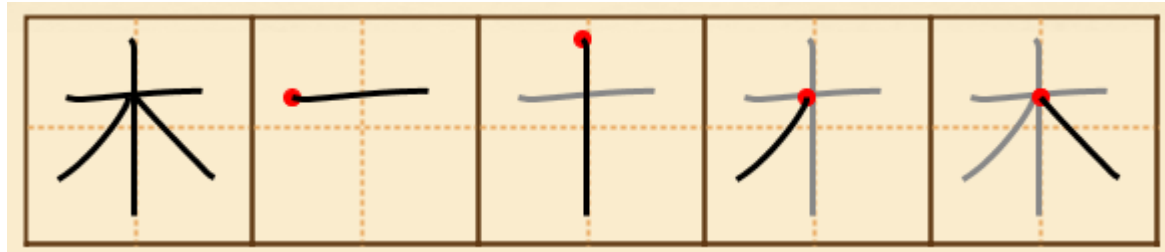


ONLINE CHINESE CHARACTER RECOGNITION

HMM VS BAYES NET

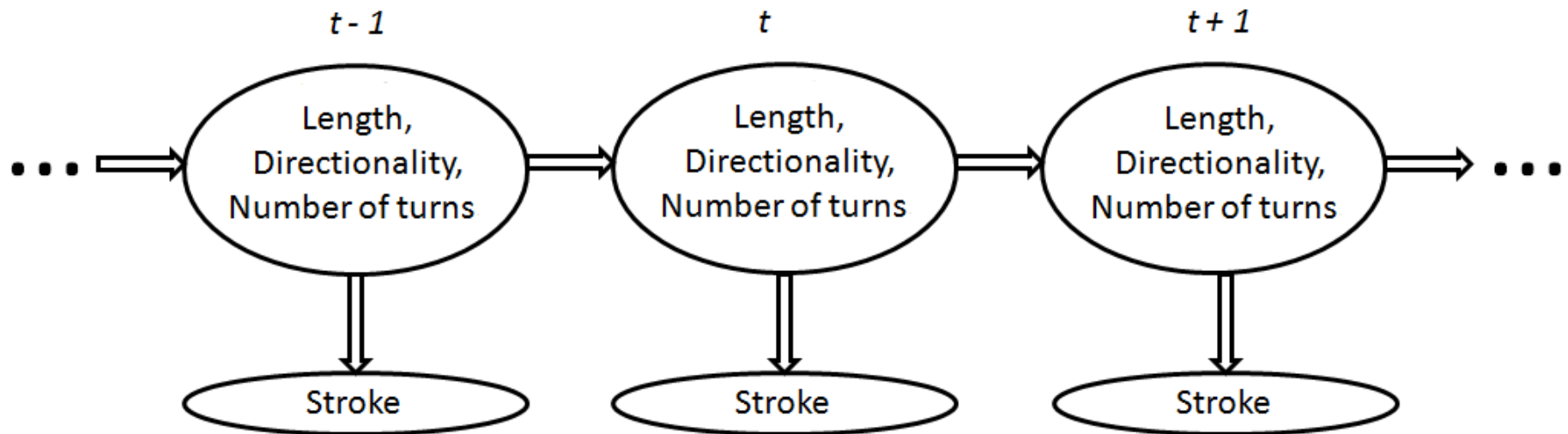
CHINESE CHARACTER RECOGNITION



- Number of strokes
- Stroke order
- Stroke direction
- Stroke shape
- Stroke length
- Stroke position

PRIOR WORK

- Hidden Markov Model with character strokes as the hidden variables



SIMPLIFICATIONS

- **Online, not OCR**
- **Small subset of dictionary**
- **User inputs strokes in the correct order**
- **User lifts pen after every stroke**
- **Training data is fully labeled**

NEW APPROACH

- **Improve recognition on poorly recognized strokes**
- **Test tolerance of model to removing simplification (stroke order correctness)**
- **Naïve (hybrid) Bayes classifier to recognize the strokes**
- **Same variables: length, directionality, number of turns**

INITIAL RESULTS

Stroke	# Correct	Wrong Length /Other	Percent Correct	# Correct	Wrong Length/ Other	Percent Correct
Long L-to-R	85	5/10	85%	83	17/0	83%
Short L-to-R	33	9/8	66%	49	1/0	98%
Short T-to-B	9	1/5	60%	14	0/1	93%
Short TL-to-BR	25	0/0	100%	24	0/1	96%
Long T-to-B	70	0/0	100%	69	0/1	99%

ASPIRATIONS VS REALITY & LIMITATIONS

- What I actually wanted to do is really big!
- Changing simplifications like whether or not the user picks up the pen makes labeling much harder
- Ideal model is probably a hybrid of the two approaches
- With a more complex dictionary, relative stroke positioning/points of intersection become very important, eg:

石 右

QUESTIONS?

INPUT DATA

...

```
<point id="e459eb54-b05f-4c2b-854c-5b10d9eb623c" name="point"
time="1240728086895" pressure="25" y="990" x="480"/>
```

...

```
<shape id="2ec3e273-208b-41fd-b6fd-af75e79b0098" name="substroke"
time="1240728087323" y="875" x="480" source="ConverterJnt"
end="8f6e79a8-5ae0-4546-8344-65b84653bd4e" start="e459eb54-b05f-
4c2b-854c-5b10d9eb623c" raster="CopyPen" penHeight="1"
penWidth="53" penTip="Ball" width="1343" height="115" color="-
16777216" type="substroke">
```

```
<arg type="point">291721e3-1c01-46df-9614-b303ddfb552b</arg>
```

...

```
</shape>
```