Coarse-to-Fine Syntactic MT using Language Projections

Slav Petrov Aria Haghighi Dan Klein

Why is MT so slow?

Why is MT so slow?

quiero salir

Why is MT so slow?

quiero

- ➡ i want
- ➡ i like
- ➡ i love
- ➡ i want to
- ➡ i like to
- ➡ i love to

salir

- ➡ leave
- ➡ to leave
- ➡ go
- ➡ to go
- ➡ exit
- ➡ to exit

Why is MT so slow?

quiero

- ➡ i want
- ➡ i like
- ➡ i love
- ➡ i want to
- ➡ i like to
- ➡ i love to

salir

- ➡ leave
- ➡ to leave
- ➡ go
- ➡ to go
- ➡ exit
- ➡ to exit

Translations:

Why is MT so slow?

quiero	salir
➡ i want	➡ leave
➡ i like	➡ to leave
➡ i love	➡ go
➡ i want to	➡ to go
➡ i like to	➡ exit
➡ i love to	➡ to exit

Translations:

Why is MT so slow?

quiero	salir
➡ i want	➡ leave
➡ i like	➡ to leave
➡ i love	➡ go
➡ i want to	➡ to go
➡ i like to	➡ exit
➡ i love to	➡ to exit

Translations:

Why is MT so slow?

quiero

- ➡ i want
- ➡ i like
- ➡ i love
- ➡ i want to
- ➡ i like to
- ➡ i love to

salir

- ➡ leave
- ➡ to leave
- ➡ go
- ➡ to go
- ➡ exit
- ➡ to exit

Translations:

- i want leave ✗

Why is MT so slow?

quiero

- ➡ i want
- ➡ i like
- ➡ i love
- ➡ i want to
- ➡ i like to
- ➡ i love to

salir

- ➡ leave
- ➡ to leave
- ➡ go
- ➡ to go
- ➡ exit
- ➡ to exit

Translations:

- i want leave ✗
- i like leave ✗

Why is MT so slow?

quiero

- ➡ i want
- ➡ i like
- ➡ i love
- ➡ i want to
- ➡ i like to
- ➡ i love to

salir

- ➡ leave
- ➡ to leave
- ➡ go
- ➡ to go
- ➡ exit
- ➡ to exit

Translations:

- i want leave ✗
- i like leave ✗
- i love leave ✗

Why is MT so slow?

quiero

- ➡ i want
- ➡ i like
- ➡ i love
- ➡ i want to
- ➡ i like to
- ➡ i love to

salir

- ➡ leave
- ➡ to leave
- ➡ go
- ➡ to go
- ➡ exit
- ➡ to exit

Translations:

- i want leave ✗
- i like leave ✗
- i love leave ✗
- i want to leave ✓

Why is MT so slow?

quiero

- ➡ i want
- ➡ i like
- ➡ i love
- ➡ i want to
- ➡ i like to
- ➡ i love to

salir

- ➡ leave
- ➡ to leave
- ➡ go
- ➡ to go
- ➡ exit
- ➡ to exit

Translations:

- i want leave ✗
- i like leave ✗
- i love leave ✗
- i want to leave ✓
- i like to leave ✗
- i love to leave ✗
- i want to leave ✗
- i like to leave ✗
- ...

Why is MT so slow?

There are too many words!

quiero

- ➡ i want
- ➡ i like
- ➡ i love
- ➡ i want to
- ➡ i like to
- ➡ i love to

salir

- ➡ leave
- ➡ to leave
- ➡ go
- ➡ to go
- ➡ exit
- ➡ to exit

Translations:

- i want leave ✗
- i like leave ✗
- i love leave ✗
- i want to leave ✓
- i like to leave ✗
- i love to leave ✗
- i want to leave ✗
- i like to leave ✗
- ...

Why is MT so slow?

There are too many words!

quiero

- ➡ i want
- ➡ i like
- ➡ i love
- ➡ i want to
- ➡ i like to
- ➡ i love to

salir

- ➡ leave
- ➡ to leave
- ➡ go
- ➡ to go
- ➡ exit
- ➡ to exit

Translations:

- i want leave ✗
- i like leave ✗
- i love leave ✗
- i want to leave ✓
- i like to leave ✗
- i love to leave ✗
- i want to leave ✗
- i like to leave ✗
- ...

Why is MT so slow?

There are too many words!

quiero

- ➡ i want
- ➡ i like
- ➡ i love
- ➡ i want to
- ➡ i like to
- ➡ i love to

salir

- ➡ leave
- ➡ to leave
- ➡ go
- ➡ to go
- ➡ exit
- ➡ to exit

Translations:

- i want leave ✗
- i like leave ✗
- i love leave ✗
- i want to leave ✓
- i like to leave ✗
- i love to leave ✗
- i want to leave ✗
- i like to leave ✗
- ...

Coarse-to-Fine Decoding

Spanish: reanudación del período de sesiones

Coarse-to-Fine Decoding

Spanish: reanudación del período de sesiones

English:

Coarse-to-Fine Decoding

Spanish: reanudación del período de sesiones

English:

Coarse-to-Fine Decoding

Spanish: reanudación del período de sesiones

English:

resumption	of	the	session

Coarse-to-Fine Decoding

Spanish: reanudación del período de sesiones

English:

resumption	of	the	session

Coarse-to-Fine Decoding

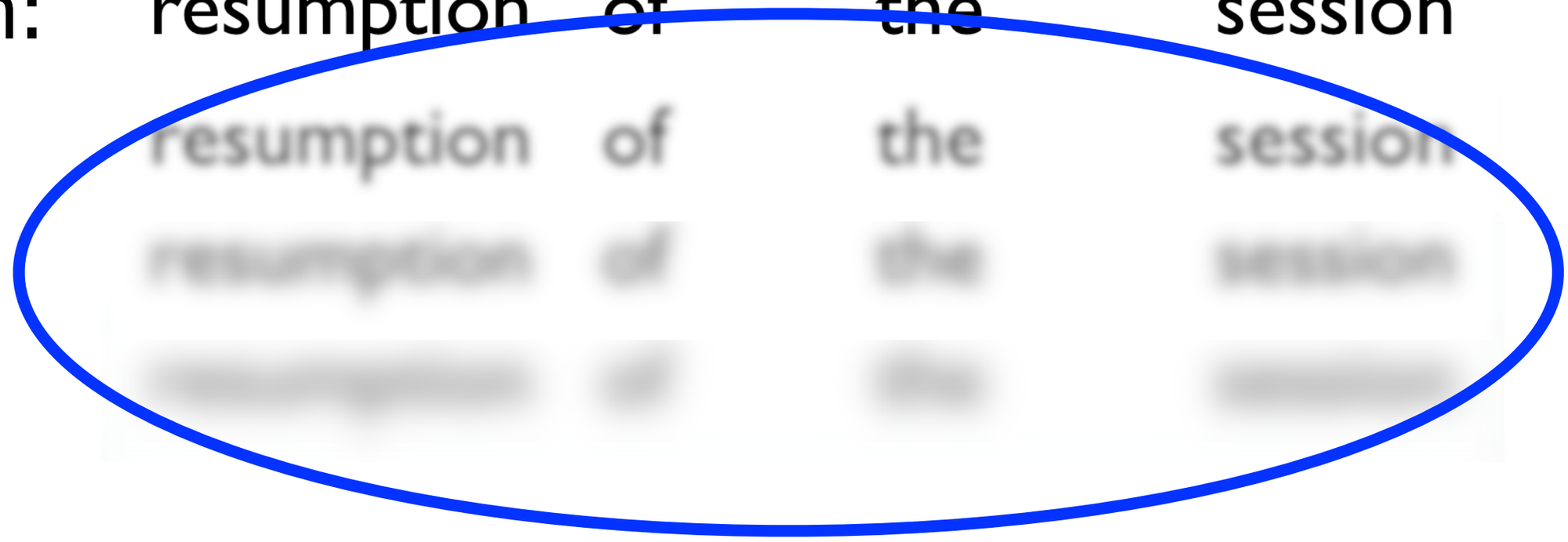
Spanish: reanudación del período de sesiones

English: resumption of the session

Coarse-to-Fine Decoding

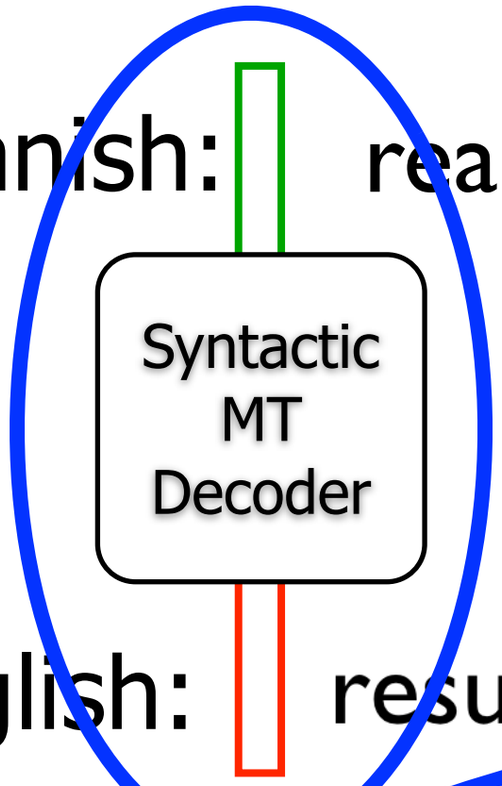
Spanish: reanudación del período de sesiones

English: resumption of the session



Coarse-to-Fine Decoding

Spanish: reanudación del período de sesiones



English: resumption of the session

resumption of the session

resumption of the session

resumption of the session

Coarse Pass

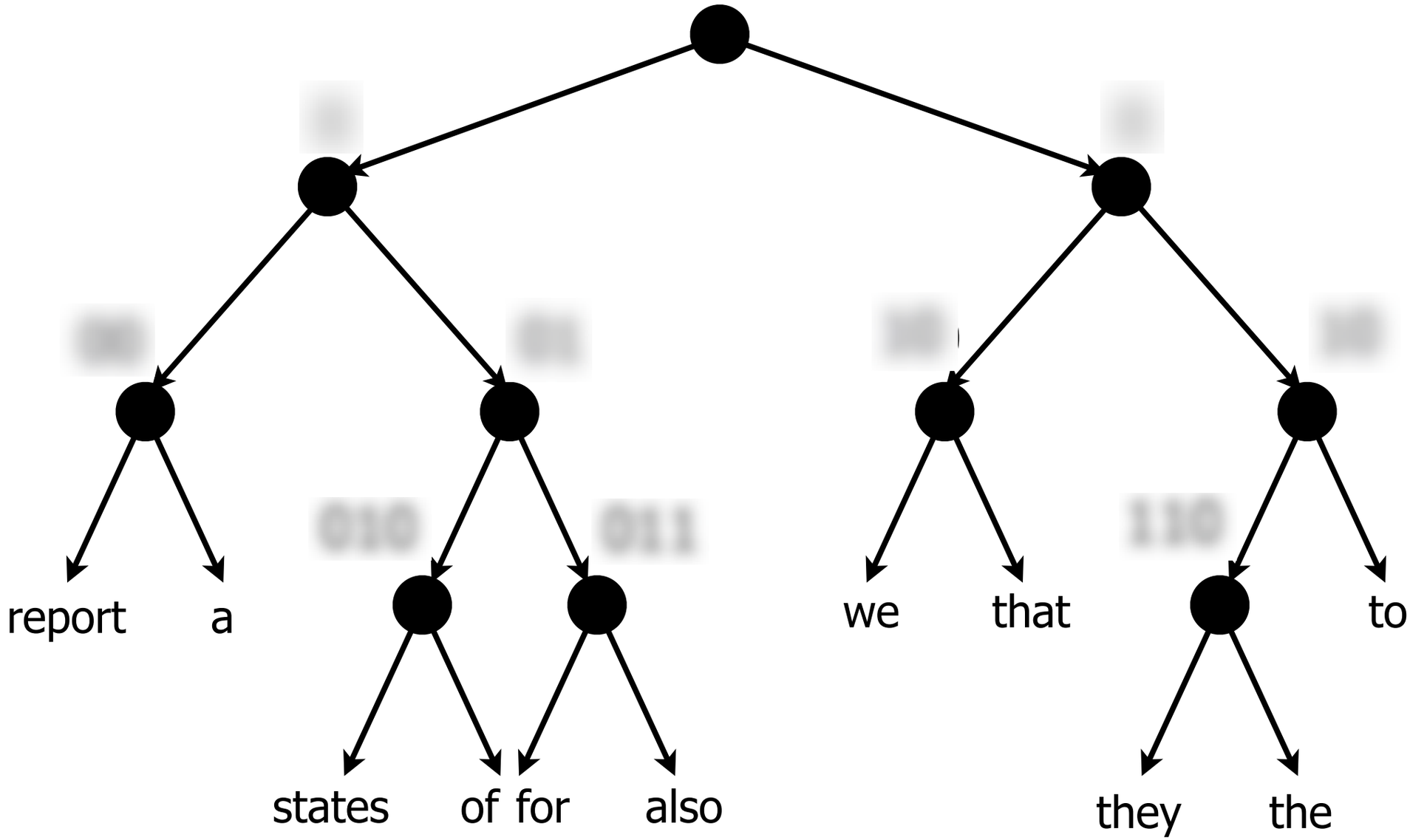
Coarse-to-Fine Syntactic MT using Language Projections

Slav Petrov Aria Haghighi Dan Klein

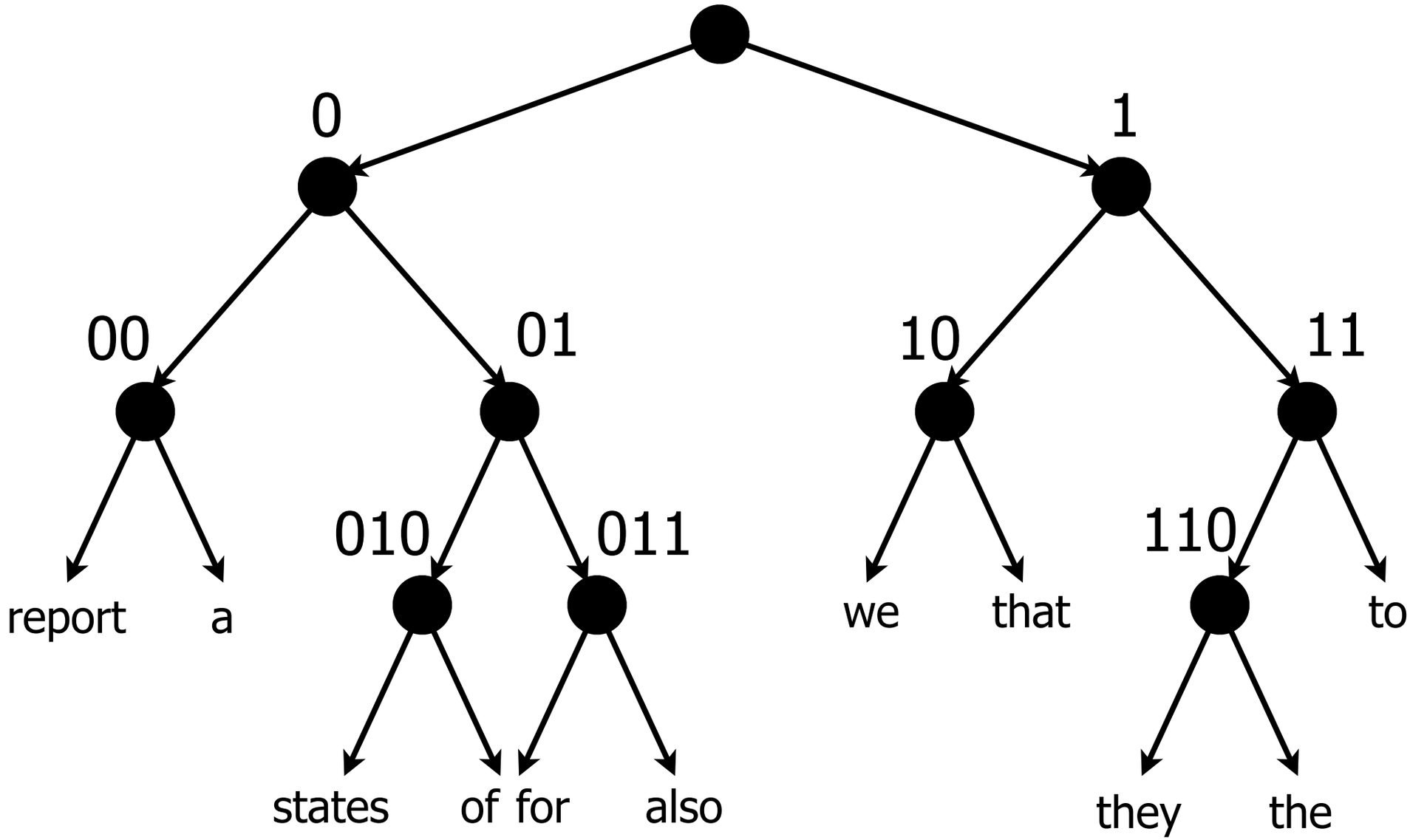
Second Pass

Fine Pass

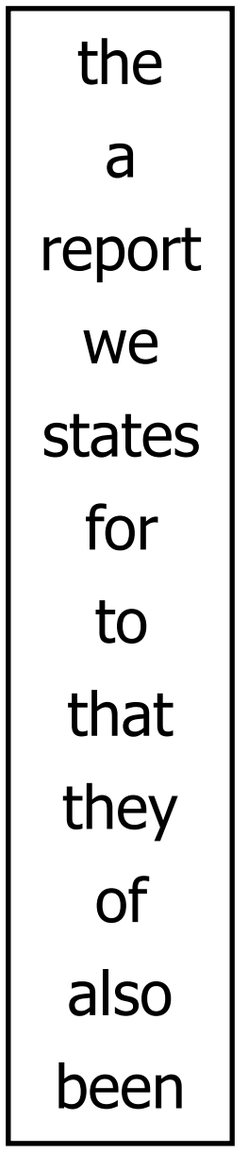
Hierarchical Clustering



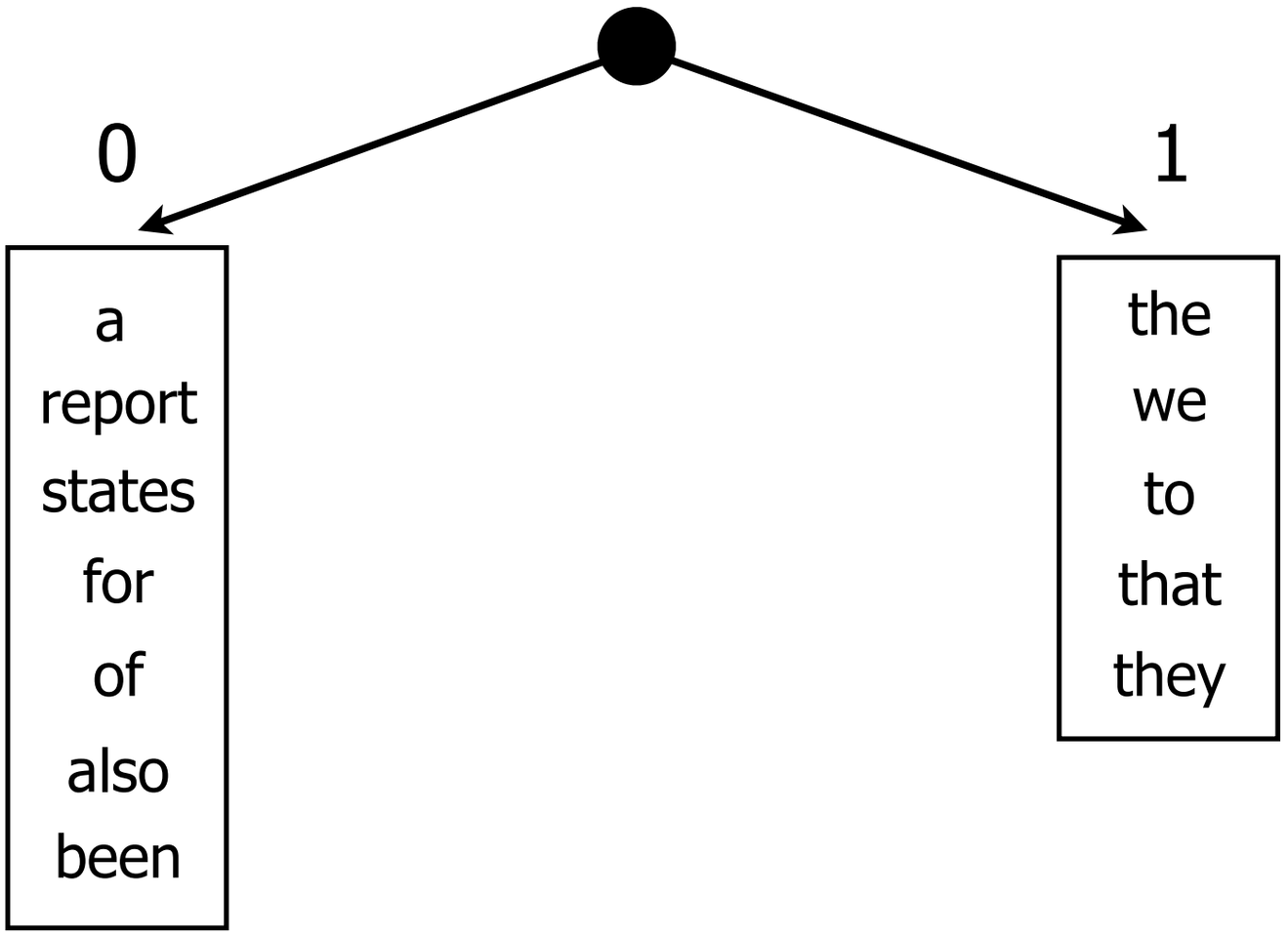
Hierarchical Clustering



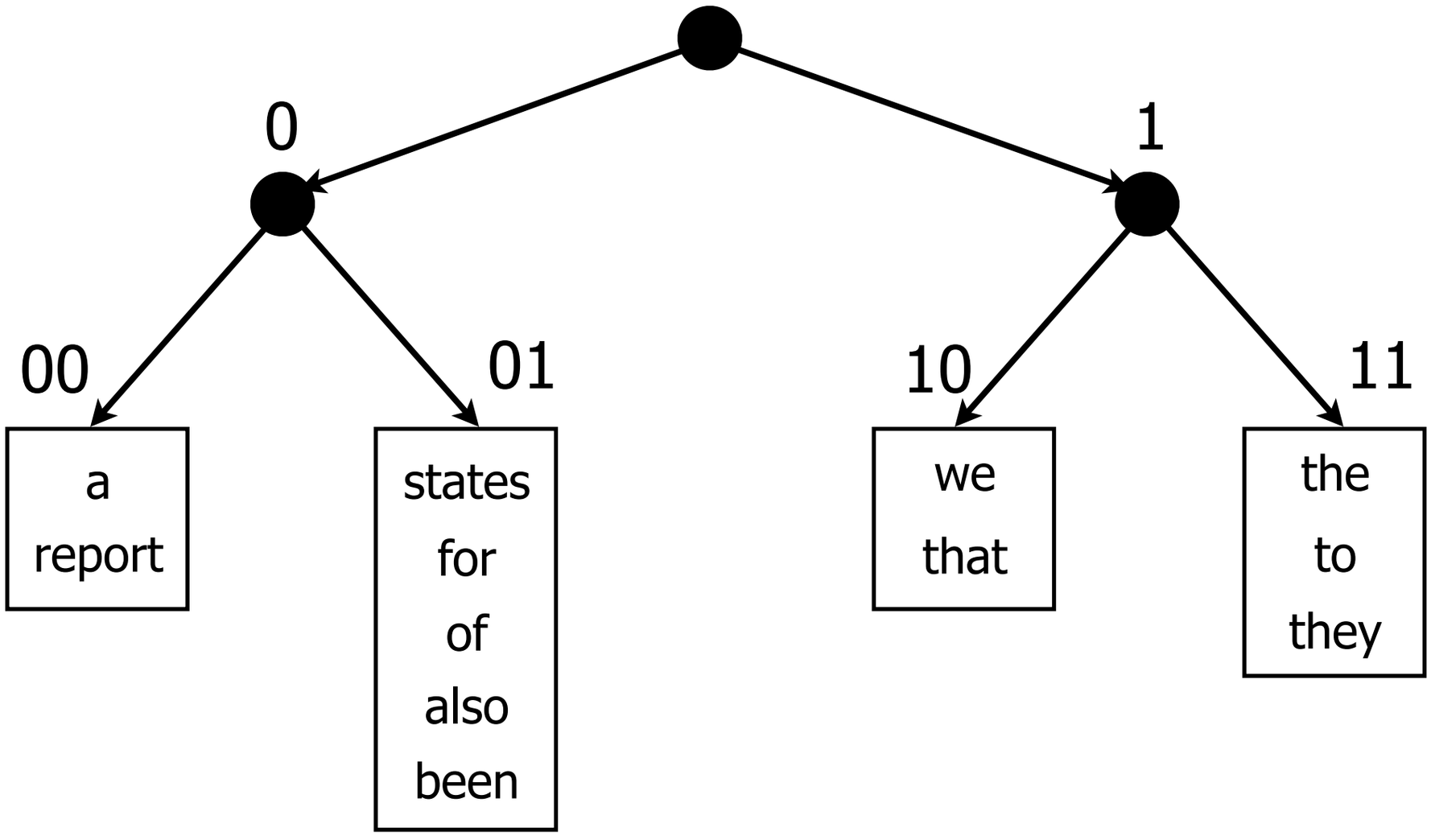
Random Clustering



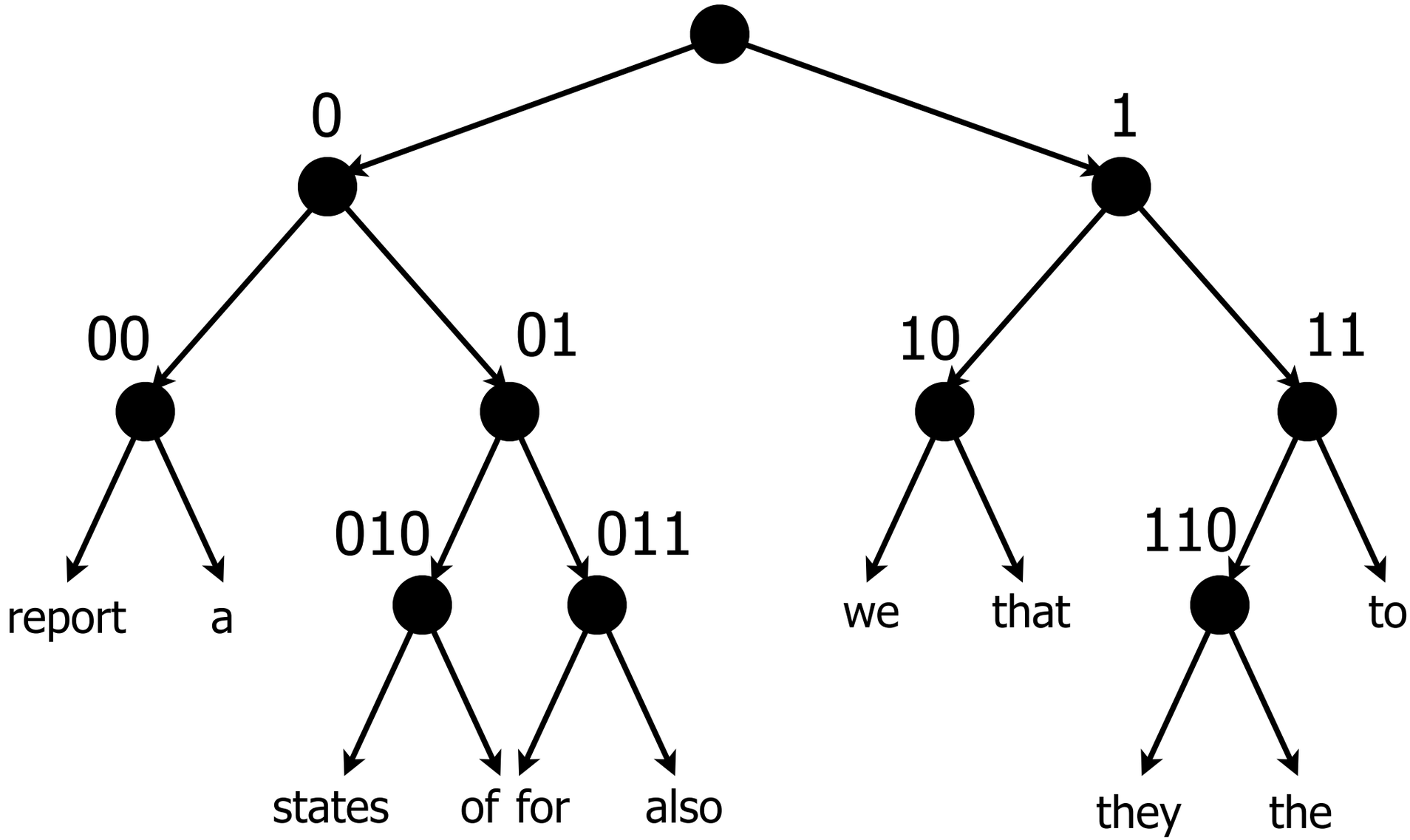
Random Clustering



Random Clustering

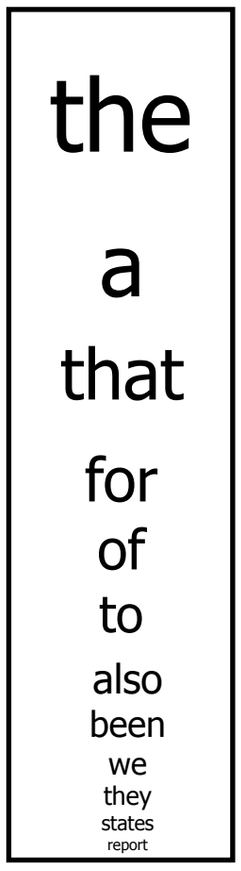


Random Clustering

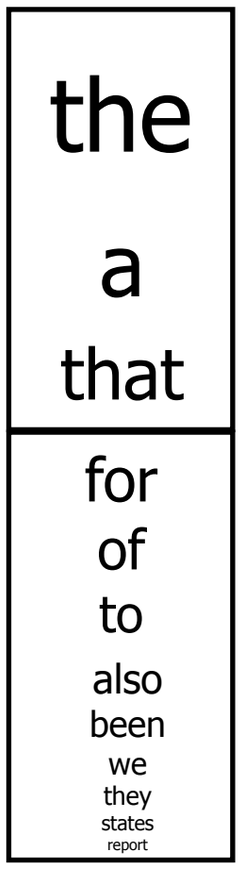


Frequency Clustering

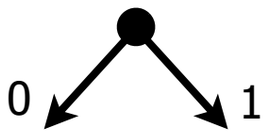
Frequency Clustering



Frequency Clustering



Frequency Clustering



0

the

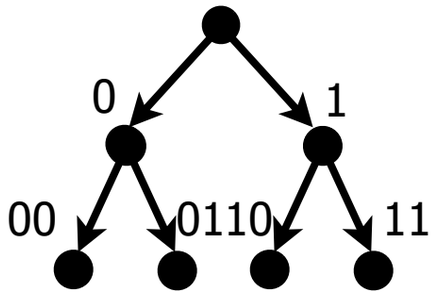
a

that

1

for
of
to
also
been
we
they
states
report

Frequency Clustering



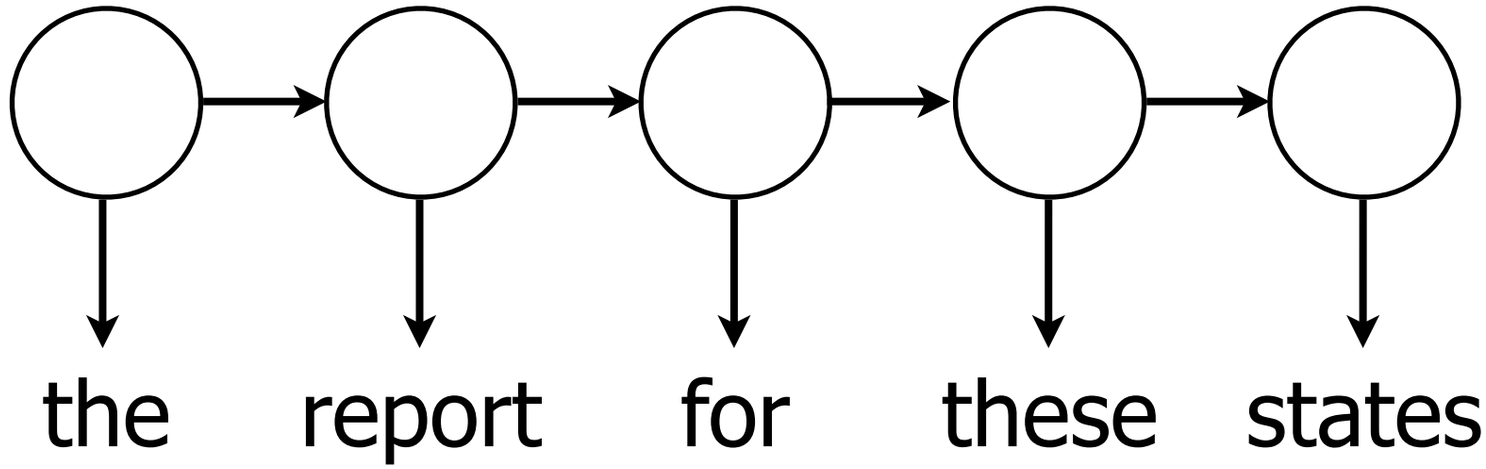
00	01	10	11
the	a that	for of to	also been we they states report

