MathWebSearch is a content-based search engine that focuses on fast query answering for interactive applications. It is currently restricted to exact formula search via unification queries, i.e. no similarity search and no full-text search.

**Formula Search with Named Wildcards on ZBMath**

Approximate $\int_{[a,b]}|f(x)|g(x)|dx \leq r$ from above? \(\Rightarrow\) Ask MathWebSearch!

It finds Hölder’s inequality with universal variables in the index

$$\int_D |f(x)|g(x)||dx \leq (\int_D |f(x)|^p dx)^{1/p} (\int_D |g(x)|^q dx)^{1/q}$$

with substitution $x \mapsto t$, $f \mapsto \sin$, $g \mapsto \cos$, $D \mapsto \mathbb{R}^2$ \(\Rightarrow\) Solution:

$$\int_{\mathbb{R}^2} |\sin(t)\cos(t)|dt \leq (\int_{\mathbb{R}^2} |\sin(t)|^p dt)^{1/p} (\int_{\mathbb{R}^2} |\cos(t)|^q dt)^{1/q}$$

Variant query $\int_{\mathbb{R}^2} |\sin(t)\cos(2t)|dt$ will not find Hölder’s inequality since that would introduce inconsistent substitutions $x \mapsto t$ and $x \mapsto 2t$.

The MathWebSearch backend is realized as a RESTful web service that keeps a formula index in memory and hit URIs in database. MathWebSearch front-ends post MathML queries via HTTP and receive XML results.

**System Architecture: Web Service with multiple Front-Ends**

MathWebSearch aims at high-quality hits only, reported 434 hits:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>145</td>
<td>67</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
</tr>
</tbody>
</table>

- a) lots of hits (100+): general queries with multiple query variables
- b) few hits: specific queries with precise expressions.
- c) no hits: errors in query or missing exact matches

**Current Work: Full-Text Search, Ranking, Extensions, Embedding**

**Searching other Corpora**

e.g. Spreadsheet formulae \(\Rightarrow\) just encode in content MathML!