MT System Combination
System Combination in MT

- Methods of machine translation
  - Rule based
  - Statistical
  - Hierarchical
  - Syntax based

- Output is different
- Make use of the individual strengths of the different systems to improve translation quality
- Selecting the best output on a sentence-by-sentence basis
- Synthetic combination of the output from the original systems
- Fixing mistakes of one system by using another system afterwards
Several MT System Outputs

Reference Translation:

hoffman was addicted to drugs, fortunately awakening in a timely manner to begin an acting career

- hoffman was obsessed timely wake up to create a career drug
- hoffman were drug fortunately awakening in a timely manner to create career
- hoffman previously enamored drug. luckily i realized create career
- hoffman was mesmerized by drug but woke up in a timely manner to create career
- hoffmann was obsessed drug, in a timely manner to create a career
- hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause

Chinese-English MT06

- Statistical Phrase Based
- Statistical Hierarchical
- Example Based

Translation hypotheses are in order of the systems testset BLEU score
Parallel Combination

Source Language Text → MT System → MT System → MT System → MT System Combination → Translation
Parallel Combination Approaches

- Selection of whole hypotheses
  - Hypothesis Selection from N-best Lists
  - Minimum Bayes Risk Re-ranking
  - Hypothesis Selection from Forests

- Synthesis of new hypotheses
  - Confusion Network based Combination
  - Joint Decoding with Flexible Ordering
Parallel Combination: Hypothesis Selection

Source Language Text → MT System → MT System → MT System → MT System Combination → Translation
Reference Translation:

hoffman was addicted to drugs, fortunately awakening in a timely manner to begin an acting career

hoffman was obsessed timely wake up to create a career drug
hoffman were drug fortunately awakening in a timely manner to create career
hoffman previously enamored drug. luckily i realized create career
hoffman was mesmerized by drug but woke up in a timely manner to create career
hoffmann was obsessed drug, in a timely manner to create a career
hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
Hypothesis Selection

Reference Translation:

**hoffman was addicted to drugs, fortunately awaking in a timely manner to begin an acting career**

hoffman was obsessed timely wake up to create a career drug

hoffman were drug fortunately awakening in a timely manner to create career

hoffman previously enamored drug. luckily i realized create career

hoffman was mesmerized by drug but woke up in a timely manner to create career

hoffmann was obsessed drug, in a timely manner to create a career

hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
How to decide which hypothesis to pick?

- Language model
- System bias / trained system weights
  - Boost hypothesis from each system according to its overall (BLEU) score on development data
- MT System confidence score
  - System tells you what it thinks how well it translated the sentence

Problematic:
- Most systems use a score, that is not a real confidence measure or probability
- not comparable between sentences within a system
- not comparable between systems
- Need to be normalized
Hypothesis Selection from N-best Lists
Hypothesis Selection from N-best Lists

Pick the best hypothesis from the different systems for each source sentence
Hypothesis Selection from N-best Lists

- Pick the best hypothesis from the different systems for each source sentence
- Use n-best list re-ranking approach
- Add n-best hypotheses from each system
- Re-rank joint n-best list
Hypothesis Selection from N-best Lists

- Pick the best hypothesis from the different systems for each source sentence
- Use n-best list re-ranking approach
- Add n-best hypotheses from each system
- Re-rank joint n-best list
- Find good features: add n-best list based features
N-best list re-ranking: Features

- Consistently calculated for the joint n-best list
  - Language model
  - Statistical word lexicon
  - System weights
  - Rank in system’s n-best list
  - N-gram overlap features
  - N-best list n-gram probability
  - N-best list sentence length features

- Minimum error rate training to determine feature weights on a development test set
N-gram Agreement Features

n-best list for one source sentence

\[ e^5 \]
N-gram Agreement Features

n-best list for one source sentence

$e_5$

n = 1 word agreement: 90%
N-gram Agreement Features

n-best list for one source sentence

$e_5$

n = 1
word agreement: 90%

n = 3
tri-gram agreement: 50%
N-gram Agreement Features

n-best list for one source sentence

e_5

<table>
<thead>
<tr>
<th>n = 1</th>
<th>word agreement: 90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 3</td>
<td>tri-gram agreement: 50%</td>
</tr>
<tr>
<td>n = 5</td>
<td>5-gram agreement: 30%</td>
</tr>
</tbody>
</table>
N-gram Agreement vs. N-gram Probability

San Francisco

n-best list for one source sentence

...500 hypotheses ...

