Abductive Inference for Interpretation of Metaphors

Ekaterina Ovchinnikova*, Ross Israel*, Suzanne Wertheim[+],

Vladimir Zaytsev*, Niloofar Montazeri*, Jerry Hobbs*

* USC ISI, 4676 Admiralty Way, CA 90292, USA
`{katya,israel,vzaytsev,niloofar,hobbs}@isi.edu`
[+] Worthwhile Research & Consulting, 430 1/2 N Genesee Av., Los Angeles, CA 90036, USA
`worthwhileresearch@gmail.com`

Abstract

This paper presents a metaphor interpretation pipeline based on abductive inference. In this framework following (Hobbs, 1992) metaphor interpretation is modelled as a part of the general discourse processing problem, such that the overall discourse coherence is supported. We present an experimental evaluation of the proposed approach using linguistic data in English and Russian.

1 Introduction

In this paper, we elaborate on a semantic processing framework based on a mode of inference called *abduction*, or inference to the best explanation. In logic, abduction is a kind of inference which arrives at an explanatory hypothesis given an observation. (Hobbs et al., 1993) describe how abduction can be applied to the discourse processing problem, viewing the process of interpreting sentences in discourse as the process of providing the best explanation of why the sentence would be true. (Hobbs et al., 1993) show that abductive reasoning as a discourse processing technique helps to solve many pragmatic problems such as reference resolution, the interpretation of noun compounds, detection of discourse relations, etc. as a by-product. (Hobbs, 1992) explains how abduction can be applied to interpretation of metaphors.

The term *conceptual metaphor* (CM) refers to the understanding of one concept or conceptual domain in terms of the properties of another (Lakoff and Johnson, 1980; Lakoff, 1987). For example, development can be understood as movement (e.g., *the economy moves forward*, *the engine of the economy*). In other words, a conceptual metaphor consists in mapping a *target* conceptual domain (e.g., economy) to a *source* domain (e.g., vehicle) by comparing their properties (e.g., an economy develops like a vehicle moves). In text, conceptual metaphors are represented by *linguistic metaphors* (LMs), i.e. natural language phrases expressing the implied comparison of two domains.

We present a metaphor interpretation approach based on abduction. We developed an end-to-end metaphor interpretation system that takes text potentially containing linguistic metaphors as input, detects linguistic metaphors, maps them to conceptual metaphors, and interprets conceptual metaphors in terms of both logical predicates and natural language expressions. Currently, the system can process linguistic metaphors mapping predefined target and source domains.

We perform an experimental evaluation of the proposed approach using linguistic data in two languages: English and Russian. We select target concepts and generate potential sources for them as described at *github.com/MetaphorExtractionTools/mokujin*. For top-ranked sources, we automatically find corresponding linguistic metaphors. These linguistic metaphors are each then validated by three expert linguists. For the validated linguistic metaphors, we generate natural language interpretations, which are also validated by three experts.

2 Related Work

Automatic interpretation of linguistic metaphors is performed using two principal approaches: 1) deriving literal paraphrases for metaphorical expressions from corpora (Shutova, 2010; Shutova et al., 2012) and 2) reasoning with manually coded knowledge (Hobbs, 1992; Narayanan, 1999; Barnden and Lee, 2002; Agerri et al., 2007; Veale and Hao, 2008).

(Shutova, 2010; Shutova et al., 2012) present methods for deriving paraphrases for linguistic metaphors from corpora. For example, the metaphorical expression "a carelessly *leaked* re-

port" is paraphrased as "a carelessly *disclosed* report". This approach currently focuses on single-word metaphors expressed by verbs only and does not explain the target–source mapping.

The KARMA (Narayanan, 1999) and the ATT-Meta (Barnden and Lee, 2002; Agerri et al., 2007) systems perform reasoning with manually coded world knowledge and operate mainly in the source domain. The ATT-Meta system takes logical expressions that are representations of a small discourse fragment as input; i.e., it does not work with natural language. KARMA focuses on dynamics and motion in space. For example, the metaphorical expression *the government is stumbling in its efforts* is interpreted in terms of motion in space: stumbling leads to falling, while falling is a conventional metaphor for failing.

(Veale and Hao, 2008) suggest to derive common-sense knowledge from WordNet and corpora in order to obtain concept properties that can be used for metaphor interpretation. Simple inference operations, i.e. insertions, deletions and substitution, allow the system to establish links between target and source concepts.

