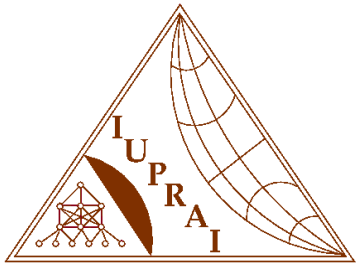
ICVGIP 2016



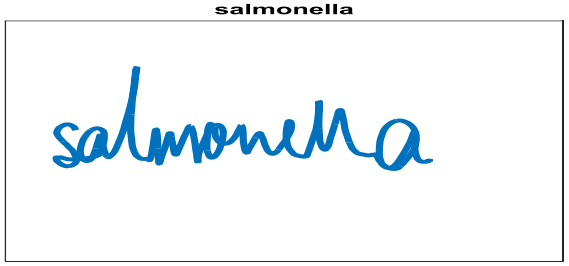
Generating Synthetic Handwriting using n-gram letter glyphs

Arka Ujjal Dey, IIT Jodhpur; Gaurav Harit, IIT Jodhpur

Presented by: Arka Ujjal Dey

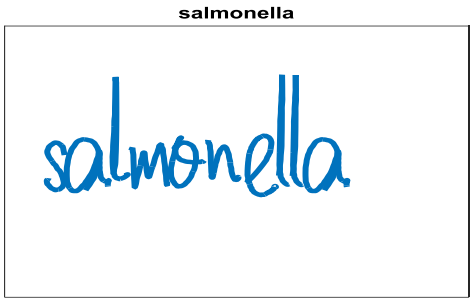
The Handwriting Synthesis task

Character model

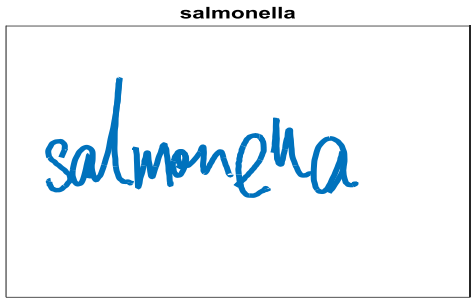


Character model + Ligature model

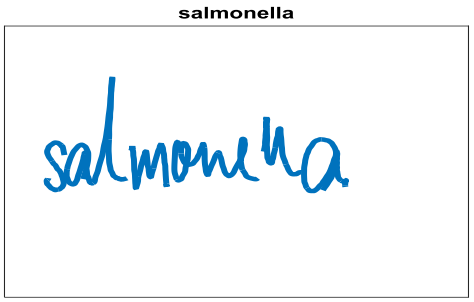
