Machine Translation

80 Years of
Machine Translation
in 1 Hour

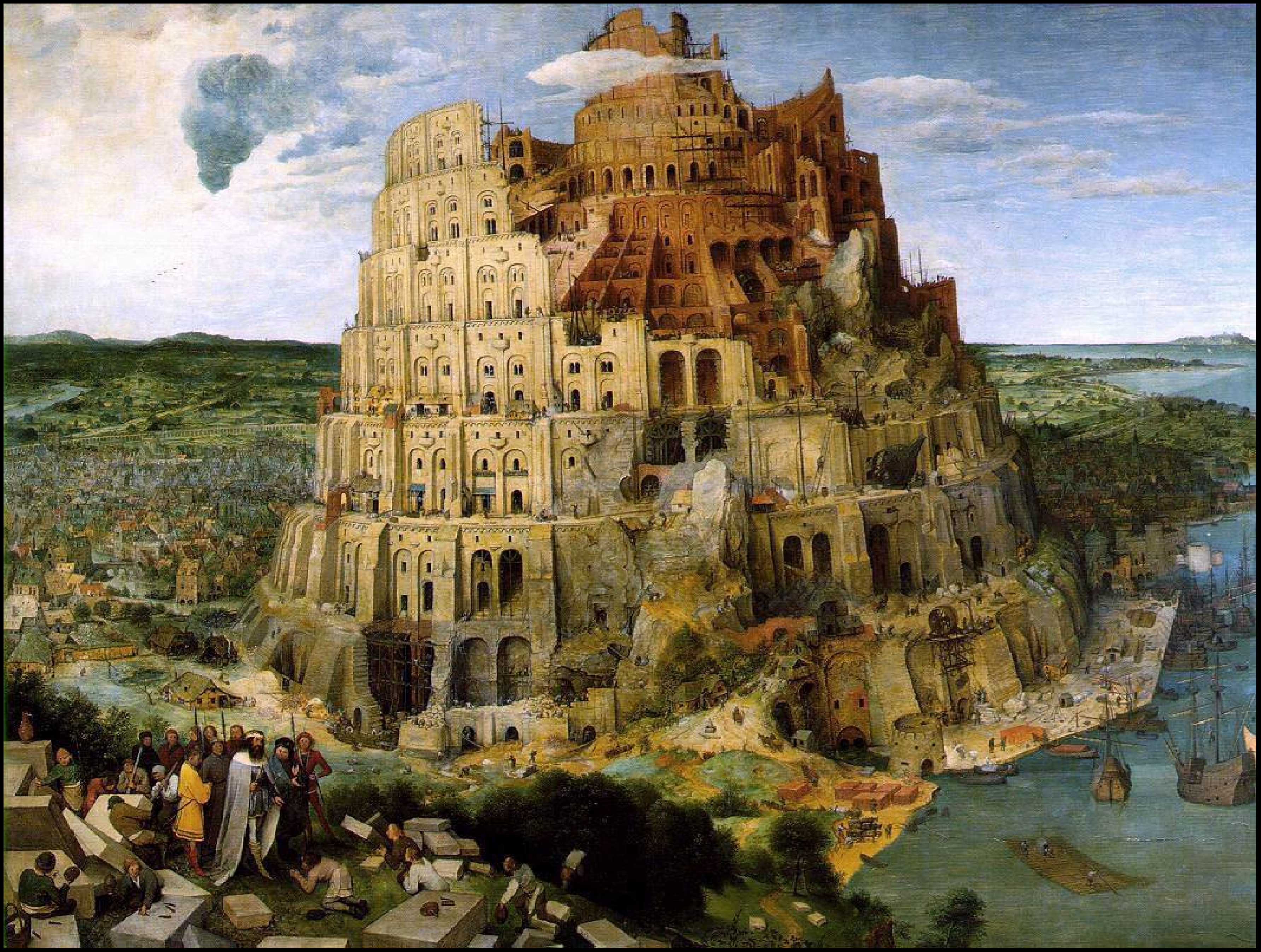
80 Years of Machine Translation in 1 Hour

For more hours: mt-class.org

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。

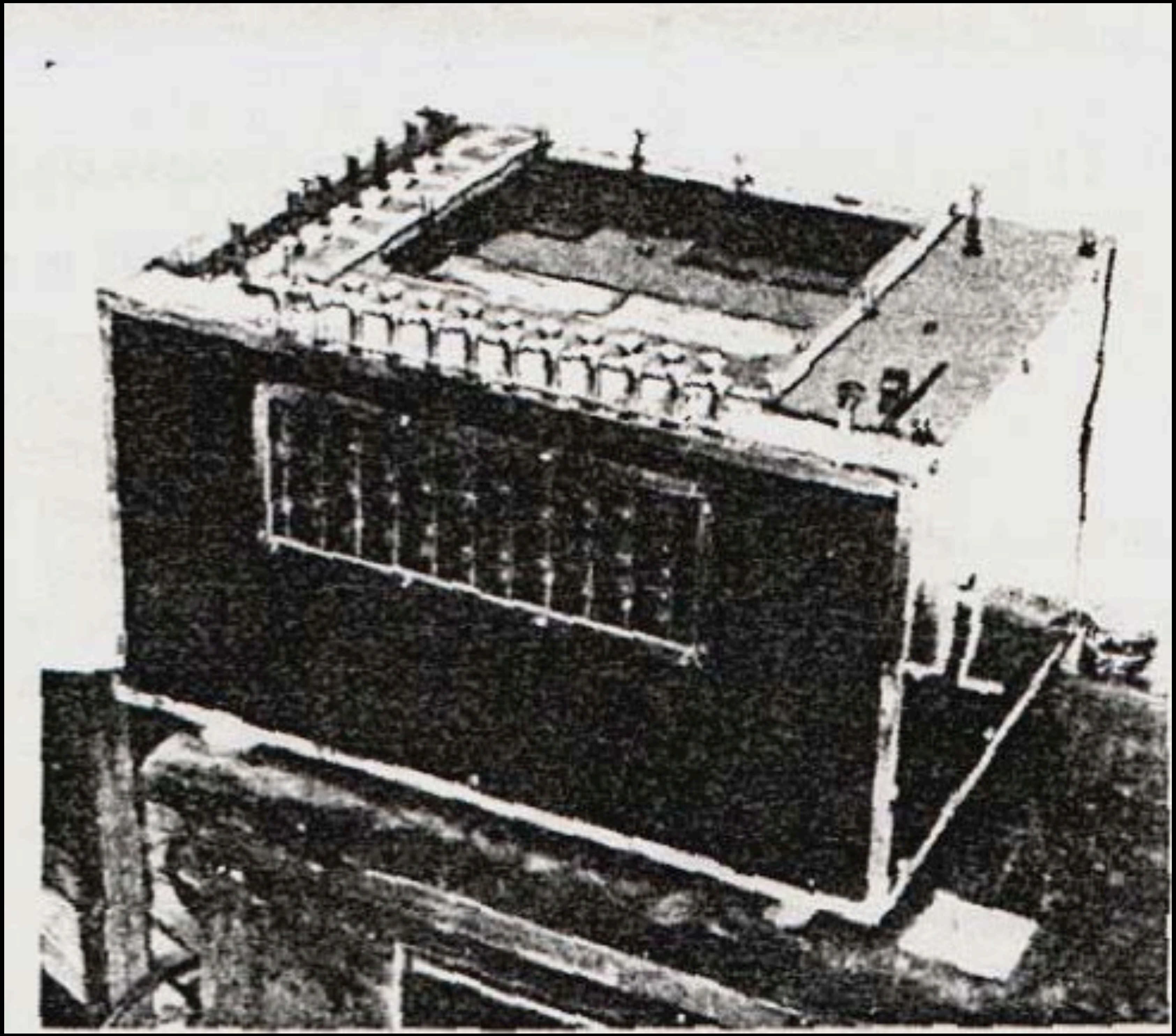
虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。

However , the sky remained clear under the strong north wind .

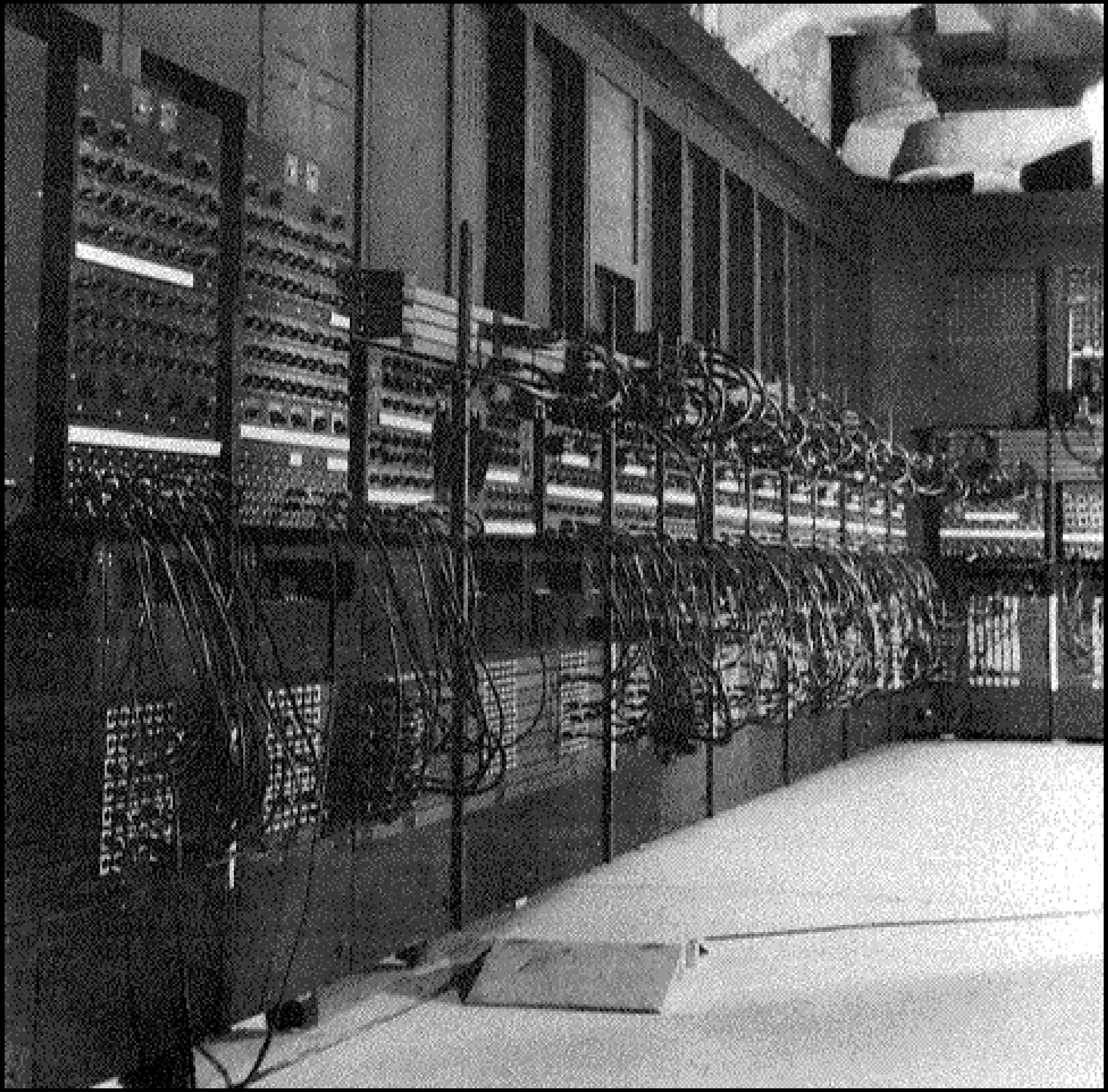


The Tower of Babel

Pieter Brueghel the Elder (1563)



Georges Artsrouni's "mechanical brain",
patented 1933 (France)

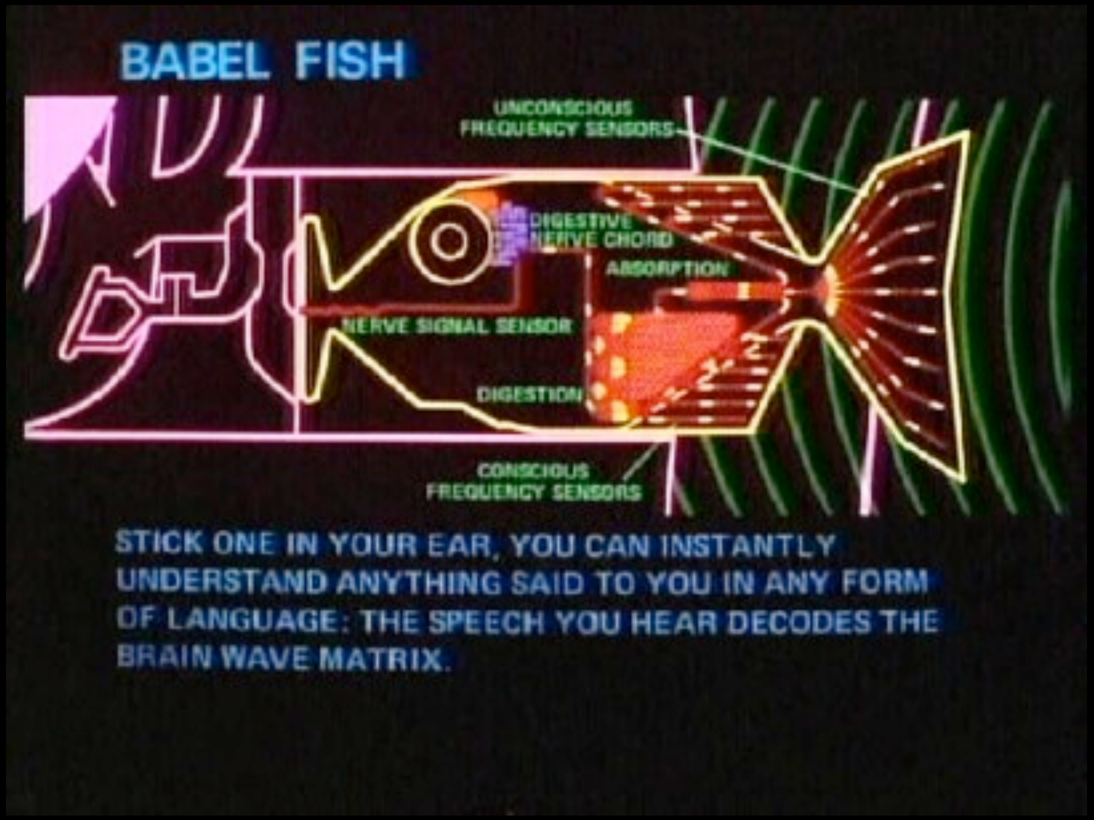


ENIAC (1946)

When I look at an article in Russian, I say: "This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode."

Warren Weaver (1949)

Popular view of MT in 2003



Popular view of MT in 2013

Google Translate

Break through language barriers.

bing™ Translator

SDL freeTranslation

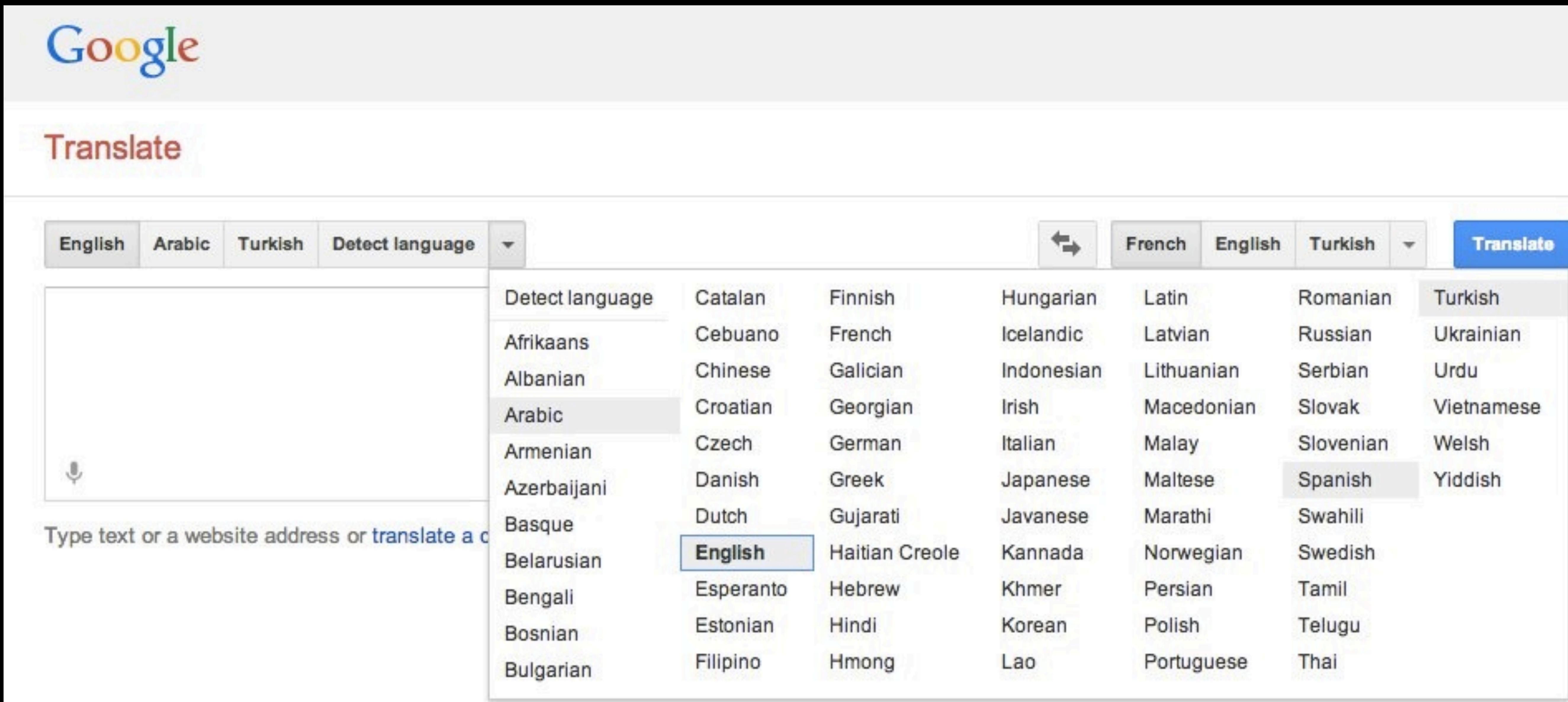
Statistical Machine Translation Live

4/28/2006 03:40:00 PM

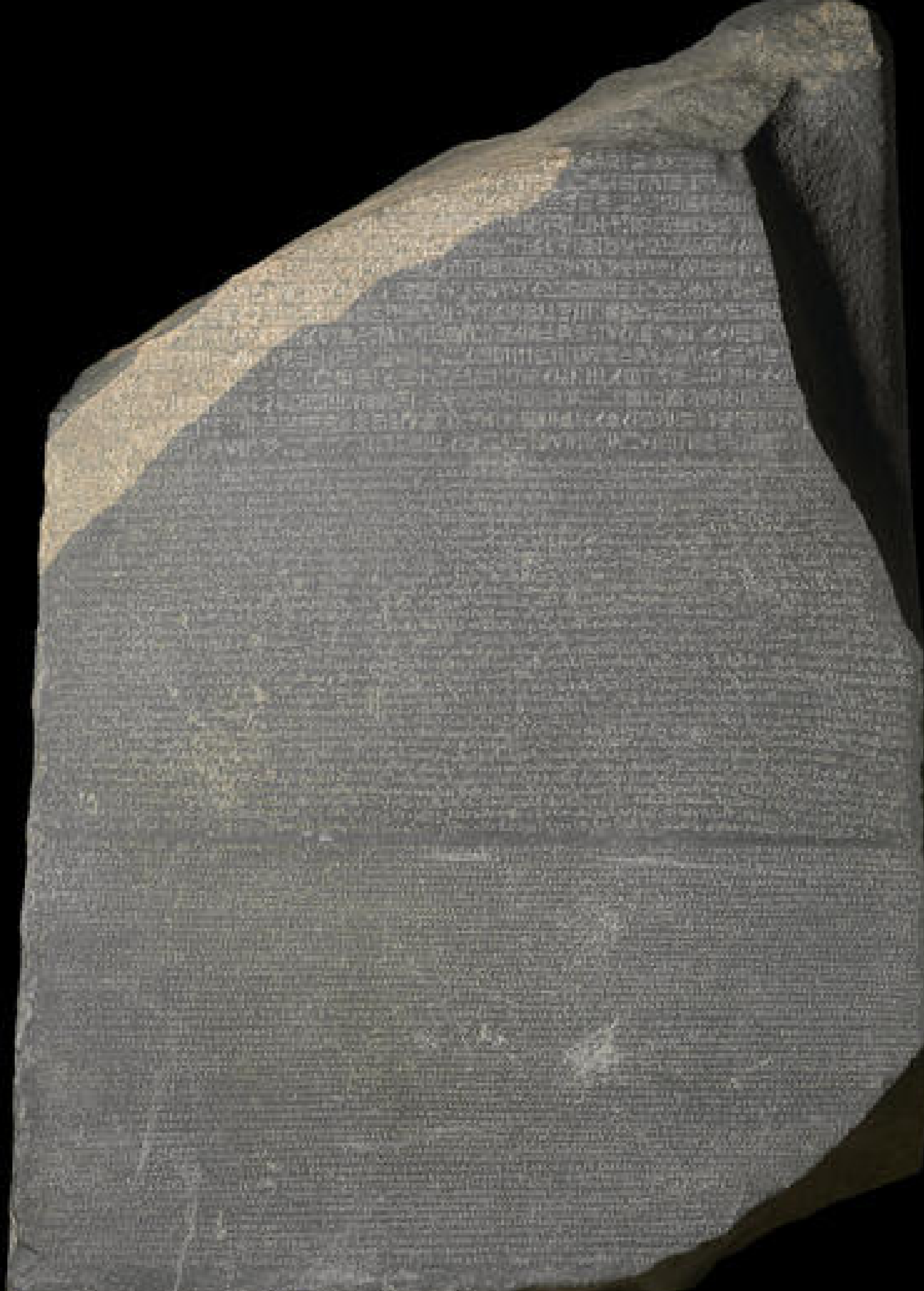
Franz Och

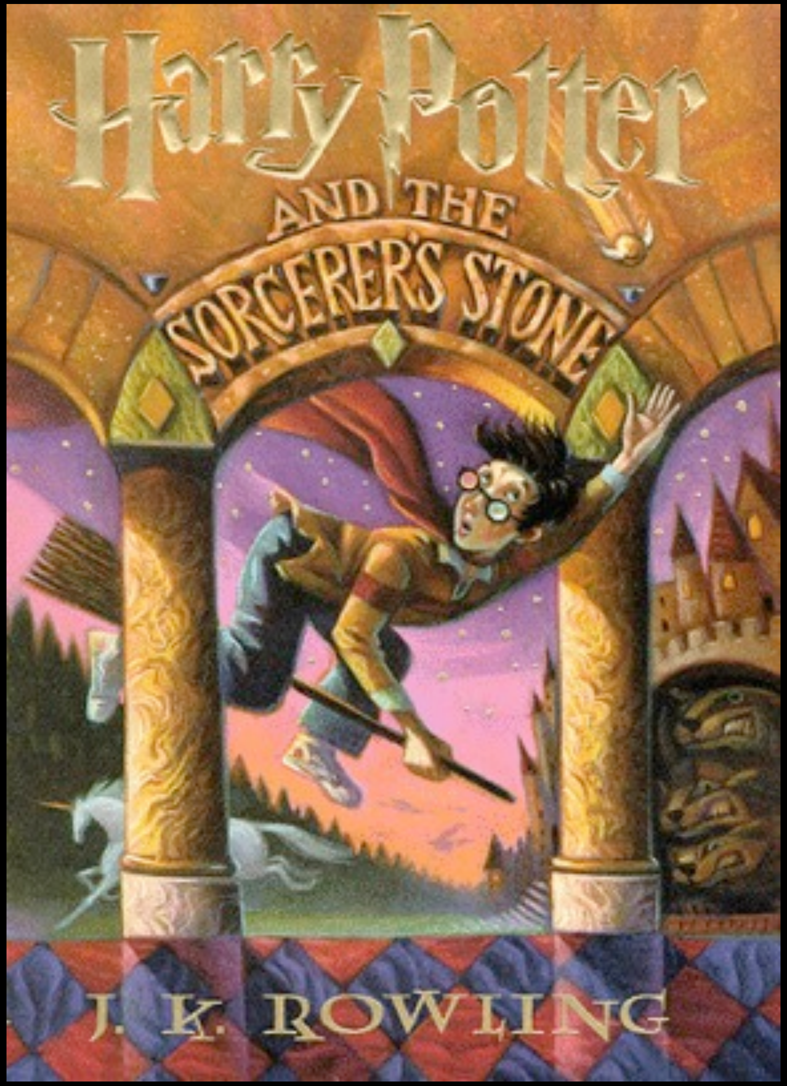
Because we want to provide everyone with access to all the world's information, including information written in every language, one of the exciting projects at Google Research is machine translation... Now you can see the results for yourself. We recently launched an online version of our system for Arabic-English and English-Arabic. Try it out!

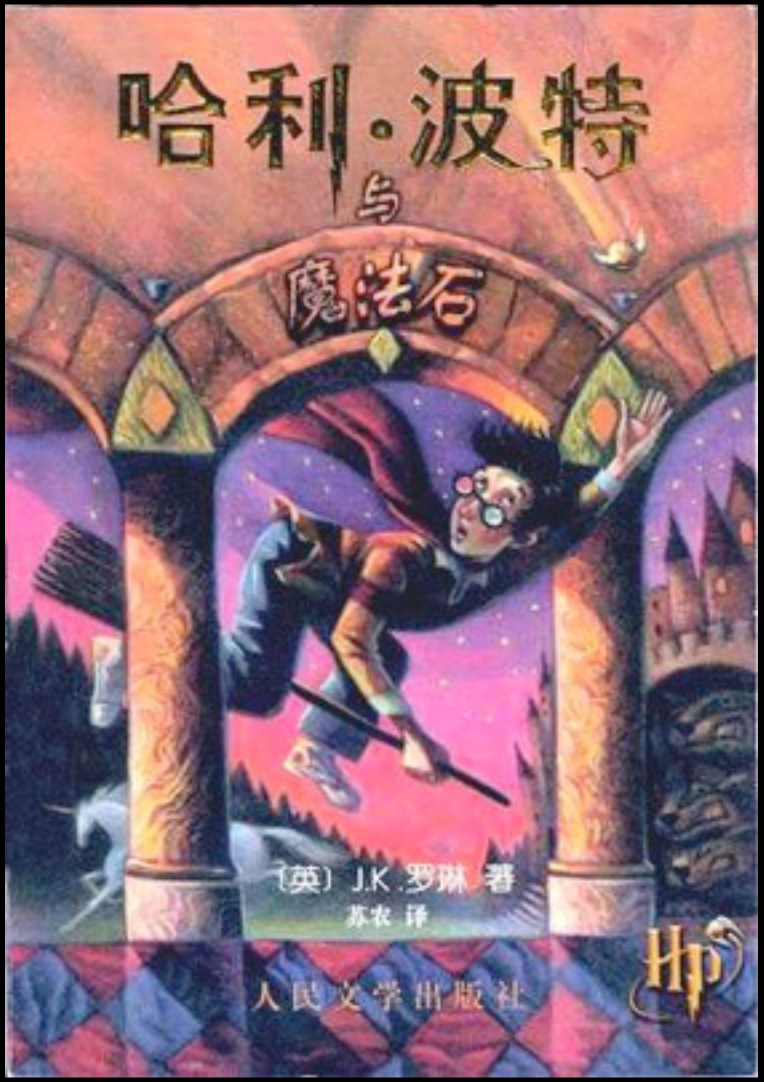
7 years later

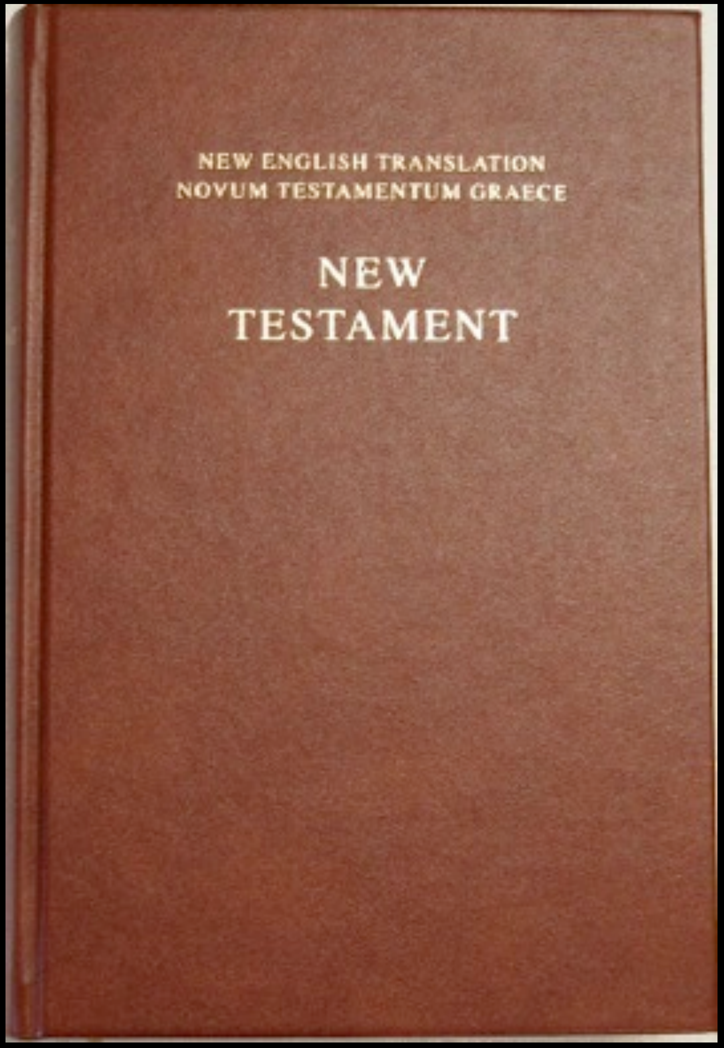


How did they do this?




Harry Potter
AND THE
SORCERER'S STONE
J. K. ROWLING

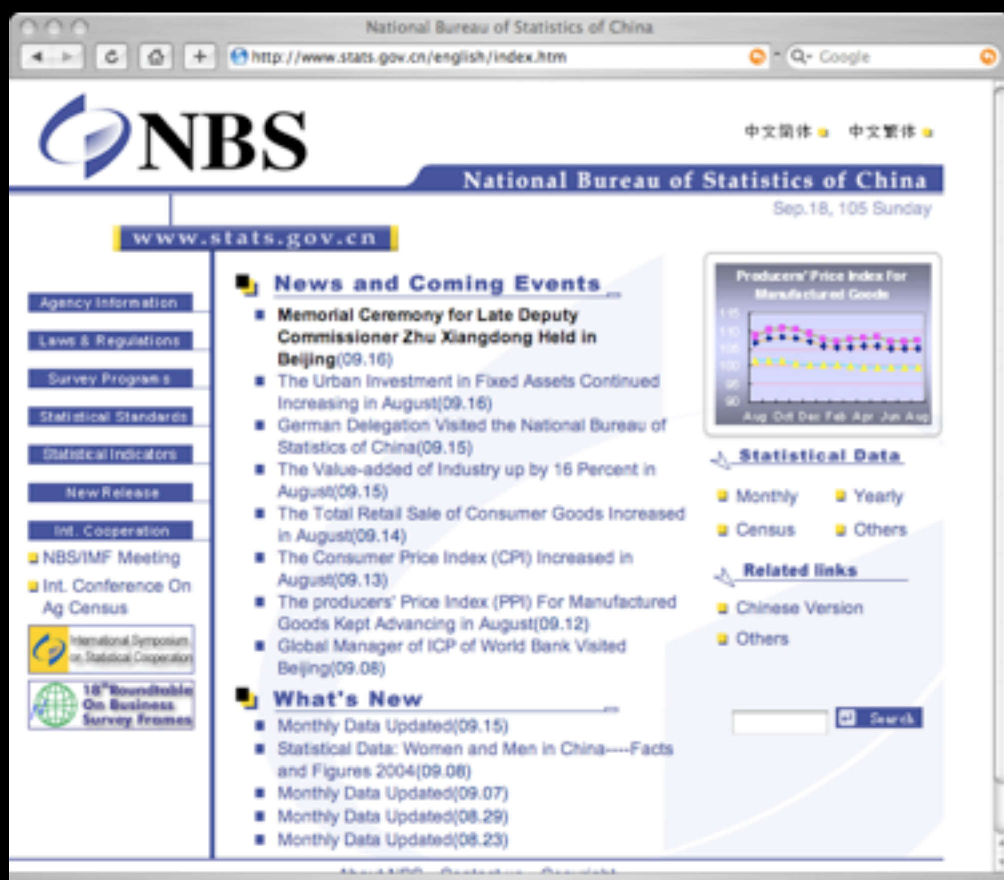

哈利·波特
与
魔法石
〔英〕J.K.罗琳 著
苏农 译
人民文学出版社
HP


NEW ENGLISH TRANSLATION
NOVUM TESTAMENTUM GRAECE
NEW
TESTAMENT

聖經
尋道本


www.un.org
http://www.un.org/english/
Google
We the peoples
Daily Briefing | Radio, TV, Photo | Documents, Maps | Publications, Stamps, Databases | UN Works | Search
Peace & Security | Economic & Social Development | Human Rights | Humanitarian Affairs | International Law
Welcome to the
United Nations
UN Millennium Development Goals
United Nations News Centre
About the United Nations
Main Bodies
Conferences & Events
Member States
General Assembly President
Secretary-General
Situation in Iraq
Mideast Roadmap
Renewing the UN
UN Action against Terrorism
Issues on the UN Agenda
Civil Society / Business
UN Webcast
CyberSchoolBus
8 September 2005 >>
Home | Recent Additions | Employment | UN Procurement | Comments | Q & A | UN System Sites | Index
عربي | 中文 | English | Français | Русский | Español
Copyright, United Nations, 2000-2005 | Use of UN60 Logo | Terms of Use | Privacy Notice | Help
[Text version]
Live and On-Demand Webcasts, 24 Hours a Day: Click on UN Webcas


联合国主页
http://www.un.org/chinese/
Google
我们人民
每日简报 | 多媒体 | 文件与地图 | 出版物 邮票 数据库 | 服务全球 | 网址搜索
和平与安全 | 经济与社会发展 | 人权 | 人道主义事务 | 国际法
欢迎来到
联合国
联合国千年发展目标
联合国新闻
联合国概况
联合国主要机关
会议与活动
联合国会员国
联合国大会主席
联合国秘书长
伊拉克局势
中东路线图
更新联合国
反恐怖主义
联合国日常议题
民间团体 / 商业
联合国网络直播
空中校车
联大第60届会议一般性辩论
新增内容 | 工作机会 | 联合国采购 | 建议 | 问题与解答 | 其他网址 | 网址索引
عربي | 中文 | English | Français | Русский | Español
联合国2000-2005年版权|联合国60周年徽标使用准则|使用条件|隐私通告|帮助
[纯文字版]


National Bureau of Statistics of China
http://www.stats.gov.cn/english/index.htm
NBS
中文简体
中文繁体
National Bureau of Statistics of China
Sep.18, 105 Sunday
www.stats.gov.cn
Agency Information
Laws & Regulations
Survey Programs
Statistical Standards
Statistical Indicators
New Release
Int. Cooperation
NBS/IMF Meeting
Int. Conference On Ag Census
News and Coming Events
Memorial Ceremony for Late Deputy Commissioner Zhu Xiangdong Held in Beijing(09.16)
The Urban Investment in Fixed Assets Continued Increasing in August(09.16)
German Delegation Visited the National Bureau of Statistics of China(09.15)
The Value-added of Industry up by 16 Percent in August(09.15)
The Total Retail Sale of Consumer Goods Increased in August(09.14)
The Consumer Price Index (CPI) Increased in August(09.13)
The producers' Price Index (PPI) For Manufactured Goods Kept Advancing in August(09.12)
Global Manager of ICP of World Bank Visited Beijing(09.08)
What's New
Monthly Data Updated(09.15)
Statistical Data: Women and Men in China----Facts and Figures 2004(09.08)
Monthly Data Updated(09.07)
Monthly Data Updated(08.29)
Monthly Data Updated(08.23)
Producers' Price Index For Manufactured Goods
Statistical Data
Monthly
Yearly
Census
Others
Related links
Chinese Version
Others
Search

中华人民共和国国家统计局
http://www.stats.gov.cn/
中华人民共和国国家统计局
National Bureau of Statistics of China
中文简体
中文繁体
ENGLISH
www.stats.gov.cn
统计公报
统计数据
统计分析
统计出版
统计法规
统计管理
数据直报
统计链接
最新统计信息
最新统计动态
工业品出厂价格指数
重要公告
统计机构
统计动态
统计标准
统计制度
统计知识


中文简体
NBS
www.stats.gov.cn
News and Coming
Agency Information
Laws & Regulations
Survey Programs
Statistical Standards
Statistical Indicators
New Release
Int. Cooperation
NBS/IMF Meeting
Int. Conference On Ag Census
Memorial Ceremony for Late Deputy Commissioner Zhu Xiangdong Held in Beijing(09.16)
The Urban Investment in Fixed Assets Continued Increasing in August(09.16)
German Delegation Visited the National Bureau of Statistics of China(09.15)
The Value-added of Industry up by 16 Percent in August(09.15)
The Total Retail Sale of Consumer Goods Increased in August(09.14)
The Consumer Price Index (CPI) Increased in August(09.13)
The producers' Price Index (PPI) For Manufactured Goods Kept Advancing in August(09.12)
Global Manager of ICP of World Bank Visited Beijing(09.08)
What's New
Monthly Data Updated(09.15)
Statistical Data: Women and Men in China----Facts and Figures 2004(09.08)
Monthly Data Updated(09.07)
Monthly Data Updated(08.29)
Monthly Data Updated(08.23)
Statistical Data
Monthly
Yearly
Census
Others
Related links
Chinese Version
Others
Search


中文繁体
ENGLISH
中华人民共和国国家
National Bureau of Statistics
统计公报
统计数据
统计分析
统计出版
统计法规
统计管理
数据直报
统计链接
最新统计信息
最新统计动态
重要公告
统计机构
统计动态
统计标准
统计制度
统计知识

CLASSIC SOUPS

			Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲吞	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲吞湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

CLASSIC SOUPS

			Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲吞	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲吞湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

CLASSIC SOUPS

			Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲吞	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲吞湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

CLASSIC SOUPS

	No.		Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲吞	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲吞湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

CLASSIC SOUPS

			Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲吞	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲吞湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

CLASSIC SOUPS

			Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲吞	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲吞湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

CLASSIC SOUPS

			Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲呑	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲呑湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

CLASSIC SOUPS

			Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲吞	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲吞湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

CLASSIC SOUPS

			Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲呑	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲呑湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

CLASSIC SOUPS

			Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲吞	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲吞湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

...But the probability that an event has happened is the same as the probability I have to guess right if I guess it has happened. Wherefore the following proposition is evident: If there be two subsequent events, the probability of the 2d b/N and the probability both together P/N, and it being 1st discovered that the 2d event has also happened, the probability I am right is P/b.

Thomas Bayes

...But the probability that an event has happened is the same as the probability I have to guess right if I guess it has happened. Wherefore the following proposition is evident: If there be two subsequent events, the probability of the 2d b/N and the probability both together P/N, and it being 1st discovered that the 2d event has also happened, the probability I am right is P/b.

Thomas Bayes

(image by Chris Dyer)

Bayes' Rule

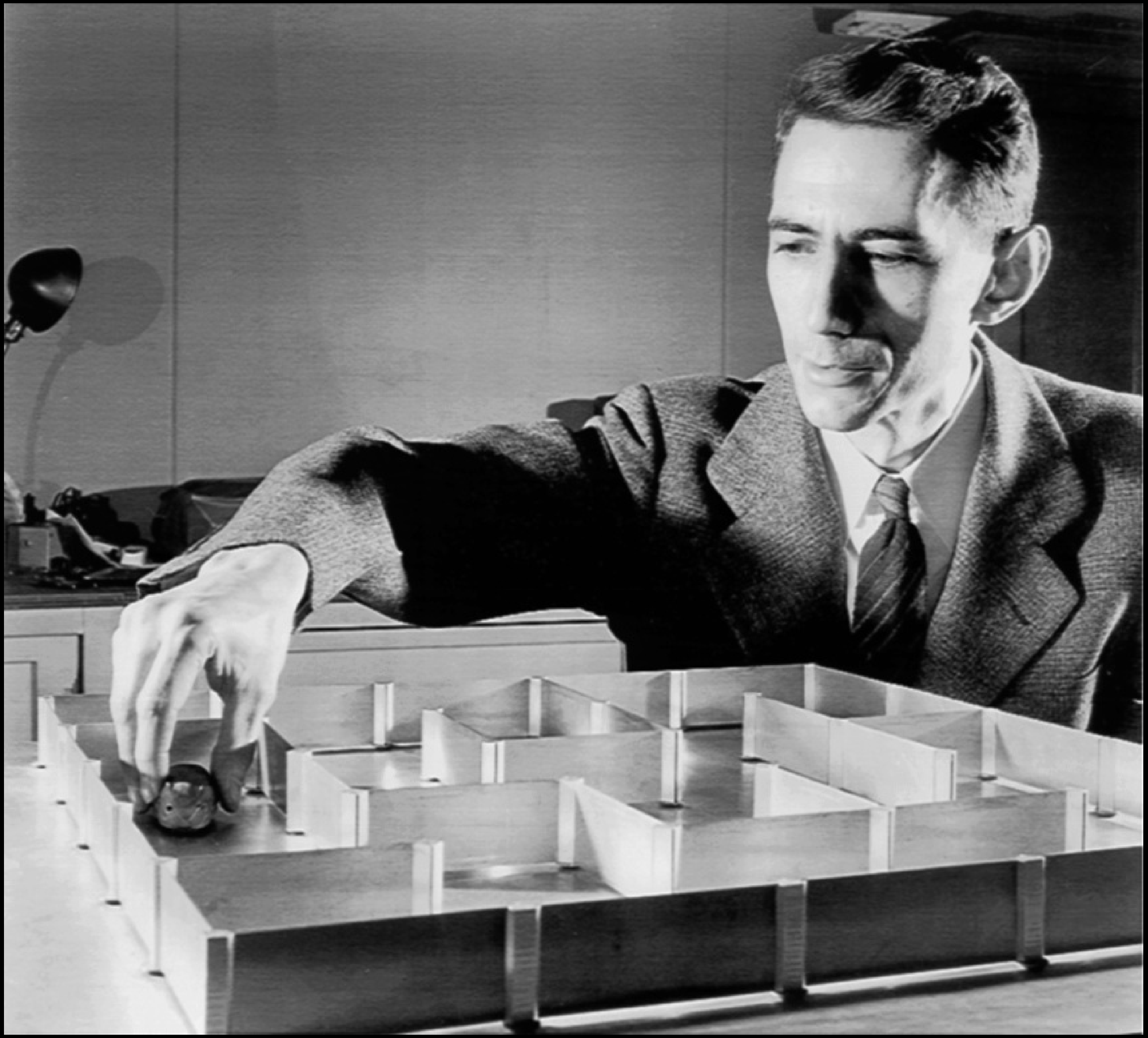
$$p(English|Chinese) = \frac{\overbrace{p(English)}^{\text{prior}} \times \overbrace{p(Chinese|English)}^{\text{likelihood}}}{\underbrace{p(Chinese)}_{\text{normalization term (ensures we're working with valid probabilities).}}}$$

When I look at an article in Russian, I say: "This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode."

Warren Weaver (1949)

THE MATHEMATICAL THEORY OF COMMUNICATION

by Claude E. Shannon and Warren Weaver



Claude Shannon

Bayes' Rule

$$p(English|Chinese) = \frac{p(English) \times p(Chinese|English)}{p(Chinese)}$$

prior

likelihood

normalization term (ensures we're working with valid probabilities).

Noisy Channel

$$p(English|Chinese) = \frac{p(English) \times p(Chinese|English)}{p(Chinese)}$$

signal model

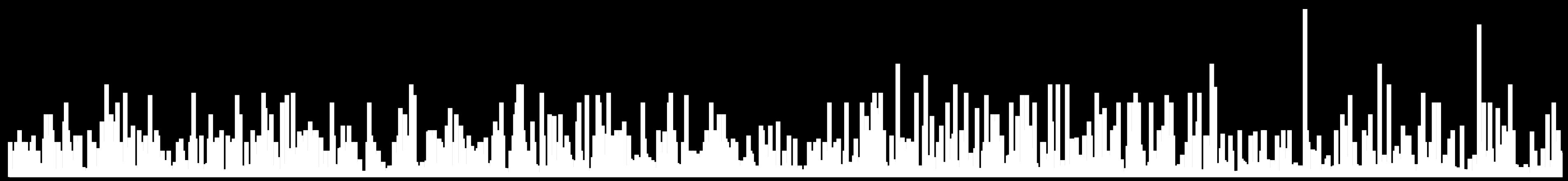
channel model

normalization term (ensures we're working with valid probabilities).

Machine Translation

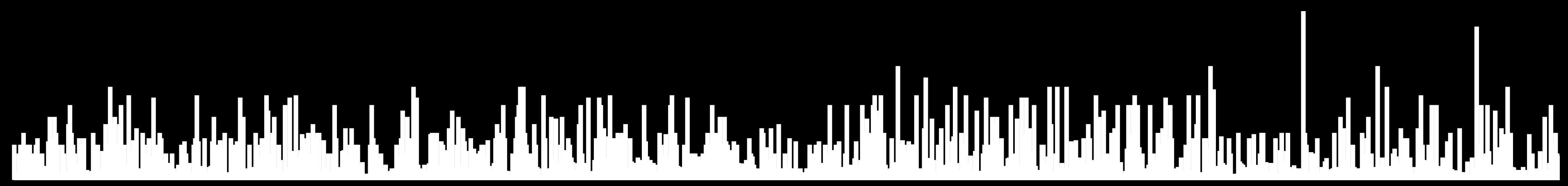
$$p(English|Chinese) = \frac{\underbrace{p(English)}_{\text{language model}} \times \underbrace{p(Chinese|English)}_{\text{translation model}}}{\underbrace{p(Chinese)}_{\text{normalization term (ensures we're working with valid probabilities).}}}$$

$$p(Chinese|English)$$

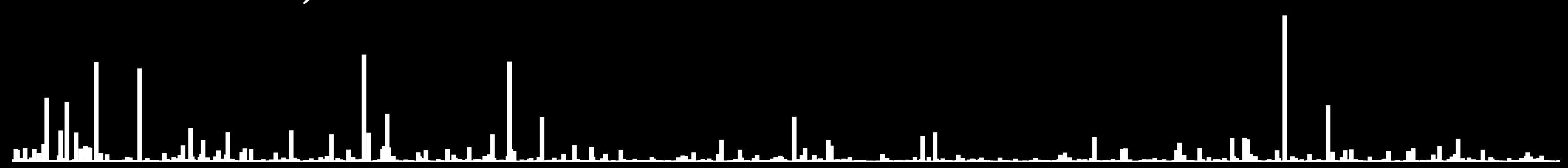


$$English$$

$$p(Chinese|English)$$

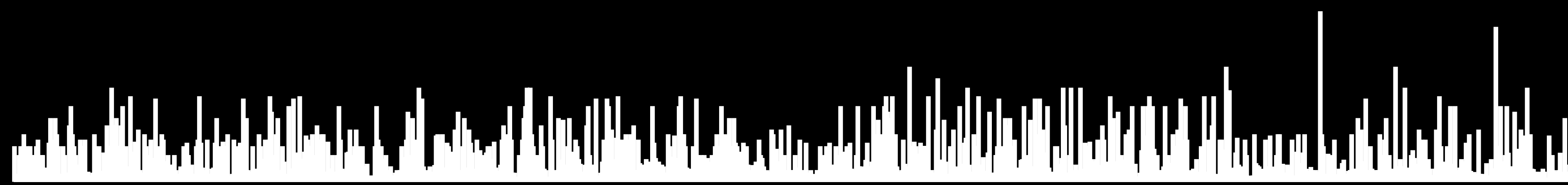


$$\times\ p(English)$$

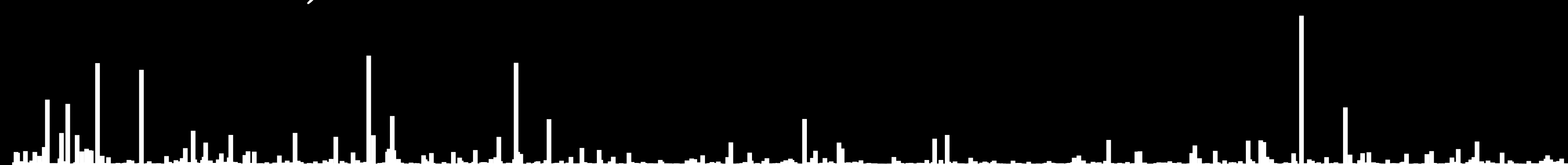


$English$

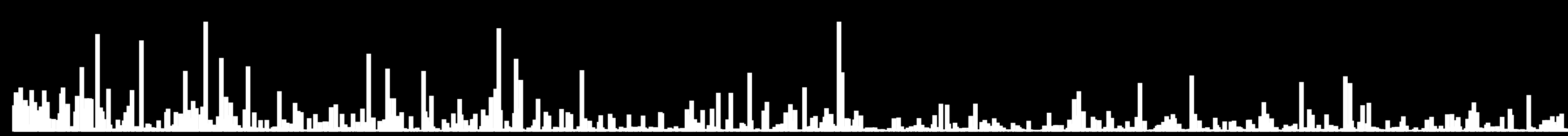
$$p(Chinese|English)$$



$$\times\ p(English)$$



$$\sim p(English|Chinese)$$



$English$

Machine Translation

$$p(English|Chinese) =$$

$$\frac{p(English) \times p(Chinese|English)}{p(Chinese)}$$

language model

translation model

normalization term (ensures we're working with valid probabilities).

Machine Translation

$$p(English|Chinese) \sim$$

$$p(English) \times p(Chinese|English)$$

Machine Translation

$$p(English|Chinese) \sim$$

$$p(English) \times p(Chinese|English)$$

What is the probability of an English sentence?

Machine Translation

$$p(English|Chinese) \sim$$

$$p(English) \times p(Chinese|English)$$

What is the probability of an English sentence?

What is the probability of a Chinese sentence, given a particular English sentence?

Language Models

Our language model must assign a probability to *every possible English sentence*.

Language Models

Our language model must assign a probability to *every possible English sentence*.

Q: What should this model look like?

Language Models

Our language model must assign a probability to *every possible English sentence*.

Q: What should this model look like?

Q2: What is the dumbest thing you can think of?

Language Models

Our language model must assign a probability to *every possible English sentence*.

Q: What should this model look like?

Q2: What is the dumbest thing you can think of?

A: An n-gram model.

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

__ __ __ __ __ __ __ __ __ __ __ __

$$p(English\ length | Chinese\ length)$$

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

$$p(English\ length|Chinese\ length)$$

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

$$p(Chinese\ word\ position)$$

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。 ε

___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

However

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

However

$$p(English\ word|Chinese\ word)$$

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

However

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

However

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

However ,

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

However ,

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

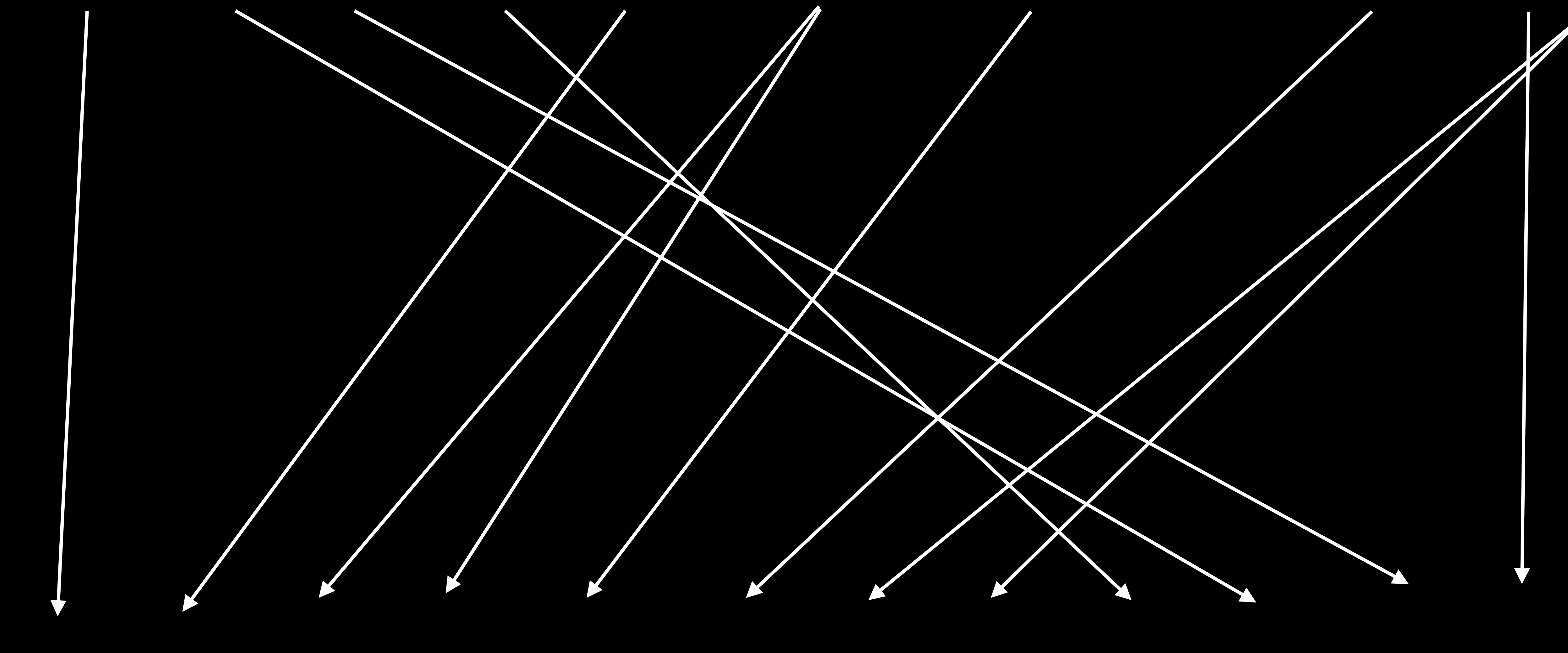
___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

However , the

IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε



However , the sky remained clear under the strong north wind .

IBM Model 1

$$p(despite|\text{虽然})$$

$$p(however|\text{虽然})$$

$$p(although|\text{虽然})$$

$$p(northern|\text{北})$$

$$p(north|\text{北})$$

IBM Model 1

$p(despite|$ 虽然 $)$???

$p(however|$ 虽然 $)$???

$p(although|$ 虽然 $)$???

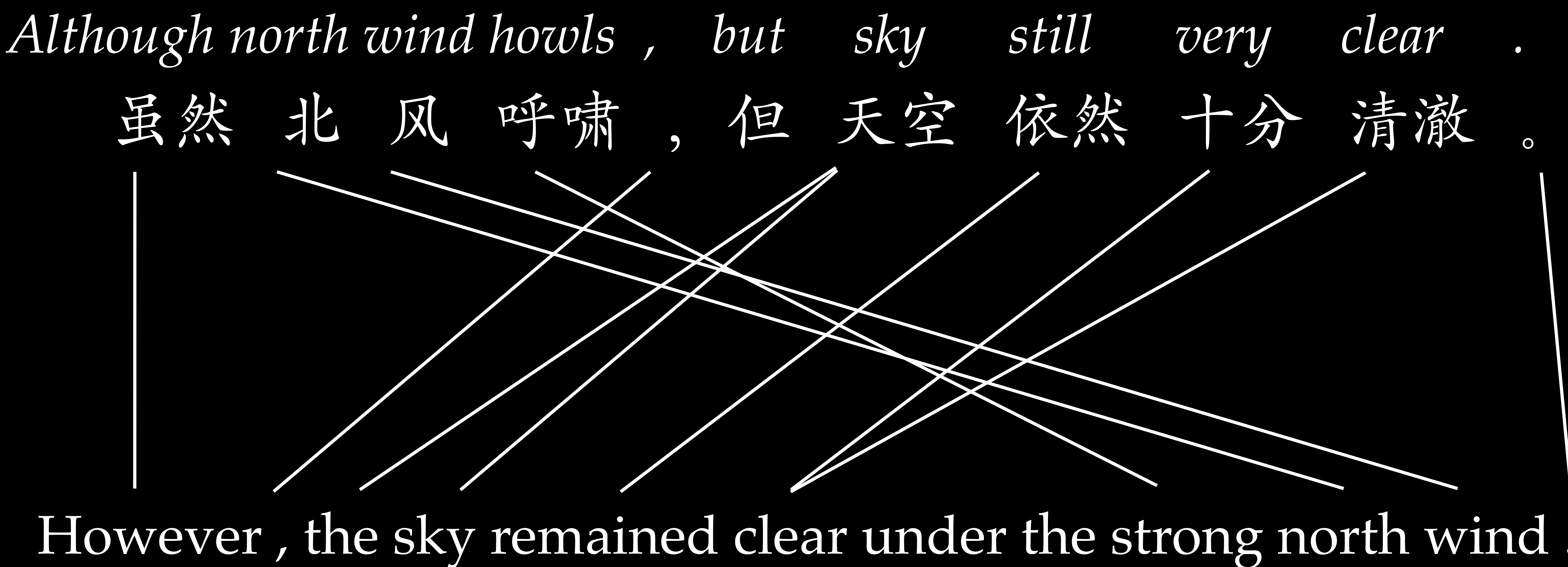
$p(northern|$ 北 $)$???

$p(north|$ 北 $)$???

IBM Model 1

$$\theta \left\{ \begin{array}{ll} p(despite|\text{虽然}) & ??? \\ p(however|\text{虽然}) & ??? \\ p(although|\text{虽然}) & ??? \\ & \\ p(northern|\text{北}) & ??? \\ p(north|\text{北}) & ??? \end{array} \right.$$

IBM Model 1



$$p(north|\ 北\ ) = ???$$

Suppose that we only ever see 北 aligned to north or northern.

北 → north

北 → northern

Suppose that we only ever see 北 aligned to north or northern.

北 → north

北 → northern

$p(north|\ 北\ )$

$p(northern|\ 北\ )$

Suppose that we only ever see 北 aligned to north or northern.

