The Complexity of Phrase Alignment Problems

John DeNero and Dan Klein
{denero, klein}@cs.berkeley.edu

University of California
Berkeley

The Phrase Alignment Problem

The Phrase Alignment Problem

Japan to freeze aid to Russia .

The Phrase Alignment Problem

Japan to freeze aid to Russia .

日本

冻结

向

俄

提供

援助

。

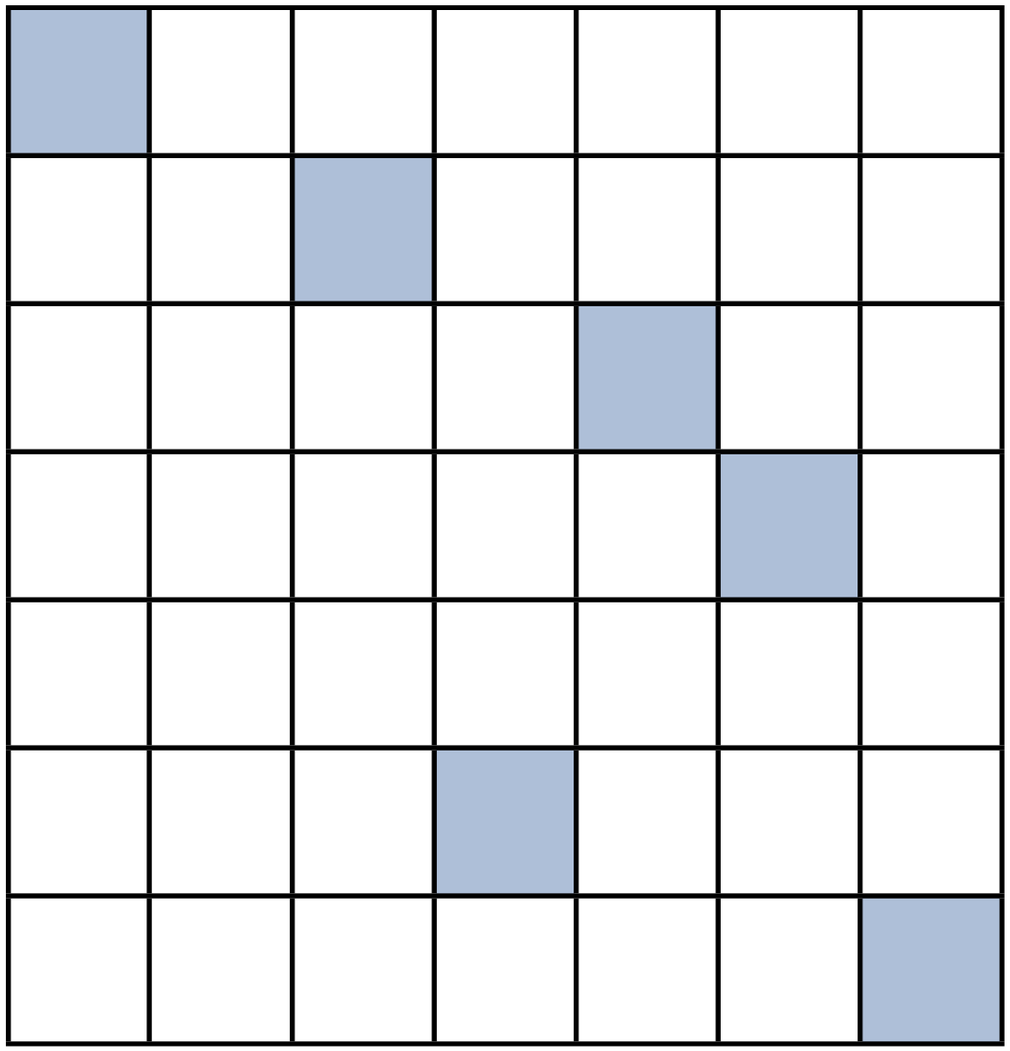
The Phrase Alignment Problem

Japan to freeze aid to Russia .

	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Phrase Alignment Problem

Japan to freeze aid to Russia .



	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Phrase Alignment Problem

Japan to freeze aid to Russia .

	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Phrase Alignment Problem

Japan | to freeze | aid | to | Russia | .

	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

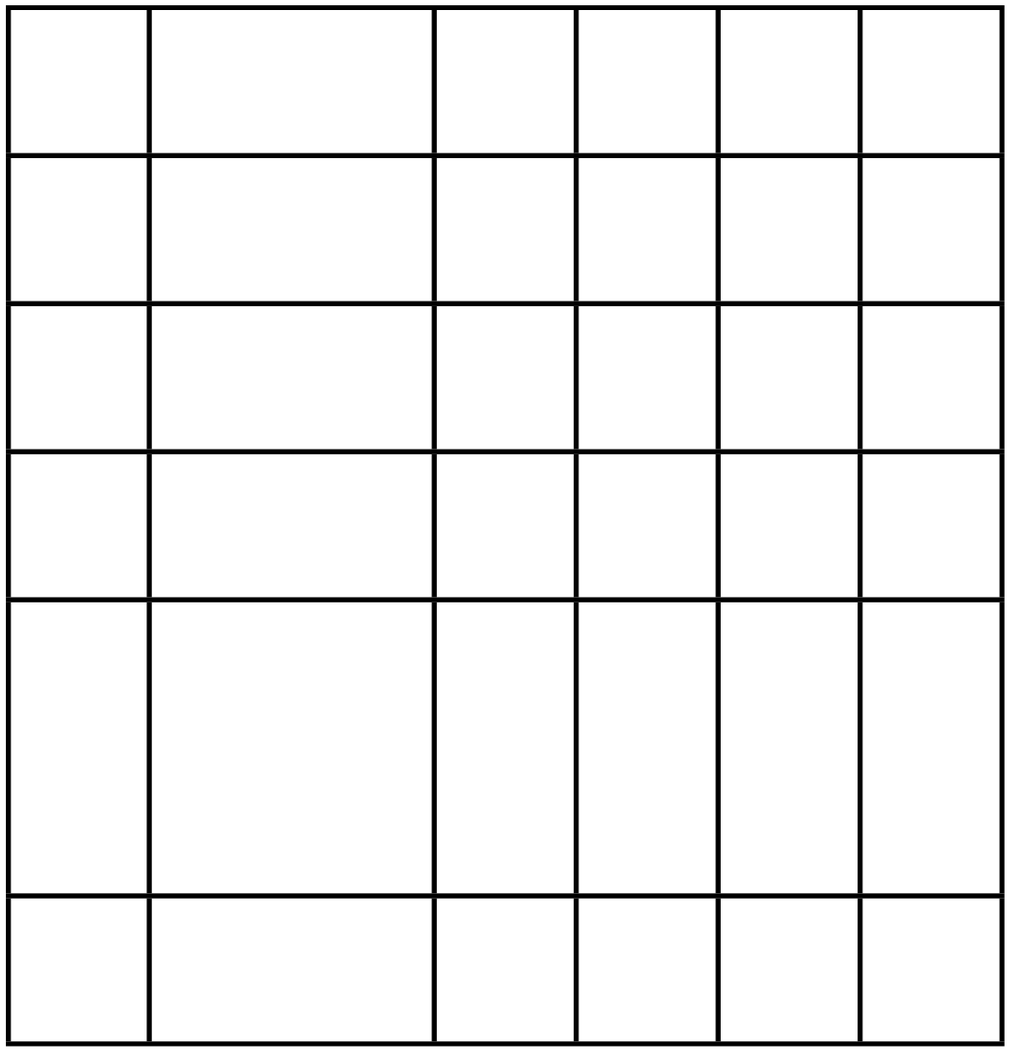
The Phrase Alignment Problem

Japan | to freeze | aid | to | Russia | .

	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Phrase Alignment Problem

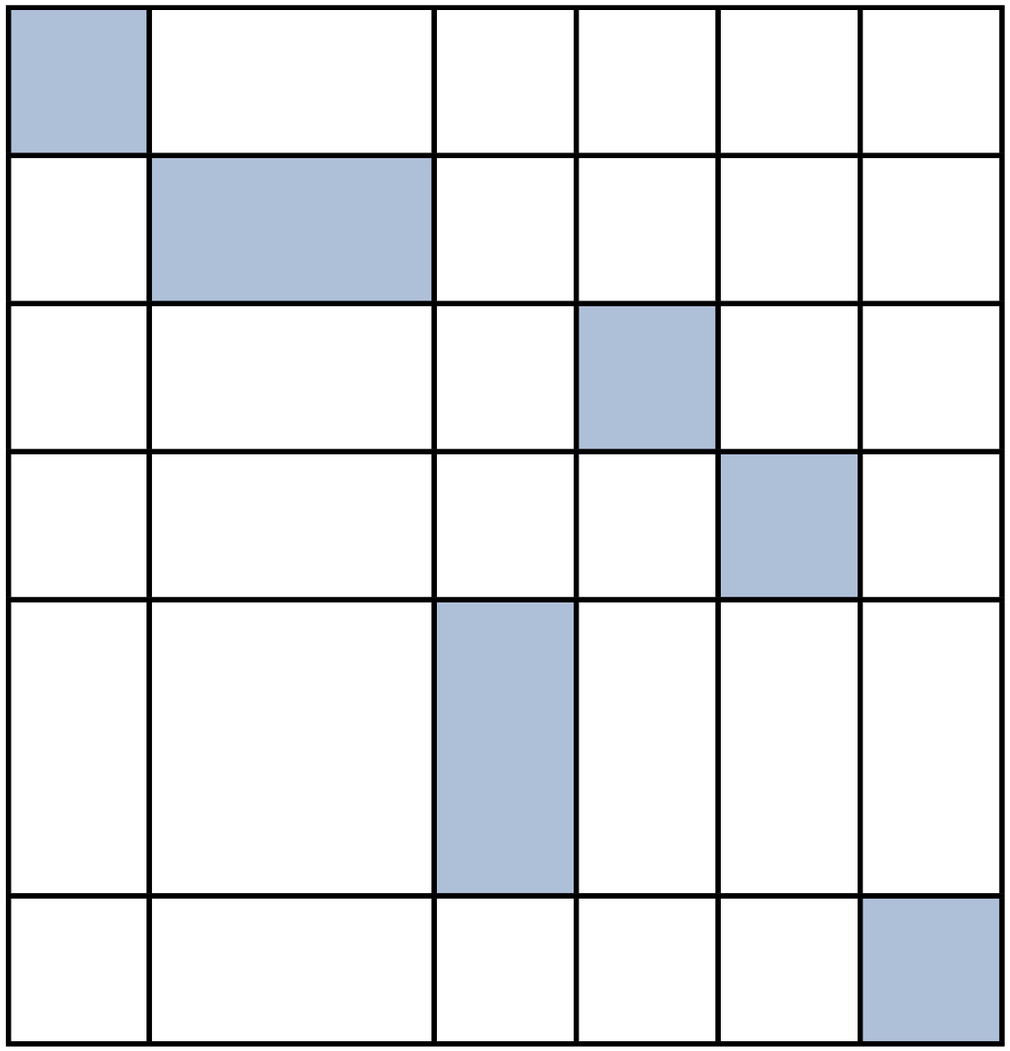
Japan | to freeze | aid | to | Russia | .



	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Phrase Alignment Problem

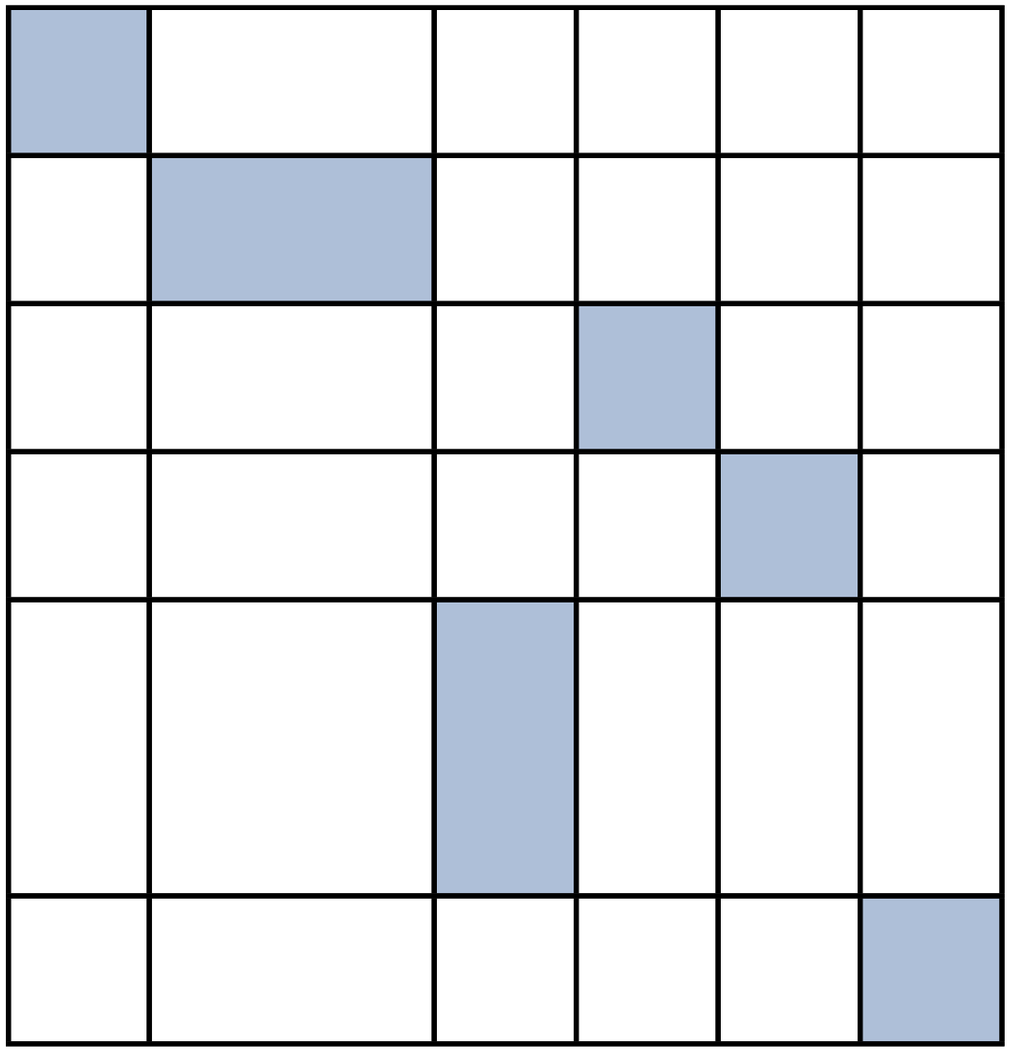
Japan | to freeze | aid | to | Russia | .



	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Phrase Alignment Problem

Japan | to freeze | aid | to | Russia | .

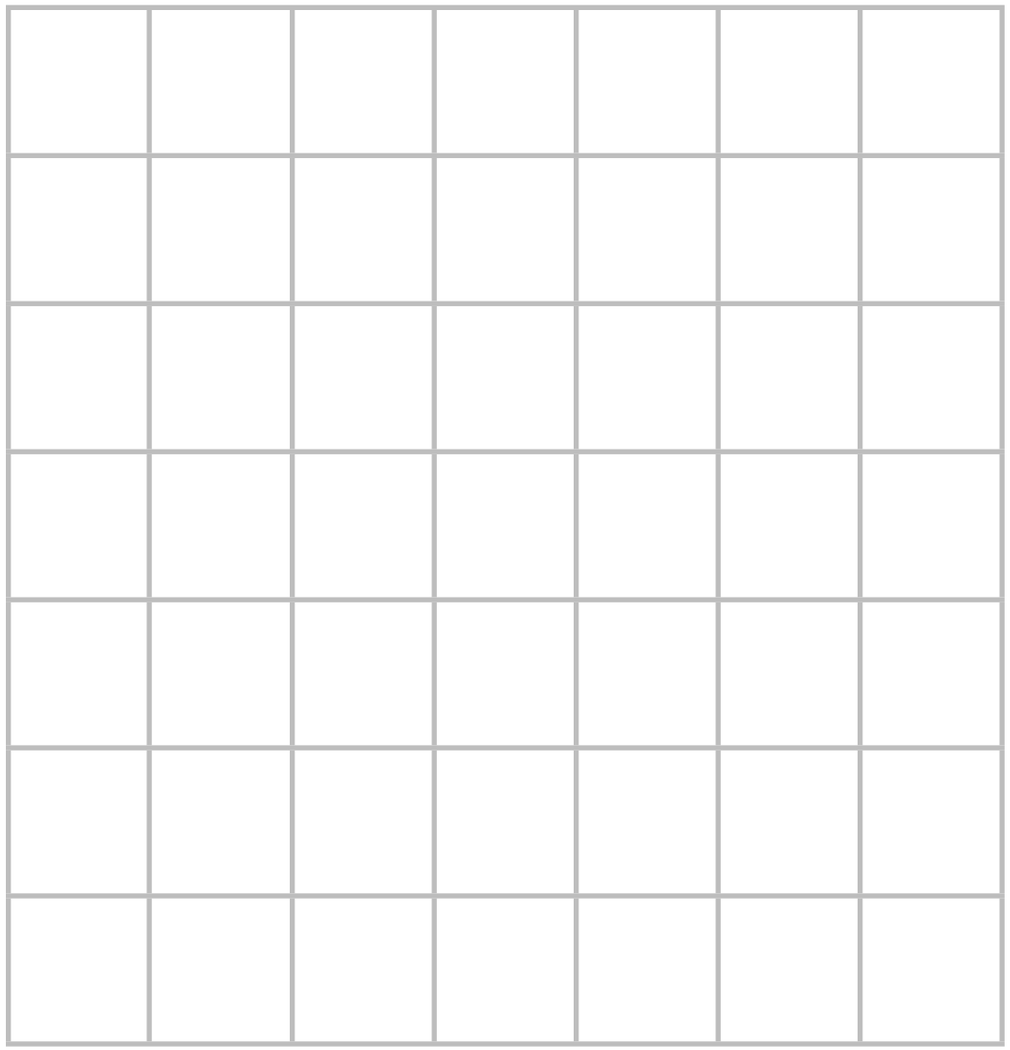


	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

Phrase alignments are one-to-one and onto

The Weighted Phrase Alignment Problem

Japan to freeze aid to Russia .



	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Weighted Phrase Alignment Problem

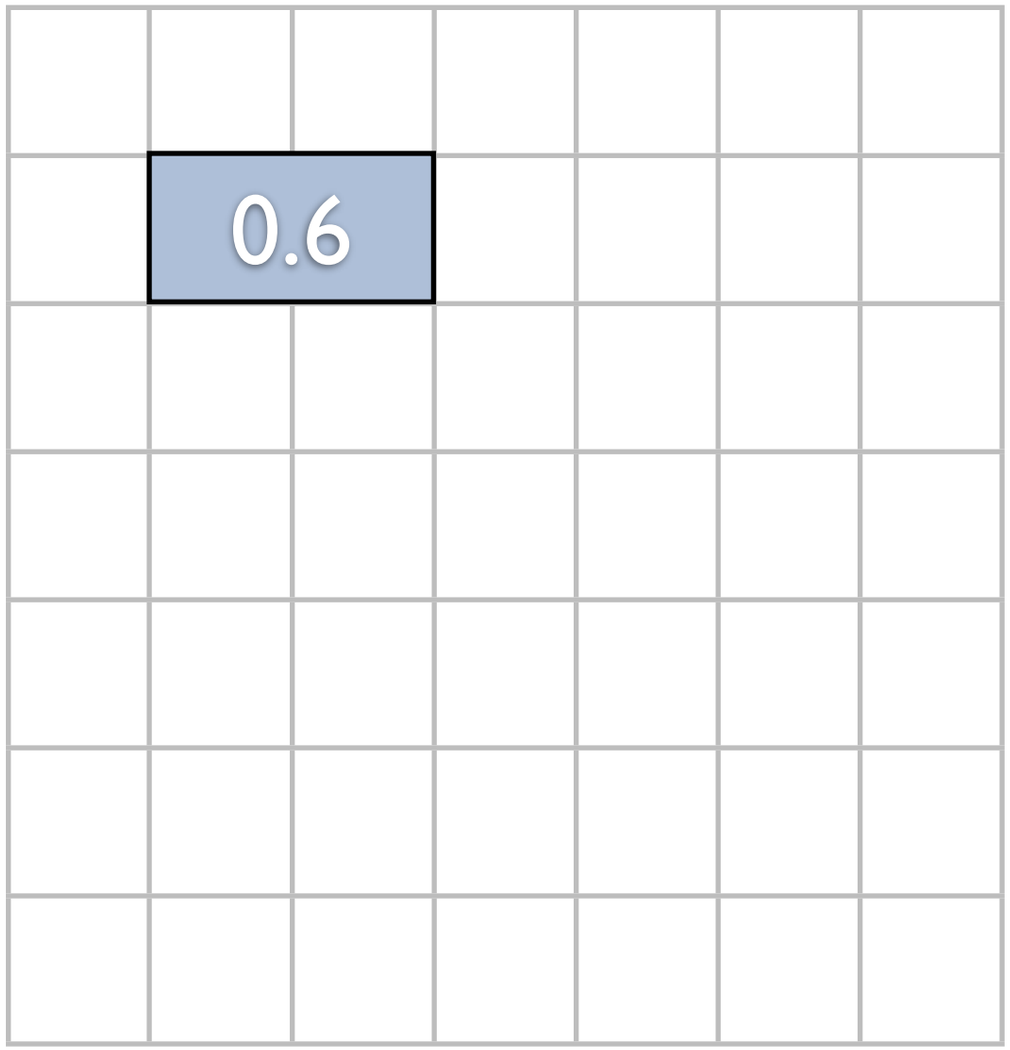
Japan to freeze aid to Russia .

0.9						

	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Weighted Phrase Alignment Problem

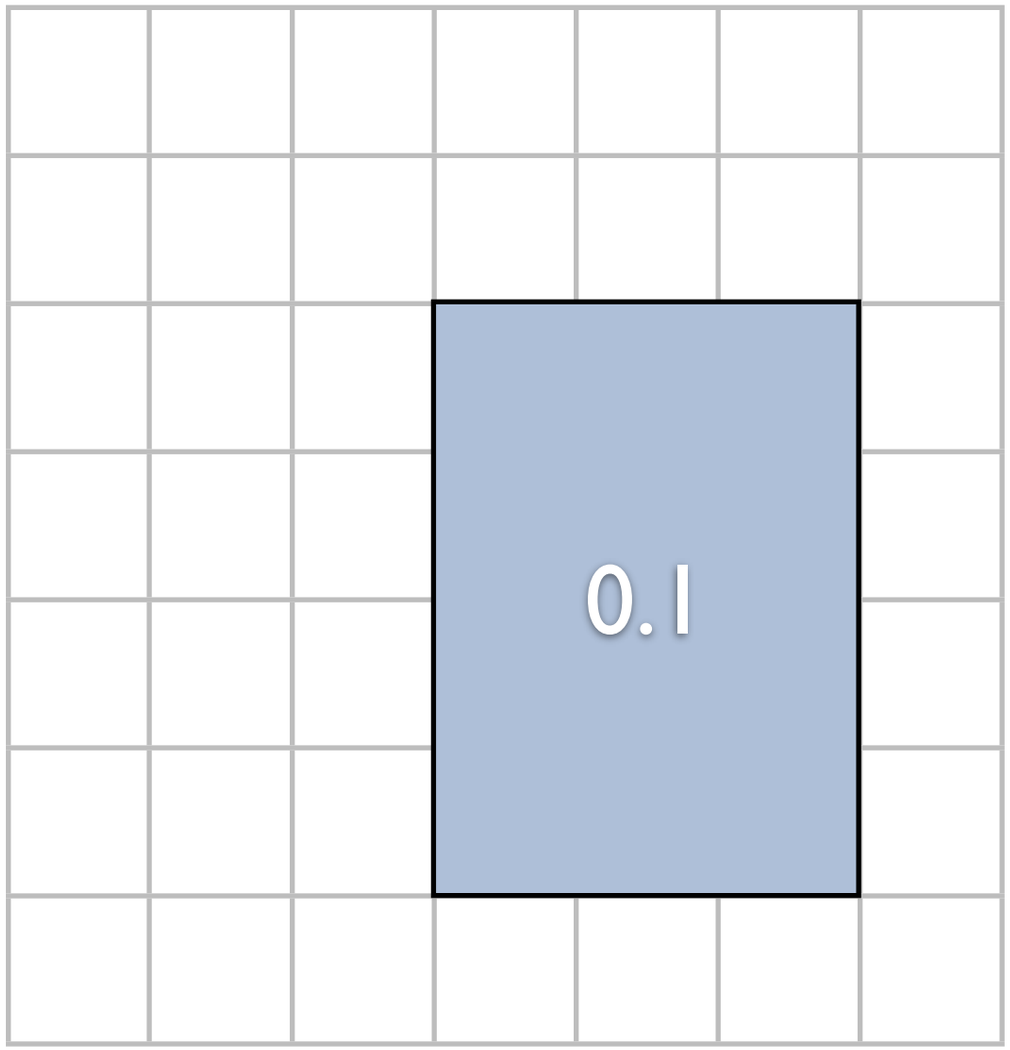
Japan to freeze aid to Russia .



	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Weighted Phrase Alignment Problem

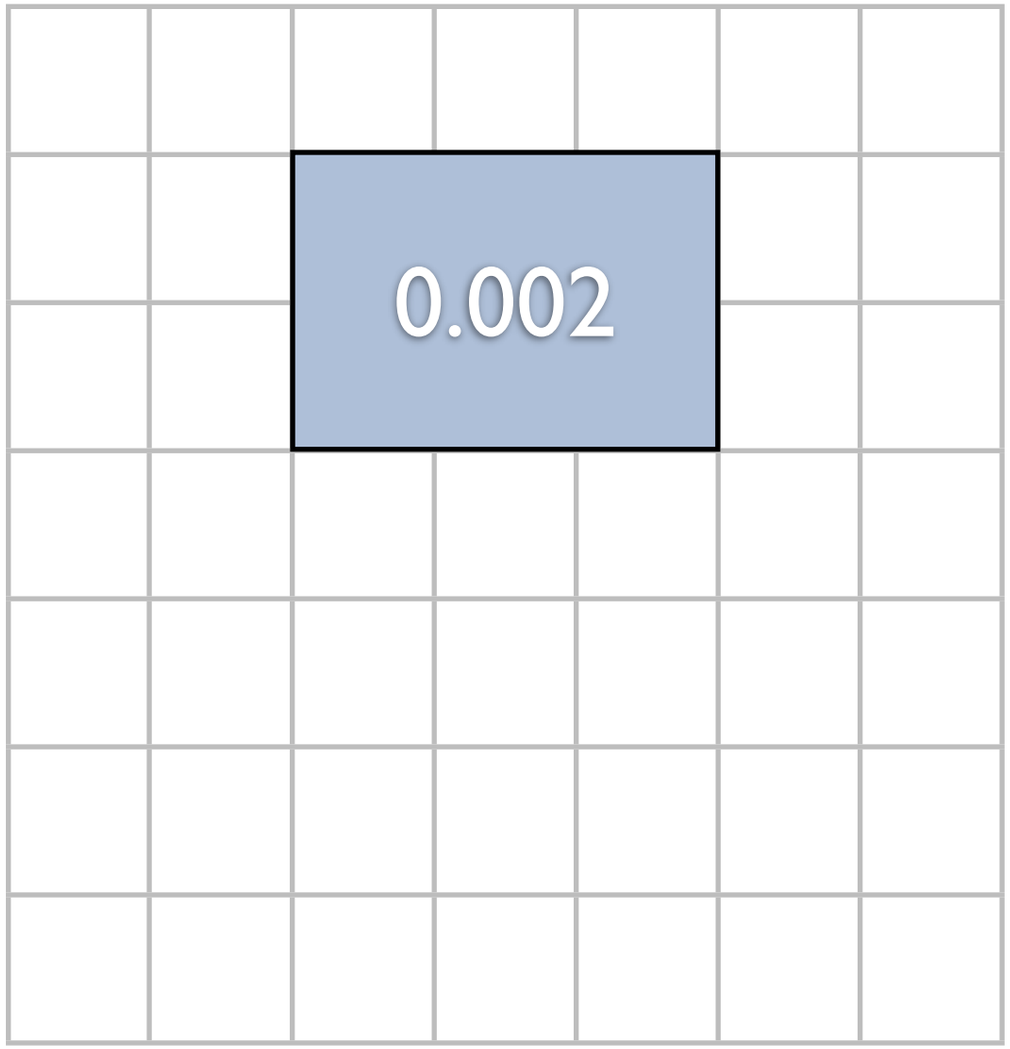
Japan to freeze aid to Russia .