HMM Clustering

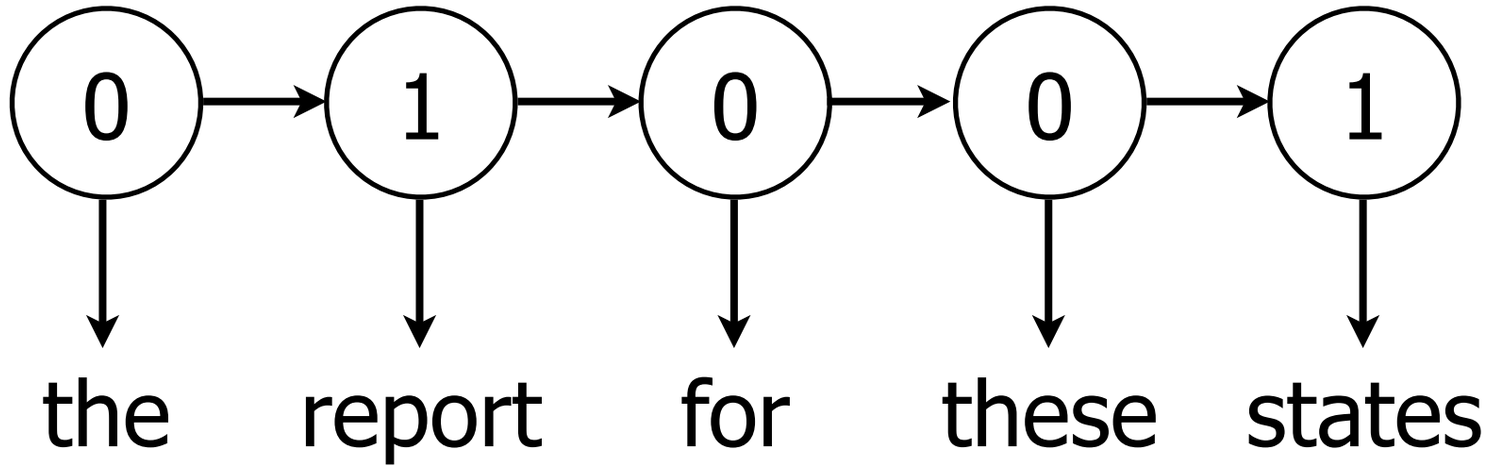
HMM Clustering

the report for these states

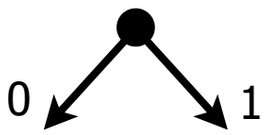
HMM Clustering

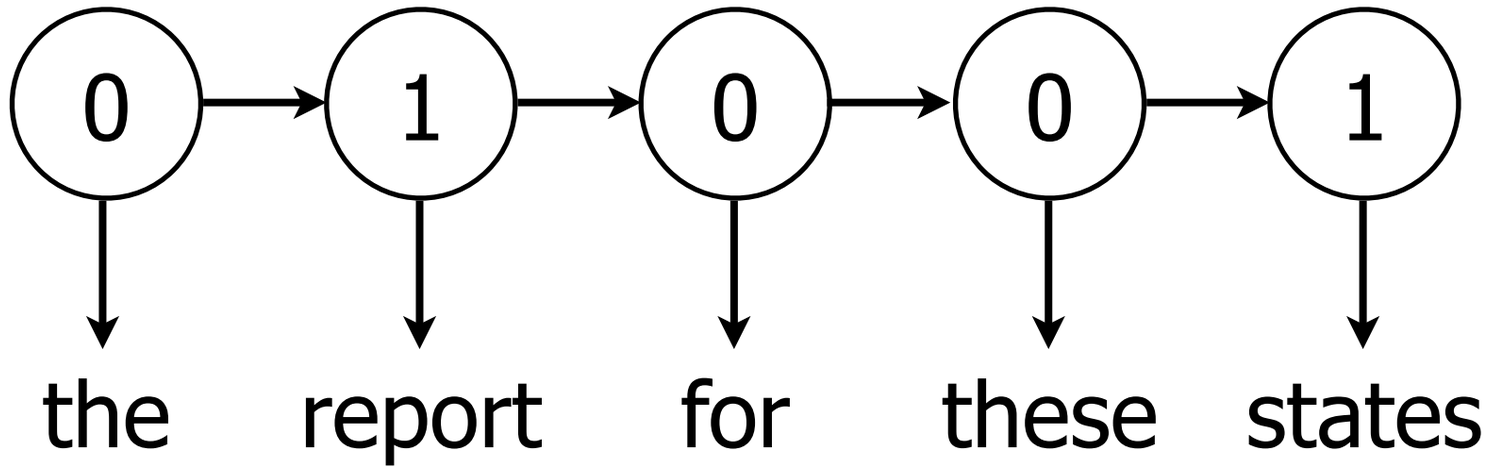


HMM Clustering

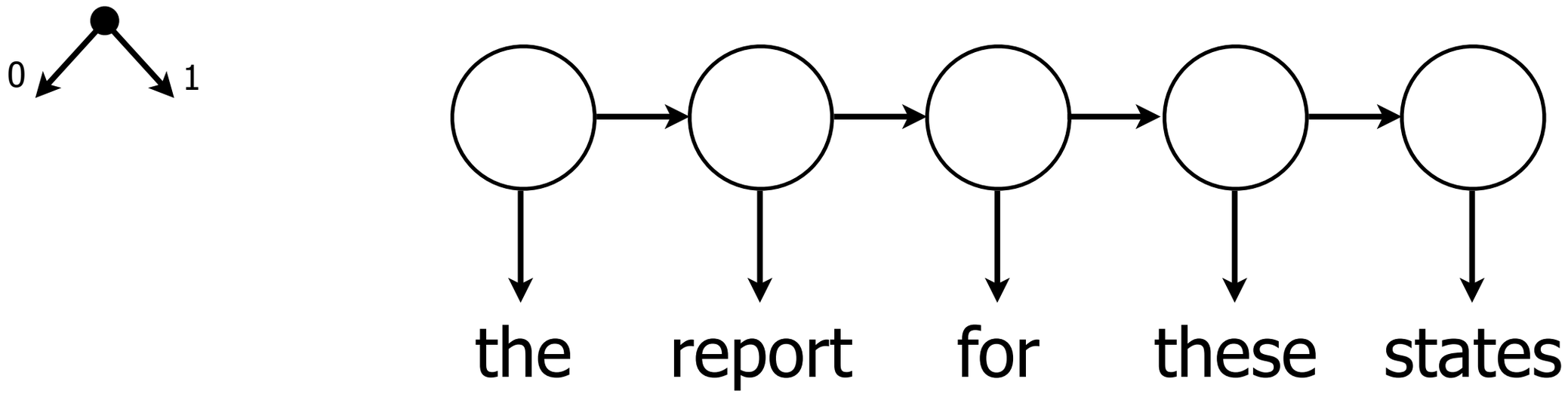


HMM Clustering

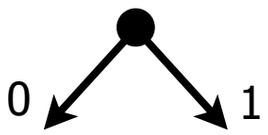


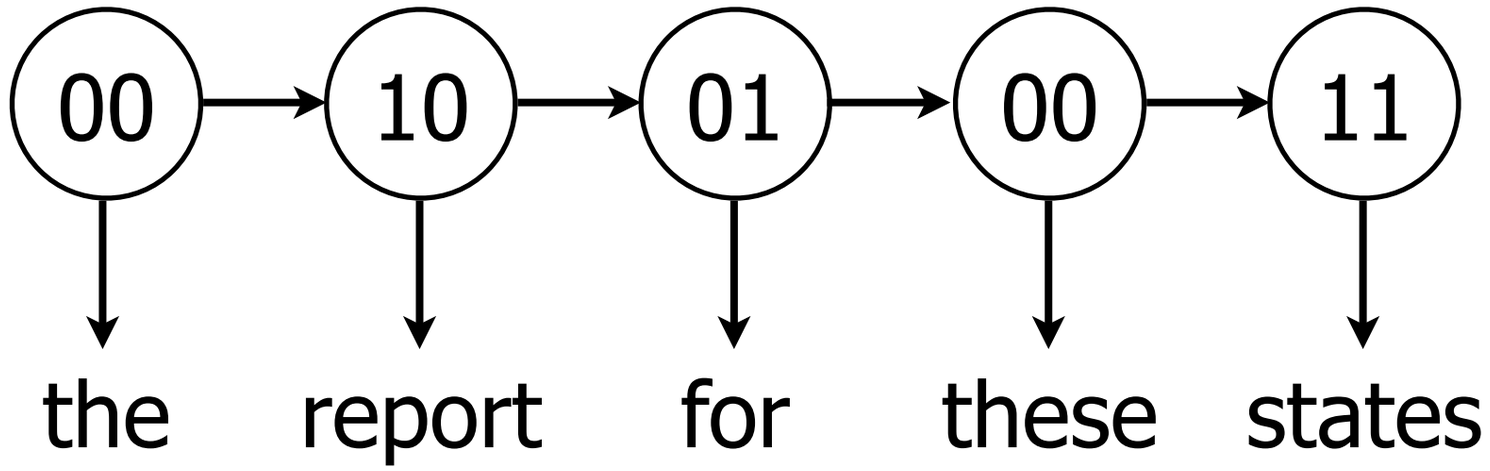


HMM Clustering

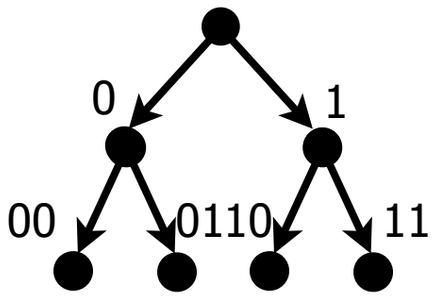


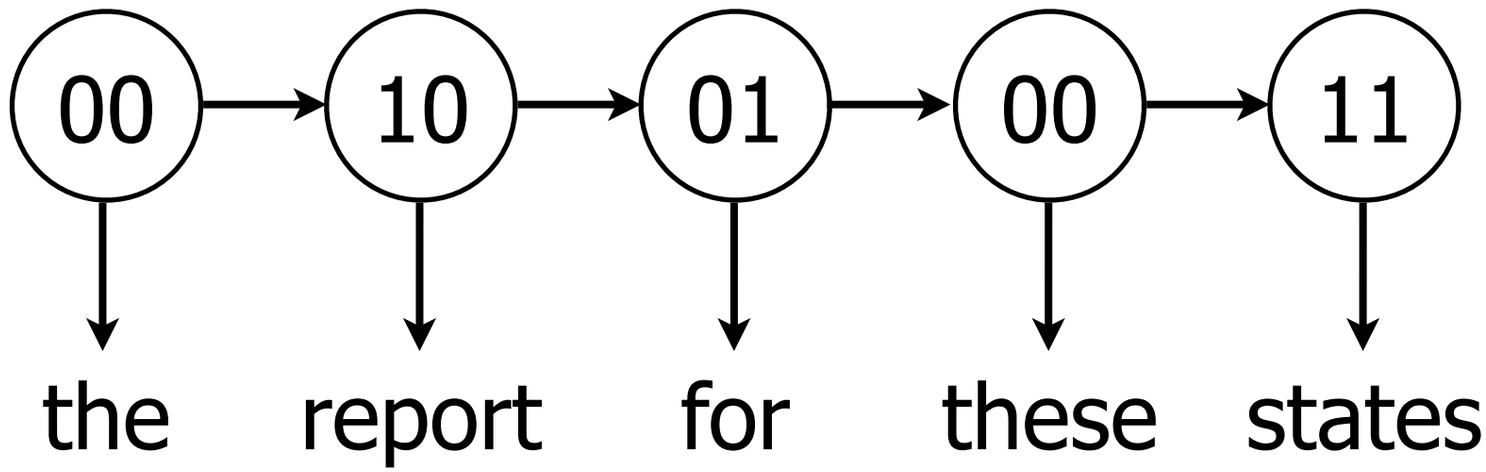
HMM Clustering





HMM Clustering

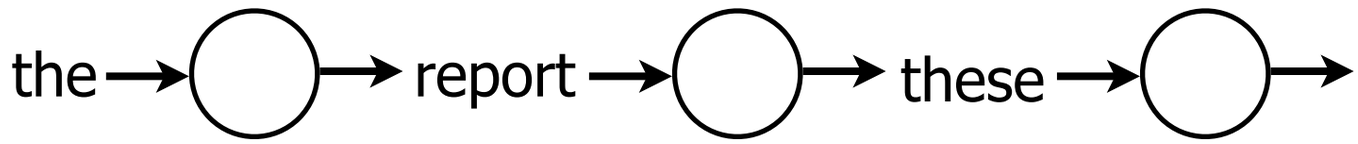




JCluster

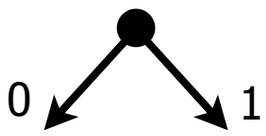
[Goodman ’01]

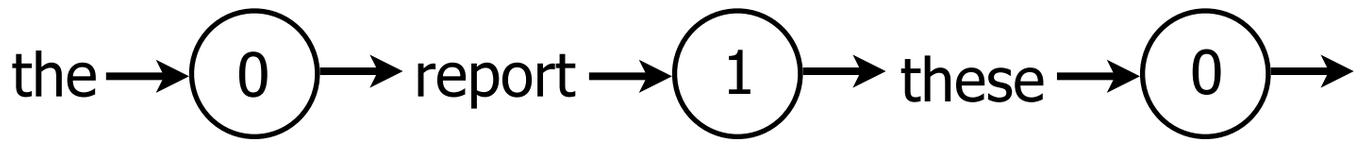
JCluster



[Goodman ’01]

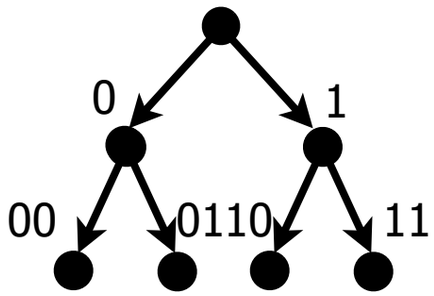
JCluster

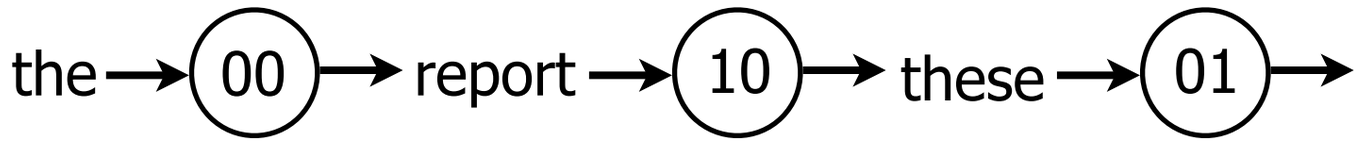




[Goodman ’01]

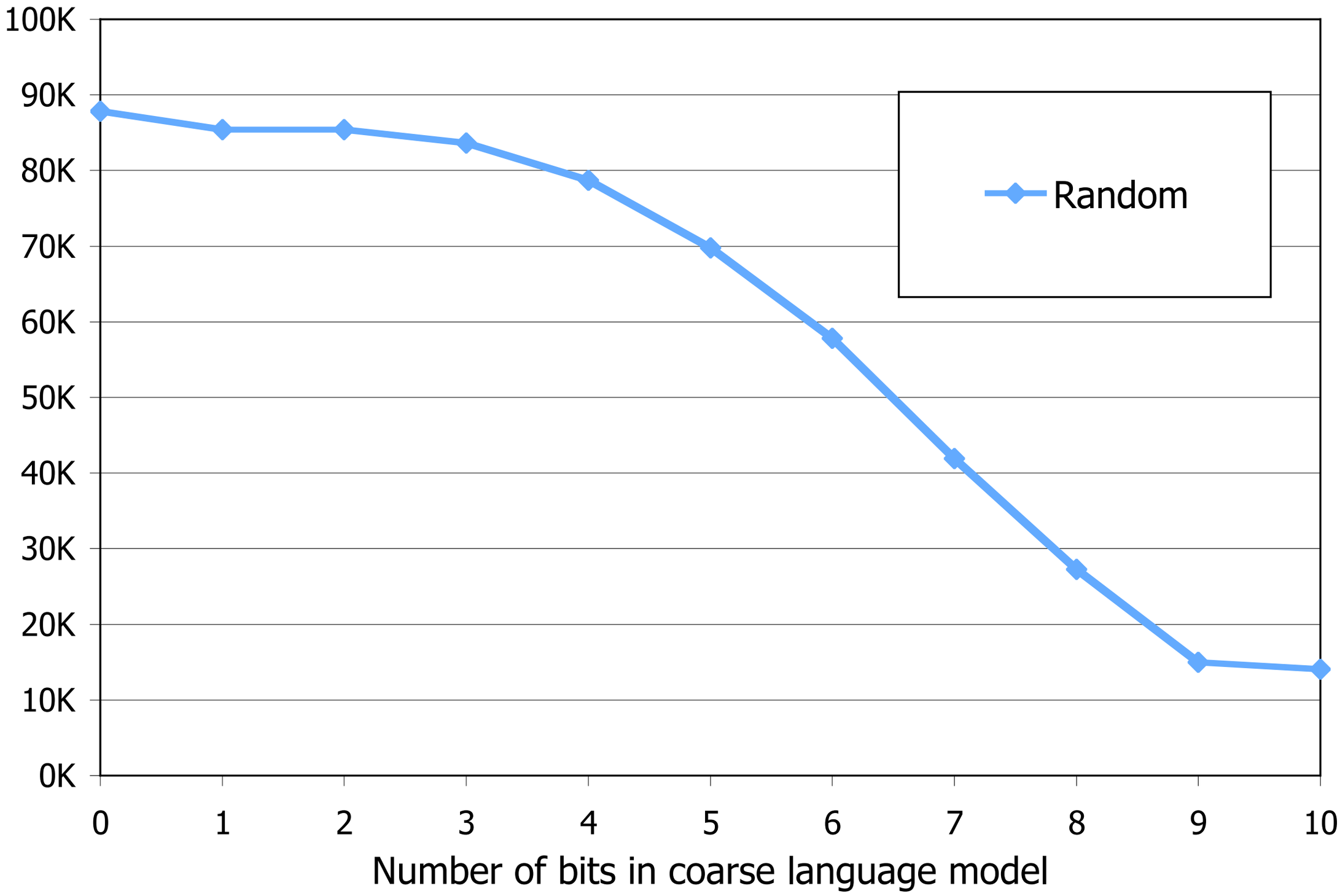
JCluster



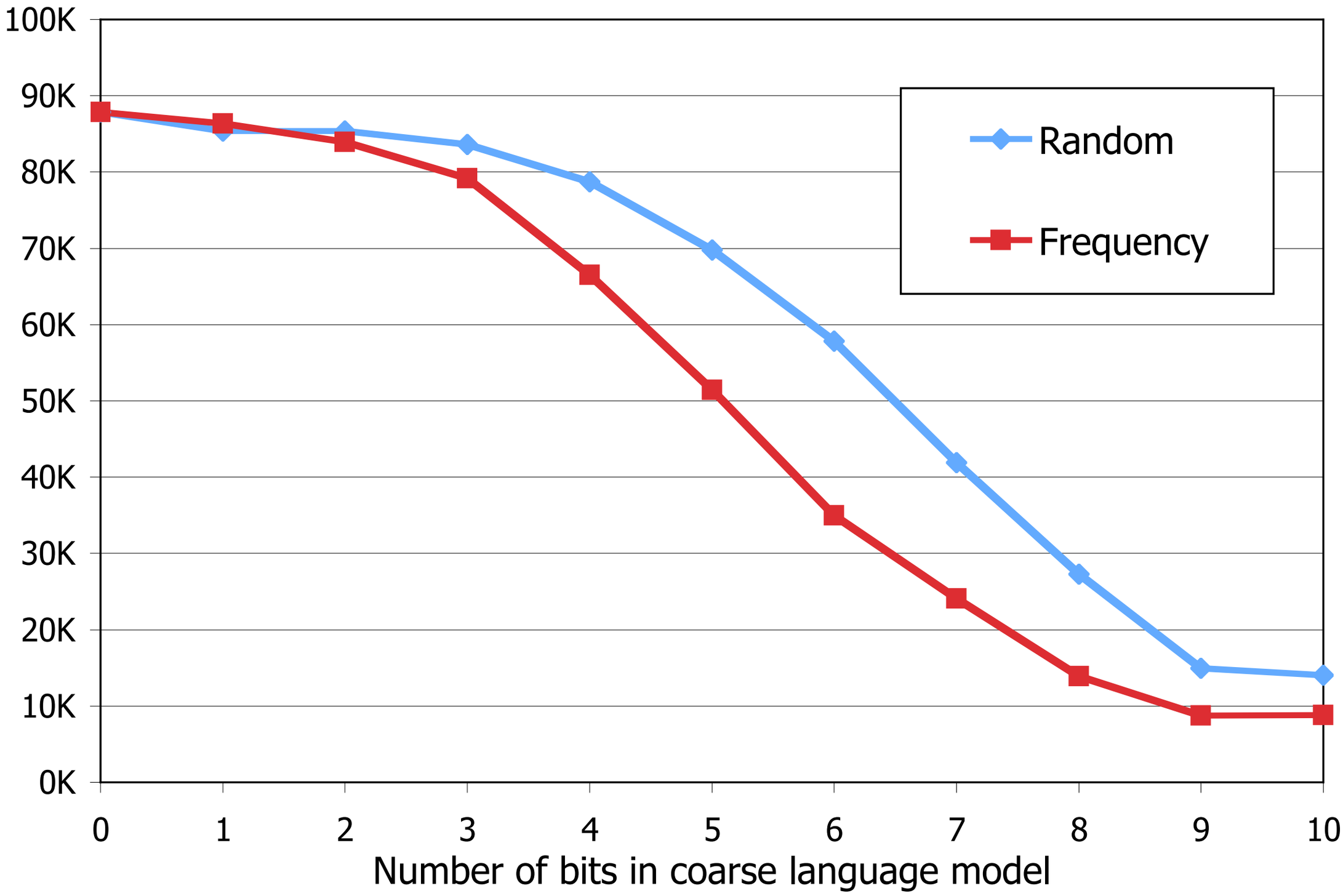


[Goodman ’01]

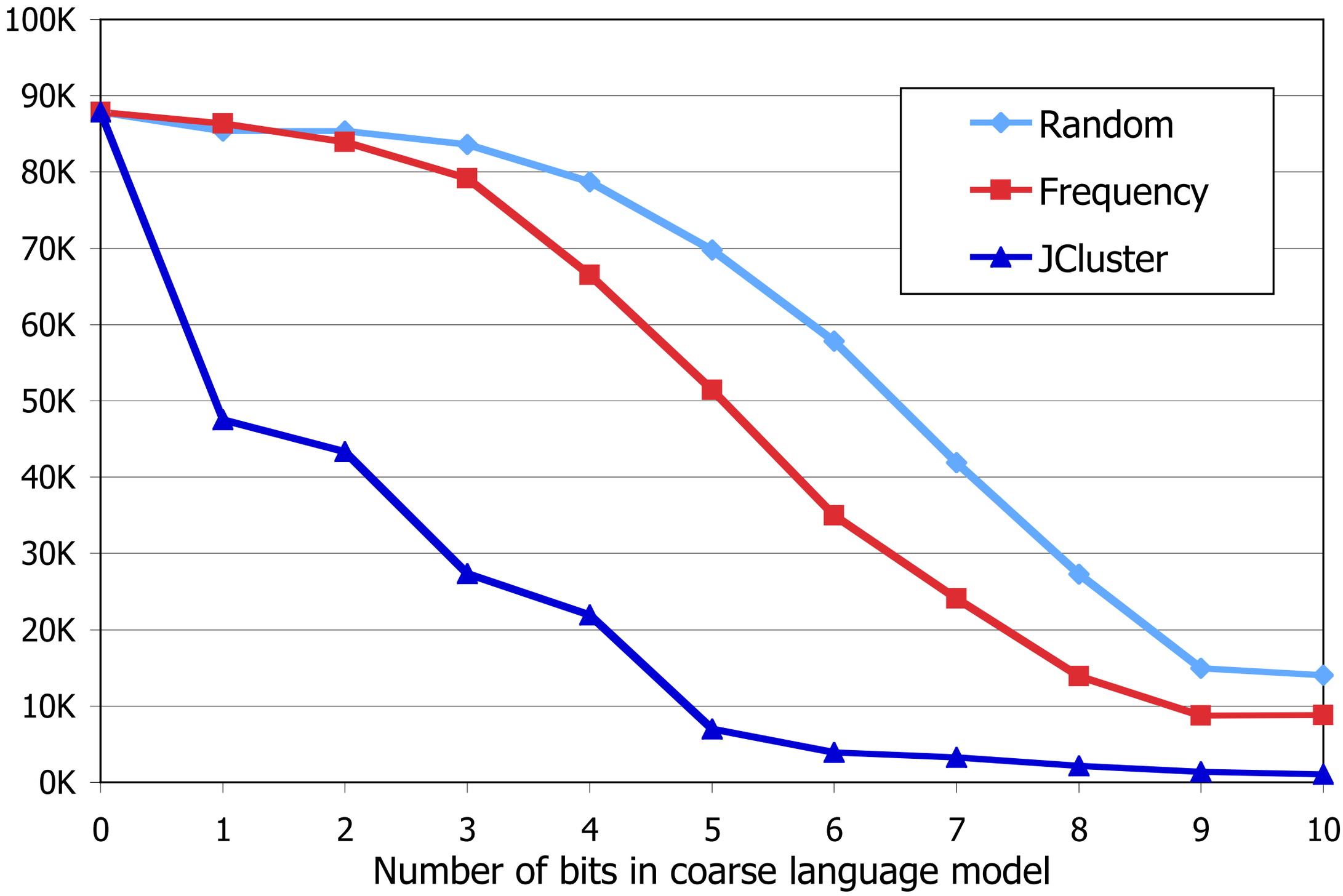
Cluster Quality (Perplexity)



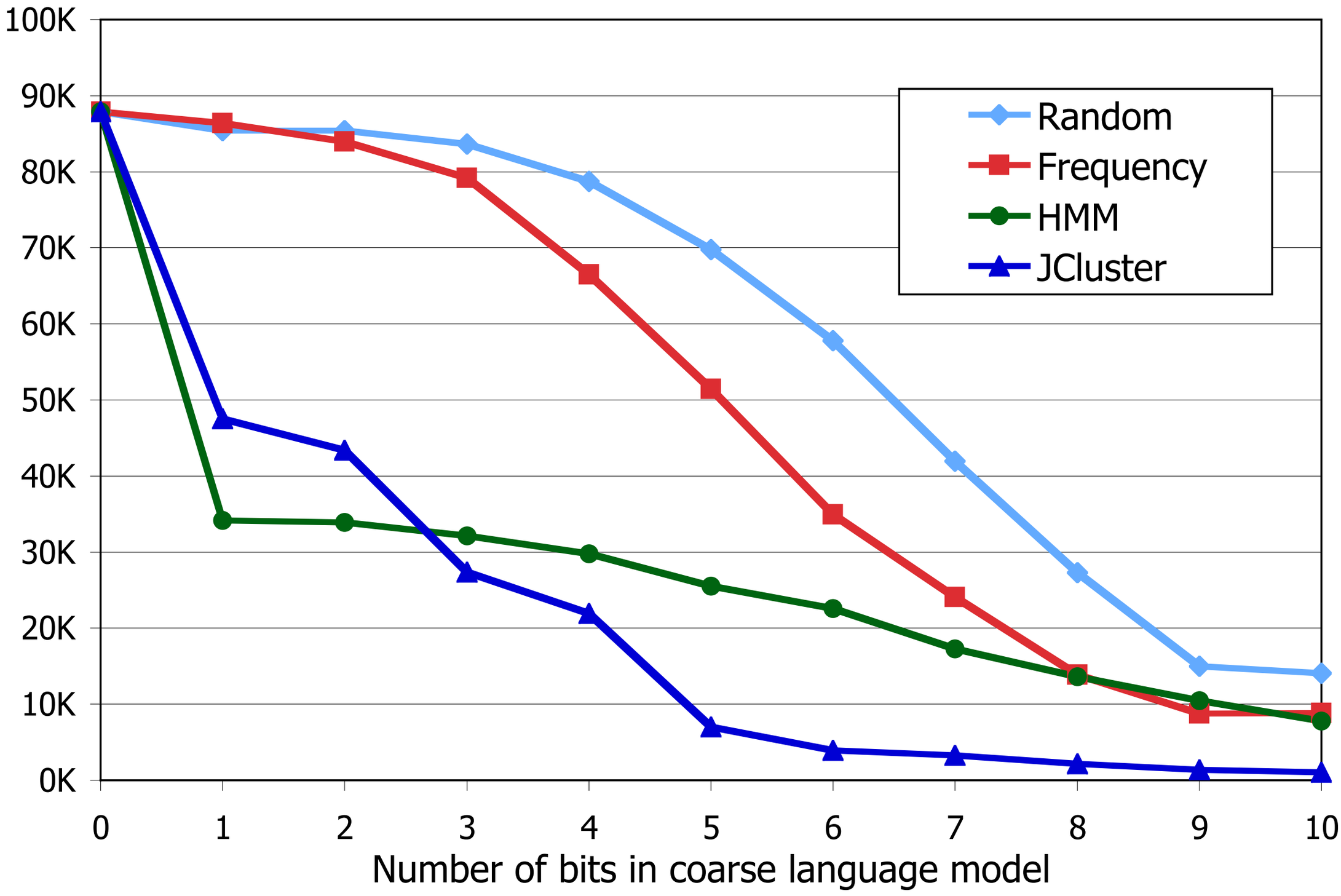
Cluster Quality (Perplexity)



Cluster Quality (Perplexity)



Cluster Quality (Perplexity)



Hierarchical Clustering

Hierarchical Clustering

fine
the
a
they
we
states
report
also
in

Hierarchical Clustering

fine	1-bit
the	0
a	0
they	0
we	0
states	1
report	1
also	1
in	1

Hierarchical Clustering

fine	1-bit	2-bit
the	0	00
a	0	00
they	0	01
we	0	01
states	1	10
report	1	10
also	1	11
in	1	11

Hierarchical Clustering

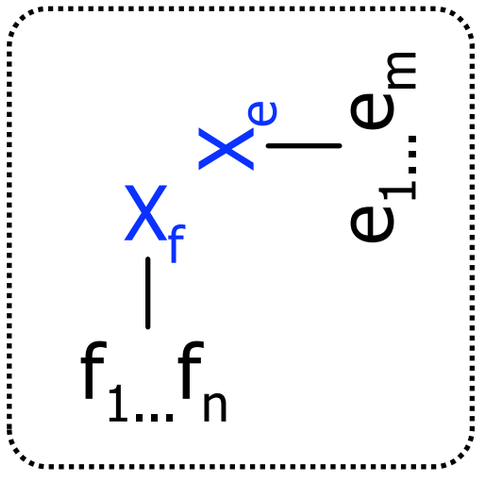
fine	1-bit	2-bit	3-bit
the	0	00	000
a	0	00	001
they	0	01	010
we	0	01	011
states	1	10	100
report	1	10	101
also	1	11	110
in	1	11	111

Inversion Transduction Grammars

[Wu ’97]

Inversion Transduction Grammars

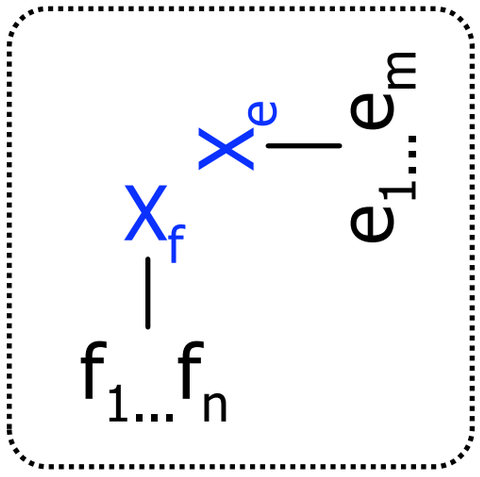
Generation



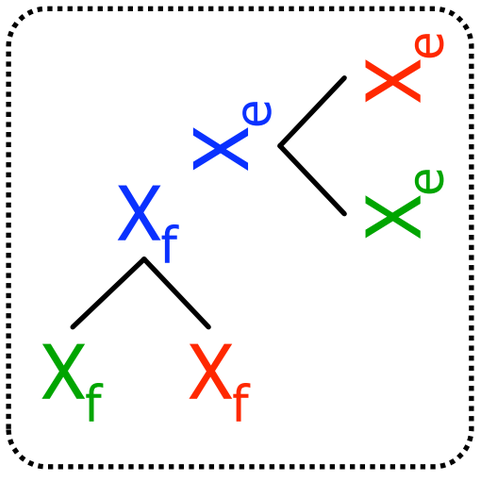
[Wu ’97]

Inversion Transduction Grammars

Generation

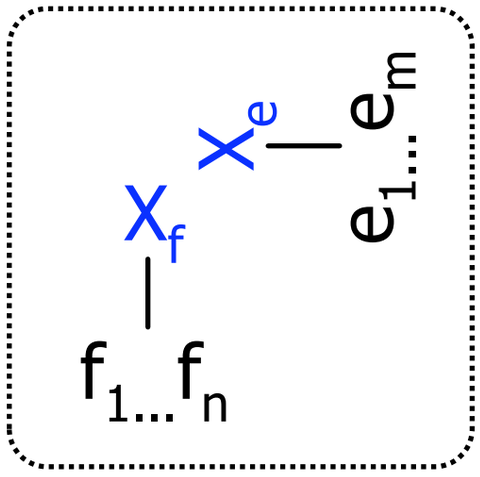
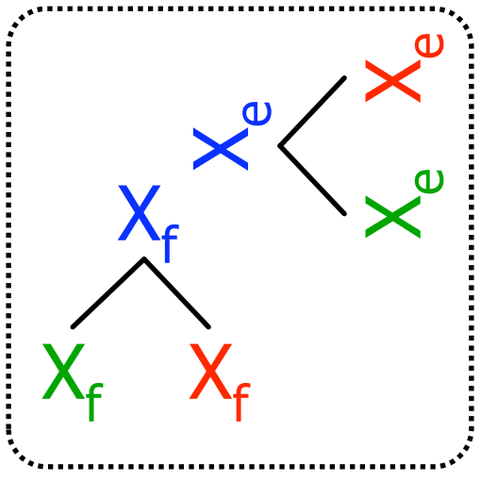
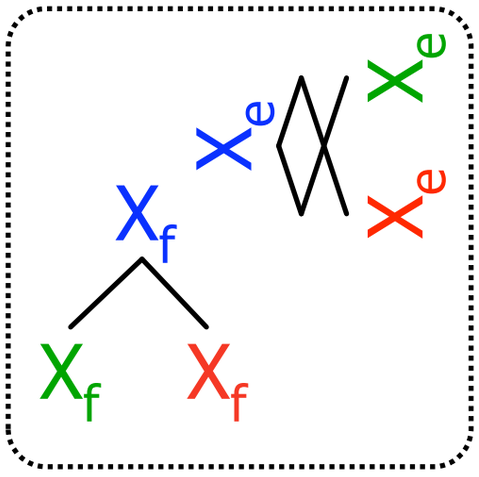


Monotonic Combination



[Wu ’97]

Inversion Transduction Grammars

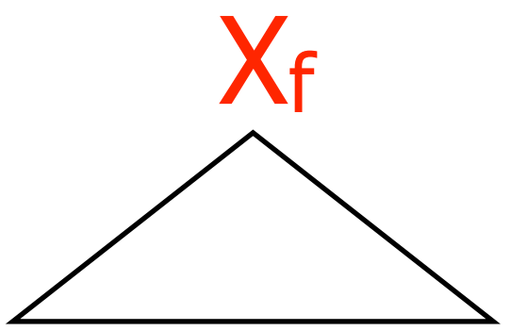
Generation	Monotonic Combination	Inverted Combination
		

[Wu ’97]

Translating with ITG (w/o LM)

reanudación del período de sesiones

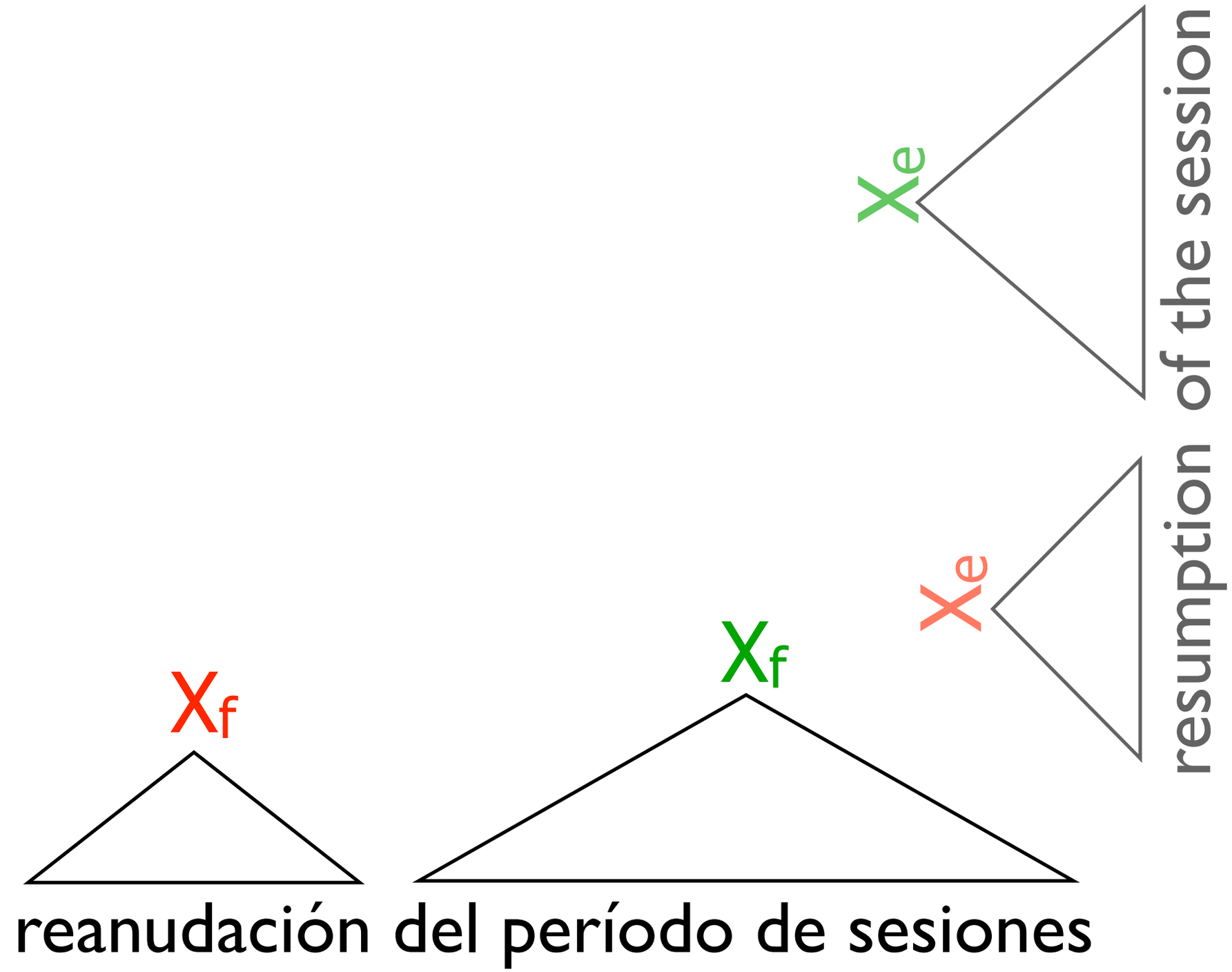
Translating with ITG (w/o LM)



