

Opinosis:

A Graph Based Approach to Abstractive Summarization of Highly Redundant Opinions

Kavita Ganesan, ChengXiang Zhai, Jiawei Han

University of Illinois @ Urbana Champaign



ILLINOIS

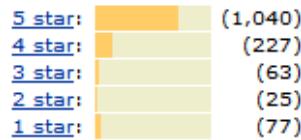
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Opinion Summarization Today...

Customer Reviews

Average Customer Rating

★★★★★ (1,432 customer reviews)



Appearance	★★★★★ (1,213)
Ease of use	★★★★★ (1,212)
Portability	★★★★★ (1,202)
Sound quality	★★★★★ (1,196)

[See and rate all 12 attributes](#)

Most Helpful Customer Reviews

3,677 of 3,770 people found the following review helpful

★★★★★ **WARNING for new 8GB 3G owners**

By [Hassan B. Bn Hadhram](#) - [See all my reviews](#)

REAL NAME

Amazon Verified Purchase ([What's this?](#))

This review is from: Apple iPod touch 8 GB (2nd Generation)

Before i start let me just tell you "what's New" with the iPod touch Third generation :

- Faster Cpu/Double the ram/Better graphic (faster Boot time/faster loading is all what i did notice)
- Double the storage for the same old price
- Voice control (I'll explain it in a second)
- Latest firmware for free
- New Earbuds with built in remote+Microphone (So you can use voice control)

And that is everything~ depends on your needs upgrading from 2G to 3G might be not worth it.

Important Note : only the New iPod touch 32GB/64GB are third generation ~ (8GB is repacked 2G) Details below.

Overall aspect based
ratings for entity

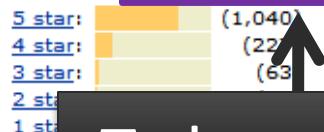
[Lu et al., 2009; Lerman et al., 2009;..]

Opinion Summarization Today...

Customer Reviews

Average Customer Rating

★★★★★ (1,432 customer reviews)



Appearance	★★★★★ (1,213)
Ease of use	★★★★★ (1,212)
Portability	★★★★★ (1,202)
Sound quality	★★★★★ (1,196)

12 attributes.

To know more: read
many sentences

3.6

★★★★★ WARNING for new 8GB 3G owners and ipod touch 3G Review, September 11, 2009

By Hassan B. Bn Hadhram - [See all my reviews](#)

REAL NAME

Amazon Verified Purchase ([What's this?](#))

This review is from: Apple iPod touch 8 GB (2nd Generation--with iPhone OS 3.1 Software Installed) [NEWEST MODEL] (Electronics)

Before i start let me just tell you "what's New" with the iPod touch Third generation :

- Faster Cpu/Double the ram/Better graphic (faster Boot time/faster loading is all what i did notice)
- Double the storage for the same old price

Structured format → useful,
but not enough!

**We need....
supporting textual summary!**

Properties of unstructured opinion summaries

► Summarize the major opinions

- What are the major complaints/praise in an aspect?



► Concise

- Easily digestible
- Viewable on smaller screens



► Readable

- Easily understood

**How to generate such
summaries?**

Extractive Summarization

- ▶ Widely studied for years

[Radev et al.2000; Erkan & Radev, 2004; Mihalcea & Tarau, 2004...]

- ▶ But, not suitable for:

1. generating concise summaries
2. summarizing highly redundant text

Extractive Summarization

Problem 1: Bias with limit on summary size

Ex: Opinion sentences about iPhone's Battery

1. *The iPhone's battery lasts long...have to charge it once every few days.*



2. *iPhone's battery is bulky but it is cheap.*

3. *iPhone's battery is bulky but it lasts long!*

Extractive Summarization

Problem 1: Bias with limit on summary size

Ex: Opinion sentences about iPhone's Battery

1. *The iPhone's battery lasts long...have to charge it once every few days.*

Tendency of missing out information



Extractive summarization

Problem 2: Verbose

- May contain irrelevant information
- Not suited for smaller devices

An Ideal Summary

*The iPhone's battery lasts long
and is cheap but its bulky.*



- ▶ Important information summarized
- ▶ Concise
- ▶ Readable

An Ideal Summary

*“The iPhone’s battery lasts long
and is cheap but its bulky.*

Extractive 
Abstractive 

Abstractive Summarization

► **HARD!!**

► **Some methods require manual effort**

[DeJong1982] [Radev and McKeown1998] [Finley and Harabagiu2002]

- Need to define templates to be filled

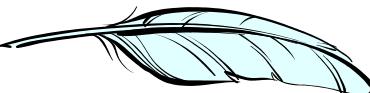
► **Some methods rely on deep NL understanding**

[Saggion and Lapalme2002] [Jing and McKeown2000]

- Domain dependent
- Impractical – high computational costs

Our Method: Opinosis

- ▶ ‘Shallow’ abstractive summarizer
- ▶ Generates **concise summaries** using:
 - existing text
 - inherent redundancies
- ▶ Uses **minimal** external knowledge
 - lightweight



Opinosis: High Level Overview

Opinosis: High Level Overview



Input

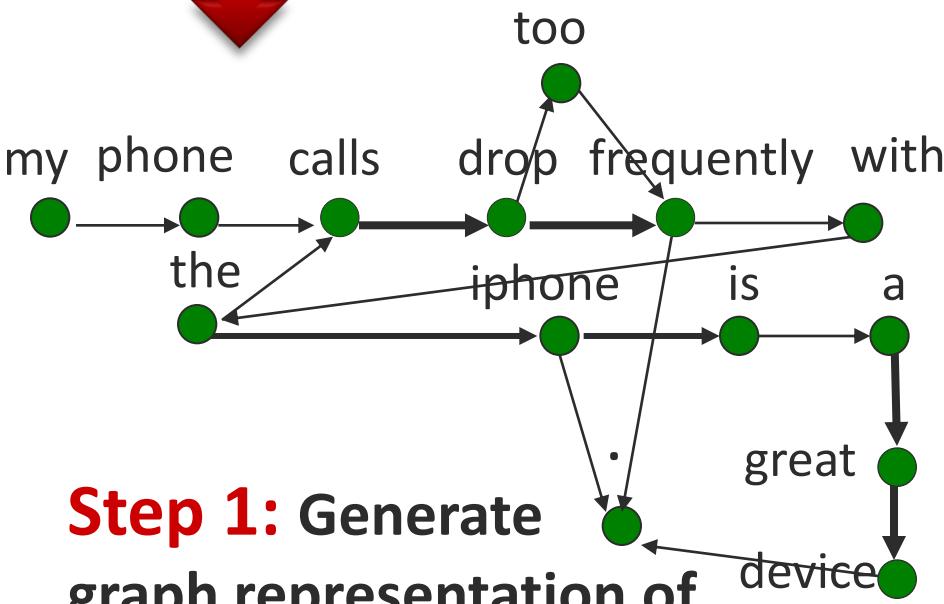
Set of sentences:

