Structural Alignment as an Abductive Integer Linear Programming Problem

Clifton McFate CJM@EC.AI
Elemental Cognition, 230 Park Ave, 3rd Floor West. NY, NY. 10169

Abstract

We present a novel implementation of analogical structural alignment as constraints in an abductive inference problem where one structured representation is justified by its alignment to another. Following Inoue & Inui (2011), we implement this solve as a 0-1 integer linear programming problem. We call our approach the Constrained Abductive Mapping Model of Analogy (CAMMA), and demonstrate how our formulation facilitates experimentation with both hard structural constraints and possible non-structural constraints. As a proof of concept, we use our CAMMA implementation to model results from Gentner & Ratterman (1991) who found that young children exhibited a bias towards matching objects based on shared attributes rather than shared relations, and further found that additional relational structure facilitated relational mapping.

1. Introduction

There is ample psychological evidence that analogy, and structural similarity more broadly, play an important role in core cognitive processes ranging from visual perception (Sagi et al., 2012) to higher-level reasoning and decision making (Markman & Medin, 2002). As such, computational models of structural alignment are valuable as research tools and as a component in applied computational agents. Indeed, existing models have demonstrated utility in and provided insight on many cognitively challenging tasks (Forbus et al., 2017).

Two prominent models of analogy are Falkenhainer et al's (1989) Structure Mapping Engine based on Gentner's (1983) Structure Mapping Theory (SMT) and Holyoak and Thagard's Analogical Constraint Matching Engine (ACME) (1989).

Both claim analogical mappings are subject to structural constraints. However, SME uses these constraints to limit initial mapping hypotheses and guide a greedy search for a global solution, while ACME encodes both structural and non-structural (e.g. pragmatic) constraints as excitatory and inhibitory connections among a larger hypothesis network.

One benefit of SME is that it allows missing structure to be projected across a mapping as a novel *candidate inference* (CI). ACME can fill in structure from a suggestion but did not produce spontaneous inferences. Furthermore, SME can generate multiple alternative analogies, which ACME's "winner-take-all" algorithm did not allow. There is also evidence that SME's stricter structural constraints better align with human judgements. (Forbus et al., 2017; Markman, 1997).

At the same time, Holyoak and Thagard's constraint based approach is alluring because it facilitates modeling interactions between structural similarity and other softer influences such as preexisting biases. While extensions to SME do allow a modeler to apply declarative match filters on viable matches, only a few kinds of filters are allowed and they do not include soft biases (Forbus et al., 2017).

One candidate for such a bias in cognition comes from Gentner's (1988) "relational shift hypothesis" which proposes that a shift occurs from alignment based on on shared attributes (e.g. color, shape) to a preference for relational matches (e.g. a shared spatial layout) throughout maturation (Gentner, 1988; Gentner & Ratterman, 1991; Hespos et al., 2020). This could be the result of an evolving bias, though, Gentner and colleagues also argue that this "relational shift" is facilitated by the acquisition of relational schemas which facilitate richer encoding.

These accounts need not be mutually exclusive, and it seems likely that as children gain relational schema, they also learn to attend to relations. Exploration of these and other hypotheses would be facilitated by a common framework that easily allows experimentation with both hard structural constraints (like SME) as well as soft biases (like ACME), and that generates testable SME-like inferences.

To that end we introduce a novel implementation of structural alignment called the Constrained Abductive Mapping Model of Analogy (CAMMA) which formulates structural alignment as a weighted abductive inference problem, the goal of which is to justify an observed propositional structure through its alignment to another. Following Inoue & Inui (2011), we ground the abductive problem as an integer linear programming problem which makes it easy to add declarative hard structural constraints as well as soft scoring biases. A CAMMA model can be solved by any existing ILP solver (we use the commercial solver Gurobi).

While we do not claim that our axiomatic implementation of structural alignment or ILP itself is a process level cognitive model, we do argue that implementing structural alignment in this way creates a convenient single framework for quickly testing predictions of competing theories of analogy (e.g. SME's hard vs ACME's soft 1-1 mapping constraint) and exploring how the principles of structural alignment may interact with constraints from other cognitive processes. We also believe that grounding structural alignment in ILP could have practical benefits because it allows the application of mature optimization technologies to the alignment problem while also bringing psychologically motivated constraints to optimization.

As a proof of concept, we use the CAMMA implementation to investigate the relational shift hypothesis by simulating Gentner and Ratterman's (1991) cross-mapping experiment which found an inhibitory effect of object attribute similarity on structural alignment. We model this experiment with opposing relational and attribute soft biases, compare predictions generated by a hard vs soft implementation of SMT's 1-1 mapping principle, and examine how the addition of relational schema facilitates relational mapping even in the absence of an explicit relational bias.