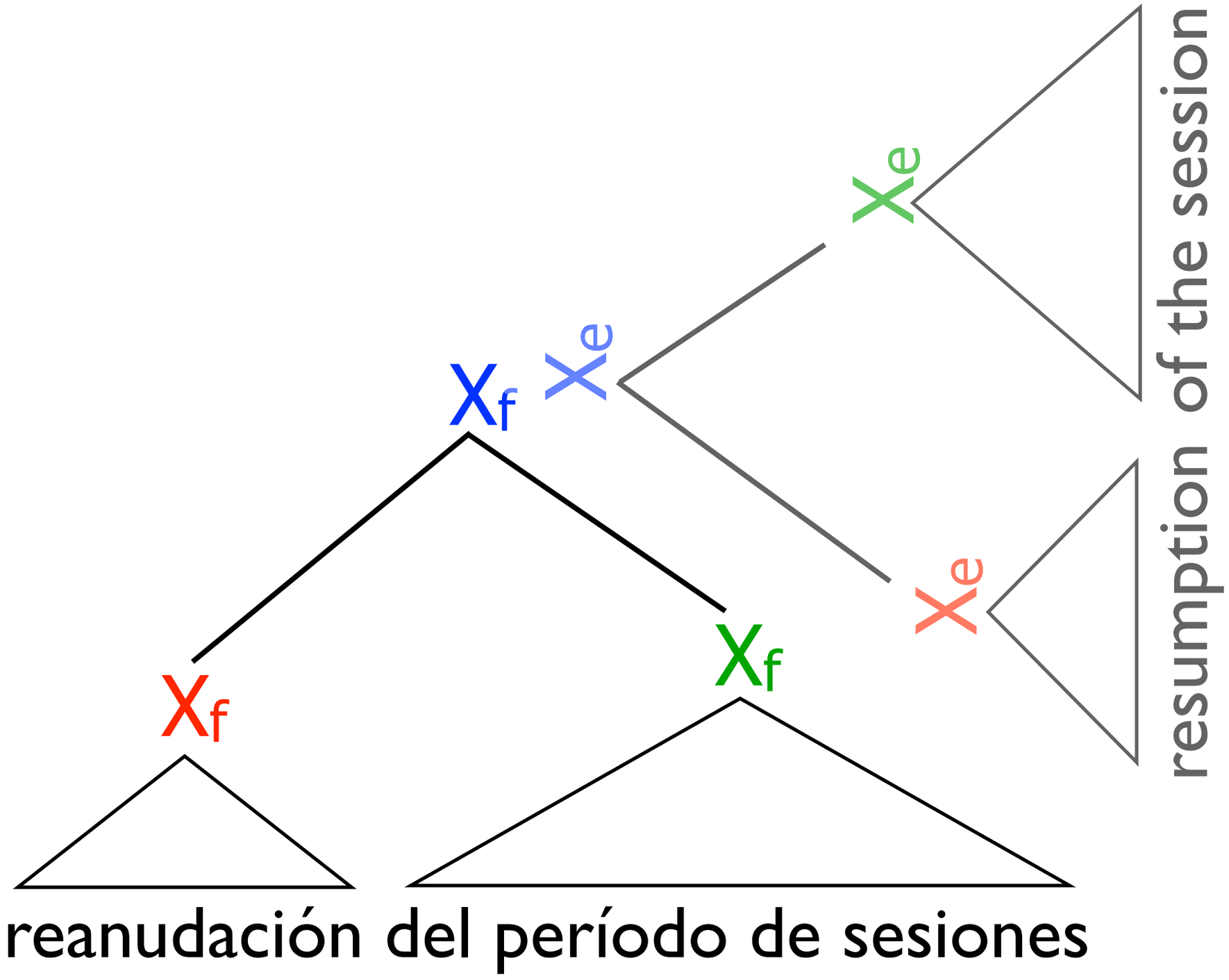
reanudación del período de sesiones

Translating with ITG (w/o LM)

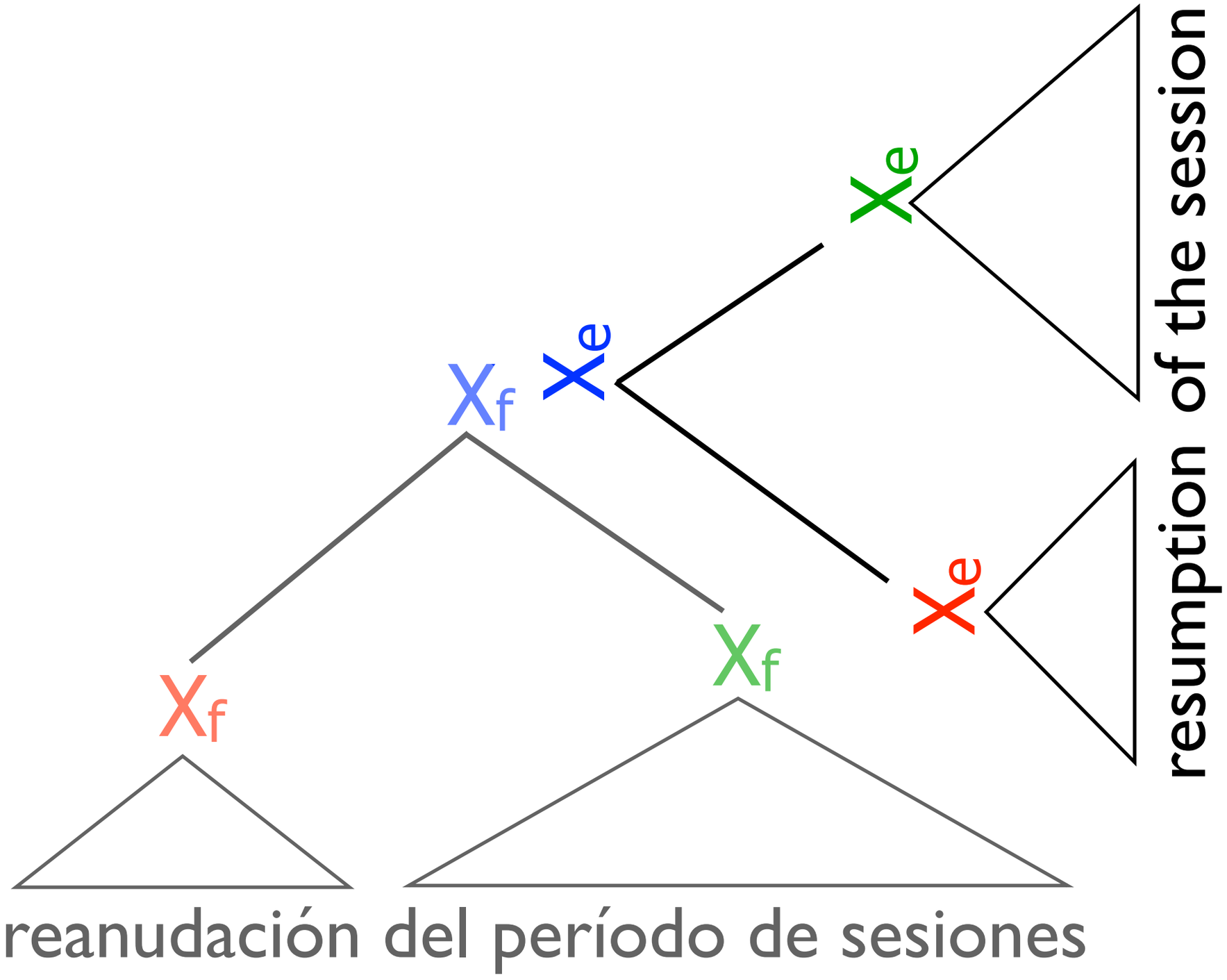
Translating with ITG (w/o LM)



Translating with ITG (w/o LM)



Translating with ITG (w/o LM)

