A Content-Constrained Spatial (CCS) Model for Layout Analysis of Mathematical Expressions

Texas A&M University
Department of Computer Science & Engineering
Xing Wang, Jyh-Charn Liu
Presenter: Xing Wang
What’s ME layout and application

• The spatial relationships, including: OVER, UNDER, HOR, SUP, SUB

\[ \delta = \sum_{j=1}^{n} (-1)^{j+1} \zeta_j d\zeta_j \]

• Applications:
  – From OCR or PDF font setting layout => Math layout => semantic parsing
  – Indexing for mathematical information retrieval
  – Enable symbolic computing for inference, proof checking
Challenges

• Varying vertical setting

• Overlapping in the feature distribution for different relative spatial relation

\[
K_{\alpha+1} \left( \sum (-1)^{|J|-1} \frac{\partial a}{\partial \zeta_k} d\zeta^J \wedge d\zeta^k \wedge d\zeta^I \right)
\]
Content-constrained Spatial Model

- Two-step approach
  - First identify the non-horizontal structure
  - Then resolve super/subscript using Horizontal Content-Constrained Spatial Model

1. Vertical/enclosed structure processing
   - Accent Processing
   - Radical Processing
   - Fraction Processing
   - BindVar Processing
   - SUP&SUB Processing
   - Fence Processing

2. CCS horizontal structure processing
   - Layout candidates generation
     - 7. Op./Rel./Punct. Constraints
     - 8. Script Level Sameness
   - Horizontal chain generation
   - Recursive sub config gen.
   - Layout candidates Ranking
     - Layout
     - CRPCs
     - P(CRPCs|HR, NVCD)
Non-horizontal Structure Identification

**<<Interface>>**

<table>
<thead>
<tr>
<th>MEGroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>- children(): vec&lt;MEGroup&gt;</td>
</tr>
<tr>
<td>- attachedObj(): MEObject</td>
</tr>
<tr>
<td>- attacherObj(): MEObject</td>
</tr>
</tbody>
</table>

**Vertical ME Structures**

1. MEAccentGroup
   - hatSymbol: MESymbol
   - underMEGroup: MEGroup

2. MEFractionGroup
   - fractionPath: MEPath
   - upMEGroup: MEGroup
   - downMEGroup: MEGroup

3. MEBindVarGroup
   - operator: MESymbol
   - upMEGroup: MEGroup
   - downMEGroup: MEGroup

4. MESupSubGroup
   - baseMEGroup: MEGroup
   - subMEGroup: MEGroup

**Horizontal ME Structures**

1. MEGroup
   - children(): vec<MEGroup>

2. MESymbolGroup
   - meSymbol: MESymbol

**Enclosed ME Structures**

1. MERadicalGroup
   - radicalSymbol: MESymbol
   - containMEGroup: MEGroup

2. MESupGroup
   - baseMEGroup: MEGroup
   - supMEGroup: MEGroup

3. MEFenceGroup
   - open/closSymbol: MESymbol
   - meGroup: MEGroup

---

**1. Vertical/enclosed structure processing**

1. Accent Processing
2. Radical Processing
3. Fraction Processing
4. BindVar Processing
5. SUP&SUB Processing
6. Fence Processing

\[ \delta = \sum_{j=1}^{n} (-1)^{j+1} \zeta_j \hat{\zeta}_j \]

\[ \alpha = \sqrt{1 - |\zeta|^2} \sqrt{1 - z^2} / (1 - \zeta \cdot z) \]

\[ h^b_{\alpha} = \sum_{q=0}^{n-1} \frac{\Gamma(n + \alpha - q - 1)}{\Gamma(n + \alpha)q} \]
Horizontal Content-Constraints Spatial Model

• What is a horizontally ME layout $HML_{ME}$?
• What properties a valid $HML_{ME}$ should satisfy?
  – Four axioms and constraints
• How to generate $HML_{ME}$ candidates?
• How to rank the candidates to find the most proper one?
  – Transformation into equivalent child-parent-relation chain set
  – Bayesian probabilistic inference
Horizontal ME layout and four axioms

- **HML**_{ME} = \{T_i\}_{i=1,...,n}
  - \( T_i = (i, p, r) \)
    represents that \( s_i \) is the right sibling to \( s_p \) with the relation of \( r \), where \( r \in \{HOR, SUB, SUP\} \).