2. Background

2.1 Structure Mapping Theory and SME

Gentner's (1983) Structure Mapping Theory (SMT) proposes that analogical comparison, and similarity more broadly, result from the alignment of a base and target set of structured representations. Table 1 provides an overview of structure mapping theory's structural alignment constraints which

Table 1. SMT Constraints

Constraint	Description
1-1 Mapping	Each base item may match to at most one target item.
Parallel Connectivity	Expressions only align if their children align.
Identicality	Non-identical predicates align only if they are arguments of aligned higher-order expressions.
Systematicity	Alignments that contain higher-order (nested) structure are preferred.

Table 2. ACME Constraints

Constraint	Description
Isomorphism	Prefer isomorphic matches (i.e. ones that obey 1-1 and parallel connectivity).
Similarity	Prefer predicates and constants to map to similar predicates and constants. Not necessarily inheritance.
Pragmatic Centrality	Pragmatically important mappings (e.g. goals) or salient mappings are preferred.

define a valid mapping. For a comprehensive discussion of these constraints and their psychological validity, see the analysis in Forbus et al. (2017).

SMT has been implemented computationally in the Structure Mapping Engine (SME) (Falkenhainer et al., 1989; Forbus et al., 2017) which has been used to model a wide array of cognitive phenomena including geometric analogies, grammatical coercion, and intent recognition. (Lovett & Forbus, 2012; McFate & Forbus, 2016; Rabkina & Forbus, 2019)

In SME, identicality is used to constrain the initial space of match hypotheses which are then coalesced into kernels that abide by 1-1 and parallel connectivity. Systematicity acts as a scoring function that guides a greedy process of merging structurally consistent kernels. Matches that violate the principles are not allowed. SME is able to produce alternative consistent analogies, and can generate inferences based on aligned structure.

CAMMA will also use predicate identity to constrain an initial set of match hypotheses. Parallel connectivity will be encoded in our axiom generation procedure, and 1-1 mapping can be encoded as either a hard constraint that makes a mapping infeasible or as a soft constraint that is allowed at a cost. Because CAMMA is implemented as an ILP solve, it is also able to consider alternative consistent mappings, and we will generate cadidate inferences as a part of mapping.

2.2 Constraint Satisfaction and ACME

Holyoak and Thagard's (1989) ACME both expands and weakens SMT's principles, modeling them in competition (or cooperation) with other constraints (see Table 2).

ACME constructs an initial network of match hypotheses between same arity concepts. Each hypothesis is linked to others by excitatory and inhibitory edges. For example, isomorphism would be enforced with inhibitory edges between alternative matches for the same entity. The matches are also connected to a semantic and a pragmatic unit which excites or inhibits the match based on a predefined function(s). Given an initial activation, the score of each node (MH) is iteratively updated given the links to its neighbors. The final alignment can be read from the state of the network.

Like ACME, CAMMA allows additional constraints to impact mapping, and it is capable of loosening the 1-1 mapping constraint. However, our approach treats identity and parallel connectivity like SME, using them to constrain the initial match hypothesis space. ACME treats inferences like goals (external) whereas SME generates novel inferences by alignment. Like SME, we generate inferences without external goals, but like ACME, we are capable of using desired inferences to guide our solution.

2.3 Weighted Abduction and ILP

Abductive inference seeks the best explanation of an observation given background knowledge. Formally, given a set of background rules B, and a set of observed literals, O, the goal of abduction is to find a hypothesis H where:

- $B \not\models O$: The rules do not already entail the observations
- $H \cup B \models O$: The hypothesis and rules entail all observations
- $H \cup B \not\models \perp$: They do not entail something false

Integer Linear Programming is an optimization paradigm which finds an assignment of integers to a set of variables given constraints on variable values and variable coefficients (Papadimitriou & Steiglitz, 1998). As a simple example, you must invest 10 dollars. Investment x_1 returns twice the input, but is limited to a 5 dollar investment. x_2 returns half and has no limit. This could be formulated as:

Maximize: $2x_1 + 0.5x_2$

Subject to: $x_1 <= 5$ and $2x_1 + 0.5x_2 <= 10$

The answer is $x_1 = 5$ and $x_2 = 5$. You should put the maximum, 5 dollars, into investment 1 and the rest into investment 2. There exist a wide array of optimization techniques and commercial solvers for integer linear problems. Our approach further limits variables to boolean values (0 or 1) making it a 0-1 ILP problem, a well studied subclass of ILP.

Inoue & Inui (2011) represent an abductive solve as a 0-1 ILP problem where B is a set of first-order Horn clauses and O and H are conjunctions of ground literals. They refer to each hypothesis

H that may justify the observations O as a candidate hypothesis and refer to each literal $h \in H$ as an elemental hypothesis. We use this terminology here. In weighted abduction each literal $h \in H$ has a real valued cost. The goal is to find the least cost candidate hypothesis H.

