

Conceptual Metaphor and Scripts in Recognizing Textual Entailment

William R. Murray

Boeing Phantom Works, P.O.Box 3707, MC 7L-66, Seattle, Washington 98124-2207, U.S.A.

Abstract. The power and pervasiveness of conceptual metaphor can be harnessed to expand the class of textual entailments that can be performed in the Recognizing Textual Entailment (RTE) task and thus improve our ability to understand human language and make the kind of textual inferences that people do. RTE is a key component for question understanding and discourse understanding. Although extensive lexicons, such as WordNet, can capture some word senses of conventionalized metaphors, a more general capability is needed to handle the considerable richness of lexical meaning based on metaphoric extensions that is found in common news articles, where writers routinely employ and extend conventional metaphors. We propose adding to RTE systems an ability to recognize a library of common conceptual metaphors, along with scripts. The role of the scripts is to allow entailments from the source to the target domain in the metaphor by describing activities in the source domain that map onto elements of the target domain. An example is the progress of an activity, such as a career or relationship, as measured by the successful or unsuccessful activities in a journey towards its destination. In particular we look at two conceptual metaphors: IDEAS AS PHYSICAL OBJECTS, which is part of the Conduit Metaphor of Communication, and ABSTRACT ACTIVITIES AS JOURNEYS. The first allows inferences that apply to physical objects to (partially) apply to ideas and communication acts (e.g., “he lobbed jibes to the comedian”). The second allows the progress of an abstract activity to be assessed by comparing it to a journey (e.g., “his career was derailed”). We provide a proof of concept where axioms for actions on physical objects, and axioms for how physical objects behave compared to communication objects, are combined to make correct RTE inferences in Prover9 for example text-hypothesis pairs. Similarly, axioms describing different states in a journey are used to infer the current progress of an activity, such as whether it is succeeding (e.g., “steaming ahead”), in trouble (e.g., “off course”), recovering (e.g., “back on track”), or irrevocably failed (e.g., “hijacked”).

1 Introduction

Recognizing textual entailment (RTE, hereafter) is the process of determining whether a text (called T) entails a hypothesis (called H). For example, given

T: Dick Cheney peppered Harry Whittington with buckshot.

H: Dick Cheney pelted Harry Whittington with bird shot.

We can infer, solely on the basis of WordNet [2] that T entails H as there is a synonym `pelt#v2` for `pepper#v2` and `bird_shot#n1` for `buckshot#n1`.

We can also use WordNet to infer that T entails H in this case:

T: Reporters tossed accusations at Cheney.

H: Reporters threw charges at Cheney.

However, “toss” is polysemous, we must select the context-appropriate sense, `toss#v1`, which applies to physical objects, and extend it to apply to communicative objects such as questions, insults, and requests.

Currently, there is no verb sense of toss in WordNet that applies to communication acts. The most appropriate senses of “toss” for the above context, `toss#v1` or `toss#v3`, both refer to tossing physical objects. Their hypernyms do not include any communicative senses of ‘throw’, such as `throw#v5` or `throw#v10`, instead they only include `throw#v1`, which clearly refers to physical objects, as made clear by the accompanying definition and in its immediate hypernym `propel#v1`.

Humans easily interpret these verbs metaphorically via a conceptual metaphor that communicative acts such as questions, requests, and statements can be viewed and talked about as if they were physical objects. Questions can be thrown and dodged. Accusations can be launched and torpedoed. Requests fired and deflected. A computational system that checks that verbs are applied according to semantic preferences will reject the second inference as a violation of semantic type constraints, expecting physical objects where communicative acts are found instead. One that enforces no constraints at all would not flag the following sentences as anomalous (we will use *T to indicate anomalous T sentences and *H to indicate anomalous entailments that humans do not make such as those made by a system that takes a literal interpretation of sentences when a non-literal interpretation is more appropriate):

*T: Reporters tossed air at Cheney.

*H: Reporters threw a gas at Cheney.

Ideally, we would like a system that would know when conceptual metaphors sanction relaxing semantic type constraints, and what parts of the source domain (in the example above, physical objects, viewed as projectiles) map to what parts of a target domain (in the example above communicative acts or ideas), and furthermore what entailments we can procure from the source domain that apply correctly to the target domain.

In this paper we look at a mechanism to extend RTE inferences to those in the T-H pairs above, and consider how world knowledge represented as scripts might also allow us to predict some metaphoric entailments, e.g., that questions may be dodged, ducked, or deflected just as physical objects that are thrown may be.