Silja Hildebrand – MT System Combination
N-gram Agreement vs. N-gram Probability

San Francisco

n-best list for one source sentence

n = 2 bi-gram agreement: 0.6%

...500 hypotheses ...

Silja Hildebrand – MT System Combination
**N-gram Agreement vs. N-gram Probability**

San Francisco

- **n-best list for one source sentence**
  - 500 hypotheses ...
  - n = 2 bi-gram agreement: 0.6%
  - $P(\text{Francisco} | \text{San}) = \frac{3}{3}$ bi-gram probability: 100%
N-gram Agreement vs. N-gram Probability

San Francisco

n-best list for one source sentence

500 hypotheses...

- n = 2 bi-gram agreement: 0.6%
- \( P( \text{Francisco} | \text{San} ) = \frac{3}{3} \) bi-gram probability: 100%

LM n-gram probability gives information on word order.
Parallel Combination

- Synthesis of new translations

Source Language Text → MT System → MT System → MT System → MT System Combination → Translation
Synthesis of new translations

hoffman was addicted to drugs, fortunately awake in a timely manner to begin an acting career

hoffman was obsessed timely wake up to create a career drug
hoffman were drug fortunately awakening in a timely manner to create career
hoffman previously enamored drug. Luckily I realized create career
hoffman was mesmerized by drug but woke up in a timely manner to create career
hoffmann was obsessed drug, in a timely manner to create a career
hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause

Chinese-English MT06
Synthesis of new translations

hoffman was addicted to drugs, fortunately awaking in a timely manner to begin an acting career

hoffman was obsessed timely wake up to create a career
drug

hoffman were drug fortunately awakening in a timely manner to create career

hoffman previously enamored drug. luckily i realized create career

hoffman was mesmerized by drug but woke up in a timely manner to create career

hoffmann was obsessed drug, in a timely manner to create a career

hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause

hoffman was mesmerized by drug
Synthesis of new translations

hoffman was addicted to drugs, fortunately awakening in a timely manner to begin an acting career
Synthesis of new translations

hoffman was addicted to drugs, fortunately awaking in a timely manner to begin an acting career

hoffman was obsessed timely wake up to create a career drug
hoffman were drug fortunately awakening in a timely manner to create career
hoffman previously enamored drug. luckily i realized create career
hoffman was mesmerized by drug but woke up in a timely manner to create career
hoffmann was obsessed drug, in a timely manner to create a career
hoffman has fortunately drug come to realize in a timely manner for performing arts
to open up the cause

hoffman was mesmerized by drug
fortunately awakening
in a timely manner to create career
for performing arts

Chinese-English MT06
Confusion Network Based Combination

Align Words, Build Confusion Network

hoffman was obsessed timely wake up to create a career drug
hoffman were drug fortunately awakening in a timely manner to create career
hoffman previously enamored drug. luckily i realized create career
hoffman was mesmerized by drug but woke up in a timely manner to create career
hoffmann was obsessed drug, in a timely manner to create a career
hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
Confusion Network Based Combination

- Align Words, Build Confusion Network

- hoffman was obsessed timely wake up to create career drug
- hoffman were drug fortunately awakening in a timely manner to create career
- hoffman previously enamored drug, luckily i realized create career
- hoffman was mesmerized by drug but woke up in a timely manner to create career
- hoffmann was obsessed drug, in a timely manner to create a career
- hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
Align Words, Build Confusion Network

hoffman was obsessed with timely wake up to create a career drug
hoffman were drug fortunately awakening in a timely manner to create career
hoffman previously enamored drug, luckily i realized create career
hoffman was mesmerized by drug but woke up in a timely manner to create career
hoffmann was obsessed drug in a timely manner to create a career
hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
Align Words, Build Confusion Network

- hoffman was obsessed, timely wake up to create a career drug
- hoffman were drug, fortunately awakening in a timely manner to create career
- hoffman previously enamored drug, luckily i realized create career
- hoffman was mesmerized by drug but woke up in a timely manner to create career
- hoffmann was obsessed drug, in a timely manner to create a career
- hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
Confusion Network Based Combination

Align Words, Build Confusion Network

- hoffman was obsessed timely wake up to create a career drug
- hoffman were drug fortunately awakening in a timely manner to create career
- hoffman previously enamored drug. luckily i realized create career
- hoffman was mesmerized by drug but woke up in a timely manner to create career
- hoffmann was obsessed drug, in a timely manner to create a career
- hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
Confusion Network Based Combination

