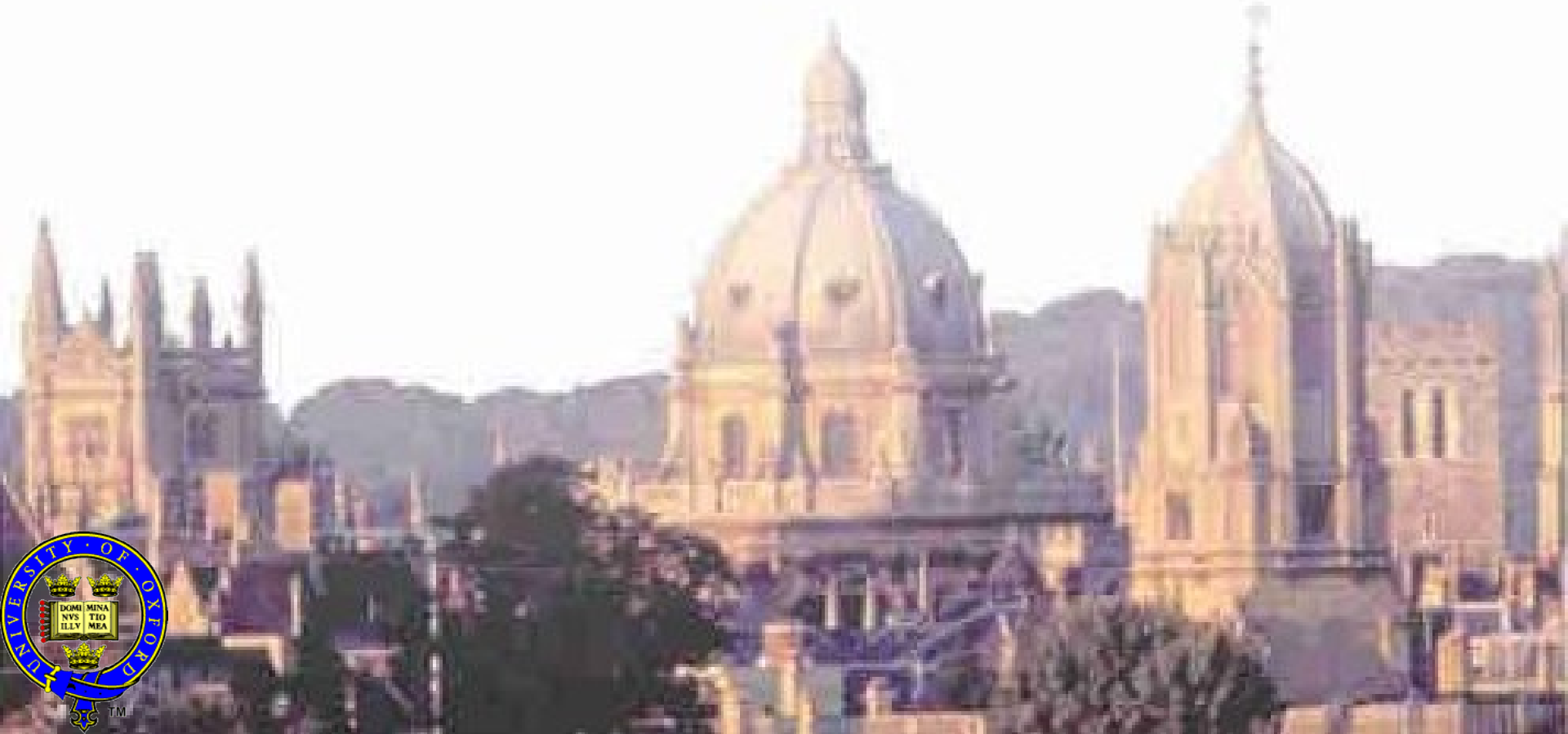
From strokelets to character interpretations

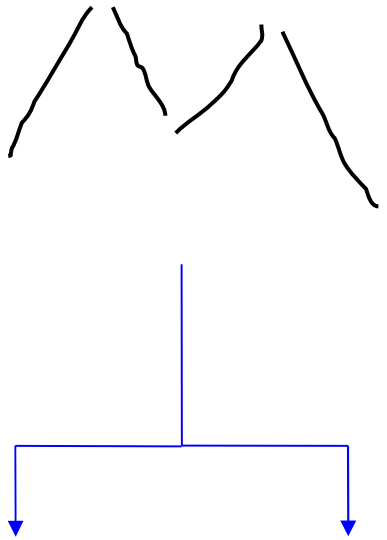
Mike Brady, Ségolène Tarte, Martin Moran



The next steps

Filling the gaps

q u e m



> "*All perceptual tasks are an effort after meaning*"
>
> G. von Helmholtz, *Optics*, 1880

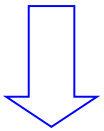
Note that (see Melissa's book), this requires a program to know the expected stylus form of "e"; rather than the ink form

