Making Mountains out of Molehills: Abstraction without Information Loss in Analogical Mapping

Tony Veale
School of Computer Science,
University College Dublin,
Dublin D4, Ireland.
tony.veale@UCD.ie

Abstract

Analogy is a cognitive process that propels many of our most creative leaps, from the cross-domain forays of scientific discovery and case-based reasoning to the poetry of metaphor, bisociation and blending. By concerning itself with the *shape* of meanings, and the structural arrangements of their parts, analogy allows us to unite different meanings with similar shapes across distant domains. Crucial to this unification, or mapping, of domains is the ability to abstract over structural representations that capture the broad sweep of an idea without getting bogged down in details. This paper presents a large new resource for analogical mapping that defines structured representations to support abstraction at multiple levels, but without information loss. This allows the mapping process to be flexible in its reconciliation of different meanings, while also preserving the distinctions that make abstraction necessary in the first place. This resource, named ATLAS, is a wide-ranging database of symbolic structures for lexical concepts (the ideas behind common words), and is designed to support explicit analogical reasoning in an era where symbolic reasoning is giving way to the statistics of LLMs.

Introduction: The Mind Has Mountains

Our most compelling and memorable arguments often rest on an analogy, even if a sharp insight can easily cut both ways. Consider this exchange from the 1993 film Jurassic Park, in which the park's owner, Hammond, defends the carnage that has ensued with an analogy to an iconic resort:

> **Hammond**: All major theme parks have delays. When they opened Disneyland in 1956, nothing worked!
>
> **Malcolm**: Yeah, but, John, if *The Pirates of the Caribbean* breaks down, the pirates don't eat the tourists.

In Hammond's analogy, Disneyland serves as the familiar *source* domain through which the implied *target*, Jurassic park, is viewed (Gentner 1983). But Malcolm inverts this analogy, and instead uses the latter as a source domain to reimagine the former. Each analogy re-frames the context to suit the speaker's own worldview: Hammond seeks to downplay the enormity of his park's failings, while Malcolm's counterfactual aims to remind him of the fatal consequences of his hubris. It is the isomorphic structure of each domain – immersive theme parks with hi-tech attractions – that enables these analogies to map so easily from one to the other, as in *animatronic* to *genetic* and *pirate* to *dinosaur*, and to suggest new candidate inferences, such as Malcolm's allusion to tourist-gobbling pirates (Gentner and Toupin 1986; Falkenhainer, Forbus, and Gentner 1989). Identifying this shared structure is the job of an analogical mapping engine, yet such structures must already be present for such a mapping to occur. Creating a large database of these structures to support analogical mapping and retrieval is our goal here.

As in metaphor and simile, analogy relies on abstraction to reconcile two different domains of experience or conceptual content. Abstraction is what allows us to get to the nub of an idea, and to equate very different meanings on the basis of a common gist (Gentner and Hoyos 2017). In analogy, this shared core is essentially structural, even if the two domains also have emotions and other features in common. There are two basic approaches to structural abstraction in analogy: in the first, the analogist identifies a subset of the content of one domain that is isomorphic to a corresponding sub-structure in the other, where isomorphism insists that both sub-structures have the same shape, and the same arrangement of labeled edges. If those edges connect different vertices in the source and target, a consistent 1-to-1 mapping from source to target vertices must also be possible. In essence, analogical structure mapping is a variant of the NP-hard problem *largest sub-graph isomorphism*, and so is itself NP-hard as a result (Veale and Keane 1997). The second approach generalizes an isomorphic mapping to sub-structures that have the same shape even if their edges have different labels, and so produces a mapping of vertices *and* relations, provided the relations in one structure are similar to those in the other. This second approach lifts abstraction to the next level, by allowing domains themselves to be re-represented during the mapping process (Hofstadter and Mitchell 1995).

(Gentner and Hoyos 2017) define "abstraction" as the decrease in specificity, and concomitant increase in scope, of a concept. Thus, as a source structure sheds specific details, it becomes applicable to a broader range of analogical targets. In this view, abstraction necessarily entails information loss, as only the most systematic core of the structure is retained. However, in an analogical mapping, only certain elements – the relational edges – are matched identically, while others – the vertices – are mapped non-identically in 1-to-1 pairings.

