Statistical Phrase-Based Speech Translation

Bill Byrne

Cambridge University Engineering Department

Machine Intelligence Laboratory

22 March 2006

Statistical Phrase-Based Speech Translation [1]

Objective: Tight Integration of a Phrase-Based SMT System with Speech-to-Text Systems

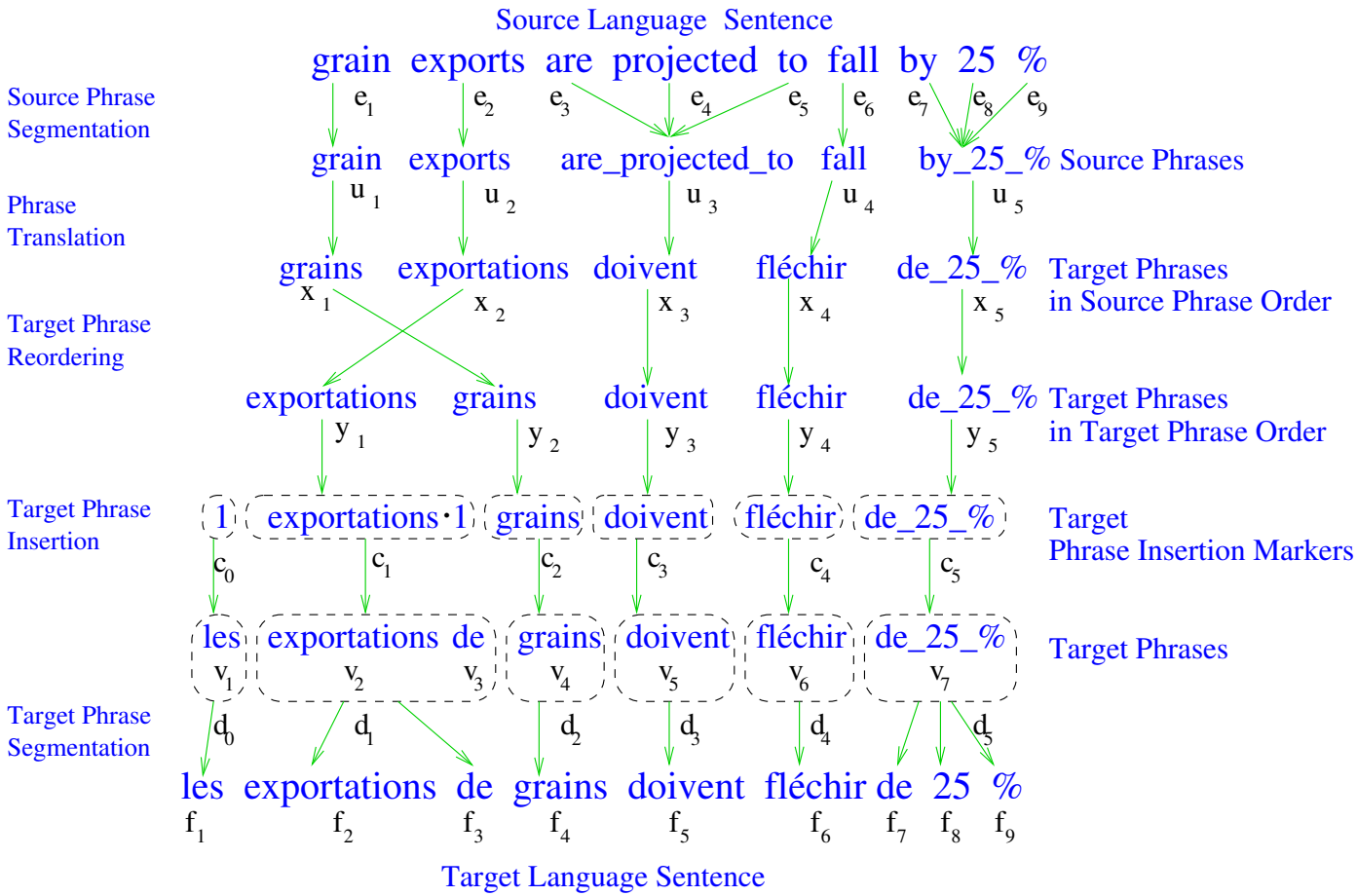
- Unified statistical modeling framework

- Straightforward translation from speech without extensive reformation of the underlying statistical models or the STT or SMT systems themselves

Problem: How to translate STT lattices ?