- **Topic** specific (ex. battery life of ipod)
- **POS** annotated

Opinosis: High Level Overview



Input

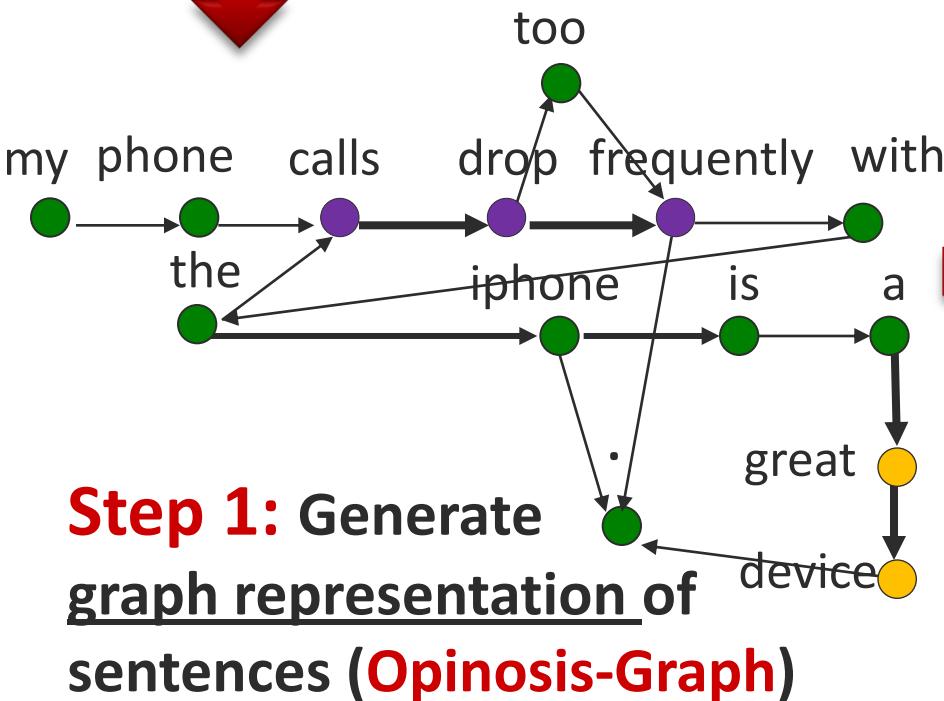


**Step 1: Generate
graph representation of
sentences (**Opinosis-Graph**)**

Opinosis: High Level Overview

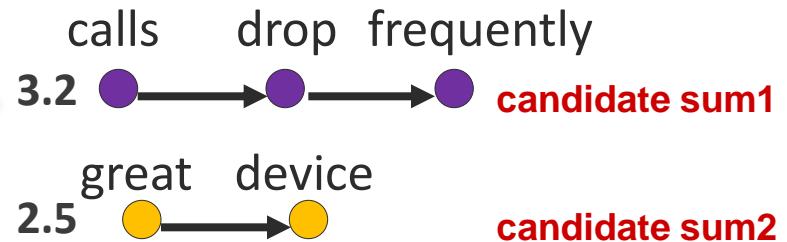


Input



Set of sentences:

- Topic specific (ex. battery life of ipod)
- POS annotated

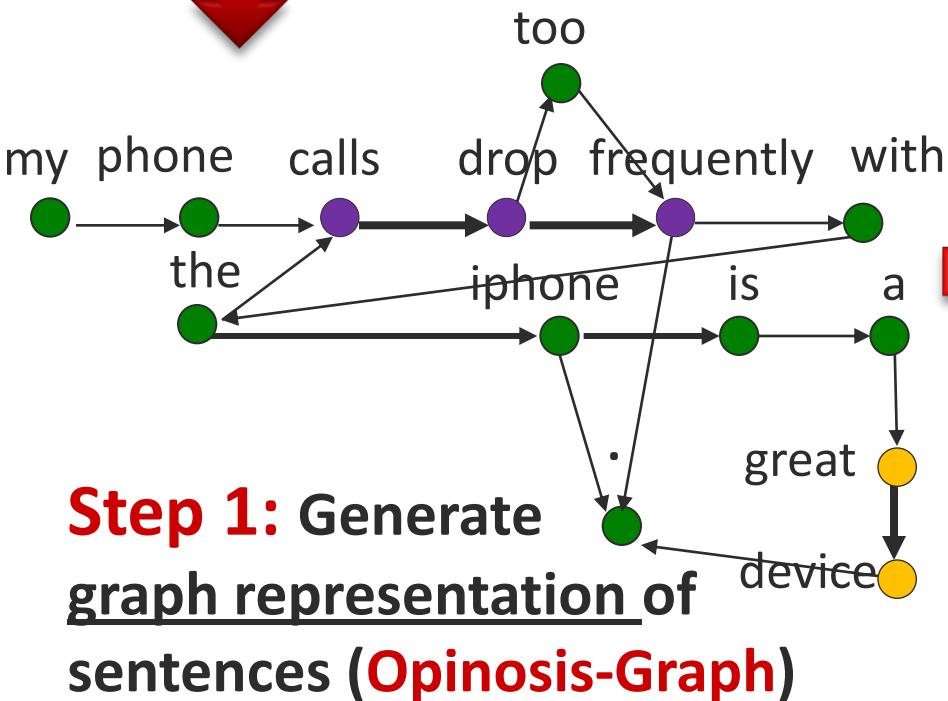


Step 2: Find promising paths (candidate summaries) & score these candidates

Opinosis: High Level Overview



Input

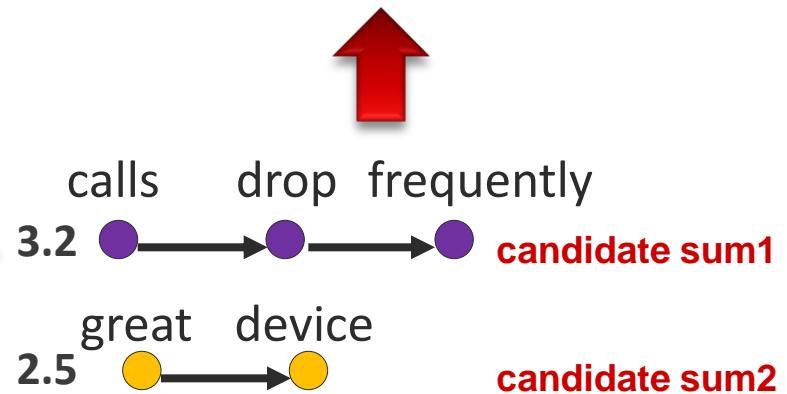


Set of sentences:

- Topic specific (ex. battery)
- POS annotated

The iPhone is a great device, but calls drop frequently.

