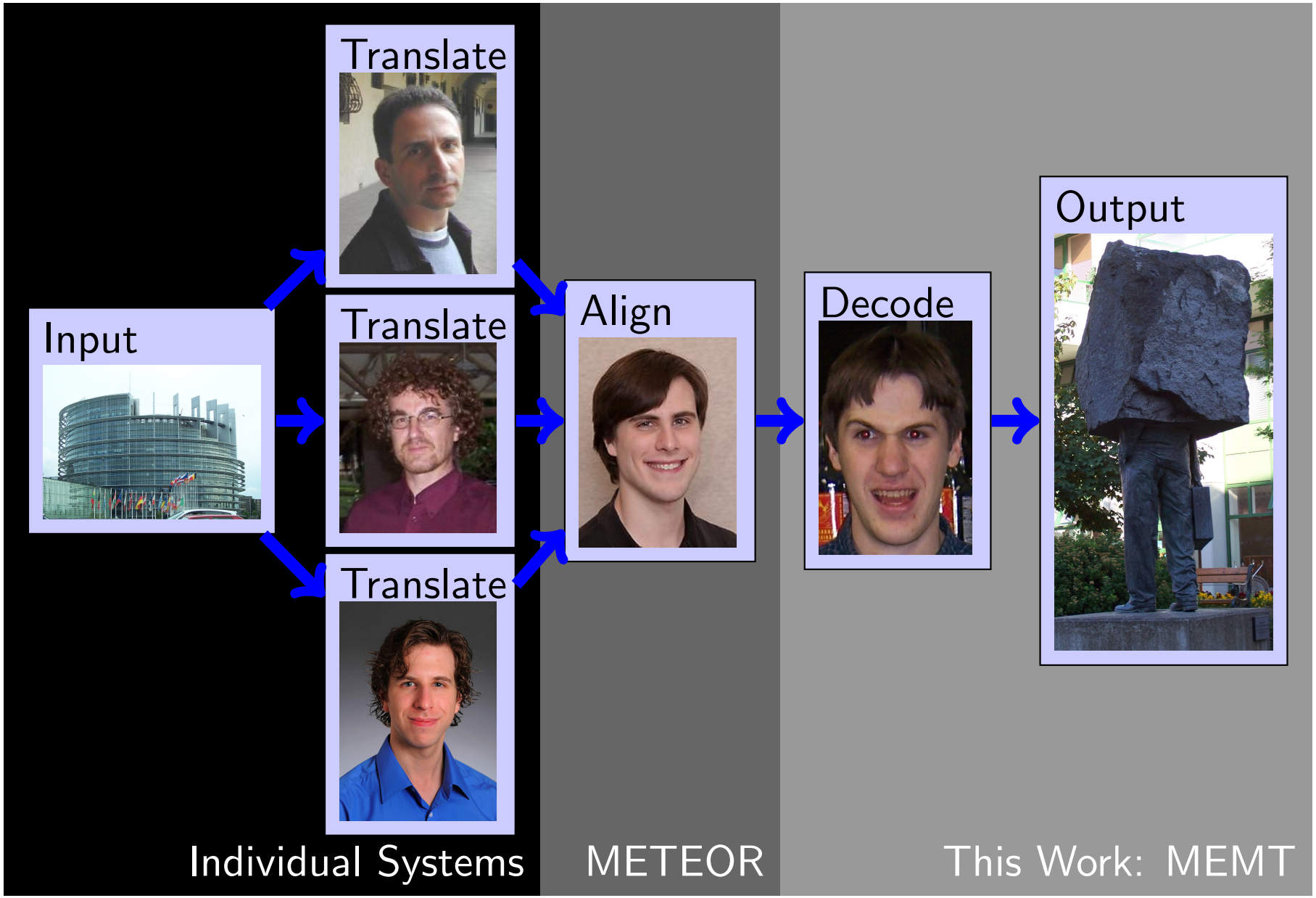
Outline


Input
Translate
Translate
Translate
Align
Decode
Output
Individual Systems
METEOR
This Work: MEMT

