

Translation Options

Maria	no	dio	una	bofetada	a	la	bruja	verde
-------	----	-----	-----	----------	---	----	-------	-------

<u>Mary</u>	<u>not</u>	<u>give</u>	<u>a</u>	<u>slap</u>	<u>to</u>	<u>the</u>	<u>witch</u>	<u>green</u>
	<u>did not</u>		<u>a slap</u>		<u>by</u>		<u>green witch</u>	
	<u>no</u>		<u>slap</u>		<u>to the</u>			
	<u>did not give</u>				<u>to</u>			
					<u>the</u>			
			<u>slap</u>			<u>the witch</u>		

- Look up possible phrase translations
 - many different ways to segment words into phrases
 - many different ways to translate each phrase

Hypothesis Expansion

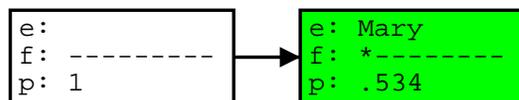
Maria	no	dio	una	bofetada	a	la	bruja	verde
<u>Mary</u>	<u>not</u>	<u>give</u>	<u>a</u>	<u>slap</u>	<u>to</u>	<u>the</u>	<u>witch</u>	<u>green</u>
	<u>did not</u>		<u>a</u>	<u>slap</u>	<u>by</u>		<u>green</u>	<u>witch</u>
	<u>no</u>		<u>slap</u>		<u>to</u>	<u>the</u>		
	<u>did not</u>	<u>give</u>			<u>to</u>			
					<u>the</u>			
			<u>slap</u>			<u>the</u>	<u>witch</u>	

```
e:
f: -----
p: 1
```

- Start with empty hypothesis
 - e: no English words
 - f: no foreign words covered
 - p: probability 1

Hypothesis Expansion

Maria	no	dio	una	bofetada	a	la	bruja	verde
Mary	<u>not</u>	<u>give</u>	<u>a</u>	<u>slap</u>	<u>to</u>	<u>the</u>	<u>witch</u>	<u>green</u>
	<u>did not</u>		<u>a slap</u>		<u>by</u>		<u>green witch</u>	
	<u>no</u>		<u>slap</u>		<u>to the</u>			
	<u>did not give</u>				<u>to</u>			
					<u>the</u>			
			<u>slap</u>			<u>the witch</u>		

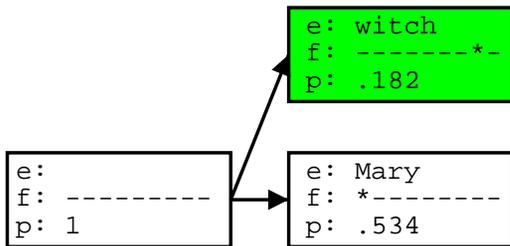


- Pick translation option
- Create hypothesis
 - e: add English phrase Mary
 - f: first foreign word covered
 - p: probability 0.534

Hypothesis Expansion

Maria	no	dio	una	bofetada	a	la	bruja	verde
-------	----	-----	-----	----------	---	----	-------	-------

Mary	not	give	a	slap	to	the	witch	green
did not			a slap		by		green witch	
no			slap		to the			
did not give					to			
					the			
			slap			the witch		

