Image Schemas and Conceptual Dependency Primitives: A Comparison

Jamie C. Macbeth
Fairfield University, jmacbeth@smith.edu

Dagmar Gromann
CSIC - Instituto de Investigacion en Inteligencia Artificial (IIIA)

Maria M. Hedblom
Free University of Bozen-Bolzano

Image Schemas and Conceptual Dependency Primitives: A Comparison

Jamie C. MACBETH [a,1], Dagmar GROMANN [b] and Maria M. HEDBLOM [c]

[a] *Department of Electrical and Computer Systems Engineering, Fairfield University, USA*

[b] *Artificial Intelligence Research Institute (IIIA-CSIC), Bellaterra, Spain*

[c] *Research Center for Knowledge and Data, Free University of Bozen-Bolzano, Italy*

Abstract. A major challenge in natural language understanding research in artificial intelligence (AI) has been and still is the grounding of symbols in a representation that allows for rich semantic interpretation, inference, and deduction. Across cognitive linguistics and other disciplines, a number of principled methods for meaning representation of natural language have been proposed that aim to emulate capacities of human cognition. However, little cross-fertilization among those methods has taken place. A joint effort of human-level meaning representation from AI research and from cognitive linguistics holds the potential of contributing new insights to this profound challenge. To this end, this paper presents a first comparison of image schemas to an AI meaning representation system called Conceptual Dependency (CD). Restricting our study to the domain of physical and spatial conceptual primitives, we find connections and mappings from a set of action primitives in CD to a remarkably similar set of image schemas. We also discuss important implications of this connection, from formalizing image schemas to improving meaning representation systems in AI.

Keywords. natural language understanding, human cognition, conceptual primitives, Conceptual Dependency, image schemas.

1. Introduction

Building systems to understand natural language has been a significant, decades-long challenge to Artificial Intelligence (AI) research. Whether these systems are composed through machine learning (see e.g. [12]), deep learning over large datasets (see e.g. [25]), or are built by hand, a major component to this challenge is the question of meaning representation behind language, also known as the symbol grounding problem. If AI systems are to achieve human-like performances on natural language understanding tasks, it stands to reason that the meaning representations they use should also be human-like. Work in cognitive AI in the area of natural language understanding will benefit from stronger links to cognitive linguistics.

There is compelling evidence from cognitive linguistics that the semantic structure of natural language reflects conceptual structure that arises from our sensorimotor expe-

[1]Corresponding Author: Jamie C. Macbeth, Department of Electrical and Computer Systems Engineering, Fairfield University, Fairfield Connecticut, 06824, USA; E-mail: jmacbeth@fairfield.edu.

rience with the world [32,4,23,6]. One theory that aims to concretize these sensorimotor experiences is the theory of image schemas, in which abstract patterns construe conceptual building blocks for higher level cognition, such as language and analogical reasoning [13,9]. This paper presents an investigation that establishes connections between image schemas and Conceptual Dependency, a theory and meaning representation system from "classical" AI research in natural language understanding. Conceptual Dependency (CD) attempts to decompose meanings into complex structures whose elements are conceptual primitives that are very similar to the building blocks of image schemas. But CD theory also presents a formalized model of a semantics of natural language that reflects human cognition and memory.

For CD and other AI research in natural language understanding systems, drawing connections between CD and image schemas serves several purposes. Firstly, insofar as image schemas are based on empirical support on how humans represent thought [14,22], it will serve to transitively ground and validate CD's conceptual primitives in empirical data. Alternatively, the methods of investigation that have been developed for image schemas may be extended to investigations into the cognitive plausibility of CD. Finally, the connection could guide processes for improving CD as a representation system—for example, by providing supporting theory and evidence for efforts to generalize CD primitives to more situations, add or remove primitives, or modify the connectives used to build CD structures. For image schemas, few approaches to their formal representation exist, and where they do, they generally focus on specific image schemas, such as CONTAINMENT [1] or SOURCE_PATH_GOAL [8], or specific problems that blend various schemas [30]. A comparison to CD holds the promise of contributing to their formalization, and enables their potential use in AI and cognitive systems. CD provides a set of action primitives that have been applied to low level sensory data [10] and thus may help to bridge the gap between the highly abstract quality of image schemas and low-level data and object manipulation instructions.