Step 3: Select top scoring candidates as final summary



Step 2: Find promising paths (candidate summaries) & score these candidates

Step 1: Building the Opinosis-Graph

Building Opinosis-Graph

Assume:

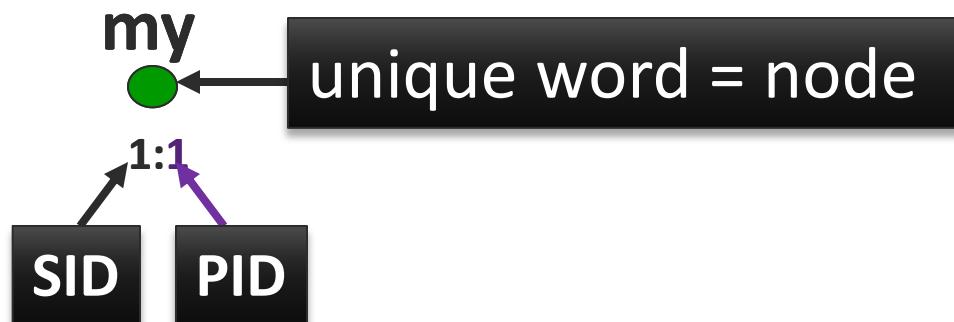
- ▶ 2 sentences about “*call quality of iphone*”
 - 1. *My phone calls drop frequently with the iPhone.*
 - 2. *Great device, but the calls drop too frequently.*
- ▶ Opinosis-Graph is empty

Building Opinosis-Graph

- 1. My phone calls drop frequently with the iPhone.*

Building Opinosis-Graph

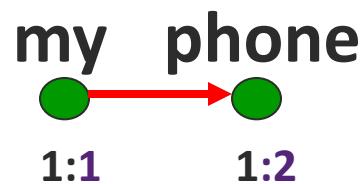
1. My phone calls drop frequently with the iPhone.