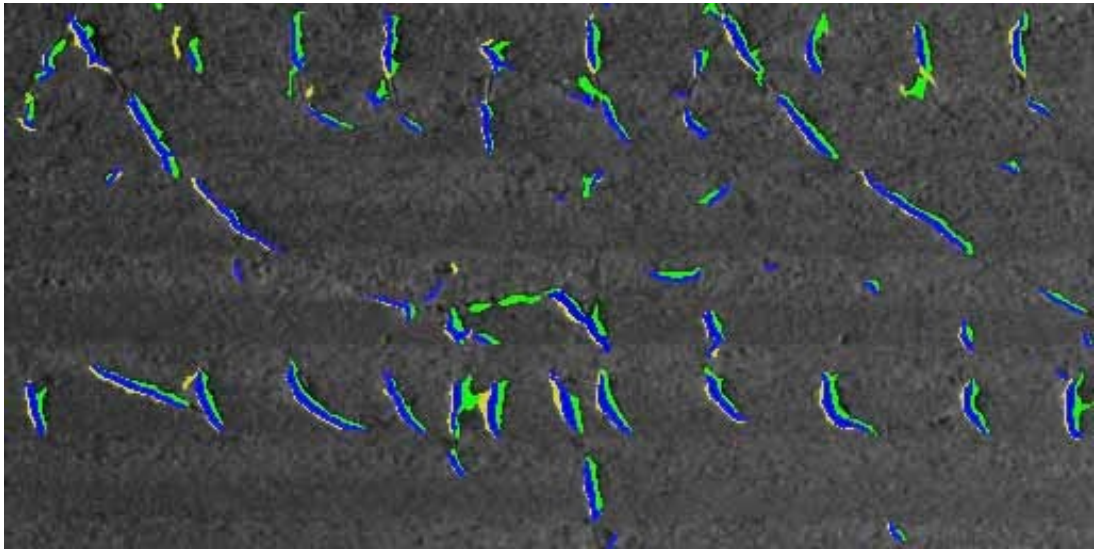
Interpretation of a scene

- Local parts of an image offer localised, and different, amounts of “constraint”
- Such localised, constrained information corresponds to a partial interpretation – a percept (Eleanor Gibson 1950)
- The human visual system combines these localised percepts to form bigger parts
  - Sometimes, we are misled in this process, as Escher so brilliantly exposed
- The processes of localised partial interpretation and their combination can be modelled in a number of ways, some more precise than others & some more explicit than others
- At each stage, prior knowledge is mobilised, effortlessly and unconsciously
  - Example: Word Superiority Effect
  - Melissa’s thesis was based on a recent (1980s) theory of reading (McClelland and Rumelhart *Distributed Parallel Processing*
- *Pace* Youtie: but experts are the least reliable at explaining their own exquisite behaviour

Statement of the problems

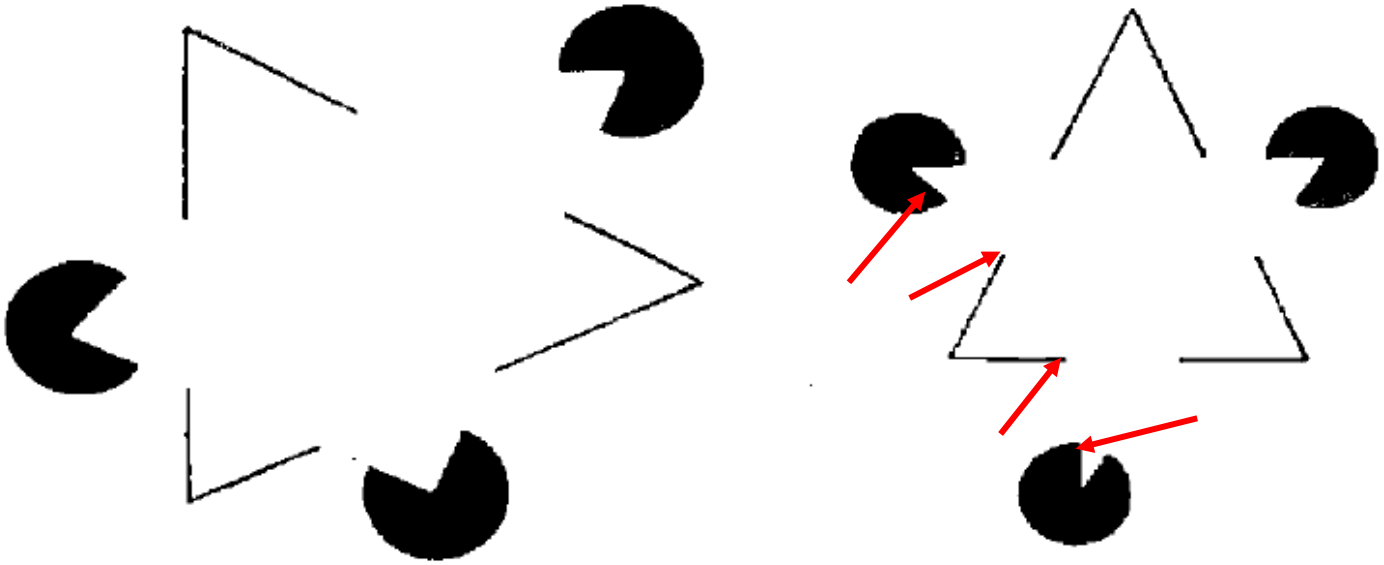
- Filling in the gaps
  - One explanation is wood grain, for which we have a representation: necessary but not sufficient
  - Subjective contours: the human visual system effortlessly fills in gaps
- Alternative explanations
  - Making them explicit
  - Keeping track of them & “reminding” the interpreter of alternative possibilities
  - Requires knowledge of *expected* written forms in different languages and on different media
- Here, we concentrate on interpretations of strokelets as possible character (sequences) – similar ideas apply to the many semantic levels at which a document is *interpreted* – see Melissa and Ségolène later





