A Creative Analogy Machine: Results and Challenges

Diarmuid P. O'Donoghue
Department of Computer Science,
National University of Ireland Maynooth,
Co. Kildare, Ireland.
diarmuid.odonoghue@nuim.ie

Mark T. Keane
Department of Computer Science and Informatics,
University College Dublin,
Ireland.
mark.keane@ucd.ie

Abstract

Are we any closer to creating an autonomous model of analogical reasoning that can generate new and creative analogical comparisons? A three-phase model of analogical reasoning is presented that encompasses the phases of *retrieval, mapping* and inference *validation*. The model of the retrieval phase maximizes its creativity by focusing on domain topology, combating the semantic locality suffered by other models. The *mapping* model builds on a standard model of the mapping phase, again making use of domain topology. A novel *validation* model helps ensure the quality of the inferences that are accepted by the model. We evaluated the ability of our tri-phase model to re-discover several *h-creative* analogies (Boden, 1992) from a background memory containing many potential source domains. The model successfully re-discovered all creative comparisons, even when given problem descriptions that more accurately reflect the original problem – rather than the standard (*post hoc*) representation of the analogy. Finally, some remaining challenges for a truly autonomous creative analogy machine are assessed.

Introduction

Analogy has a long and illustrious history within creativity, particularly within scientific and intellectual contexts (Brown, 2003). Many episodes of scientific creativity are driven by analogical comparisons (Dunbar and Blanchette, 2001), often involving image related analogies (Clement, 2008). Much progress has been made in cognitive science on modeling this analogical reasoning process (see below), prompting the following questions. Are we any closer to creating an autonomous model of the analogical reasoning that can generate new creative analogies? What progress has been made towards such a creative analogy model? What are the main challenges that lie ahead?

In this paper we envisage a creative process that can take any given target description and using a pre-stored collection of domain descriptions, identify potentially creative source domains with which to re-interpret the given problem. This paper explores and evaluates the potential for a model of analogy to act as a creativity engine.

While Boden (1992) argues that analogy is effectively the lowest form of creativity (*improbable*), we argue that analogical creativity should be seen a part of a cohesive human reasoning system. If the inferences mandated by an analogy contradicts a fundamental belief, especially one that has accrued many consequent implications, then resolving this contradiction might well involve the *"shock and amazement"* of transformational creativity. As such, it appears that analogies may drive creativity at any of Boden's levels of creativity. Our creativity model is domain independent and does not include a pragmatic component or domain context. So, as our model does not use domain-specific knowledge, arguably it cannot be easily cast as *improbable, exploratory* or *transformational* creativity (Boden, 1992).

The current work was driven by three main aims. Firstly, we wished to assess the creative potential of a three-phase model of analogy. Secondly, we wished to assess the impact of using differing knowledge bases upon the creative potential of our analogy model. Finally, we wished to assess the wider implications of analogical models for computational creativity. Is a three-phase model either necessary or sufficient to function as an engine of creativity? Can such a model re-discover analogies considered to be creative by people? Since people often overlook analogies (Gick and Holyoak, 1980) even when they are present, will such a model uncover many creative analogies or are creative analogies, in some way, different and rare?

We see the current model as being potentially useful in three distinct ways, but for now we do not commit to using it in one particular manner. Firstly, it could be used as a simple model of creativity, yielding creative interpretations for a presented problem. Secondly, it could be used as a tool to assist human creativity; suggesting source domains to people, to enable them to re-interpret a given target problem. Finally, it could be used as one possible model of how people analogize in a creative means.

The paper is structured as follows: first we describe the Kilaza[1] model for generating creative analogies, briefly illustrating its operation on the famous *atom:solar-system*

[1] Kilaza is not an acronym.

analogy. Then we present results that reflect the model's ability to re-generate some well-known *h-creative* analogies (Boden, 1992). Finally, the implications of these results are assessed and some remaining challenges are discussed.

Analogy as an Engine of Creativity

An analogy is a conceptual comparison between two collections of concepts, a *source* and *target* (Gentner, 1983), such that the source highlights particular aspects of the target, possibly suggesting some new inferences about it. In creative analogies, an productive source domain conjures up a new and revolutionary interpretations of the target domain, triggering novel inferences that help explain some previously incongruous phenomena or that help integrate some seemingly unrelated phenomena (Boden, 1992; Eysenck and Keane, 1995). Creative analogies differ from "ordinary" analogies primarily in the conceptual "distance" between the source and target domains (i.e., these two domains may never have been linked before) and the usefulness of the resulting comparison. Both creative and mundane analogies appear to use the same analogical reasoning process, as described in the following section, but different in their inputs and outputs.

