MELT
Matching EvaLuation Toolkit

Sven Hertling; Jan Portisch; Heiko Paulheim
SEMANTiCS 2019 – Karlsruhe – 2019/09/11
Joint Work

Sven Hertling  
Data and Web Science Group, University of Mannheim  
sven@informatik.uni-mannheim.de

Jan Portisch  
Data and Web Science Group, University of Mannheim / SAP SE  
jan@informatik.uni-mannheim.de

Heiko Paulheim  
Data and Web Science Group, University of Mannheim  
heiko@informatik.uni-mannheim.de
Agenda

• Motivation
• What is MELT?
• Usage Example
• Q&A
MOTIVATION
Ontology Alignment Evaluation Initiative (OAEI)

Ontology Alignment Evaluation Initiative
- running campaigns since 2005
- structured in tracks (similar to task sets)
- researchers submit their implementation
  - centrally evaluated by track organizers
  - results published
Tooling

Semantic Evaluation at Large Scale
- (among others) packaging and evaluation
- OAEI support since 2010

Holistic Benchmarking of Big Linked Data
- (among others) packaging and evaluation
- OAEI support since 2017
- OAEI 2018: 6/19 matchers support HOBBIT
Pain Points

- **Limited evaluation capabilities** in SEALS, HOBBIT, and the Alignment API
- No easy-to-use **parameter tuning**
- Packaging process might be **complicated** for new entrants to the community
- Tooling **Java-focused** (no Python)
- Implementation of the Alignment API **not maven-based**
- **Tool breaks**
What is MELT?
What is MELT?

- **Easy** matcher development
- **Non-Java** matcher development
- **Maven** support

- Facilitate **matcher packaging**
- Facilitate **matcher submission**

- **Streamlined** development process
- **Integration** with existing tooling
- **OAEI support**
- **Extensibility**

- Allow for **parameter optimization**

- **Advanced evaluation** capabilities
- **Evaluation before packaging**
- Allow for **interactive visualization**

MOTIVATION | WHAT IS MELT? | USAGE EXAMPLE
What is MELT?

Matcher Development

Matcher Submission

Matcher Fine-Tuning

Matcher Evaluation
What is MELT?

Matcher Development

Matcher Submission

Matcher Fine-Tuning

Matcher Evaluation
Yet Another Alignment API (YAAA)
Full SEALS/HOBBIT Support

MOTIVATION | WHAT IS MELT? | USAGE EXAMPLE
External Matching

- Simple wrapping
- Packageable for HOBBIT and SEALS
- Matcher can still be evaluated in MELT
- Documentation and demo project available on GitHub
What is MELT?

Matcher Development

Matcher Fine-Tuning

Matcher Submission

Matcher Evaluation
Parameter Tuning

- Run matcher configurations in **parallel** (i.e., multi-threaded)
- Hand over `ExecutionResultSet` to `Evaluator` and pick best value according to what you want to optimize
- **Out-of-the-box classes** that assist you
What is MELT?

Matcher Development

Matcher Fine-Tuning

Matcher Submission

Matcher Evaluation
Evaluation

**Full OAEI support**: All tracks available (one-time automated download)

```java
Track track = TrackRepository.Multifarm.getSpecificMultifarmTrack("ar", "cn");

ExecutionResultSet ers = new ExecutionResultSet();

ers.addAll(Executor.run(track.getTestCases(), new Matcher(), "Matcher");

EvaluatorCSV evaluatorCSV = new EvaluatorCSV(ers);

evaluatorCSV.write();
```

Multiple evaluators available, extensible.
Exemplary Evaluation

“Show me the false positive class-class mappings for Multifarm on track en-de for matcher WiktionaryMatcher.”
Exemplary Evaluation

“Show me the false positive class-class mappings for Multifarm on track en-de for matcher WiktionaryMatcher.”
“Show me the false positive class-class mappings for *Multifarm* on track *ende* for matcher *WiktionaryMatcher.*” → Just filter the correspondences!
More Evaluation

Filter for Scores or Correspondences

• **Micro Average Precision** and **Macro Average Precision** over track *Conference*

• All *residual true positives* for track *Anatomy*

• **Macro Average Class-F$_1$** for all tracks

• ...
What is MELT?