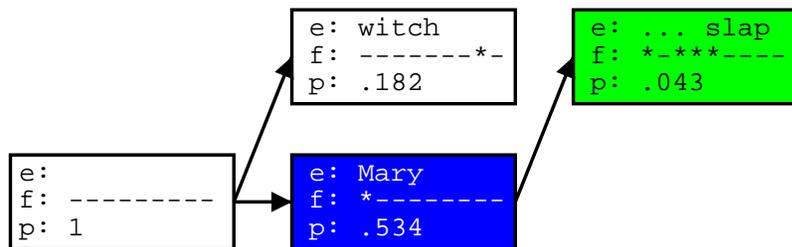


- Add another hypothesis

Hypothesis Expansion

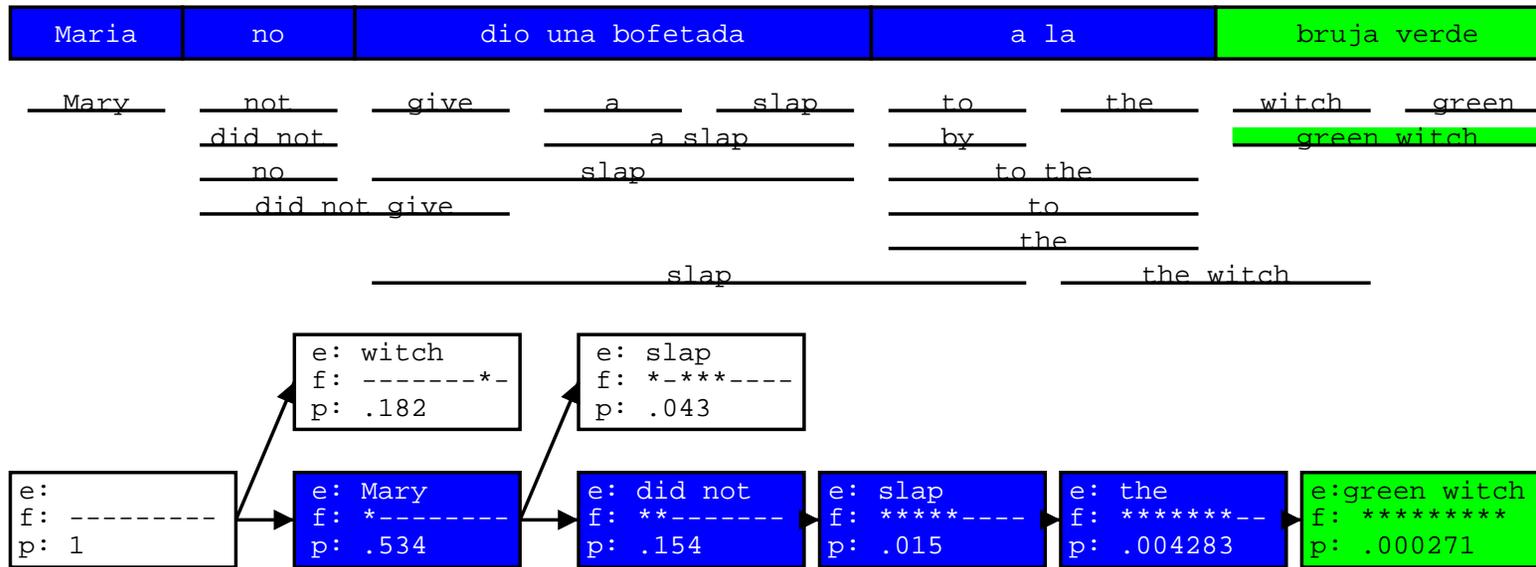
Maria	no	dio una bofetada	a	la	bruja	verde
-------	----	------------------	---	----	-------	-------

<u>Mary</u>	<u>not</u>	<u>give</u>	<u>a</u>	<u>slap</u>	<u>to</u>	<u>the</u>	<u>witch</u>	<u>green</u>
<u>did not</u>			<u>a</u>	<u>slap</u>	<u>by</u>			<u>green</u>
<u>no</u>		<u>slap</u>			<u>to</u>	<u>the</u>		
<u>did not give</u>					<u>to</u>	<u>the</u>		
					<u>the</u>			
			<u>slap</u>			<u>the</u>	<u>witch</u>	