Arabic-English Example Combination

System 1: So even if that was meaningful , it is because you were late

System 2: Even if feasible , it is because you have been delayed

↓ Combine

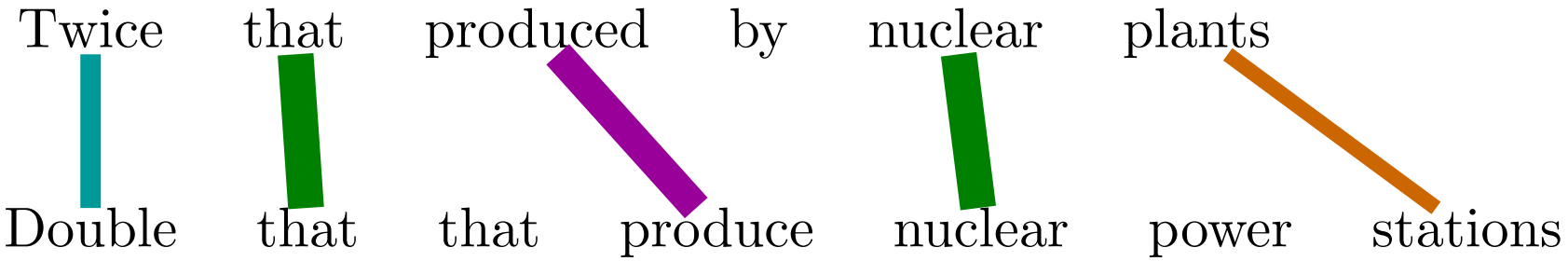
Combined: Even if feasible , it is because you were late

≠ Compare

Reference: And even if that was useful , it was because you were late

Sentence Pair Alignment

Match surface, stems, WordNet synsets, and automatic paraphrases
Minimize crossing alignments



Lavie and Agarwal, METEOR: An Automatic Metric for MT Evaluation with High Levels of Correlation with Human Judgments, WMT 2007.

Overall Alignment: Urdu-English Example

1 Russian President Putin Mir ولادی it for a big success .

2 The Russian president ولادی the result of a big victory for Putin .

Overall Alignment: Urdu-English Example

1 Russian President Putin Mir ولادی it for a big success .

2 The Russian president ولادی the result of a big victory for Putin .

1 Russian President Putin Mir ولادی it for a big success .

3 For the result Russian President ولادی Mir Putin is a great success .

2 The Russian president ولادی the result of a big victory for Putin .

3 For the result Russian President ولادی Mir Putin is a great success .

Search Space

Algorithm

Start at the beginning of each sentence

Branch by appending the first unused word from a system

Example

System 1: Now can know why .

System 2: Now we can now know why .

Partial Hypothesis

Now
Now

Search Space

Algorithm

Start at the beginning of each sentence

Branch by appending the first unused word from a system

Use the appended word and those aligned with it

Example

System 1: Now can know why .

System 2: Now we can now know why .

Partial Hypothesis

Now { can / we }

Search Space

Algorithm

Start at the beginning of each sentence

Branch by appending the first unused word from a system

Use the appended word and those aligned with it

Loop until all hypotheses reach end of sentence

Example

System 1: Now can know why .

System 2: Now we can now know why .

Partial Hypothesis

Now we { can / can

Search Space

Algorithm

Start at the beginning of each sentence

Branch by appending the first unused word from a system

Use the appended word and those aligned with it

Loop until all hypotheses reach end of sentence

Example

System 1: Now can know why .

System 2: Now we can now know why .

Partial Hypothesis

Now we can { know / now }

Outline

Features

Length

Length of hypothesis

Language Model

Model: log probability from an ARPA language model
OOV: count of words not found in the model

Match

Count of *n*-grams matching each system

Feature Rationale

Length

Length of hypothesis

Compensate for length's impact on other features

Language Model

Model: log probability from an ARPA language model
OOV: count of words not found in the model

Fluent output with tuned OOV penalty

Match

Count of *n*-grams matching each system

Agreement with translation systems

Match Features

System 1: Supported Proposal of France

System 2: Support for the Proposal of France

↓ Hypothesis

Hypothesis: Support for Proposal of France

↓ Count

	Unigram	**Bigram**	**Trigram**	**Quadgram**
System 1	4	2	1	0
System 2	5	3	1	0

What's in a match?