To build the ILP problem, chains of elemental hypotheses are enumerated by back-chaining over B given O. The least cost hypothesis in the chain is paid (assumed). Hypotheses can be unified to pay only the cost of one. As an example, consider the following simple set of rules B that define unsafe driving conditions. A road is unsafe if it is wet and obscured, both of which are caused by snow. Assume each rule has a cost of 10.

`unsafe(?road)` $\Leftarrow$ `wet(?road)` $\wedge$ `obscured(?road)`

`wet(?road)` $\Leftarrow$ `snow(?road)`

`obscured(?road)` $\Leftarrow$ `snow(?road)`

Given an observation O that the road, I95, is unsafe `unsafe(I95)`, we back-chain to generate the hypothesis space. The ground axioms are shown below.

`unsafe(I95)` $\Leftarrow$ `wet(I95)` $\wedge$ `obscured(I95)`

`wet(I95)` $\Leftarrow$ `snow(I95)`

`obscured(I95)` $\Leftarrow$ `snow(I95)`

Any literal can be assumed at cost. Here, the observation can be assumed directly at the high cost of 40. Figure 1 shows the hypothesis space for this problem. Using the terminology from Inoue and Inui, the rows are candidate hypotheses and the columns are the elemental hypotheses. For each candidate hypothesis, an elemental hypothesis can be derived true, assumed (assm), or not included. Note that `snow(I95)` appears twice. In their approach, the elemental hypotheses resulting from different rule applications remain separate but can be unified to pay their cost only once.

	unsafe(I95)	wet(I95)	obscured(I95)	snow(I95)	snow(I95)	Cost
H1	assm					40
H2	true	assm	assm			20
H3	true	true	assm	assm		20
H4	true	true	assm	assm		20
H5	true	assm	true		assm	20
H6	true	true	true	assm	assm	20
H7	true	true	true	UNIFIED (assm)		10

Figure 1. Propositional Hypothesis Space

We can leave I95's unsafe conditions unexplained by just assuming `unsafe(I95)` at a cost of 40. Alternatively we could assume its causal antecedents, that I95 is unsafe because it is presumably wet and obscured. Either of these conditions can be caused by snow, and so in the lowest cost solution, we unify the shared cause into a single expression `snow(I95)`, thus paying its cost only once. Given that I95 is unsafe, in this model it is most reasonable to assume it has snowed.

Inoue & Inui (2011) transform this hypothesis space in to a 0-1 ILP problem by assigning the following variables for each elemental hypothesis $h \in H$:

- $h\{0,1\}$ 1 if elemental hypothesis is a part of a candidate hypothesis
- $r\{0,1\}$: 1 if cost of elemental hypothesis is NOT paid
- $u_{p,q}\{0,1\}$: 1 if elemental hypotheses p and q unify

In ILP, the parameters of a valid solution are defined by constraints. Inoue & Inui (2011) define several such constraints for abduction. We summarize the relevant constraints for our problem below and illustrate the ILP problem space for the unsafe roads example in Figure 2

- $h_p\{1\}$ `for` $p \in O$: Observations must be in a candidate hypothesis.
- $r_p\{1\}$ `iff` $h_{p1}\{1\}$ `and` $p \Leftarrow p1$ `or` $u_{p1,p}\{1\}$: cost of elemental hypothesis is not paid only if it is unified with or justified by another elemental hypothesis.
- $u_{p1,p}\{1\}$ `iff` $h_p\{1\}$ `and` $h_{p1}\{1\}$: literals can unify only if they are both in the candidate hypothesis.

	h_u	r_u	h_w	r_w	h_o	r_0	h_{s1}	r_{s1}	h_{s2}	r_{s2}	$U_{s1,s2}$	Cost
H1	1	0	0	0	0	0	0	0	0	0	0	40
H2	1	1	1	0	1	0	0	0	0	0	0	20
H3												
H4												
H5							...					
H6												
H7	1	1	1	1	1	1	1	1	1	0	1	10

Figure 2. ILP Hypothesis Space

In Figure 2 each variable is sub-scripted by first letter of the literal it applies to(e.g. h_u corresponds to including the elemental hypothesis `unsafe(I95)`). In the final solution, we pay only the cost of one of the `snow(I95)` assertions because they unify. Inoue & Inui (2011) extend their approach to allow axioms with existentially quantified variables by introducing a substitution ILP variable and corresponding constraints. In our model, we will treat unbound variables resulting from back-chaining as inferred entities that can be "bound" through proposition unification as above. See Figure 3, our analogical mapping formulation, for an example.