Kekulé's is famous for his analogy between the carbon-chain and a snake biting its own tail. But this analogy could have been triggered by many alternative and more mundane source domains – from tying his own shoe-lace to buckling his belt. While many source domains could have generated the creative carbon-ring structure, Gick and Holyoak (1980) have shown most people (including Kekulé) frequently fail to notice many potential analogies. This highlights one potential advantage of a computational model, in that a model can tirelessly explore all potential analogies, returning only the most promising comparisons to a user for more detailed consideration. Thus, computational models could potentially act a tools helping people overcome one barrier; namely, their failure to perceive analogies when they are present.

Kilaza Analogical Creativity Engine

Keane (1994) presented a five-phase model of the analogical reasoning process, which recognises the distinct phases of *representation, retrieval, mapping, validation* and *induction*. While other authors describe slightly different subdivisions of this process, there is broad agreement on these phases. Our computational model encompasses the three central phases of analogy (see Figure 1). We highlight that Walls & Hadamard subdivide creativity into the phases of *preparation, incubation, illumination* and *verification* (Boden, 1992), which is reminiscent of several multi-phase models of analogy.

The heart of our creativity model is the central mapping phase and this borrows heavily from Keane and Brayshaw's (1988) IAM model (see also Keane, Ledgeway & Duff, 1994). Our model of the retrieval phase attempts to overcome the semantic bias suffered by many previous models, improving the diversity of the source domains that are returned. It was intended that this diversity might address the quality of *novelty* (Ritchie, 2001) associated with creativity, retrieving more "unexpected" and potentially creative sources. Finally, our model of the validation phase attempts to filter out invalid inferences, addressing the *quality* (Ritchie, 2001) factor associated with computational creativity.

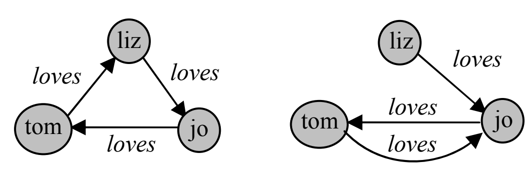
Ritchie (2001) identifies the essential properties of creativity as being *directed, novel* and *useful*. We argue that our model is *directed* in that it focuses on re-interpreting some given target domain. Our model addresses the *novelty* property by its ability to retrieve potentially useful but semantically distant, even disconnected, source domains. Finally, the *useful* property is addressed through a validation process that imposes a quality measure on the inferences that are accepted by the model.

retrieval → *mapping* → *validation*

Figure 1: Kilaza is a three-phase model of Analogy

Analogical Retrieval Phase-Model

Existing models for analogical retrieval suffer from the limitations in the range of possible retrievals because their they either (i) focus exclusively on domain semantics (like MAC/FAC; Forbus, Gentner and Law, 1995) or (ii) focus primarily on domain semantics (like HRR; Plate, 1998). Other models -- such as ARCS (Thagard *et al*, 1990) and Rebuilder (Gomes *et al*, 2006) - supplement domain representations by elaboration from external sources (like WordNet) to widen the net to include more semantically non-identical sources. However, all of these approaches arguably over-constraint retrieval for the the proposes of creativity. We argue that a creative retrieval process must allow semantically distant and even semantically disconnected sources to be retrieved, ideally without overwhelming the subsequent phase-models with irrelevant domains.



(a) Love Triangle (b) Unrequited Love

Figure 2: Topology is a key characteristic in retrieving creative source domains

Gentner (1983) mentions two specific qualities are required of analogical comparisons: semantic similarity and structural similarity. The model presented in this paper performs retrieval based exclusively on structural similarity, performing retrieval based exclusively on the graph structure (or topology) of each domain description. This design decision was taken to overcome the semantic narrowness that constrains existing models, with the hope that this would increase the possibility of retrieving surprising and creative source domains. As the example in Figure 2 illustrates, semantics and domain topology are often intertwined.

Each domain description is mapped onto a location in an *n*-dimensional structure space (Figure 3), where each dimension represents a particular topological quality of that domain. Structure space is somewhat akin to feature vectors (Yanner and Goel, 2006; Davies, Goel and Yanner, 2008). Image related analogies are often involved in creative comparisons (Clement, 2008) and a variety of image-based analogy models has been developed, focusing on specific topics such as; geometric proportional (IQ type) analogies (Evans, 1967; Bohan and O'Donoghue, 2000), geo-spatial comparisons (O'Donoghue *et al*, 2006), spatial representations of conceptual analogies (Davies et al, 2008; Yanner *et al*, 2008) and reasoning about sketch diagrams (Forbus *et al*, 2011). Our model performs a single retrieval process for each presented target, in contrast to the iterative retrieval and spreading activation phases employed by KDSA to retrieve semantically distant sources (Wolverton and Hayes-Roth, 1994).