(Hobbs, 1992) understands metaphor interpretation as a part of the general discourse processing problem. According to Hobbs, a metaphorical expression should be interpreted in context. For example, *John is an elephant* can be best interpreted as "John is clumsy" in the context *Mary is graceful, but John is an elephant*. In order to obtain context-dependent interpretations, (Hobbs, 1992) uses abductive inference linking parts of the discourse and ensuring discourse coherence.

3 Metaphor Interpretation System

Our abduction-based metaphor interpretation system is shown in Fig. 1. Text fragments possibly containing linguistic metaphors are given as input to the pipeline. The text fragments are parsed and converted into logical forms (section 3.1). The logical forms are input to the abductive reasoner (section 3.2) that is informed by a knowledge base (section 4). The processing component labelled "CM extractor & scorer" extracts conceptual metaphors from the logical abductive interpretations and outputs scored CMs and Target-Source mappings (section 3.3). The Target-Source mappings are then translated into natural language expressions by the NL generator module (section 3.4).

3.1 Logical Form Generation

A *logical form* (LF) is a conjunction of propositions which have argument links showing relationships among phrase constituents. We use logical representations of natural language texts as described in (Hobbs, 1985). In order to obtain LFs we convert dependency parses into logical representations in two steps: 1) assign arguments to each lemma, 2) apply rules to dependencies in order to link arguments.

Consider the dependency structure for the sentence, *John decided to leave*: `[PRED decide [SUBJ John] [OBJ leave]]`. First, we generate unlinked predicates for this structure: $John(e_1, x_1) \wedge decide(e_2, x_2, x_3) \wedge leave(e_3, x_4)$. Then, based on the dependency labels, we link argument x_1 with x_2, x_3 with e_3, and x_1 with x_4 to obtain the following LF: $John(e_1, x_1) \wedge decide(e_2, x_1, e_3) \wedge leave(e_3, x_1)$.

LFs are preferable to dependency structures in this case because they generalize over syntax and link arguments using long-distance dependencies. Furthermore, we need logical representations in order to apply abductive inference.

In order to produce logical forms for English, we use the *Boxer* semantic parser (Bos et al., 2004). As one of the possible formats, *Boxer* outputs logical forms of sentences in the style of (Hobbs, 1985). For Russian, we use the *Malt* dependency parser (Nivre et al., 2006). We developed a converter turning Malt dependencies into logical forms in the style of (Hobbs, 1985).[1]

3.2 Abductive Inference

In order to detect conceptual metaphors and infer explicit mappings between target and source domains, we employ a mode of inference called weighted abduction (Hobbs et al., 1993). This framework is appealing because it is a realization of the observation that we understand new material by linking it with what we already know.

Abduction is inference to the best explanation. Formally, logical abduction is defined as follows:

Given: Background knowledge B, observations O, where both B and O are sets of first-order logical formulas,
Find: A hypothesis H such that $H \cup B \models O$, $H \cup B \not\models \bot$, where H is a set of first-order logical formulas.

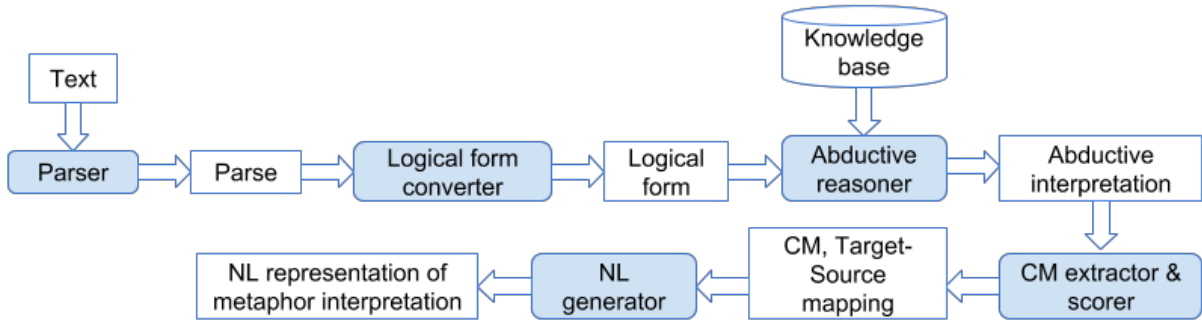[1]The converter is freely available at *https://github.com/eovchinn/Metaphor-ADP*.

Figure 1: Abduction-based metaphor interpretation system.

Typically, there exist several hypotheses H explaining O. To rank hypotheses according to plausibility and select the best hypothesis, we use the framework of weighted abduction (Hobbs et al., 1993). Frequently, the best interpretation results from identifying two entities with each other, so that their common properties only need to be proved or assumed once. Weighted abduction favors those interpretations that link parts of observations together and supports discourse coherence, which is crucial for discourse interpretation.