Matcher Development

Matcher Submission

Matcher Fine-Tuning

Matcher Evaluation
Matcher Submission

(pom xml)

(fill out a template – no maven wizardry required)

> mvn install
> mvn deploy

• Creates SEALS zip file → can be submitted right away
• Creates HOBBIT docker container
• Uploads and deploys container directly in HOBBIT automatically
Usage Example
Usage Example

Motivation

“Which matchers are interesting candidates for combination?” → Matchers with a high $F_1$ score and high “diversity”.
Quantitative Analysis I

Analysis OAEI 2018 results for Conference and Anatomy: Jaccard overlap of Alignments rendered as heatmap in LaTex.

\[ J(a_1, a_2) = \frac{|corr(a_1) \cap corr(a_2)|}{|corr(a_1) \cup corr(a_2)|} \]

Executor.loadFromAnatomyResultsFolder("myPath");
// few lines of other code (available on GitHub as // example)
# Results for Anatomy

Table 1. OAEI Anatomy 2018 Alignment Similarity

<table>
<thead>
<tr>
<th></th>
<th>ALIN</th>
<th>ALOD2Vec</th>
<th>AML</th>
<th>DOME</th>
<th>FCAMapX</th>
<th>Holontology</th>
<th>KEPLER</th>
<th>Lily</th>
<th>LogMap</th>
<th>LogMapBio</th>
<th>LogMapLt</th>
<th>POMAP++</th>
<th>SANOM</th>
<th>XMap</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALIN</td>
<td>1</td>
<td>0.93</td>
<td>0.62</td>
<td>0.97</td>
<td>0.72</td>
<td>0.47</td>
<td>0.79</td>
<td>0.63</td>
<td>0.66</td>
<td>0.6</td>
<td>0.81</td>
<td>0.63</td>
<td>0.62</td>
<td>0.65</td>
</tr>
<tr>
<td>ALOD2Vec</td>
<td>0.93</td>
<td>1</td>
<td>0.65</td>
<td>0.94</td>
<td>0.77</td>
<td>0.45</td>
<td>0.81</td>
<td>0.67</td>
<td>0.7</td>
<td>0.63</td>
<td>0.84</td>
<td>0.66</td>
<td>0.64</td>
<td>0.68</td>
</tr>
<tr>
<td>AML</td>
<td>0.62</td>
<td>0.65</td>
<td>1</td>
<td>0.62</td>
<td>0.76</td>
<td>0.3</td>
<td>0.74</td>
<td>0.72</td>
<td>0.8</td>
<td>0.82</td>
<td>0.72</td>
<td>0.83</td>
<td>0.79</td>
<td>0.83</td>
</tr>
<tr>
<td>DOME</td>
<td>0.97</td>
<td>0.94</td>
<td>0.62</td>
<td>1</td>
<td>0.73</td>
<td>0.47</td>
<td>0.79</td>
<td>0.64</td>
<td>0.66</td>
<td>0.6</td>
<td>0.81</td>
<td>0.63</td>
<td>0.62</td>
<td>0.66</td>
</tr>
<tr>
<td>FCAMapX</td>
<td>0.72</td>
<td>0.77</td>
<td>0.76</td>
<td>0.73</td>