Features detected on the fragment.
Note the gaps in perceptual strokes resulting from woodgrain (and its removal)

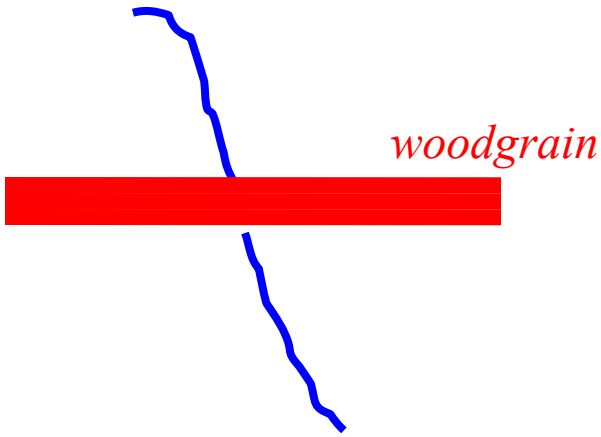
Subjective contours



The bright white “triangular” shapes seen effortlessly by the human visual system do not exist, yet we cannot help ourselves from seeing them. The shapes of these illusory contours are determined by the local constraints.

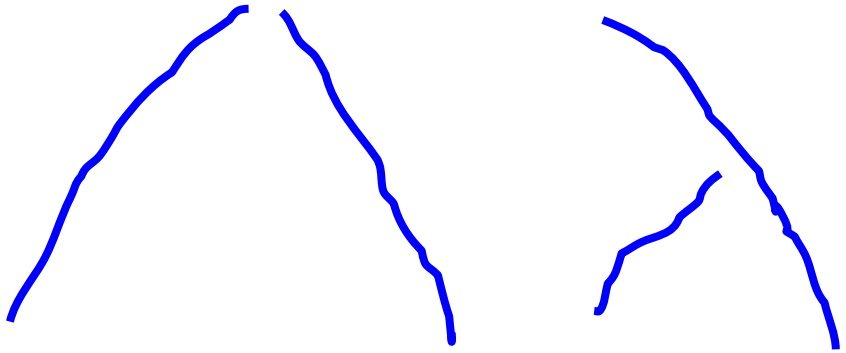
Our stroke completion algorithm is based on an idea proposed by Ullman (1984) to model human perception of subjective contours “*Filling in the gaps” Bio. Cyb.*

Filling in the gaps in strokes



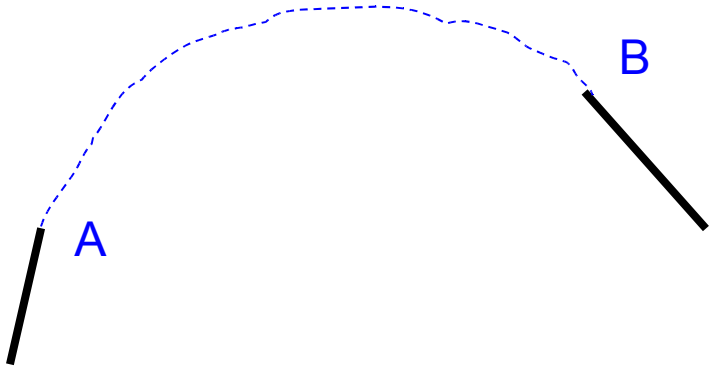
Collinearity and proximity makes connection probable

If the gap overlaps a likely woodgrain region, then it is more probable



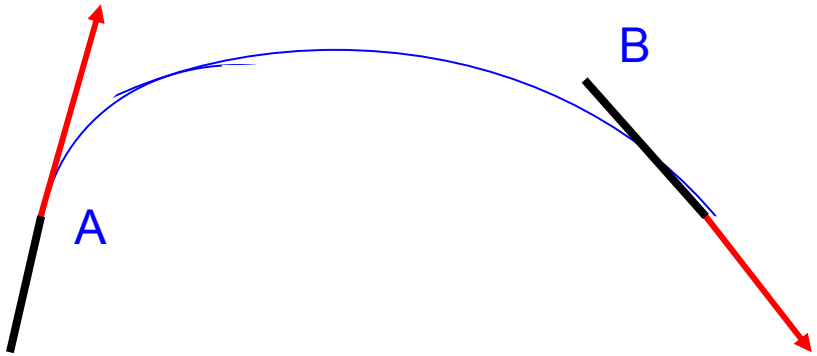
Junctions can be treated in the same way

The process



The task is to complete the gap from A to B, where the slope at A and the slope at B are shown by the black lines and the arrows.