A loss-free process of abstraction via re-representation can preserve information if it shifts components that are matched identically into positions that are mapped non-identically. We present such a process in this paper, to build a resource, named ATLAS, that provides conceptual structures at multiple levels of abstraction for a wide range of words and ideas.

We must first define what we mean by a conceptual structure, and what it means to abstract over such a structure. To populate ATLAS at scale, we will leverage existing resources that were previously built to enable the generation of metaphors, blends and stories, such as the *talking points* model of (Veale 2014; 2015). As the semantic triples defined by this model are too shallow to support structure mapping as defined in (Gentner 1983; Falkenhainer, Forbus, and Gentner 1989), we shall present a rule-based means of constructing richly nested structures from these rather flat ingredients. As the new structures will vary in depth and breadth, we will define a measure of richness that will allow a mapping engine to rank the structures that define a concept by how well they embody the *systematicity principle* of (Gentner 1983).

We will also present a mapping algorithm and a retrieval algorithm to work efficiently with the structures of ATLAS. The analogy literature is replete with alternatives (Falkenhainer, Forbus, and Gentner 1989; Holyoak and Thagard 1989; Hofstadter and Mitchell 1995; Veale and Keane 1997), but we shall define variants of the highly efficient structure-hashing approach of (Veale 2005) that work especially well with our model of loss-free abstraction. ATLAS defines over two million structures across more than 9000 word concepts, making efficient storage, retrieval and mapping a necessity. But we must start at the beginning, to consider the assumptions that are made by structure-mapping models of analogy.

That would be an ecumenical matter

Representation is always a matter of choice: of what to represent, of which viewpoint to adopt, of what level of detail to capture, and with what feature set to encode it. An open representation can utilize an unlimited set of symbols, while a small number of primitives may suffice in a closed system. These choices introduce a great deal of wiggle room into any symbolic representation that is used for analogical mapping.

Every analogical system will bring its own assumptions to bear upon its representation of the world. Consider SME, the Structure-Mapping Engine of (Falkenhainer, Forbus, and Gentner 1989), which is a computational realization of the structure-mapping theory (SMT) of (Gentner 1983). While SME/SMT is open with respect to the choice of symbols for relations and attributes in any given domain, it assumes that the most systematic relations – such as those for causality – are shared across domains, and that these relations form hierarchical structures corresponding to nested propositions. The graph representation of a domain is thus a forest of interconnected tree structures, and a mapping can be constructed for the whole, or the largest sub-graph of the whole, by first finding partial mappings for the largest and deepest trees. A global mapping is then produced by merging a consistent selection of these partial mappings into a coherent whole.

Nesting is another choice that a representation may use, or not. The Sapper model of (Veale and Keane 1997) represents domains as a semantic-network of vertices and labeled, directed edges. An edge encodes a relation between two entities, and takes its label from a closed set of relation types. Sapper does not define distinct domains for different ideas. Rather, a single large graph encodes all concepts, so that the local region around a given node is the domain of that node. With no pockets of nested structure to exploit, Sapper seeks out pathways in the network that originate at the nodes for the two ideas that one is aiming to map. When pathways that comprise the same sequence of relations meet at a common association, they are placed into analogical correspondence to form a partial mapping. As in SME, a global mapping is constructed by merging partial maps into a coherent whole.

Each model works rather well on its own representations, but Sapper simply cannot use SME's representations, while SME finds no nesting to exploit in Sapper's graph structures. SME senses systematicity in the depth of nested structures, but Sapper sees it in the length of extended pathways. Each is blind to what makes the other's mappings systematic. If a common representation, and a shared database of structures usable by both approaches, is thus not feasible, we might, instead, map from one model's worldview onto the other's. We can, for instance, flatten SME's nested representations into a semantic network by introducing new nodes to represent branching points in its structures. Conversely, we can add causal nesting to Sapper-like graphs, to represent the higher-order connections between flat node-to-node relations.