	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Weighted Phrase Alignment Problem

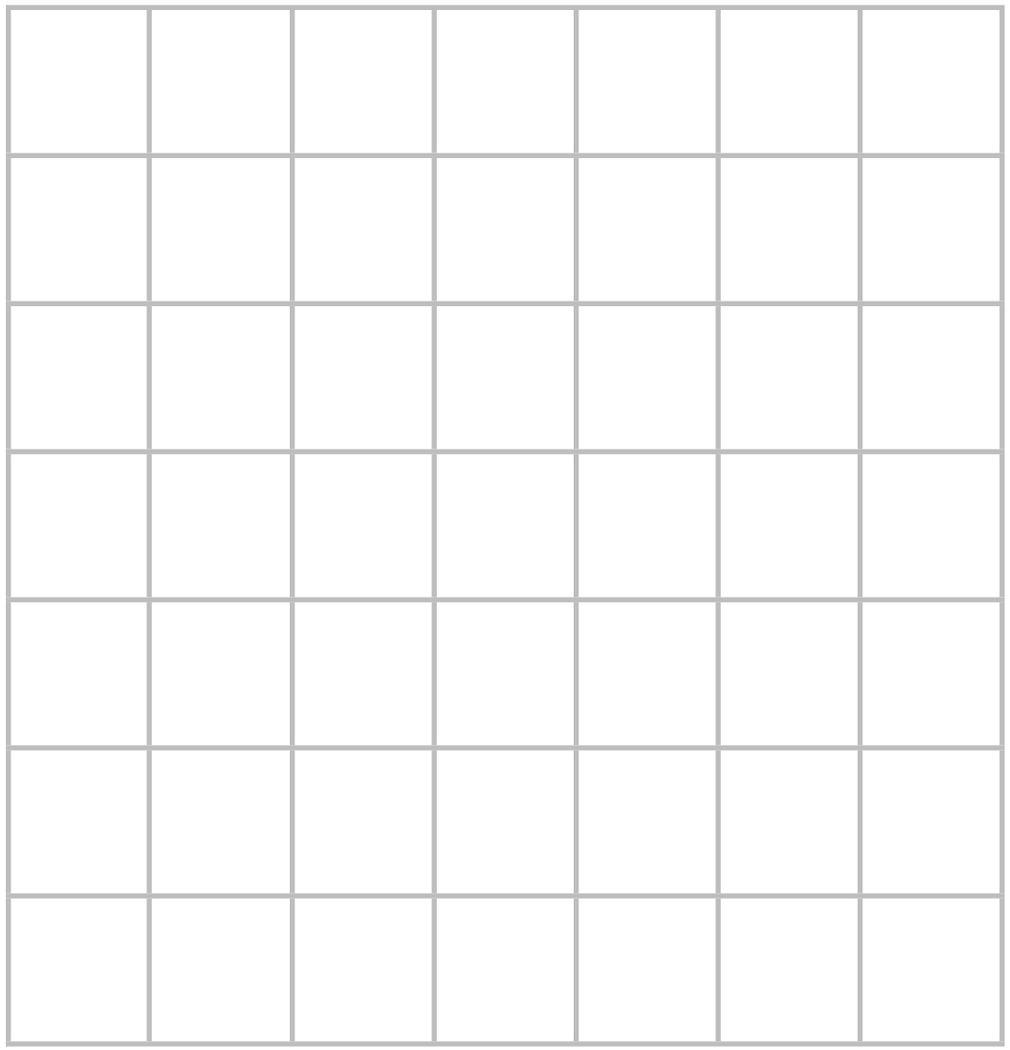
Japan to freeze aid to Russia .



	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Weighted Phrase Alignment Problem

Japan to freeze aid to Russia .



	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Weighted Phrase Alignment Problem

Japan to freeze aid to Russia .

0.9					
	0.6				
			0.8		
				0.9	
		0.7			
					0.9

	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

The Weighted Phrase Alignment Problem

Japan to freeze aid to Russia .

0.9					
	0.6				
			0.8		
				0.9	
		0.7			
					0.9

	Pinyin	Gloss
日本	ri4 ben3	*Japan*
冻结	dong4 jie2	*freeze*
向	xiang4	*to*
俄	e2	*Russia*
提供	ti2 gong1	*supply*
援助	yuan2 zhu4	*assistance*
。	。	。

$0.9 \cdot 0.6 \cdot 0.7 \cdot 0.8 \cdot 0.9 \cdot 0.9 = 0.24$

Applications Involving Phrase Alignments

Applications Involving Phrase Alignments

- Inference under a phrase alignment model

Applications Involving Phrase Alignments

- Inference under a phrase alignment model
- Viterbi Training for a phrase alignment model

Applications Involving Phrase Alignments

- Inference under a phrase alignment model
- Viterbi Training for a phrase alignment model
- Forced decoding for phrase-based systems

Applications Involving Phrase Alignments

- Inference under a phrase alignment model
- Viterbi Training for a phrase alignment model
- Forced decoding for phrase-based systems
- Improved decoding for word alignment models

University of California
C A L
N L P
Berkeley

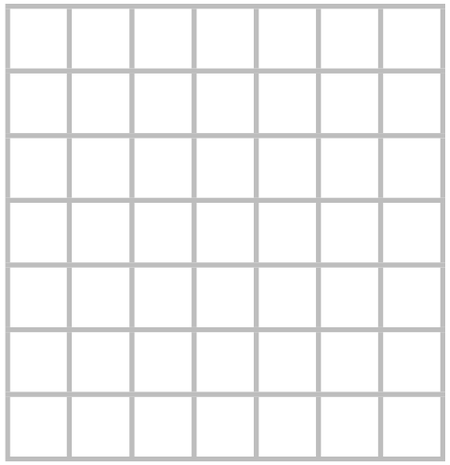
Related Problems are Polynomial

Related Problems are Polynomial

MATCHING: Given segmentations, find the maximal matching.

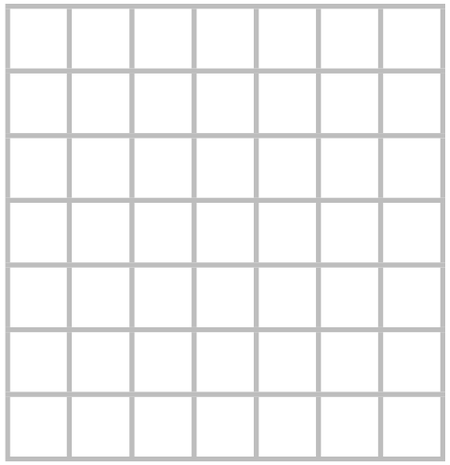
Related Problems are Polynomial

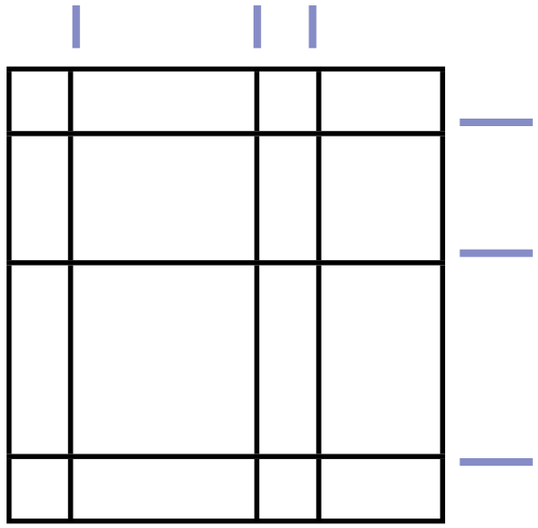
MATCHING: Given segmentations, find the maximal matching.



Related Problems are Polynomial

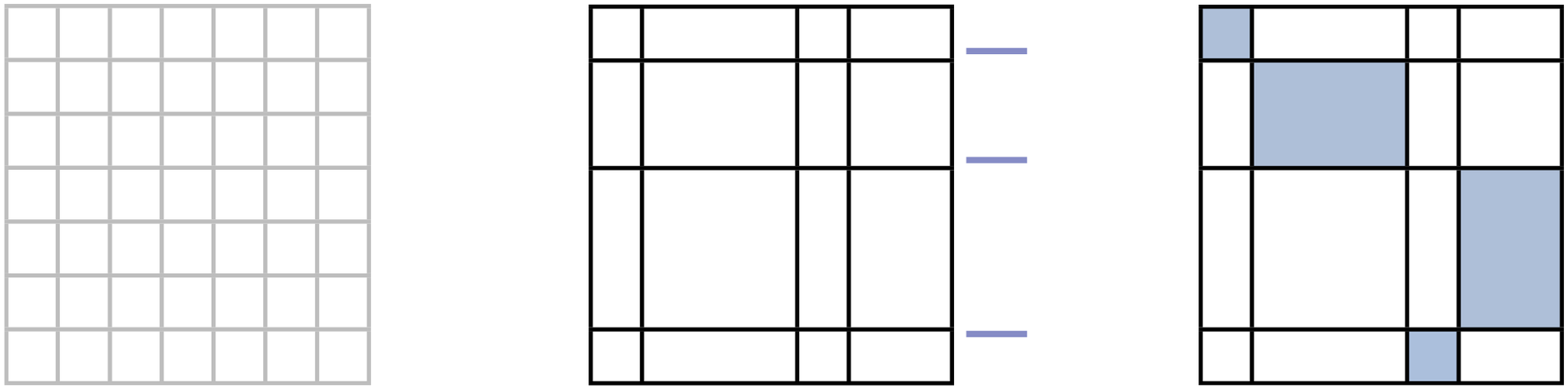
MATCHING: Given segmentations, find the maximal matching.





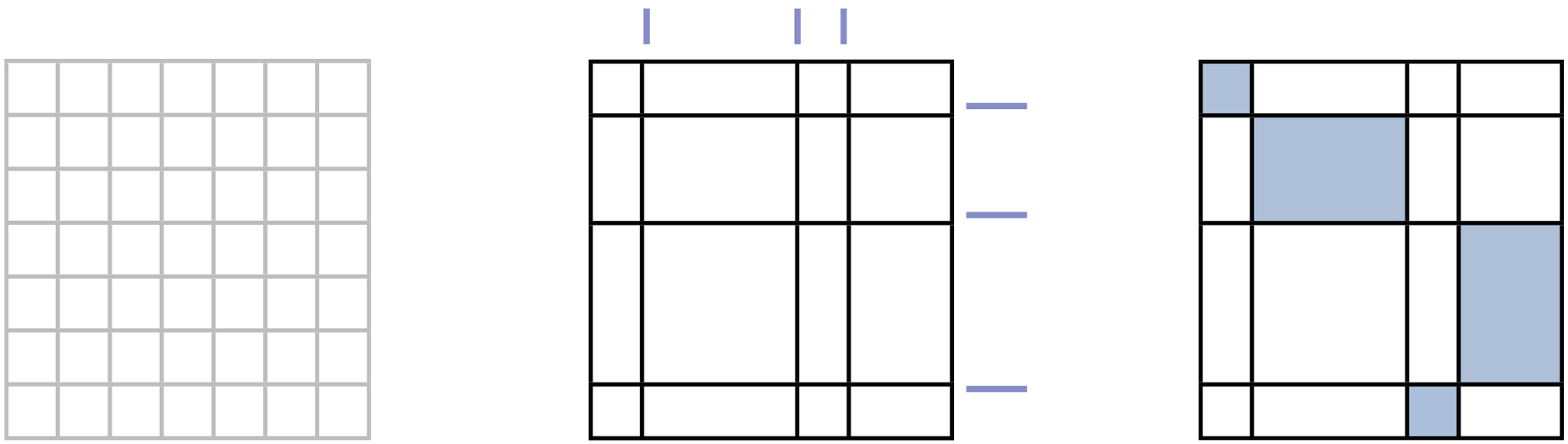
Related Problems are Polynomial

MATCHING: Given segmentations, find the maximal matching.



Related Problems are Polynomial

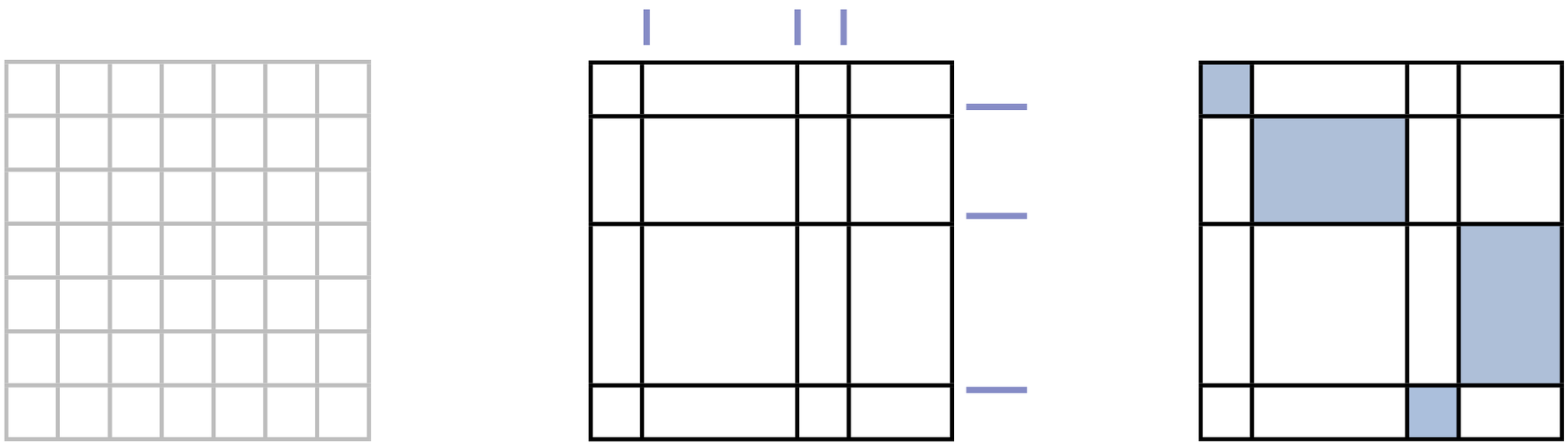
MATCHING: Given segmentations, find the maximal matching.



PARTITIONING: Given phrase weights, find the max segmentation.

Related Problems are Polynomial

MATCHING: Given segmentations, find the maximal matching.

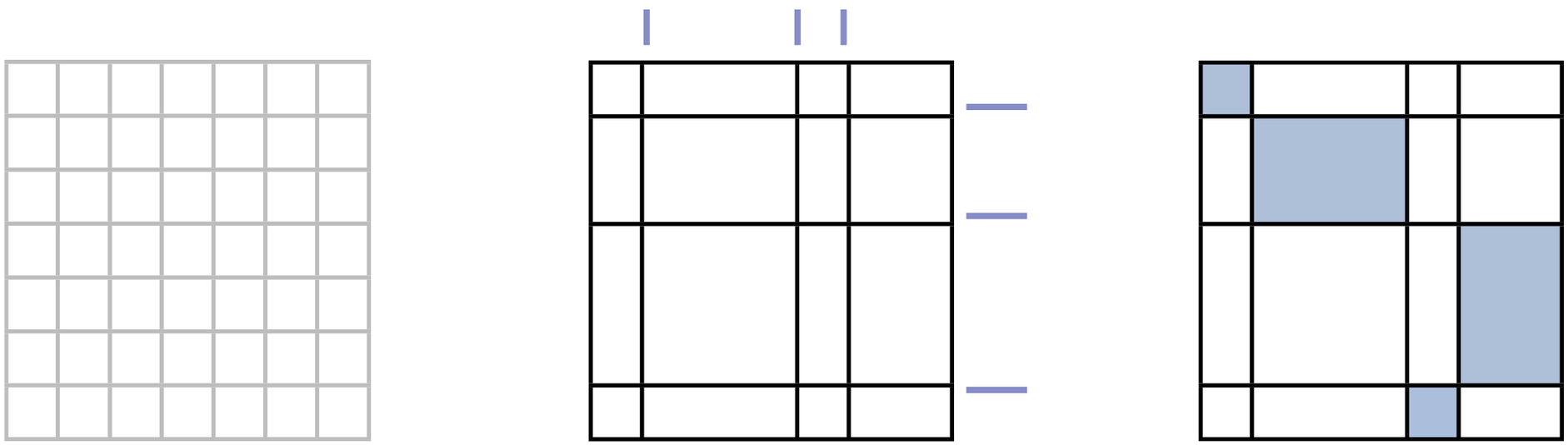


PARTITIONING: Given phrase weights, find the max segmentation.

Japan to freeze aid to Russia .

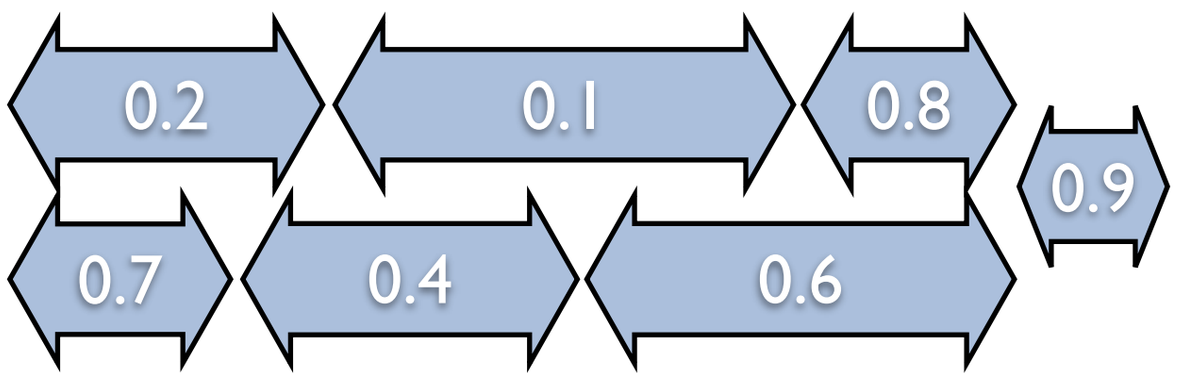
Related Problems are Polynomial

MATCHING: Given segmentations, find the maximal matching.



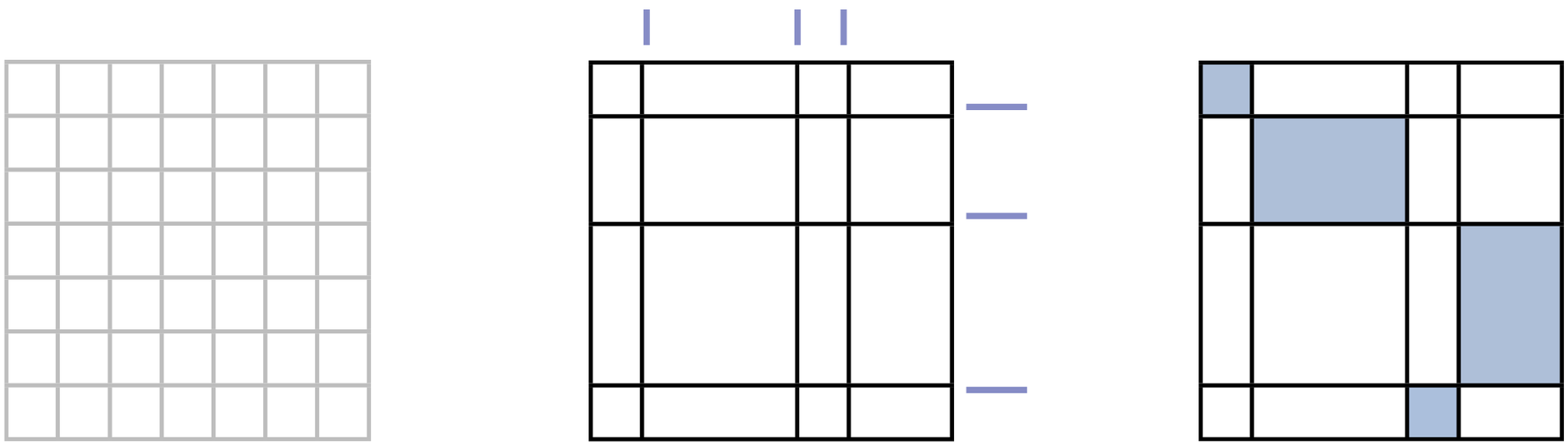
PARTITIONING: Given phrase weights, find the max segmentation.

Japan to freeze aid to Russia .

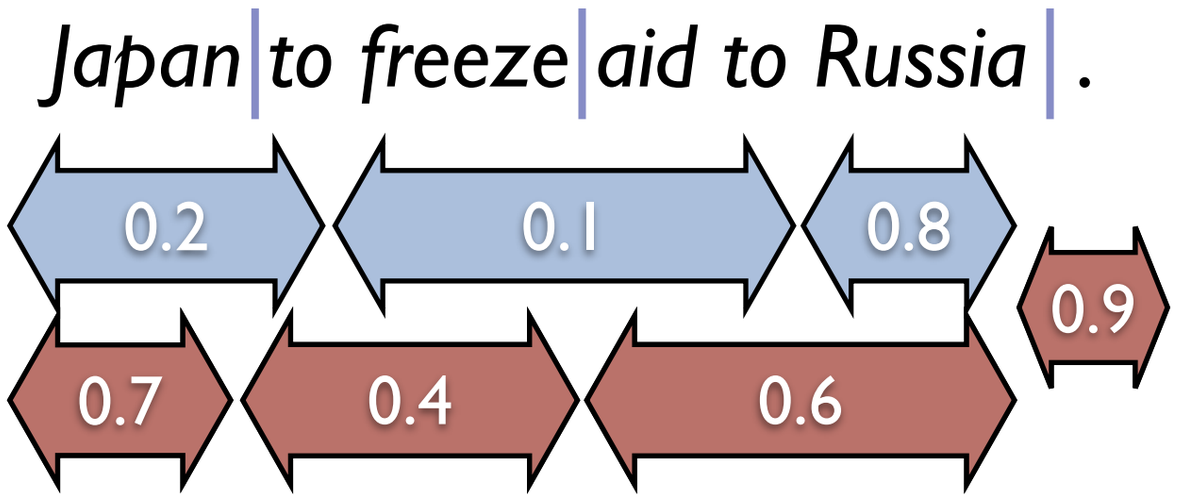


Related Problems are Polynomial

MATCHING: Given segmentations, find the maximal matching.



PARTITIONING: Given phrase weights, find the max segmentation.



The Phrase Alignment Problems

Given a sentence pair and scores for all phrase pairs:

The Phrase Alignment Problems

Given a sentence pair and scores for all phrase pairs:

- PHRASE OPTIMIZATION: Find the highest scoring phrase alignment.

The Phrase Alignment Problems

Given a sentence pair and scores for all phrase pairs:

- PHRASE OPTIMIZATION: Find the highest scoring phrase alignment.

- PHRASE DECISION: Determine if there is a phrase alignment with score $\geq$ t.

Reducing 3-SAT to PHRASE DECISION

Anatomy of a reduction

- Choose an old problem that is known to be NP-hard.
- Show that we can solve that old problem easily if we can solve our new problem.
- Conclude that if the new problem were in P, the old problem would be too (which it's not, we think).

Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$

Reducing 3-SAT to PHRASE DECISION

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$

Satisfying assignment:

v_1 is *true*

v_2 is *false*

v_3 is *false*

Reducing 3-SAT to PHRASE DECISION

$$\boxed{v_1} \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \boxed{\bar{v}_3}$$

$$\bar{v}_1 \vee \boxed{\bar{v}_2} \vee \bar{v}_3$$

$$\bar{v}_1 \vee \boxed{\bar{v}_2} \vee v_3$$

Satisfying assignment:

v_1 is *true*

v_2 is *false*

v_3 is *false*

Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$

Reducing 3-SAT to Phrase Decision

$$v_1 \quad \bar{v}_1 \quad v_2 \quad \bar{v}_2 \quad v_3 \quad \bar{v}_3$$

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$

Reducing 3-SAT to Phrase Decision

	v_1	$\bar{v}_1$	v_2	$\bar{v}_2$	v_3	$\bar{v}_3$
$v_1 \vee v_2 \vee v_3$						
$\bar{v}_1 \vee v_2 \vee \bar{v}_3$						
$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$						
$\bar{v}_1 \vee \bar{v}_2 \vee v_3$						

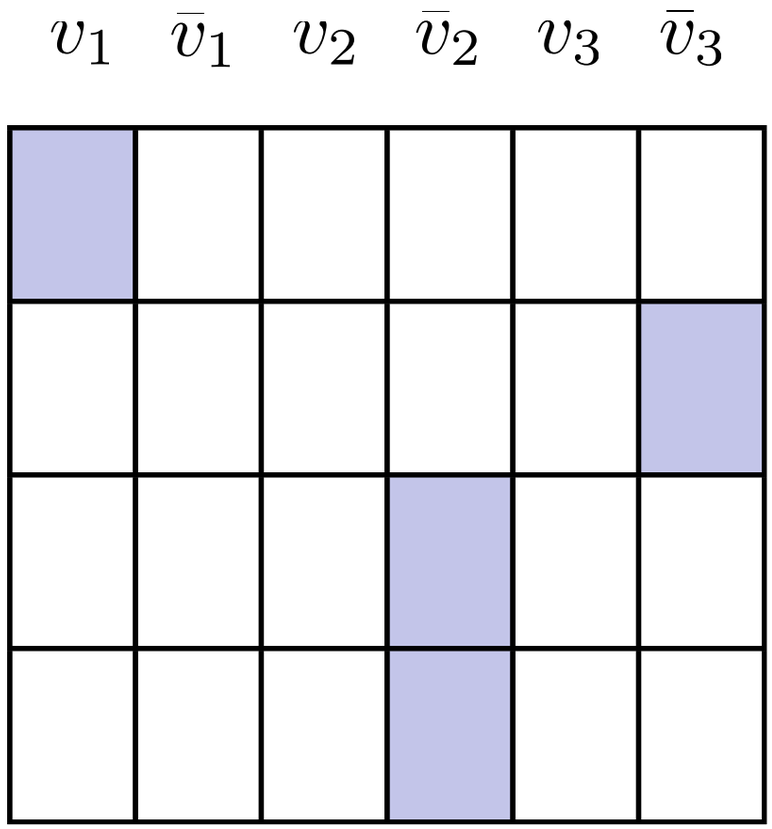
Reducing 3-SAT to Phrase Decision

$v_1 \vee v_2 \vee v_3$

$\bar{v}_1 \vee v_2 \vee \bar{v}_3$

$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$

$\bar{v}_1 \vee \bar{v}_2 \vee v_3$



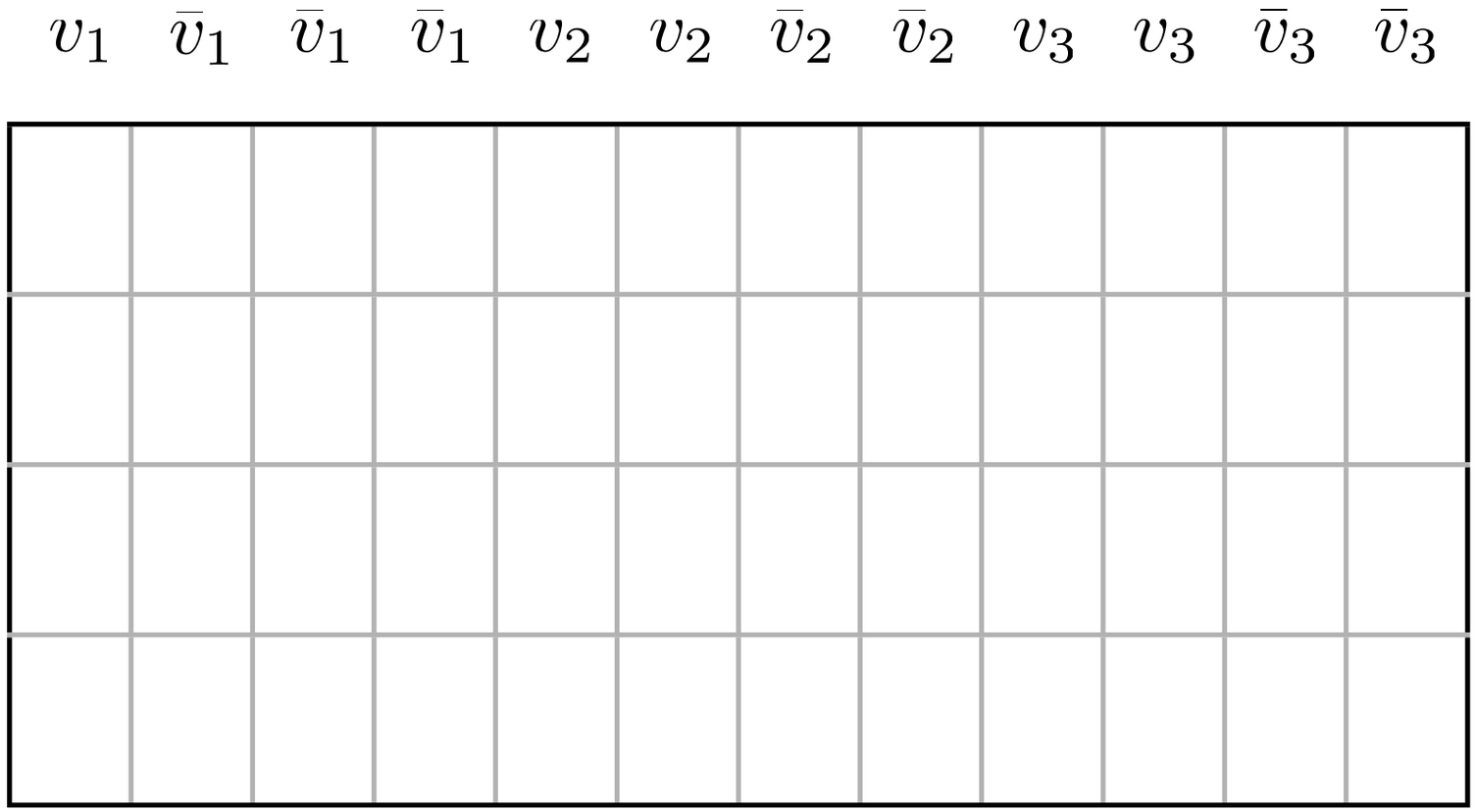
Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