Align Words, Build Confusion Network

- hoffman was obsessed timely wake up to create a career drug
- hoffman were drug fortunately awakening in a timely manner to create career
- hoffman previously enamored drug, luckily i realized create career
- hoffman was mesmerized by drug but woke up in a timely manner to create career
- hoffmann was obsessed drug, in a timely manner to create a career
- hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
Silja Hildebrand – MT System Combination

Confusion Network Based Combination

Align Words, Build Confusion Network

- hoffman was obsessed timely wake up to create a career drug
- hoffman were drug fortunately awakening in a timely manner to create career
- hoffman previously enamored drug. Luckily I realized create career
- hoffman was mesmerized by drug but woke up in a timely manner to create career
- hoffmann was obsessed drug. in a timely manner to create a career
- hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
Silja Hildebrand – MT System Combination

Confusion Network Based Combination

Align Words, Build Confusion Network

hoffman was obsessed timely wake up to create a career drug
hoffman were drug fortunately awakening in a timely manner to create career
hoffman previously enamored drug. luckily i realized create career
hoffman was mesmerized by drug but woke up in a timely manner to create career
hoffmann was obsessed drug, in a timely manner to create a career
hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause

hoffman (5)  
1  2  

hoffmann (1)
Confusion Network Based Combination

Align Words, Build Confusion Network

hoffman were drug fortunately awakening in a timely manner to create career

hoffman were drug, luckily I realized to create career

hoffman was obsessed by drug, but woke up in a timely manner to create career

hoffman was mesmerized by drug, but woke up in a timely manner to create career

hoffman was obsessed with drug, in a timely manner to create career

hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
Confusion Network Based Combination

Align Words, Build Confusion Network

hoffman was obsessed timely wake up to create a career drug
hoffman were drug fortunately awakening in a timely manner to create career
hoffman previously enamored drug, luckily i realized create career
hoffman was mesmerized by drug but woke up in a timely manner to create career
hoffmann was obsessed drug, in a timely manner to create a career
hoffman has fortunately drug come to realize in a timely manner for performing arts to open up the cause
Confusion Network Based Combination

Align Words, Build Confusion Network

- hoffman
- was
- obsessed
- timely
- wake up
- to
- create
- a
- career
- drug

- hoffman
- were
- drug
- fortunately
- awakening
- in
- a
- timely
- manner
- to
- create
- career

- hoffman
- previously
- enamored
- drug
- luckily
- realized
- create
- career

- hoffman
- was
- mesmerized
- by
- drug
- but
- woke up
- in
- a
- timely
- manner
- to
- create
- career

- hoffmann
- was
- obsessed
- drug
- in
- a
- timely
- manner
- to
- create
- a
- career

- hoffman
- has
- fortunately
- drug
- come
- to
- realize
- in
- a
- timely
- manner
- for
- performing
- arts
- to
- open
- up
- the
- cause
Silja Hildebrand – MT System Combination

Confusion Network Based Combination

Align Words, Build Confusion Network

hoffman were drugFortunately been to realize in a timely manner to create career
hoffman previously enamored drug, luckily i realized create career
hoffman was mesmerized by drug, but woke up in a timely manner to create career
huffmann was obsessed drug, in a timely manner to create a career
huffman has fortunately drug come to realize in a timely manner for performing arts

to open up the cause
Confusion Network Decoding

- Choose one translation hypothesis as skeleton
  - Determines word order
- Align each hypothesis to the skeleton using
  - TER (Translation Edit Rate)
  - ITG (Inversion Transduction Grammar) based WER
  - Statistical word alignment
- Build confusion network with consensus votes
- Add LM score into network
- Train system weights, add into network
- Choose best path through the network (decode)
- Output consensus translation
I like eating chocolate ice-cream.
I like to eat vanilla ice-cream.
I like to eat ice-cream with chocolate.
I like ice-cream.
Confusion Network Decoding: Example

I like eating chocolate ice-cream.
I like to eat vanilla ice-cream.
I like to eat ice-cream with chocolate.
I like ice-cream.

choose as skeleton
Confusion Network Decoding: Example

I like eating chocolate ice-cream.
I like to eat vanilla ice-cream.
I like to eat ice-cream with chocolate.
I like ice-cream.