According to (Hobbs, 1985), metaphor interpretation can be modelled as abductive inference revealing conceptual overlap between the target and the source domain. Consider the abductive interpretation produced for the sentence *We intend to cure poverty*, Fig. 2. In the top line of the figure, we have the LF (cf. Sec. 3.1), where we can see that a *person* (x_1) is the agent for the verbs *intend* (e_1) and *cure* (e_2) and that *poverty* (x_2) is the object of *cure*. In the first box in the next row, we see that *cure* invokes the source concepts of DISEASE, CURE, and DOCTOR, where DISEASE is the object of CURE, and DOCTOR is the subject. In the same row, we see that *poverty* invokes the POVERTY concept in the target domain. Importantly, POVERTY and DISEASE share the same argument (x_2), which refers to *poverty*.

The next row contains two boxes with ellipses, representing long chains of common-sense inferences in the source and target domains of DISEASE and POVERTY, respectively. For DISEASE we know that linguistic tokens such as *illness, sick, disease*, etc. cause the afflicted to experience loss of health, loss of energy, and a general lack of productivity. For POVERTY, we know that tokens such as *poor, broke, poverty* mean that the experiencer of poverty lacks money to buy things, take care of basic needs, or have access to trans-

portation. The end result of both of these frameworks is that the affected individuals (or communities) cannot function at a normal level, with respect to unaffected peers. We can use this common meaning of causing the individual to not function to link the target to the source.

The next three rows provide the mapping from the meaning of the source (CURE, DOCTOR, DISEASE) concepts to the target concept (POVERTY). As explained above, we can consider DISEASE as a CAUSING-AGENT that can CAUSE NOT FUNCTION; POVERTY can be explained the same way, at a certain level of abstraction. Essentially, the interpretation of *poverty* in this sentence is that it causes some entity not to function, which is what a DISEASE does as well. For CURE, we see that *cure* can CAUSE NOT EXIST, while looking for a CAUSING-AGENT (*person*) and an EXISTING DISEASE (*poverty*).

In our system, we use the implementation of weighted abduction based on Integer Linear Programming (ILP) (Inoue and Inui, 2012), which makes the inference scalable.

3.3 CM Extractor and Scorer

The abductive reasoning system produces an interpretation that contains mappings of lexical items into Target and Source domains. Any Target-Source pair detected in a text fragment constitutes a potential CM. For some text fragments, the system identifies multiple CMs. We score Target-Source pairs according to the length of the dependency path linking them in the predicate-argument structure. Consider the following text fragment:

opponents argue that any state attempting to force an out-of-state business to do its dirty work of tax collection violates another state's right to regulate its own corporate residents and their commerce
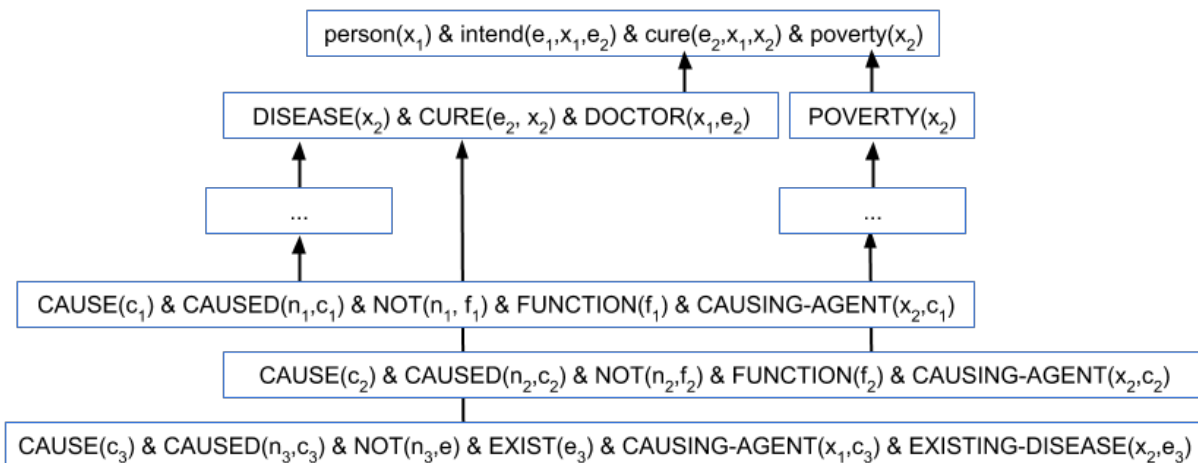
35

Figure 2: Abductive interpretation for the sentence *We intend to cure poverty*.