Exact matches

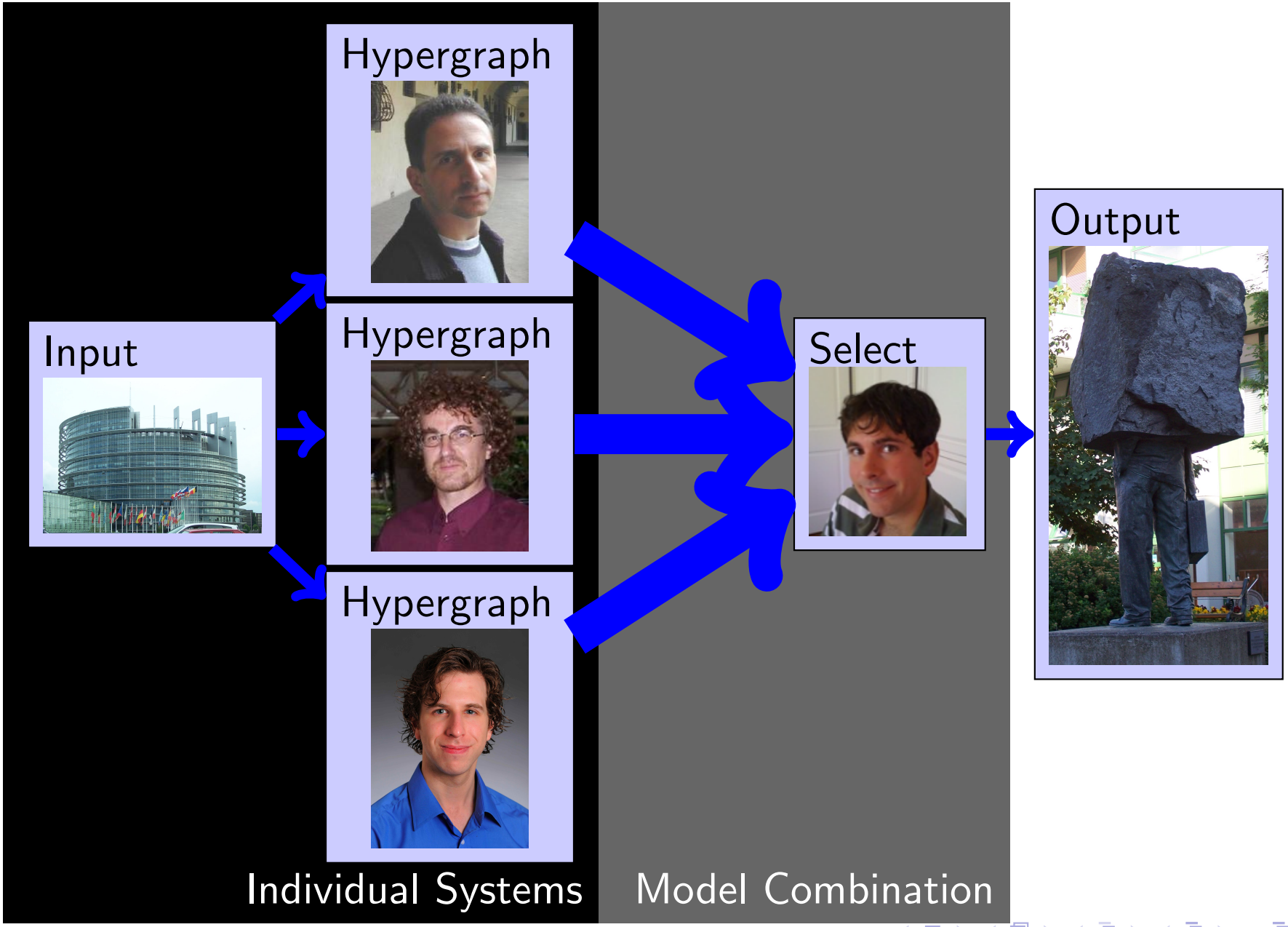
- Lexical choice
- Choosing between aligned alternatives

Approximate matches

- Vote to include/exclude text
- Word order

Answer

- Use both types of features
- Exact matches effectively get a tunable bonus


Input
Hypergraph
Hypergraph
Hypergraph
Select
Output
Individual Systems
Model Combination

Model Combination is Hypothesis Selection

The Search Space

- Union of search spaces from each system
- Combined sentence must be in one system's hypergraph

Formally

- Every system outputs a hypergraph
- Phrasal lattice is just a special-case hypergraph
- Add a root node and an edge to each system root

Model Combination is Hypothesis Selection

The Search Space

- Union of search spaces from each system
- Combined sentence must be in one system's hypergraph

Formally

- Every system outputs a hypergraph
- Phrasal lattice is just a special-case hypergraph
- Add a root node and an edge to each system root

Source Alignment

Hypergraphs retain alignment to source

Features

Length

Length of hypothesis

Model score

Score given by the underlying system

System indicator

Each system has a feature:

- 1 if derived from that system
- 0 otherwise

N-gram support

Support from each system for n-grams

N-gram support

Posterior of n-gram

What fraction of system i's translations include "crack rocks"?