In this paper we present a sophisticated and rich collection of correspondences between CD and image schemas. In particular, by narrowing the scope of our investigation to structures representing physical and spatial concepts, we find a strong mapping between six action primitives and "picture-producing structures" of CD and a corresponding set of image schemas and the spatial primitives that are theorized to compose them. We support our conclusions with linguistic evidence and evidence from a human-subjects study connecting conceptual primitives to language comprehension. Although this comparison did not result in an ideal one-to-one mapping between the two systems, the links we found have a broad range of implications, from formalizing image schemas to engineering meaning representation systems that are grounded in cognitive linguistics.

2. Background

Prior to an overall comparison, we individually and separately provide a background introduction to CD and image schemas.

2.1. CD Primitives

Conceptual Dependency (CD) theory is a meaning representation theory developed for computer-based natural language understanding systems as an alternative to meaning

Table 1. The six physical primitive acts of CD, and the PART and CONTAIN primitives used to specify properties of objects in conceptualizations. Descriptions along with example sentences and examples of how PART and CONTAIN are used to specify physical objects in conceptualizations are also provided (`<=>` indicates a mutual dependency relationship in CD). For the six primitive acts, the descriptions and example sentences correspond to those used in a recent human-subjects study on CD [19].

Primitive	Description	Example/"Target" Sentences
PTRANS	A person, object, or thing changes physical position or location.	"Matthew flew home from Los Angeles." "Robert returned home from downtown."
MOVE	A person, object, or thing moves a part of its body or part of itself.	"Kevin crossed his arms." "Joe swung his fist at David."
GRASP	A person, object, or thing grabs hold of another person, object, or thing, or becomes attached to another person, object, or thing.	"Jim held on to the railing." "The gecko stuck to the wall."
PROPEL	A person, object, or thing applies a force to another person, object, or thing, or a moving person, object, or thing strikes or impacts another person, object, or thing.	"Lisa kicked the ball." "Bill was hit by a car."
EXPEL	A person, object, or thing, is taken from or comes from inside another person, object, or thing and is forced out.	"Michelle threw up her lunch." "Stephanie bled from a cut on her leg."
INGEST	A person, object, or thing is forced (or forces itself) to go inside of another person, object, or thing.	"Charles ate a hamburger." "Amy took a deep breath."
PART	Denotes inalienable possession relations (i.e. body parts) for objects in CD conceptualizations.	`hand <=> PART(John)`
CONTAIN	Denotes containment relations for objects in CD conceptualizations.	`frog <=> CONTAIN(box)`

representation theories based on formal, mainstream linguistics [26,27,28,18]. CD is in a unique class of AI systems that intend to be *non-linguistic* meaning representations to enable systems with an in-depth, human-like understanding of language that mirrors human cognition and the imagery it generates in the understanding process. CD was applied successfully over several decades of natural language understanding systems research [28,29,18].

CD decomposes meaning into primitive physical acts in a way that renders it particularly apt for action-based representations as favored by image schemas. As such, is suitable as a representation system to correspond to image schemas, since image schemas are suggested to be generalizations made from sensorimotor experiences, and thus are more abstract than lexical concepts. Although extant, hand-built CD systems are currently rare following a turn in natural language processing research towards big data, cognitively-inspired meaning representation systems like CD continue to be relevant for machine learning in relation to in-depth natural language understanding tasks [17,16,10,20].

CD has a number of primitives used to represent thought, perception, social interaction and communication (see [27]). In this paper, however, we narrow our focus to CD's physical, spatial, and object-defining primitives, whose names and descriptions are given in Table 1. Although many of the names of the primitives are chosen from the English lexicon, the concepts that they identify differ from the typical meanings corresponding to their English names, and are usually broader and more abstract. For example, although the English word "ingest" is typically used to describe events where animate beings consume food or drink, the INGEST primitive of CD is broader and is intended for use in a broad variety of acts where a substance or object enters the body of an animate being, such as a person breathing air, a person injecting themselves or someone else with a substance using a hypodermic needle, or a single-cell organism absorbing a molecule through its cell wall. Some of the primitives have abbreviated names; for example PTRANS is short for "Physical TRANSfer"[2].