Suppose our target domain is TAXATION, triggered by *tax collection* in the sentence above. In our corpus, we find realizations of the CM TAXATION is an ENEMY (*fight against taxes*). The lexeme *opponent* triggers the STRUGGLE/ENEMY domain. However, the sentence does not trigger the CM TAXATION is an ENEMY. Instead, it instantiates the CM TAXATION is DIRT (*dirty work of tax collection*). The length of the dependency path between *dirty* and *tax* is equal to 2, whereas the path between *opponent* and *tax* is equal to 9. Therefore, our procedure ranks TAXATION is DIRT higher, which corresponds to the intuition that target and source words should constitute a syntactic phrase in order to trigger a CM.

3.4 NL Representation of Metaphor Interpretation

The output of the abduction engine is similar to the logical forms provided in Fig. 2. In order to make the output more reader friendly, we produce a natural language representation of the metaphor interpretation using templates for each CM. For example, the text *their rivers of money mean they can offer far more than a single vote* would invoke the WEALTH is WATER CM, and the abduction engine would output: LARGE-AMOUNT[river], THING-LARGE-AMOUNT[money]. We then take this information and use it as input for the NL generation module to produce: *"river" implies that there is a large amount of "money"*.

4 Knowledge Base

In order to process metaphors with abduction, we need a knowledge base that encodes the informa-

tion about the source domain, the target domain, and the relationships between sources and targets. We develop two distinct sets of axioms: lexical axioms that encode lexical items triggering domains, and mapping axioms that encode knowledge used to link source and target domains. We will discuss the details of each axiom type next.

4.1 Lexical Axioms

Every content word or phrase that can be expected to trigger a source or target domain is included as a lexical axiom in the knowledge base. For example, the STRUGGLE domain contains words like *war, fight, combat, conquer, weapon,* etc. An example of how a lexical axiom encodes the system logic is given in (1). On the left side, we have the linguistic token, *fight*, along with its part-of-speech, *vb*, and the argument structure for verbs where e_0 is the eventuality (see (Hobbs, 1985)) of the action of fighting, x is the subject of the verb, and y is the object. On the right side, STRUGGLE is linked to the action of fighting, the subject is marked as the AGENT, and the object is marked as the ENEMY.

(1) $fight\text{-}vb(e_0, x, y) \rightarrow STRUGGLE(e_0) \land$
$AGENT(x, e_0) \land ENEMY(y, e_0)$

The lexicon is not limited to single-token entries; phrases can be included as single entries; For example, the ABYSS domain has phrases such as *climb out of* as a single entry. Encoding phrases often proves useful, as function words can often help to distinguish one domain from others. In this case, climbing *out of* something usually denotes an abyss, whereas climbing *up* or *on* usually does not. The lexical axioms also include the POS

for each word. Thus a word like *fight* can be entered as both a noun and a verb. In cases where a single lexical axiom could be applied to multiple domains, one can create multiple entries for the axiom with different domains and assign weights so that a certain domain is preferred over others.

Initial lexical axioms for each domain were developed based on intuition about each domain. We then utilize ConceptNet (Havasi et al., 2007) as a source for semi-automatically extracting a large-scale lexicon. ConceptNet is a multilingual semantic network that establishes links between words and phrases. We query ConceptNet for our initial lexical axioms to return a list of related words and expressions.

4.2 Mapping Axioms

Mapping axioms provide the underlying meanings for metaphors and link source and target domains. All of these axioms are written by hand based on common-sense world knowledge about each target-source pair. For each CM, we consider a set of LMs that are realizations of this CM in an effort to capture inferences that are common for all of the LMs. We consider the linguistic contexts of the LMs and overlapping properties of the target and source domains derived from corpora as described in section 5.1.

We will outline the process of axiomatizing the STRUGGLE domain here. We know that a verb like *fight* includes concepts for the struggle itself, an agent, and an enemy. In the context of a STRUGGLE, an enemy can be viewed as some entity a that attempts to, or actually does, inhibit the functioning of some entity b, often through actual physical means, but also psychologically, economically, etc. The struggle, or fight, itself then, is an attempt by a to rid itself of b so that a can ensure normal functionality. So, given a phrase like *poverty is our enemy*, the intended meaning is that poverty is hindering the functionality of some entity (an individual, a community, a country, etc.) and is seen as a problem that must be fought, i.e. eliminated. In a phrase like *the war against poverty*, war refers to an effort to stop the existence of poverty. These inferences are supported by the overlapping property propositions extracted from English Gigaword as described in Sec. 5.1, e.g., *scourge of X, country fights X, country pulls of X, suffer from X, fight against X*.