This approach also makes it easy to declaratively define hard and soft constraints in background knowledge. A hard constraint, if violated, results in an infeasible solution. A soft constraint, if violated, instead adds a penalty to the overall cost. We can implement both through axiom generation. For example, to implement 1-1 mapping as a hard constraint we can add the following to our background knowledge:

$$\emptyset \Leftarrow align(?x, ?y) \wedge align(?x, ?z)$$

Whereas a soft constraint could conclude a cost modification.

$$cost(10) \Leftarrow align(?x, ?y) \wedge align(?x, ?z)$$

3. Analogy as Weighted Abduction

Abstractly, the goal is to treat a mapping as a solution that justifies the base and target observations. Therefore, each expression in the base and target can be justified by an alignment that holds between them, and each expression alignment is justified by the alignment of its arguments. Figure 3 illustrates an example analogy between a base and target case.

3.1 Phase 1: Calculating Match Hypotheses and Potential Candidate Inferences

CAMMA begins like SME, creating match hypotheses (MH) between pairs of expressions that meet the identicality function. It then recursively generates match hypotheses for each argument, discarding those that violate parallel connectivity.

Unlike SME, CAMMA also pre-generates candidate inferences (CI) that justify an unmapped expression through the alignment of its arguments (see phase 2). Inferences will be generated through unification as a part of the abductive solve.

Figure 3 illustrates a simple alignment where the base consists of a conjunction of propositions, `P(a)` and `Q(a)`, and the target consists of the proposition `P(c)`. `P(a)` in the base will align with `P(c)` in the target based on identity, and so they form a match hypothesis.

3.2 Phase 2: Axiom Generation

The initial set of match hypotheses is used to construct the set of ground axioms for the solve. For each match hypothesis, the base and target observations are justified by the assumption of an `ExpAlign` proposition representing their alignment. In Figure 3, the observations `P(a)` and `P(c)` can be derived true (justified) if their `ExpAlign` proposition is assumed true (at a cost of 10). CAMMA justifies each such proposition by the alignment of it arguments ultimately grounding out in a set of entity alignments (`EntAlign`). We can conclude that `P(a)` and `P(c)` are aligned if a and c align (at a lower cost of 1).

An expression can also be justified by a candidate inference proposition (`CI`). A candidate inference is justified by the alignment of its child entities to unbound variables. These variable

	Observations
40	Base: `P(a)` ∧ `Q(a)`
40	Target: `P(c)`
Cost	Axioms
10	`P(a) ∧ P(c) <= ExpAlign(P(a),P(c))`
5	`Q(a) <= CI(Q(?a))`
1	`CI(Q(?a))) <= EntAlign(?a, a)`
1	`ExpAlign(P(a),P(c))<= EntAlign(a, c)`

Hypothesis Space Propositional

	P(a)	Q(a)	Exp(P_a,P_c)	CI($Q_{?a}$)	Ent(a,c)	Ent(?a,a)	Cost
H_1	assm	assm					80
H_2	true	assm	assm				50
H_3	assm	true		assm			46
H_4	true	assm	true		assm		41
H_5	true	true	assm	assm			15
...							
H_{n-1}	true	true	true	true	assm	assm	2
H_n	true	true	true	true	UNIFIED ?a = a		1

Hypothesis Space ILP

	h_p	r_p	h_q	r_q	h_{exp}	r_{exp}	h_{ci}	r_{ci}	h_{ent1}	r_{ent1}	h_{ent2}	r_{ent2}	$U_{ent1,ent2}$	Cost
H1	1	0	1	0	0	0	0	0	0	0	0	0	0	80
H2	1	1	1	0	1	0	0	0	0	0	0	0	0	50
...														
H7	1	1	1	1	1	1	1	1	1	1	1	0	1	1

Figure 3. Axioms are generated from initial match hypotheses. Note that $Q(a)$ can be justified by the candidate inference (CI) that $Q(?a)$ holds in the target with an inferred or unified filler. The hypothesis space shows the abductive cost for each solution, where an assumption (assm) means the cost of the proposition is paid. In the lowest cost solution, CAMMA unifies the entity alignment (EntAlign) expression containing the inferred (variable) entity and the one that aligns a to c, thus paying only one cost. This produces a bound candidate inference: $Q(c)$. In the ILP formulation, $h\{1\}$ means the elemental hypothesis is included in the candidate hypothesis, $r\{1\}$ means its cost is paid, and $u\{1\}$ represents a unification.

entities can be bound to existing entities when the expressions they participate in are unified with fully bound expressions. If no unification is possible, they become inferred entities. In Figure 3, we can justify `Q(a)` with a candidate inference (at a cost of 5). To summarize:

For each $\mathtt{MH}(b_n, t_n)$

justify literals with expression alignment

$b_n \Leftarrow \mathtt{exprAlign}(b_n, t_n)$

$t_n \Leftarrow \mathtt{exprAlign}(b_n, t_n)$

For each $\mathtt{exprAlign}(b_n, t_n)$

justify with entity alignment of arguments

$$\mathtt{exprAlign}(b(x_1...x_n), t(y_1...y_n)) \Leftarrow \mathtt{entAlign}(x_1, y_1...x_n, y_n)$$

For each literal p in `CIs`

justify by an entity alignment to variable arguments

$p(x_1...x_n) \Leftarrow \texttt{CI}(?x_1...?x_n)$

$\texttt{CI}(?x_1...?x_n) \Leftarrow$

$\quad \texttt{entAlign}(?x_1, x_1...?x_n, x_n))$

3.3 Phase 3: Model Generation

The first step of model generation is to back-chain from observations using the ground abductive axioms generated in phase 2 in order to generate the hypothesis space. For analogy, the observations are the base and target literals. The resulting model consists of a set of elemental hypotheses (propositions) that can justify the observations connected by the set of axioms invoked when back-chaining.

Recall that unbound variables will be ground as inferred entities. Inferred entities can become resolved to existing entities as a result of unification during the solve. Thus candidate inferences opportunistically reduce cost as a part of a solve.

As described above, a model can also include hard and soft constraints, both represented in background knowledge and enumerated as axioms. Hard constraints are implemented as Boolean constraints that are generated given a logical implication. Soft constraints are also implications, but they result in a modification of the cost function for a hypothesis rather than a Boolean constraint.

3.4 Phase 4: ILP Solve

We use the commercial ILP solver (Gurobi) to produce solutions of our mapping model, though there are a wide array of commercial and non-commercial solvers available. Following the solve, the alignment can be read off as the set of assumed entity alignments and candidate inferences. In Figure 3, the least cost solution assumed an alignment between the entities a and c, thus justifying the expression alignments and candidate inference, and therefore the observations.

4. Proof of Concept: Modeling Attribute vs Relational Biases

Gentner's (1988) relational shift hypothesis proposes that object/attribute similarity precedes relational similarity. Formally, object/attribute similarity is similarity on the basis of unary predicates (e.g. `red(ball)` whereas relational similarity holds on the basis of shared nary or higher order predicates (e.g. causal relations). Gentner and colleagues further argue that this "relational shift" is facilitated by a focus on relations and the acquisition of relational schema which facilitate a richer relational encoding (e.g. Christie et al., 2007; Loewenstein & Gentner, 2005).

Gentner & Ratterman (1991) provide evidence for this hypothesis in a series of experiments. In each, a child and experimenter both had a set of three objects that monotonically changed in size (e.g. Large-Medium-Small). The experimenter placed a sticker under one of their objects, and asked the child to find a sticker under their corresponding object.

In what we call the "consistent" condition, the child's occluding objects were identical to the experimenters. They then introduced a "cross-map" condition where the child's occluding objects retained the montonic size change, but started with a larger object (e.g. Large-Medium-Small vs

Extra Large-Large-Medium). This created a cross-mapping between shared attributes (size) and shared relations (relative size). The sticker was always under the object that shared the same relation as the experimenter's object.

As expected, young children were unable to consistently produce the relational match in the cross-mapping condition, and further experiments revealed that performance degraded even more when the simple objects from the first experiment were substituted with feature-rich (e.g. visually complex) objects. We call this manipulation the "feature-rich" condition (as opposed to the feature-sparse original objects). The first two columns of Figure 4 graphically depict the feature-sparse and feature-rich consistent and cross-map conditions. In each condition, the location of the sticker (object 2) is highlighted with a box.