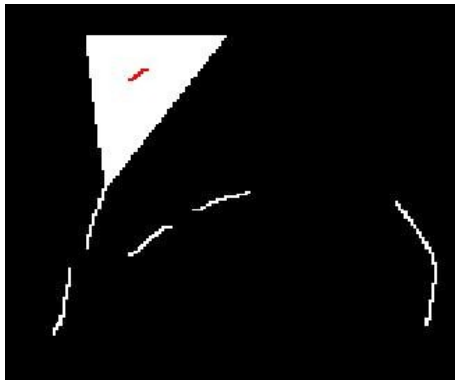
The process



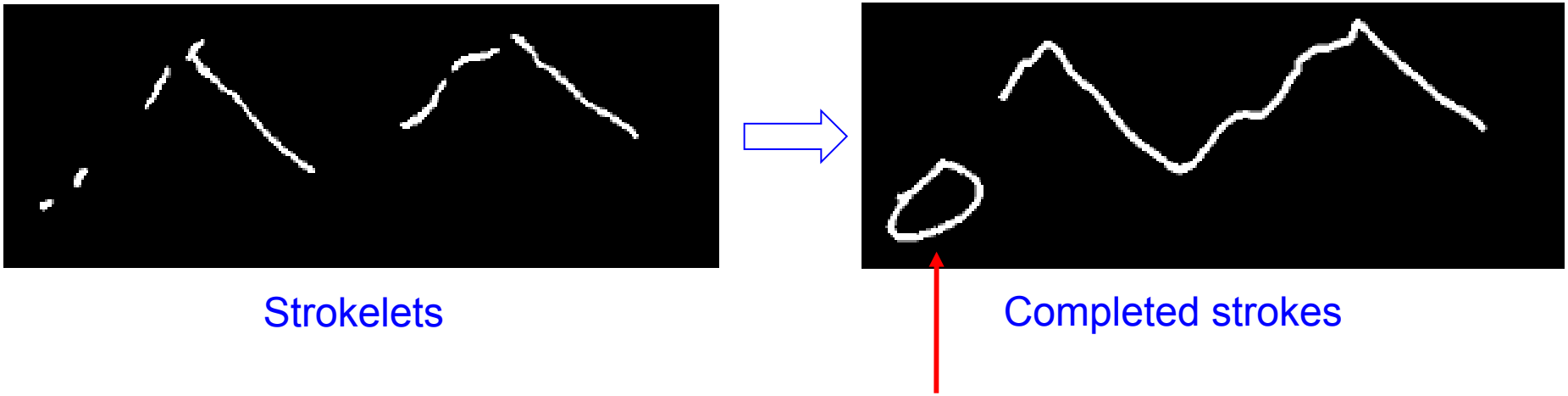
The curve that fills the gap is a cubic spline, constrained by the end point positions and directions

The task is to complete the gap from A to B, where the slope at A and the slope at B are shown by the black lines and the arrows.

Of course, the directions at A and B are only approximate, and so we construct a set of completions in slightly different directions at each end.



If only it were that easy….



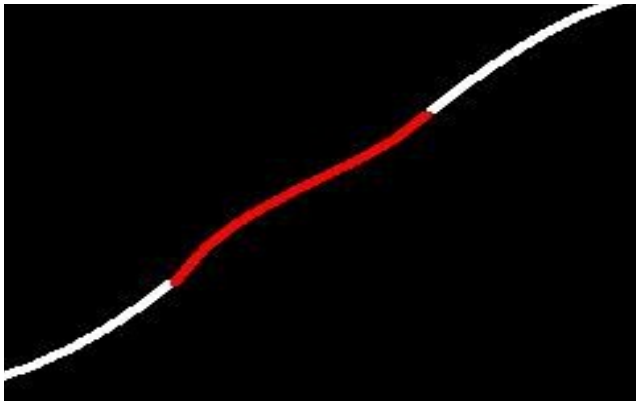
Strokelets

Completed strokes

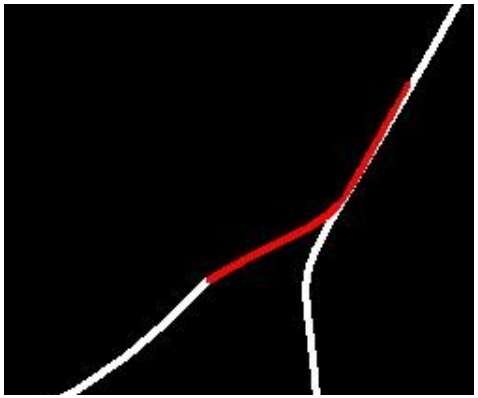
The moral of the story:

Mathematics is wonderful; but don't rely on it totally when it comes to human perception!!