<td>1</td>
<td>0.35</td>
<td>0.75</td>
<td>0.69</td>
<td>0.82</td>
<td>0.77</td>
<td>0.89</td>
<td>0.77</td>
<td>0.75</td>
<td>0.78</td>
</tr>
<tr>
<td>Holontology</td>
<td>0.47</td>
<td>0.45</td>
<td>0.3</td>
<td>0.47</td>
<td>0.35</td>
<td>1</td>
<td>0.38</td>
<td>0.3</td>
<td>0.32</td>
<td>0.29</td>
<td>0.39</td>
<td>0.31</td>
<td>0.3</td>
<td>0.31</td>
</tr>
<tr>
<td>KEPLER</td>
<td>0.79</td>
<td>0.81</td>
<td>0.74</td>
<td>0.79</td>
<td>0.75</td>
<td>0.38</td>
<td>1</td>
<td>0.69</td>
<td>0.78</td>
<td>0.72</td>
<td>0.75</td>
<td>0.76</td>
<td>0.71</td>
<td>0.76</td>
</tr>
<tr>
<td>Lily</td>
<td>0.63</td>
<td>0.67</td>
<td>0.72</td>
<td>0.64</td>
<td>0.69</td>
<td>0.3</td>
<td>0.69</td>
<td>1</td>
<td>0.7</td>
<td>0.68</td>
<td>0.69</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>LogMap</td>
<td>0.66</td>
<td>0.7</td>
<td>0.8</td>
<td>0.66</td>
<td>0.82</td>
<td>0.32</td>
<td>0.78</td>
<td>0.7</td>
<td>1</td>
<td>0.9</td>
<td>0.81</td>
<td>0.81</td>
<td>0.8</td>
<td>0.81</td>
</tr>
<tr>
<td>LogMapBio</td>
<td>0.6</td>
<td>0.63</td>
<td>0.82</td>
<td>0.6</td>
<td>0.77</td>
<td>0.29</td>
<td>0.72</td>
<td>0.68</td>
<td>0.9</td>
<td>1</td>
<td>0.74</td>
<td>0.8</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>LogMapLt</td>
<td>0.81</td>
<td>0.84</td>
<td>0.72</td>
<td>0.81</td>
<td>0.89</td>
<td>0.39</td>
<td>0.75</td>
<td>0.69</td>
<td>0.81</td>
<td>0.74</td>
<td>1</td>
<td>0.74</td>
<td>0.74</td>
<td>0.75</td>
</tr>
<tr>
<td>POMAP++</td>
<td>0.63</td>
<td>0.66</td>
<td>0.83</td>
<td>0.63</td>
<td>0.77</td>
<td>0.31</td>
<td>0.76</td>
<td>0.72</td>
<td>0.8</td>
<td>0.74</td>
<td>1</td>
<td>0.79</td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>SANOM</td>
<td>0.62</td>
<td>0.64</td>
<td>0.79</td>
<td>0.62</td>
<td>0.75</td>
<td>0.3</td>
<td>0.71</td>
<td>0.72</td>
<td>0.8</td>
<td>0.78</td>
<td>0.74</td>
<td>0.79</td>
<td>1</td>
<td>0.78</td>
</tr>
<tr>
<td>XMap</td>
<td>0.65</td>
<td>0.68</td>
<td>0.83</td>
<td>0.66</td>
<td>0.78</td>
<td>0.31</td>
<td>0.76</td>
<td>0.72</td>
<td>0.81</td>
<td>0.78</td>
<td>0.75</td>
<td>0.83</td>
<td>0.78</td>
<td>1</td>
</tr>
</tbody>
</table>
### Results for Conference

