ImageSchemaNet: A framester graph for embodied commonsense knowledge

Stefano De Giorgis [a,*], Aldo Gangemi [b] and Dagmar Gromann [c]

[a] *FICLIT, University of Bologna, Italy*
E-mail: stefano.degiorgis2@unibo.it
[b] *CNR, Consiglio Nazionale delle Ricerche, Italy*
E-mail: aldo.gangemi@unibo.it
[c] *Center for Translation Studies, University of Vienna, Austria*
E-mail: dagmar.gromann@univie.ac.at

Abstract. Commonsense knowledge is a broad and challenging area of research which investigates our understanding of the world as well as human assumptions about reality. Deriving directly from the subjective perception of the external world, it is intrinsically intertwined with embodied cognition. Commonsense reasoning is linked to human sense-making, pattern recognition and knowledge framing abilities. This work presents a new resource that formalizes the cognitive theory of image schemas. Image schemas are dynamic conceptual building blocks originating from our sensorimotor interactions with the physical world, and enable our sense-making cognitive activity to assign coherence and structure to entities, events and situations we experience everyday. ImageSchemaNet is an ontology that aligns pre-existing resources, such as FrameNet, VerbNet, WordNet and MetaNet from the Framester hub, to image schema theory. This article describes an empirical application of ImageSchemaNet, combined with semantic parsers, on the task of annotating natural language sentences with image schemas.

Keywords: Image schemas, cognitive semantics, frame semantics, commonsense reasoning

1. Introduction

Extracting and representing commonsense knowledge is a broad and challenging area of research, in order to solve reasoning problems related to everyday situations. Commonsense knowledge that we deal with everyday includes topicalized knowledge about socio-cultural dynamics [21] e.g. *water is sold in bottles*, jointly with its general patterns, also known as semantic frames, scripts, scenarios, etc. [22]. For instance, the `CommercialTransaction` frame, involves some buyer, some seller, a possible storing place (e.g. a fridge), a way of payment, etc. The approach of utilizing frames to represent a type of object, event or situation builds upon Fillmore's frame semantics [18], the foundation for FrameNet [3], later formalised in various resources, notably the Framester hub [23] that integrates multimodal knowledge under a formal hat.

---

*Corresponding author. E-mail: stefano.degiorgis2@unibo.it.

Furthermore, everyday commonsense is bound to naive inferential physical patterns in order to operate in open world environments [67], e.g. *water is in the fridge* usually implies that *water is in a bottle, which is in the fridge* [39]. Given the dynamic and flexible character of natural language, automatically understanding those implications for concrete, physical situations is very challenging [6,13,44]. However, automated understanding is further aggravated when the situation is in abstract language, typically through metaphors. For instance, *bottle up*, meaning *to keep emotions inside*, relates to the conceptual metaphors FEELINGS ARE LIQUIDS and BODY IS A CONTAINER, as proposed by cognitive semantics [52] and partially listed in [51]. The state-of-the-art approach to motivate and describe metaphorical projection from the physical realm, e.g. the "bottle", to the non-physical, abstract world, e.g. *bottle up*, is the image schema theory by Johnson [42] and Lakoff [50].

Image schemas have been proposed within the tradition of embodied cognition as conceptual structures that capture sensorimotor experiences and shape abstract cognition, including commonsense reasoning and semantic structures of natural language (see e.g. [58,80]). Image schemas are internally structured gestalts, that is, composed by spatial primitives that make up more complex image schemas as unified wholes of meaning [34,37,58]. For instance, a "bottle" is a CONTAINER with an inside, an outside, and a border containing liquids in the physical world. By way of metaphorical projection these characteristics are captured by the image schema CONTAINMENT that is transferred to the abstract realm of emotions inside a body and linguistically expressed, e.g. to *bottle up*. While their existence in natural language has been studied by means of corpus-based (e.g. [66,69]) and machine learning methods (e.g. [32,33,83]), few approaches to formalize image schemas (e.g. Image Schema Logic [36]) and connect them to existing resources to capture semantics exist.

In this article, we present a formal representation of image schemas as a new layer of the Framester hub, called ImageSchemaNet. Since a major flaw in current image schema theory is the lack of agreement about the lexical coverage of image schemas, we introduce an image-schematic layer linked to FrameNet [4], WordNet [60], VerbNet [74], etc., thereby creating a formal, lexicalized integration of cognitive semantics, enactive theories, and frame semantics. The main contributions of this approach are as follows:

- An image-schematic layer in the Framester hub called ImageSchemaNet that is easy to access by means of a SPARQL endpoint, linking image schemas to existing resources
- A formal and re-usable representation of image schemas as Semantic Web technology in form of an ontological layer
- An explicit representation of the interplay of existing (lexical) semantic and formal resources to interlink commonsense knowledge represented as image schemas to natural language and vice versa
- An empirically evaluated method for semi-automatically identifying image schemas in natural language sentences.

The paper is organized as follows: in Section 2 we summarise embodied cognitive semantics theories, with a focus on image schemas (IS), spatial primitives (SP), and their operationalization; in Section 3 we introduce the basics about IS literature and Frame Semantics; in Section 4 we describe ImageSchemaNet and its vocabulary; in Section 5 we explain the SPARQL queries workflow used to populate ImageSchemaNet; in Section 6 we provide an evaluation setting; in Section 7 we discuss the results of annotating natural language sentences with image schemas using ImageSchemaNet and semantic parsers; finally in Section 8 we draw some conclusion and future developments.

2. Related work

Previous work on image schemas and ontologies focuses on formalizing specific IS, e.g. CONTAINMENT [35], or a specific perspective, e.g. IS as families of micro-theories [8,37], where authors exemplify their perspective or theory based on (possible combinations of) specific image schemas.

In terms of dynamic aspects, Galton [20] and Steedman [76] investigate affordances in relation to image schemas. Langacker's works [54,55] about spatial dimensio, grammar and cognition relate to image-schematic spatial structuring. Affordances as defined by Gibson [31] concern commonsense about the opportunities for action offered by real world objects, environments and roles. Schorlemmer et al. [73] investigate the characterisation of creative

processes in conceptual blending by means of diagrams of image schemas. A diagram is intended here within the context of category theory, and a means to model the internal structure of a categorical object. Such framing of image-schematic diagrams within a category-theoretic model of creative processes seeks to provide a mathematically rigorous and computationally feasible model of image-schematic structures.

The work by Kuhn [48] is relevant in using WordNet to extract the image schematic structure from expressions and concepts, followed by formally representing the extracted image schemas using the Haskell programming language. Walton and Worboys [84] advance on this work by aiming to express how image schemas are not only connected to one another, but can be combined to form complex conceptualisations.

Several approaches have used image schemas to model events and scenarios (e.g. [1,7]) starting from compositionality of IS, like OBJECT, PATH and CONTACT to obtain more complex ones like BLOCKAGE, BOUNCING and BLOCKED_MOVEMENT, introducing temporal dimension.

Other formal work includes how to structure IS as families or clusters of similar concepts (e.g. [10,37,72]). Bennett and Cialone [5] take up this idea of clustering and analyze occurrences of CONTAINMENT in biological textbooks in order to propose method based on sense clusters for semi-automatic construction of a spatial ontology from natural language. It focuses on senses as contextualized interpretations, and expresses them within RCC-8 [70] to formally represent different spatial configurations of spatial primitives within the context of CONTAINMENT.

The formal language Image Schema Logic ISL^M [36] provides a more complete formalization of image schema theory. It brings together RCC-8 [70], Qualitative Trajectory Calculus [85], cardinal direction, and linear temporal logic [47,71]. ISL^M was exemplified with the formalization of SUPPORT and CONTACT, and later applied to CONTAINMENT [35] as well. While ISL^M provides strong formalization for (combinations of) image schemas [49], a direct link to natural language and proper lexical coverage was still missing.

ImageSchemaNet is built on ISAAC [15], the Image Schema Abstraction And Cognition ontology[1] as described in Section 4. While ISAAC creates an OWL network of existing IS theories, ImageSchemaNet integrates IS and SP entities within a new ontology, linking them to lexical and factual knowledge bases. In this work, we perform an experimental evaluation of ImageSchemaNet by automatically annotating natural language sentences with image schemas.

Related work in this direction has mostly focused on identifying image schemas in natural language by means of clustering verb-preposition pairs with noun vectors [33], also in a multilingual setting [32]. An extension of this traditional machine learning approach to include word embeddings has been proposed by Wachowiak [82]. One approach that relies on the Image Schema Repository [41], also used in the experimental setting of this article, is a fully automated method of classifying natural language expressions into image schema categories by fine-tuning a pre-trained neural language model [83]. While the results, especially of transferring the learned knowledge to other languages, are promising, there is still room for improvement. For instance, one short-coming of the previous approach is that it can only predict one image schema per natural language expression because of the nature of the dataset, while multiple image schemas frequently co-occur in a natural language sequence.

Our work departs from previous research in linking to Framester, since it operationalizes image schemas as a new layer on top of frame-based knowledge extracted from text. We include testing on a small evaluation dataset, using full-fledged Semantic Web techniques to design ImageSchemaNet. In contrast to clustering and neural approaches, the method for annotating natural language sequences with image schemas is fully explainable, since we keep track of lexical units, their related frames, and the links between frames and formalized image schemas. Furthermore, we can identify more than one image schema per sequence when applicable.

3. Preliminaries

Two important theoretical pillars for ImageSchemaNet are image schemas and frames. Prior to detailing our approach, we first define IS and their relation to spatial primitives. Secondly, we specify frames and their representation in the Framester hub.

---

[1]Available on GitHub here: https://github.com/StenDoipanni/ISAAC.

3.1. Defining image schemas

According to Johnson's famous definition, "an image schema is a recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience" [42]. For instance, playing with shape puzzles[2] as an infant, represents early experiences of spatial boundedness and CONTAINMENT. IS are directly meaningful experiential gestalts [43], that is, they are internally structured compositions of parts to form coherent, uniform wholes. These repeatedly experienced structures are considered to shape higher-level abstract cognition, such as language and commonsense reasoning. Although previous works, also providing forms of neural grounding [16], have proposed an IS inventory [11,58], a generally agreed upon final list is still open to debate and Johnson's [42] caveat still stands:

> What is important is that these recurrent patterns are relatively few in number, they are not propositional in the Objectivist sense, and yet they have sufficient internal structure to generate entailments and to constrain inferences (and thus to be propositionally elaborated).

To give an example, *He just sails through life* depicts life as a CONTAINER, *through* which we travel on our PATH. By way of metaphorical projection, structures of physical source domains can be mapped onto abstract target domains, e.g. the inside (being alive), outside (not being alive), and boundary (birth, death) [46]. Johnson [42] and Lakoff [50,53] provided many linguistic and sensorimotor examples, as well as related high-level entailments without, however, formalizing their theory.

To provide a more formal account, Hedblom et al. [37] propose the unified metalanguage Distributed Ontology, Modeling and Specification Language (DOL) [62] to represent shared gestalt structures of seemingly unrelated image schemas as a family, that is, a set of interlinked theories. Such a gestalt grouping of experiential structures implies a distinction between primitive and complex types. To this end, the theory of Mandler and Pagán Cánovas [58], rooted in developmental psychology, was adopted to distinguish between spatial primitives, image schemas, and conceptual integrations. Spatial primitives are the very first preverbal building blocks that infants form to quickly compose more complex structures, i.e., spatial primitives are the parts that compose coherent unified wholes. These wholes (or spatial events) built from spatial primitives are image schemas. Finally, conceptual integration refers to the inclusion of non-spatial elements, such as emotions. Hedblom et al. [37] take up this initial definition and depict spatial primitives as roles participating in a frame's image schema (e.g. the spatial primitives SOURCE, PATH and GOAL are roles of the SOURCE_PATH_GOAL image schema). Thereby, image schematic-structures, primitive or not, can be grouped based on experiential gestalt family resemblances. This initial formalisation is later extended as Image Schema Logic (ISL^M) [36].