• When filling the gaps in strokes, there are three types of connection/junction that need to be considered

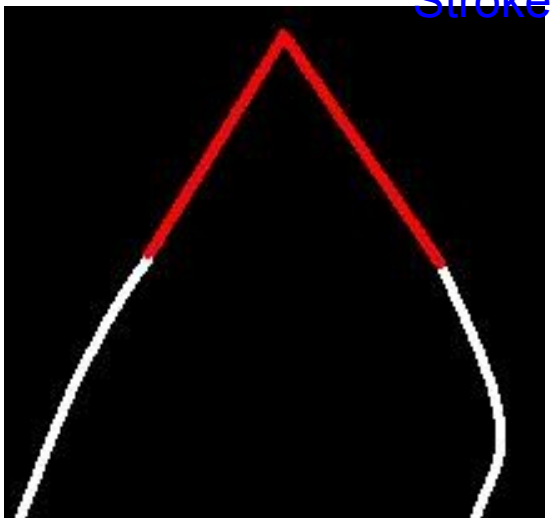


• Smooth Stroke End to Stroke End

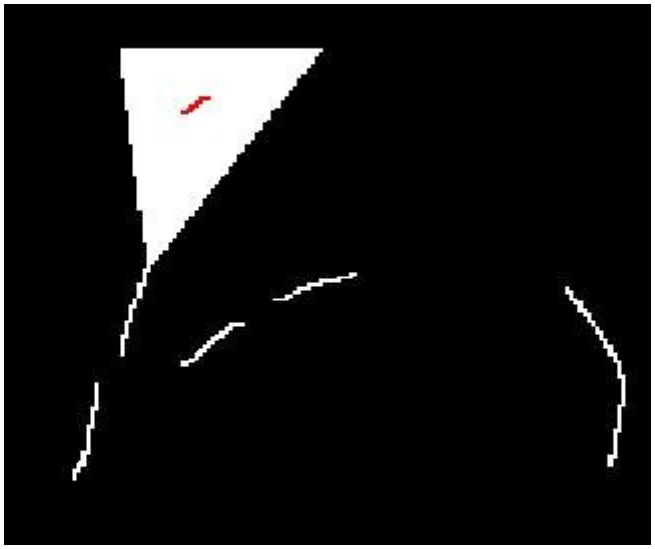


• Smooth Stroke End to middle of Stroke

The white portions are strokelets found manually or automatically; the red portions are the completions to be proposed (and which may be right or wrong)



• Non-smooth discontinuous Junction



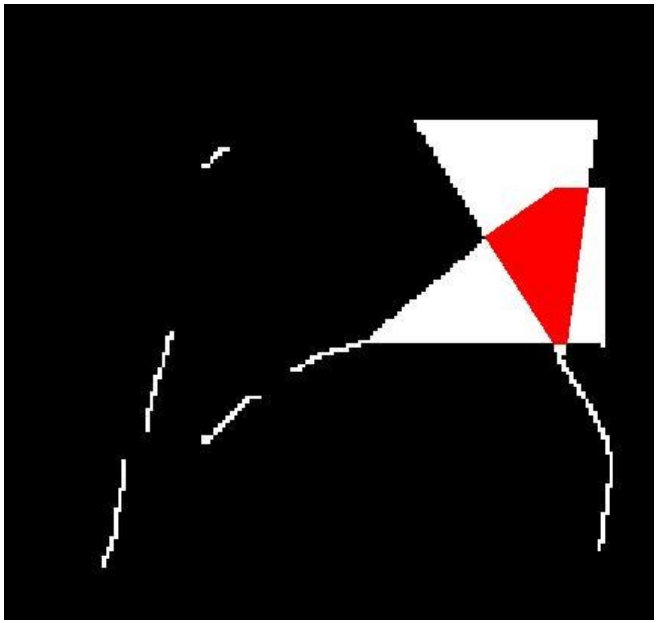
• A search area is produced by measuring the average direction of each “strokelet”

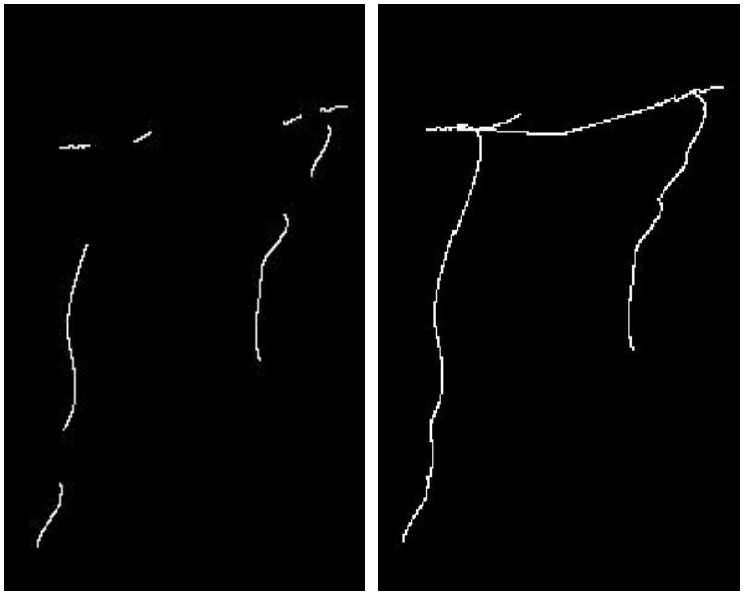
• If there is a part of another strokelet within this area, a curve interpolation is produced to smoothly match the directions of each strokelet

• If two or more search area overlap, this suggests there may be a discontinuous junction