Synthesis Using only glyphs



s-a-l-m-o-n-e-l-l-a



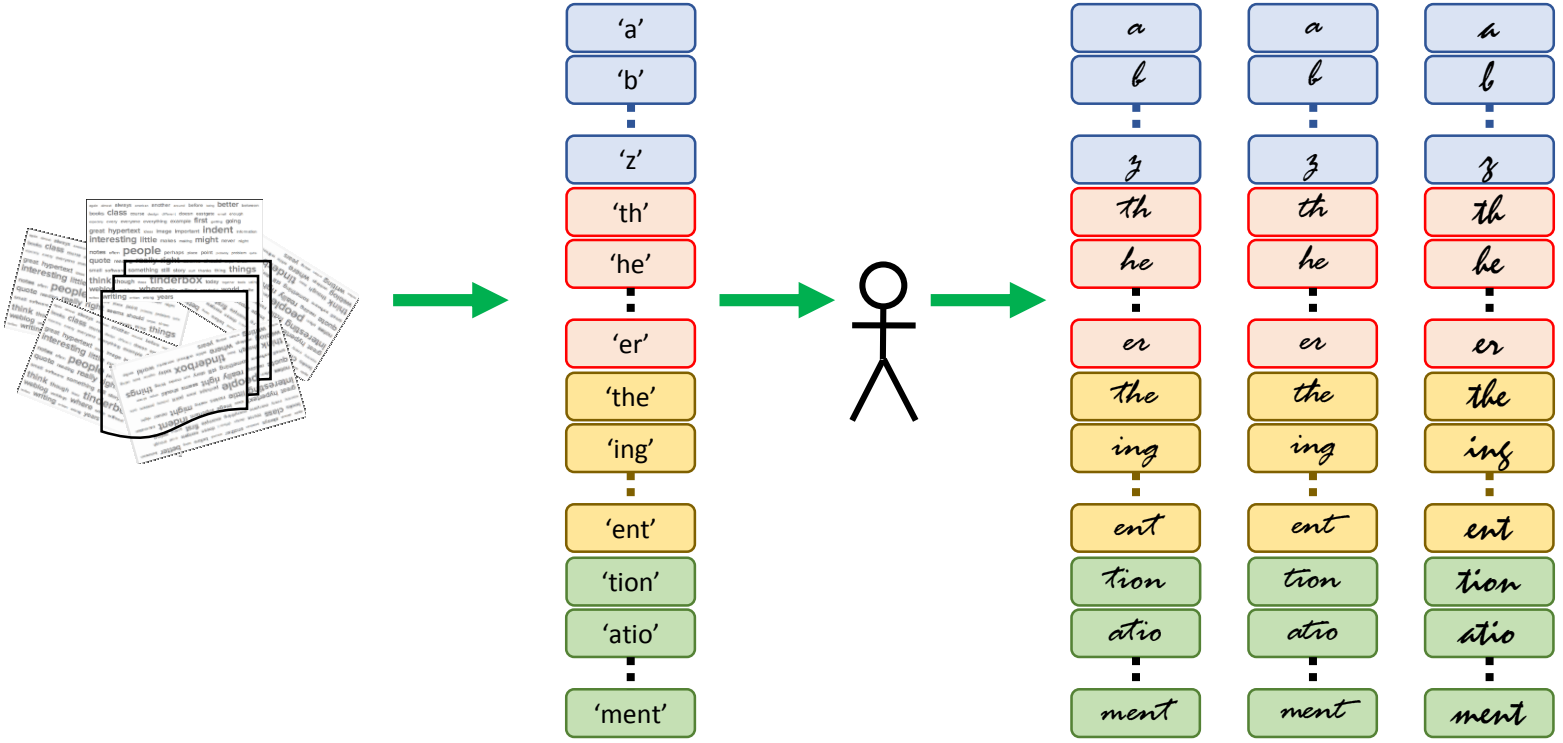
s-al-m-on-e-ll-a



s-al-m-one-ll-a

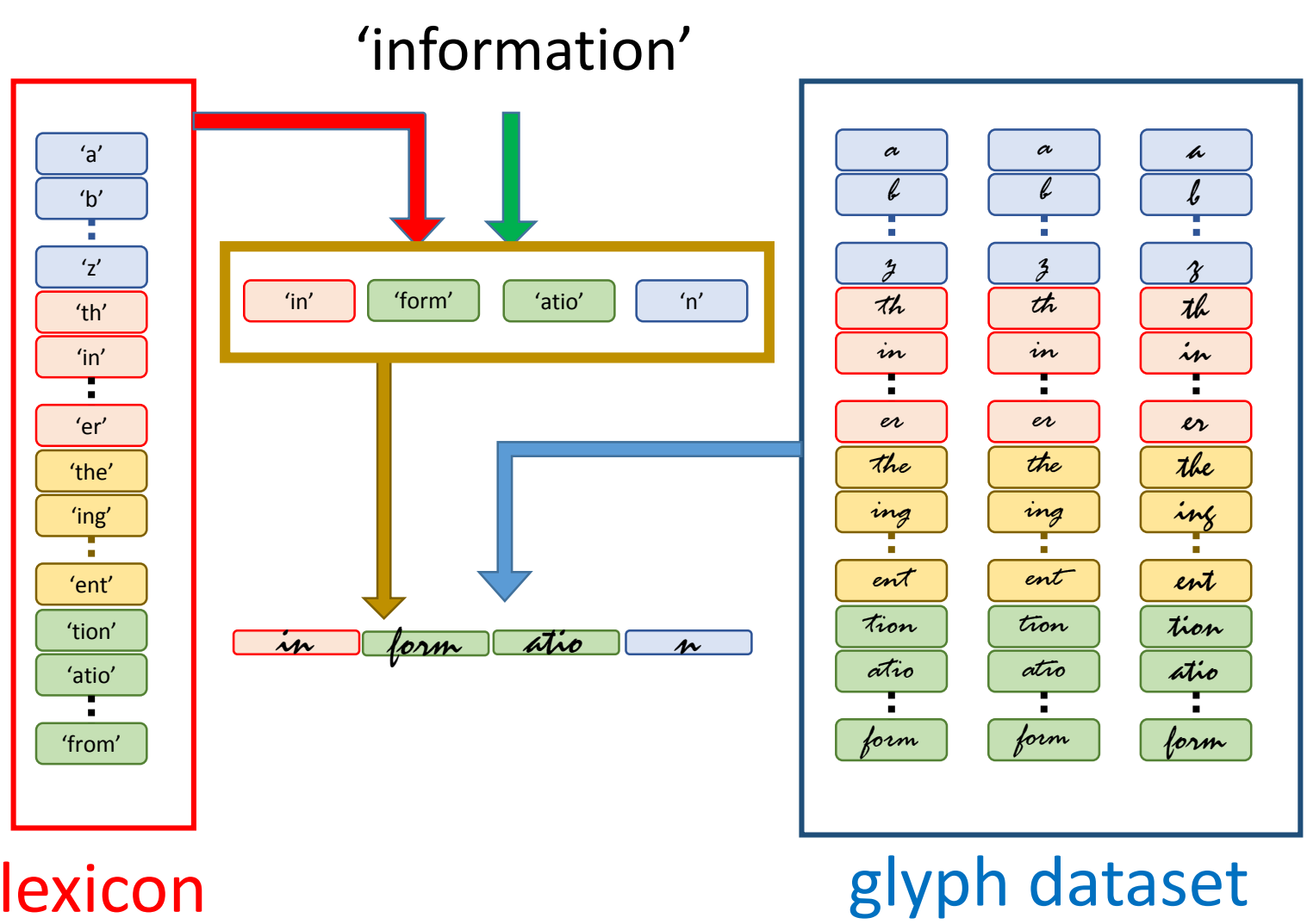
With multi-letter glyphs ligature comes for free