Researchers in CD purposefully endeavored to keep the number of primitives in the system small, and to develop the system to represent a large number of surface linguistic expressions by decomposing each into a unique but complex structure of primitive events, actors, objects, and connectives between events (such as causality). Having a small number of primitives was intended to make the representations as unambiguous as possible, by reducing cases where more than one primitive or more than one combination of primitives represented the same meaning [29]. Reducing ambiguity in the meaning representation is of paramount importance to natural language understanding and story understanding systems in AI, since much of the understanding process involves inference and deduction operations performed on a meaning instance, or using the meaning instance to search for knowledge in a knowledge base [33,3].

2.2. *Image Schemas and Spatial Primitives*

Image schemas were first introduced by Lakoff [13] and Johnson [9] as abstract patterns that, most often, can be described as spatio-temporal relations arising from bodily experience. Within the framework of embodied cognition theory [31], the sensorimotor experiences humans make in interacting with the world, even in early infancy, are mentally represented as collections of perceptual states [4,24]. For instance, if a child learns that a glass can contain water, it can, with enough exposure, infer that a cup can contain any liquid. Through further exposure and generalization of similar situations of 'objects contained in other objects' the cognitive understanding represented as the image schema CONTAINMENT emerges as part of the child's conceptual world. In contrast to purely perceptual information, image schemas are stored in an explicit and accessible format [24]. Cognitive linguistics has provided compelling evidence that image schemas underlie lexical concepts externalized in specific lexical forms [32,2].

Mandler and Pagán Cánovas [23] study the development of image schemas in infancy and differentiate three degrees of complexity in embodied cognitive structures: spatial primitives, image schemas, and schematic integrations. Spatial primitives are initial conceptual buildings blocks that allow us to understand what we perceive, image schemas are combined to simple spatial events using those primitives, and schematic integrations are built from the first two and represent conceptual blends structured by

[2]The abbreviation is likely due to the primitive names being invented in resource-limited computing environments of the 1960s and 1970s where implementations abbreviated symbol names to conserve memory.

Table 2. Some common image schemas and selected spatial primitives.

Image Schema	Spatial Primitives	Description
CONTAINMENT	IN, OUT, BOUNDARY, CONTAINER	Boundary, enclosed area or volume, or excluded area or volume [9]
SOURCE_PATH_GOAL	SOURCE, GOAL, PATH, LINK, MOVE, DIRECTION	Source or starting point, goal or endpoint, a series of contiguous locations connecting those two, and movement [9,13]
PART-WHOLE	PARTS, WHOLE, CONFIGURATION	Whole(s) consisting of parts and a configuration of the parts [13]
SUPPORT	CONTACT	Contact between two objects in the vertical domain [21]
FORCE	SOURCE, GOAL, PATH, DIRECTION, MOVE, SCALE	A source, target, direction, and intensity of the force, path of motion of the source and/or target, and a sequence of causation [9]

image schemas. For instance, they suggest PATH, MOVE, START_PATH, and END_PATH among others as primitives that structure the image schema SOURCE_PATH_GOAL. Findings from linguistic analyses support the claim that image schemas are structured by spatial primitives that can be combined in different ways to make up an image schema. For instance, English children and adults were found to more frequently use PATH_GOAL (PATH, MOVE, END_PATH) over SOURCE_PATH (PATH, MOVE, START_PATH) in language [15].

Some of the most common image schemas that are relevant for an initial mapping to primitives of CD are presented in Table 2, which specifies their spatial primitives and a description derived from the indicated existing literature. The distinction between image schemas and spatial primitives is not yet well defined. For instance, Johnson considers LINK an image schema in its own right, since the type of connection between two OBJECTs can be of temporal, causal or functional nature, whereas its use as a primitive here only relates to its spatial connection between OBJECTs.