Formally

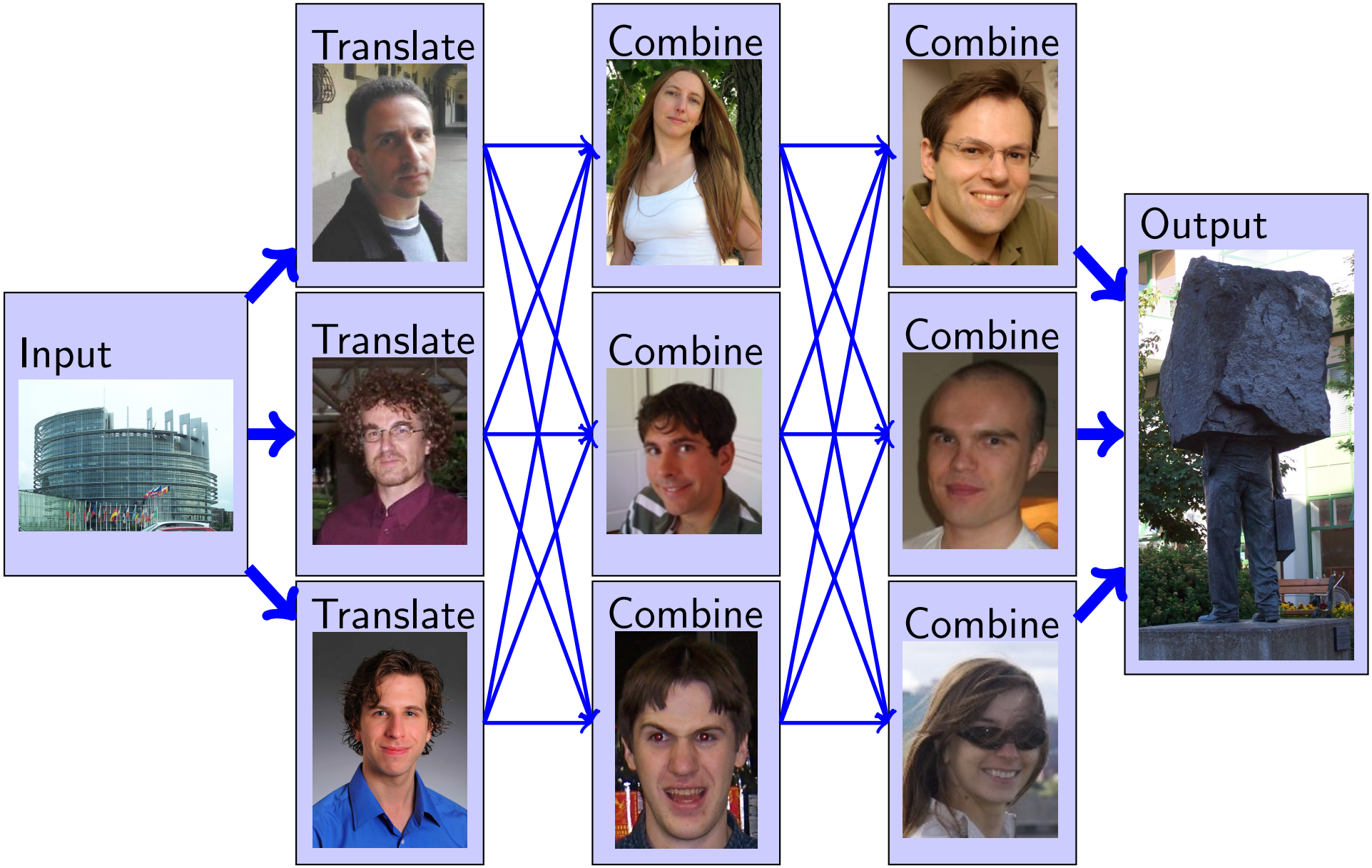
$$v_i^n(g) = \mathbb{E}_{P_i(d|f)} h(d, g)$$

$v_i^n(g)$	System i's vote for n-gram g
$P_i(d\|f)$	Probability of a derivation d in hypergraph f from system i
$h(d, g)$	1 if the derivation d contains n-gram g; 0 otherwise