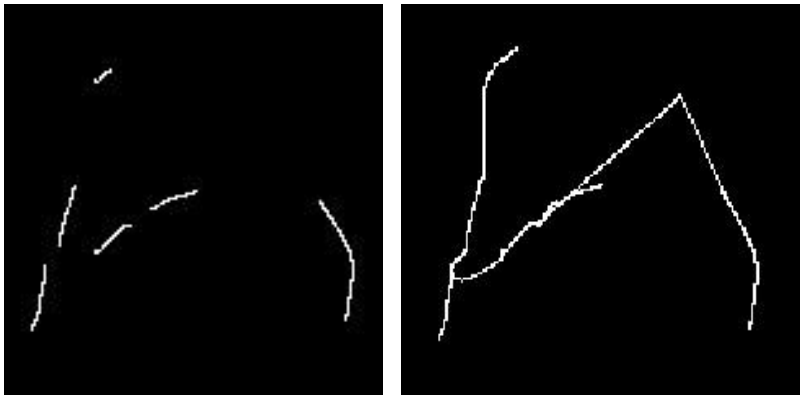
• This junction is produced by creating a straight line from the centre of the overlapping area to the end of the corresponding strokes

•With these two methods, the three junction cases shown previously can be observed.

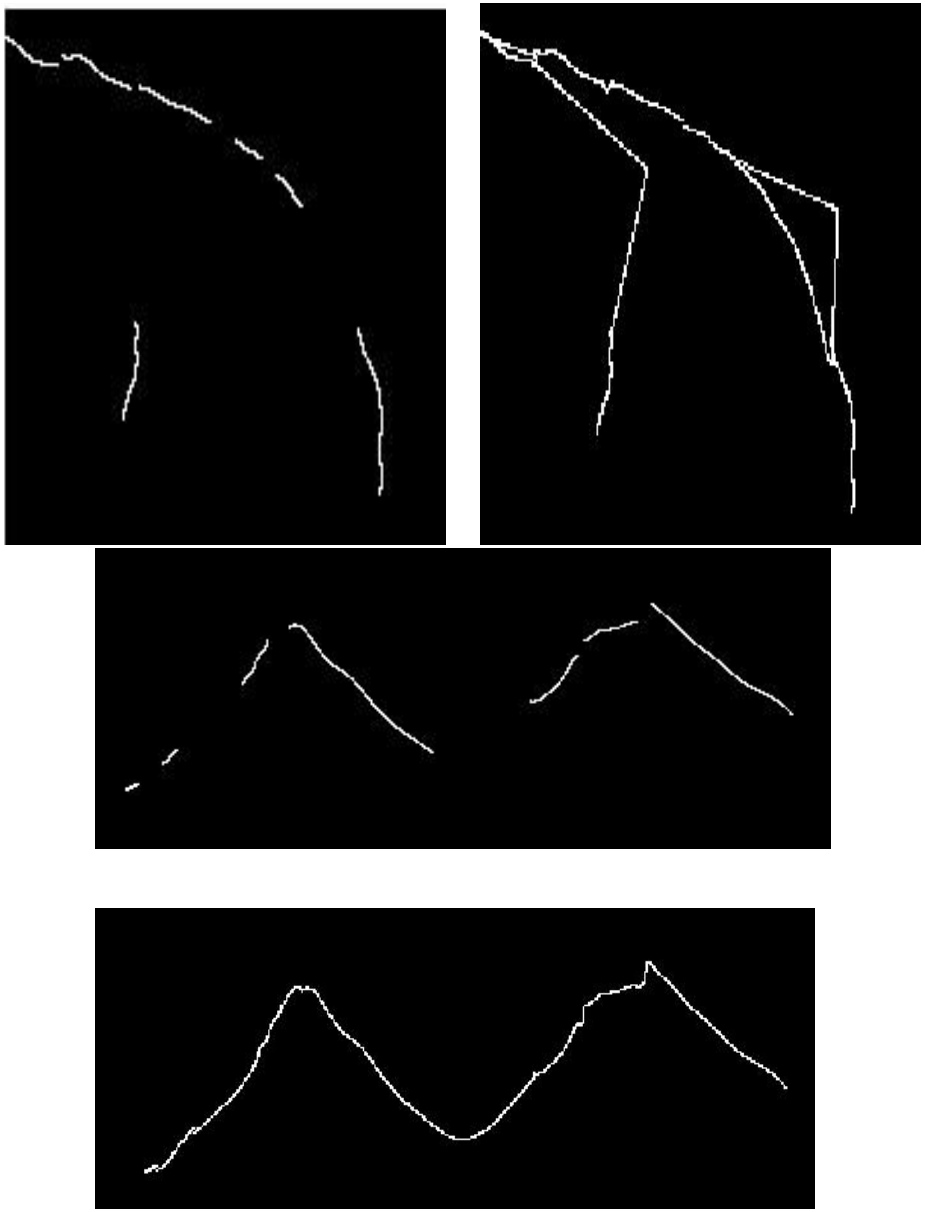




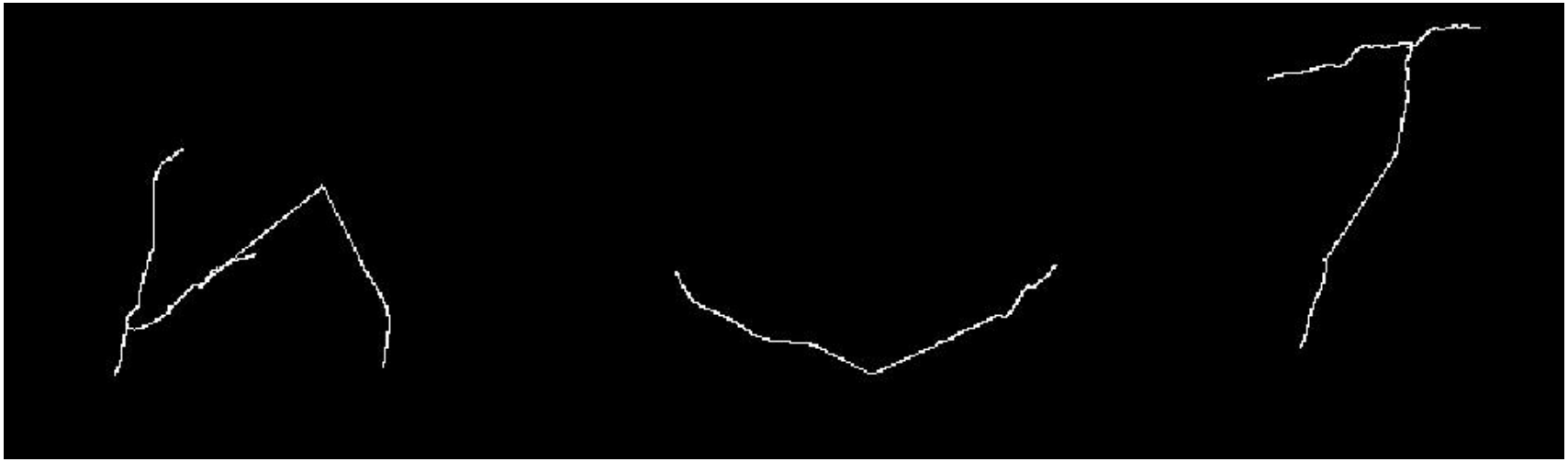
Left: before filling in, right: after the process is complete.



These examples are from those hand-labelled by Ségolène



We identify two or more completed strokes that “explain” the same stroke, then eliminate them.



Top: before; bottom: after

Top: before; bottom: after

Interpretation support tool

- The set of partial interpretations are represented
  - Explicitly
