A New View of the
Output from Word Recognition
Dr. Marc-Peter Schambach, Siemens AG
A New View of the Output from Word Recognition

Introduction

My standard example:

Hamburg

Hamburg: 95%
Homburg: 92%

What does the output of word recognition mean?

- Probability
- Confidence
- A number
- ???
A New View of the Output from Word Recognition

Overview

- Word recognition
- Motivation for my work
- Goals

- Definition of recognition scores
- Estimation of recognition scores
- Performance metrics
- Experiments and results
A New View of the Output from Word Recognition

Word Recognition (Traditional Interface)

- Hamburg → 87%
- Hameln → 8%
- reject → 5%

- One Confidence value for each result
- Confidences are probabilities (Σ=100%)
A New View of the Output from Word Recognition

Motivation

Deficiencies

- Distinction of recognition tasks

**Reference**

**Nominal**

**Lists**

- Gummersbach
- Gundelfingen
- Hagen
- Halberstadt
- Halle
- Hamburg
- Hameln
- Hanau
- Hannover
- Heidelberg
- Helgoland
- Hennef

**Regular expressions**

\([0-9]\{1,5\}\)

**Language models**

(n-gram)
A New View of the Output from Word Recognition

Motivation

Deficiencies

- Distinction of recognition tasks
- No quantification of redundancy effects
A New View of the Output from Word Recognition

Motivation

Deficiencies

- Distinction of recognition tasks
- No quantification of redundancy effects
- Result is dependent on reference patterns
- Difficult comparability of results
- No interpretation of result quality
- Difficult tuning of reject behavior
- Difficult tuning of out-of-vocabulary behavior
Goals

- Improve result combination
- Improve post-processing
- Improve reject / out-of-vocabulary behavior

Requirements

- Results are independent of reference patterns
A New View of the Output from Word Recognition
Writing and recognition

Writer intention
Misspelling
New York
Recognition result
Correction
Nominal hypothesis

Nominal image label
Writing Style
Written word

Scanned image
Nominal image label
Classification

New Yak
A New View of the Output from Word Recognition
Two output values

Writer intention
Misspelling
New York
Recognition result
Correction
Similarity
New Yak
Nominal hypothesis
Quality

Nominal image label
Written word
Writing Style
Classification
Scanned image
Correlation

A New View of the Output from Word Recognition Interface: Quality and similarity

set_image(image)

22083 Hamburg

set_Pattern(regexp)

[0-9]{1,5}\w[a-zA-Z]*

get_Quality()

0.95

get_Similarity(name)

22083 Hamburg → 0.96
22083 Homburg → 0.95
22085 Hamburg → 0.93
...

A New View of the Output from Word Recognition

Definition of recognition scores

Nominal image label: \( w_{\text{nom}} \)  
Nominal hypothesis: \( w_{\text{hyp}} \)  
Result alternative: \( w_i \)  
String distance: \( d(\cdot) \)  

Quality: \( q_{\text{ref}} = d(w_{\text{nom}}, w_{\text{hyp}}) \)
Similarity: \( s_{\text{ref}} = d(w_{\text{nom}}, w_i) \)
A New View of the Output from Word Recognition
Longest common subsequence

\[
\text{Similarity} = \frac{7}{(18+12)/2}
\]
A New View of the Output from Word Recognition
Estimation of recognition scores

Character based word recognition

- **Quality**
  \[ q_{\text{est}} = \text{avg}(c_{i}^{\text{nom}}) \]

- **Similarity**
  \[ s_{\text{est}} = 1 - \frac{1}{n} \left( \sum_{\text{sub}} (c_{i}^{\text{nom}} - c_{i}^{\text{ref}}) + \sum_{\text{ins}} k_{\text{ins}} + \sum_{\text{del}} k_{\text{del}} \right) \]

Script word recognition (HMM)

- **Quality**
  \[ q_{\text{est}} = \text{map}(p_{\text{joker}}^{\text{ref}}) \]

- **Similarity**
  \[ s_{\text{est}} = \text{map}(p_{\text{ref}} - p_{\text{joker}}^{\text{ref}}) \]
A New View of the Output from Word Recognition

Performance metrics

Provide **two** metrics:

- How well does the recognition engine **perform**?
- How well does the recognition engine **predict** its reliability?

**Performance:**

- Correctness of similarity
  \[ C_{\text{sim}} = 1 - (s_{\text{est}} - s_{\text{ref}})^2 \]
- Average correctness (alternatives \( i \))
  \[ P_n = \text{avg}(C^i_{\text{sim}}) \]
- Reader performance (test cases \( n \))
  \[ P = \text{avg}(P_n) \]

**Predictability**

- Correctness of quality:
  \[ Q = 1 - \text{avg}(P_n - q_n) \]
A New View of the Output from Word Recognition Experiments /1
A New View of the Output from Word Recognition Experiments /2

Word recognition: Performance and predictability (HMM, us_cy, mapped)

Performance

Predictability

Data points (images)
Linear approximation
A New View of the Output from Word Recognition

Summary

What is new?

- Definition of Interface (Quality and similarity)
- Definition of Metrics (Performance and predictability)
- Estimation (Character based, HMM recognition)
- Evaluation (HMM recognition)

Next steps

- New estimation methods
- Adapt post-processing
A New View of the Output from Word Recognition

Thank you!