[1] L. Mathias and W. Byrne. *Statistical Phrase-Based Speech Translation*. ICASSP 2006

TTM : Translation with Moving Target Phrase Order [2]



- Transformations via stochastic models implemented as WFSTs
- Implementation is direct, using standard WFST operations

[2] S. Kumar and W. Byrne, Local phrase reordering models for statistical machine translation, HLT-EMNLP, 2005.

Statistical Modeling for Text Translation

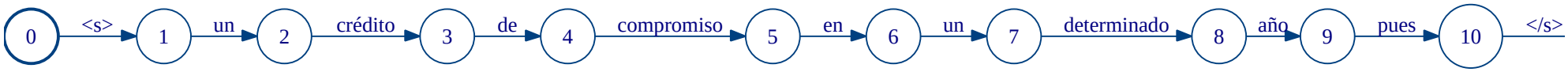
Foreign Sentence		Foreign Phrases		English Phrases		English Sentence
F	$\longleftarrow$	**V**	$\longleftarrow$	**U**	$\longleftarrow$	**E**
Models $\longrightarrow$	$P(\mathbf{F}\|\mathbf{V})$		$P(\mathbf{V}\|\mathbf{U})$		$P(\mathbf{U}\|\mathbf{E})$	$P(\mathbf{E})$
FSMs $\longrightarrow$	Ω		Φ		W	G
	Target Phrase Mapping		Phrase Translation, Reordering		Source Phrase Mapping	Source Language Model

Ω : All foreign phrase sequences that could have generated the foreign text

The translation system effectively translates phrase sequences, rather than word sequences

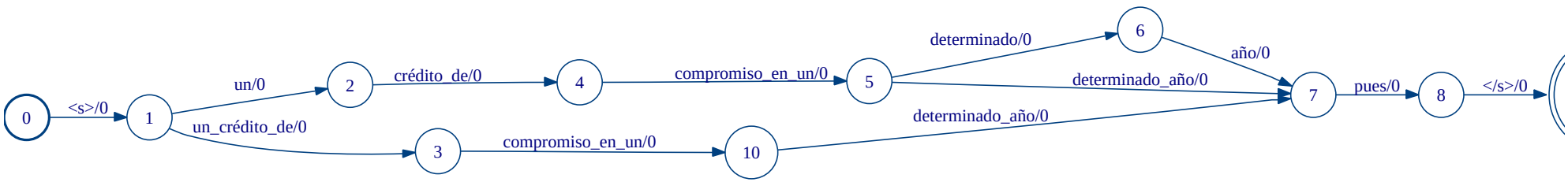
- This is done by first mapping the sentence into all its phrase sequences