  - Visually
  - Annotated by the historian
- Elaborations of partial interpretations correspond to the mobilisation of knowledge
- The dependency of partial interpretations upon each other, the evidence, and the knowledge mobilised is made explicit
- The set of partial interpretations and dependencies can inform debate between scholars
- Equally, it can enable a group of scholars, at the beginning of a study session, to recall where they were up to last time, and why

r u s

t e

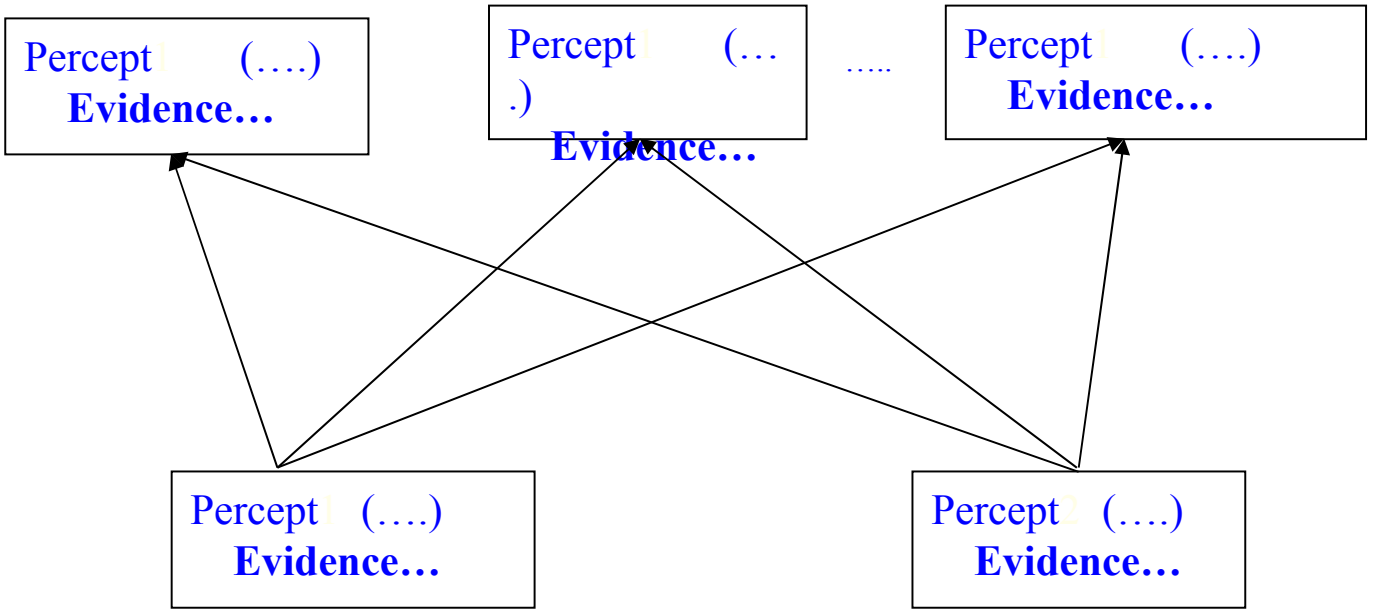
r u s

p e

p/t e

r u s

A set of evolving partial interpretations



- Why should we believe that it is possible to develop a system to support historians keep track of the set of evolving interpretations?

