Finite State Translation:
Word and Phrase-based Models

Trevor Cohn
University of Sheffield
Melbourne, 2011

Based on slides by Adam Lopez, Philipp Koehn, Kevin Knight, Chris Callison-Burch, David Chiang and Phil Blunsom
Course Content

• Day 1: Introduction
  • historical context
  • overview of statistical MT
  • n-gram language models

• Day 2: Word-based approaches
  • learning translations using EM
  • alignment of sentences, words and phrases
Course Content

• Day 3: Decoding
  • Finite state ‘phrase based’ decoding
  • Grammar based decoding

• Day 4: Acquiring translation rules
  • Learning richer models without word-alignments

• Day 5: Training and Evaluation
  • Discriminative training of feature-based models
  • Human and Automatic Evaluation
Quick Recap

Develop a statistical model of translation that can be learned from data and used to predict the correct English translation of new Chinese sentences.
Quick Recap

- *Minimally*, our model must account for:
  - Lexical ambiguity.
  - One-to-many translation.
  - Many-to-many translation.
  - Untranslated words.
  - Word reordering.
Quick Recap

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。

However, the sky remained clear under the strong north wind.
Quick Recap
Quick Recap

training data
(parallel text)
Quick Recap

training data (parallel text) → learner
Quick Recap

training data (parallel text) → learner → model
Quick Recap

training data
(parallel text)

learner

model

decoder

聯合國 安全 理事會 的
五个 常任 理事 国都
Quick Recap

training data
(parallel text)

learner

model

decoder

聯合国 安全 理事会 的
五个 常任 理事 国都

However, the sky remained clear under the strong north wind.
Quick Recap

\[ p(English|Chinese) \sim p(English) \times p(Chinese|English) \]
Quick Recap

\[ p(English|Chinese) \sim p(English) \times p(Chinese|English) \]

Language Model

What is the probability of an English sentence?
Quick Recap

\[ p(English|\text{Chinese}) \sim \]

\[ p(English) \times p(\text{Chinese}|English) \]

Language Model
What is the probability of an English sentence?

Translation Model
What is the probability of a Chinese sentence, given a particular English sentence?
Translation Models

What is a good story about how a Chinese sentence came into being, given that we already have an English sentence?
Translation Models

What is a good story about how a Chinese sentence came into being, given that we already have an English sentence?

Note: in this example I’ll show you an English sentence, conditioned on a Chinese sentence. Note that we can apply the same technique in either direction.
Translation Models

雖然北風呼嘯，但天空依然十分清澈。

\[ p(English|Chinese) \]
Translation Models

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。

\[ p(\text{English}|\text{Chinese}) \]
Translation Models

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。

However, the sky remained clear under the strong north wind.

$p(English|Chinese)$
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。

However, the sky remained clear under the strong north wind.
Translation Models

Although north wind howls, but sky still very clear.

However, the sky remained clear under the strong north wind.

$p(English|Chinese)$?
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Although north wind howls, but sky still very clear.
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Although north wind howls, but sky still very clear.

雖然北風呼嘯，但天空依然十分清澈。
IBM Model 1

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Although north wind howls, but sky still very clear.

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
IBM Model 1

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Although north wind howls, but sky still very clear.

However
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
IBM Model 1

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。ε

However
IBM Model 1

Although north wind howls, but sky still very clear.

However,

Thursday, 24 November 2011
IBM Model 1

Although north wind howls, but sky still very clear.

However,
Although north wind howls, but sky still very clear.

However, the
Although north wind howls, but sky still very clear.

However, the sky remained clear under the strong north wind.
IBM Model 4

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
IBM Model 4

Although north wind howls, but sky still very clear.

雖然 北 風 呼嘯，但 天空 依然 十分 清澈。
IBM Model 4

Although north wind howls, but sky still very clear.

雖然北風呼嘯，但天空依然十分清澈。

$\text{pf}(1|\text{雖然})$
Although north wind howls, but sky still very clear.

Although north wind howls, but sky still very clear.

IBM Model 4

Thursday, 24 November 2011
Although north wind howls, but sky still very clear.
Although north wind howls, but sky still very clear.
Although north wind howls, but sky still very clear.

IBM Model 4

Thursday, 24 November 2011
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Although north wind howls, but sky still very clear.

IBM Model 4

Thursday, 24 November 2011
Although north wind howls, but sky still very clear.