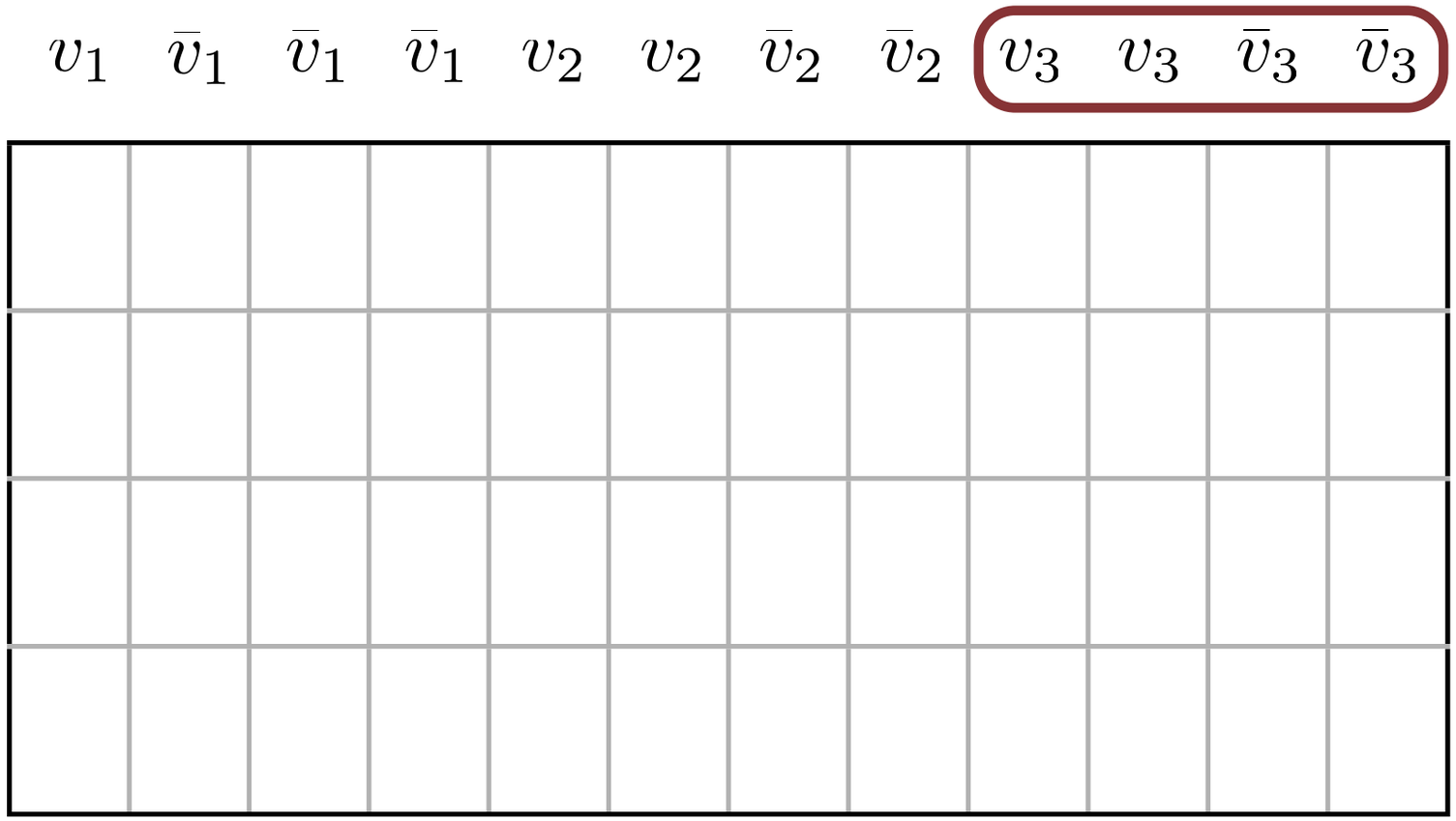
$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$



Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$
$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$
$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$
$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$



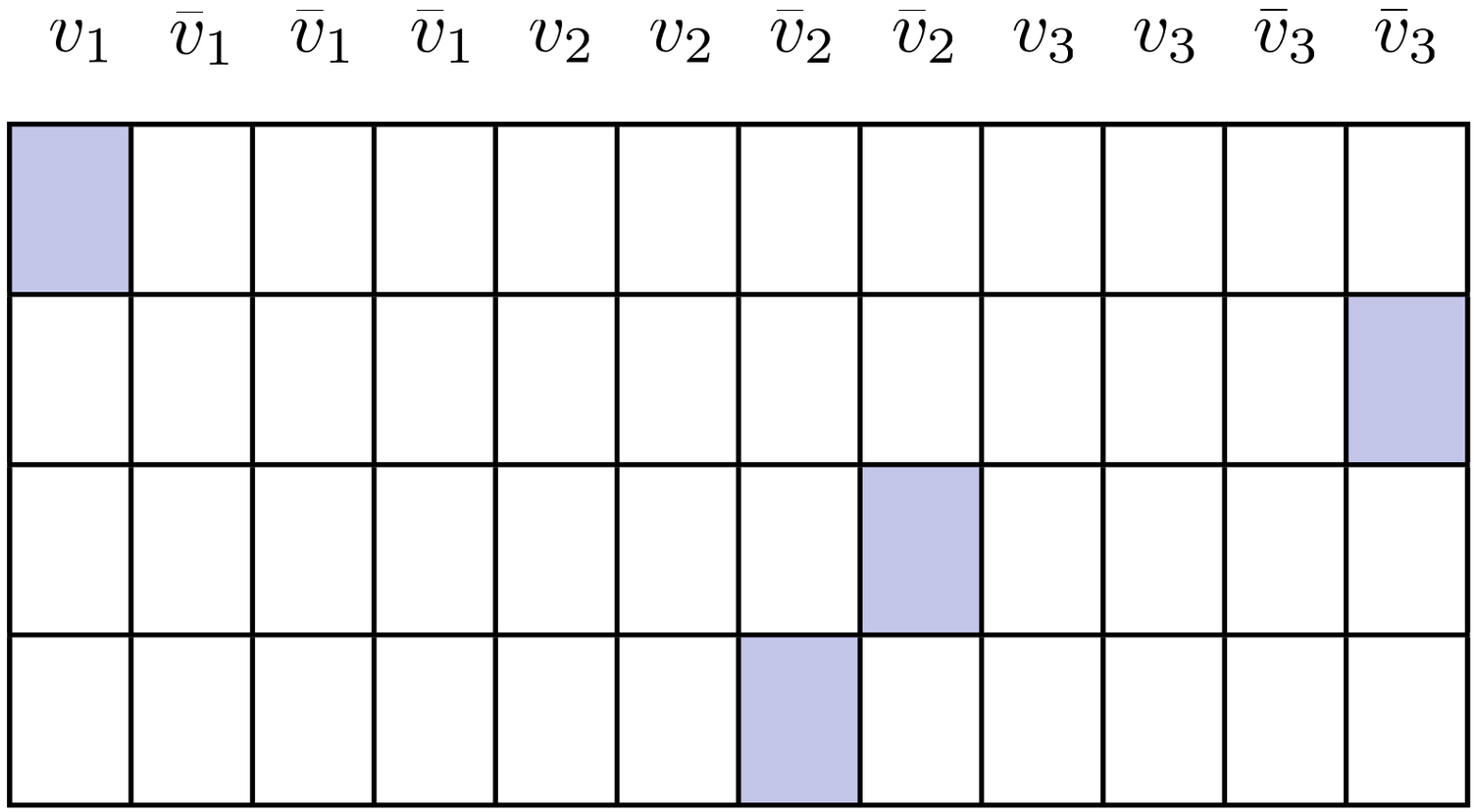
Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$



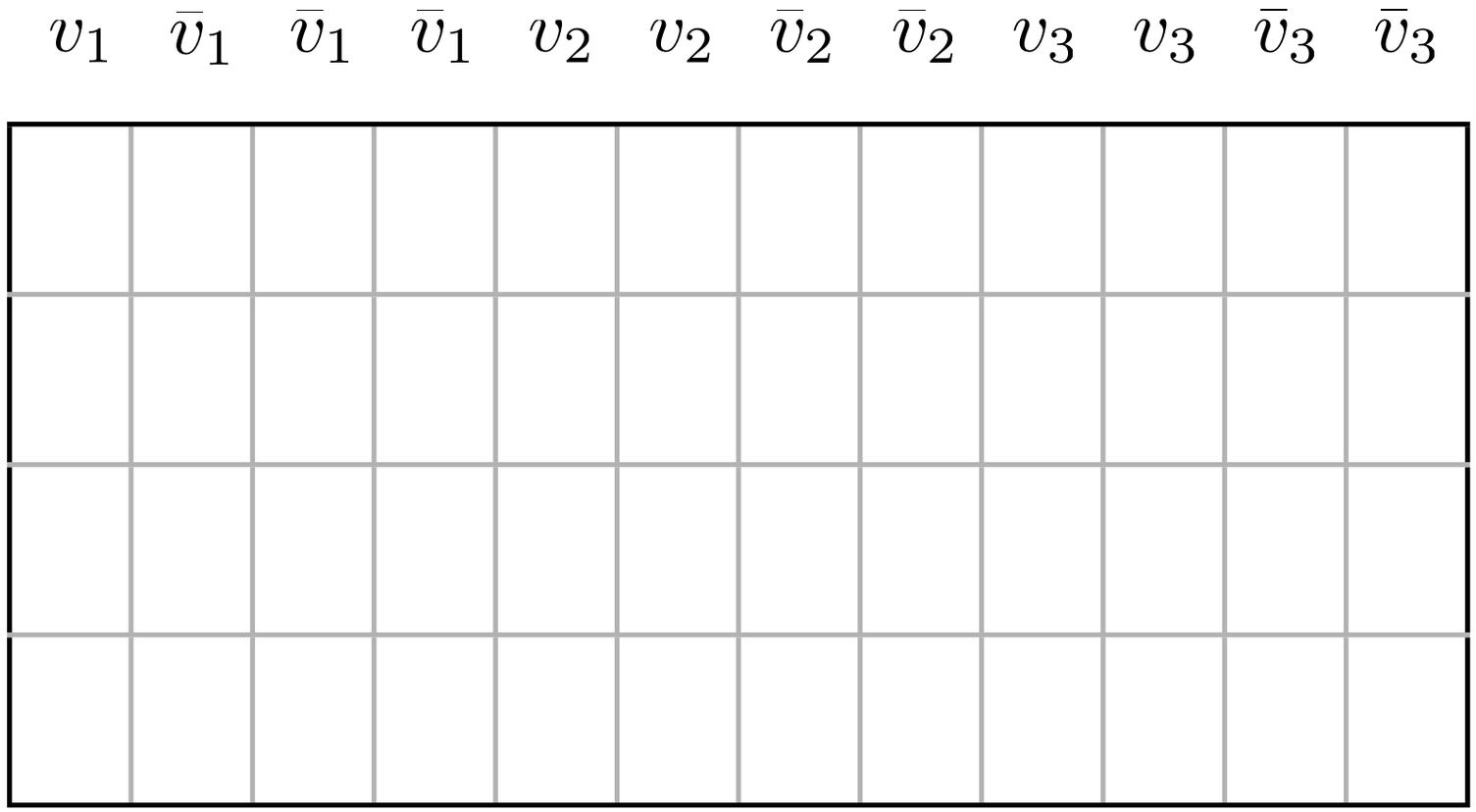
Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$



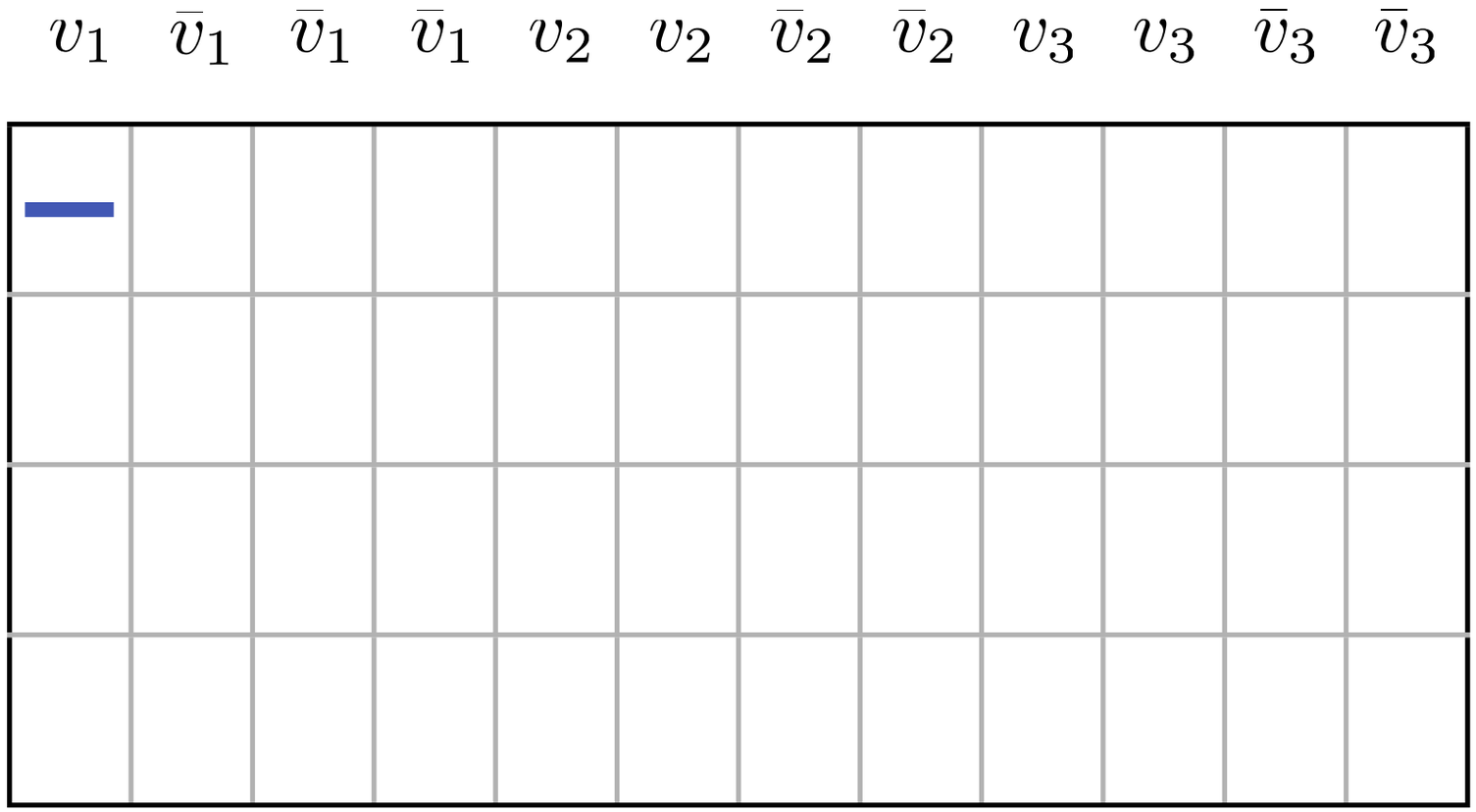
Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$



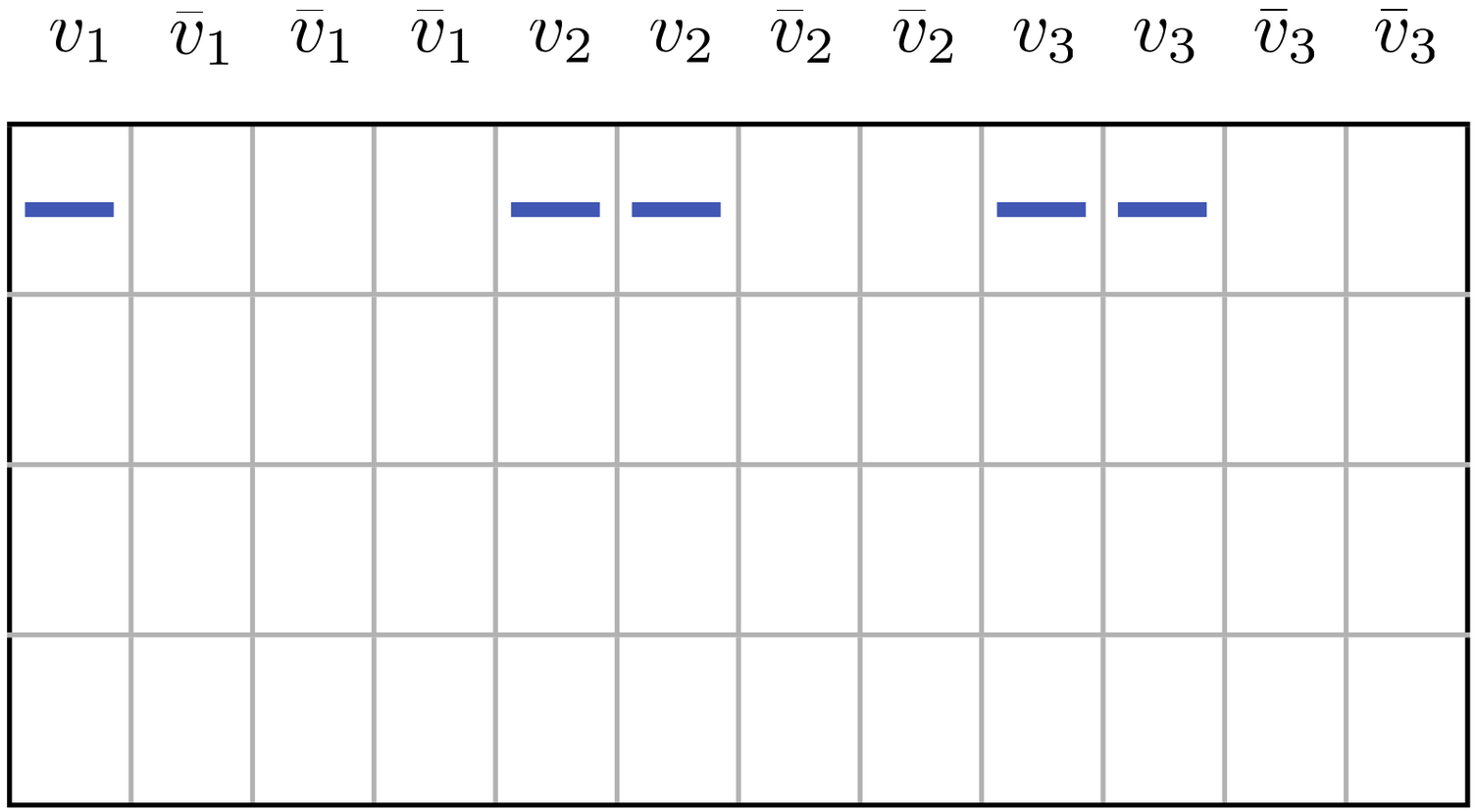
Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$



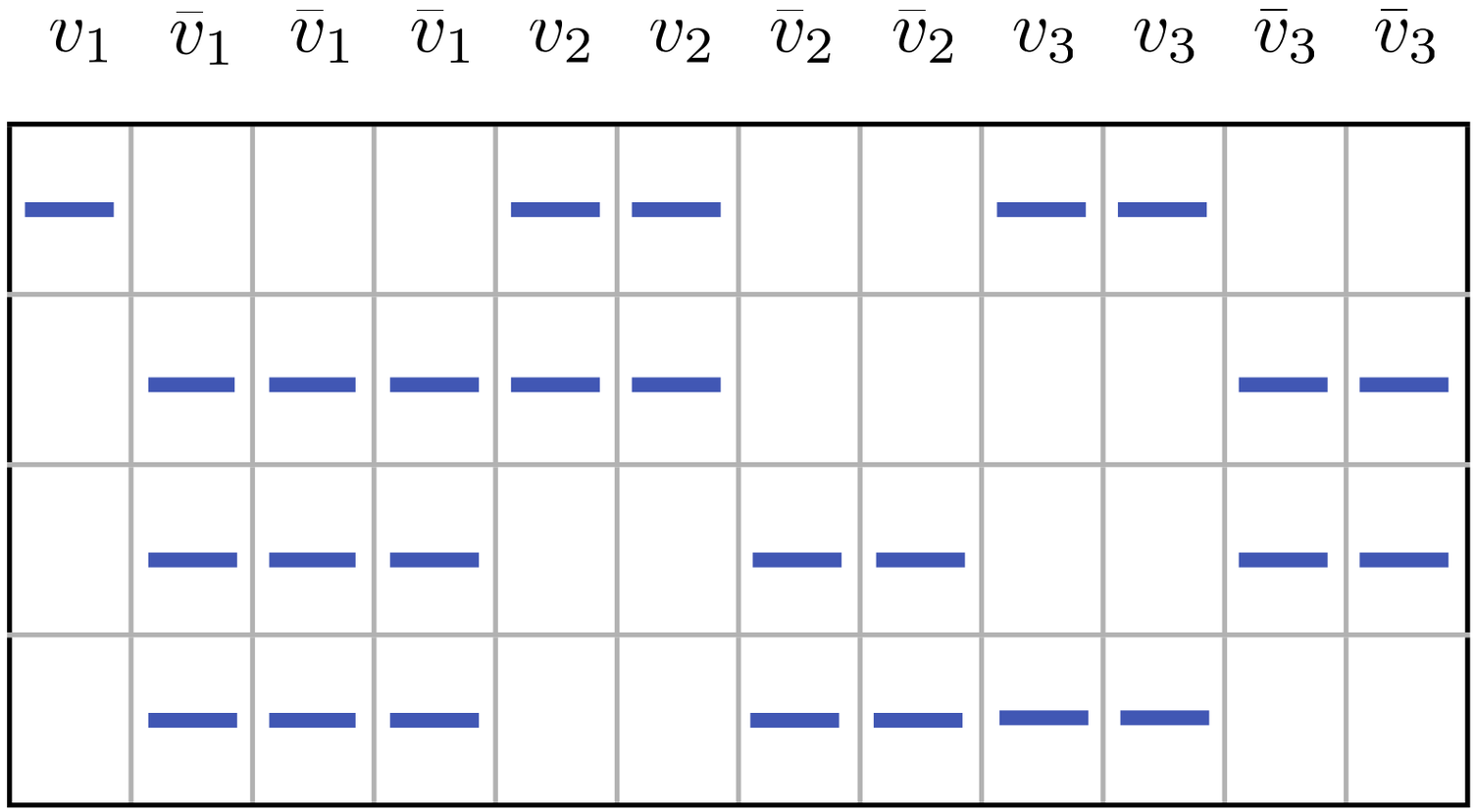
Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$



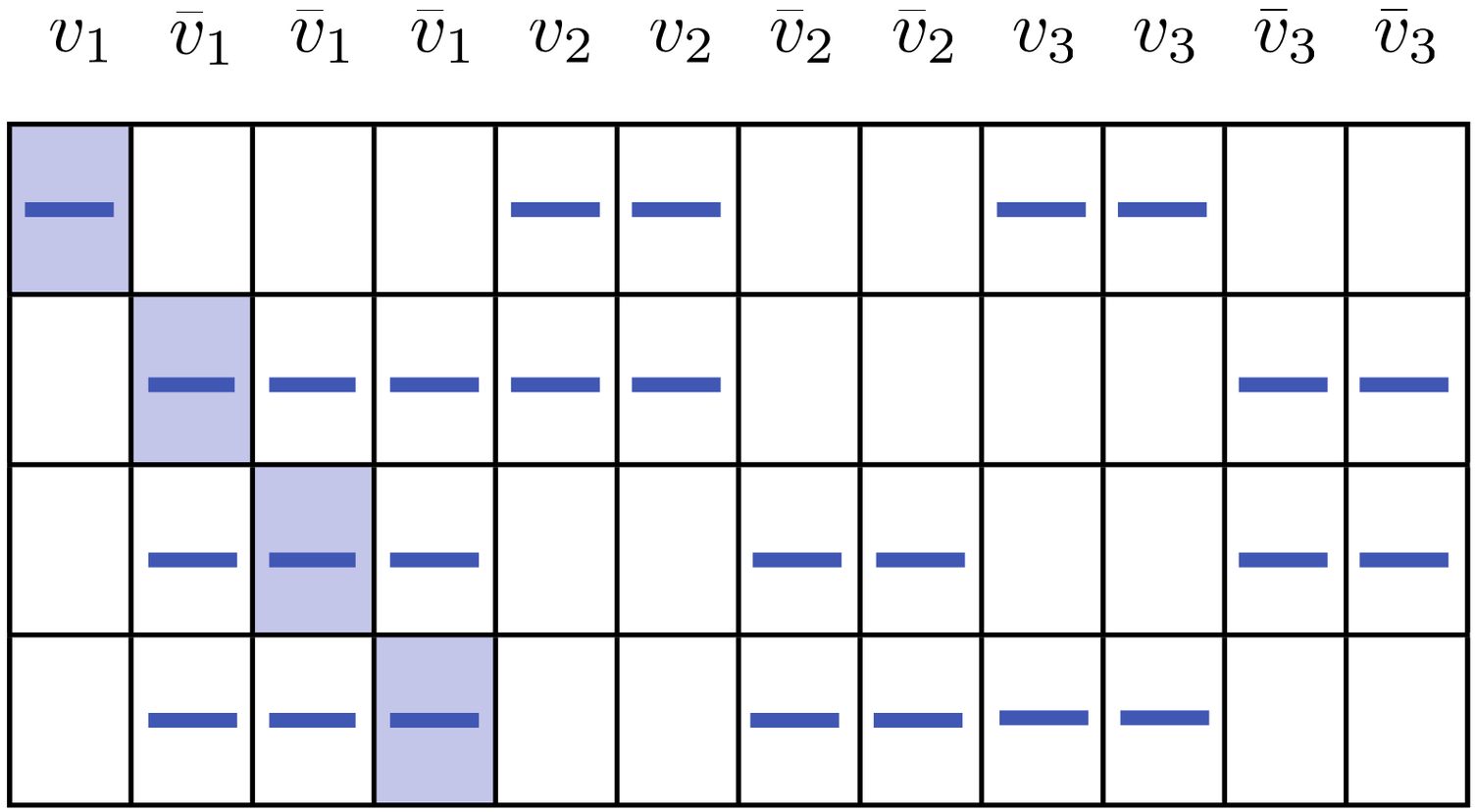
Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$