The natural language definitions provided in the works introduced in this section were taken as the basis to propose our ontological representation of image schemas. Instead of relying on first order logic or different calculi for this representation, we focus on a practical frame-based structure and a strong connection to lexical resources. The following presents a list of IS, for which ImageSchemaNet currently provides both a formalization and a lexical coverage, with natural language definitions from IS literature:

– CONTAINMENT: an experience of boundedness, entailing an interior, exterior and a boundary [42].
– CENTER_PERIPHERY: the experience of objects or events as central, while others are peripheral or even outside [29,30]. The periphery depends on the center but not vice versa [50].
– SOURCE_PATH_GOAL: a source or starting point, goal or endpoint, a series of contiguous locations connecting those two, and movement [42].
– PART_WHOLE: wholes consisting of parts and a configuration of parts [50].
– SUPPORT: CONTACT between two objects in the vertical dimension [57].
– BLOCKAGE: obstacles that block or resist our force; a force vector encountering a barrier and then taking any number of directions [42].

---

[2]A game where objects of specific shapes (e.g. triangle, square) have to be inserted into openings of the same shape, learning in this way which opening is shaped to *contain* which object.

Examples of other frequently discussed IS are among others: CONTACT, SCALE, LINK, BALANCE, OBJECT, SUBSTANCE and CYCLE. These IS are not yet in ImageSchemaNet because, compared to the above list, they are less documented in literature. This is due to various reasons. For example, OBJECT is the abstraction of any sensorimotor experience of any bounded entity of the world, and this would result in considering all physical and, by metaphorical projection, all non-physical entities as activator of OBJECT, which would not generate distinct, relevant knowledge. For other IS there is an ongoing debate on whether they should be considered IS in their own right, e.g. CONTACT is at times seen as a spatial primitive interacting with SUPPORT [57,58] or its right as being an IS itself is debated [36]. Thus, the list above represents the most agreed upon selection of IS from literature at the moment of writing.

3.2. Frame semantics and Framester

To connect our ontologies to linguistic examples of image schemas, we rely on formal representations of frames from FrameNet [64] and MetaNet [24] in Framester. Frames in a most general notion are (cognitive) representations of typical features of a situation. Fillmore's frame semantics [18,19] has been most influential as a combination of linguistic descriptions and characterisation of related knowledge structures to describe cognitive phenomena [12]. Words or phrases are associated with frames based on the common scene they evoke. In FrameNet, frames are also explained as *situation types* and they have been used by Dodge to improve embodied construction grammar coverage [17]. In Framester semantics [22] observed/recalled/anticipated/imagined situations are consequently occurrences of frames.

Fillmore explicitly compares frames to other notions, such as experiential gestalt [52], stating that frames can refer to a unified framework of knowledge or a coherent schematization of experience. Thus, widely acknowledged frames provide a theoretically well founded and practically validated basis for commonsense knowledge patterns.

Framester [22,23] provides a formal semantics for frames in a curated linked data version of multiple linguistic resources (e.g. besides FrameNet, WordNet [60], VerbNet [45], BabelNet [63], etc.), factual knowledge bases (e.g. DBpedia [2], YAGO [78], etc.), and ontology schemas (e.g. DOLCE-Zero [26,27]), with formal links between them, resulting in a strongly connected RDF/OWL knowledge graph.

Framester can be used to jointly query (via a SPARQL endpoint[3]) all the resources aligned to its formal frame ontology.[4] Framester has been used [24] to formalize the MetaNet resource of conceptual metaphors,[5] based on FrameNet frames as metaphor sources and targets, as well as to uncover semantic puzzles emerging from a logical treatment of frame-based metaphors. Yet, an image-schematic analysis of MetaNet was lacking, and can be enabled by a refinement of the FrameNet imagistic foundation, which has only been sketched with respect to Framester's prepositional knowledge [25] starting from [56].

4. ImageSchemaNet structure

ImageSchemaNet relies on ISAAC, the Image Schema Abstraction And Cognition ontology. ISAAC models both formal and semi- or unstructured state-of-the-art IS theories, and proposes an integrated theory combining Johnson's definition [42] of image schemas as gestalt structures, Mandler and Pagán Cánovas spatial primitives conception [58] as "first conceptual building blocks", and Hedblom's IS compositionality [7]. ISAAC uses Framester (and derivatively Fillmore's Frame Semantics) to deliver a reified representation of situations evoked in natural language as occurrences of frames and their foundational IS. It is developed following the NeOn methodology [77] and includes reusage of Ontology Design Patterns [68] as good modeling practices.

ImageSchemaNet reuses the `:bibRef` property from Exuviae – a formal "exoskeleton" for representing epistemological choices when comparing ontologies [14] developed considering OPLA Protégé plugin [75] – which is meant to keep precise reference of the bibliographical and theoretical provenance of each entity and property with the original definition and formal dependencies. In particular, ImageSchemaNet focuses on the `:ImageSchema`,

---

:SpatialPrimitive and :IS_Profile classes from the ISAAC ontology, and introduces the :activates property in order to declare assertions about the activation (i.e., a bodily-schematic evocation) of some image schema or spatial primitive from any entity in the Framester resource.

The ImageSchemaNet ontology is available and can be queried from the Framester endpoint.[6] A detailed documentation about the structure, querying, and evaluation is provided in the following sections and in the Appendix, as well as on the ImageSchemaNet GitHub repository.[7] Albeit importing ISAAC ontology, ImageSchemaNet specifically focuses on providing lexical coverage to the Image Schema Theory, via :activation assertions, which currently cover the following image schemas: CONTAINMENT, CENTER_PERIPHERY, SOURCE_PATH_GOAL, PART_WHOLE, BLOCKAGE, and SUPPORT.

4.1. ImageSchemaNet classes

:ImageSchema   The :ImageSchema class represents the general concept of Image Schema, it is defined using the :bibRef property, quoting literature definitions, and it takes as instances image schemas whose activation is covered in the ImageSchemaNet ontology. Each IS is axiomatized as a gestalt structure, composed by at least 2 spatial primitives, and it is modeled as a kind of conceptual frame.

:SpatialPrimitive   The :SpatialPrimitive class takes as instances the "first conceptual building blocks formed in infancy" as in [58], and represents them as semantic roles. The labels used respectively for IS and SP refers to well established and documented names used in literature, as for the SUPPORT IS, quoting their definition and provenance. When specific "official" names were not already given to entities, which existence was nonetheless implicitly or explicitly stated, we used labels extracted from empirical use case. For instance, in the aforementioned SUPPORT case, while literature is often mentioning examples involving its spatial primitives, no official name was available, and for this reason the SUPPORTER and SUPPORTED SP were introduced from anew.

:IS_Profile   The :IS_Profile class is used as in [38] and [65] to describe the collection of IS which are activated by some entity, sentence, situation or event. One of the relevant future developments stemming from our work is the automatic extraction of the image schema profile and the investigation of the conceptual nature of relations among IS in such a collection. The prominence of one particular IS in a set generated from, for example, a text string, refers to a form of frame compositionality as in [22], which could be determined by syntax as well as discourse structure, depending on term, sentence and text compositionality. The :IS_Profile class is particularly relevant here since it's the class used for our evaluation system as described in Section 6.

4.2. ImageSchemaNet properties

All the activation declarations in ImageSchemaNet are realized via the :activates object property or its subproperties, which specify details about the way, layer, resource and type of activation. The meaning of :activates refers to some element that *activates* the cognitive substratum that is associated with an image schema. For instance, the verb *to contain*, the noun *container*, the frame Containment, and the frame element Container all activate the image schema CONTAINMENT.

For this reasons, the following sub-properties were introduced in the graph:

– :activates: declares the activation from a Framester or Framenet frame to an IS. It is the super-property to all the following properties
– :closeMatchActivation: used for the activation of some IS from entities which have a skos: closeMatch (close alignment declarations from Framester) to a FrameNet frame that activates an IS
– :coreSPActivation, :peripheralSPActivation, :extraThematicSPActivation: used for the activation of spatial primitives from FrameNet frame elements, which are distinguished into core (necessary), peripheral (optional), and extra-thematic (not frame-specific)

- :lexicalSenseActivation: used for lexical entities directly evoking spatial primitives or image schemas. This property represents activation based on: 1) very accurate manually verified alignments; 2) alignments inferred from logical rules. For example, the IS activation from WordNet synsets and Framester frames is realized by the query proposed in Appendix A, paragraph "Lexical Entity Activation", which encodes the following rule: if a synset s evokes a frame f that activates an IS i, then s activates i
- :semTypeActivation: used for semantic types used e.g. in FrameNet or VerbNet as selectional restrictions, which activate image schemas or spatial primitives
- :semanticRoleActivation: used for VerbNet arguments, FrameNet frame elements and PropBank roles activating spatial primitives
- :gestaltActivation: activation of an image schema through its spatial primitives

In Section 7 we provide examples of cases from Section 6, in which the assertion of some synset as activator could be acceptable or debatable, and some others in which a specific type of activation has been crucial for detecting the correct IS. Some useful queries to explore the resource using the aforementioned properties can be found in Appendix B. In the following section we describe the SPARQL queries used to retrieve Framester entities activating an image schema or a spatial primitive.

5. ImageSchemaNet grounding pipeline

There is no repository that aligns entities from different semantic layers (lexical units, semantic roles, framal structures, factual entities, etc.), to image schemas and spatial primitives. Moreover, albeit a few references could be found in FrameNet, no lexical grounding has been provided for image schemas.

In order to operationalise ImageSchemaNet, we have created a lexical and factual grounding with a *heuristic abstraction* method. The Framester hub is appropriate to heuristic abstraction, since it implements a formal frame semantics in OWL2, creating interoperability across lexical and factual resources that have been reengineered as knowledge graphs (or directly reused), and aligned to frames and foundational ontologies. The overall architecture of Framester provides then inheritance and unification within the resources integrated in the hub.

Since ImageSchemaNet is an extension of Framester, and image schemas are represented as a special kind of frames activated by other Framester elements, that grounding is straightforwardly performed according to the heuristic abstraction method presented in the following. We firstly provide a simple example of how ImageSchemaNet can be used after being grounded, in order to make the process more intuitive to the reader.

Given the sentence *The Obama administration had entered into an agreement with Iran*, we can: (a) tokenize the sentence into its main elements (*Obama administration*, *enter into agreement with*, *Iran*), (b) collect their senses and (c) disambiguate the contextually valid ones (e.g. Obama_Presidency from DBpedia entity linking, Enter_51010000 from VerbNet disambiguation, Joint Comprehensive Plan of Action from DBpedia, Iran State from DBpedia), (d) retrieve the frames evoked by the senses (Organization, Path_shape, Be_in_agreement_on_Paction, Political_locales, all from FrameNet), and finally (e) retrieve the image schema activated by an entity, a sense, or a frame (nil, SOURCE_PATH_GOAL, nil, nil).[8]

In practice, the heuristic abstraction method reveals that the main image schema activated by the sentence is SOURCE_PATH_GOAL. The inferential structure of SOURCE_PATH_GOAL can further lead us to infer the roles played by an organization, an observed situation, and a political locale.

The exemplified heuristic abstraction can be performed with automated tools, which are evaluated in Section 6. Here the focus is on presenting the hybrid grounding procedure used to populate ImageSchemaNet on top of Framester. We have used the queries listed in Appendix A, following the workflow shown in Appendix E in Fig. 6, together with manual revision. The queries can be reproduced on the Framester endpoint by substituting (manually or programmatically) the insert_variable element with the corresponding entity, as specified in the query description, and by providing the correct prefix as stated in Appendix A. Figure 6 in Appendix E shows the relations among IS/SP activators in different resources in Framester. Entities activating IS/SP, represented as rectangular

---

[8]The "nil" values could be further populated by looking for possible activated spatial primitives.

boxes, are retrieved via SPARQL queries, represented as oval shapes, listed in Appendix A and described in the next sections.

Each query returns a different number of results depending on the entity introduced as input_variable. Figure 6 shows how some entities, being retrieved by some query, are used as input for other queries. Rectangular boxes with no incoming output and only varFor arrow represent those steps that need a human in the loop (e.g. the semantic type query, to produce meaningful results, requires some domain expert which analyzes all results and filters them manually). Ovals with no varFor incoming arrow represent those queries which need a human in the loop providing some input_variable. All other steps can be automatized, although, due to the great amount of knowledge in Framester resource, a manual check could result in higher quality data.