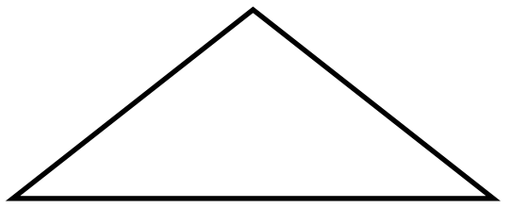


Syntactic MT w/ Language Model

reanudación del período de sesiones

Syntactic MT w/ Language Model

resumption-X_f-resumption



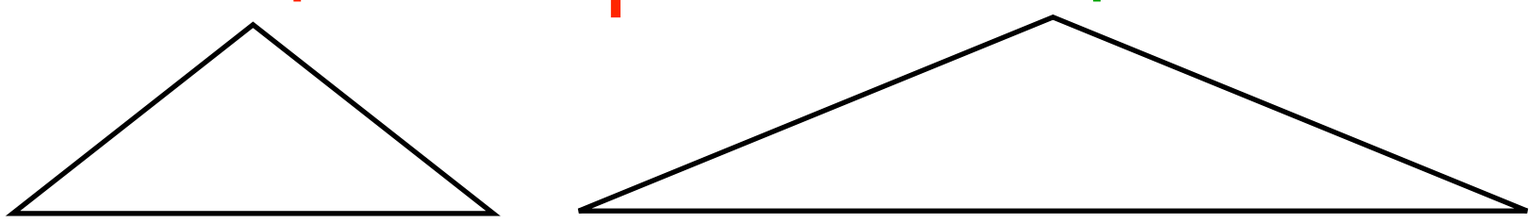
reanudación del período de sesiones

resumption

Syntactic MT w/ Language Model

resumption of the session

resumption-X_f-resumption of-X_f-session

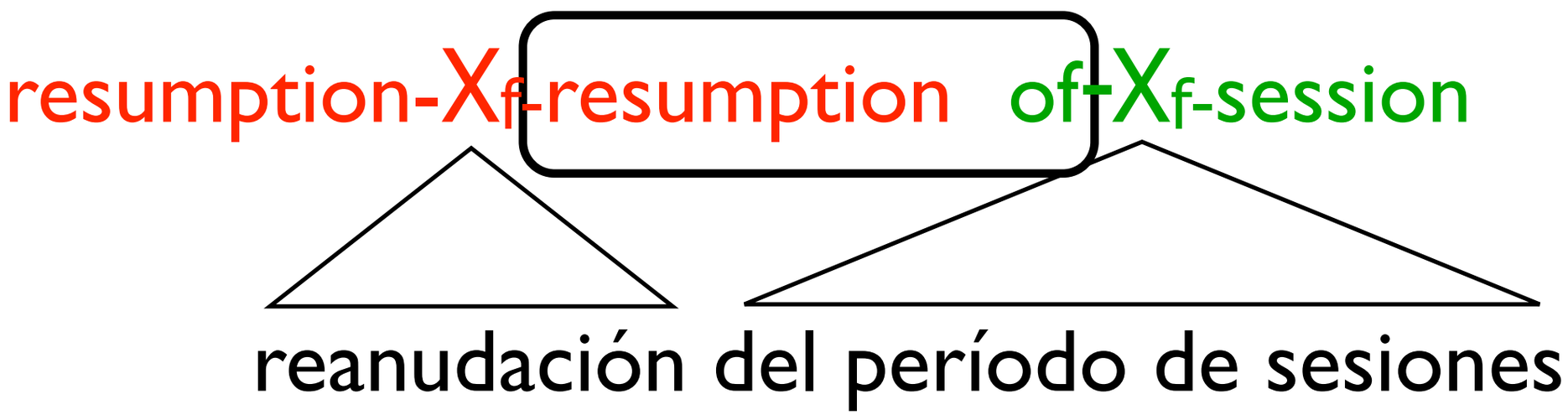


reanudación del período de sesiones

Syntactic MT w/ Language Model

Score with Language Model

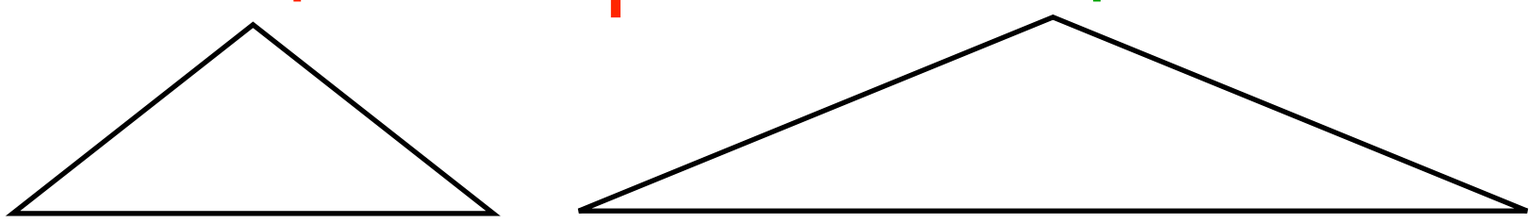
resumption of the session



Syntactic MT w/ Language Model

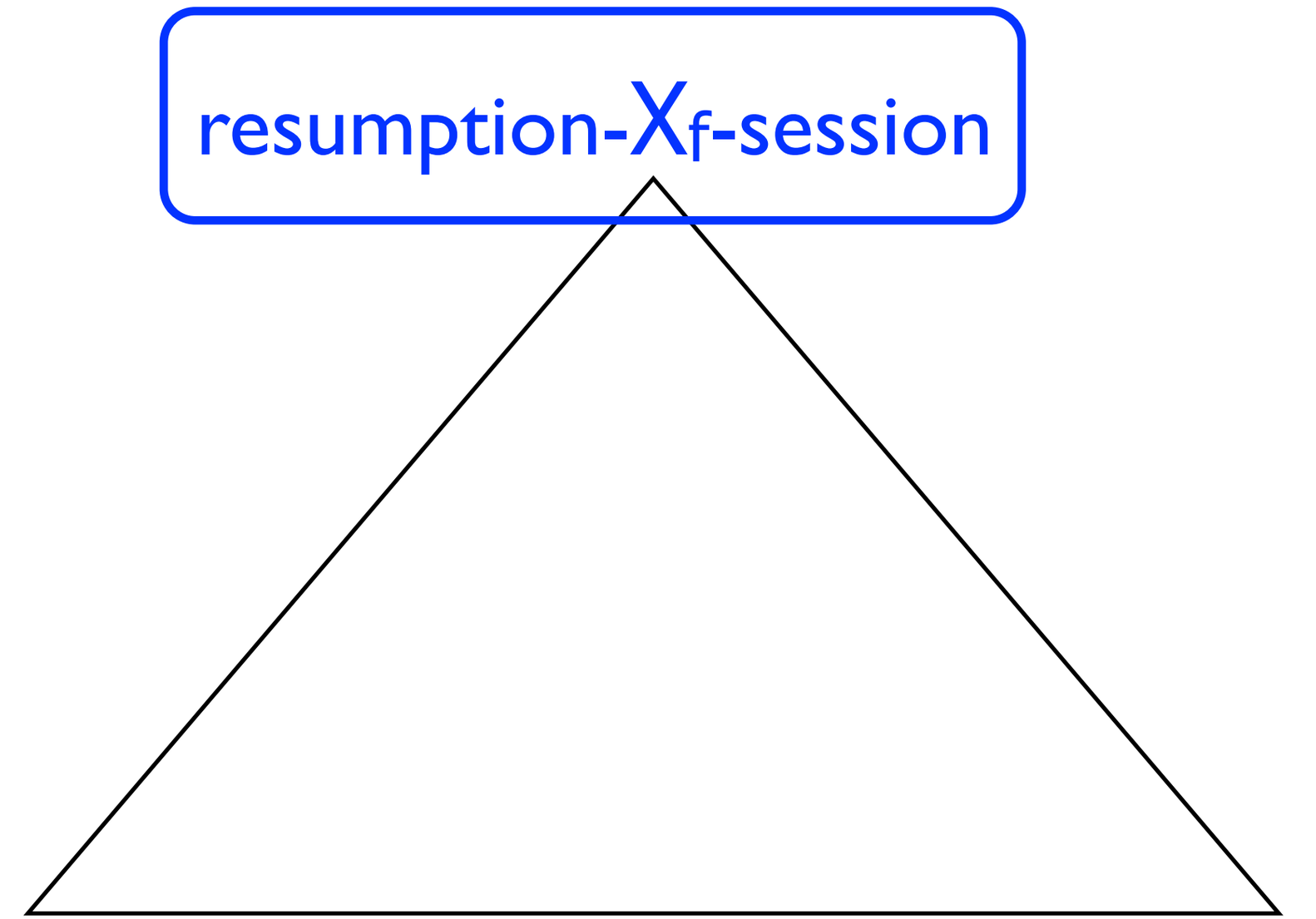
resumption of the session

resumption-X_f-resumption of-X_f-session



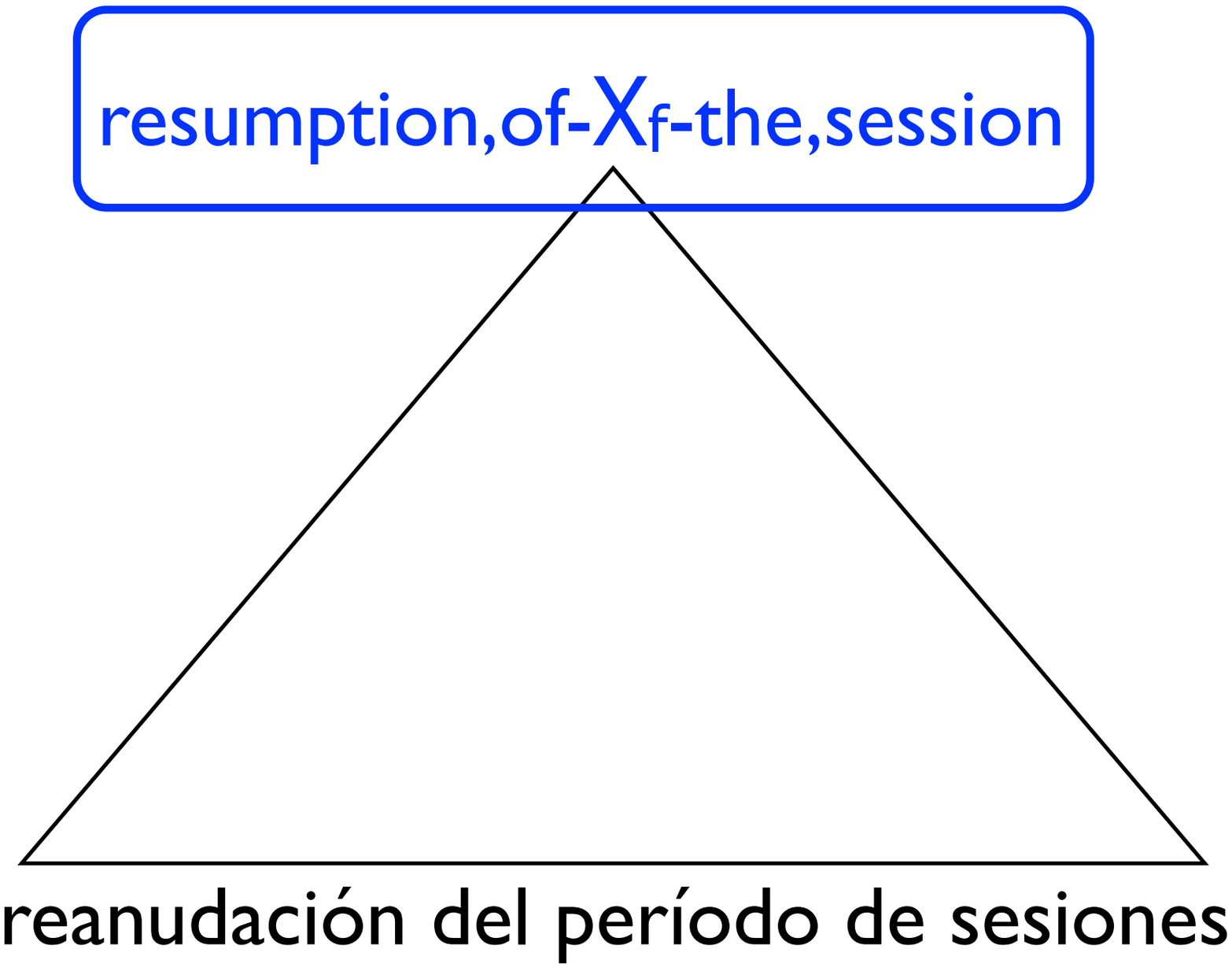
reanudación del período de sesiones

Syntactic MT w/ Language Model

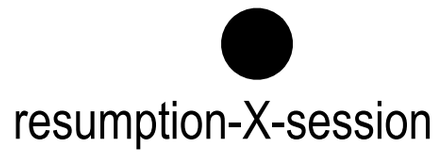


Syntactic MT w/ Language Model

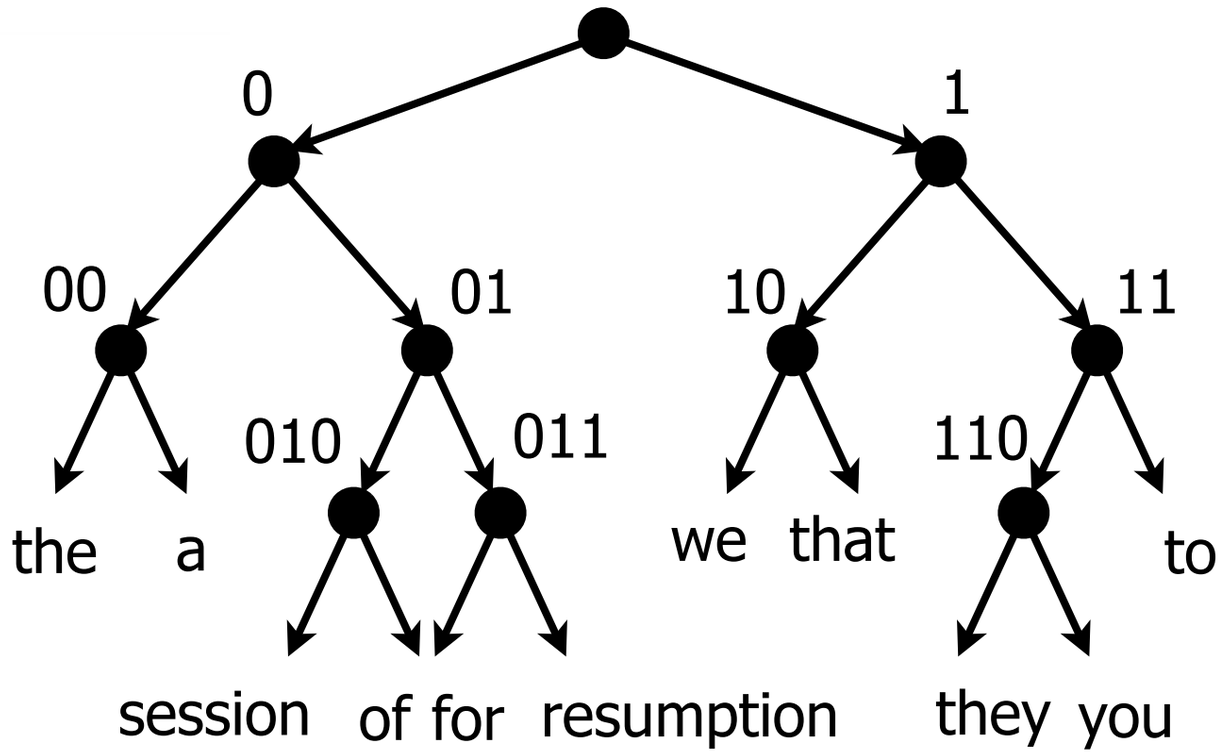
Lexical Explosion

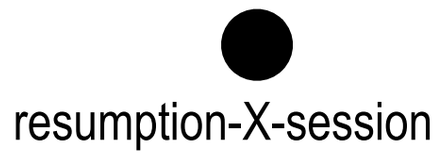


State Space Projections

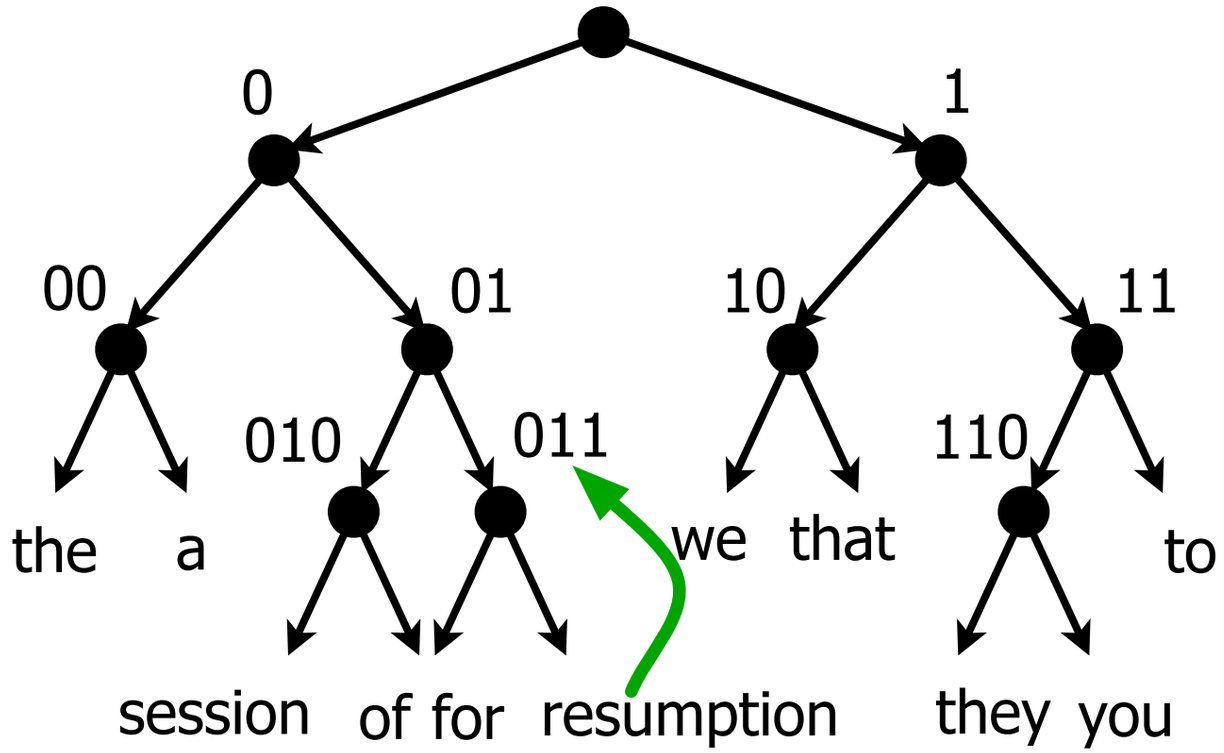


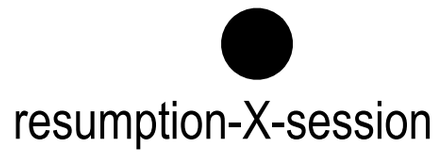
State Space Projections



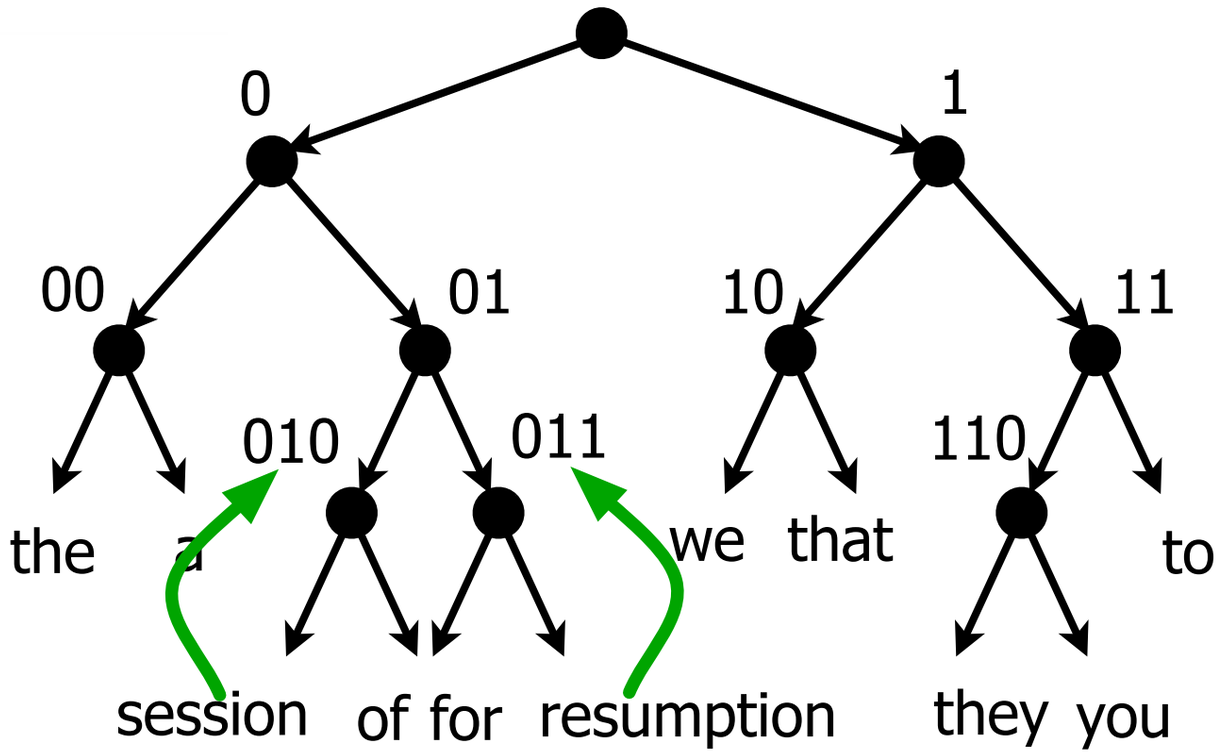


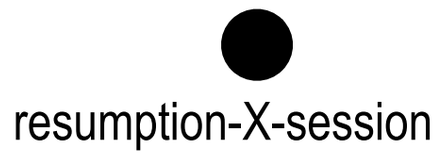
State Space Projections



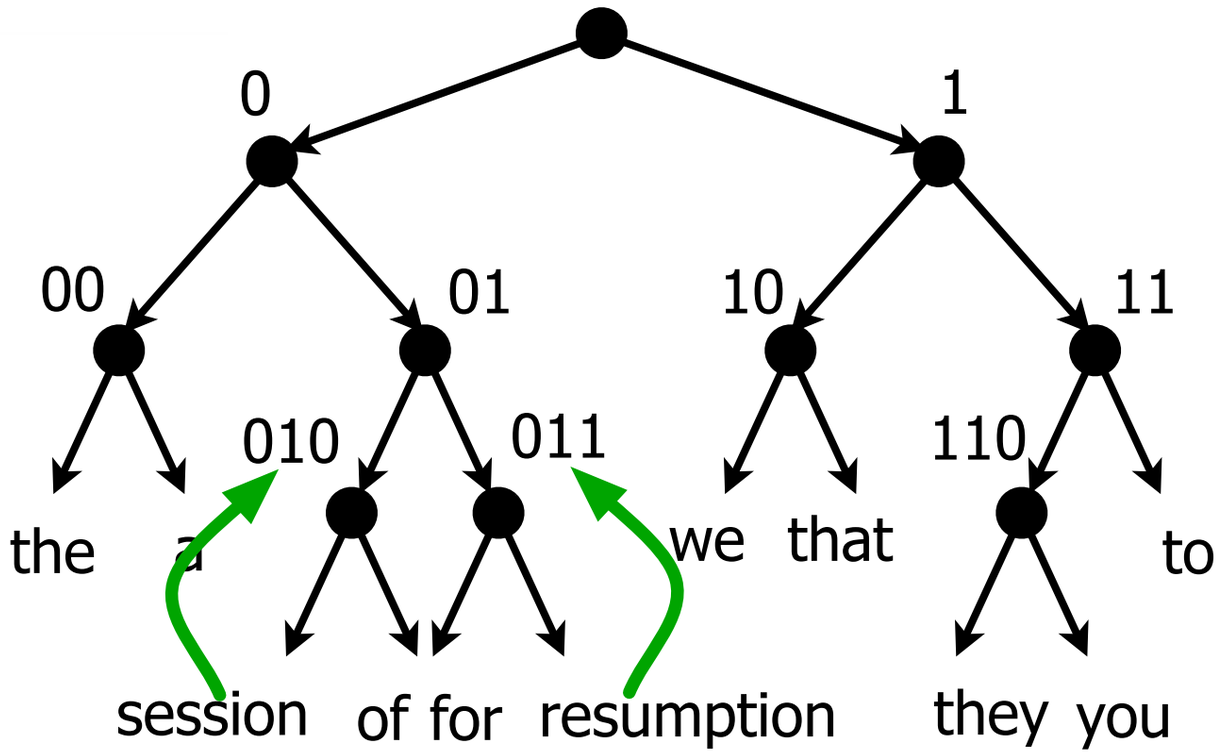


State Space Projections





State Space Projections

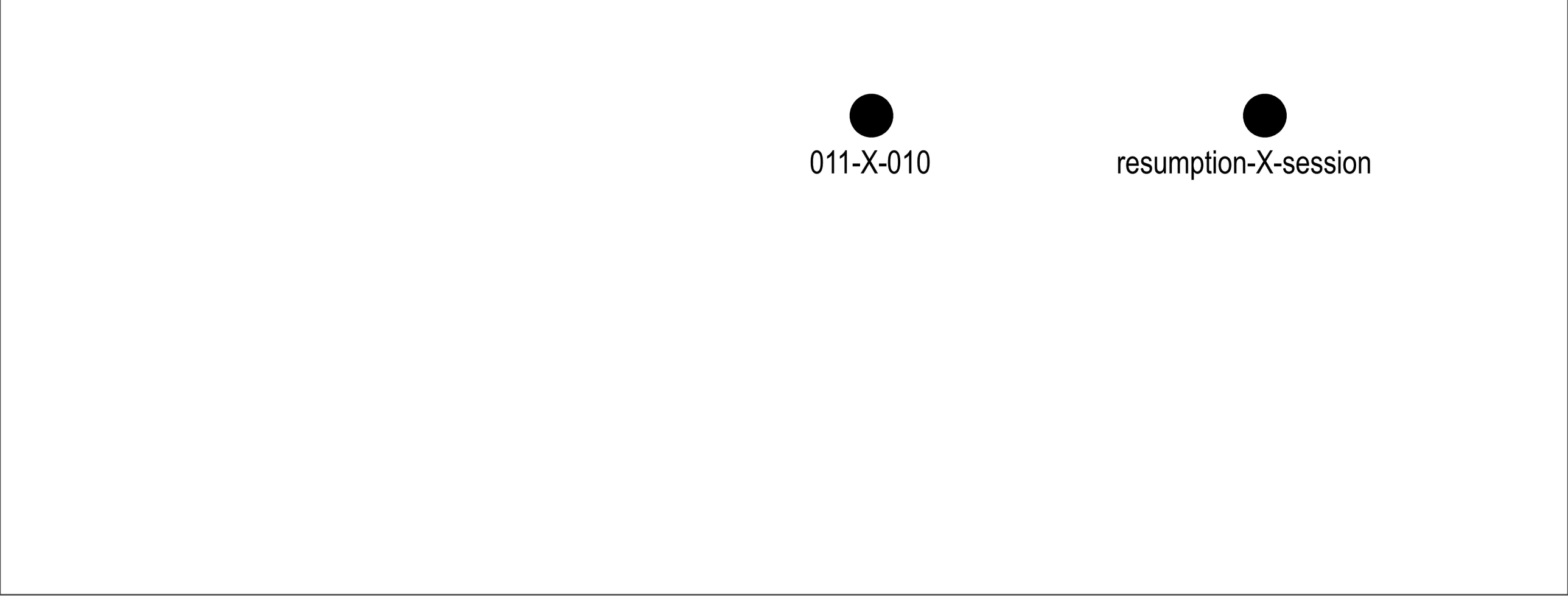


-X-

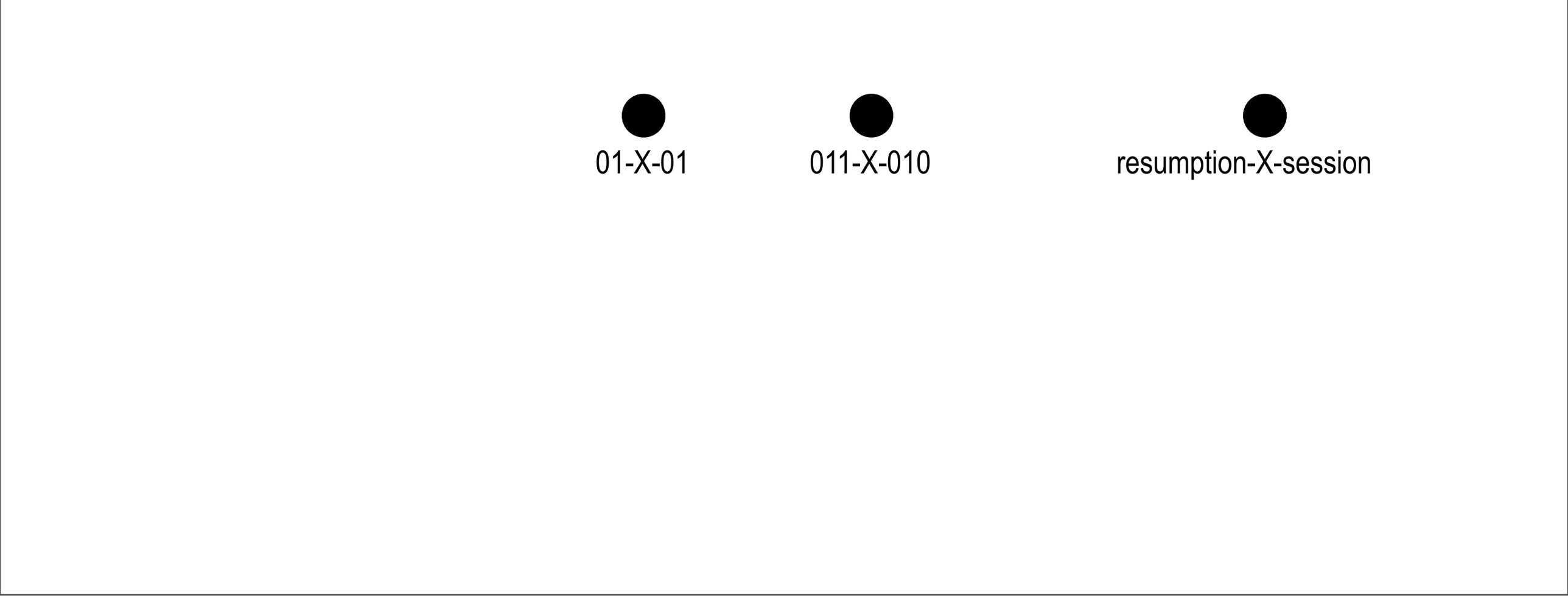
resumption-X-session

State Space Projections

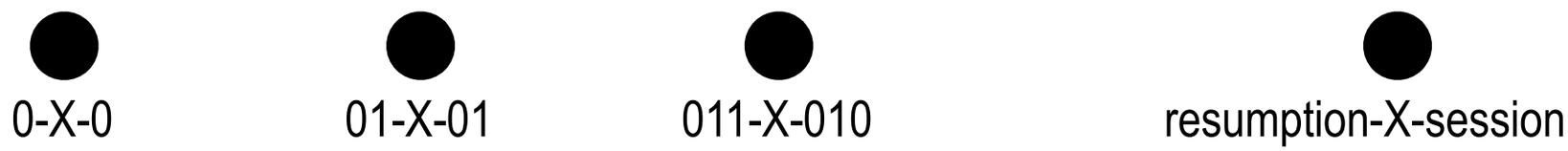
State Space Projections



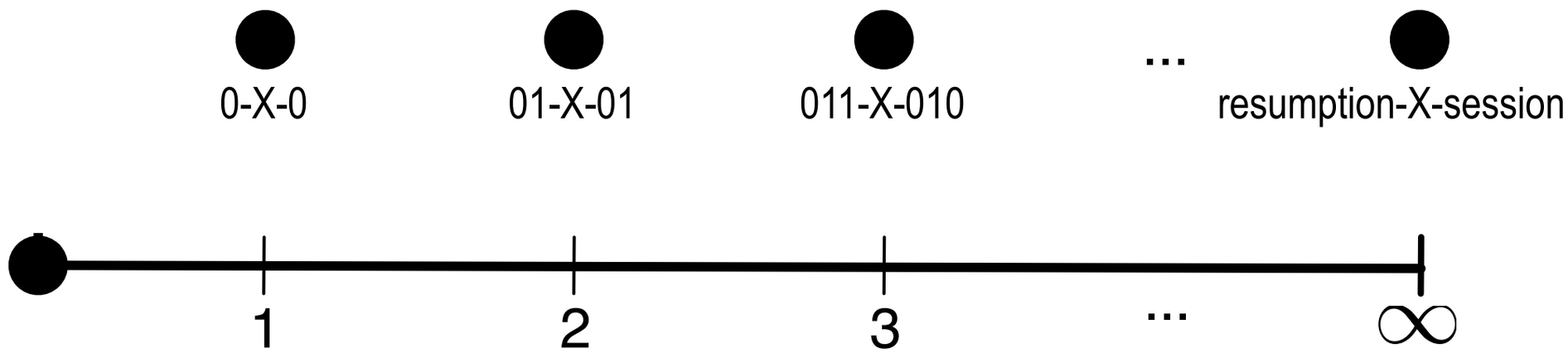
State Space Projections



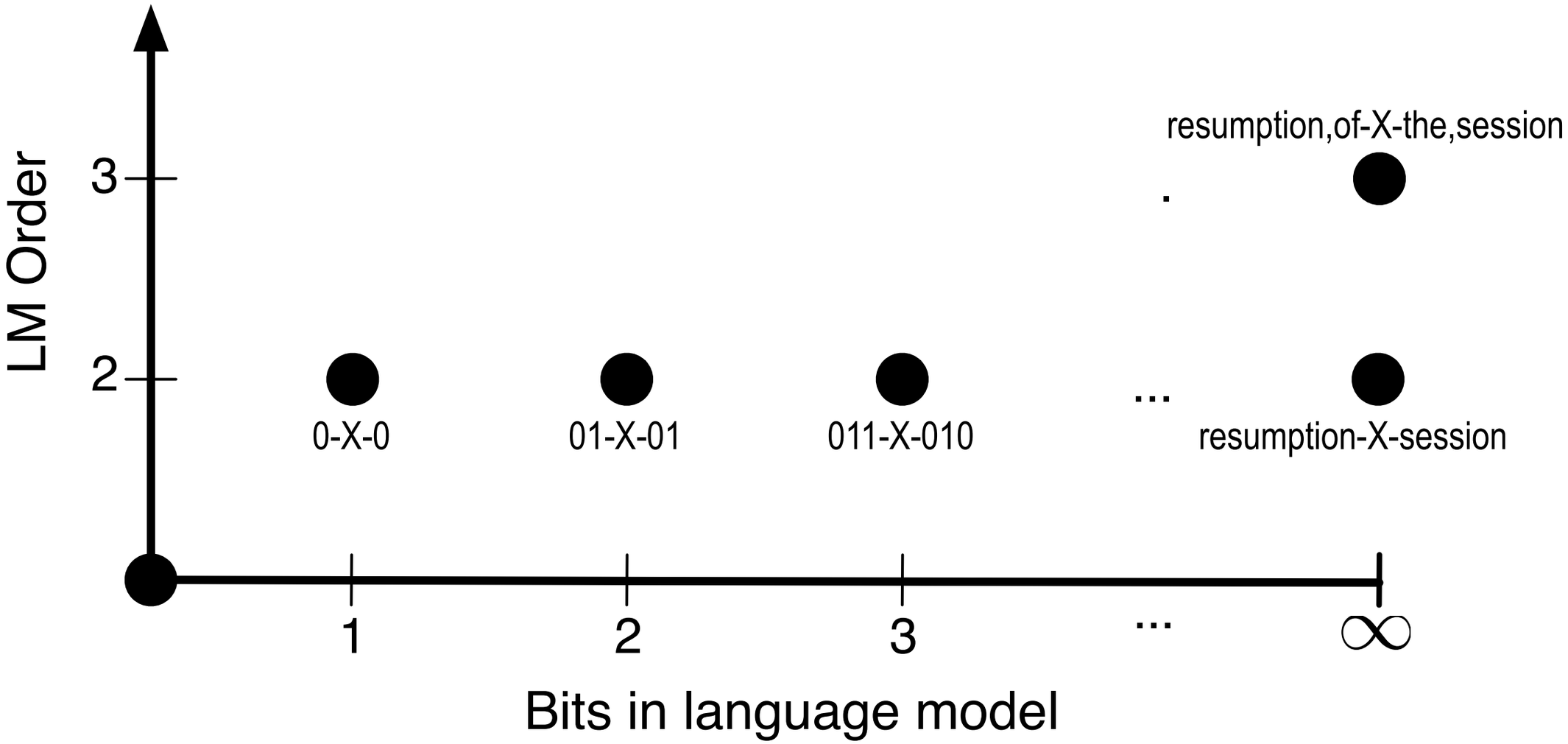
State Space Projections



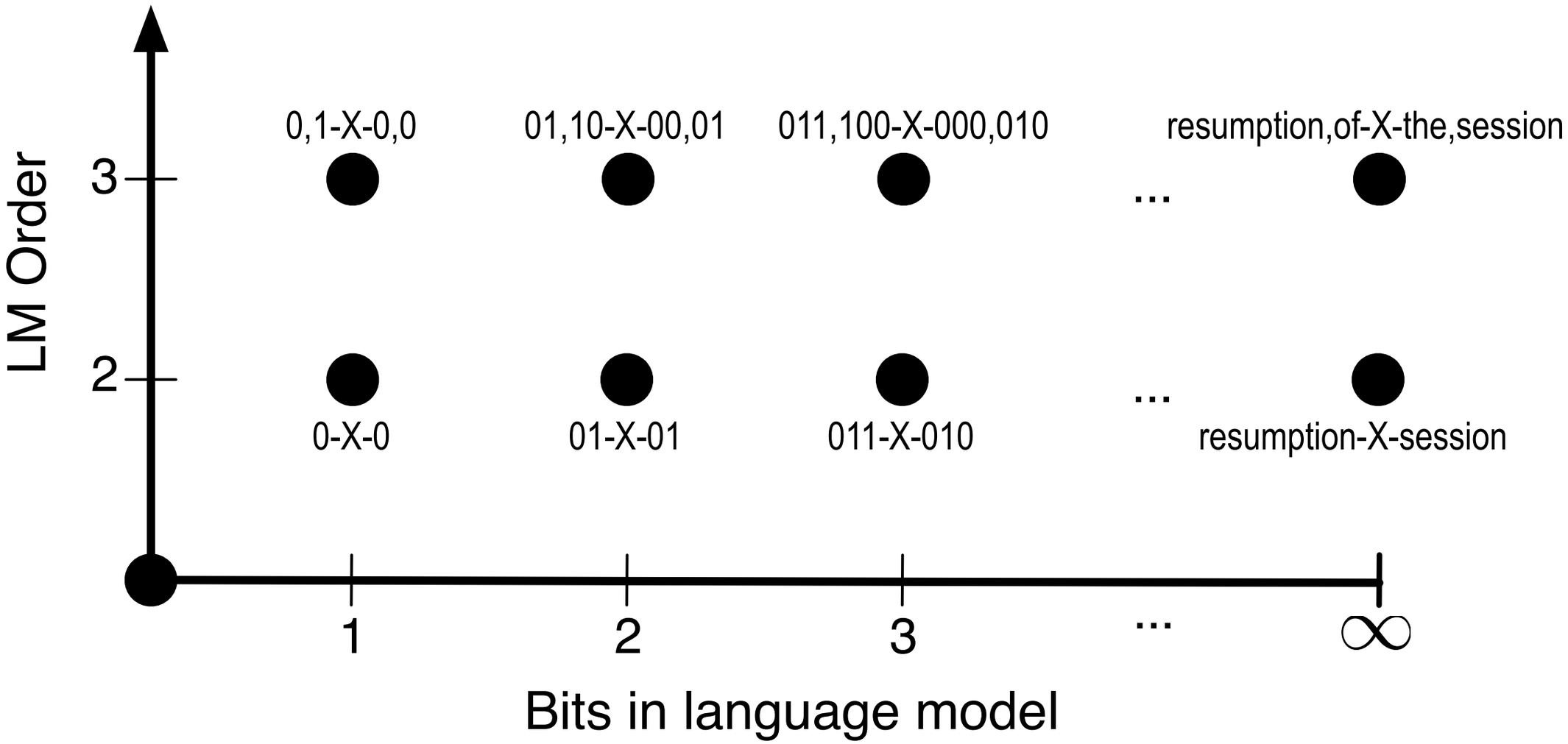
State Space Projections



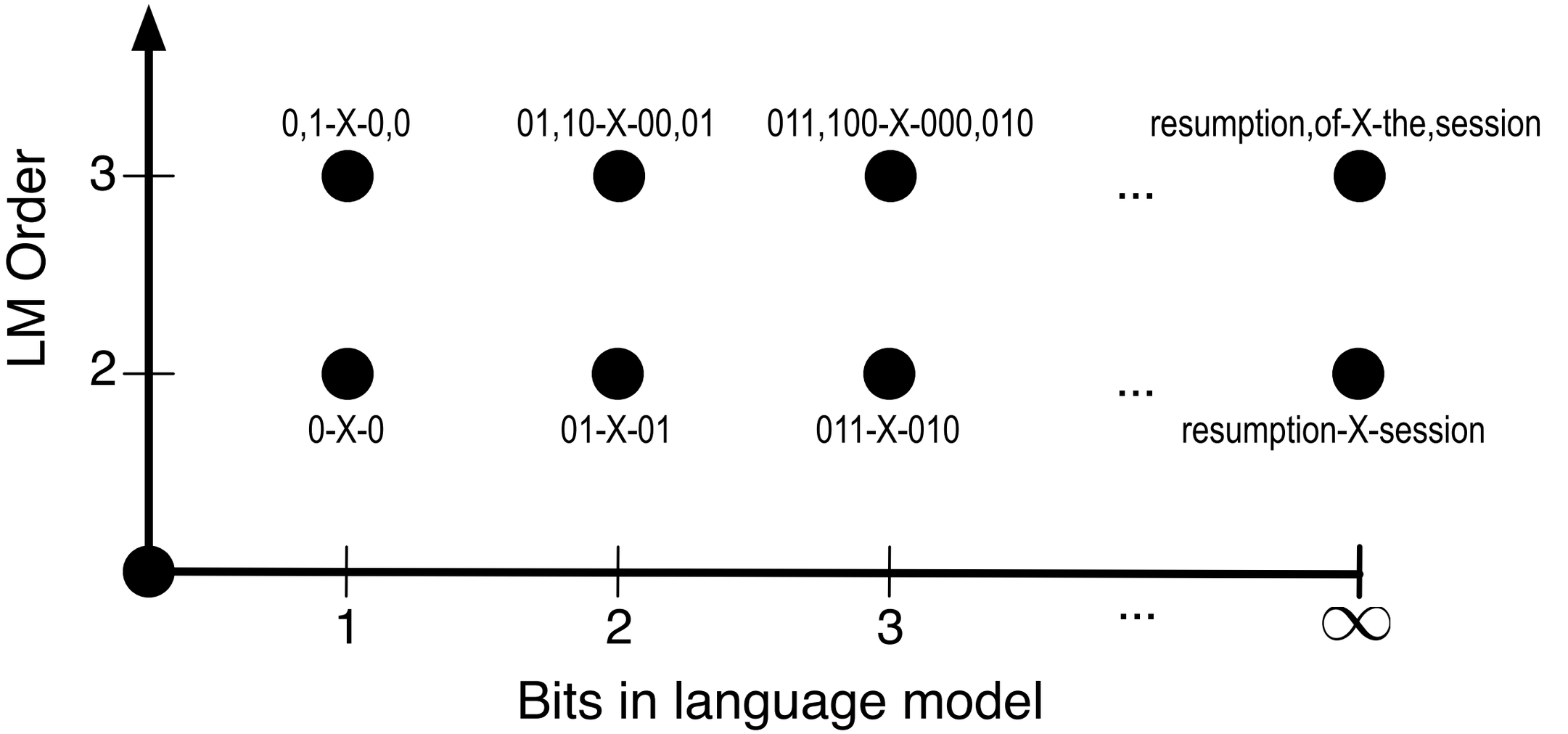
State Space Projections



State Space Projections

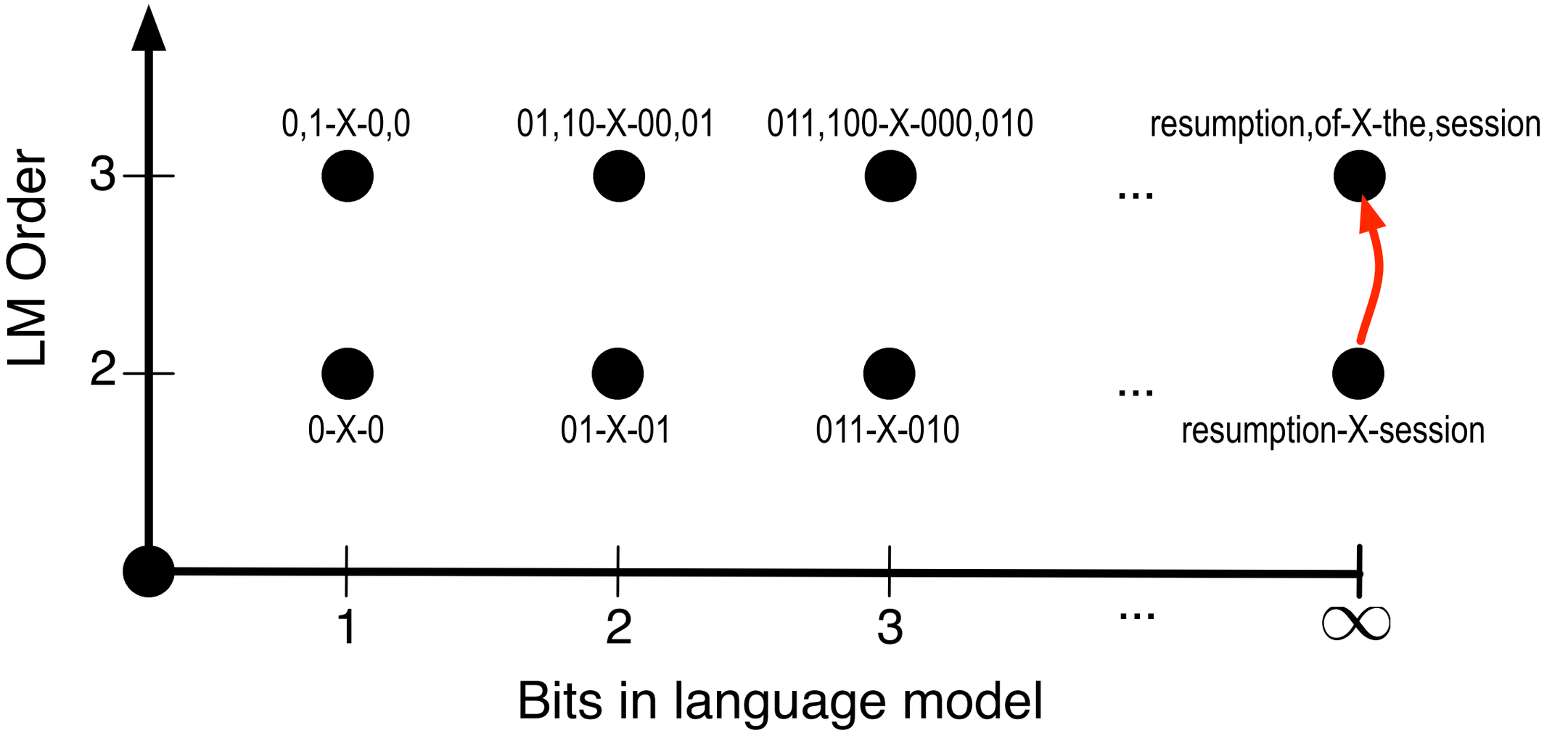


Coarse-to-Fine Inference

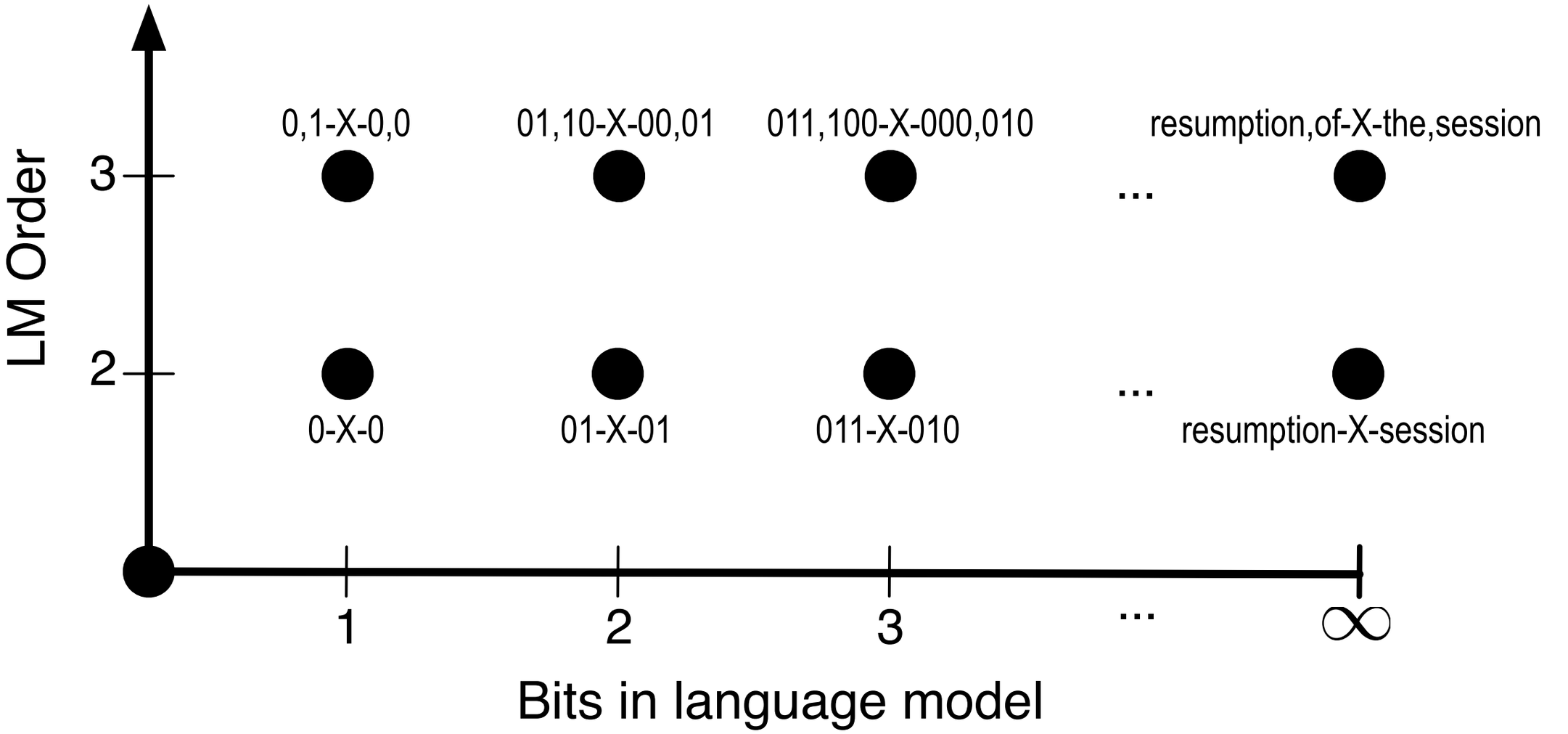


Coarse-to-Fine Inference

Order based [Zhang & Gildea '08]

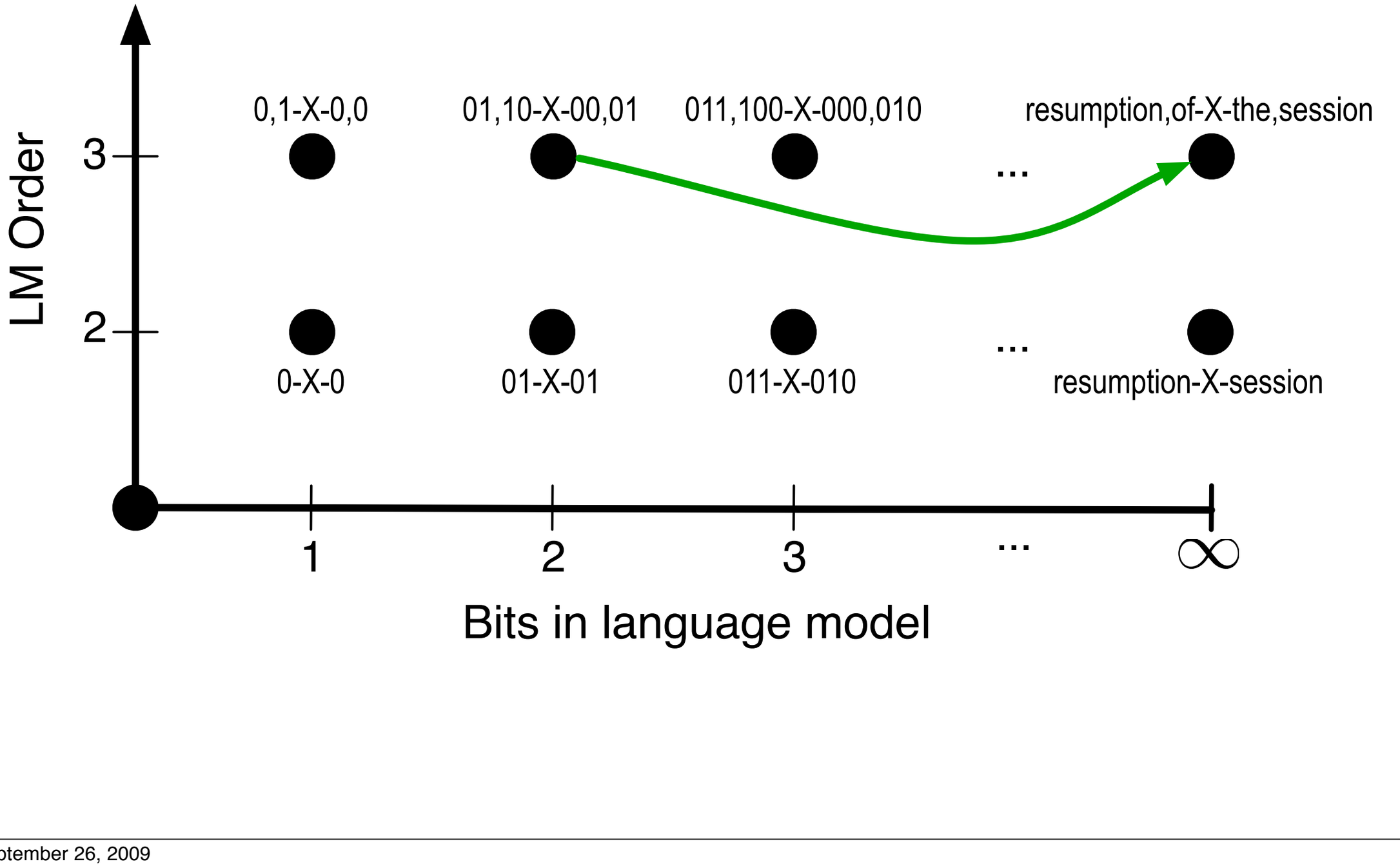


Coarse-to-Fine Inference

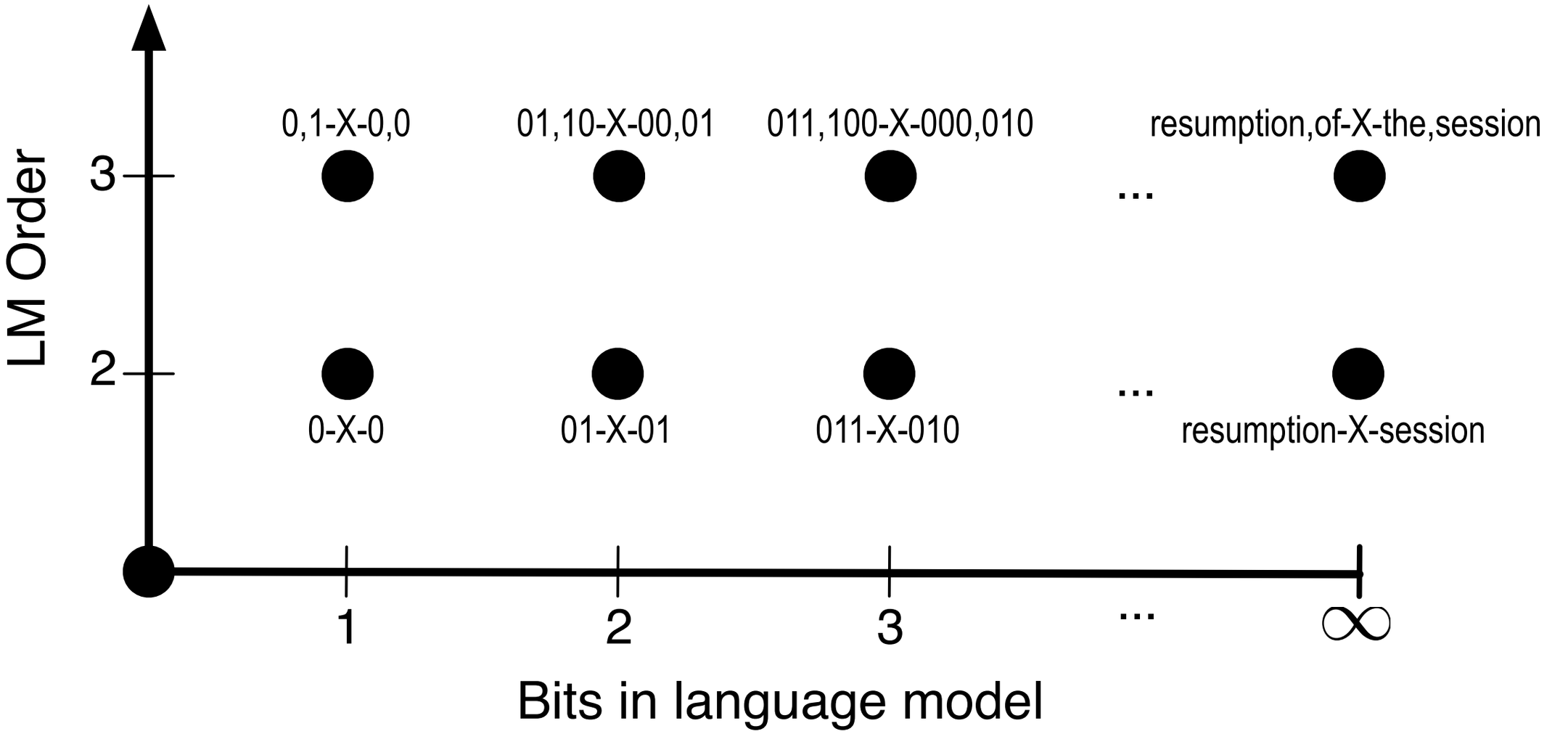


Coarse-to-Fine Inference

Encoding based

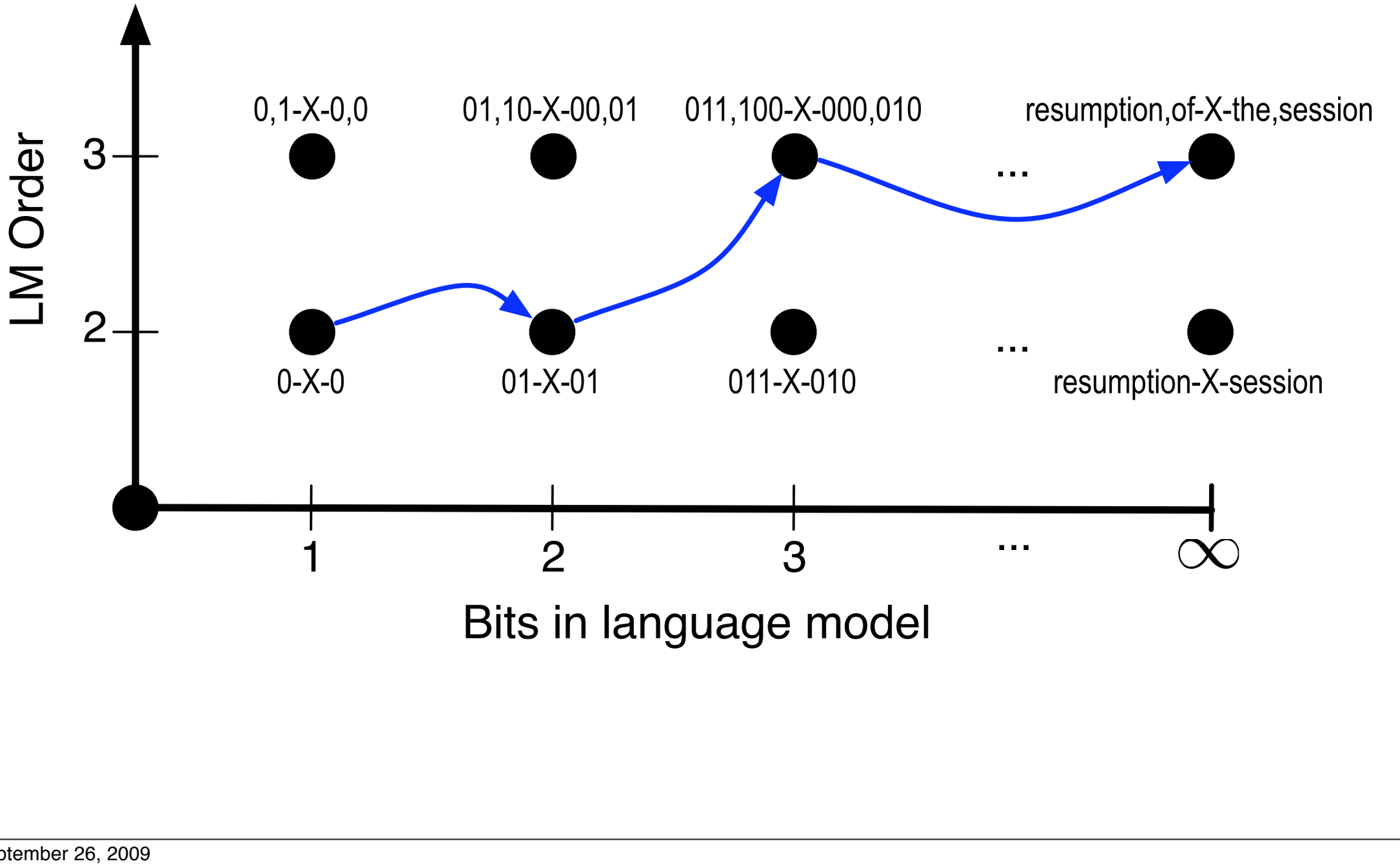


Coarse-to-Fine Inference



Coarse-to-Fine Inference

Combine both?