  Not surprisingly, we have built a system to support medical decision making

- What are some of the necessary ingredients?

  A suitable ontology, a GUI, …

What knowledge will populate our ontology?

- The expected shapes of letters, as written on stylus tablets (etc). This will, in turn, be defined in terms of strokes and junctions
- Certain sequences of letters, corresponding to the most common digrams, trigrams, and word endings
- A lexicon of words that are expected to feature in the document, perhaps with grammatical variations (endings, cases, tenses, …)
- An additional lexicon of proper nouns that refer, for example to places, regions, tribes, …
- …

The fundamental problem we face is that historians know a vast amount, and, like all experts, they mobilise their knowledge effortlessly, and try as they may, introspection is not a reliable source of what is mobilised!!

Historians know too much

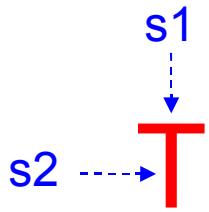
- This set of thin line fragments instantiates a single stroke of the stylus
- This set of stroke instantiates a character which may be interpreted as : in the case it is interpreted as , this stroke is the downstroke at left, this stroke is …; in the case it is interpreted as this stroke is …. Etc
- This sequence of letters may be interpreted as the word “soldier”, “bracelet”, ….
- This sequence of word percepts may be interpreted as a noun syntactic structure *adjective-qualifying-noun*, as in big red rooster;
- This sequence of word percepts may be interpreted, semantically, as “favourite ???? dress”, where ???? could be “mother’s”, “purple”, “M&S”,…
- It could be the ***batavians*** because there is evidence that members of this tribe were conscripted to this region and they tended to use a lancea to kill their opponents and we seem to have read batavians elsewhere on the document
- This piece of weaponry must have been brought to the archaeological site later because it was not used in this region before 400AD

Alternatives from Vindolanda*

*From Melissa Terras *From Image to Interpretation*

Representations of letter shapes

- Characters are composed of *strokes* and *junctions*
- A *stroke* has a qualitative *orientation* (horizontal, vertical, forward and backward diagonal); and a *relative position* (left-of, right-of, above, below). There may be a set of orientations.
- A *junction* has a type (end-end, end-middle); and both have an associated *angle* of intersection (the lower for end-middle)



The ideal case, rarely, if ever, seen!!

T *isa* Character

has-strokes

has-junctions

{s1,s2}

Orientation: vertical

Position: below s2

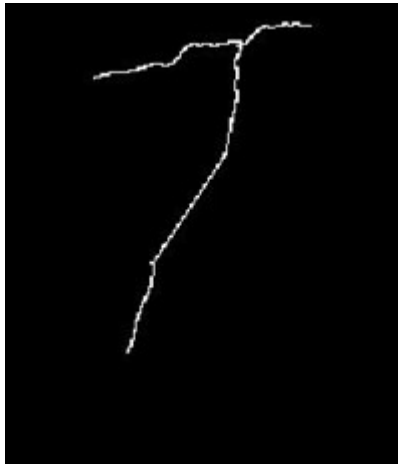
s1 *end-middle s2*

angle $\pi / 2$

Orientation: horizontal

Position: above s1

recognition



entity

Has-strokes

Has-junction:

u end-middle v

Angle : $\pi / 2$

u,v

Orientation: {vertical, forward diagonal}

Below: v

Orientation: horizontal

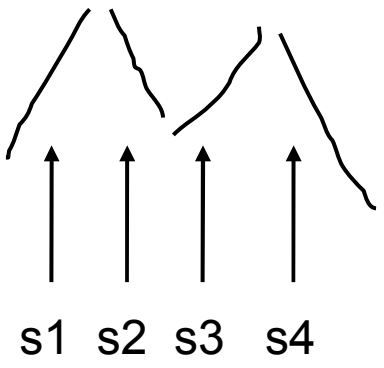
Above: u

It is trivial to see this entity as an instance of a T, by matching *u* to *s1*, and *v* to *s2* and interpreting *u* as vertical

Example recognitions

q u e m

Ambiguous readings:
the interpretation tree



start

s1=left of A

s1=left E

s1=left H

s1=left M

s1=left P

s1 isnt in this character

s2 right of A (weak)

s2 2 left M

s2=right P

s3 2 right M

s4 right M

We can detect alternative readings, and explain why they are alternative, and assign likelihoods

Next steps

1. Couple the character interpreter to Ségolène's strokelet proposer
2. Extend the algorithm to other types of text/document
3. Extend to incorporate additional knowledge
   - Bigram/trigram frequencies
   - Latin words