5.1. Frame-driven activation

We have started looking for the frames activating an IS. The first search uses a non-disambiguated lexical unit (e.g. *contain* for the CONTAINMENT IS) to retrieve all the senses and frames evoked by a lexical unit in isolation. For example, for *contain*, the searching process can collect all its senses, and their evoked frames. Based on sense inheritance hierarchies (as available in OWL versions of WordNet and other lexical resources), the search is extended to more specific or more generic senses of e.g. *contain*, so extending the set of evoked frames, and potentially activated IS. This kind of query is exemplified on Framester and can be found in Appendix A at the "Frames Activation Query" paragraph, in Appendix E, Fig. 6 as "Frame activation query" starting node, and in the OWL file as annotation of the :activates object property using the :operationalizedVia annotation property.

However, the amount of senses and related frames can be large, and we need *contextual disambiguation* in order to make it more precise. After performing the query, the selection of frames activating an IS is done manually, and after the iteration of the query for all synonyms and hyponyms, the first phase of frames activation search is closed, and we move to the frame element activation search.

5.2. Frame element-driven activation

Frame element activation concerns the activation of a spatial primitive, and can be performed similarly as with frames. This kind of query is exemplified by focusing on retrieving FrameNet frame elements of type "Core", "Extra-Thematic" and "Peripheral". After performing the query, the selection of frame elements is done manually, using as pivotal the set of frames selected in the step before, possibly enriching the set with further frames, not retrieved by the query in the first step. The query is available in Appendix A, in the "Frame Elements Activation Query" paragraph.

5.3. Lexical unit-driven activation

Activation from lexical material is a substantial part of the heuristic abstraction, and it is generated by automatically querying Framester knowledge base, asking for all the elements (typically WordNet synsets or VerbNet verb senses) that evoke a frame. The query is performed for all the frames retrieved and selected as activators by the Frame Activation query. The heuristic rule here is: if an entity evokes a frame, which activates an IS, than that entity should have some form of activation for the IS. The amount of elements retrieved may be considerable (for some IS, thousands of WordNet synsets). As a consequence, the synsets in the knowledge base are very useful for the coverage they provide in the populated ImageSchemaNet knowledge graph.

Of course, this coverage may contain some noise, since synsets are retrieved by making an inference from previous existing alignments in Framester, which may have different levels of confidence on their turn (there is a manually curated kernel of alignments, while other sets have used various heuristic rules). For example, both the "vase" and "absolutism" terms end up activating the CONTAINER image schema, because some synsets for theoretical concepts of philosophical doctrines or behavioural attitudes (e.g. "absolutism") are aligned in Framester with a skos:closeMatch to the FrameNet frame Containing (because of a lower alignment confidence). In practice, "absolutism as a container" could be considered valid only when conceptual metaphors, e.g. IDEAS ARE OBJECTS, THINKING IS OBJECT MANIPULATION or CATEGORIES ARE BOUNDED REGIONS, are

taken into account. A part of Section 7 discusses this and other critical examples. This query could be found in Appendix A at paragraph "Lexical Elements Activation Query".

5.4. Semantic role-driven activation

Activation assertions to FrameNet frame elements are extended through the multiple sources of semantic roles present in Framester (VerbNet arguments, PropBank roles, WordNet tropes, etc.). Semantic roles in Framester are organized as a complex taxonomy with a small top level that helps integrating them, and getting to the activated IS. The activation of spatial primitives, modelled as semantic roles, is materialised via the `:semantiRoleActivation` property. Roles are retrieved with two queries, starting from top nodes of different graphs, in order to declare the activation of both general and specific roles. The queries are available in Appendix A in the "General Semantic Roles Activation" and "Specific Semantic Roles Activation" sections.

5.5. Semantic type-driven activation

A final important aspect of populating the image schematic activation graph is constituted by the semantics of entity types. For example, FrameNet semantic types `Lateral`, `Leftish`, and `Motion_based_orientation` activate CENTER_PERIPHERY IS, while the frame element `Goal` in frames like `Attaching`, `Body_movement` or `Bringing` has the FrameNet semantic type `Goal`, which activates the GOAL spatial primitive.

Further examples are provided in Section 7. The queries used for semantic type activation assertions include an initial query listing all existing semantic types, followed by a manual exploration of their differences and coverage, resulting in a selection of semantic types activating some IS or SP. Then, a second query is performed, looking for entities filtered by the aforementioned iteration of non-disambiguated lexical units from synsets and their hyponyms, also extracting their semantic type, ending in a final coherence checking between the entities retrieved, their semantic type, and their semantic type activation of an IS or SP. The queries are available at Appendix A, in paragraph "Semantic Type Activation Query".

6. Evaluation

Devising an evaluation method for ImageSchemaNet is not an easy task, since there is no available formal resource featuring IS activation, and no tool able to automatically detect and extract IS from text. Due to this cold start problem, no baseline is proposed. Starting from a corpus of manually annotated sentences, we have performed an evaluation of ImageSchemaNet as an extension to existing automated methods: the end-to-end OpenSesame frame parser, and the hybrid FRED frame-based machine reader. In practice, we have taken an existing corpus of sentences, manually annotated with IS and SP, and we have measured the accuracy of automatically inferring IS based on mapping frames – detected by OpenSesame and FRED – to IS/SP.

6.1. Evaluation setting

The evaluation setting uses an excerpt of the ISCAT dataset[9] [40], and state-of-the-art tools for frame detection from natural language. The ISCAT excerpt has been taken from a cleaned version[10] of the ISCAT online resource. ISCAT is a repository of image schema sentences taken from a large variety of original sources, mainly from literature (e.g. [29,42]), but also from some online sources (e.g. MetaNet, newspaper articles), which are listed in the cleaned repository. The sentences from the excerpt were manually annotated with one IS per sentence.

In this evaluation run, we selected 99 out of 2,478 sample sentences from the cleaned ISCAT excerpt. The main reasons for this extreme reduction of the evaluation set is due to the fact that the sample size per image schema varied considerably and the original dataset only annotates one image schema per sentence. The gold standard is

---

[9]Image Schema Database procured by Jörn Hurtienne.
[10]Available here https://github.com/dgromann/ImageSchemaRepository.

Table 1

Distribution of sentences per image schema

Image Schema	Count
CONTAINMENT	33
CENTER_PERIPHERY	19
SOURCE_PATH_GOAL	17
PART_WHOLE	14
BLOCKAGE	10
SUPPORT	6
Total	99

therefore limited to this unique annotation, but image schemas often co-occur in a single sentence or even phrase, and we were interested in whether the image schemas resulting from the evaluation pipeline would at all be possible for the sentence at hand. For that reason, we had to manually analyze all results, so providing a customised manual evaluation in addition to the automated standard evaluation.

To avoid the introduction of bias, further criteria for selecting the set of sample sentences from the larger cleaned repository were (1) variety of original sources, (2) distribution of image schemas, (3) only image schemas already covered in ImageSchemaNet, (4) mixture of concrete (literal) and abstract (metaphoric) examples, (5) English language only, (6) no syntax linearity restrictions, and (7) no minimum or maximum length of the sentence. In terms of variety of sources, we wanted to ensure that not all samples are derived from the same authors, addressing similar ideas or scenarios.

The evaluation setting uses two frame parsers with entirely different architectures, in order to get a finer assessment of the effect of ImageSchemaNet in the process. The parsers include OpenSesame [59] and FRED [28]. OpenSesame is an end-to-end system focused on frame (and semantic role) detection. Its trained model is based on softmax-margin segmental recurrent neural nets. As with most NLP tools, OpenSesame labels extracted textual segments rather then trying to abstract them as entities and their relations in a knowledge graph. FRED is a hybrid knowledge extraction system with a pipeline including both statistical and rule-based components, aimed at producing RDF and OWL knowledge graphs, with embedded entity linking, word-sense disambiguation, and frame/semantic role detection. The big differences between the two parsers are supposed to make evaluation nuances emerge across parsing paradigms (string-centric vs. entity-centric, informal vs. logical representation).

In order to evaluate ImageSchemaNet, we automatically parse natural language sentences in order to annotate them with frames from FrameNet, and we use these frames to get the activated image schemas as encoded in ImageSchemaNet. We then compare the automated annotations to the manual ones, in order to estimate the accuracy of the process, so providing the first results for explainable image-schema detection in natural language texts. Explainability is granted by the heuristic abstraction applied in ImageSchemaNet and in its usage with the parsers. Example of explainability of IS and SP activation from a FRED graph is provided in Appendix C.

The image schemas covered in Framester and their frequency in the evaluation dataset are reported in Table 1, where we can notice that considerably more examples for CONTAINMENT were included than for the other image schemas. This distribution was selected to reflect the image schema frequency in the original dataset, with by far fewest examples for SUPPORT. Finally, both concrete, i.e., directly relating to a physical or real scenario, and non-physical, i.e., transferring physical aspects to a more abstract scenario, such as MIND AS A CONTAINER, sentences have been selected. The evaluation corpus is available in Appendix D as well as on the ImageSchemaNet GitHub.[11]

6.2. Evaluation procedure

In order to measure the coverage of ImageSchemaNet, an initial trigger frame is required, to evaluate whether that frame leads to the correct image schema profile. To this end, we used natural language sentences as initial

---

[11] https://github.com/StenDoipanni/ISAAC/tree/main/ImageSchemaNet

Table 2

Comparison of retrieved image schemas and frames by OpenSesame and FRED

Parser	Frame types	Frame tokens	IS Types	IS tokens	SP Types	SP Tokens	IS-Annnotated Sentences
OpenSesame	15	57	6	78	2	11	53 / 99
FRED	43	124	6	126	6	20	75 / 99

Table 3

Comparison of weighted F1 scores by parser

Parser	Precision	Recall	Weighted F1	Processed Sentences
OpenSesame	33.95	24.24	26.89	86 / 99
FRED	78.90	39.80	46.06	98 / 99

frame triggers and implemented a two-step pipeline. First, we parse natural language sentences with OpenSesame [79] and FRED [28], which return frames for sentences in the test set. Second, frames are in turn used to query ImageSchemaNet, and identify potentially activated image schema profiles.

To evaluate the final result set from this approach, we first performed automated evaluation (against the original manual IS annotation) applying standard information extraction measures: precision, recall, and weighted F1 score. Each natural language sentence in the evaluation set is annotated with exactly one image schema. However, as discussed in Section 7, and shown in Table 2, multiple image schemas are often co-activated in individual sentences or even phrases, and ImageSchemaNet enables the detection of more than one IS per sentence. For that reason, we have performed a second manual evaluation process to establish whether any returned image schema(s) is plausible for a given sentence.

6.3. Evaluation results

The dataset described in Section 6.1 and listed in Appendix D was used to test our pipeline approach for correct linking between frames detected in natural language and underlying image schema. In Table 2 we present data about IS and SP activation from the selected corpus, noting a better performance from FRED except for the IS type, which was limited by default by the current ImageSchemaNet coverage of six image schemas. In Table 3 we present weighted F1 scores for each frame parser as well as for their confusion matrices in Fig. 1 and Fig. 2.

Table 3 compares the final results of the two parsers on the evaluation set of 99 sentences, where the last column represents the number of sentences which were actually processed. In fact, due to the brevity of sentences, or the metaphorical language or non-linear syntax, in some cases no graph/parsing is produced, resulting in blank result sets, and consequently no image schema detection is possible.