We take the latter route here, but we need a starting point: a database of content on which to apply our transformations. The AnalogyKB of (Yuan et al. 2024) is a large knowledge-base of structured symbolic content derived from *WikiData* (Vrandečćíc and Krötzsch 2014) and *ConceptNet* (Speer, Chin, and Havasi 2017). Each resource expresses its knowledge in the form of predicate-subject-object triples; *WikiData* triples convey dry, objective facts (such as the CEO of company X is Y, or the capital of country A is B), while *ConceptNet* offers more subjective claims about the world. This kind of knowledge supports proportional reasoning of the *A is to B as C is to D* variety (A:B::C:D), as in Sam Altman is to OpenAI as Donald Trump is to USA. Proportional analogy, first studied by Aristotle in his *Poetics*, was once a cornerstone of the *Scholastic Aptitude Test* (or SAT), and computational solvers have been presented by (Veale 2004) and (Turney and Littman 2005), among others. These analogies are sometimes the tip of a structural iceberg, implying mappings beyond that of A to B and C to D, but we need to excavate that structure to support richer, deeper analogies.

The knowledge defined in (Veale 2014; 2015) to support story generation systems also has a flat triple structure, but since these triples are designed to work together, or indeed to *click* together like pieces of track, it is easier to imagine how a nested causal structure can be imposed on top. Moreover, these story-oriented triples concern commonsense ideas and stereotypical associations, not contingent facts, and can thus support analogies at the level of ideas rather than facts. In the next section we describe how rich causal structures can be assembled at scale from these flat triples, to support analogical mapping and retrieval at various levels of abstraction. We must start with abstractions, in the form of linking rules.

The Shape Of Meaning

The following tuples are representative of the 85,000 or so *predicate-subject-object* triples provided by (Veale 2014):

(*work_in* scientist lab)
(*conduct* lab experiment)
...
(*serve* priest congregation)
(*perform* congregation worship)

There is an intuitive linkage between *work_in* and *conduct* in the first case, and between *serve* and *perform* in the second: scientists work in labs *that* conduct experiments, and priests serve congregations *that* perform worship. So, our first task is to merge these disparate facts into composite structures:

(*work_in* **scientist** (*some* lab (*that* (*conduct* experiment))))
...
(*serve* **priest** (*some* congregation (*that* (*perform* worship))))

We first codify these intuitions with a pair of abstractions:

(*work_in* $person_X$ (*some* $place_Y$ (*that* (*conduct* $action_Z$))))
...
(*serve* $person_X$ (*some* $place_Y$ (*that* (*perform* $action_Z$))))

These hand-crafted abstractions have the function of rules: they allow pairs of triples that relate a $person_X$, a $place_Y$ and an $action_Z$ to be coalesced into a nested structure. With enough abstractions like these, we can impose new structure on many of the flat, disjoint triples of the original database.

These abstractions have the same shape, yet they cannot be aligned without violating the SMT condition of predicate identicality, as *conduct* $\neq$ *perform* and *work_in* $\neq$ *serve*. However, if we rewrite the original facts as follows, using a smaller, canonical set of predicates, a mapping is possible:

(*by* working (*perform* **scientist**
(*some* work (*for* lab
(*that* (*conduct* experiment))))))
(*by* serving (*perform* **priest**
(*some* service (*for* congregation
(*that* (*conduct* worship))))))

The rewriter maps the original, open-ended set of predicates onto a reduced set of just 12, such as *perform* and *enhance*. The rewritten structures are more general, since they use a small set of predicates, yet they incur no loss of information, as they preserve the distinction between *work_in* and *serve*. Not only are these structures alignable under SMT, they produce a richer mapping, of *scientist:priest*, *lab:congregation*, *experiment:worship*, *working:serving* and *work:service*.

By further abstracting over these canonical variants, using numbers as variables, we obtain this unique generalization:

(*by* 0 (*perform* 1 (*some* 2 (*for* 3 (*that* (*conduct* 4))))))

Following (Veale 2005) we call this a *structural hash*, which serves as a key for *all* structures that produce the same hash. So, to find potential analogues for a given structure, we simply abstract to its hash, and directly look up all other structures in the ATLAS database with the same abstraction key.

