The Imagination Machine V: On Abstraction and Analogy

Mark Tracy

1 Overview

Analogy is the bedrock of communication. Even that sentence makes use of analogy: as bedrock underlies and supports structures, so too does analogy underlie and support communication, allowing us to coordinate activity and manipulate our environment. Analogy allows a reasoner to transfer previously learned structure to a new situation, generating hypotheses and thereby facilitating new understanding. So fundamental is analogy to language that it proves challenging to articulate the abstract structure of analogy and to codify valid analogical reasoning. Nonetheless, it remains a fundamental endeavor for any interested in understanding mentation. In the foregoing, I introduce and augment one popular model of analogy, and I utilize the formalism thus achieved to attempt a definition of valid analogical reasoning.

2 Classical Theories of Analogy

A domain may be defined as a tuple:[1]

$$D = (O, A, R, S, T)$$

- O = set of objects
- A = set of attributes (unary operators: $a \in A \implies a : O \to S$)
- R = set of relations (n-ary operators: $r \in R \implies \exists\, n \in \mathbb{N}, \quad r : O^n \to S$)
- S = set of statements
- T = set of statements believed to be true (belief set)

Note that attributes are a special case of relations: each $a \in A$ is simply a unary relation, so formally $A \subseteq R$.

[1] This definition follows the standard treatment of domains in analogy and relational reasoning literature (cf. 1), but extends it to include a set of statements S and a belief set T, corresponding respectively to the expressible and the held-to-be-true propositions within the domain.

2.1 Structure-Mapping Theory of Analogy

In the landmark paper "Structure-Mapping: A Theoretical Framework for Analogy," Gentner argues that an analogy is a mapping between objects in a base domain and objects in a target domain that does not necessarily carry over object-level attributes but which carries over some relational predicates.[1]

2.2 A Formal Definition of Analogy

An analogy between a source domain $D_s = (O_s, A_s, R_s, S_s, T_s)$ and a target domain $D_t = (O_t, A_t, R_t, S_t, T_t)$ is defined by a tuple:

$$A = (X, Y, M, P)$$

- $X \subset O_s$: a collection of objects in the source domain
- $Y \subset O_t$: a collection of objects in the target domain
- $M : X \to Y$: a mapping of objects from source to target domain
- $P \subset \{r \mid r \in R_s \cap R_t \text{ and } \exists \mathbf{x} \in X^k \text{ for some } k \in \mathbb{N} \text{ such that } r(\mathbf{x}) \in T_s \text{ and } r(M(\mathbf{x})) \in T_t\}$: a set of relations that are present in the source and target domains, are true of some tuple of objects in the source domain, and are preserved in the target domain via the mapping M. As a notational convention, we consider $M(\mathbf{x})$ to be the component-wise application of the mapping M to the tuple $\mathbf{x}$, i.e. $\mathbf{x} = (x_1, \ldots, x_n) \implies M(\mathbf{x}) = (M(x_1), \ldots, M(x_n))$.

2.3 Analogical Reasoning

Let D_s be a source domain and D_t a target domain. Suppose:

- $X_1 \subset O_s$ is a subset of objects in the source domain. Let $|X_1| = n$.
- $Y_1 \subset O_t$ is a subset of objects in the target domain.
- $M : X_1 \to Y_1$ is a mapping of the source domain subset to the target domain subset.
- P is a set of relations preserved by the mapping M.

This establishes an analogy between D_s and D_t. Now suppose that some further fact (of a particular form to be specified below) holds in the source domain; we formally define an **analogical reasoning step** to be the positing of a corresponding form of further fact in the target domain. Formally:

Suppose there exists a superset of X_1 called X_0:

$$\begin{aligned} X_1 &\subseteq X_0 \\ |X_0| &= m \geq n \end{aligned}$$

and suppose that

$$r(\mathbf{x}^*) \in T_s$$

for some tuple $\mathbf{x}^* \in X_0^k$ for some $k \in \mathbb{N}$ and for some relational predicate $r \in (R_s \cap R_t)$.

Then an analogical reasoning step is to hypothesize that there exists a mapping M' that preserves and extends the original analogical mapping M and preserves the further observed relation in the source domain, r. In particular, the hypothesis is as follows:

$$\begin{aligned} \exists\, Y_2 \subset O_t \quad & \text{and} \\ \exists\, M' : X_0 &\rightarrow Y_1 \cup Y_2 \quad \text{such that} \\ M'(x) &= M(x) \quad \forall x \in X_1 \quad \text{and} \\ r(M'(\mathbf{x}^*)) &\in T_t, \end{aligned}$$

where $M'(\mathbf{x}^*)$ is the component-wise application of the mapping M' to the tuple $\mathbf{x}^*$ identified above.