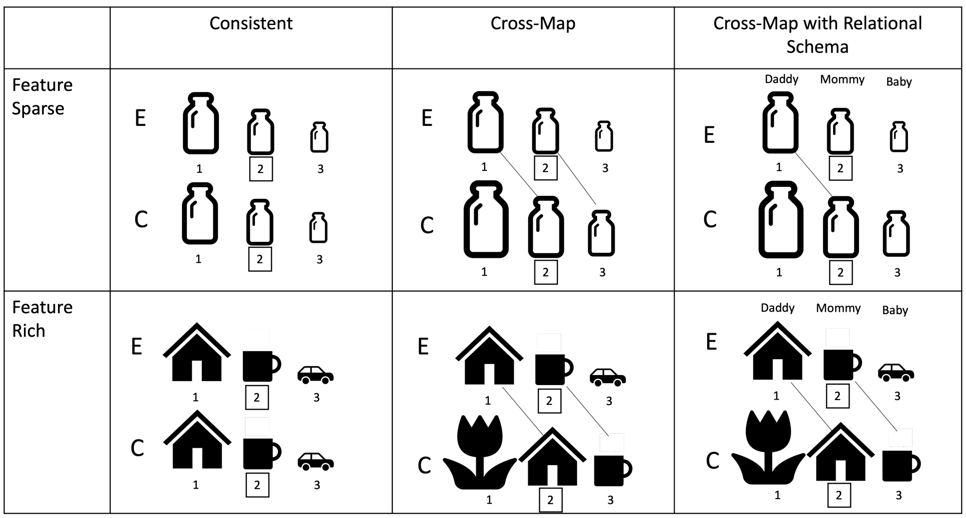

Figure 4. Each cell visually depicts an experimental condition. The top row of each condition depicts the experimenter's objects. The bottom depicts the child's objects. The location of the hidden sticker (object 2) is highlighted in each condition.

To test the hypothesis that relational schema facilitate relational alignment, they provided the children with familiar relational labels that mirrored the monotonic change relationship (Daddy-Mommy-Baby). With the relational labels, the children were able to match on the basis of shared relations rather than attributes. We call this the "cross-map with relational schema" condition as shown in the last column of Figure 4.

In the following sections, we use the CAMMA model to simulate the judgements in the "consistent", "cross-map", and "feature-rich cross map" conditions from Gentner & Ratterman (1991), with and without a relational schema. We further show how a cognitive bias (an ILP soft constraint) for mapping attributes or relations can result in the maturing judgements predicted by the relational shift hypothesis, and we model the interaction between the strength of that bias and the added structural support provided by the schema.

4.1 Experiment Setup

We represented the Gentner & Ratterman (1991) experiments as an inference task where the goal was to infer the location of the target sticker given the base sticker's location. In our experiments, a relational response corresponds to a candidate inference that the sticker is under the the median object (e.g. object 2) on the basis of aligning the shared relations between the base and target.

Each task was represented in predicate calculus; attributes were represented as unary predicates, and the monotonic increase in size was represented with a binary relation between each container. To model stimuli in the feature rich condition (e.g. a red flower) we included additional attributes. For example, the experimenter "feature-sparse consistent" condition base case was represented as:

```
Jar(object1), Jar(object2), Jar(object3) Sticker(sticker)

small(object1), medium(object2), large(object3)

largerThan(object2,object1), largerThan(object3,object2)

under(object2,sticker)
```

In the schema condition, we modeled the Daddy-Mommy-Baby relational schema as a ternary relation that was the consequent of a causal relation from the `largerThan` relations:

```
causes(

    &(largerThan(object2,object1), largerThan(object3,object2))
    familySchema(object1,object2,object3))
```

4.2 Modeling

CAMMA encodes the identicality principle from structure mapping theory (SMT) into its initial match hypothesis construction. As discussed above, this differs from ACME which instead uses excitatory and inhibitory edges to prefer mappings between similar predicates. 1-1 mapping can be either a hard constraint (makes the model infeasible) or a soft constraint (increases the cost of the solution). We model both in our experiments. The SMT parallel connectivity constraint falls out of the justification of expression alignments (they are justified by the alignment of their arguments). We give expression alignments an overwhelmingly high cost to assume so that they cannot be assumed in the absence of parallel connectivity.

Systematicity is somewhat enforced as a part of the problem formulation, since expressions that justify a larger structure will nullify the cost of assuming that structure. However, one could also add a negative cost soft constraint.

Across experiments we assume a cost of

- 1000 for `expressionAlignment`
- 1 for `entityAlignment`
- 2 for `candidateInference`.

Like ACME, we can introduce cognitive (pragmatic) biases. In these experiments we do so through soft constraints which modify the cost of a solution. To model no bias, a fixed score of 10 is used to assume expressions without alignment. To simulate a relational bias, we assign an additional penalty of 20 for each argument, thus making relations more costly to leave unaligned. For an attribute bias, we apply a flat cost of 10 to nary predicates and 20 to unary attributes.

The specific values (2 vs 20) are arbitrary. What matters is the cost relative to other solutions. Thus in our attribute bias condition, it was twice as expensive to leave an attribute unmatched as it was to leave a relation unmatched.

4.3 Results

Table 3 shows the modeling results for the sparse "consistent" and "cross-map" conditions as well as the "feature-rich cross-map" condition. These results do not include the relational schema (e.g. Daddy-Mommy-Baby). Each condition was modeled with no bias, an attribute bias, and a relational bias. Assuming a hard 1-1 mapping constraint, our experimental results are consistent with the human-subjects findings.