Mapping & Retrieval of Analogous Structures

The domain of an idea such as *priest* or *magic* is the set of all structures associated with the term in the ATLAS database. To flesh out an analogy between a source set of structures S and a target T, the systematicity principle of (Gentner 1983; Falkenhainer, Forbus, and Gentner 1989) dictates that the most comprehensive structures are mapped first. As a guide, we define $r(\sigma)$ as a measure of the richness of a structure σ:

$$r(\sigma) = \log_{10} \sum_{i=0}^{n} count_i(\sigma) 10^i$$

The deepest level of nesting in σ is n, and $count_i(\sigma)$ is the number of σ's sub-parts at level i. For our canonical rewriting of the claim that scientists perform work for a lab that conducts experiments, $r(\sigma) = 5.05$. For the facts (*work_in* scientist lab) and (*serve* priest congregation), $r(\sigma) = 0$.

Structural hashing makes the mapping process highly efficient. For each structure σ_S in S, we calculate $hash(\sigma_S)$ and index σ_S under this key. For each structure σ_T in T, we calculate $hash(\sigma_T)$ and use this key to lookup the set of equivalent structures $\{\sigma_S\}$ that were just indexed for S. We use $r(\sigma_T)$ to order our search of the structures in T, so that the first mapping produced, $\sigma_T : \sigma_S$, is the most systematic. As we proceed through other structures in T of successively lower richness, we greedily merge any other mapping that is produced into the cumulative whole if it is coherent to do so.

The retrieval of an idea/domain S to analogically describe a target T proceeds in a similar fashion, and follows the "many are called but few are chosen" principle of (Forbus, Gentner, and Law 1995): using r to rank the structures $\{\sigma_T\}$ of T from most to least rich, we calculate $hash(\sigma_T)$ for each structure in turn, and retrieve all other structures indexed under this key in the ATLAS database. From these we extract a set $\{S_0, S_1, ..., S_n\}$ of candidate sources for T. For a candidate S_i, the quality of the mapping of S_i to T is a function of both the number and the richness of the alignable structures that unite T and S_i. Should we reward sources with many alignable structures of low richness – unconnected facts – over sources with a few rich, or just one very rich, alignment? The quality q of S_i as a source for T is defined over the alignable structures $\{\sigma_0, \sigma_1, ..., \sigma_n\}$ of S_i as follows:

$$q(S_i, T) = \sum_{j=0}^{n} r(\sigma_j)^\beta$$

where β is our non-linear reward for systematicity. So when $\beta = 3$, which it is by default, $q(S_i, T)$ becomes the sum of the cubes of the richness of each structure σ_j in S_i that can be mapped onto an equivalent structure σ_k in T. The choice of β allows us to reward analogies with a single systematic core over those with many disjoint, low-level similarities.

The use of abstraction to rewrite structures in a canonical form, and then turn those forms into generic hashes, allows for the efficient mapping between two sets, or domains, of structures, as well as for the efficient retrieval of analogies. Every structure in every domain is indexed by these abstract hashes, making abstraction central to the organization of the ATLAS knowledge-base and the processes that act upon it.

Abstraction as an Organizing Principle

Abstraction is achieved through a suitable choice of structured representation (Forbus, Liang, and Rabkina 2017). In the case of ATLAS, this is constructed from the content of a flatter, more disjoint and less mapping-rich knowledge-base. To achieve scale, diversity and structural depth, we use over two-thousand abstraction rules to link separate triples in the representation of (Veale 2014) into causal combinations. We have seen a *thru*-rule that links a person in a place to the activity conducted there, but most rules are *cross*-rules that link two facts about the same subject. Some cross-rules link goals to outcomes, to build structures such as the following:

(*if* (*can* (*by* seeking (*seek* **scientist** advance (*as* outcome))))
(*can* (*by* advancing (*enhance* **scientist** sophistication
(*of* science)))))