1.1 Conceptual versus Linguistic Metaphor

Both Kövecses [4] and Lakoff & Johnson [5] make a clear distinction between conceptual metaphors and their linguistic expression. The former are partial mappings, typically from more abstract concepts to simpler concepts, e.g., LIFE IS A GAME.

The latter are linguistic expressions of these underlying concepts, e.g., “The odds are against me.”

Conceptual metaphors are so ubiquitous that many of their linguistic expressions are conventionalized. Consider the ontological metaphor IDEAS ARE PHYSICAL OBJECTS that allows us to talk about ideas as if they were physical objects using verbs like ‘bat’ and ‘shoot’. So we can bat an idea around and someone else may shoot it down. Indeed, we can find a communicative sense of shoot down in WordNet (shoot down#v3) but not a communicative sense of bat. These lexical synset holes are inevitable as no one lexicon can account for the infinite generative capability of language and its ability to produce unpredictable new linguistic variants of conceptual metaphors.

IDEAS ARE PHYSICAL OBJECTS is part of what is called the Conduit Metaphor of Communication by Lakoff & Johnson [5]. In this way of understanding the abstract act of communication we can visualize the sending of objects over some physical conduit. They are sent from one mind to another mind. Each mind acts as a container holding ideas, and can also send forth or receive new ideas in the form of containers (speech acts or messages) that hold the meaning of the ideas.

We believe that it is important for a natural language understanding system to handle conceptual metaphors, and their unconventional linguistic expressions of them, as these routinely occur in everyday discourse and writing. Some concepts, such as arguments, are difficult to conceptualize without metaphors [4] such as ARGUMENT IS WAR (he shot my points down) or ARGUMENT AS BUILDING (he built his argument carefully from the ground up, but the foundation was weak and his thesis advisor tore it down).

Finally, conceptual metaphors are a key part of understanding noun-noun constructions. For example, a *decapitation strike* is one aimed at removing the head of an enemy’s armed forces, where the top leader(s) of the armed forces are conceptualized as the head of a body, and where the body is the armed forces. Other constructed terms can also be understood with conceptual metaphor. For example, *emotional baggage* can be left over from a previous relationship, using the RELATIONSHIP AS A JOURNEY conceptual metaphor, but this word sense of baggage does not appear in WordNet. We have not yet extended this work to handle these kind of constructs, but this work is a step in that direction.

2 Examples

We work through an illustrative scenario to show how our system handles conceptual metaphor in RTE in a proof of concept implementation. We want our RTE-reasoner to be able to infer, using metaphor, that entailment holds in the case of

T: Joe lobbed questions at Rob.

H: Joe asked Rob questions.

similar to the counterpart T-H pair with physical objects:

T: Joe lobbed tennis balls at Rob.

H: Joe threw tennis balls at Rob.

In one case we have a communicative act (asking) specific to the communicative objects (questions) being used whereas in the other we have physical acts (throwing) specific to the physical objects (tennis balls) being used.

In contrast,

T: Joe dropped hints to Rob.

*H: Joe no longer has hints.

is not entailed, even though its corresponding expression with physical objects would be, e.g.,

T: Joe dropped the food to Rob.

H: Joe no longer has the food.

What we see is that we need to preserve coherence in mapping from one domain to another and block inferences that would be incoherent.

A similar example of an entailment we do not want to sanction would be:

T: Joe gave the idea to IBM.

*H: Joe no longer has the idea.

which would be true for physical objects:

T: Joe gave the hard drive to IBM.

H: Joe no longer has the hard drive.

So we see that we cannot one-to-one replace physical objects with conceptual objects, but instead need to ensure that our claims in the mapped domain do not violate any common-sense rules. This principle in cognitive science is called the Principle of Invariance [4]. We will model this principle by having different axioms for different verbs (e.g., drop versus toss) and different axioms for different kinds of objects (e.g., physical objects versus conceptual objects).

The latter two kinds of axioms will also allow us to also take into account inferences that apply to conceptual objects but not physical objects:

T: Joe had an idea. He forgot it.

H: Joe no longer has the idea.

...compared to...

T: Joe had an apple. He forgot it.

*H: Joe no longer has the apple.

We also want to apply scripts [8] from source domains to target domains when a source-to-target metaphoric mapping occurs. These scripts are simplified temporal models of the source domain states. In our proof of concept implementation, these are akin to finite-state machines where states are possible states of an activity, such as embarking, being on board, moving forward, seeing the destination, and arrival (see Figure 1). Each mapping will have conventionalized mappings that restrict what will be transferred. So in comparing an argument to a building, we might say “The foundations of his theory were weak, and it eventually toppled under repeated attacks.” In this case what is mapped is the strength of a building and the building process itself in constructing an argument, and not details such as the cost of the building, its number of rooms, and highway access [4].