The formalization of image schemas in algebraic and logical theories (e.g. [11,1, 8]) has been a concern across disciplines. Any formal representation of image schemas needs to not only represent static relationships, but also temporal change and movement, while simultaneously following logical transitivity and the *gestalt* laws identified with a particular image schema. From a logical perspective, Bennett and Cialone [1] formally specified different cases of CONTAINMENT as found in natural language text. Galton [5] formalized some transitive aspects of CONTAINMENT in a study on how to formally approach affordances. Hedblom, Kutz and Neuhaus [8] proposed a formalized family of micro-theories for SOURCE_PATH_GOAL. In [7] the authors introduce a method to logically represent image schemas based on a combination of different calculi and linear temporal logic inspired by previous formalizations, which allows for the representation of events and complex image schemas. However, no complete formalization of all image schemas has yet been achieved.

3. Case Study Evidence

To facilitate a discussion comparing and contrasting image schemas and CD theory, we present qualitative evidence from data collected in a recent human subjects study on the coherence of CD primitives [19]. In this CD focused study, 50 subjects were presented with descriptions of the six physical primitive acts of CD as well as 12 short, simple "target" sentences which described human characters performing simple acts. The descriptions of the primitives and the 12 sentences that participants were presented with are illustrated in Table 1. The sentences were designed so that the major act in each sentence was a match for the broad category of acts represented by one of the conceptual primitives, and among the 12 sentences there were two sentences corresponding to each of the six primitives. For each sentence, subjects were asked to freely choose the primitive which was the best match for the "main action" in their understanding and conceptualization of the sentence. They were also asked to give brief, one-sentence explanations of their selections. The CD primitives were identified by numbers instead of their English-word names so that subjects would not be biased by the typical meanings attached to their names.

The study found that the subjects' answers occasionally disagreed with the expected matching between primitives and sentences, and that these answers and explanations were informative of ways that the primitive representation could be improved, either for applying a primitive to new situations, or for modifying the way the primitives, as a complementary set, cover the space of conceptualizations. Throughout this paper we use particular cases from this study as evidence to support our arguments, views, and reflections regarding connections between CD and image schemas.

4. Contrasting CD and Image Schemas

The method of the comparison of CD and image schemas presented in detail in this section is based on an element-wise comparison of CD primitives to image schemas and their spatial primitives. To support the resulting mapping (shown in Table 3), we provide natural language example sentences as well as their interpretation by participants in a case study described in Section 3.

4.1. Motion and PATH*s*

A majority of CD physical primitives incorporate some sort of motion through space. While PTRANS is typically used to build conceptualizations corresponding to an object or actor in its entirety changing location from one place to another by whatever means, the MOVE primitive is reserved for situations in which an animate actor moves a part of their (or its) body, while the rest of their body remains at rest. This distinction is useful for story understanding systems in AI, because they are often required to keep track of the locations of story characters and objects as a story progresses, while at the same time processing movements of actors' body parts that do not cause a change in their location, such as moving a foot to kick a ball or moving a hand to grasp a glass of liquid.

However, moving a body part does also imply a change of location of that body part; if a character is near a door and MOVEs their left hand to the location of the doorknob, the story understander system must be able to track or represent that the character's left hand has changed its location so that it is now at the doorknob, and not, say, in the character's pants pocket. A further examination indicates that, for all practical purposes, the only difference between PTRANS and MOVE is that MOVE represents a change in location of a body part, while PTRANS represents a change in the location of a character's entire body, as well as any other kind of object.

The CD primitives PTRANS and MOVE most resemble the image schema SOURCE_PATH_GOAL and could also be interpreted image-schematically in terms of the spatial primitives found in SOURCE_PATH_GOAL. One of the spatial primitives of SOURCE_PATH_GOAL is actually called MOVE (see Table 2), only it does not have the constraint to the movement of body parts that CD's MOVE relies on. Both PTRANS and MOVE in CD describe the movement of an entity along a PATH from one location, a SOURCE, to another location, a GOAL, passing a number of contiguous locations in between those two points. One fundamental difference is that the image schema SOURCE_PATH_GOAL is not limited to physical locations but also includes abstract, metaphorical relocations, such as "We want to continue along the successful path to growth". Continuing the previous example on MOVE, one could consider the "hand" as an OBJECT image schema that moves along a PATH towards a final destination, the GOAL. When we consider a movement to the doorknob, the doorknob would be its GOAL. This would give us a PATH_GOAL image schema for the movement of the hand to the doorknob, which is similar to MOVE.