Specific topological features used by our retrieval model include quantifying the number of objects and predicates (first order and higher order) and number of root predicates *etc*. Thus, the representation in Figure 4 might be mapped onto the location (4 0 2 2 0 0 1) in structure space – 4 object references, 0 high-order predicates, 2 unique first-order relations, 2 first-order relations and 2 root predicates *etc*. The distinction between unique and non-unique relations, for example, distinguishes between domains repeatedly using a small number of relations and domains that typically have one instance of each relation in its description. One advantage of this scheme is that the distance between domains is not impacted by the number of domains contained in memory so the retrieval system should scale reasonably well. For the retrieval results presented later in this paper a maximum retrieval distance of 10 is imposed – and only candidate source inside this threshold are considered.

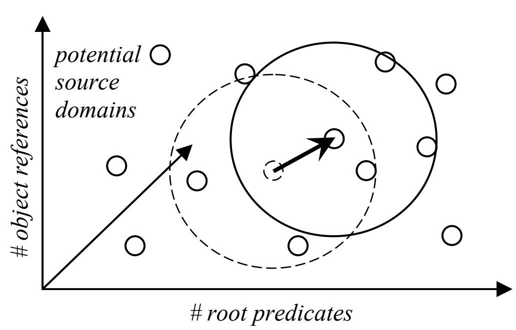


Figure 3: Displacing the Locus of Retrieval within a 3D representation of *n*-dimensional Structure Space. Only source domains within the displaced boundary are retrieved and passed to the remaining phases of analogy.

Topologically similar (i.e., homomorphic as well as isomorphic) domains are mapped onto similar locations within this topology-based structure space (O'Donoghue and Crean, 2002). To account for the inferences that were sought from any inspiring source domain, the locus of retrieval was slightly offset to account for this additional source domain material. Included in this offset is the desire for sources containing additional first-order relations and high-order relations. However, this offset has relatively little impact on the final results.

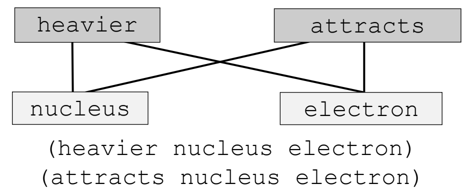


Figure 4: Simplified Model of Rutherford's Problem

Analogical Mapping Phase-Model

The model for the mapping phase is based on the Incremental Analogy Machine (IAM) model (Keane & Bradshaw, 1988; Keane et al, 1994). It consists of the three sub-processes of *root-selection, root-elaboration* and *inference generation*. Mapping proceeds as a sequence of root-selection and root-elaboration activities, gradually building up a single inter-domain mapping. Typically a domain description will consist of a small number of root predicates, each controlling a large number of (partly overlapping) lower-order predicates.

Root selection Root selection identifies "root predicates" within a representation, which are typically the controlling causal relations in that domain. Each root predicate lies at the root of a tree of predicates and each root is be seen as "controlling" the relations lower down the tree. In our implementation of IAM, the root-selection process examines the "order" of each predicate. Objects are defined as order zero and first-order relations that connect two objects are defined as order one. The order of a causal relation is defined as one plus the maximum order of its arguments. Mapping begins with the highest order relations and maps any unmapped low-order root-predicates last.

Root elaboration Root elaboration extends each root-mapping, placing the corresponding arguments of these relations in alignment. If these arguments are themselves relations, then their arguments are mapped in turn and so on until object arguments are mapped. Items are only added to the inter-domain mapping when they conform to the 1-to-1 mapping constraint (Gentner, 1983).

Inference Generation Each analogical comparison is passed to the inference generation sub-process. Analogical inferences are generated using the standard algorithm for pattern completion CWSG – Copy With Substitution and

Generation (Holyoak *et al*, 1994). In effect, additional information contained in the source domain is carried over to the target, creating a more cohesive understanding of that target problem.

Analogical Validation Phase-Model

The third part of our tri-phase model is focused on analogical validation. Validation attempts to ensure that the analogical inferences that are produced are correct and useful.

O'Donoghue (2007) discusses the accuracy of this validation process, using human raters to assess the goodness of inferences that were rated as either valid or invalid. However, this paper did not assess the models ability to discover creative analogies.

Phineas (Falkenhainer, 1990) is a multi-phase model of analogy that incorporates a post-mapping verification process. To achieve this Phineas incorporates a model of the target domain – qualitative physics simulation - illustrating the power of embedding an analogy model within a specific problem domain. However, this qualitative-simulation process effectively limits Phineas to reasoning only about physical and physics-related analogies.