- Further hypothesis expansion

Hypothesis Expansion

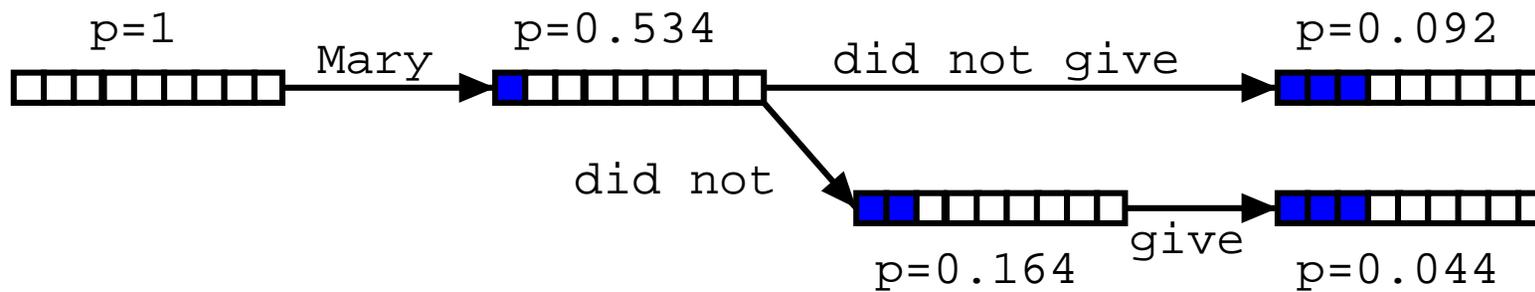


- ... until all foreign words covered
 - find best hypothesis that covers all foreign words
 - backtrack to read off translation

Explosion of Search Space

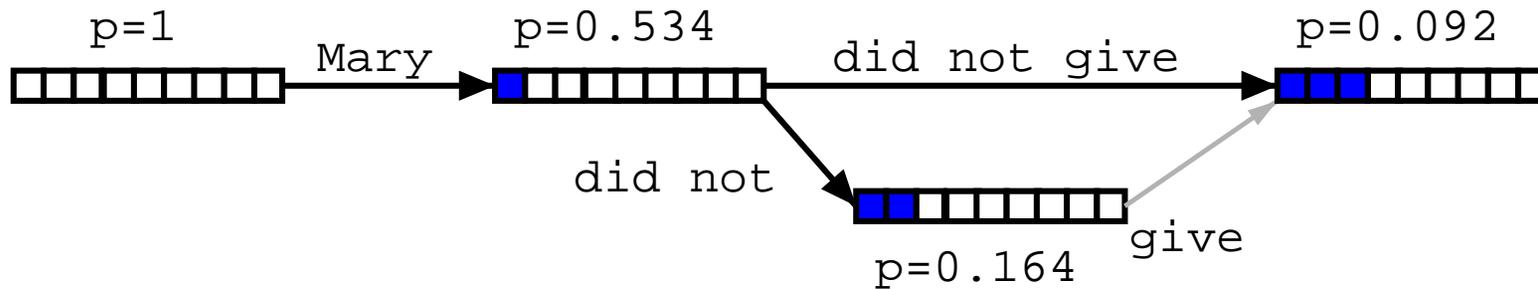
- Number of hypotheses is exponential with respect to sentence length
- ⇒ Decoding is NP-complete [Knight, 1999]
- ⇒ Need to reduce search space
- risk free: hypothesis recombination
 - risky: histogram/threshold pruning