Offline pipeline

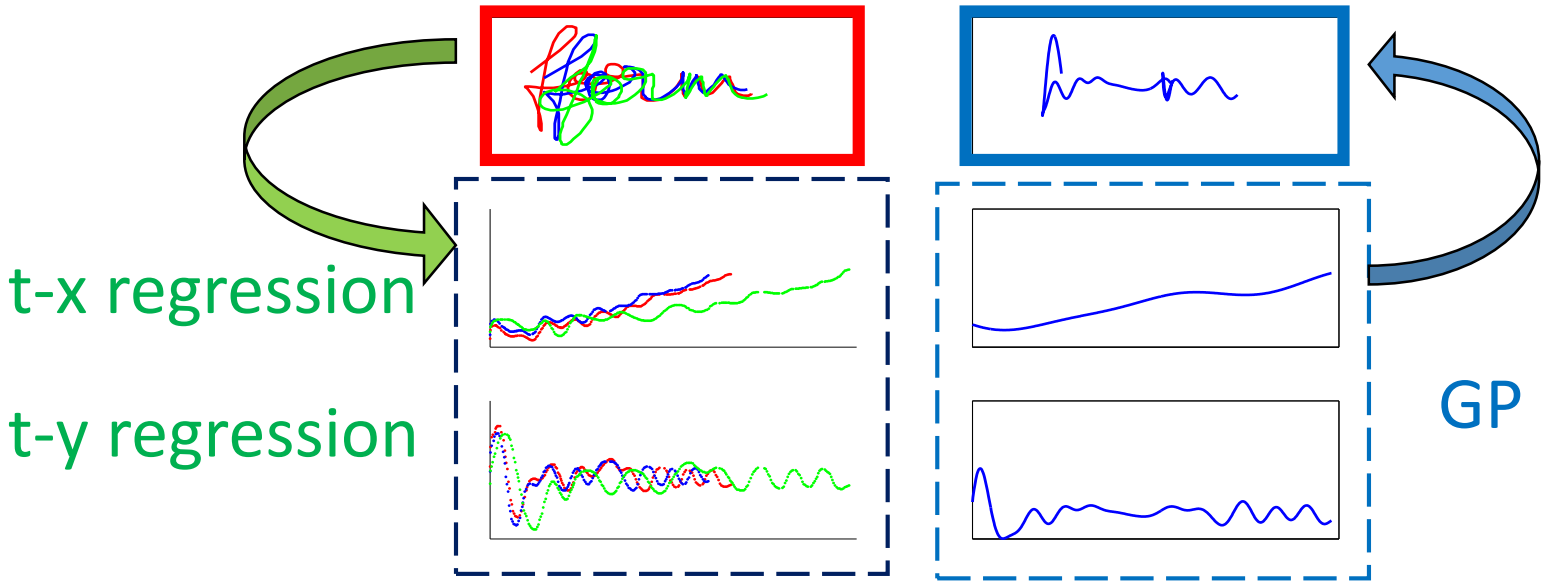


corpus lexicon glyph dataset

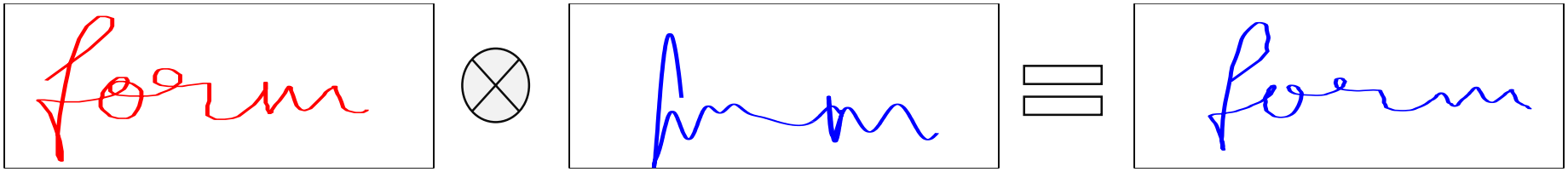
example


text
'information'
'a'
'b'
'z'
'th'
'in'
'er'
'the'
'ing'
'ent'
'tion'
'atio'
'from'
'in'
'form'
'atio'
'n'
lexicon
glyph dataset