However
Although north wind howls, but sky still very clear.

However

$p_t(\text{However} | \text{虽然})$
Although north wind howls, but sky still very clear.

However north wind strong, the sky remained clear. under the
Although north wind howls, but sky still very clear.

Although north wind howls, but sky still very clear.

However north wind strong, the sky remained clear. under the
Although north wind howls, but sky still very clear.

However north wind strong, the sky remained clear. under the
Although north wind howls, but sky still very clear.

However north wind strong, the sky remained clear. under the

\[ p_d(0|\text{However}) \]
Although north wind howls, but sky still very clear.

However north wind strong, the sky remained clear. under the
Although north wind howls, but sky still very clear.

However north wind strong, the sky remained clear. under the

$p_d(8|\text{north})$
Although north wind howls, but sky still very clear.

However, north wind strong, the sky remained clear. Under the
Although north wind howls, but sky still very clear.

However, north wind strong, the sky remained clear under the strong north wind.
Although north wind howls, but sky still very clear.

Although north wind howls, but sky still very clear.

However, the sky remained clear under the strong north wind.

\[ p(English, alignment|Chinese) = \prod_{p_f} \prod_{p_t} \prod_{p_d} \]
However, the sky remained clear under the strong north wind.

\[ p(English, alignment|Chinese) = \prod^{p_f} \prod^{p_t} \prod^{p_d} \]
However, the sky remained clear under the strong north wind.

\[ p(English|Chinese) = \sum_{alignments} \prod_{pf} \prod_{pt} \prod_{pd} \]
The IBM Models
The IBM Models

- Fertility probabilities.
The IBM Models

• Fertility probabilities.

• Word translation probabilities.
The IBM Models

- Fertility probabilities.
- Word translation probabilities.
- Distortion probabilities.
The IBM Models

• Fertility probabilities.
• Word translation probabilities.
• Distortion probabilities.

Some problems:
• Weak reordering model -- output is not fluent.
• Many decisions -- many things can go wrong.
• Not symmetric, considers only many-t-1 but not one-to-many alignments
The IBM Models

• Fertility probabilities.
• Word translation probabilities.
• Distortion probabilities.
• Some problems:
  • Weak reordering model -- output is not fluent.
  • Many decisions -- many things can go wrong.
  • Not symmetric, considers only many-t-1 but not one-to-many alignments.
The IBM Models

- Fertility probabilities.
- Word translation probabilities.
- Distortion probabilities.

Some problems:

- Weak reordering model -- output is not fluent.
- Many decisions -- many things can go wrong.
- Not symmetric, considers only many-t-1 but not one-to-many alignments.
Training Model 1

Although north wind howls, but sky still very clear.

雖然北風呼嘯,但天空依然十分清澈。
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Training Model 1

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Although north wind howls, but sky still very clear.

雖然 北 風 呼嘯，但 天空 依然 十分 清澈。
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Although north wind howls, but sky still very clear.

雖然 北 風 呼嘯，但 天空 依然 十分 清澈。
Training Model 1

Although north wind howls, but sky still very clear.

Although north wind howls, but sky still very clear.

$p(\text{Chinese word position})$
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Although north wind howls, but sky still very clear.

Training Model 1

Although north wind howls, but sky still very clear.
Training Model 1

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。

However

\[ p(\text{English word} | \text{Chinese word}) \]
Training Model 1

Although north wind howls, but sky still very clear.

However
Training Model 1

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。

However
Although north wind howls, but sky still very clear.

Training Model 1

However,
Training Model 1

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。

However,
Training Model 1

Although north wind howls, but sky still very clear.

Although north wind howls, but sky still very clear.

Thursday, 24 November 2011
Training Model 1

Although north wind howls, but sky still very clear.

However, the sky remained clear under the strong north wind.
Training Model 1

Although north wind howls, but sky still very clear.

However, the sky remained clear under the strong north wind.

$p(\text{English length}|\text{Chinese length})$ observed
Although north wind howls, but sky still very clear.

However, the sky remained clear under the strong north wind.