$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$



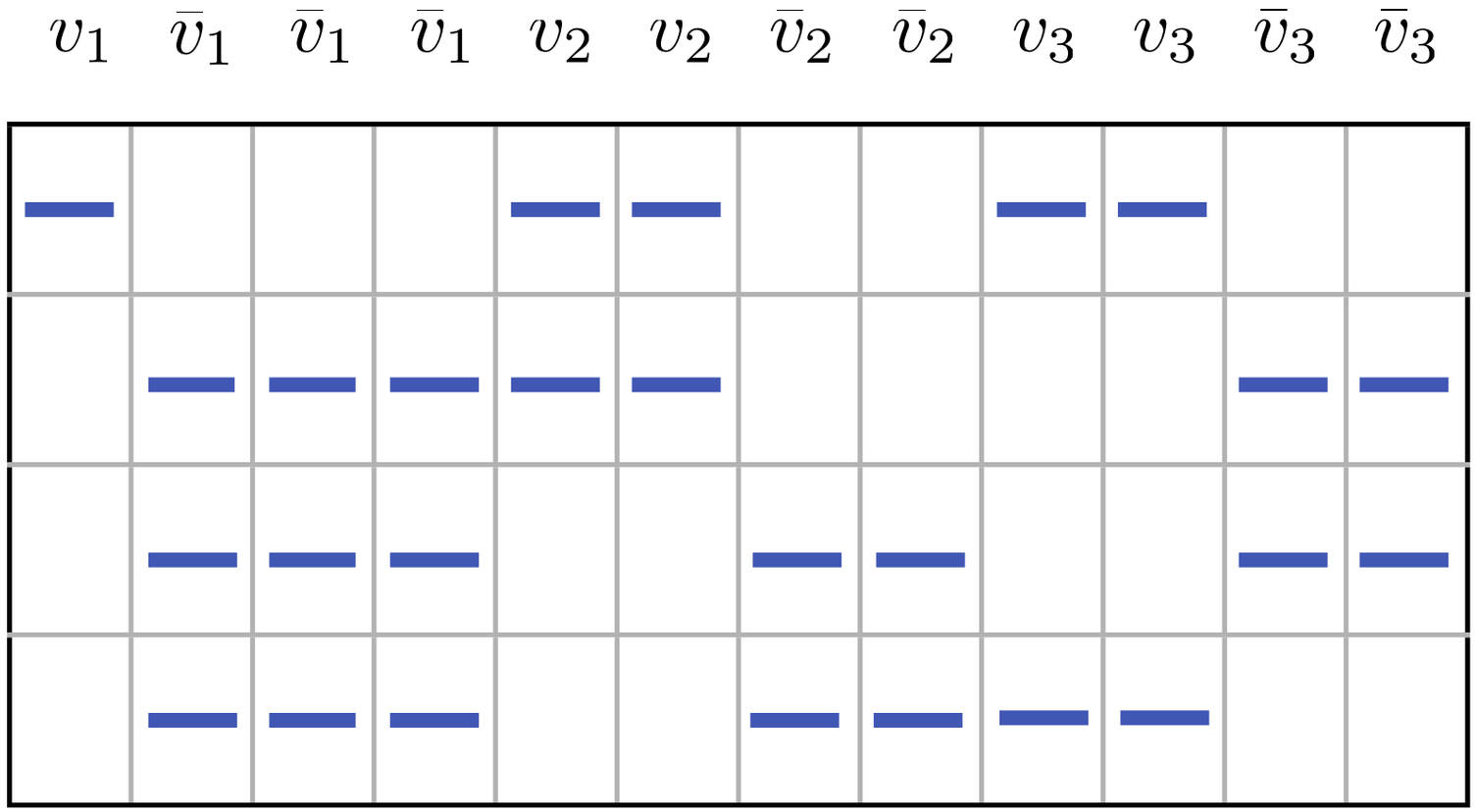
Reducing 3-SAT to Phrase Decision

$$v_1 \vee v_2 \vee v_3$$

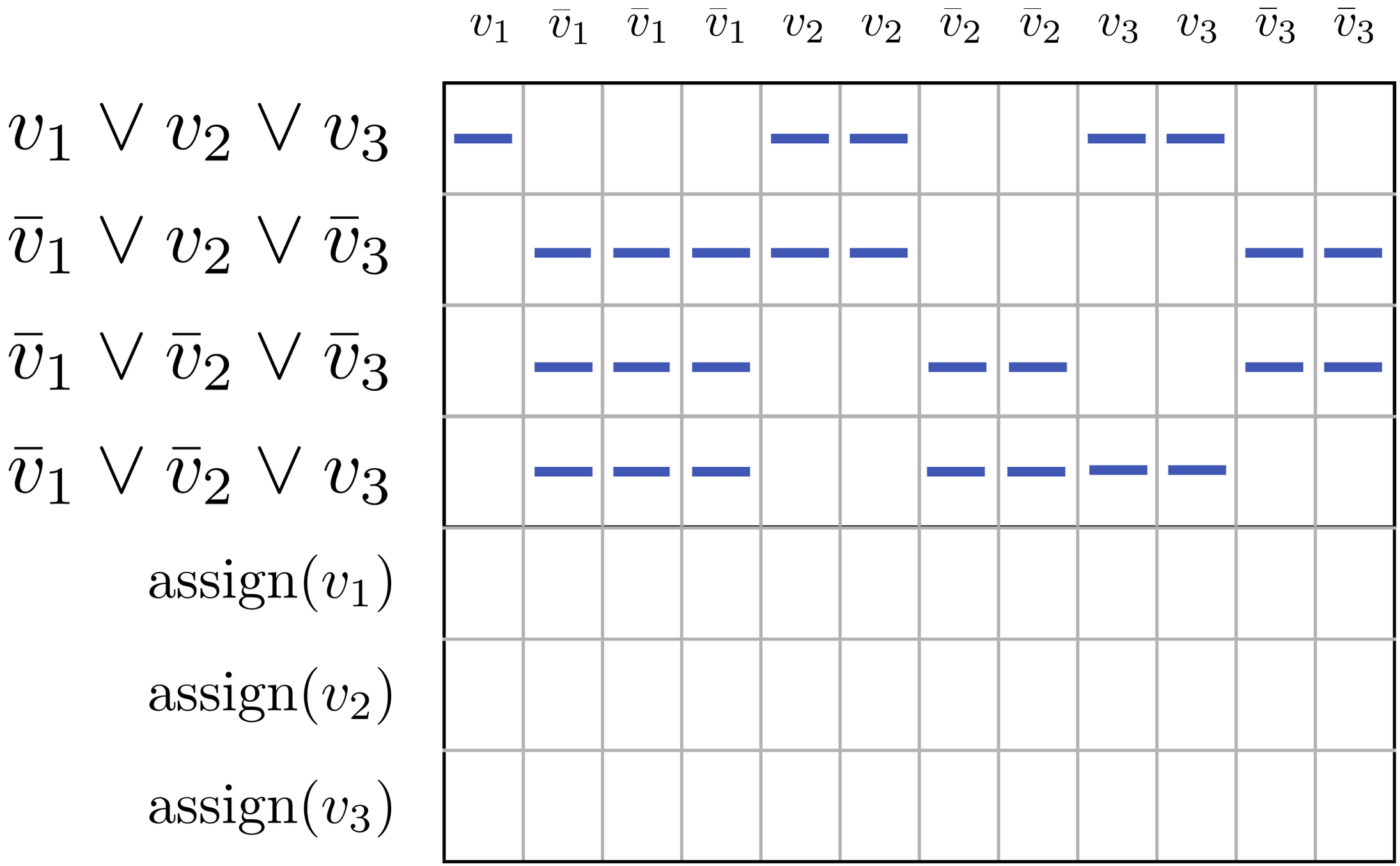
$$\bar{v}_1 \vee v_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$$

$$\bar{v}_1 \vee \bar{v}_2 \vee v_3$$



Reducing 3-SAT to Phrase Decision



Reducing 3-SAT to Phrase Decision

$v_1 \vee v_2 \vee v_3$

$\bar{v}_1 \vee v_2 \vee \bar{v}_3$

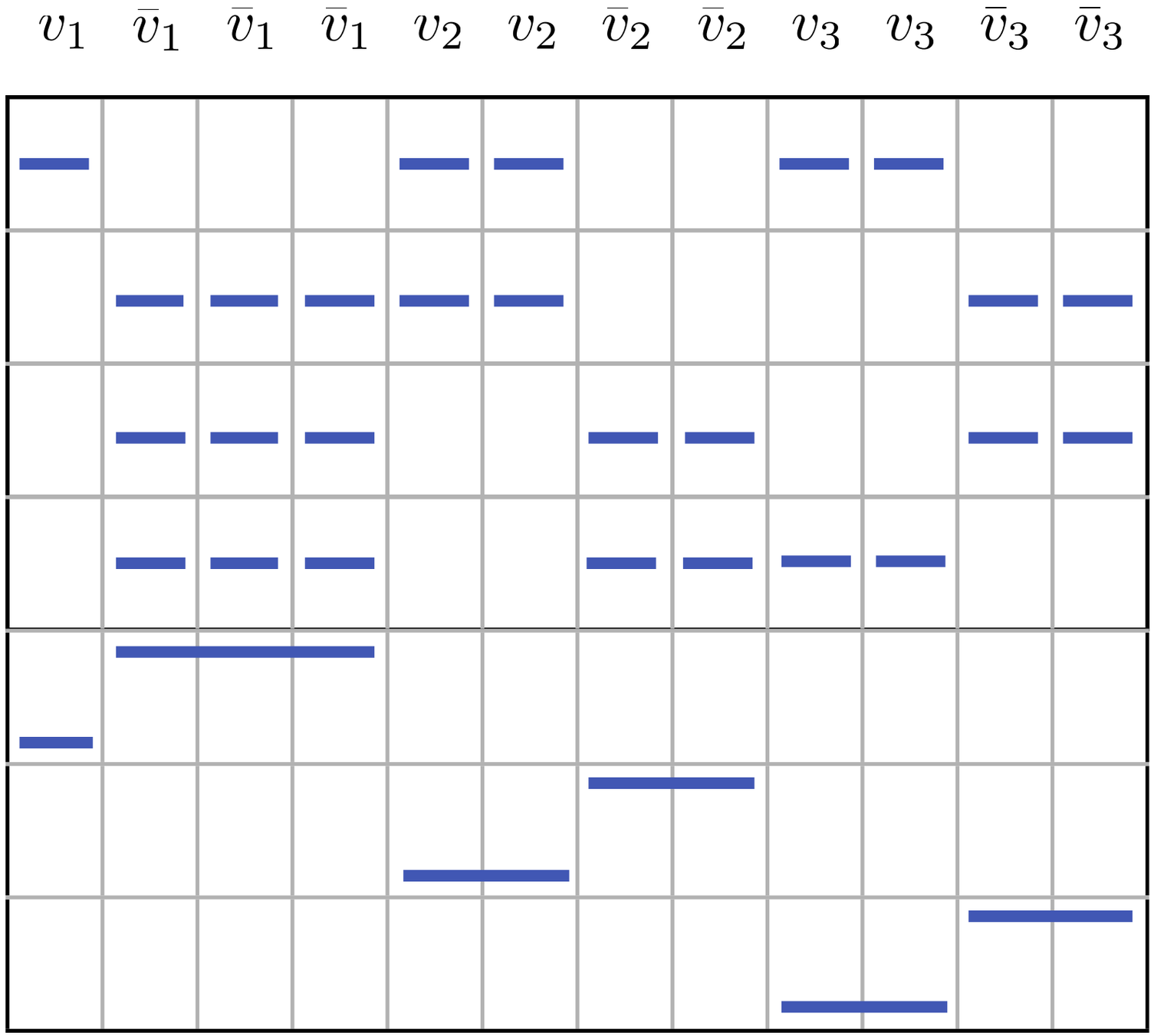
$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$

$\bar{v}_1 \vee \bar{v}_2 \vee v_3$

$\text{assign}(v_1)$

$\text{assign}(v_2)$

$\text{assign}(v_3)$



Reducing 3-SAT to Phrase Decision

$v_1 \vee v_2 \vee v_3$

$\bar{v}_1 \vee v_2 \vee \bar{v}_3$

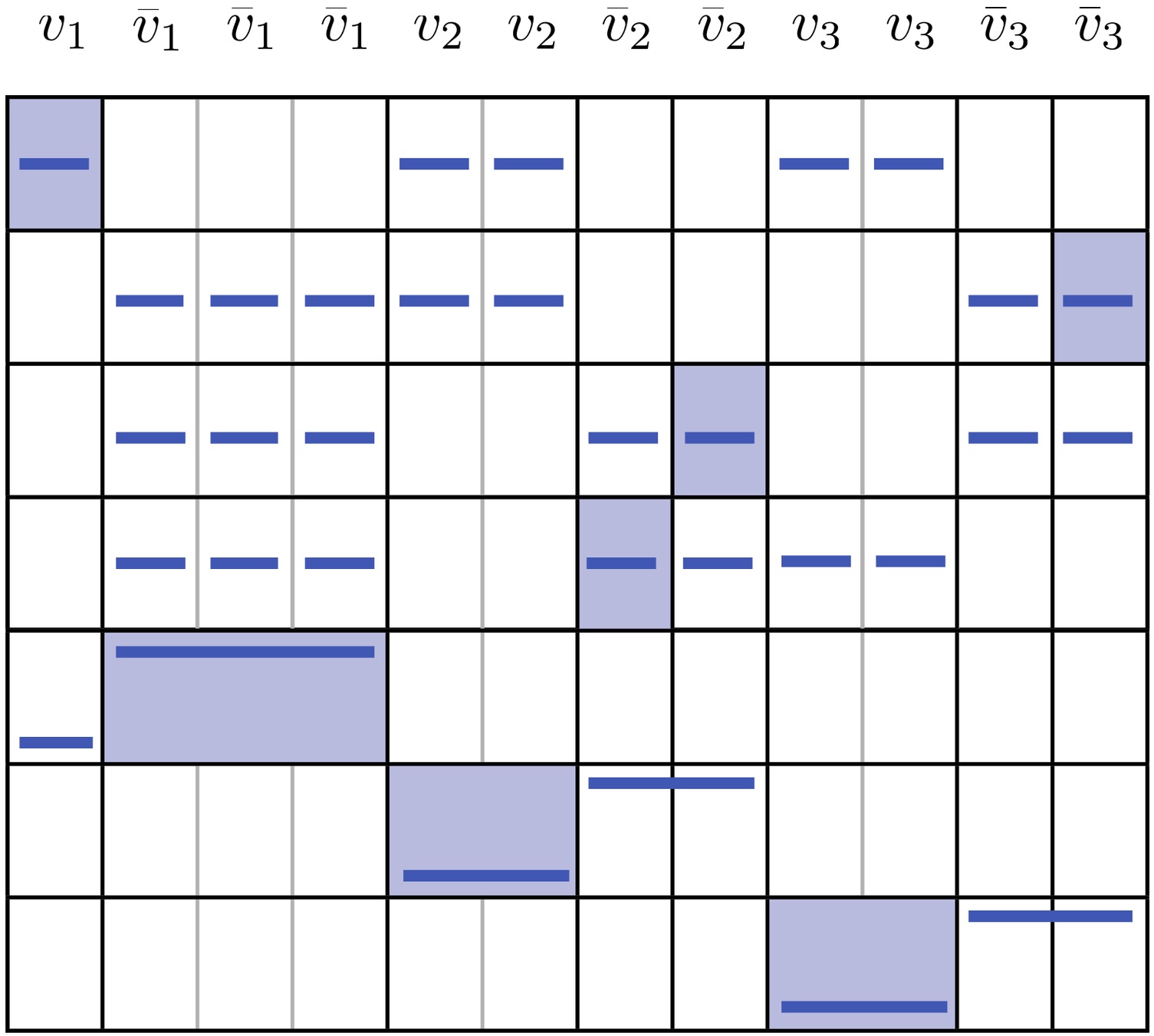
$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$

$\bar{v}_1 \vee \bar{v}_2 \vee v_3$

$\text{assign}(v_1)$

$\text{assign}(v_2)$

$\text{assign}(v_3)$



Reducing 3-SAT to Phrase Decision

$v_1 \vee v_2 \vee v_3$

$\bar{v}_1 \vee v_2 \vee \bar{v}_3$

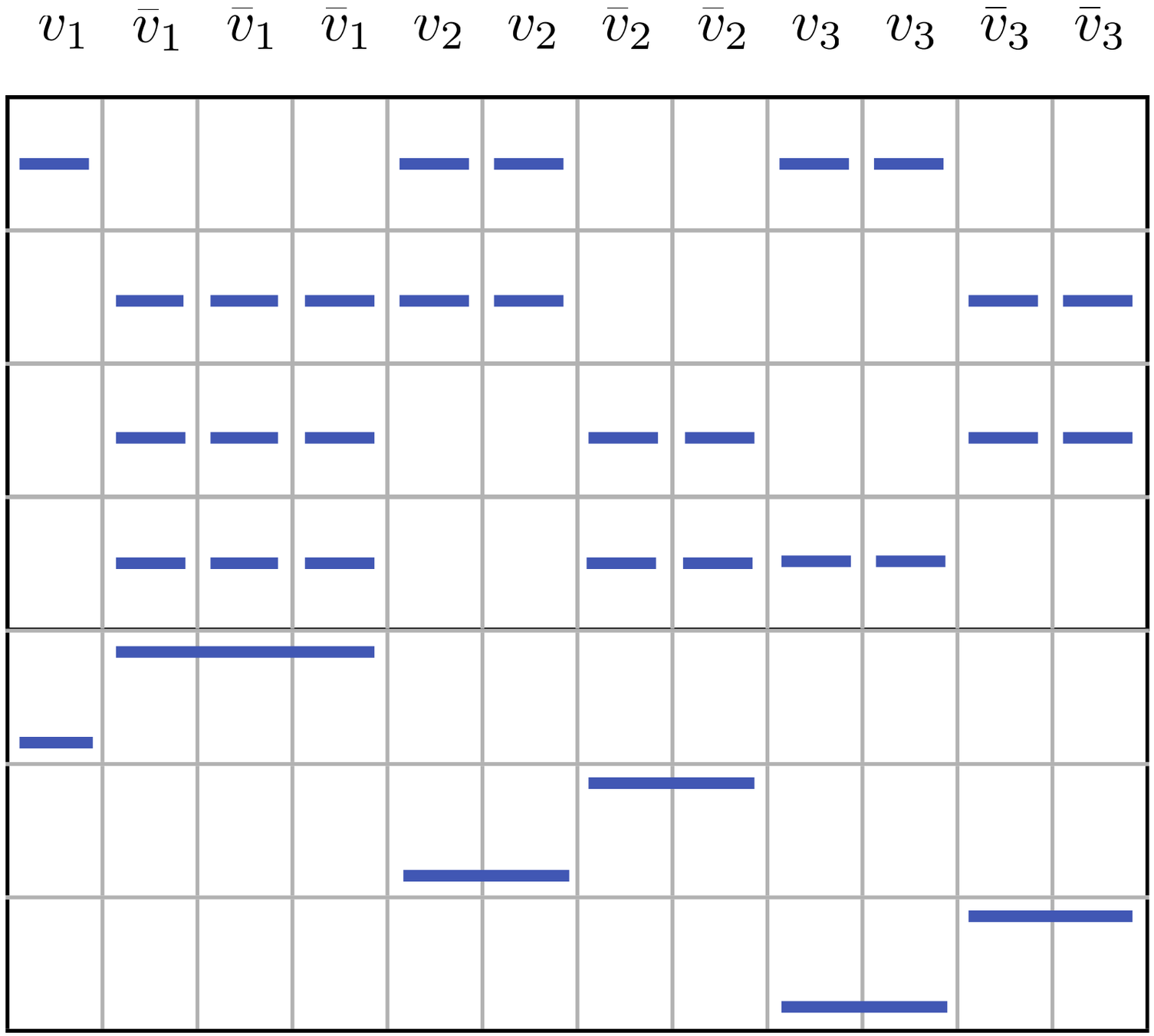
$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$

$\bar{v}_1 \vee \bar{v}_2 \vee v_3$

$\text{assign}(v_1)$

$\text{assign}(v_2)$

$\text{assign}(v_3)$



Reducing 3-SAT to Phrase Decision

$v_1 \vee v_2 \vee v_3$

$\bar{v}_1 \vee v_2 \vee \bar{v}_3$

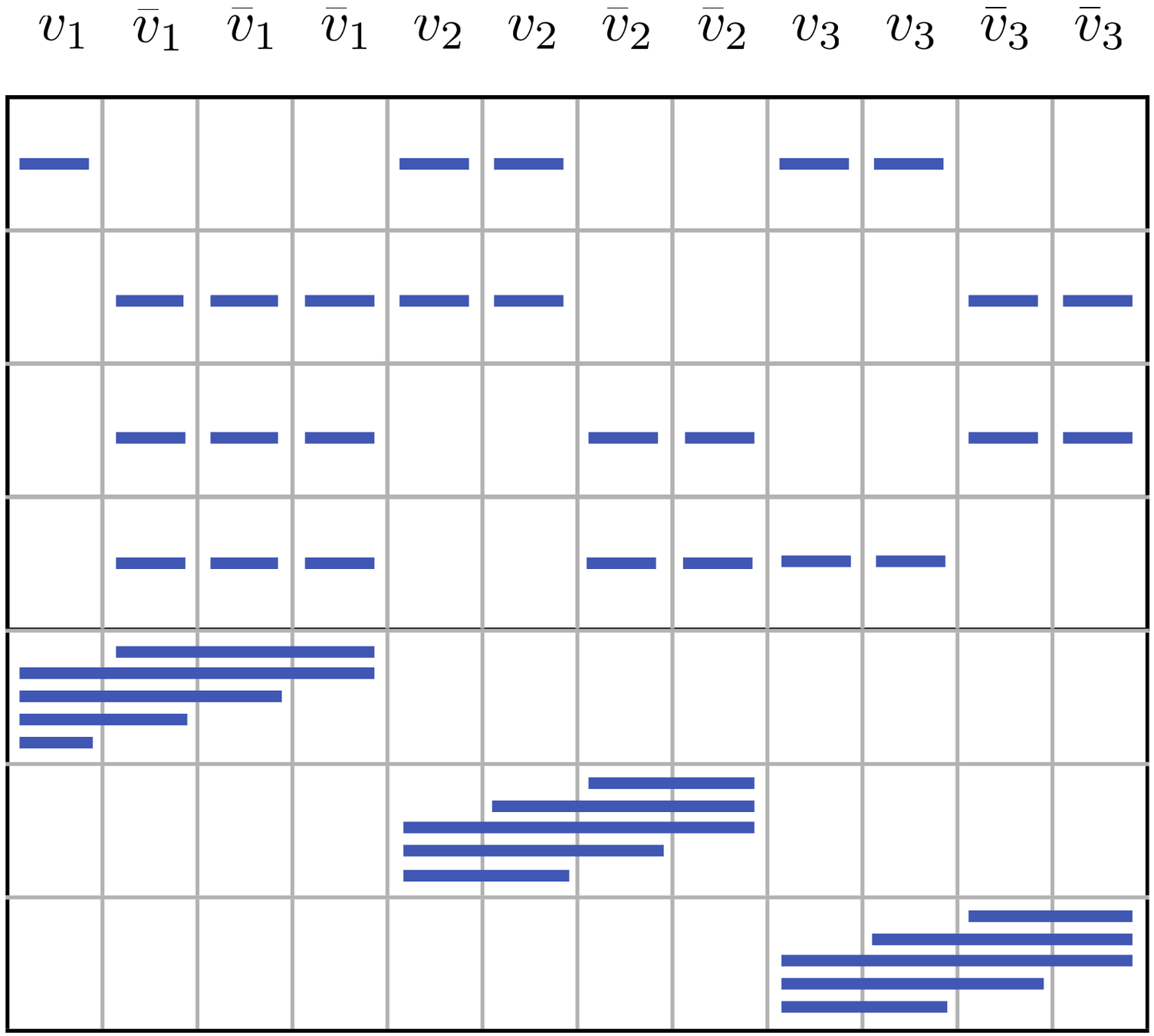
$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$

$\bar{v}_1 \vee \bar{v}_2 \vee v_3$

$\text{assign}(v_1)$

$\text{assign}(v_2)$

$\text{assign}(v_3)$



Reducing 3-SAT to Phrase Decision

$v_1 \vee v_2 \vee v_3$

$\bar{v}_1 \vee v_2 \vee \bar{v}_3$

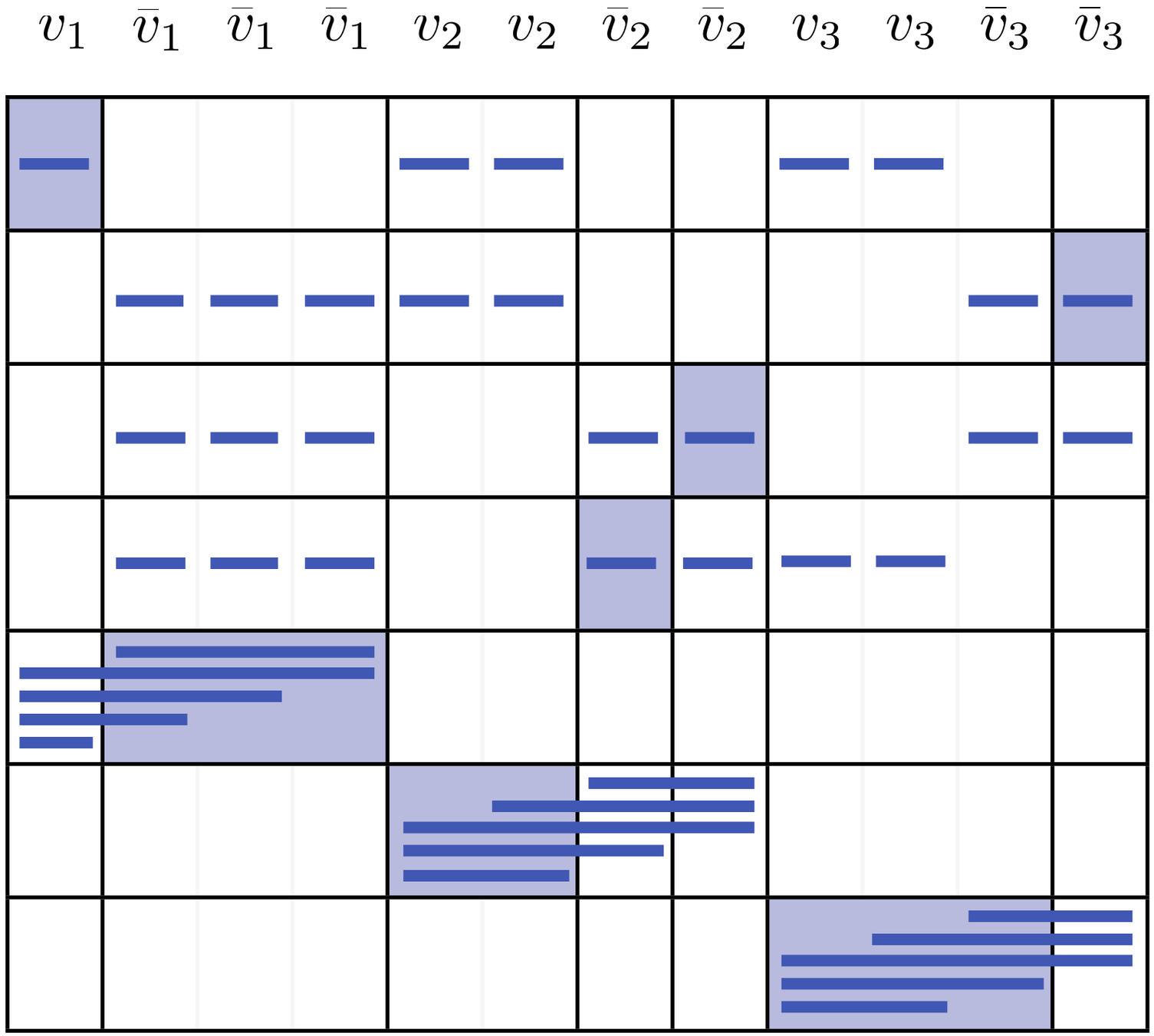
$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$

