, cover cut, extinguish, (12 more)
What can you use (as instrument) to –	
cut firewood	ax, axe, backsaw, battle-axe, broad hatchet, (25 more)
contain a liquid	alembic, amphora, ampulla, aspersorium, autoclave, (162 more)
cover a bowl	adhesive tape, book, duct tape, foil, newspaper, (3 more)
What can you –	
open with scissors	envelope, letter, packet
scoop with a ladle	alphabet soup, applejack, aqua vitae, aquavit, armagnac, (75 more)

Table 1: Disposition queries, informally stated; concept names shortened for space, responses in alphabetical order

6 Conclusion and Future Work

We have presented a knowledge graph that contains information about object dispositions and possible combinations of objects that can be used to achieve a task. The graph focuses on knowledge useful for everyday activities, and can be employed by a computational agent to select appropriate tools or patients (objects to act on) for a task, or to understand a scene in terms of what is possible for the objects in it to do together. Our graph differs both in its focus – everyday activity knowledge – and in its purpose – logical reasoning, as opposed to information retrieval – from previous projects.

We will add support for non-monotonic inference, because it captures better the default with open-ended exception list way that human knowledge is often organized, and will integrate knowledge about causal relations between events.

Acknowledgments

This work was funded by the by the FET-Open Project #951846 "MUHAI – Meaning and Understanding for Human-centric AI" by the EU Pathfinder and Horizon 2020 Program and by the German Research Foundation (DFG) as part of Collaborative Research Center (SFB) 1320 EASE – Everyday Activity Science and Engineering, University of Bremen in subprojects P01 and P05.

References

[Agre, 1997] Philip Agre. *Computation and Human Experience*. Learning in doing. Cambridge University Press, New York, 1997. 00550.

[Awaad *et al.*, 2014] Iman Awaad, Gerhard Kraetzschmar, and Joachim Hertzberg. Challenges in finding ways to get the job done. In *2nd Planning and Robotics (PlanRob) Workshop at 24th International Conference on Automated Planning and Scheduling (ICAPS)*, 06 2014.

[Beßler *et al.*, 2020] Daniel Beßler, Robert Porzel, Mihai Pomarlan, Michael Beetz, Rainer Malaka, and John Bateman. A Formal Model of Affordances for Flexible Robotic Task Execution. In *European Conference on Artificial Intelligence (ECAI)*, 2020.

[Beßler *et al.*, 2021] Daniel Beßler, Robert Porzel, Mihai Pomarlan, Abhijit Vyas, Sebastian Höffner, Michael Beetz, Rainer Malaka, and John Bateman. Foundations of the socio-physical model of activities (soma) for autonomous robotic agents. In Boyan Brodaric and Fabian Neuhaus, editors, *Formal Ontology in Information Systems - Proceedings of the 12th International Conference, FOIS 2021, Bozen-Bolzano, Italy, September 13-16, 2021*, Frontiers in Artificial Intelligence and Applications. IOS Press, 2021.

[Chemero, 2003] Anthony Chemero. An outline of a theory of affordances. *Ecological Psychology*, 15(2):181–195, 2003.

[Davis, 2017] Ernest Davis. Logical formalizations of commonsense reasoning: A survey. *J. Artif. Intell. Res.*, 59:651–723, 2017.

[Deng *et al.*, 2009] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[Detry *et al.*, 2009] Renaud Detry, E. Baseski, M. Popovic, Y. Touati, N. Kruger, Oliver Kroemer, Jan Peters, and Justus Piater. Learning object-specific grasp affordance densities. 01 2009.

[Fellbaum, 1998] Christiane Fellbaum, editor. *WordNet: an electronic lexical database*. MIT Press, 1998.

[Gibson, 1979] James J Gibson. *The Ecological Approach to Visual Perception*. Psychology Press Classic Editions, 1979.