For several sentences, the results of both parsers lead to a co-location of image schemas, i.e., the result set contained more than one image schema, which we counted as correct if the set contained the correct image schema. Please be reminded that multiple image schemas might be correct for a single sentence, and no inconsistency or incompatibility could come from co-location of different image schemas in the same sentence, even if each sentence in this dataset is manually annotated with only one. For the sentences where the result set of the parser contained multiple image schemas, 11 result sets contained the correct manually annotated image schema when using OpenSesame and 15 result sets when using FRED. The overall results of FRED are significantly higher than those of OpenSesame, due to the fact that the former identifies more frames and synsets activating a correct image schema. This is also reflected by the absolute counts of sentences for which frames activating some IS (at least one) were detected, which are 75 for FRED and only 53 for OpenSesame, as shown in last column of Table 2. Table 2 shows that FRED individuates 124 occurrences of frames activating some IS, for 43 different frame types, while OpenSesame activates 57 frame occurrences for 15 types. Note that this number is not the total number of evoked frames, but only those that activate some IS. There are 126 IS activation occurrences (tokens) for FRED and 78 for OpenSesame, while both parsers retrieve at least one IS from each of the six types covered in ImageSchemaNet. Spatial Primitives are mainly activated by verb arguments in OpenSesame: 11 tokens for 2 types, both SP of SOURCE_PATH_GOAL; and by WordNet synsets in FRED: 20 tokens for 6 types, from CONTAINMENT, SOURCE_PATH_GOAL, PART_WHOLE
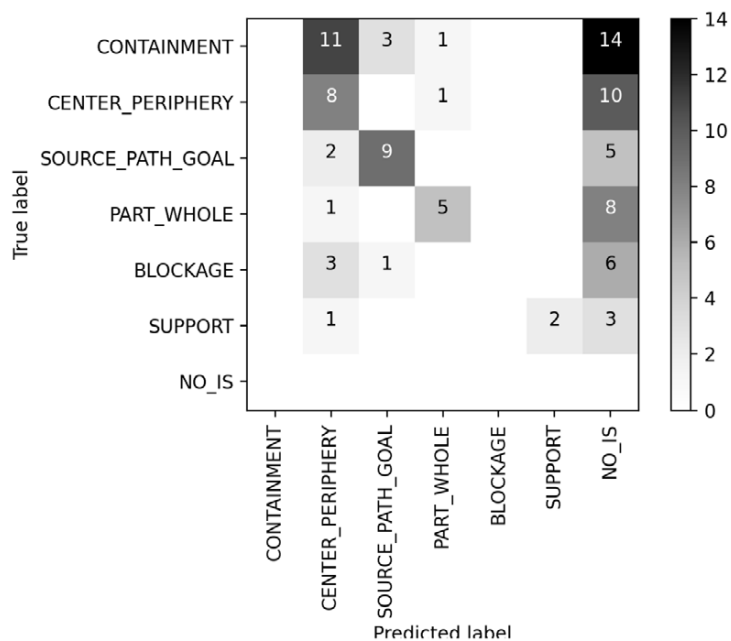
Fig. 1. Confusion matrix for true and predicted image schema labels using OpenSesame.
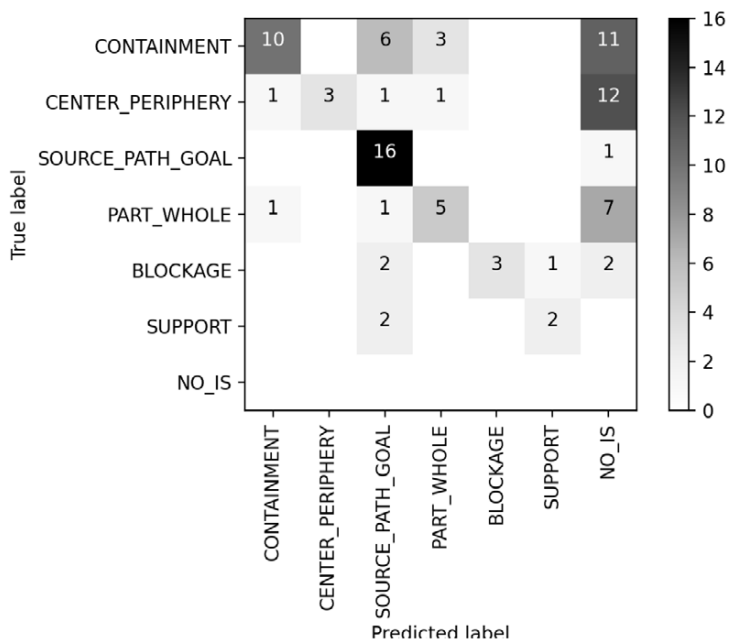

Fig. 2. Confusion matrix for true and predicted image schema labels using FRED.

and CENTER_PERIPHERY. Tables with IS and SP activation by both parsers are available on the ImageSchemaNet Github,[12] jointly with the FRED knowledge graphs for all the processed sentences.

---

[12] https://github.com/StenDoipanni/ISAAC/tree/main/ImageSchemaNet

A confusion matrix is provided for the results derived from each parser. Figure 1 for OpenSesame and Fig. 2 for FRED provide the true labels on the vertical axis and the predicted labels on the horizontal axis. Zero confusion is represented as white space for a clearer visualization and there is no true label for the class "NO_IS" since all sentences in the dataset were annotated with one image schema, however, this class is used to show when no image schema could be returned from the pipeline. When using OpenSesame to detect frames and derive IS (Fig. 1), a total of 94 examples are represented, since 5 sentences lead to a result set of more than one image schema that did not contain the true label. *NO_IS* therefore means that no image schema was returned from the detection process (46 sentences with OpenSesame). The most confusing image schema is CONTAINMENT, possibly since, as shown in Table 1, it is the most common in the corpus. For 11 sentences, frames linking to CENTER_PERIPHERY were returned, e.g. `Locative_relation` for the sentence *There was passion in her eyes*. Overall, a tendency to confuse other image schemas for CENTER_PERIPHERY can be observed. A possible explanation for this tendency comes from the correlation between the semantic extension of these broad IS, and their activators in ImageSchemaNet. Apparently, the semantics of CONTAINMENT, CENTER_PERIPHERY and SOURCE_PATH_GOAL overlaps in examples like *He stepped in the middle of a difficult situation*: "difficult situation" can be conceived as a CONTAINER for the agent; as an area, in which the agent lies in its CENTER; or as a destination of the agent as a moving entity along a SOURCE_PATH_GOAL. This simultaneous activation leads to IS combination (cf. [38]) like CROSSING_OPENING, GOING_IN or GOING_THROUGH.

When using FRED to detect frames and derive IS, the highest number of correct results is obtained for SOURCE_PATH_GOAL (Fig. 2). However, there was understandable confusion across this image schema and CONTAINMENT, when frames, such as `Motion`, were returned, e.g. for the sentence *The whole situation spiraled out of control*. This confusion is understandable, since CONTAINMENT and SOURCE_PATH_GOAL are frequently collocated as movements in and out of a CONTAINER. Spatial primitives are activated by WordNet synsets: 76 tokens for 45 types; and FrameNet frame elements: 15 tokens for 8 types activating SOURCE_PATH_GOAL and PART_WHOLE; all of them are core frame elements. Out of 99, only 91 samples are represented in this confusion matrix, because the remaining 8 sentences lead to a result set of more than one image schema that does not contain the true label.

Even though there is room for improvement, as shown in Table 3 these results show that this idea of interlinking frames with an image-schematic layer in Framester is promising. The main bottleneck at the moment is the frame parser. For instance, a strong preposition to frame detection component in FRED could drastically improve these results, since prepositions are currently considered in the tool only as features to detect or generate roles, however, they provide a very strong indicator for spatial language and type of image schema (see also [25]).

The evaluation dataset is available in Appendix D, while the OpenSesame parsing file, FRED knowledge graphs generated from text, and manual IS and SP detection files can be found at the ImageSchemaNet GitHub.[13]

We have manually inspected the returned image schemas with respect to whether (a) the returned image schema that does not correspond to the original gold standard label could be correct, and (b) whether several returned image schemas actually apply to the sentence at hand. For instance for (a), the expression *We are approaching the end of the year* is labeled with CENTER_PERIPHERY, however, clearly shows a collocation with SOURCE_PATH_GOAL. And for (b), for instance, *My symptoms went away* is labeled as CENTER_PERIPHERY. FRED parser, as shown in Fig. 3, detects three frames: `Motion`, `Travel` and `Departing`. All of them activate SOURCE_PATH_GOAL but `Departing` also activates CENTER_PERIPHERY, which is the label from the ISCAT repository. OpenSesame, on the contrary, as shown in Fig. 4, detects only `Motion` from the verb *go*, but recognizes the `Motion` frame element `Distance`, which has a `:coreSPActivation` towards PATH and PERIPHERY. Consequently, the `:IS_Profile` according both to FRED and OpenSesame shows a co-activation of SOURCE_PATH_GOAL and CENTER_PERIPHERY.

The method based on OpenSesame showed a preference for CENTER_PERIPHERY irrespective of the gold standard label. In 40 of 53 annotated sequences the method returned one image schema, of which 13 were correct and

---
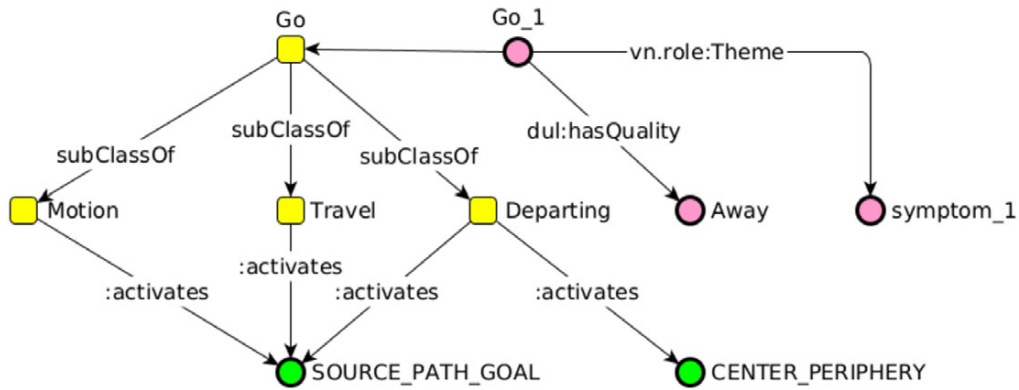
Fig. 3. FRED graph with image schemas activation for *my symptoms went away*.
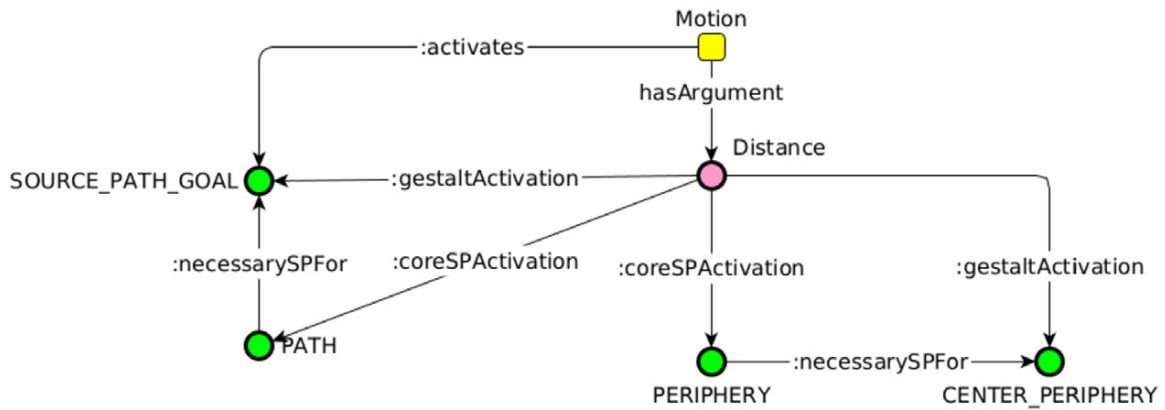
Fig. 4. OpenSesame graph with image schemas activation for *my symptoms went away*.

16 returned reasonable image schemas even though not corresponding to the gold standard label. The remaining 13 sequences of 53 were annotated with more than one image schema, out of which 10 contained the gold standard label, and 9 out of the 13 represented correct image schema collocations.

The method based on FRED frequently returned SOURCE_PATH_GOAL when the label is CONTAINMENT, where in all of these six cases a collocation of both could be observed, e.g. *Try to get out of those commitments*. Interestingly, slight lexical variations would result in the same set of image schemas, e.g. *He took the problem apart piece by piece* and *He tore the problem apart looking for its solution* would both be annotated with BLOCKAGE, SOURCE_PATH_GOAL, and PART_WHOLE, whereas the gold label only considered PART_WHOLE. For FRED, 44 natural language sequences out of 65 annotated only resulted in one image schema, of which 24 corresponded to the gold label and 13 where image schemas that can be considered also correct. In 21 cases the method relying on FRED returned more than one label, of which 14 contained the gold standard label and 11 provided reasonable collocations of up to three image schemas.