Target Phrase Mapping - Text Translation



Sentence Acceptor

⇓ Target Phrase Segmentation Transducer

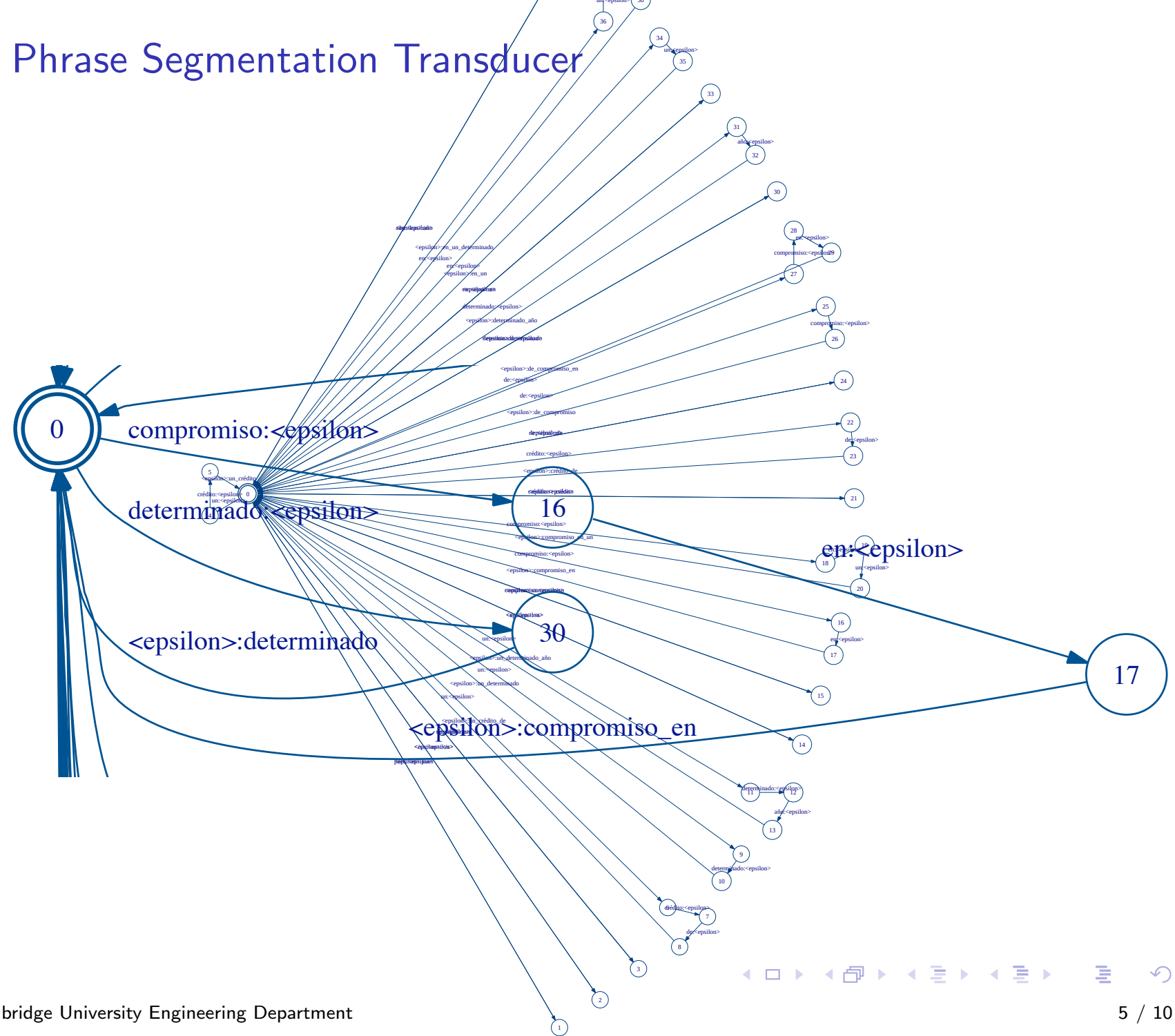


Phrase Sequence Lattice Ω

Phrase Sequence Lattice contains the phrase sequences that can be extracted from the text

- all phrase sequences correspond to the unique foreign sentence
- here, a phrase is a sequence of words which can be translated
- different phrase sequences lead to different translations
- the lattice is unweighted

Target Phrase Segmentation Transducer



Statistical Modeling for Speech Translation

Foreign Speech		Foreign Sentence		Foreign Phrases		English Phrases		English Sentence
A	$\longleftarrow$	$\mathbf{F}$	$\longleftarrow$	$\mathbf{V}$	$\longleftarrow$	$\mathbf{U}$	$\longleftarrow$	$\mathbf{E}$
Models $\longrightarrow$	$P(\mathbf{A}\|\mathbf{F})$		$P(\mathbf{F}\|\mathbf{V})$		$P(\mathbf{V}\|\mathbf{U})$		$P(\mathbf{U}\|\mathbf{E})$	$P(\mathbf{E})$
FSMs $\longrightarrow$	$\mathcal{L}$		Ω		Φ		W	G
	STT Word Lattice		Target Phrase Mapping		Phrase Translation, Reordering		Source Phrase Mapping	Source Language Model

The Target Phrase Mapping transducer is applied to the foreign language STT word lattice

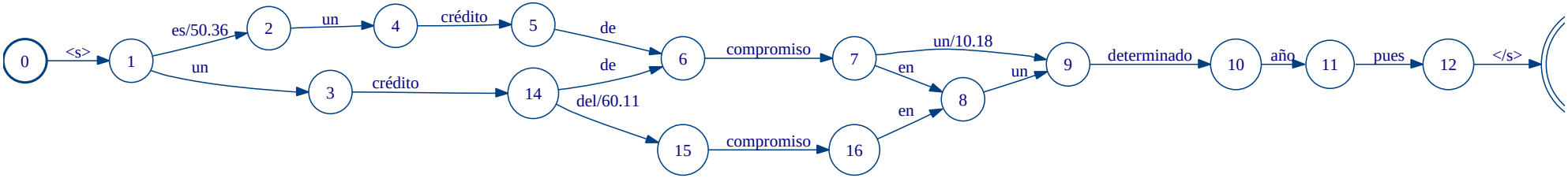
- $\mathcal{L} \cdot \Omega$: the likely foreign phrase sequences that could have generated the foreign speech

The translation system still effectively translates phrase sequences, rather than word sequences