$\bar{v}_1 \vee \bar{v}_2 \vee v_3$

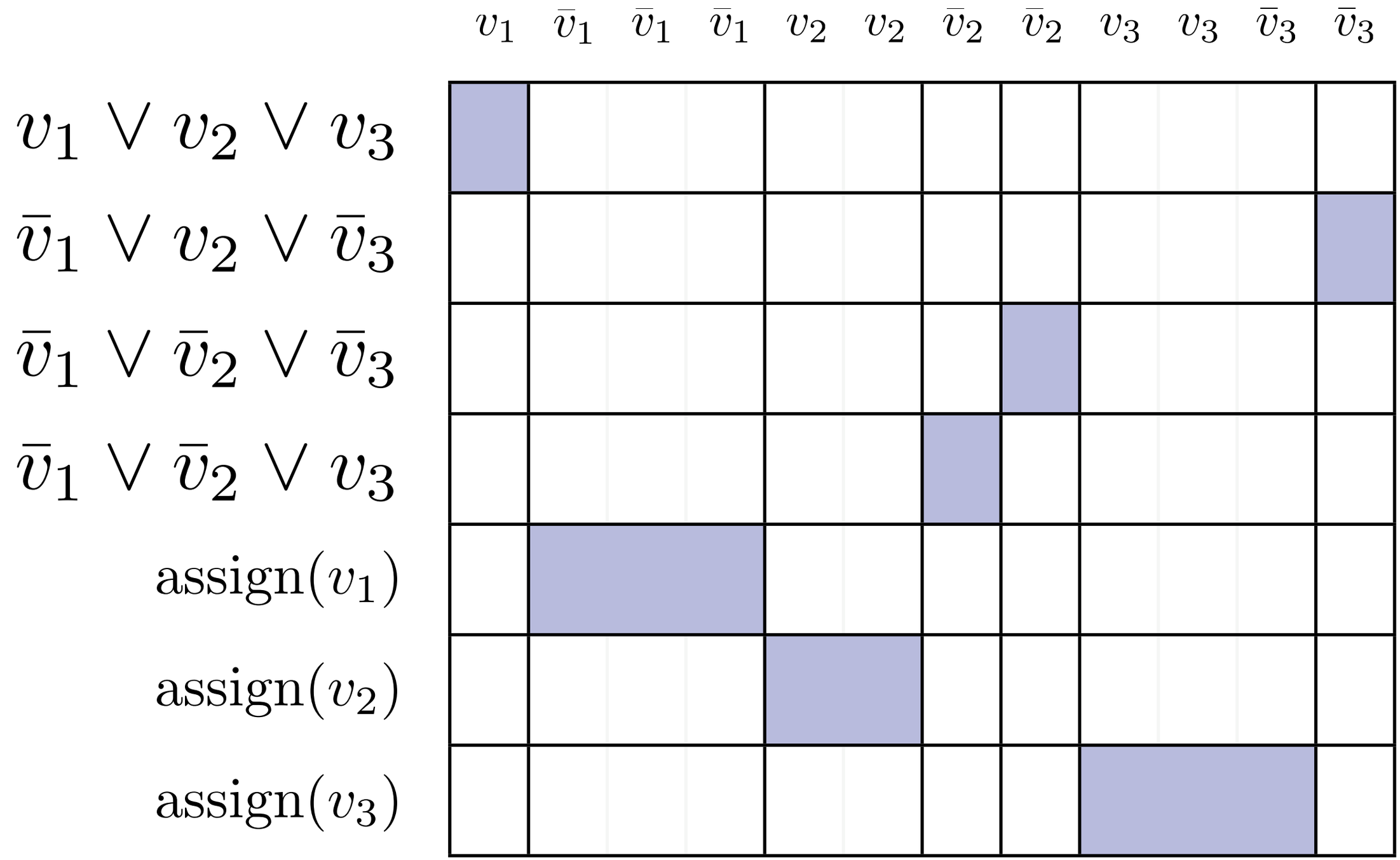
$\text{assign}(v_1)$

$\text{assign}(v_2)$

$\text{assign}(v_3)$



Reducing 3-SAT to Phrase Decision



Reducing 3-SAT to Phrase Decision

$v_1 \vee v_2 \vee v_3$

$\bar{v}_1 \vee v_2 \vee \bar{v}_3$

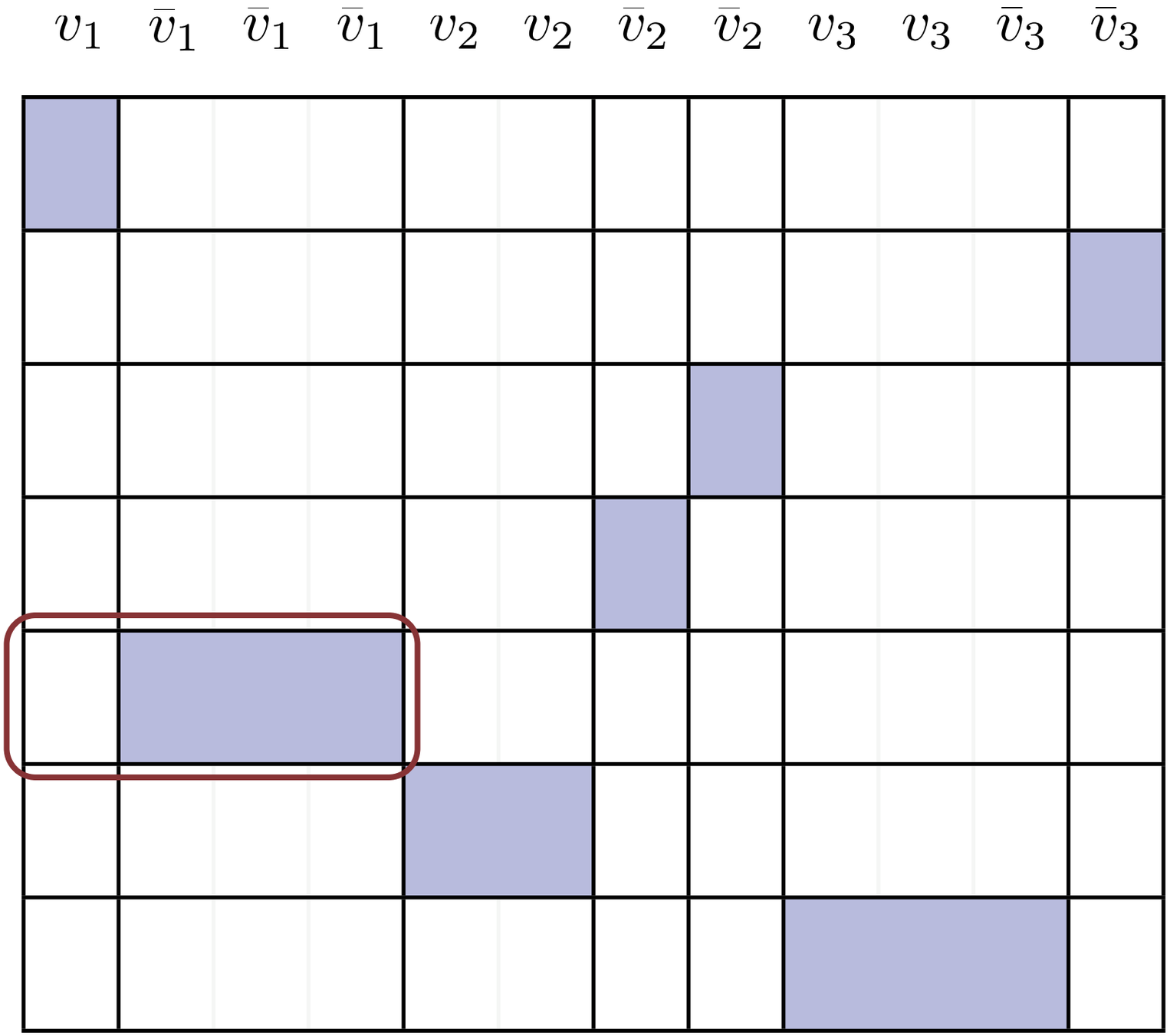
$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$

$\bar{v}_1 \vee \bar{v}_2 \vee v_3$

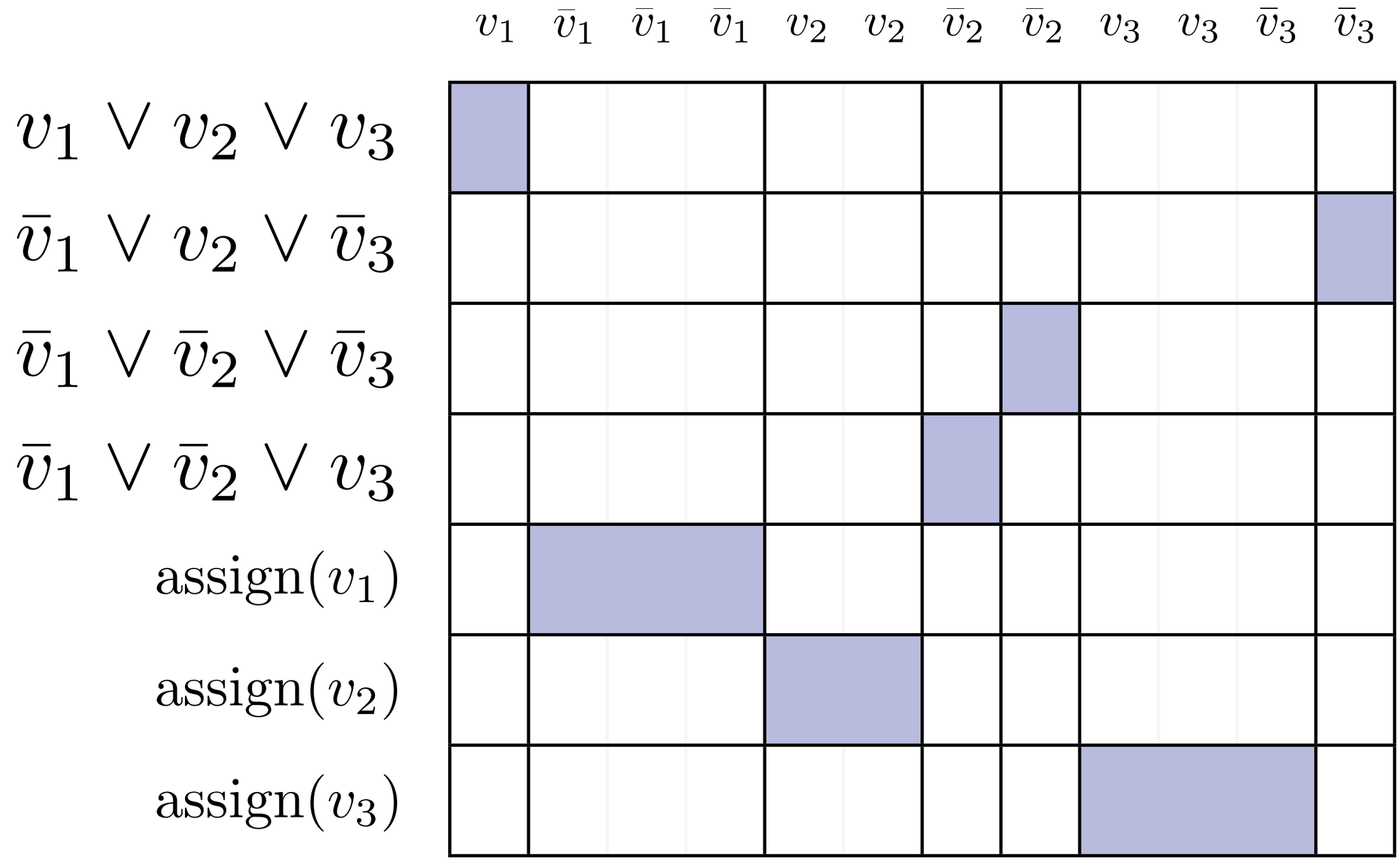
assign(v_1)

assign(v_2)

assign(v_3)



Reducing 3-SAT to Phrase Decision



Reducing 3-SAT to Phrase Decision

$v_1 \vee v_2 \vee v_3$

$\bar{v}_1 \vee v_2 \vee \bar{v}_3$

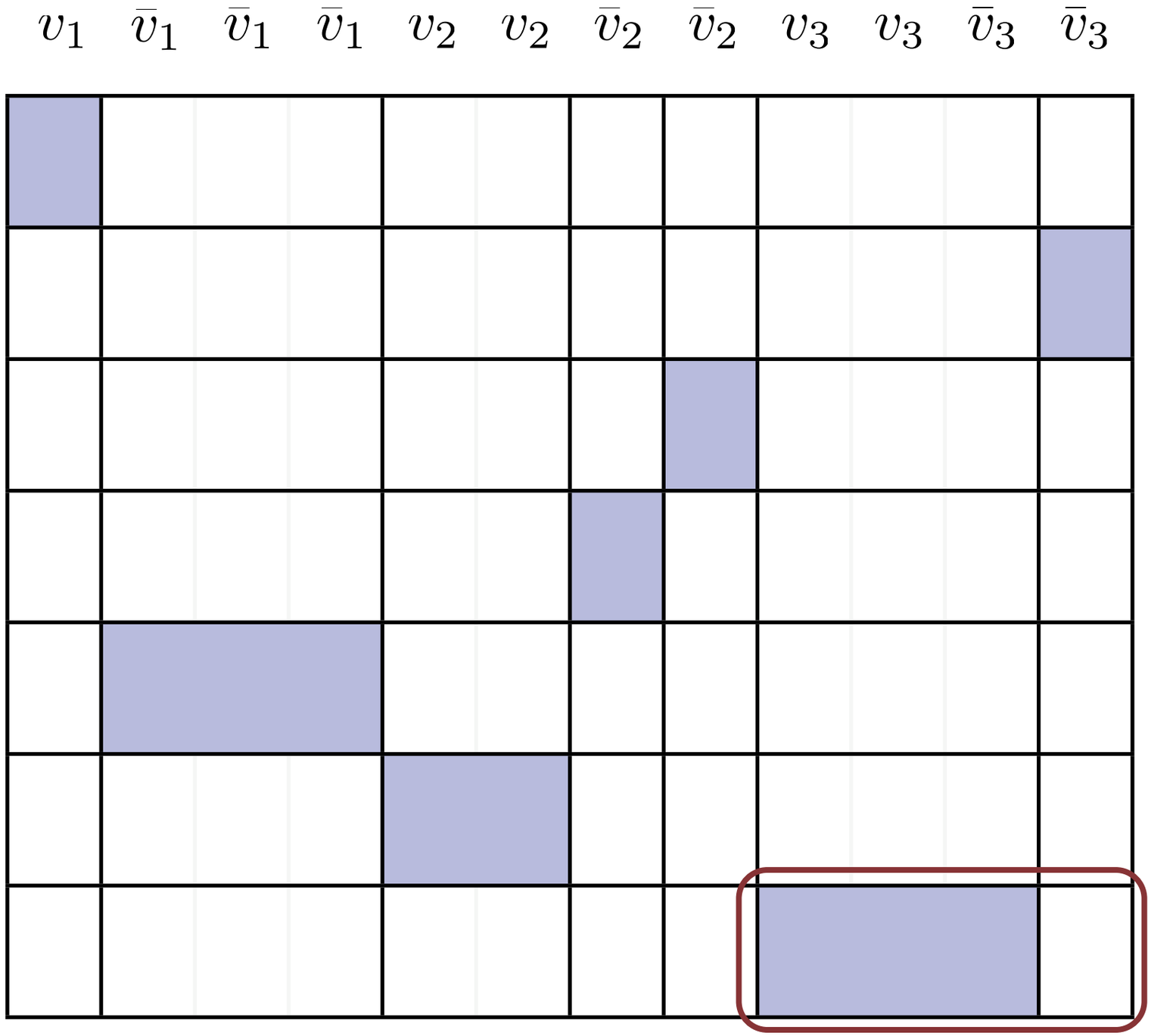
$\bar{v}_1 \vee \bar{v}_2 \vee \bar{v}_3$

$\bar{v}_1 \vee \bar{v}_2 \vee v_3$

$\text{assign}(v_1)$

$\text{assign}(v_2)$

$\text{assign}(v_3)$

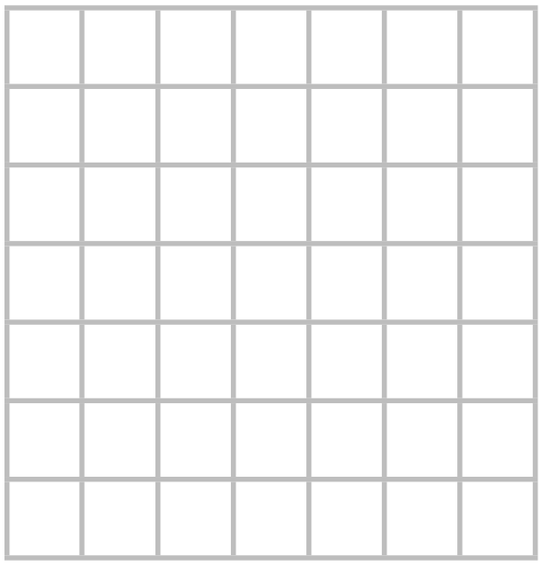


University of California
C A L
N L P
Berkeley

Taking Expectations is #P-hard

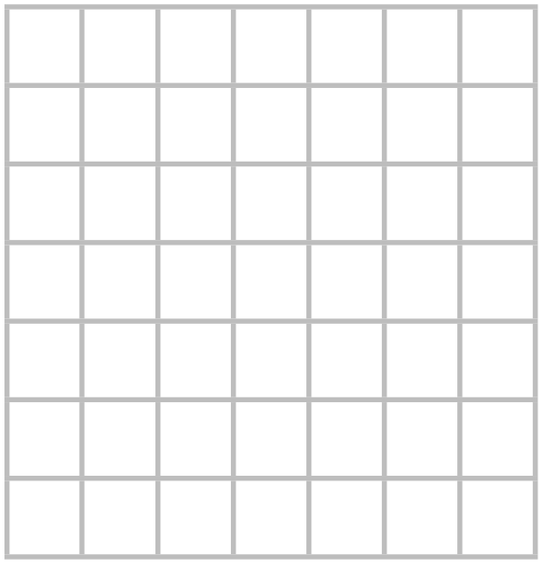
Taking Expectations is #P-hard

Weighted grid:



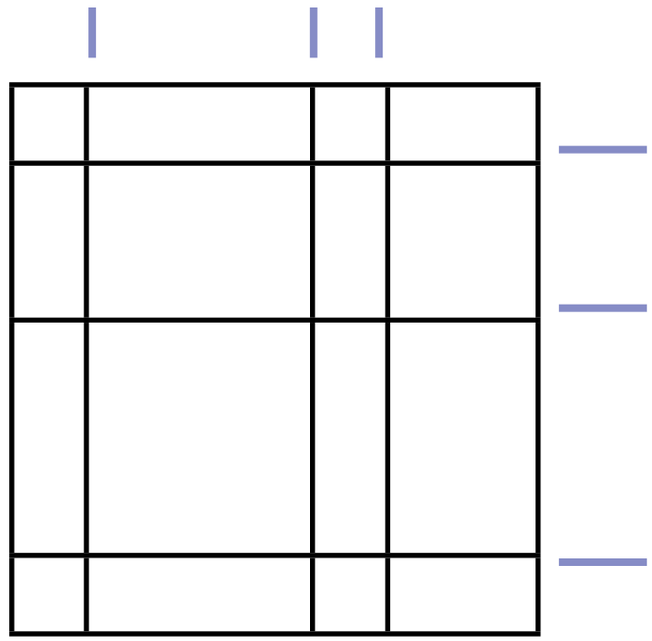
Taking Expectations is #P-hard

Weighted grid:



+

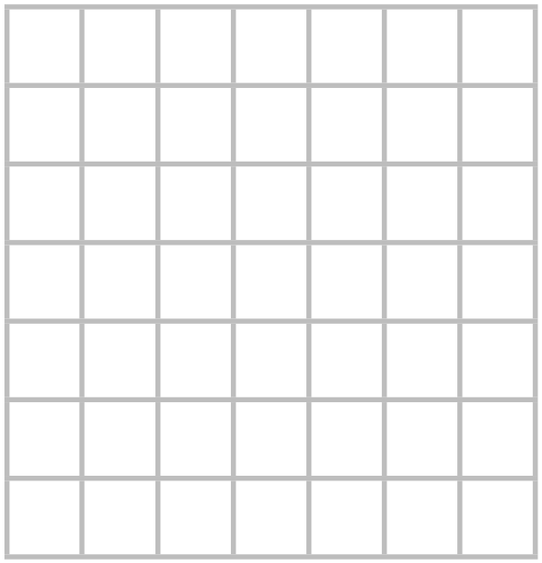
Segmentation:



University of California
Berkeley

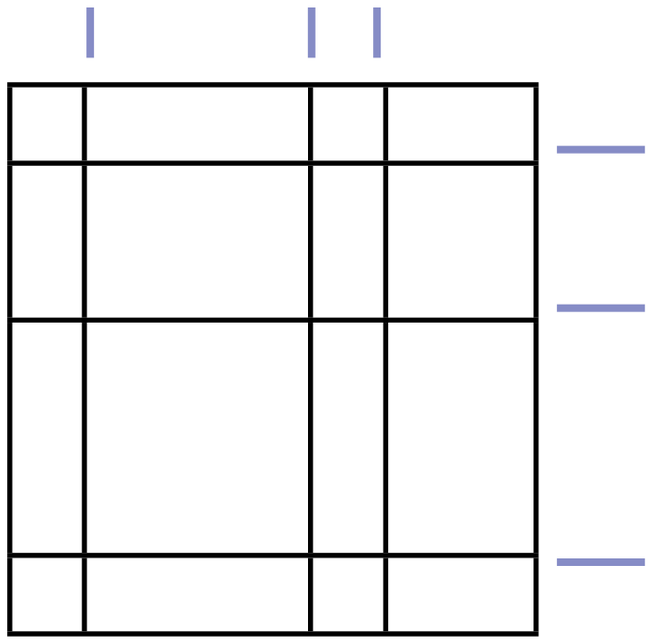
Taking Expectations is #P-hard

Weighted grid:

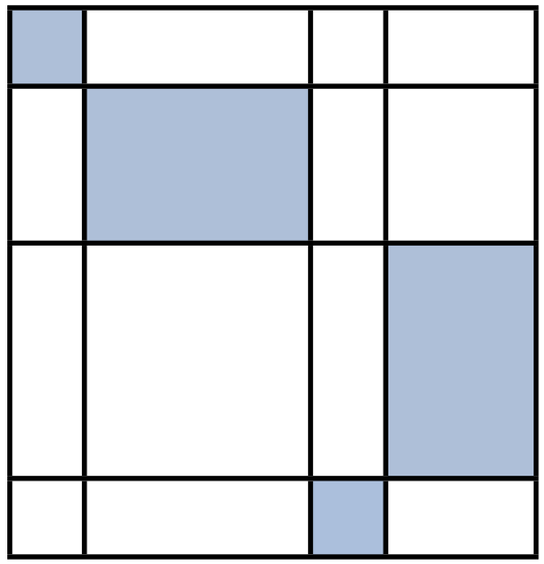


+

Segmentation:

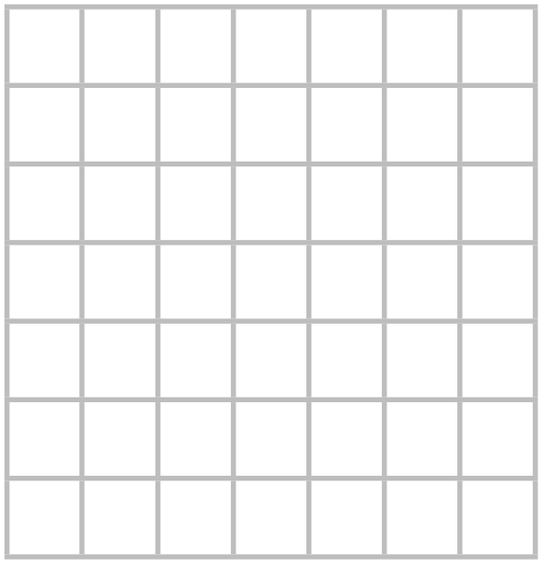


Find Maximal:



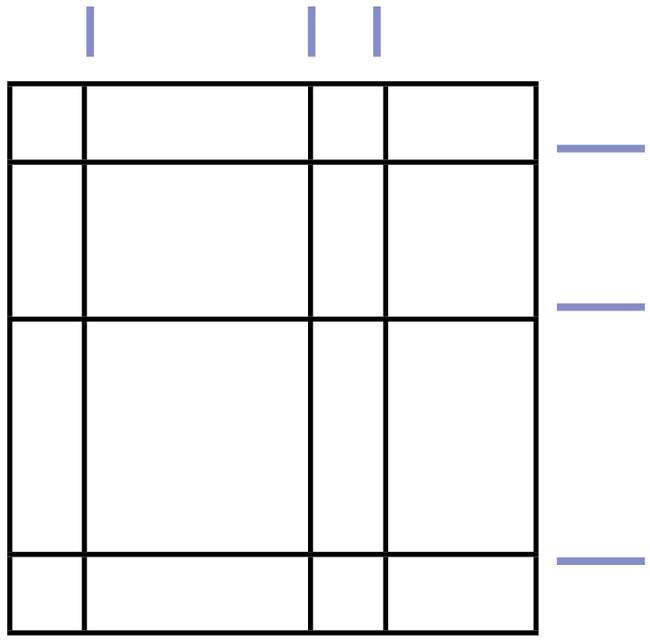
Taking Expectations is #P-hard

Weighted grid:

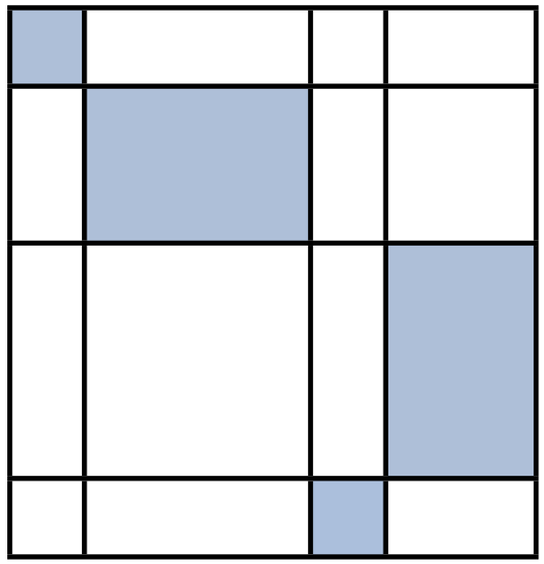


+

Segmentation:



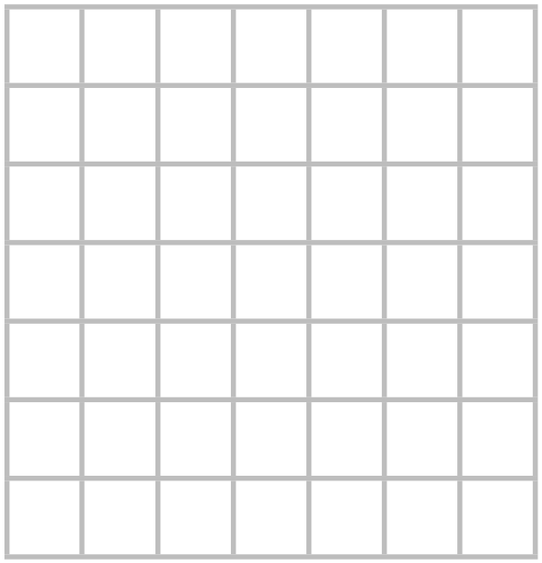
Find Maximal:



Summing over matching is #P-hard

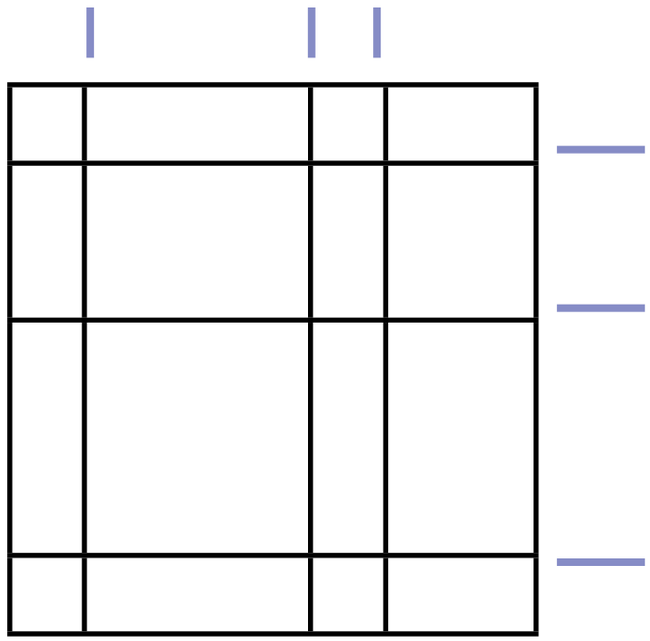
Taking Expectations is #P-hard

Weighted grid:

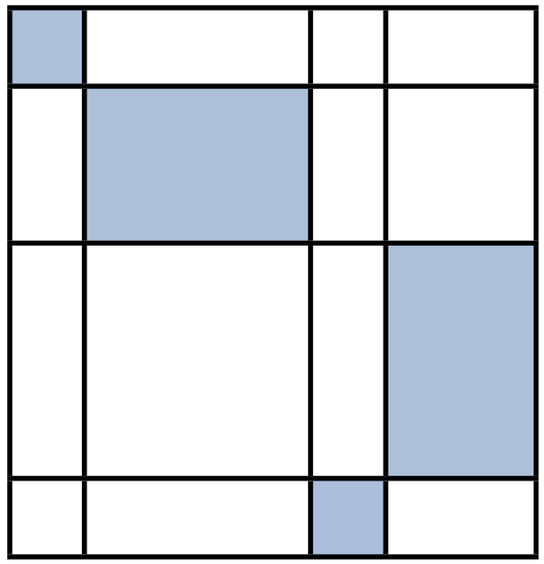


+

Segmentation:



Find Maximal:

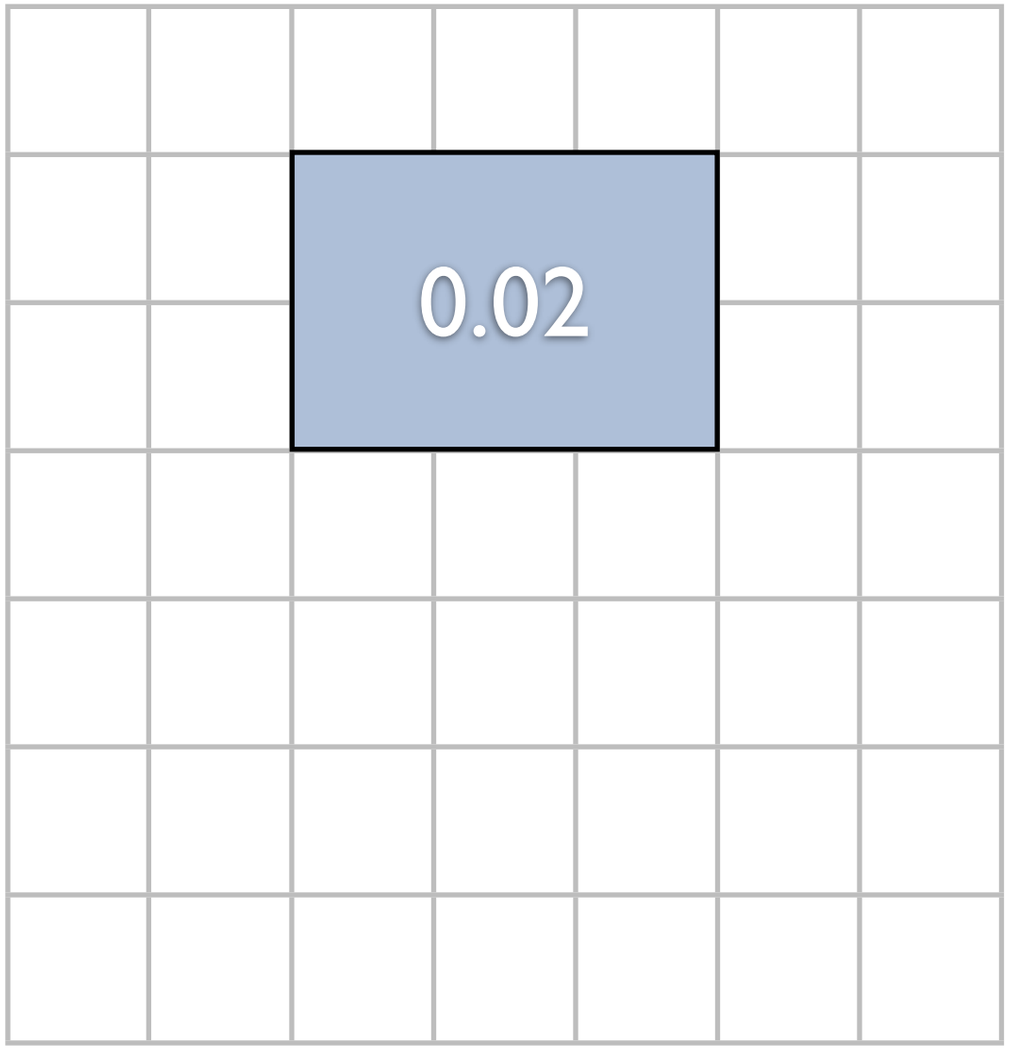


Summing over matching is #P-hard

Expectations of phrase alignments is at least as hard

Phrase Alignment as Integer Programming

Japan to freeze aid to Russia .