北 → north

北 → northern

$$p(north|\ 北\ ) \qquad 1 - p(north|\ 北\ )$$

北	→	north	北	→	north
北	→	northern	北	→	north
北	→	north	北	→	north
北	→	northern	北	→	north
北	→	north	北	→	northern

北 → north	北 → north
北 → northern	北 → north
北 → north	北 → north
北 → northern	北 → north
北 → north	北 → northern

$$p(north|\text{北})?$$

北 → north
北 → northern
北 → north
北 → northern
北 → north

北 → north
北 → north
北 → north
北 → north
北 → northern

$$p(north|\text{ 北 })?$$

$$p(data) = p(north|\text{ 北 })^7 + p(northern|\text{ 北 })^3$$

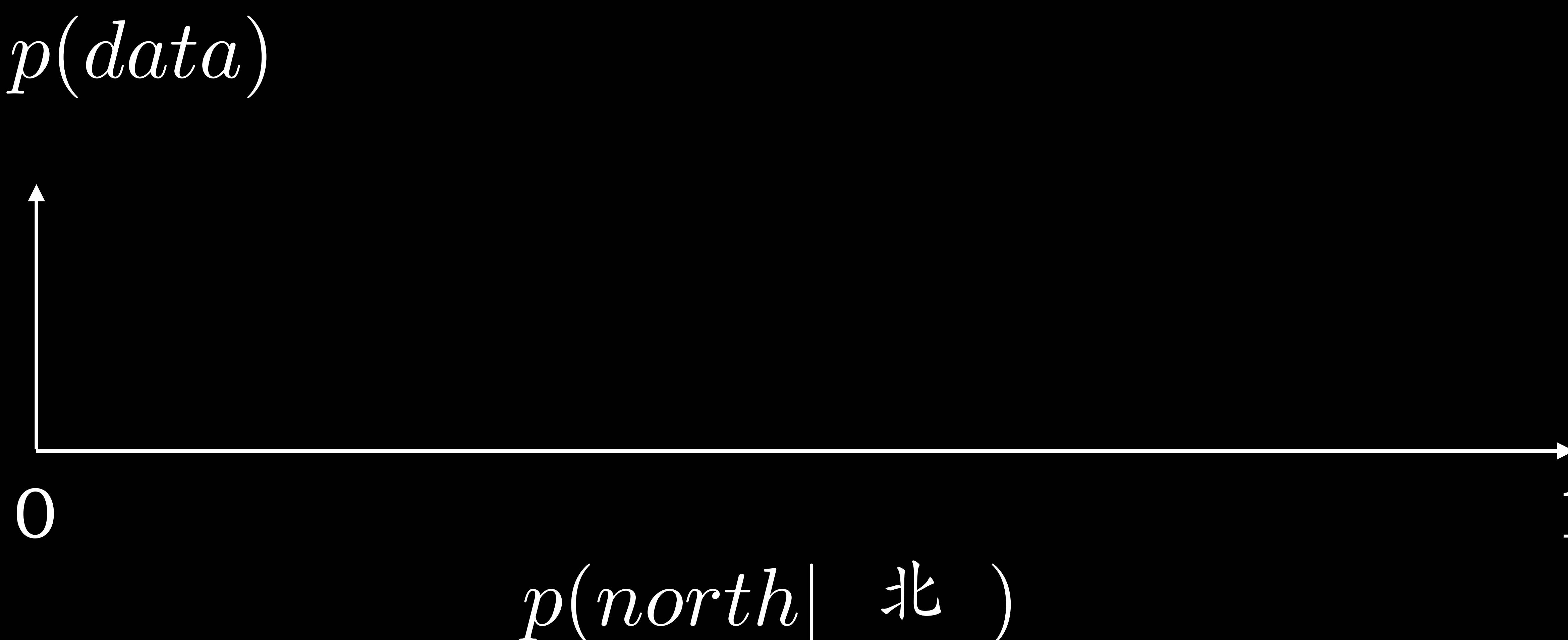
北 → north
北 → northern
北 → north
北 → northern
北 → north

北 → north
北 → north
北 → north
北 → north
北 → northern

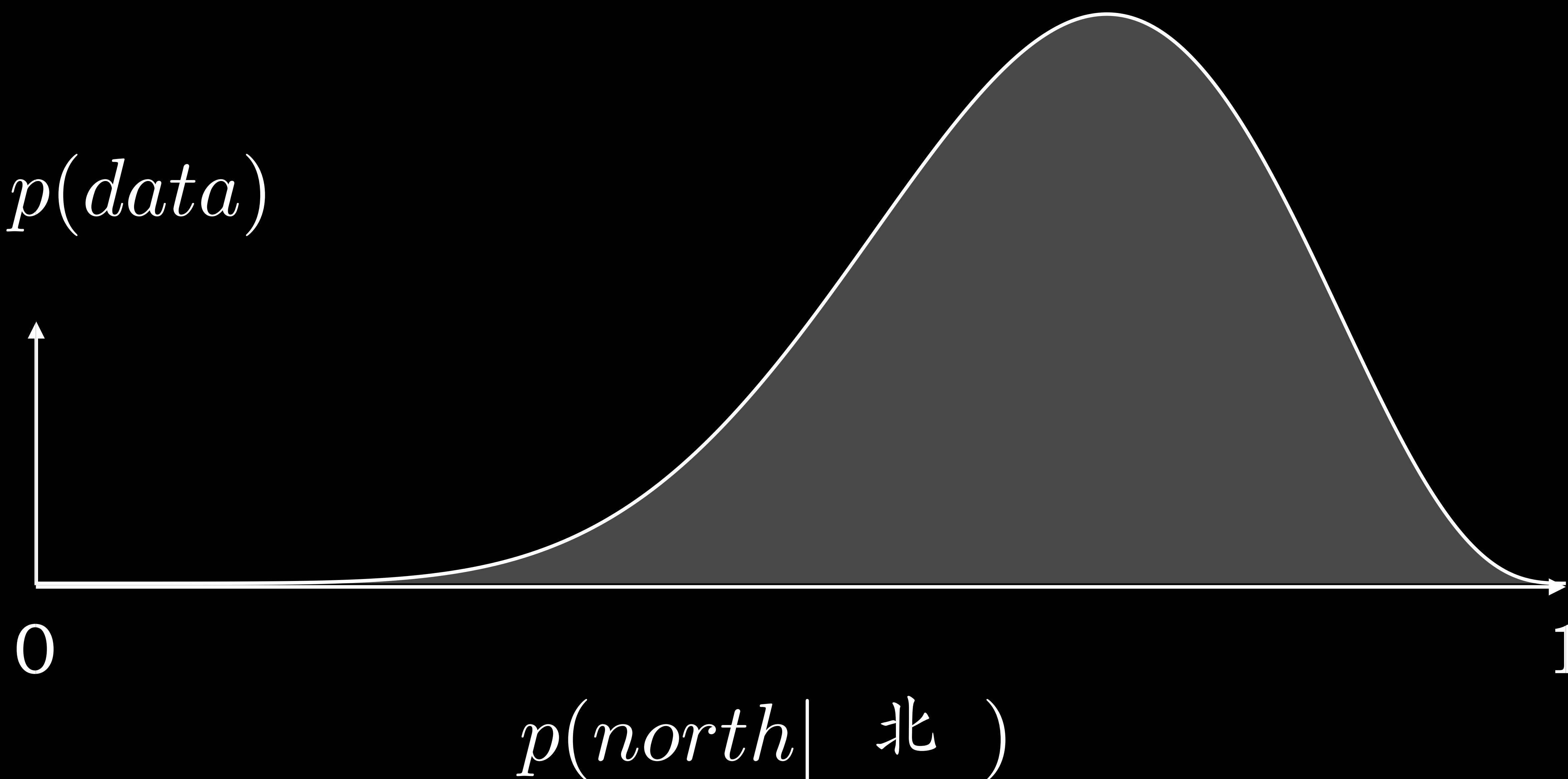
$$p(north|\text{北})?$$

$$p(data) = p(north|\text{北})^7 + [1 - p(north|\text{北})]^3$$

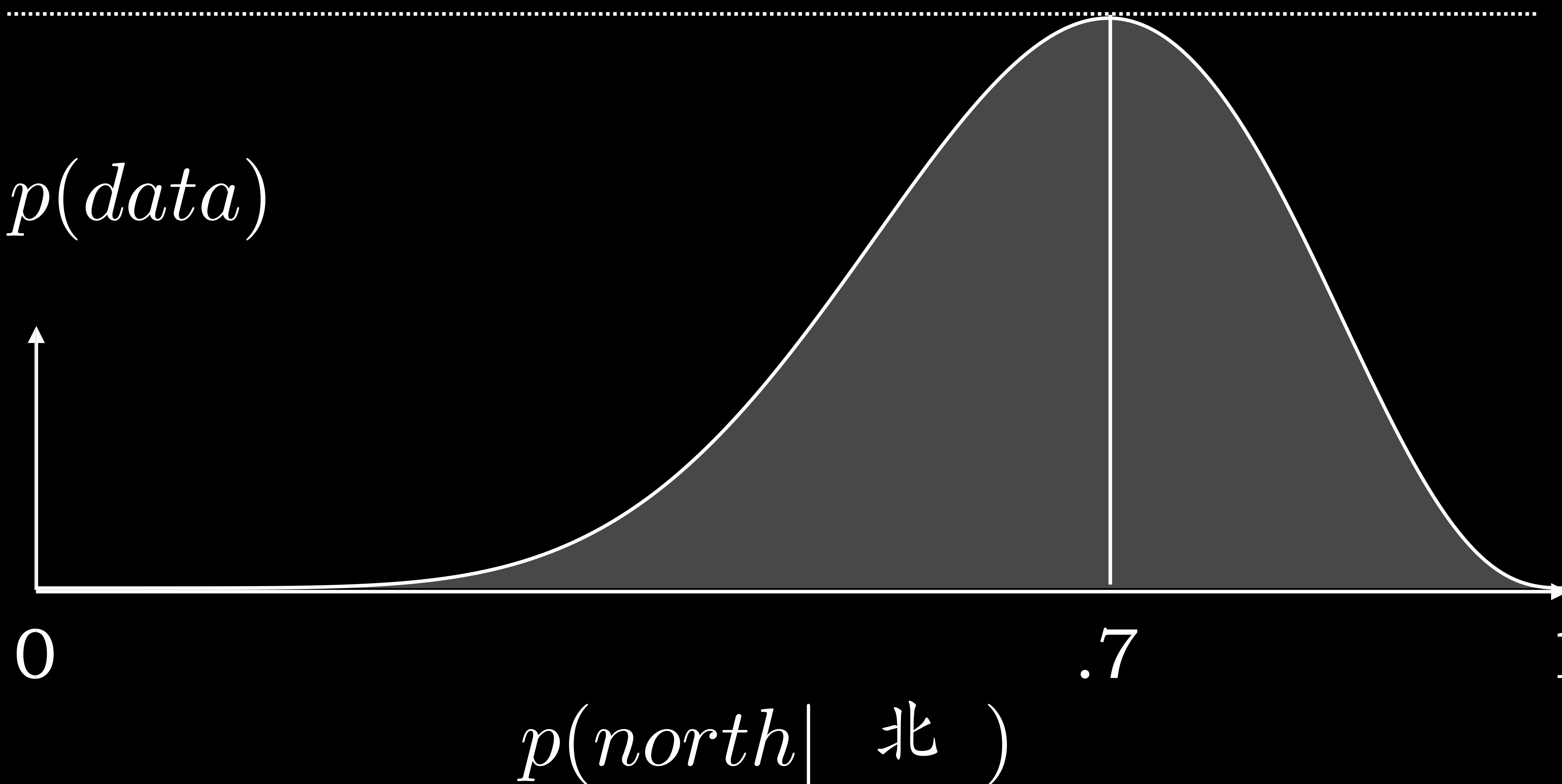
$$p(data) = p(north|\ 北\ )^7 + [1 - p(north|\ 北\ )]^3$$



$$p(data) = p(north|\text{北})^7 + [1 - p(north|\text{北})]^3$$



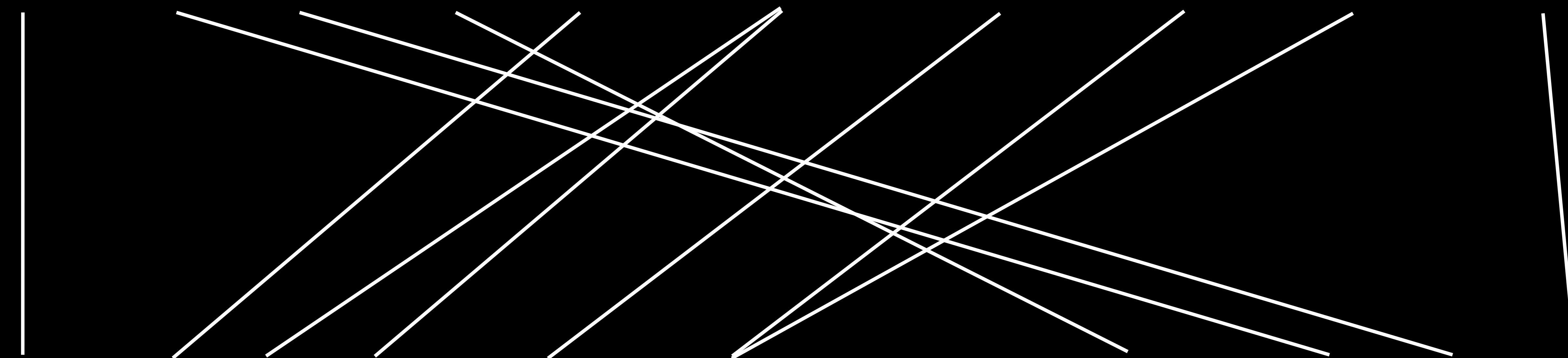
$$p(data) = p(north|\text{北})^7 + [1 - p(north|\text{北})]^3$$



IBM Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。



However , the sky remained clear under the strong north wind .

$$p(north|\text{ 北 }) = \frac{\text{\# of times 北 aligns to north}}{\text{\# of times 北 occurs}}$$

- Optimization

MLE for IBM Model 1 (observed)

$$\hat{\theta} = \arg\max_{\theta} \prod_{n=1}^{N} \left(p(I^{(n)}|J^{(n)}) \prod_{i=1}^{I^{(n)}} p(a_i^{(n)}|J^{(n)}) \cdot p(f_i^{(n)}|e_{a_i}^{(n)}) \right)$$

MLE for IBM Model 1 (observed)

number of sentences

alignment of French word at position *i*

$$\hat{\theta} = \arg\max_{\theta} \prod_{n=1}^{N} \left(p(I^{(n)}|J^{(n)}) \prod_{i=1}^{I^{(n)}} p(a_i^{(n)}|J^{(n)}) \cdot p(f_i^{(n)}|e_{a_i}^{(n)}) \right)$$

French, English sentence lengths

French, English word pair

MLE for IBM Model 1 (observed)

$$\hat{\theta} = \arg\max_{\theta} \prod_{n=1}^{N} \left(\underbrace{p(I^{(n)} | J^{(n)}) \prod_{i=1}^{I^{(n)}} p(a_i^{(n)} | J^{(n)})}_{\text{constant!}} \cdot p(f_i^{(n)} | e_{a_i}^{(n)}) \right)$$

MLE for IBM Model 1 (observed)

$$\hat{\theta} = \arg\max_{\theta} C \prod_{n=1}^{N} \prod_{i=1}^{I^{(n)}} p(f_i^{(n)} | e_{a_i}^{(n)})$$

MLE for IBM Model 1 (observed)

$$\hat{\theta} = \arg\max_{\theta} \log \left(C \prod_{n=1}^{N} \prod_{i=1}^{I^{(n)}} p(f_i^{(n)} | e_{a_i}^{(n)}) \right)$$

$$\log(a) < \log(b) \iff a < b$$

MLE for IBM Model 1 (observed)

$$\hat{\theta} = \arg\max_{\theta} \log \left(C \cdot \prod_{f,e} p(f|e)^{count(\langle f,e \rangle)} \right)$$

MLE for IBM Model 1 (observed)

$$\hat{\theta} = \arg\max_{\theta} \log C + \sum_{f,e} count(\langle f, e \rangle) \log p(f|e)$$

log of product = sum of logs

MLE for IBM Model 1 (observed)

$$\Lambda(\theta, \lambda) = \log C + \sum_{f,e} count(\langle f, e \rangle) \log p(f|e) \underbrace{- \sum_{e} \lambda_e \left(\sum_{f} p(f|e) - 1 \right)}_{\text{Lagrange multiplier expresses normalization constraint}}$$

MLE for IBM Model 1 (observed)

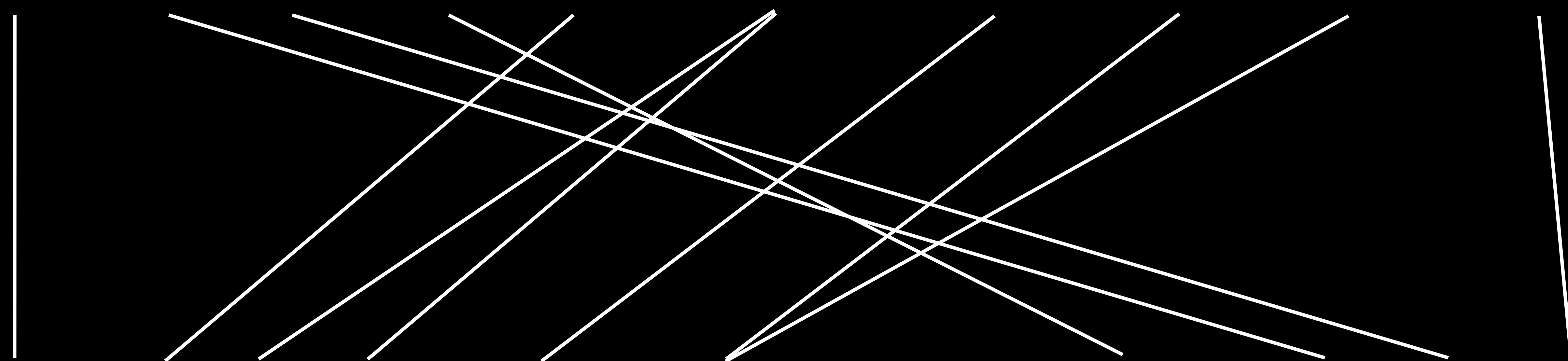
$$\Lambda(\theta, \lambda) = \log C + \sum_{f,e} count(\langle f, e \rangle) \log p(f|e) - \sum_{e} \lambda_e \left(\sum_{f} p(f|e) - 1 \right)$$

derivative $$\frac{\partial \Lambda(\theta, \lambda)}{\partial p(f|e)} = \frac{count(\langle f, e \rangle)}{p(f|e)} - \lambda_e$$

MLE for IBM Model 1 (observed)

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。



However , the sky remained clear under the strong north wind .

$$p(north|\text{北}) = \frac{\text{\# of times 北 aligns to north}}{\text{\# of times 北 occurs}}$$

MLE for IBM Model 1 (unobserved)

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。

However , the sky remained clear under the strong north wind .

$p(north|$ 北 $)$ = ???

MLE for IBM Model 1 (observed)

$$\hat{\theta} = \arg\max_{\theta} \log \left(C \prod_{n=1}^{N} \prod_{i=1}^{I^{(n)}} p(f_i^{(n)} | e_{a_i}^{(n)}) \right)$$

MLE for IBM Model 1 (unobserved)

$$\hat{\theta} = \arg\max_{\theta} \log \left(C \prod_{n=1}^{N} \sum_{a} \prod_{i=1}^{I^{(n)}} p(f_i^{(n)} | e_{a_i}^{(n)}) \right)$$

marginalize over alignments:

$$p(f|e) = \sum_{a} p(f, a|e)$$

MLE for IBM Model 1 (unobserved)

$$\hat{\theta} = \arg\max_{\theta} \log \left(C \cdot \prod_{f,e} p(f|e)^{\mathbb{E}[count(\langle f,e \rangle)]} \right)$$

MLE for IBM Model 1 (unobserved)

$$\hat{\theta} = \arg\max_{\theta} \log \left(C \cdot \prod_{f,e} p(f|e)^{\mathbb{E}[count(\langle f,e \rangle)]} \right)$$

Not constant! Depends on parameters, no analytic solution.

MLE for IBM Model 1 (unobserved)

$$\hat{\theta} = \arg\max_{\theta} \log \left(C \cdot \prod_{f,e} p(f|e)^{\mathbb{E}[count(\langle f,e \rangle)]} \right)$$

Not constant! Depends on parameters, no analytic solution.

But it does strongly imply an iterative solution.

Likelihood Estimation for Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

Parameters and alignments are both unknown.

However , the sky remained clear under the strong north wind .

$p(English\ word|Chinese\ word)$ unobserved!

Likelihood Estimation for Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

Parameters and alignments are both unknown.

If we knew the alignments, we could calculate the values of the parameters.

However , the sky remained clear under the strong north wind .

$p(English\ word|Chinese\ word)$ unobserved!

Likelihood Estimation for Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。 ε

Parameters and alignments are both unknown.

If we knew the alignments, we could calculate the values of the parameters.

If we knew the parameters, we could calculate the likelihood of the data.

However , the sky remained clear under the strong north wind .

$p(English\ word|Chinese\ word)$ unobserved!

Likelihood Estimation for Model 1

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。 ε

Parameters and alignments are both unknown.

If we knew the alignments, we could calculate the values of the parameters.

If we knew the parameters, we could calculate the likelihood of the data.

However , the sky remained clear under the strong north wind .

$p(English\ word|Chinese\ word)$ unobserved!

The Plan: Bootstrapping

- Arbitrarily select a set of parameters (say, uniform).
- Calculate ***expected counts*** of the unseen events.
- Choose new parameters to maximize likelihood, using expected counts as proxy for observed counts.
- Iterate.
- Guarantee: likelihood will be monotonically nondecreasing.

The Plan: Bootstrapping

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

However , the sky remained clear under the strong north wind .

The Plan: Bootstrapping

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

if we had observed the alignment, this line would either be here (count 1) or it wouldn't (count 0).

However , the sky remained clear under the strong north wind .

The Plan: Bootstrapping

Although north wind howls , but sky still very clear .

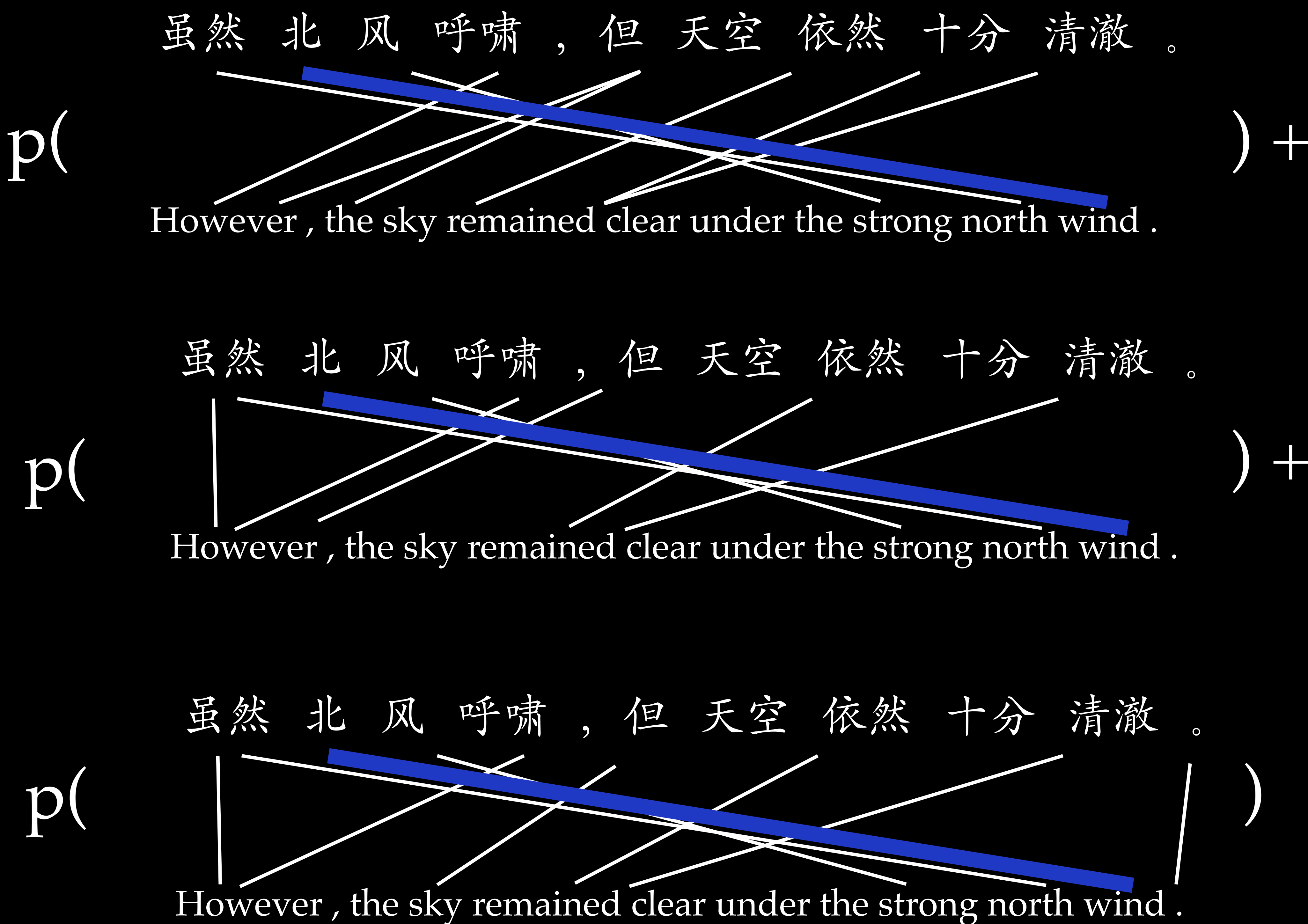
虽然 北 风 呼啸 ，但 天空 依然 十分 清澈 。ε

if we had observed the alignment, this line would either be here (count 1) or it wouldn't (count 0).

since we didn't observe the alignment, we calculate the probability that it's there.

However , the sky remained clear under the strong north wind .

Marginalize: sum all alignments containing the link



Divide by sum of all *possible* alignments

p(虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。

However , the sky remained clear under the strong north wind .) +

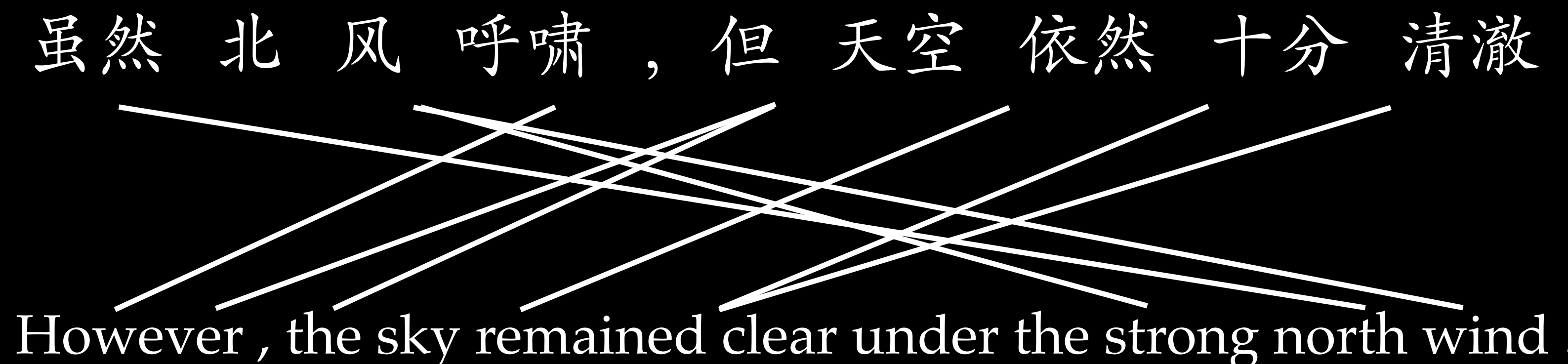
p(虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。

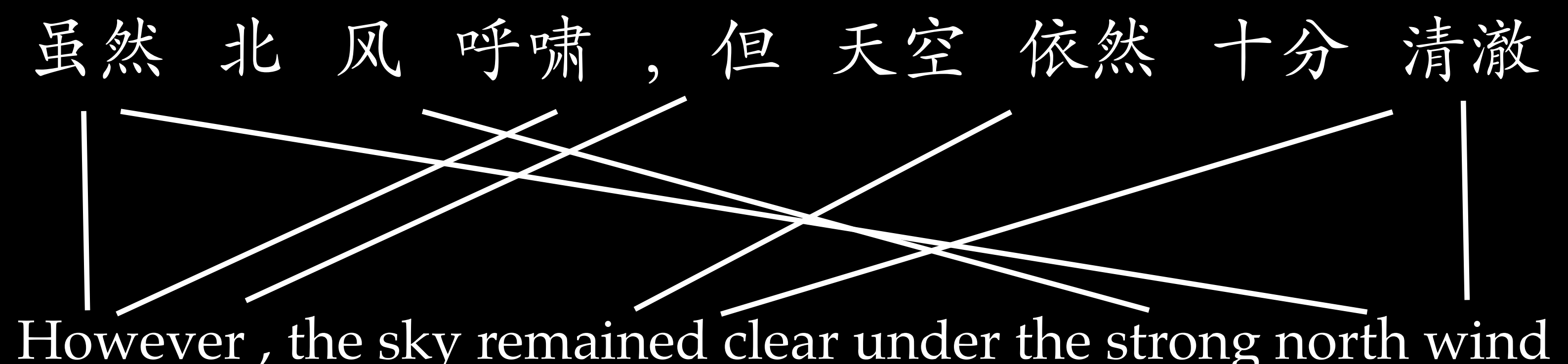
However , the sky remained clear under the strong north wind .) +

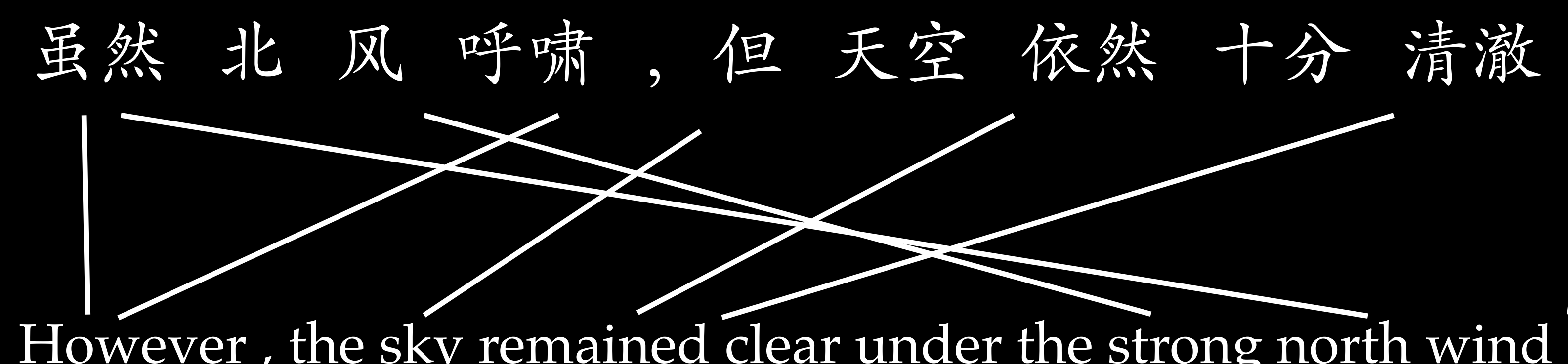
p(虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。

However , the sky remained clear under the strong north wind .)

Divide by sum of all *possible* alignments

p(
) +

p(
) +

p(
)

Is this hard? How many alignments are there?

Expectation Maximization

probability of an alignment.

$$p(F, A|E) = p(I|J) \prod_{a_i} p(a_i = j) p(f_i|e_j)$$

Expectation Maximization

probability of an alignment.

$$p(F, A|E) = p(I|J) \prod_{a_i} p(a_i = j) p(f_i|e_j)$$

observed

uniform

Expectation Maximization

probability of an alignment.

factors across words.

$$p(F, A|E) = p(I|J) \prod_{a_i} p(a_i = j) \boxed{p(f_i|e_j)}$$

observed

uniform

Expectation Maximization

marginal probability of
alignments containing link

$$\sum_{a \in A : 北 \leftrightarrow north} p(north|北) \cdot p(rest\ of\ a)$$

Expectation Maximization

marginal probability of
alignments containing link

$$p(north|北) \sum_{a \in A: 北 \leftrightarrow north} p(rest\ of\ a)$$

Expectation Maximization

marginal probability of
alignments containing link

$$\frac{p(north|北) \sum_{a \in A: 北 \leftrightarrow north} p(rest\ of\ a)}{\sum_{c \in Chinese\ words} p(north|c) \sum_{a \in A: \mathrm{c} \leftrightarrow north} p(rest\ of\ a)}$$

marginal probability of all
alignments

Expectation Maximization

marginal probability of
alignments containing link

$$\frac{p(north|北) \sum_{a \in A: 北 \leftrightarrow north} p(rest\ of\ a)}{\sum_{c \in Chinese\ words} p(north|c) \sum_{a \in A: c \leftrightarrow north} p(rest\ of\ a)}$$

marginal probability of all
alignments

Expectation Maximization

marginal probability of
alignments containing link

$$\frac{p(north|北) \sum_{a \in A: 北 \leftrightarrow north} p(rest\ of\ a)}{\sum_{c \in Chinese\ words} p(north|c) \sum_{a \in A: c \leftrightarrow north} p(rest\ of\ a)}$$

identical!

marginal probability of all
alignments

Expectation Maximization

$$\frac{p(north|\text{北})}{\sum_{c \in Chinese\ words} p(north|c)}$$

Expectation Maximization

marginal probability (expected count) of an alignment containing the link

$$\frac{p(north|\text{北})}{\sum_{c \in Chinese\ words} p(north|c)}$$

Expectation Maximization

marginal probability (expected count) of an alignment containing the link

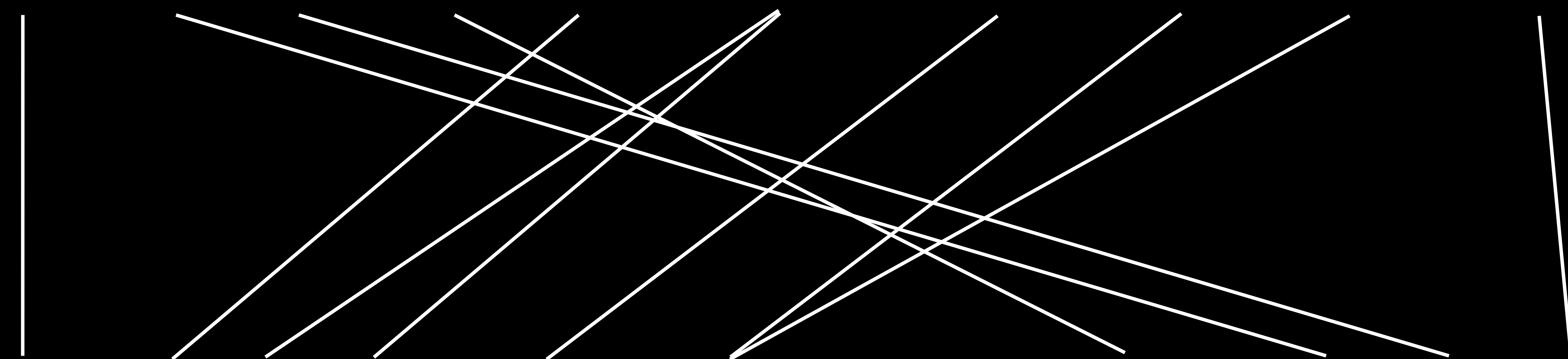
$$\frac{p(north|\text{北})}{\sum_{c \in Chinese\ words} p(north|c)}$$

For each sentence, use this quantity instead of 0 or 1

Translation Models

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。



However , the sky remained clear under the strong north wind .

$$p(however \mid \text{虽然}) = \frac{\text{\# of times 虽然 aligns to However}}{\text{\# of times 虽然 occurs}}$$

Translation Models

Although north wind howls , but sky still very clear .

虽然 北 风 呼啸 ， 但 天空 依然 十分 清澈 。

However , the sky remained clear under the strong north wind .

$$p(however \mid \text{虽然}) = \frac{\textit{Expected}\text{ \# of times 虽然 aligns to However}}{\text{\# of times 虽然 occurs}}$$

Expectation Maximization

Why does this even work?

$$\frac{p(north|\text{北})}{\sum_{c \in Chinese\ words} p(north|c)}$$

Expectation Maximization

Observation 1: We are still solving a maximum likelihood estimation problem.

Expectation Maximization

Observation 1: We are still solving a maximum likelihood estimation problem.

$$p(Chinese|English) = \sum_{alignments} p(Chinese, alignment|English)$$

Expectation Maximization

Observation 1: We are still solving a maximum likelihood estimation problem.

$$p(Chinese|English) = \sum_{alignments} p(Chinese, alignment|English)$$

MLE: choose parameters that maximize this expression.

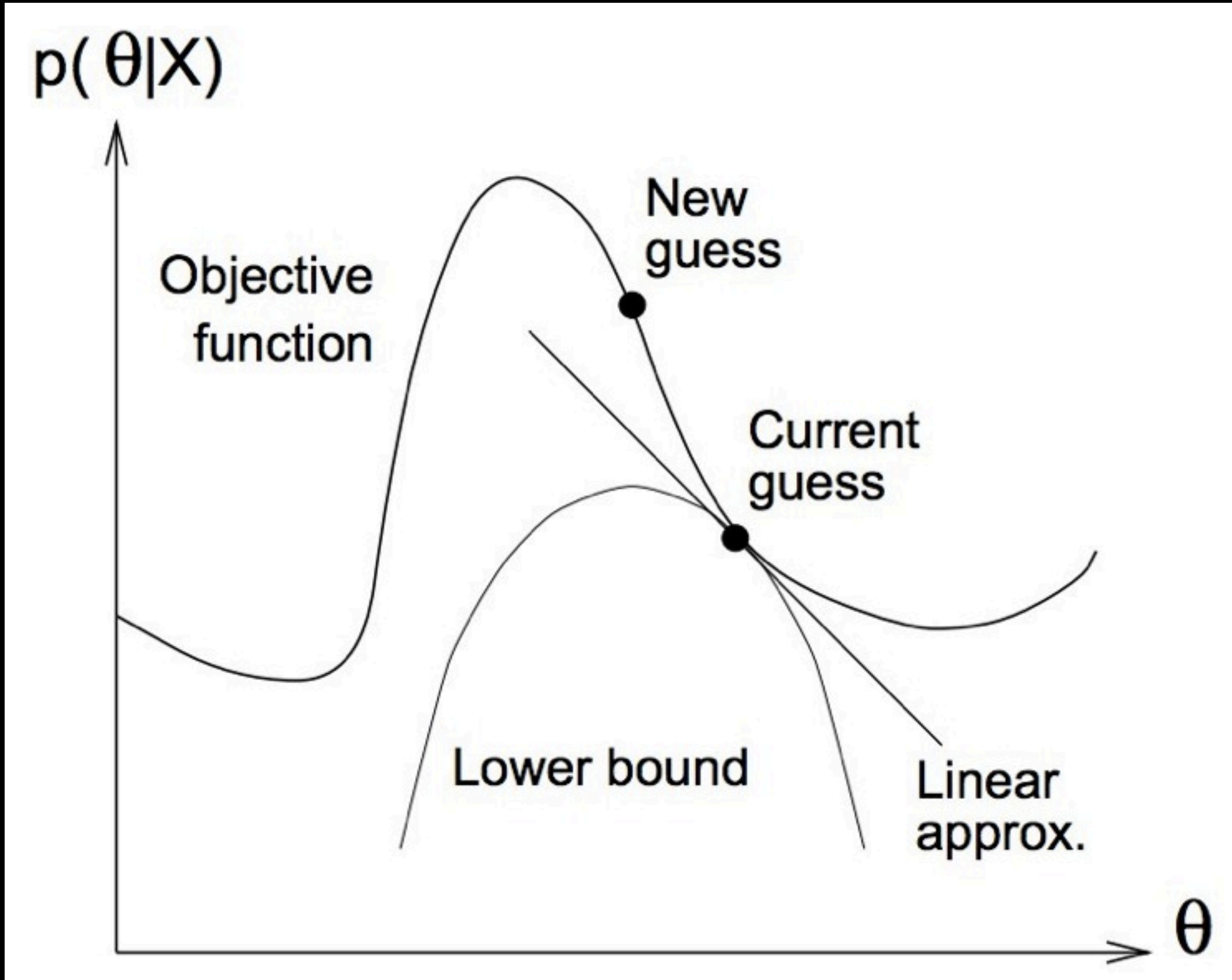
Expectation Maximization

Observation 1: We are still solving a maximum likelihood estimation problem.

$$p(Chinese|English) = \sum_{alignments} p(Chinese, alignment|English)$$

MLE: choose parameters that maximize this expression.

Minor problem: there is no analytic solution.



(from Minka ’98)

Decoding

Probability models enable us to ***make predictions***:
Given a particular Chinese sentence, what is the most probable English sentence corresponding to it?

Decoding

Probability models enable us to ***make predictions***:
Given a particular Chinese sentence, what is the most probable English sentence corresponding to it?

In math, we want to solve:

$$\operatorname{argmax}_{English}\, p(English|Chinese)$$

Decoding

Probability models enable us to ***make predictions***:
Given a particular Chinese sentence, what is the most probable English sentence corresponding to it?

In math, we want to solve:

$$\text{argmax}_{English}\, p(English|Chinese)$$

problem: there are a lot of English sentences to choose from!


THE WØRD

- Optimization

北 风 呼啸 。

北 风 呼啸 。

substitutions
permutations

北 风 呼啸 。

substitutions $O(5^n)$

permutations

北 风 呼啸 。

substitutions	$O(5^n)$
permutations	$O(n!)$

北 风 呼啸 。

substitutions $O(5^n)$

permutations $O(n!)$

15,000 possibilities!

北 风 呼啸 。

Can we do this without enumerating $O(5^n n!)$ pairs?

北 风 呼啸 。

the strong north wind .

Can we do this without enumerating $O(5^n n!)$ pairs?

北 风 呼啸 。

the strong north wind .

Can we do this without enumerating $O(5^n n!)$ pairs?

北 风 呼啸 。

the strong north wind .

Given a sentence pair and an alignment, we can easily calculate

$p(English, alignment|Chinese)$

Can we do this without enumerating $O(5^n n!)$ pairs?

Key Idea

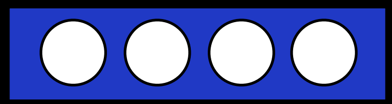
There are $O(5^n n!)$ target sentences.

But there are only $O(5^n)$ ways to start them.

Key Idea

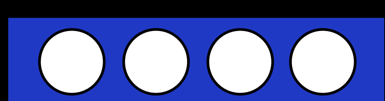
北 风 呼啸 。

Key Idea



北 风 呼啸 。

Key Idea



coverage vector

北 风 呼啸 。

Key Idea

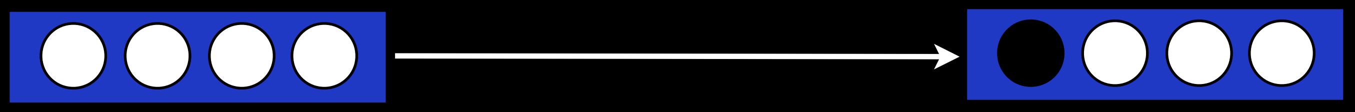
coverage vector

北 风 呼啸 。

Key Idea

$$p(north|START) \cdot p(\text{北}|north)$$

north

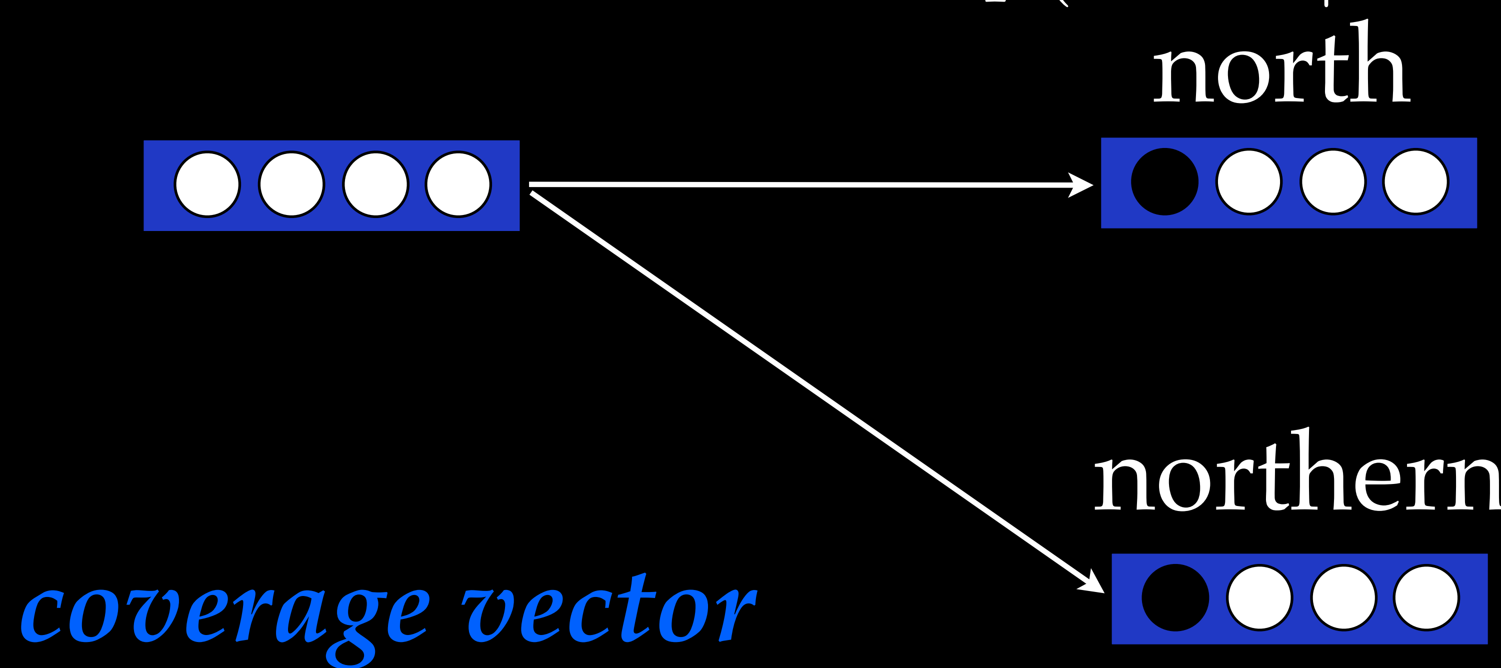


coverage vector

北 风 呼啸 。

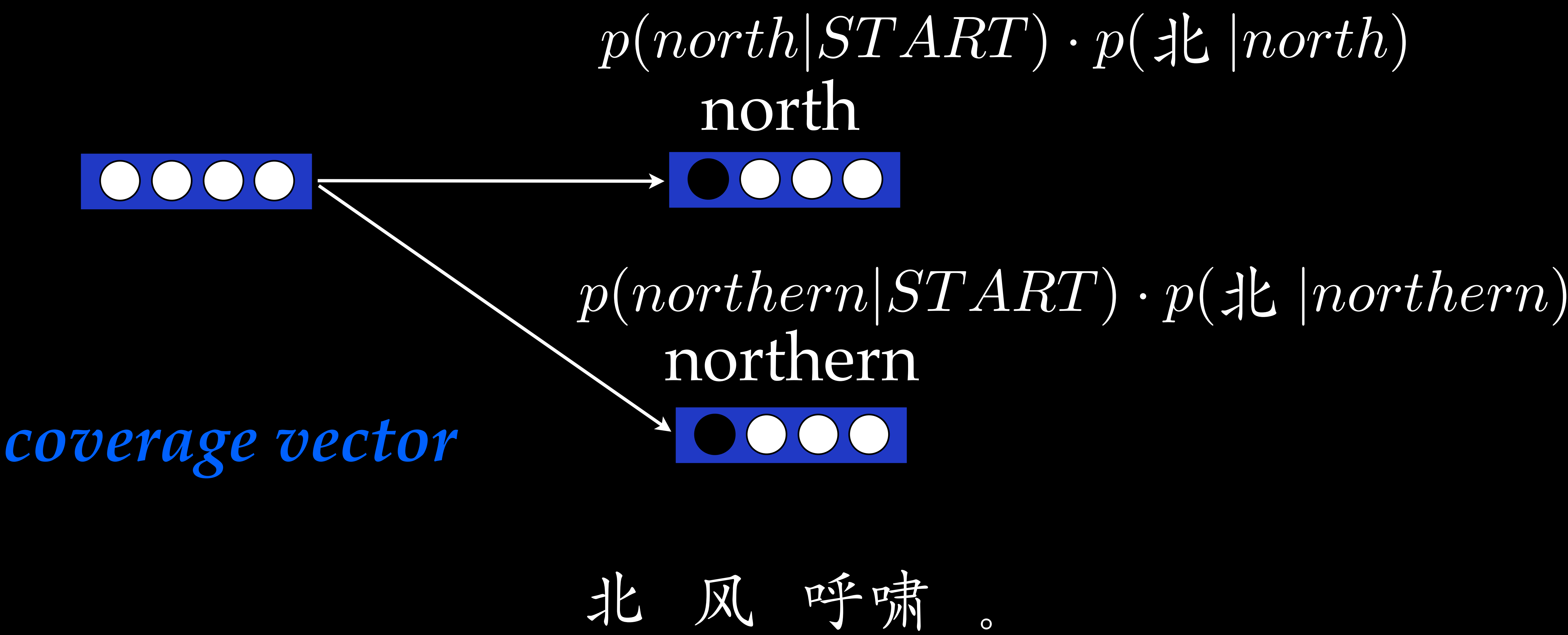
Key Idea

$$p(north|START) \cdot p(北\,|north)$$



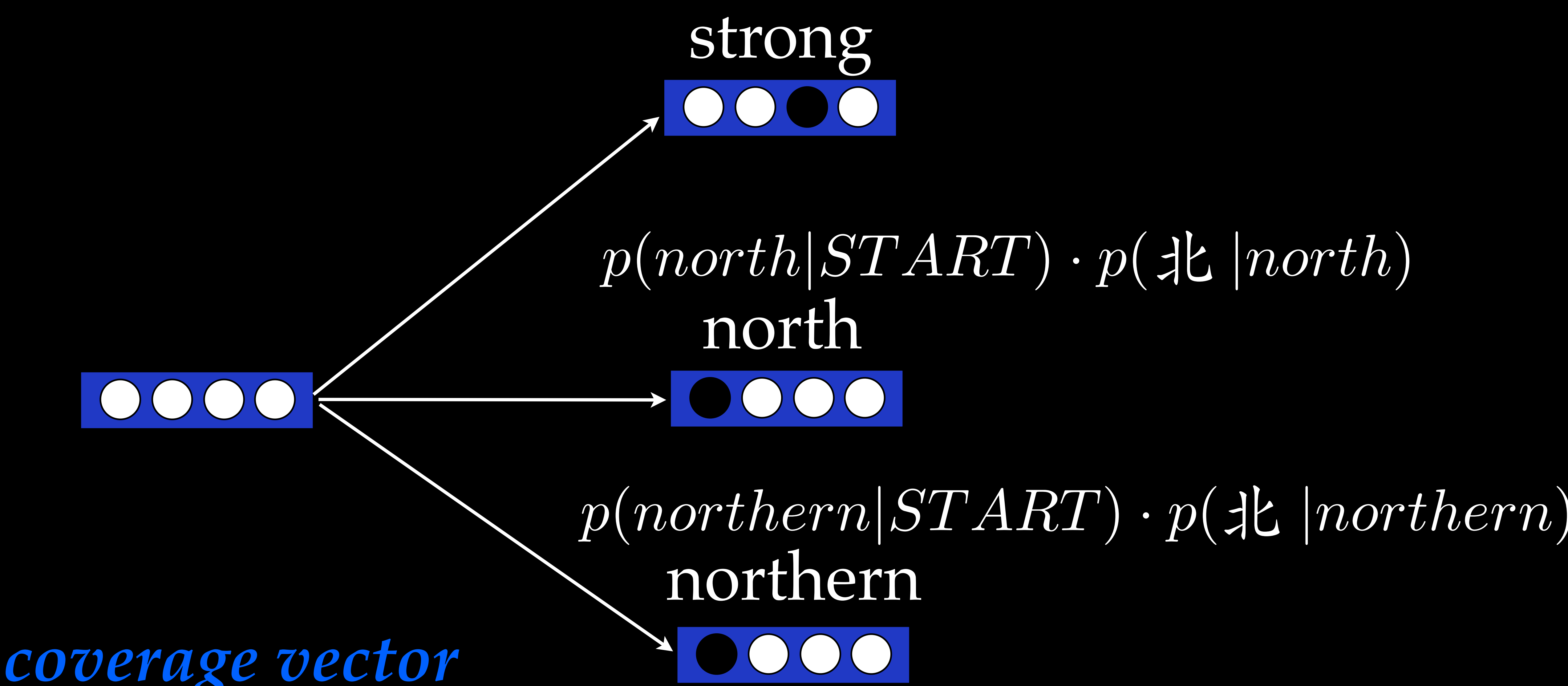
北 风 呼啸 。

Key Idea



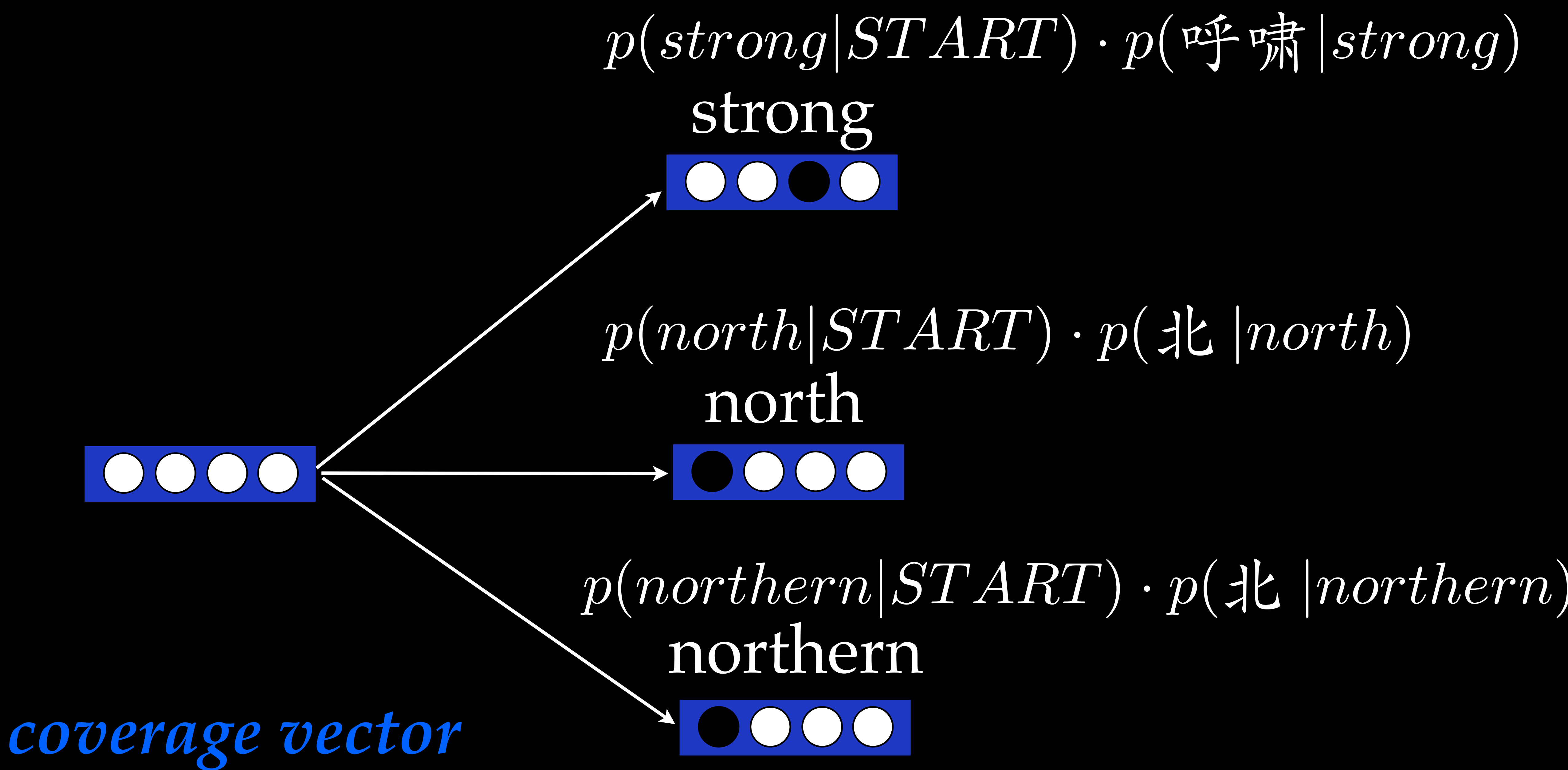
$p(north|START) \cdot p(北|north)$
north
$p(northern|START) \cdot p(北|northern)$
northern
coverage vector
北 风 呼啸 。

Key Idea



北 风 呼啸 。

Key Idea

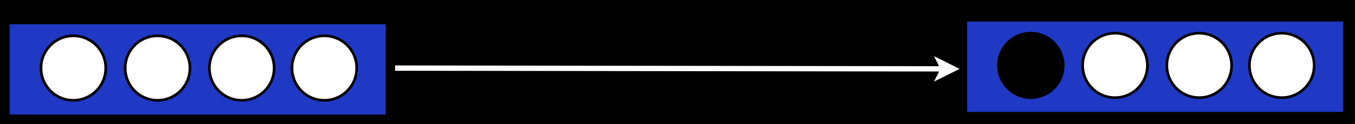


北 风 呼啸 。

Key Idea

$$p(north|START) \cdot p(\text{北}\,|north)$$

north

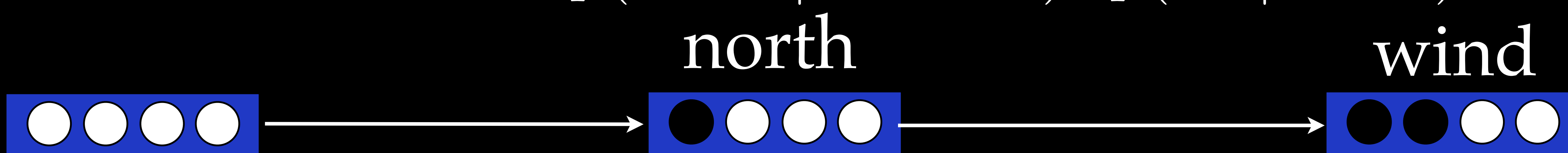


coverage vector

北 风 呼啸 。

Key Idea

$$p(north|START) \cdot p(\text{北}\,|north)$$



coverage vector

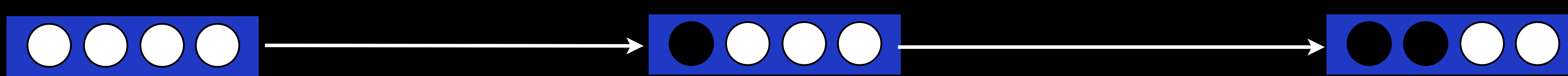
北 风 呼啸 。

Key Idea

$$p(north|START) \cdot p(\text{北}|north)$$

north

wind

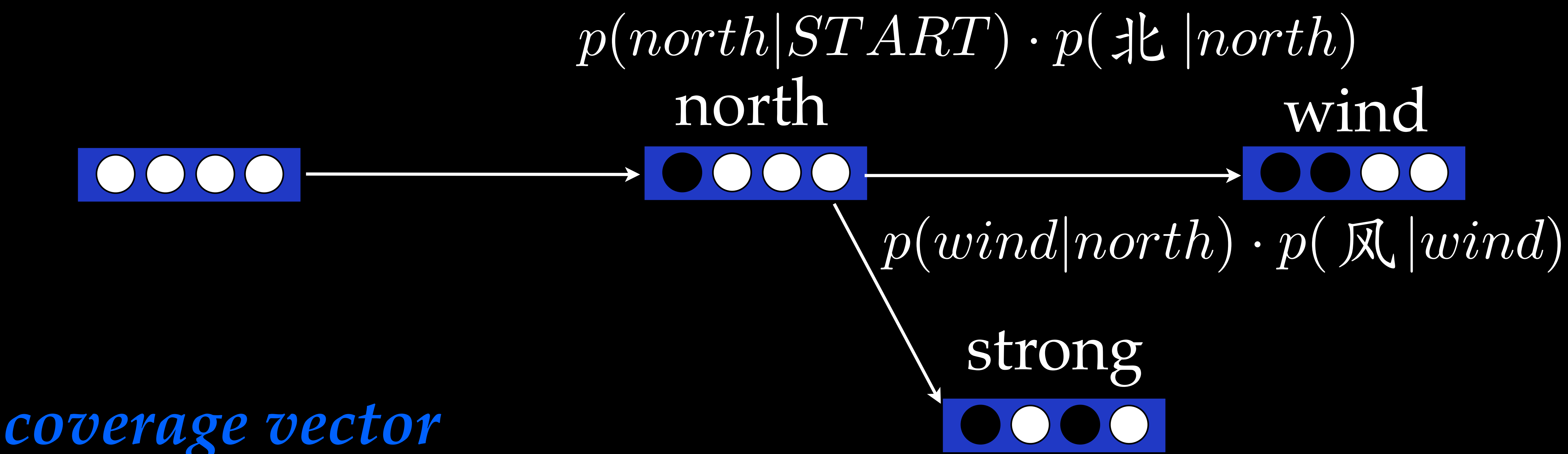


$$p(wind|north) \cdot p(\text{风}|wind)$$

coverage vector

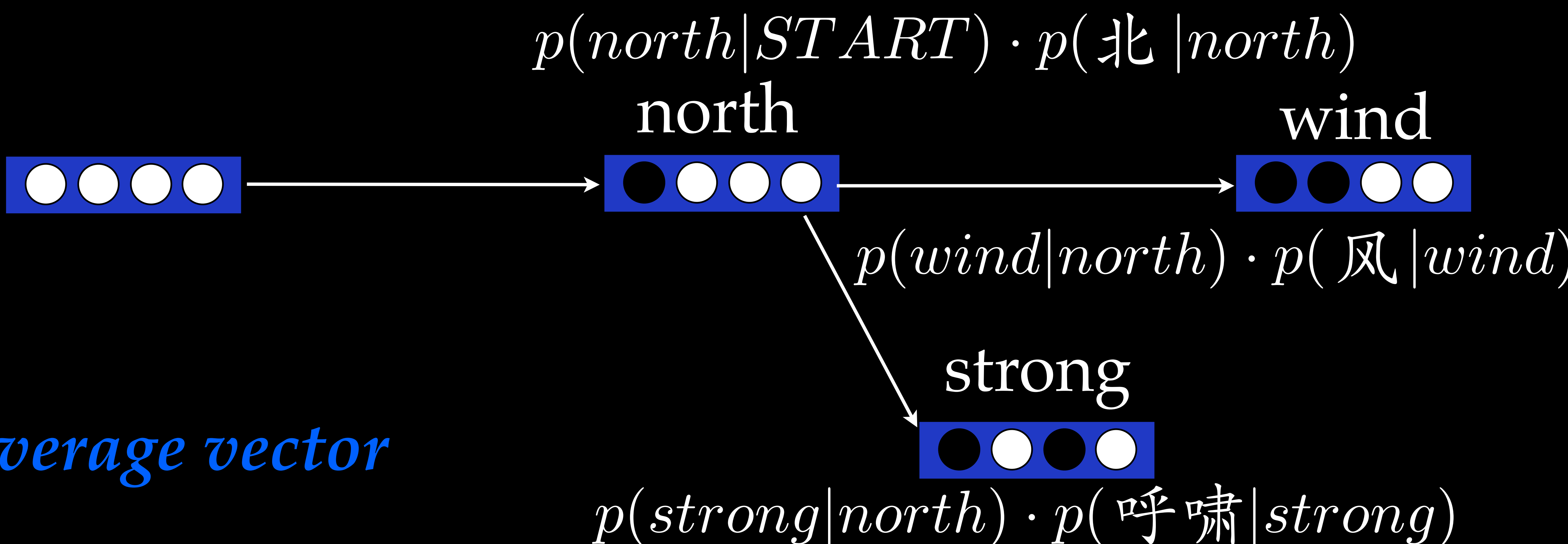
北 风 呼啸 。

Key Idea



北 风 呼啸 。

Key Idea

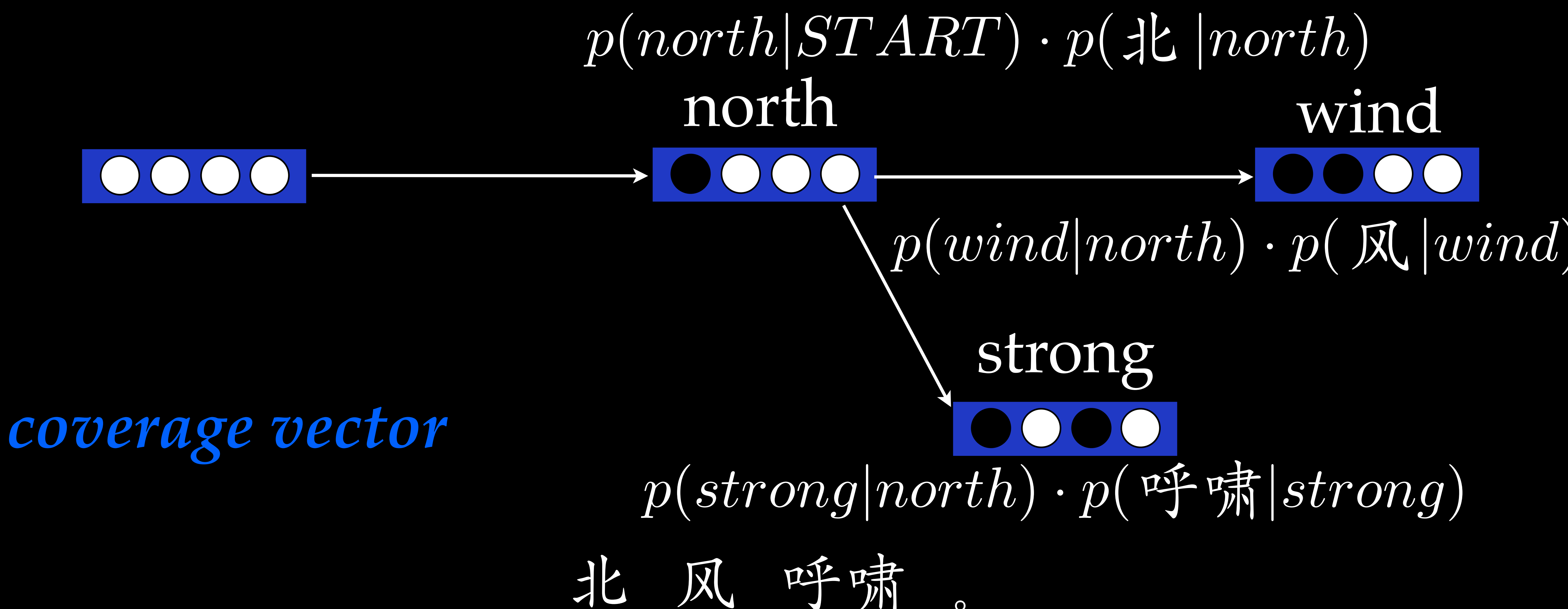


coverage vector

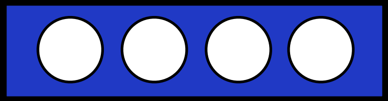
北 风 呼啸 。

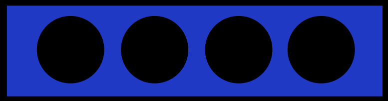
Key Idea

Work done at sentence beginnings is shared across many possible output sentences!