To extend the example in (1), consider (2).

Here, we encode a STRUGGLE action, e.g. *fight*, as CAUSE NOT EXIST, the AGENT of the fight as CAUSING-AGENT, and the ENEMY as EXISTING-THING. Then, for a verb phrase like *we fight poverty*, *we* is the AGENT that engages in causing *poverty*, the ENEMY, to not exist.

(2) $STRUGGLE(e_0) \quad \wedge \quad AGENT(x, e_0) \quad \wedge$
$ENEMY(y, 2_0) \rightarrow CAUSE(e_0) \wedge CAUSED(n, e_0) \wedge$
$NOT(n, ex) \quad \wedge \quad EXIST(ex) \quad \wedge \quad CAUSING -$
$AGENT(x, e_0) \wedge EXISTING - THING(y, ex)$

We use 75 mapping axioms to cover the valid LMs discussed in Sec. 5.2. Some interesting trends emerge when examining the core meanings of the LMs. Following (Hobbs, 2005), we found that over 65% of the valid LMs in this study could be explained in terms of causality. The next most prevalent aspect that these metaphors touch upon is that of functionality (nearly 35%), with some of these overlapping with the causality aspect where the meaning has to do with *X causing Y* to *function* or *not function*.

Many of the CMs covered in this study have fairly transparent interpretations based on these ideas of causality and functionality, such as POVERTY is DISEASE, where the main underlying meaning is that a disease causes the sufferer not to function properly. However, for some CMs, the interpretation can be more difficult to pin down. For example, the interpretation of WEALTH is a GAME is quite opaque. Given a sentence such as, *Wealth is a game and you better start playing the game*, there are no obvious connections to concepts such as causality or functionality. Instead, *game* raises such ideas as competition, winning, and losing. In the literal context of a game, the competition itself, who the competitors are, and what it means to win or lose are usually clearly defined, but this is not so when speaking metaphorically about wealth. To derive a meaning of *game* that can apply to wealth, we must look at a higher level of abstraction and define *game* as the instantiation of a positive or negative outcome, i.e. to *win* is to achieve a positive outcome, or gain *wealth*. In the same sentence *play* implies that some voluntary action must be taken to achieve a positive outcome.

For some metaphors, a simple transfer of the source properties to the target does not result in a coherent interpretation at all. Given, for example, the CM POVERTY is a PRICE, one LM from this study is, *poverty is the price of peace*. In this case, the meaning has to do with some notion of

an exchange, where a negative consequence must be accepted in order to achieve a desired outcome. However, the metaphorical meaning of *price* differs from the literal meaning of the word. In literal contexts, *price* refers to an amount of money or goods with inherent value that must be given to acquire something; the buyer has a supply of money or goods that they willingly exchange for their desired item. In the metaphorical sense, though, there often is no buyer, and there is certainly not an inherent value that can be assigned to *poverty*, nor can one use a supply of it to acquire *peace*.

Another issue concerns cultural differences. While writing the axioms to deal with English and Russian source-target pairs we noticed that a majority of the axioms applied equally well to both languages. However, there are some subtle differences of aspect that impact the interpretation of similar CMs across the two languages. Looking again at the WEALTH is a GAME metaphor, the Russian interpretation involves some nuance of a lack of importance about the subject that does not seem to be present in English when using words like *game* and *play*. Note that there may be some notion of carelessness for English (see Sec. 5.3), but for Russian, the notion of being carefree, which is not the same as careless, about wealth has a strong prevalence.

5 Experimental Validation

5.1 Source Generation

Following from the definition of metaphor, the target and the source domain share certain properties. In natural language, concepts and properties are represented by words and phrases. There is a long-standing tradition for considering computational models derived from word co-occurrence statistics as being capable of producing reasonable property-based descriptions of concepts (Baroni and Lenci, 2008). We use proposition stores to derive salient properties of concepts that can be potentially compared in a metaphor.

A *proposition store* is a collection of propositions such that each proposition is assigned its frequency in a corpus. Propositions are tuples of words that have a determined pattern of syntactic relations among them (Clark and Harrison, 2009; Peñas and Hovy, 2010; Tsao and Wible, 2013). For example, the following propositions can be extracted from the sentence *John decided to go to school*:

(*NV John decide*)
(*NV John go*)
(*NVPN John go to school*)
...