Positional Reference
Information

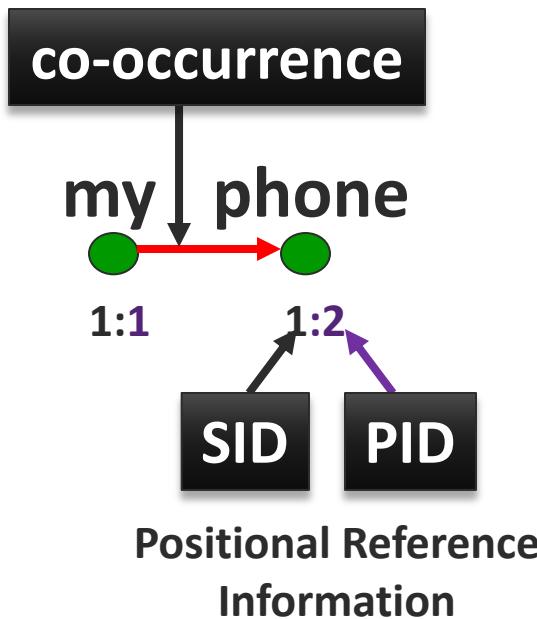
Building Opinosis-Graph

1. *My phone calls drop frequently with the iPhone.*



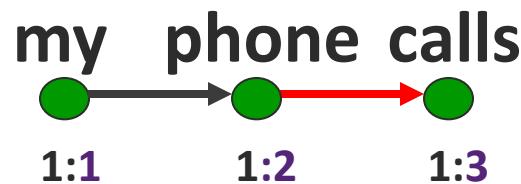
Building Opinosis-Graph

1. *My phone calls drop frequently with the iPhone.*



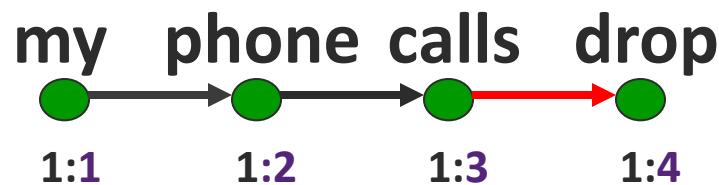
Building Opinosis-Graph

1. *My phone calls drop frequently with the iPhone.*