Mapping scripts allows us to make RTE entailments we would not be able to make otherwise. A simplified script for journeys is shown in Figure 1. Even with such a simplified representation, we can make inferences when the conceptual metaphor ABSTRACT ACTIVITIES ARE JOURNEYS has been applied. We can fill in missing steps if we know that a step has occurred and rule out subsequent steps if we know a step has not (or cannot) occur.

Our big picture goal is to reduce the impedance mismatch between what we have discovered in cognitive science about how people think and the formal representations that are used in most computational linguistics programs today. In the former case, cognitive science has empirically shown that people think using prototypes, recognize concepts via family resemblances, and use metaphors in day-to-day thinking, not just as a linguistic construction. The field of artificial intelligence and computational linguistics lags behind these findings, and continues to model sentence meanings as First-Order Logic constructions where concepts are defined with hard-and-fast necessary-and-sufficient conditions and where metaphor is viewed largely as linguistic phenomena. Basically, humans do not think logically, but largely metaphorically, so we need to better model metaphoric reasoning to better model human reasoning and improve RTE. A better modeling does not require us to abandon logic and deductive processes, but to instead better account for the looser matching that humans are capable in recognizing category instances and also provide a capability to handle conceptual metaphors. The cognitive linguistics view also recognizes that the most basic and universal conceptual metaphors are grounded in our bodily experience and experience moving through the world and achieving our goals.

In our implementation we axiomatize the preconditions and postconditions of the verbs and we similarly characterize the physical or communicative objects. Thus we can distinguish between properties of physical and communicative objects, for example, that communicative objects can be duplicated and present at more than one place, whereas this is not true of a physical object. That is why a person can drop a hint in a meeting and still retain (the idea behind) the hint. When we are discussing communication acts (e.g., hints, questions, accusations) in this way we are conceptualizing them as physical containers for meanings or ideas. These ideas can result in any number of these meaning containers being sent without the idea being lost, unlike the sending of money or objects like tennis balls where there is a reduction in amount or quantity.

Our axioms for physical objects include:

1. A physical object is at one place at a time. Thus transferring it to a new location removes it from the old location, which will not be the case for communication objects.
2. Physical objects as projectiles can cause physical injury.

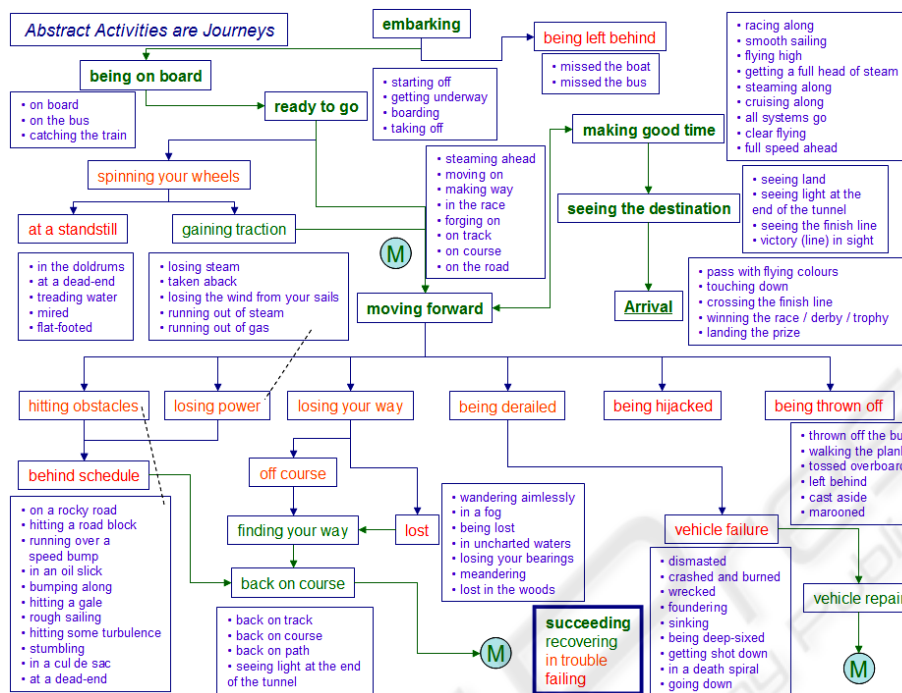


Fig. 1. Script for Journeys for use in making inferences when the ABSTRACT ACTIVITIES ARE JOURNEYS Conceptual Metaphor is applied.