Hypothesis Recombination



- Different paths to the same partial translation

Hypothesis Recombination

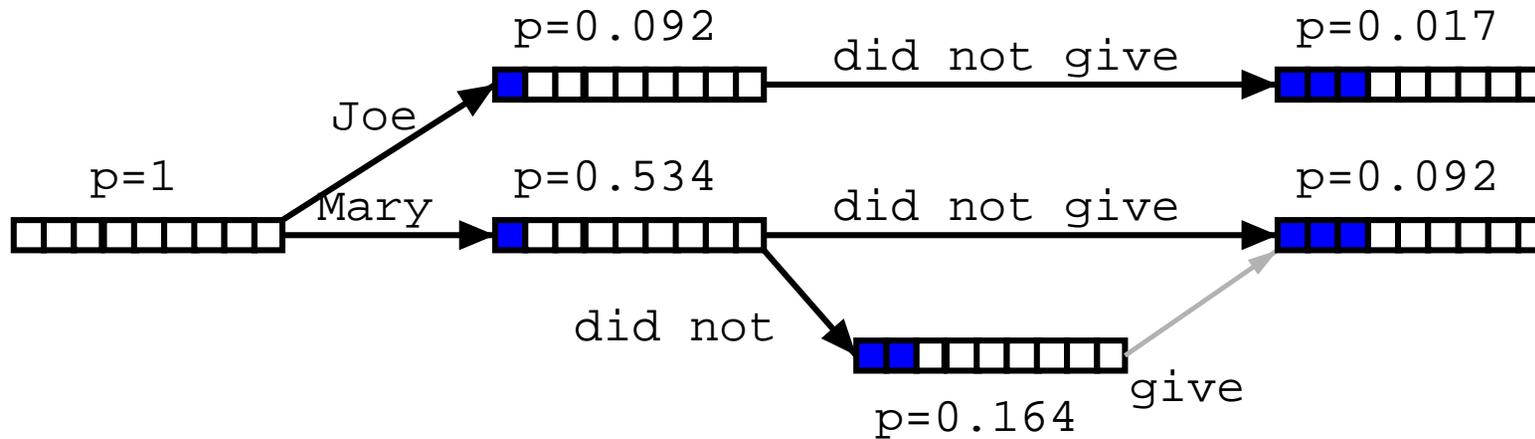


- Different paths to the same partial translation

⇒ Combine paths

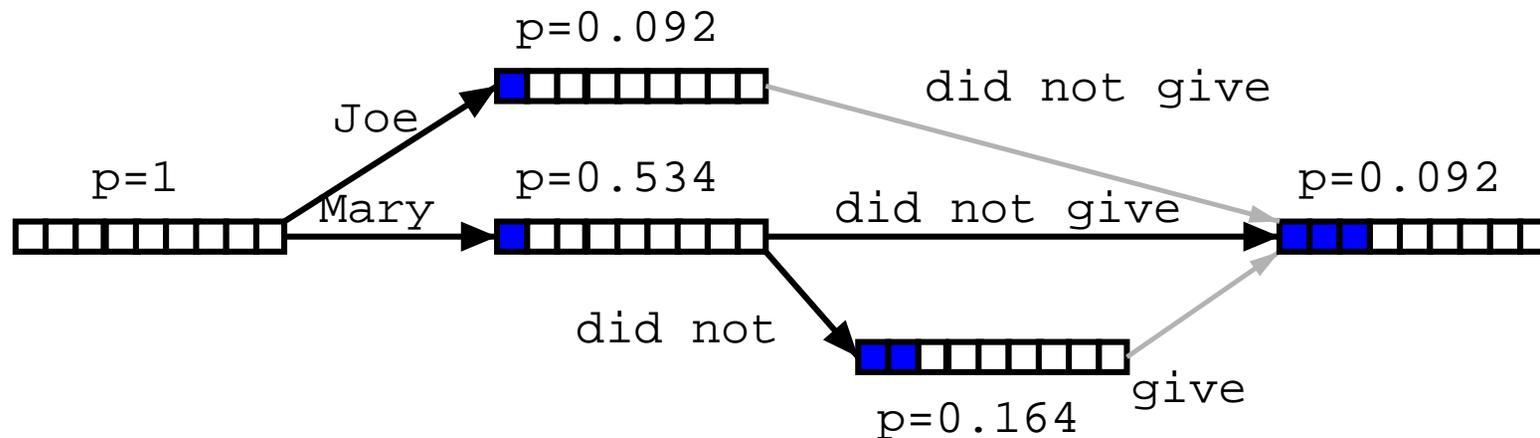
- drop weaker hypothesis
- keep pointer from worse path

Hypothesis Recombination



- Recombined hypotheses do not have to match completely
- No matter what is added, weaker path can be dropped, if:
 - last two English words match (matters for language model)
 - foreign word coverage vectors match (effects future path)

Hypothesis Recombination



- Recombined hypotheses do not have to match completely
- No matter what is added, weaker path can be dropped, if:
 - last two English words match (matters for language model)
 - foreign word coverage vectors match (effects future path)

⇒ Combine paths

Pruning

- Hypothesis recombination is not sufficient

⇒ Heuristically discard weak hypotheses

- Organize Hypothesis in stacks, e.g. by
 - same foreign words covered
 - same number of foreign words covered (Pharaoh does this)
 - same number of English words produced
- Compare hypotheses in stacks, discard bad ones
 - histogram pruning: keep top n hypotheses in each stack (e.g., $n=100$)
 - threshold pruning: keep hypotheses that are at most α times the cost of best hypothesis in stack (e.g., $\alpha = 0.001$)