The validation model presented in this paper is relatively simple, aimed at rejecting those predicates that are deemed invalid – rather than guaranteeing the validity of those inferences that are accepted. This approach helped maximise the creative potential of this model, by resisting the rejection of potentially plausible inferences. Of course, a more complex validation process could make use of problem-specific domain knowledge (where available). In the absence of such domain-specific knowledge verification and validation of the analogy could be carried using user feedback, employing Kilaza in a tool-like way.

The validation phase-model is composed to two main parts. The first performs validation by comparing the newly generated inference to predicates already stored somewhere in memory. The second mode of validation is more general and driven in part by the functionally relevant attributes that play a role in analogical inference (Keane, 1985).

Validation by Predicate Comparison The validation process compares newly inferred predicates (produced by CWSG) to the previous contents of memory. Inferences are firstly compared to predicates in memory, with both the agent and patient roles potentially being validated independently. This validation mechanism thus has access to the entire contents of memory, accessing predicates from any of the domains stored in that memory. This model of validation captures the advantages of simplicity and generality, but it does of course mean that dependencies between arguments are not captured. This limitation was deemed acceptable within the context of our desire for a creativity engine. While many simple inferences were validated by this mechanism, many creative inferences were not. This may be partly attributed to the relatively small number of predicates contained in memory and to the novelty associated with creative inferences. To address this shortcoming validation using functional attributes was introduced.

Validation with Functional Attributes Functional attributes specify necessary attribute requirements for each role of a predicate – being inspired by the *functionally relevant attributes* of Keane (1985). Functional attributes are intra-predicate constraints that ensure each predicate appears to be a plausible combination of a relation coupled with each of its arguments.

It should be pointed out that functional attributes have only been used with first-order predicates – those whose arguments are objects. Although validating higher-order (causal) relations might make use of the spatio-temporal contiguity associated with causality, but this cannot be relied upon (Pazzani, 1991) and is not enforced by our model. Thus our model treats all causal inferences as implicitly valid.

Functional attribute definitions connect each role of a predicate directly into an attribute hierarchy, whereby arguments filling those roles must conform to these attribute constraints. Kilaza stores functional attributes for both the agent and patient arguments of each relation independently. More general relations (*part-of, next-to*) typically have few functional attributes, whereas more specific relations (*hit, eat*) possess a greater number of attribute restrictions. For example the agent role of *hit* might require the hitter to be a physical object, whereas the agent of an *eat* relation might have to be a living organism or an animal. Relations that are more specific are seen to be more amenable to the validation process, while their more general counterparts are more difficult to validate accurately.

In addition, functional attributes have also been used to support a form of inference adaptation. This allows an inferred relation to be adapted to a semantically similar relation that better suits the arguments that pre-existed within in the target domain. Adaptation uses the functional attributes to conduct a local search of the taxonomy, to identify a more semantically suitable relation that better fits the given arguments.

Data Sets

Three datasets were used to conduct experiments using the described model. These are referred to as the Professions dataset, the Assorted dataset and an Alphanumeric dataset. The dataset contained a total of 158 domains and our creativity engine attempted to find creative source analogues for a given number of target problems. It was hoped that the differing natures of these collections would provide a reasonable grounds on which to evaluate the computational model – and to assess its potential to act as a creativity engine.

Professions Dataset consists of descriptions of fourteen professions, including *accountant, butcher, priest* and *scientist*. These are rather large domain descriptions created by Veale (1995) and range in size from 10 to 105 predicates (M=55.4, SD=29.3). One important feature of the Professions dataset is its reliance on many different instances of a small number of relational predicates, including *control, affect, depend,* and *part*. The domains range from using just 6 distinct relational predicates (ignoring duplicates) to the most diverse domain that uses 15 (M=8.9, SD=2.2). Another important feature is that this dataset does not appear to use a set of clearly identifiable high-order relations (such as a *cause, result-in* or *inhibit*) between first-order predicates.

Assorted Dataset consists of a large number of smaller and more varied domain descriptions, including many of the frequently referenced domains in the analogy literature; such as the *solar-system, atom, heat-flow* and *water-flow* domains. It also includes an assortment of other domains describing *golf, soccer* and *story-telling*. The 81 domains of the Assorted dataset use 108 distinct (*ie* non-repeated) relations. Each of these domains contains between 1 and 15 predicates (M=4.16, SD = 2.9). The average number of distinct relational predicates in each domain is M=3.48, indicating that most relational predicates are used just once in each of the Assorted domains.