Table 3. Results without relational schema: Object2 (the median sized object) is the correct (relational) answer. Object1 in the cross-map conditions reflects a match based on shared attributes. Each condition was modeled with no bias (None), an attribute bias (Attribute), or a relational bias (Relation). The * indicates a violation of 1-1 mapping occurred.

	1-1 Hard	1-1 Soft
Consistent: None	Object2	Object2
Consistent: Attribute	Object2	Object2
Consistent: Relation	Object2	Object2
Cross-Map: None	Object1	Object3*
Cross-Map: Attribute	Object1	Object3*
Cross-Map: Relation	Object2	Object3*
Feature Rich Cross-Map: None	Object1	Object1
Feature Rich Cross-Map: Attribute	Object1	Object1
Feature Rich Cross-Map: Relation	Object1	Object1
Feature Rich Cross-Map: 3xRelation	Object2	Object1

With no cross-mapping, our model selected object 2, which shared an attribute (e.g. small, medium, large) and relational (largerThan) mapping. In the cross-mapping condition, the shared attributes overwhelmed the relational mapping, leading to the selection of object 1 based on shared size category alone. This could be countered by augmenting the scoring function to reflect a relational bias as described above.

In the feature-rich cross-mapping condition, the attribute matches were chosen with much lower costs than the relational matches, so much so that the relational bias multiplier had to be increased by 3X, penalizing unaligned relations with a cost of 60, in order to prefer the relational match.

When modeling 1-1 as a soft constraint, the results are much the same in the first and third experiments, but the model predicts object 3 in the cross-map condition. Here, the model found a cheaper solution by doubly aligning object 2 in the base with the first and third object in the target and accepting the cost of the 1-1 violation. This is of course nonsensical, providing some evidence that 1-1 mapping ought to be modeled as a hard constraint.

Consistent with expectations from the human subjects experiment, adding a relational schema facilitated the relational mapping. Table 4 shows the modeling results for the sparse "consistent" and "cross-map" conditions as well as the "feature-rich cross-map" condition with the addition of the relational schema.

With a causal relationship between monotonic-size and the family schema, the hard 1-1 model now made the relational choice with no bias, and was able to make the relational choice in the feature-rich cross-map condition with the significantly smaller relational bias. However, the feature rich cross-map was unsuccessful with no bias, and an attribute bias could still overcome the relational schema. In short, our model predicts that a relational schema facilitates relational mapping, but that this effect is relative to the strength of cognitive bias and the complexity of the entities in the task.

The soft 1-1 mapping model preferred the relational match in the rich condition, but only by virtue of an unrealistic multiple-mapping, again suggesting 1-1 as a hard constraint.

Table 4. Results with relational schema: Object2 (the median sized object) is the correct (relational) answer. Object1 in cross-map conditions reflects a match based on shared attributes. Each condition was modeled with no bias (None), an attribute bias (Attribute), or a relational bias (Relation). The * indicates a violation of 1-1 mapping occurred.

	1-1 Hard	1-1 Soft
Consistent: None	Object2	Object2
Consistent: Attribute	Object2	Object2
Consistent: Relation	Object2	Object2
Cross-Map: None	Object2	Object2*
Cross-Map: Attribute	Object1	Object1
Cross-Map: Relation	Object2	Object2*
Feature Rich Cross-Map: None	Object1	Object2*
Feature Rich Cross-Map: Attribute	Object1	Object2*
Feature Rich Cross-Map: Relation	Object2	Object2*
Feature Rich Cross-Map: 3xRelation	Object2	Object2*

5. Related Work

CAMMA implements structural alignment as constraints in an abductive solve. Prior sections have discussed how CAMMA relates to its ancestors, SME and ACME (Falkenhainer et al., 1989; Forbus et al., 2017; Holyoak & Thagard, 1989). Space precludes an exhaustive account of existing similarity models; instead we classify CAMMA along the dimensions used in Forbus and Gentner's (2011) far more complete review of existing approaches.

Like SME, our approach is symbolic, as opposed to connectionist (e.g. ACME or LISA (Hummel & Holyoak, 2005)). Though we note that both CAMMA's cost function and soft constraint scorers could incorporate a neural model.

Gentner & Forbus (2011) characterize mapping algorithms as bottom-up, top-down, or local-global. The first begins with only entity matches, the second with target expressions, and the last starts with local potential matches and then finds a globally constrained mapping of expressions and entities.

Like ACME and SME, our approach is local to global. However, we implement the global phase as a generic optimization problem and can use a variety of ILP strategies. Unlike SME and ACME, we don't claim that CAMMA provides a process level account of the global optimization step, but we do argue that optimization is a useful abstraction for emulating and combining alternative theories, as we have done.