We generated proposition stores from parsed English Gigaword (Parker et al., 2011) and Russian ruWac (Sharoff and Nivre, 2011). Given the proposition stores, we generate potential sources for a seed target lexeme l in three steps:

1. Find all propositions P_l containing l.

2. Find all potential source lexemes S such that for each $s \in S$ there are propositions p, p' in the proposition store such that l occurs at position i in p and s occurs at position i in p'. The set of propositions containing l and s at the same positions is denoted by $P_{l,s}$.

3. Weight potential sources $s \in S$ using the following equation:

$$weight_l(s) = \sum_{p \in P_{l,s}} weight_l(t), \quad (1)$$

The source generation procedure and its validations are described in detail at *github.com/MetaphorExtractionTools/mokujin*.[2] In the experiment described below, we generated potential sources for the target domains of POVERTY and WEALTH.

5.2 Linguistic Metaphors Extraction and Validation

For each potential CM, we look for supporting LMs in corpora. A a large number of LMs supporting a particular CM suggests that this CM might be cognitively plausible. We use a simple method for finding LMs. If a target lexeme and a source lexeme are connected by a dependency relation in a sentence, then we assume that this dependency structure contains a LM. For example, in the phrases *medicine against poverty* and *chronic poverty*, the target word (*poverty*) is related via dependency arc with the source words (*medicine*, *chronic*). LMs were extracted from English Gigaword (Parker et al., 2011) and Russian ruWac (Sharoff and Nivre, 2011).

For the generated CMs, we select seed lexemes for target and source domains. We expand the

---

[2]The tools for generating proposition stores and the obtained resources are freely available at *https://ovchinnikova.me/proj/metaphor.html*.

sets of these target and source lexemes with se-
mantically related lexemes using English and Rus-
sian ConceptNet (Speer and Havasi, 2013) and top
ranked patterns from the proposition stores. For
example, the expansion of the lexeme *disease* re-
sults in the following set of lexemes: {*disease,
symptom, syndrome, illness, unwellness, sickness,
sick, medicine, treatment, treat, cure, doctor, ...* }

For each language, we select 20 top-ranked
sources per target. Then we randomly select at
most 10 sentences per each target-source pair.
These sentences are validated by 3 linguist experts
each. For each sentence, the experts are asked if
it contains a metaphor comparing an indicated tar-
get domain with an indicated source domain. The
inter-annotator agreement on the validation task is
defined as the percentage of judgements on which
the three experts agree. Agreement is 81% for En-
glish and 80% for Russian.

Tables 1 and 2 show 10 potential sources per
target with the best agreement. Column ALL pro-
vides the number of sentences per a proposed CM
such that all experts agreed that the sentence con-
tains a metaphor. Column TWO provides the num-
ber of sentences such that any two experts agreed
on, and Column ONE shows the number of sen-
tences such that a single expert thought it con-
tained a metaphor.

target	source	ALL	TWO	ONE
wealth	blood	10	10	10
	water	9	10	10
	drug	9	10	10
	food	9	9	10
	body	9	9	10
	power	8	9	10
	game	8	9	9
	security	7	9	10
	resource	7	7	9
	disease	7	8	9
poverty	war	10	10	10
	abyss	10	10	10
	violence	9	9	10
	price	8	9	9
	location	7	8	8
	disease	7	7	7
	crime	4	5	6
	crop	3	7	9
	terrorism	3	3	5
	cost	2	3	7

Table 1: Validation of English linguistic
metaphors found for potential sources.

5.3 Metaphor Interpretation Validation

Metaphor interpretations were generated for posi-
tively validated linguistic metaphors, as described

		ALL	TWO	ONE
богатство (wealth)	энергия (energy)	10	10	10
	вода (water)	10	10	10
	свобода (freedom)	10	10	10
	власть (power)	9	10	10
	бог (god)	9	10	10
	кровь (blood)	9	10	10
	путь (way)	9	10	10
	игра (game)	8	10	10
	слава (glory)	4	5	5
	товар (ware)	3	8	10
бедность (poverty)	пропасть (abyss)	10	10	10
	враг (enemy)	9	10	10
	болезнь (disease)	9	9	9
	власть (power)	8	10	10
	тело (body)	6	6	6
	боль (pain)	5	10	10
	отчаяние (despair)	5	10	10
	цена (price)	4	4	4
	смерть (death)	3	5	6
	страх (fear)	3	9	10

Table 2: Validation of Russian linguistic
metaphors found for potential sources.

in Sec. 3.4. Each interpretation was validated by
three expert linguists. We calculated strict and
relaxed agreement for the validated data. Strict
agreement is calculated over three categories: cor-
rect (C), partially correct (P), and incorrect (I). Re-
laxed agreement is calculated over two categories:
C/P and I. Partially correct means that the valida-
tor felt that something was missing from the inter-
pretation, but that what was there was not wrong.
Table 3 presents the validation results for both lan-
guages. As can be seen in the table, strict agree-
ment (AgrS) is 62% and 52% and strict system
accuracy (AccS ALL) is 62% and 50% for En-
glish and Russian, respectively. Relaxed agree-
ment (AgrR) results is 93% and 83%, and relaxed
accuracy (AccR ALL) is 91% and 78%.