7. Discussion

Evaluation results, even from this small corpus, open ample room for discussion. Firstly, some IS, once associated with their lexical grounding, seem to be more intelligible than others. E.g., PART_WHOLE is the foundational gestalt structure at the core of image schemas, to the point that it can lead to its pervasive detection. PART_WHOLE is primarily activated by the conceptual frame `PartWhole`, but other frames seem eligible, e.g., `BodyParts`. In this

case, the large lexical evocation of `BodyParts` would activate PART_WHOLE, as described in Section 5, however, this activation would result in cases like "liver" activating PART_WHOLE. This might be acceptable, because a liver is a part of the human body, but it might sound less relevant to include being a body part as a core schematic semantics of all sentences about livers. A strategy we implement is, depending on the case, to use as activators both a frame and its lexical grounding, or only the frame, or even selected synsets only. This is possible thanks to the ImageSchemaNet object property diversification.

Another example for discussion is CONTAINMENT. Some abstract concepts and doctrines, e.g., *humanism*, evoke the `Containing` frame, hence activating CONTAINMENT IS from ImageSchemaNet. In this experiment, we have followed the FrameNet-based heuristics, and indeed such abstract concepts might be likely used as CONTAINERs in many contexts. This hypothesis could be tested on a larger lexical corpus, including also longer texts, and lexical data can be analyzed in their different aspects, considering also e.g., their semantic type.

Referring to the initial example "water is sold in bottles", an autonomous agent operating in uncertain conditions being able to make inferences starting from the semantic type of an entity, could be able to make, starting from the lexical unit, the inference that, if the WordNet synset `water-noun-1` has a `:semTypeActivation` of SUBSTANCE, then, in order to be moved on purpose and in its integrity, it is necessarily contained in some CONTAINER. In this specific case, the `waterbottle-noun-1` synset has a `:lexicalSenseActivation` to CONTAINER.

Finally, some activations are intrinsic to the commonsense semantics of a frame or lexical unit. It is the case of the `Storing` and `Ingestion` frames, which in three occasions are the sole elements that allow the correct detection of CONTAINMENT. However, in the sentence *He tore the problem apart looking for its solution* we face a false positive, since the correct activation of PART_WHOLE is not due to "tear" or "apart" or a combination of both, but stems from a wrong disambiguation of "solution" in a chemistry-related sense of a compound of particles.

8. Conclusions and future work

We presented ImageSchemaNet, a resource of more than 40,000 triples, which formalizes image schemas with a Framester semantics, so providing an image-schematic layer to FrameNet, MetaNet, WordNet, VerbNet, and other resources in the Framester hub. ImageSchemaNet has been built starting from image schema definitions and examples in literature, and provides lexical grounding (i.e., lexical activators) for the detection of image schemas or spatial primitives. Activation is retrieved via SPARQL queries on the Framester hub. ImageSchemaNet allows nontrivial image schema profile extraction from various semantic layers, including disambiguated natural language units from multiple semantic resources, semantic roles, frames, semantic types, and individual entities. This extraction has been exemplified in an empirical evaluation of annotating natural language sentences with frame parsers and ImageSchemaNet.

As future work, we plan to extend the coverage of ImageSchemaNet to all image schemas in literature, e.g. VERTICALITY, SCALE, etc. As a direct consequence, other than a quantitative improvement of the resource, this extension would enable further investigation on relations among image schemas, in order to clarify possible taxonomic, lexical, functional, mereological and usage relations between IS, bringing greater clarity on frame compositionality and the related underlying commonsense reasoning. Furthermore, the ontological structure allows to include multimodal data about image schematic knowledge, such as Mittelberg's work [61].

From an operational perspective, we plan to realize a fine tuning of the FRED tool, with a focus on image schema detection and image schema profile extraction from natural language. One such improvement would be the consideration of prepositions in the parsing process, which are currently underexploited. We also envisage to integrate recently proposed BERT-based frame detection algorithms (e.g. [81]). Second, we intend to provide a tool that directly proposes image schemas for natural language sequences without a two-step pipeline process as done in our empirical evaluation. This would allow non-trivial spatial commonsense inferences starting from image-schematic reasoning, such as a frame representation for space transitions proposed in [9], with application in the robOntics field (robotics with ontology-based reasoning systems).

Appendix A. Building the resource

A.1. ImageSchemaNet building queries

The following queries can be performed at Framester endpoint: http://etna.istc.cnr.it/framester2/sparql using the uri: <http://www.ontologydesignpatterns.org/ont/is/isnet.owl#>

Frames activation query

```
SELECT DISTINCT ?frame
WHERE {
?frame rdf:type fschema:ConceptualFrame , owl:Class ;
rdfs:subClassOf fschema:FrameOccurrence ;
owl:sameAs ?fnframe .
?fnframe skos:closeMatch ?syn ; a fn15schema:Frame .
?syn wn30schema:senseLabel "insert_variable"@en-us
}
```

Frame elements activation query

```
SELECT DISTINCT ?corefe ?etfe ?perife
WHERE {
{ ?frame1 fn15schema:hasFrameElement ?corefe .
?corefe
<https://w3id.org/framester/framenet/tbox/FE_coreType> "Core" ^^xsd:string .
FILTER(regex(?corefe, "insert_variable", "i")) }
UNION
{ ?frame2 fn15schema:hasFrameElement ?etfe .
?etfe <https://w3id.org/framester/framenet/tbox/FE_coreType>
"Extra-Thematic"^^xsd:string .
FILTER(regex(?etfe, "insert_variable", "i")) }
UNION
{ ?frame3 fn15schema:hasFrameElement ?perife .
?perife <https://w3id.org/framester/framenet/tbox/FE_coreType>
"Peripheral"^^xsd:string .
FILTER(regex(?perife, "insert_variable", "i")) }
}
```

Lexical elements activation query

```
SELECT DISTINCT ?framestersyn ?wnsyn
WHERE { <insert_frame_uri>
<https://w3id.org/framester/schema/subsumes> ?framestersyn .
?framestersyn a <https://w3id.org/framester/schema/WnSynsetFrame> ;
<https://w3id.org/framester/schema/unaryProjection> ?wnsyn .
}
```

General semantic roles activation query

```
SELECT DISTINCT ?argument ?fe ?gfe ?genRole ?genArg ?tropeRole ?semRole
{ GRAPH ?g {
{ ?argument a <https://w3id.org/framester/vn/schema/Argument>.
FILTER(regex(?argument, "insert_variable", "i")) }
UNION
{ ?fe a <https://w3id.org/framester/framenet/tbox/FrameElement> .
FILTER(regex(?fe, "insert_variable", "i")) }
UNION
{ ?gfe a <https://w3id.org/framester/framenet/tbox/GenericFE> .
FILTER(regex(?gfe, "insert_variable", "i")) }
UNION
{?genRole a <https://w3id.org/framester/schema/GenericRole> .
FILTER(regex(?genRole, "insert_variable", "i")) }
UNION
{ ?genArg a <https://w3id.org/framester/vn/schema/GenericArgument> .
FILTER(regex(?genArg, "insert_variable", "i")) }
UNION
{ ?tropeRole a
<https://w3id.org/framester/wn/wn30/wordnet-verbnountropes/TropeRole> .
FILTER(regex(?tropeRole, "insert_variable", "i")) }
UNION
{ ?semRole a <https://w3id.org/framester/schema/semanticRole> .
FILTER(regex(?semRole, "insert_variable", "i")) }
} }
```

Specific semantic roles activation query

```
SELECT DISTINCT ?x ?coreRole ?y ?arg ?z ?fe ?k ?role ?s ?necRole ?q ?optRole
?r ?vnRole
{ GRAPH ?g {
{ ?x <https://w3id.org/framester/schema/coreRole> ?coreRole .
FILTER(regex(?coreRole, "insert_variable", "i")) }
UNION
{ ?y <https://w3id.org/framester/vn/schema/hasArgument> ?arg .
FILTER(regex(?arg, "insert_variable", "i")) }
UNION
{ ?z <https://w3id.org/framester/framenet/tbox/hasFrameElement> ?fe .
FILTER(regex(?fe, "insert_variable", "i")) }
UNION
{ ?k <https://w3id.org/framester/pb/pbschema/hasRole> ?role .
FILTER(regex(?role, "insert_variable", "i")) }
UNION
{ ?s <https://w3id.org/framester/schema/necessaryRole> ?necRole .
FILTER(regex(?necRole, "insert_variable", "i")) }
UNION
{ ?q <https://w3id.org/framester/schema/optionalRole> ?optRole .
FILTER(regex(?optRole, "insert_variable", "i")) }
UNION
{ ?r <https://w3id.org/framester/schema/vnRole> ?vnRole .
FILTER(regex(?vnRole, "insert_variable", "i")) }
```

```
} }
```

Semantic type query

```
SELECT DISTINCT ?entity ?semanticType
WHERE {
?entity
<https://w3id.org/framester/framenet/tbox/hasSemType> ?semanticType .
FILTER(regex(?entity, "insert_variable", "i"))
}
```

Appendix B. Exploring the resource

Some useful queries which show how the resource can be explored. A desired prefix can be substituted to "isnet:" declaring the uri: <http://www.ontologydesignpatterns.org/ont/is/isnet.owl#> on the Framester enpoint.

ASK   Query to ask if some entity is a lexical activator of some spatial primitive which is necessary to the image schema CONTAINMENT.

```
ASK
?entity isnet:lexicalSenseActivation ?sp .
?sp ^isnet:necessarySP isnet:CONTAINMENT .
```

SELECT   Query to retrieve all the image schemas and spatial primitives in ISAAC ontological module.

```
SELECT DISTINCT ?is ?sp
WHERE {
{ ?is a isaac:ImageSchema . }
UNION
{ ?sp a isaac:SpatialPrimitive . }
}
```

Query to retrieve all the entities, image schemas and spatial primitives for which some entity is a lexical activator of some SP which is a necessary SP to some IS.

```
SELECT DISTINCT ?entity ?is ?sp
WHERE {
?entity isnet:lexicalSenseActivation ?sp .
?sp ^isnet:necessarySP ?is .
FILTER(regex(?entity, "insert_variable", "i")) }
```

CONSTRUCT   Query to simulate the image-schema-profile extraction starting from a single non disambiguated lexical unit. The query can be executed by replacing each "insert_variable" with the same lexical unit.

```
CONSTRUCT { [] isnet:ISProfile ?isLex , ?coresp , ?perisp , ?etsp , ?semis ,
?rolesp . }
WHERE {
{ ?x1 isnet:lexicalSenseActivation ?isLex .
FILTER(regex(?x1, "insert_variable", "i")) }
UNION
{ ?x2 isnet:coreSPActivation ?coresp .
FILTER(regex(?x2, "insert_variable", "i")) }
UNION
{ ?x3 isnet:peripheralSPActivation ?perisp .
```

```
FILTER(regex(?x3, "insert_variable", "i")) }
UNION
{ ?x4 isnet:extraThematicSPActivation ?etsp .
FILTER(regex(?x4, "insert_variable", "i")) }
UNION
{ ?x5 isnet:semTypeActivation ?semis .
FILTER(regex(?x5, "insert_variable", "i")) }
UNION
{ ?x6 isnet:semanticRoleActivation ?rolesp .
FILTER(regex(?x6, "insert_variable", "i")) }
}
```

Appendix C. Explainable image schema detection

Here you find the graph produced by FRED for a sample sentence enriched with image schemas and spatial primitives activation. The last three triples of the graph are automatically produced via SPARQL querying ImageSchemaNet and attached to original FRED graph. The locus of activation is in this way clear and explainable via linking the IS or SP activation directly to the lexical unit from natural language text.

