1. **Justification of analogical reasoning**
   - an argument that it is reasonable to give *serious consideration* to hypotheses supported by analogies that meet *certain standards*

   *Serious consideration:* non-negligible prior probability (or better)

   *Certain standards:* what standards?

2. **Representation**
   - Last time: schema for all analogical arguments
   - Desideratum: a justifiable rule for analogical reasoning

   Comparison: simplicity of the “straight rule” for enumerative induction

   In the absence of reasons for thinking otherwise, if out of $n$ observed A, $m$ have been B, then infer that the next A will be B with probability $m/n$ (or infer that the proportion of A that are B in the general population is $m/n$).

   No comparable rule for analogical reasoning.
Two conceptions of analogical reasoning

- Each conception involves different representations, different criteria of goodness, and a different justification.

I. Structuralist/deductive:

   Analogy relation is, ideally, an isomorphism of two domains (laws stated in same mathematical form).

   We judge an analogical argument by considering how closely the analogy approximates isomorphism.

   **Examples:** fluid models // metal stress
               Electric circuits // fluid flow in systems of pipes

II. Sampling/inductive:

   The analogy relation is a conjunction of material resemblances between domains

   We judge an analogical argument by comparing the extent of material similarity and difference (a ‘competition’: Mill)

   **Examples:** life on other planets, cultural origin of artifacts

Common procedural assumption:

Common to both conceptions is the idea that what matters is the overall similarity of two domains

Procedure for evaluating analogical arguments:

   First assess the overall similarity of two domains (the analogy relation).
   Use this to evaluate whether the argument succeeds.
I. Structuralist conception

A. Nomic isomorphism (Hempel)

The laws governing the two domains have the same mathematical form.

\[
\begin{align*}
\text{Ohm’s Law} & \quad \Delta v = iR \\
\text{Poiseulle’s Law} & \quad \Delta p = Vc
\end{align*}
\]

Results obtained for one system can be transferred to the other (e.g., circuit model for a system of pipes).

Assessment:

- The assumption of isomorphism transforms analogical inference into unproblematic deductive inference.
- But suitable only given very detailed background knowledge; cannot account for most analogical inferences.
- Isomorphism is harder to justify than just the conclusion of an analogical argument.
B. Determinations (Russell, Weitzenfeld)

- Mill, section 1
  
  If two cases are shown to resemble in all material circumstances, the argument “cannot be resisted”.

- Russell, section 1
  
  **Determination rule**: the value of a variable (or set of variables) $Q$ is completely determined by the values of the variables $P_1, \ldots, P_m$ (abbreviated as $P$).

  **Example**: the value of a car is determined by its make, model, year, mileage, and condition.

  Use in analogical arguments:

<table>
<thead>
<tr>
<th>A’s car</th>
<th>B’s car</th>
</tr>
</thead>
<tbody>
<tr>
<td>’88 Corolla</td>
<td>’88 Corolla</td>
</tr>
<tr>
<td>100,000 km</td>
<td>100,000 km</td>
</tr>
<tr>
<td>Fair condition</td>
<td>Fair condition</td>
</tr>
<tr>
<td>sold for $5,000$</td>
<td>value is $5,000$</td>
</tr>
</tbody>
</table>

  **General form**:

  $$ P(x, y) \rightarrow Q(x, z) $$

  $$ \forall wxyz \ P(w, y) \land P(x, y) \land Q(w, z) \Rightarrow Q(x, z) $$

  If S and T agree on y, and S has z, then T has z.

  **Analysis**:

  The assumption of a determination transforms analogical argument into a valid deductive inference.

  The use of source information is indispensable. (The determination rule is weaker than a generalization.)

  The proposal helps to bring out background knowledge often employed in analogical reasoning

  **Problems**:

  To know a determination rule is to have a complete list of which attributes are relevant. This is too restrictive to apply to most analogical arguments (e.g., animal testing).

  A determination rule is harder to justify than the original analogical inference.
Analogical arguments are enthymematic arguments (Weitzenfeld)

a) Similarity alone is inadequate to justify any conclusion.

   Argument: Multiple sources can lead to conflicting conclusions.
   (Cf: legal reasoning)

b) Isomorphic structures (or determining structures) support valid analogical inferences.

   Ex: the game 15 and tic-tac-toe are formally isomorphic
   Ex: real estate values

   Determining structures are best defined as for Russell:
   \{a_1, \ldots, a_n, a_t\} constitutes a determining structure if the value of a_t is
determined by the values of a_1 through a_n.

c) Strategy for analogical reasoning:

   i) Establish the existence of an isomorphism/determining structure.
      - enumeration, surveillance, inference

   ii) Find a source analogue that differs from the target only on variables whose
effect can be calculated (using the determination rules).