<table>
<thead>
<tr>
<th></th>
<th>ALIN</th>
<th>ALOD2Vec</th>
<th>AML</th>
<th>DOME</th>
<th>FCAMapX</th>
<th>Holontology</th>
<th>KEPLER</th>
<th>Lily</th>
<th>LogMap</th>
<th>LogMapLt</th>
<th>SANOM</th>
<th>XMap</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ALIN</strong></td>
<td>1</td>
<td>0.75</td>
<td>0.65</td>
<td>0.84</td>
<td>0.63</td>
<td>0.77</td>
<td>0.53</td>
<td>0.43</td>
<td>0.72</td>
<td>0.76</td>
<td>0.52</td>
<td>0.6</td>
</tr>
<tr>
<td><strong>ALOD2Vec</strong></td>
<td>0.75</td>
<td>1</td>
<td>0.58</td>
<td>0.87</td>
<td>0.67</td>
<td>0.75</td>
<td>0.51</td>
<td>0.37</td>
<td>0.67</td>
<td>0.86</td>
<td>0.5</td>
<td>0.54</td>
</tr>
<tr>
<td><strong>AML</strong></td>
<td>0.65</td>
<td>0.58</td>
<td>1</td>
<td>0.61</td>
<td>0.58</td>
<td>0.56</td>
<td>0.53</td>
<td>0.45</td>
<td>0.71</td>
<td>0.67</td>
<td>0.59</td>
<td>0.63</td>
</tr>
<tr>
<td><strong>DOME</strong></td>
<td>0.84</td>
<td>0.87</td>
<td>0.61</td>
<td>1</td>
<td>0.67</td>
<td>0.81</td>
<td>0.59</td>
<td>0.39</td>
<td>0.7</td>
<td>0.86</td>
<td>0.52</td>
<td>0.56</td>
</tr>
<tr>
<td><strong>FCAMapX</strong></td>
<td>0.63</td>
<td>0.67</td>
<td>0.58</td>
<td>0.67</td>
<td>1</td>
<td>0.6</td>
<td>0.55</td>
<td>0.41</td>
<td>0.62</td>
<td>0.66</td>
<td>0.51</td>
<td>0.53</td>
</tr>
<tr>
<td><strong>Holontology</strong></td>
<td>0.77</td>
<td>0.75</td>
<td>0.56</td>
<td>0.81</td>
<td>0.6</td>
<td>1</td>
<td>0.53</td>
<td>0.37</td>
<td>0.64</td>
<td>0.72</td>
<td>0.49</td>
<td>0.52</td>
</tr>
<tr>
<td><strong>KEPLER</strong></td>
<td>0.53</td>
<td>0.61</td>
<td>0.53</td>
<td>0.59</td>
<td>0.55</td>
<td>0.53</td>
<td>1</td>
<td>0.41</td>
<td>0.57</td>
<td>0.62</td>
<td>0.5</td>
<td>0.54</td>
</tr>
<tr>
<td><strong>Lily</strong></td>
<td>0.43</td>
<td>0.37</td>
<td>0.45</td>
<td>0.39</td>
<td>0.41</td>
<td>0.37</td>
<td>1</td>
<td>0.46</td>
<td>0.39</td>
<td>0.48</td>
<td>0.5</td>
<td>0.51</td>
</tr>
<tr>
<td><strong>LogMap</strong></td>
<td>0.72</td>
<td>0.67</td>
<td>0.71</td>
<td>0.7</td>
<td>0.62</td>
<td>0.64</td>
<td>0.57</td>
<td>0.46</td>
<td>1</td>
<td>0.7</td>
<td>0.63</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>LogMapLt</strong></td>
<td>0.76</td>
<td>0.86</td>
<td>0.59</td>
<td>0.86</td>
<td>0.66</td>
<td>0.72</td>
<td>0.62</td>
<td>0.39</td>
<td>0.7</td>
<td>1</td>
<td>0.51</td>
<td>0.56</td>
</tr>
<tr>
<td><strong>SANOM</strong></td>
<td>0.52</td>
<td>0.5</td>
<td>0.63</td>
<td>0.52</td>
<td>0.51</td>
<td>0.49</td>
<td>0.5</td>
<td>0.48</td>
<td>0.63</td>
<td>0.51</td>
<td>1</td>
<td>0.61</td>
</tr>
<tr>
<td><strong>XMap</strong></td>
<td>0.6</td>
<td>0.54</td>
<td>0.64</td>
<td>0.56</td>
<td>0.53</td>
<td>0.52</td>
<td>0.54</td>
<td>0.51</td>
<td>0.66</td>
<td>0.56</td>
<td>0.61</td>
<td>1</td>
</tr>
</tbody>
</table>
Mean Absolute Deviation (MAD) of Similarities plotted against $F_1$.

$$MAD = \frac{1}{n} \sum_{i=1}^{n} |x_i - \text{mean}(X)|$$
Results for Anatomy

Fig. 2. Matcher comparison using MAD and $F_1$ on the Anatomy data set
Fig. 3. Matcher comparison using MAD and $F_1$ on the Conference data set
There is MUCH more to MELT

Ontology **Caching** Services

Multi-Threaded Matcher Execution

Baseline **Matchers**

Execution of SEALS Packages from within MELT

**OAIE-Track Organizer Tools**

ExecutionResult Indexing

**TRY IT!**

One-Time **Auto-Download** of OAEI Tracks

Matcher **Pipelining**

Alignment **Refinement**

Automatic Reading of OAEI Result Alignments
Thank you!

Sven Hertling
Data and Web Science Group, University of Mannheim
sven@informatik.uni-mannheim.de

Jan Portisch
Data and Web Science Group, University of Mannheim
jan@informatik.uni-mannheim.de

Heiko Paulheim
Data and Web Science Group, University of Mannheim
heiko@informatik.uni-mannheim.de