Alphanumeric Dataset One final dataset contained 62 semantically constrained domains. However, these domains contained a great deal of topological diversity. It was hoped that this mixture of topologies might support some novel comparisons and inferences and provided a counterpoint to the semantic richness of the other domains.

Example: p-Creative Re-Discovery of Rutherford's Analogy

Before presenting detailed results, we will first see how Kilaza can re-discover Rutherford's famous *solar-system:atom* analogy. We highlight that this is a test for the p-creativity (Boden 1992) of our model – though not necessarily a model how Ernest Rutherford actually conducted his own reasoning.

The traditional representation of this analogy (Figure 5) is heavily based on a *post hoc* description of the domains involved. These descriptions and are heavily influenced by the analogy itself. We shall first look at the traditional representation of this domain, before examining how our model can also deal with more realistic version of how Rutherford might have thought of the target problem *before* arriving at his famous comparison.

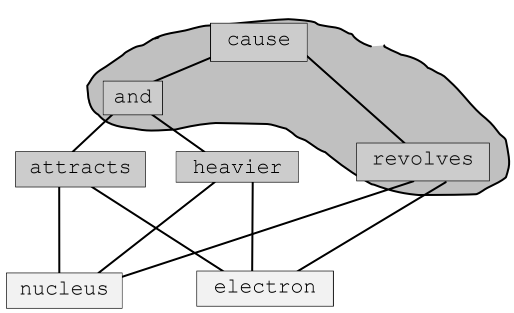


Figure 5: Traditional representation of Rutherford's Solution

First, the semantically impoverished target problem (Figure 4) is mapped onto its location in structure space. We highlight that the "locus of retrieval" is slightly displaced from the targets original location to account for the additional information that one expects to be found in a useful source domain. In this instance the desired source was retrieved at a distance of just over 6 "units" in structure space. The desired source (the *solar-system* domain) and all other candidate sources near the locus of retrieval were passed in turn to the mapping and validation phases of the model.

In total 10 other candidate source domains that were retrieved also generated inferences, most yielding only one inference each. Three domains generated more than one candidate inference – but all three were different versions of the *solar-system* domain. We point out that our semantic "free" retrieval process can also trigger identification of the same source, even if it was represented in a number of alternate ways (O'Donoghue, 2007). Our mapping model successfully generated the correct inter-domain mapping and CWSG generated the desired inferences without adaptation.

Representation Issues in *de novo* Discovery of p-creative analogies

We argue that the traditional presentation of Rutherford's analogy is a simplified pedagogical device (Figure 5). This description of the target problem effectively removes much of the complexity of the real discovery task as encountered by Rutherford. The description of the target problem uses terminology specifically designed to accentuate the semantic (and structural) similarity that is the *result* of Rutherford's comparison – and should not be treated as an input when re-creating this creative episode.

This distinction between the problem domain as it would have existed *before* the creative analogy and its subsequent representation *after* discovering that analogy is a serious problem - one that is easily overlooked. Any model that attempts to re-discover known creative analogies must address the original problem, not just the representation that accentuates the desired similarity. Differences in domain terminology and topology are central to the distinction between elaborating a given analogy, and the much more difficult task of generating a novel *h-creative* (or *p-creative*) analogy (Boden, 1992).

We argue that generating Rutherford's analogy using the representation in Figure 6 is a far better test of a models creative ability, than the normal *post hoc* representation in Figure 5. Terminological differences are particularly prevalent in distant between-domains analogies as the first-order relationships describing the problem domains originate in different disciplines. When modeling analogical creativity, we must expect to encounter these differences in terminology, and our models of retrieval, mapping and validation must be able to overcome these problems.

Ernest Rutherford would most likely have thought of the target relation between the nucleus and electron as `electromagnetic-attraction`, and not the more generic `attracts` relation. The corresponding relationship between source's sun and planet is `gravitation`. It is only after he found the analogy (which involved mapping electromagnetic-attraction with gravitation) that these relationships can be generalized to a common super-class like attracts (Gentner, 1983).

We point out that our model can operate successfully on either the simplified or more realistic domain descriptions. This is primarily the result of our retrieval and mapping models using domain topology, rather than using identicality (or similarity) between the predicates in both domains.

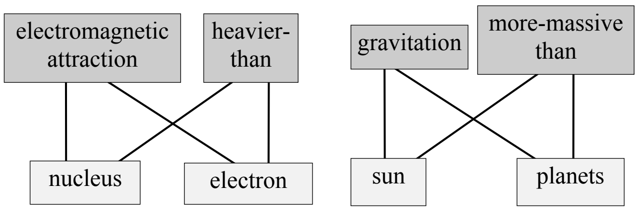


Figure 6: More realistic representation of Rutherford's Analogy

Results of Individual Phase Models

We shall first briefly examine the performance of the *retrieval* and *validation* models in isolation, before looking at their combined performance in the next section. We shall briefly examine the results of the *mapping* model, but our focus will remain on the inferences that it produced. Results were produced from a memory containing the three previously described datasets.