Assessment:

Finding a determining structure transforms an analogical argument into a valid argument.
This does represent the final stage in many analogical arguments.

Problems:

1) Existence of determining structures must established by informal plausibility
   arguments (148).
   So what is gained by replacing the original analogical argument with a deductive
   argument containing a dubious premise that, in any case, has to be supported
   with plausibility arguments?

2) Not generally suitable for analogical arguments used in the early part of an inquiry.
II. Sampling/inductive conception

- Three important ways to represent (and justify) an analogical inference as in inductive argument

a) Single-case generalization

Aristotle’s *paradeigma*:

Enthymemes based upon example are those which proceed from one or more similar cases, arrive at a general proposition, and then argue deductively to a particular inference. *Rhetoric* (1402b15)

Schematically:

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>((x)(P(x) \supset Q(x)))</td>
<td>(P(S), Q(S))</td>
</tr>
</tbody>
</table>

Assessment:

The problem of analogy becomes the problem of single-case generalization (since the final deduction is unproblematic)

Problems:

The argument is weaker than just passing to the next case

Generally, only applicable to arguments with multiple sources, but then the argument becomes a straight induction rather than an argument by analogy

- purest problem of analogy: single source

Single-case generalization is sometimes possible in well-understood domains.

Ex: Testing one sample of a chemical compound is enough to determine boiling point, solubility, etc. of all samples. Here, P is a natural kind.

But analogical inferences are typically made where we lack this rich background knowledge.

Deeper issue: this analysis does nothing to help us distinguish between relevant and irrelevant similarities, or between good and bad analogical arguments
Agassi on analogies and generalization

Ref: “Analogy Hard and Soft”, in D. Helman, Analogical Reasoning

1. All analogies are either hard or soft.
   
   Hard = mathematical isomorphism
   
   Soft = similarity that falls short of isomorphism

2. All soft analogical arguments are either ad hoc or redundant.
   (Hard analogical arguments are deductively valid, and unproblematic.)

Argument for thesis 2:

2a. All analogical arguments must be analyzed as single-case generalization (followed by a deductive inference).

2b. Such generalizations are either already accepted on the basis of independent evidence (in which case the analogy is redundant) or based on a single case (and thus ad hoc).

Argument for 2a: (ironically, the argument is a generalization from a single example!)

<table>
<thead>
<tr>
<th>Ex:</th>
<th>Known</th>
<th>Conjectured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tom</td>
<td>English, Tall</td>
<td></td>
</tr>
<tr>
<td>Dick</td>
<td>English</td>
<td>Tall</td>
</tr>
<tr>
<td>Harry</td>
<td>English</td>
<td></td>
</tr>
</tbody>
</table>

- If the inference “Dick is tall” is warranted by analogy, then so is “Harry is tall” (or any other Englishman), i.e., so is the generalization.

- If not, we must distinguish between Dick and Harry: cite some property R (“drank milk as a child”) common to Tom and Dick, but not Harry. But then the argument is just a narrower generalization, “All Englishmen who drank milk as children are tall.”

- Continuing in this way, we eventually exhaust the characteristics of Tom and Dick and reach a generalization that we must accept, or make an arbitrary discrimination between Dick and Harry.
Assessment of Agassi:

i) Formal fallacy.

A warrant for each instance may not be a warrant for the generalization. (Each instance may have positive probability, but the generalization might have 0 probability.)

Ex: random number generator; fair die.

ii) Supporting generalizations.

A supporting generalization is a generalization which, if true, would guarantee a successful analogical argument.

Supporting: \[(x)(P(x) \supset Q(x)) \land (x)((P(x) \cdot C(x)) \supset Q(x)); C \text{ is common to } S \text{ and } T \text{ (but unknown)}\]

Non-supporting: \[(x)((P(x) \cdot A(x)) \supset Q(x)); \text{ we know } A(S), \neg A(T)\]

We need not commit ourselves to any specific generalization when we reason by analogy. At most, we need to suppose there is some supporting generalization, or a positive probability that some supporting generalization is true.

Objection: What justification for assuming a positive probability for any supporting generalization?

Quick answer: with no compelling evidence that the negative analogy is essential to the conclusion Q, to suppose that the relevant law must be a non-supporting generalization involves raising this probability (that the relevant generalization is non-supporting) to 1. That is a very non-conservative thing to do. So conservatism requires open-mindedness, i.e., giving some of the probability to a supporting generalization.

iii) Expected utility

Objection: even if existence of a supporting generalization has a small positive probability \(p\), the decision to adopt a hypothesis supported by analogy is always unwise because of the bad consequences of adopting a false hypothesis.