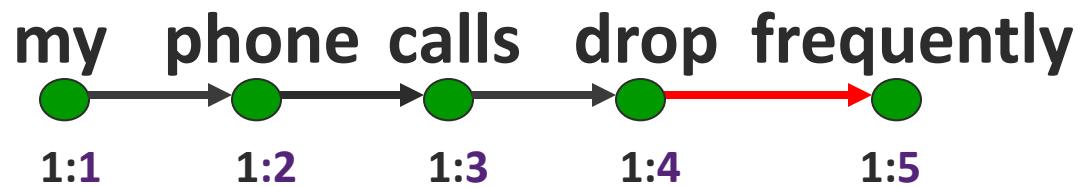
Building Opinosis-Graph

1. *My phone calls drop frequently with the iPhone.*



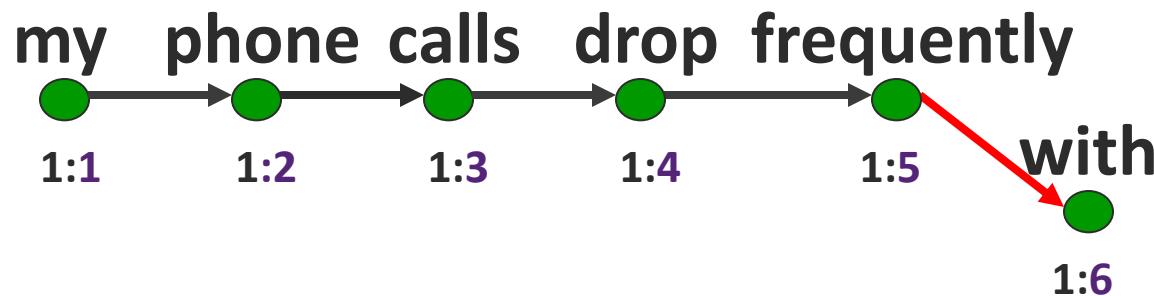
Building Opinosis-Graph

1. *My phone calls drop frequently with the iPhone.*



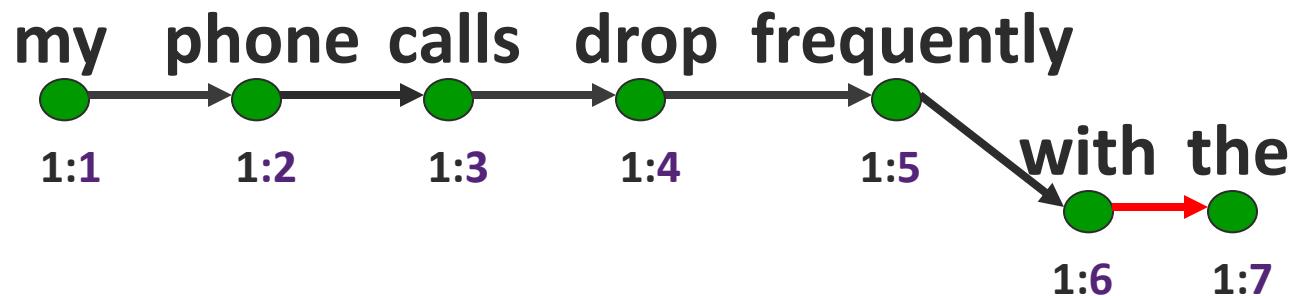
Building Opinosis-Graph

1. *My phone calls drop frequently with the iPhone.*



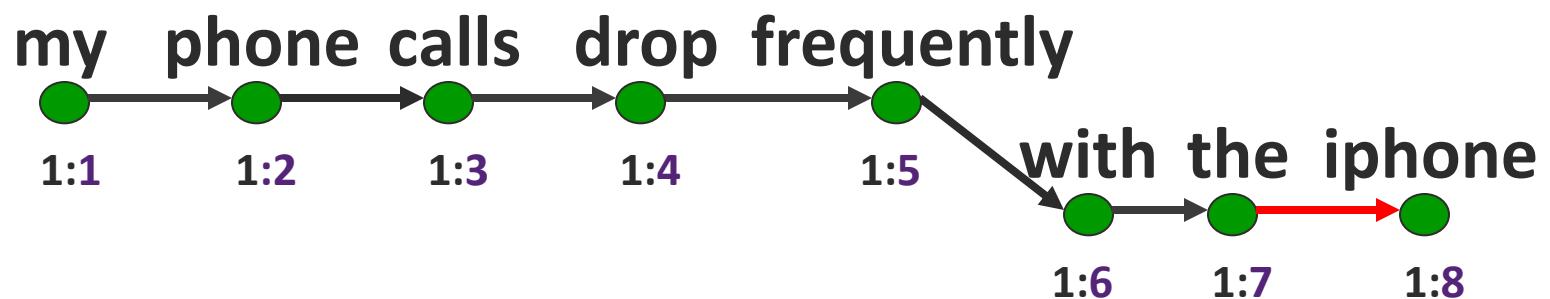