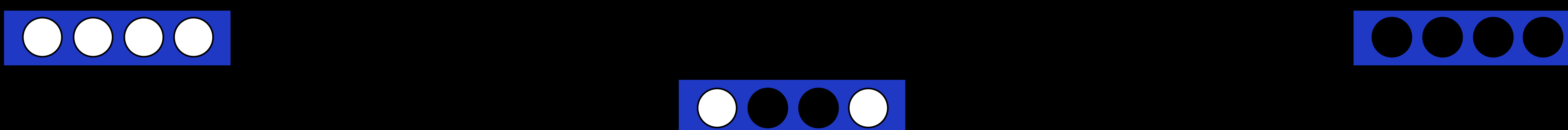


Key Idea

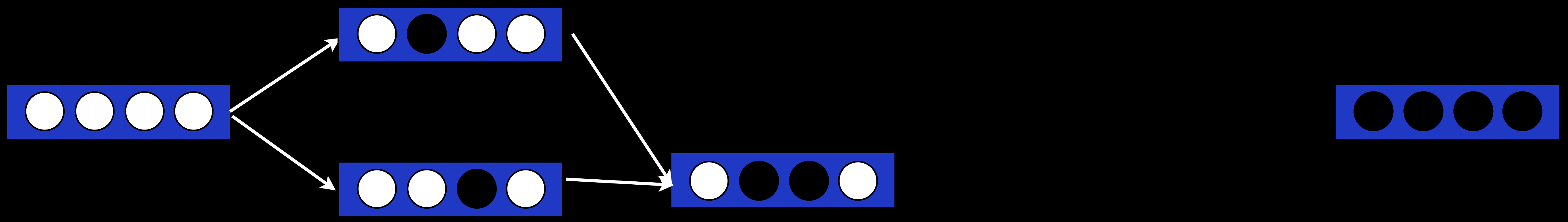




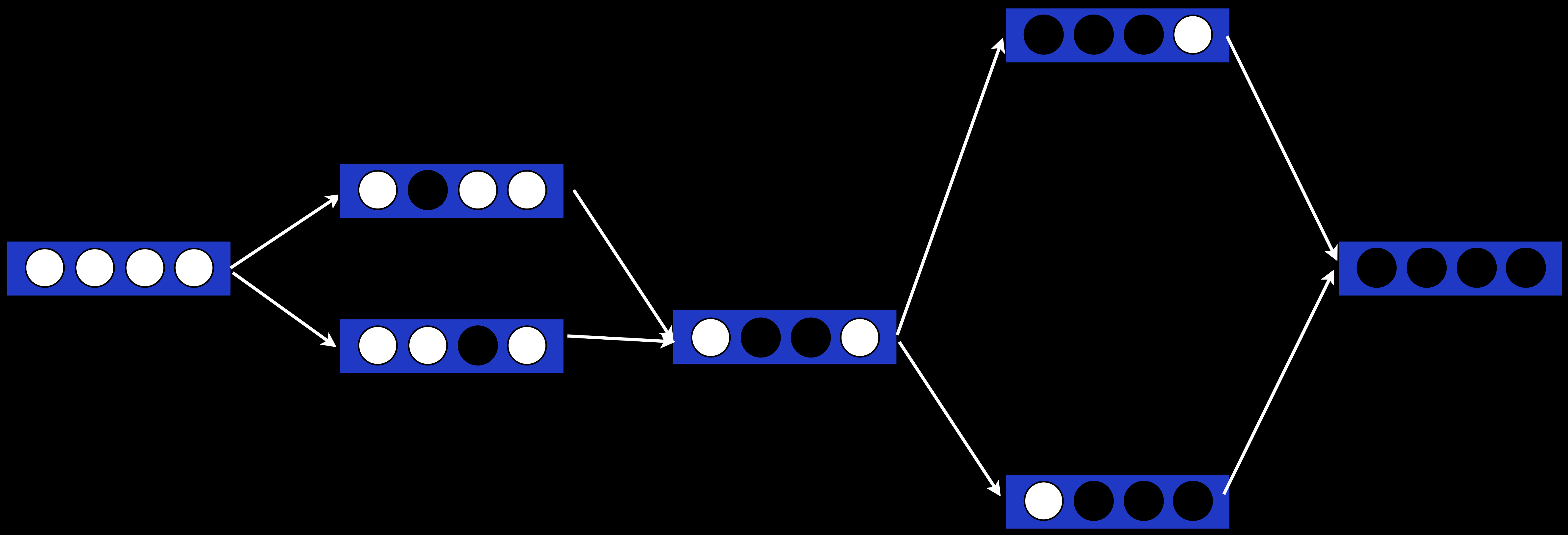
Key Idea



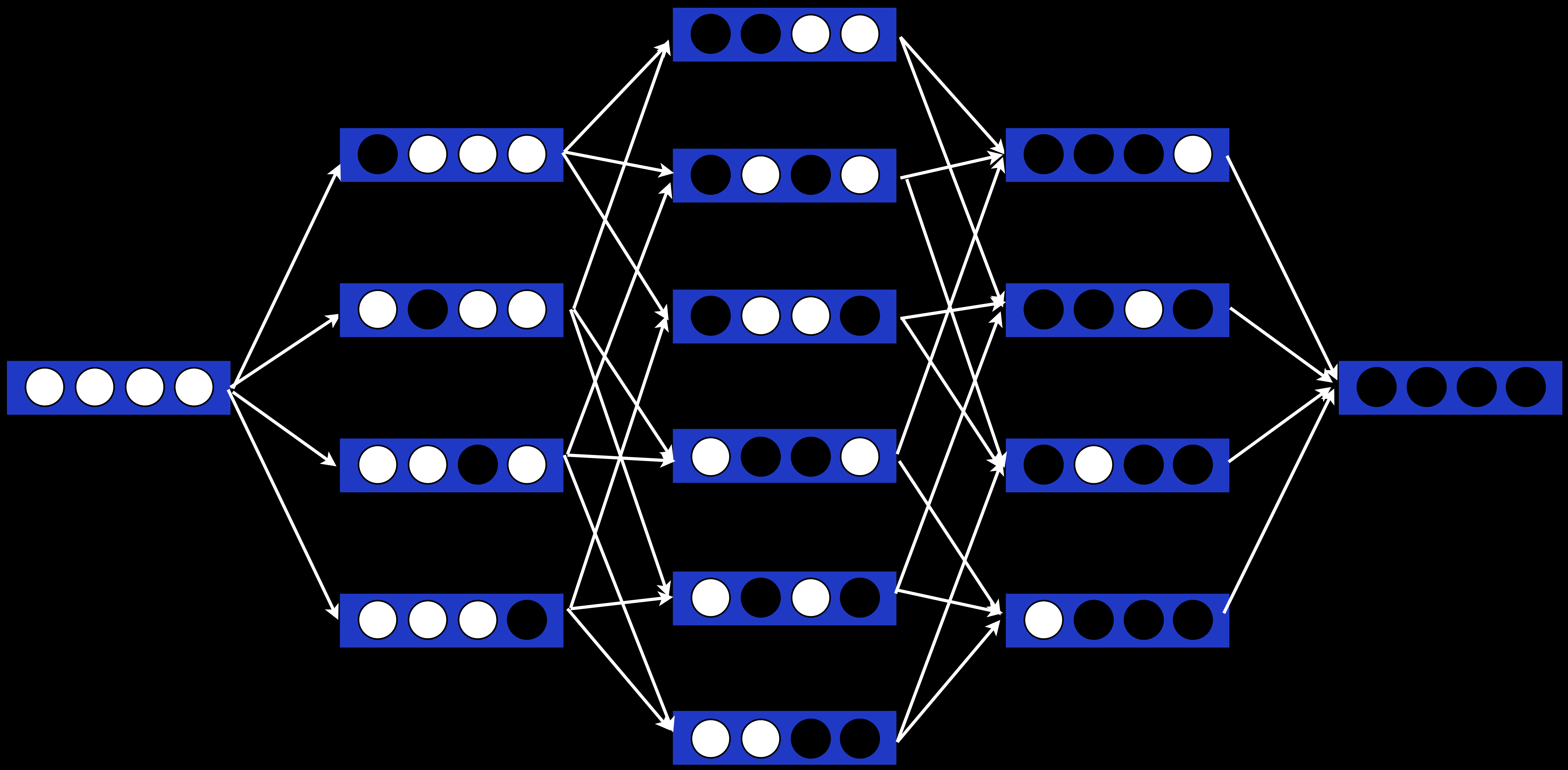
Key Idea



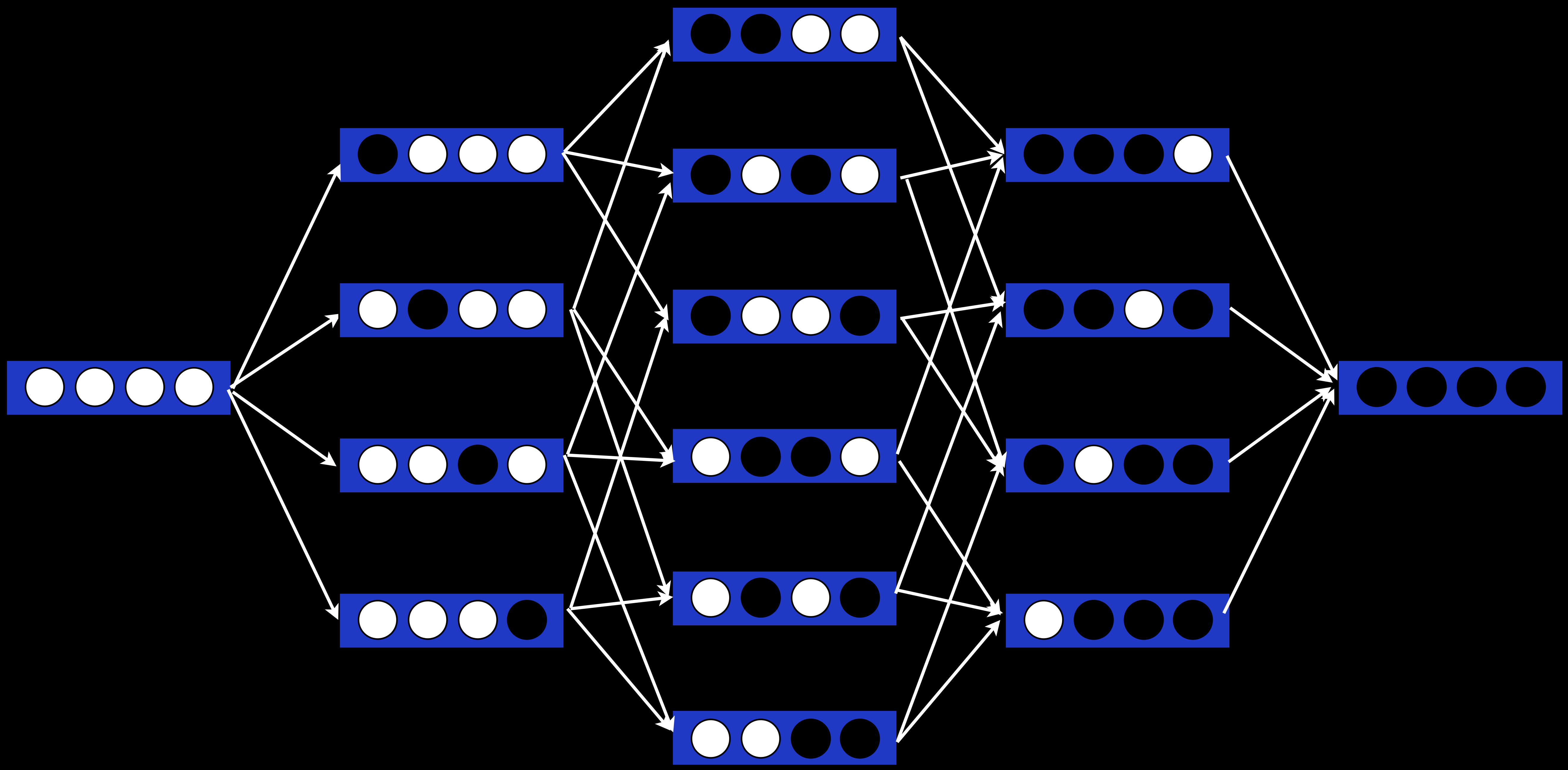
Key Idea



Key Idea



Key Idea

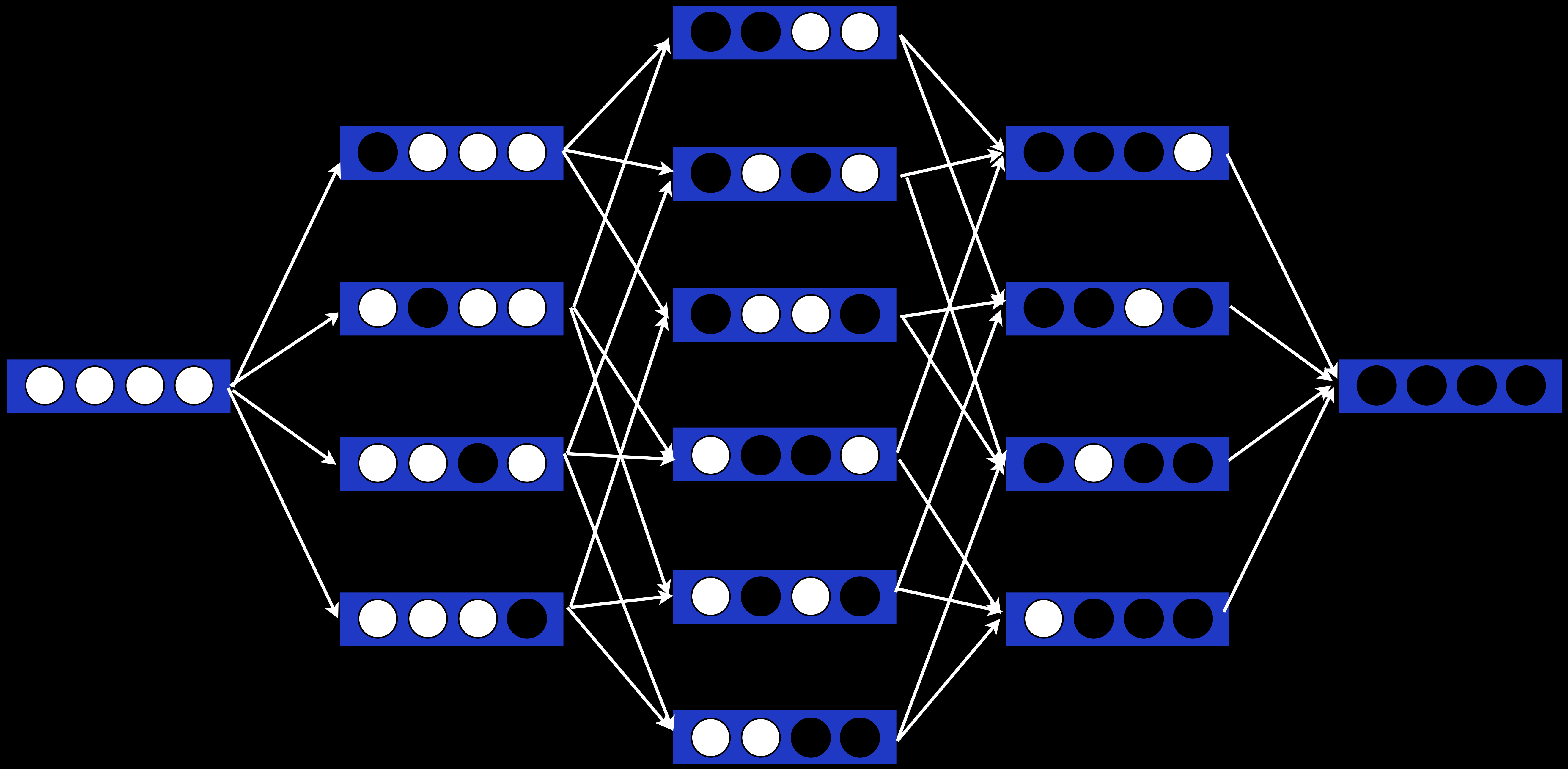


Dynamic Programming

Key Idea

amount of work:

$O(5^n 2^n)$

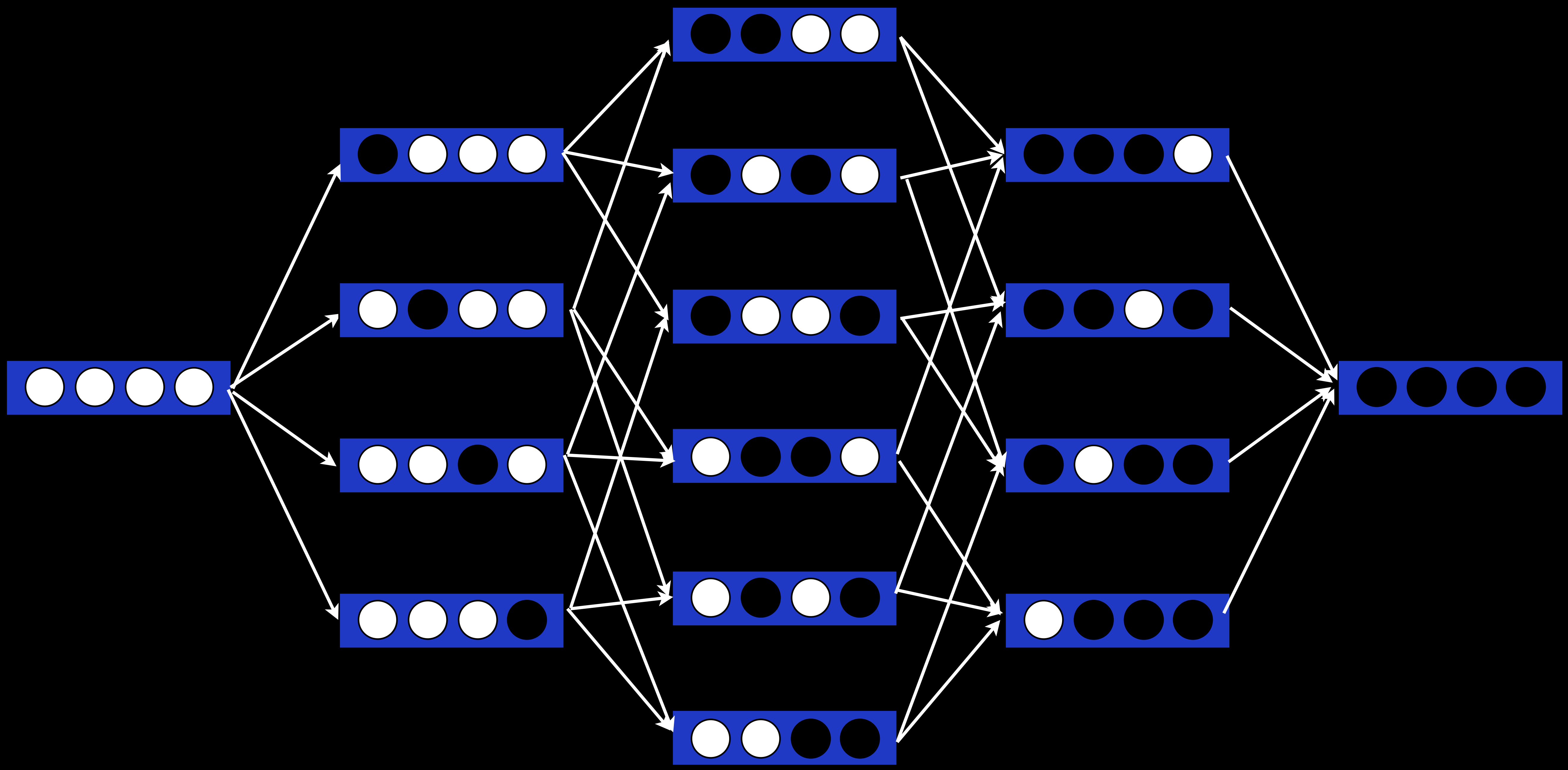


Dynamic Programming

Key Idea

amount of work: $O(5^n 2^n)$

bad, but much better than $O(5^n n!)$

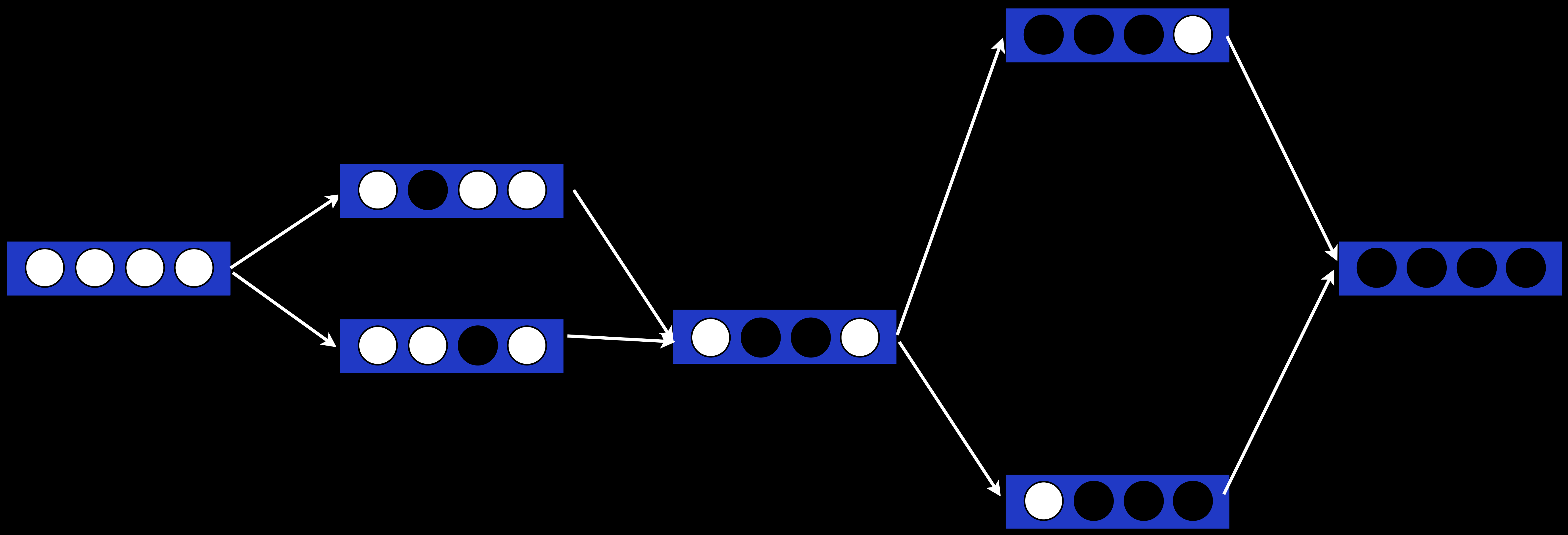


Dynamic Programming

Key Idea

amount of work:
$O(5^n 2^n)$

bad, but much better than
$O(5^n n!)$

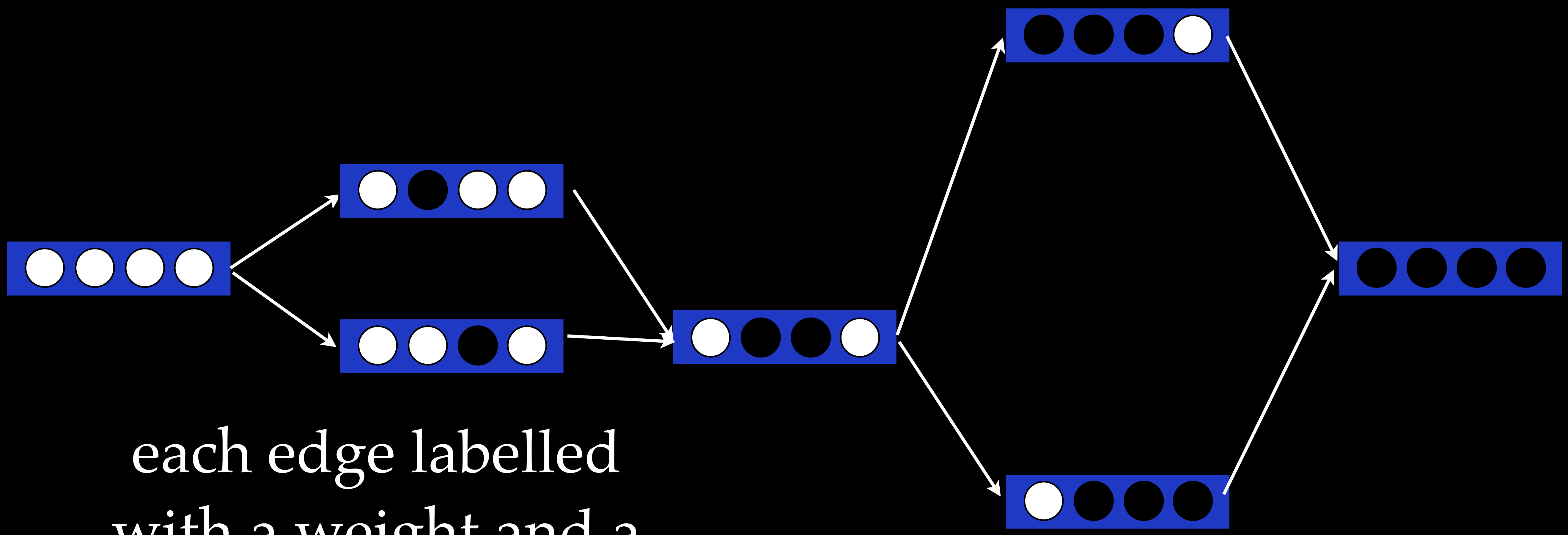


Dynamic Programming

Key Idea

amount of work:
$O(5^n 2^n)$

bad, but much better than $O(5^n n!)$



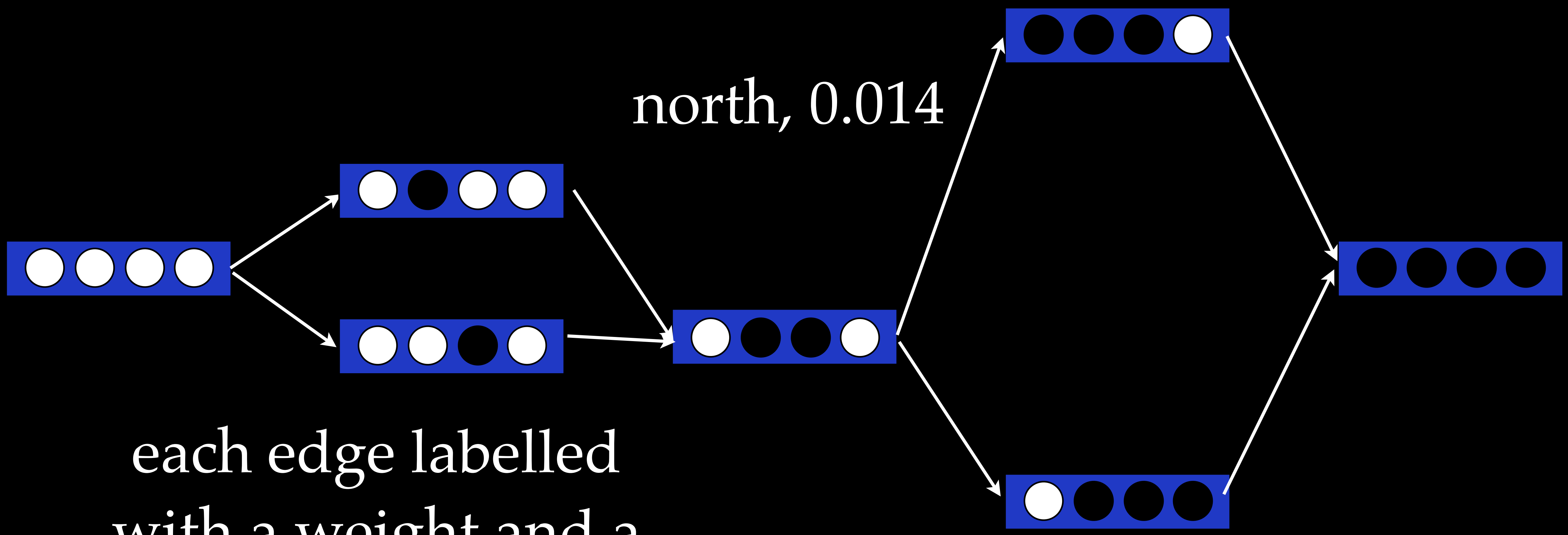
each edge labelled with a weight and a word (or words)

Dynamic Programming

Key Idea

amount of work: $O(5^n 2^n)$

bad, but much better than $O(5^n n!)$



each edge labelled with a weight and a word (or words)

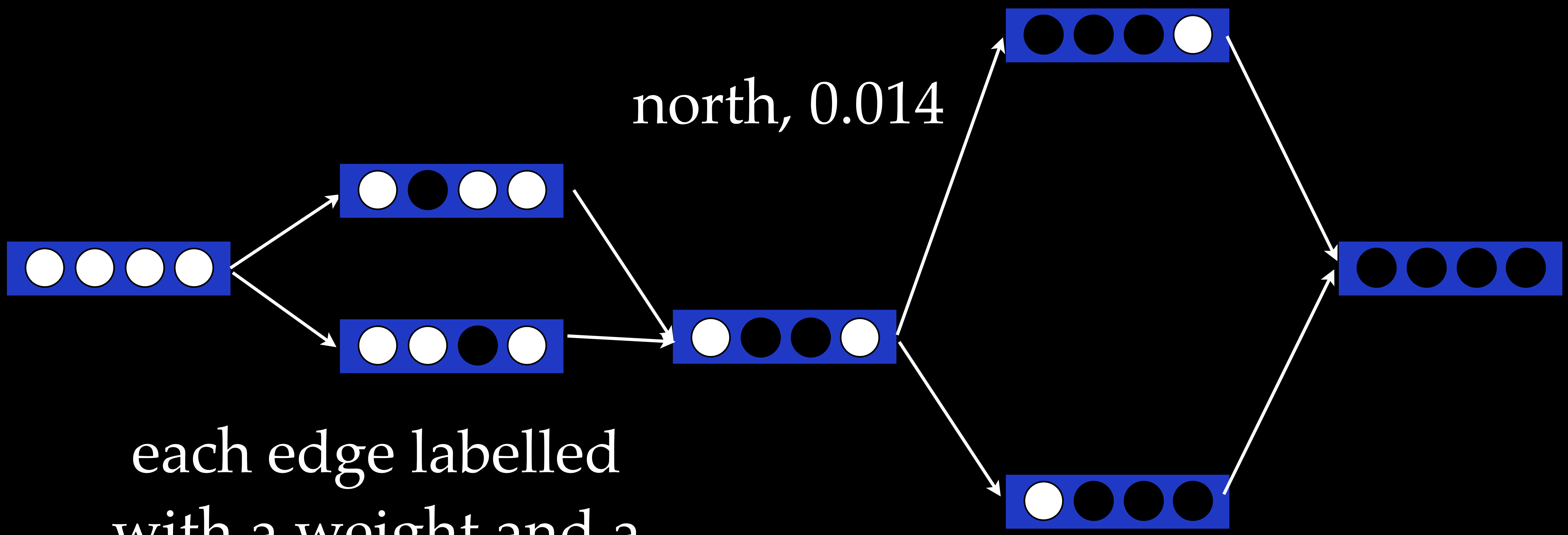
Dynamic Programming

Key Idea

amount of work: $O(5^n 2^n)$

bad, but much better than $O(5^n n!)$

weighted finite-state automata



each edge labelled with a weight and a word (or words)

Dynamic Programming

Weighted languages

- The lattice describing the set of all possible translations is a ***weighted finite state automaton***.
- So is the language model.
- Since regular languages are closed under intersection, we can intersect the devices and run shortest path graph algorithms.
- Taking their intersection is equivalent to computing the probability under Bayes' rule.

Practical Issues

$O(5^n 2^n)$ is still far too much work.

Practical Issues

$O(5^n 2^n)$ is still far too much work.

Can we do better?

Can we do better?

北 风 呼啸 。

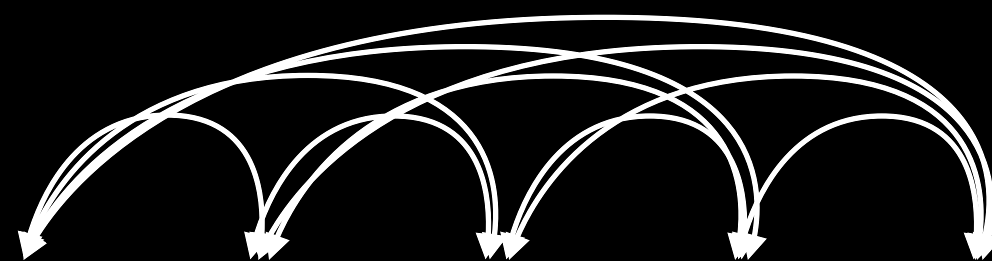
Can we do better?

北　风　呼啸　。

north wind the strong .

Can we do better?

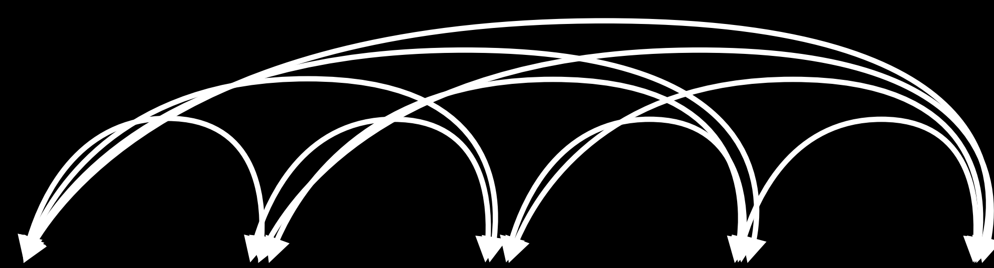
北 风 呼啸 。



north wind the strong .

Can we do better?

北 风 呼啸 。

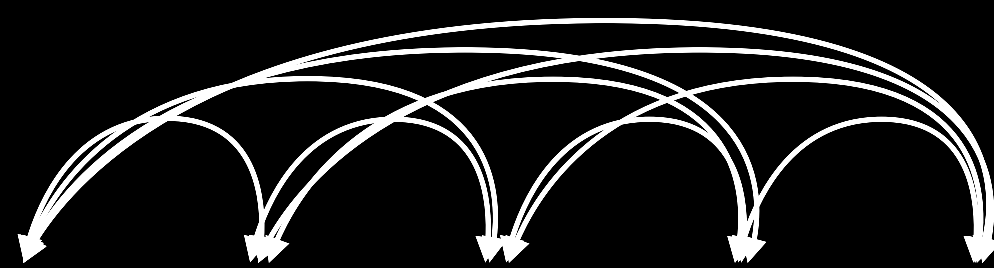


north wind the strong .

Each arc weighted by
translation probability +
bigram probability

Can we do better?

北 风 呼啸 。



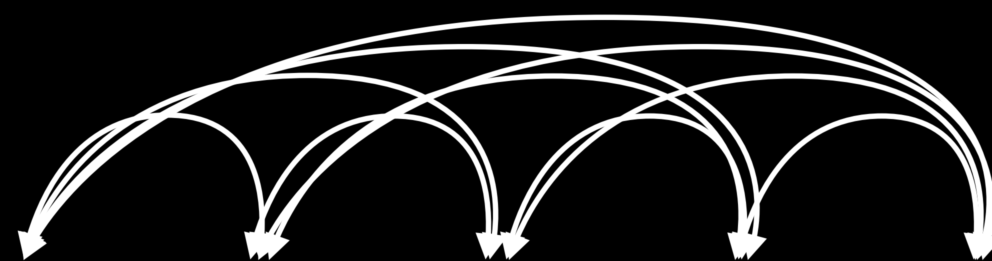
north wind the strong .

Each arc weighted by
translation probability +
bigram probability

Objective: find shortest path that visits each word once.

Can we do better?

北 风 呼啸 。



London Paris NY Tokyo .

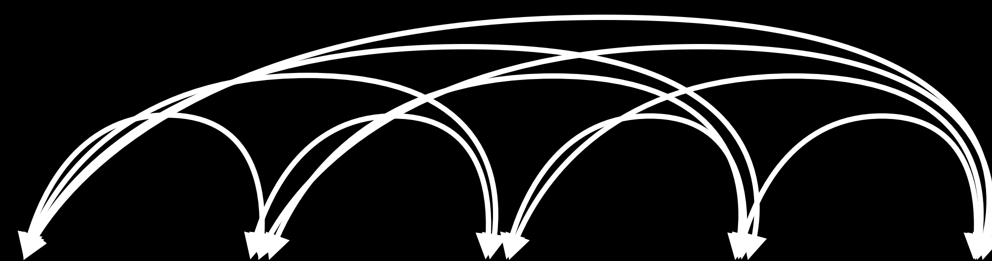
Each arc weighted by
translation probability +
bigram probability

Objective: find shortest path that visits each word once.

Can we do better?

Probably not: this is the traveling salesman problem.

北 风 呼啸 。



London Paris NY Tokyo .

Each arc weighted by
translation probability +
bigram probability

Objective: find shortest path that visits each word once.

Two Problems

- Exact decoding requires exponential time.
  - This is a consequence of arbitrary permutation.
  - But in translation reordering is not arbitrary!
- Parameterization of reordering is weak.
  - No generalization!

la empresa tiene enemigos fuertes en Europa .

the company has strong enemies in Europe .

Garcia and associates .

Garcia y asociados .

Carlos Garcia has three associates .

Carlos Garcia tiene tres asociados .

his associates are not strong .

sus asociados no son fuertes .

Garcia has a company also .

Garcia tambien tiene una empresa .

its clients are angry .

sus clientes estan enfadados .

the associates are also angry .

los asociados tambien estan enfadados .

the clients and the associates are enemies .

los clientes y los asociados son enemigos .

the company has three groups .

la empresa tiene tres grupos .

its groups are in Europe .

sus grupos estan en Europa .

the modern groups sell strong pharmaceuticals .

los grupos modernos venden medicinas fuertes .

the groups do not sell zanzanine .

los grupos no venden zanzanina .

the small groups are not modern .

los grupos pequenos no son modernos .

la empresa tiene enemigos fuertes en Europa .

the company has strong enemies in Europe .

Garcia and associates .

Garcia y asociados .

Carlos Garcia has three associates .

Carlos Garcia tiene tres asociados .

his associates are not strong .

sus asociados no son fuertes .

Garcia has a company also .

Garcia tambien tiene una empresa .

its clients are angry .

sus clientes estan enfadados .

the associates are also angry .

los asociados tambien estan enfadados .

the clients and the associates are enemies .

los clientes y los asociados son enemigos .

the company has three groups .

la empresa tiene tres grupos .

its groups are in Europe .

sus grupos estan en Europa .

the modern groups sell strong pharmaceuticals .

los grupos modernos venden medicinas fuertes .

the groups do not sell zanzanine .

los grupos no venden zanzanina .

the small groups are not modern .

los grupos pequenos no son modernos .

la empresa tiene enemigos fuertes en Europa .

the company has strong enemies in Europe .

Same pattern:

NN JJ → JJ NN

sus asociados no son fuertes .

sus grupos estan en Europa .

Garcia has a company also .

Garcia tambien tiene una empresa .

the modern groups sell strong pharmaceuticals .

los grupos modernos venden medicinas fuertes .

its clients are angry .

sus clientes estan enfadados .

the groups do not sell zanzanine .

los grupos no venden zanzanina .

the associates are also angry .

los asociados tambien estan enfadados .

the small groups are not modern .

los grupos pequenos no son modernos .

la empresa tiene enemigos fuertes en Europa .

the company has strong enemies in Europe .

Same pattern:

NN JJ → JJ NN

Finite-state models do not capture this generalization.

sus asociados no son fuertes .

Garcia has a company also .

Garcia tambien tiene una empresa .

its clients are angry .

sus clientes estan enfadados .

the associates are also angry .

los asociados tambien estan enfadados .

sus grupos estan en Europa .

the modern groups sell strong pharmaceuticals .

los grupos modernos venden medicinas fuertes .

the groups do not sell zanzanine .

los grupos no venden zanzanina .

the small groups are not modern .

los grupos pequenos no son modernos .