<table>
<thead>
<tr>
<th>Supporting generalization (p)</th>
<th>None (1 - p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adopt the hypothesis</td>
<td>benefit (a)</td>
</tr>
<tr>
<td>Don’t adopt</td>
<td>0</td>
</tr>
</tbody>
</table>
Agassi’s argument would require $pa + (1-p)b < 0$, or $p/1-p < -b/a$. But in context of considering working hypotheses, $b$ may well be rather small, and $a$ rather large.
b) Property sampling

i. Representation

**Analogical inference:**

\[
P(S) \quad P^*(T) \quad (n \text{ shared attributes } P_1, \ldots, P_n)
\]

\[
A(S) \quad \neg A^*(T) \quad (r \text{ differences})
\]

\[
\neg B(S) \quad B^*(T) \quad (s \text{ differences})
\]

\[
Q(S) \quad \text{prob}(Q^*(T)) = p
\]

where \( p > 0 \), \( p \) increases with larger \( n \) and decreases with larger \( r \) and \( s \).

Mill, section 2:

“Every resemblance proved between B and A, in any point not known to be immaterial with respect to \( m \), forms some additional reason for presuming that B has the attribute \( m \)... [and]... every dissimilarity which can be proved between them furnishes a counter-probability of the same nature on the other side.” (366)

“There will be, therefore, a competition between the known points of agreement and the known points of difference in A and B; and according as the one or the other may be deemed to preponderate, the probability derived from analogy will be for or against B’s having the property \( m \).” (366)

Important constraints:

1. No connection is known to exist between \( m \) and the points of agreement (i.e., between P and Q), but neither are they known to be unconnected.

2. No property known to be indispensable to \( m \) belongs to the negative analogy, A.
ii. Philosophical basis for the inference

- Mill, section 2:
  
  “The supposition must be that \( m \) is an effect really dependent on some property of A, but we know not on which.”

  Derivative and ultimate properties: “every resemblance affords grounds for thinking that A and B share more ultimate properties” (and hence, possibly, \( m \)).

  Proportionality: if A and B agree in nine out of ten known properties, then “we may conclude with a probability of nine to one that [B] will possess any derivative property of A”.

Qu: What assumptions about probability here? The known properties of S are taken to be a random sample of all S’s properties — random, that is, with respect to the attribute of also belonging to T.

Objections:

1. How do we ‘count’ properties?
2. What reason to believe we have a random sample?
   - drawings from an urn vs. presenting an analogical argument
3. The connection between derivative and ultimate properties needs clarification.
4. Why should all similarities and differences – the overall similarity relation – be used to assess the analogical argument?

- Russell

  Goal: clear postulation of what underlies the similarity-based reasoning

  Two possible justifications for “general-purpose” similarity reasoning:

  a) naturalistic/empirical approach:

  - large resemblance justifies analogical inference; justification is empirical
  - explanation: humans have evolved useful similarity metrics

Russell’s objection: does not justify the metrics used in AI programs

My Objections:

i) no reason to think those metrics would automatically be useful for, say, advanced mathematics or science

ii) objection to modularized approach to similarity (developed below)
b) logical approach:

Consider the case where you don’t know the relevant determination(s), and you don’t know which of the known facts might be relevant to Q.

**Idea:** under ignorance, maximally similar source is most likely to match the target on the relevant facts (even though they are unknown).

Suppose we have sources and a target described by setting values to attributes $P_1, \ldots, P_m$ and we know the value of Q for the sources, but not the target.

Two assumptions:

(A) All features relevant to the value of Q appear among $P_1, \ldots, P_m$.

(B) Given no information about which of these features are relevant, we may assume that all are equally likely.

**Argument in simplest case:**

1. Suppose S and T match on the values of $s$ of the attributes $P_1, \ldots, P_m$.
2. Suppose that $r$ of the attributes $P_1, \ldots, P_m$ are actually relevant to Q.

Let $p(s, r) = \text{probability that S matches T on those } r \text{ relevant attributes.}$ Then by (B),

$$p(s, r) = p(s, r) = \binom{s}{r} \binom{m}{r}.$$

The higher the number of similarities, the greater the chance of a total match on the relevant attributes. So $p(s, r)$ is also the probability that the value of Q is the same in S and T.

More complex: allow $r$ to vary according to a probability distribution $q(r)$. Can still compute the probability of success by a weighted sum of $p(s, r)$.

Most complex: don’t assume that each property is equally likely to be one of the relevant ones. But need to assume that these properties derive from ultimate properties that are equally likely to be relevant; and we’d need to justify that assumption.