Building Opinosis-Graph

1. *My phone calls drop frequently with the iPhone.*



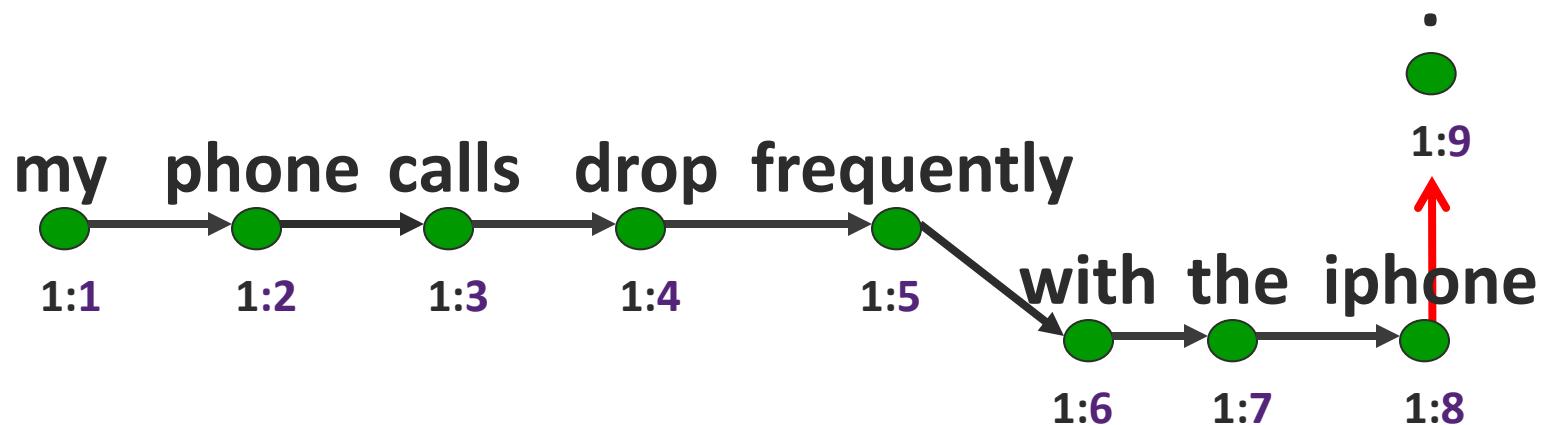
Building Opinosis-Graph

1. *My phone calls drop frequently with the iPhone.*



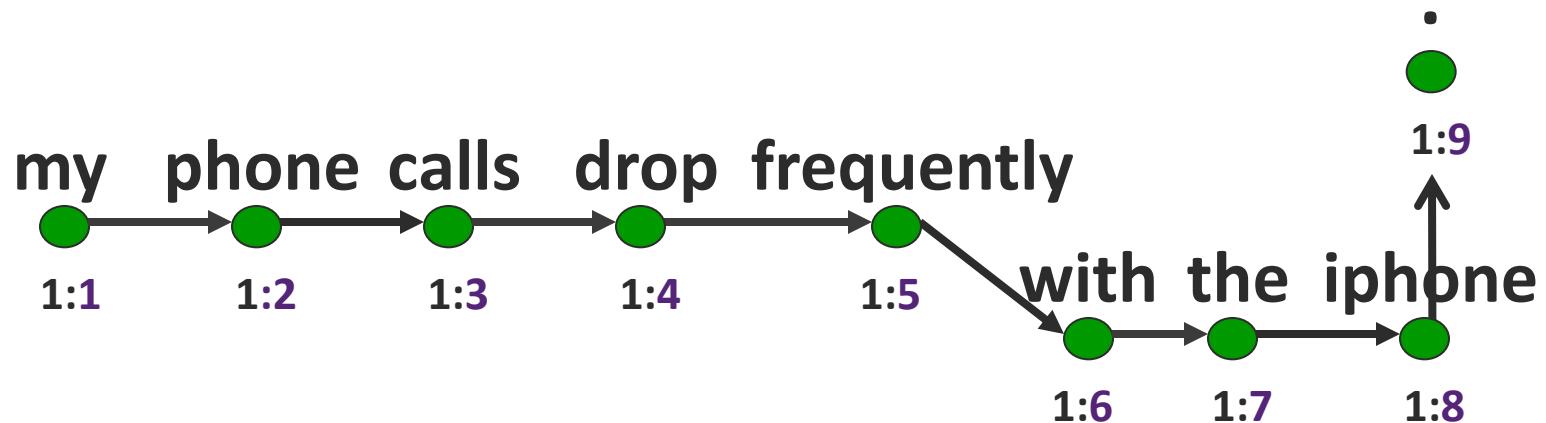
Building Opinosis-Graph

1. *My phone calls drop frequently with the iPhone.*



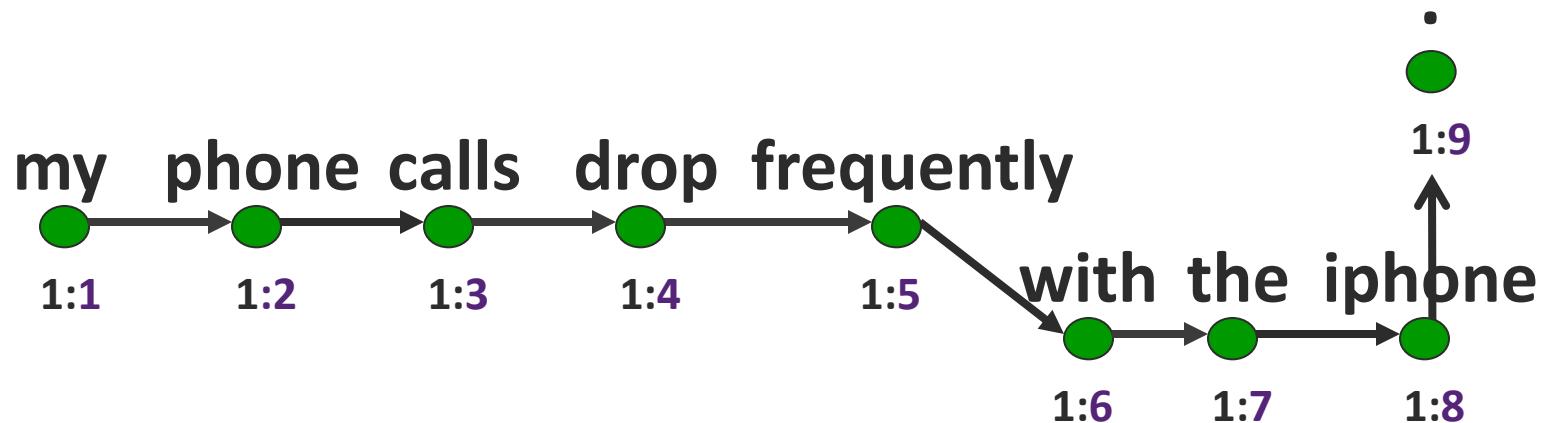
Building Opinosis-Graph