From samples to Perturbation: GP regression

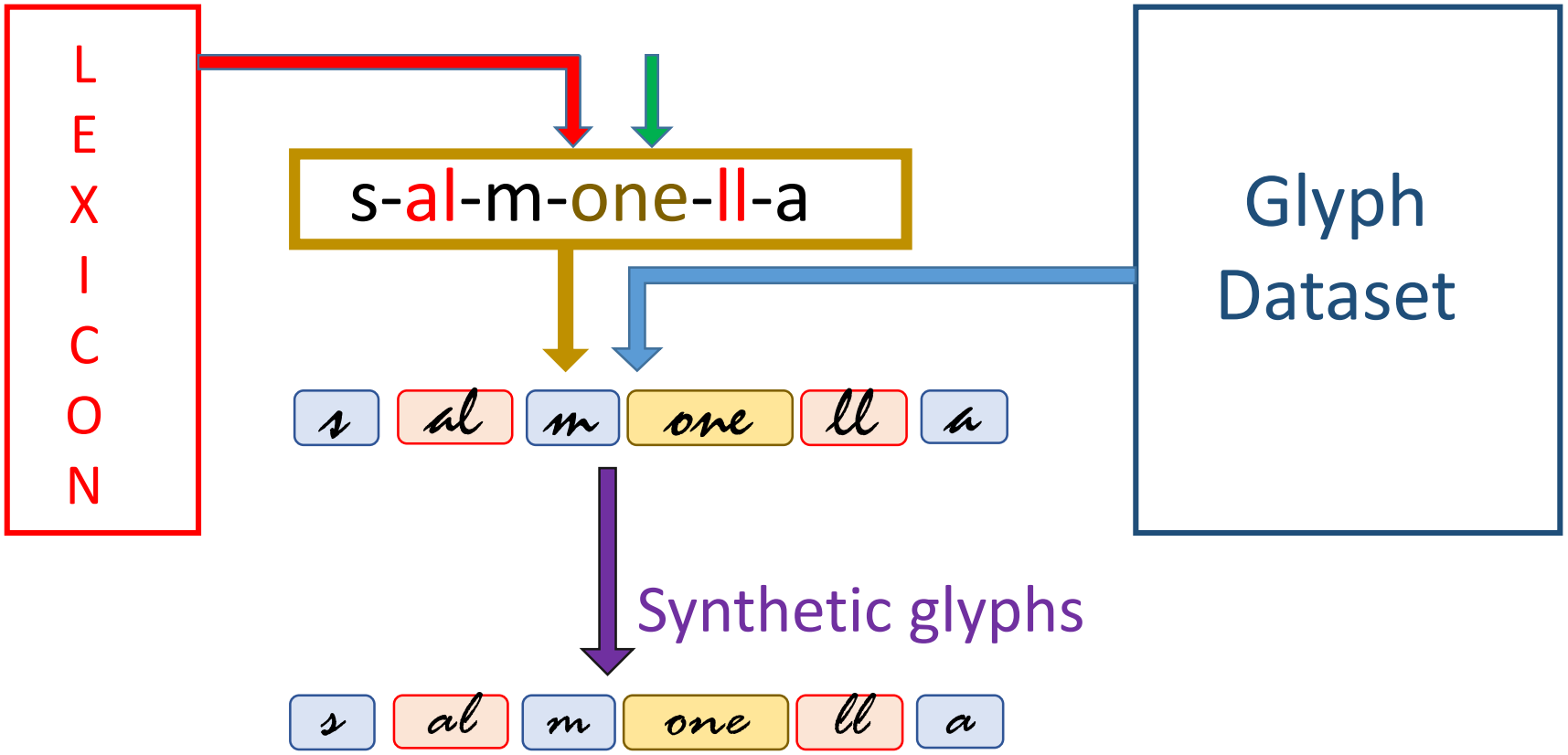


From Perturbation to Generation



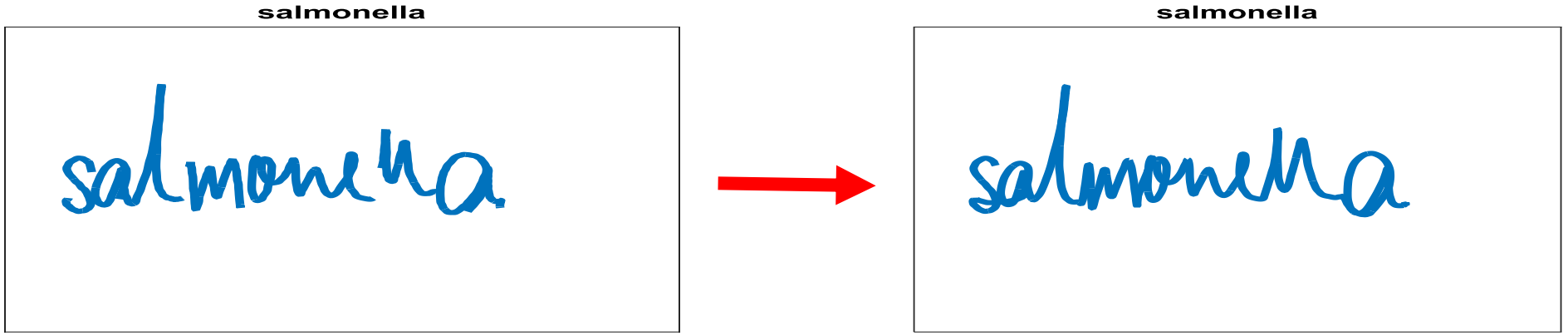
random seed | Perturbation | Synthetic glyph

Text: salmonella



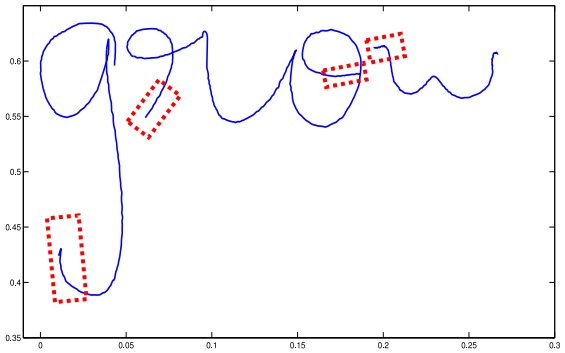
E.g.