$\hat{p}(\text{Chinese word position})$ uniform, no need to estimate.
Training Model 1

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。

However, the sky remained clear under the strong north wind.

\[ p(\text{English word}|\text{Chinese word}) \quad \text{unobserved!} \]
Training Model 1

Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。

the missing alignment is a \textit{latent variable}

However, the sky remained clear under the strong north wind.

\[ p(\text{English word}|\text{Chinese word}) \quad \text{unobserved!} \]
Likelihood Estimation for Model 1

Although north wind howls, but sky still very clear.

Parameters and alignments are both unknown.

However, the sky remained clear under the strong north wind.

\[ p(\text{English word}|\text{Chinese word}) \quad \text{unobserved!} \]
Likelihood Estimation for Model 1

Although north wind howls, but sky still very clear.

Parameters and alignments are both unknown.

If we knew the alignments, we could calculate the values of the parameters.

However, the sky remained clear under the strong north wind.

$p(\text{English word}|\text{Chinese word})$ unobserved!
Although north wind howls, but sky still very clear.

Parameters and alignments are both unknown.

If we knew the alignments, we could calculate the values of the parameters.

If we knew the parameters, we could calculate the alignments.

However, the sky remained clear under the strong north wind.

$p(English\ word|Chinese\ word)$ unobserved!
Although north wind howls, but sky still very clear.

Although north wind howls, but sky still very clear.

Parameters and alignments are both unknown.

If we knew the alignments, we could calculate the values of the parameters.

If we knew the parameters, we could calculate the alignments.

However, the sky remained clear under the strong north wind.

$p(\text{English word}|\text{Chinese word})$ unobserved!
The Plan: Bootstrapping

- Arbitrarily select a set of parameters (say, uniform).
- Calculate *expected counts* of the unseen events.
- Choose new parameters to maximize likelihood, using expected counts as proxy for observed counts.
- Iterate.
- Guaranteed that likelihood is monotonically nondecreasing.
Learning

la maison ... la maison bleu ... la fleur

the house ... the blue house ... the flower
Learning

la maison ... la maison bleu ... la fleur

the house ... the blue house ... the flower

Chicken and egg problem:
use Expectation Maximisation
Learning

la maison  ...  la maison bleu  ...  la fleur

the house  ...  the blue house  ...  the flower

Start with *uniform* model parameters
Learning

la maison ... la maison bleu ... la fleur
the house ... the blue house ... the flower

E-Step: Estimate alignments
Learning

la maison  ...  la maison bleu  ...  la fleur

the house  ...  the blue house  ...  the flower

M-Step: estimate the model parameters

the-la entry receives high weight
Learning

la maison  la maison bleu  la fleur

the house  the blue house  the flower

E-Step: reestimate alignments
Learning

la maison ... la maison bleu ... la fleur

the house ... the blue house ... the flower

Continue iterating convergence on inherent hidden alignment
EM algorithm

• *Expectation-Step*: applies the model to the data
  • uses model to calculate the probabilities of all different values of hidden structure (alignments)

• *Maximisation-Step*: estimate model from data
  • treat the (soft) assignments of hidden values as fact
  • estimate model parameters (collect counts and normalise)

• Iterate these steps until convergence
IBM Model 1

\[ p(e, a | f) = \frac{\epsilon}{(l + 1)^m} \prod_{j=1}^{m} t(e_j | f_{a_j}) \]

- Where
  - the foreign sentence is \( f = f_1, f_2, \ldots, f_l \)
  - the English sentence is \( e = e_1, e_2, \ldots, e_m \)
  - \( \epsilon \) is a normalisation constant
  - \( t(e | f) \) are the word translation probabilities
Model 1: E-Step

- Need the *expectations* over the alignments

\[
p(a|e, f) = \frac{p(a,e|f)}{p(e|f)} = \prod_{j=1}^{m} \frac{t(e_j | f_{a_j})}{\sum_{i=0}^{l} t(e_j | f_i)}
\]

The probabilities of each alignment \(a_j\) are independent. They’re just the normalised translation probabilities

\(f_0 = \varepsilon\) is special ‘null’ symbol
Model 1: M-step

- After collecting the expected counts for each alignment position in E-step

- Just count up the translation pairs

\[
t(e|f) = \frac{c(e,f)}{\sum_{e'} c(e',f)}
\]

\[
c(e, f) = \sum_{(e',f')} \sum_{j=1}^{m} \sum_{i=0}^{l} \delta(e'_j, e) \delta(f'_i, f) \frac{t(e_j|f_i)}{\sum_{k=0}^{m} t(e_j|f_k)}
\]

frequency of seeing seen \((e, f)\) aligned together

weighted by partial counts from expectation step
IBM Models 2-5

- Model 1 naive in many ways:
  - M2 adds absolute *reordering* model
  - M3 adds *fertility* model
  - M4 changes to relative *reordering* model
  - M5 fixes *deficiency*
- Each getting increasingly complex mathematically and intractable to solve
Why so many models?