Context-Free Grammar

Context-Free Grammar

$S \rightarrow NP\ VP$

$NP \rightarrow$ watashi wa

$NP \rightarrow$ hako wo

$VP \rightarrow NP\ V$

$V \rightarrow$ akemasu

Context-Free Grammar

S

S → NP VP
NP → watashi wa
NP → hako wo
VP → NP V
V → akemasu

Context-Free Grammar

S

$S \rightarrow NP\ VP$

$NP \rightarrow$ watashi wa

$NP \rightarrow$ hako wo

$VP \rightarrow NP\ V$

$V \rightarrow$ akemasu

Context-Free Grammar

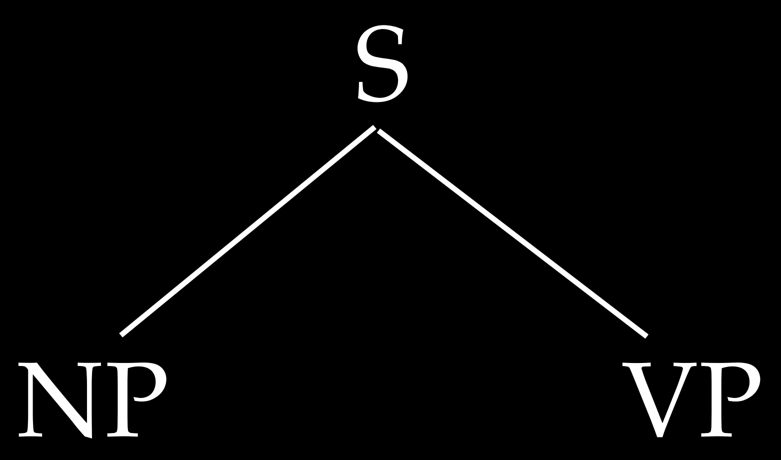
S → NP VP

NP → watashi wa

NP → hako wo

VP → NP V

V → akemasu



Context-Free Grammar

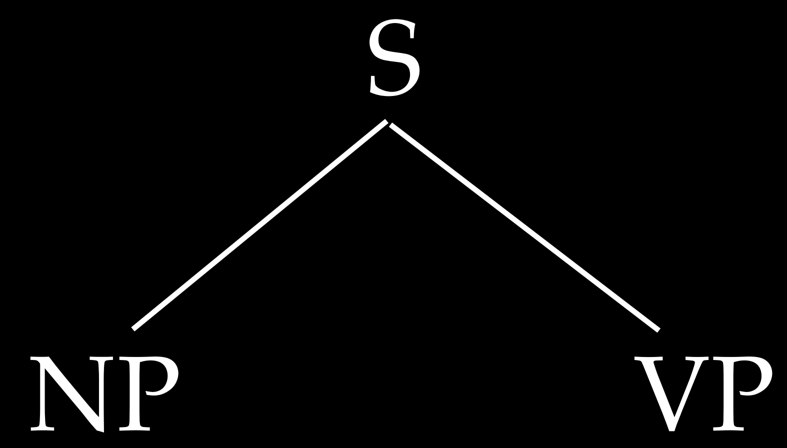
S → NP VP

NP → watashi wa

NP → hako wo

VP → NP V

V → akemasu



Context-Free Grammar

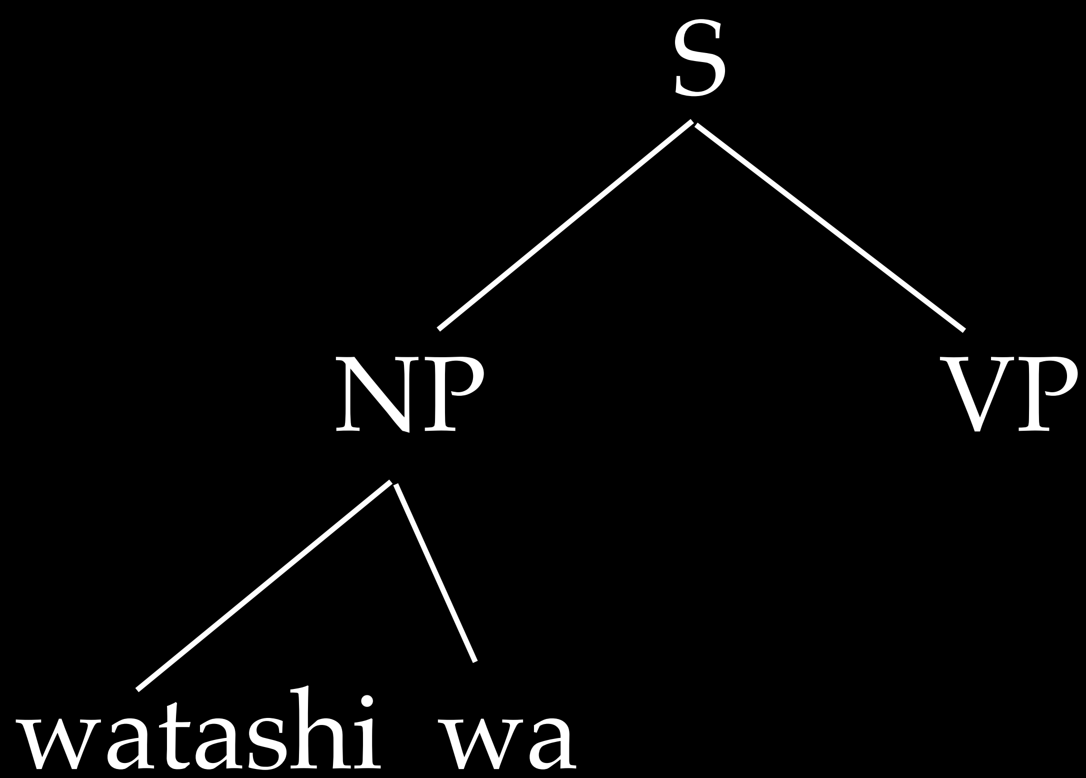
S → NP VP

NP → watashi wa

NP → hako wo

VP → NP V

V → akemasu



Context-Free Grammar

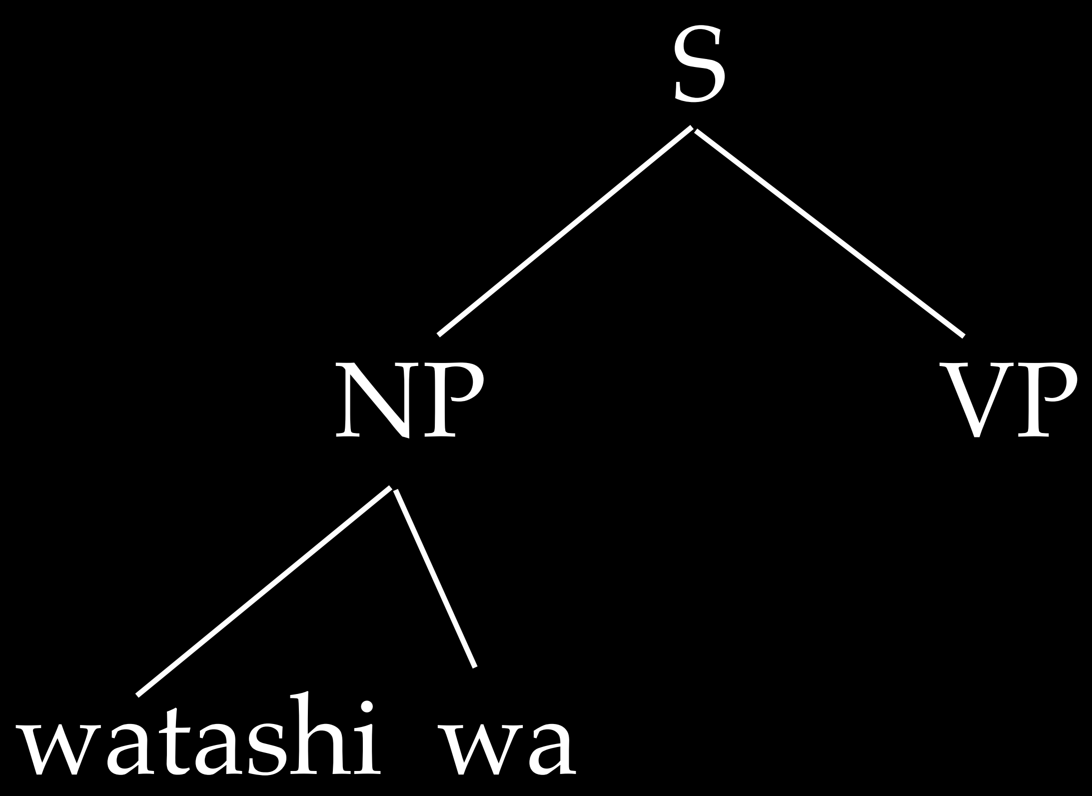
S → NP VP

NP → watashi wa

NP → hako wo

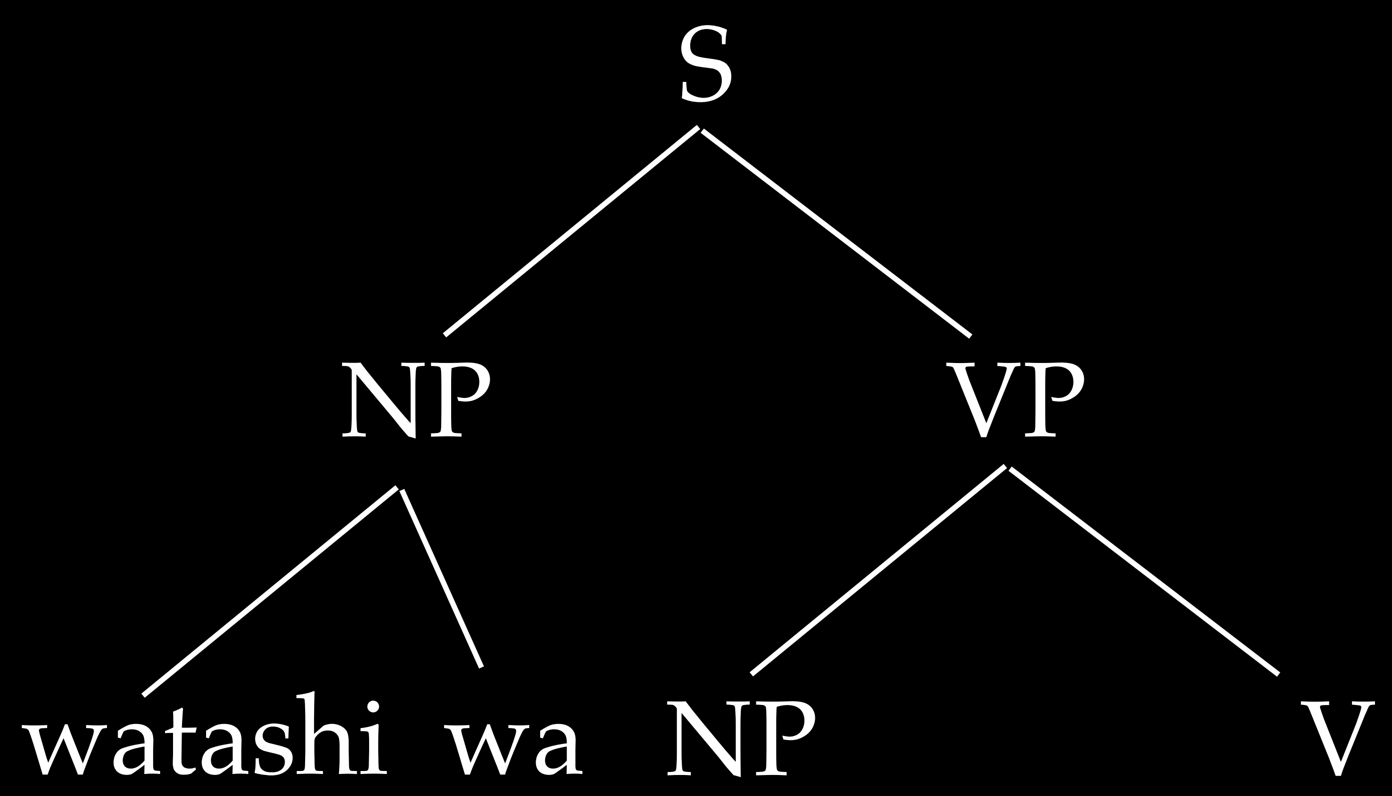
VP → NP V

V → akemasu



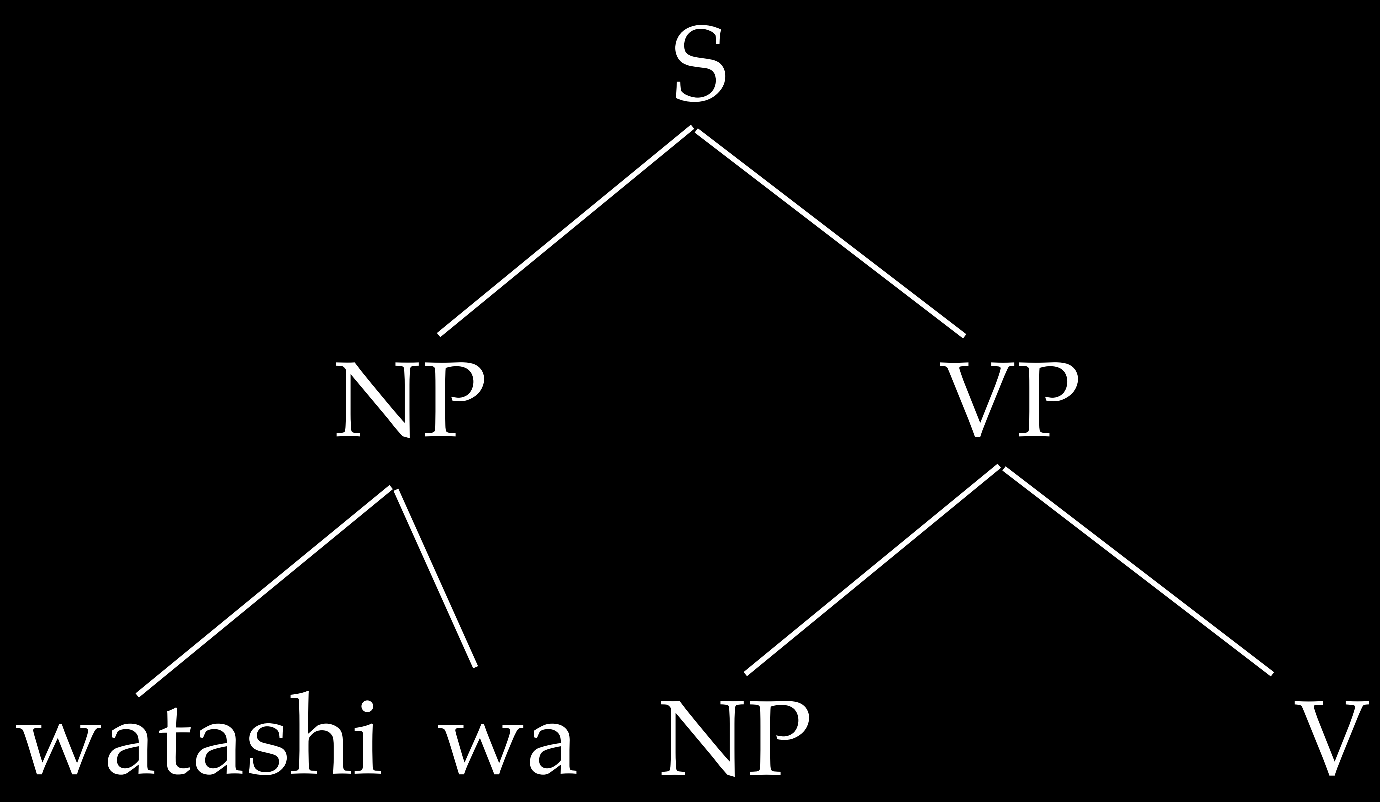
Context-Free Grammar

S → NP VP
NP → watashi wa
NP → hako wo
VP → NP V
V → akemasu



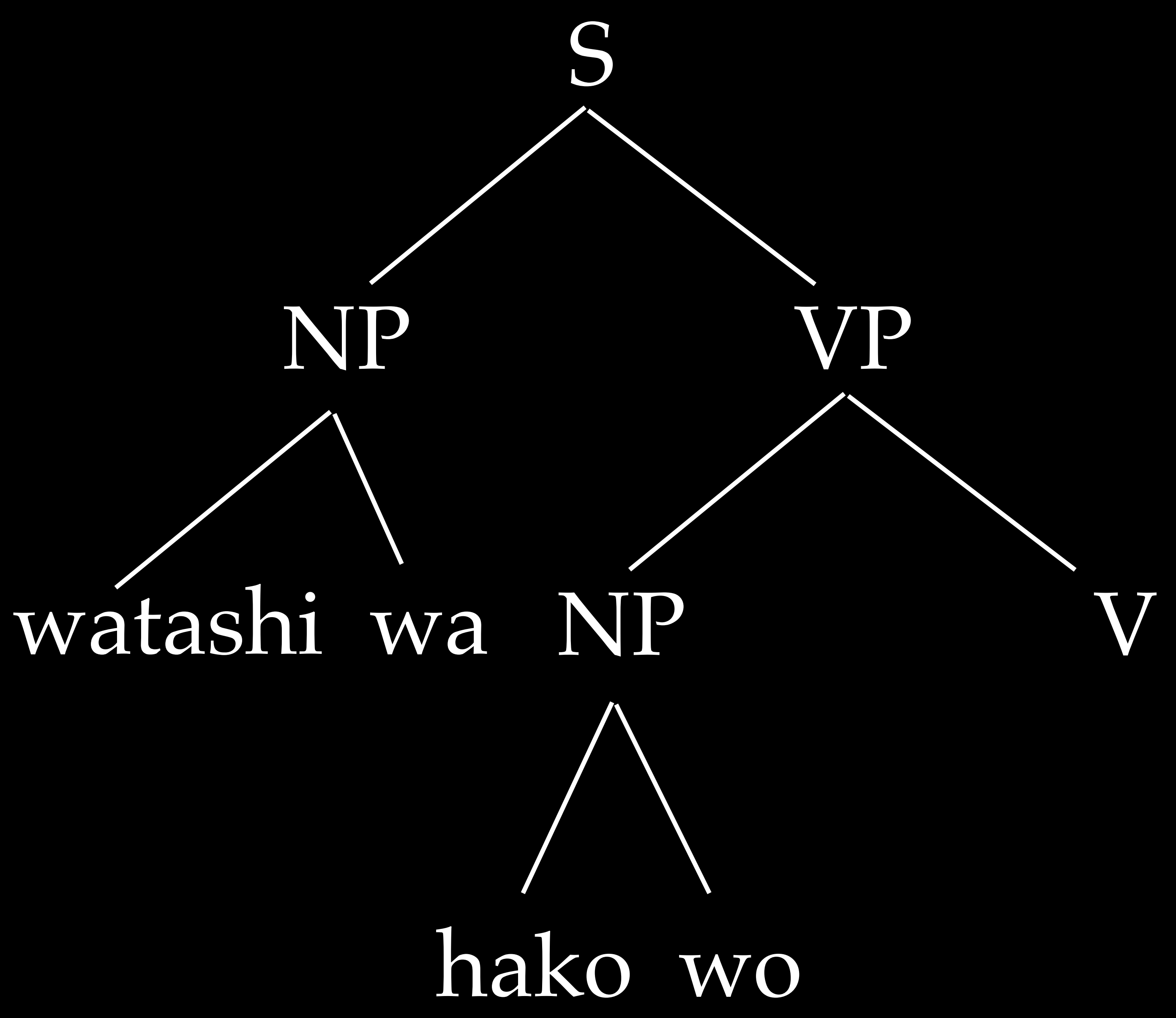
Context-Free Grammar

S → NP VP
NP → watashi wa
NP → hako wo
VP → NP V
V → akemasu



Context-Free Grammar

S → NP VP
NP → watashi wa
NP → hako wo
VP → NP V
V → akemasu



Context-Free Grammar

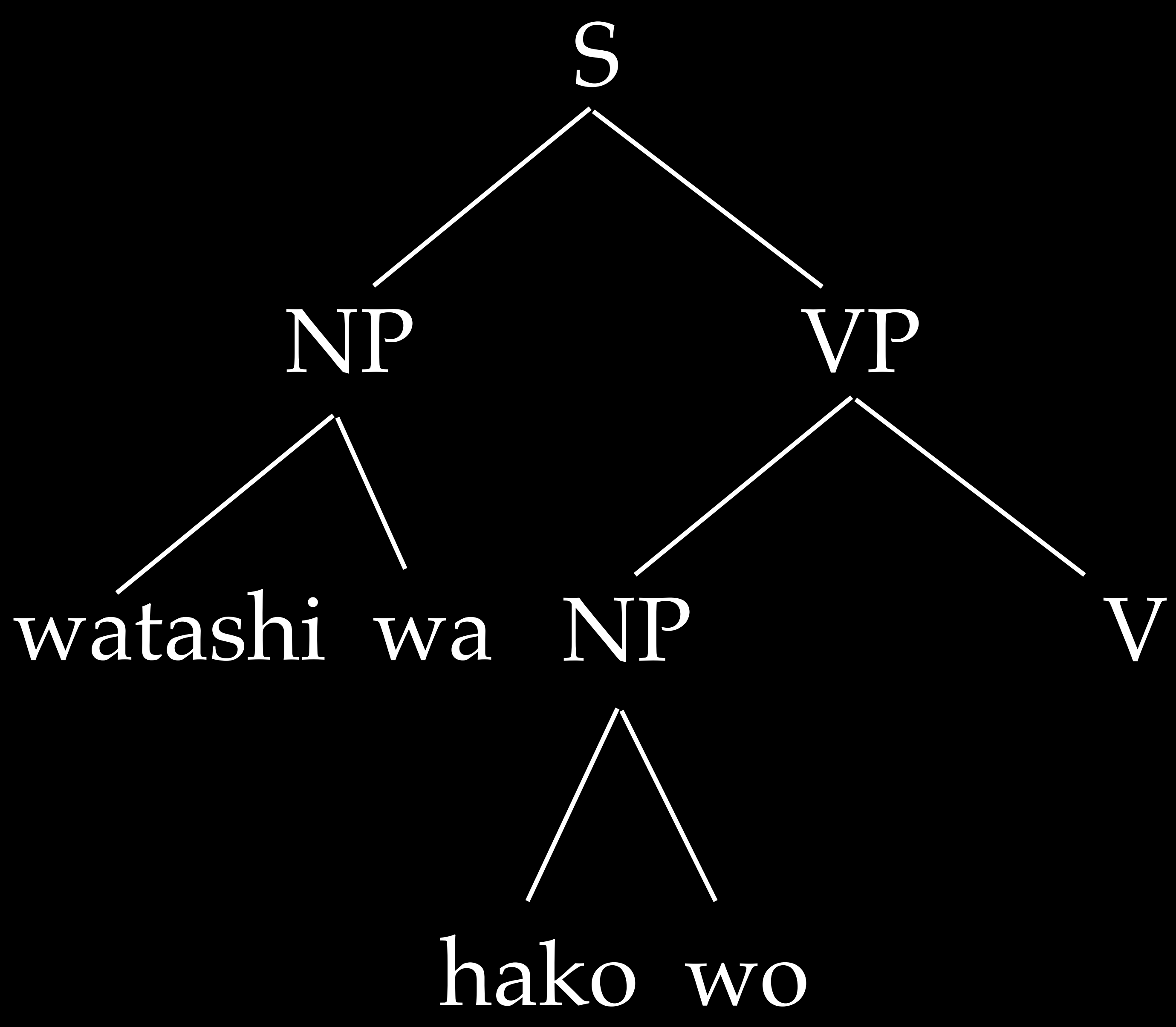
$S \rightarrow NP\ VP$

$NP \rightarrow$ watashi wa

$NP \rightarrow$ hako wo

$VP \rightarrow NP\ V$

$V \rightarrow$ akemasu



Context-Free Grammar

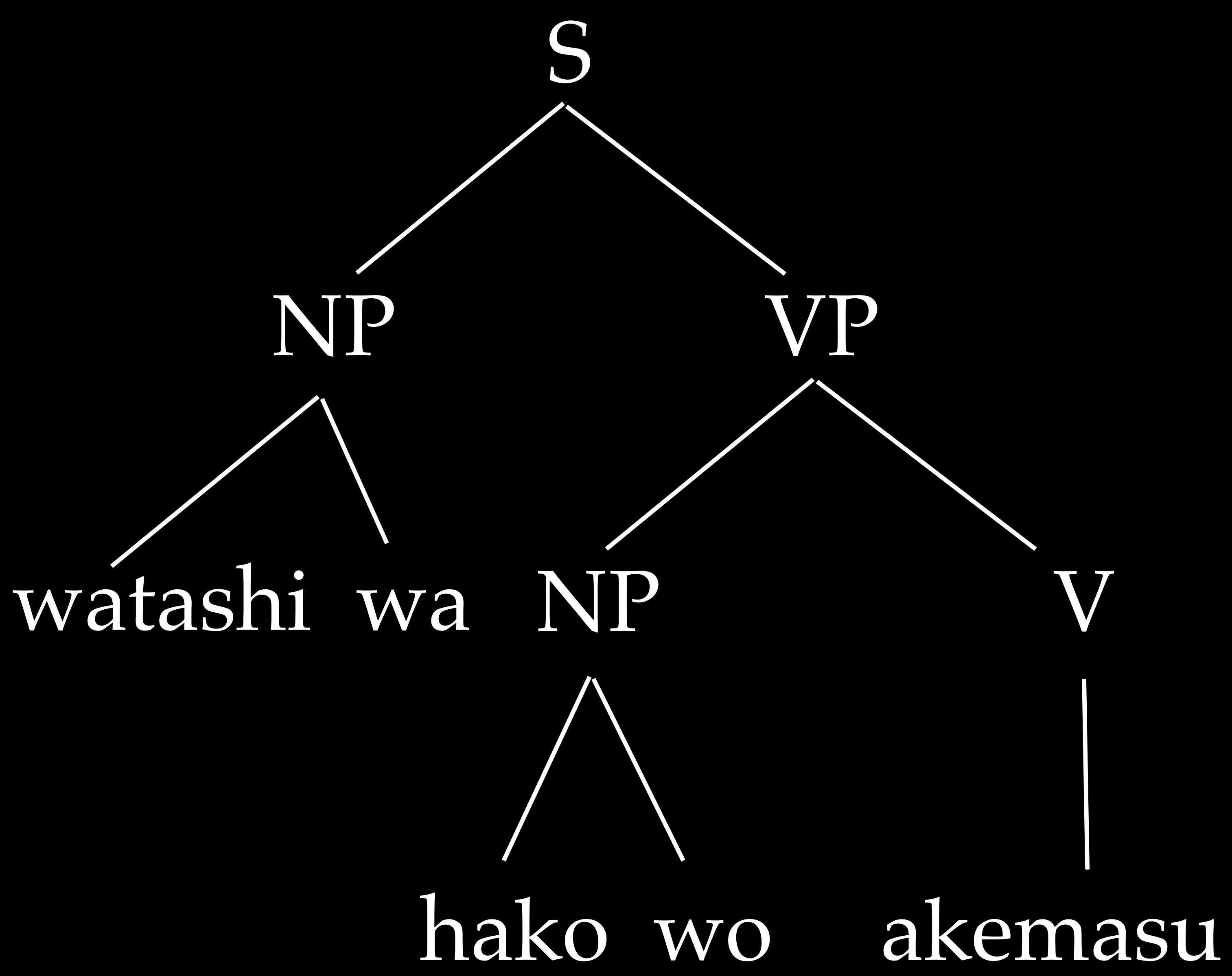
S → NP VP

NP → watashi wa

NP → hako wo

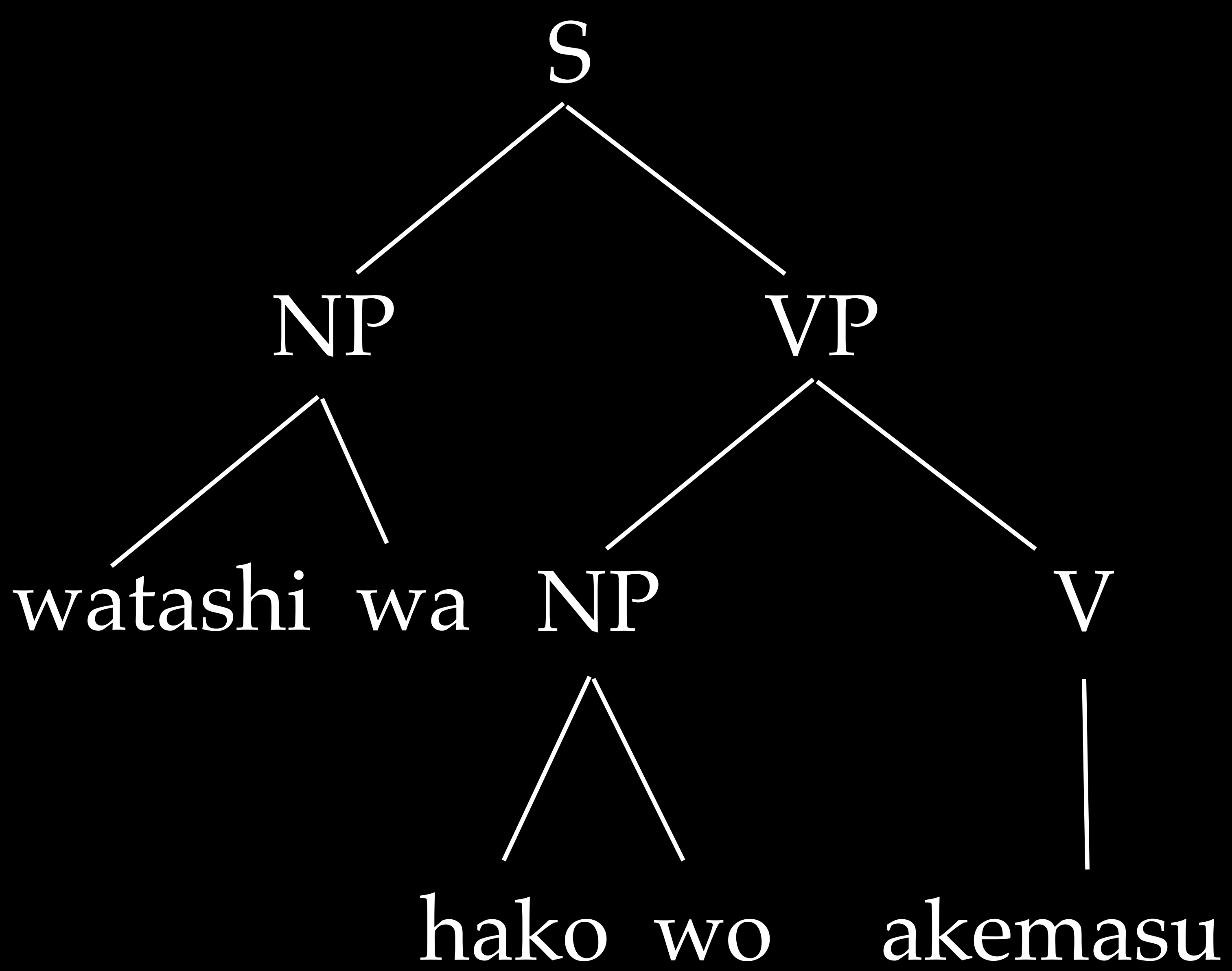
VP → NP V

V → akemasu



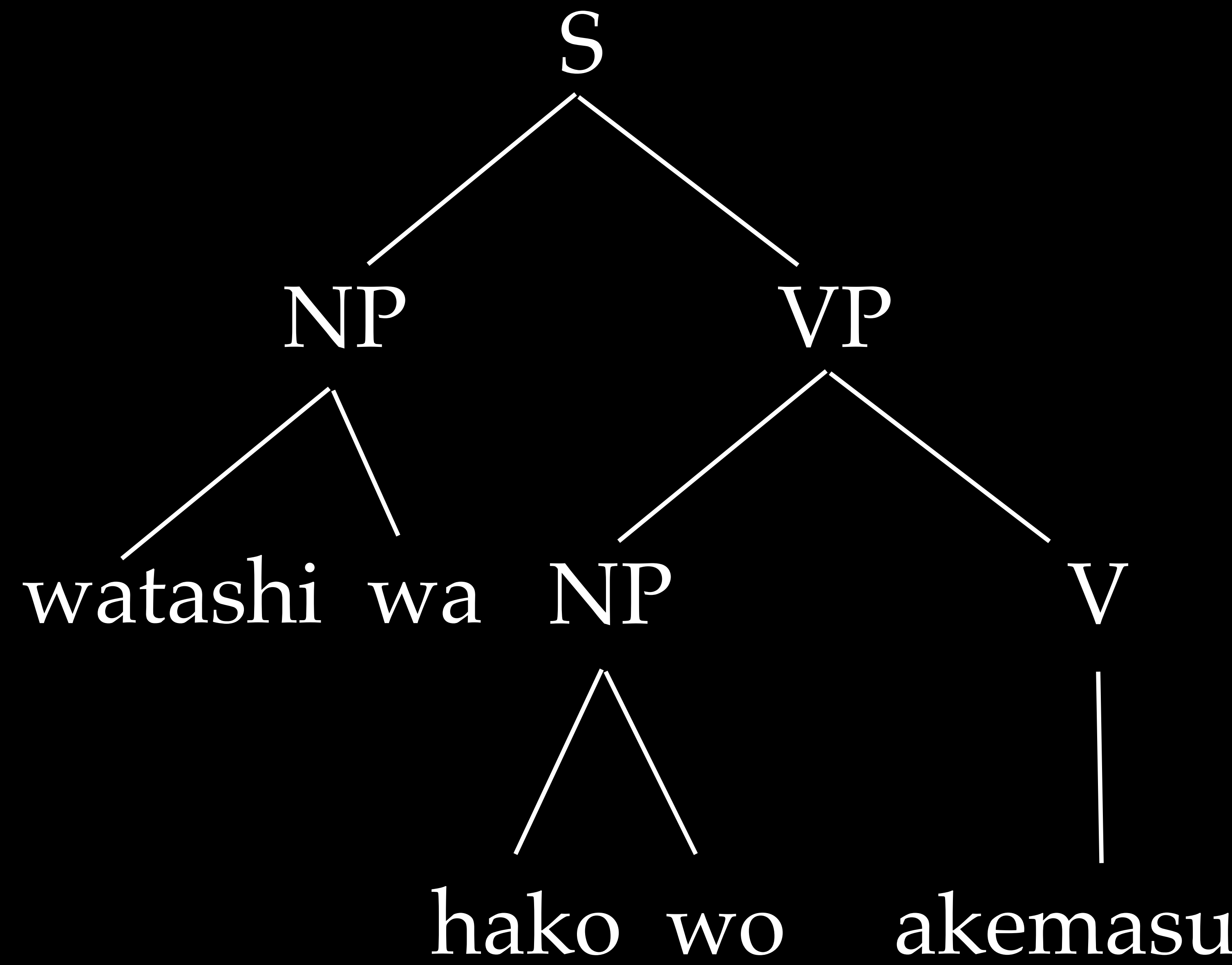
Context-Free Grammar

S → NP VP
NP → watashi wa
NP → hako wo
VP → NP V
V → akemasu



Context-Free Grammar

- S → NP VP
- NP → watashi wa
- NP → hako wo
- VP → NP V
- V → akemasu



watashi wa hako wo akemasu

Context-Free Grammar

$S \rightarrow NP\ VP$
$NP \rightarrow$ watashi wa
$NP \rightarrow$ hako wo
$VP \rightarrow NP\ V$
$V \rightarrow$ akemasu

Note: this particular grammar is finite, hence regular.

$$\left\{ \begin{array}{c} \text{watashi wa watashi wa akemasu} \\ \text{watashi wa hako wo akemasu} \\ \text{hako wo hako wo akemasu} \\ \text{hako wo watashi wa akemasu} \end{array} \right\}$$

Context-Free Grammar

$S \rightarrow A\ B$

$S \rightarrow A\ S\ B$

$A \rightarrow a$

$B \rightarrow b$

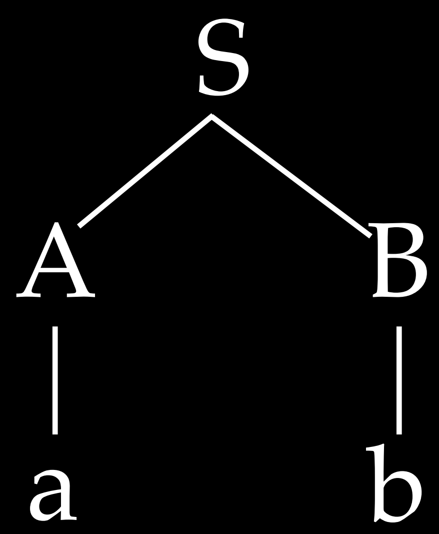
Context-Free Grammar

$S \rightarrow A\ B$

$S \rightarrow A\ S\ B$

$A \rightarrow a$

$B \rightarrow b$



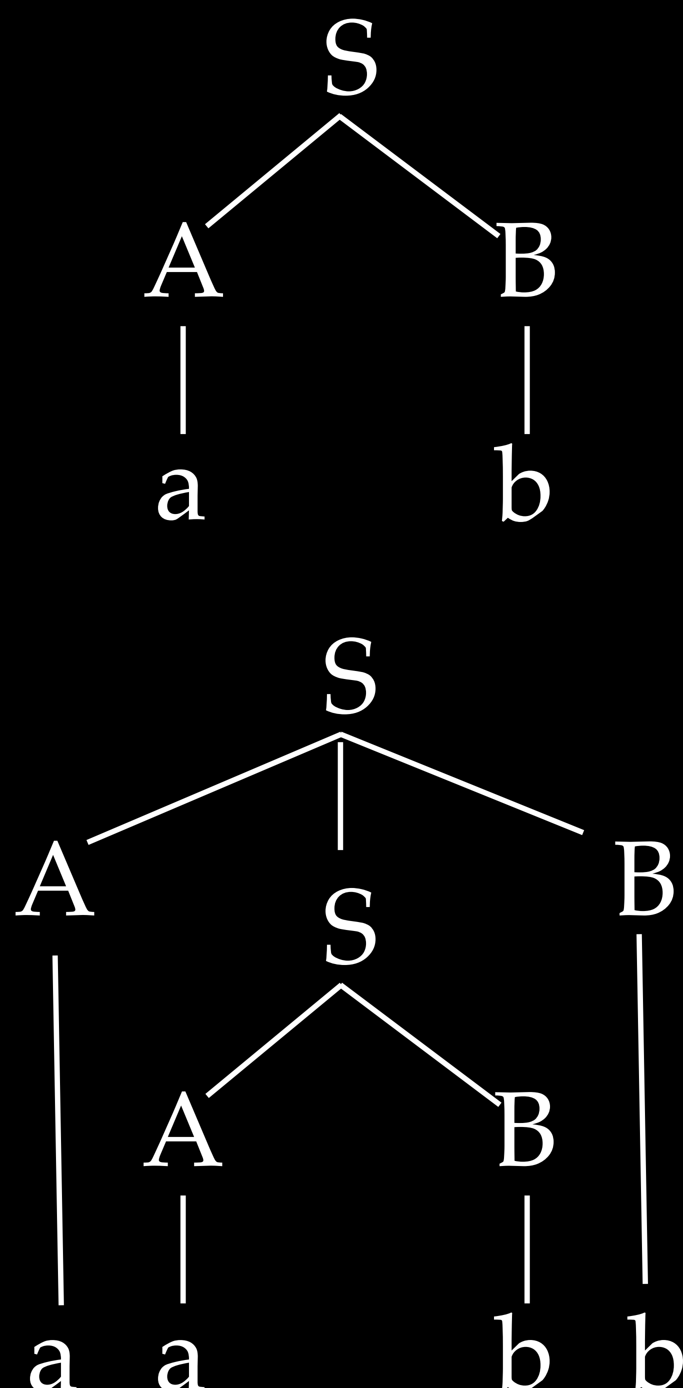
Context-Free Grammar

$S \rightarrow A\ B$

$S \rightarrow A\ S\ B$

$A \rightarrow a$

$B \rightarrow b$



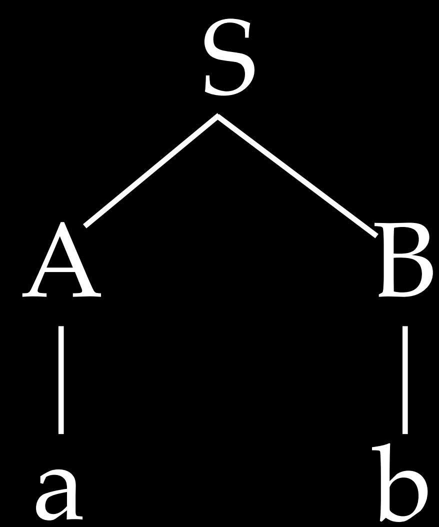
Context-Free Grammar

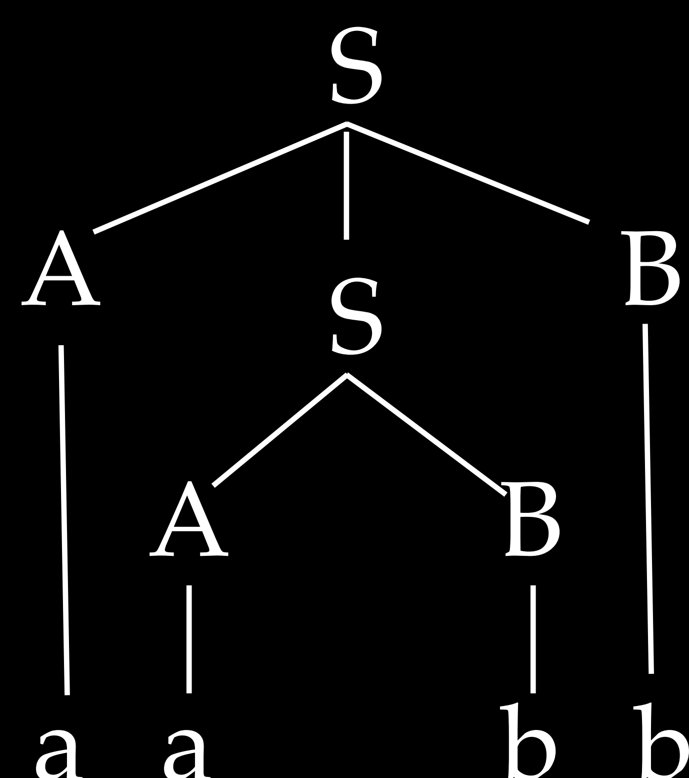
$S \rightarrow A\ B$

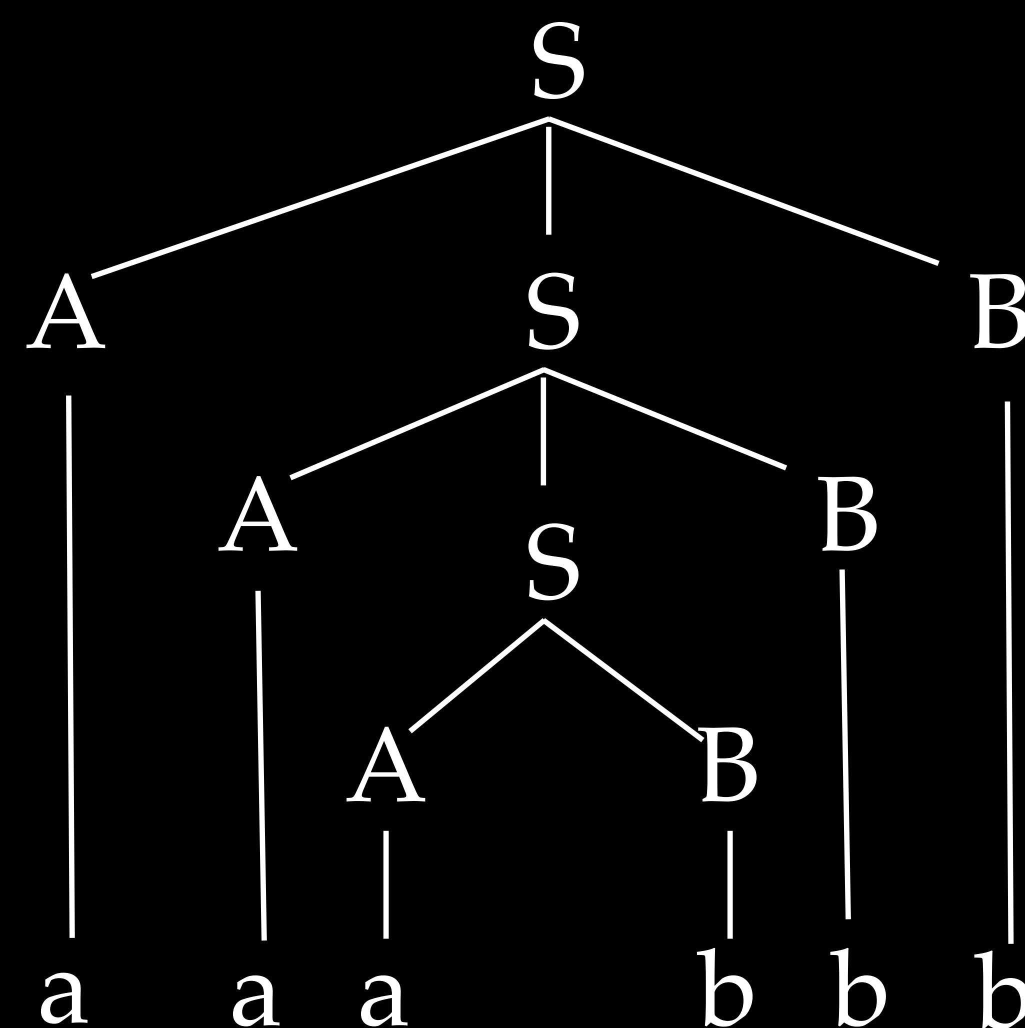
$S \rightarrow A\ S\ B$

$A \rightarrow a$

$B \rightarrow b$







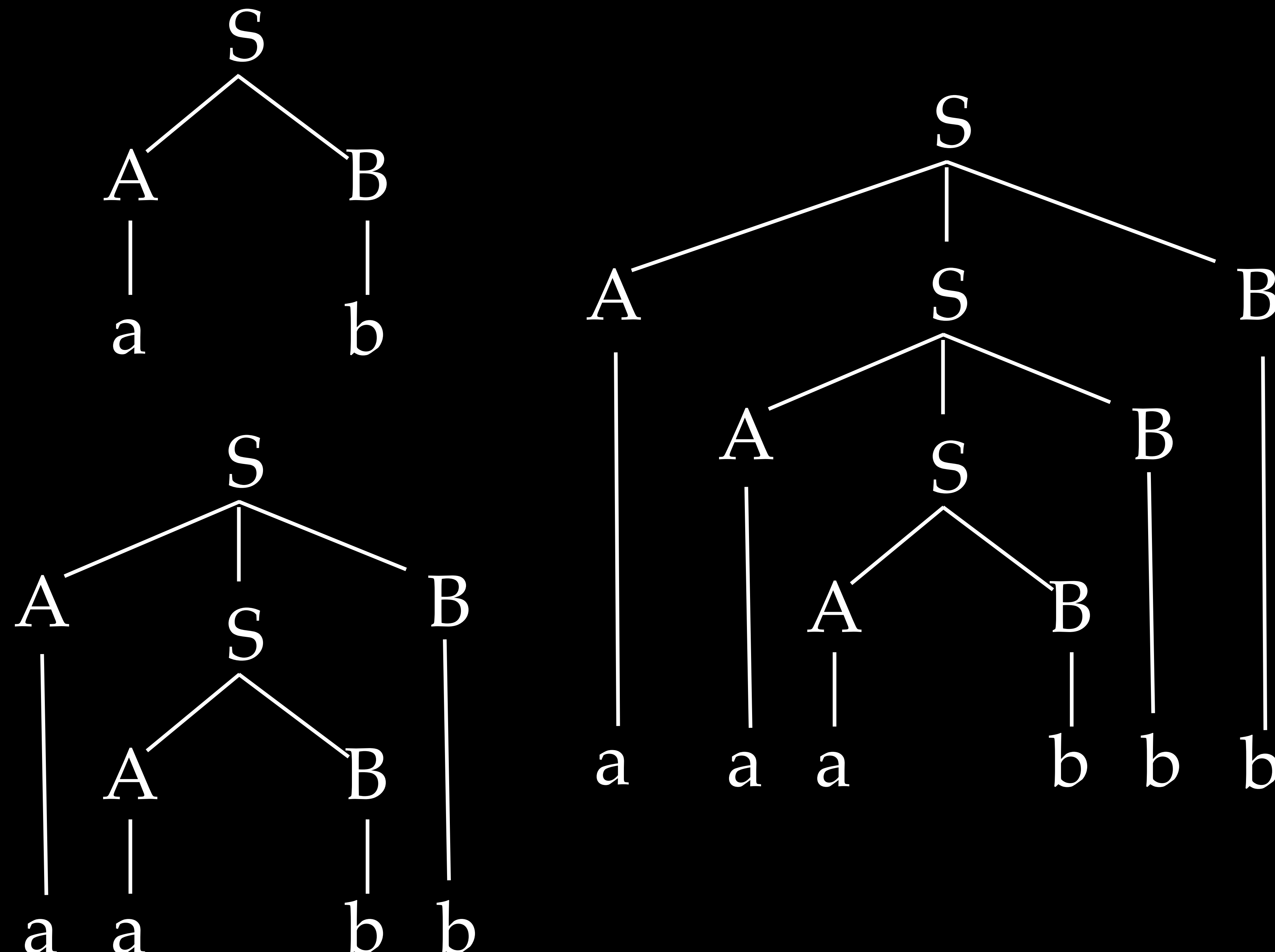
Context-Free Grammar

S → A B

S → A S B

A → a

B → b



$$\mathcal{L}_4 = \{ab, aabb, aaabbb, ...\} = \forall_{n \in [1, \text{inf})} a^n b^n$$

Context-Free vs. Regular

Context-Free vs. Regular

- Regular languages ⊂ Context-free languages

Context-Free vs. Regular

- Regular languages ⊂ Context-free languages
- Composition of languages:

Context-Free vs. Regular

- Regular languages ⊂ Context-free languages
- Composition of languages:
  - Regular ∩ Regular = Regular

Context-Free vs. Regular

- Regular languages $\subset$ Context-free languages
- Composition of languages:
  - Regular $\cap$ Regular = Regular
  - Regular $\cap$ Context-free = Context-free

Context-Free vs. Regular

- Regular languages ⊂ Context-free languages
- Composition of languages:
  - Regular ∩ Regular = Regular
  - Regular ∩ Context-free = Context-free

$$A \rightarrow BC \in \mathcal{G}_{CFL}$$

Context-Free vs. Regular

- Regular languages $\subset$ Context-free languages
- Composition of languages:
  - Regular $\cap$ Regular = Regular
  - Regular $\cap$ Context-free = Context-free

$A \rightarrow BC \in \mathcal{G}_{CFL}$ $\qquad s, r, t \in states(\mathcal{G}_{RL})$

Context-Free vs. Regular

- Regular languages ⊂ Context-free languages
- Composition of languages:
  - Regular ∩ Regular = Regular
  - Regular ∩ Context-free = Context-free

$$A \rightarrow BC \in \mathcal{G}_{CFL} \qquad\qquad s, r, t \in states(\mathcal{G}_{RL})$$

$$_sA_t \rightarrow {_sB_r}{_rC_t} \in \mathcal{G}_{CFL} \cap \mathcal{G}_{RL}$$

Context-Free vs. Regular

- Regular languages ⊂ Context-free languages
- Composition of languages:
  - Regular ∩ Regular = Regular
  - Regular ∩ Context-free = Context-free

$$A \rightarrow BC \in \mathcal{G}_{CFL} \qquad\qquad s, r, t \in states(\mathcal{G}_{RL})$$

$$_sA_t \rightarrow {_sB_r}{_rC_t} \in \mathcal{G}_{CFL} \cap \mathcal{G}_{RL}$$

Bar-Hillel 1964

Context-Free vs. Regular

- Regular languages ⊂ Context-free languages
- Composition of languages:
  - Regular ∩ Regular = Regular
  - Regular ∩ Context-free = Context-free
  - Context-free ∩ Context-free = Undecidable

$$A \rightarrow BC \in \mathcal{G}_{CFL} \qquad\qquad s, r, t \in states(\mathcal{G}_{RL})$$

$$_sA_t \rightarrow {}_sB_{rr}C_t \in \mathcal{G}_{CFL} \cap \mathcal{G}_{RL}$$

Bar-Hillel 1964

Synchronous Context-Free Grammar

$$
\begin{aligned}
S &\rightarrow NP\ VP \\
NP &\rightarrow \text{watashi wa} \\
NP &\rightarrow \text{hako wo} \\
VP &\rightarrow NP\ V \\
V &\rightarrow \text{akemasu}
\end{aligned}
$$

Synchronous Context-Free Grammar

S → NP VP	S → NP VP
NP → watashi wa	NP → I
NP → hako wo	NP → the box
VP → NP V	VP → V NP
V → akemasu	V → open

Synchronous Context-Free Grammar

$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar

$$\mathrm{S} \rightarrow \mathrm{NP}_1\ \mathrm{VP}_2 \ /\ \mathrm{NP}_1\ \mathrm{VP}_2$$

$$\mathrm{NP} \rightarrow \text{watashi wa} \ /\ \text{I}$$

$$\mathrm{NP} \rightarrow \text{hako wo} \ /\ \text{the box}$$

$$\mathrm{VP} \rightarrow \mathrm{NP}_1\ \mathrm{V}_2 \ /\ \mathrm{V}_2\ \mathrm{NP}_1$$

$$\mathrm{V} \rightarrow \text{akemasu} \ /\ \text{open}$$

Synchronous Context-Free Grammar

S S

$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar

S S

$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar

S .... S

$\boxed{S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2}$

$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar

$\boxed{S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2}$

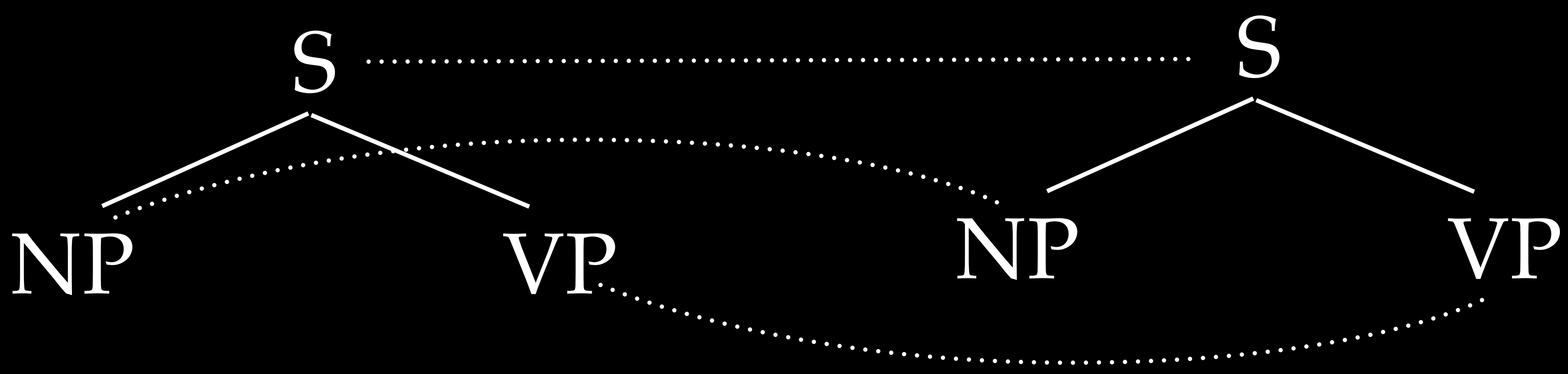
$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar



$\boxed{S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2}$

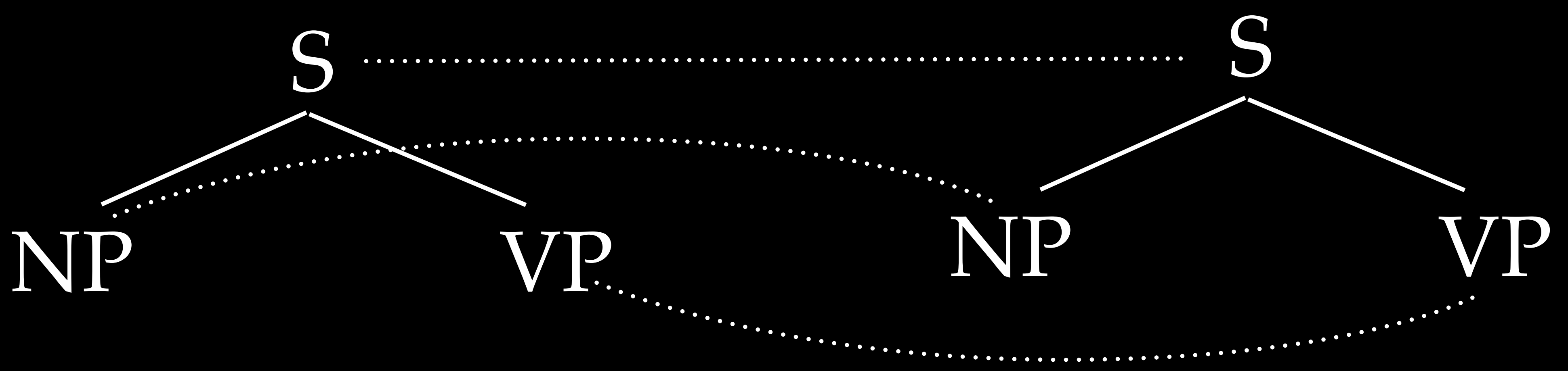
NP → watashi wa / I

NP → hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

V → akemasu / open

Synchronous Context-Free Grammar



$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

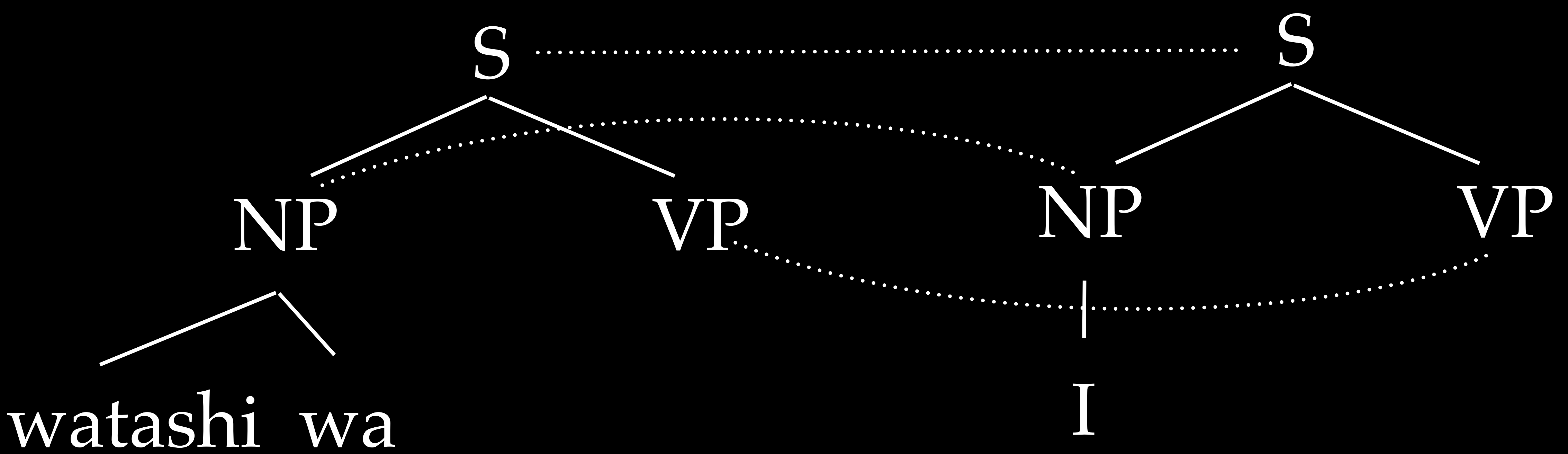
$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar



$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

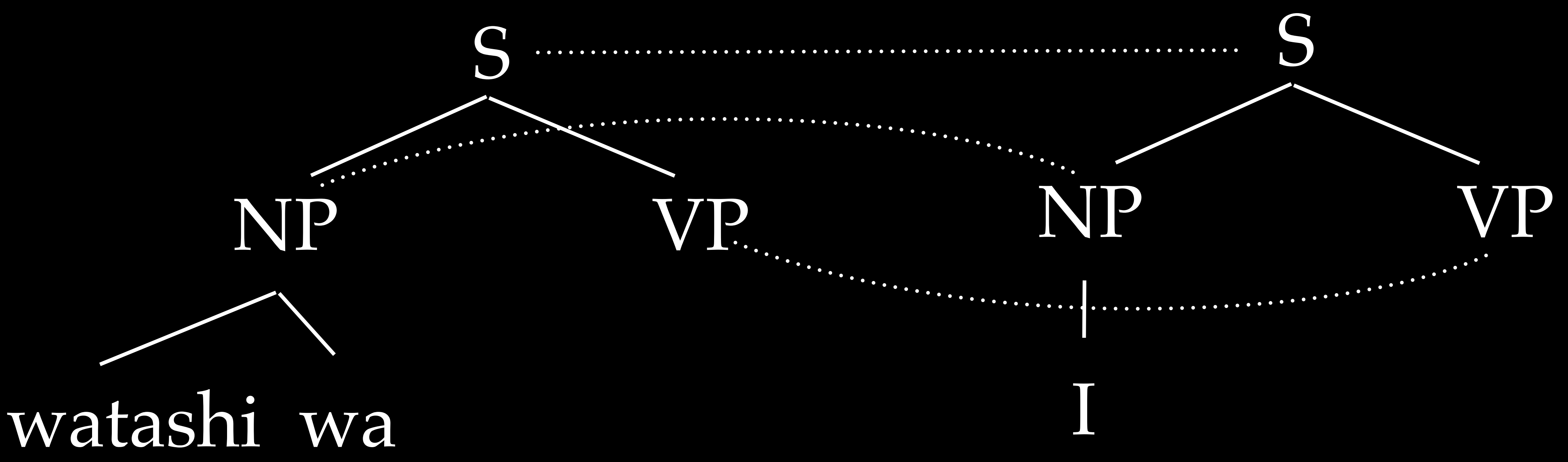
$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar



$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

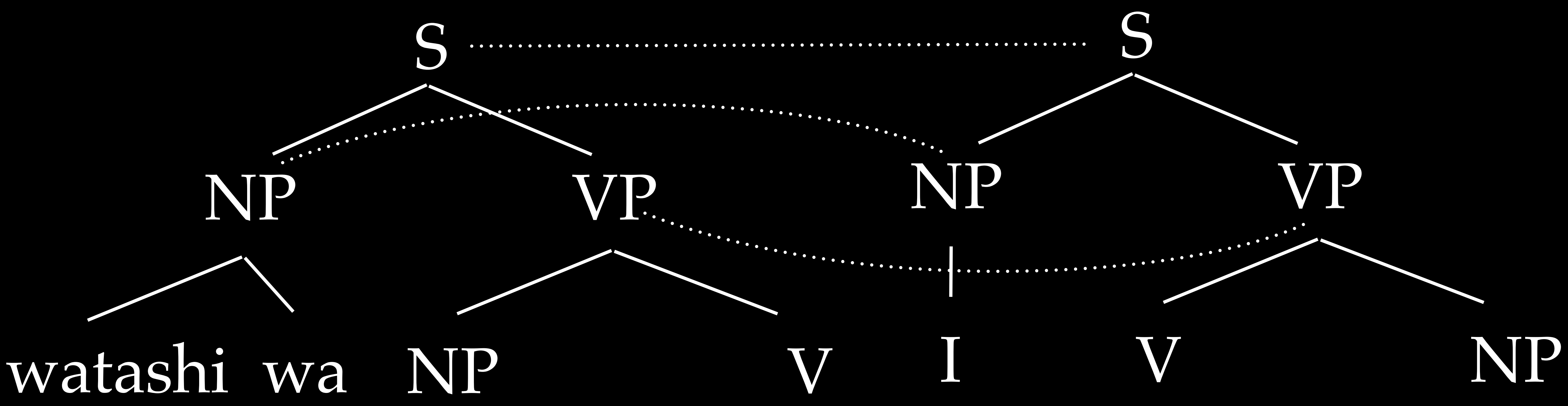
$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$\boxed{VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1}$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar



$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

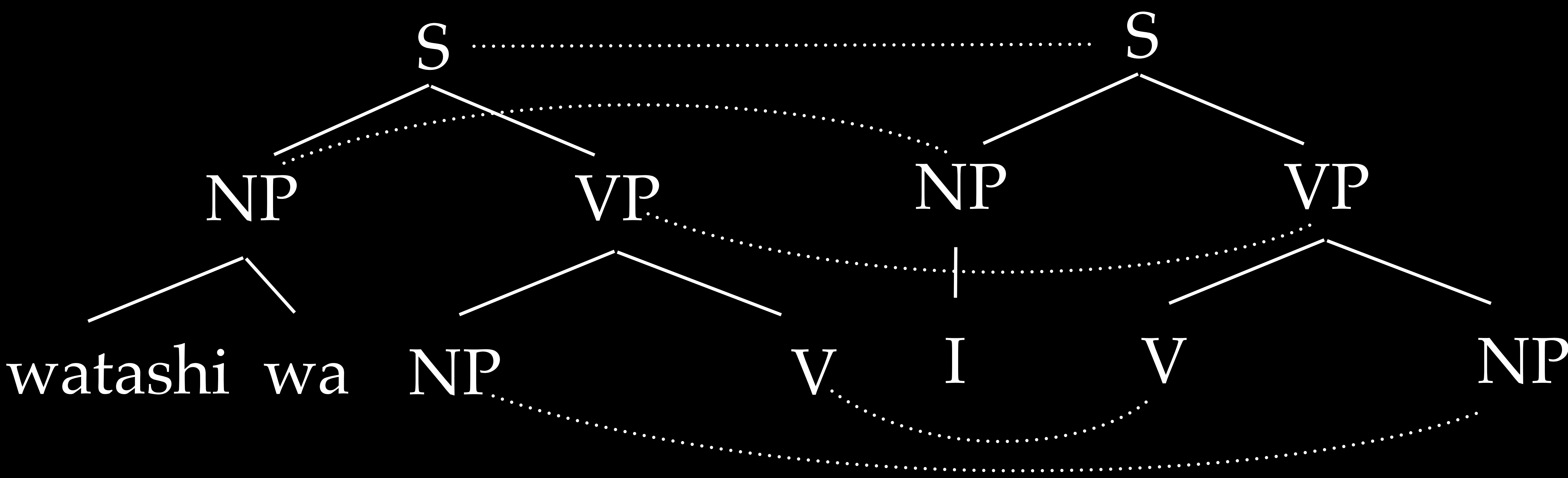
$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar



$\mathrm{S} \rightarrow \mathrm{NP}_1\ \mathrm{VP}_2\ /\ \mathrm{NP}_1\ \mathrm{VP}_2$

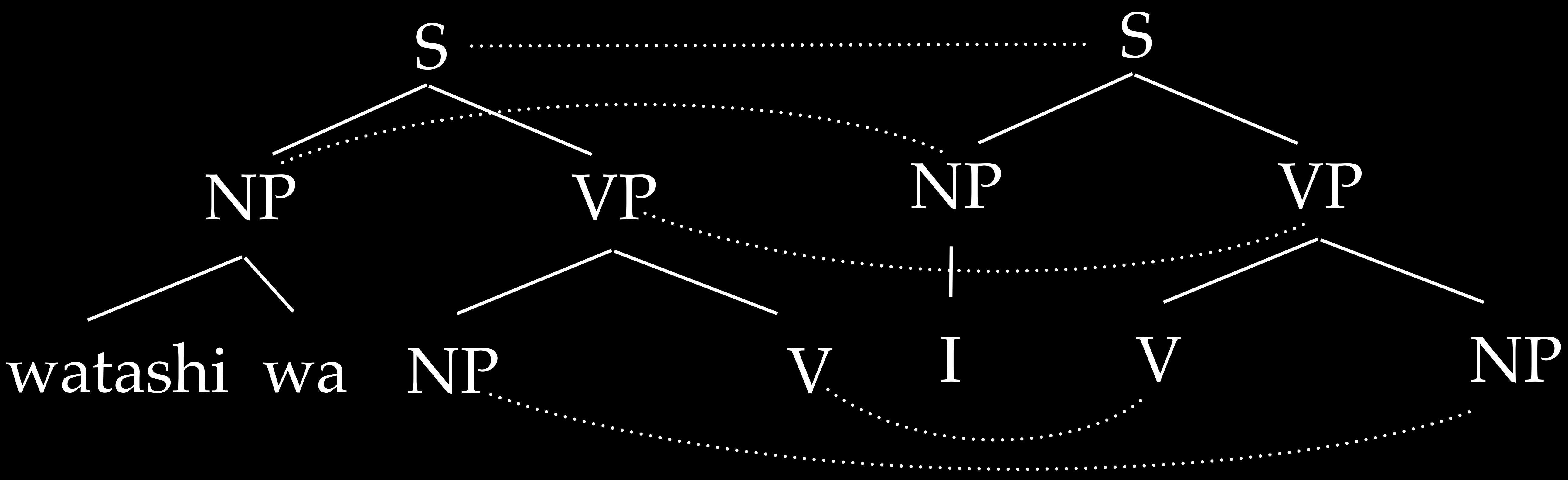
NP → watashi wa / I

NP → hako wo / the box

$\mathrm{VP} \rightarrow \mathrm{NP}_1\ \mathrm{V}_2\ /\ \mathrm{V}_2\ \mathrm{NP}_1$

V → akemasu / open

Synchronous Context-Free Grammar



$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

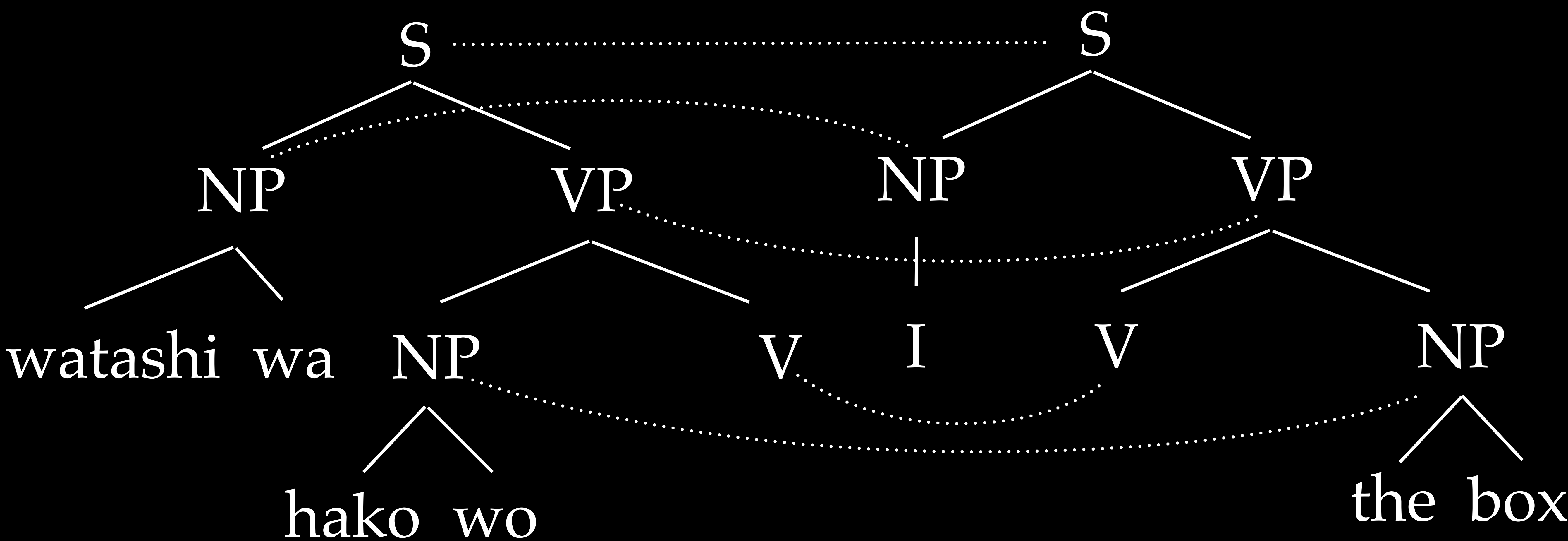
$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar



$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

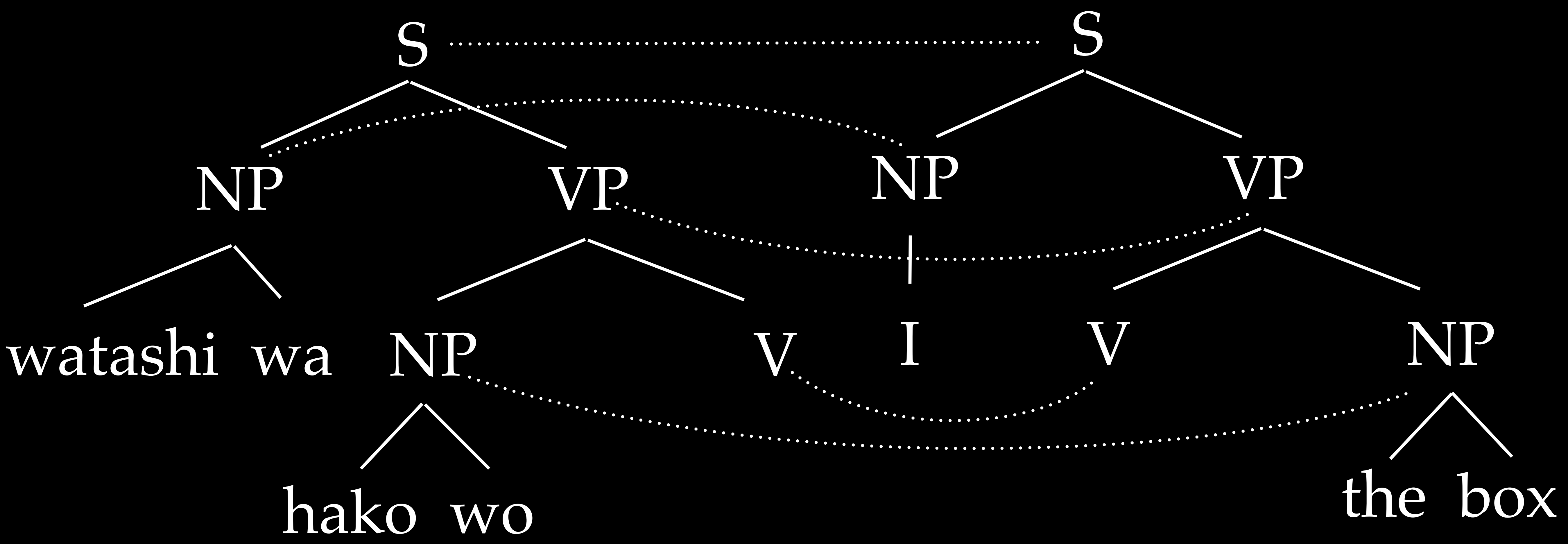
$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar



$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

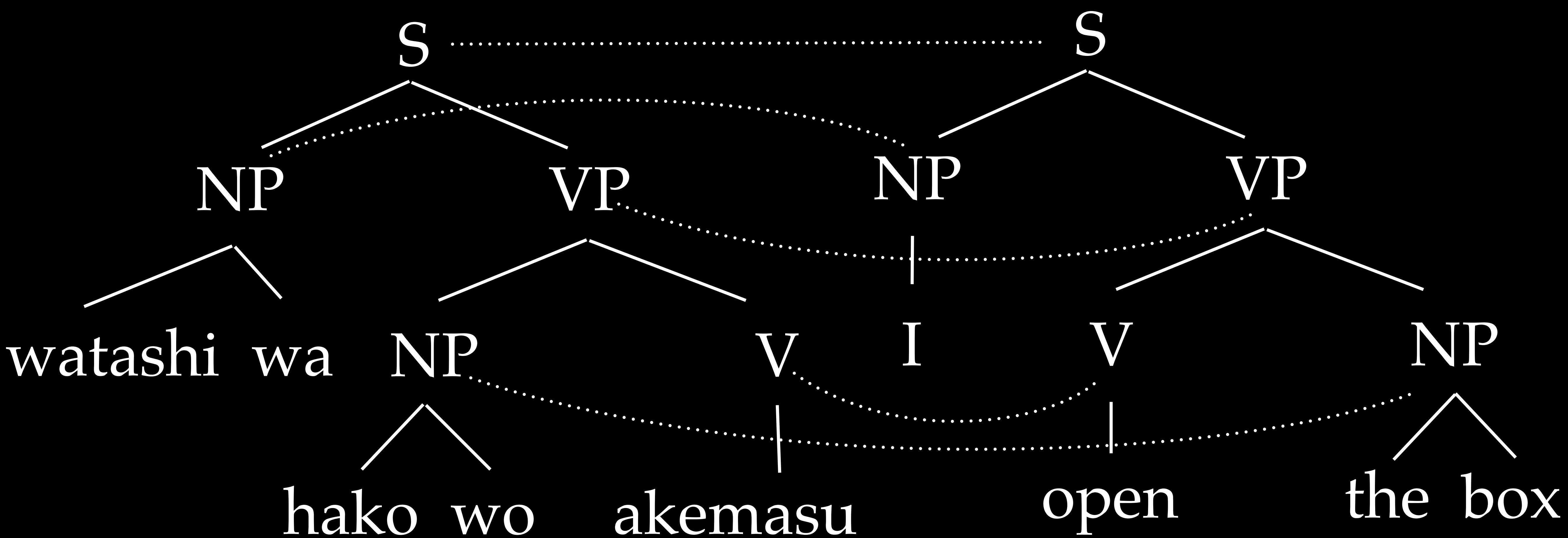
$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar



$S \rightarrow NP_1\ VP_2\ /\ NP_1\ VP_2$

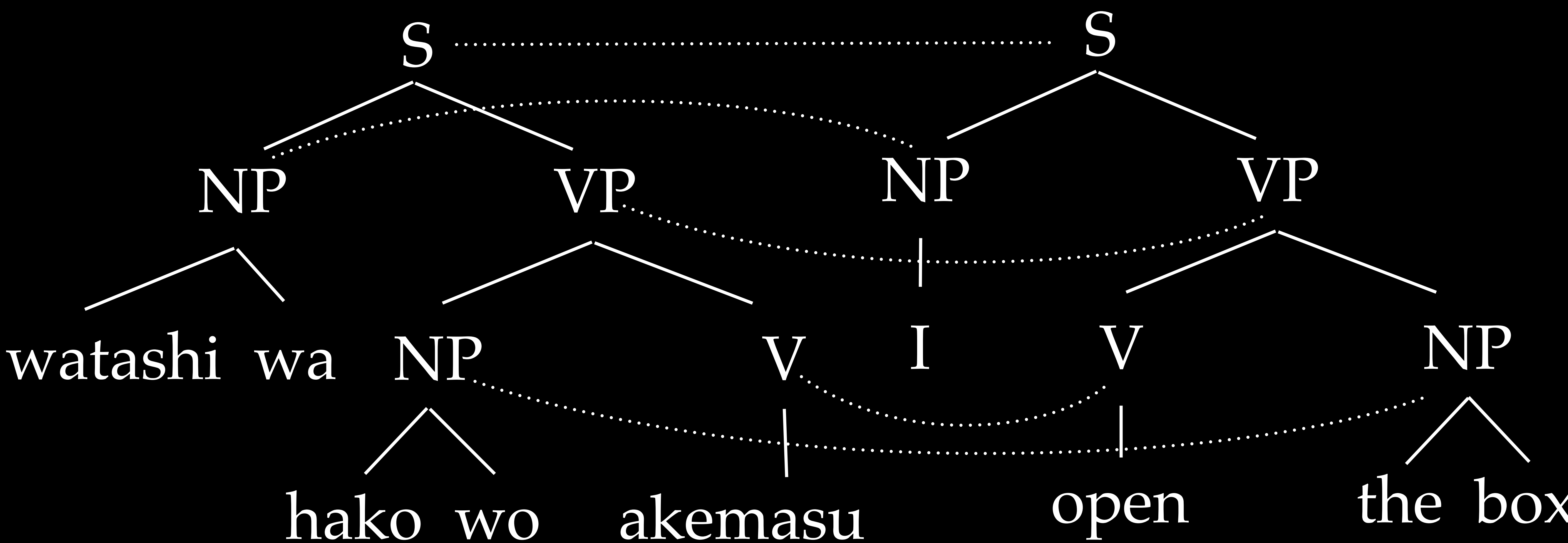
$NP \rightarrow$ watashi wa / I

$NP \rightarrow$ hako wo / the box

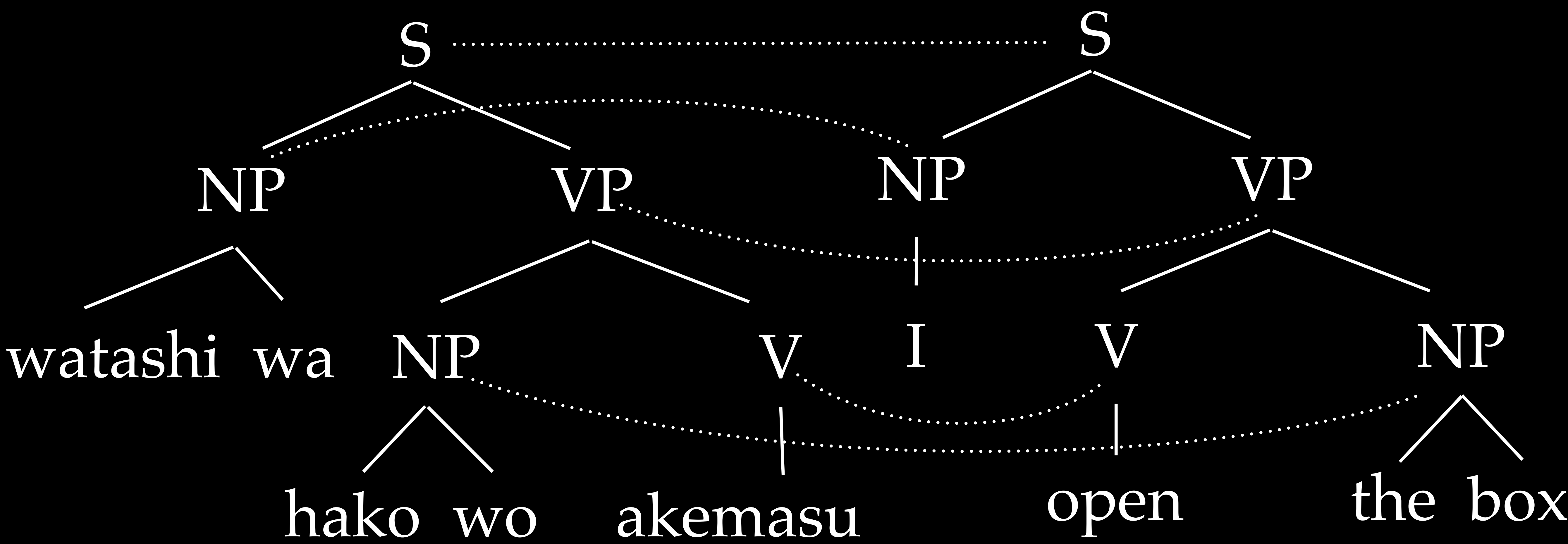
$VP \rightarrow NP_1\ V_2\ /\ V_2\ NP_1$