Retrieval Results Retrieval was performed in structure space. The distance between domains in structure space varied from 2.645 to 230 (M= 80, SD=57.3), with a large number of domains being given a unique structural index in this space. A small number of locations contained multiple domains – these mostly involved small domains of just a few predicates from the Assorted dataset.

Retrieval and Mapping A broad tendency was identified between structure-based retrieval and the size of the resulting inter-domain mapping, although the correlation was low. A range effect was identified between structure space and the size of the resulting mapping, indicating that larger distances between domains in structure space tend to produce smaller inter-domain mappings. This indicates a weak connection between structure-based retrieval and the size of any resulting mappings.

Validation Results Although the validation model was very simplistic, it proved surprisingly effective. For example with the inferences generated on the Professions dataset, the average (human) rating awarded to predicates that Kilaza categorized as valid was M=2.62 (SD=2.09), while the average rating awarded to the invalid predicates was M=1.57 (SD=1.23). As ratings were given between 1 and 7 with 7 representing clearly valid inferences, this indicates that many of the generated inferences were of rather poor quality.

Adaptation Results In addition, 24 inferences were passed to the adaptation process and 20 of these were adapted. While we *cannot* realistically assess if these adapted inferences matched what was "intended" by our analogy model, we did assess the validity of these inferences using two human raters.

When we look at human ratings for the 20 adapted predicates before and after adaptation, we see that the average ratings were increased by the adaptation process - from 1.57 (SD=1.23) to 2.57 (SD=1.70). The average ratings of the adapted predicates was broadly in line with the predicates from Kilaza's valid category above (M=2.62, SD=2.09). Before adaptation, 18 of the 20 (90%) predicates were given rated as invalid and after adaptation just 12 (60%) were rated as *invalid*. Thus, adaptation has a distinct influence on improving the ratings of the rejected inferences.

It may well be argued that this adaptation process is itself somewhat creative – identifying new relations that better fit the available target arguments. In contrast to the top-down nature of the creative analogy approach, predicate adaption is a very much a bottom-up process that is motivated by the detection of a potential analogical comparison.

Creativity Test Results

To assess the creative potential of our model, we assess its performance at the *p-creative* task of re-creating some well-known *h-creative* analogies (Boden, 1992). These include some of the famous examples of creative analogical comparisons including the Rutherford's *solar-system:atom* analogy, the *heat-flow:water-flow* and the *tumour:fortress* analogies. Our descriptions are based on the standard representation of these domains as found in the analogy literature.

Creative Retrieval We now examine the performance of our model on the creative retrieval task. We presented our model with the target domain of each of 10 creative analogies, together with a memory of 158 source domains. From this memory of 158 potential sources, the retrieval model selected a number of these domains as candidate sources. Only the selected candidate sources were passed to the mapping and validation phase-models. Evaluating only the selected source domains was necessary in order to avoid an exhaustive search through all possible analogical comparisons. While computationally feasibly in this instance, an exhaustive search would be impractical on a larger collection of domains.

Before looking at the results, we point out that many comparisons did not generate a viable inter-domain mapping. Furthermore, most analogies did not generate any valid inferences. The following results ignore these unproductive comparisons and we focus only on the productive analogies.

All of the desired creative sources were among the candidate sources that were retrieved by the model. This gives our retrieval model a *recall* value of 100% on this creative retrieval task. While a large number of other candidate sources were also retrieved, this was still a pleasantly surprising result. The distance within structure space between the target and the creative sources ranged from 3.1 to 7.9, suggesting that structure based retrieval was reasonably accurate in locating candidate sources.

The *precision* of the retrieval processing is summarised in Figure 7. As can be seen, precision was above 0.2 for two problems showing that few other sources were located near the structural index of those targets. However, precision was much lower for most problems, indicating that the desired source was merely one of a larger number of candidate sources that had to be explored.

