Mathematicians’ Perceptions of “Generalisations” in Mathematics

Ned Wontner (ILLC, UvA)

11th July 2021
Mathematical Cultures and Practice XI, St. Andrews
Why care?

We should care as generalisations occur all over mathematics:

- **Theories**: finite arithmetic to cardinal arithmetic
- **Objects**: $\mathbb{N}$ to cardinals, $\mathbb{R}$ to $\mathbb{R}_\kappa$, analytic to arbitrary functions

But! What do we know about them?
Plan for the Talk

**Data:** two small studies on mathematicians’ use of generalisation

1. corpus study
2. interviews

**Theory:** the start of a theory of generalisations. End goal: an account of

1. distinctive features
2. metrics for evaluation
3. internal structure
1 Corpus Study

2 Interviews

3 Start of a theory of generalisations
**Point of the Study**

**Broad question:** how is “generali*” used in print?

**Hypothesis:** papers often use “generali*” to contextualise

For each paper, we:

- **recorded the number of ‘hits’ of the text string** `generali*`. This picks up word forms containing this string, e.g. `generalised`, etc.
- **calculated a “processed” value**, $P$, to remove junk (e.g. hits in page titles, acknowledgements)

<table>
<thead>
<tr>
<th></th>
<th>Processed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>840</td>
</tr>
<tr>
<td>Mean</td>
<td>1.79</td>
</tr>
<tr>
<td>Median</td>
<td>1</td>
</tr>
<tr>
<td>Mode</td>
<td>0</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>4.48</td>
</tr>
</tbody>
</table>

**Table**: Hits-per-paper statistics
Evidence that “generalisation” is used to contextualise

First, exclude 0s

**Broad evidence:** most papers have only 1 or 2 counts (low modal $P$)

**Specific evidence:** (linear regression test) 4.4% chance of no negative linear relationship between page length and hits per page

- **Conclusion:** there *is* a negative linear relationship, i.e. longer paper have fewer hits-per-page
Limitations and Solutions

- **Limitation 1**: evidence for special rôle of generalisation, not necessarily contextualisation

- **Partial Solution**: new text analysis methods: word embedding model, topic modelling, coincidental nouns

- **Limitation 2**: uses ‘self-reporting’ - i.e. some intuitive examples of generalisations not so-called, possible inclusion of non-examples
Corpus Study

Interviews

Start of a theory of generalisations
Interviews

- Exploratory ethnographic study, professionals from various subfields - no logicians.
- They were asked to list examples of good, bad, fruitful, successful generalisations.
- Suggest subfields are systematic similar/homogeneous in use of “generalisation” & measures of success.
- Responses fit with pre-experiment expectations.
- Limitation: small sample size, not representative in sociological sense.
Corpus Study

Interviews

Start of a theory of generalisations
Other meta-theoretic processes. Argue:

1. Distinct extension from gen’n
2. Distinct intension from gen’n

Examples:

1. Easy: Scientific Gen’n - gen’n with weak background assumptions?
2. Hard: Abstraction - gen’n in class with constant level of abstraction? E.g. geometrical/analytic to arbitrary functions
3. also: domain expansion, concept expansion, idealisation, etc.
Classify generalisations in (at least) 2 dimensions:

1. **Ontology**: objects, theorems, proofs, structures, concepts, ...
   - ‘everyday’ ontology of mathematics?

2. **Exactness**: exact vs. analogous

Not just syntax! (Contra, e.g., Propp)

1. formal or natural language syntax? Canonical form and similarity relation **underdetermined**
2. No “recipe” for analogous generalisation - **unpredictable analogies**
3. What is *the* defining property? **Contingent properties** and structure *relativity*
4. Thoughts welcome!
Evaluation of Generalisations

1. (fruitful) generalisation is limited: no ‘generalisation without end’ - contra Kronecker, Chasles, Leibniz

2. Generalised proofs need substantially new ideas for success

3. Kitcher (Cournot, Mancosu?): generalisations are fruitful if explanatory
   - Not full story?
   - More abstract mathematics not explanatory, e.g. Cardinal arithmetic does not explain finite arithmetic (though it may unify)
   - Object generalisation - objects not explanatory alone!
   - Thoughts welcome!
Thank you!
References


We recorded 7 values for each of the paper.

A length of the paper,
B frequency of generali* anywhere in the text,
C frequency of without loss of generality anywhere in the text,
D frequency of generali* in the title,
E frequency of generali* in the references,
F frequency of generali* in page headers,
G frequency of generali* in the acknowledgements.
Figure: Graph: Length in pages (len) against Processed value ($P$)
Non-integer results are rounded to two decimal places.

<table>
<thead>
<tr>
<th></th>
<th>length (A)</th>
<th>generali (B)</th>
<th>wlog (C)</th>
<th>ref (D)</th>
<th>title (E)</th>
<th>header (F)</th>
<th>acknw (G)</th>
<th>processed (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>5185</td>
<td>1208</td>
<td>198</td>
<td>149</td>
<td>5</td>
<td>19</td>
<td>2</td>
<td>840</td>
</tr>
<tr>
<td>Mean</td>
<td>11.08</td>
<td>2.58</td>
<td>1.52</td>
<td>1.35</td>
<td>0.45</td>
<td>1.90</td>
<td>1.79</td>
<td>1.79</td>
</tr>
<tr>
<td>Median</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mode</td>
<td>11</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>SD</td>
<td>3.44</td>
<td>4.90</td>
<td>1.03</td>
<td>0.77</td>
<td>0.52</td>
<td>2.77</td>
<td>0.00</td>
<td>4.48</td>
</tr>
</tbody>
</table>

**Table:** total, mean, median, mode, and standard deviation of the data

Overall B/C ratio
Average of the individual B/C ratios

**Table:** general/without loss of generality ratios
Not in Talk: Corpus Study Observations

On data:

- More “generali*” occurred than initially speculated
- Fewer “without loss of generality” than initially speculated
- Outliers, $P > 10$, involved names, e.g. “Generalized torsion and decomposition of 3–manifolds”

Further limitation:

1. **Limitation 1**: no data on where generali* occurs in papers
2. **Solution**: richer corpus data
3. **Limitation 2**: published papers not representative of practice
4. **Partial Solution**: Informal corpora: “the great unread” e.g. FOM, Mathoverflow?
   - FOM and HO/GM/LO in Mathoverflow are still not representative: somehow about maths not within maths
5. **Limitation 3** We miss variant phrases without generali*, e.g. “$X$ is a more general version of $Y$” and “$X$, or more generally $Y$”.
No mention of ‘false’ generalisation

The responses suggest “generalisation” is a kind in mathematics, not a façon de parler
- Concept expansion - only similar to **object gen’n**?
- Idealisation - some idealisations **add information**
- Domain expansion - **distinct domain and results** in gen’n of analysis to higher reals
**Maths education view** (Ellis, Strachota) of mental processes only about *methods* - so only suitable for *proof* gen’n?