[Hobbs, 1995] Jerry R. Hobbs. Sketch of an ontology underlying the way we talk about the world. *Int. J. Hum. Comput. Stud.*, 43(5-6):819–830, 1995.

[Ilievski *et al.*, 2020] Filip Ilievski, Pedro A. Szekely, and Bin Zhang. CSKG: the commonsense knowledge graph. *CoRR*, abs/2012.11490, 2020.

[Kahneman, 2011] Daniel Kahneman. *Thinking, fast and slow*. Farrar, Straus and Giroux, New York, 2011.

[Lenat, 1995] Douglas B. Lenat. Cyc: a large-scale investment in knowledge infrastructure. *Commun. ACM*, 38:32–38, 1995.

[Mascardi *et al.*, 2007] Viviana Mascardi, Valentina Cordì, and Paolo Rosso. A comparison of upper ontologies. pages 55–64, 01 2007.

[Moldovan *et al.*, 2012] Bogdan Moldovan, Plinio Moreno, Martijn Van Otterlo, José Santos-Victor, and Luc De Raedt. Learning relational affordance models for robots in multi-object manipulation tasks. In *2012 IEEE International Conference on Robotics and Automation*, pages 4373–4378. IEEE, 2012.

[Moralez, 2016] Larry A Moralez. Affordance ontology: towards a unified description of affordances as events. *Res. Cogitans*, 7(1):35–45, 2016.

[Ortmann and Kuhn, 2010] Jens Ortmann and Werner Kuhn. Affordances as qualities. In *Proceedings of the 2010 Conference on Formal Ontology in Information Systems: Proceedings of the Sixth International Conference (FOIS 2010)*, pages 117–130, Amsterdam, The Netherlands, The Netherlands, 2010. IOS Press.

[Pfau and Malaka, 2020] Johannes Pfau and Rainer Malaka. We asked 100 people: How would you train our robot? In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*, pages 335–339, 2020.

[Porzel, 2021] Robert Porzel. On formalizing narratives. In *Proceedings of the JOWO – Ontology Workshops*, Bolzano, Italy, 2021.

[Ruppenhofer *et al.*, 2010] Josef Ruppenhofer, Michael Ellsworth, Miriam Petruck, Christopher Johnson, and Jan Scheffczyk. Framenet ii: Extended theory and practice. 01 2010.

[Sakaguchi *et al.*, 2020] Keisuke Sakaguchi, Ronan Le Bras, Chandra Bhagavatula, and Yejin Choi. Winogrande: An adversarial winograd schema challenge at scale. *ArXiv*, abs/1907.10641, 2020.

[Sap *et al.*, 2019] Maarten Sap, Ronan Le Bras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A. Smith, and Yejin Choi. ATOMIC: an atlas of machine commonsense for if-then reasoning. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*, pages 3027–3035. AAAI Press, 2019.

[Schuler, 2006] Karin Kipper Schuler. *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. PhD thesis, University of Pennsylvania, 2006.

[Speer *et al.*, 2016] Robert Speer, Joshua Chin, and Catherine Havasi. Conceptnet 5.5: An open multilingual graph of general knowledge. In *AAAI Conference on Artificial Intelligence*, 2016.

[Toyoshima *et al.*, 2022] Fumiaki Toyoshima, Adrien Barton, and Jean-Francois Ethier. Affordances and their ontological core. *Journal of Applied Ontology*, 17(2):285–320, 2022.

[Toyoshima, 2018] Fumiaki Toyoshima. Modeling affordances with dispositions. In *Proceedings of the Joint Ontology Workshops*, 2018.

[Turvey, 1992] M. T. Turvey. Ecological foundations of cognition: Invariants of perception and action. *American Psychological Association*, 1992.

[Ugur and Piater, 2017] Emre Ugur and Justus H. Piater. Emergent structuring of interdependent affordance learning tasks using intrinsic motivation and empirical feature selection. *IEEE Trans. Cogn. Dev. Syst.*, 9(4):328–340, 2017.