Spline based *ligature* model ..(1)

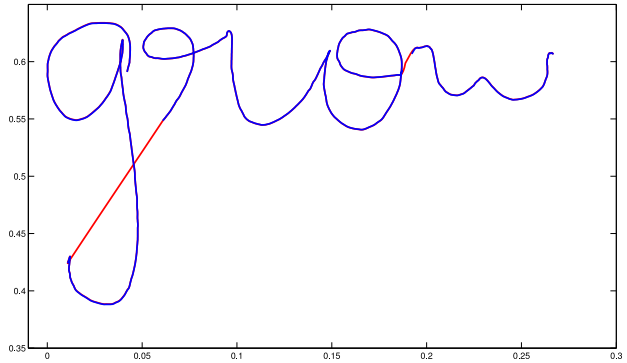


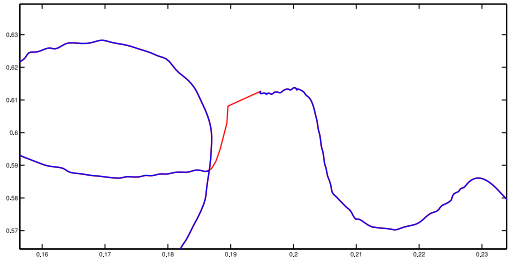
from simple glyph alignment to ligature

Spline based ligature model ..(2)