   "My mind was racing and I wasn't able to stop it!"

```
@prefix dul: <http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#> .
@prefix isaac:
  <http://www.ontologydesignpatterns.org/ont/is/isaac_vanilla.owl#> .
@prefix isnet: <http://www.ontologydesignpatterns.org/ont/is/isnet.owl#> .
@prefix ns1: <http://www.ontologydesignpatterns.org/ont/vn/abox/role/> .
@prefix ns2: <http://www.ontologydesignpatterns.org/ont/boxer/boxer.owl#> .
@prefix ns3: <http://www.ontologydesignpatterns.org/ont/boxer/boxing.owl#> .
@prefix ns4: <http://www.ontologydesignpatterns.org/ont/fred/domain.owl#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix vn: <https://w3id.org/framester/vn/vn31/data/> .

ns4:able_1 a ns4:Able ;
  ns2:agent ns4:person_1 ;
  ns2:theme ns4:stop_1 ;
  ns3:hasTruthValue ns3:False .

ns4:race_1 a ns4:Race ;
  ns1:Theme ns4:mind_1 .

ns4:Able rdfs:subClassOf dul:Event .
ns4:Mind rdfs:subClassOf
  <http://www.ontologydesignpatterns.org/ont/d0.owl#CognitiveEntity>,
  <https://w3id.org/framester/data/framestercore/MentalProperty>,
  <https://w3id.org/framester/wn/wn30/instances/supersense-noun_cognition> ;
  owl:equivalentClass
  <https://w3id.org/framester/wn/wn30/instances/synset-mind-noun-1> .

ns4:Race rdfs:subClassOf dul:Event,
  <https://w3id.org/framester/data/framestercore/Motion>,
  <https://w3id.org/framester/data/framestercore/SelfMotion> ;
```

```
  owl:equivalentClass vn:Race_51032000 .

ns4:Stop rdfs:subClassOf dul:Event,
  <https://w3id.org/framester/data/framestercore/ActivityStop>,
  <https://w3id.org/framester/data/framestercore/Halt>,
  <https://w3id.org/framester/data/framestercore/ProcessStop> ;
  owl:equivalentClass vn:Stop_55040100,
  <https://w3id.org/framester/wn/wn30/instances/synset-discontinue-verb-1> .

ns4:stop_1 a ns4:Stop ;
  ns1:Agent ns4:person_1 ;
  ns1:Theme ns4:mind_1 .

ns4:mind_1 a ns4:Mind, owl:Thing ;
  ns4:mindOf ns4:person_1 .
ns4:person_1 a ns4:Person .

<https://w3id.org/framester/data/framestercore/Halt>
  isnet:activates isaac:BLOCKAGE .
<https://w3id.org/framester/data/framestercore/Motion>
  isnet:activates isaac:SOURCE_PATH_GOAL .
<https://w3id.org/framester/data/framestercore/SelfMotion>
  isnet:activates isaac:SOURCE_PATH_GOAL .
```

Appendix D. Evaluation corpus

Figure 5 show the full evaluation corpus with image schema activation, sentence and provenance.

Appendix E. ImageSchemaNet building workflow

Figure 6 shows how ImageSchemaNet was built with a workflow of SPARQL queries extracting image schemas and spatial primitives activators. Rectangular boxes are ontological entities from different resources, which are retrieved via queries as classes of inputs/outputs. Oval shapes denote queries, dashed arrows labeled as varFor denote the usage of entities as a variable in the query that uses them. Dashed arrows labeled as output start from a query and end at entities which are retrieved as output via that query. The vertical black line denotes the end of processes for the final step of populating ImageSchemaNet.

|	IMAGE_SCHEMA_ANNOTATION	LinguisticExamples	[SOURCE]
1			
2	CENTER-PERIPHERY	That's just a peripheral issue.	[Lakoff et al., 1991:13]
3	CENTER-PERIPHERY	This is the core of the matter.	[Lakoff et al., 1991:13]
4	CENTER-PERIPHERY	Put aside for a moment the fact that the man has been in prison.	[Jäkel:158-159]
5	CENTER-PERIPHERY	In the midst of the nation's grief.	[Verbytska 2017]
6	CENTER-PERIPHERY	The chairman has managed to weave together quite a coalition	[Grady 1998]
7	CENTER-PERIPHERY	I can't tell them apart.	[Lakoff et al., 1991:60]
8	CENTER-PERIPHERY	These ideas are closely related.	MetaNet
9	CENTER-PERIPHERY	France has close ties to former colony Tunisia	www.allvoices.com/.../78525367-france-has-close-ties-to-former-col...
10	CENTER-PERIPHERY	Far be it for me to try and teach you your job, but…	[BNC, 2007]
11	CENTER-PERIPHERY	That's a far-fetched idea.	[Tolaas, 1991:216]
12	CENTER-PERIPHERY	I feel close to him.	[Lakoff et al., 1991:155]
13	CENTER-PERIPHERY	He distances himself.	[Lakoff et al., 1991:155]
14	CENTER-PERIPHERY	People who are dear to us should be near us.	[Krzeszowski, 1997:118]
15	CENTER-PERIPHERY	Who's with me?	MetaNet
16	CENTER-PERIPHERY	We are approaching the end of the year.	[Núñez & Sweetser 2006:438]
17	CENTER-PERIPHERY	It is close on midnight.	[Hurtienne & Meschke, 2016]
18	CENTER-PERIPHERY	He was close to tears.	[Hurtienne & Meschke, 2016]
19	CENTER-PERIPHERY	My symptoms went away.	MetaNet
20	CENTER-PERIPHERY	Canada is by far the closest of all our neighbors.	http://www.historians.org/projects/GIRoundtable/Canada/Canada_Intro.htm
21	CONTAINMENT	What obligations have you gotten yourself into?	[Lakoff et al., 1991:33,207]
22	CONTAINMENT	Try to get out of those committments, don't let your boss box you in.	[Lakoff et al., 1991:33,207]
23	CONTAINMENT	Have they entered into an agreement yet?	[Lakoff et al., 1991:33]
24	CONTAINMENT	My morning has been very full.	[Grady 1998]
25	CONTAINMENT	Who put that idea in your head?	[Jäkel:156-157]
26	CONTAINMENT	Keep it in the back of your mind	MetaNet
27	CONTAINMENT	Open your mind to some new thoughts.	[Jäkel:158]
28	CONTAINMENT	I've kept this in my memory for years.	[Jäkel, 2003:168]
29	CONTAINMENT	There was passion in her eyes.	[Lakoff & Johnson, 1980:50]
30	CONTAINMENT	The ship is coming into view.	[Lakoff & Johnson, 1980:30]
31	CONTAINMENT	He's out of sight now.	[Lakoff & Johnson, 1980:30]
32	CONTAINMENT	He's a young man trapped in an old man's body.	[Lakoff et al., 1991:209]
33	CONTAINMENT	He's a beautiful person inside.	[Lakoff et al., 1991:209]
34	CONTAINMENT	She was filled with hatred.	[Raymond:43]
35	CONTAINMENT	The truth had to be dragged out of him.	[Lingua]
36	CONTAINMENT	We're in a mess.	[Lakoff et al., 1991:75]
37	CONTAINMENT	We're in a lot of trouble now.	[Lakoff et al., 1991:75]
38	CONTAINMENT	The solution finally was brought to light.	[Jäkel:164]
39	CONTAINMENT	We fished out a number of unpleasant facts.	[Jäkel:164]
40	CONTAINMENT	I put a lot of energy into washing the windows.	[Lakoff & Johnson, 1980:31]
41	CONTAINMENT	He's in love.	[Lakoff & Johnson, 1980:32]
42	CONTAINMENT	I'm slowly getting into shape.	[Lakoff & Johnson, 1980:32]
43	CONTAINMENT	The limit must be drawn somewhere.	[Tolaas, 1991:216]
44	CONTAINMENT	This goes beyond all bounds.	[Tolaas, 1991:216]
45	CONTAINMENT	The whole situation spiraled out of control.	Ellen
46	CONTAINMENT	There are twenty members in this group.	MetaNet
47	CONTAINMENT	He kept a lid on anger	[Stefanowitsch:19]
48	CONTAINMENT	He's bursting with excitement.	MetaNet
49	CONTAINMENT	There is both cotton and polyester in that shirt.	[Grady 1998:121]
50	CONTAINMENT	let the cat out of the bag	[Clausner & Croft, 1997:265]
51	CONTAINMENT	The wind bellied out the sails.	[Lingua]
52	CONTAINMENT	You're getting off the subject.	[Lakoff & Johnson, 1980:91]
53	CONTAINMENT	They had a falling out.	MetaNet
54	PART-WHOLE	They pieced a theory together.	[Lakoff et al., 1991:114]
55	PART-WHOLE	I'm going to pieces.	[Lakoff & Johnson, 1980:28]
56	PART-WHOLE	He took the problem apart piece by piece.	[Lakoff et al., 1991:200]
57	PART-WHOLE	He tore the problem apart looking for its solution.	[Lakoff et al., 1991:200]
58	PART-WHOLE	Something is missing in that argument.	[Lakoff et al., 1991:138]
59	PART-WHOLE	His thoughts are scattered.	[Lakoff et al., 1991:138]
60	PART-WHOLE	I'm missing a piece of the puzzle.	[Lakoff et al., 1991:138]
61	PART-WHOLE	We are one.	[Lakoff et al., 1991:154]
62	PART-WHOLE	She is my other half.	[Lakoff et al., 1991:154]
63	PART-WHOLE	Five is made up of two plus three.	[Lakoff & Nunez, 2000:65]
64	PART-WHOLE	The Milky Way belongs to a cluster of galaxies.	MetaNet
65	PART-WHOLE	Our galaxy has at least 100 billion planets.	MetaNet
66	PART-WHOLE	The eye possesses several parts.	MetaNet
67	PART-WHOLE	Which car part belongs to which car?	MetaNet
68	SOURCE_PATH_GOAL	Let's meet the future head-on.	[Lakoff & Johnson, 1980:43]
69	SOURCE_PATH_GOAL	The deadline is approaching.	[Lakoff & Johnson, 1999:143]
70	SOURCE_PATH_GOAL	There's going to be trouble along the road.	[Kövecses, 2002:34, Lakoff et al.:146]
71	SOURCE_PATH_GOAL	Things are going my way.	[Lakoff & Johnson, 1999:192]
72	SOURCE_PATH_GOAL	Let's keep moving forward.	[Lakoff, 1990:60]
73	SOURCE_PATH_GOAL	You are falling behind on your studies.	[Macaranas et al. 2012]
74	SOURCE_PATH_GOAL	He went over the edge.	[Lakoff et al., 1991:148]
75	SOURCE_PATH_GOAL	He finally reached his goals.	[Kövecses, 2002:135]
76	SOURCE_PATH_GOAL	It took him hours to reach a state of perfect concentration.	[Lakoff et al., 1991:8]
77	SOURCE_PATH_GOAL	As we travel down life's path. . .	[Lakoff et al., 1991:36]
78	SOURCE_PATH_GOAL	His life has taken a good course.	[Grady 1998:113]
79	SOURCE_PATH_GOAL	I go where my path leads me.	[Grady 1998:113]
80	SOURCE_PATH_GOAL	As I've traveled through life, I've made a lot of friends along the way.	[Grady 1998:113]
81	SOURCE_PATH_GOAL	He has changed his direction in life, and taken a more spiritual path.	[Grady 1998:113]
82	SOURCE_PATH_GOAL	He started going down a bad life path.	MetaNet
83	SOURCE_PATH_GOAL	He's no longer with us.	MetaNet
84	SOURCE_PATH_GOAL	They stopped me from leaving.	[Lakoff & Johnson, 1999:204]
85	SUPPORT	Our conservation program needs your support.	[Grady 1998]
86	SUPPORT	I've really been leaning on my friends the past few months.	[Grady 1998]
87	SUPPORT	The poor in our country need a boost up.	MetaNet
88	SUPPORT	In times of need, the impoverished should be able to lean on our government for support.	MetaNet
89	SUPPORT	Our conservation program needs your support.	[Grady 1998]
90	SUPPORT	I've really been leaning on my friends the past few months.	[Grady 1998]
91	FORCE	Let's try to get around this problem.	[Kövecses, 2002:74]
92	FORCE	We ran into a brick wall.	[Lakoff & Johnson, 1999:189]
93	FORCE	We hit a roadblock.	[Lakoff 1990:58]
94	FORCE	Our project has hit a roadblock.	MetaNet
95	FORCE	We have a lot of hurdles to get over in our relationship.	MetaNet
96	FORCE	Recently enacted policies provide cover for companies engaged in illicit activities.	MetaNet
97	FORCE	My life is going nowhere - I'm just spinning my wheels.	MetaNet
98	FORCE	Break out of your daily routine.	[Lakoff & Johnson, 1999:188]
99	FORCE	Workers of the world throw off your chains!	[Lakoff & Johnson, 1999:188]
100	FORCE	All these rules and regulations are obstacles to innovation.	MetaNet