This formulation captures the logic of projecting relational structures from the source domain into the target domain, conditioned on preserved analogical structure. It highlights how analogy can support hypothesizing about unseen objects, roles, or relations in the target domain by structurally mapping known relations in the source.

2.4 Analogy as Mediated by Abstraction

Abstraction, in the broadest sense, refers to the process or result of mapping a collection of objects, attributes, or relations to a single representation, typically to retain only information which is relevant for a particular purpose.

There is a connection between abstraction and analogy that is insufficiently explored in Genter's 1983 paper. If, as Gentner convincingly argues, an analogy is a mapping between objects in a base domain and objects in a target domain that does not necessarily carry over object-level attributes but which carries over some relational predicates [1], then for any analogy there exists an abstract domain that implicitly mediates the analogy. In particular, the domain that mediates an analogy $A = (X, Y, M, P)$ between a source domain $D_s = (O_s, A_s, R_s, S_s, T_s)$ and a target domain $D_t = (O_t, A_t, R_t, S_t, T_t)$ consists of:

- **A new set of objects, $O_{\mathbf{abs}}$:**
  - Call them symbols.
  - $\forall x \in X, (x, M(x)) \in O_{\text{abs}}$.
  - Notational convention: for a $k-$tuple of objects in the source domain, $\mathbf{x} \in X^k$, we denote the corresponding tuple of symbols as $(\mathbf{x}, M(\mathbf{x})) \in O_{\text{abs}}^k$, where $M(\mathbf{x})$ is the component-wise application of M to $\mathbf{x}$. In particular:

$$\begin{aligned} \mathbf{x} &= (x_1, \ldots, x_k) \implies \\ M(\mathbf{x}) &= (M(x_1), \ldots, M(x_k)) \text{ and} \\ (\mathbf{x}, M(\mathbf{x})) &= ((x_1, M(x_1)), \ldots, (x_k, M(x_k))) \end{aligned}$$

- **A set of predicate attributes, $A_{\mathbf{abs}}$:**
  - $A_{\text{abs}} = P \cap A_s$
  - The set of unary relations preserved by the analogy, if any.
- **A set of predicate relations, $R_{\mathbf{abs}}$:**
  - Call them abstract relations.
  - $r \in P \iff r \in R_{\text{abs}}$
  - $r(\mathbf{x}) \in T_s$ for some $\mathbf{x} \in X^k$ with $k \in \mathbb{N} \implies r((\mathbf{x}, M(\mathbf{x}))) \in T_{\text{abs}}$.
- **A statement set, $S_{\mathbf{abs}}$:**
  - All possible combinations from the collections of objects, attributes, and relations specified above.
- **A belief set, $T_{\mathbf{abs}}$**
  - A subset of S_{abs}, populated as specified above.

2.4.1 An example

Take the analogy, "An atomic nucleus is like the solar system." [1] At an earlier point in scientific history, the analogical mapping may have looked like this:

$$M : X \to Y$$

$$\text{NUCLEUS} \mapsto \text{SUN}$$

$$\text{ELECTRON} \mapsto \text{PLANET}$$

And the relationships preserved include:

$$\{\text{ORBITS}, \text{IS\_MOVING}\} \subset P.$$

Now, in recognizing a mediating abstract domain we may synthesize new symbols with carried-over attributes and abstract relations, thereby forming a mediating abstract domain that both source and target instantiate:

$$\begin{aligned}
\{\text{NUCLEUS}, \text{SUN}\} &\mapsto \text{CENTRAL_BODY} \\
\{\text{ELECTRON}, \text{PLANET}\} &\mapsto \text{SATELLITE} \\
\text{ORBITS} &\in R_{\text{abs}} \\
\text{IS_MOVING} &\in A_{\text{abs}} \subset R_{\text{abs}}
\end{aligned}$$

Now, obviously each instance of SATELLITE and of CENTRAL_BODY in the two original domains has attributes (mass, charge, etc.) whose values determine how the abstract relation

$$\text{ORBITS}(\text{SATELLITE}, \text{CENTRAL\_BODY})$$

manifests in these two distinct domains. Note that the statement IS_MOVING(SATELLITE) $\in S_{\text{abs}}$ happens to carry over into the belief set of this abstract domain, T_{abs}, since relative motion is characteristic of a classical satellite in both original domains.

Analogy is not simply recognizing, "D_s is like D_t". Instead, analogy is mediated by abstraction: it is to say, "D_s is like D_t because there exists an abstract domain D_{abs} of which both D_s and D_t are instances." Or, in other words, to recognize an analogy is to say, "This pattern of relations in D_s is like that pattern of relations in D_t—and there's a higher-order domain D_{abs} that generalizes both."

References

[1] Dedre Gentner. Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2):155–170, 1983. doi: 10.1207/s15516709cog0702_3.