Our axioms for communication objects:

1. A communication object represents some meaning.
2. The recipient must receive the communication object to grasp its meaning.
3. Communication objects as projectiles can cause emotional injury, but not physical injury (thus they cannot kill in the physical sense)

We would also like to have scripts for actions that affect physical objects. These scripts will be similar to macro (composite) actions, in that they will also have preconditions and postconditions, just as primitive actions do. They will differ in that they will predict different possible results of actions, according to our experience with these actions. Thus when we toss an object to someone else possible results are that they will catch it, dodge it, fumble it, deflect it, or ignore it, that is “let it pass”, but some objects are harder to ignore than others of course! Different people may have different scripts for actions, and experts in fields would be expected to have more refined and detailed scripts than lay people. So a baseball player will know more about catching a ball and a sailor more about sailing a boat. We do not address the acquisition of scripts in this paper as that is a separate problem in machine-learning and knowledge acquisition. Automated machine knowledge acquisition also raises questions as to whether an agent must be embodied to understand scripts that depend on human motor actions, such as running, jumping, doing a handstand, etc. Instead, we will assume that our scripts only model (and do not completely capture) this hu-

man knowledge, that they are incomplete and that human scripts may differ from person to person in their particulars.

Since most of these communicative scripts are very verb-specific and have only a small number of likely states we have instead elected to use the Journey Script in Figure 1 to better illustrate the utility of scripts when the conceptual metaphor ABSTRACT ACTIVITIES ARE JOURNEYS is applied.

3 Implementation

We tested our implementation using the Prover9 [6] first-order logic theorem prover on a small set of RTE examples. The examples are encoded in logic as would be provided ideally from a Natural Language system. We currently have an NL system that can provide some of these encodings but further work is needed in the logic generation portion of our NL system to provide these kind of outputs. Thus the inputs can be viewed as simulations of the kind of output we eventually hope to get and the reasoning we hope to do in our own NL system. They are illustrations of the kinds of inferences that would then be possible using conceptual metaphors with scripts applied from source domains onto target domains to make metaphor entailments.

3.1 Axioms

We give a shortened listing of the axioms we use in Prover9 for the RTE examples for the ABSTRACT ACTIVITIES ARE JOURNEYS metaphor. We group together states that lead to a similar outlook for the activity, e.g., a trip is currently succeeding if we are (in a state of) embarking on it or moving forward...

```
member(State, [embarking, being_on_board, ready_to_go,
  moving_forward, making_good_time, seeing_the_destination])
-> outlook(State, Time, Trip, succeeding) .
```

...and it is currently in trouble if we are hitting obstacles or have gone off course...

```
member(State, [spinning_your_wheels, hitting_obstacles,
  losing_power, losing_your_way, being_derailed, off_course])
-> outlook(State, Time, Trip, in_trouble) .
```

We can reason backwards for a current state to a previous state, if the current state only has one predecessor and the current state is either definitely true (definite) or definitely false (impossible):

```
certainty(State, Trip, Time, definite)
  & (prior(State) = Previous)
-> certainty(Previous, Trip, pred(Time), definite) .

certainty(State, Trip, Time, impossible)
  & (prior(State) = Previous)
```

```
-> certainty(Previous, Trip, pred(Time), impossible) .
```

Evidence is what we know absolutely about any given state of a trip:

```
evidence(State, Time, Trip)
-> certainty(State, Trip, Time, definite) .
```

```
-evidence(State, Time, Trip)
-> certainty(State, Trip, Time, impossible) .
```

The states of Figure 1 are enumerated in the axioms, only a few shown here...

```
state(embarking) .
state(being_on_board) .
```

...along with preceding and subsequent states...

```
subsequent_states(ready_to_go) =
  [spinning_your_wheels, moving_forward] .
preceding_states(ready_to_go) = [being_on_board] .
```

...most again omitted here for brevity. Here is the definition for member:

```
member(Item, [Other:Rest])
  <- (Other = Item) | member(Item, Rest) .
```

Test cases are given as evidence for states on specific trips or activities viewed metaphorically as journeys:

```
evidence(being_derailed, time1, career) .
evidence(moving_forward, t1, relationship) .
evidence(behind_schedule, succ(succ(t1)), relationship) .
evidence(vehicle_repair, time_x, dissertation) .
```

```
% predecessor (pred) and successor (succ)
pred(succ(X))=X.
% We are using an overly simple theory of time right now.
```

RTE Encodings and Results. Below we show some of the RTE examples we have tried, giving both their English and formal logic equivalents. The latter are hand-coded at present. In each case H is entailed and the corresponding Prover9 goal is shown.