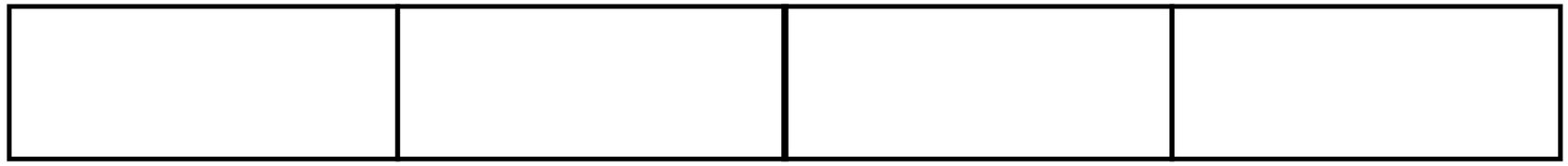
Experimental Setup

- Europarl corpus:
  - Spanish, German, French → English
  - About 8M words of parallel text
  - Average sentence length 30
  - 2K development sentences
- Pre-Processing
  - Standard GIZA++ word alignments and phrase extraction
  - SRILM trigram language model
- ITG translation model
  - Extracted phrases / scores used as ITG terminal productions
  - Simplistic distortion model: 1 monotonic and 1 inverted rule
  - Less than 1 BLEU point below Moses on same setup

Coarse-to-Fine Pruning

For each span:

Coarse Pass:



Coarse-to-Fine Pruning

For each span:

Coarse Pass:

0-X-0	0-X-1	1-X-0	1-X-1

Coarse-to-Fine Pruning

For each span:

Coarse Pass:

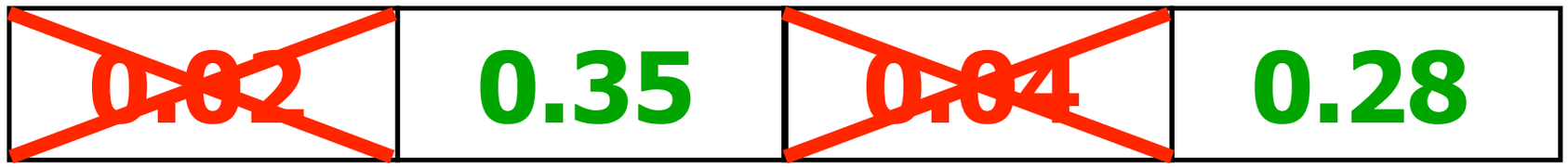
Posteriors	0.02	0.35	0.04	0.28

Coarse-to-Fine Pruning

For each span:

Coarse Pass:

Posteriors



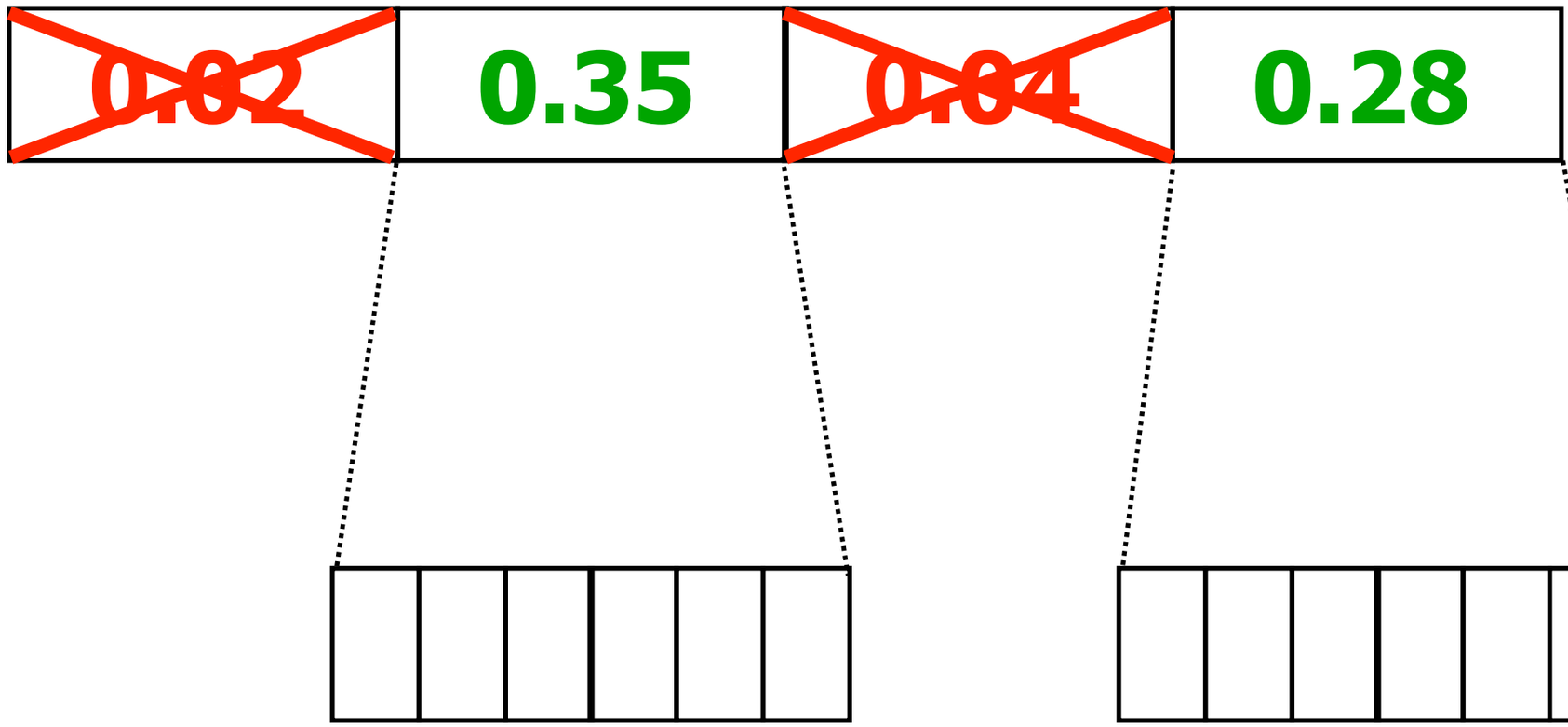
< t

Coarse-to-Fine Pruning

For each span:

Coarse Pass:

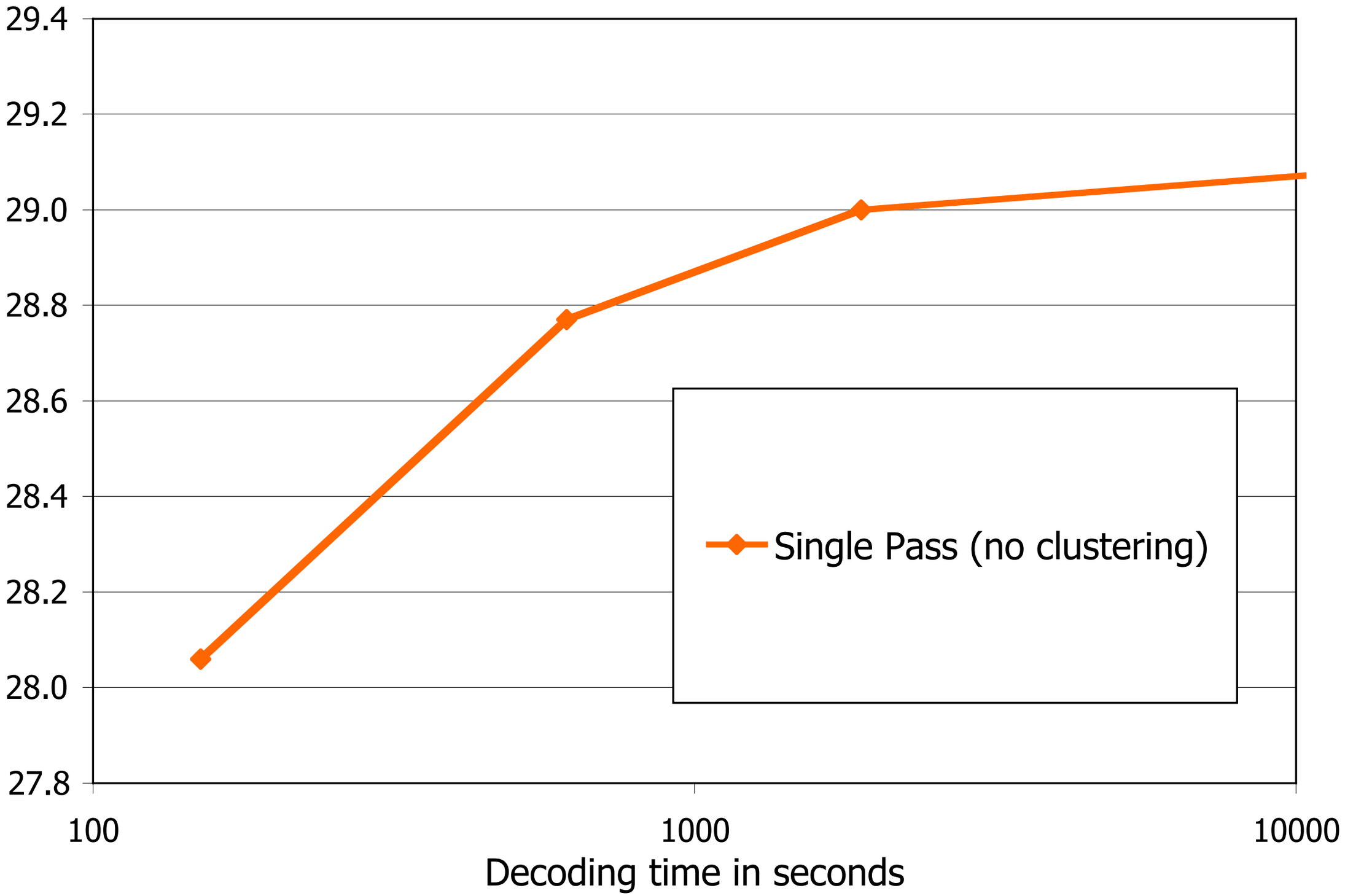
Posteriors



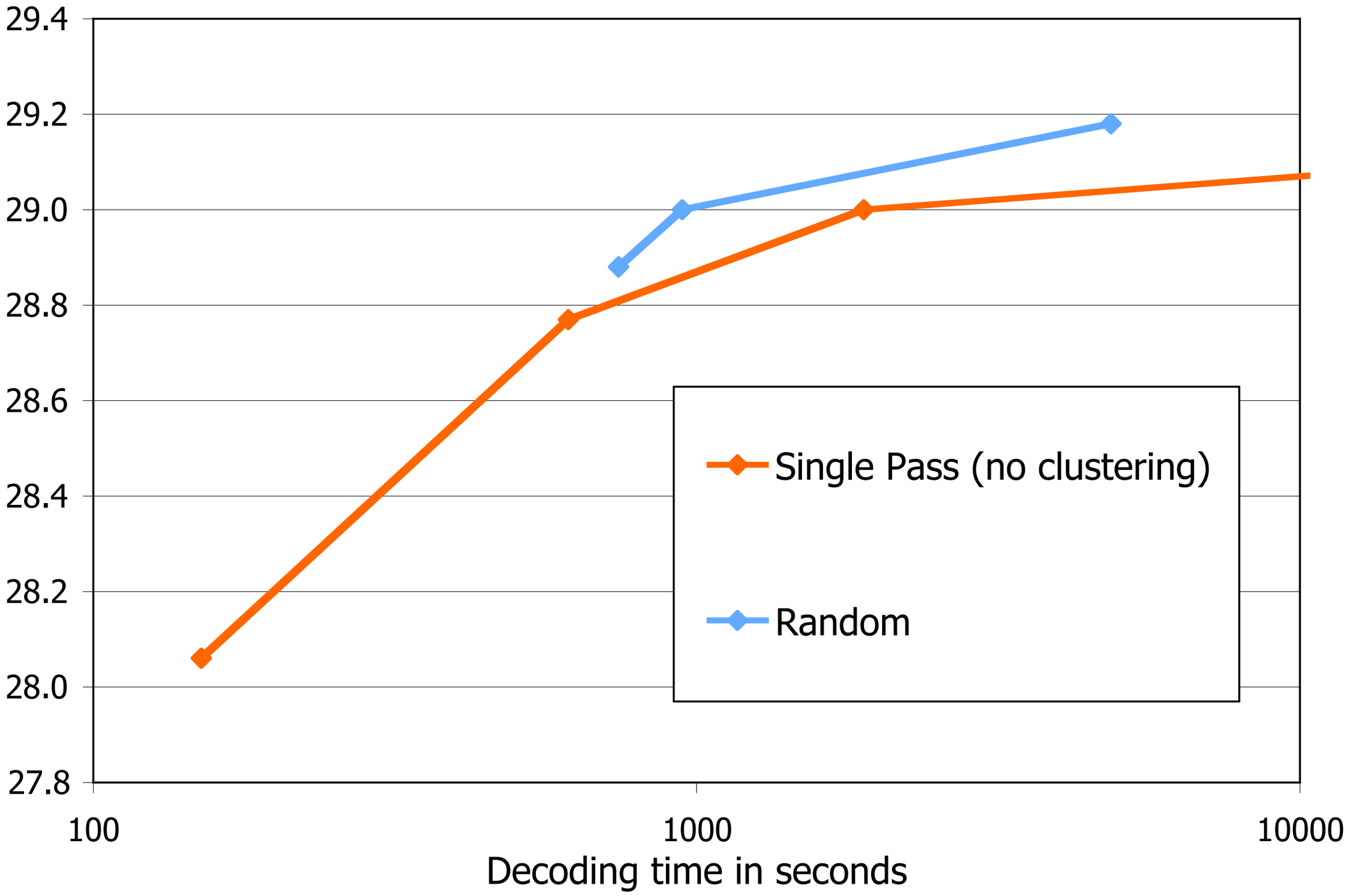
< t

Finer Pass:

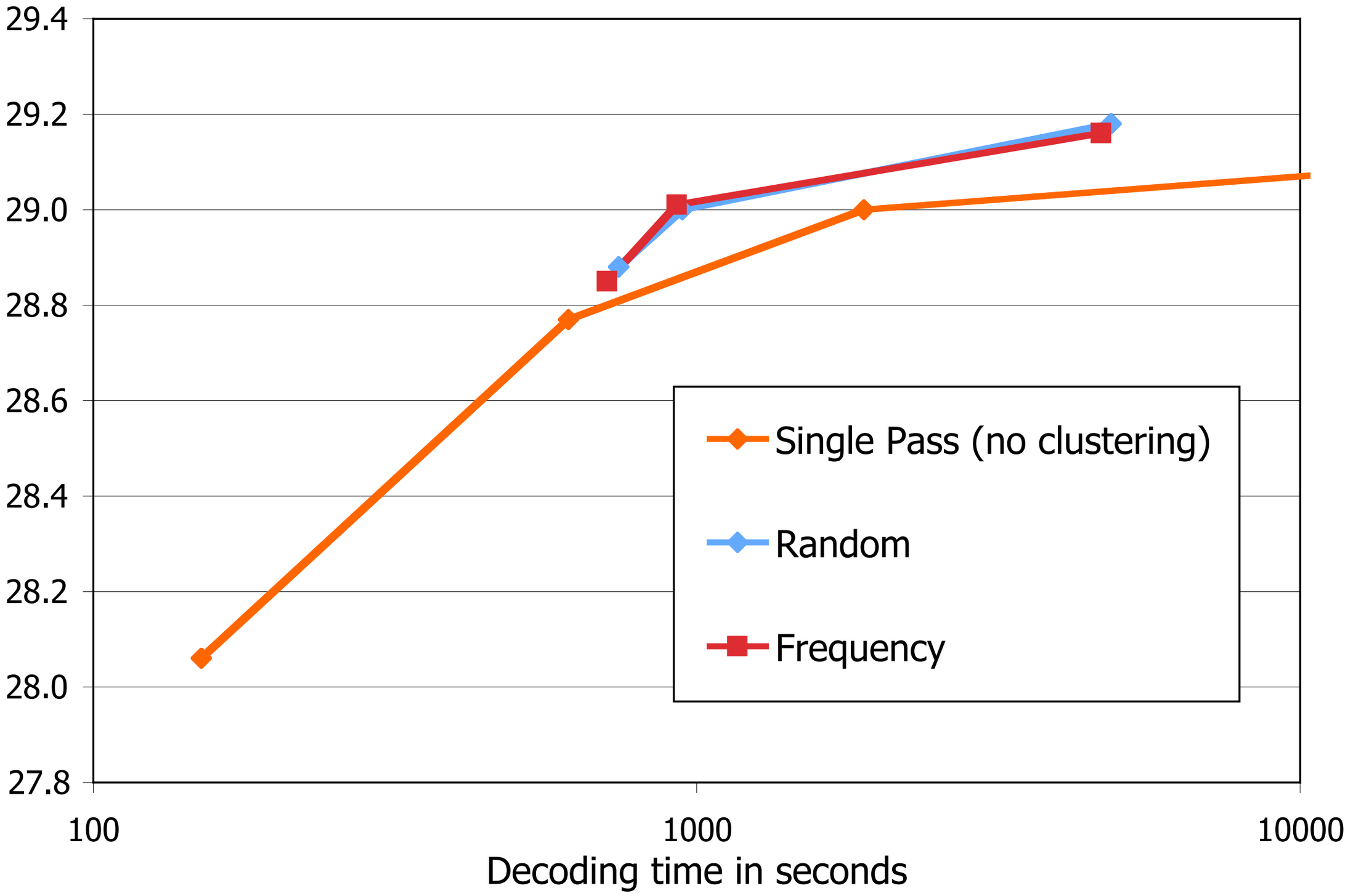
Cluster Quality (BLEU)



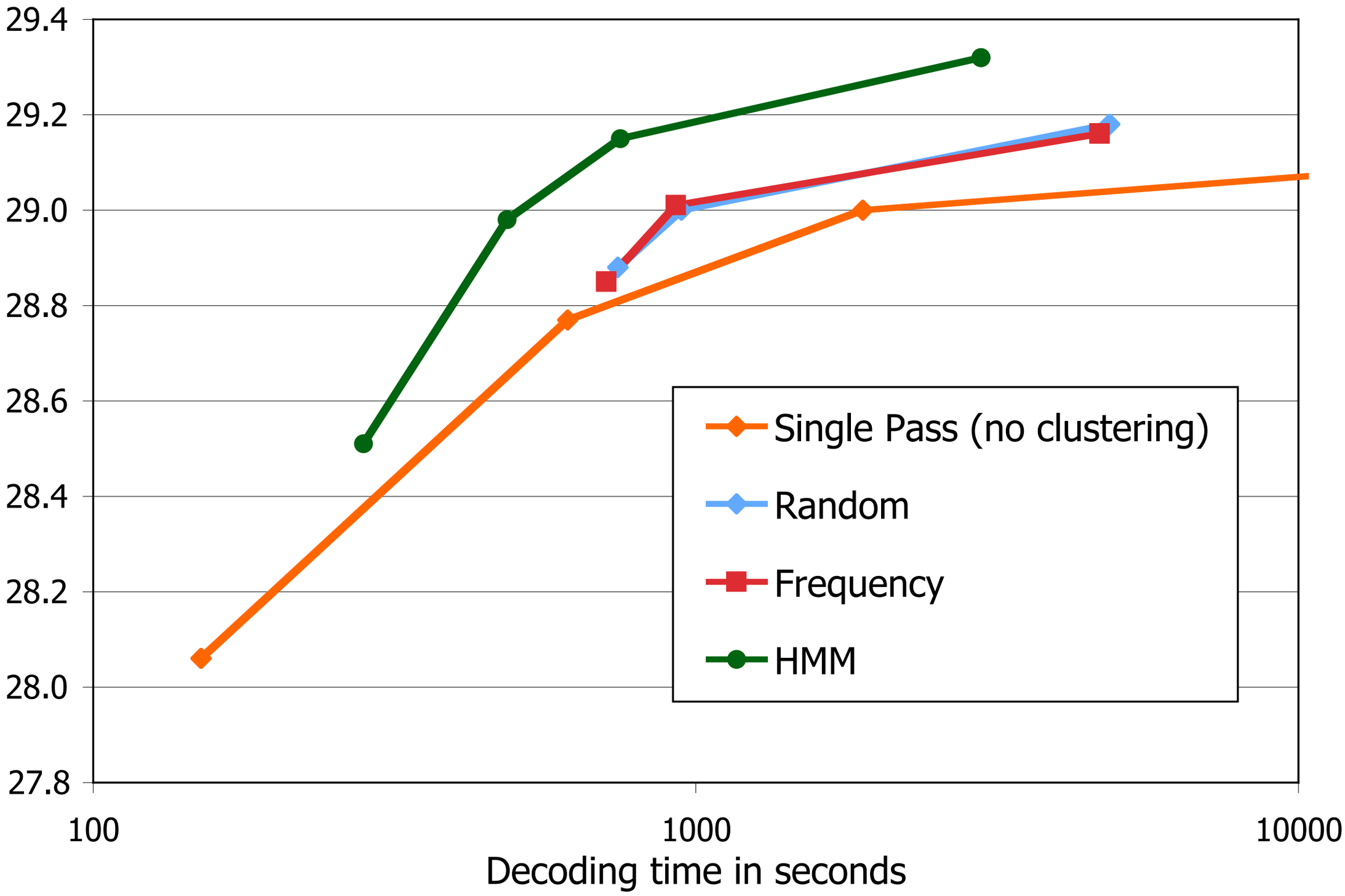
Cluster Quality (BLEU)



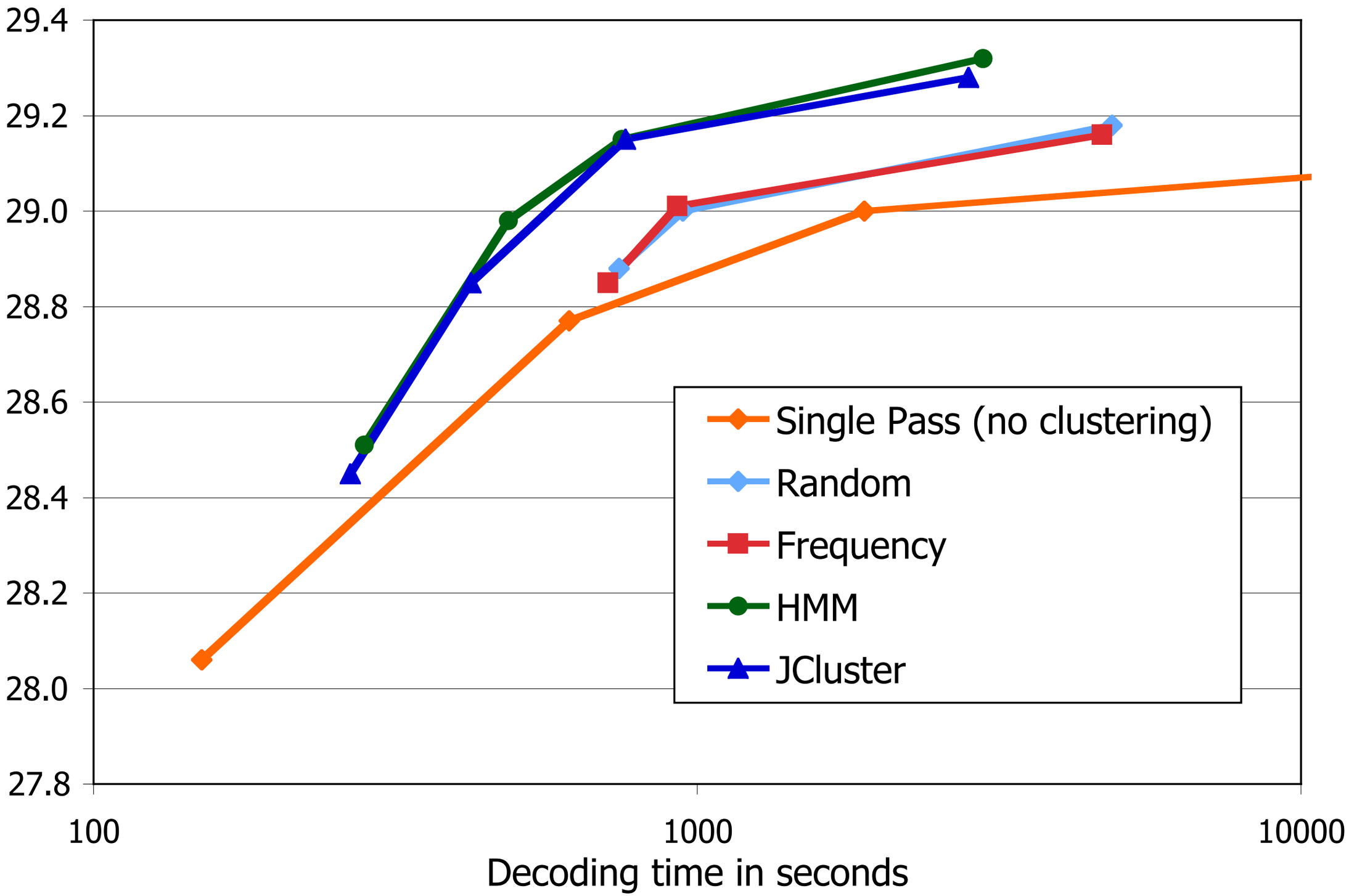
Cluster Quality (BLEU)



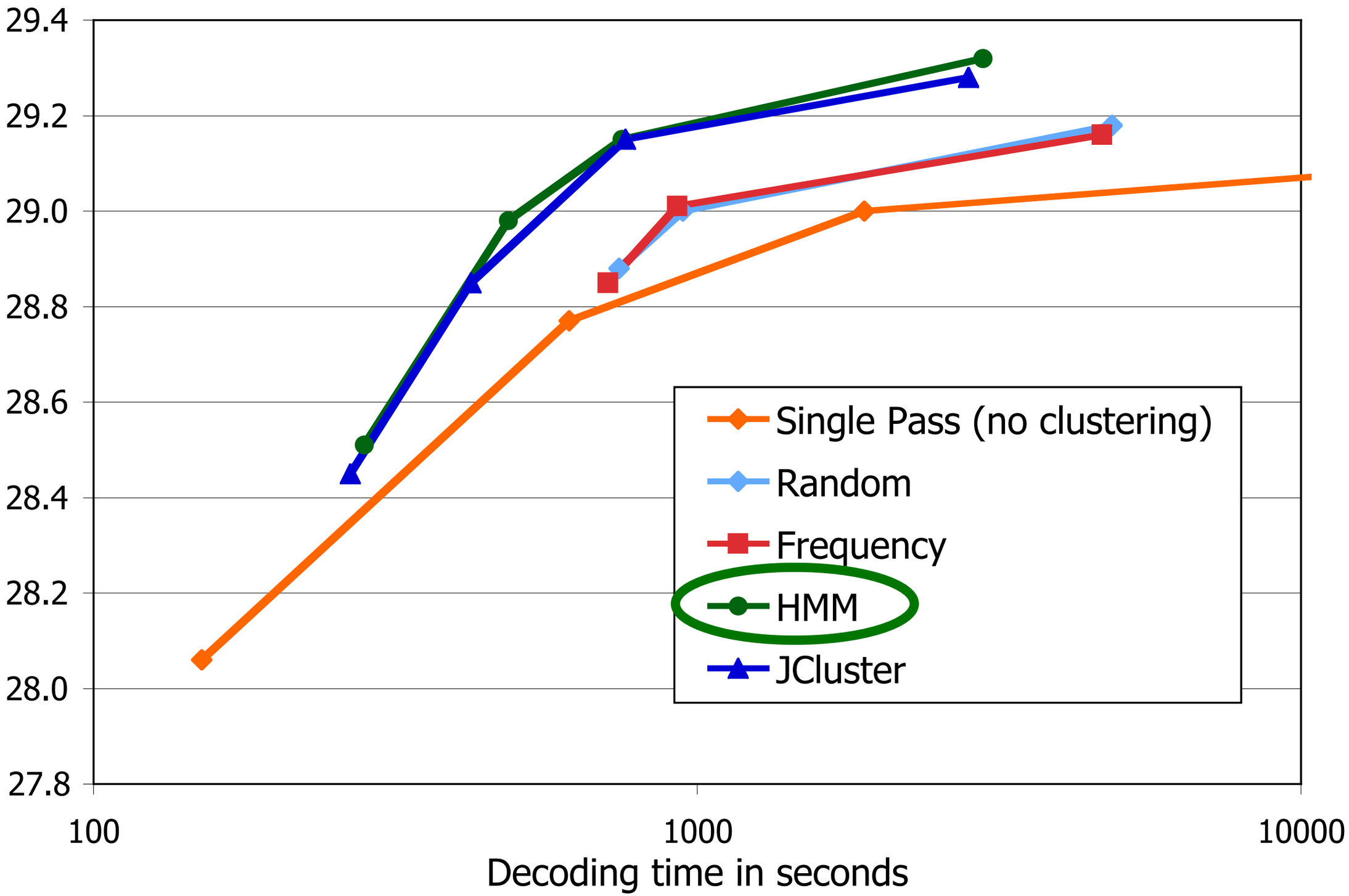
Cluster Quality (BLEU)



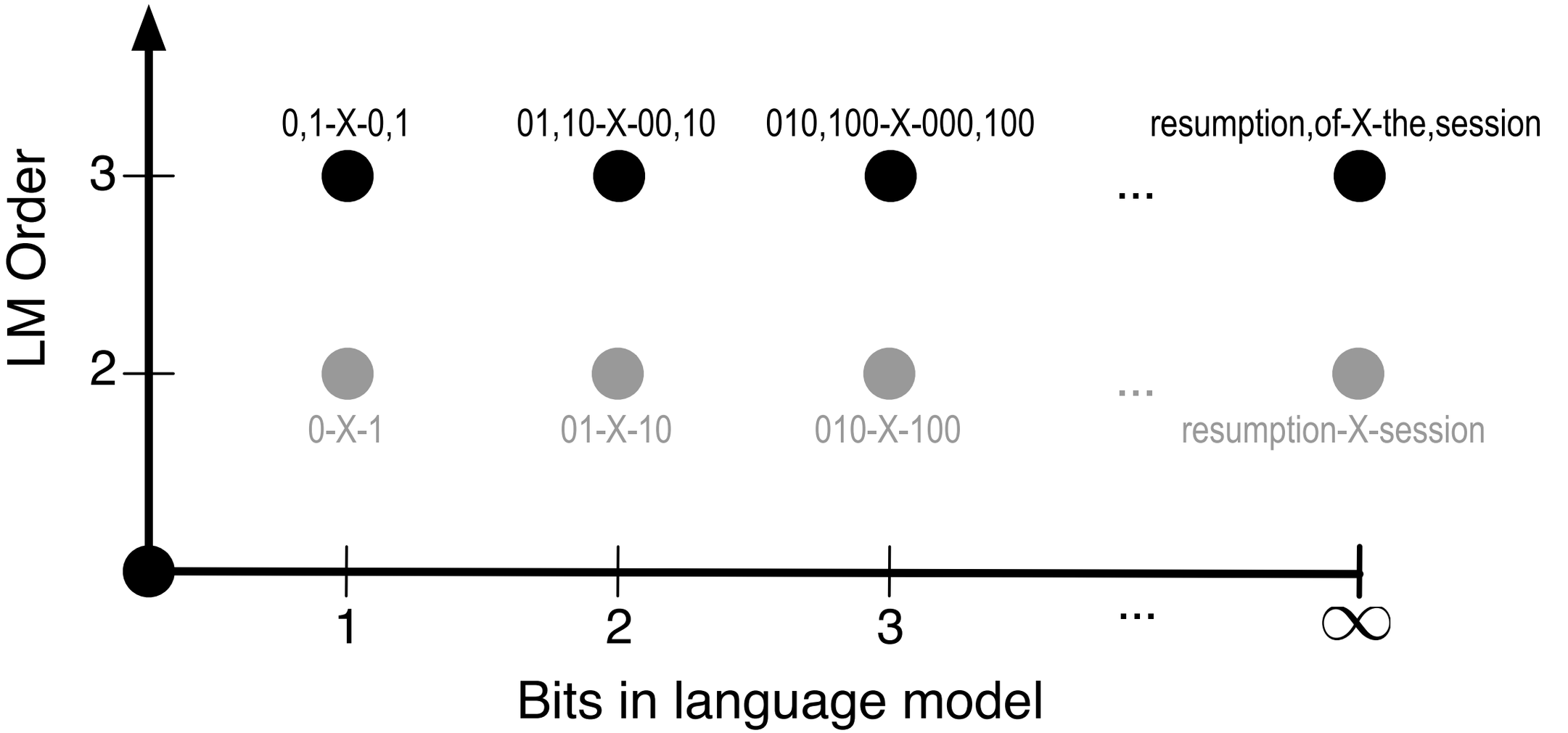
Cluster Quality (BLEU)



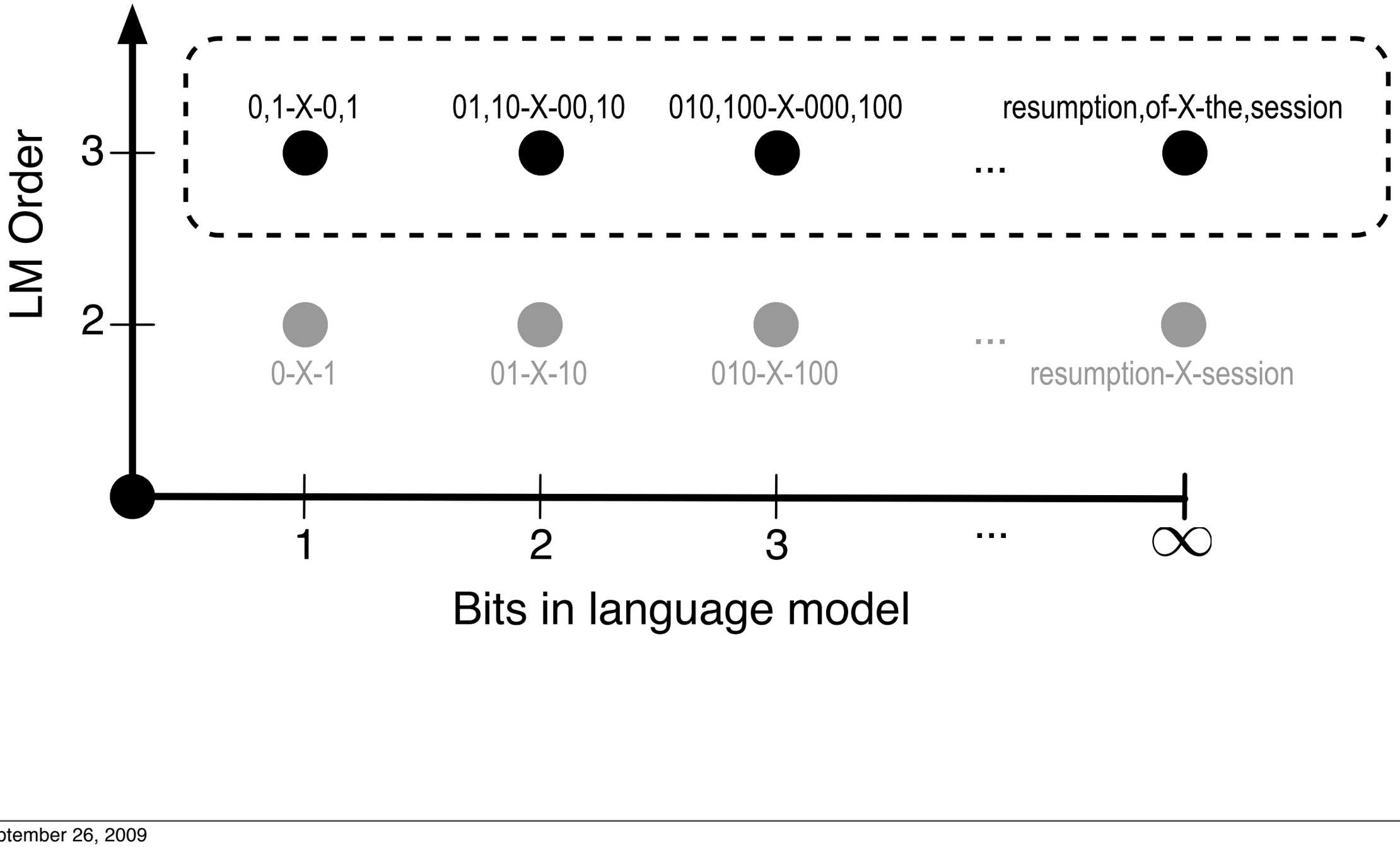
Cluster Quality (BLEU)



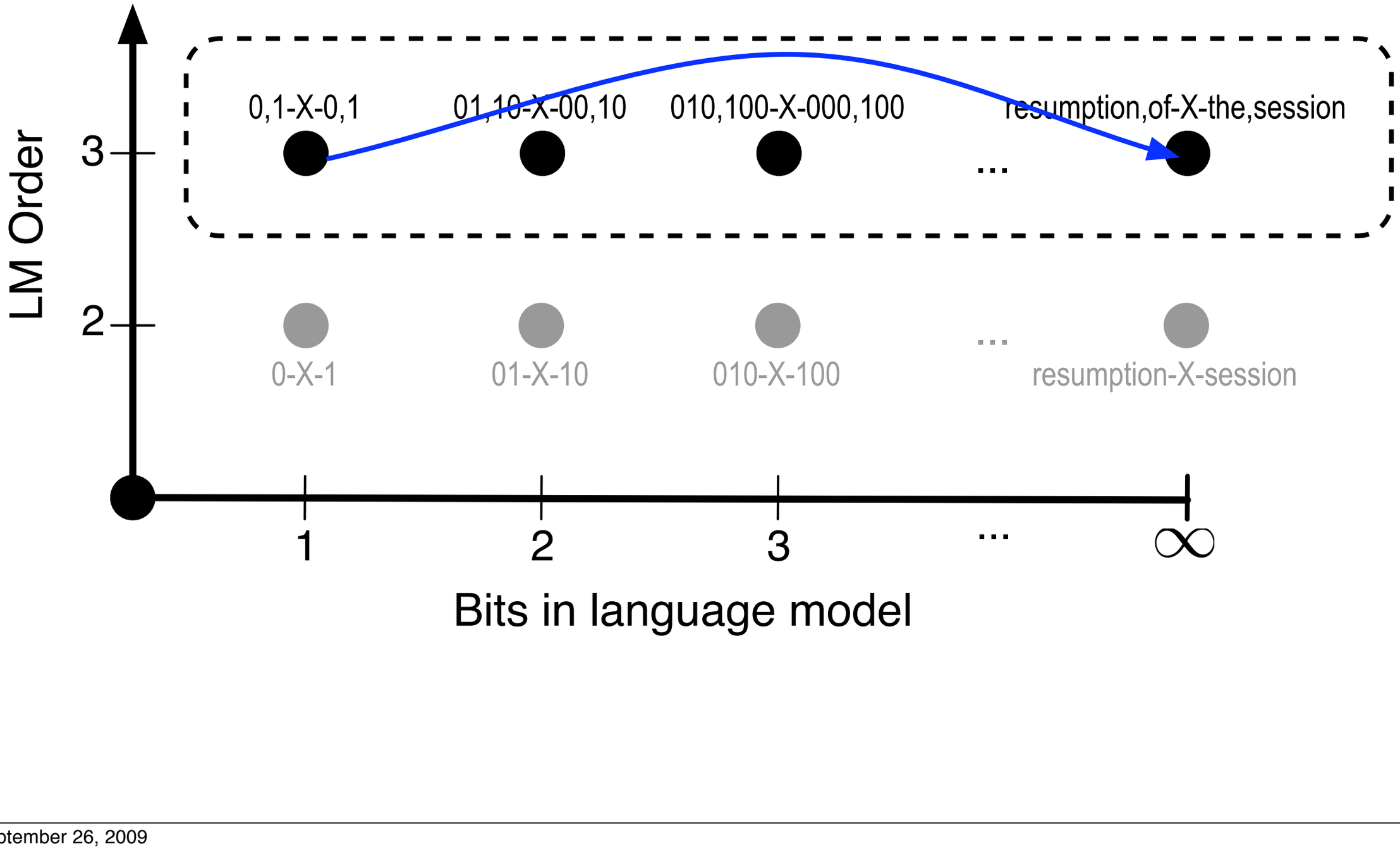
Where to start?



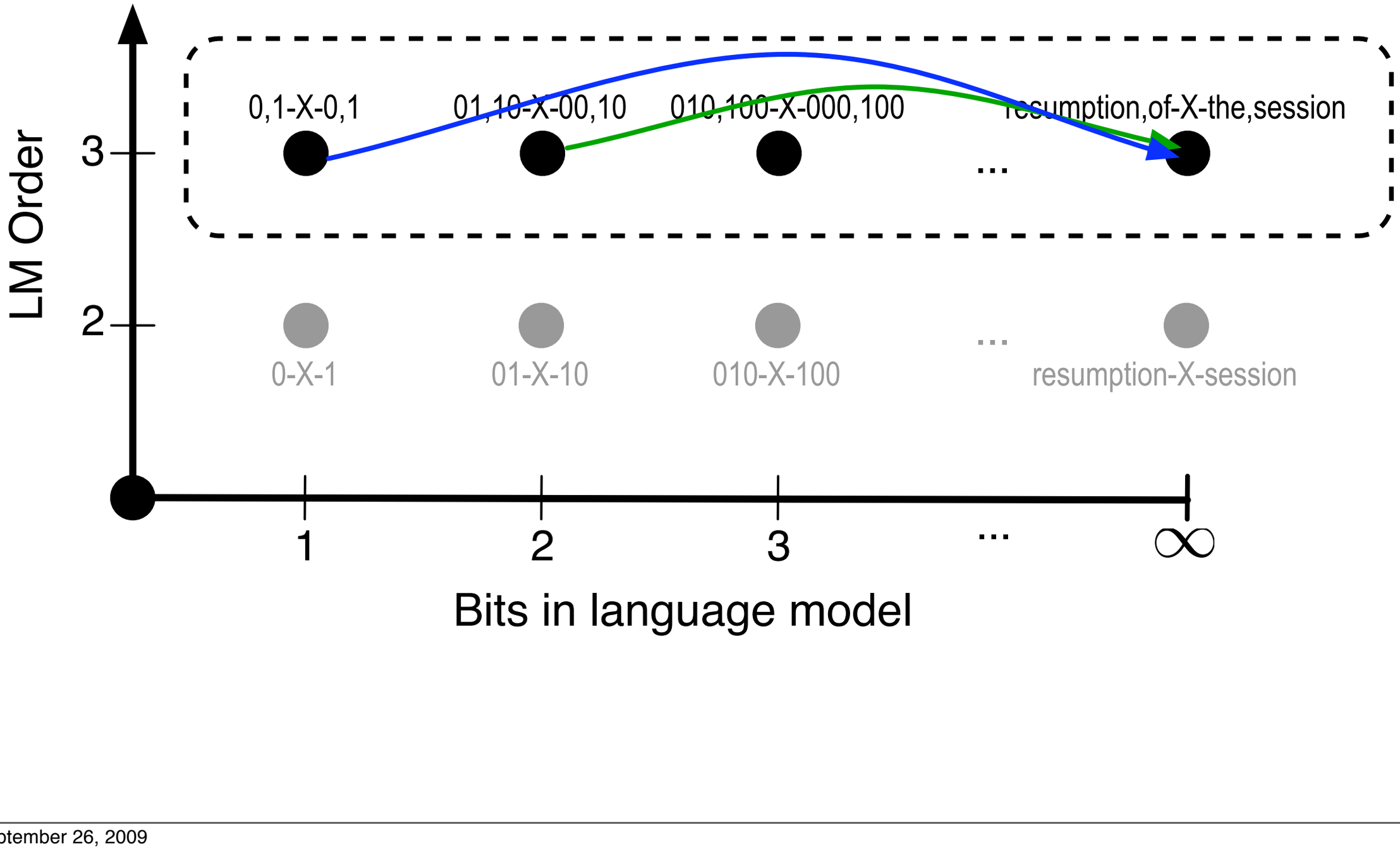
Where to start?