日本

冻结

向

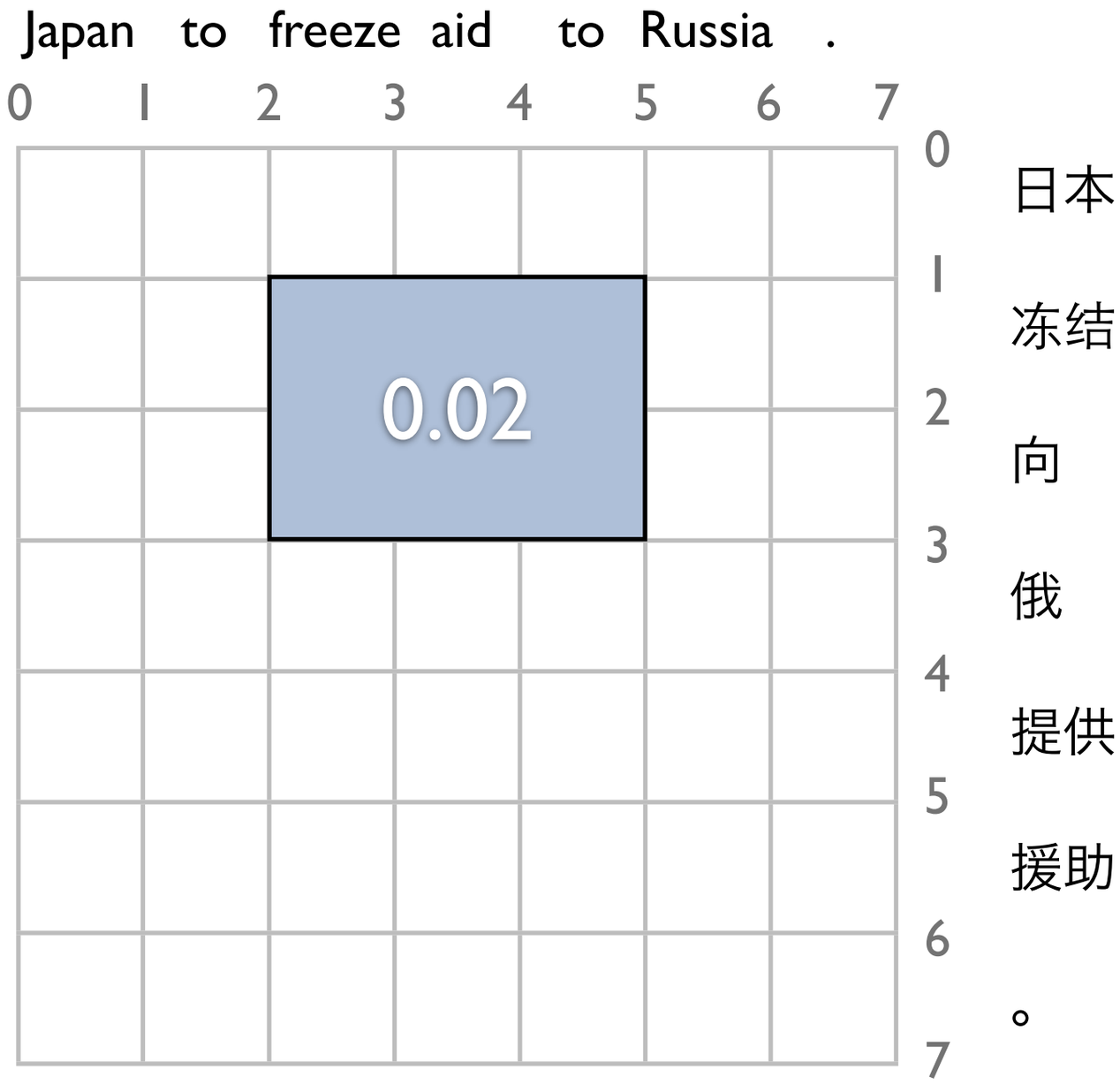
俄

提供

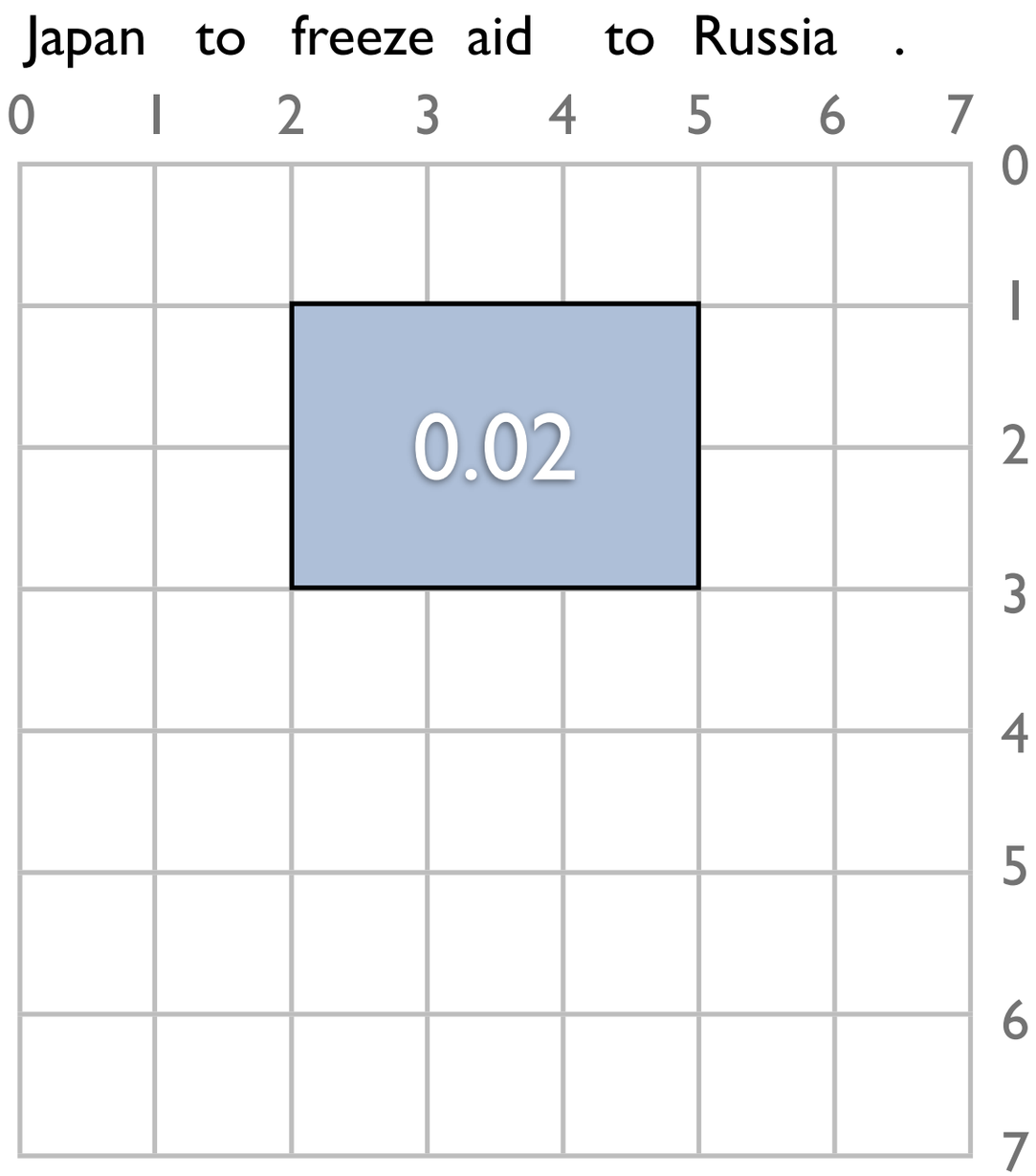
援助

。

Phrase Alignment as Integer Programming



Phrase Alignment as Integer Programming



日本

冻结

向

俄

提供

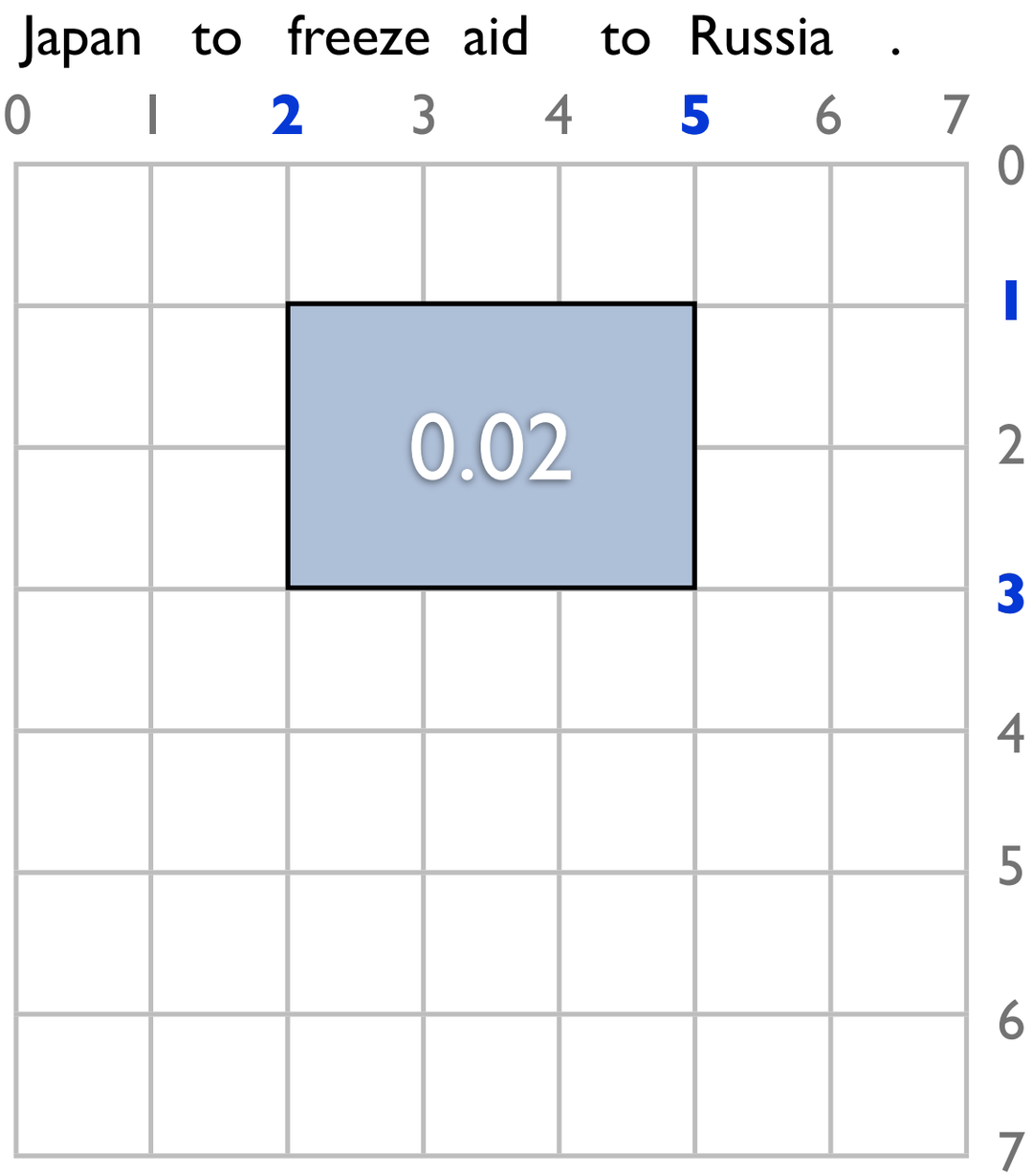
援助

。

Constants:

$$w_{2,5,1,3} = \log 0.02$$

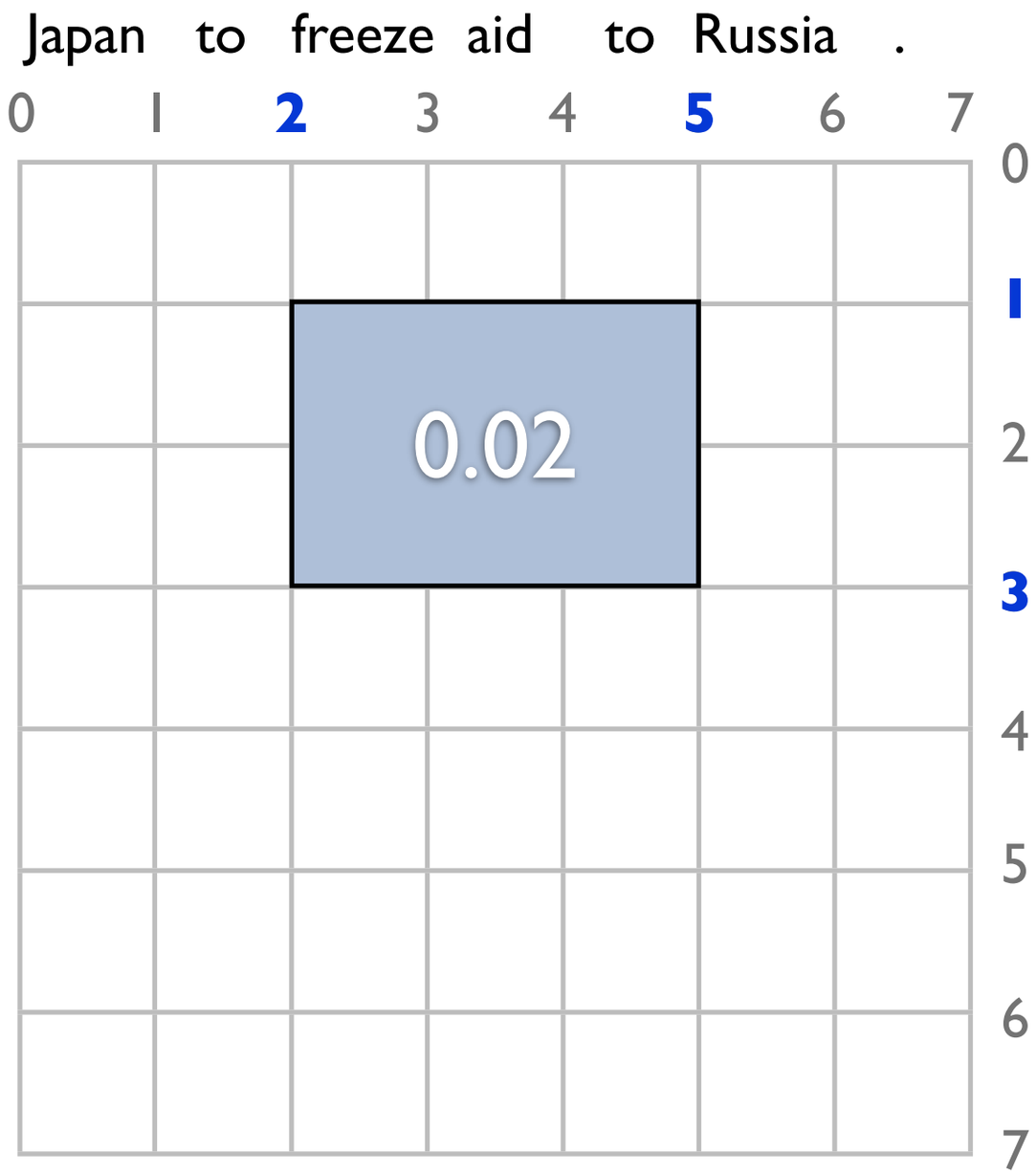
Phrase Alignment as Integer Programming



Constants:

$$w_{2,5,1,3} = \log 0.02$$

Phrase Alignment as Integer Programming



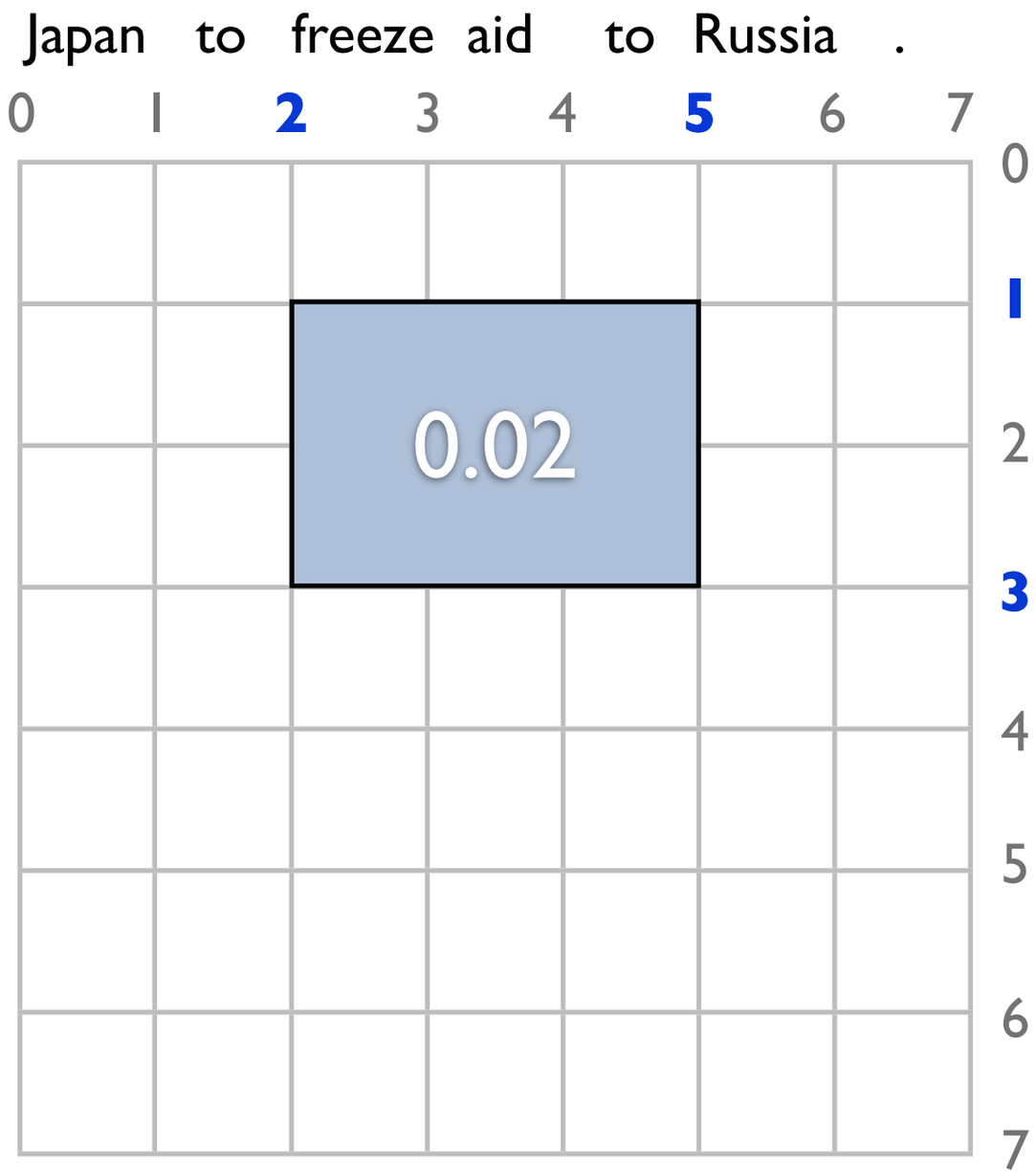
Constants:

$$w_{2,5,1,3} = \log 0.02$$

Indicator Variables:

$$a_{2,5,1,3} = 1$$

Phrase Alignment as Integer Programming



Constants:

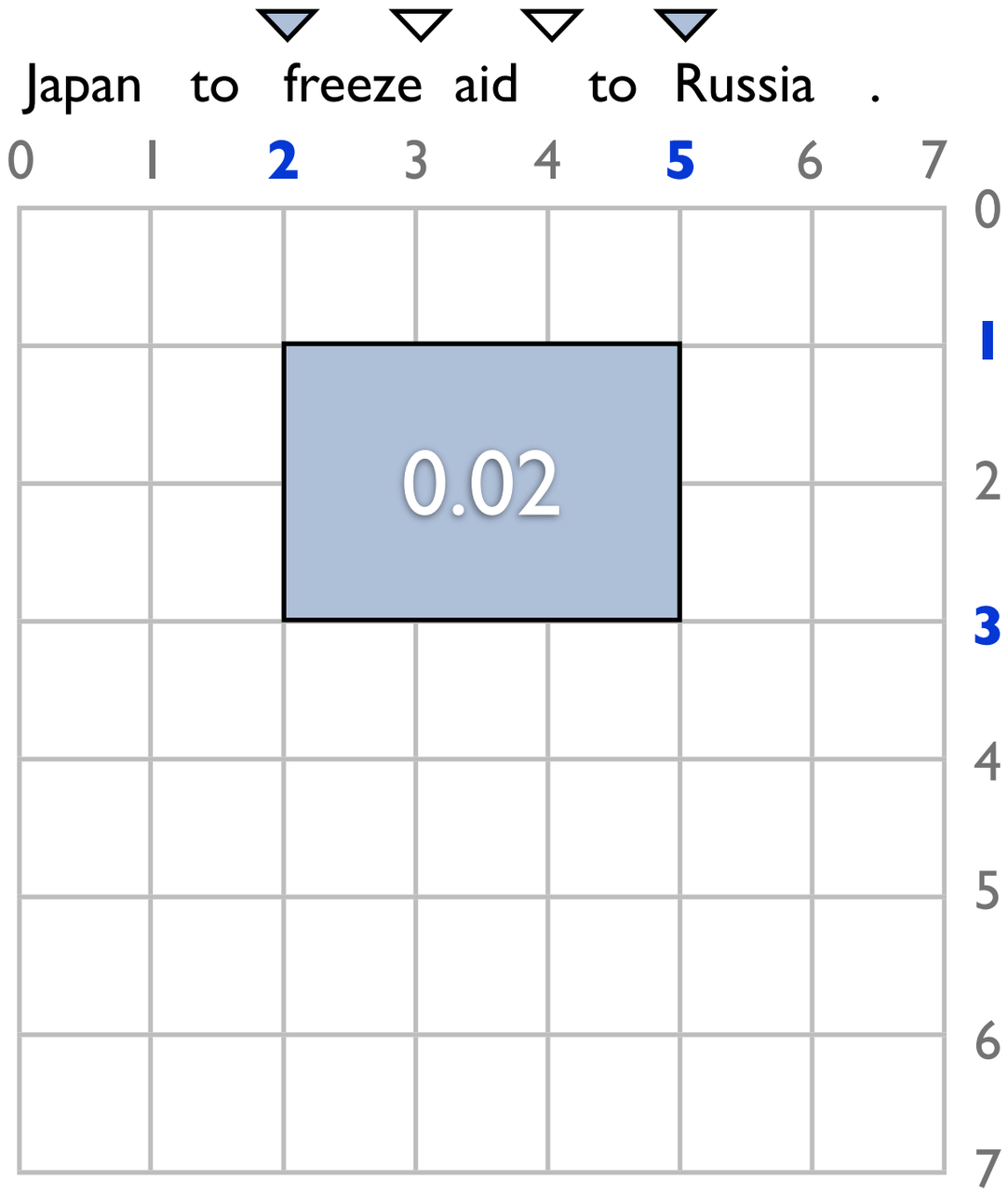
$$w_{2,5,1,3} = \log 0.02$$

Indicator Variables:

$$a_{2,5,1,3} = 1$$

$$e_{2,5} = 1$$

Phrase Alignment as Integer Programming



Constants:

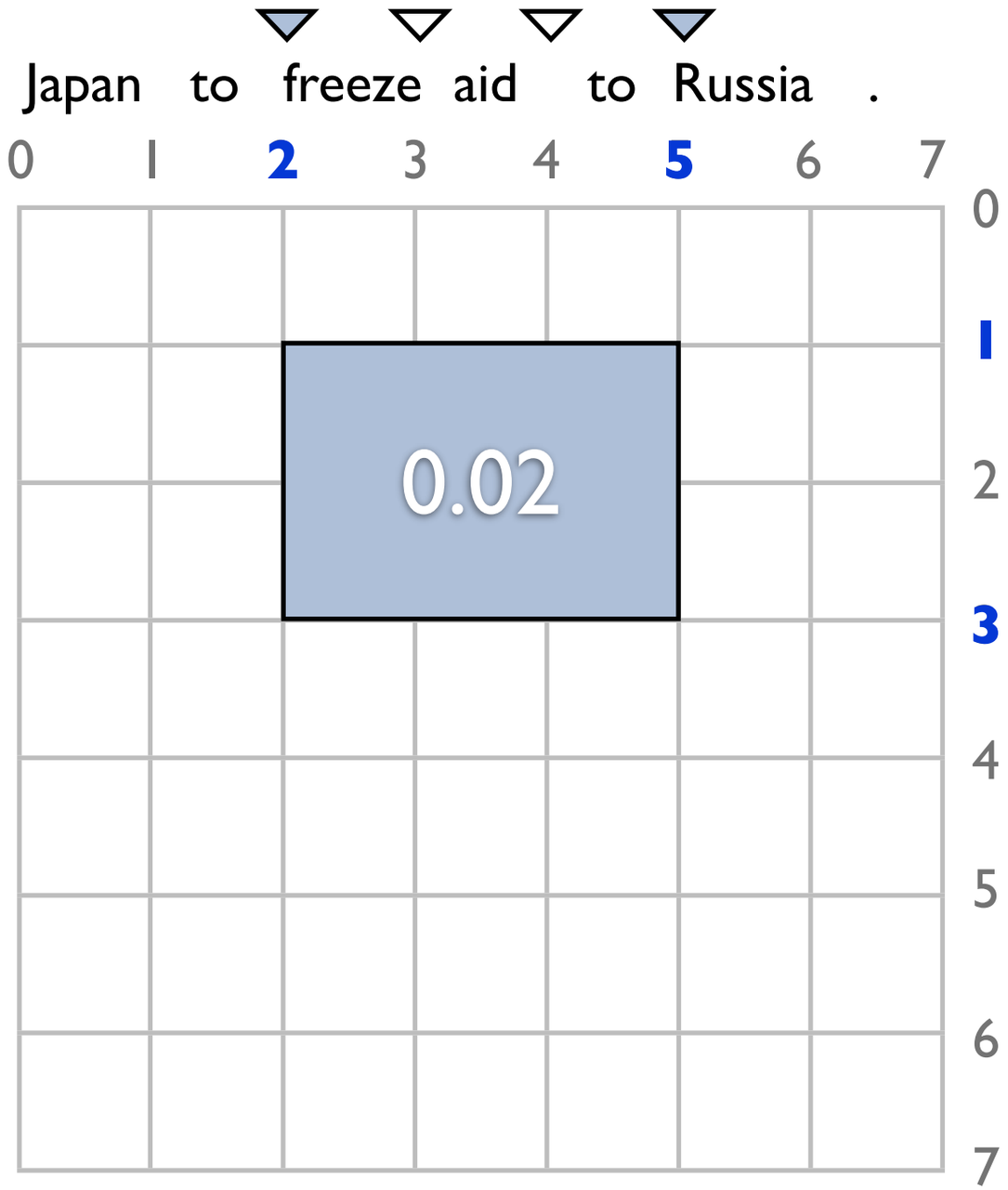
$$w_{2,5,1,3} = \log 0.02$$

Indicator Variables:

$$a_{2,5,1,3} = 1$$

$$e_{2,5} = 1$$

Phrase Alignment as Integer Programming



Constants:

$$w_{2,5,1,3} = \log 0.02$$

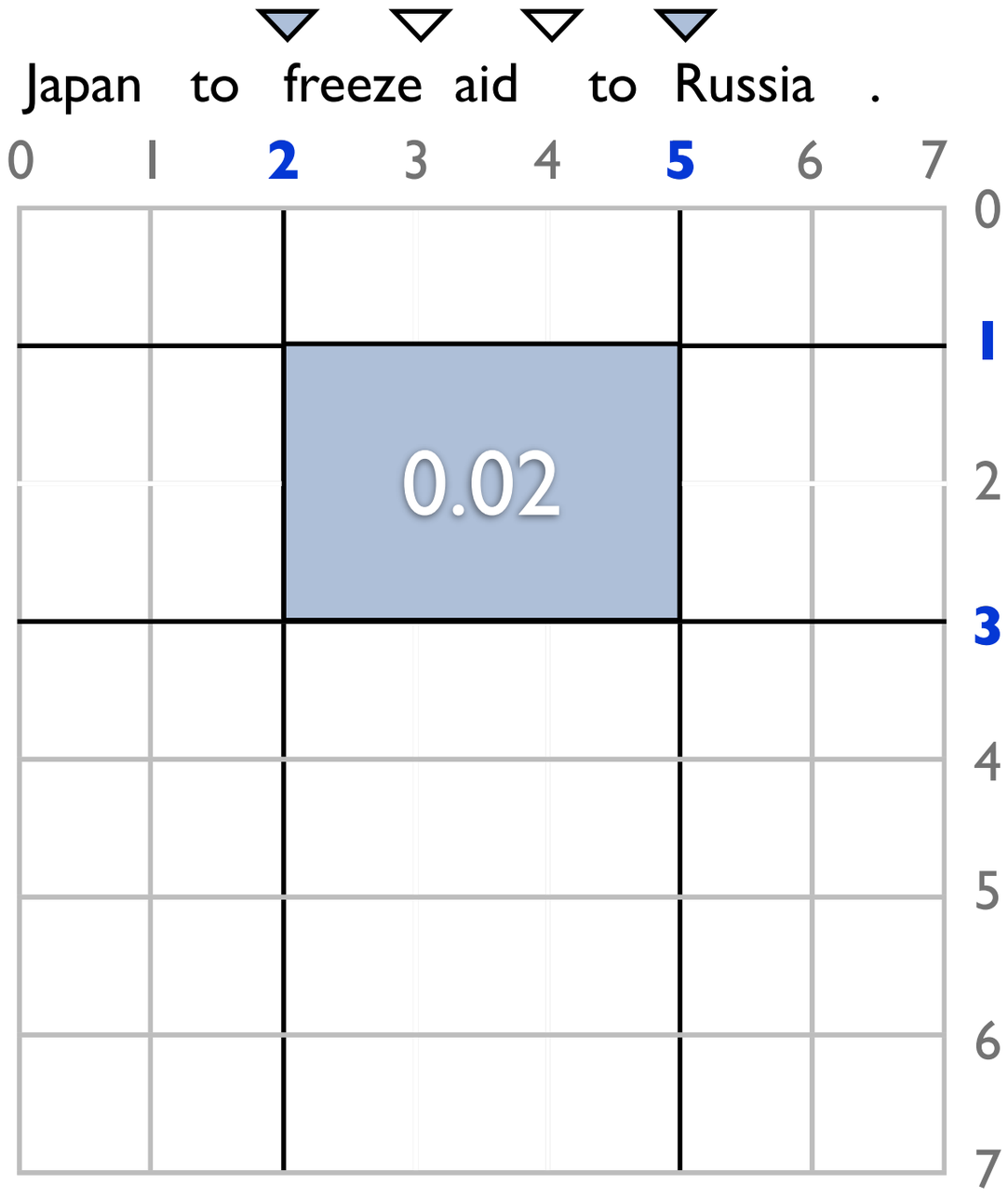
Indicator Variables:

$$a_{2,5,1,3} = 1$$

$$e_{2,5} = 1$$

$$f_{1,3} = 1$$

Phrase Alignment as Integer Programming



日本
冻结
向
俄
提供
援助
。

Constants:

$$w_{2,5,1,3} = \log 0.02$$

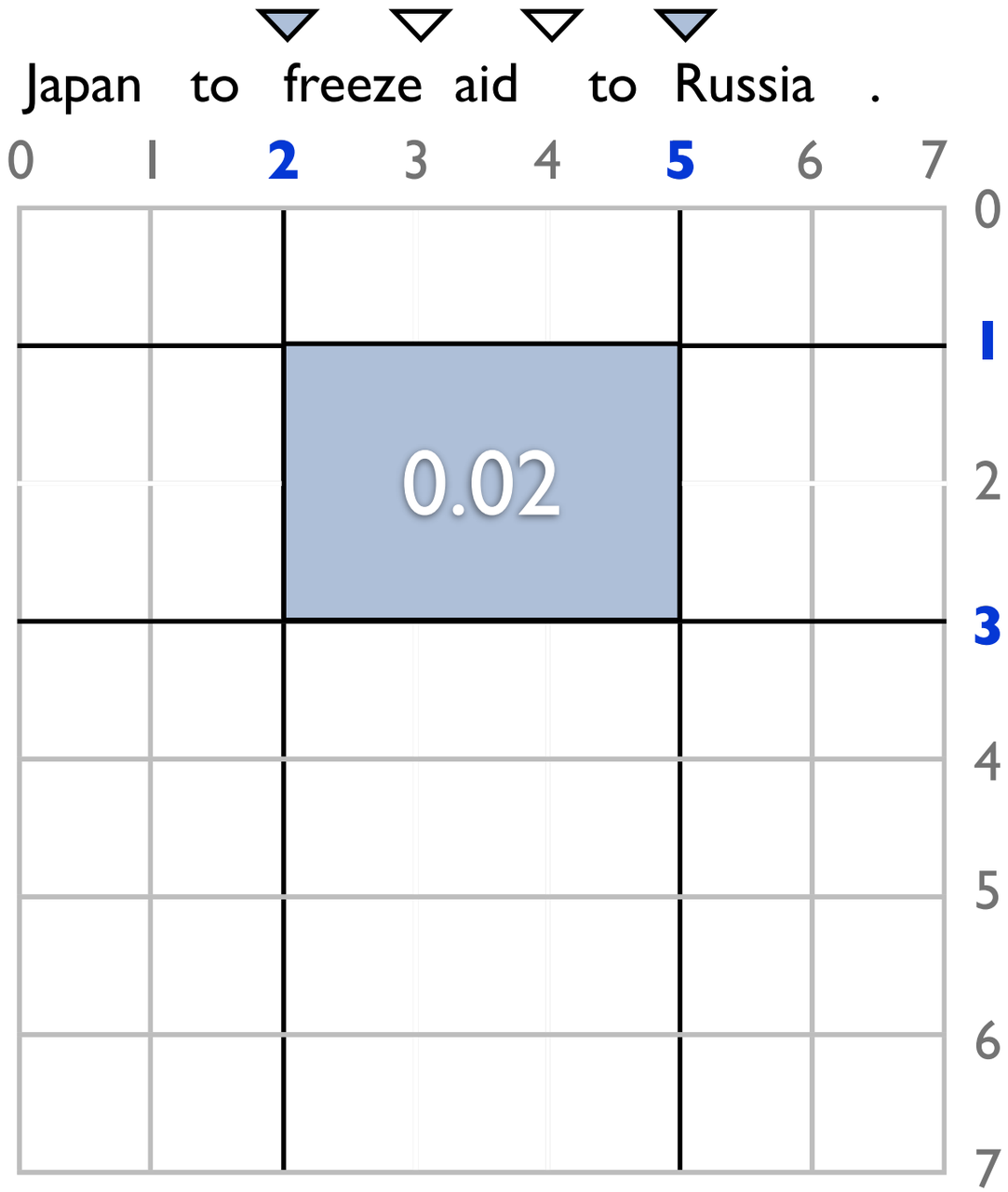
Indicator Variables:

$$a_{2,5,1,3} = 1$$

$$e_{2,5} = 1$$

$$f_{1,3} = 1$$

Phrase Alignment as Integer Programming



Constants:

$$w_{2,5,1,3} = \log 0.02$$

Indicator Variables:

$$a_{2,5,1,3} = 1$$

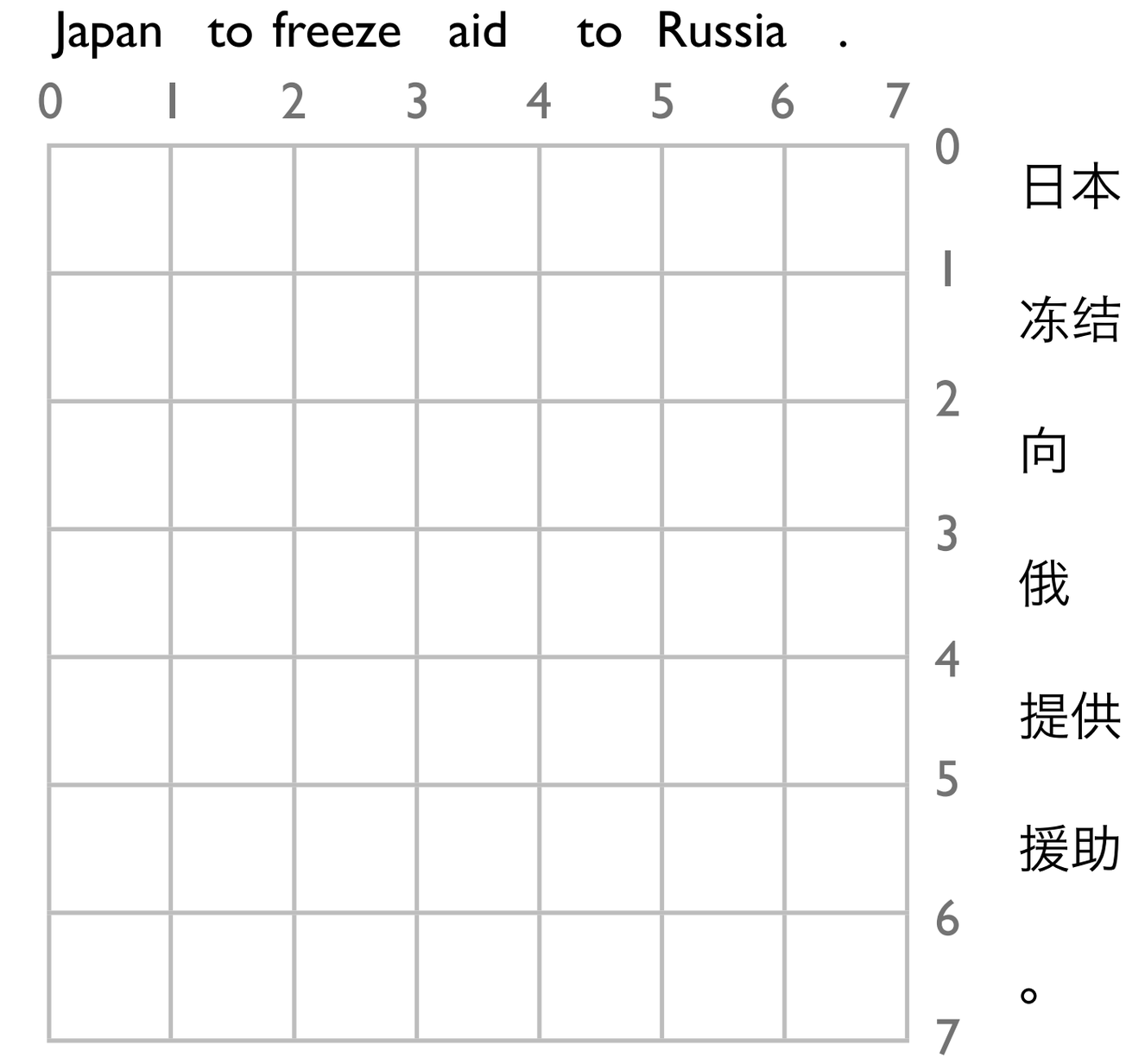
$$e_{2,5} = 1$$

$$f_{1,3} = 1$$

Indicator Variables:
a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:
w: weights

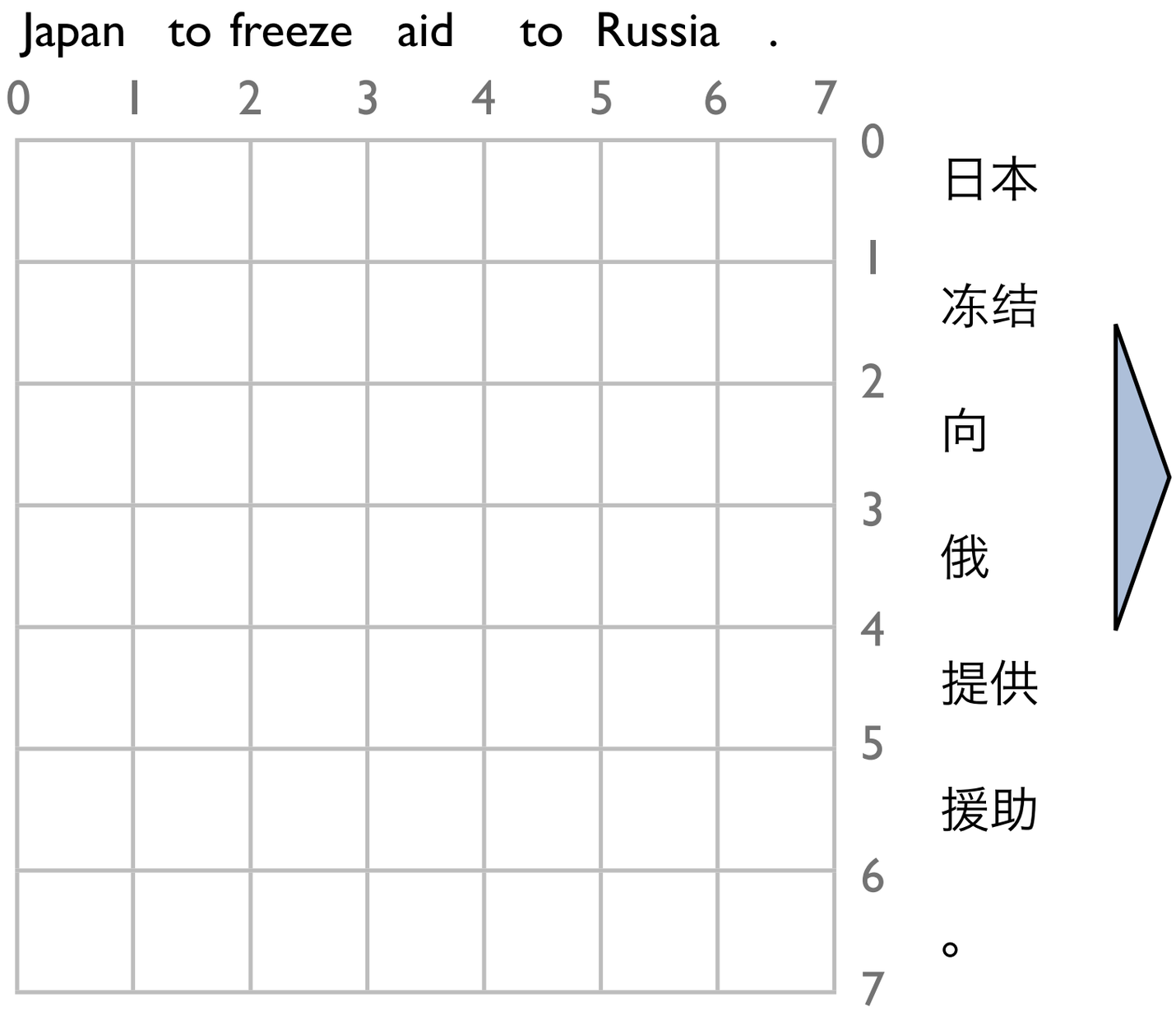
Phrase Alignment as Integer Programming



Indicator Variables:
a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:
w: weights

Phrase Alignment as Integer Programming



$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

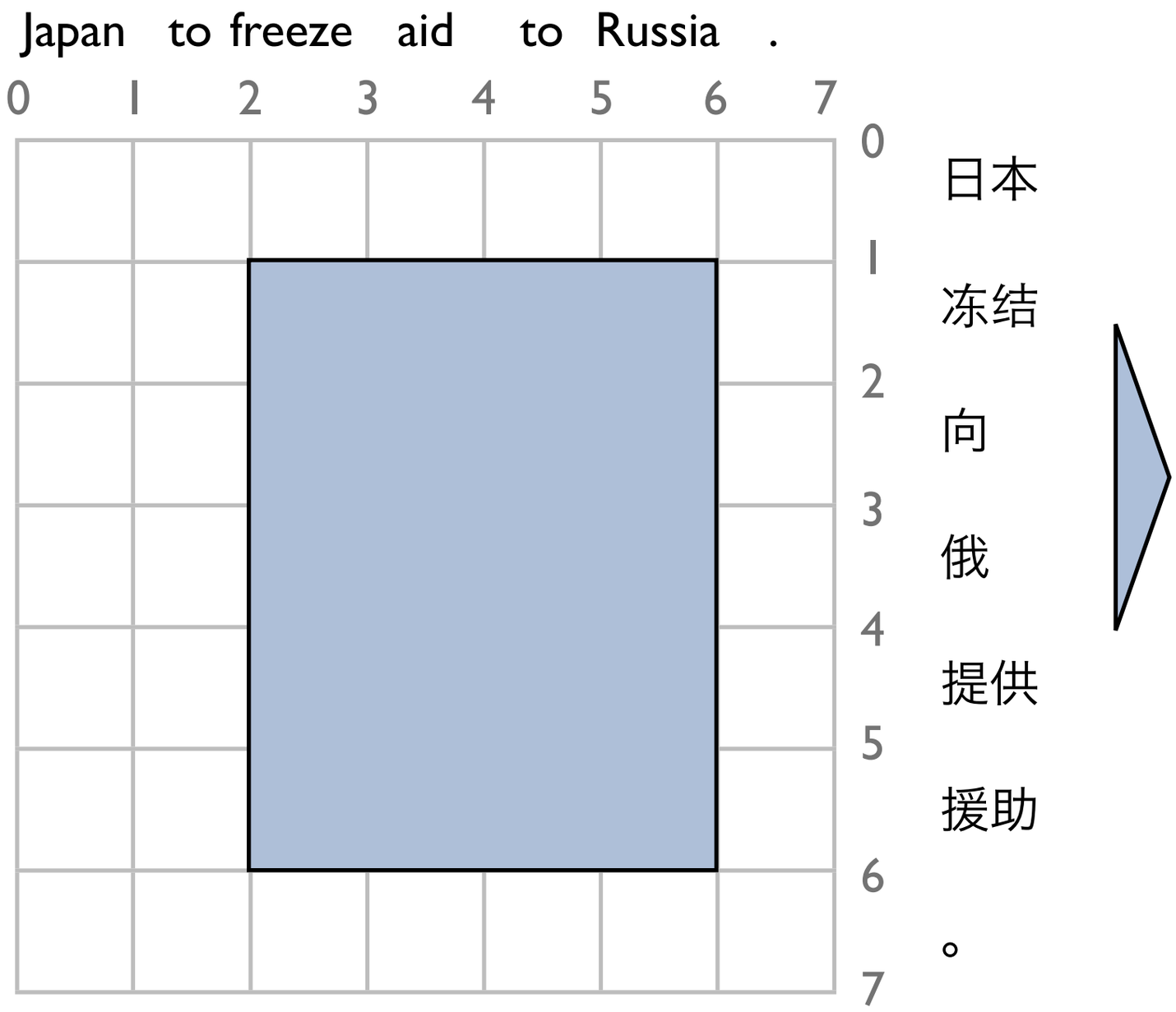
Indicator Variables:

a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:

w: weights

Phrase Alignment as Integer Programming



$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

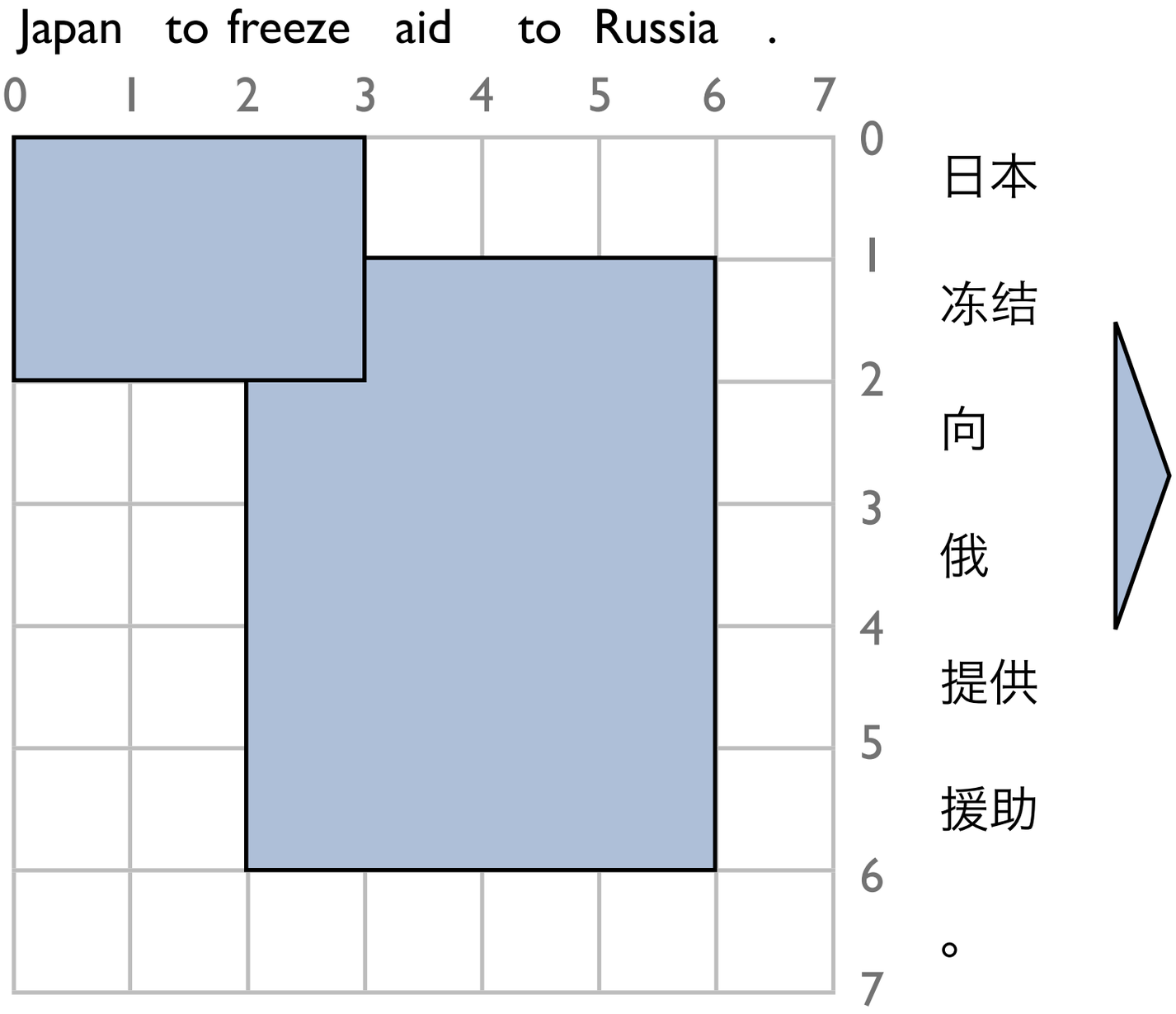
Indicator Variables:

a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:

w: weights

Phrase Alignment as Integer Programming



$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

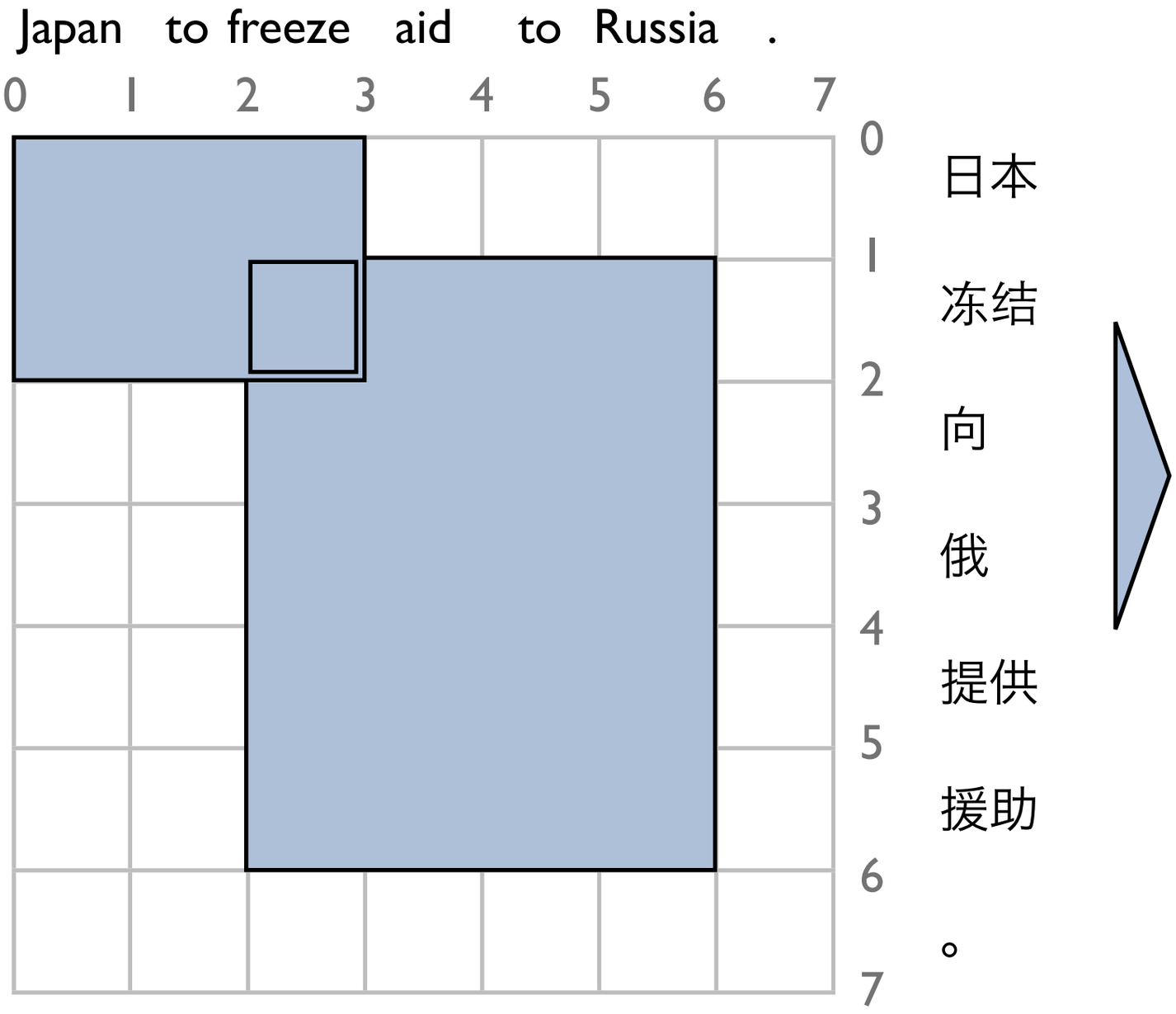
Indicator Variables:

a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:

w: weights

Phrase Alignment as Integer Programming



$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

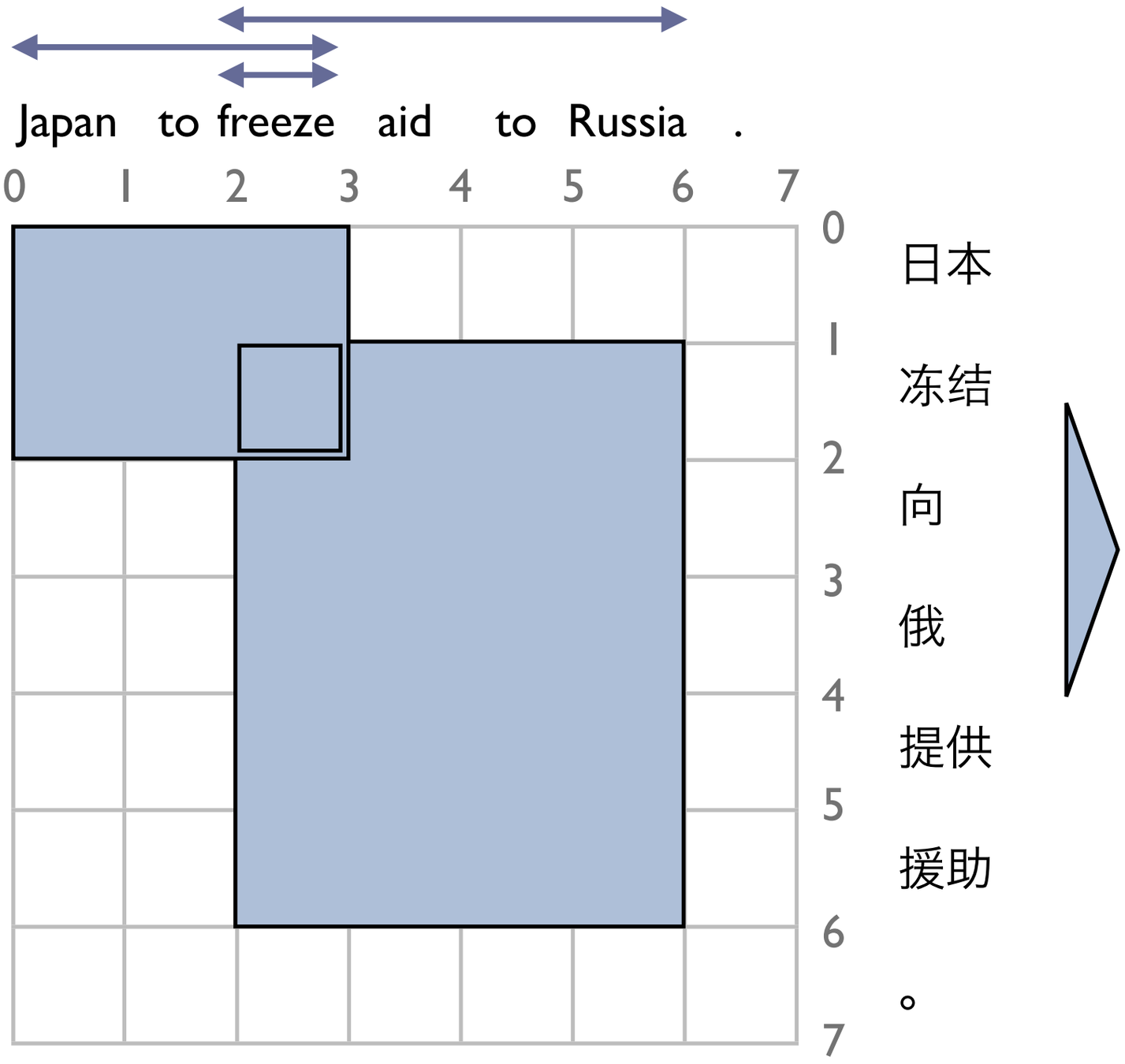
Indicator Variables:

a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:

w: weights

Phrase Alignment as Integer Programming



$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

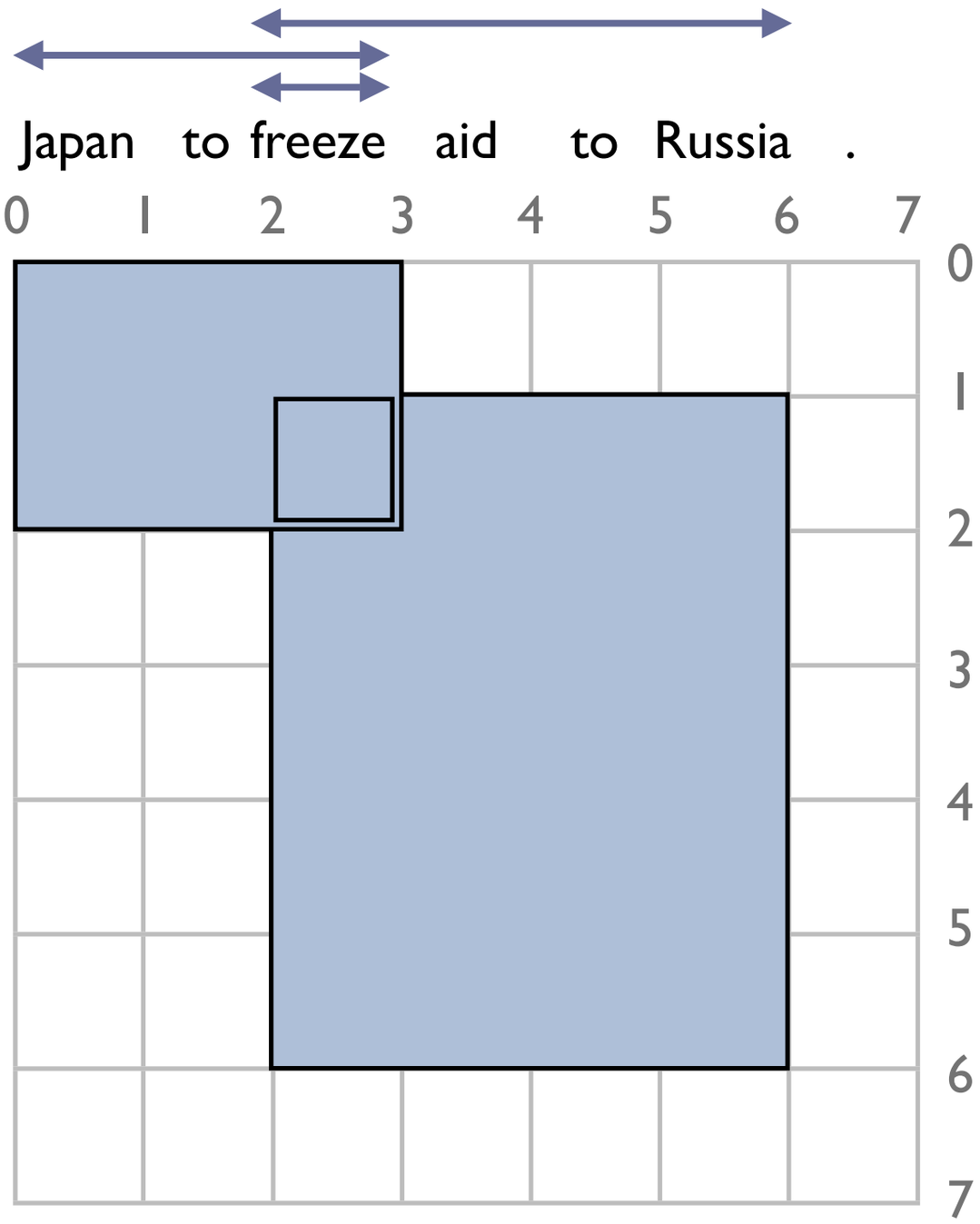
Indicator Variables:

a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:

w: weights

Phrase Alignment as Integer Programming



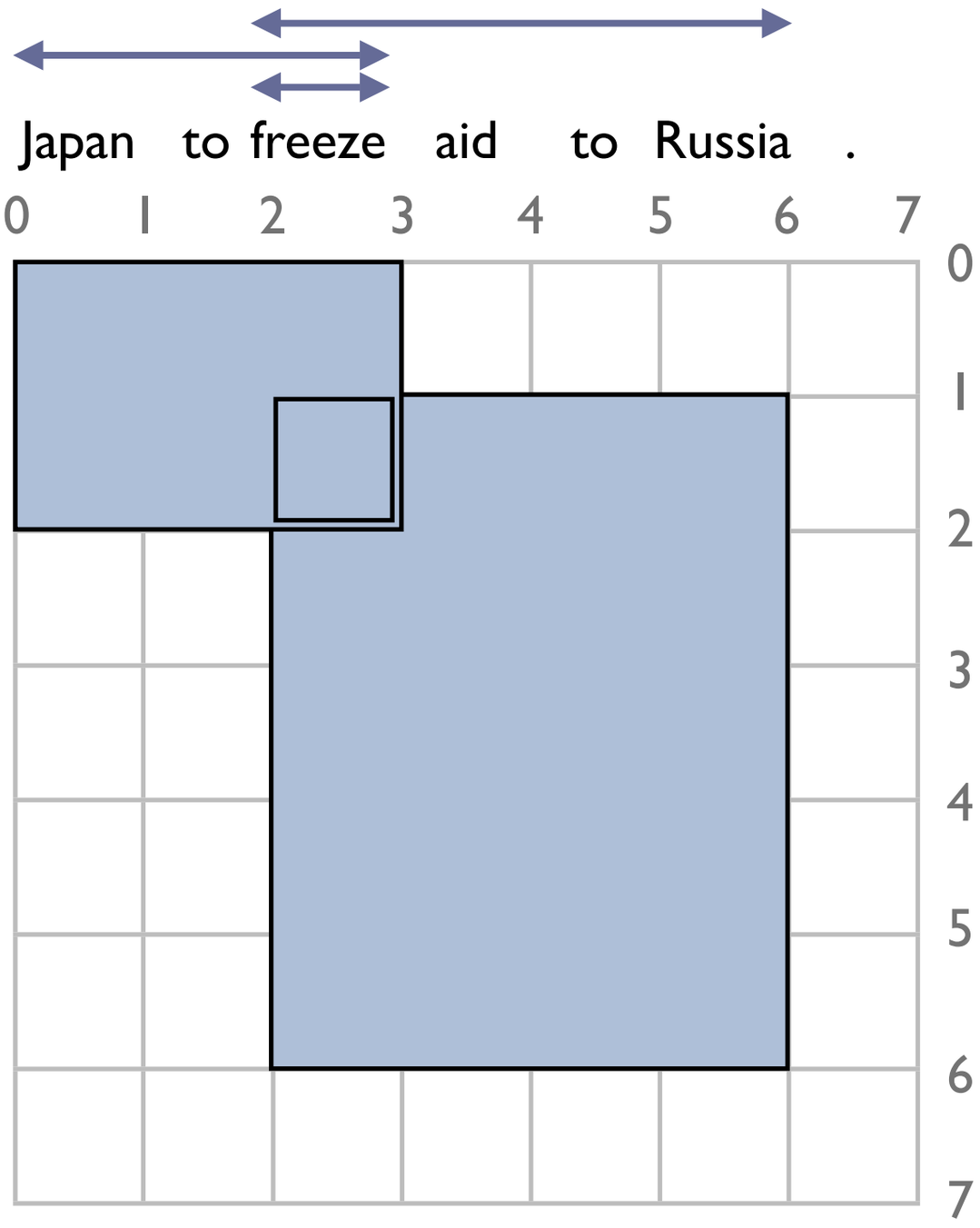
$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

$$\text{s.t.} \sum_{i,j:i<x\leq j} e_{i,j} = 1$$

Indicator Variables:
a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:
w: weights

Phrase Alignment as Integer Programming



$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

$$\text{s.t.} \sum_{i,j:i<x\leq j} e_{i,j} = 1$$

$$\sum_{k,l:k<y\leq l} f_{k,l} = 1$$

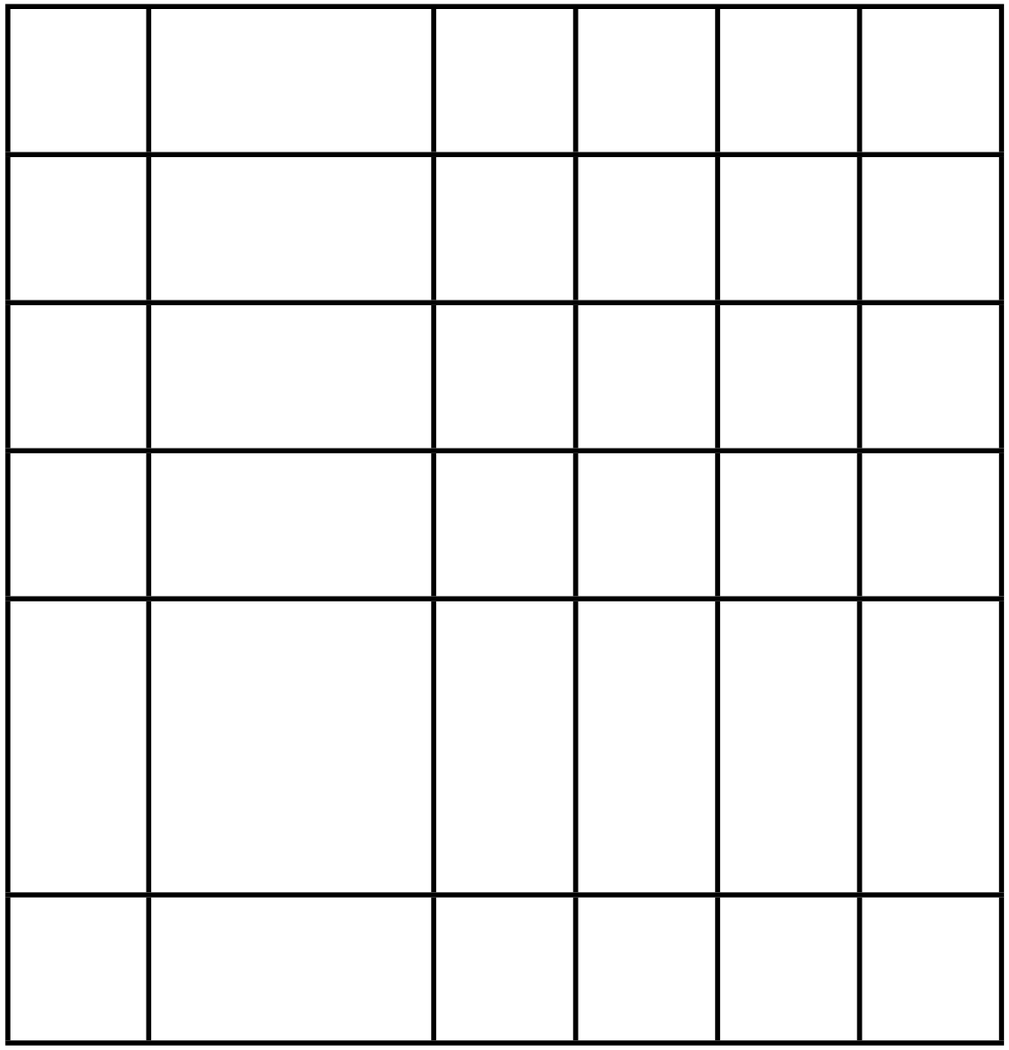
Indicator Variables:
a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:
w: weights

Phrase Alignment as Integer Programming

Japan to freeze aid to Russia .

1 2 3 4 5 6



1 日本

2 冻结

3 向

4 俄

5 提供

6 援助

。

$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

$$\text{s.t.} \sum_{i,j:i<x\leq j} e_{i,j} = 1$$

$$\sum_{k,l:k<y\leq l} f_{k,l} = 1$$

Indicator Variables:

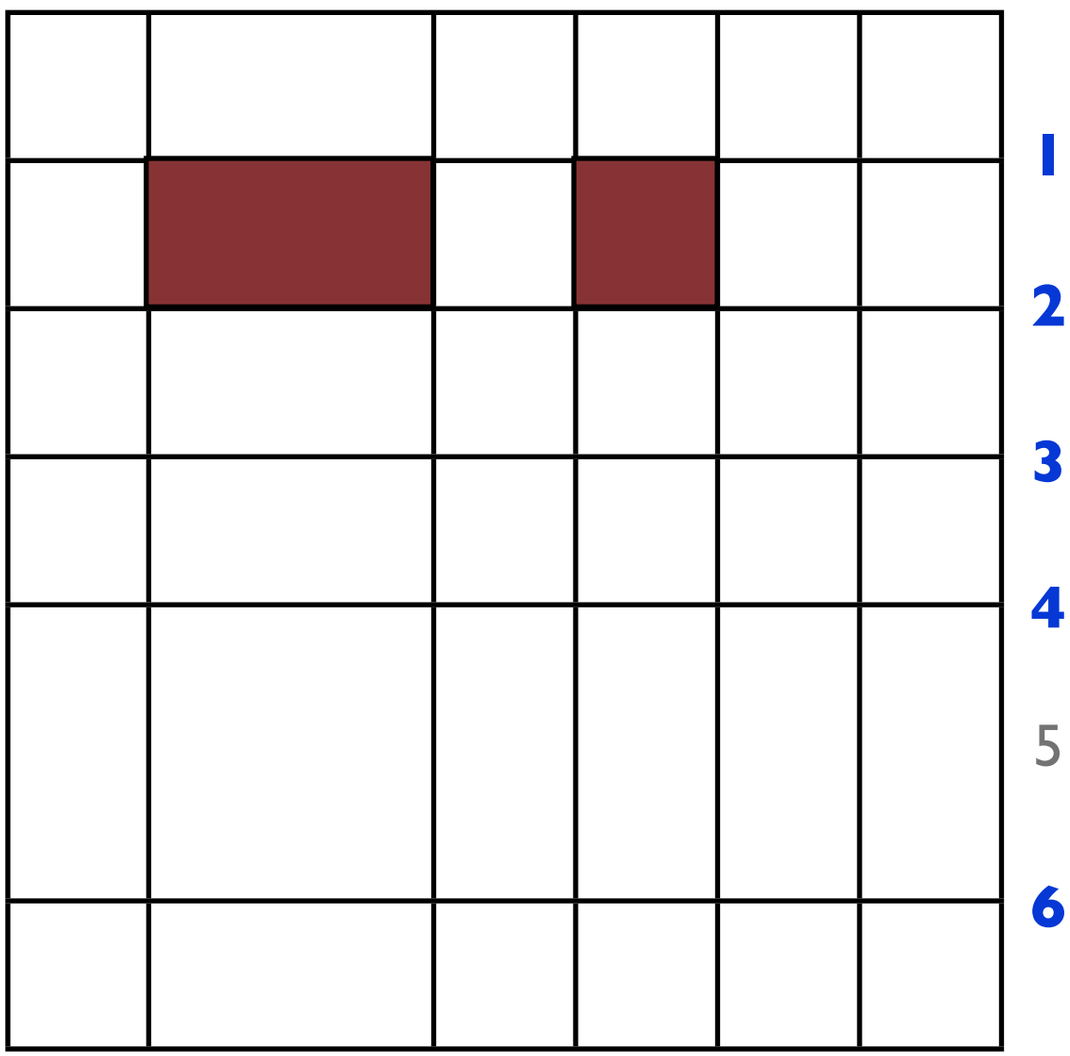
a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:

w: weights

Phrase Alignment as Integer Programming

Japan to freeze aid to Russia .



日本
冻结
向
俄
提供
援助
。

$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

$$\text{s.t.} \sum_{i,j:i<x\leq j} e_{i,j} = 1$$

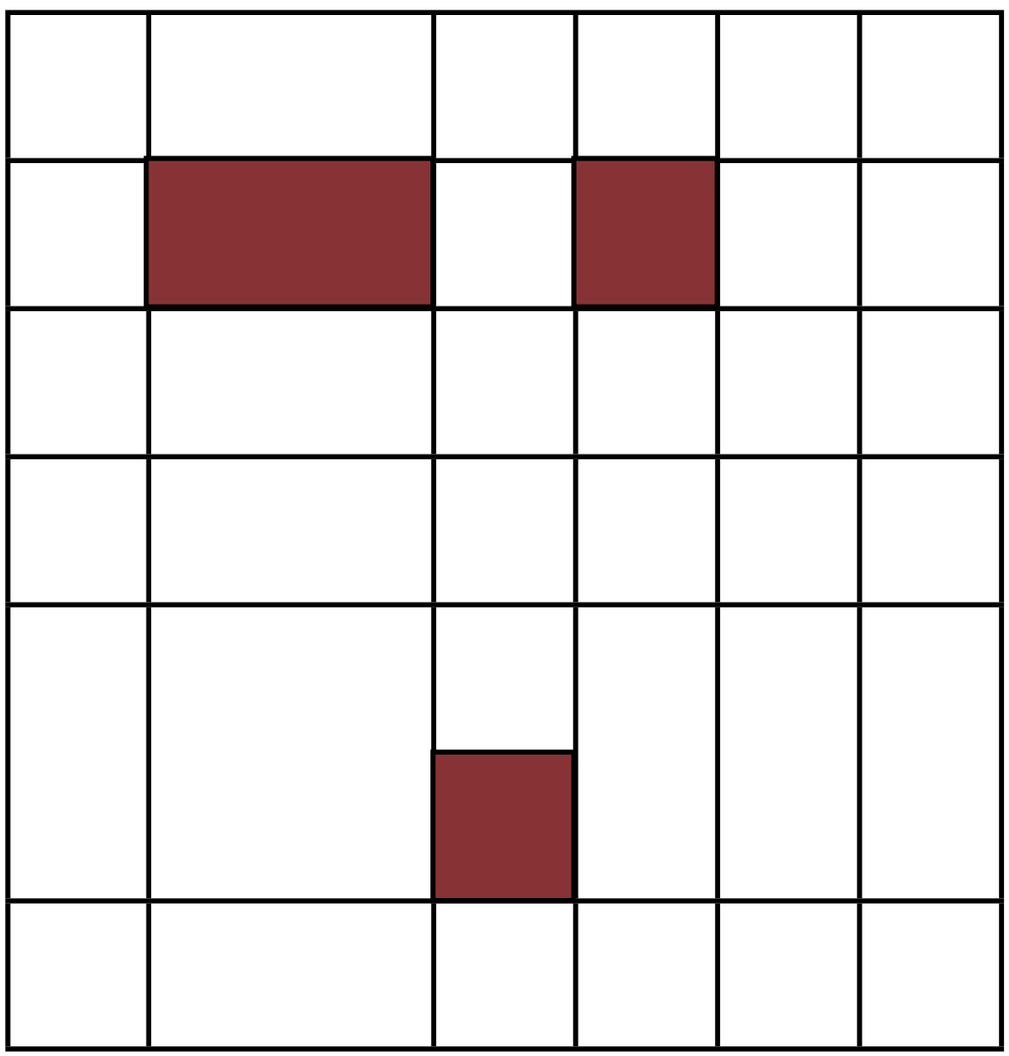
$$\sum_{k,l:k<y\leq l} f_{k,l} = 1$$

Indicator Variables:
a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:
w: weights

Phrase Alignment as Integer Programming

Japan to freeze aid to Russia .



1 日本
2 冻结
3 向
4 俄
5 提供
6 援助
。

$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

$$\text{s.t.} \sum_{i,j:i<x\leq j} e_{i,j} = 1$$

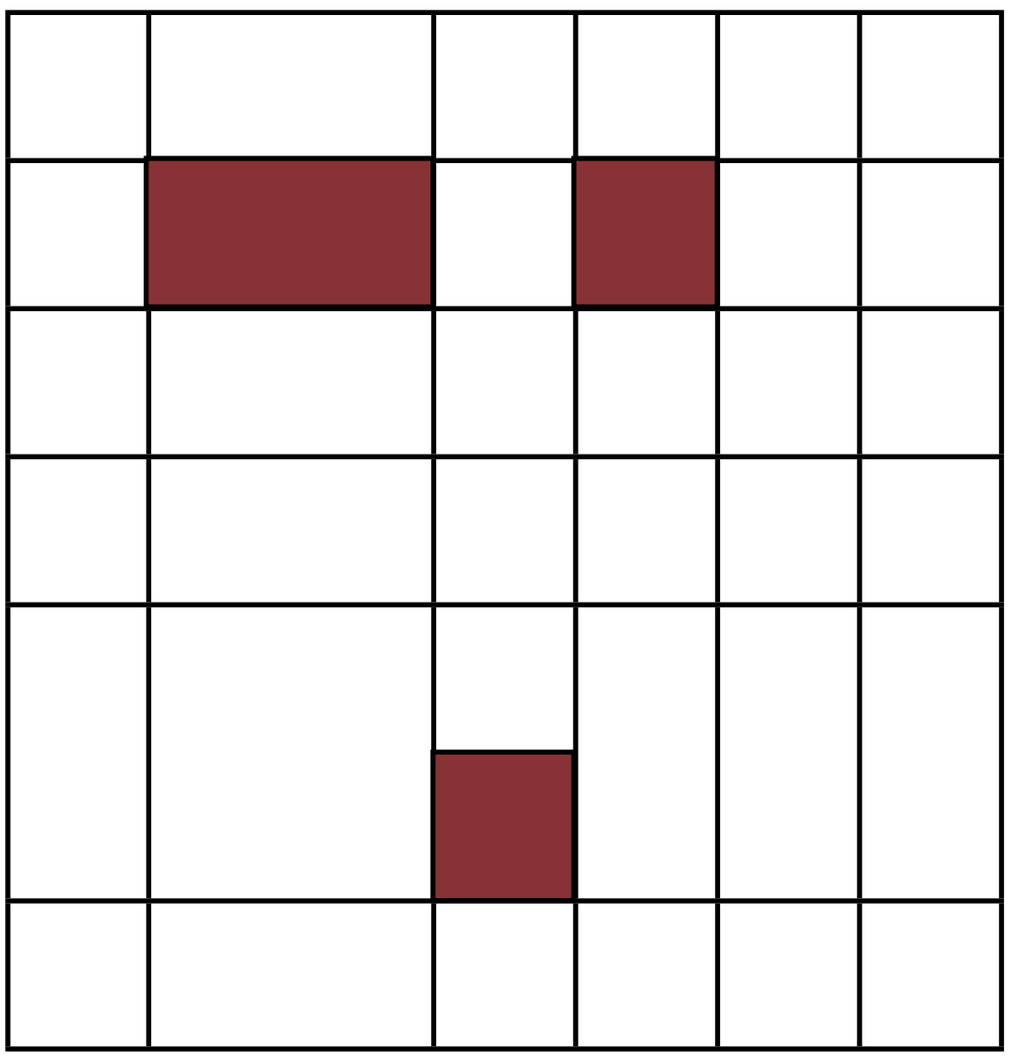
$$\sum_{k,l:k<y\leq l} f_{k,l} = 1$$

Indicator Variables:
a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:
w: weights

Phrase Alignment as Integer Programming

Japan to freeze aid to Russia .



日本
冻结
向
俄
提供
援助
。

$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

$$\text{s.t.} \sum_{i,j:i<x\leq j} e_{i,j} = 1$$

$$\sum_{k,l:k<y\leq l} f_{k,l} = 1$$

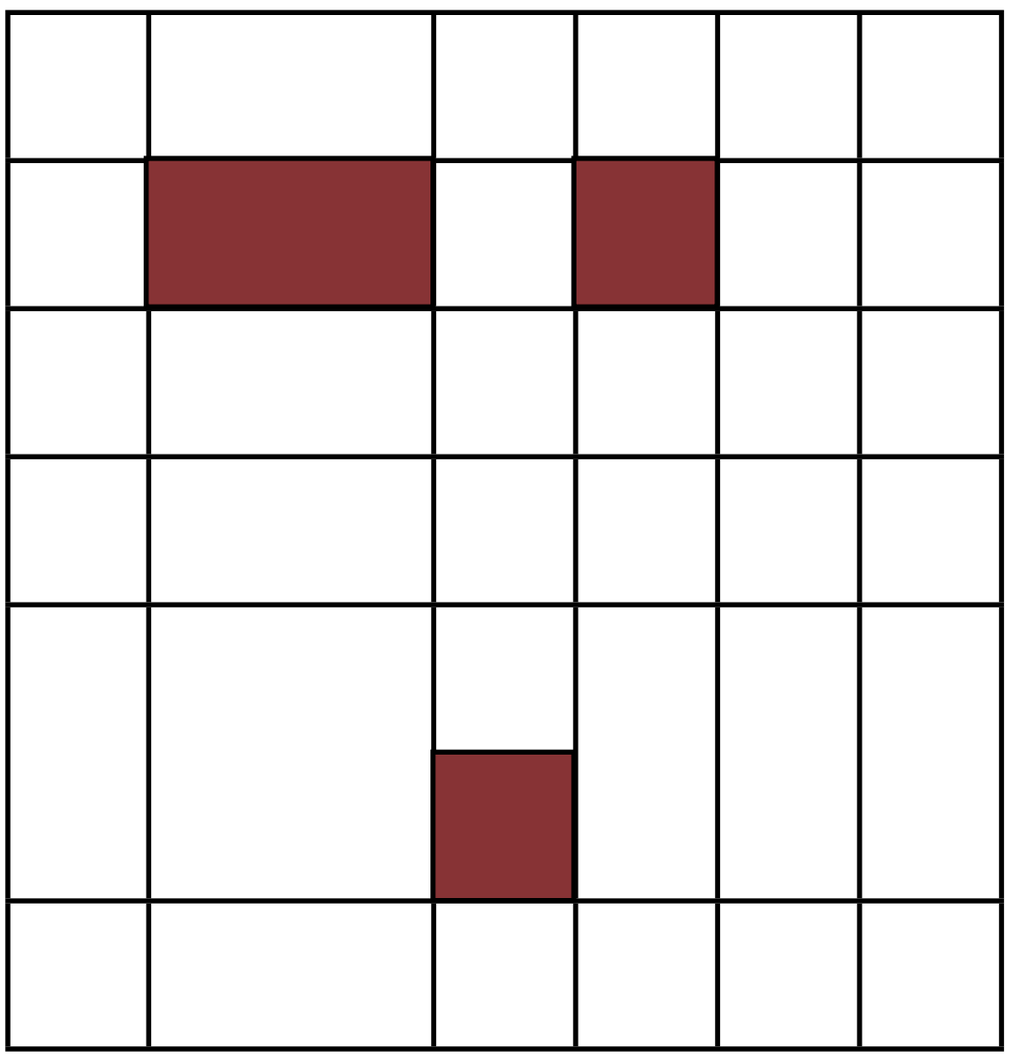
$$\sum_{k,l} a_{i,j,k,l} = e_{i,j}$$

Indicator Variables:
a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:
w: weights

Phrase Alignment as Integer Programming

Japan to freeze aid to Russia .



日本
冻结
向
俄
提供
援助
。

$$\max \sum_{i,j,k,l} w_{i,j,k,l} \cdot a_{i,j,k,l}$$

$$\text{s.t.} \sum_{i,j:i<x\leq j} e_{i,j} = 1$$

$$\sum_{k,l:k<y\leq l} f_{k,l} = 1$$

$$\sum_{k,l} a_{i,j,k,l} = e_{i,j}$$

$$\sum_{i,j} a_{i,j,k,l} = f_{k,l}$$

Indicator Variables:
a: phrase alignment **e**: English segmentation **f**: foreign segmentation

Constants:
w: weights