Validators often marked things as only partially
correct if they felt that the interpretation was lack-
ing some aspect that was critical to the meaning of
the metaphor. A common feeling amongst the val-
idators, for example, is that the interpretation for
people who are terrorized by poverty should in-
clude some mention of "fear" as a crucial aspect
of the metaphor, as the interpretation provided
states only that *"terrorize" implies that "poverty"
is causing "people" not to function*. However, the
end result of "fear" itself is often that the experi-
encer cannot function, as in *paralyzed by fear*.

Tables 4 and 5 contain interpretation system ac-
curacy results by CM. We calculated the percent-
age of LMs evoking this CM that were validated
as C vs. I (strict) or P/C vs. I (relaxed) by all three

	AgrS	AgrR	AccS ALL	AccS TWO	AccS ONE	AccR ALL	AccR TWO	AccR ONE
English	0.62	0.93	0.62	0.84	0.98	0.91	0.97	0.99
Russian	0.52	0.83	0.50	0.76	0.96	0.78	0.93	0.99

Table 3: Validation results for metaphor interpretation for English and Russian.

(ALL), or just two (TWO) validators. In most of the cases, the system performs well on "simple" CMs related to the concepts of causation and functioning (e.g., WEALTH is POWER), cf. section 4, whereas its accuracy is lower for richer metaphors (e.g., WEALTH is a GAME).

target	source	ALL		TWO	
		S	R	S	R
	blood	0.8	1	1	1
	water	1	1	1	1
	drug	0.44	0.78	0.89	0.89
	food	0.89	1	1	1
wealth	body	0.67	0.78	0.78	0.78
	power	1	1	1	1
	game	0.63	1	1	1
	security	0.14	0.88	0.71	1
	resource	1	1	1	1
	disease	0	1	1	1
	war	0.9	0.9	1	1
	abyss	0	0.5	0.4	1
	violence	0	1	0.11	1
	price	0.88	0.88	0.88	1
	location	1	1	1	1
poverty	disease	0.43	0.86	0.86	0.86
	crime	0.75	1	1	1
	crop	1	1	1	1
	terrorism	0	1	0.33	1
	cost	1	1	1	1

Table 4: Accuracy of English interpretations for each CM.

The data used in the described experiments, system output, and expert validations are available at *http://ovchinnikova.me/suppl/AbductionSystem-Metaphor-Validation.7z.*

6   Conclusion and Future Work

The developed abduction-based metaphor interpretation pipeline is available at *https://github.com/eovchinn/Metaphor-ADP* as a free open-source project. This pipeline produces favorable results, with metaphor interpretations that are rated as at least partially correct, for over 90% of all valid metaphors it is given for English, and close to 80% for Russian. Granted, the current research is performed using a small, controlled set of metaphors, so these results could prove difficult to reproduce on a large scale where any metaphor is possible. Still, the high accuracies achieved on both languages indicate

T	source	ALL		TWO	
		S	R	S	R
	энергия (energy)	0.4	0.8	0.9	1
	вода (water)	0	0.9	0.6	0.9
	свобода (freedom)	1	1	1	1
	власть (power)	1	1	1	1
богатство (wealth)	бог (god)	0.67	1	0.89	1
	кровь (blood)	1	1	1	1
	путь (way)	0.78	0.78	0.89	0.89
	игра (game)	0.1	0.2	0.2	0.3
	слава (glory)	0	0.75	0.75	1
	товар (ware)	0	0	0	1
	пропасть (abyss)	0.7	1	1	1
	враг (enemy)	0.56	1	1	1
	болезнь (disease)	0.33	0.89	0.67	1
	власть (power)	0.5	0.5	1	1
бедность (poverty)	тело (body)	0.17	0.17	0.17	0.83
	боль (pain)	1	1	1	1
	отчаяние (despair)	0.6	0.6	1	1
	цена (price)	0.75	0.75	1	1
	смерть (death)	0	0	0.33	1
	страх (fear)	0	1	0.67	1

Table 5: Accuracy of Russian interpretations for each CM.

that the approach is sound and there is potential for future work.