$V \rightarrow$ akemasu / open

Synchronous Context-Free Grammar

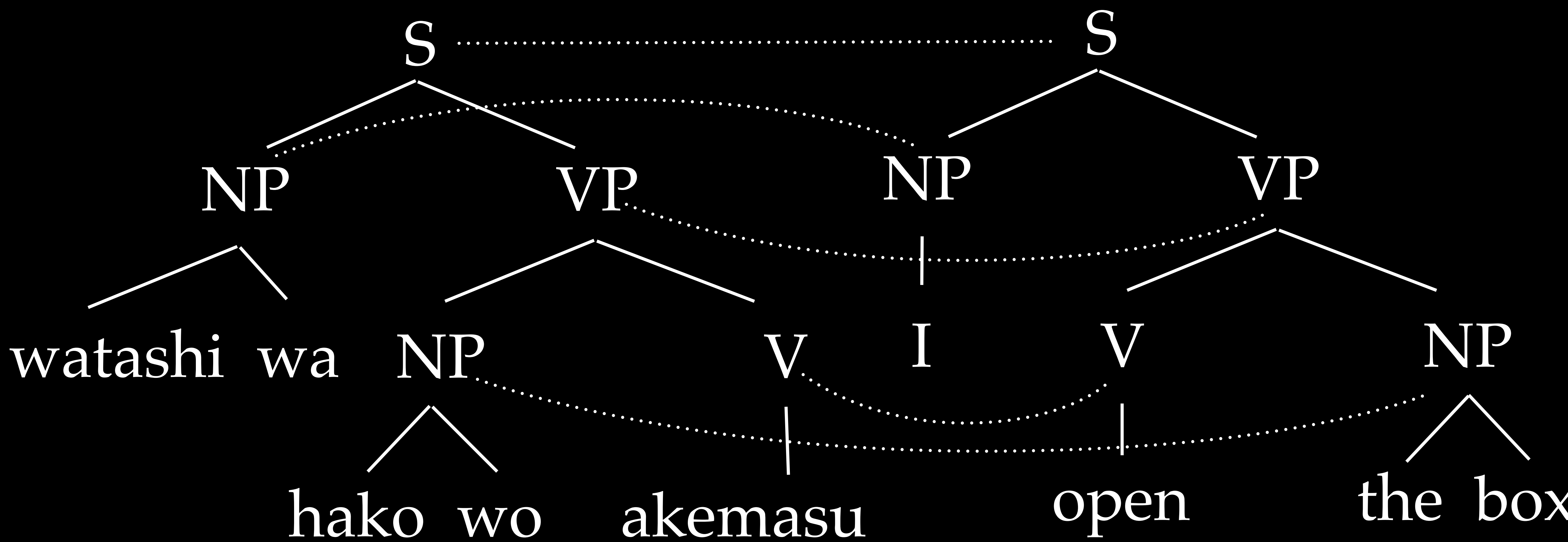


Synchronous Context-Free Grammar



watashi wa hako wo akemasu

Synchronous Context-Free Grammar



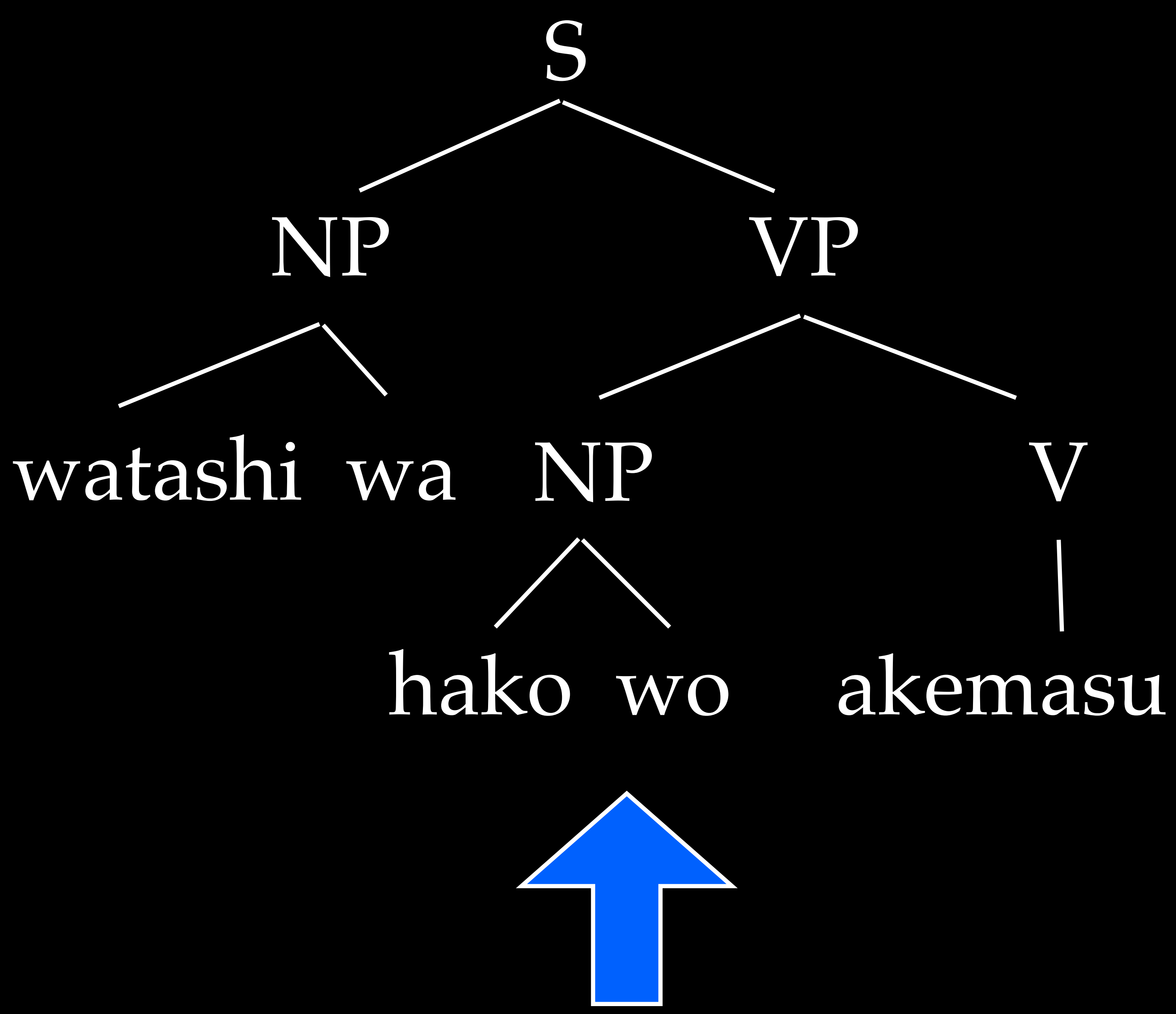
watashi wa hako wo akemasu

I open the box

Translation is Parsing

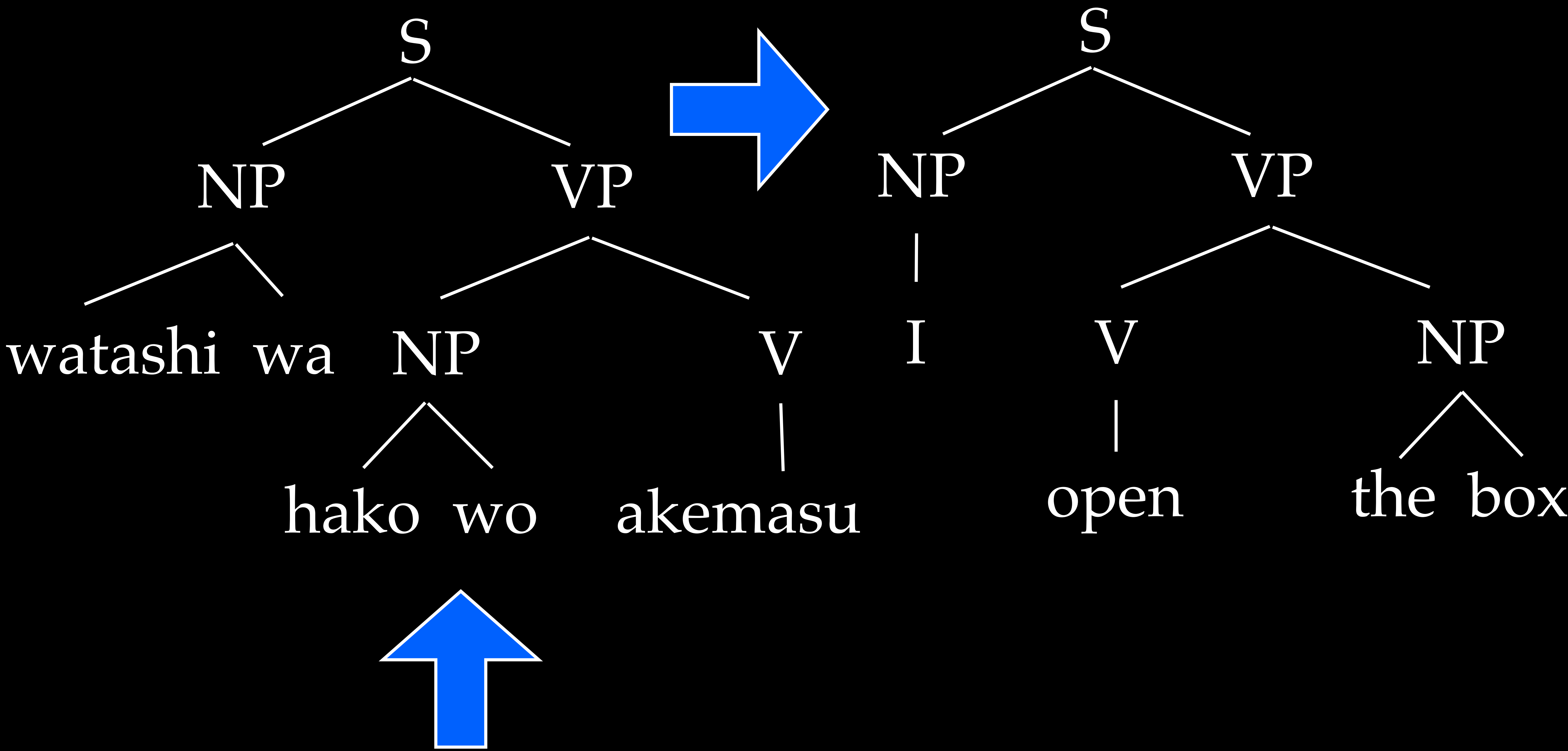
watashi wa hako wo akemasu

Translation is Parsing



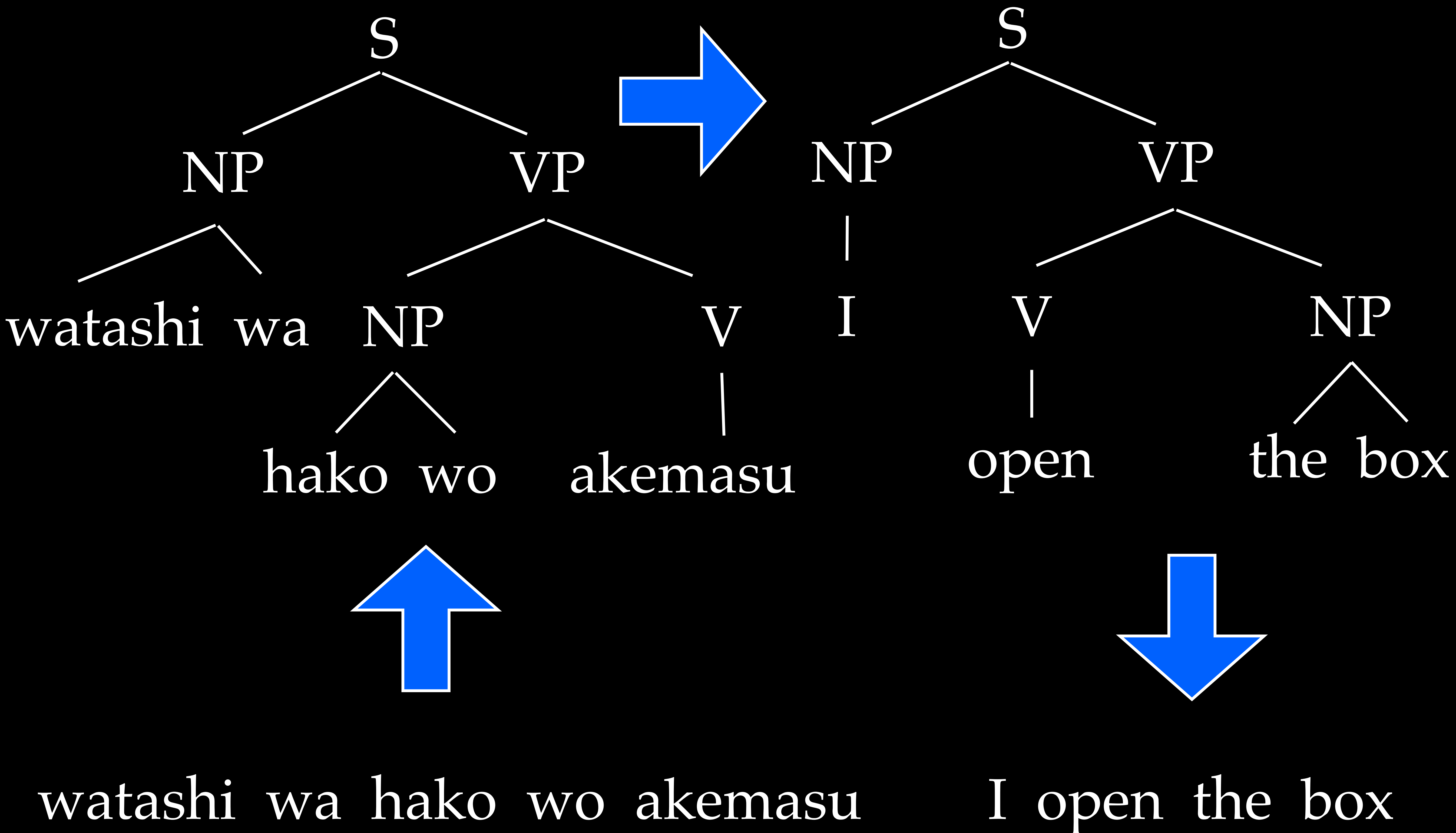
watashi wa hako wo akemasu

Translation is Parsing



watashi wa hako wo akemasu

Translation is Parsing



Translation is Parsing

Translation is Parsing

- How many parses of a sentence are there?

Translation is Parsing

- How many parses of a sentence are there?
  - For binary grammar: Catalan number.

Translation is Parsing

- How many parses of a sentence are there?
  - For binary grammar: Catalan number. $O(\frac{(2n)!}{(n+1)!n!})$

Translation is Parsing

- How many parses of a sentence are there?
  - For binary grammar: Catalan number. $O(\frac{(2n)!}{(n+1)!n!})$
- Dynamic programming to the rescue!

Parsing

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

I_1 saw_2 her_3 duck_4

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

I_1 saw_2 her_3 duck_4

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$\boxed{\mathrm{I}_1}$ saw_2 her_3 duck_4

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

I_1 saw_2 her_3 duck_4

Parsing

NN → duck

NP → PRP$ NN

PRP → her

$\boxed{\text{PRP} \rightarrow \text{I}}$

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$\boxed{PRP_{0,1} \leftarrow (w_1 = \mathrm{I}) \wedge (PRP \rightarrow \mathrm{I})}$$

$\boxed{\mathrm{I}_1}$ saw_2 her_3 duck_4

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$\boxed{PRP_{0,1} \leftarrow (w_1 = \mathrm{I}) \wedge (PRP \rightarrow \mathrm{I})}$$

$\boxed{PRP_{0,1}}$

$\boxed{\mathrm{I}_1}$ saw_2 her_3 duck_4

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$PRP_{0,1}$$

$\uparrow$

I$_1$ saw$_2$ her$_3$ duck$_4$

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$VBD_{1,2}$

$PRP_{0,1}$

I$_1$ saw$_2$ her$_3$ duck$_4$

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$VBD_{1,2}$

$PRP_{0,1}$ $PRP\$_{2,3}$

I_1 saw_2 her_3 duck_4

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$VBD_{1,2}$ $PRP_{2,3}$

$PRP_{0,1}$ $PRP\$_{2,3}$

I_1 saw_2 her_3 duck_4

Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

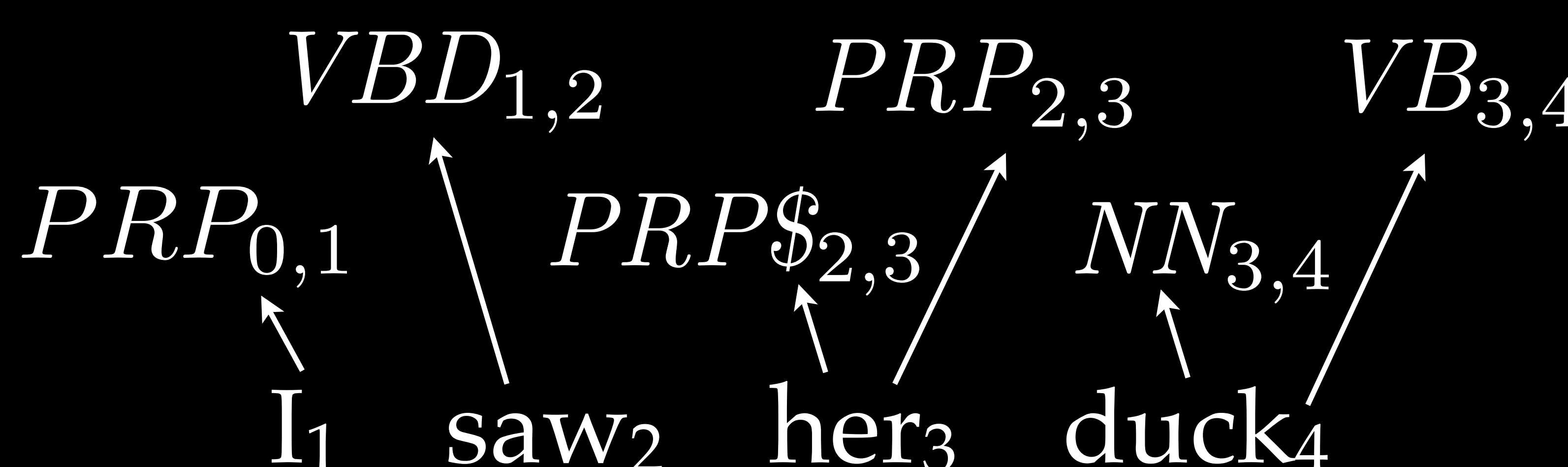
VB → duck

VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

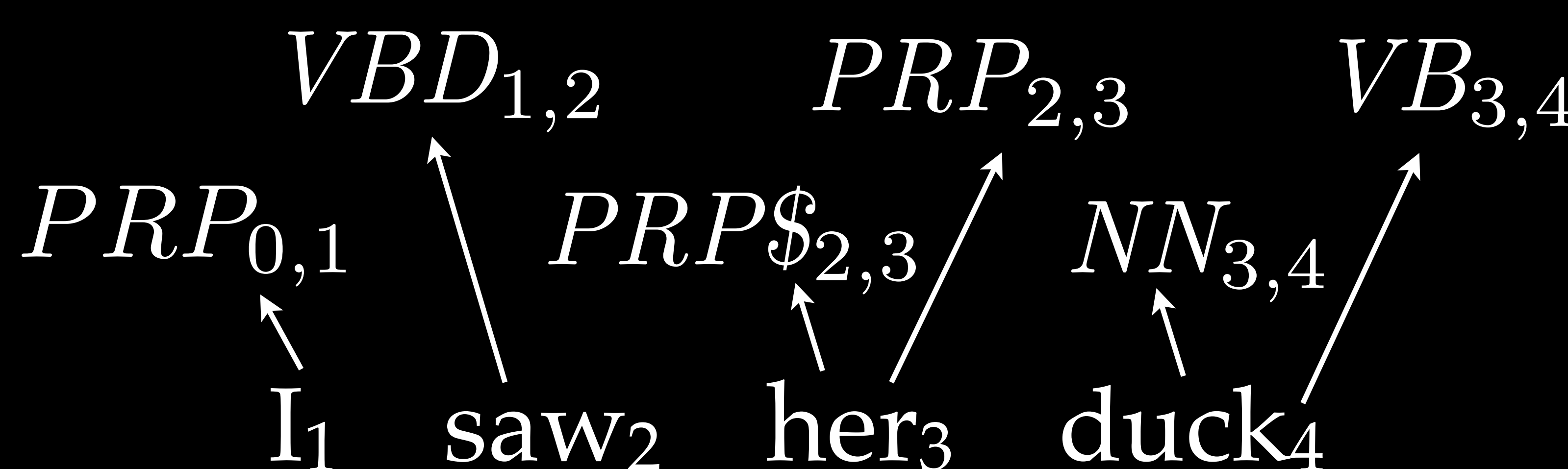
VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow YZ)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

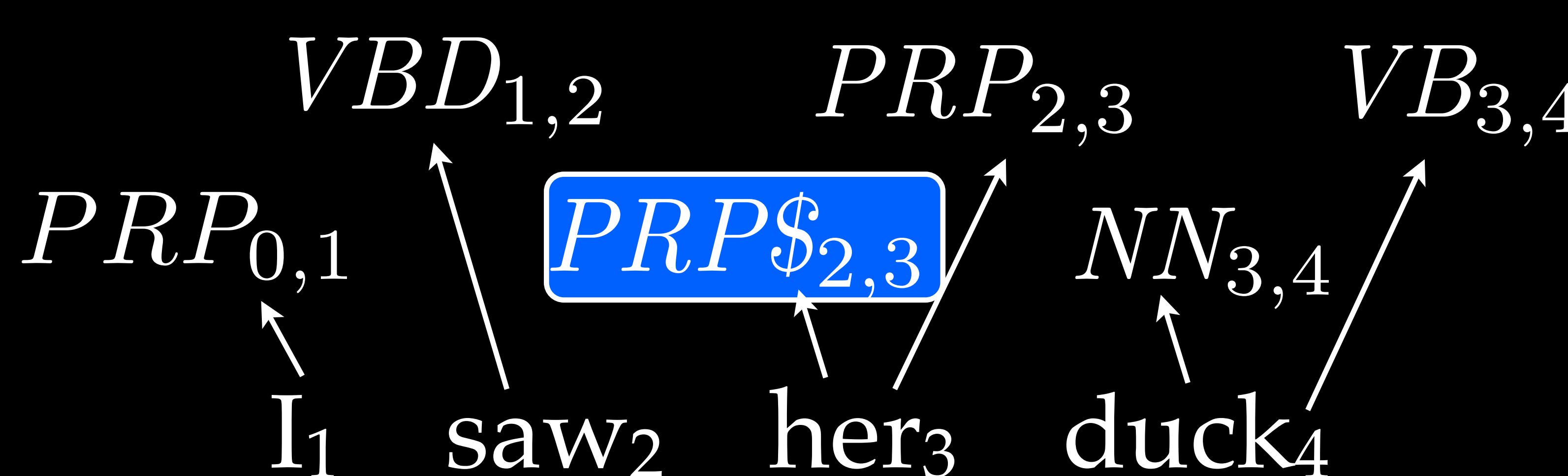
VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y Z)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

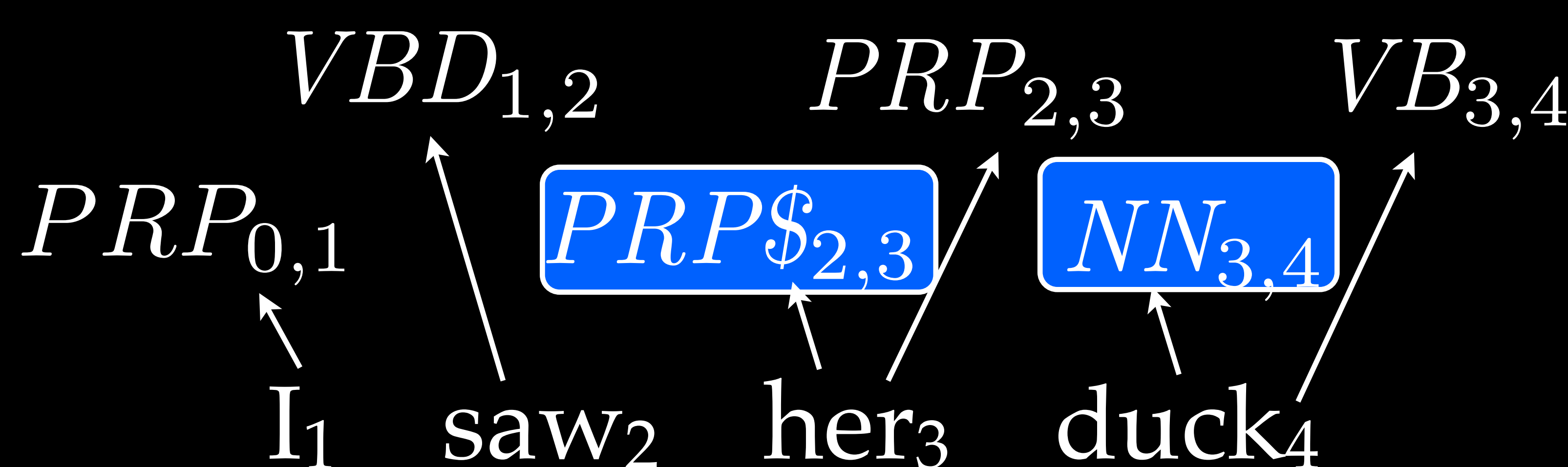
VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow YZ)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

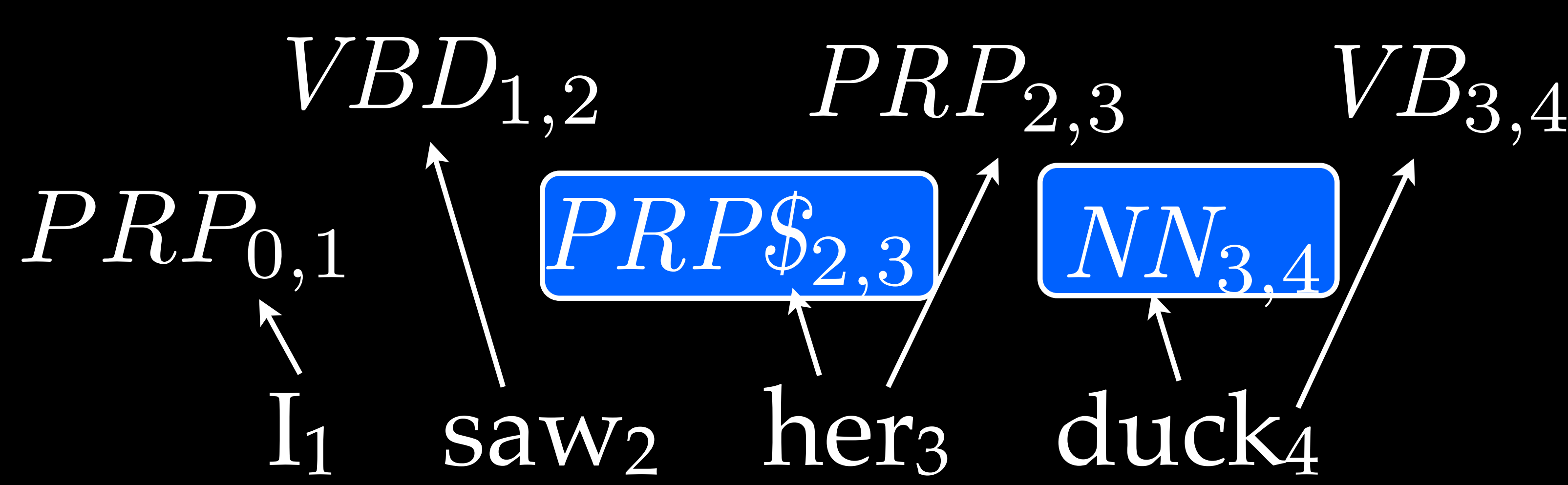
VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y Z)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

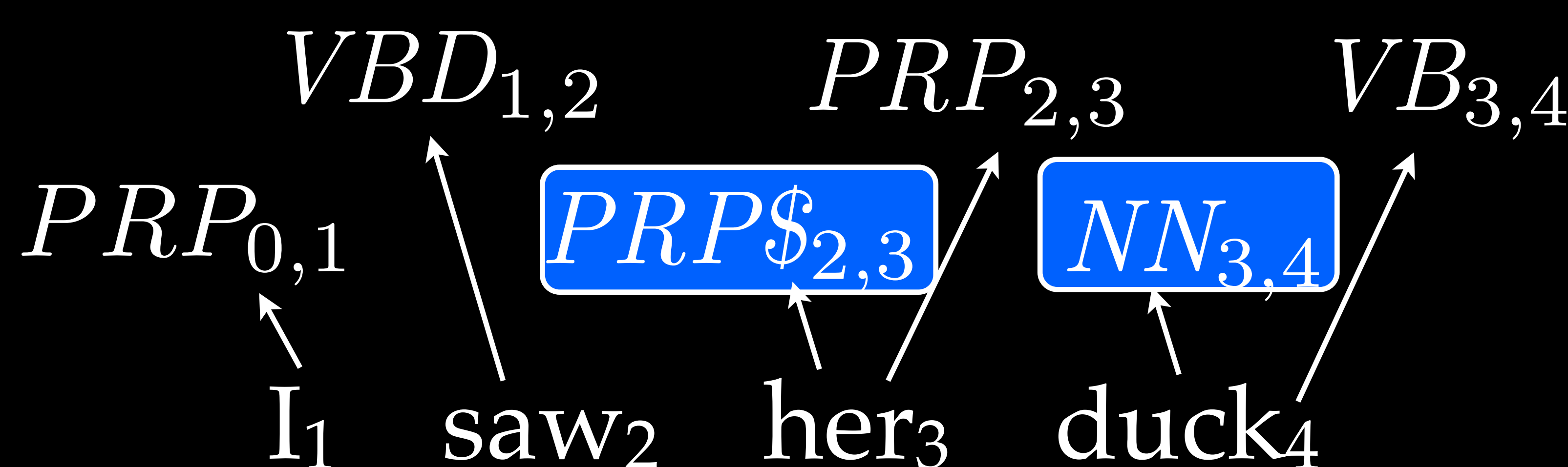
VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y Z)$$

$$NP_{2,4} \leftarrow PRP\$_{2,3} \wedge NN_{3,4} \wedge (NP \rightarrow PRP\$\ NN)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

VP → VBD NP

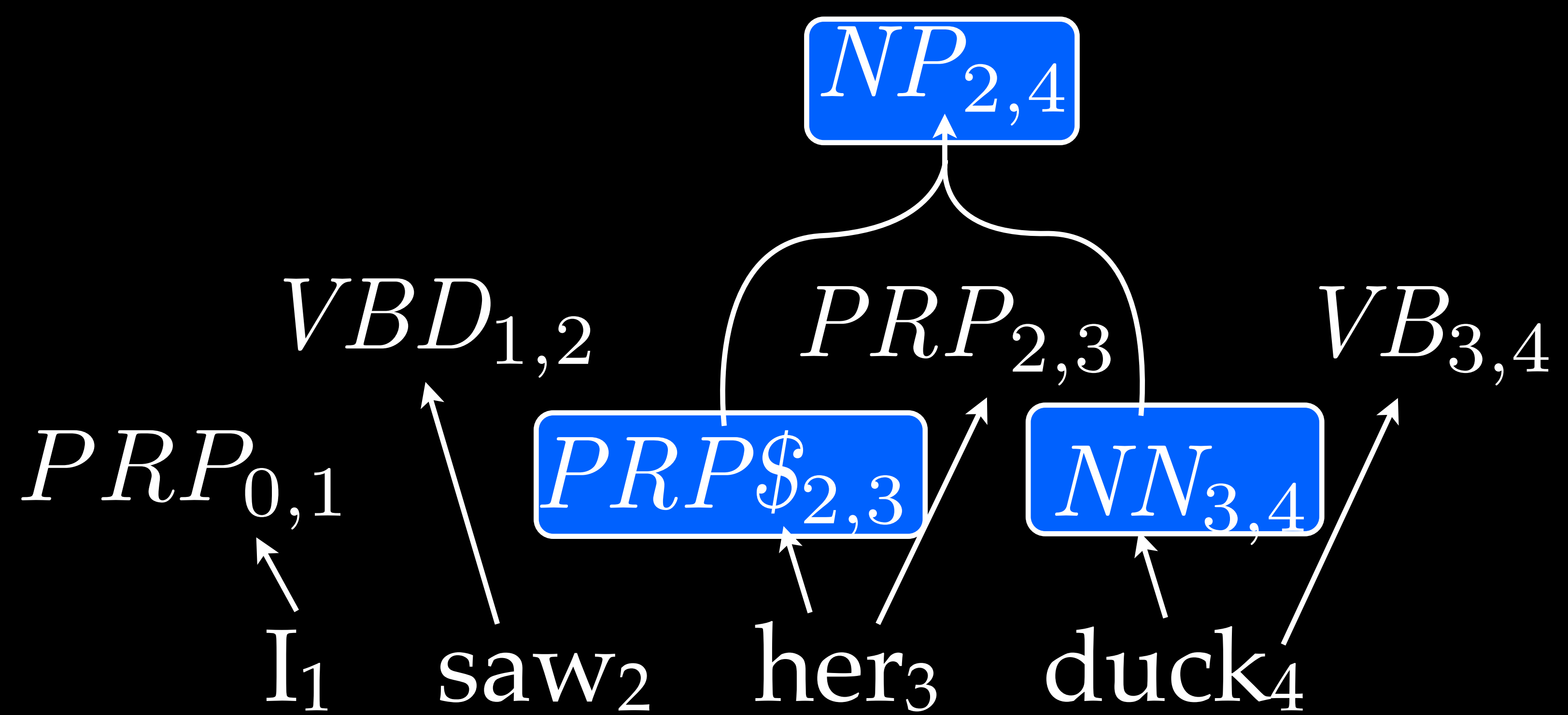
VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y Z)$$

$$NP_{2,4} \leftarrow PRP\$_{2,3} \wedge NN_{3,4} \wedge (NP \rightarrow PRP\$\ NN)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

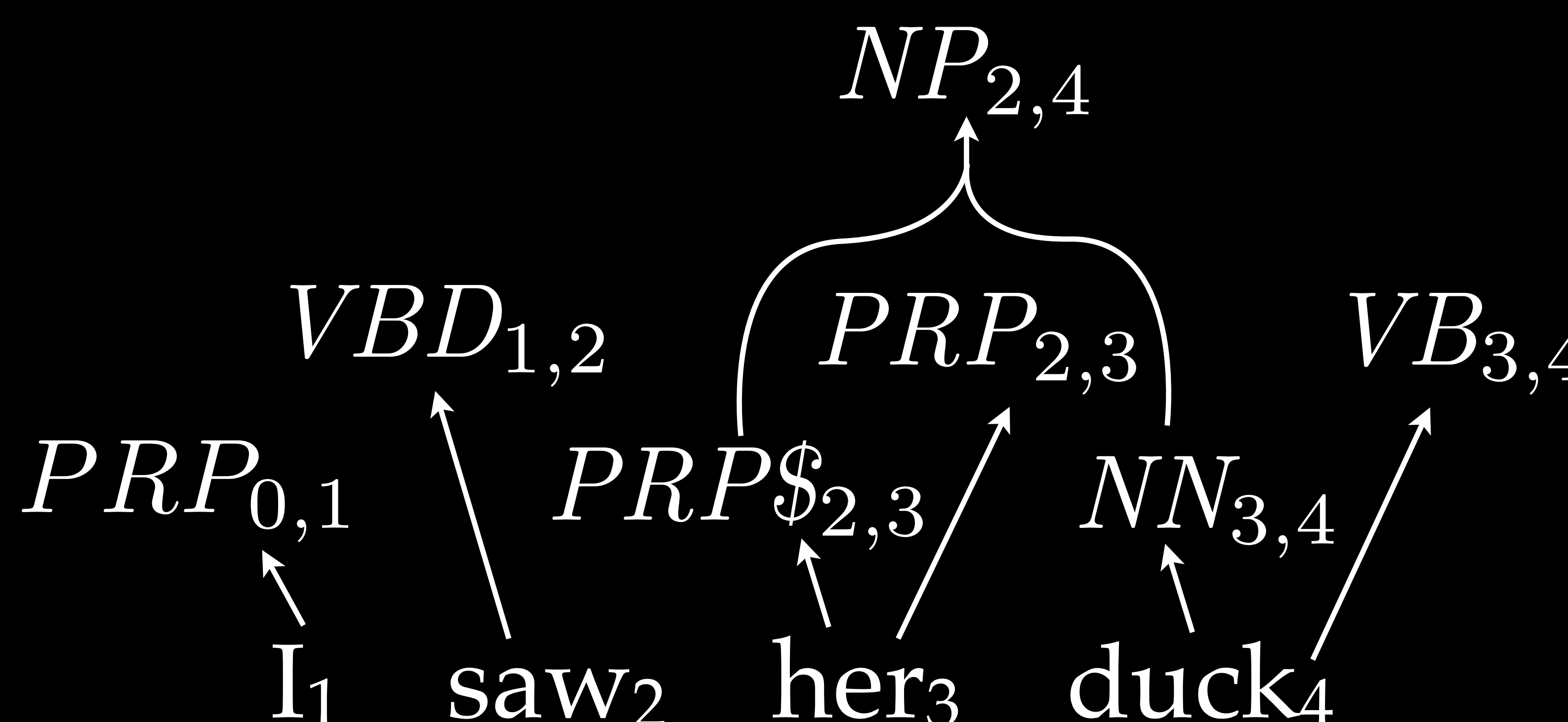
VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y Z)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

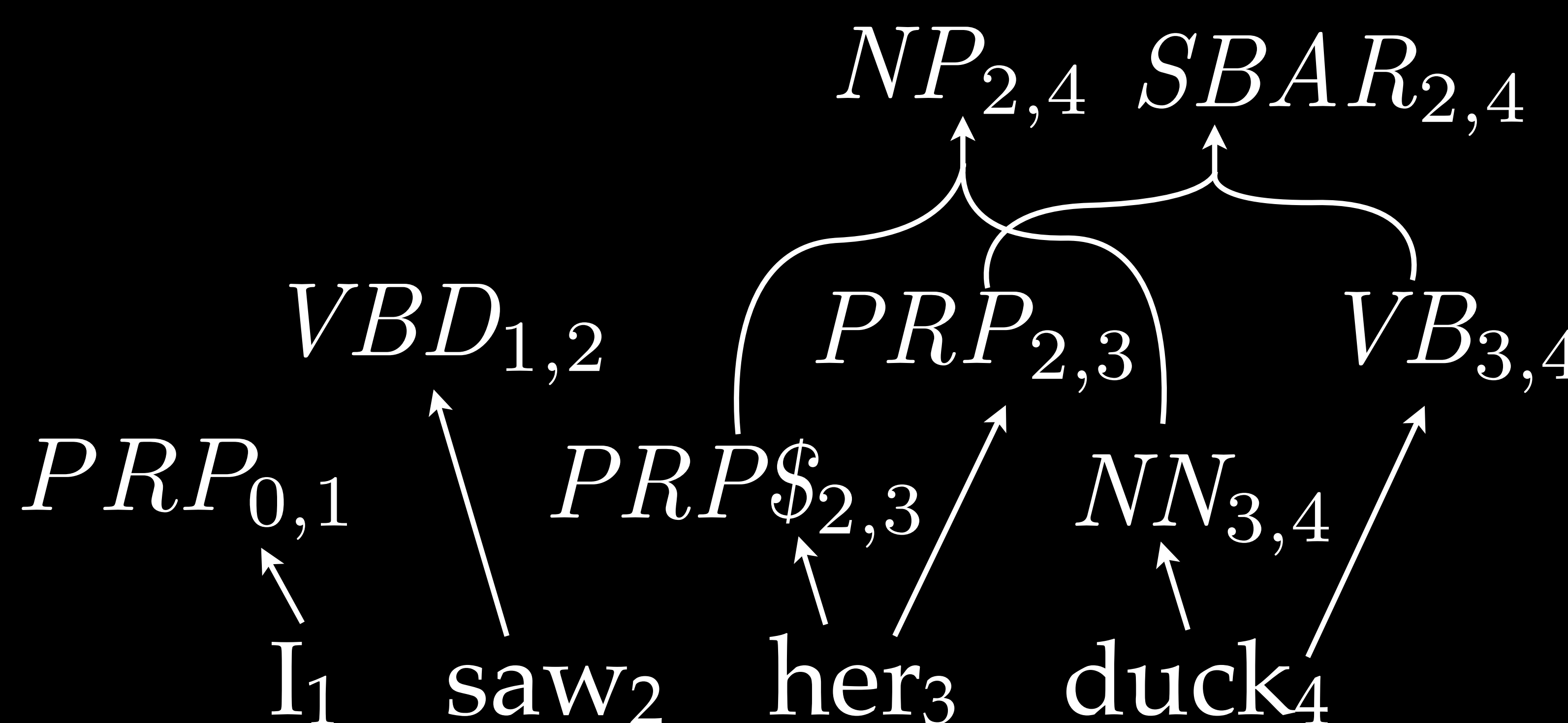
VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow YZ)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

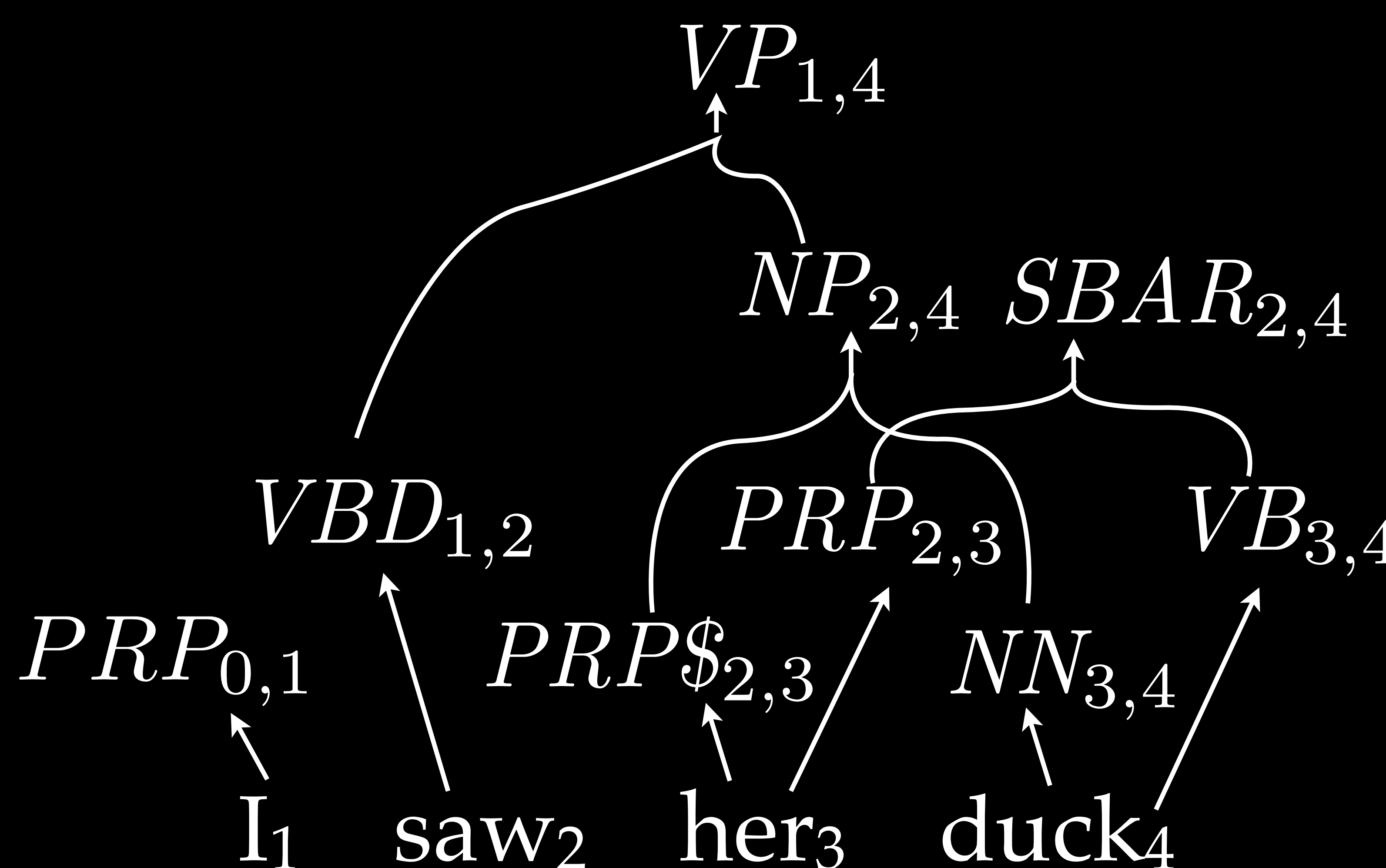
VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow YZ)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

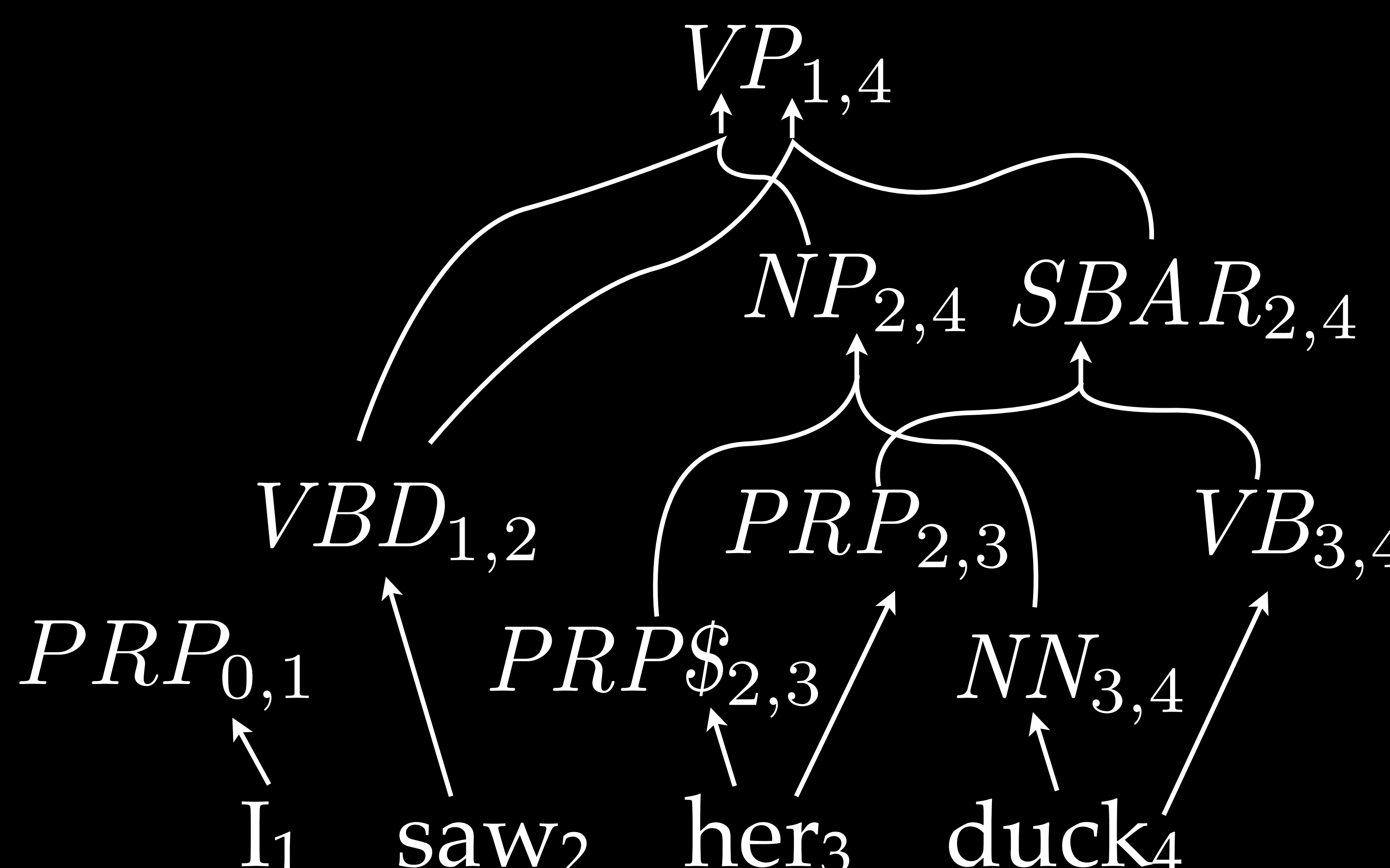
VP → VBD NP

VP → VBD SBAR

VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow YZ)$$



Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

SBAR → PRP VB

VB → duck

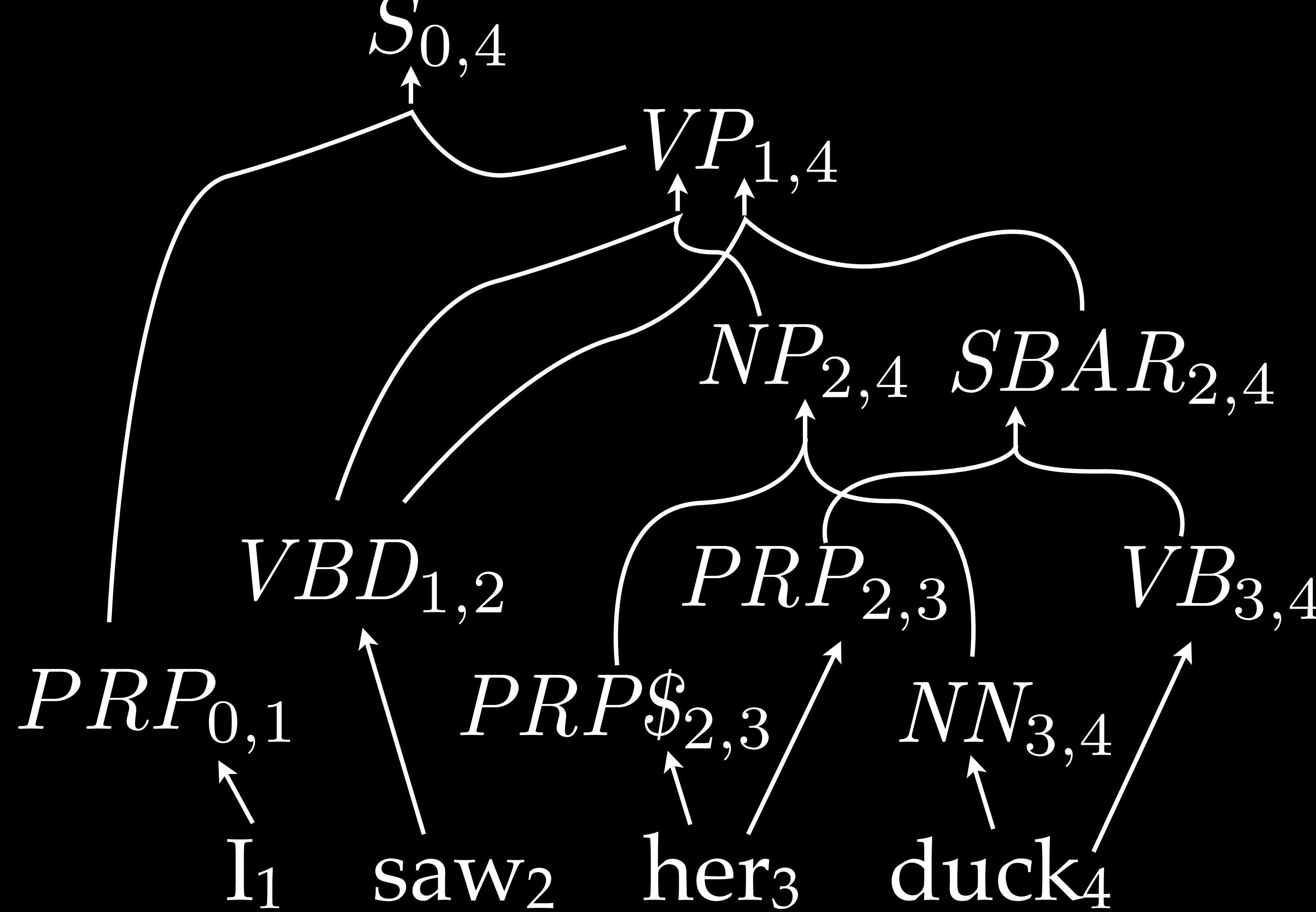
VP → VBD NP

VP → VBD SBAR

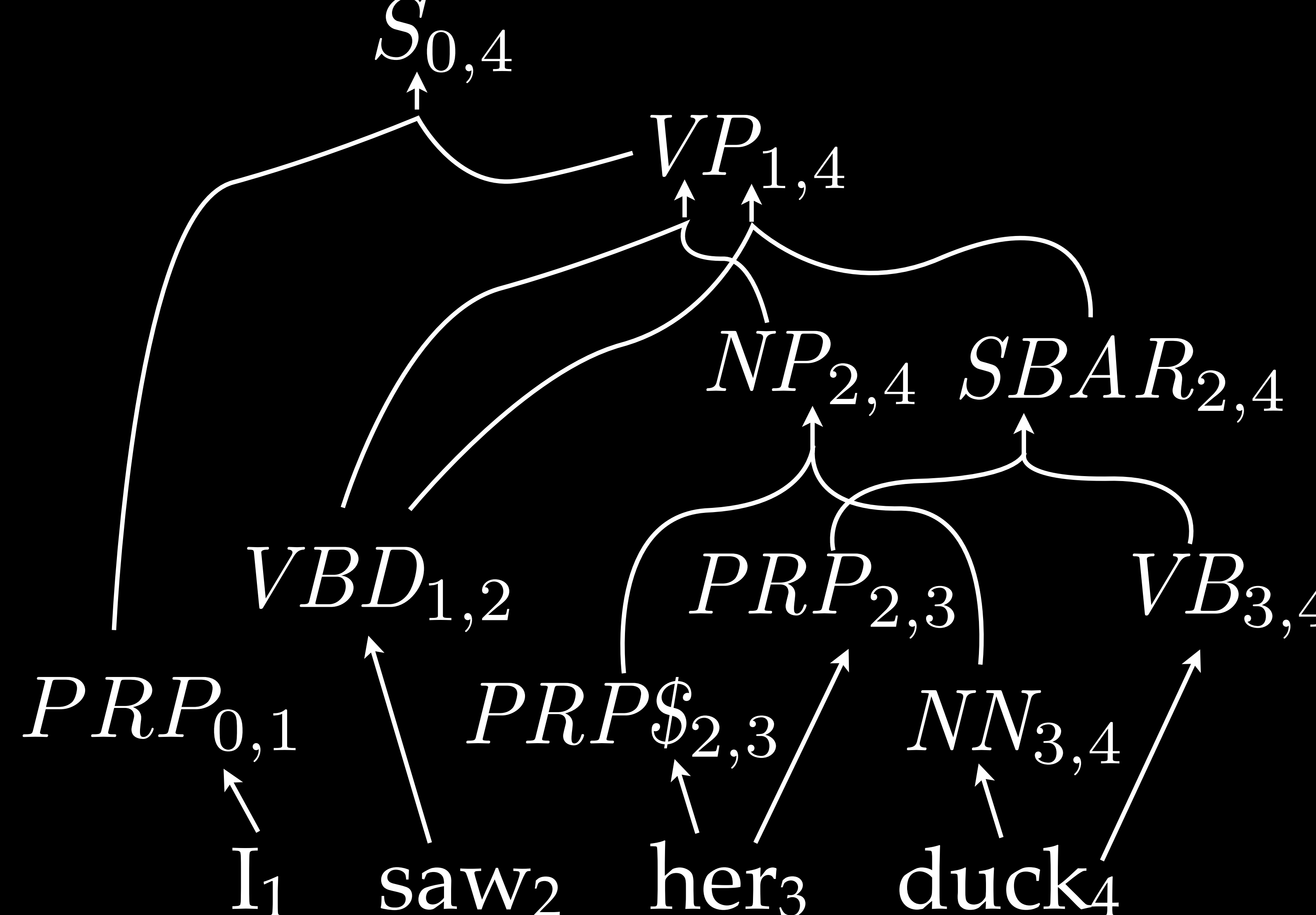
VBD → saw

$$X_{i,i+1} \leftarrow (w_{i+1} = w) \wedge (X \rightarrow w)$$

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow YZ)$$

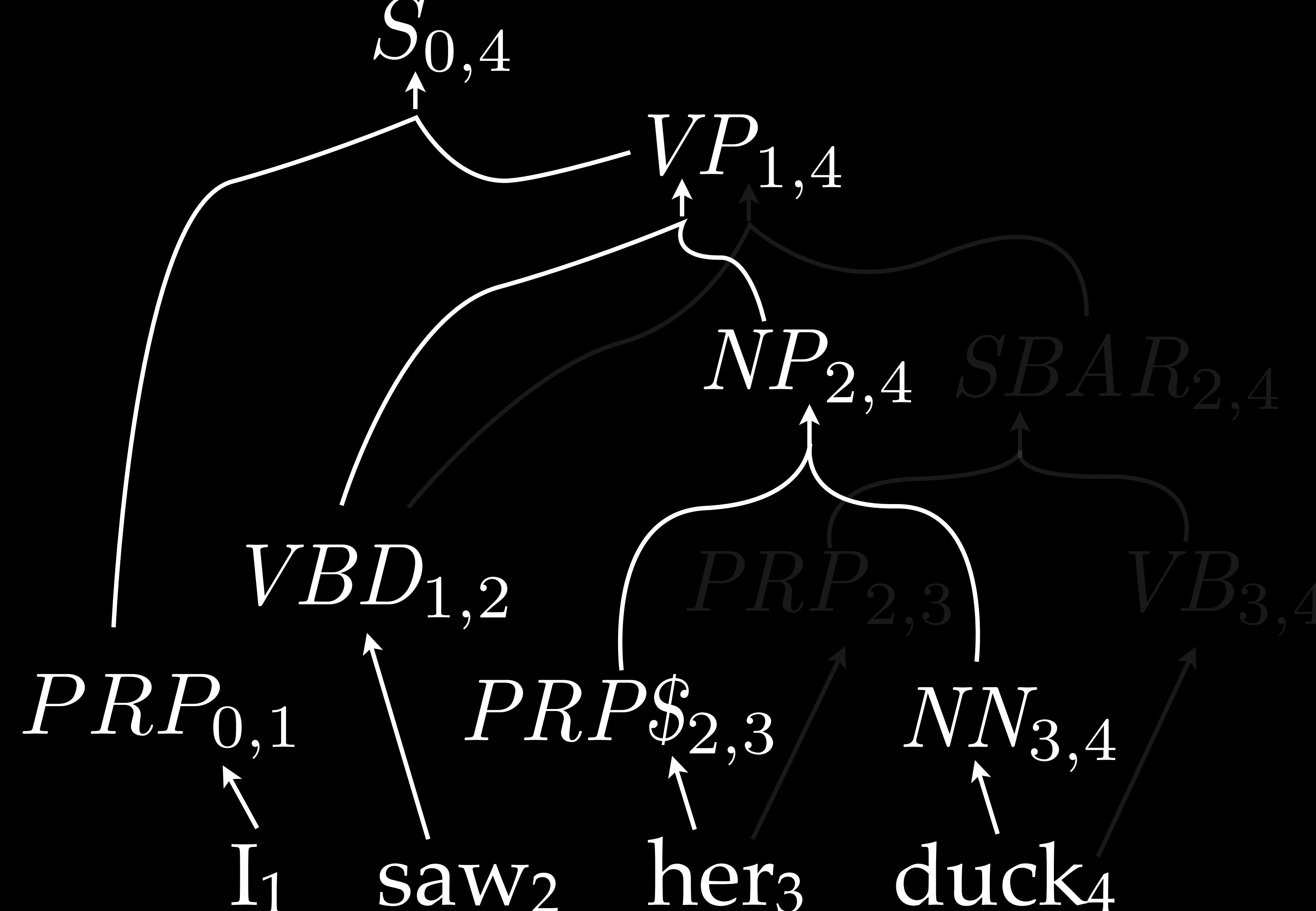


Parsing



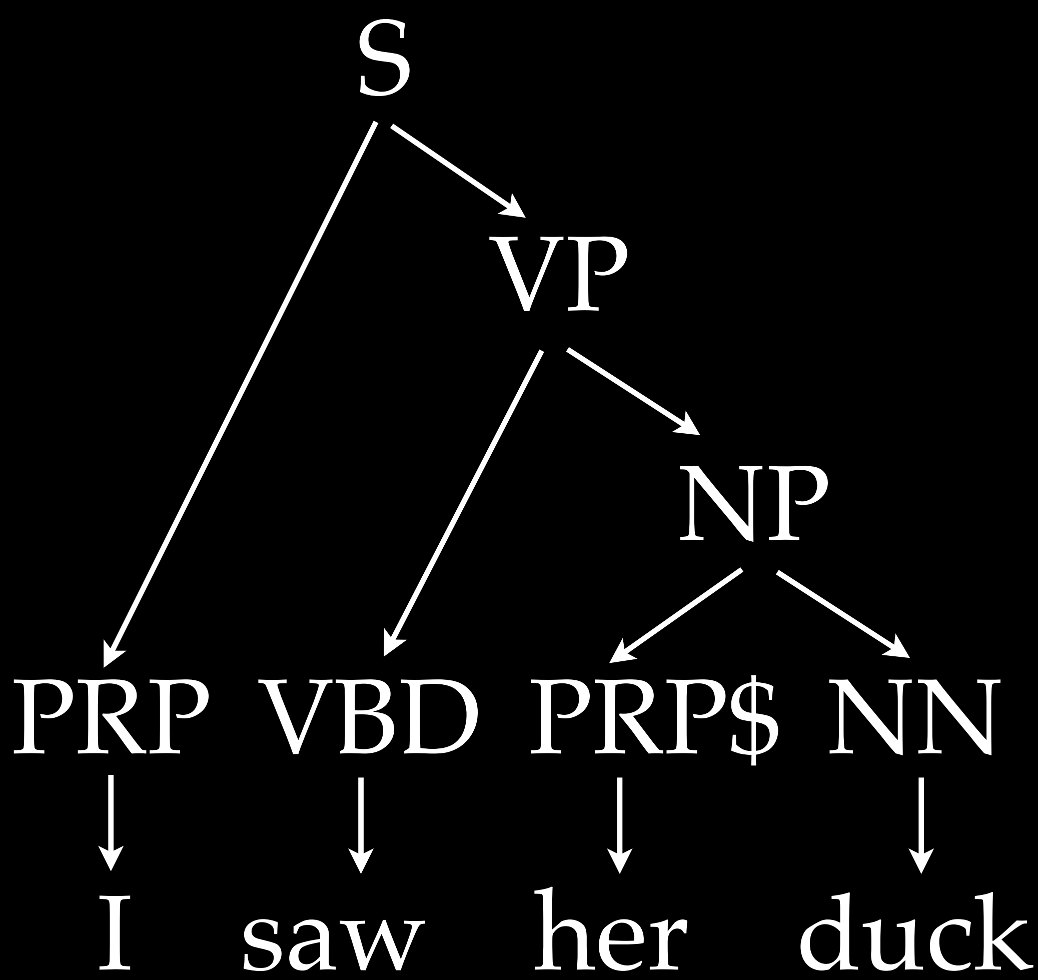
$S_{0,4}$
$VP_{1,4}$
$NP_{2,4}$ $SBAR_{2,4}$
$VBD_{1,2}$
$PRP_{2,3}$
$VB_{3,4}$
$PRP_{0,1}$
$PRP\$_{2,3}$
$NN_{3,4}$
I₁ saw₂ her₃ duck₄