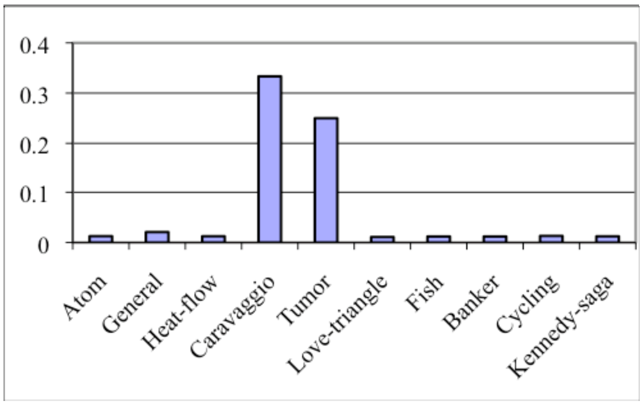


Figure 7 –*Precision of retrieval for 10 Creative Analogies*

Creative Inferences Next we summarise the inferences that were generated by each of these comparisons (Table 1). These results implicitly encompass a productive inter-domain mapping between the target and each candidate source in turn. Kilaza generated and validated the correct inferences for 9 (70%) of the creative analogies. The *cycling:driving* analogy correctly generated no inferences.

Target	Correct Inferences	Validated Inferences
Atom: Solar-System	y	4
Atom-Falkenhainer: Solar-System-Falkenhainer	y	3
General: Surgeon	y	4
Heat-flow: Water-Flow	y	4
Leadbelly : Caravaggio	y	4
Love-triangle: Triangle-Directed	y	0
Requited-love: Love Triangle	y	3
Fish : Bird	y	4
Vampire : Banker	y	3
Cycling: Driving	n	0

Table 1 – Number of Inferences generated by different analogies

One of these analogies also required one inference to be adapted. The `bird:fish` analogy generated the inference `(flies-through fish water)`, which was correctly adapted to `(swim fish water)`.

Conclusion

We presented a three-phase model of analogy, adapting it to function as a tool for discovering creative analogies. This model encompasses the three central phases of analogy, namely *retrieval, mapping* and *validation*. We argue that a model encompassing these three core phases of analogy is the minimum required to be considered a model of analogical creativity.

Our *retrieval* model overcomes the semantic bias of previous retrieval models, helping retrieve new and surprising source domains. This helps to improve the *novelty* of the source domains identified by our creativity engine. Our model of the post-mapping *validation* phase attempts to filter out any clearly invalid inferences, thereby improving the *quality* of the analogies identified as being creative. We note that novelty and quality are two attributes strongly associated with creativity (Ritchie, 2001).

Our three-phase model of analogy successfully rediscovered 10 examples of creative analogies, including the *heat-flow:water-flow* and *solar-system-atom* analogies. In doing so, the model retrieved the correct source from a large memory of potential sources. It then developed the correct mapping and successfully validated (and adapted) the resulting inferences. We point out that these analogical comparisons, if produced by a human analogizer, would be considered creative.

Our focus on creative analogies rather that the more normal (or pedagogical) analogies had a far-reaching impact on the model. Terminological differences are particularly prevalent in creative between-domains analogies, as the first-order relations describing each domain originate in different disciplines. When modeling analogical creativity, we must expect to encounter these differences and cannot rely heavily on the presence of identical relations. Our model successfully created Rutherford's famous *solar-system:atom* analogy, even when the target was represented in a more realistic and challenging from. Our model shows that very significant progress has been made towards an autonomous creativity machine, re-discovering many creative analogies.

We briefly outline three remaining challenges to analogical creativity, beginning with the issue of knowledge representation. Our results illustrate a trade-off between the *specifity* and the *generality* of domain descriptions. Overly specific representations make comparisons more difficult to discover, but overly general representations appear too profligate and can overwhelm the validation (and subsequent) processes. Perhaps multiple representations of each domain might offer a useful avenue for progress. Multiple representations might also help explain why exerts are more fluent in their use of analogy within their own domains (Dunbar and Blanchette, 2001). Our model does not currently include an explicit re-representation process highlighting "*tiered identicality*" (Gentner and Kurtz, 2006).

It seems that the greatest challenge to computational analogizing might lie with the post-mapping phases. Challenges include assessing analogical inferences for validity, evaluating the significance of an analogy and considering the implications of creative comparisons. Surprisingly little attention has been given to this phase – partly because of its ultimate dependency on the target problem domain. Phineas (Falkenhainer, 1990) and also Rebuilder (Gomes *et al*, 2006) showed that integration of the analogy and case-based reasoning within the target domain can have very positive effects. While tight integration of all target domains into an analogy model seems most unlikely, Kilaza has show that a generic validation model can play a part improving the quality of the inferences that are accepted.

Overall, the results presented in this paper highlight that a three-phase model of analogical reasoning can operate successfully as a model of analogical creativity. Our results highlight the improbability of finding a suitable source domain to re-interpret a given target in a creative manner. Extending this model will necessitate a tighter integration of the analogy process with other facets of intelligence.

References

Boden, M. 1992. *The Creative Mind*. London: Abacus.