Fig. 5. Evaluation corpus.
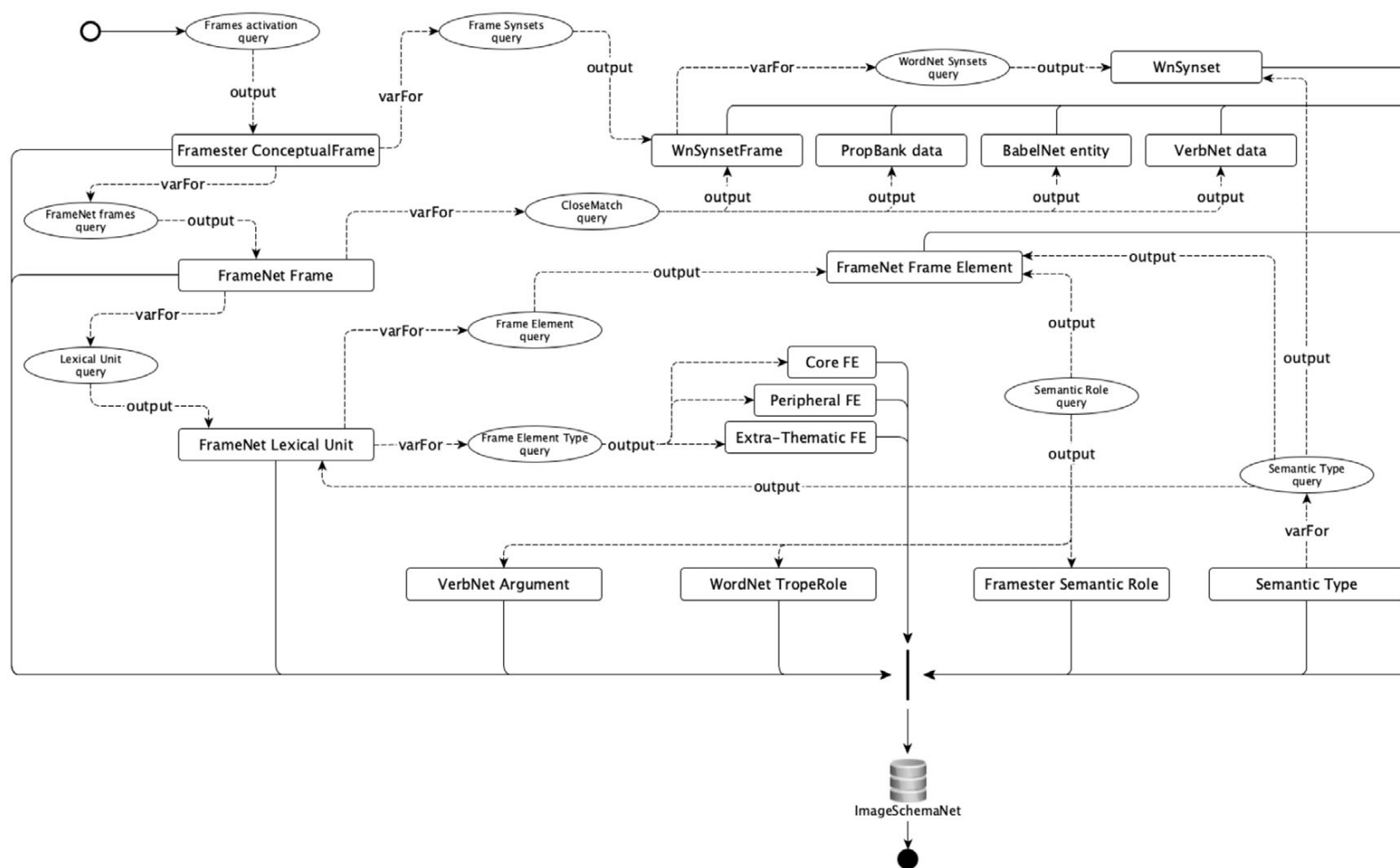
Fig. 6. ImageSchemaNet building workflow. Ovals denote queries, rectangles denote classes of inputs/outputs, dashed arrows represents operations on data reused as variable for another query (`varFor`), or being the output of some query (`output`).

References

[1] R.S. Amant, C.T. Morrison, Y.-H. Chang, P.R. Cohen and C. Beal, An image schema language, in: *Proc. of the 7th Int. Conf. on Cognitive Modeling (ICCM)*, 2006, pp. 292–297.

[2] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak and Z. Ives, Dbpedia: A nucleus for a web of open data, in: *The Semantic Web*, Springer, 2007, pp. 722–735. doi:10.1007/978-3-540-76298-0_52.

[3] C. Baker, FrameNet, present and future, in: *The First International Conference on Global Interoperability for Language Resources*, 2008, pp. 12–17.

[4] C.F. Baker, C.J. Fillmore and J.B. Lowe, The Berkeley framenet project, in: *Proceedings of the 17th International Conference on Computational Linguistics*, Vol. 1, Association for Computational Linguistics, 1998, pp. 86–90. doi:10.3115/980451.980860.

[5] B. Bennett and C. Cialone, Corpus guided sense cluster analysis: A methodology for ontology development (with examples from the spatial domain), in: *FOIS*, 2014, pp. 213–226. doi:10.3233/978-1-61499-438-1-213.

[6] B. Bergen and N. Chang, Embodied construction grammar, in: *The Oxford Handbook of Construction Grammar*, 2013. doi:10.1093/oxfordhb/9780195396683.013.0010.

[7] T.R. Besold, M.M. Hedblom and O. Kutz, A narrative in three acts: Using combinations of image schemas to model events, *Biologically Inspired Cognitive Architectures* **19** (2017), 10–20. doi:10.1016/j.bica.2016.11.001.

[8] K. Bicknell and E. Dodge, Image schemas and force-dynamics in FrameNet, 2004, https://www.klintonbicknell.com/dl/tr.pdf.

[9] G.C. Borchardt, *Thinking Between the Lines: Computers and the Comprehension of Causal Descriptions*, MIT Press, 1994. ISBN 9780262023740.

[10] S. Borgo, P. Hitzler and O. Kutz (eds), *Formal Ontology in Information Systems – Proceedings of the 10th International Conference, FOIS 2018, Cape Town, South Africa, 19–21 September 2018*, Frontiers in Artificial Intelligence and Applications, Vol. 306, IOS Press, 2018, http://ebooks.iospress.nl/volume/formal-ontology-in-information-systems-proceedings-of-the-10th-international-conference-fois-2018. ISBN 978-1-61499-909-6.

[11] A. Cienki, Some properties and groupings of image schemas, in: *Lexical and Syntactical Constructions and the Construction of Meaning*, 1997, pp. 3–15. doi:10.1075/cilt.150.04cie.

[12] A. Cienki, Frames, idealized cognitive models, and domains, in: *The Oxford Handbook of Cognitive Linguistics*, 2007. doi:10.1093/oxfordhb/9780199738632.013.0007.

[13] A. Clark, *Being There: Putting Brain, Body, and World Together Again*, MIT Press, 1998. doi:10.7551/mitpress/1552.001.0001.

[14] S. De Giorgis and A. Gangemi, Exuviae: An ontology for conceptual epistemic comparison, in: *Graphs and Networks in the Humanities 2022 – Proceedings of the 6th International Conference*, Amsterdam, Netherlands, 2022, Accepted for publication, forthcoming.

[15] S. De Giorgis, A. Gangemi and D. Gromann, Introducing ISAAC: The image schema abstraction and cognition modular ontology, in: *Proceedings of the Joint Ontology Workshops 2022*, CEUR Workshop Proceedings, CEUR-WS.org, 2022.

[16] E. Dodge and G. Lakoff, Image schemas: From linguistic analysis to neural grounding, in: *From Perception to Meaning: Image Schemas in Cognitive Linguistics*, 2005, pp. 57–91. doi:10.1515/9783110197532.1.57.

[17] E.K. Dodge, S. Trott, L. Gilardi and E. Stickles, Grammar scaling: Leveraging FrameNet data to increase embodied construction grammar coverage, in: *2017 AAAI Spring Symposium Series*, 2017, https://www.aaai.org/ocs/index.php/SSS/SSS17/paper/viewFile/15298/14529.

[18] C.J. Fillmore, Frame semantics, in: *Linguistics in the Morning Calm*, Hanshin, Seoul, 1982, pp. 111–138. doi:10.1016/b0-08-044854-2/00424-7.

[19] C.J. Fillmore, Frames and the semantics of understanding, *Quaderni di semantica* **6**(2) (1985), 222–254, http://www.icsi.berkeley.edu/pubs/ai/framesand85.pdf.

[20] A. Galton, The formalities of affordance, in: *Proc. of Workshop Spatio-Temporal Dynamics*, M. Bhatt, H. Guesgen and S. Hazarika, eds, 2010, pp. 1–6, https://empslocal.ex.ac.uk/people/staff/apgalton/papers/2010-STeDy-Galton.pdf.

[21] A. Gangemi, Norms and plans as unification criteria for social collectives, *Autonomous Agents and Multi-Agent Systems* **17**(1) (2008), 70–112. doi:10.1007/s10458-008-9038-9.

[22] A. Gangemi, Closing the loop between knowledge patterns in cognition and the semantic web, *Semantic Web* **11**(1) (2020), 139–151. doi:10.3233/SW-190383.

[23] A. Gangemi, M. Alam, L. Asprino, V. Presutti and D.R. Recupero, Framester: A wide coverage linguistic linked data hub, in: *European Knowledge Acquisition Workshop*, Springer, 2016, pp. 239–254. doi:10.1007/978-3-319-49004-5_16.

[24] A. Gangemi, M. Alam and V. Presutti, Amnestic forgery: An ontology of conceptual metaphors, in: *Proceedings of the 10th International Conference, FOIS 2018*, S. Borgo, P. Hitzler and O. Kutz, eds, Frontiers in Artificial Intelligence and Applications, Vol. 306, IOS Press, 2018, pp. 159–172. ISBN 978-1-61499-909-6. doi:10.3233/978-1-61499-910-2-159.

[25] A. Gangemi and D. Gromann, Analyzing the imagistic foundation of framality via prepositions, in: *Proceedings of the Joint Ontology Workshops 2019*, A. Barton, S. Seppälä and D. Porello, eds, CEUR Workshop Proceedings, Vol. 2518, CEUR-WS.org, 2019, http://ceur-ws.org/Vol-2518/paper-CAOS3.pdf.

[26] A. Gangemi, N. Guarino, C. Masolo and A. Oltramari, Sweetening wordnet with dolce, *AI Magazine* **24**(3) (2003), 13. doi:10.1609/aimag.v24i3.1715.

[27] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari and L. Schneider, Sweetening ontologies with DOLCE, in: *International Conference on Knowledge Engineering and Knowledge Management*, Springer, 2002, pp. 166–181. doi:10.1007/3-540-45810-7_18.

[28] A. Gangemi, V. Presutti, D.R. Recupero, A.G. Nuzzolese, F. Draicchio and M. Mongiovì, Semantic web machine reading with FRED, *Semantic Web* **8**(6) (2017), 873–893. doi:10.3233/SW-160240.

[29] R.W. Gibbs Jr., D.A. Beitel, M. Harrington and P.E. Sanders, Taking a stand on the meanings of stand: Bodily experience as motivation for polysemy, *Journal of Semantics* **11**(4) (1994), 231–251. doi:10.1093/jos/11.4.231.

[30] R.W. Gibbs Jr. and H.L. Colston, The cognitive psychological reality of image schemas and their transformations, *Cognitive Linguistics* (1995). doi:10.1515/cogl.1995.6.4.347.

[31] J.J. Gibson, The theory of affordances, in: *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, R. Shaw and J. Bransford, eds, Lawrence Erlbaum, Hillsdale, NJ, 1977, pp. 67–82. doi:10.4324/9781315467931.

[32] D. Gromann and M.M. Hedblom, Body-mind-language: Multilingual knowledge extraction based on embodied cognition, in: *AIC*, 2017, pp. 20–33, http://ceur-ws.org/Vol-2090/paper2.pdf.