T: John's career was derailed on Monday.

H: John's career is in trouble.

```
outlook(being_derailed, time1, career, in_trouble) .
```

Here the goal (formula above) is to show that John's career is in trouble if the activity of his career is currently in the state of being derailed. Prover9 shows this is true in our formulation by proving the goal formula correct.

T: Jane's relationship with Jim was behind schedule.
 H: Jane's relationship with Jim was hitting obstacles.

```
certainty(hitting_obstacles,
  relationship, succ(t1), definite) .
```

If we know Jane's relationship is behind schedule now at time $t+2$ but was moving forward at time t and we assume it was not losing steam at $t+1$ then it must have started hitting obstacles at $t+1$.

T: Joe's dissertation, previously a leaky ship, was in good repair after plugging the holes his dissertation advisor warned against.
 H: Joe had been moving forward on his dissertation.

```
cetainty(moving_forward, dissertation,
  pred(pred(pred(time_x))), definite) .
```

In this last case, Prover9 can prove that if Joe's activity (dissertation) is in a state of vehicle repair at a time t then at time $t - 3$ it was moving forward (see Figure 1).

4 Related Work

Dan Fass meta5 [1] implements a metaphor and metonymy finding system that relies on first checking semantic type preferences, and then finding violations, seeking for alternate explanations through part-whole metonymies, metaphors where these constraints are relaxed, or in the last resort declaring a sentence an anomaly. The key difference between that work and ours is that the constraints we seek to impose are primarily motivated by cognitive science findings, whereas in Fass' work structural matching is the key driver and cognitive science plays a secondary role, more as a catalog of some potential metaphors to consider. For example, in metonymy, Fass' system could equally well prefer 'rudder', 'hull', or 'sheet', as a reference to a sailboat, as well as 'sail', ignoring that a sail is more characteristic of a sailboat, and thus more central to a metonymy [4].

Deidre Gentner's work [3], too, emphasizes structural matching over a cognitive basis for conceptual metaphors. The SME (Structure Mapping Engine) that provides a simulation of Genter's theories relies heavily on structural graph matching. Again, there is no notion that metaphors are ultimately motivated by human experiences (e.g., correlation of anger with feelings of pressure leading to the ANGER IS A PRESSURIZED FLUID IN A CONTAINER conceptual metaphor shared across many cultures [4]).

Srini Narayanan's work [7] is most similar to ours, as it provides a means of making inferences via metaphors, but it does not emphasize the use of scripts or applying metaphoric entailment to RTE. Instead, it more directly models embodied actions

such as movement using Petri Nets and uses Belief Networks for its inference mechanism.

5 Summary

We have shown that conceptual metaphor can play a role in extending the capabilities of RTE systems, so that they may make entailments that rely on conceptual metaphors. Furthermore, we have posited that world knowledge as expressed as scripts can be usefully mapped along the lines of these metaphors, except where contradictions are encountered, to provide metaphoric entailments. These entailments include predictions of future events, denials of past events that are not consistent with what is known, and inferred expectations for events that were elided in the discourse. Thus, the ability to apply conceptual metaphor extends the ability of a natural language processing system to more closely model the human ability to reason about language and make inferences from language, by providing a better model of human reasoning than systems lacking conceptual metaphor capabilities and locked into only handling the literal representation of sentences.

Acknowledgements

I would like to thank Christiane Fellbaum and Peter Clark for very helpful suggestions on this work.

References

1. Fass, D.: Processing Metonymy and Metaphor. Ablex Publishing, Connecticut (1997).
2. Fellbaum, C.: WordNet: An Electronic Lexical Database. The MIT Press (1998).
3. Gentner, D.: Why're We're So Smart. In D. Gentner, S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195-235). MIT Press, Cambridge, MA (2003).
4. Kovecses, Z.: *Metaphor: A Practical Introduction*. Oxford University Press, USA (2002)
5. Lakoff, G., Johnson, M.: *Metaphors We Live By*. University Of Chicago Press (1980)
6. McCune, W. Prover9, <http://www.cs.unm.edu/~mccune/prover9/>
7. Narayanan, S.: Moving Right Along: A Computational Model of Metaphoric Reasoning about Events. Proc. of the Nat. Conf. on Art. Intelligence, pp. 121-128. AAAI Press. (1999).
8. Schank.: *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*, Cambridge University Press; 1st edition (1983).