The current axiomatization methodology is based mainly on manually writing mapping axioms based on the axiom author's intuition. Obviously, this approach is subject to scrutiny regarding the appropriateness of the metaphors and faces scalability issues. Thus, developing new automatic methods to construct the domain knowledge bases is a main area for future consideration.

The mapping axioms present a significant challenge as far producing reliable output automatically. One area for consideration is the aforementioned prevalence of certain underlying meanings such as causality and functionality. Gathering enough examples of these by hand could lead to generalizations in argument structure that could then be applied to metaphorical phrases in corpora to extract new metaphors with similar meanings. Crowd-sourcing is another option that could be applied to both axiom writing tasks in order to develop a large-scale knowledge base in considerably less time and at a lower cost than having experts build the knowledge base manually.

References

R. Agerri, J.A. Barnden, M.G. Lee, and A.M. Wallington. 2007. Metaphor, inference and domain-independent mappings. In *Proc. of RANLP'07*, pages 17–23.

J. A. Barnden and M. G. Lee. 2002. An artificial intelligence approach to metaphor understanding. *Theoria et Historia Scientiarum*, 6(1):399–412.

M. Baroni and A. Lenci. 2008. Concepts and properties in word spaces. *Italian Journal of Linguistics*, 20(1):55–88.

J. Bos, S. Clark, M. Steedman, J. R. Curran, and J. Hockenmaier. 2004. Wide-coverage semantic representations from a ccg parser. In *Proc. of COLING'04*, pages 1240–1246.

P. Clark and P. Harrison. 2009. Large-scale extraction and use of knowledge from text. In *Proc. of the 5th international conference on Knowledge capture*, pages 153–160. ACM.

Catherine Havasi, Robert Speer, and Jason Alonso. 2007. Conceptnet 3: a flexible, multilingual semantic network for common sense knowledge. In *Recent Advances in Natural Language Processing*, Borovets, Bulgaria, September.

J. R. Hobbs, M. Stickel, P. Martin, and D. Edwards. 1993. Interpretation as abduction. *Artificial Intelligence*, 63:69–142.

J. R. Hobbs. 1985. Ontological promiscuity. In *Proc. of ACL*, pages 61–69, Chicago, Illinois.

J. R. Hobbs. 1992. Metaphor and abduction. In A. Ortony, J. Slack, and O. Stock, editors, *Communication from an Artificial Intelligence Perspective: Theoretical and Applied Issues*, pages 35–58. Springer, Berlin, Heidelberg.

Jerry R. Hobbs. 2005. Toward a useful concept of causality for lexical semantics. *Journal of Semantics*, 22(2):181–209.

N. Inoue and K. Inui. 2012. Large-scale cost-based abduction in full-fledged first-order predicate logic with cutting plane inference. In *Proc. of JELIA*, pages 281–293.

G. Lakoff and M. Johnson. 1980. *Metaphors we Live by*. University of Chicago Press.

G. Lakoff. 1987. *Women, fire, and dangerous things: what categories reveal about the mind*. University of Chicago Press.

S. Narayanan. 1999. Moving right along: A computational model of metaphoric reasoning about events. In *Proc. of AAAI/IAAI*, pages 121–127.

J. Nivre, J. Hall, and J. Nilsson. 2006. Maltparser: A data-driven parser-generator for dependency parsing. In *Proc. of LREC'06*, volume 6, pages 2216–2219.

R. Parker, D. Graff, J. Kong, K. Chen, and K. Maeda. 2011. English gigaword fifth edition. *LDC*.

A. Peñas and E. H. Hovy. 2010. Filling knowledge gaps in text for machine reading. In *Proc. of COLING'10*, pages 979–987.

S. Sharoff and J. Nivre. 2011. The proper place of men and machines in language technology: Processing Russian without any linguistic knowledge. In *Proc. Dialogue 2011, Russian Conference on Computational Linguistics*.

E. Shutova, T. Van de Cruys, and A. Korhonen. 2012. Unsupervised metaphor paraphrasing using a vector space model. In *COLING (Posters)*, pages 1121–1130.

E. Shutova. 2010. Automatic metaphor interpretation as a paraphrasing task. In *Proc. of NAACL'10*.

R. Speer and C. Havasi. 2013. Conceptnet 5: A large semantic network for relational knowledge. In *The People's Web Meets NLP*, pages 161–176. Springer.

N. Tsao and D. Wible. 2013. Word similarity using constructions as contextual features. In *Proc. JSSP'13*, pages 51–59.

T. Veale and Y. Hao. 2008. A fluid knowledge representation for understanding and generating creative metaphors. In *Proc. of COLING'08*, pages 945–952. ACL.