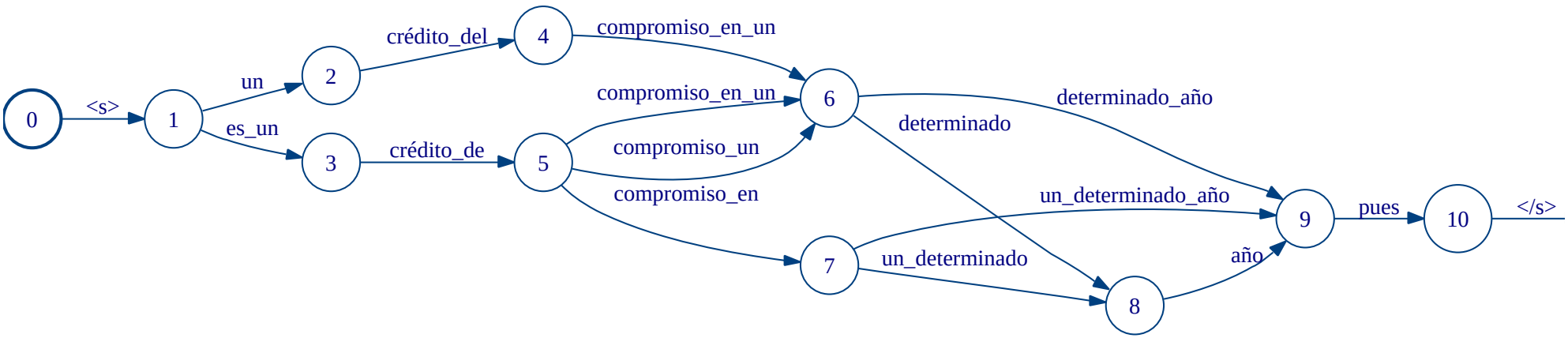
- These are extracted from the STT lattice, with STT scores, rather than from a text sentence

Target Phrase Mapping - Speech Translation

STT Word Lattice



⇓ Target Phrase Segmentation Transducer



STT Phrase Sequence Lattice

Phrase Sequence Lattice contains the phrase sequences that can be extracted from the text

- phrase sequences correspond to the translatable word sequences in the lattice
- the lattice contains weights from the STT system
- translating this foreign phrase lattice is MAP translation of the foreign speech under the generative model

Direct Translation of STT Word Lattices

Original Problem: How to translate STT lattices ?

New Problem: How to extract phrases from lattices ?

'Speech Translation' is recast as an STT analysis problem in which the goal is to extract translatable foreign language phrases from STT word lattices

Step 1. Perform foreign language STT to generate a foreign language word lattice $\mathcal{L}$

Step 2. Analyze the foreign language word lattice and extract the phrases to be translated

Step 3. Build the target language phrase mapping transducer Ω

Step 4. Compose $\mathcal{L}$ and Ω to create the foreign language STT Phrase Lattice

Step 5. Translate the foreign language phrase lattice using the TTM as if it were a phrase lattice extracted from a foreign language text sentence.

Tight coupling of STT and SMT models and systems :

- Step 1 is a 'standard' STT operation
- Steps 3, 4, and 5 are 'standard' TTM operations

Experimental Results

First, do no harm ....

- Translation from lattices should be no worse than translation from a single transcription

- Prior to recently developed techniques[3] the ambiguity introduced by STT lattices was reported to degrade translation performance

- Posterior-based pruning is used to control the amount of ambiguity presented to the translation system

Initial experiments in phrase-based speech translation using the TTM

TC-STAR 2005 Chinese-English Broadcast News Translation Task (BLEU)

	Mandarin Source	DEV	EVAL
Monotone Phrase Order	Ref. Transcription	16.1	18.8
	STT 1-Best	14.8	13.6
	STT lattice	15.0	13.8
MJ-1 VT Phrase Reordering	Ref. Transcription	16.1	19.3
	STT 1-Best	15. 0	13.8
	STT lattice	15.1	14.0

[3] For example, E. Matusov, S.Kanthak, H. Ney, On the integration of speech recognition and statistical machine translation, InterSpeech, 2005.
N. Bertoldi M. Federico, A new decoder for spoken language translation based on confusion networks, ASRU, 2005.

Extracting Translatable Phrases from STT Lattices

Modeling problems :

- Proper inclusion of the foreign language model (in progress ...)
- Phrase sequences extracted from STT lattices can be hard to find in parallel text
  - STT errors
  - spoken language phenomena, disfluencies, silences, ...
  - the usual genre mismatches
- The STT and Translation systems must be very compatible for this approach to work
- Possibly apply Rich Text transcription methods developed to 'clean' STT output
  - Inserting SU markers, phrase boundaries, etc. into STT lattices