• Aim is to learn model 5

• However, complex models are difficult to train
  • Inference is intractable
  • Expressive models have many local optima

• Solution: use simple models initialise complex models
  • M1 - M2 - HMM - M3 - M4 - M5 ...
  • Initialises with a good guess of the optimal solution
Model 2

• Translations typically follow a similar word order, but M1 doesn’t model this

• Word 1 in Chinese should probably align to word 1 in English...

• M2 adds a new term to M1, based on the alignment position

\[ A(a_j | j, l, m) \]

• This way the model can learn to prefer points near the diagonal \( a_j \sim j \) over off-diagonal entries
HMM

- Words rarely move independently, but in groups
- True process of translation uses larger units than words
- HMM extends M1 with following term over adjacent pairs of alignments
  \[ A(a_j | a_{j-1}, m) \]
- Requires dynamic programming in E-step

See Vogel 96 for details
Alignment Indices
Symmetrization

- Training $p(e \mid f)$ and $p(f \mid e)$ learns quite different alignments

- Model assymmetric, can only many-to-1 but not 1-to-many

- A rough-and-ready solution: take their intersection, union etc

- More on this later for phrase-based and grammar-based models...
Evaluating Alignments

- A number of testing datasets have been made
  - Humans annotated the alignments for each pair of sentences
  - Can measure precision, recall, F1, AER etc of system predictions

Resources

- [http://www.cse.unt.edu/~rada/wpt05/](http://www.cse.unt.edu/~rada/wpt05/)
Supervised Alignment

How about we use part of the word aligned test set for training?

No longer merely hoping unsupervised induction finds good alignments... **But** there’s not much data!
Supervised Alignment

- Could learn IBM Model 4 $\equiv$ a single M-step
- Or, better, could develop feature-based model
- using the predictions of Model 1-5
- using parts-of-speech, syntax
- using a bilingual dictionary, etc.
- More difficult estimation but only done once
- a big improvement in the alignments and end translations
CRF Alignment

\[ p_\Lambda(a|e, f) = \frac{\exp \sum_t \sum_k \lambda_k h_k(t, a_{t-1}, a_t, e, f)}{Z_\Lambda(e, f)} \]

- No longer a ‘generative story’
- Doesn’t describe english e, only the alignment a
- Allows the use of vast numbers of features, h_k, and associated parameters, \( \lambda_k \)
- Very powerful model...
- Caveat: only applicable in supervised setting
Effect of Markov Feature
Results

Alignment Results:

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>French → English</td>
<td>0.97</td>
<td>0.86</td>
<td>0.91</td>
</tr>
<tr>
<td>French ← English</td>
<td>0.98</td>
<td>0.83</td>
<td>0.91</td>
</tr>
<tr>
<td>French ↔ English</td>
<td>0.96</td>
<td>0.90</td>
<td>0.93</td>
</tr>
<tr>
<td>French → English (+ibm model4)</td>
<td>0.98</td>
<td>0.88</td>
<td>0.93</td>
</tr>
<tr>
<td>French ← English (+ibm model4)</td>
<td>0.98</td>
<td>0.87</td>
<td>0.93</td>
</tr>
<tr>
<td>French ↔ English (+ibm model4)</td>
<td>0.98</td>
<td>0.91</td>
<td>0.95</td>
</tr>
<tr>
<td>GIZA++ (French ↔ English)</td>
<td>0.87</td>
<td><strong>0.95</strong></td>
<td>0.91</td>
</tr>
</tbody>
</table>

Translation Results:

<table>
<thead>
<tr>
<th></th>
<th>unigram</th>
<th>bigram</th>
<th>trigram</th>
<th>quadgram</th>
<th>BLEU</th>
</tr>
</thead>
<tbody>
<tr>
<td>F ↔ E (+ibm m4)</td>
<td>65.1</td>
<td>40.3</td>
<td>27.8</td>
<td>19.6</td>
<td><strong>34.54</strong></td>
</tr>
<tr>
<td>GIZA++ (F ↔ E)</td>
<td><strong>65.4</strong></td>
<td>40.0</td>
<td>27.2</td>
<td>18.9</td>
<td>34.05</td>
</tr>
</tbody>
</table>

Thursday, 24 November 2011
Uses for Alignments

- Machine translation
- core part of word-based models
- used to define *translation units* in phrase-based and syntax-based MT and for backoff
- Can *project* annotations between languages
  - tagging, parsing, named entity, coreference...
- Can *pivot* via foreign language to create thesauri
IBM Model 4 Again

Although north wind howls, but sky still very clear.