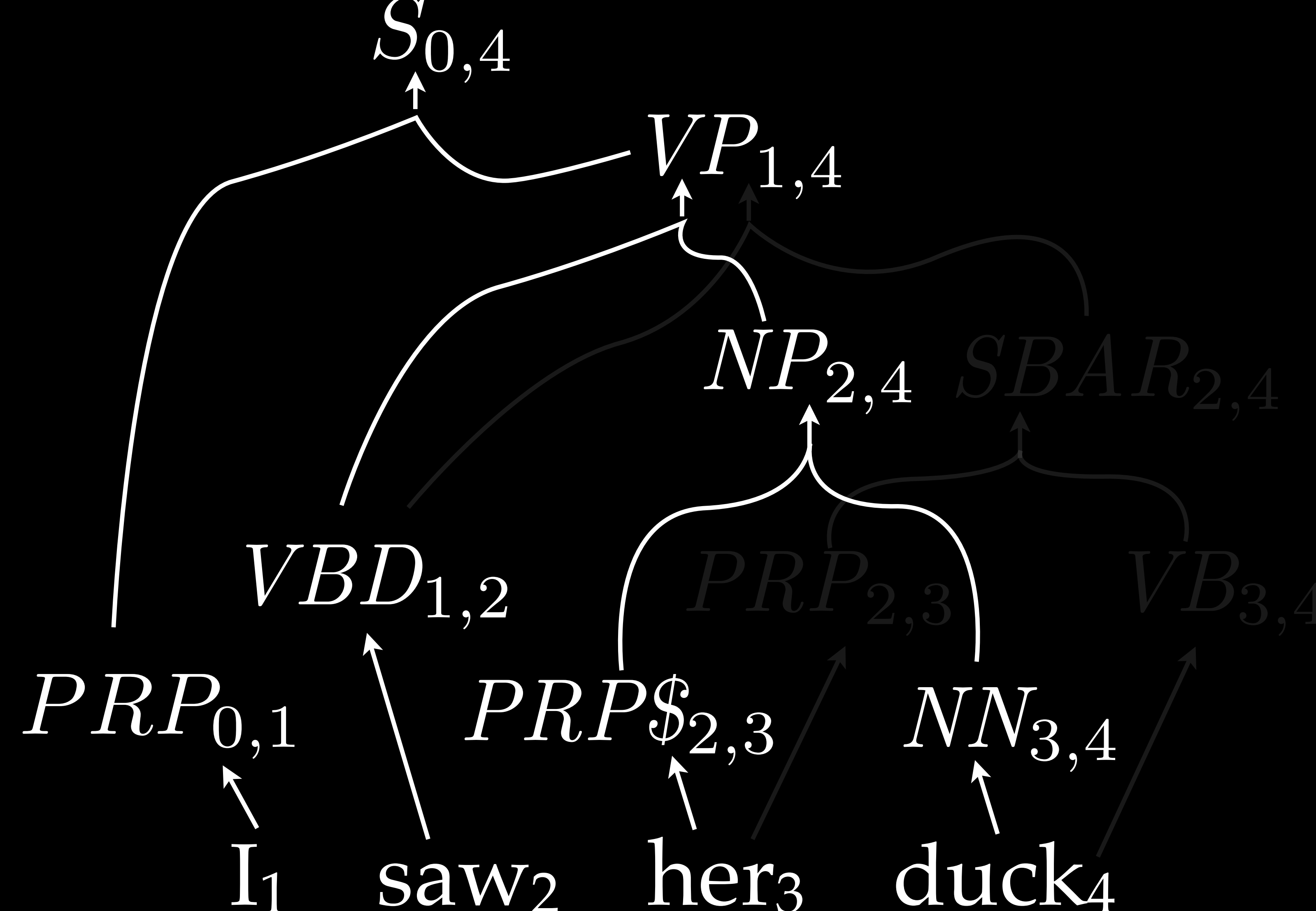
Parsing



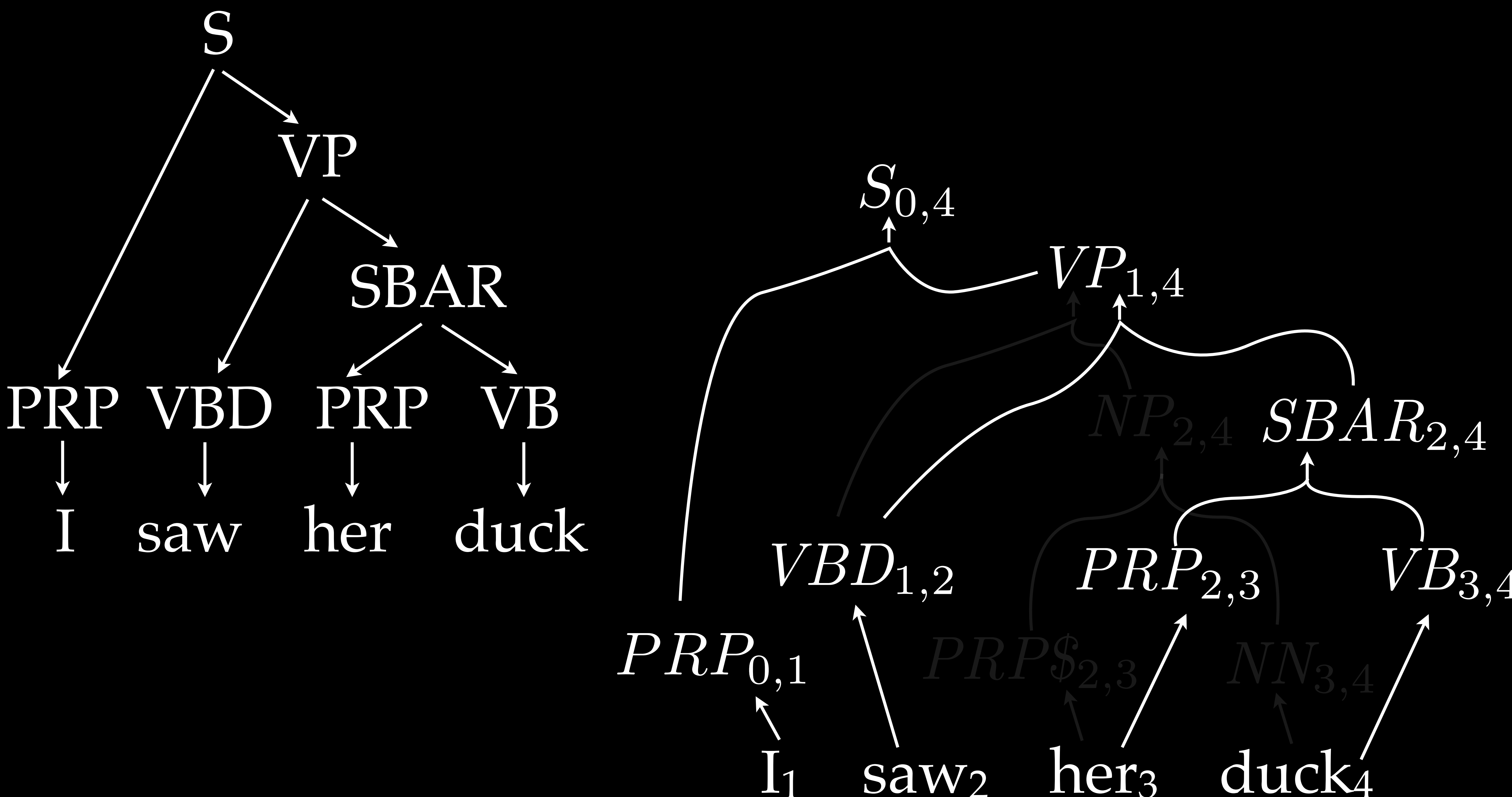
$S_{0,4}$
$VP_{1,4}$
$NP_{2,4}$
$SBAR_{2,4}$
$VBD_{1,2}$
$PRP_{2,3}$
$VB_{3,4}$
$PRP_{0,1}$
$PRP\$_{2,3}$
$NN_{3,4}$
I1 saw2 her3 duck4

Parsing


S
VP
NP
PRP VBD PRP$ NN
I saw her duck

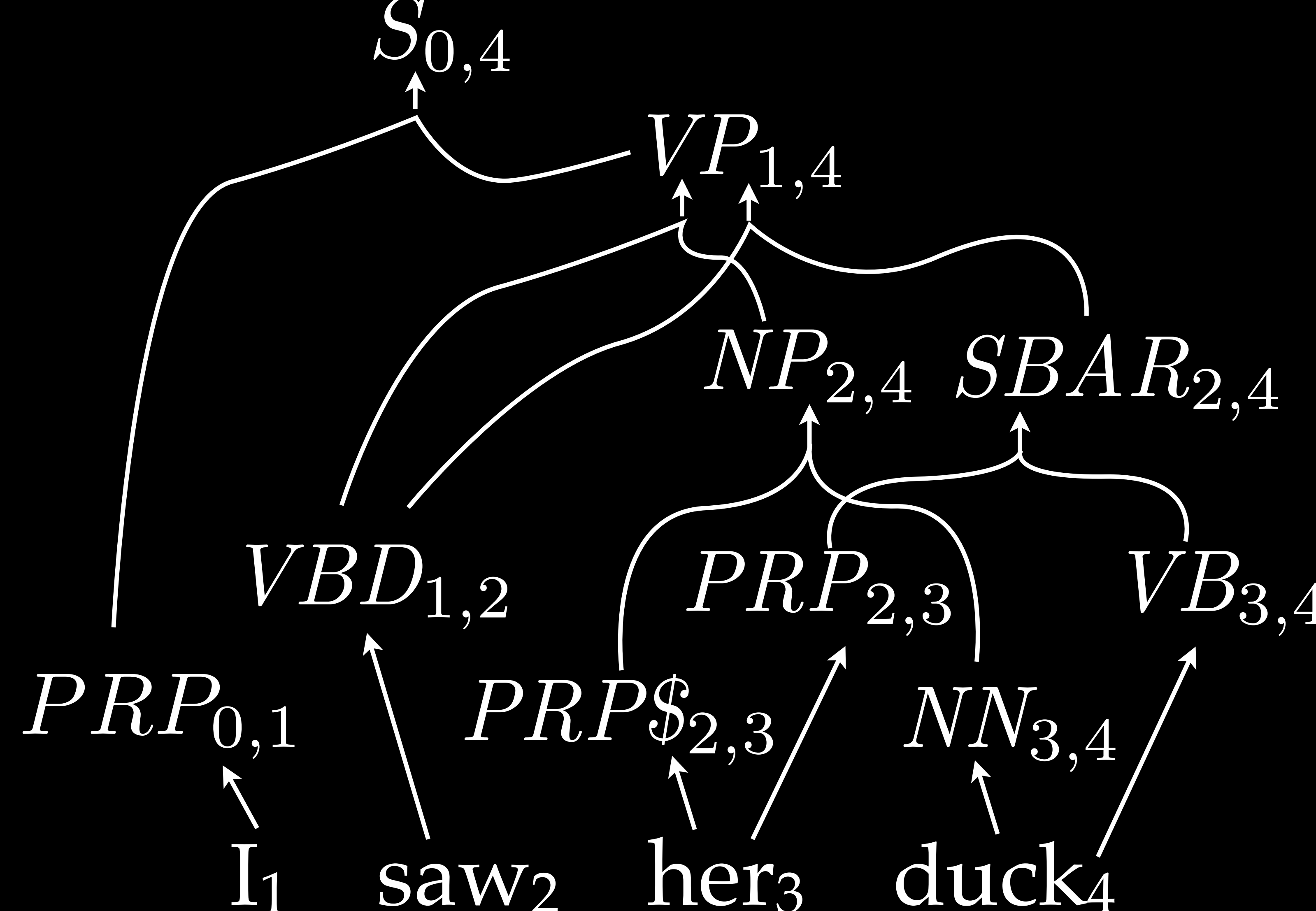

$S_{0,4}$
$VP_{1,4}$
$NP_{2,4}$ $SBAR_{2,4}$
$VBD_{1,2}$ $PRP_{2,3}$ $VB_{3,4}$
$PRP_{0,1}$ $PRP\$_{2,3}$ $NN_{3,4}$
I_1 saw_2 her_3 $duck_4$

Parsing


S
VP
SBAR
PRP VBD PRP VB
I saw her duck
$S_{0,4}$
$VP_{1,4}$
$NP_{2,4}$
$SBAR_{2,4}$
$VBD_{1,2}$
$PRP_{2,3}$
$VB_{3,4}$
$PRP_{0,1}$
$PRP\$_{2,3}$
$NN_{3,4}$
I_1 saw_2 her_3 $duck_4$

Parsing

Analysis

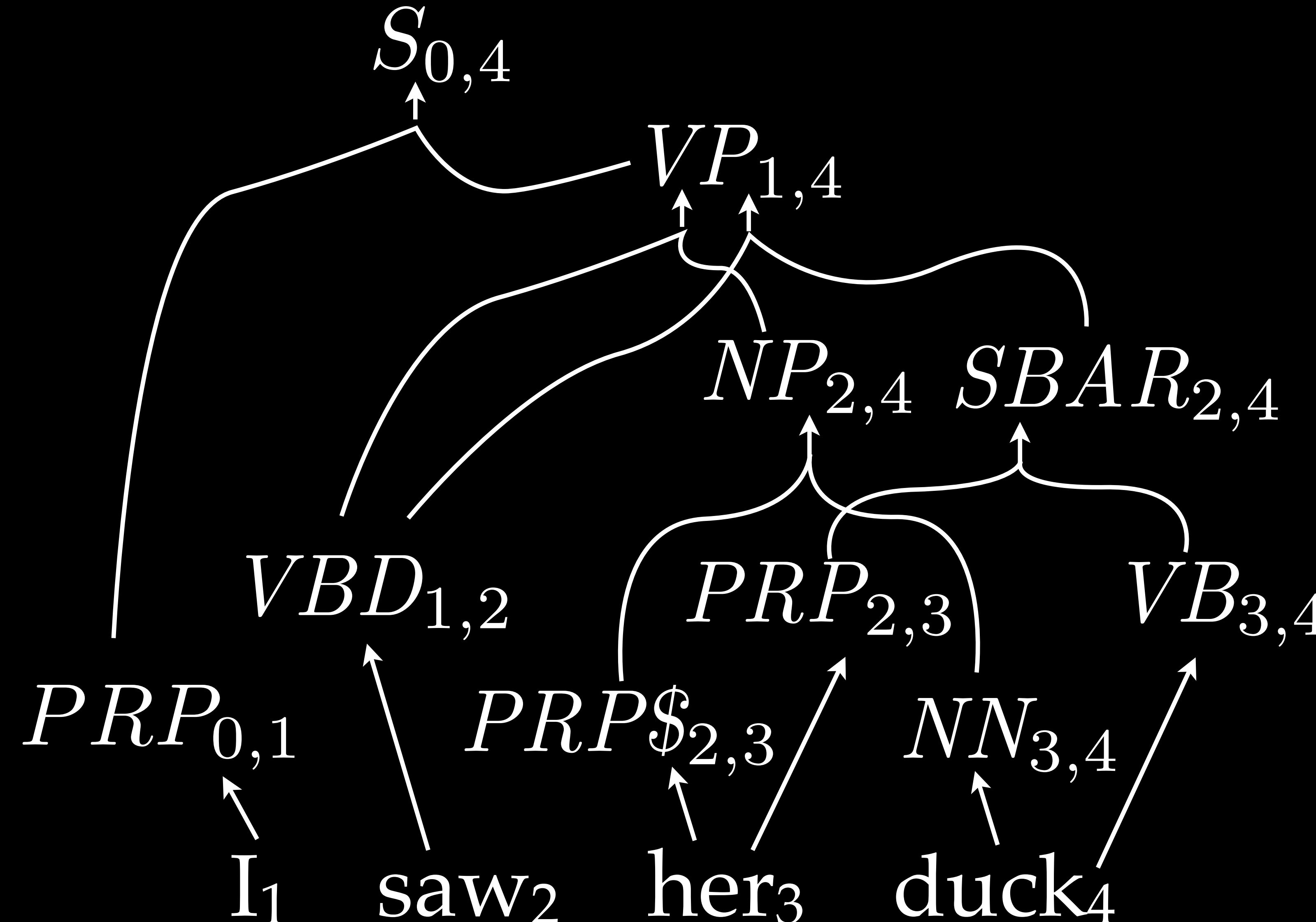


Parsing

Analysis

$O(Nn^2)$ nodes

$O(Gn^3)$ edges



Probabilistic Parsing

NN → duck

NP → PRP$ NN

PRP → her

PRP → I

PRP$ → her

S → PRP VP

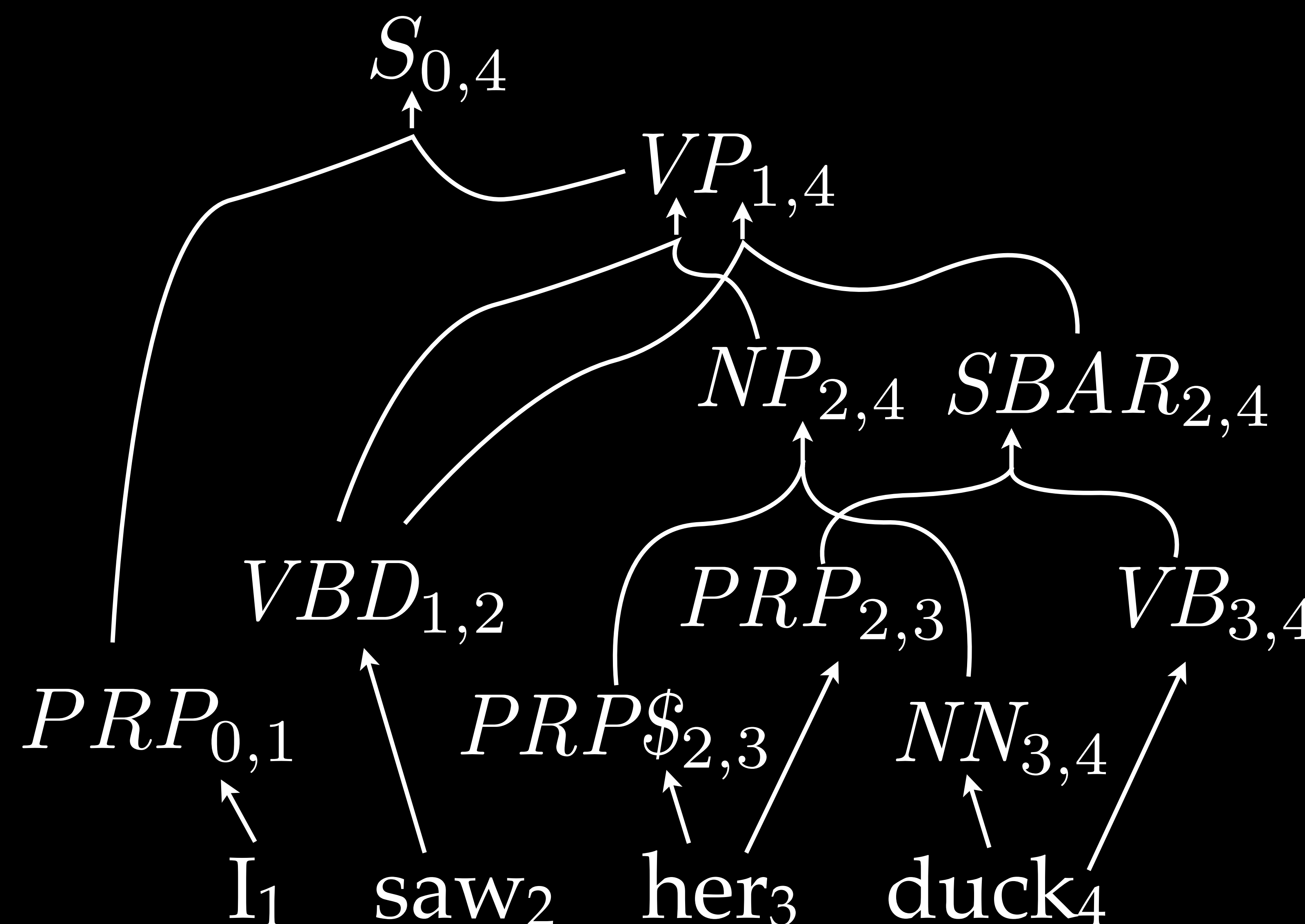
SBAR → PRP VB

VB → duck

VP → VBD NP

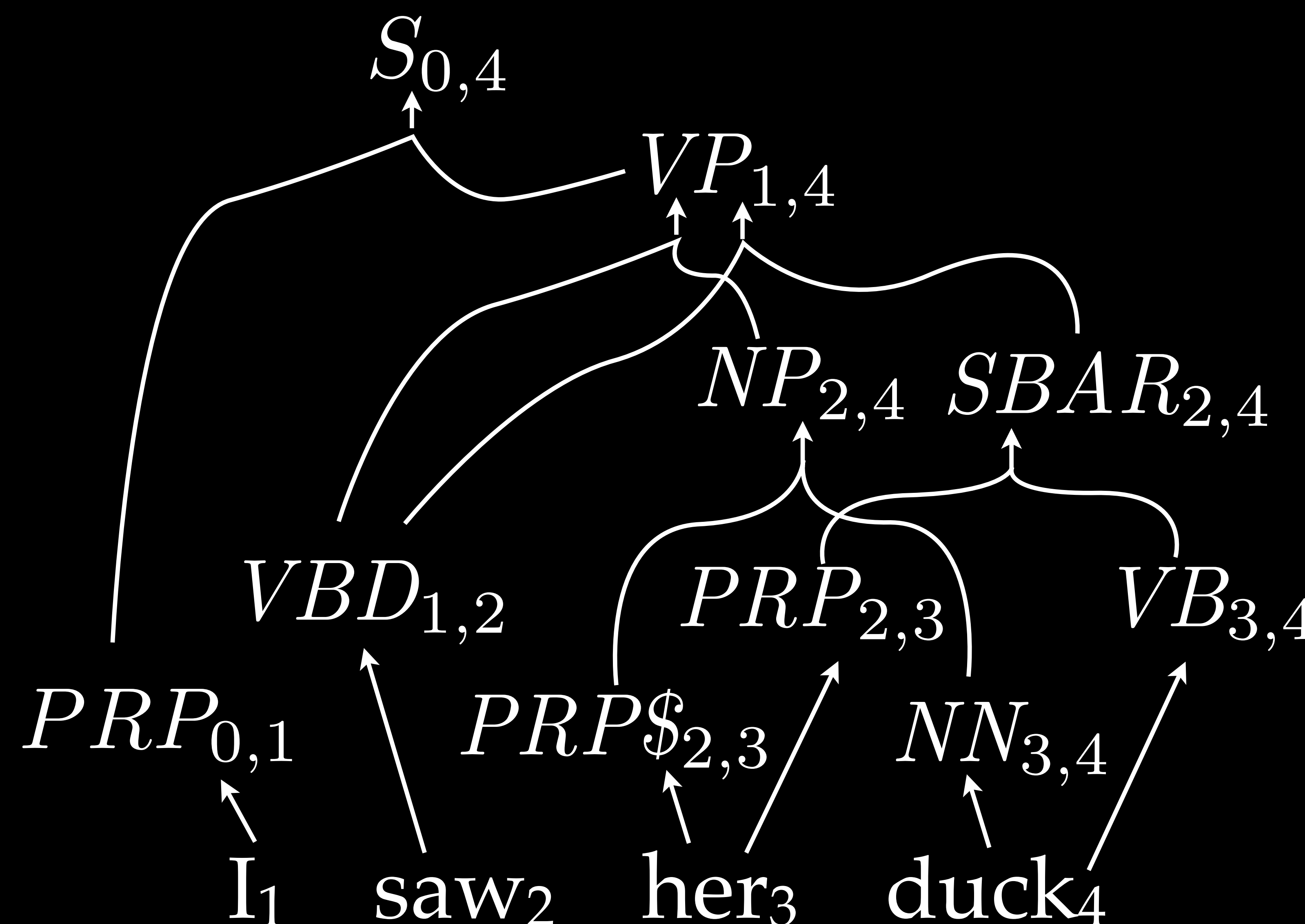
VP → VBD SBAR

VBD → saw



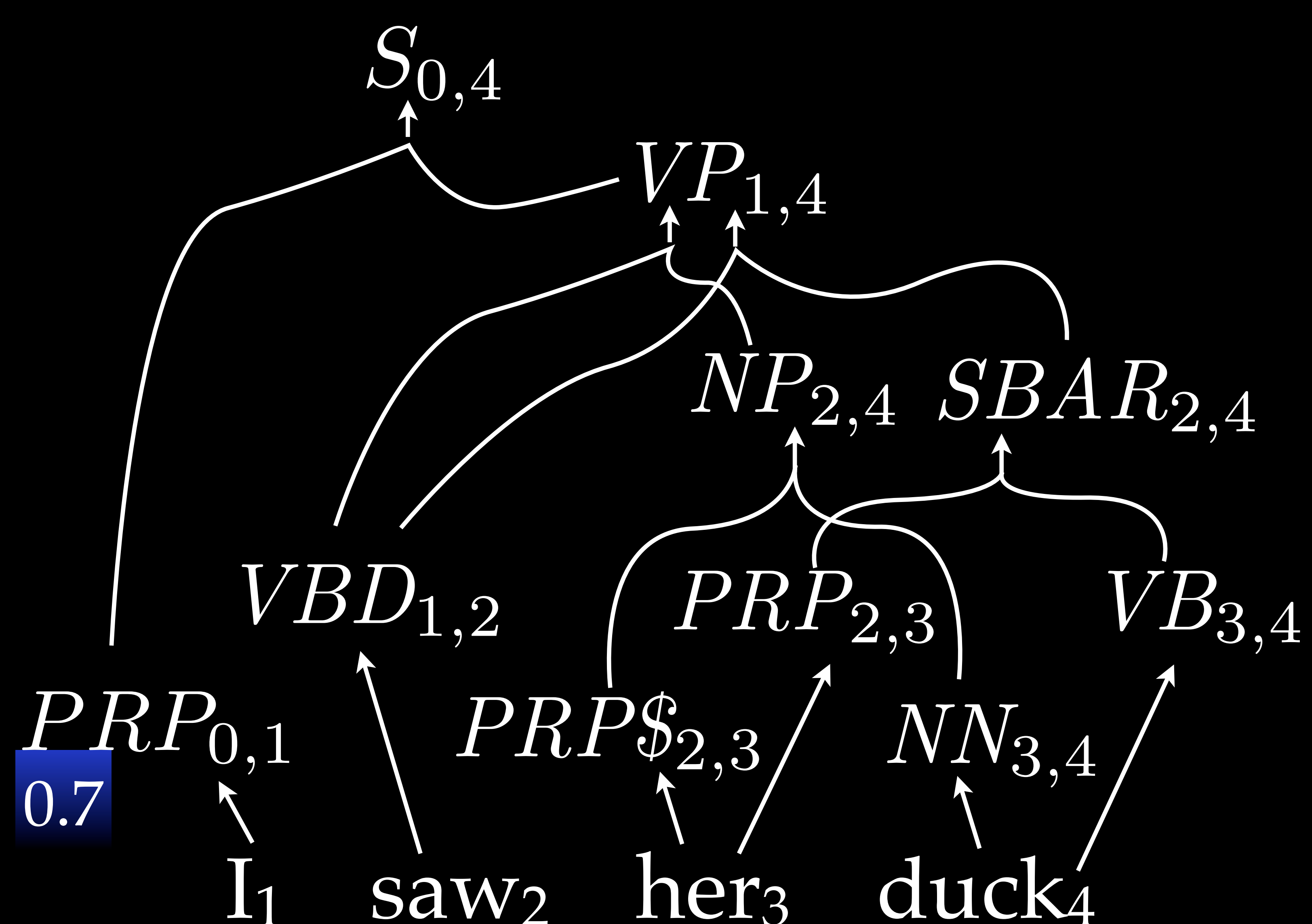
Probabilistic Parsing

NN → duck	(1.0)
NP → PRP\$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP\$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



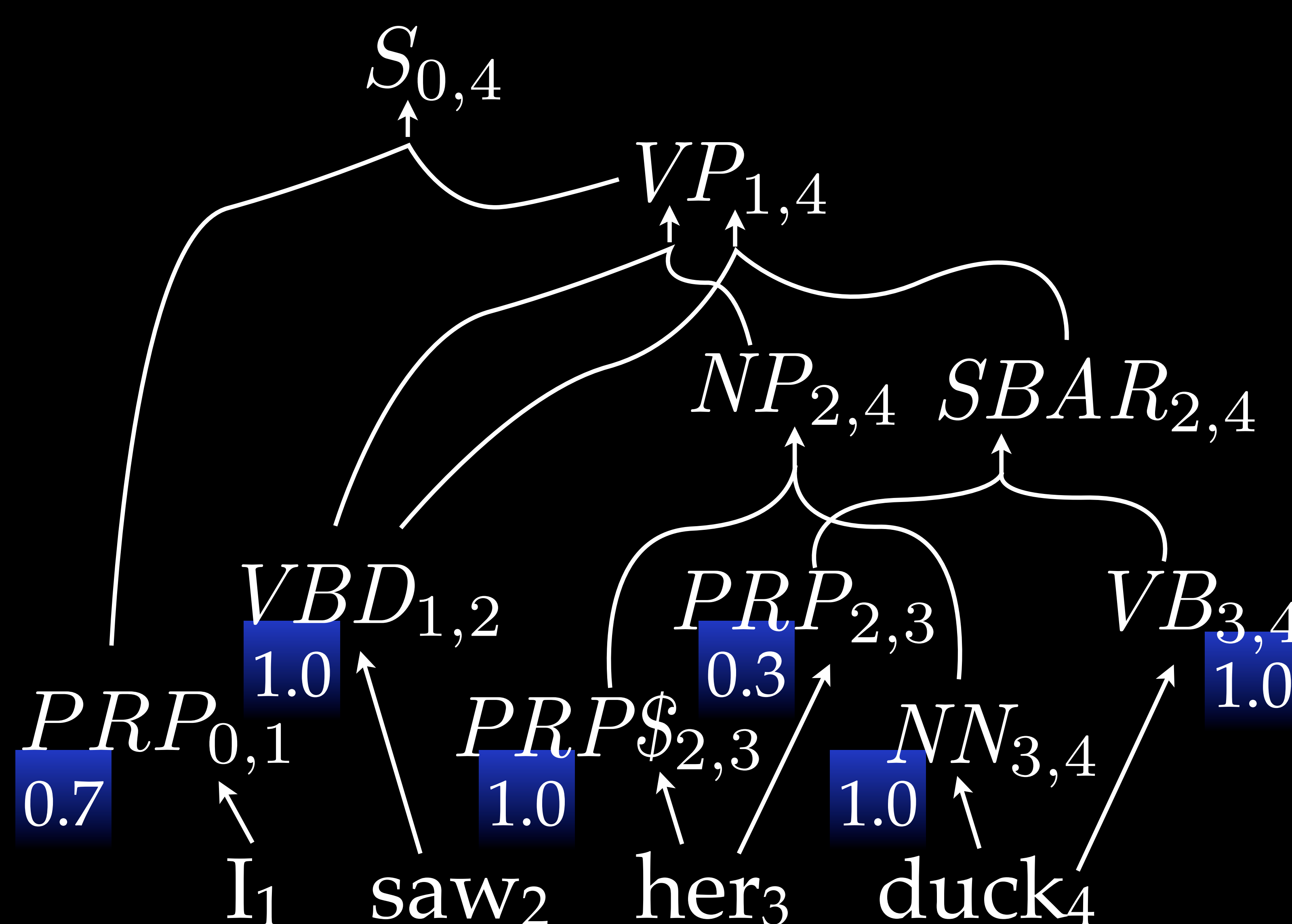
Probabilistic Parsing

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



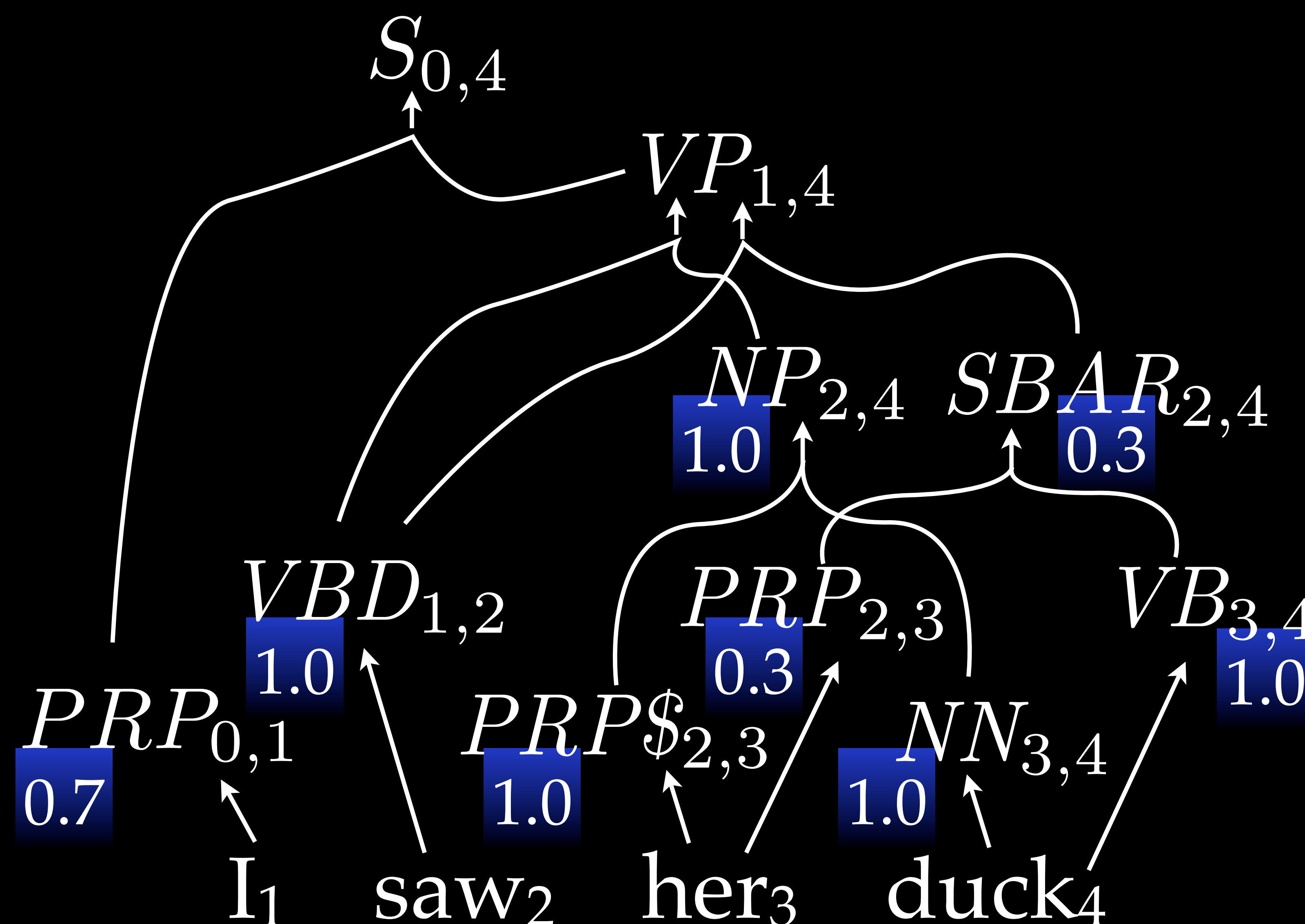
Probabilistic Parsing

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



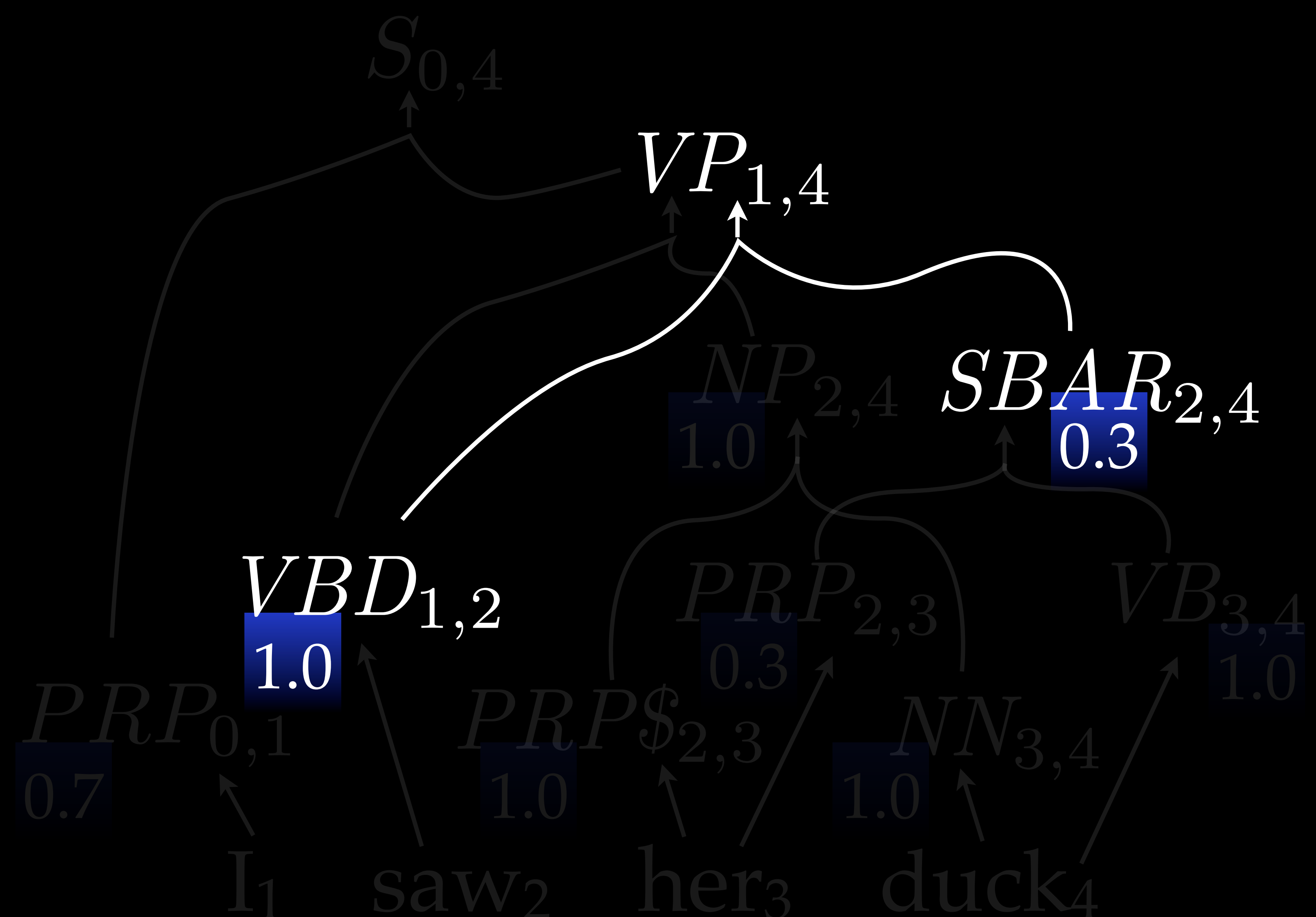
Probabilistic Parsing

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



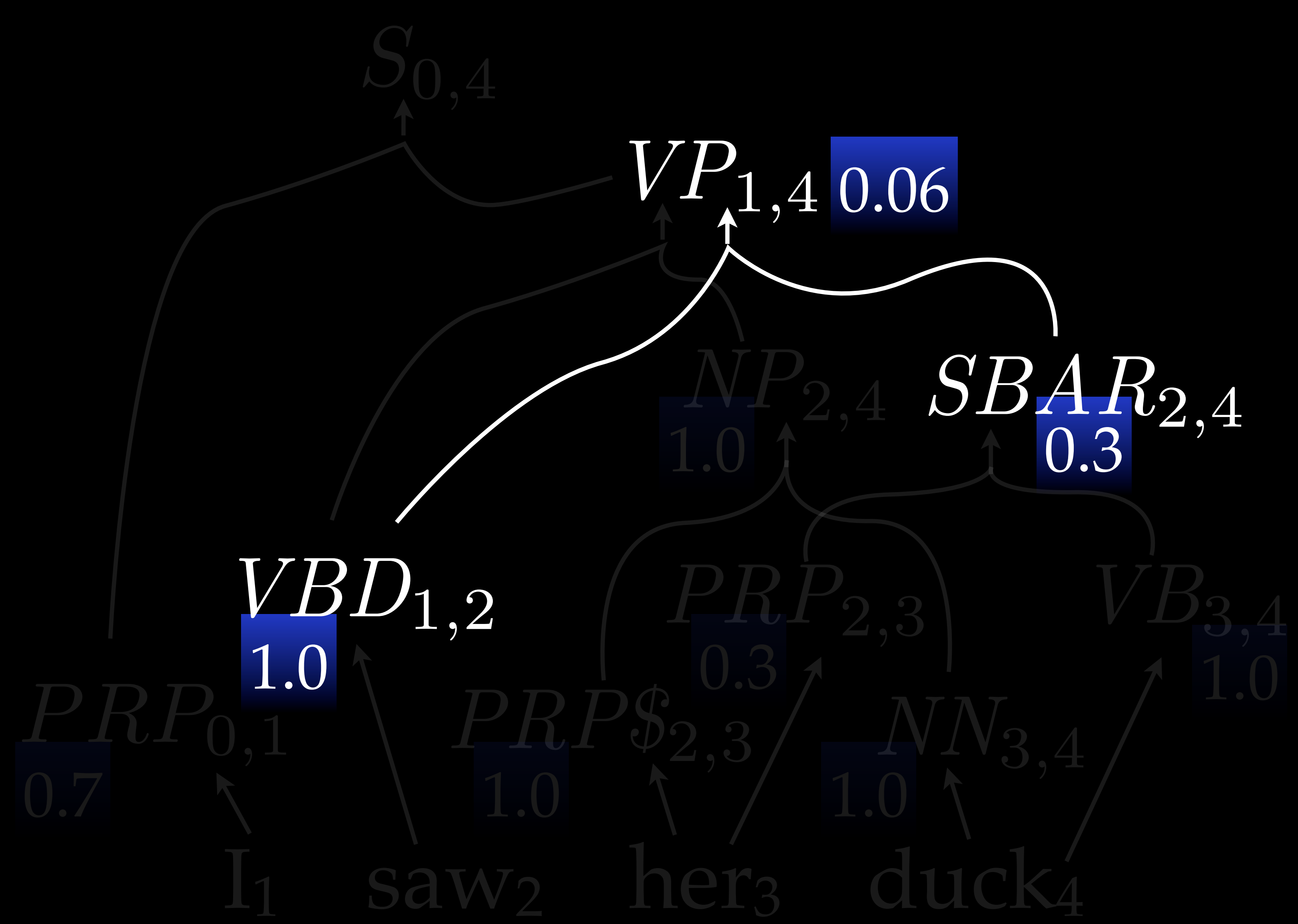
Probabilistic Parsing

NN → duck	(1.0)
NP → PRP\$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP\$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	**(0.2)**
VBD → saw	(1.0)



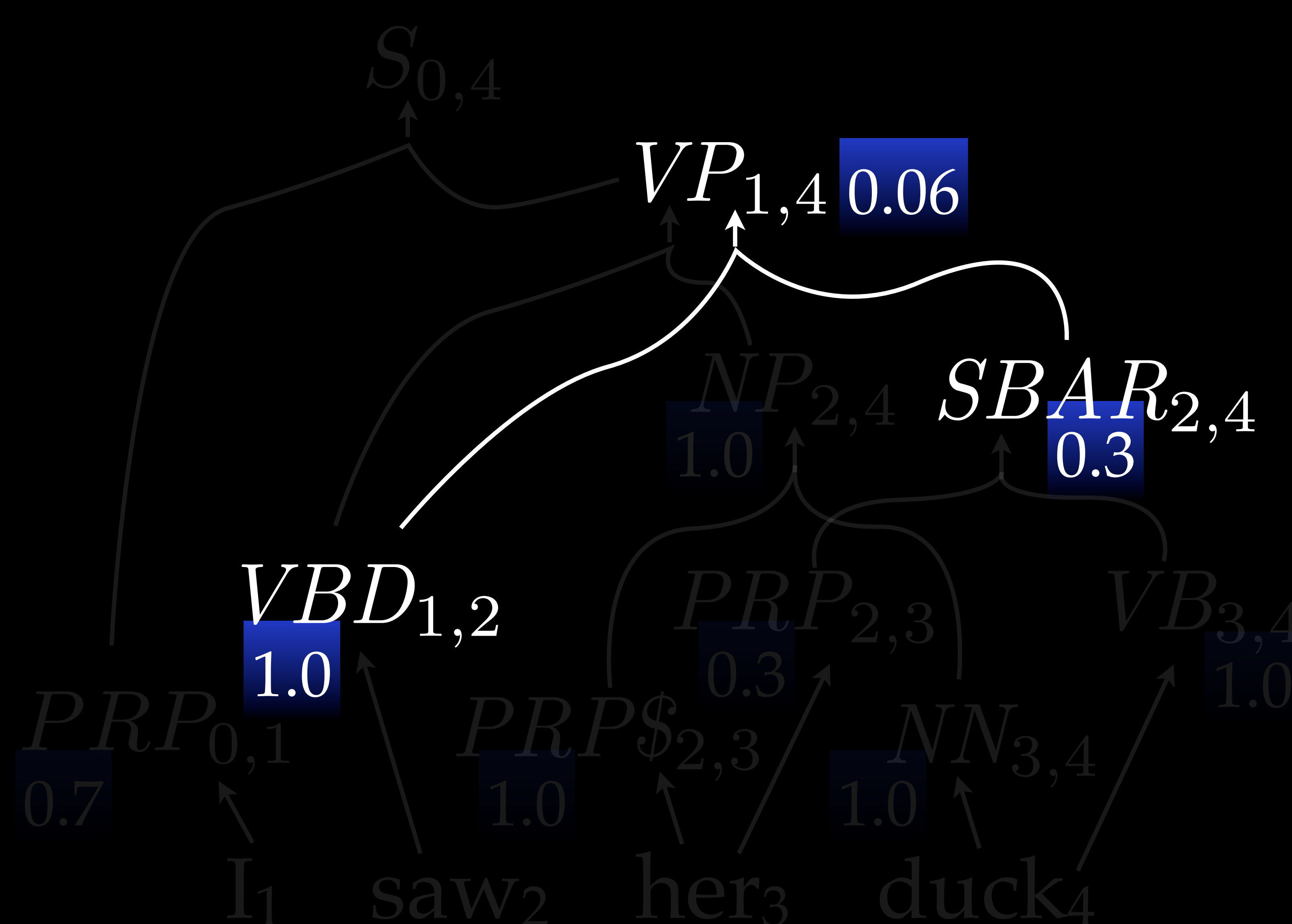
Probabilistic Parsing

NN → duck	(1.0)
NP → PRP\$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP\$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	**(0.2)**
VBD → saw	(1.0)



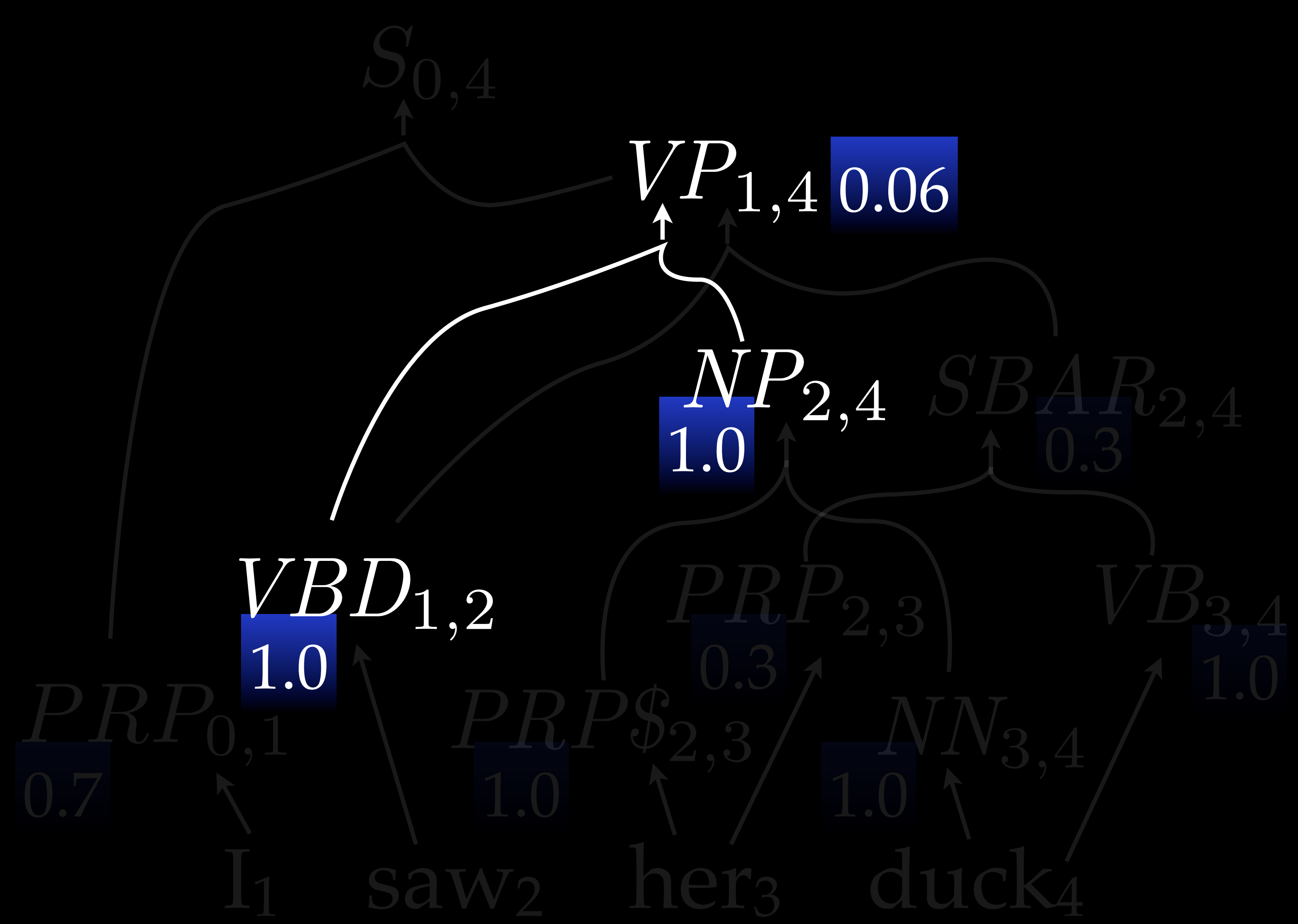
Probabilistic Parsing

NN $\rightarrow$ duck	(1.0)
NP $\rightarrow$ PRP$ NN	(1.0)
PRP $\rightarrow$ her	(0.3)
PRP $\rightarrow$ I	(0.7)
PRP$ $\rightarrow$ her	(1.0)
S $\rightarrow$ PRP VP	(1.0)
SBAR $\rightarrow$ PRP VB	(1.0)
VB $\rightarrow$ duck	(1.0)
VP $\rightarrow$ VBD NP	(0.8)
VP $\rightarrow$ VBD SBAR	(0.2)
VBD $\rightarrow$ saw	(1.0)



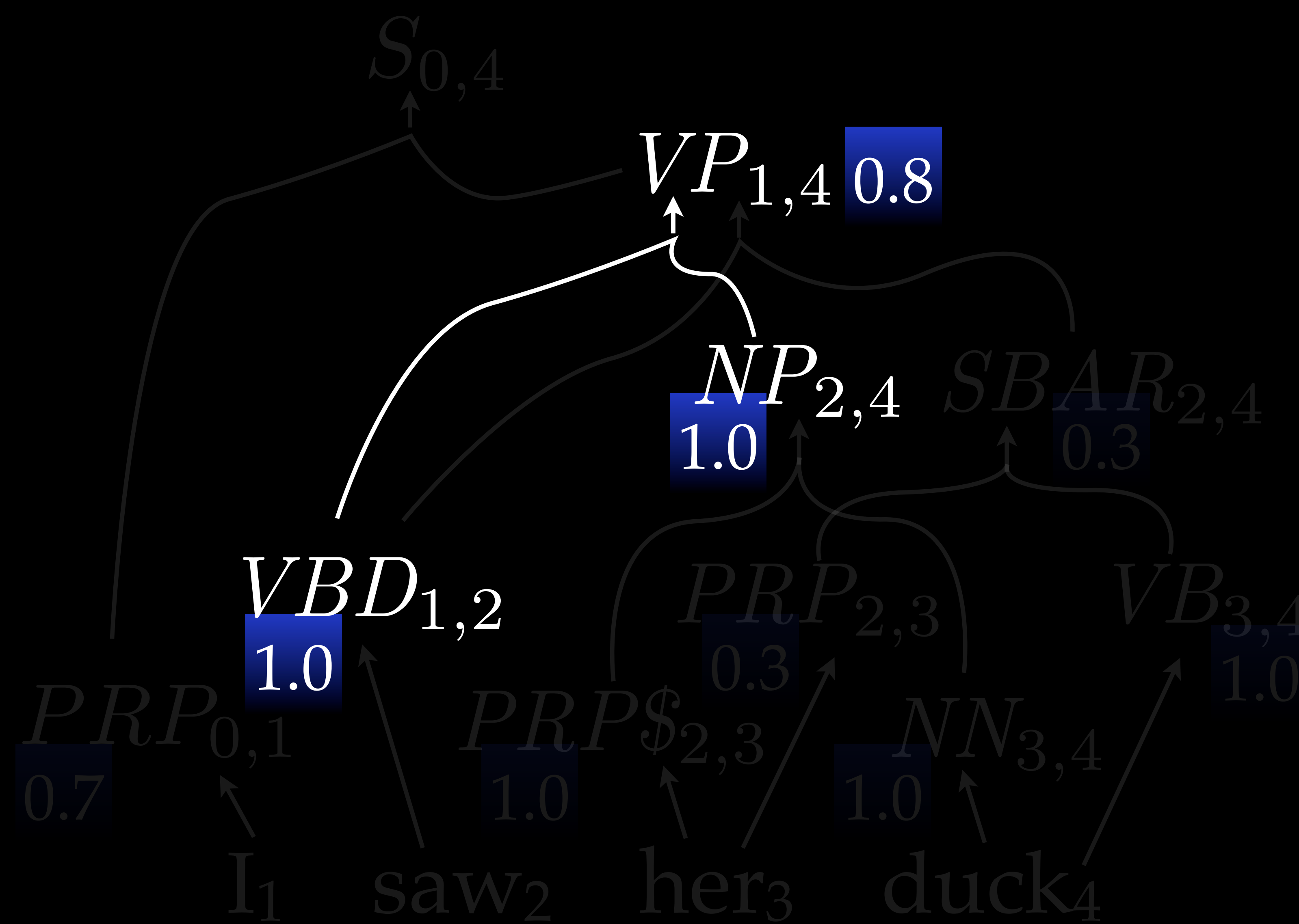
Probabilistic Parsing

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	**(0.8)**
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



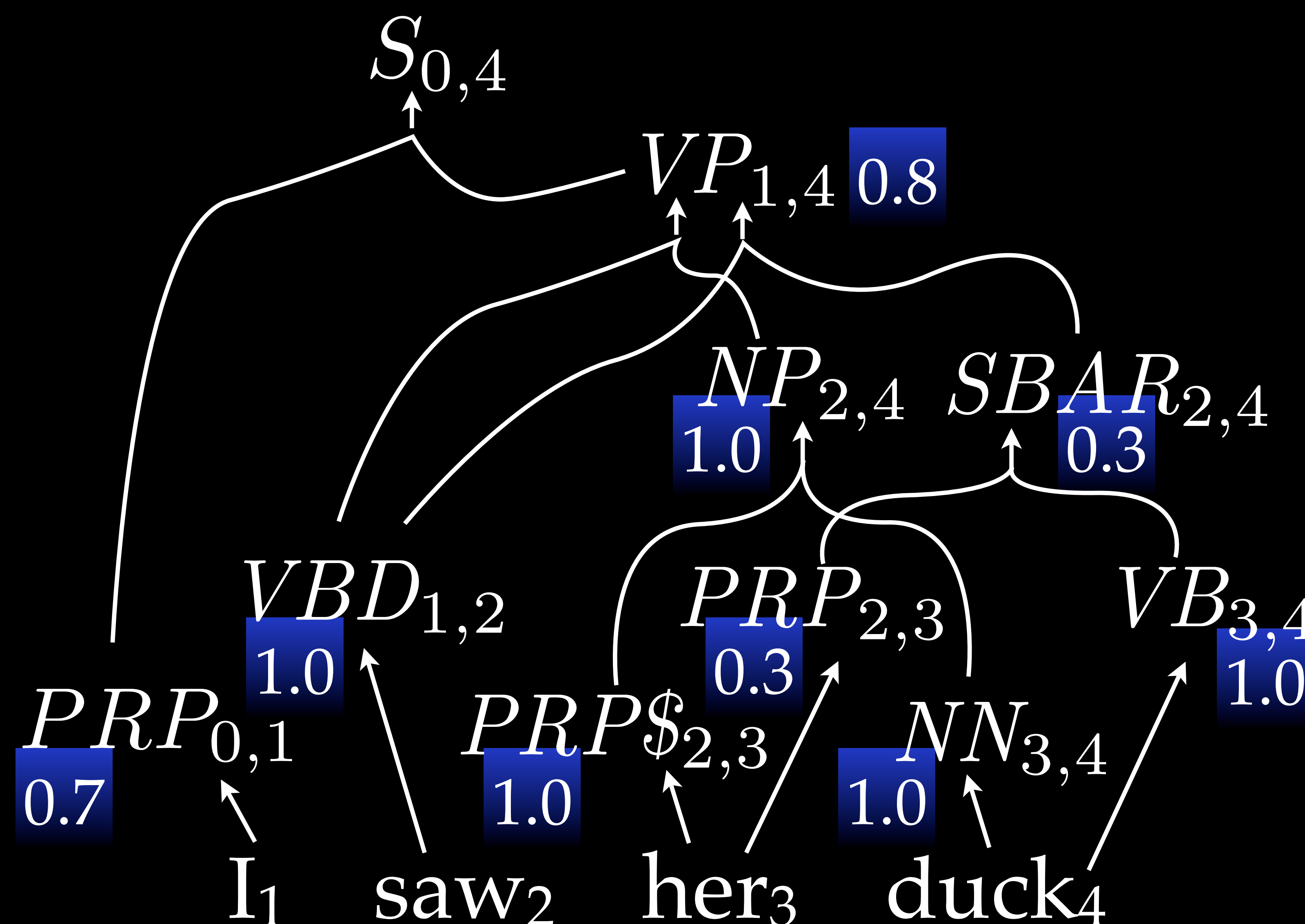
Probabilistic Parsing

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	**(0.8)**
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