1. *My phone calls drop frequently with the iPhone.*



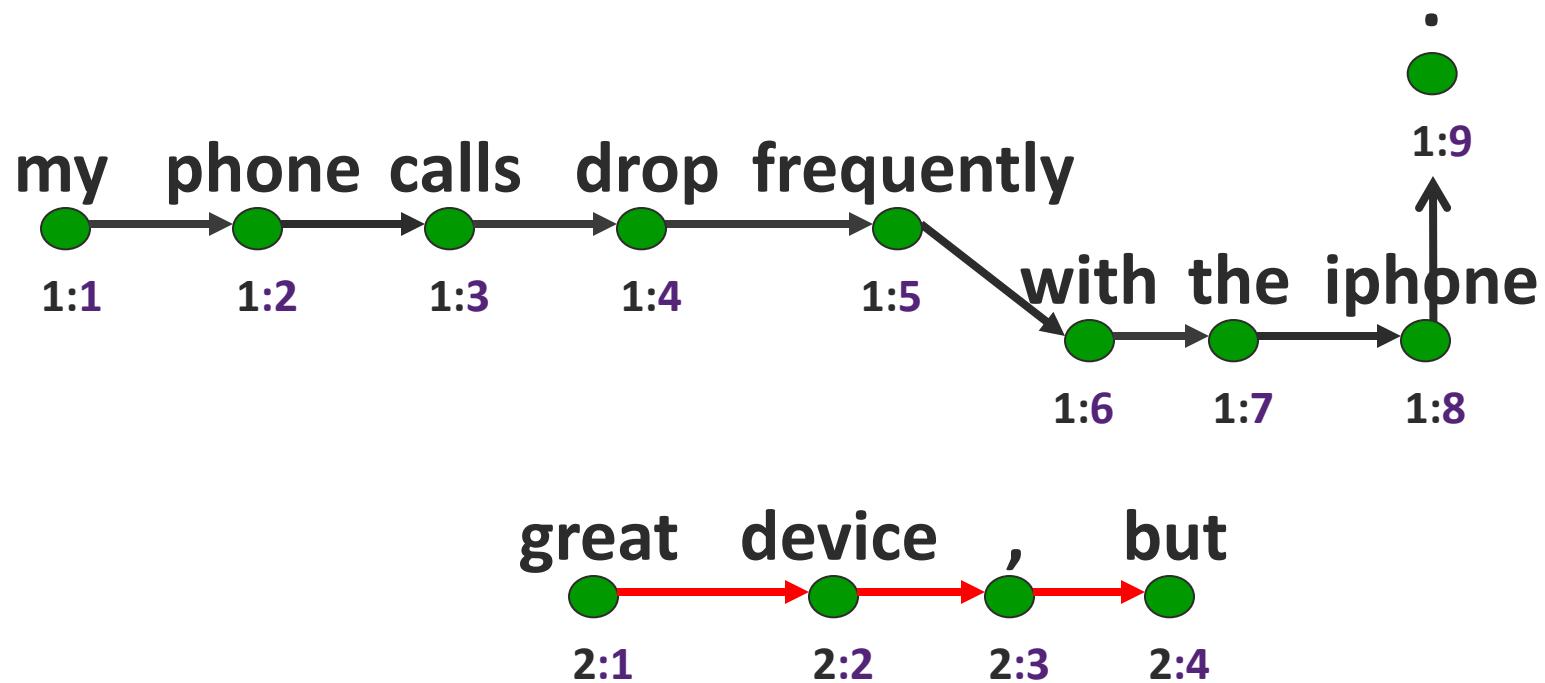
Building Opinosis-Graph

2. Great device, but the calls drop too frequently.



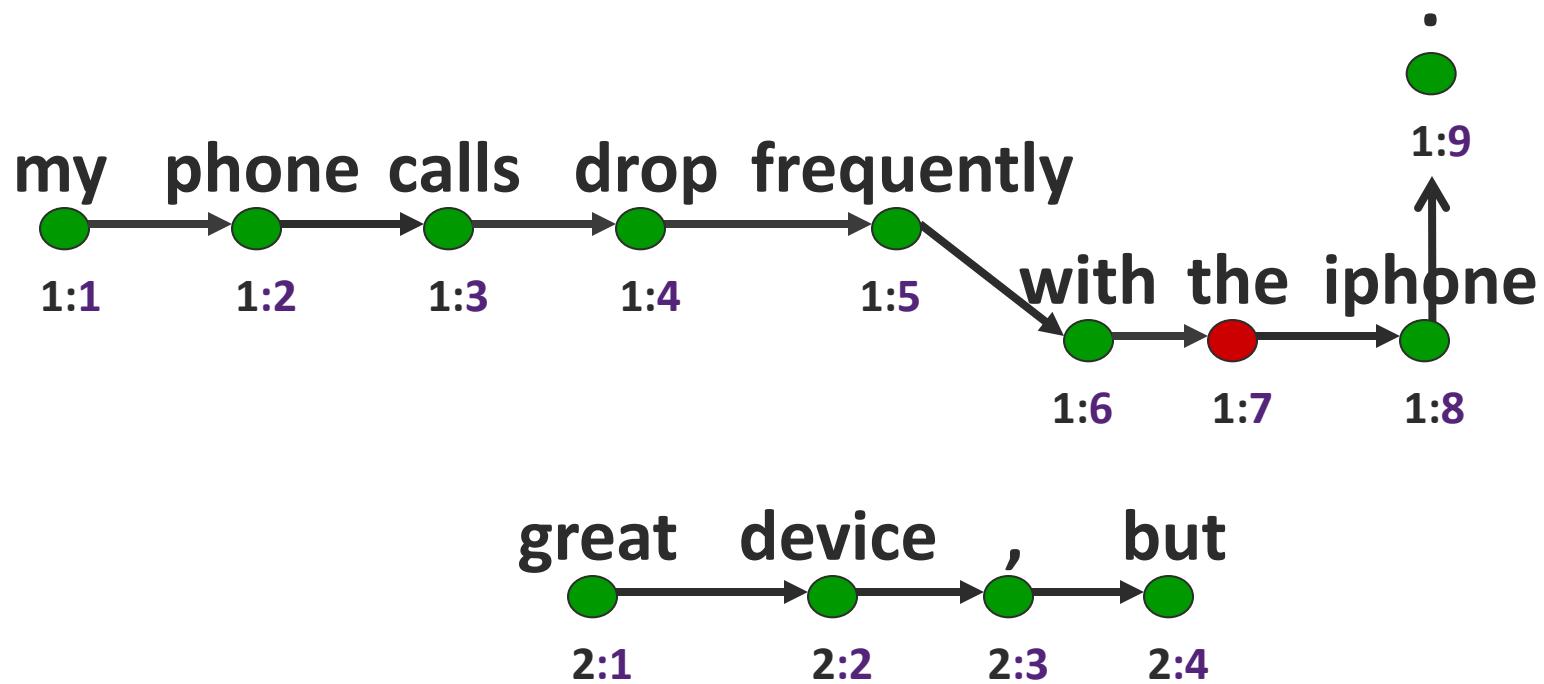
Building Opinosis-Graph

2. Great device, but the calls drop too frequently.



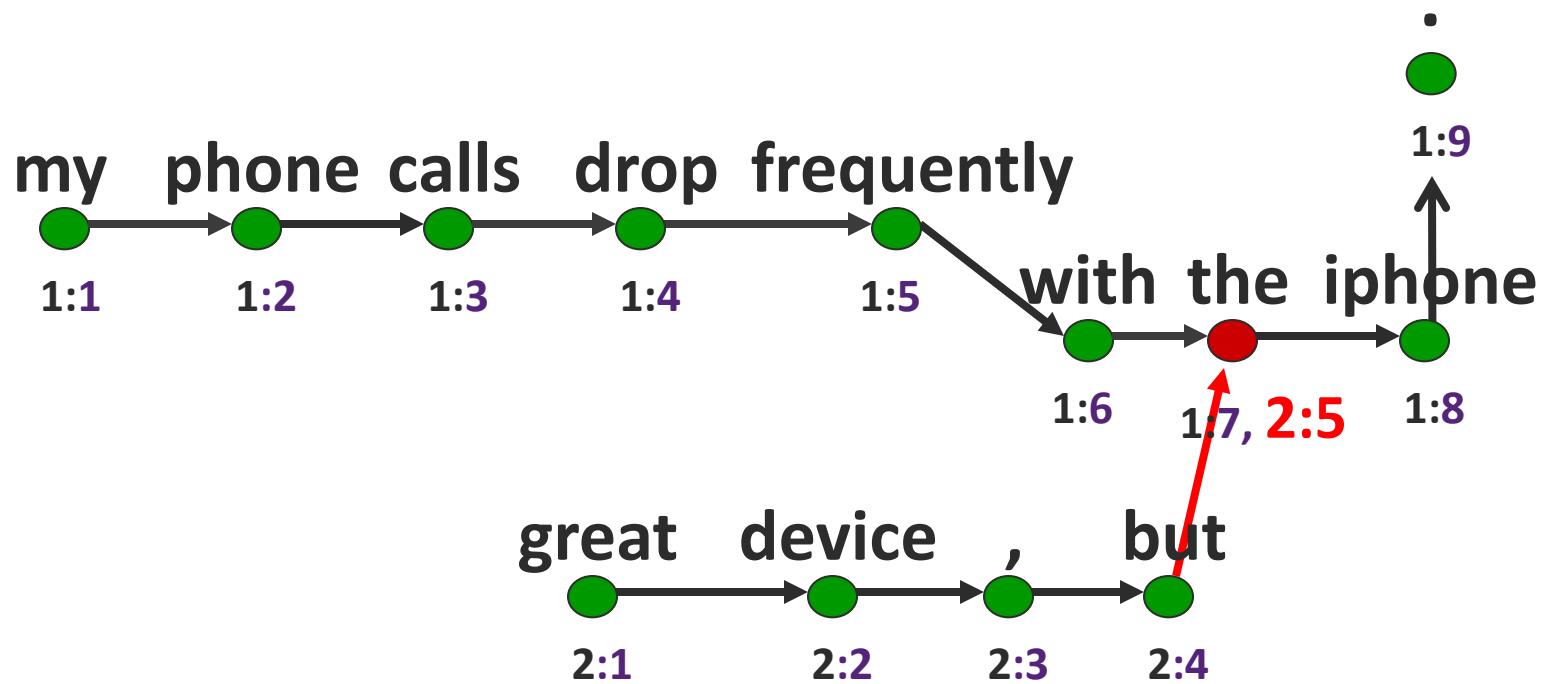
Building Opinosis-Graph