However, the sky remained clear under the strong north wind.
IBM Model 4 Again

Although north wind howls, but sky still very clear.

However, the sky remained clear under the strong north wind.
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。
Phrase-based Models

Although north wind howls, but sky still very clear.
Phrase-based Models

Although north wind howls, but sky still very clear.

雖然北風呼嘯，但天空依然十分清澈。
Although north wind howls, but sky still very clear.

However the strong north wind, the sky remained clear under.
Although north wind howls, but sky still very clear.
Although north wind howls, but the sky still very clear.

However, the strong north wind, the sky remained clear under.
Although north wind howls, but sky still very clear.

However, the strong north wind, the sky remained clear under.

Although 北风 呼啸, 但 天空 依然 十分 清澈.

虽然 北风 呼啸, 但 天空 依然 十分 清澈。

Thursday, 24 November 2011
Although north wind howls, but sky still very clear.

虽然北风呼啸，但天空依然十分清澈。

$$p(\text{English, alignment}|\text{Chinese}) = p(\text{segmentation}) \cdot p(\text{translations}) \cdot p(\text{reorderings})$$
Phrase-based Models
Phrase-based Models

- Segmentation probabilities.
Phrase-based Models

- Segmentation probabilities.
- Phrase translation probabilities.
Phrase-based Models

- Segmentation probabilities.
- Phrase translation probabilities.
- Distortion probabilities.
Phrase-based Models

- Segmentation probabilities.
- Phrase translation probabilities.
- Distortion probabilities.

Some problems:

- Weak reordering model -- output is not fluent.
- Many decisions -- many things can go wrong.
Phrase-based Models

- Segmentation probabilities.
- Phrase translation probabilities.
- Distortion probabilities.

Some problems:
- Weak reordering model -- output is not fluent.
- Many decisions -- many things can go wrong.
Phrase-based Models

- Segmentation probabilities.
- Phrase translation probabilities.
- Distortion probabilities.
- Some problems:
  - Weak reordering model -- output is not fluent.
  - Many decisions -- many things can go wrong.
Phrase-based Models
Phrase-based Models

• Phrase-based Models are dumb.
Phrase-based Models

• Phrase-based Models are dumb.
• But they are widely regarded as state-of-the-art.
Phrase-based Models

- Phrase-based Models are dumb.
- But they are widely regarded as state-of-the-art.
- Why? Simple models are easier to learn and deploy.
Phrase-based Models

• Phrase-based Models are dumb.

• But they are widely regarded as state-of-the-art.

• Why? Simple models are easier to learn and deploy.

• With lots of training data, will have seen big chunks of the test sentence - just need to join together their translations.
Phrase-based Models

- Phrase-based Models are dumb.
- But they are widely regarded as state-of-the-art.
- Why? Simple models are easier to learn and deploy.
- With lots of training data, will have seen big chunks of the test sentence - just need to join together their translations.
- Need proof? Google uses a phrase-based model.
However, the sky remained clear under the strong north wind.
Overview

However, the sky remained clear under the strong north wind.

Thursday, 24 November 2011
Decoding

Probability models allow us to make predictions: Given a particular Chinese sentence, what is the most probable English sentence corresponding to it?
Decoding

Probability models allow us to make predictions: Given a particular Chinese sentence, what is the most probable English sentence corresponding to it?

In math:

$$\arg\max_{\text{English}} p(\text{English}|\text{Chinese})$$
Decoding

Probability models allow us to make predictions: Given a particular Chinese sentence, what is the most probable English sentence corresponding to it?