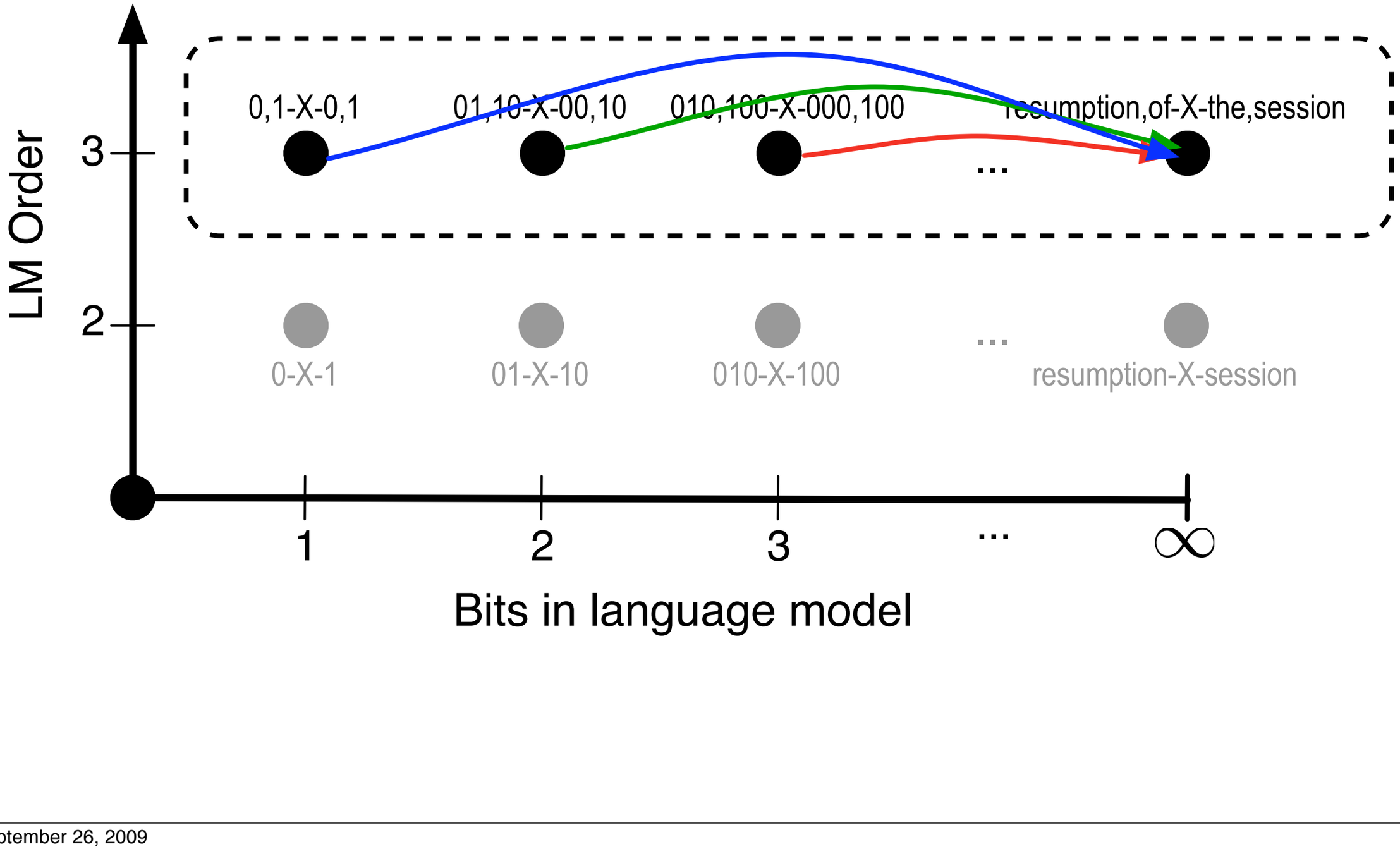
Where to start?



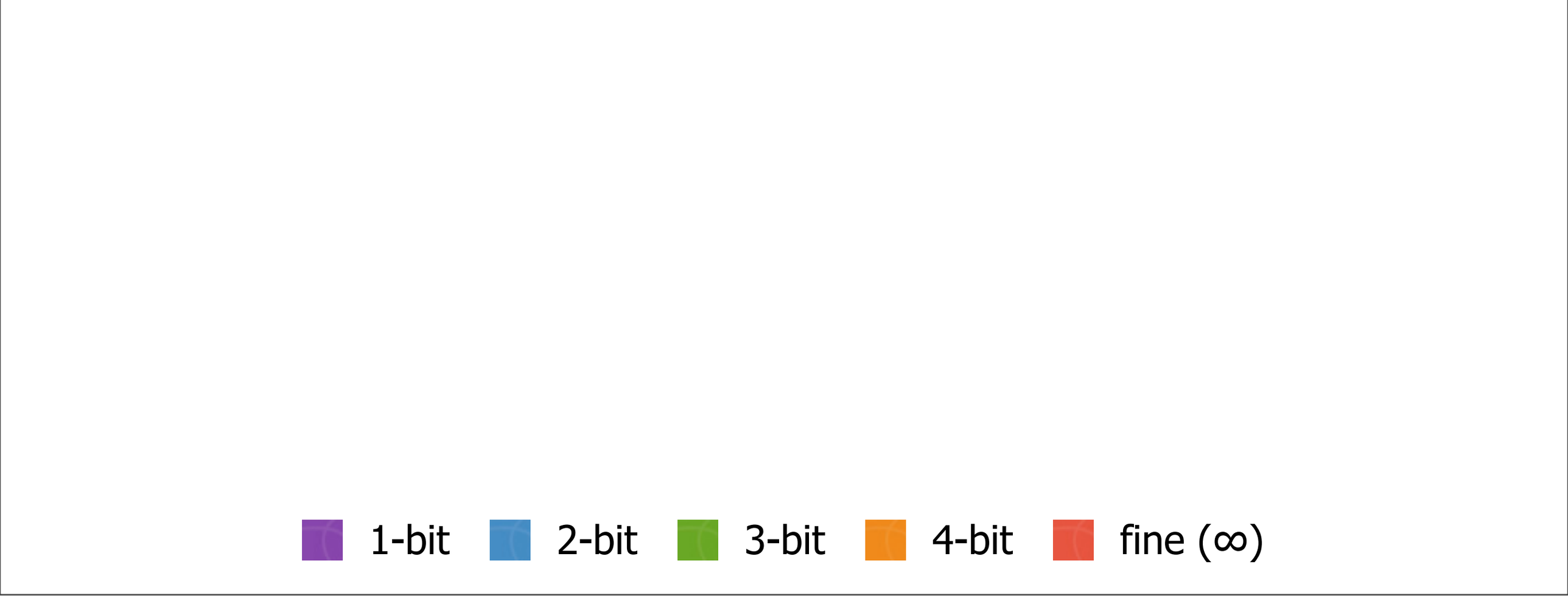
Where to start?



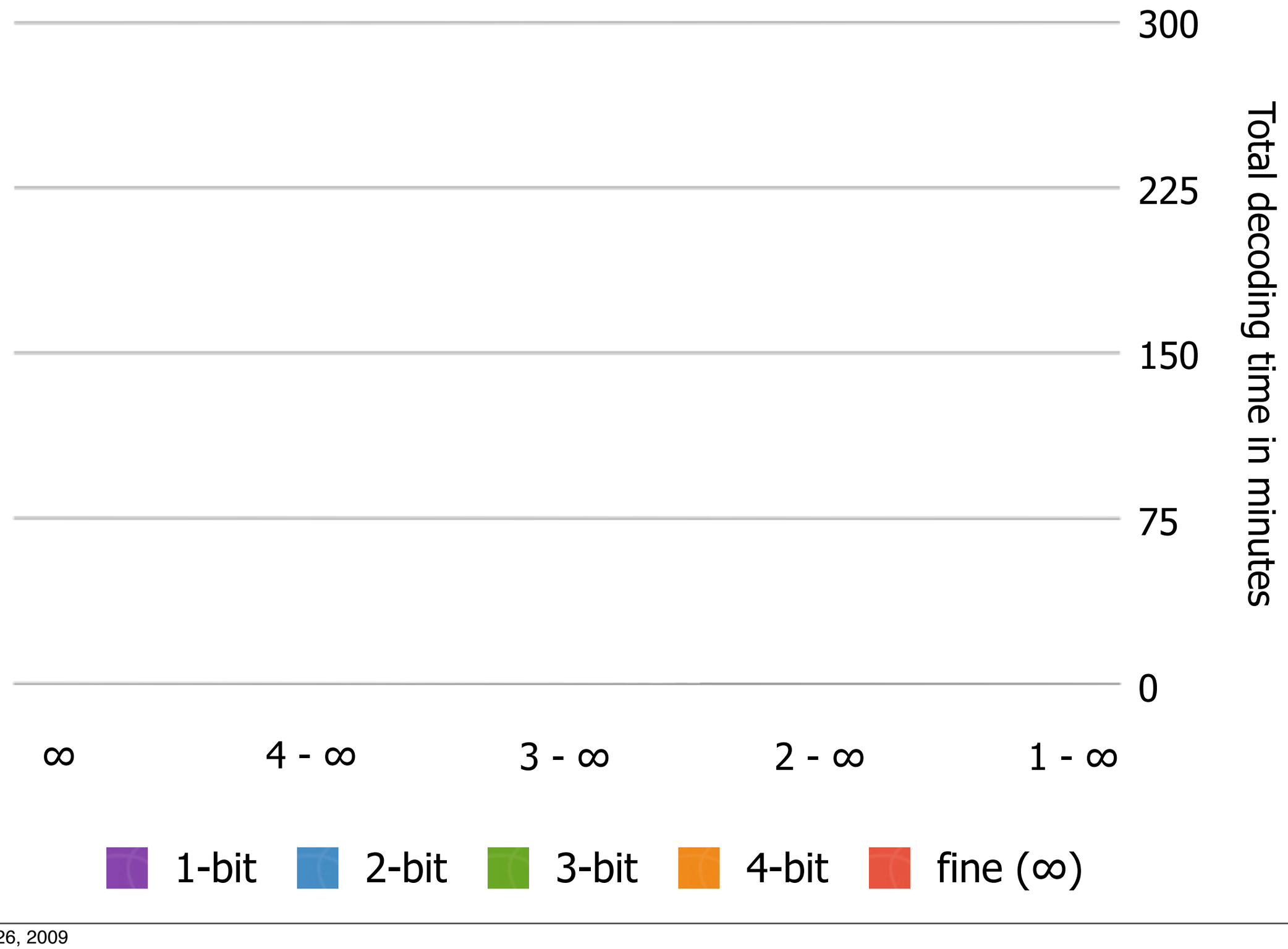
Where to start?



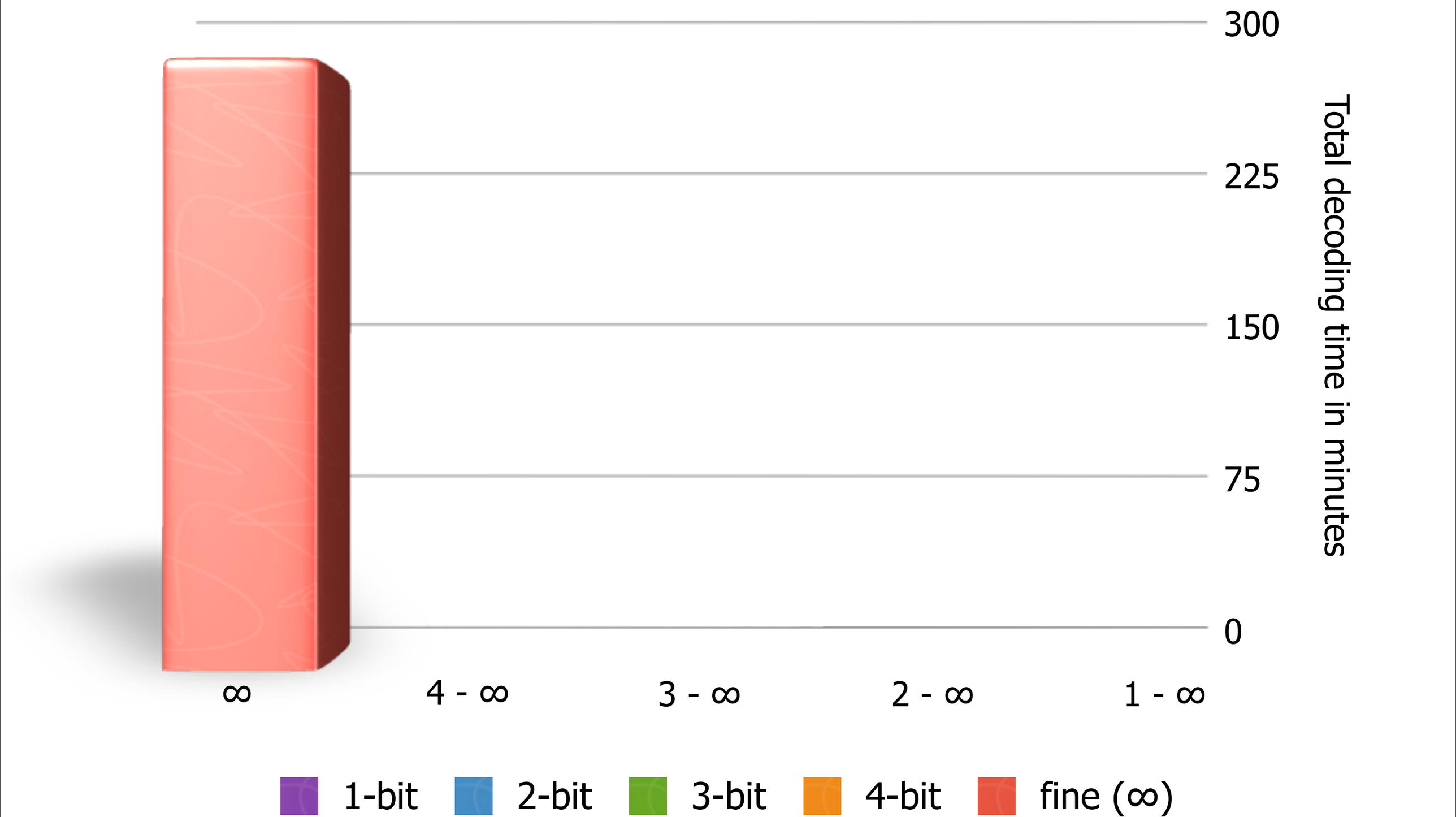
Where to start?



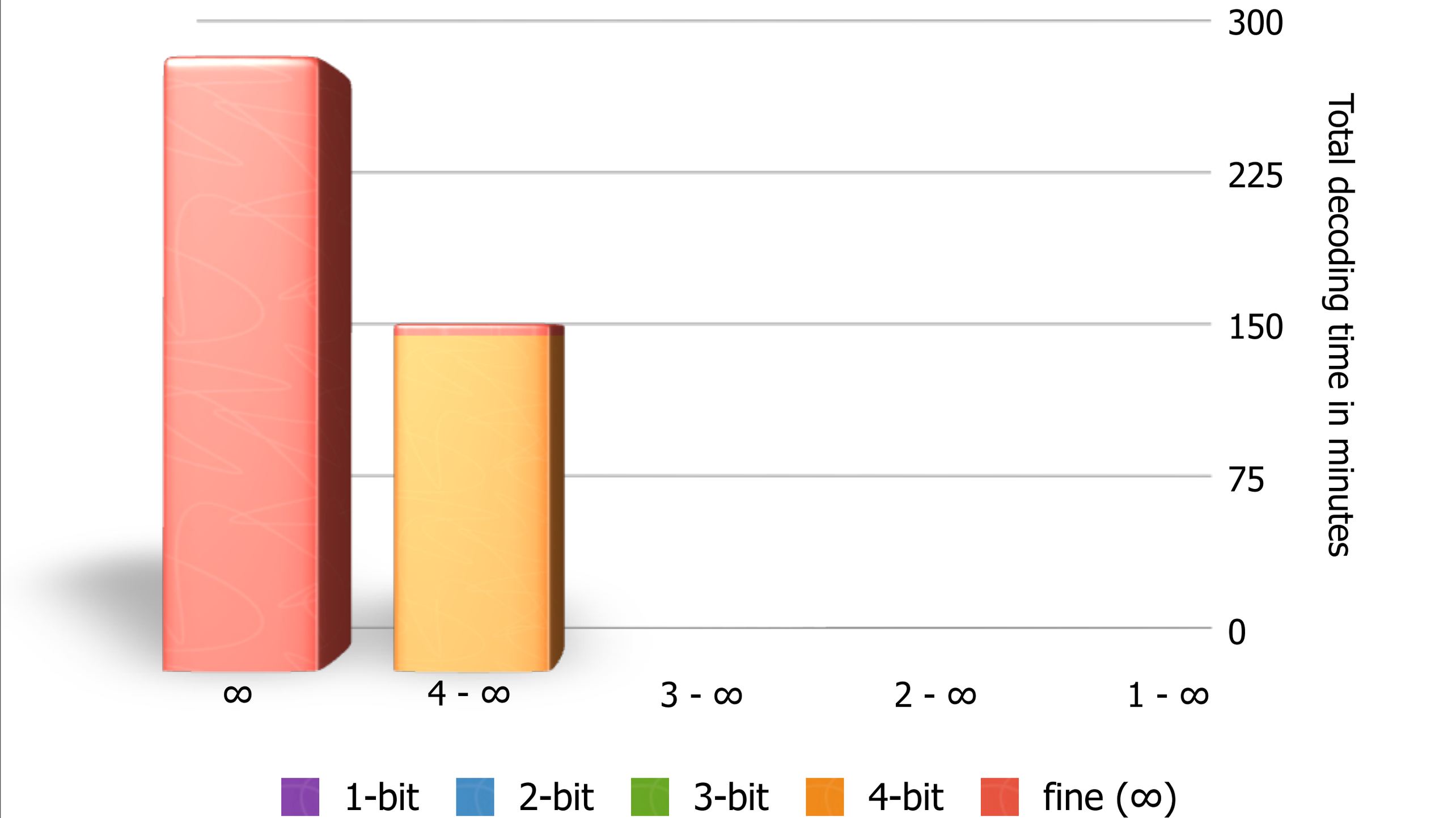
Where to start?



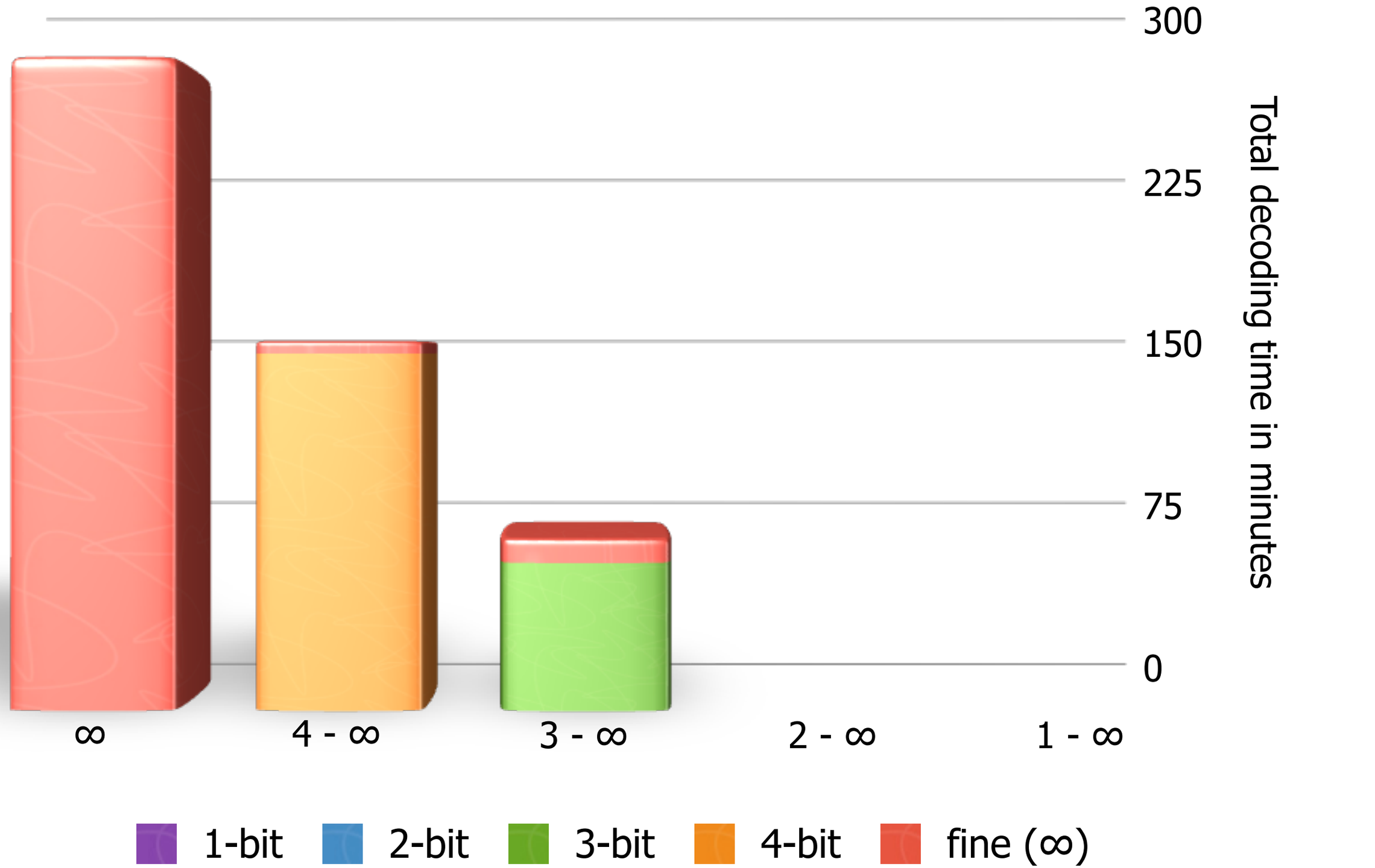
Where to start?



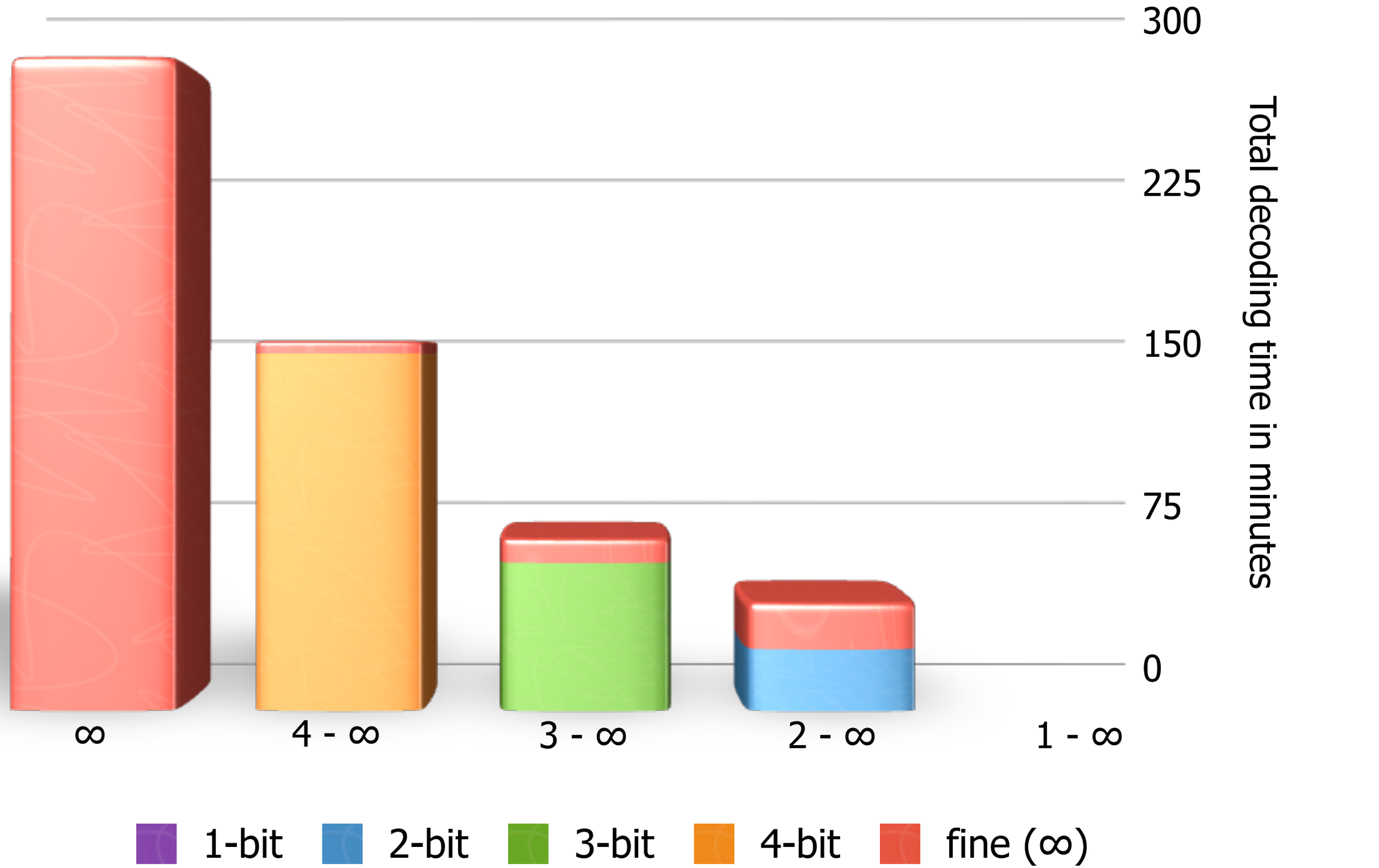
Where to start?



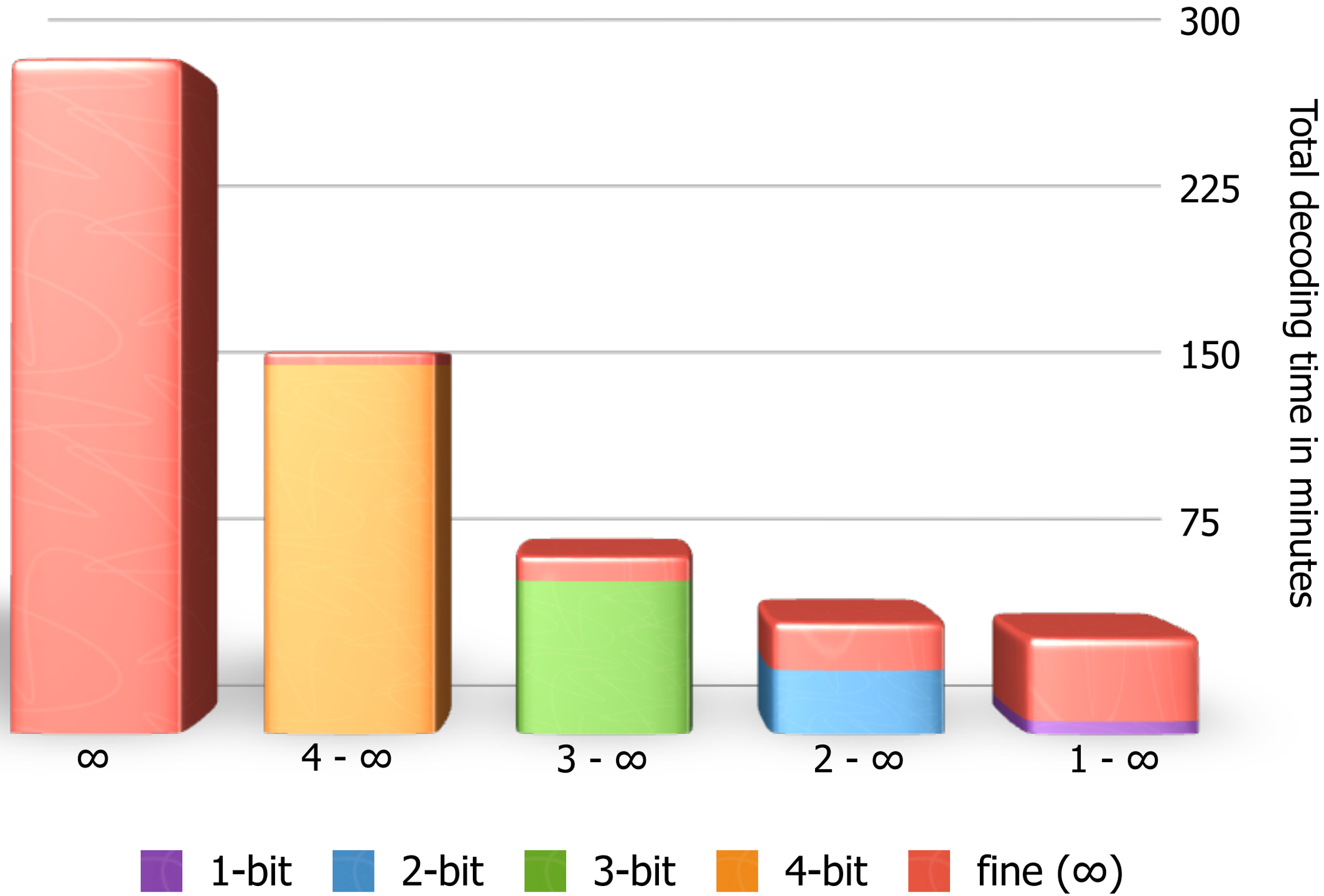
Where to start?



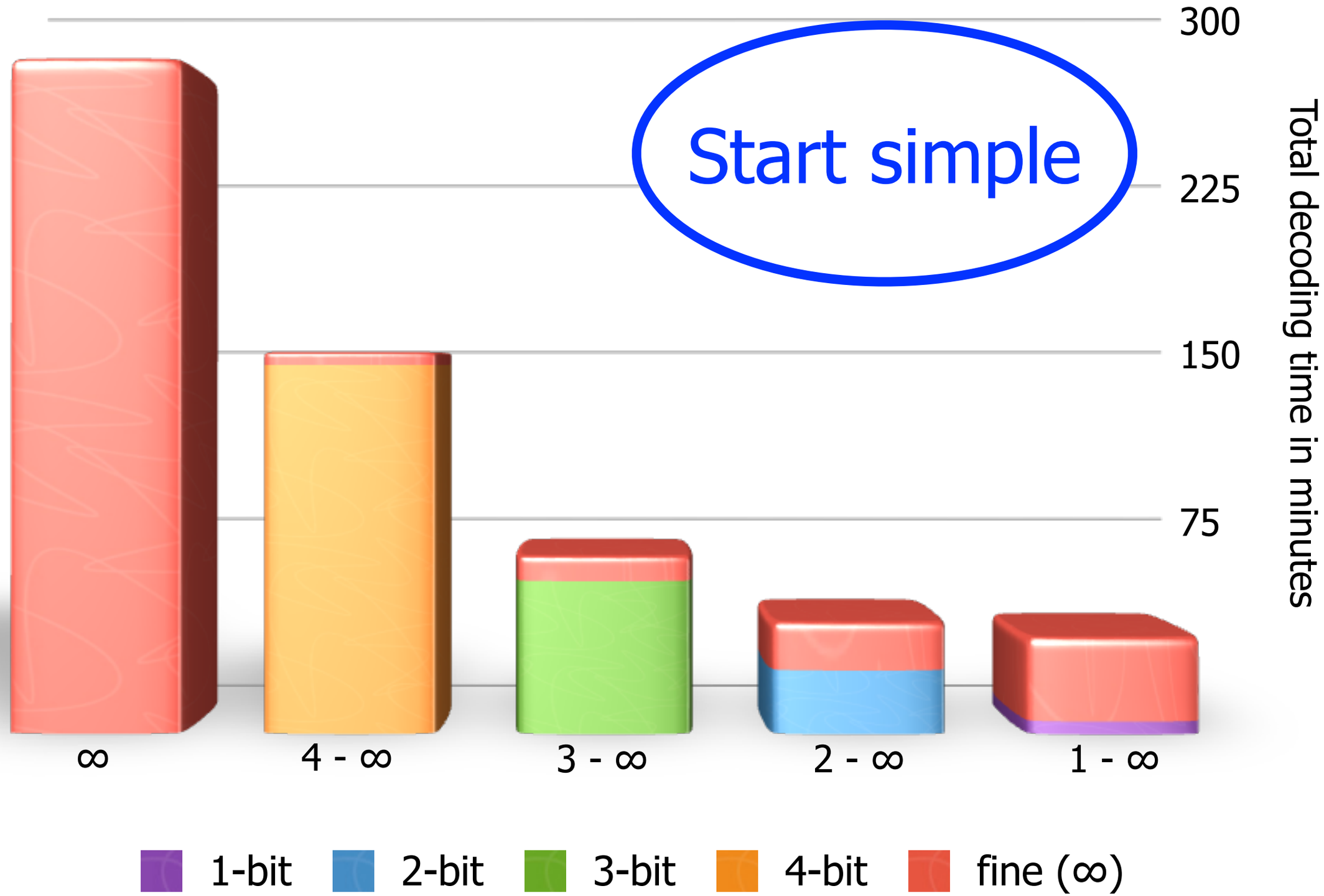
Where to start?



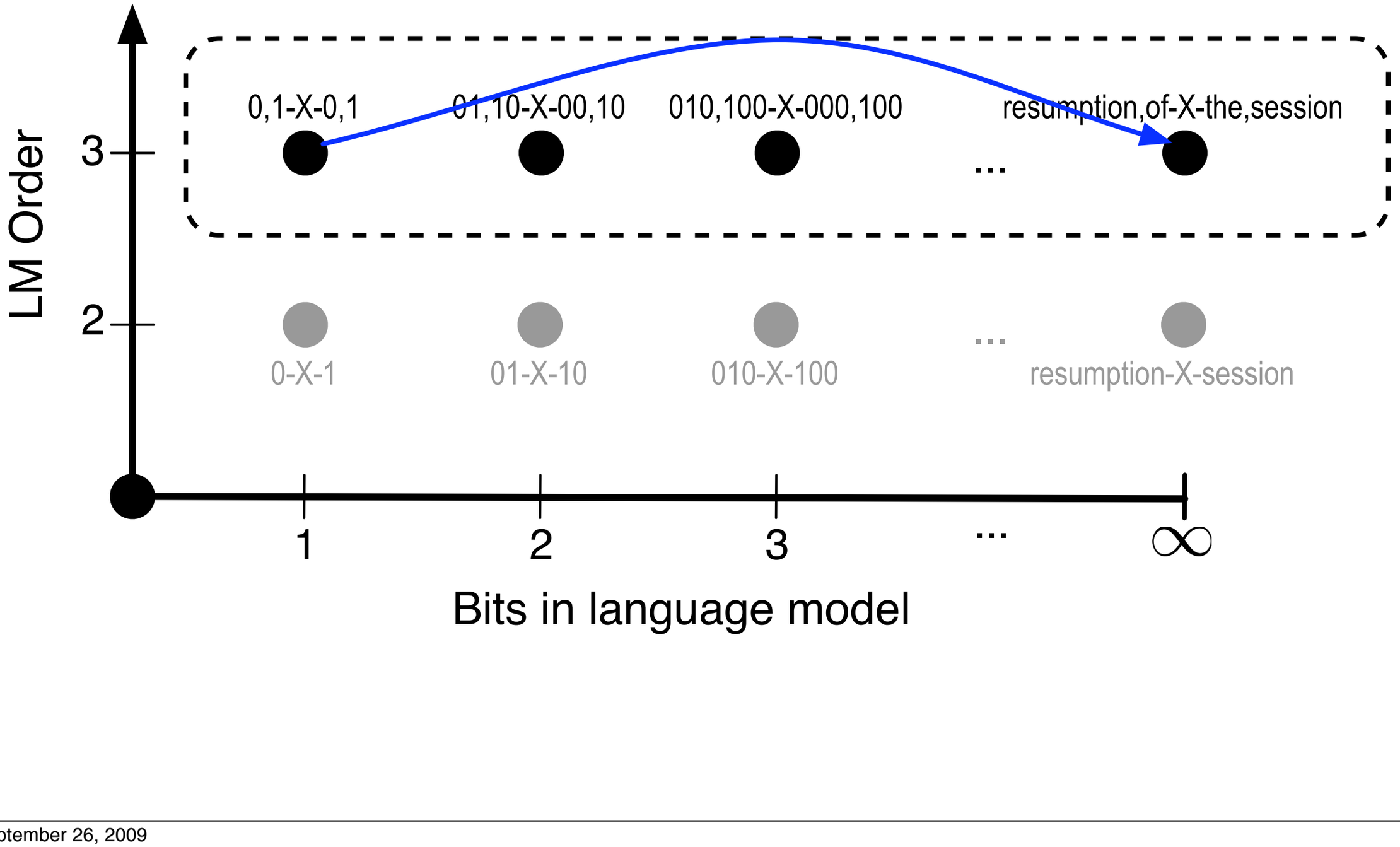
Where to start?