**Assessment:**
- no real advance in either the counting problem or the randomness assumption problem
- no clear reason to think that overall similarity is the relevant basis for assessment.

c) Keynes

Main failure of sampling argument: does not come to grips with relevance

Keynes’ construction is useful because it reveals where the argument goes wrong, and points the way to a better approach.

Construction:

1) Generator properties

a) Independent: any two generator properties are causally and probabilistically independent.

\[ \text{Prob}(\phi_1|\phi_2) = \text{Prob}(\phi_1) \]

b) Comprehensive: the generator properties produce all other (derivative) properties.
(Each derivative property is associated with a necessary and sufficient group of generators, or with a disjunction of such groups.)

Each group is associated with a set of derivative properties; each derivative property is associated with one or more groups.

Ex: Three generators, \( \phi_1, \phi_2, \phi_3 \). \( \phi_1 \) and \( \phi_2 \) sufficient to generate \( P \), with or without \( \phi_3 \). Then the groups associated with \( P \) are \{\( \phi_1, \phi_2 \}\} and \{\( \phi_1, \phi_2, \phi_3 \}\}.

\( n \) generators \( \rightarrow 2^n \) groups

infinitely many generators \( \rightarrow \) infinitely many groups

**Independent variety** of a system = \# of groups

2) Assumptions

i) Positiveness: each object has a positive probability of having each generator property.

ii) Limitation of Independent Variety:

The number of groups is finite (version 1); equivalently, the number of generators is finite.

From i) and ii), every generalization has positive probability of holding for each group (and hence overall), and each confirming instance increases the probability. So these assumptions justify enumerative induction.
Fails for analogical reasoning: the probability of the generalization is raised by showing that a group satisfies it, but we know that T and S belong to different groups (because of the negative analogy).

[If the groups are $G_1, \ldots, G_k$, with probability $p_k$ that the generalization $(x)(P(x) \supset Q(x))$ holds for $G_k$, then the probability of the whole generalization is $p_1 \cdots p_k$. Finding that it holds for $G_k$ raises this probability by $1/p_k$. But the probability that it holds for one of the other groups $G_i$ remains unchanged at $p_i$, by the assumption of independence.]

This exposes the real weakness of the Sampling Argument:

It could be that the similarities P between S and T derive from $\phi_1 \ldots \phi_m$, and Q derives from different generators $\Psi_1 \ldots \Psi_n$, and the groups for S and T overlap only in $\phi_1 \ldots \phi_m$. In such a case, the similarities would all be irrelevant to Q.

*Indeed, we should reject a justification that justifies every analogical argument that fits our schema. Analogies made without any knowledge of co-occurrence have no persuasive power.

Analogical arguments can be justified within a Keynesian framework if we suppose:

(1) There is evidence of some overlap between the groups that generate P and Q.

(2) There is no compelling evidence that the intersection of the groups that generate P and Q is sufficient to generate A (the negative analogy would then be essential to produce Q).

(3) There is no compelling evidence that the intersection of the groups that generate P and Q is inconsistent with B (the negative analogy would rule out Q in the target).
III. Rejecting the two models

1. Extreme cases

Both models work occasionally (but note the lack of examples in these articles)

Isomorphism or determining relations: presuppose in-depth knowledge of the relevance relations

Sampling argument: presuppose total ignorance of the relevance relations (cf. drug testing on a corpse)

In a typical analogical argument we have *some knowledge* of the relevance relations, but not of any determining relations. Is it not better to seek a theory that models this situation, and perhaps handles the others as limiting cases?

2. Approach to similarity

Overall similarity is neither meaningful nor useful, since only similarities viewed as relevant to the conclusion of interest are important

Focus of attention: the connection between the known similarities (P) and the projected similarity (Q) in the source domain, and whether a similar connection could hold in the target.

Similarity should always be assessed *relative to the particular analogical argument*, not prior to such arguments (as with the sampling approach)

Relevance figures appropriately in neither the deductive nor the inductive models

3. Outline of an alternative approach

The connection to generalization, noted by many authors, is indeed the key.

Poincare: prelude to generalization.

A good analogical argument must satisfy two requirements:

a) *Prior association.* There must be a definite connection between the known positive analogy P(S) and the hypothesized analogy Q(S). The connection must be explicitly represented in the analogical argument. (Note: it has nothing to do with the analogy per se, since it pertains to the vertical relations within the source domain.)

b) *Potential for generalization.* It must be possible for the prior association to be generalized to cover both source and target domains. The known similarities must provide evidence for, and the known differences must not block, such a generalization.