In math:

$$\argmax_{English} p(English|Chinese)$$

problem: there are a lot of English sentences to choose from!
北风呼啸。
segmentations
substitutions
permutations
北 风 呼啸。

segmentations  $O(2^n)$
substitutions
permutations
segmentations $O(2^n)$
substitutions $O(5^n)$
permutations
segmentations $O(2^n)$
substitutions $O(5^n)$
permutations $O(n!)$
Key Idea
Key Idea
Key Idea
Key Idea
Key Idea
Key Idea

Dynamic Programming
虽然北风呼啸，但天空依然十分清澈。
虽然北风呼啸，但天空依然十分清澈。
Thursday, 24 November 2011

START Although

虽然北风呼啸，但天空依然十分清澈。

crystal clear

虽然北风呼啸，但天空依然十分清澈。
Although crystal clear

虽然北风呼啸，但天空依然十分清澈。

However

虽然北风呼啸，但天空依然十分清澈。
wind screamed

wind shrieked

north wind

雖然北風呼嘯，但天空依然十分清澈。

雖然北風呼嘯，但天空依然十分清澈。
wind screamed

雖然北風呼啸，但天空依然十分清澈。

wind shrieked

雖然北風呼啸，但天空依然十分清澈。

north wind

虽然北风呼啸，但天空依然十分清澈。
shrieked,

the sky

yet

Thursday, 24 November 2011
shrieked,

the sky

yet

Thursday, 24 November 2011
sky,

虽然北风呼啸，但天空依然十分清澈。
虽然北风呼啸，但天空依然十分清澈。
still quite clear.

blue.

虽然北风呼啸，但天空依然十分清澈。
still quite clear.

blue.

clear.

blue.

Thursday, 24 November 2011
Although the northern wind shrieked across the sky, but was still very clear.

虽然北风呼啸，但天空依然十分清澈。
the sky

虽然北风呼啸，但天空依然十分清澈。
number of vertices: $O(2^n)$

the sky

雖然北風呼嘯，但天空依然十分清澈。
number of vertices: $O(2^n)$

the sky

雖然北風呼嘯，但天空依然十分清澈。

\[ d = 4 \]

window
number of vertices: $O(2^n)$

the sky

虽然北风呼啸，但天空依然十分清澈。

outside window to left: covered

d = 4

outside window to right: uncovered
number of vertices: $O(n^{2^d})$

the sky

虽然北风呼啸，但天空依然十分清澈。

outside window
to left: covered
d = 4
window
to right: uncovered

Thursday, 24 November 2011
Observations

- The lattice describing the set of all possible translations is a *weighted finite state automaton*.
- So is the language model.
- Since regular languages are closed under intersection, we can intersect the devices and run shortest path graph algorithms.
- Taking their intersection is equivalent to computing the probability under Bayes’ rule.
Recap

• Probability theory enables us to learn from data.
Recap

- Probability theory enables us to learn from data.
- Very simple models get us pretty far!
Recap

• Probability theory enables us to learn from data.
• Very simple models get us pretty far!
• There’s no data like more data.
Recap

• Probability theory enables us to learn from data.
• Very simple models get us pretty far!
• There’s no data like more data.
• Word-based models follow our basic intuitions.
Recap

- Probability theory enables us to learn from data.
- Very simple models get us pretty far!
- There’s no data like more data.
- Word-based models follow our basic intuitions.
- Phrase-based models are dumb but effective.
Recap

- Probability theory enables us to learn from data.
- Very simple models get us pretty far!
- There’s no data like more data.
- Word-based models follow our basic intuitions.
- Phrase-based models are dumb but effective.
- *All* of these models are weighted regular languages.
Recap

• Probability theory enables us to learn from data.
• Very simple models get us pretty far!
• There’s no data like more data.
• Word-based models follow our basic intuitions.
• Phrase-based models are dumb but effective.
• All of these models are weighted regular languages.
• Is this the best we can do? Find out tomorrow!
Implementations

- Phrase-based Translation
  - Moses -- [www.statmt.org/moses/](http://www.statmt.org/moses/)
  - cdec -- [www.cdec-decoder.org](http://www.cdec-decoder.org)

- Language models
  - RandLM -- [sourceforge.net/projects/randlm](http://sourceforge.net/projects/randlm)
ارتفاع عجز الميزان التجاري الأردني

قيمة الواردات الأردنية بلغت 7.39 مليارات دولار في النصف الأول من العام (الجزيرة نت)

أفادت بيانات رسمية بأن العجز في الميزان التجاري الأردني ارتفع في النصف الأول من العام الحالي بنسبة 18.1% نتيجة زيادة حجم الواردات مقابل الصادرات، كما تراجع حجم القروض الائتمانية التي قدمتها البنوك الأردنية بنسبة 11% خلال الفترة نفسها.