Generally the comparison between CD and image schemas suggests that a modified version of CD could combine PTRANS and MOVE together into a single primitive representing change in location for both objects and body parts of animate actors. There is evidence from studies of human subjects conceiving CD primitive acts in the meanings of simple sentences in support of combining PTRANS and MOVE in a way that is similar to the single SOURCE_PATH_GOAL image schema. When subjects were asked which CD primitive was the best match for the main act in the sentence "Kevin crossed his arms", two subjects responded that PTRANS was the best match, in spite of the fact that "arms" should probably be conceived as a part of the actor's body, implying the MOVE CD primitive act. Their explanations were also consistent with viewing the main act as a PTRANS:

> "The location of his arms changed."
> "Physical position of his arms changed."

For the sentence "Joe swung his fist at David," we again expected subjects to respond with the MOVE primitive, but one subject answered PTRANS. This subject explained:

> "Physical position of his fist changed and no force is implied."

4.2. Motion and CONTAINMENT

The act of one object going into another object as well as the act of an object exiting from another object are represented in the CD system through the pair of conceptual primitives, INGEST and EXPEL. In representing meaning in AI systems based on CD, the INGEST primitive was typically used to represent some object or substance, either animate or inanimate, going inside of an animate object. For example, INGEST was used to describe humans or other animate beings consuming food or drink, breathing, or inhaling smoke [27]. Likewise the EXPEL primitive was used to represent animate beings performing acts like expelling or vomiting, sweating, bleeding from a wound, or exhaling air or smoke.

The image schema CONTAINMENT is structured in terms of an inside, an outside, and a border that separates the two. Compared to the definition of INGEST in Table 1, going into an object (or region) corresponds to crossing the border and entering a CONTAINER. Since the CD is closely tied to motion, the CONTAINER represents the GOAL of a SOURCE_PATH_GOAL and based on the definition of INGEST, FORCE is applied to the motion of an OBJECT. To map from the CD INGEST we require the three image schemas CONTAINMENT, FORCE, and SOURCE_PATH_GOAL. In case the type of force happens to be passive, that is, something is forced rather than forces, we could specify the FORCE image schema to COMPULSION in the conceptual blend or schematic integration, where COMPULSION stands for one entity being moved by external force with a certain magnitude and a certain direction along a path. For EXPEL, only the DIRECTION primitive of the SOURCE_PATH_GOAL image schema changes, but not the conceptual blend as such, which means we can map EXPEL and INGEST to the same conceptual blend. The image schemas also encourage extending the application of CD's EXPEL and INGEST to inanimate objects performing these acts.

4.3. *GRASP and* CONTACT

While the CD primitive GRASP does not appear to be about motion, it may be used in stories where actors grasp some object and then move or change location while they are still grasping that object. This is important for story understanding because story characters may grab an object with their hand, and then MOVE their hand. The story understanding system needs to make the inference that the object, grasped in the actor's hand, also changed location to wherever the hand moved. GRASP does not have to be performed with human hands, it could be a person pinching a newspaper between their elbow and the side of their body, a dog holding something in its teeth, or an animal using a prehensile tail. For animate objects performing a GRASP, GRASP may be generalized to also contain the typical meaning of the verb "to attach".

From an image-schematic perspective, the two different situations described by GRASP, namely the attachment on the one hand and the active grabbing on the other hand, need to be distinguished. While both represent a CONTACT between two OBJECTs, the attachment corresponds to a unidirectional ATTRACTION. In other words, one OBJECT applies force to achieve a continuous CONTACT with the other OBJECT. This can be seen particularly well from the example sentence for GRASP in Table 1, which is "The gecko stuck to the wall". The case of active grabbing is more complex and requires conceptual blends in most situations. In the example "Jim held on to the railing" a mapping to SUPPORT could be argued if we accept a non-vertical dimension of the entity being supported as a possibility for this schema. However, if one OBJECT encloses the other OBJECT, we would consider it a CONTAINER. Should the situation involve motion, it could be a blend with SOURCE_PATH_GOAL.

Table 3. Mappings between the physical primitive acts, the PP, PART, and CONTAIN primitives of CD, and the related image schemas and spatial primitives.