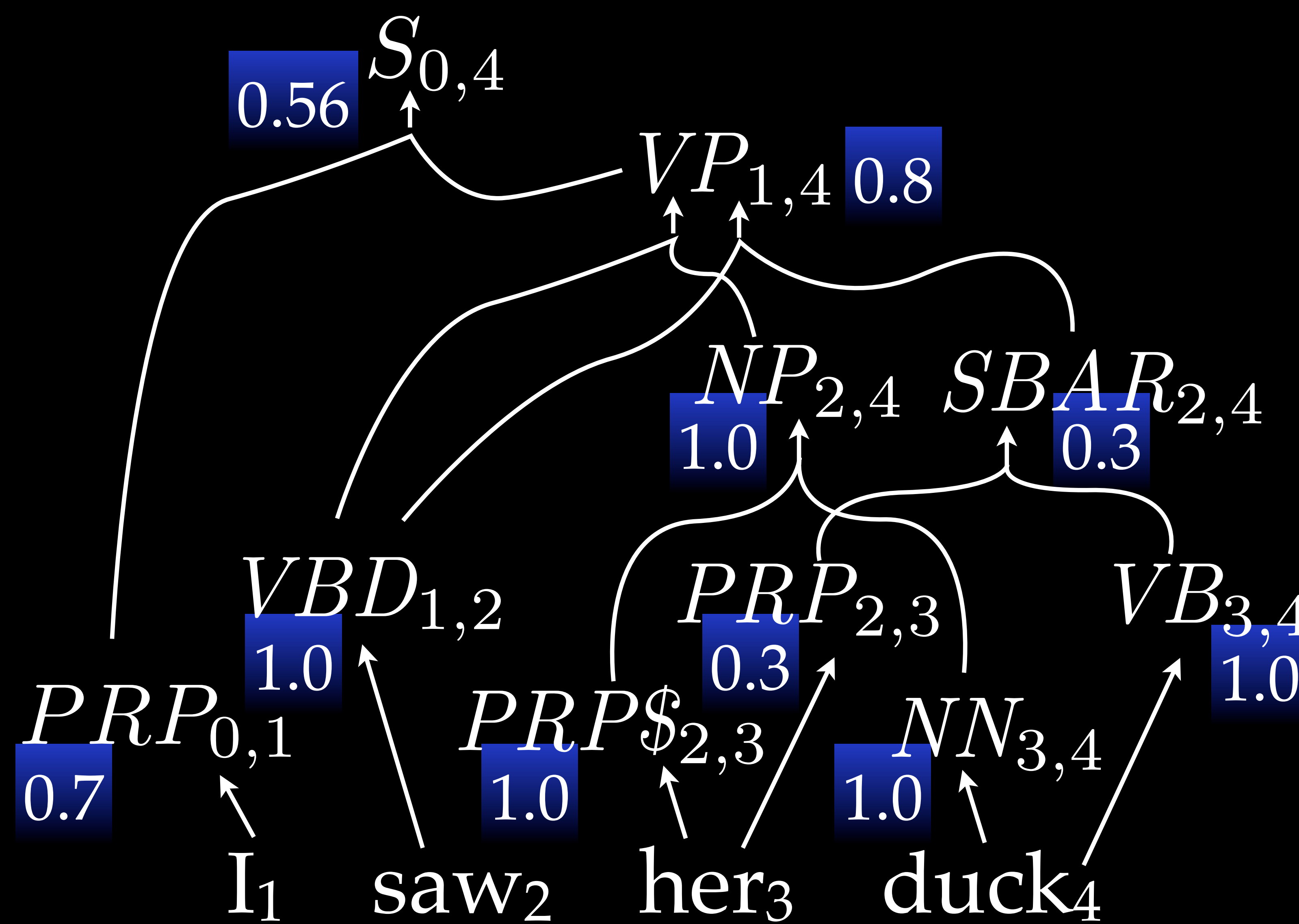
Probabilistic Parsing

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



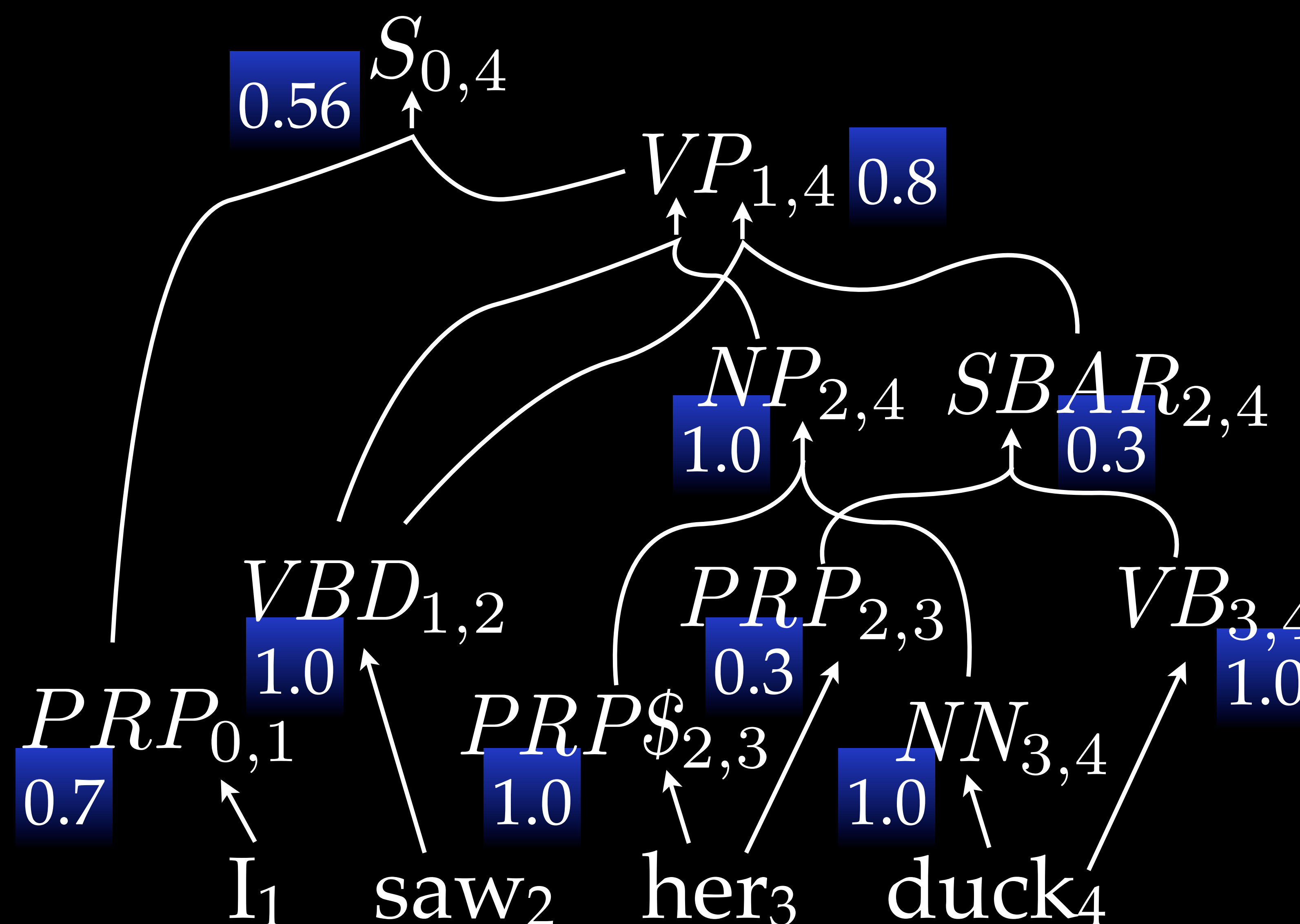
Probabilistic Parsing

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



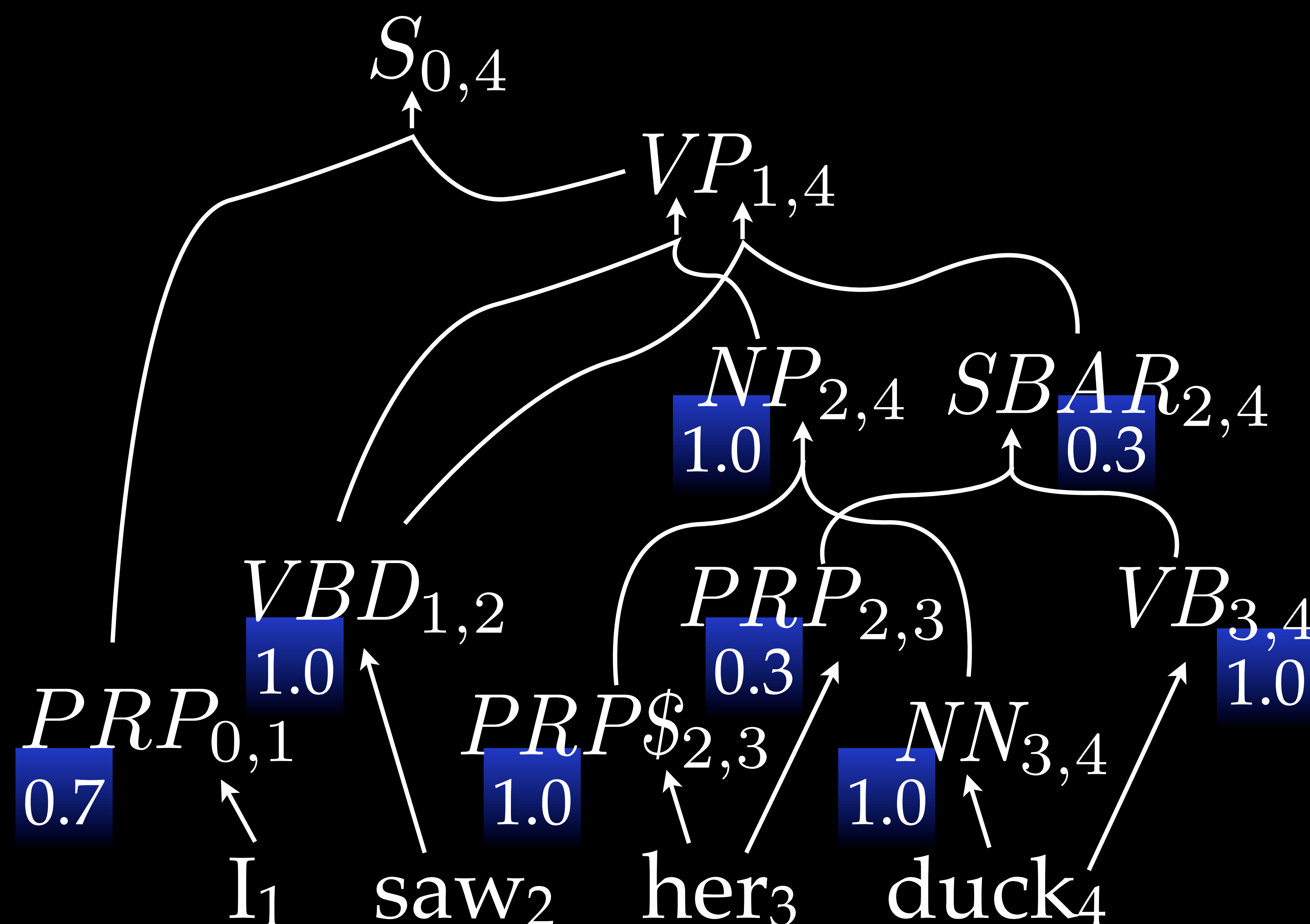
Probabilistic Parsing

$$X_{i,j} = \max(X_{i,j}, Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$



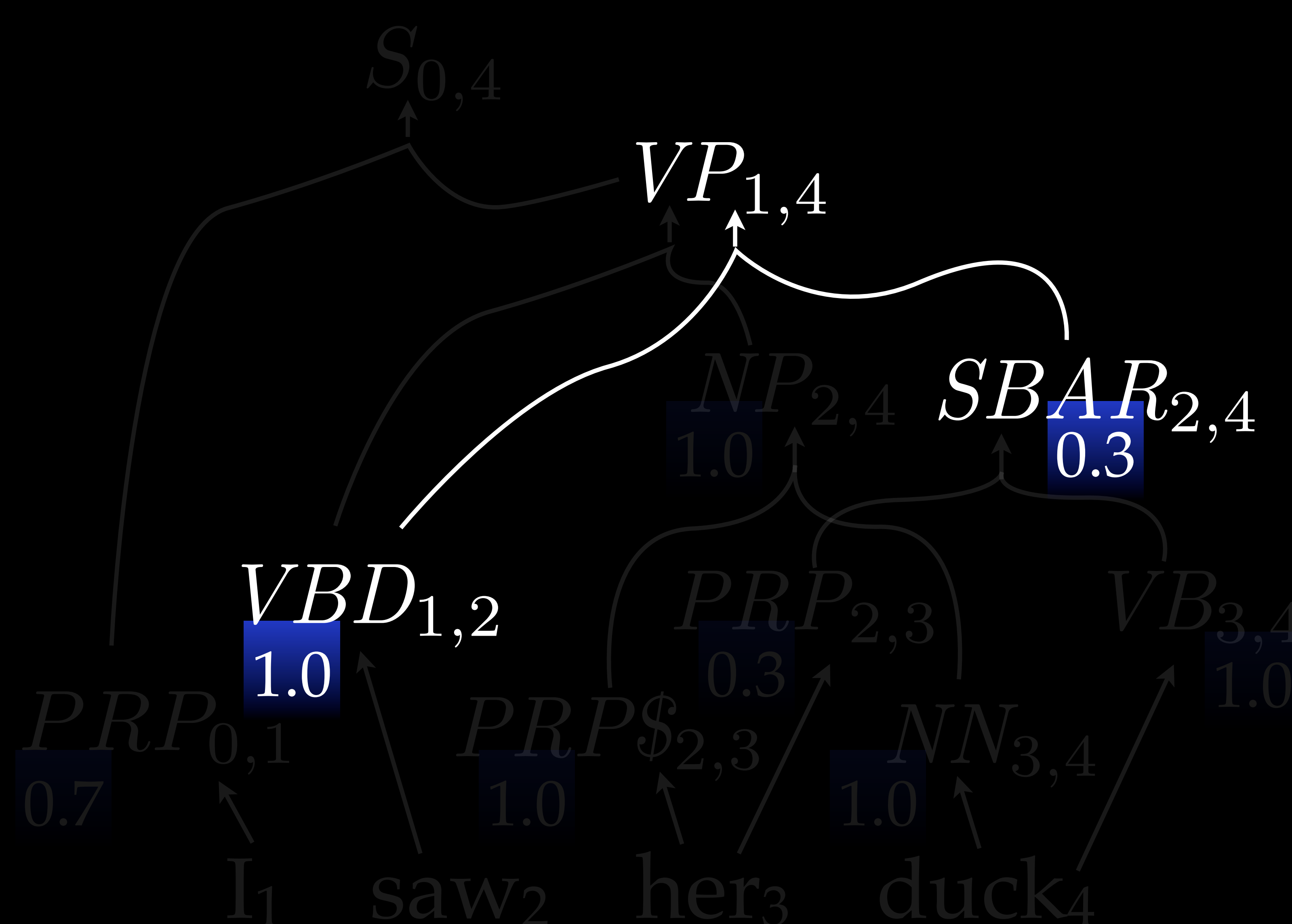
Computing Expectations

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



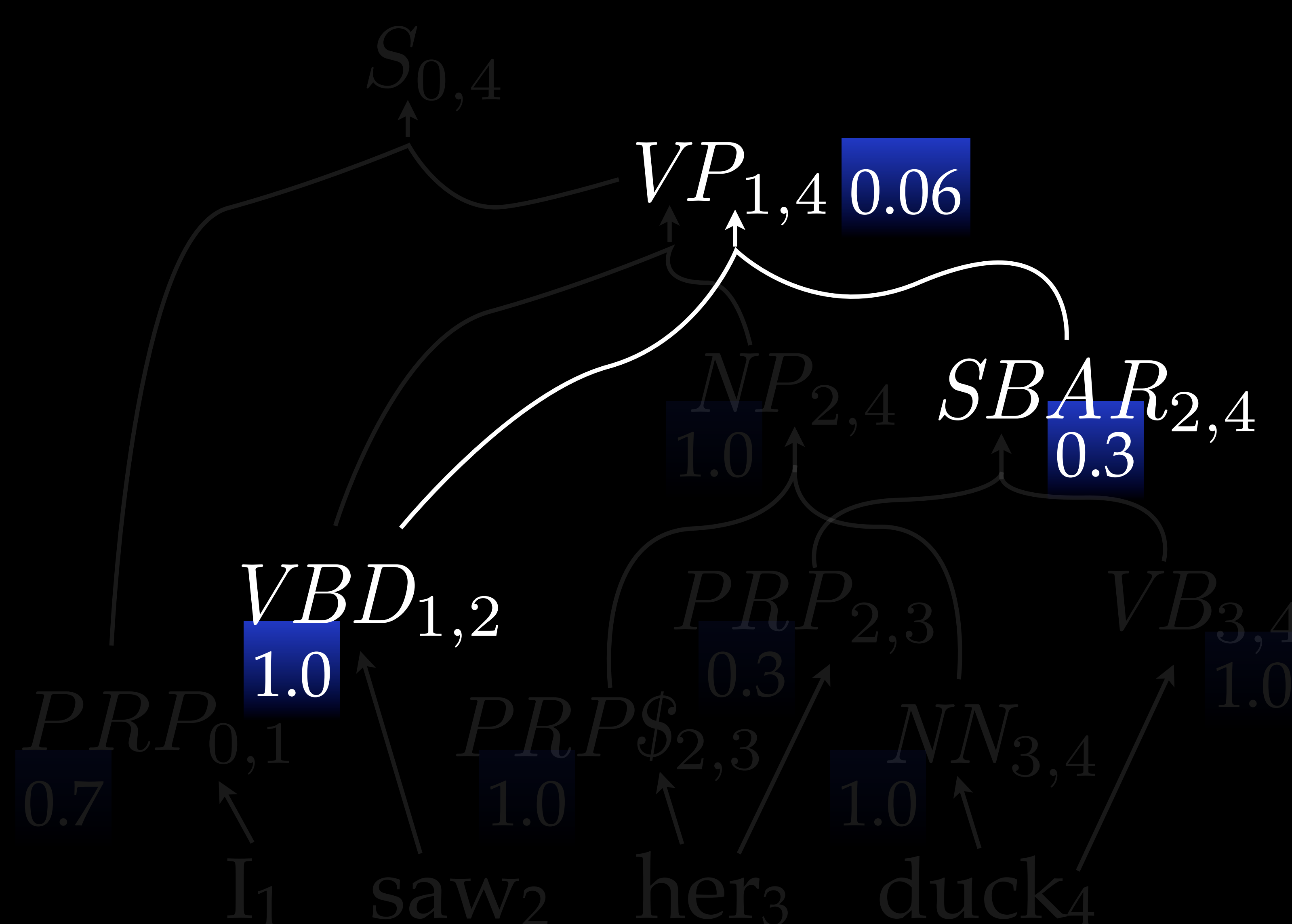
Computing Expectations

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



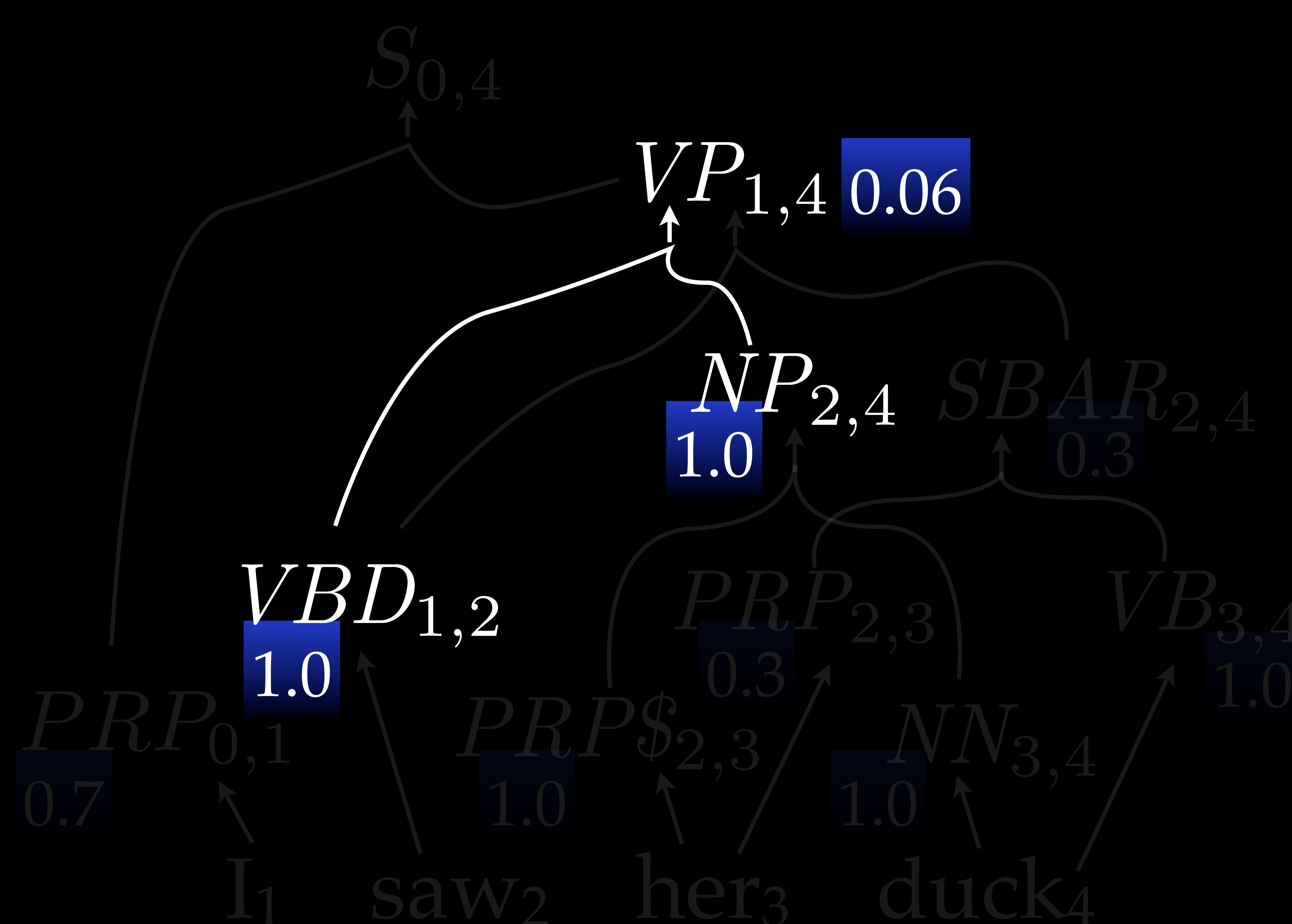
Computing Expectations

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



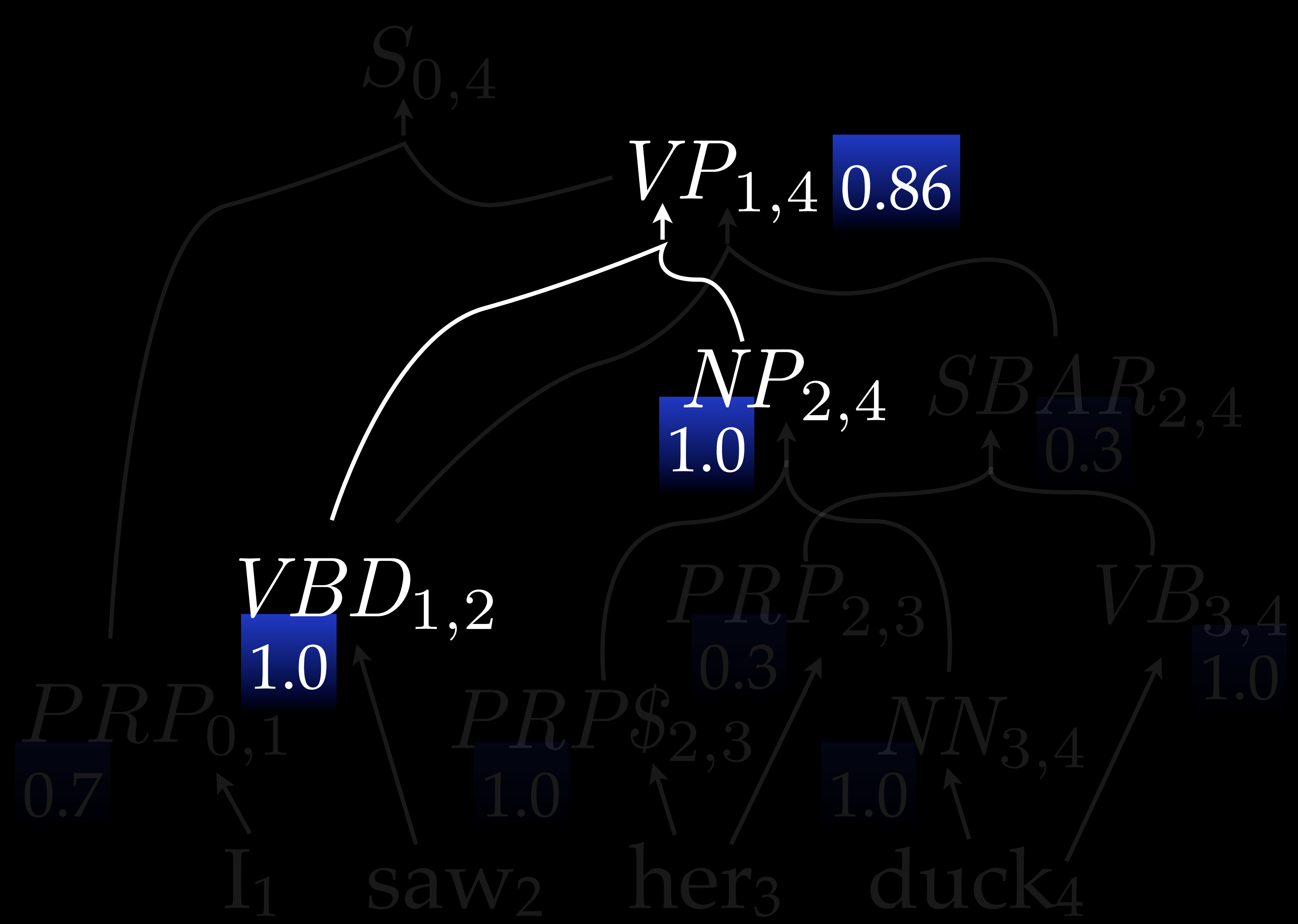
Computing Expectations

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



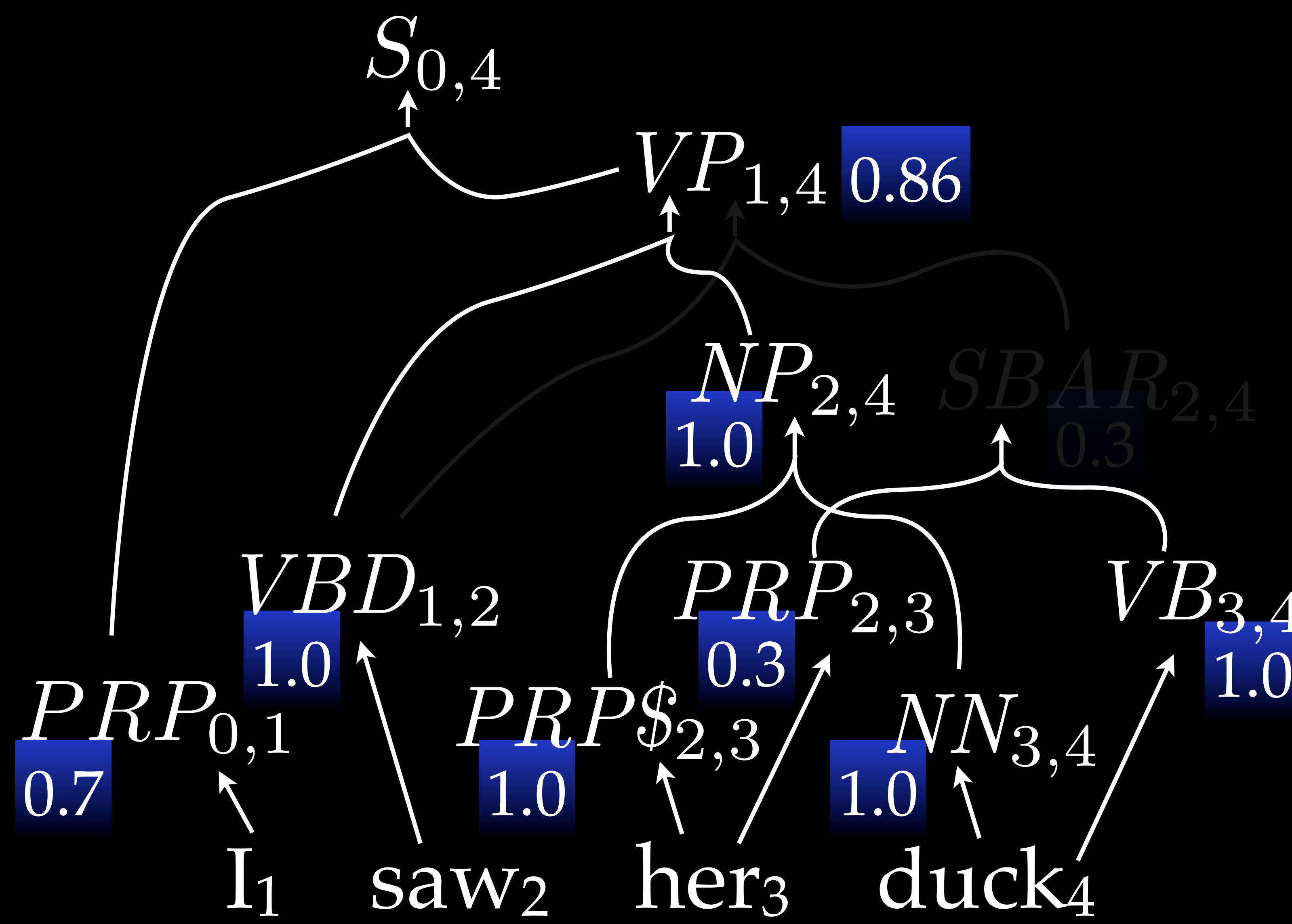
Computing Expectations

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



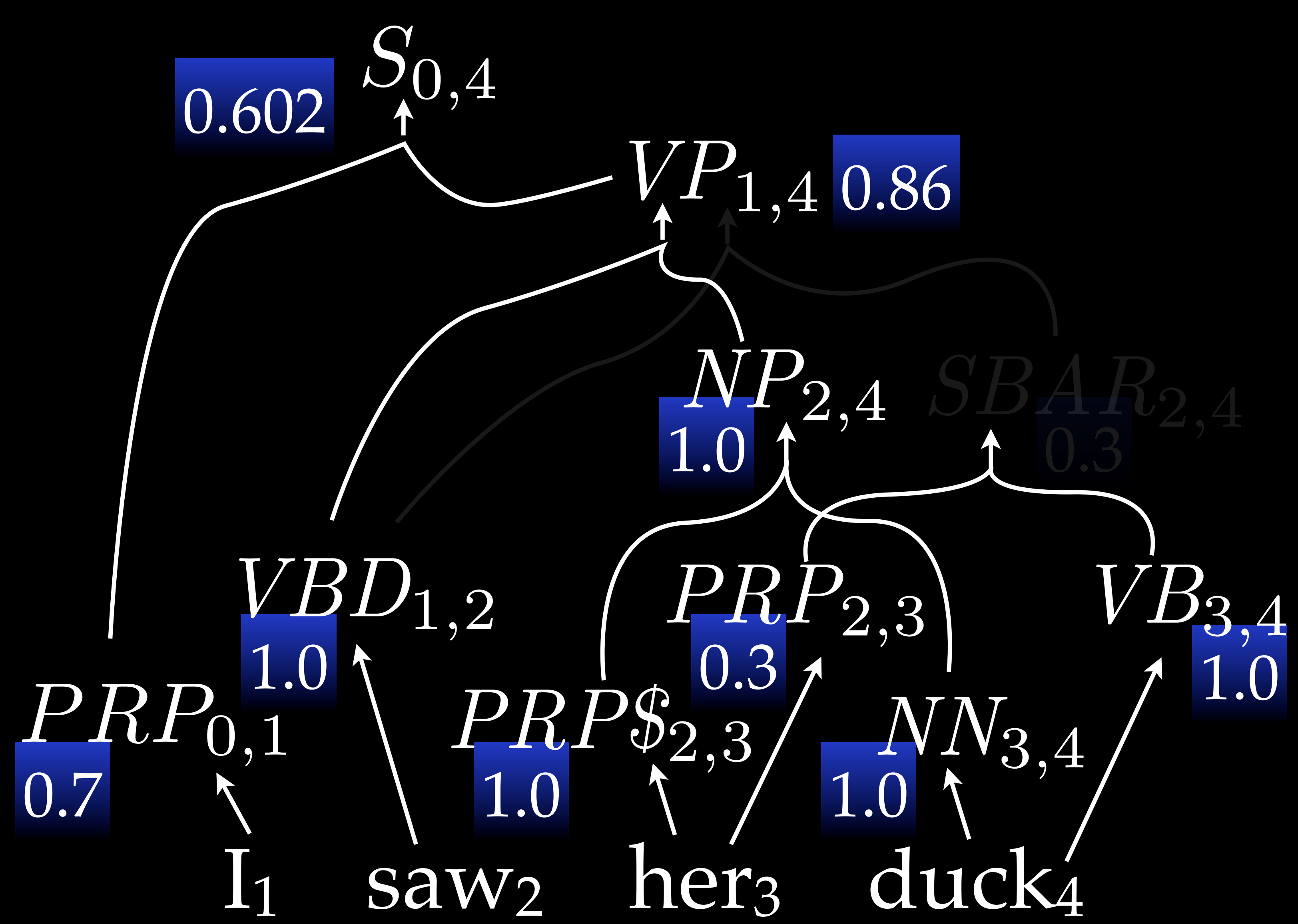
Computing Expectations

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



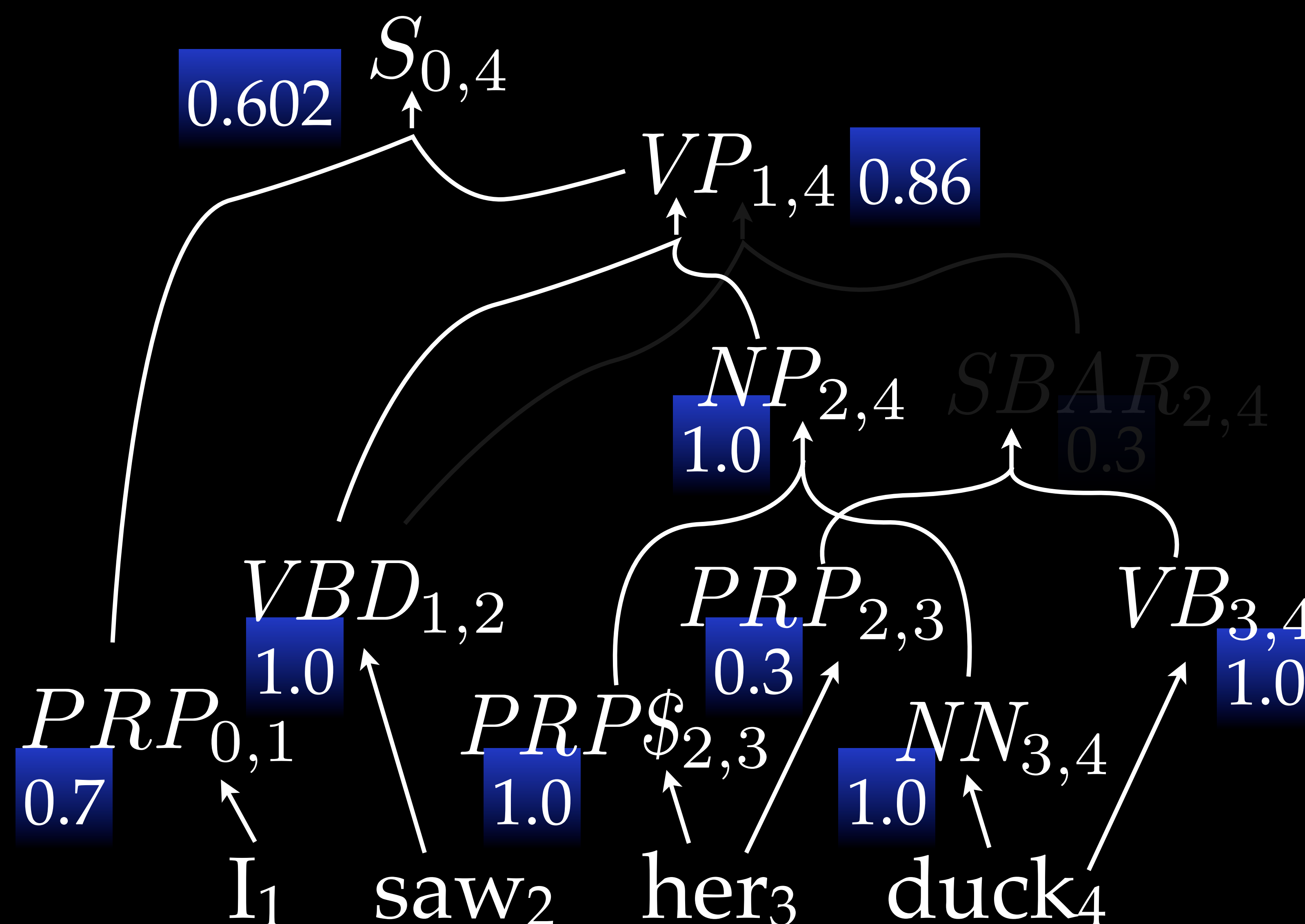
Computing Expectations

NN → duck	(1.0)
NP → PRP$ NN	(1.0)
PRP → her	(0.3)
PRP → I	(0.7)
PRP$ → her	(1.0)
S → PRP VP	(1.0)
SBAR → PRP VB	(1.0)
VB → duck	(1.0)
VP → VBD NP	(0.8)
VP → VBD SBAR	(0.2)
VBD → saw	(1.0)



Computing Expectations

$$X_{i,j} = X_{i,j} + (Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$



Semiring Parsing

$$X_{i,j} \leftarrow Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y\,Z)$$

$$X_{i,j} = \max(X_{i,j}, Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$

$$X_{i,j} = X_{i,j} + (Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$

Semiring Parsing

$$X_{i,j} = X_{i,j} \vee (Y_{i,k} \wedge Z_{k,j} \wedge (X \to Y Z))$$

$$X_{i,j} = \max(X_{i,j}, Y_{i,k} \times Z_{k,j} \times p(X \to Y Z))$$

$$X_{i,j} = X_{i,j} + (Y_{i,k} \times Z_{k,j} \times p(X \to Y Z))$$

Semiring Parsing

$$X_{i,j} = X_{i,j} \vee (Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y\,Z))$$

$$\langle \{T, F\}, \vee, \wedge \rangle$$

$$X_{i,j} = \max(X_{i,j}, Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$

$$X_{i,j} = X_{i,j} + (Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$

Semiring Parsing

$$X_{i,j} = X_{i,j} \vee (Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y\,Z))$$

$$\langle \{T, F\}, \vee, \wedge \rangle$$

$$X_{i,j} = \max(X_{i,j}, Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$

$$\langle \mathbb{R}, \max, \times \rangle$$

$$X_{i,j} = X_{i,j} + (Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$

Semiring Parsing

$$X_{i,j} = X_{i,j} \vee (Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y Z))$$

$$\langle \{T, F\}, \vee, \wedge \rangle$$

$$X_{i,j} = \max(X_{i,j}, Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y Z))$$

$$\langle \mathbb{R}, \max, \times \rangle$$

$$X_{i,j} = X_{i,j} + (Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y Z))$$

$$\langle \mathbb{R}, +, \times \rangle$$

Semiring Parsing

$$X_{i,j} = X_{i,j} \vee (Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y\,Z))$$

$$\langle \{T, F\}, \vee, \wedge \rangle$$

$$X_{i,j} = \max(X_{i,j}, Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$

$$\langle \mathbb{R}, \max, \times \rangle$$

$$X_{i,j} = X_{i,j} + (Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$

$$\langle \mathbb{R}, +, \times \rangle$$

$$X_{i,j} = X_{i,j} \oplus (Y_{i,k} \otimes Z_{k,j} \otimes R(X \rightarrow Y\,Z))$$

Semiring Parsing

$$X_{i,j} = X_{i,j} \vee (Y_{i,k} \wedge Z_{k,j} \wedge (X \rightarrow Y\,Z))$$

boolean $\langle \{T, F\}, \vee, \wedge \rangle$

$$X_{i,j} = \max(X_{i,j}, Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$

Viterbi $\langle \mathbb{R}, \max, \times \rangle$

$$X_{i,j} = X_{i,j} + (Y_{i,k} \times Z_{k,j} \times p(X \rightarrow Y\,Z))$$

inside $\langle \mathbb{R}, +, \times \rangle$

$$X_{i,j} = X_{i,j} \oplus (Y_{i,k} \otimes Z_{k,j} \otimes R(X \rightarrow Y\,Z))$$

Parsing

Is Intersection!

$NN_{3,4} \rightarrow$ duck

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ her

$PRP_{0,1} \rightarrow$ I

$PRP\$_{2,3} \rightarrow$ her

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

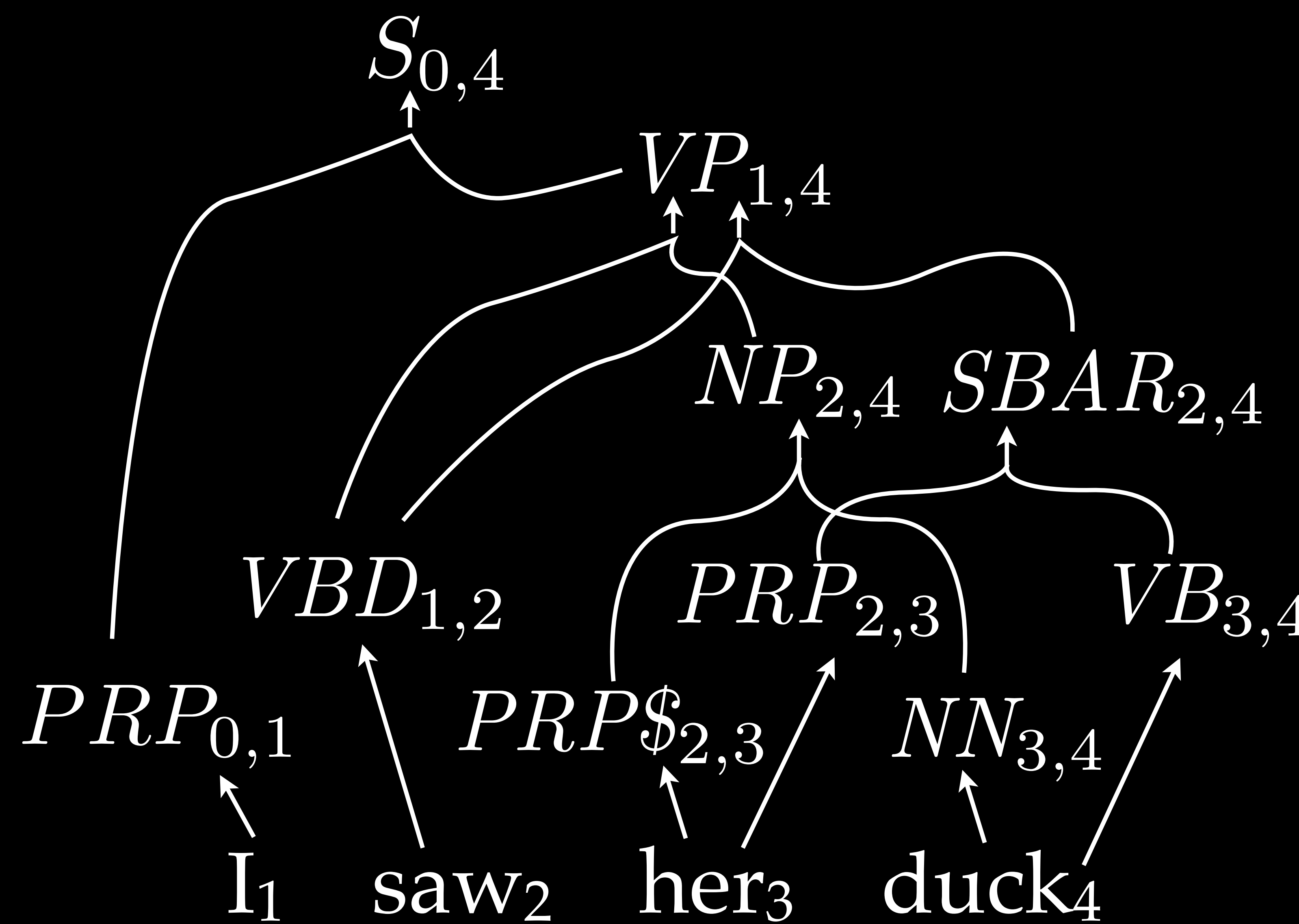
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ duck

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

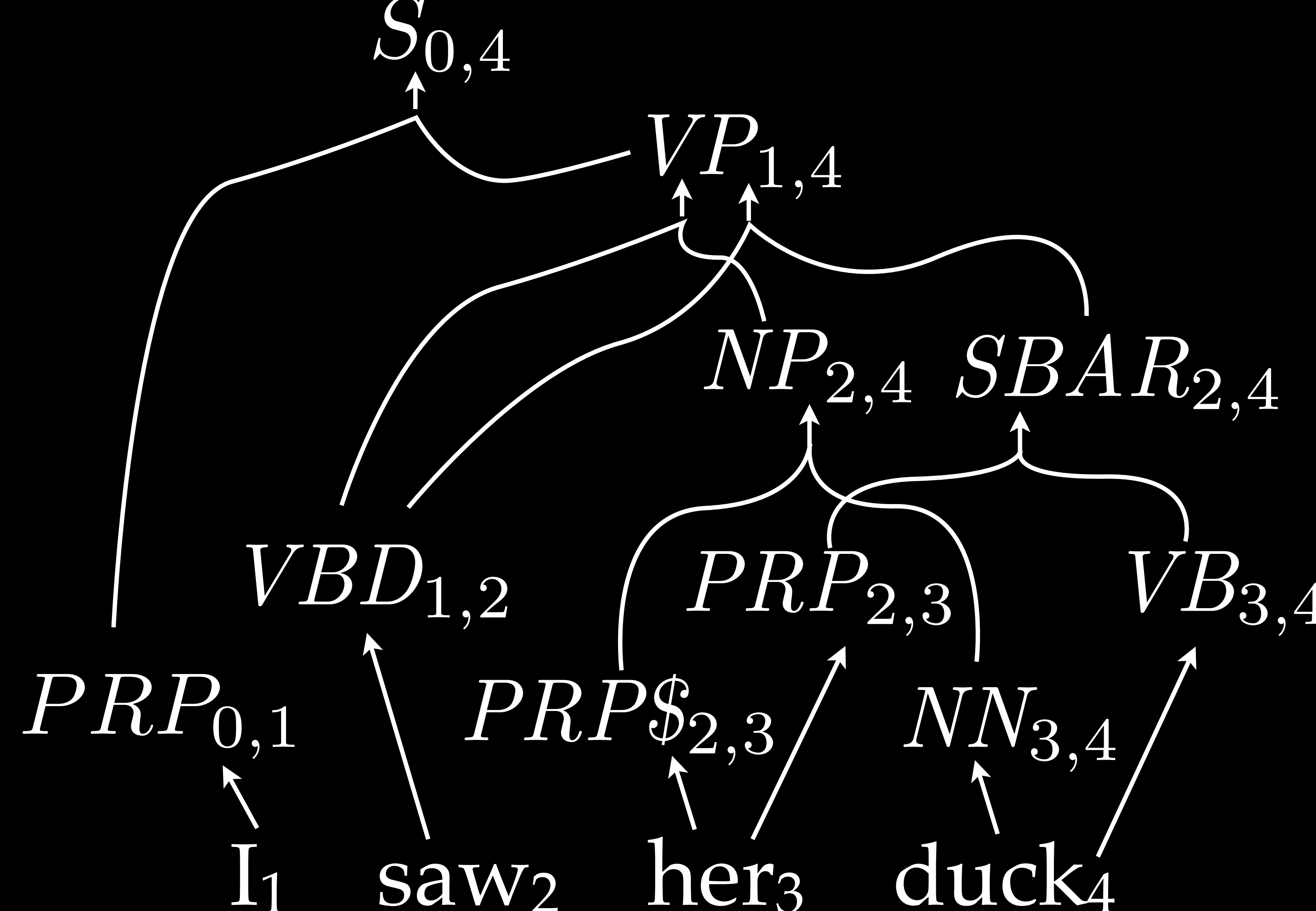
$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ saw



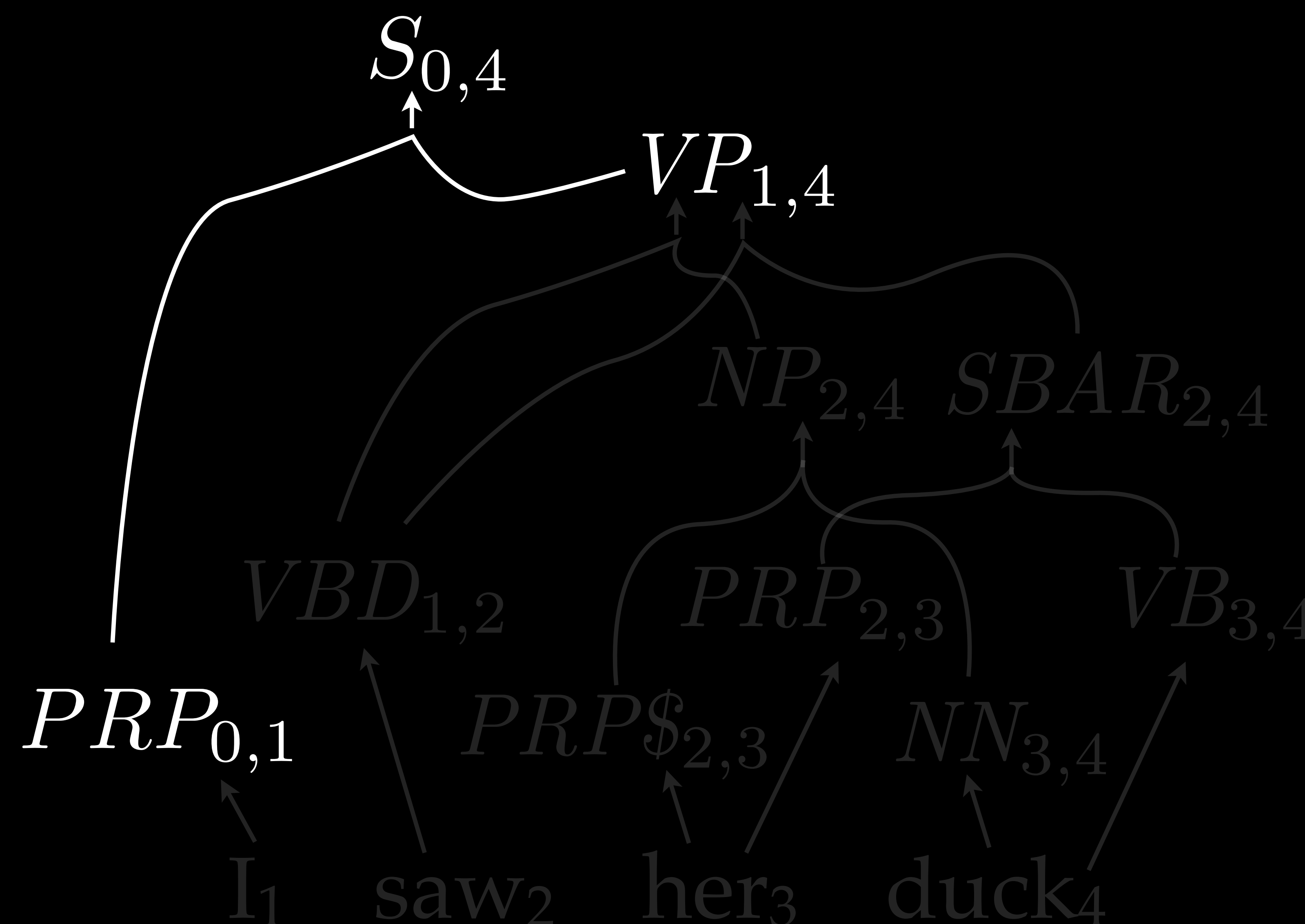
Parsing

Is Intersection!



Parsing

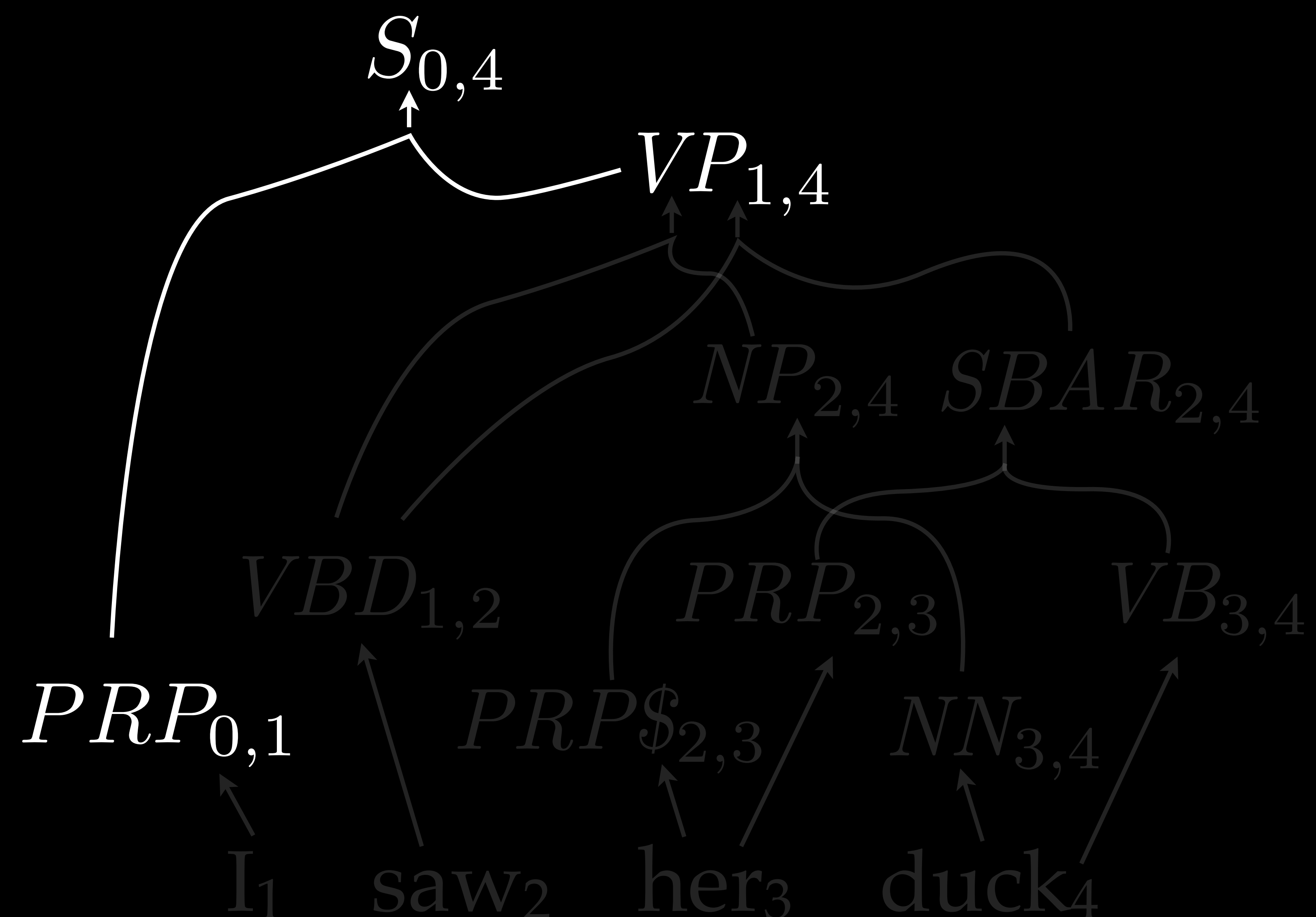
Is Intersection!



Parsing

Is Intersection!

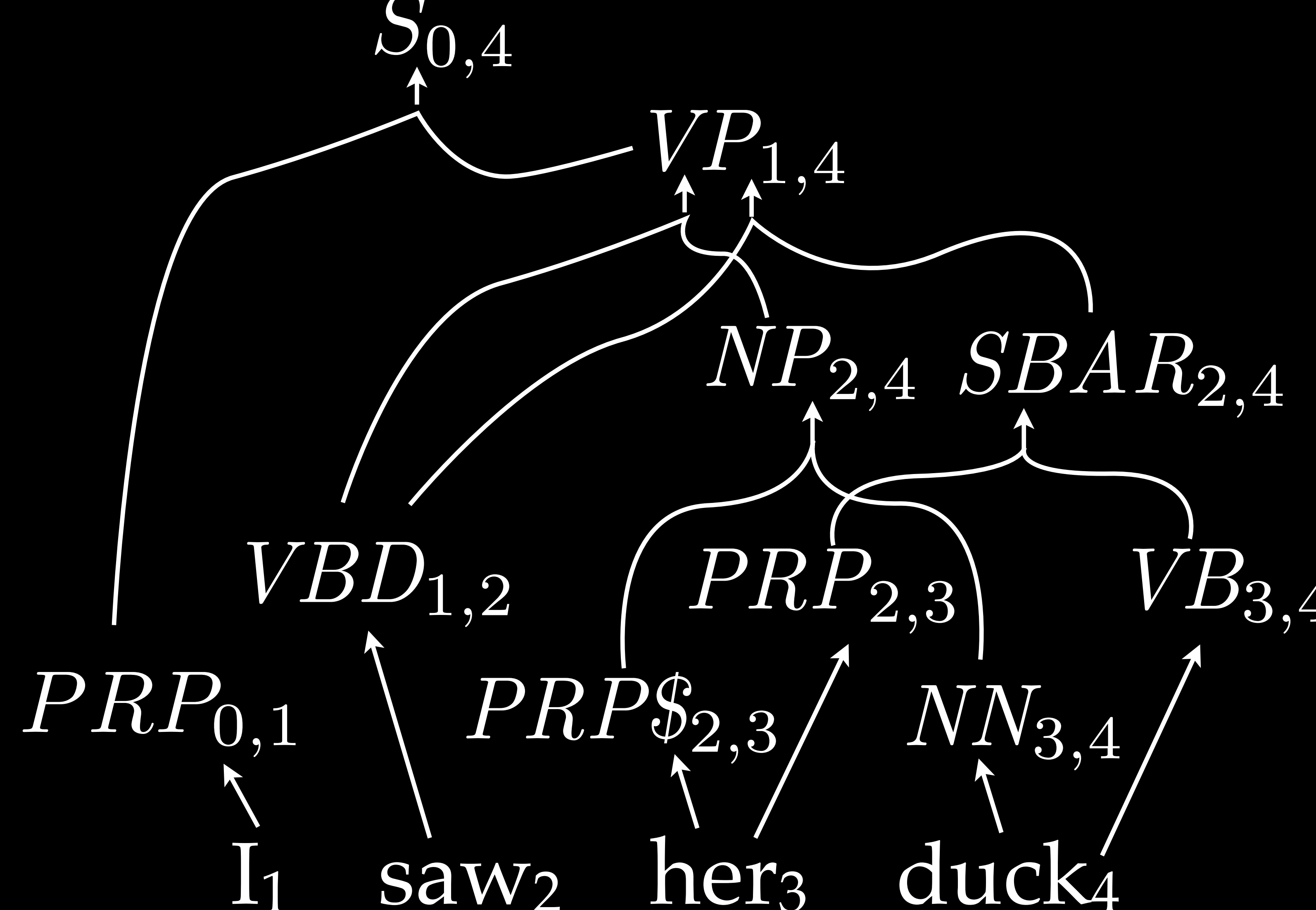
$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$



Parsing

Is Intersection!

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$



Parsing

Is Intersection!

$NN_{3,4} \rightarrow$ duck

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ her

$PRP_{0,1} \rightarrow$ I

$PRP\$_{2,3} \rightarrow$ her

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

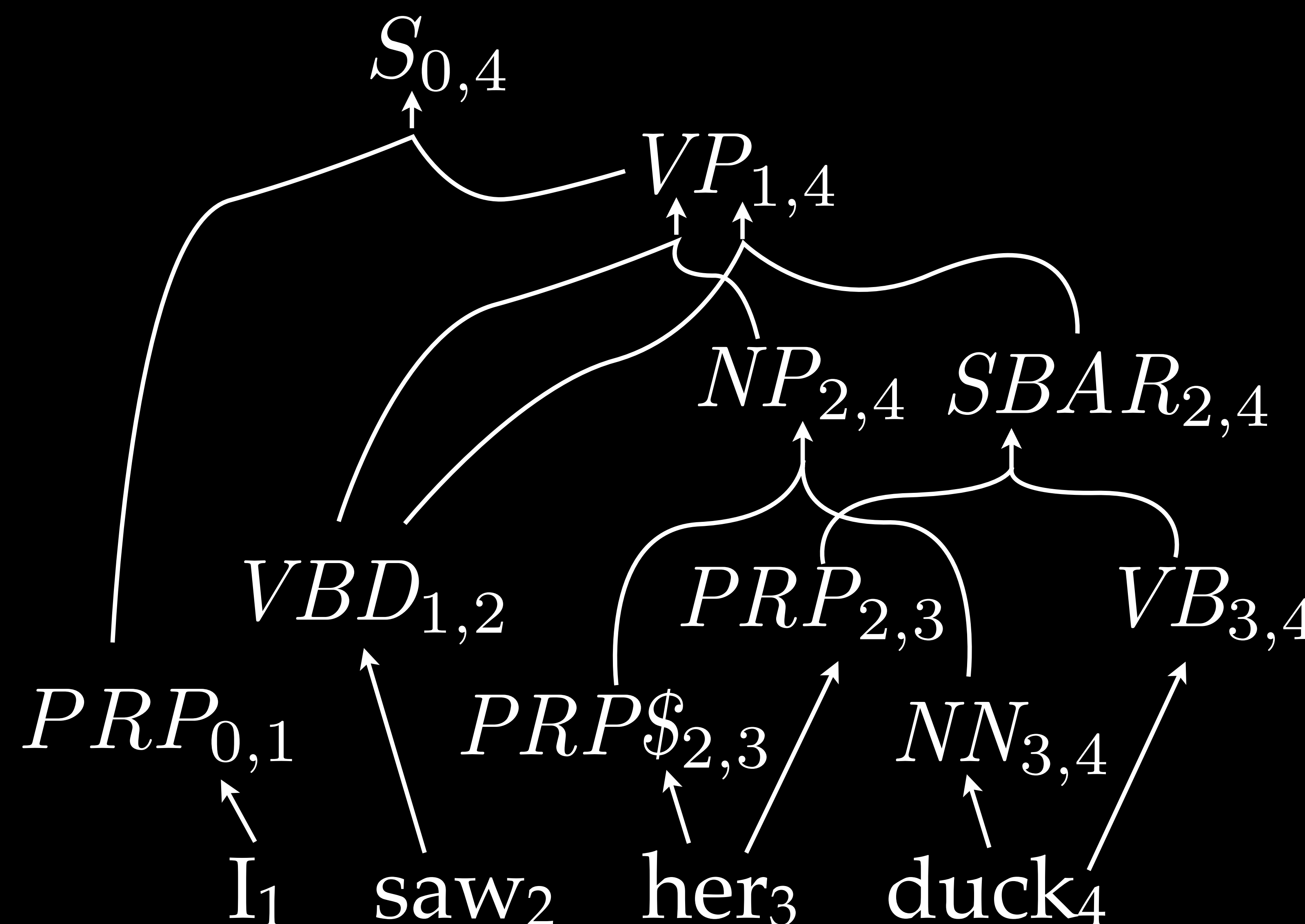
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ duck

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ saw



Parsing

Is Intersection!

$NN_{3,4} \rightarrow$ duck	$NN_{3,4} \rightarrow$ pato
$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$	$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$
$PRP_{2,3} \rightarrow$ her	$PRP_{2,3} \rightarrow$ su
$PRP_{0,1} \rightarrow$ I	$PRP_{0,1} \rightarrow$ yo
$PRP\$_{2,3} \rightarrow$ her	$PRP\$_{2,3} \rightarrow$ ella
$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$	$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$	$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$
$VB_{3,4} \rightarrow$ duck	$VB_{3,4} \rightarrow$ agacharse
$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$	$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$
$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$	$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$
$VBD_{1,2} \rightarrow$ saw	$VBD_{1,2} \rightarrow$ vi

Parsing

Is Intersection!

$NN_{3,4} \rightarrow$ duck
$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$
$PRP_{2,3} \rightarrow$ her
$PRP_{0,1} \rightarrow$ I
$PRP\$_{2,3} \rightarrow$ her
$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$
$VB_{3,4} \rightarrow$ duck
$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$
$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$
$VBD_{1,2} \rightarrow$ saw

$NN_{3,4} \rightarrow$ pato
$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$
$PRP_{2,3} \rightarrow$ su
$PRP_{0,1} \rightarrow$ yo
$PRP\$_{2,3} \rightarrow$ ella
$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$
$VB_{3,4} \rightarrow$ agacharse
$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$
$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$
$VBD_{1,2} \rightarrow$ vi

yo vi ella agacharse

yo vi su pato

Synchronous Parsing as Intersection

Synchronous Parsing as Intersection

- Parse the English sentence (intersection).

Synchronous Parsing as Intersection

- Parse the English sentence (intersection).
- Project grammar into French.

Synchronous Parsing as Intersection

- Parse the English sentence (intersection).
- Project grammar into French.
- Parse the French sentence (intersection).

Synchronous Parsing as Intersection

- Parse the English sentence (intersection).
- Project grammar into French.
- Parse the French sentence (intersection).

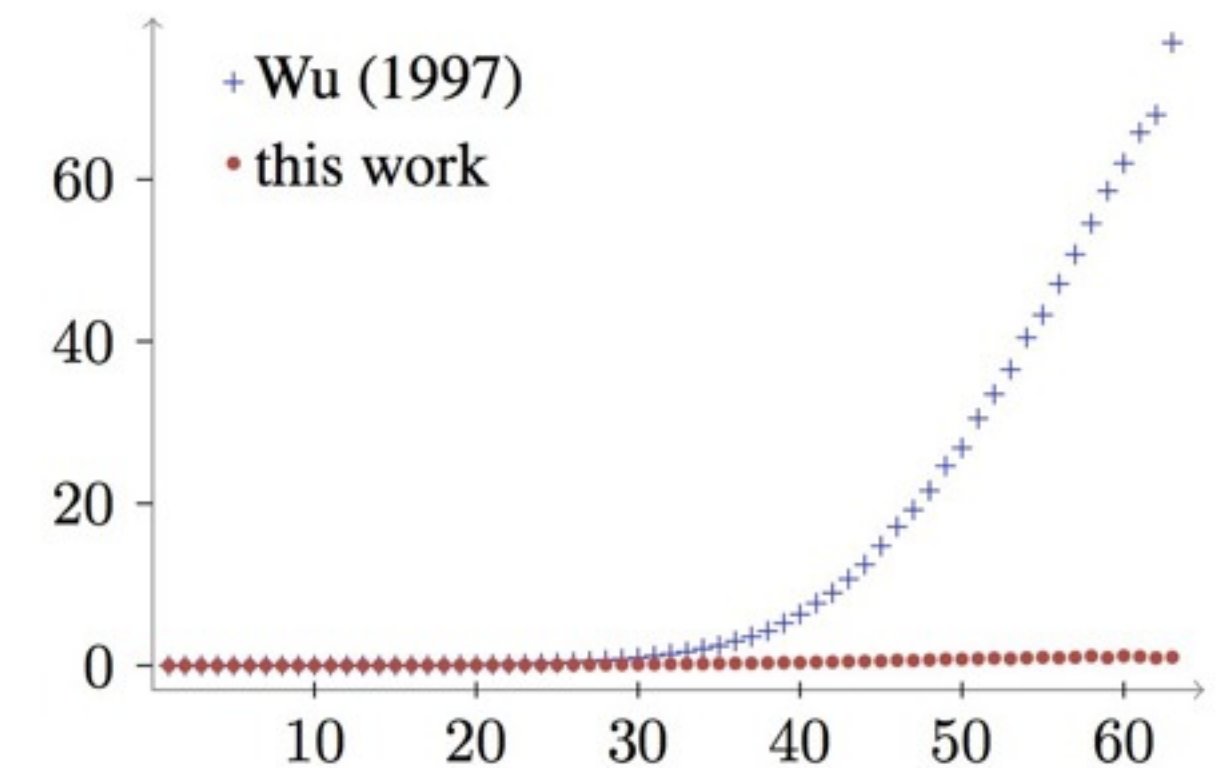


Figure 3: Average synchronous parser run-time (in seconds) as a function of Arabic sentence length (in words).