I like ε eating chocolate ε ice-cream .
I like to eat vanilla ε ice-cream .
Confusion Network Decoding: Example

I like eating chocolate ice-cream.
I like to eat vanilla ice-cream.
I like to eat ice-cream with chocolate.
I like ice-cream.

I like ε eating chocolate ε ice-cream .
I like to eat vanilla ε ice-cream .
I like to eat chocolate with ice-cream .

choose as skeleton
skeleton determines word order
Confusion Network Decoding: Example

I like eating chocolate ice-cream.
I like to eat vanilla ice-cream.
I like to eat ice-cream with chocolate.
I like ice-cream.

I like ε eating chocolate ε ice-cream .
I like to eat vanilla ε ice-cream .
I like to eat chocolate with ice-cream .
I like ε ε ε ε ε ice-cream .

choose as skeleton

skeleton determines word order
Confusion Network Decoding: Example

I like eating chocolate ice-cream.
I like to eat vanilla ice-cream.
I like to eat ice-cream with chocolate.
I like ice-cream.

I like ε eating chocolate ε ice-cream .
I like to eat vanilla ε ice-cream .
I like to eat chocolate with ice-cream .
I like ε ε ε ε ice-cream .

choose as skeleton

skeleton determines word order
Confusion Network Decoding: Example

I like eating chocolate ice-cream.
I like to eat vanilla ice-cream.
I like to eat ice-cream with chocolate.
I like ice-cream.

I (4) like (4) to (2) eat (2) chocolate (2) ice-cream (4) . (4)
I like ε eating chocolate ε ice-cream .
I like to eat vanilla ε ice-cream .
I like to eat chocolate with ice-cream .
I like ε ε ε ε ice-cream .

choose as skeleton
skeleton determines word order

I like to eat chocolate ice-cream.
Challenges

- **Word alignment**
  - TER alignment (Translation Edit Rate)
  - ITG based alignment (Inversion Transduction Grammar) - invWER
  - Use morphology, synonyms, POS tag
  - Go to phrases
    - Difficult without source-target phrase alignment available

- **Double translations**

- **Dropped words**

- **Pairwise vs. incremental alignment**
  - Next hypothesis is aligned to the existing network, not to the skeleton
  - Order of adding hypothesis does make a difference, e.g. use increasing TER/decreasing BLEU of the system
Challenges: Double Translations

I like big blue balloons
I like blue kites
I like balloons
Challenges: Double Translations

I like big blue balloons
I like blue kites
I like balloons

I like big blue balloons
Challenges: Double Translations

I like big blue balloons
I like blue kites
I like balloons

like big blue balloons
like blue
Challenges: Double Translations

I like big blue balloons
I like blue kites
I like balloons

I like big blue balloons
I like kites blue NULL
Challenges: Double Translations

<table>
<thead>
<tr>
<th>I like big blue balloons</th>
</tr>
</thead>
<tbody>
<tr>
<td>I like blue kites</td>
</tr>
<tr>
<td>I like balloons</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I like big blue balloons</th>
</tr>
</thead>
<tbody>
<tr>
<td>I like kites blue NULL</td>
</tr>
<tr>
<td>I like NULL NULL balloons</td>
</tr>
</tbody>
</table>
Challenges: Double Translations

I like big blue balloons
I like blue kites
I like balloons

I like big blue balloons
I like kites blue NULL
I like NULL NULL balloons

I (3) like (3) big (1) kites (1) ε (1) balloons (2)

1 -> 2 -> 3 -> 4 -> 5 -> 6
Challenges: Double Translations

I like big blue balloons
I like blue kites
I like balloons

I like big blue balloons
I like kites blue NULL
I like NULL NULL balloons

I like kites blue balloons
Challenges: Dropped Words

I like blue balloons
I like my blue
I like blue kites
Challenges: Dropped Words

I like blue balloons
I like my blue
I like blue kites

I like NULL blue balloons
I like my blue NULL
I like NULL NULL kites
Challenges: Dropped Words