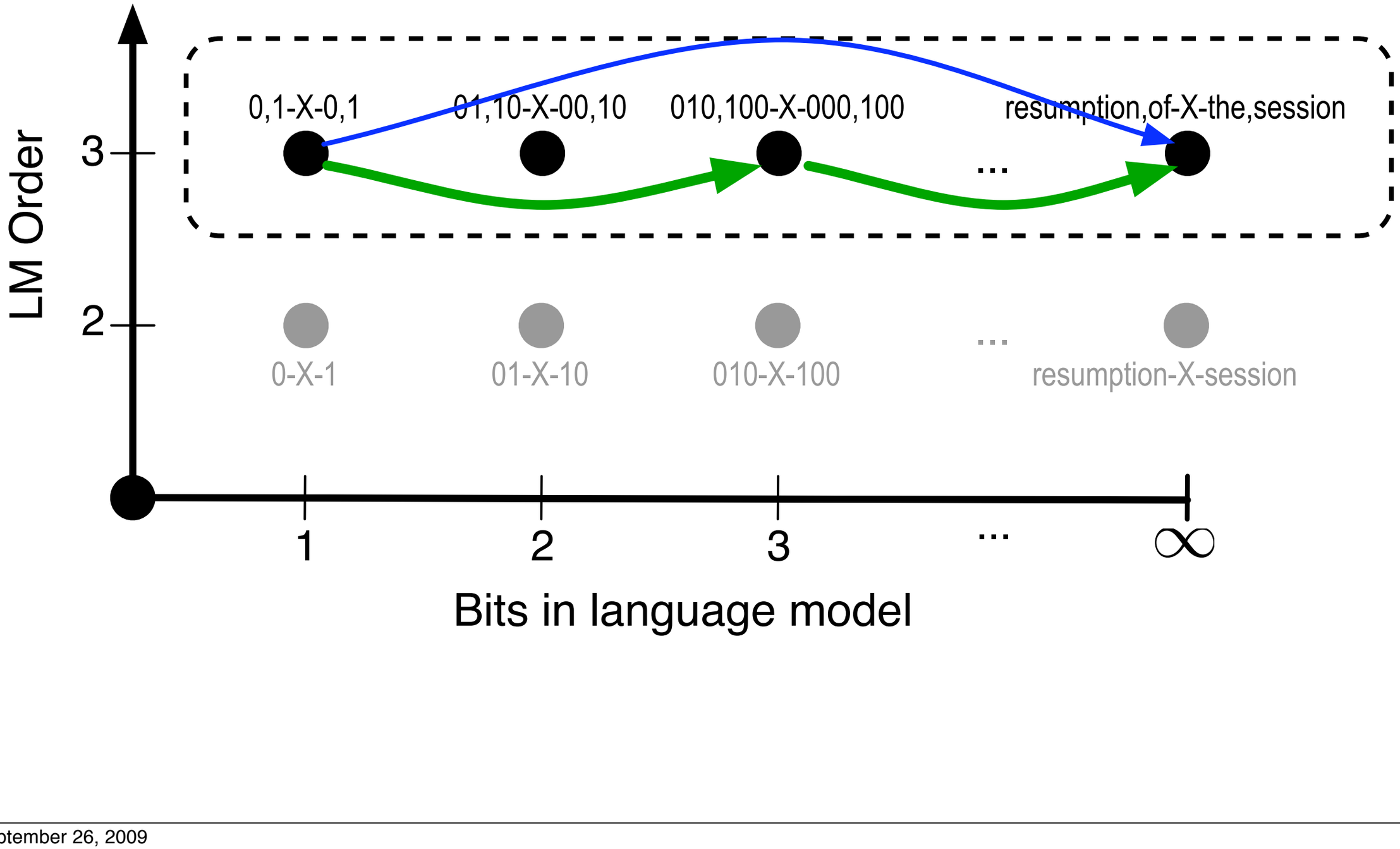
Where to start?



How many passes?

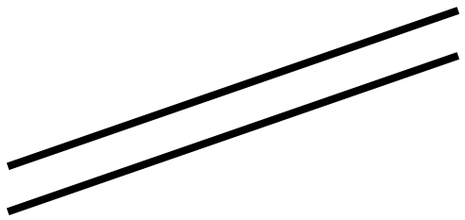


How many passes?



How many passes?

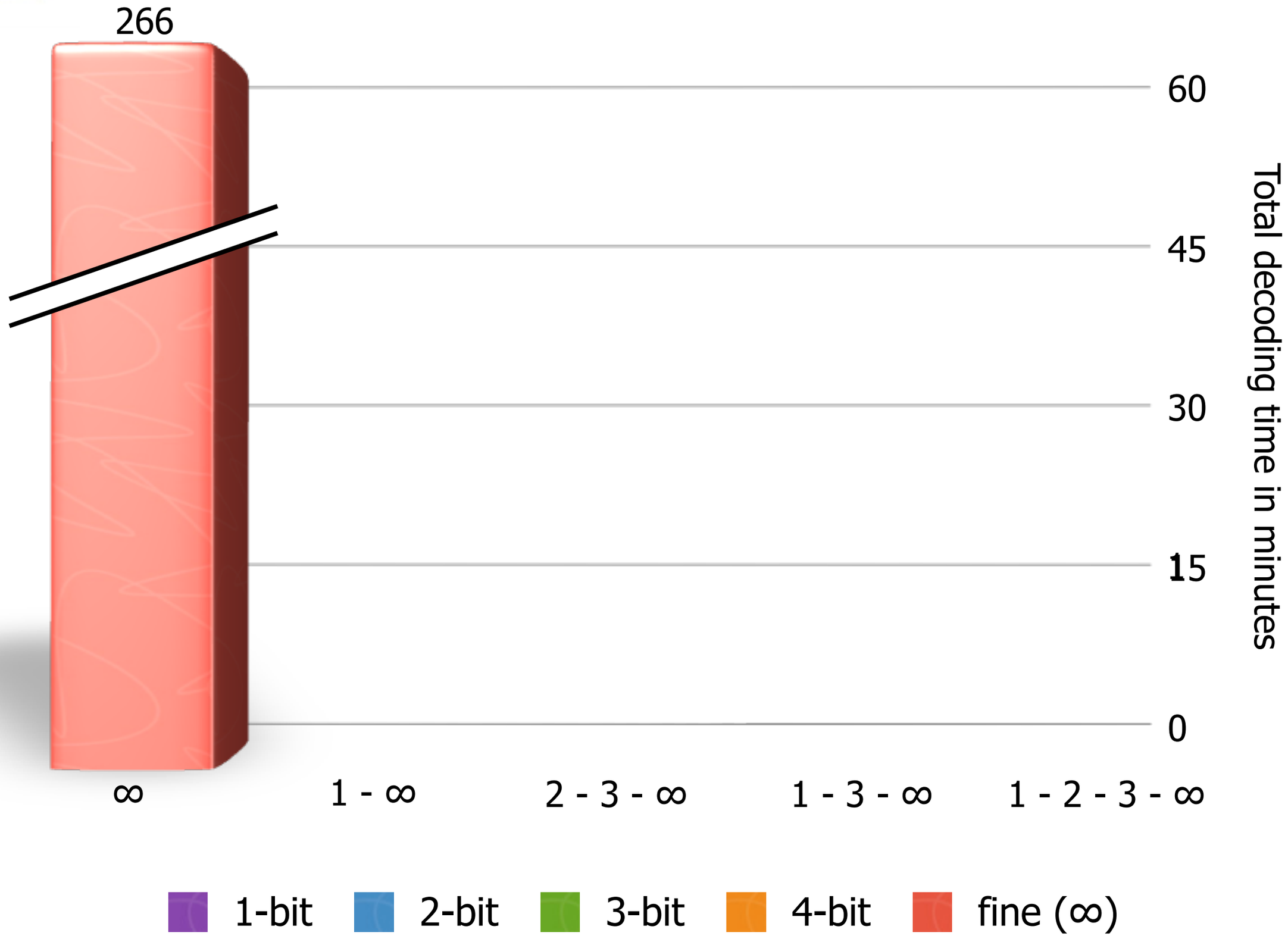
266



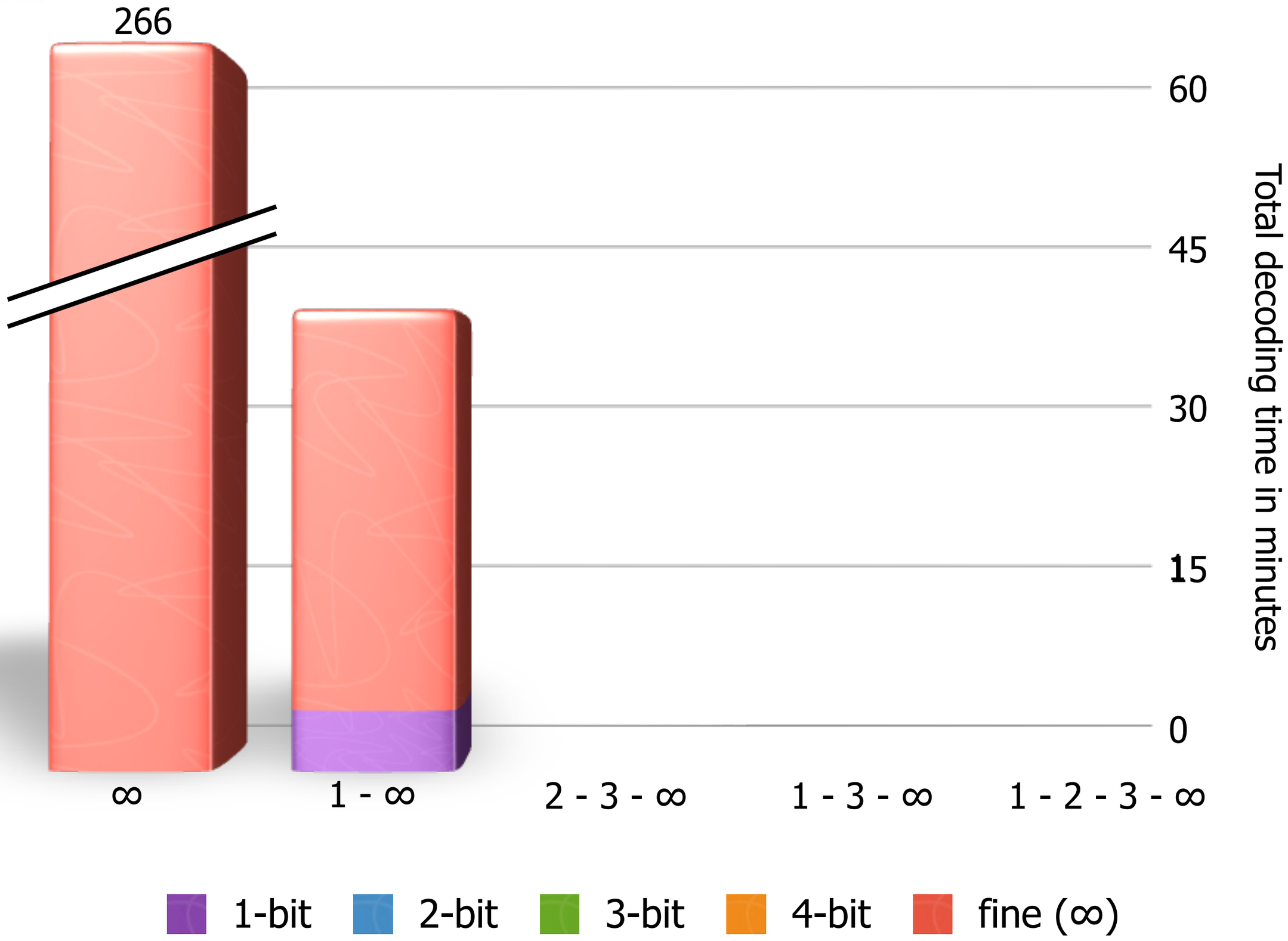
1

1-bit 2-bit 3-bit 4-bit fine (∞)

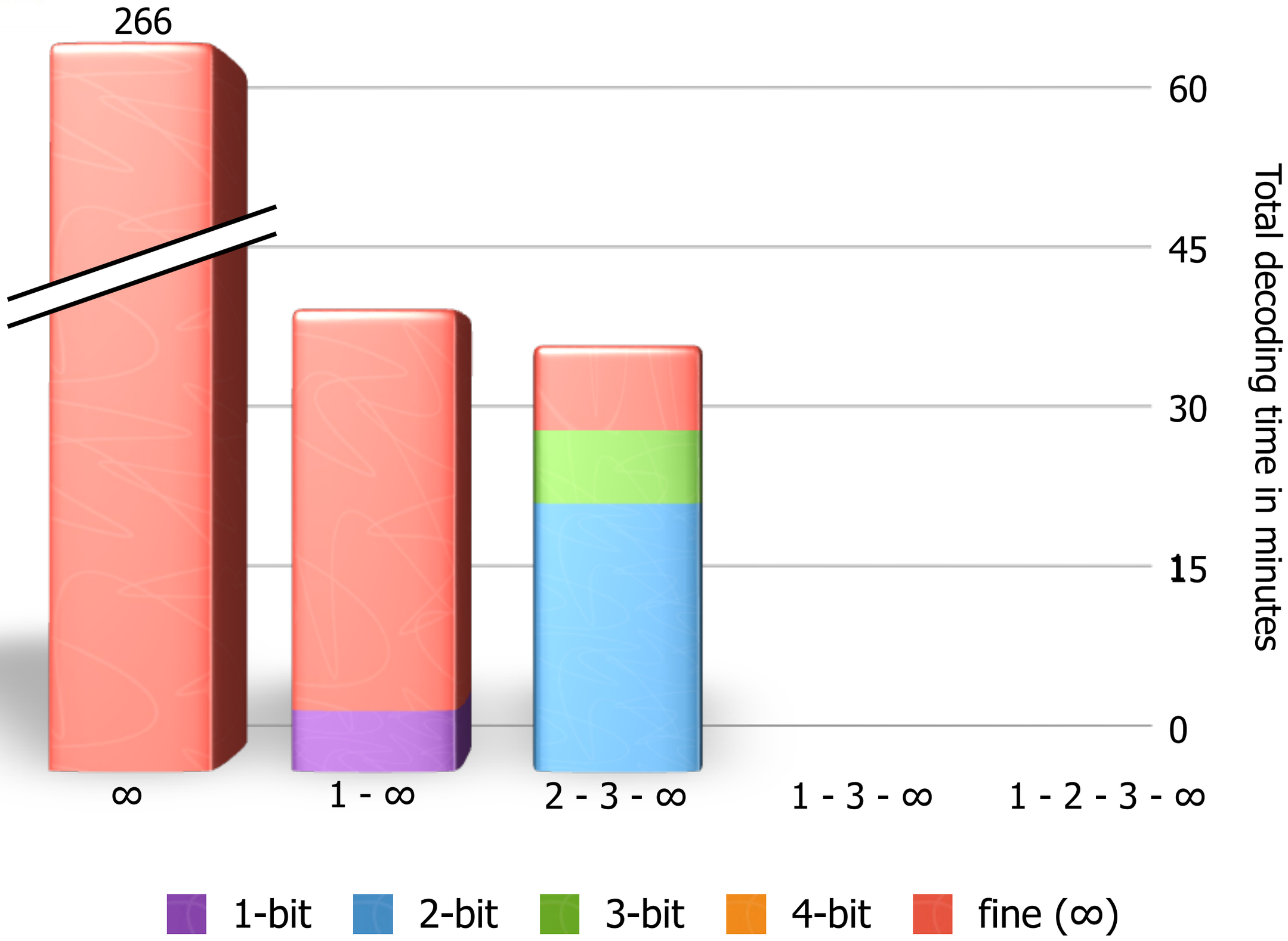
How many passes?



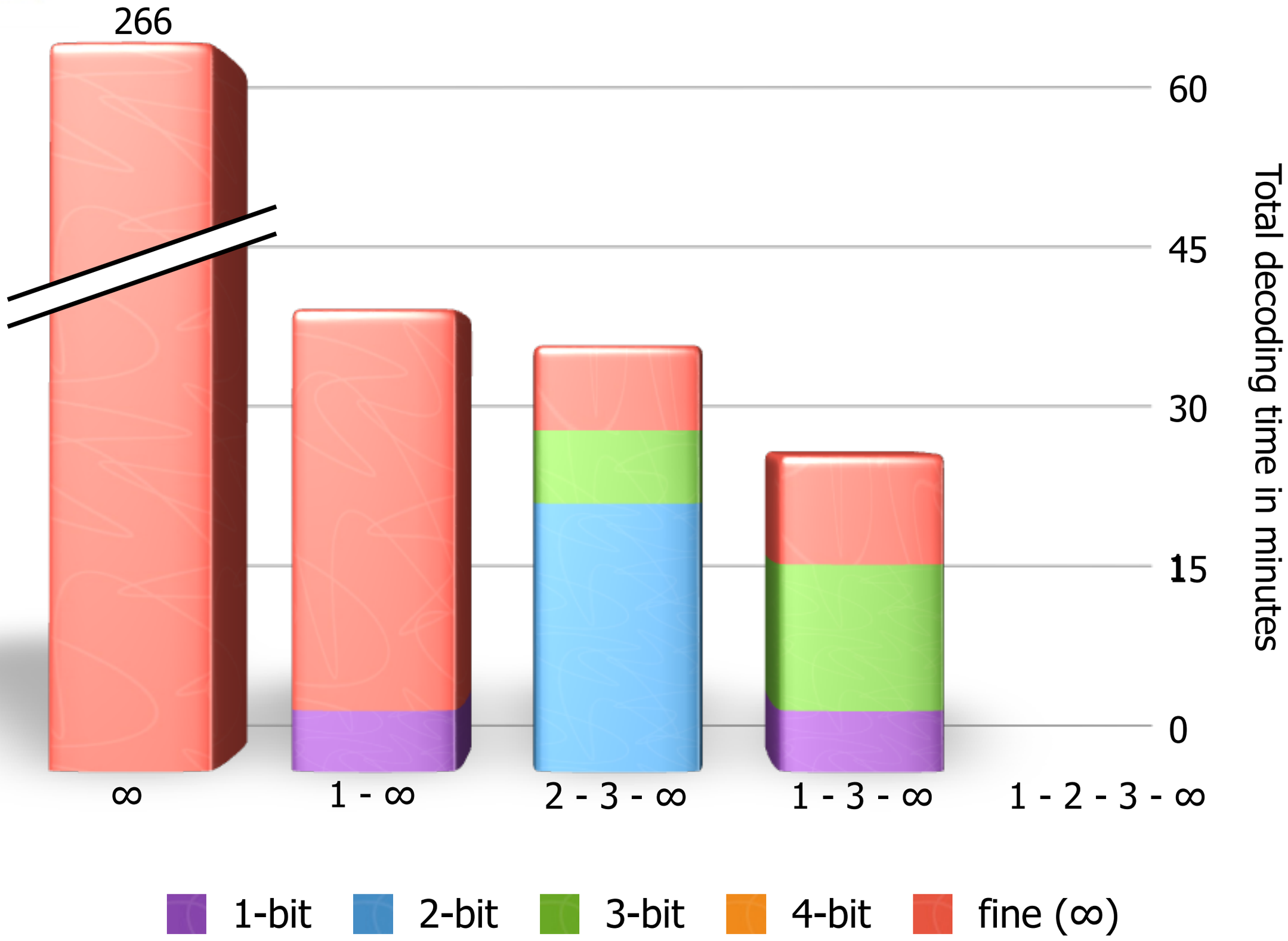
How many passes?



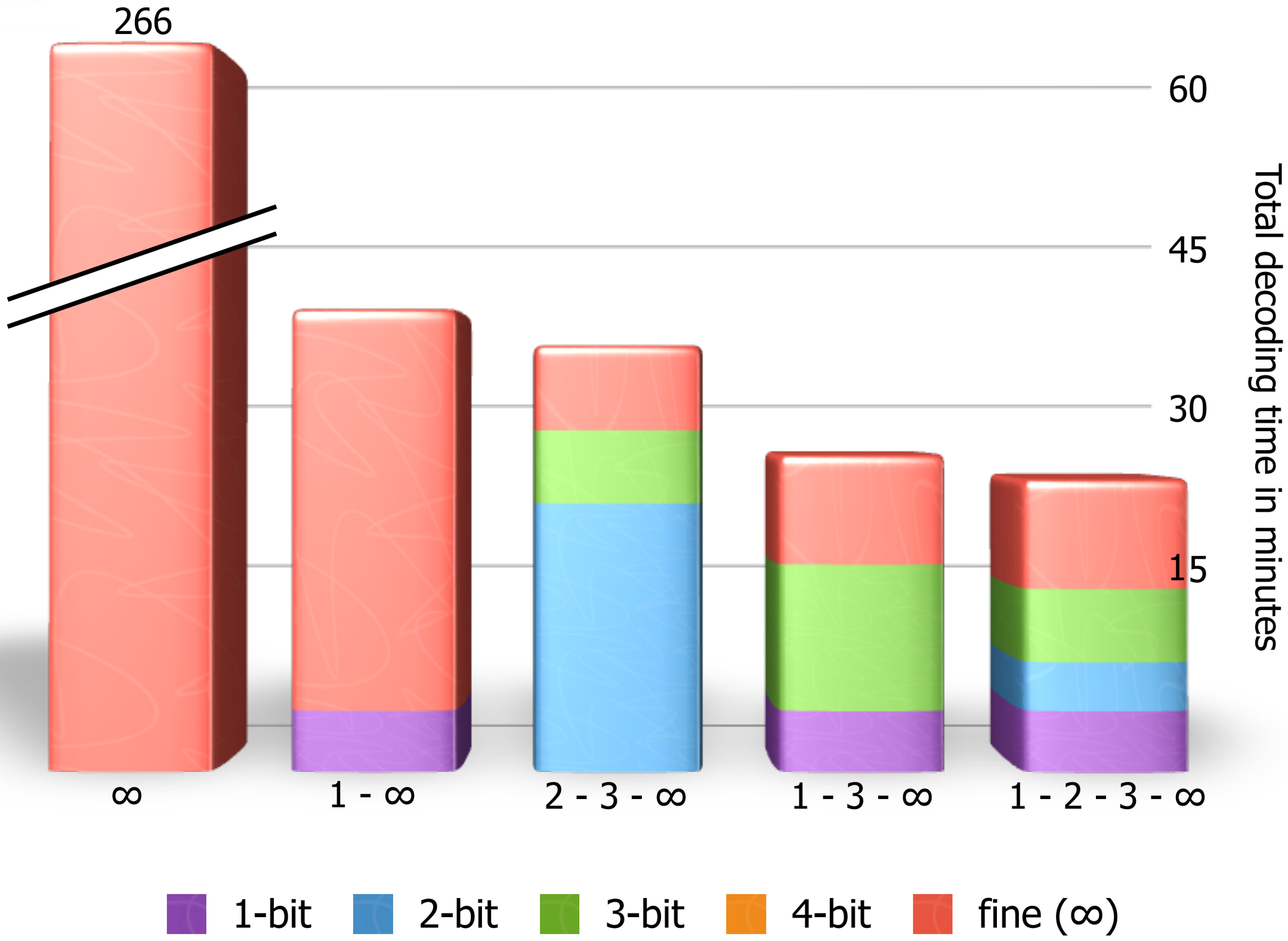
How many passes?



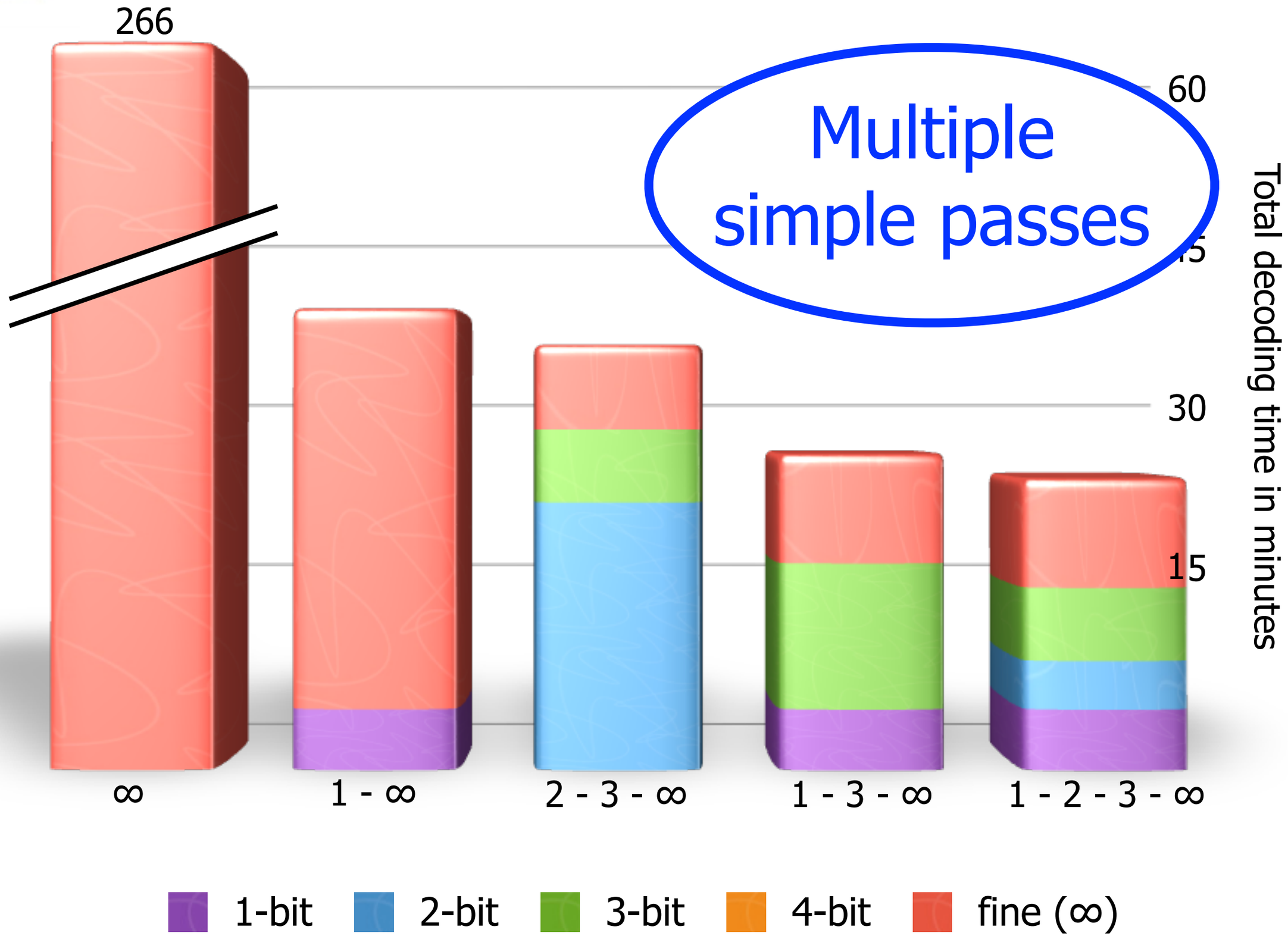
How many passes?



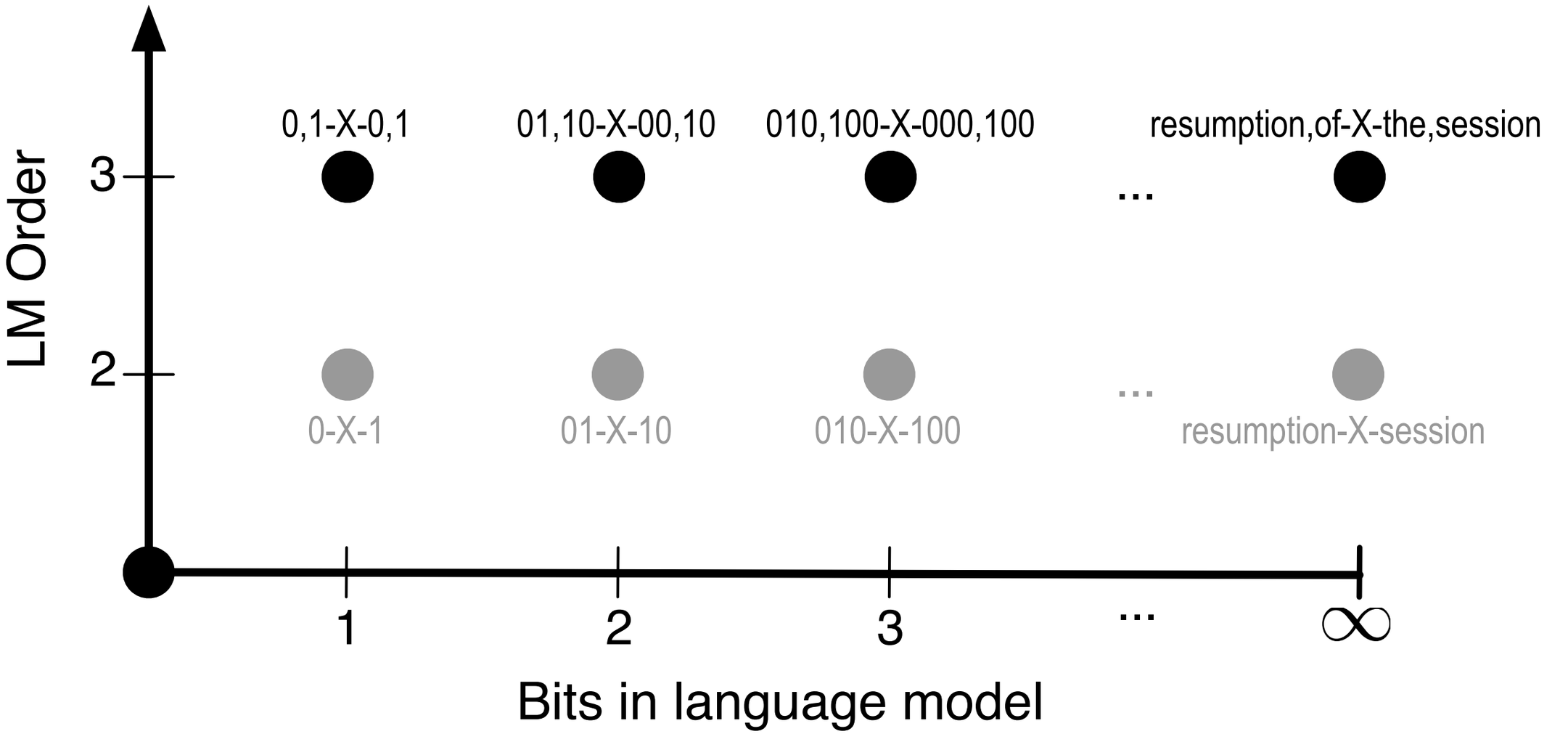
How many passes?



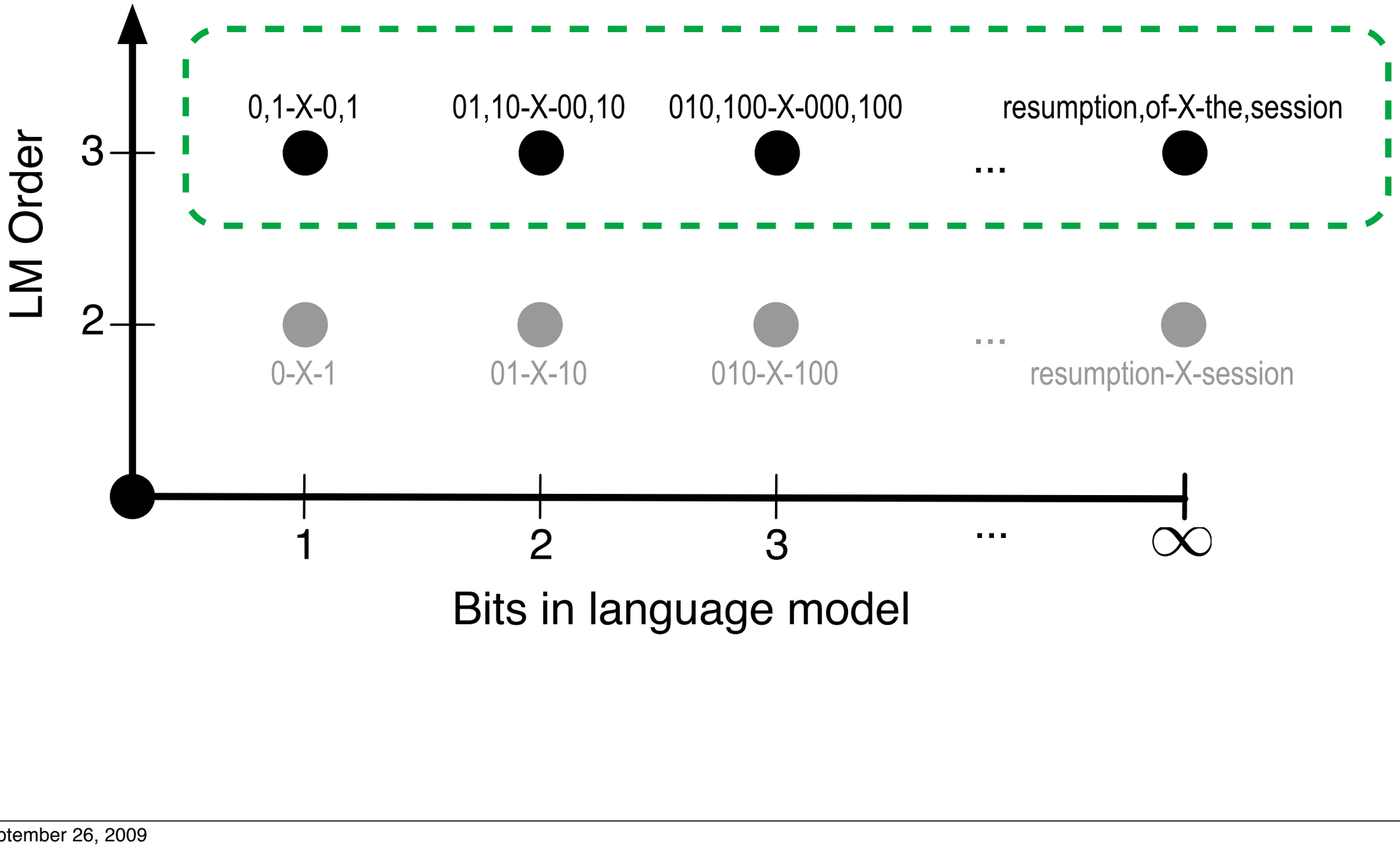
How many passes?



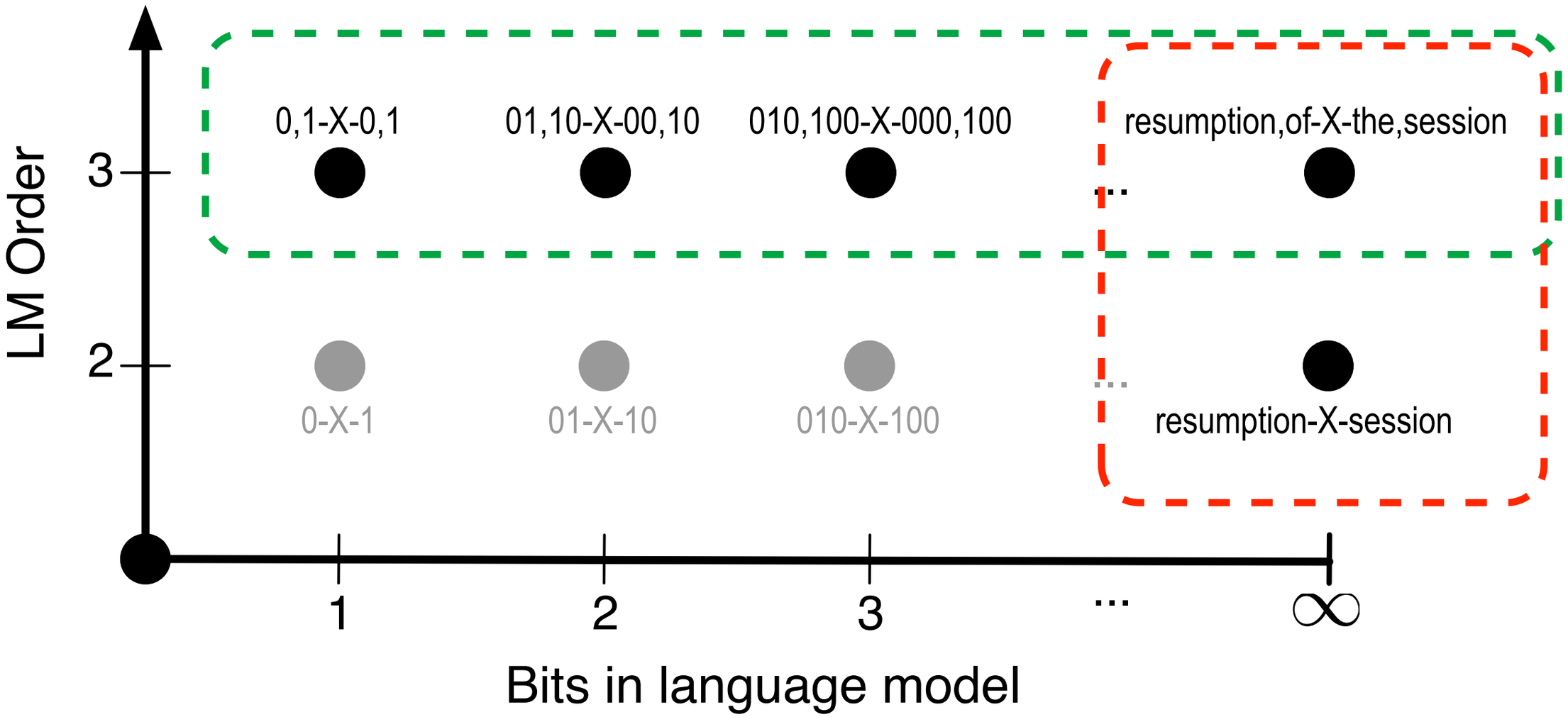
Use Encoding or Order?