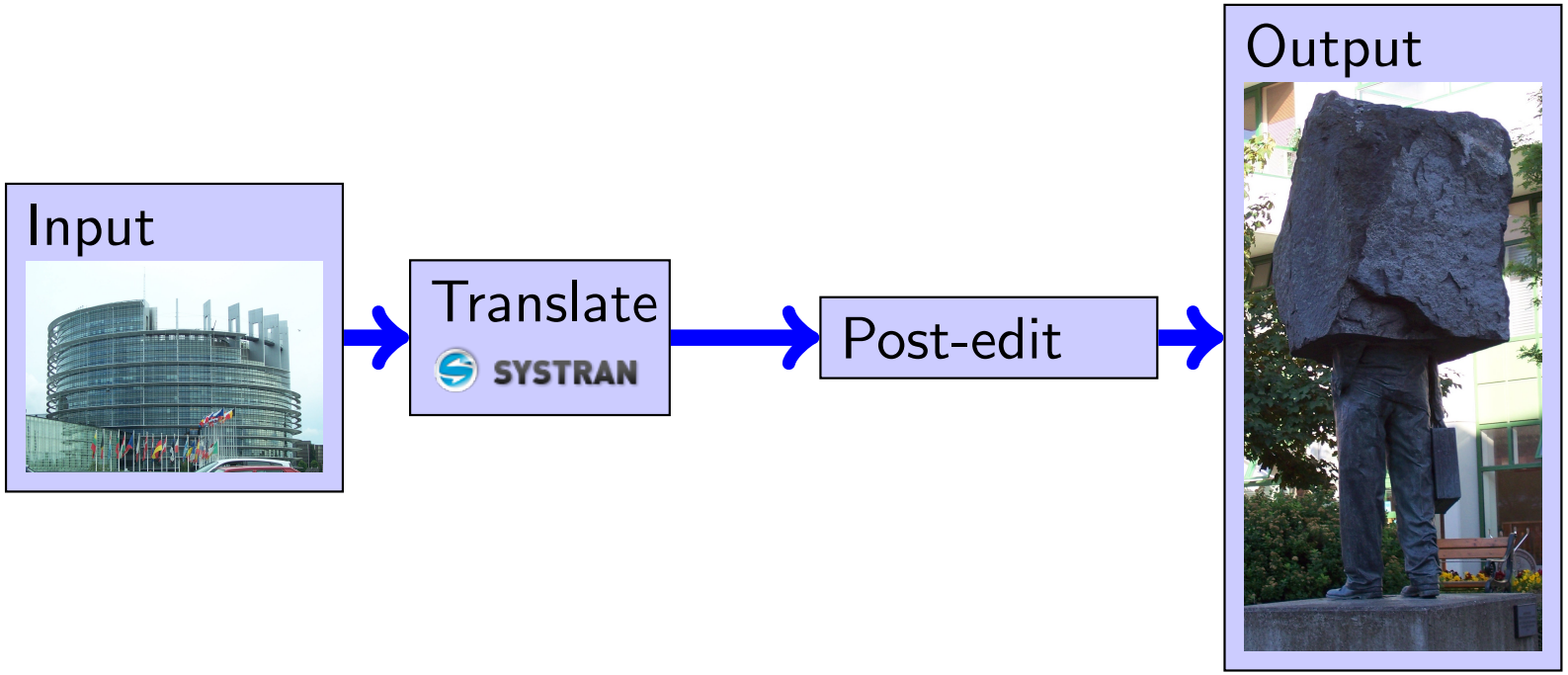
Performance

	ar-en	zh-en
Best individual	43.9	28.4
Combined	45.3	29.0

Table: Performance (BLEU) on NIST 2008 task using three systems


Input
Translate
Translate
Translate
Combine
Combine
Combine
Combine
Combine
Combine
Output