\[
\delta = \sum_{j=1}^{n} (-1)^{j+1} \xi_j d \xi_j
\]
Content-based constraints

Exist One to attach to with Horizontal relationship

\[ \delta = \sum_{j=1}^{n} (-1)^{j+1} \xi_j d\xi_j \]
$HML_{ME}$ candidates generation

HOR chain
enumeration

1 2 3
1 3 + sub problem of \{2\}
1 2 + sub problem of \{3\}
1 + sub problem of \{2,3\}
Constrained \( \text{HML}_{ME} \) candidate generation

HOR chain enumeration

1 2 3

1 + sub problem of \{2,3\}

• If the 2\textsuperscript{nd} symbol is ‘-’ sign, it could be a unary operator, which requires the symbol ‘-’ to be in HOR relation with 3\textsuperscript{rd} symbol.
Probabilistic Inference (transformation)

• To avoid local errors propagation, each layout $L = \{\langle i, i_p, r \rangle \}$ is transformed into an equivalent list of child-parent-relation chain, $\Theta_L = \{\theta_{ij}\}$ for all pairs of elements.

<table>
<thead>
<tr>
<th>Pair</th>
<th>Relation Chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,2)</td>
<td>SUP</td>
</tr>
<tr>
<td>(1,3)</td>
<td>HOR</td>
</tr>
<tr>
<td>(2,3)</td>
<td>REV_SUP, HOR</td>
</tr>
</tbody>
</table>

$\langle 2,1, SUP \rangle, \langle 1,3, HOR \rangle$
Probabilistic inference (formulation & features)

- The problem is transformed into:
  
  \[ \text{argmax}_{L \in L(S)} \: P(\Theta_L | O), \]

where:

  - \( L(S) \) is the set of all \( \text{HML}_{\text{ME}} \) candidates
  - \( O \) is the observation feature values

- Height ratio \( \Phi_{12} = \frac{h_2}{h_1} \)
- Normalize vertical center difference \( \Psi_{12} = \frac{(c_2 - c_1)}{h_1} \)

\[
P(\Theta_L | O) = \prod_{\theta_{ij} \in \Theta_L} P(\theta_{ij} | O)
= \prod_{\theta_{ij} \in \Theta_L} P(O | \theta_{ij}) P(\theta_{ij}) / P(O)
\approx \prod_{\theta_{ij} \in \Theta_L} P(O | \theta_{ij})
\]
Recursive distribution estimation for $P(O|\theta_{ij})$

- Given a CPRC $\theta_{i_1i_m}$ corresponding to a series of $m$ triples, $\langle i_{k+1}, i_k, r_k \rangle$, we have:
  
  $- HR \Phi_{i_1i_m} = \frac{H_{im}}{H_{i_1}} = \frac{H_{im}}{H_{ims}} \frac{H_{ims}}{H_{i_1}} = \Phi_{ims,i_m} \Phi_{i_1,i_m}$

  $- NVCD$ feature as $\Psi_{i_1,i_m} = \Psi_{i_1,i_m} + \Phi_{ims,i_m} \Phi_{i_1,i_m}$

A non-parametric sampling approach
Experiment Result

- Dataset: InftyCDB
- Evaluation criteria: the F1 score at the layout triple level \( \{i, i_p, r\} \)
- A overall F1 score of 0.98 is achieved.
- Accent structure affect the performance most
## Experiment Result - Case Study

<table>
<thead>
<tr>
<th>Type</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Right item of FP</strong></td>
<td>LeftPar 13, plus 25, minus 31, comma 45, equal 52, RightPar 70</td>
</tr>
<tr>
<td><strong>Left item of FP</strong></td>
<td>vert 14, zero 16, p 16, f 21, P 22, one 23, overline 25</td>
</tr>
<tr>
<td><strong>Right item of FN</strong></td>
<td>plus 25, minus 31, comma 45, equal 52, LeftPar 69</td>
</tr>
<tr>
<td><strong>Left item of FN</strong></td>
<td>L 14, tilde 15, RightPar 16, P 18, f 18, overline 50</td>
</tr>
</tbody>
</table>
Conclusion

• We propose a content-constrained spatial model that tries to overcome the varieties in the vertical setting of the symbols.
• The two-step approach could resolved the structure of an ME very efficiently and avoid propagation of local errors.
• Limitation: the sampling based distribution estimation is computational bottleneck -> approximation or elimination of redundant computation.
Thanks!

Q&A