An alignable structure in a different domain is the following:

(*if* (*can* (*by* seeking (*seek* **priest** proselyte (*as* supporter))))
(*can* (*by* spreading (*enhance* **priest** spread
(*of* doctrine)))))

Other rules integrate (Veale 2011)'s stereotypical properties dataset to add highly salient features to these structures:

(*by* working (*perform* **scientist** work
(*in* (*some* science
(*when* exacting (*lack* flexibility))))))

(*by* serving (*perform* **priest** service
(*in* (*some* church
(*when* dogmatic (*lack* flexibility))))))

By coalescing and cross-combining representations in this way, we can populate ATLAS with structures for over 9,000 concepts, from *science* to *politics* to *magicians* to *war*. For these, ATLAS defines a total of 1.7 million canonical structures which, in turn, are abstracted into 114,000 structural hashes. Canonical rewriting increases the alignability of ATLAS's contents: less than 2% of canonical structures cannot be mapped to, or aligned with, at least one other structure.

As a good analogy may hinge on the mapping of a single systematic structure, Table 1 shows the mean number of candidate sources for a target if we impose a minimum richness (r) for their core alignable structures. So, although the very deepest structures ($r \geq 8.0$) are not commonplace, supporting analogies for just 1 in 8 targets, ATLAS also supports a great many analogies at lower levels of structural richness.

Minimum r value	Mean sources per target
2.0	2396.55
3.0	558.25
4.0	503.21
5.0	218.89
6.0	40.88
7.0	6.74
8.0	0.12

Table 1: The mean number of sources retrieved per target domain with alignable structures of a minimum richness r.

Conclusion: Abstraction Finds a Way

Analogy is both a starting and an end-point in the cycle of inspiration. For instance, the author Michael Crichton based *Jurassic Park* on *Westworld*, a film that he directed in 1973, in which the robotic cowboys of a Western-themed park run amok. To motivate their homicidal rampage, Crichton would invent the analogy of a *computer virus*, more than a decade before hackers would make it a literal reality. Crichton had been inspired to write *Westworld* after a trip to Disneyworld, where the mechanical pirates about which Malcolm would later joke sparked an abiding interest in science gone wrong. With advances in genetics and cloning, Westworld's killer robots would later become the test-tube dinosaurs of *Jurassic Park*, where Malcolm's *Pirates of the Caribbean* retort would reconnect Crichton's analogy with its Disney origins.

Abstraction is key to successful, far-reaching analogies (Gentner and Hoyos 2017; Mitchell 2021), and so we put abstraction at the core of ATLAS's representation of the world. It is abstraction that allows us to relate superficially different experiences so as to perceive a deeper level of similarity. In recent years, advances in Large Language Models (LLMs) have made analogy one of the many "language games" (Wittgenstein 1953) that these models can play with great fluency and aplomb. Since LLMs produce sequences of language tokens, they conflate the tasks of inventing, presenting and talking about an analogy; they do not deal directly with, or produce, explicitly structural representations. Nonetheless, they can go far beyond the simple proportional analogies of the SAT test. For instance, when asked to generate an explanatory metaphor for *quantum mechanics*, GPT-4 (Achiam et al. 2023) responds with a richly systematic (and rather insightful) use of the *theatre* as a source domain. In its mapping, which it elaborates in detail, the stage is classical physics, while the unseen ropes, pulleys and trapdoors behind and below the stage comprise the quantum realm.

When compared to the nimbleness and generative reach of LLMs, even on tasks such as analogical invention, the structured representations of symbolic AI can seem like relics of a bygone era. But even the dinosaurs of *Jurassic Park* could be reinvented for a new age by splicing in some novel DNA. A long history of ideas in structure mapping is packed into ATLAS not to compete against LLMs but to work with them. ATLAS is an open repository[1] with a scale comparable to that of AnalogyKB (Yuan et al. 2024), but no such resource is ever complete enough or robust enough by itself. We thus anticipate two scenarioså in which ATLAS can work with an LLM, to either guide the LLM or to be enriched by it. In the first, relevant ATLAS structures can be retrieved for a target, to drive *few-shot* prompting for analogies in an LLM. These few-shot exemplars can guide an LLM to not just invent and discuss a new analogy, but to formalize a comparable representation for T and a source domain S. These symbolic representations can then be checked for consistency and added, in the second scenario, to ATLAS if they pass logical muster.