Serial System Combination

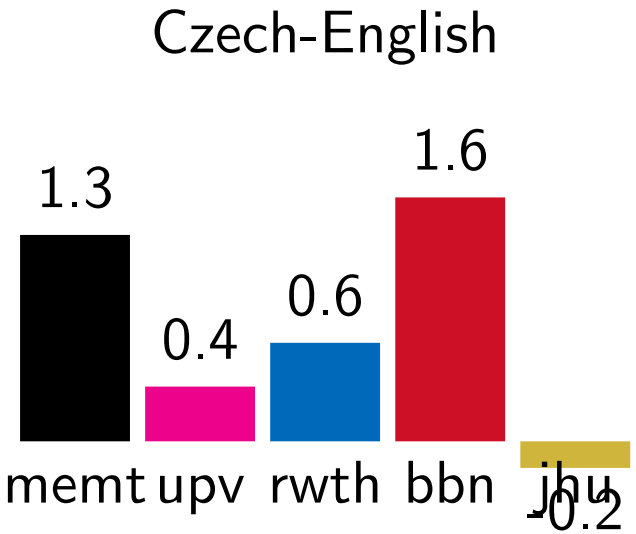


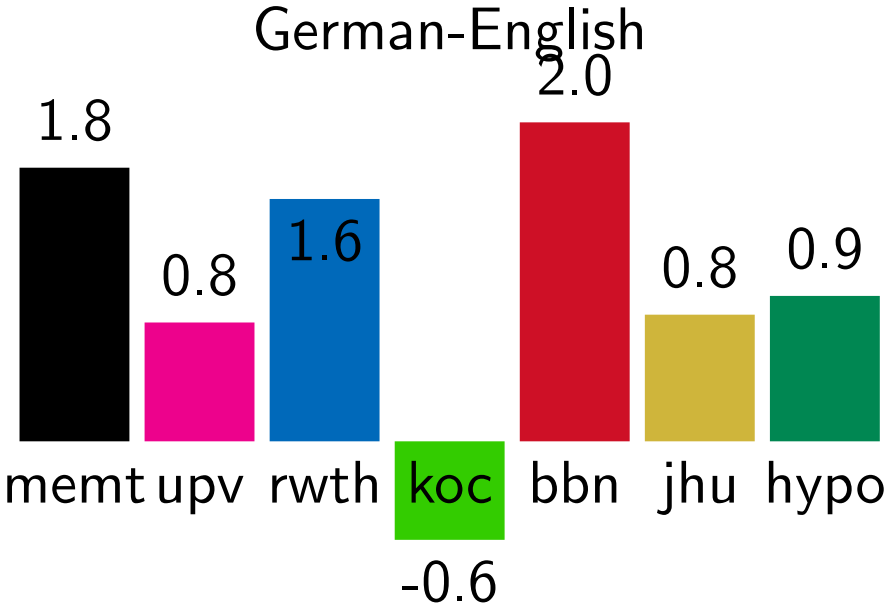
Input Comparison

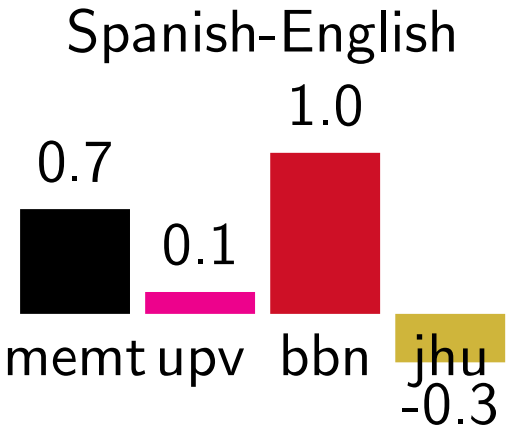
Input to System Combination

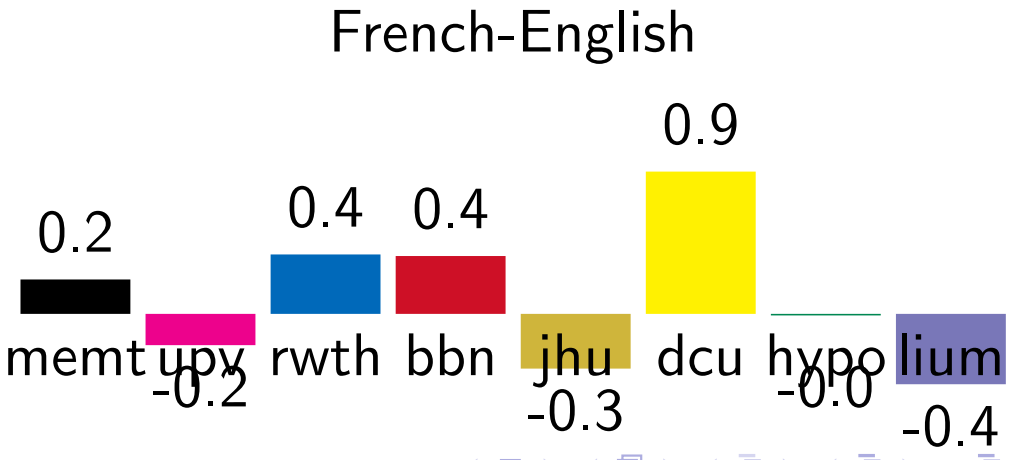
MBR	N-best list
Hyposel	N-best list
Confusion Networks	N-best list
MEMT	1-best
Model Combination	Hypergraph
Serial System Combination	Single output

Results Into English









Results From English