CD Primitive	Related Image Schema(s)	Related Spatial Primitives
PTRANS	SOURCE_PATH_GOAL	SOURCE, GOAL, PATH, MOVE, DIRECTION
MOVE	SOURCE_PATH_GOAL, PART-WHOLE	SOURCE, GOAL, PATH, MOVE, DIRECTION, PARTS, WHOLE,
GRASP	SUPPORT	ATTRACTION, COMPULSION
PROPEL	FORCE	CONTACT
INGEST	SOURCE_PATH_GOAL, CONTAINMENT, FORCE	IN, BOUNDARY, CONTAINER, SOURCE, GOAL, PATH, MOVE, DIRECTION
EXPEL	SOURCE_PATH_GOAL, CONTAINMENT, FORCE	OUT, BOUNDARY, CONTAINER, SOURCE, GOAL, PATH, MOVE, DIRECTION
PP	OBJECT	
PART	PART-WHOLE	PARTS, WHOLE, CONFIGURATION
CONTAIN	CONTAINMENT	CONTAINER, IN, OUT, BOUNDARY

4.4. PROPEL and Impact

In examining the CD primitive act PROPEL, the general description of an entity applying force to another entity can be seen as corresponding to the general definition of the FORCE schema, which implies a physical interaction between two entities. Based on previous research, we believe it is reasonable to take a broad stance on FORCE, as it appears to be part of many image schematic structures. FORCE is an especially complicated case: some argue it is a non-spatial conceptual structure that impacts image schemas and therefore is not an image schema in its own right [23], while others utilize FORCE as a conceptual primitive [7].

Still others argue that force is available as a group of image schemas and does not appear as an individual case [9,13], and because PROPEL encompasses different types of FORCEs and impacts of one OBJECT on another, it can be generally mapped to the FORCE schema group. In the case of one entity striking or kicking another, a mapping to COMPULSION—one entity being moved by external force with a certain magnitude and a certain direction along a path—could be achieved. In the event "Lisa kicked the ball," the "ball" will be moved by the external force of Lisa's kick, unless it is made of steel or lead and cannot be moved by the impact of the kick; this would result in BLOCKAGE of the movement—one entity encountering another entity or obstacle that blocks or resists the movement of the first—and probably pain or an injury for Lisa.

4.5. Picture Producers and OBJECT*s*

Recognition of objects and parts of objects is essential to all forms of conceptualization of physical acts. While much discussion of CD focuses on the primitive acts [27], CD conceptualizations also have dependent cases and sub-conceptualizations that are objects. For example, CD specifies that a simple conceptualization is specified by one of the primitive acts, but also specifies an actor that performs the act, and often an object of the act. CD calls these actor and object "cases" of the conceptualization "picture producers," and specifies that they, unlike nouns in surface language expressions, must always represent physical objects. Picture producers (PPs) correspond strongly to OBJECTs in image schemas.

Additionally, CD retains a primitive called PART that is used to specify that one object (or PP) is a part or sub-object of another object (or PP). An expression such as `hand <=> PART(John)` in a CD conceptualization modifies the PP `hand` to state that it is an inalienable possession or body part of the PP `John`. PART corresponds strongly to the image schema PART-WHOLE, and is particularly important in specifying conceptualizations involving MOVE, which requires that the object of a MOVE is a body part of the actor of a MOVE. The CD primitive CONTAIN is used to specify containment relations for PPs in CD conceptualizations: an expression such as `frog <=> CONTAIN(box)` represents the state of affairs in which the PP `frog` is contained in the PP `box`. CONTAIN, both in its name and in its function, has strong correspondence with the image schema CONTAINMENT.

5. Conclusion and Future Work

In comparing and contrasting image schemas with CD, we examined the linguistic expressions that constructs in each system are designed to represent, as well as evidence from studies in which human subjects connected physical primitives to their comprehension of sentences. We generally have found strong links between CD primitive acts and both image schemas and the spatial primitives that are theorized as building blocks for composing them.