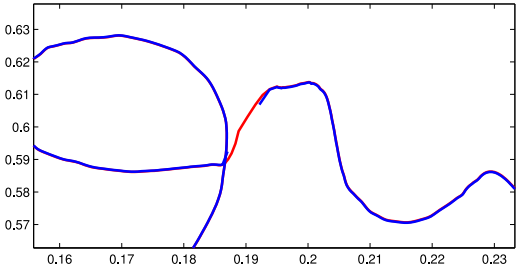


stitching

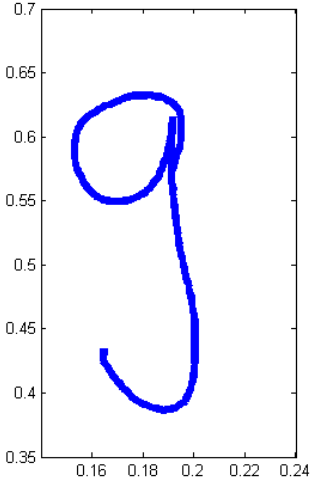


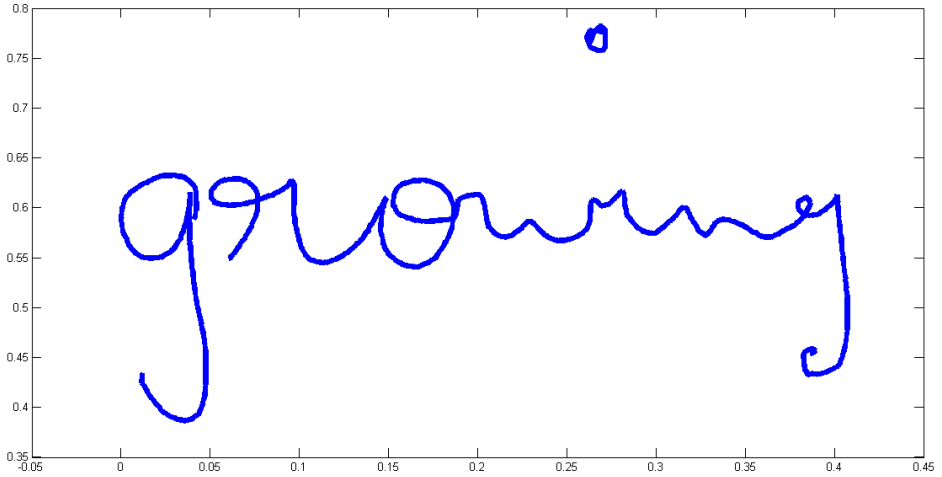


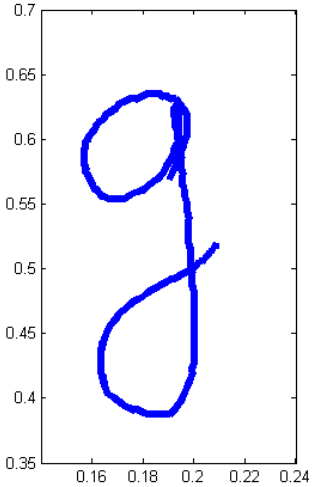
smoothing

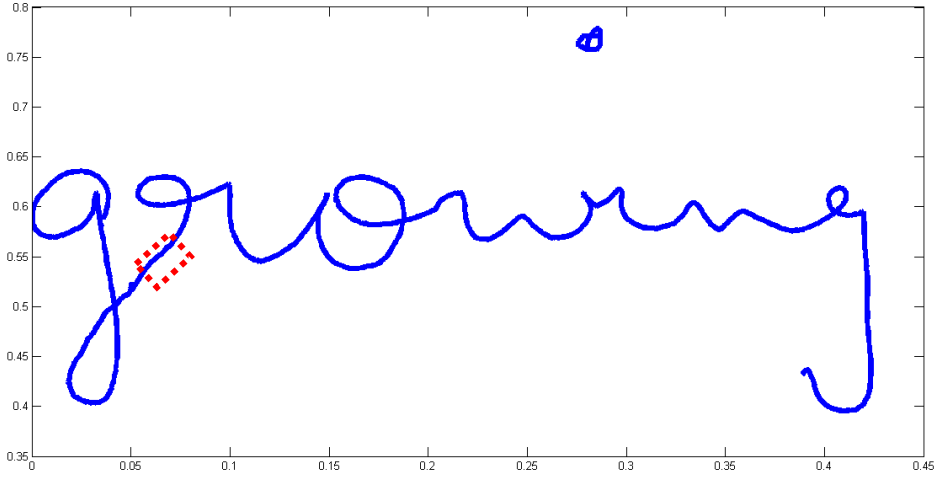


Pen lift, when **not** to connect

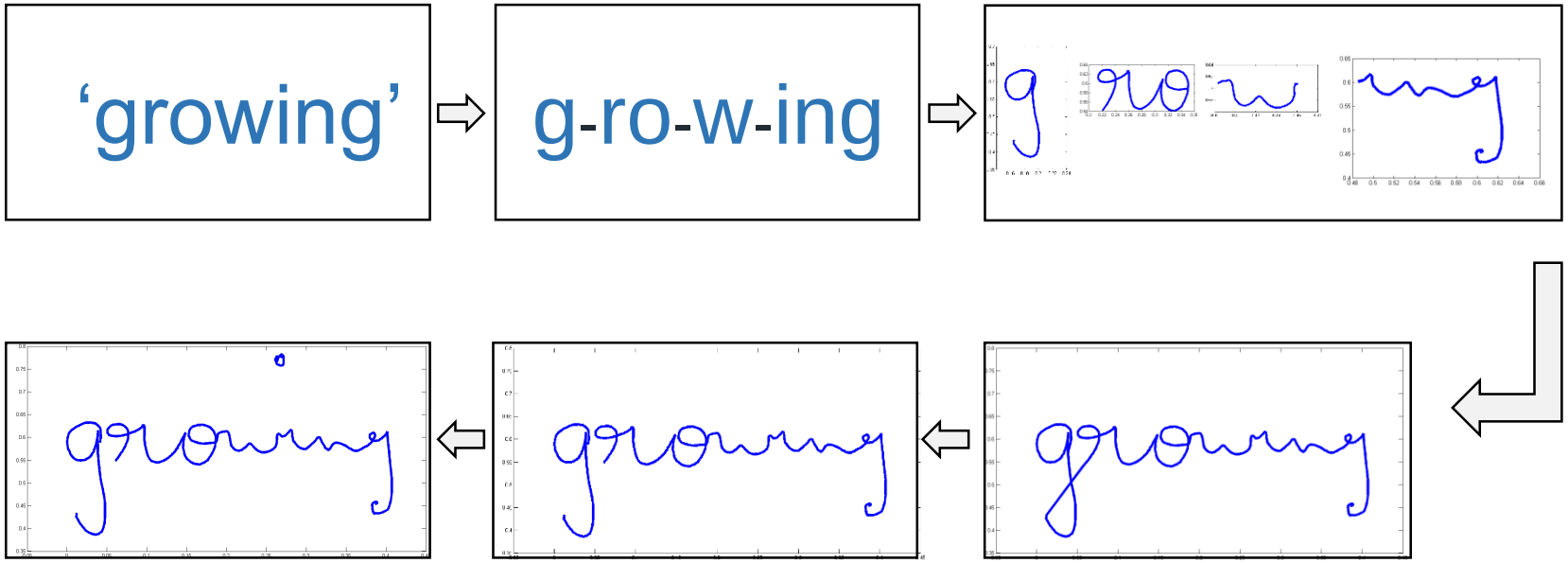




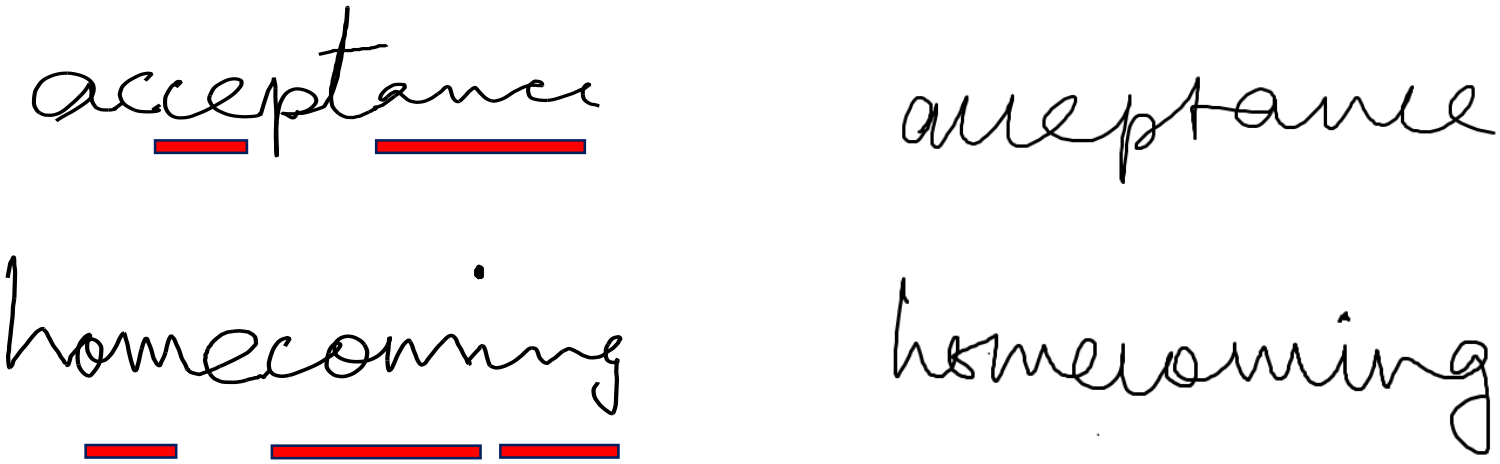




Online Pipeline



Can sum of parts be the whole?



Constraints

- Glyph input method
- Alignment issues
- Ligature is basic

Summary

- Extended lexicon captures **language statistics**
- Data driven synthetic glyphs are **consistent**
- N-gram glyphs minimizes ligature
- lexicon size, and ligatures required are inversely related
- Basic ligature scheme
- Natural looking semi cursive script