وقالت بيانات دائرة الإحصاءات العامة نشرتها اليوم الاثنين إن قيمة العجز في الميزان التجاري بلغت 2.79 مليار دينار أردني (3.94 مليار دولار أمريكي).

وأشارت البيانات إلى ارتفاع حجم الصادرات خلال النصف الأول من العام بنسبة 16.6% حيث بلغت 2.05 مليار دينار (2.91 مليار دولار)، كما ارتفعت قيمة الواردات بنسبة 11.7% لتبلغ 5.23 مليارات دينار (7.39 مليارات دولار).
Value of Jordanian imports amounted to $7.39 billion dollars in the first half of the year (Island Net).

According to official statements that the trade balance deficit rose Jordan in the first half of this year by 18.1% due to increased volume of imports versus exports, as the decline in the volume of credit provided by banks of Jordan by 11% during the same period.

The data for the Department of Statistics published on Monday that the value of the trade balance deficit amounted to $2.79 billion Jordanian dinars ($3.94 billion U.S. dollars). The data indicated the high volume of exports during the first half of the year by 16.6%, reaching $2.05 billion dinars ($2.91 billion dollars), as imports rose by 11.7% to $5.23 billion dinars ($7.39 billion dollars).

The export value increased significantly for the countries of the Greater Arab Free Trade, including Saudi Arabia, and the countries of the free trade agreement for North America, including the United States, as well as non-Arab Asian
จัดระเบียบปูย...เพิ่มคุณภาพ

วันที่ 17 ธันวาคม 2553 เวลา 5:00 น.

เนื้อหาเรื่อง:

ปัญจุบันมีผู้ประกอบการที่ทำธุรกิจเกี่ยวกับปูยมากขึ้น ทั้งเป็นผู้นำเข้า ผู้ผลิต และผู้จ้างเหมา จึงจำเป็นต้องมีการจัดระบบปูยเพื่อตรวจสอบและรับรองคุณภาพจากกรมวิชาการเกษตร เพื่อแก้ปัญหา ปูยปลอม หรือ ปูยต่อคุณภาพ เป็นการปกป้องผลประโยชน์ให้กับพื้นที่เกษตรกร

นายสมชาย ชายแย้ม รองศาสตราจารย์ในอธิบดีกรมวิชาการเกษตรกล่าวว่า ปัญจุบันมีการดำเนินการธุรกิจเกี่ยวกับปูยที่ไม่ถูกต้องสร้างความเสียหายให้กับเกษตรกร ทั้งนี้ปูยที่จ้างเหมาในท้องตลาดมี 2 ลักษณะคือ 1. ปูยสินค้า เบลนด์ (Bulk Blending) คือปูยที่นำมาผสมผสานกันเพื่อให้ได้สูตร และจุดเด่นที่เกิดขึ้นที่เป็นประโยชน์แก่เกษตรกร

แล้วนำ มาจำหน่าย 2. ปูยซิมเดียว คือ ปูยที่เป็นเม็ดแล้วส่งให้เกษตรกรสามารถตรวจสอบได้ตามมาตรฐานของกรมวิชาการเกษตรที่เกี่ยวข้อง ทั้งนี้ถ้าใช้ในนำเข้า ปูยพอสพรีส ปูยโฟเทลเข้ม
Organize ... enhance the quality fertilizer.

There are currently entrepreneurs doing business with more fertilizer. One of the leading producer and marketer. Is required to verify the registration and certification from the Department of Agriculture To resolve fake fertilizers or fertilizers second Protect the benefits to farmers brothers.

Mr. Somchai Channarong Kun-General, Department of Agriculture said the current business operations on the wrong fertilizer damage to the farmers. The commercial fertilizer has two characteristics: 1. Fertilizer Bulk Blend (Bulk Blending) is the parent fertilizer fertilizer combinations to get the formula Nutrients as a registered already be sold 2. Polycultural fertilizer is fertilizer tablets and statues of nutrient standards for registration as the nitrogen fertilizer such as potassium, phosphorus fertilizer.