These comparisons suggest a number of modifications to the arrangement of CD physical primitives. Part of the original goals of CD (and one factor in its design which made it controversial, see [29]) was its designers' insistence on keeping the number of conceptual primitives small, making them abstract and general, and working to devise representations of meaning which combined the primitives in complex ways to represent human perception and cognition of events and situations. This allows natural language understanding systems built from CD to establish relations between natural language expressions and concepts by comparing the complex structures instead of relying on "surface level" relations between words. A redesign which reduces the number of primitives while allowing them to represent meanings through more complex combinations allows for richer relations between conceptual structures. Less primitives may also reduce ambiguity, because it makes it less likely that more than one decomposition represents a particular meaning. These comparisons between CD and image schemas suggest that a redesign of CD could combine the PTRANS and MOVE primitives, or simply eliminate MOVE and use only PTRANS, in correspondence with the SOURCE_PATH_GOAL image schema. It also suggests eliminating INGEST and EXPEL in favor of simply using PTRANS and using CD's CONTAIN primitive to specify a containing object that is the source or the goal of the PTRANS. This would correspond with the image schemas SOURCE_PATH_GOAL and CONTAINMENT.

Findings from this comparison highlight the need to consider how animate and inanimate OBJECTs are discriminated conceptually in image schema research as they are in CD. Furthermore, a difference in the level of granularity of both theories in the representation of meaning can be beneficial to specifying complex and controversial image schemas, especially FORCE. CD addresses some specific types of FORCE that are not considered in the group described in the image schema, and contributes a substantial range of examples from story telling in which those types have been analyzed. In general, the comparison showed the difficulty and the potential of bringing the highly abstract image schemas to the level of physical motion and action. Additionally, while we did find links between PROPEL and GRASP and image schemas such as SUPPORT and FORCE, these connections did not immediately suggest improvements to the CD primitives; further investigation of PROPEL and the family of FORCE schemas will be necessary.

Because this investigation focused on the physical and spatial primitives of CD, extending the correspondence to CD primitives representing thought, memory, and perception (such as MTRANS, "Mental TRANSfer") is an obvious topic of future inquiry. Further investigation will also be required to determine if these modifications to CD actually improve CD as a meaning representation as it is used in natural language understanding and story understanding systems in AI. To provide further supporting evidence, of greater or lesser strength, for the validity of the mapping between CD and image schemas that we have derived, a future study similar to that in [19] could have human subjects perform crowdsourcing tasks with similar sentences, but categorize the main acts and events as image schemas instead of CD primitives. Other future work could apply both CD and image schemas to a larger corpus of natural language examples—automatically or manually—to establish a larger body of evidence of the correspondence between them.

References

[1] B. Bennett and C. Cialone. Corpus guided sense cluster analysis: a methodology for ontology development (with examples from the spatial domain). In *8th Int. Conf. on Formal Ontology in Information Systems (FOIS)*, volume 267, pages 213–226. IOS Press, 2014.

[2] Ellen Dodge and George Lakoff. Image schemas: From linguistic analysis to neural grounding. *From perception to meaning: Image schemas in cognitive linguistics*, pages 57–91, 2005.

[3] Michael G Dyer. *In-Depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension*. MIT Press, Cambridge, MA, 1982.

[4] Vyvyan Evans and Melanie Green. *Cognitive linguistics: an introduction*. Edinburgh University Press, 2006.

[5] Antony Galton. The Formalities of Affordance. In Mehul Bhatt, Hans Guesgen, and Shyamanta Hazarika, editors, *Proc. of workshop Spatio-Temporal Dynamics*, pages 1–6, 2010.

[6] Dagmar Gromann and Maria M. Hedblom. Kinesthetic mind reader: A method to identify image schemas in natural language. In *Proceedings of Advancements in Cogntivie Systems*, 2017.

[7] Maria M. Hedblom, Oliver Kutz, Till Mossakowski, and Fabian Neuhaus. Between contact and support: Introducing a logic for image schemas and directed movement. In *Proceedings of AIXIA*, 2017. Forthcoming.

[8] Maria M. Hedblom, Oliver Kutz, and Fabian Neuhaus. Choosing the Right Path: Image Schema Theory as a Foundation for Concept Invention. *Journal of Artificial General Intelligence*, 6(1):22–54, 2015.

[9] Mark Johnson. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. The University of Chicago Press, Chicago and London, 1987.

[10] Eun-Sol Kim, Kyoung-Woon On, Byoung-Tak Zhang, and Cognitive Robotics Artificial Intelligence Center. Deepschema: Automatic schema acquisition from wearable sensor data in restaurant situations. In *IJCAI*, pages 834–840, 2016.