Dyer, NAACL 2010

Translation as Intersection?

$NN_{3,4} \rightarrow$ duck	$NN_{3,4} \rightarrow$ pato
$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$	$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$
$PRP_{2,3} \rightarrow$ her	$PRP_{2,3} \rightarrow$ su
$PRP_{0,1} \rightarrow$ I	$PRP_{0,1} \rightarrow$ yo
$PRP\$_{2,3} \rightarrow$ her	$PRP\$_{2,3} \rightarrow$ ella
$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$	$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$	$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$
$VB_{3,4} \rightarrow$ duck	$VB_{3,4} \rightarrow$ agacharse
$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$	$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$
$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$	$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$
$VBD_{1,2} \rightarrow$ saw	$VBD_{1,2} \rightarrow$ vi

yo vi ella agacharse

yo vi su pato

Translation as Intersection?

Observation: target grammar generates a *finite language*

$NN_{3,4} \rightarrow$ duck	$NN_{3,4} \rightarrow$ pato
$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$	$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$
$PRP_{2,3} \rightarrow$ her	$PRP_{2,3} \rightarrow$ su
$PRP_{0,1} \rightarrow$ I	$PRP_{0,1} \rightarrow$ yo
$PRP\$_{2,3} \rightarrow$ her	$PRP\$_{2,3} \rightarrow$ ella
$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$	$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$	$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$
$VB_{3,4} \rightarrow$ duck	$VB_{3,4} \rightarrow$ agacharse
$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$	$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$
$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$	$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$
$VBD_{1,2} \rightarrow$ saw	$VBD_{1,2} \rightarrow$ vi

yo vi ella agacharse

yo vi su pato

Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi

Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$\boxed{S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}}$

$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi

Translation as Intersection?

$NN_{3,4} \rightarrow \text{pato}$

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow \text{su}$

$PRP_{0,1} \rightarrow \text{yo}$

$PRP\$_{2,3} \rightarrow \text{ella}$

$\boxed{S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}}$

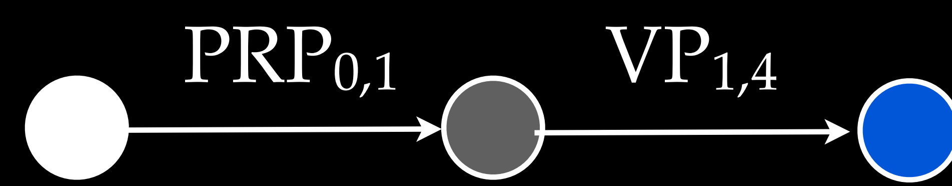
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow \text{agacharse}$

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow \text{vi}$



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

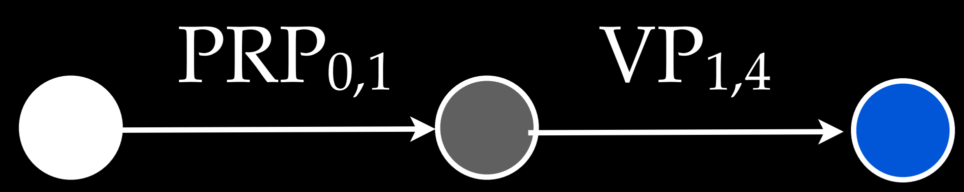
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

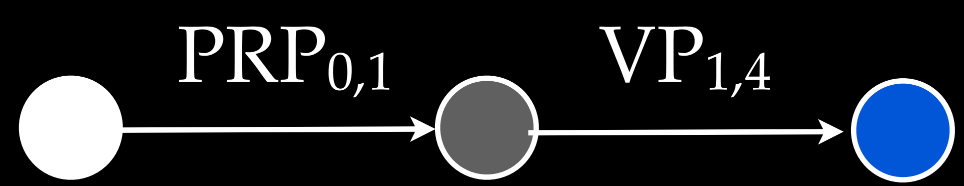
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

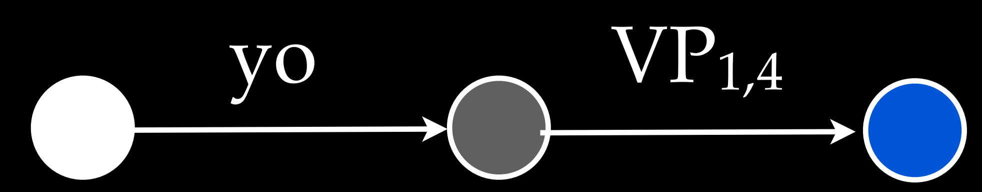
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

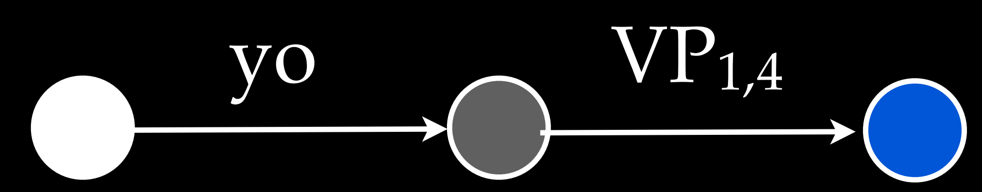
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

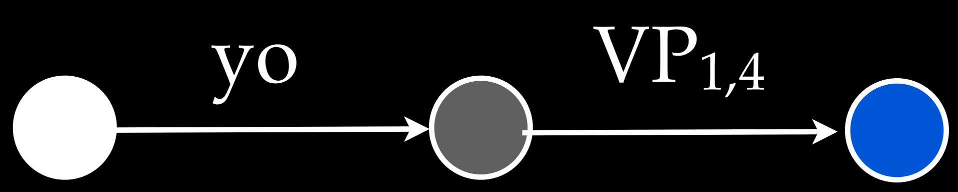
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

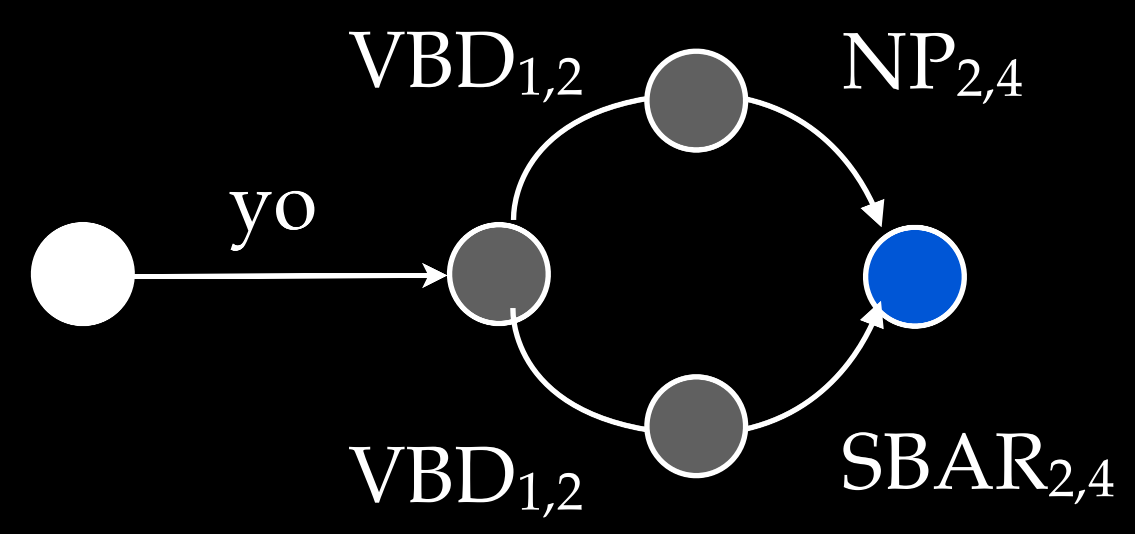
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

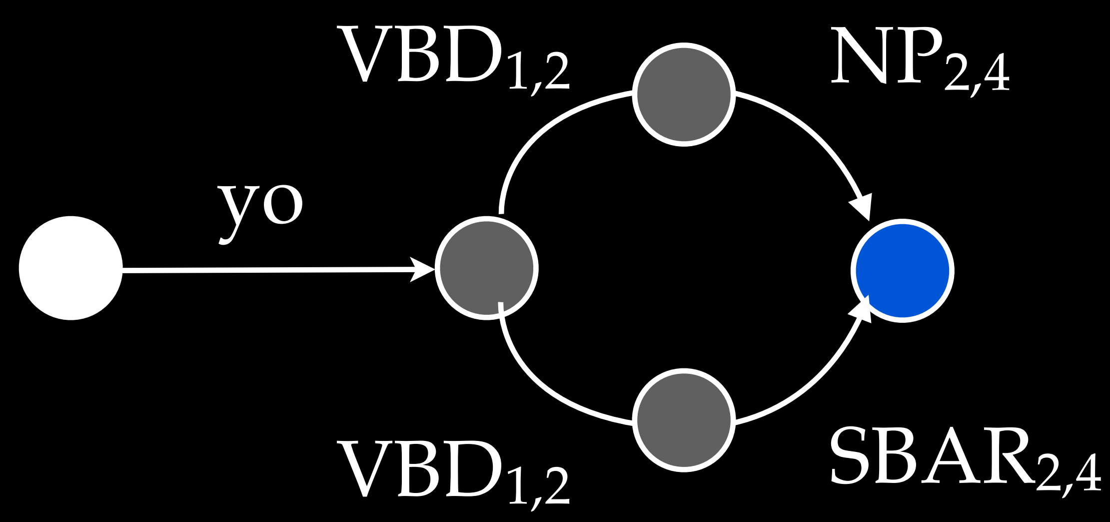
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

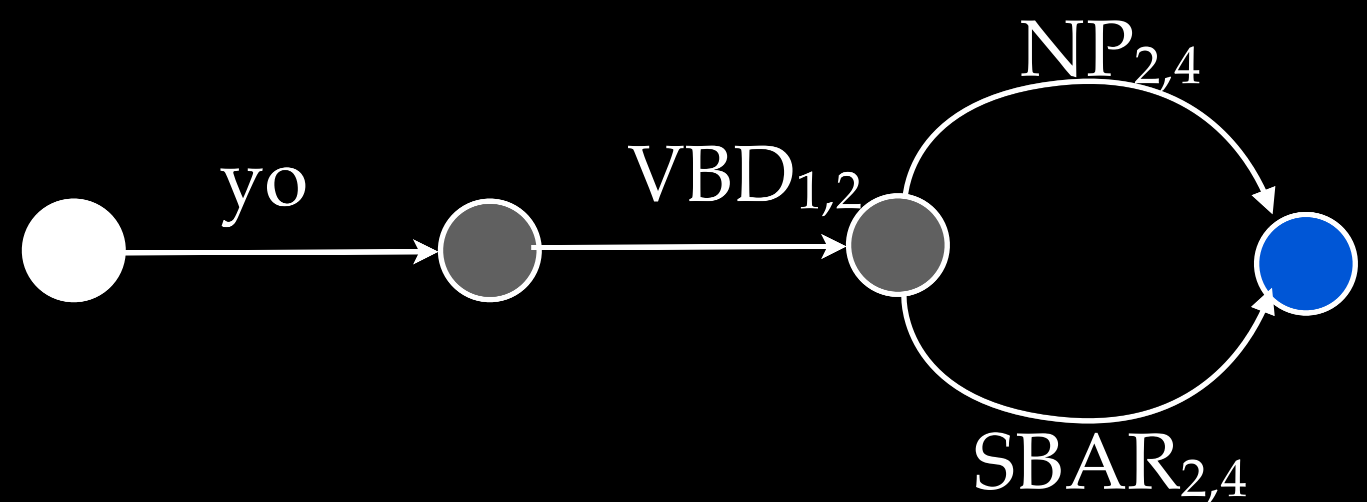
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

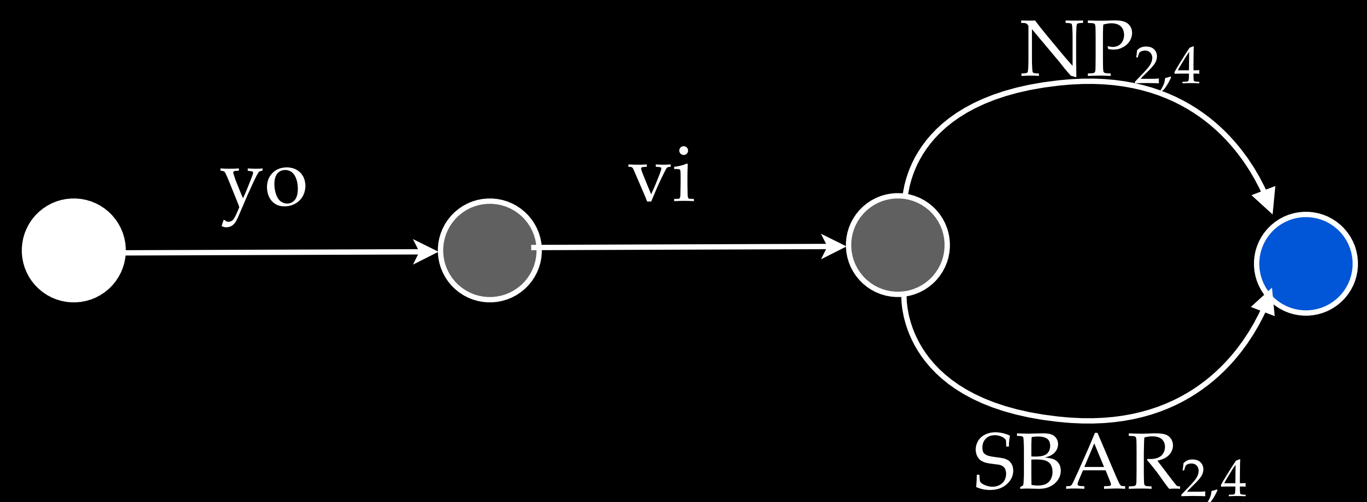
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

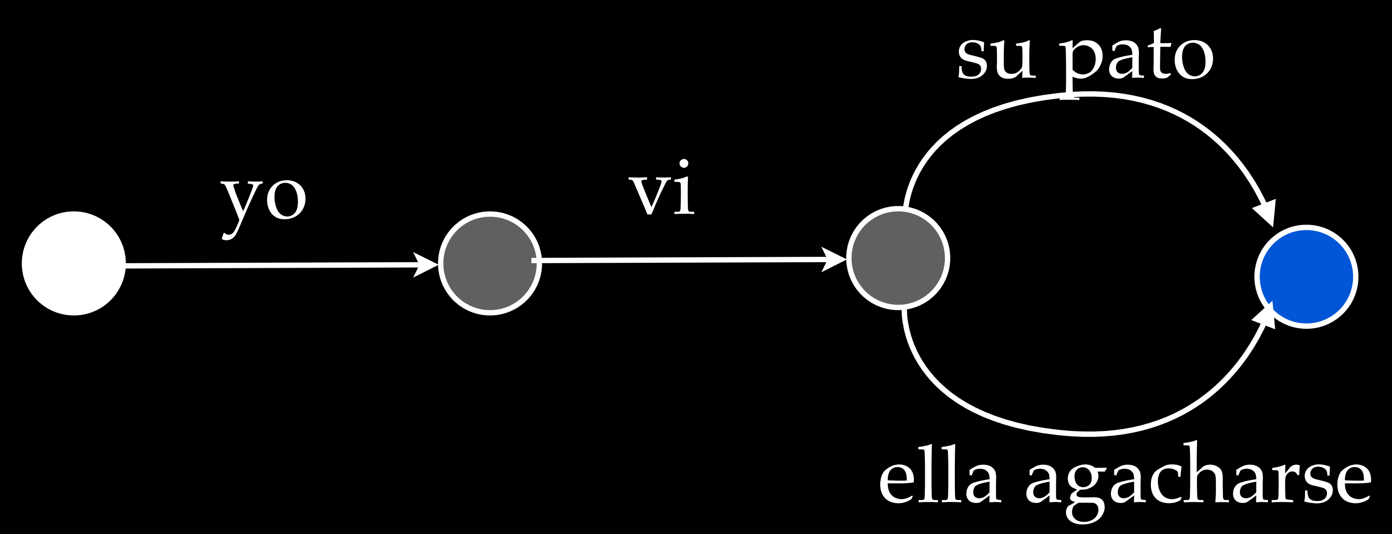
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

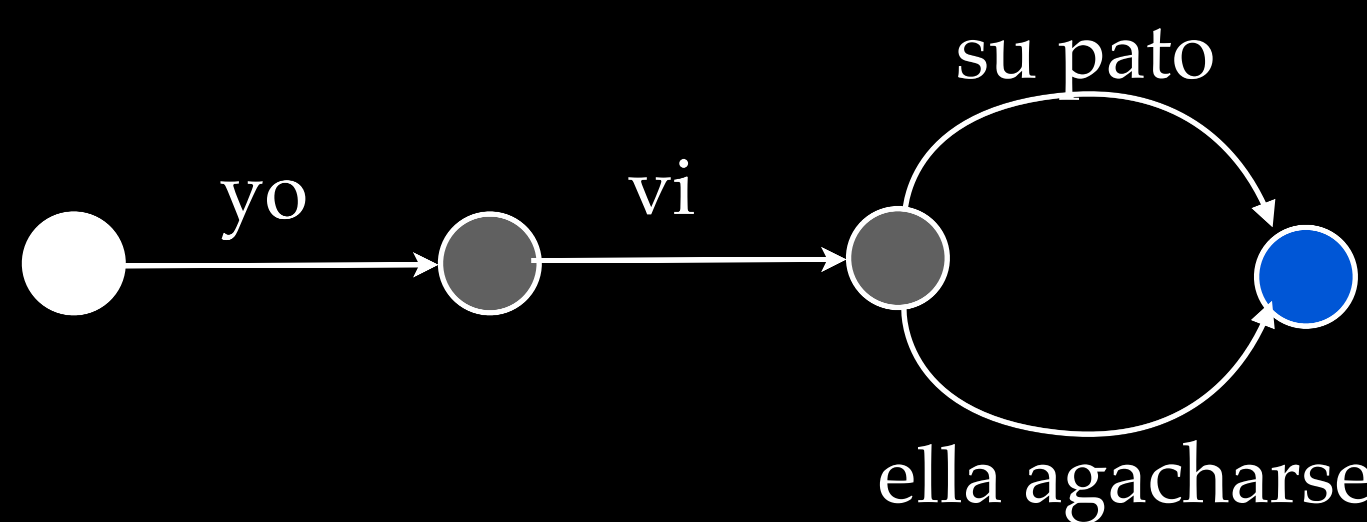
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Better: lazy algorithm

Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

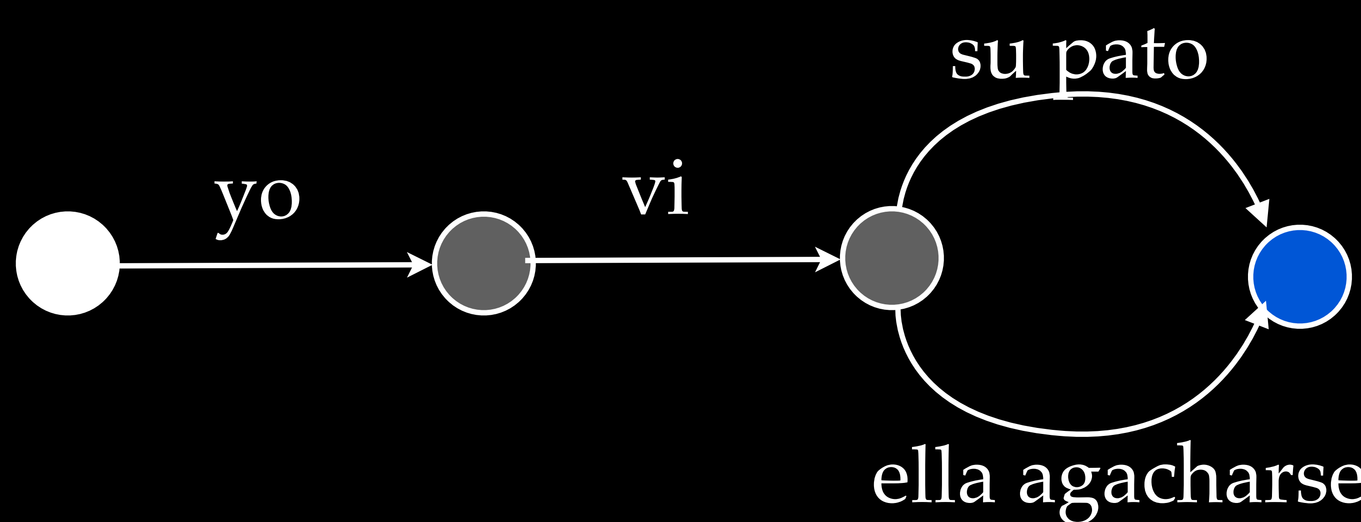
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi



Better: lazy algorithm

Even better: convert to PDA

Translation as Intersection?

$NN_{3,4} \rightarrow$ pato

$NP_{2,4} \rightarrow PRP\$_{2,3}\ NN_{3,4}$

$PRP_{2,3} \rightarrow$ su

$PRP_{0,1} \rightarrow$ yo

$PRP\$_{2,3} \rightarrow$ ella

$S_{0,4} \rightarrow PRP_{0,1}\ VP_{1,4}$

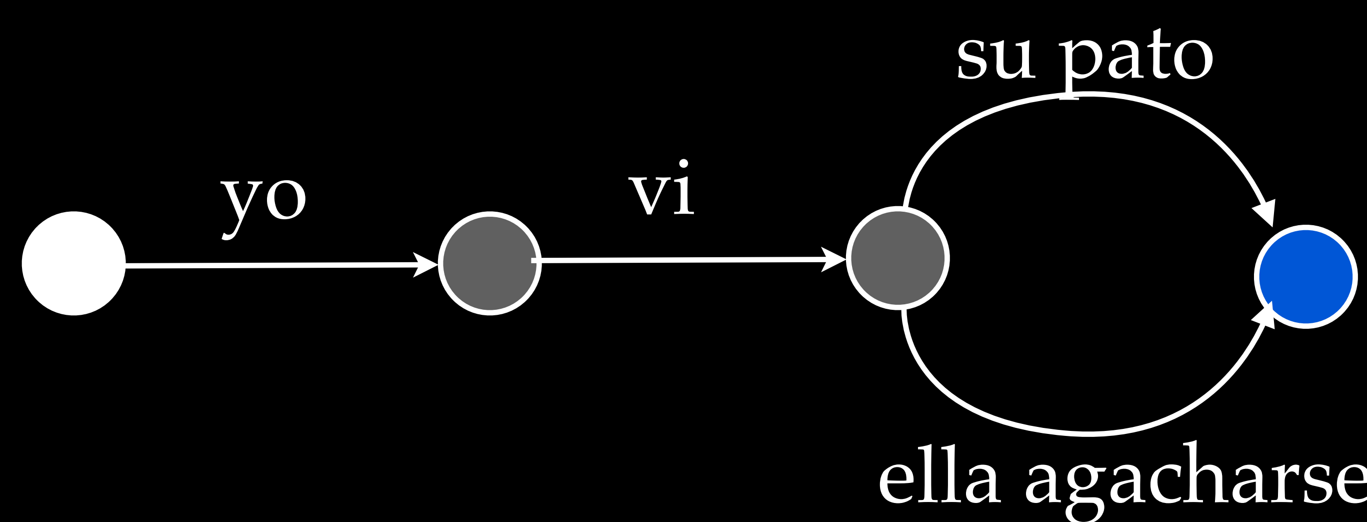
$SBAR_{2,4} \rightarrow PRP_{2,3}\ VB_{3,4}$

$VB_{3,4} \rightarrow$ agacharse

$VP_{1,4} \rightarrow VBD_{1,2}\ NP_{2,4}$

$VP_{1,4} \rightarrow VBD_{1,2}\ SBAR_{2,4}$

$VBD_{1,2} \rightarrow$ vi

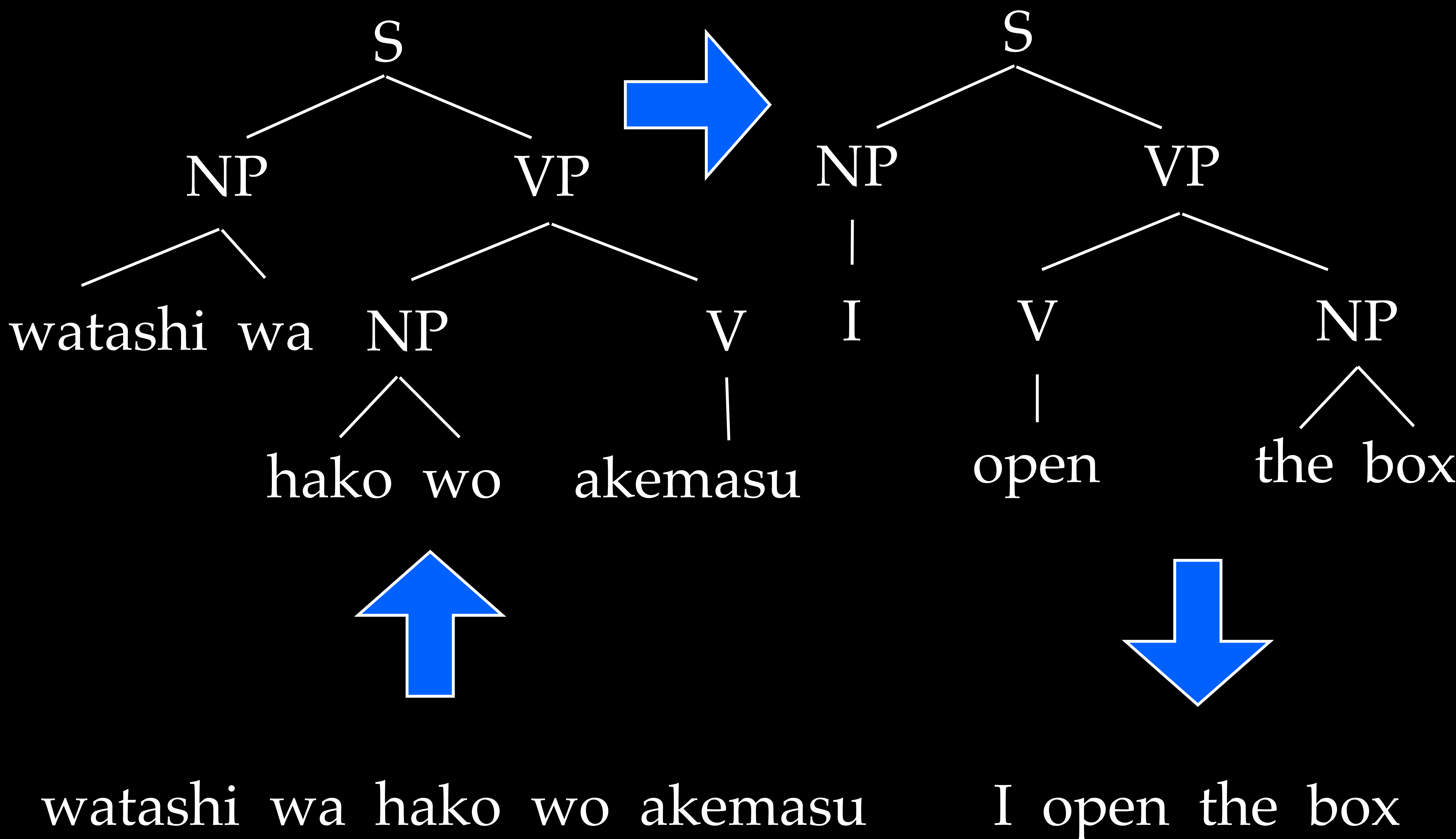


Better: lazy algorithm

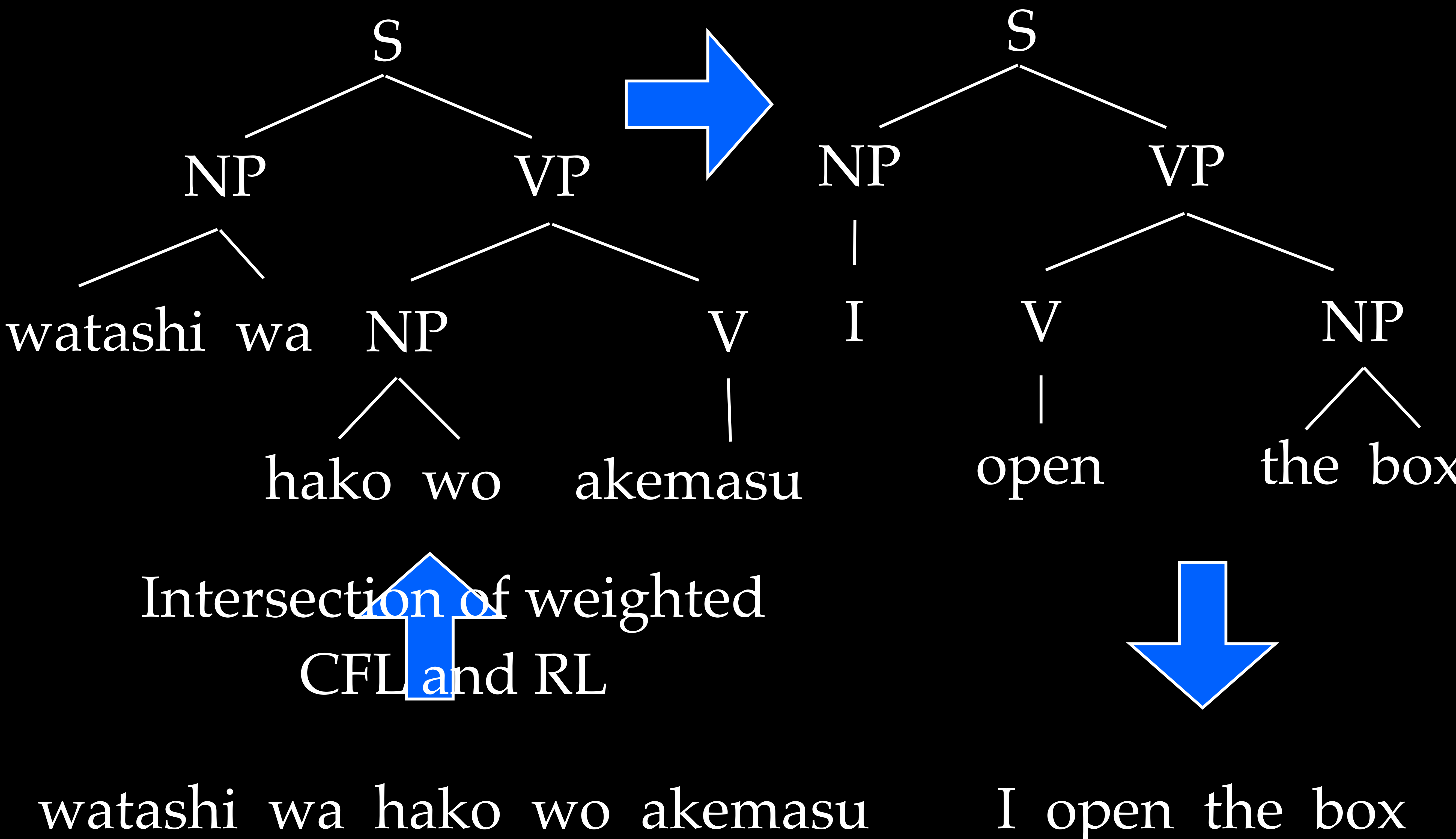
Even better: convert to PDA

Cambridge: best NIST 2009 Arabic system

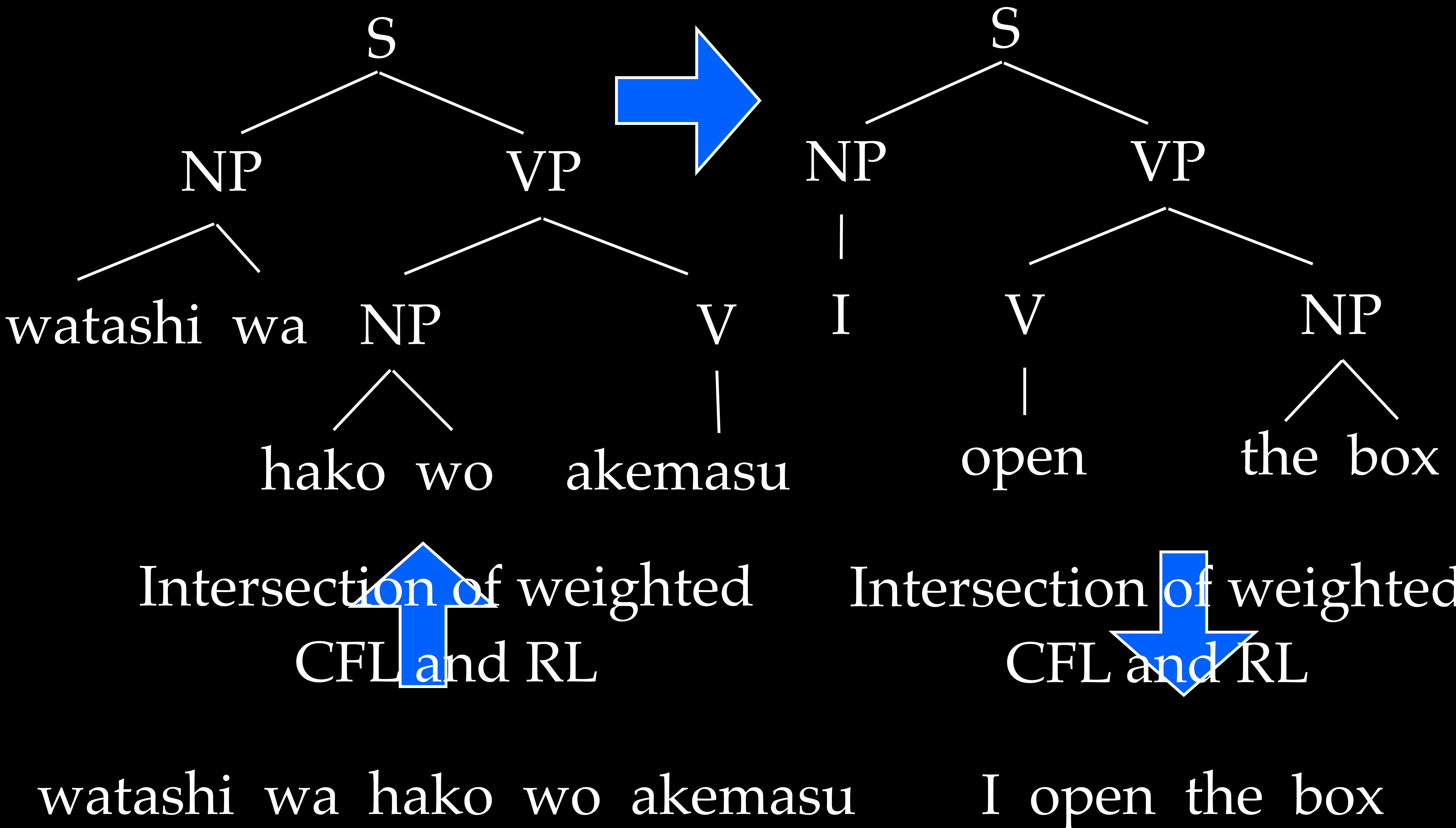
Translation as Intersection?



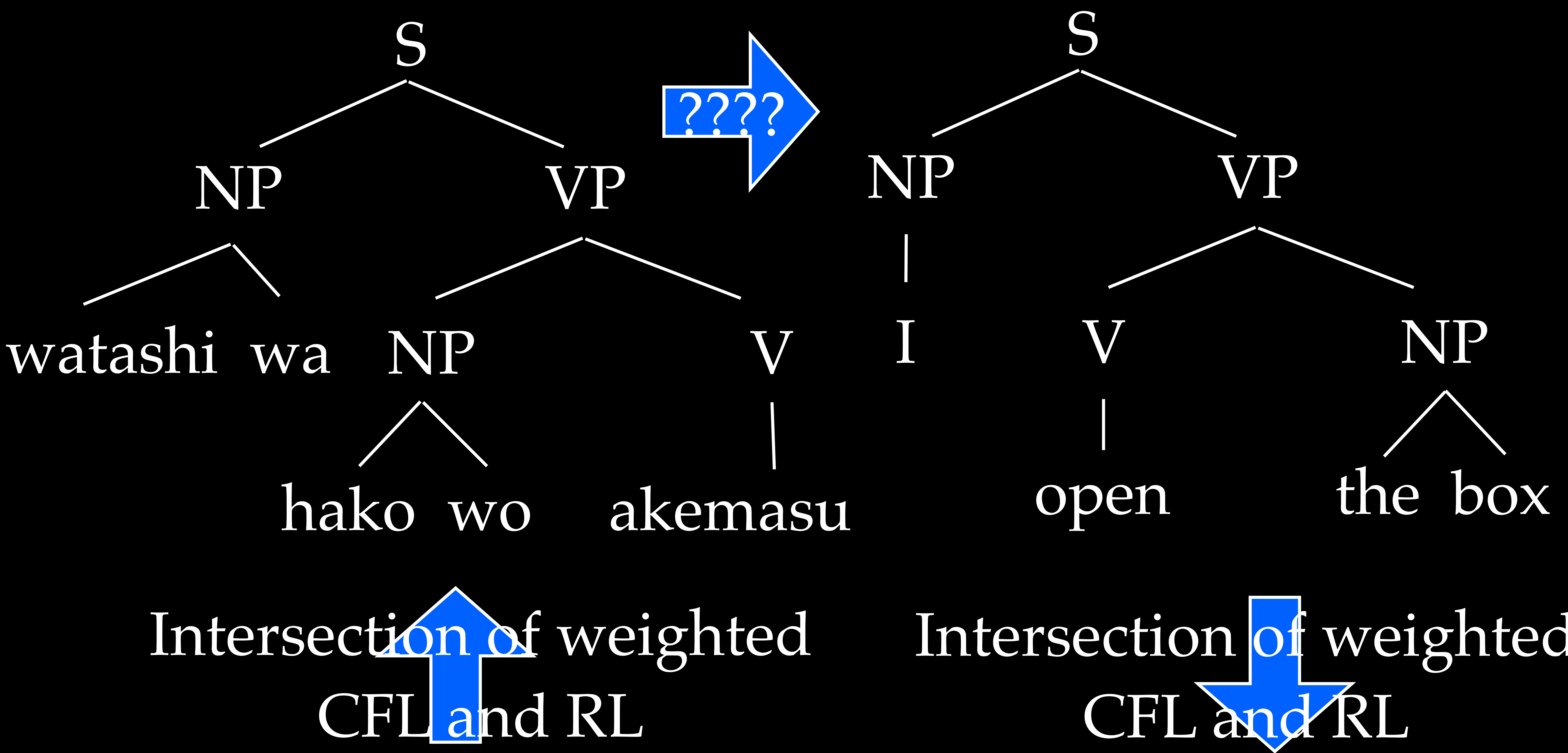
Translation as Intersection?



Translation as Intersection?



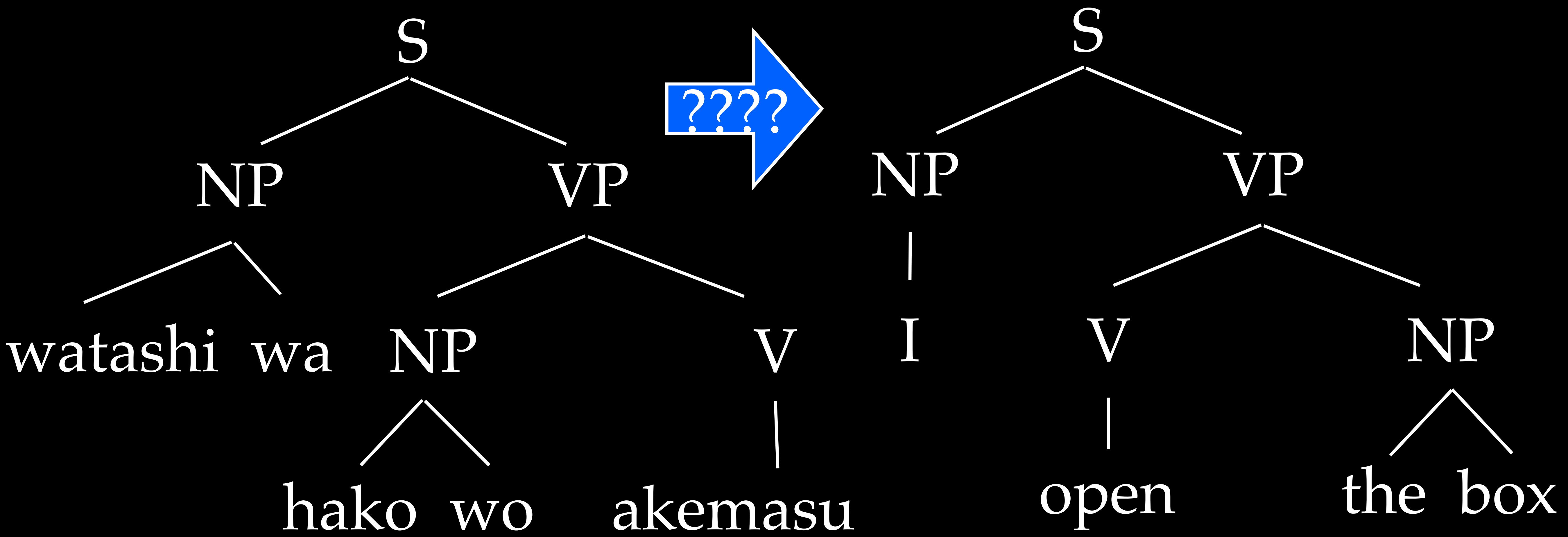
Translation as Intersection?



watashi wa hako wo akemasu

I open the box

Translation as Intersection?



Intersection of weighted CFL and RL

Intersection of weighted CFL and RL

watashi wa hako wo akemasu

I open the box

Weighted *tree* languages, automata, and transducers.

Bayes' Rule

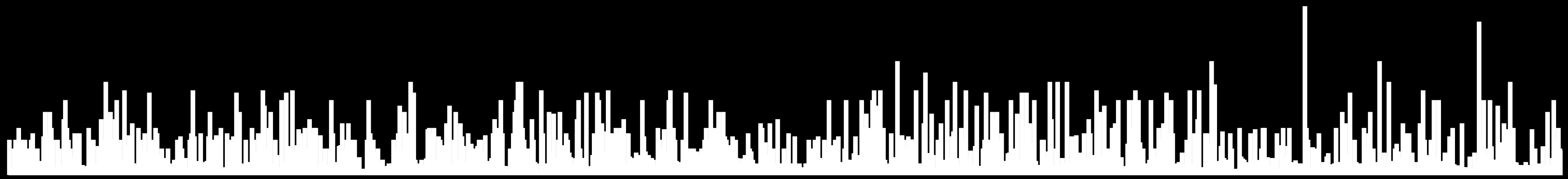
$$p(English|Chinese) \sim$$

$$p(English) \times p(Chinese|English)$$

language model

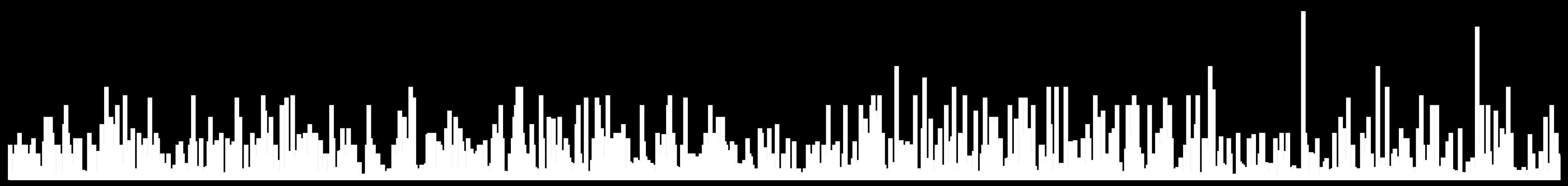
translation model

$$p(Chinese|English)$$

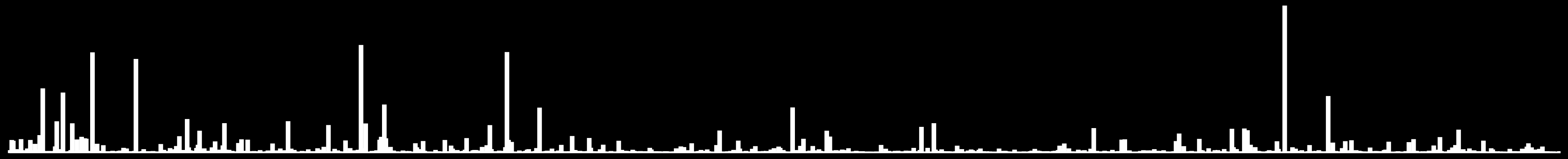


$English$

$$p(Chinese|English)$$

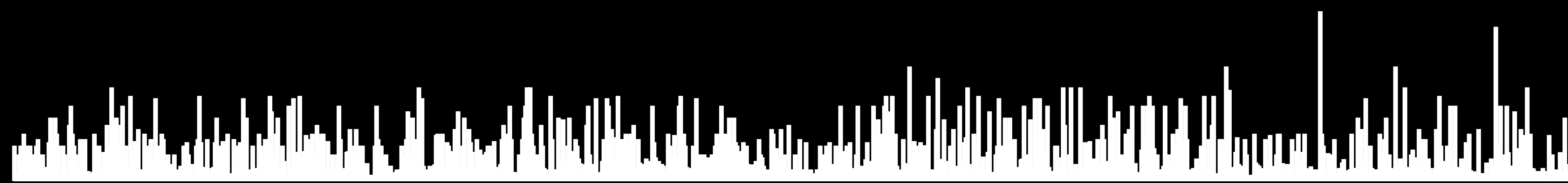


$$\times\ p(English)$$

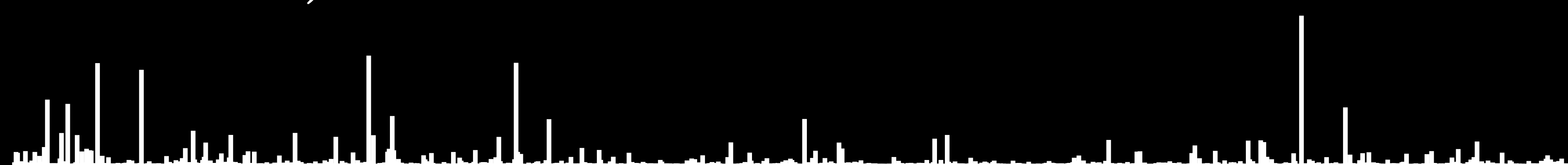


$$English$$

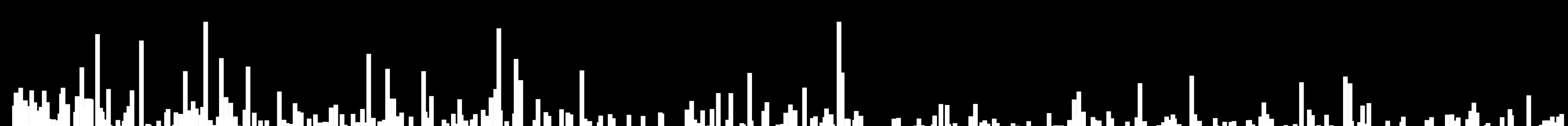
$$p(Chinese|English)$$



$$\times\ p(English)$$

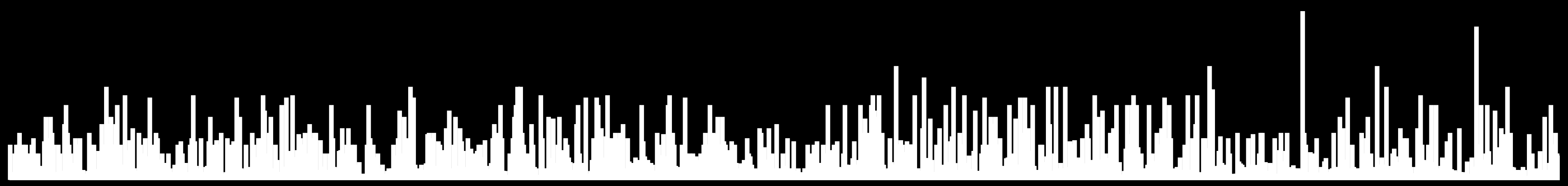


$$\sim p(English|Chinese)$$

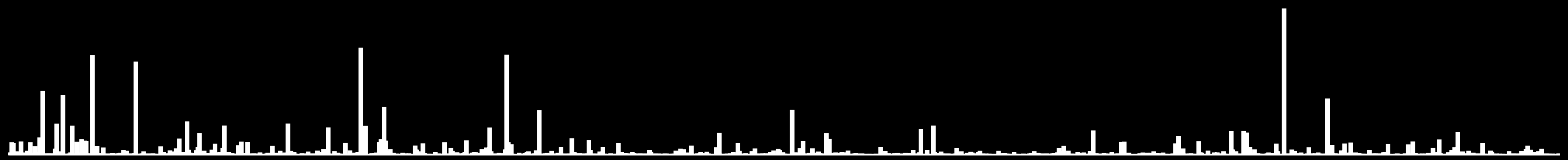


$English$

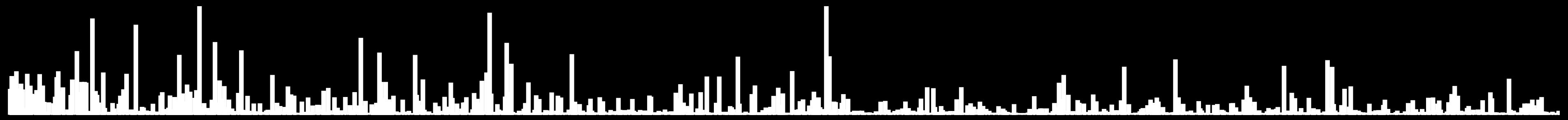
$$p(Chinese|English)^1$$



$$\times\ p(English)^1$$

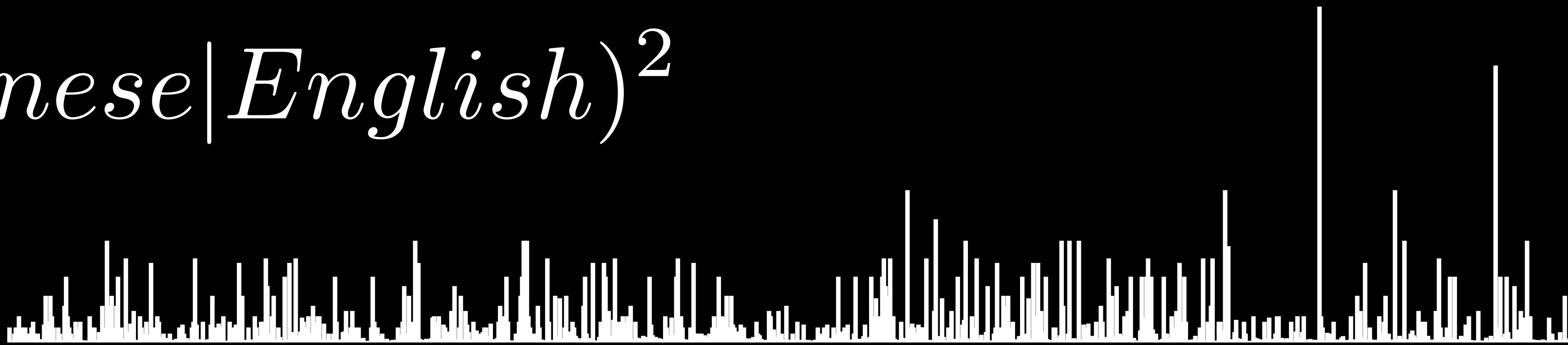


$$\sim p(English|Chinese)$$

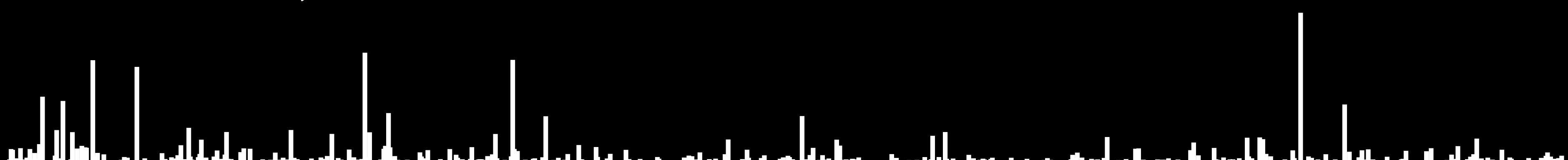


$English$

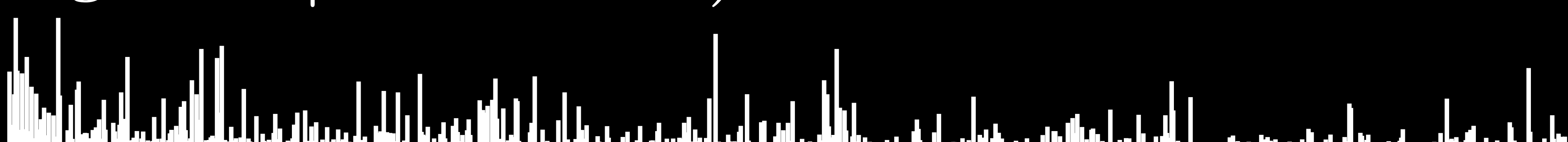
$$p(Chinese|English)^2$$



$$\times\ p(English)^1$$

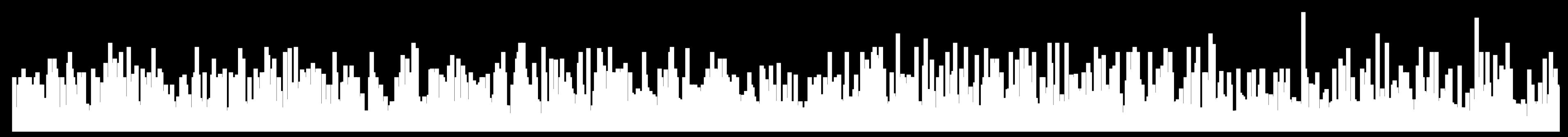


$$\sim p(English|Chinese)$$

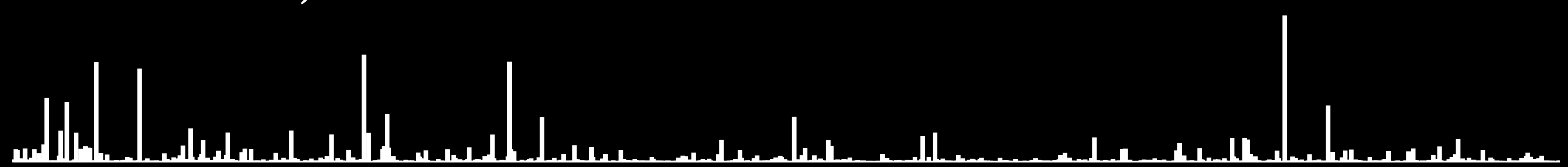


$English$

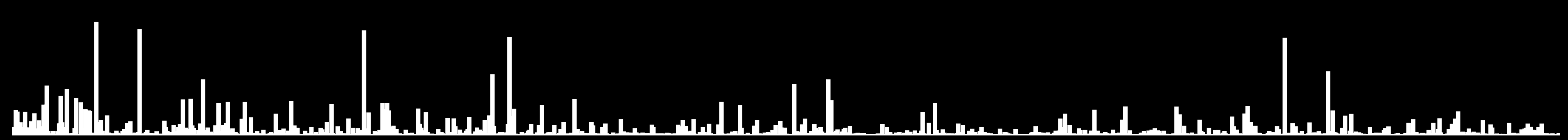
$$p(Chinese|English)^{1/2}$$



$$\times\ p(English)^{1}$$



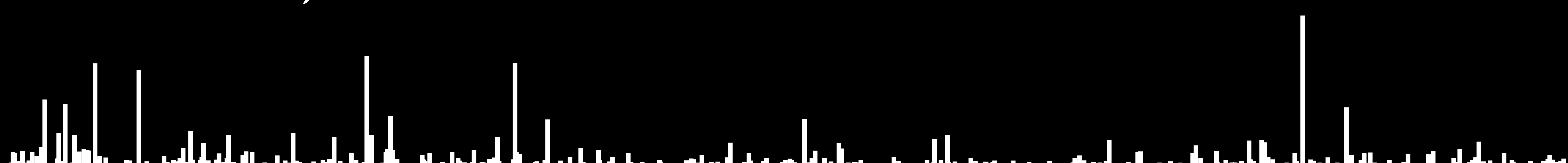
$$\sim p(English|Chinese)$$



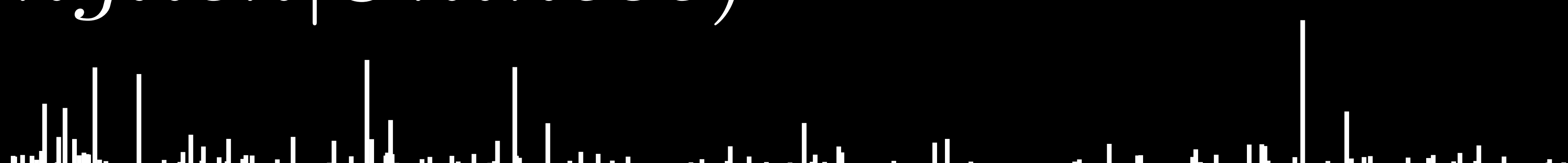
$English$

$$p(Chinese|English)^0$$

$$\times\ p(English)^1$$



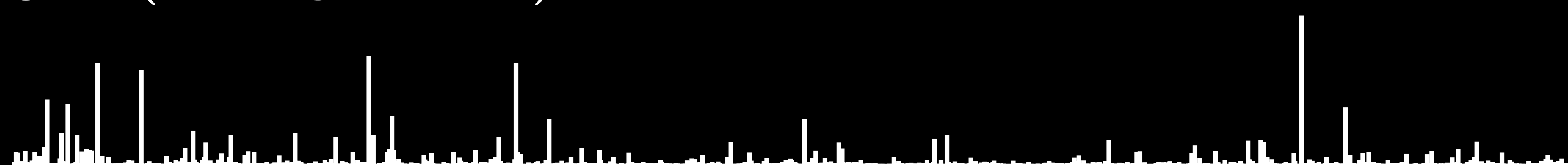
$$\sim p(English|Chinese)$$



$English$

$$0 \cdot \log p(Chinese|English)$$

$$+1 \cdot \log p(English)$$



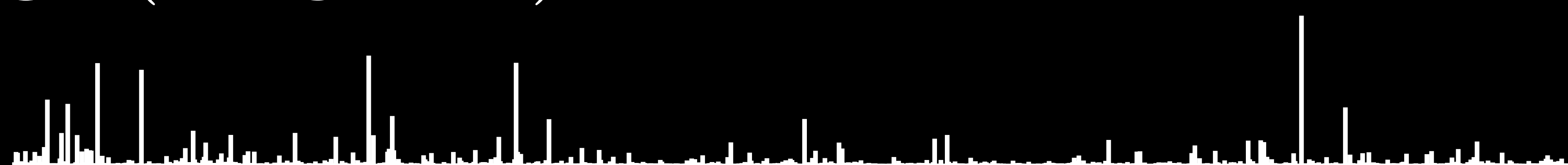
$$\sim \log p(English|Chinese)$$

$English$

$$\log(a) < \log(b) \iff a < b$$

$$0 \cdot \log p(Chinese|English)$$

$$+1 \cdot \log p(English)$$

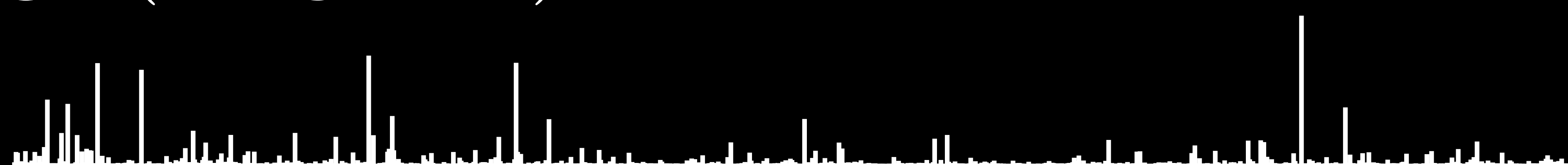


$$\sim \log p(English|Chinese)$$

$English$

$$0 \cdot \log p(Chinese|English)$$

$$+1 \cdot \log p(English)$$



$$= score(English|Chinese)$$

$English$

$$score(English|Chinese) =$$

$$\lambda_1 \log p(Chinese|English) + \lambda_2 \log p(English)$$

$$score(English|Chinese) =$$

$$\exp(\lambda_1 \log p(Chinese|English) + \lambda_2 \log p(English))$$

$$p(English|Chinese) = \frac{\exp(\lambda_1 \log p(Chinese|English) + \lambda_2 \log p(English))}{\sum_{nglish} \exp(\lambda_1 \log p(Chinese|English) + \lambda_2 \log p(English}$$

$$p(English|Chinese) =$$

$$\frac{\exp(\lambda_1 \log p(Chinese|English) + \lambda_2 \log p(English))}{\sum_{nglish} \exp(\lambda_1 \log p(Chinese|English) + \lambda_2 \log p(English)}$$

log-linear model

maximum entropy model

conditional model

undirected model

$$p(English|Chinese) =$$

$$p(English) \times p(Chinese|English)$$

Note: Original model is a special case of this model!

log-linear model
maximum entropy model
conditional model
undirected model

$$p(English|Chinese) =$$

$$\frac{\exp(\lambda_1 \log p(Chinese|English) + \lambda_2 \log p(English))}{\sum_{nglish} \exp(\lambda_1 \log p(Chinese|English) + \lambda_2 \log p(English)}$$

log-linear model

maximum entropy model

conditional model

undirected model

$$p(English|Chinese) = \frac{\exp\left\{\sum_k \lambda_k h_k(English, Chinese)\right\}}{\sum_{English'} \exp\left\{\sum_k \lambda_k h_k(English', Chinese)\right\}}$$

log-linear model
maximum entropy model
conditional model
undirected model

Spring 2014
mt-class.org