Bohan, A. O'Donoghue, D.P. 2000. A Model for Geometric Analogies using Attribute Matching, *Proc AICS*, Galway, Ireland, pp 110-119.

Brown, T.L. 2003. *Making Truth: Metaphor in Science*. University of Illinois Press, New York, USA.

Clement, J. 2008. *Creative Model Construction in Scientists and Students: The Role of Imagery, Analogy, and Mental Simulation*. Dordrecht: Springer.

Davies, J. Goel, A.K. Yaner P.W. 2008. Proteus:Visuospatial analogy in problem-solving, *Knowledge-Based Systems* 21(7): 636-654.

Dunbar, K. and Blanchette, I. 2001. The in vivo/in vitro approach to cognition: The case of analogy, *Trends in Cognitive Sciences*, 5(8): 334–339.

Falkenhainer, B. 1990 *A Unified Approach to Explanation and Theory Formation*, in Shrager, J. & Langley, P. (eds.) Computational Models of Scientific Discovery and Theory Formation, Morgan Kaufman: CA. pp 157-196.

Forbus, K. Gentner, D. Law K. 1995. MAC/FAC: A Model of Similarity-based Retrieval, *Cognitive Science*, 19(2): 141-205.

Forbus, K. Usher, J. Lovett, A Lockwood, K. Wetzel J. 2011. CogSketch: Sketch Understanding for Cognitive Science Research and for Education, *Topics in Cognitive Science*, 3(4): 648–666.

Gentner, D. Kurtz, J. 2006. Relations, Objects, and the Composition of Analogies, *Cognitive Science,* 30(4): 609–642.

Gentner, D. 1983. Structure-mapping: A theoretical framework for analogy, *Cognitive Science,* 7(2): 155–170.

Gomes P, Seco N, Pereira FC, Paiva P, Carreiro P, Ferreira JL, Bento C. 2006. The importance of retrieval in creative design analogies, *Knowledge-Based Systems*, 19(7): 480–488.

Holyoak K. J. Novick L. Melz E. 1994. *Component Processes in Analogical Transfer: Mapping, Pattern Completion and Adaptation*, in Analogy, Metaphor and Reminding, Eds. Barnden and Holyoak, Ablex, Norwood, NJ.

Keane, M.T. 1985. On Drawing Analogies When Solving Problem, *British Journal of Psychology*, 76: 449-458.

Keane, M.T., Bradshaw, M. 1988. The Incremental Analogical machine: In D. Sleeman (Ed.) *3rd European Working Session on Machine Learning*, Kaufmann CA: 53–62.

Keane, M.T., Ledgeway, T. Duff, S. (1994). Constraints on analogical mapping, *Cognitive Science,* 18: 387-438.

O'Donoghue, D.P. 2007. Statistical Evaluation of Process-Centric Computational Creativity: *4th International Joint Workshop on Computational Creativity (IJWCC),* Goldsmiths, University of London, 17-19 June.

O'Donoghue, D.P, Bohan, A. Keane M.T, 2006. Seeing Things: Inventive Reasoning with Geometric Analogies and Topographic Maps, *New Generation Computing*, Ohmsha Ltd. and Springer 24(3): 267-288.

O'Donoghue, D.P. Crean, B. 2002. Searching for Serendipitous Analogies, *ECAI - Workshop on Creative Systems,* Lyon, France, 21-26 July.

Pazzani, M. 1991. A Computational Theory of Learning Causal Relationships, *Cognitive Science*, 15(3): 401-424.

Plate T. 1998. *Structured Operations with Distributed Vector Representations*, in Advances in Analogy Research, New Bulgarian University, Sofia, Bulgaria.

Ritchie G. 2007. Some Empirical Criteria for Attributing Creativity to a Computer Program, *Minds & Machines,* 17: 67–99.

Ritchie G. 2001. Assessing Creativity, *Proc. AISB Symposium on AI and Creativity,* York, March.

Thagard, P. Holyoak K. J. Nelson, G. Gochfeld, D. 1990. Analogue Retrieval by Constraint Satisfaction, *Artificial Intelligence*, 46: 259-10.

Veale, T. 1995. *Metaphor, Memory and Meaning*, PhD Thesis, Trinity College, Dublin, Ireland.

Wolverton M. Hayes-Roth, B. 1994. Retrieving Semantically Distant Analogies with Knowledge-Directed Spreading Activation, *Proceedings AAAI*, pp 56-61.

Yaner, P.W. Goel A.K. 2006 Visual analogy: Viewing analogical retrieval and mapping as constraint satisfaction problems, *Applied Intelligence,* 25(1): 91-105.