Acknowledgments

This paper and resource are wholly the work of the author.

[1]www.kaggle.com/datasets/mtatlas/atlas-analogy-structures

References

Achiam, J.; Adler, S.; Agarwal, S.; Ahmad, L.; and et al. 2023. GPT-4 Technical Report. *arXiv preprint 2303.08774*.

Falkenhainer, B.; Forbus, K. D.; and Gentner, D. 1989. Structure-mapping engine: Algorithm and examples. *Artificial Intelligence* 41(1):1–63.

Forbus, K. D.; Gentner, D.; and Law, K. 1995. Mac/fac: A model of similarity-based retrieval. *Cognitive science* 19(2):141–205.

Forbus, K. D.; Liang, C.; and Rabkina, I. 2017. Representation and computation in cognitive models. *Topics in cognitive science* 9(3):694–718.

Gentner, D., and Hoyos, C. 2017. Analogy and Abstraction. *Topics in Cognitive Science* 9(3):672–693.

Gentner, D., and Toupin, C. 1986. Systematicity and surface similarity in the development of analogy. *Cognitive Science* 10(3):277–300.

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

Hofstadter, D. R., and Mitchell, M. 1995. *Fluid Concepts and Creative Analogies. Computer Models of the Fundamental Mechanisms of Thought*. New York: Basic Books. chapter The copycat project: A model of mental fluidity and analogy-making.

Holyoak, K. J., and Thagard, P. 1989. Analogical mapping by constraint satisfaction. *Cognitive Science* 13(3):295–355.

Mitchell, M. 2021. Abstraction and analogy-making in artificial intelligence. *Annals of the New York Academy of Sciences* 1505(1):79–101.

Speer, R.; Chin, J.; and Havasi, C. 2017. Conceptnet 5.5: An open multilingual graph of general knowledge. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 4444–51. AAAI Press.

Turney, P., and Littman, M. L. 2005. Corpus-based learning of analogies and semantic relations. *Machine Learning Journal* 60(1-3):251–278.

Veale, T., and Keane, M. T. 1997. The Competence of Sub-Optimal Structure Mapping on 'Hard' Analogies. In *Proceedings of IJCAI'97, the 15th International Joint Conference on Artificial Intelligence*, 21–26. Morgan Kaufmann.

Veale, T. 2004. WordNet sits the S.A.T.: A Knowledge-based Approach to Lexical Analogy. In *Proceedings of the 16th European Conference on Artificial Intelligence*, ECAI'2004. IOS Press.

Veale, T. 2005. Analogy Generation with HowNet. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, IJCAI'05, 1148–53. Lawrence Erlbaum Associates.

Veale, T. 2011. Creative language retrieval: A robust hybrid of information retrieval and linguistic creativity. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL'2011)*.

Veale, T. 2014. Coming good and breaking bad: Generating transformative character arcs for use in compelling stories. In *Proceedings of ICCC'14, the 5th International Conference on Computational Creativity*.

Veale, T. 2015. Fighting words and antagonistic worlds. In *Proceedings of the Third Workshop on Metaphor in NLP, held as part of NAACL HLT 2015, Denver, Colorado*, 87–94.

Vrandecčíc, D., and Krötzsch, M. 2014. Wiki-data: A free collaborative knowledgebase. *Communications of the ACM* 57(10):78–85.

Wittgenstein, L. 1953. *Philosophical Investigations (Philosophische Untersuchungen)*. London, UK: Blackwell Publishing. Translated by G.E.M. Anscombe.

Yuan, S.; Chen, J.; Sun, C.; Liang, J.; Xiao, Y.; and Yang, D. 2024. AnalogyKB: Unlocking analogical reasoning of language models with a million-scale knowledge base. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Vol 1)*, 1249–65.