I like blue balloons
I like my blue
I like blue kites

I like NULL blue balloons
I like my blue NULL
I like NULL NULL kites

I (3) like (3) my (1) blue (3) baloons (1)
1 → 2 → 3 → 4 → 5 → 6

my (1) blue (3) ε (2)

kites (1) baloons (1)
Challenges: Dropped Words

I like blue balloons
I like my blue
I like blue kites

I like NULL blue balloons
I like my blue NULL
I like NULL NULL kites

1  (3) like  (3)  ε  (2)  blue  (3)  balloons  (1)
2  3  4  5  6

I like blue
Pairwise vs. Incremental Alignment: Example

| I like balloons
I like big blue balloons
I like blue kites |
| pairwise | incremental |
## Pairwise vs. Incremental Alignment: Example

| Pairwise          | I like NULL balloons
<table>
<thead>
<tr>
<th></th>
<th>I like big blue balloons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incremental</td>
<td>I like blue kites</td>
</tr>
<tr>
<td></td>
<td>I like big blue balloons</td>
</tr>
</tbody>
</table>
Pairwise vs. Incremental Alignment: Example

pairwise

I like NULL balloons
I like big blue balloons
I like balloons NULL
I like blue kites

I like balloons
I like big blue balloons
I like blue kites

incremental
Pairwise vs. Incremental Alignment: Example

Pairwise

<table>
<thead>
<tr>
<th>I like</th>
<th>NULL</th>
<th>balloons</th>
<th>NULL</th>
</tr>
</thead>
<tbody>
<tr>
<td>I like</td>
<td>big blue</td>
<td>balloons</td>
<td>NULL</td>
</tr>
<tr>
<td>I like</td>
<td>NULL</td>
<td>blue</td>
<td>kites</td>
</tr>
</tbody>
</table>

Incremental
Pairwise vs. Incremental Alignment: Example

I like balloons
I like big blue balloons
I like blue kites

pairwise

<table>
<thead>
<tr>
<th>1</th>
<th>like</th>
<th>NULL</th>
<th>balloons</th>
<th>NULL</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>like</td>
<td>big blue</td>
<td>balloons</td>
<td>NULL</td>
</tr>
<tr>
<td>3</td>
<td>like</td>
<td>NULL</td>
<td>blue</td>
<td>kites</td>
</tr>
</tbody>
</table>

corpora

I (3) like (3) big blue (1) balloons (2) blue (1) ε (2) like (3) ε (2) balloons (2) blue (1) kites (1)

incremental
I like balloons
I like big blue balloons
I like blue kites

Pairwise vs. Incremental Alignment: Example

I (3)
like (3)
1 2 3
big blue (1) blue (1) kites (1)

e (2) balloons (2) e (2)

I like balloons
Pairwise vs. Incremental Alignment: Example

I like balloons
I like big blue balloons
I like blue kites

pairwise

1. I like NULL balloons NULL
2. I like big blue balloons NULL
3. I like NULL blue kites

incremental

1. I like NULL NULL balloons
2. I like big blue blue balloons

I like balloons
Pairwise vs. Incremental Alignment: Example

I like balloons
I like big blue balloons
I like blue kites

pairwise

<table>
<thead>
<tr>
<th>I like</th>
<th>NULL</th>
<th>balloons</th>
<th>NULL</th>
</tr>
</thead>
<tbody>
<tr>
<td>I like</td>
<td>big blue</td>
<td>balloons</td>
<td>NULL</td>
</tr>
<tr>
<td>I like</td>
<td>NULL</td>
<td>blue</td>
<td>kites</td>
</tr>
</tbody>
</table>

incremental

<table>
<thead>
<tr>
<th>I like</th>
<th>NULL</th>
<th>NULL</th>
<th>balloons</th>
</tr>
</thead>
<tbody>
<tr>
<td>I like</td>
<td>big</td>
<td>blue</td>
<td>balloons</td>
</tr>
<tr>
<td>I like</td>
<td>NULL</td>
<td>blue</td>
<td>kites</td>
</tr>
</tbody>
</table>

I like balloons
I like balloons
I like big blue balloons
I like blue kites

Pairwise vs. Incremental Alignment: Example

I like NULL balloons NULL
I like big blue balloons NULL
I like NULL blue kites

I (3) like (3)
big blue (1) blue (1) kites (1)

I like NULL NULL balloons
I like big blue balloons
I like NULL blue kites

I (3) like (3)
big (1) blue (2) kites (1)
Pairwise vs. Incremental Alignment: Example

Pairwise:

I like NULL balloons NULL
I like big blue balloons NULL
I like NULL blue kites

Incremental:

I like NULL NULL balloons
I like big blue blue balloons
I like NULL blue kites

I like balloons
I like big blue balloons
I like blue kites