AUTOMATIC ACCESS TO LEGAL TERMINOLOGY APPLYING TWO DIFFERENT ATR METHODS

MARÍA JOSÉ MARÍN CAMINO REA
1. INTRODUCTION

- Importance of identifying the terms in a specialised corpus.

- Terms are “textual realisations of specialised concepts” (Spasic et al. 2005).

- They are employed to communicate amongst specialists (Rea, 2008).

- They are mono-referential and have a univocal character (Cabré, 1993): form → content.

- Potential applications of automatic term recognition (ATR): building dictionaries and glossaries; machine translation; ontology building, etc.
1. INTRODUCTION

- This paper evaluates the efficiency of two ATR methods: *Keywords* (2008) and Chung’s (2003) on a 2.6m word legal corpus *(UKSCC).*

- Both will be validated in terms of precision and recall.
2. **UKSCC and LACELL: the study and reference corpora**

- **UKSCC (United Kingdom Supreme Court Corpus):**
  - Legal corpus of 192 judicial decisions issued by the Supreme Court of the United Kingdom (2008-2010).
  - Compiled *ad hoc* according to CL standards (Sánchez, 1995; Wynne, 2005; Pearson, 1998; Rea, 2010).
  - Monololingual and synchronic.
  - The Supreme Court chosen as source of texts due to its importance as a judicial institution: touches upon all branches of law and greater geographical scope.
  - Judicial decisions appear as the main source of law in common law countries.
2. **UKSCC and LACELL: the study and reference corpora**

- **LACELL (Lingüística Aplicada Computacional, Enseñanza de Lenguas y Lexicografía)**:  
  - Balanced general English corpus of 20m words: written (newspapers, books, magazines, brochures, letters, etc.) and oral language samples.  
  - Compiled by the LACELL research group at Murcia University (English Dept.).
3. ATR method description

Automatic term recognition (ATR) methods date back to the 1980s: they allow handling large amounts of data automatically spotting the most relevant terms in a specialised corpus. They have been profusely reviewed (Maynard and Ananiadou, 2000; Cabré et al., 2001; Drouin, 2003; Lemay et al., 2005; Vivaldi et al., 2012, etc.)

- **Keywords** (Scott, 2008):
  - Not an ATR method proper, however, it has proved to identify legal terms more efficiently than other methods designed to that purpose.
3. ATR method description

- Chung’s method (2003)
  - Singled out due to high rate of success recorded by the author (86% precision on average).
  - Chung compares a qualitative term recognition method: the rating scale approach with her own quantitative one. She concludes that terms displaying a $> 50$ ratio of occurrence are terms.
  - How to calculate it:

\[ Wr = \frac{SF(w)}{RF(w)} \] (freq. counts must be normalised)
4. Validation process and results

4.1. Validation process

- **Precision**: Percentage of true terms out of candidate terms extracted.

- **Recall**: Percentage of true terms out of total amount of terms in the whole corpus.

- We resorted to automatic validation: 10,000 entry legal electronic glossary compiled by authors used as gold standard (human validation poses problems due to subjectivity).

- Terms confirmed as true if found in glossary.

- **Keywords** implemented automatically; Chung’s method applied using spreadsheet (data obtained with Wordsmith too).
4. Validation process and results

4.2. Results

Fig. 1 Overall precision and recall achieved by each method
4. Validation process and results

4.2. Results

Fig. 2 Cumulative precision attained on top 2000 candidates
4. Validation process and results

4.2. Results

- *Keywords* excels Chung’s method both in terms of precision and recall.

- *Keywords* decreases its efficiency smoothly and constantly whereas Chung’s method performance is much more irregular.

- Chung’s method only performs better within candidates 1600-1800 dropping sharply to 3.5% precision afterwards.

- Chung’s bad results may be due to the automatic inclusion of words not in the reference corpus in the terms group, especially proper names so typical of judicial decisions.
# 4. Validation process and results

## 4.2. Results

Table 1. *Top 25 candidate terms extracted by each method*

<table>
<thead>
<tr>
<th>Method</th>
<th>Ratio</th>
<th>Keywords (2008)</th>
<th>Keyness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chung’s method</td>
<td>∞</td>
<td>COURT</td>
<td>28955.793</td>
</tr>
<tr>
<td>EHRR</td>
<td>∞</td>
<td>SECTION</td>
<td>27627.5586</td>
</tr>
<tr>
<td>EWCA</td>
<td>∞</td>
<td>PARA (paragraph)</td>
<td>25311.1152</td>
</tr>
<tr>
<td>UKHL</td>
<td>∞</td>
<td>LORD</td>
<td>25155.4434</td>
</tr>
<tr>
<td>MANCE</td>
<td>∞</td>
<td>V (versus)</td>
<td>22486.0918</td>
</tr>
<tr>
<td>SIAC</td>
<td>∞</td>
<td>APPEAL</td>
<td>21236.8652</td>
</tr>
<tr>
<td>ECHR</td>
<td>∞</td>
<td>ARTICLE</td>
<td>19301.6328</td>
</tr>
<tr>
<td>EWHC</td>
<td>∞</td>
<td>ACT</td>
<td>18577.8652</td>
</tr>
<tr>
<td>BAILII</td>
<td>∞</td>
<td>CASE</td>
<td>18328.9512</td>
</tr>
<tr>
<td>GESTINGTHORPE</td>
<td>∞</td>
<td>LAW</td>
<td>10458.0918</td>
</tr>
<tr>
<td>FOSCOTE</td>
<td>∞</td>
<td>JUDGMENT</td>
<td>9297.75</td>
</tr>
<tr>
<td>EARLSFERRY</td>
<td>∞</td>
<td>APPELLANT</td>
<td>8048.33496</td>
</tr>
<tr>
<td>JFS</td>
<td>∞</td>
<td>PROCEEDINGS</td>
<td>7787.61963</td>
</tr>
<tr>
<td>ECHR</td>
<td>∞</td>
<td>CONVENTION</td>
<td>7764.64355</td>
</tr>
<tr>
<td>STOJEVIC</td>
<td>∞</td>
<td>WHETHER</td>
<td>7716.16992</td>
</tr>
<tr>
<td>TURPI</td>
<td>∞</td>
<td>LJ</td>
<td>7707.0918</td>
</tr>
<tr>
<td>LJ</td>
<td>∞</td>
<td>RIGHTS</td>
<td>7023.53613</td>
</tr>
<tr>
<td>DALLAH</td>
<td>∞</td>
<td>DECISION</td>
<td>6950.50488</td>
</tr>
<tr>
<td>SUMPTION</td>
<td>∞</td>
<td>ORDER</td>
<td>6632.18164</td>
</tr>
<tr>
<td>SEISED</td>
<td>∞</td>
<td>JURISDICTION</td>
<td>6374.33105</td>
</tr>
</tbody>
</table>
5. Conclusion

- Evaluating the efficiency of ATR methods is highly recommendable to select the ones that suit our corpus best, especially due to the fact that some of them are domain-dependent like Chung’s.

- Nevertheless, as put forward by Lemay (2005: 245), “much still remains on the terminologist’s ability to differentiate a relevant unit from a non-relevant one. Lists must be scanned to remove irrelevant units”. Actually, “fine-grained semantic distinctions still rely ... on terminologists”.
References