[33] D. Gromann and M.M. Hedblom, Kinesthetic mind reader: A method to identify image schemas in natural language, in: *Proceedings of Advancements in Cogntivie Systems*, 2017, https://www.essence-network.com/wp-content/plugins/really-static/static/wp-content/uploads/2015/08/gromann-kinesthetic.pdf.

[34] B. Hampe, Image schemas in cognitive linguistics: Introduction, in: *From Perception to Meaning: Image Schemas in Cognitive Linguistics*, Vol. 29, 2005, pp. 1–14. doi:10.1515/9783110197532.0.1.

[35] M.M. Hedblom, D. Gromann and O. Kutz, In, out and through: Formalising some dynamic aspects of the image schema containment, in: *Proceedings of the Knowledge Representation and Reasoning Track (KRR) at the Symposium of Applied Computing (SAC)*, S. Bistarelli, M. Ceberio, F. Santini and E. Monfroy, eds, 2018, pp. 918–925. doi:10.1145/3167132.3167233.

[36] M.M. Hedblom, O. Kutz, T. Mossakowski and F. Neuhaus, Between contact and support: Introducing a logic for image schemas and directed movement, in: *Proceedings of the 16th International Conference of the Italian Association for Artificial Intelligence (AI*IA 2017)*, LNAI, Springer, Bari, Italy, 2017, pp. 256–268. doi:10.1007/978-3-319-70169-1_19.

[37] M.M. Hedblom, O. Kutz and F. Neuhaus, Choosing the right path: Image schema theory as a foundation for concept invention, *Journal of Artificial General Intelligence* **6**(1) (2015), 21–54. doi:10.1515/jagi-2015-0003.

[38] M.M. Hedblom, O. Kutz, R. Peñaloza and G. Guizzardi, Image schema combinations and complex events, *KI-Künstliche Intelligenz* **33**(3) (2019), 279–291. doi:10.1007/s13218-019-00605-1.

[39] M.M. Hedblom, M. Pomarlan, R. Porzel, R. Malaka and M. Beetz, Dynamic action selection using image schema-based reasoning for robots, 2021, http://ceur-ws.org/Vol-2969/paper33-CAOS.pdf.

[40] J. Hurtienne, Image schemas and design for intuitive use, 2011. doi:10.14279/depositonce-2753.

[41] J. Hurtienne and J.H. Israel, Image schemas and their metaphorical extensions: Intuitive patterns for tangible interaction, in: *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, 2007, pp. 127–134. doi:10.1145/1226969.1226996.

[42] M. Johnson, *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*, The University of Chicago Press, Chicago and London, 1987. ISBN 0-226-40318-1.

[43] G. Kanizsa, *Grammatica del vedere: saggi su percezione e gestalt*, Il mulino, 1980.

[44] M. Kimmel, Analyzing image schemas in literature, *Cognitive Semiotics* **5**(1–2) (2009), 159–188. doi:10.3726/81609_159.

[45] K. Kipper, H.T. Dang, M. Palmer et al., Class-based construction of a verb lexicon, in: *AAAI/IAAI 691*, 2000, p. 696, https://www.aaai.org/Papers/AAAI/2000/AAAI00-106.pdf.

[46] Z. Kovecses, *Metaphor: A Practical Introduction*, Oxford University Press, 2010. doi:10.1017/S0047404503254051.

[47] F. Kröger and S. Merz, *Temporal Logic and State Systems*, 1st edn, Texts in Theoretical Computer Science. An EATCS Series, Springer Publishing Company, Incorporated, 2008. ISBN 3540674012, 9783540674016.

[48] W. Kuhn, An image-schematic account of spatial categories, in: *Spatial Information Theory*, S. Winter, M. Duckham, L. Kulik and B. Kuipers, eds, LNCS, Vol. 4736, Springer, 2007, pp. 152–168. doi:10.1007/978-3-540-74788-8-10.

[49] O. Kutz, N. Troquard, S. Borgo, M.M. Hedblom and D. Porello, The mouse and the ball: Towards a cognitively-based and ontologically-grounded logic of agency, in: *SAC Track: Cognitive Computing (COCO)*, 2017, Submitted. doi:10.3233/978-1-61499-910-2-141.

[50] G. Lakoff, *Women, Fire, and Dangerous Things. What Categories Reveal About the Mind*, The University of Chicago Press, 1987.

[51] G. Lakoff, J. Espenson and A. Schwartz, Master metaphor list, Berkeley, CA: Cognitive Linguistics Group, 1991, http://araw.mede.uic.edu/~alansz/metaphor/METAPHORLIST.pdf.

[52] G. Lakoff and M. Johnson, *Metaphors We Live by*, University of Chicago Press, 1980.

[53] G. Lakoff, M. Johnson et al., *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, Vol. 640, Basic Books, New York, 1999. doi:10.1590/S0102-44502001000100008.

[54] R.W. Langacker, *Foundations of Cognitive Grammar: Theoretical Prerequisites*, Vol. 1, Stanford University Press, 1987. doi:10.1016/0024-3841(90)90017-F.

[55] R.W. Langacker, *Concept, Image, and Symbol: The Cognitive Basis of Grammar*, Mouton de Gruyter, 1990.

[56] K.C. Litkowski and O. Hargraves, The preposition project, in: *Proceedings of the Second ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and Their Use in Computational Linguistics Formalisms and Applications*, 2005, pp. 171–179, https://arxiv.org/ftp/arxiv/papers/2104/2104.08922.pdf.

[57] J.M. Mandler, How to build a baby: II. Conceptual primitives, *Psychological Review* **99**(4) (1992), 587. doi:10.1037/0033-295X.99.4.587.

[58] J.M. Mandler and C. Pagán Cánovas, On defining image schemas, *Language and Cognition* **6**(4) (2014), 510–532. doi:10.1017/langcog.2014.14.

[59] S. Mathôt, D. Schreij and J. Theeuwes, OpenSesame: An open-source, graphical experiment builder for the social sciences, *Behavior Research Methods* **44**(2) (2012), 314–324. doi:10.3758/s13428-011-0168-7.

[60] G.A. Miller, *WordNet: An Electronic Lexical Database*, MIT Press, 1998. doi:10.7551/mitpress/7287.001.0001.

[61] I. Mittelberg, Gestures as image schemas and force gestalts: A dynamic systems approach augmented with motion-capture data analyses, *Cognitive Semiotics* **11**(1) (2018). doi:10.1515/cogsem-2018-0002.

[62] T. Mossakowski, M. Codescu, F. Neuhaus and O. Kutz, The distributed ontology, modeling and specification language – DOL, in: *The Road to Universal Logic*, Springer, 2015, pp. 489–520. doi:10.1007/978-3-319-15368-1_21.

[63] R. Navigli and S.P. Ponzetto, BabelNet: Building a very large multilingual semantic network, in: *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, 2010, pp. 216–225, https://aclanthology.org/P10-1023.

[64] A.G. Nuzzolese, A. Gangemi and V. Presutti, Gathering lexical linked data and knowledge patterns from FrameNet, in: *Proceedings of the Sixth International Conference on Knowledge Capture*, ACM, 2011, pp. 41–48. doi:10.1145/1999676.1999685.

[65] T. Oakley, Image schemas, in: *The Oxford Handbook of Cognitive Linguistics*, 2007, pp. 214–235. doi:10.1093/oxfordhb/9780199738632.013.0009.

[66] A. Papafragou, C. Massey and L. Gleitman, When English proposes what Greek presupposes: The cross-linguistic encoding of motion events, *Cognition* **98**(3) (2006), B75–B87. doi:10.1016/j.cognition.2005.05.005.

[67] M. Pomarlan and J.A. Bateman, Embodied functional relations: A formal account combining abstract logical theory with grounding in simulation, in: *FOIS*, 2020, pp. 155–168. doi:10.3233/FAIA200668.

[68] V. Presutti and A. Gangemi, Content ontology design patterns as practical building blocks for web ontologies, in: *International Conference on Conceptual Modeling*, Springer, 2008, pp. 128–141. doi:10.1007/978-3-540-87877-3_11.

[69] J.A. Prieto Velasco and M. Tercedor Sánchez, The embodied nature of medical concepts: Image schemas and language for PAIN, *Cognitive Processing* (2014). doi:10.1007/s10339-013-0594-9.

[70] D.A. Randell, Z. Cui and A.G. Cohn, A spatial logic based on regions and connection, in: *Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning*, 1992, https://dblp.org/rec/conf/kr/RandellCC92.

[71] M. Reynolds, The complexity of temporal logic over the reals, *Annals of Pure and Applied Logic* **161**(8) (2010), 1063–1096. doi:10.1016/j.apal.2010.01.002.

[72] F. Santibáñez, The object image-schema and other dependent schemas, *Atlantis* **24**(2) (2002), 183–201, https://dialnet.unirioja.es/descarga/articulo/759850.pdf.

[73] M. Schorlemmer, R. Confalonieri and E. Plaza, The Yoneda path to the buddhist monk blend, in: *Proceedings of the Joint Ontology Workshops 2016 Episode 2*, 2016, http://ceur-ws.org/Vol-1660/caos-paper4.pdf.

[74] K.K. Schuler, VerbNet: A broad-coverage, comprehensive verb lexicon, University of Pennsylvania, 2005, http://verbs.colorado.edu/~kipper/Papers/dissertation.pdf.

[75] C. Shimizu, Q. Hirt and P. Hitzler, A Protégé plug-in for annotating OWL ontologies with OPLa, in: *European Semantic Web Conference*, Springer, 2018, pp. 23–27, https://people.cs.ksu.edu/~hitzler/pub2/opla-eswc2018.pdf. doi:10.1007/978-3-319-98192-5_5.

[76] M. Steedman, Formalizing affordance, in: *Proc. of the 24th Annual Meeting of the Cognitive Science Society*, 2002, pp. 834–839. doi:10.4324/9781315782379.

[77] M.C. Suárez-Figueroa, A. Gómez-Pérez and M. Fernández-López, The NeOn methodology for ontology engineering, in: *Ontology Engineering in a Networked World*, Springer, 2012, pp. 9–34. doi:10.1007/978-3-642-24794-1_2.

[78] F.M. Suchanek, G. Kasneci and G. Weikum, Yago: A core of semantic knowledge, in: *Proceedings of the 16th International Conference on World Wide Web*, 2007, pp. 697–706. doi:10.1145/1242572.1242667.

[79] S. Swayamdipta, S. Thomson, C. Dyer and N.A. Smith, Frame-semantic parsing with softmax-margin segmental RNNs and a syntactic scaffold, 2017, arXiv preprint arXiv:1706.09528. doi:10.48550/arXiv.1706.09528.

[80] L. Talmy, The fundamental system of spatial schemas in language, in: *From Perception to Meaning: Image Schemas in Cognitive Linguistics*, B. Hampe and J.E. Grady, eds, Cognitive Linguistics Research, Vol. 29, Walter de Gruyter, 2005, pp. 199–234. doi:10.1515/9783110197532.3.199.

[81] S.-S. Tan and J.-C. Na, Positional attention-based frame identification with BERT: A deep learning approach to target disambiguation and semantic frame selection, 2019, https://arxiv.org/abs/1910.1454.

[82] L. Wachowiak, Semi-automatic extraction of image schemas from natural language, in: *Proceedings of the MEi: CogSci Conference 2020*, Comenius University, Bratislava, 2020, p. 105, https://www.meicogsci.eu/downloads/MEiCogSci-Conference-2020-Proceedings.pdf.

[83] L. Wachowiak and D. Gromann, Systematic analysis of image schemas in NaturalLanguage through explainable MultilingualNeural language processing, in: *Proceedings of the 56th Linguistic Colloquium*, Peter Lang Group, forthcoming.

[84] L. Walton and M. Worboys, An algebraic approach to image schemas for geographic space, in: *Spatial Information Theory*, Lecture Notes in Computer Science, Vol. 5756, Springer, 2009, pp. 357–370, https://link.springer.com/content/pdf/10.1007/978-3-642-03832-7.pdf. doi:10.1007/978-3-642-03832-7_22.

[85] N.V.D. Weghe, A.G. Cohn, G.D. Tré and P.D. Maeyer, A qualitative trajectory calculus as a basis for representing moving objects in geographical information systems, *Control and Cybernetics* **35**(1) (2006), 97–119, http://matwbn.icm.edu.pl/ksiazki/cc/cc35/cc3516.pdf.