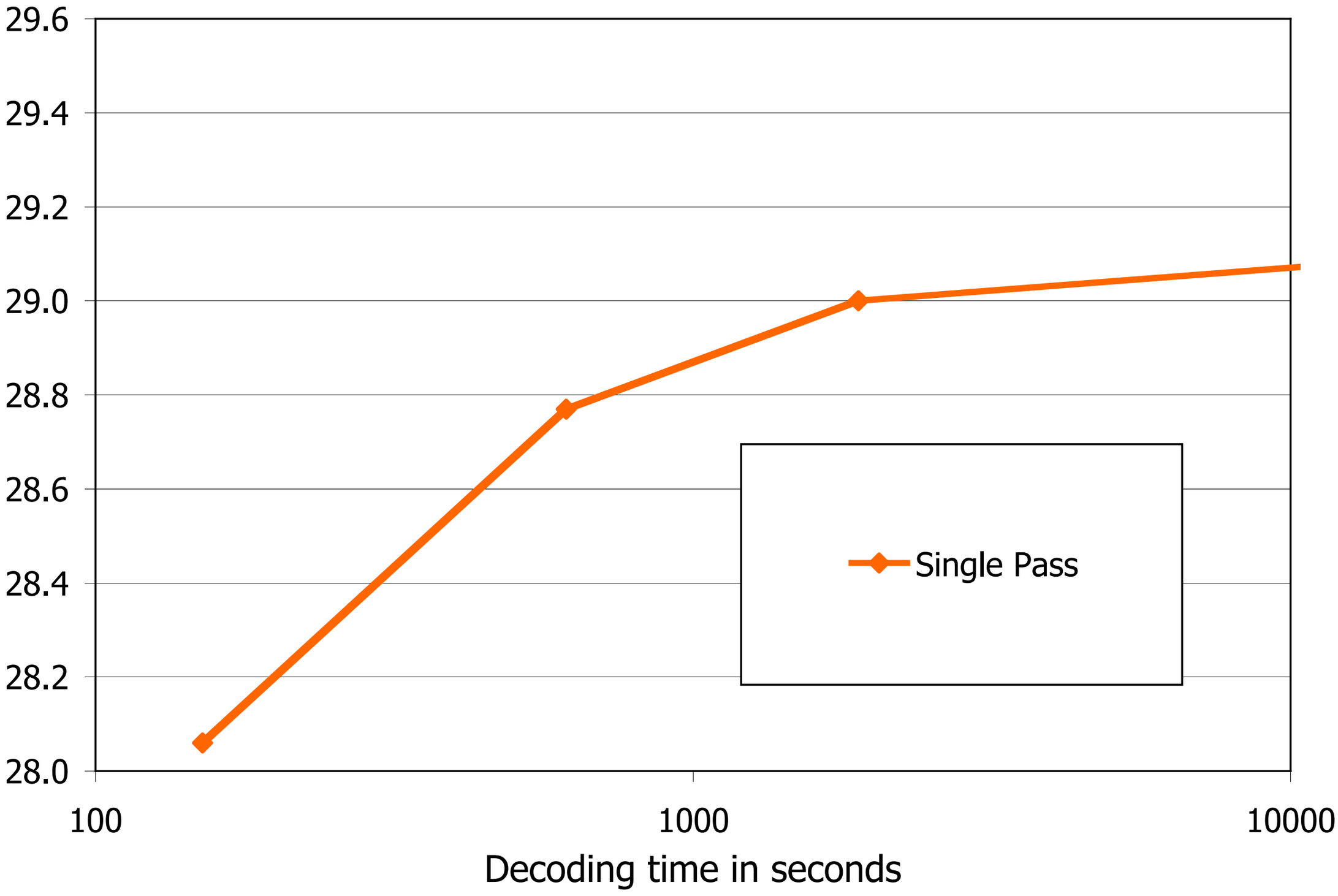
Use Encoding or Order?



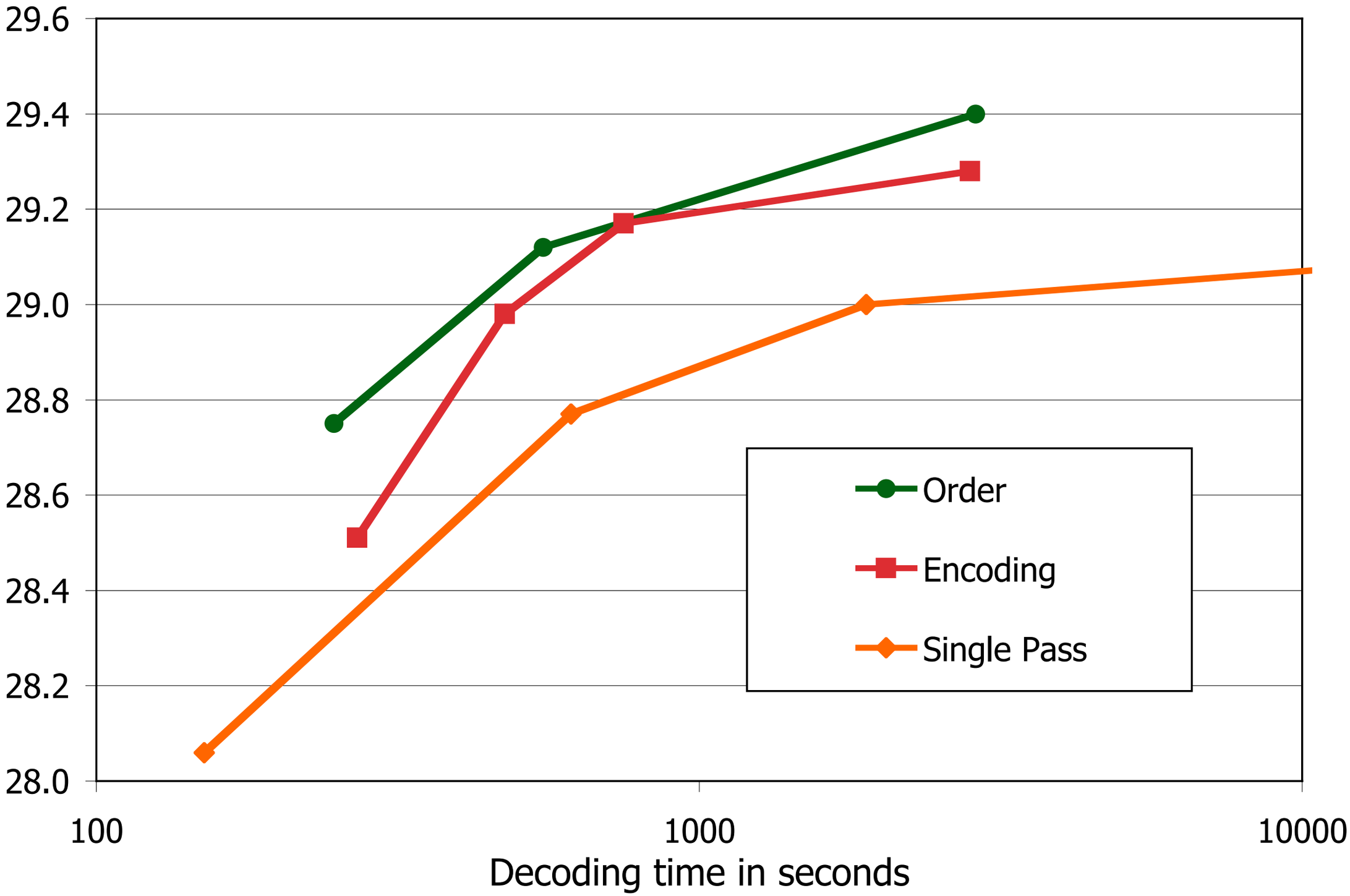
Use Encoding or Order?



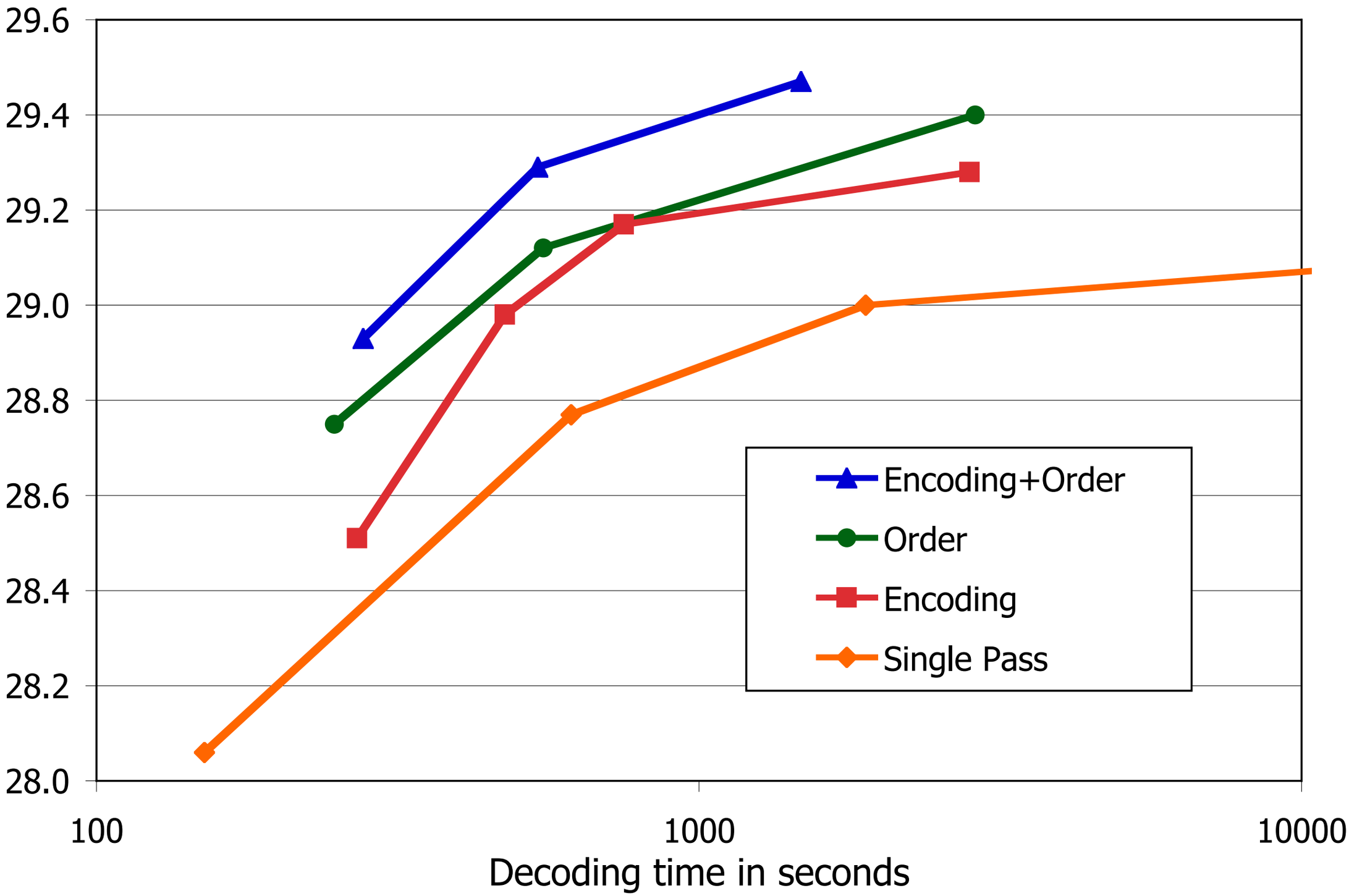
Use Encoding or Order?



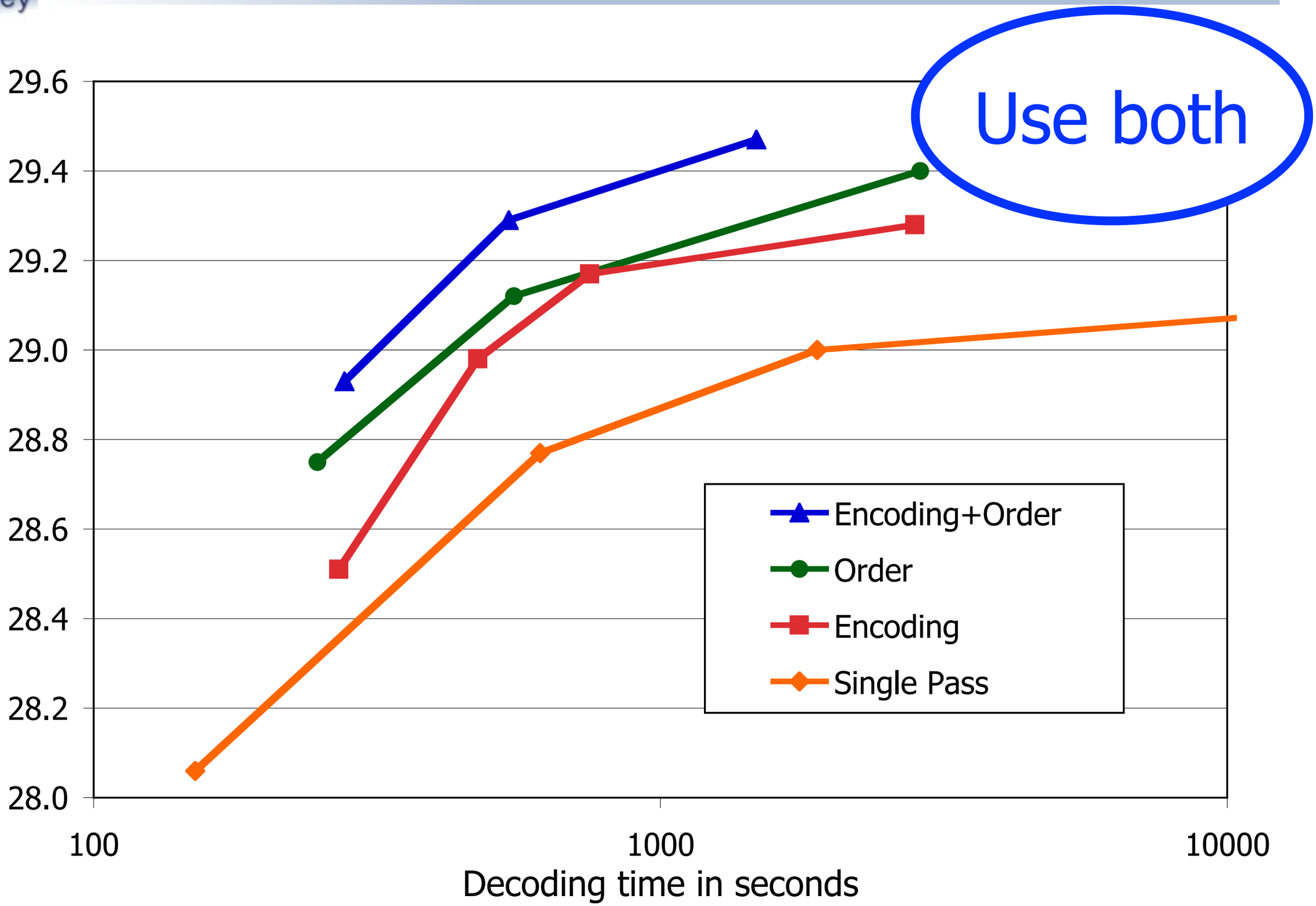
Use Encoding or Order?



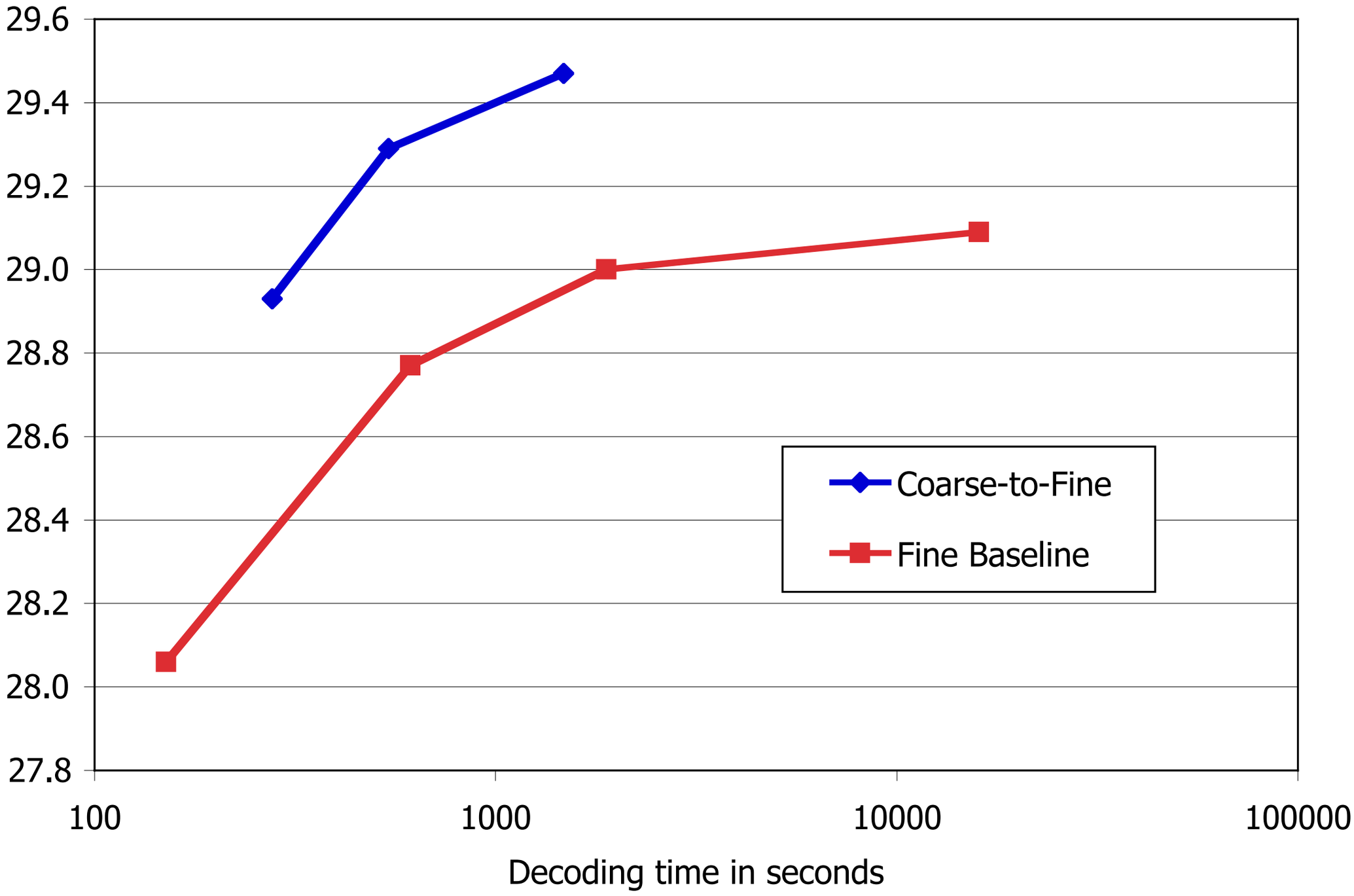
Use Encoding or Order?



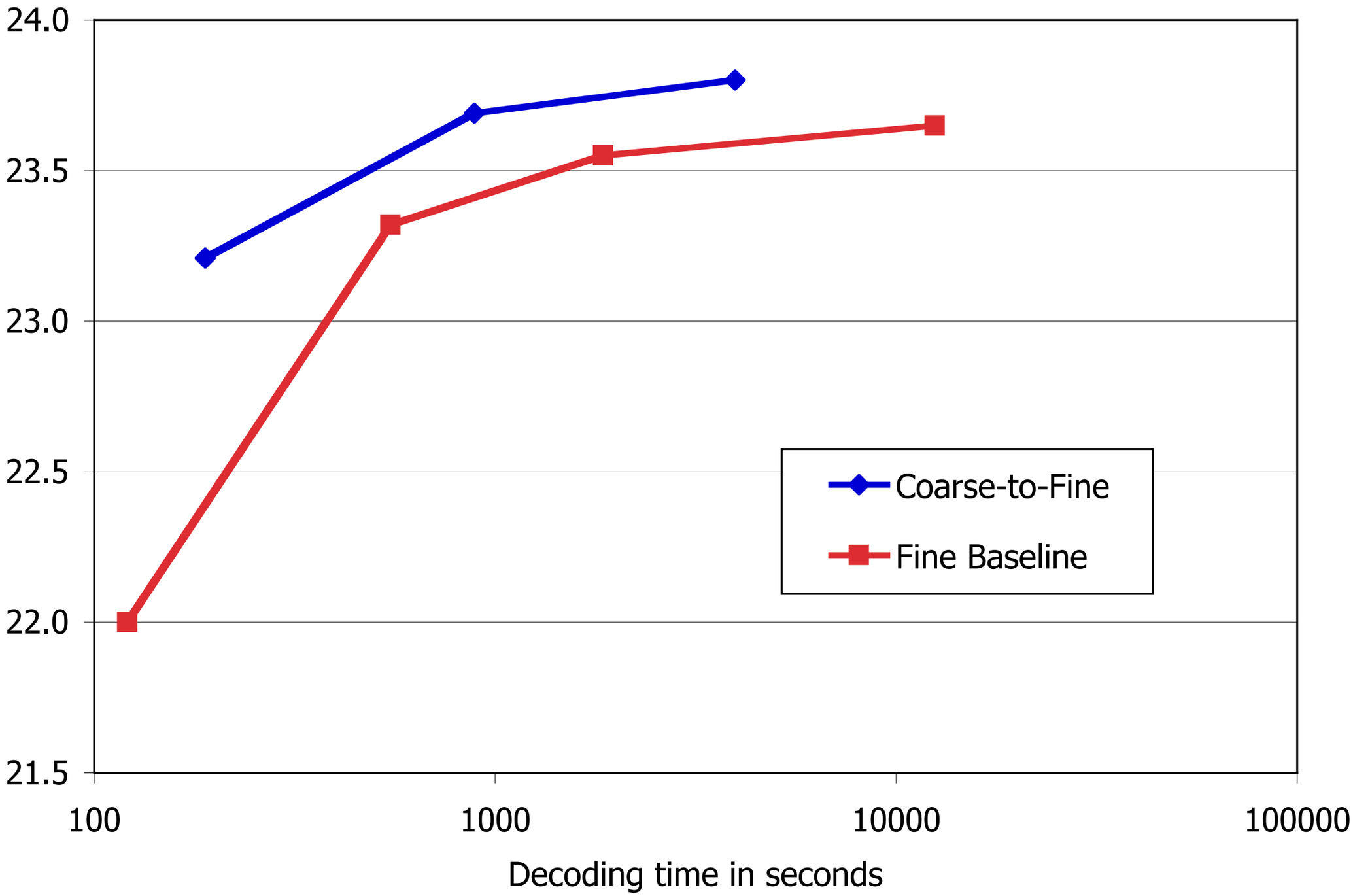
Use Encoding or Order?



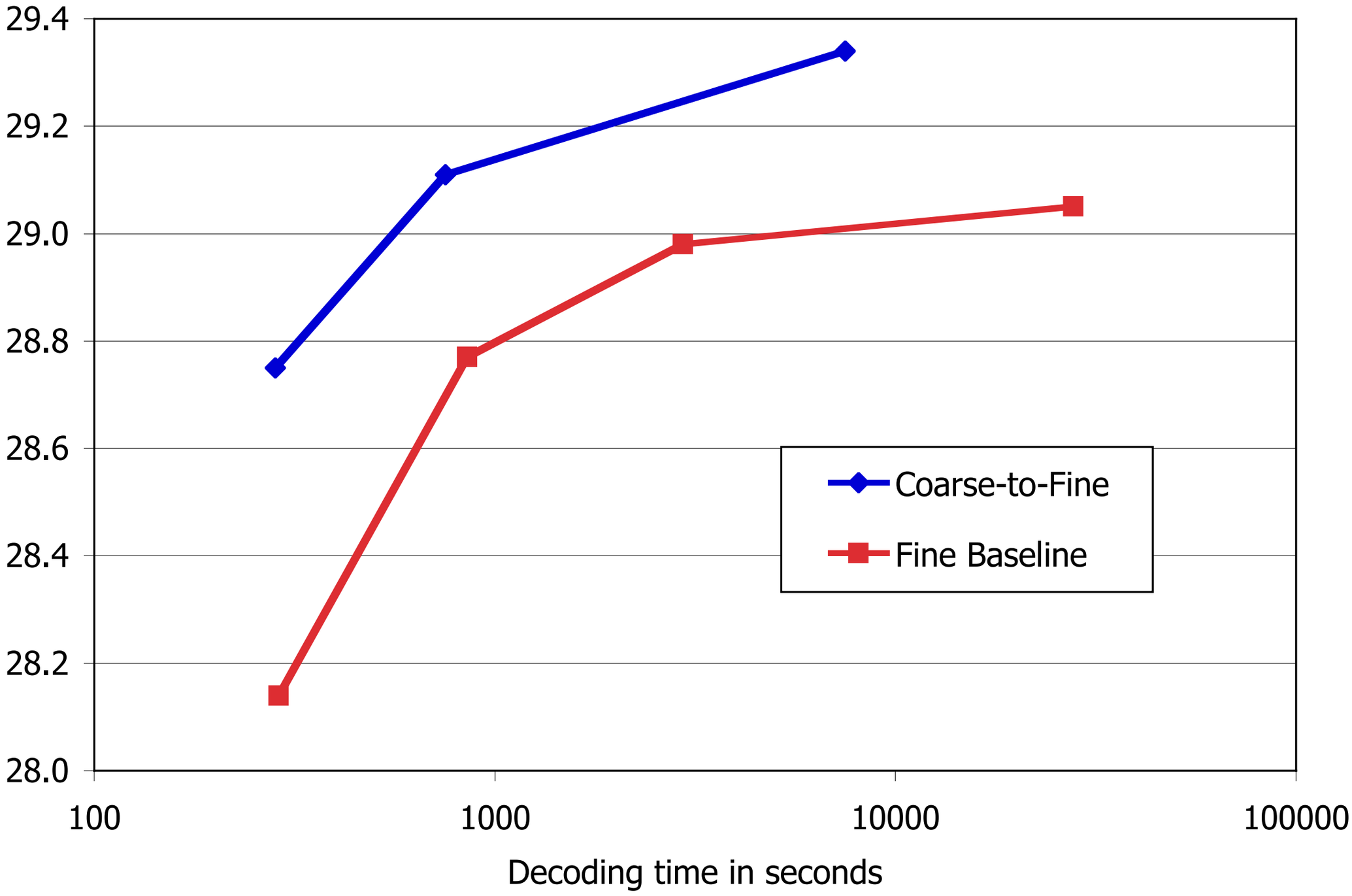
Final Results (Spanish)



Final Results (German)



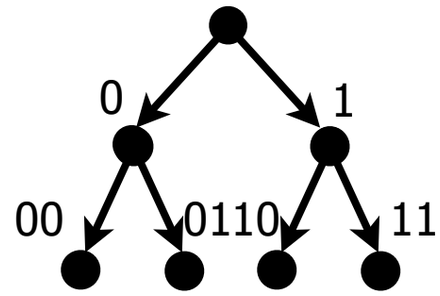
Final Results (French)



Conclusions

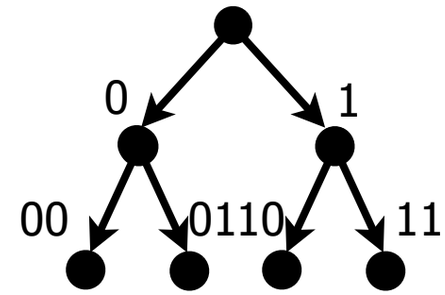
Conclusions

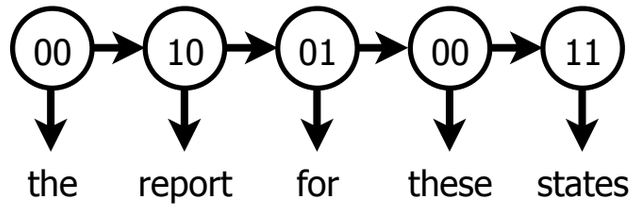
- Hierarchical clusters enable fast coarse-to-fine translation by reducing the vocabulary size



Conclusions

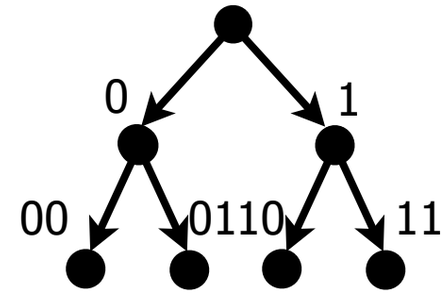
- Hierarchical clusters enable fast coarse-to-fine translation by reducing the vocabulary size
- Better clusters = better pruning

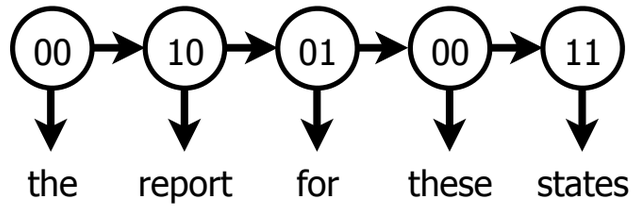




Conclusions

- Hierarchical clusters enable fast coarse-to-fine translation by reducing the vocabulary size

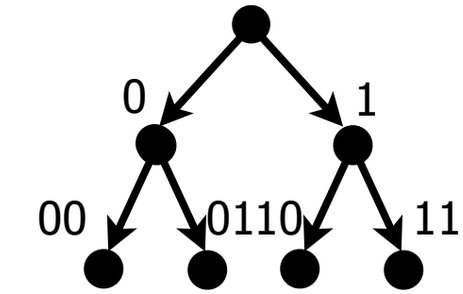


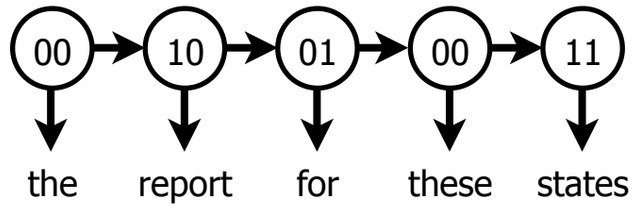


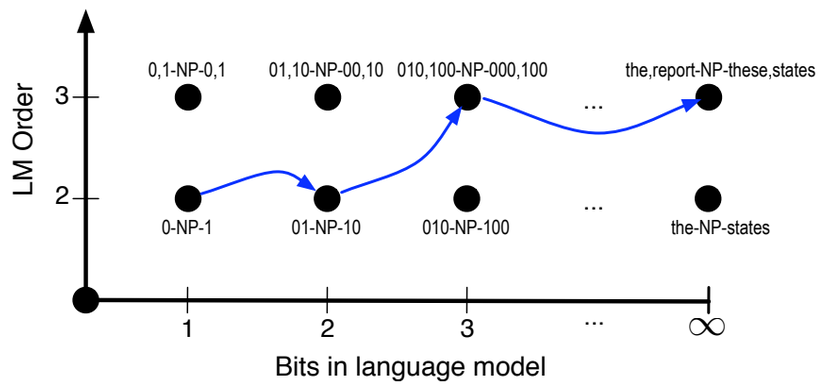
- Better clusters = better pruning
- Multiple, simple passes give the largest speed-ups
- Cluster size and n-gram order projections are complementary

Conclusions

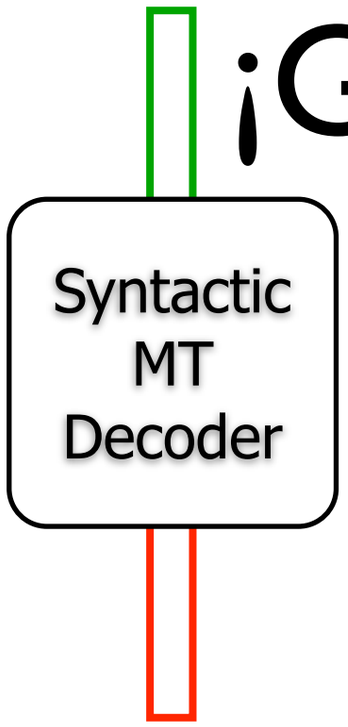
- Hierarchical clusters enable fast coarse-to-fine translation by reducing the vocabulary size
- Better clusters = better pruning
- Multiple, simple passes give the largest speed-ups
- Cluster size and n-gram order projections are complementary
- Minimal memory requirements







The End



¡Gracias!

http://nlp.cs.berkeley.edu

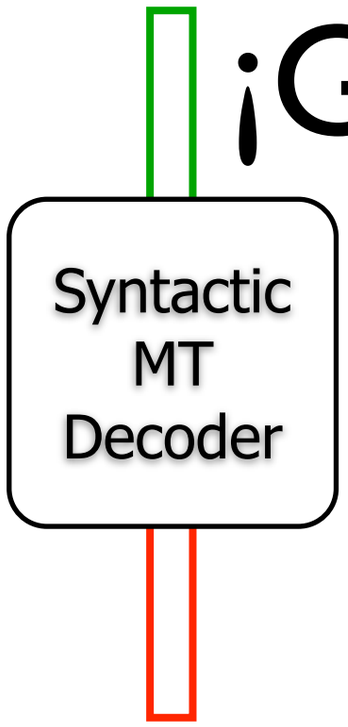
The End

¡Gracias!

Thanks!

http://nlp.cs.berkeley.edu

The End

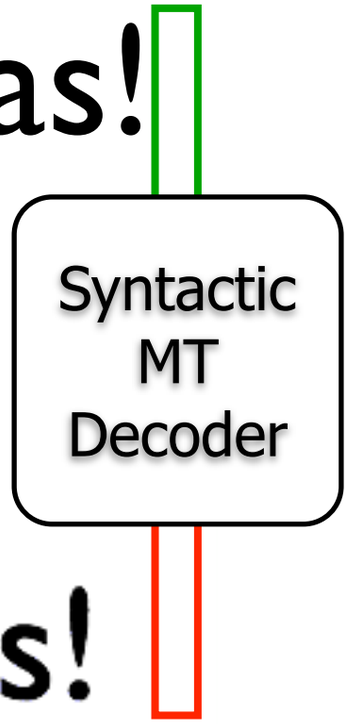


¡Gracias!

http://nlp.cs.berkeley.edu

The End

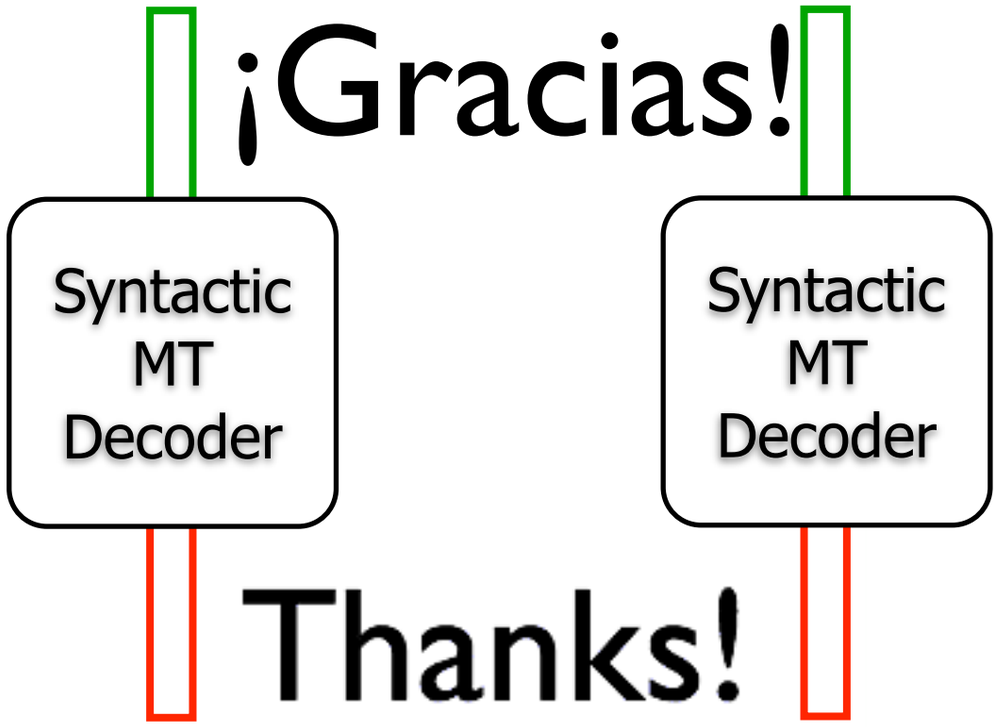
¡Gracias!



Thanks!

http://nlp.cs.berkeley.edu

The End



http://nlp.cs.berkeley.edu