[11] Werner Kuhn. An image-schematic account of spatial categories. In Stephan Winter, Matt Duckham, Lars Kulik, and Ben Kuipers, editors, *Spatial information theory*, pages 152–168. Springer Berlin Heidelberg, 2007.

[12] Ankit Kumar, Ozan Irsoy, Peter Ondruska, Mohit Iyyer, James Bradbury, Ishaan Gulrajani, Victor Zhong, Romain Paulus, and Richard Socher. Ask me anything: Dynamic memory networks for natural language processing. In *International Conference on Machine Learning*, pages 1378–1387, 2016.

[13] George Lakoff. *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. The University of Chicago Press, 1987.

[14] George Lakoff and Rafael Núñez. *Where Mathematics Comes from: How the Embodied Mind Brings Mathematics Into Being*. Basic Books, New York, 2000.

[15] Laura Lakusta and Barbara Landau. Starting at the end: the importance of goals in spatial language. *Cognition*, 96(1):1–33, 2005.

[16] Pat Langley. Intelligent behavior in humans and machines. *Advances in Cognitive Systems*, 2:3–12, 2012.

[17] Hector J Levesque. On our best behaviour. *Artificial Intelligence*, 212:27–35, 2014.

[18] Steven L Lytinen. Conceptual dependency and its descendants. *Computers & Mathematics with Applications*, 23(2):51–73, 1992.

[19] Jamie C. Macbeth and Marydjina Barionnette. The coherence of conceptual primitives. In *Proceedings of the Fourth Annual Conference on Advances in Cognitive Systems*. The Cognitive Systems Foundation, June 2016.

[20] Jamie C. Macbeth and Sandra Grandic. Crowdsourcing a parallel corpus for conceptual analysis of natural language. In *Proceedings of The Fifth AAAI Conference on Human Computation and Crowdsourcing*. The Association for the Advancement of Artificial Intelligence, October 2017.

[21] Jean M Mandler. How to build a baby: Ii. conceptual primitives. *Psychological review*, 99(4):587, 1992.

[22] Jean M. Mandler. The spatial foundations of the conceptual system. *Language and Cognition*, 2(1):21–44, 2010.

[23] Jean M. Mandler and Cristóbal Pagán Cánovas. On defining image schemas. *Language and Cognition*, pages 1–23, 2014.

[24] Jean Matter Mandler. *The foundations of mind: Origins of conceptual thought*. Oxford University Press, 2004.

[25] Christopher D Manning. Computational linguistics and deep learning. *Computational Linguistics*, 41, 2016.

[26] Roger C Schank. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 3(4):552–631, 1972.

[27] Roger C Schank. *Conceptual Information Processing*. Elsevier, New York, NY, 1975.

[28] Roger C. Schank and Robert P. Abelson. *Scripts, plans, goals and understanding : an inquiry into human knowledge structures*. L. Erlbaum Associates, Hillsdale, N.J., 1977.

[29] Roger C Schank and Christopher K Riesbeck. *Inside Computer Understanding: Five Programs Plus Miniatures*. L. Erlbaum Associates Inc., Hillsdale, NJ, 1982.

[30] Marco Schorlemmer, Roberto Confalonieri, and Enric Plaza. The yoneda path to the buddhist monk blend. In *Proceedings of the Joint Ontology Workshops 2016 Episode 2: The French Summer of Ontology co-located with the 9th International Conference on Formal Ontology in Information Systems (FOIS 2016), Annecy, France, July 6-9, 2016.*, 2016.

[31] Lawrence Shapiro. *Embodied cognition*. New problems of philosophy. Routledge, London and New York, 2011.

[32] Leonard Talmy. The fundamental system of spatial schemas in language. In Beate Hampe and Joseph E Grady, editors, *From perception to meaning: Image schemas in cognitive linguistics*, volume 29 of *Cognitive Linguistics Research*, pages 199–234. Walter de Gruyter, 2005.

[33] Patrick Henry Winston. The genesis story understanding and story telling system: A 21st century step toward artificial intelligence. Technical report, Center for Brains, Minds and Machines (CBMM), 2014.