Gentner & Forbus (2011) also classify models based on whether they generate inferences following alignment or project the base onto the target (e.g. LISA). CAMMA occupies a middle ground. Inferences are projected during axiom generation, but they are bound as a part of the global optimization. While we don't model it in this paper, our approach could allow inference desirability to effect the final mapping through hard or soft constraints.

Finally, ACRE has been incorporated into the ARCS model of analogical retrieval (Thagard et al., 1990) and SME has been incorporated as a component in the MAC/FAC model of retrieval used by the SAGE model of analogical generalization (Forbus et al., 1995; McLure et al., 2010).

Currently, CAMMA is only a mapping model, however it could be a component in existing retrieval/generalization approaches. Furthermore, since inferences play a role in the solve, there may be opportunity to encode re-representation and generalization as a part of the hypothesis space.

6. Conclusions and Future Work

In this paper we have presented a novel implementation of analogy as constraints in an abductive solve, which, following Inoue & Inui (2011), we represent as an ILP problem.

Our resulting framework, called CAMMA, combines elements of two popular analogical frameworks, ACME and SME, under one paradigm. It further facilitates easy experimentation through declarative hard and soft constraints, and can use a variety of off the shelf ILP solvers for its global optimization.

As a demonstration, we used CAMMA to simulate the results of Gentner and Ratterman's (1991) classic cross-mapping experiments, finding that shared attributes can overwhelm relational matches, but that this can be overcome with a bias and/or richer schematic representation.

There are several directions for future work. First, we plan to integrate CAMMA with external models of biases and entity similarity through soft scoring constraints. We will also further investigate how world knowledge can be integrated to constrain alignment using background knowledge axioms. Finally, we will examine how axioms can encode re-representation strategies as a part of the solve.

Acknowledgements

Thank you to Andy Beck and Aditya Kalyanpur for their thoughtful feedback and guidance.

References

Christie, S., Gentner, D., Vosniadou, S., & Kayser, D. (2007). Relational similarity in identity relation: The role of language. *Proceedings of the Second European Cognitive Science Conference*.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial intelligence*, *41*, 1–63.

Forbus, K. D., Ferguson, R. W., Lovett, A., & Gentner, D. (2017). Extending sme to handle large-scale cognitive modeling. *Cognitive Science*, *41*, 1152–1201.

Forbus, K. D., Gentner, D., & Law, K. (1995). Mac/fac: A model of similarity-based retrieval. *Cognitive science*, *19*, 141–205.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive science*, *7*, 155–170.

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child development*, (pp. 47–59).

Gentner, D., & Forbus, K. D. (2011). Computational models of analogy. *Wiley interdisciplinary reviews: cognitive science*, *2*, 266–276.

Gentner, D., & Ratterman, M. (1991). *Language and the career of similarity*, (pp. 225–277). Cambridge University Press.

Hespos, S. J., Anderson, E., & Gentner, D. (2020). Structure-mapping processes enable infants' learning across domains including language. In *Language and concept acquisition from infancy through childhood*, 79–104. Springer.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive science*, *13*, 295–355.

Hummel, J. E., & Holyoak, K. J. (2005). Relational reasoning in a neurally plausible cognitive architecture: An overview of the lisa project. *Current Directions in Psychological Science*, *14*, 153–157.

Inoue, N., & Inui, K. (2011). Ilp-based reasoning for weighted abduction. *Plan, Activity, and Intent Recognition* (pp. 25–32).

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive psychology*, *50*, 315–353.

Lovett, A., & Forbus, K. (2012). Modeling multiple strategies for solving geometric analogy problems. *Cogsci 2012*.

Markman, A. B. (1997). Constraints on analogical inference. *Cognitive science*, *21*, 373–418.

Markman, A. B., & Medin, D. L. (2002). Decision making. *Stevens' Handbook of Experimental Psychology*.

McFate, C. J., & Forbus, K. D. (2016). Analogical generalization and retrieval for denominal verb interpretation. *CogSci 2016*.

McLure, M., Friedman, S. E., & Forbus, K. D. (2010). Combining progressive alignment and near-misses to learn concepts from sketches. *Proceedings of the 24th International Workshop on Qualitative Reasoning*.

Papadimitriou, C. H., & Steiglitz, K. (1998). *Combinatorial optimization: algorithms and complexity*. Courier Corporation.

Rabkina, I., & Forbus, K. D. (2019). Analogical reasoning for intent recognition and action prediction in multi-agent systems. *Proceedings of the Seventh Annual Conference on Advances in Cognitive Systems* (pp. 504–517).

Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive science*, *36*, 1019–1050.

Thagard, P., Holyoak, K. J., Nelson, G., & Gochfeld, D. (1990). Analog retrieval by constraint satisfaction. *Artificial intelligence*, *46*, 259–310.