2. *Great device, but the calls drop too frequently.*



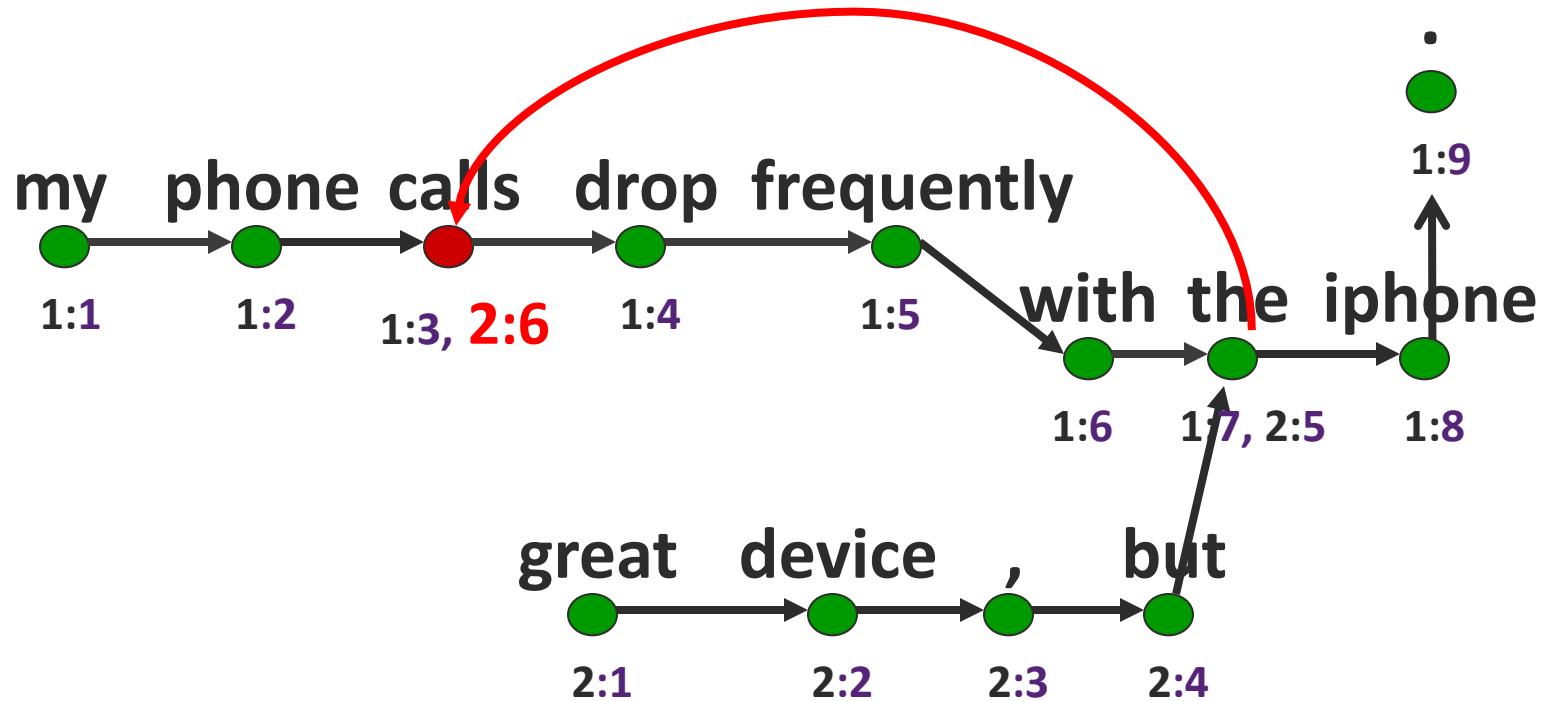
Building Opinosis-Graph

2. Great device, but the calls drop too frequently.



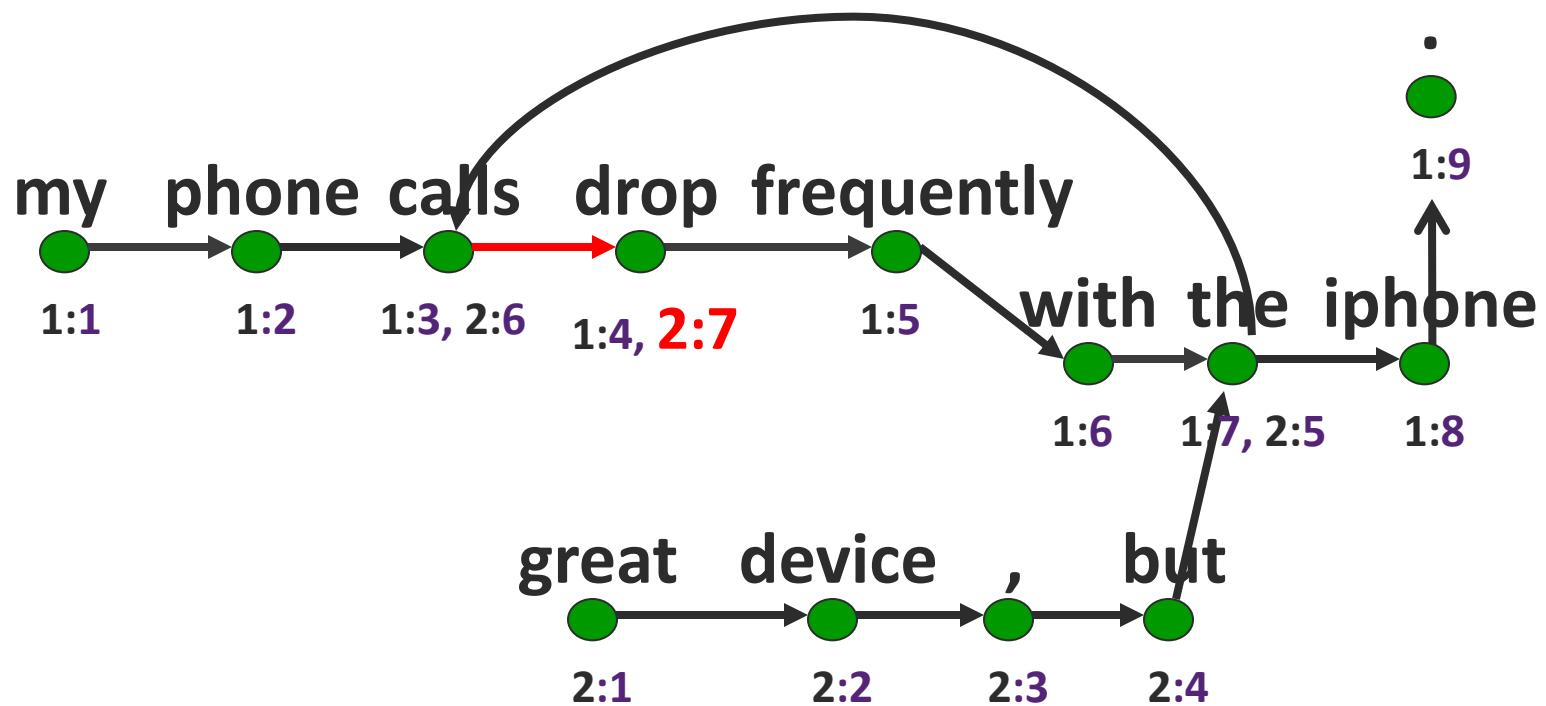
Building Opinosis-Graph

2. Great device, but the calls drop too frequently.



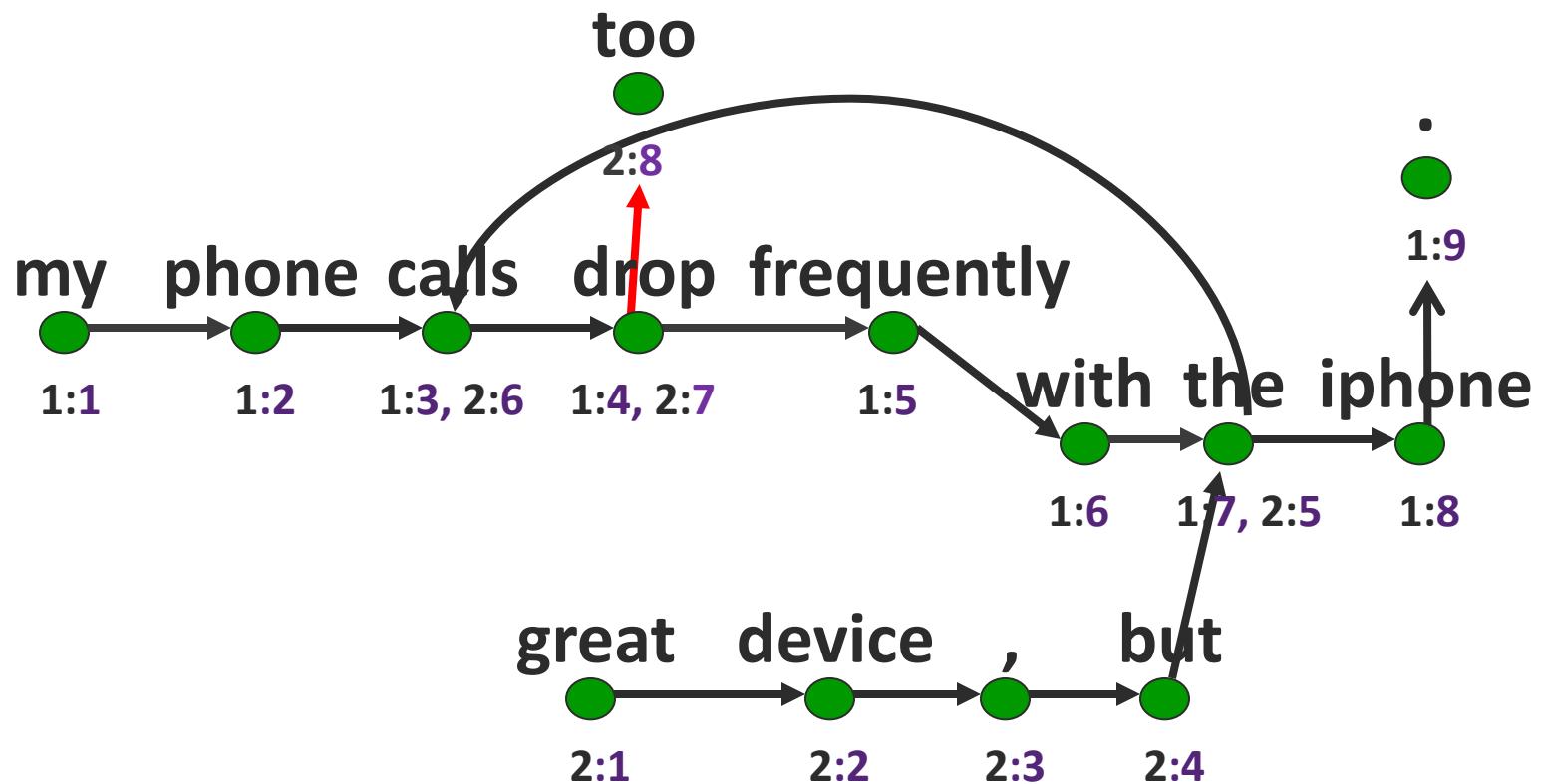
Building Opinosis-Graph

2. *Great device, but the calls drop too frequently.*



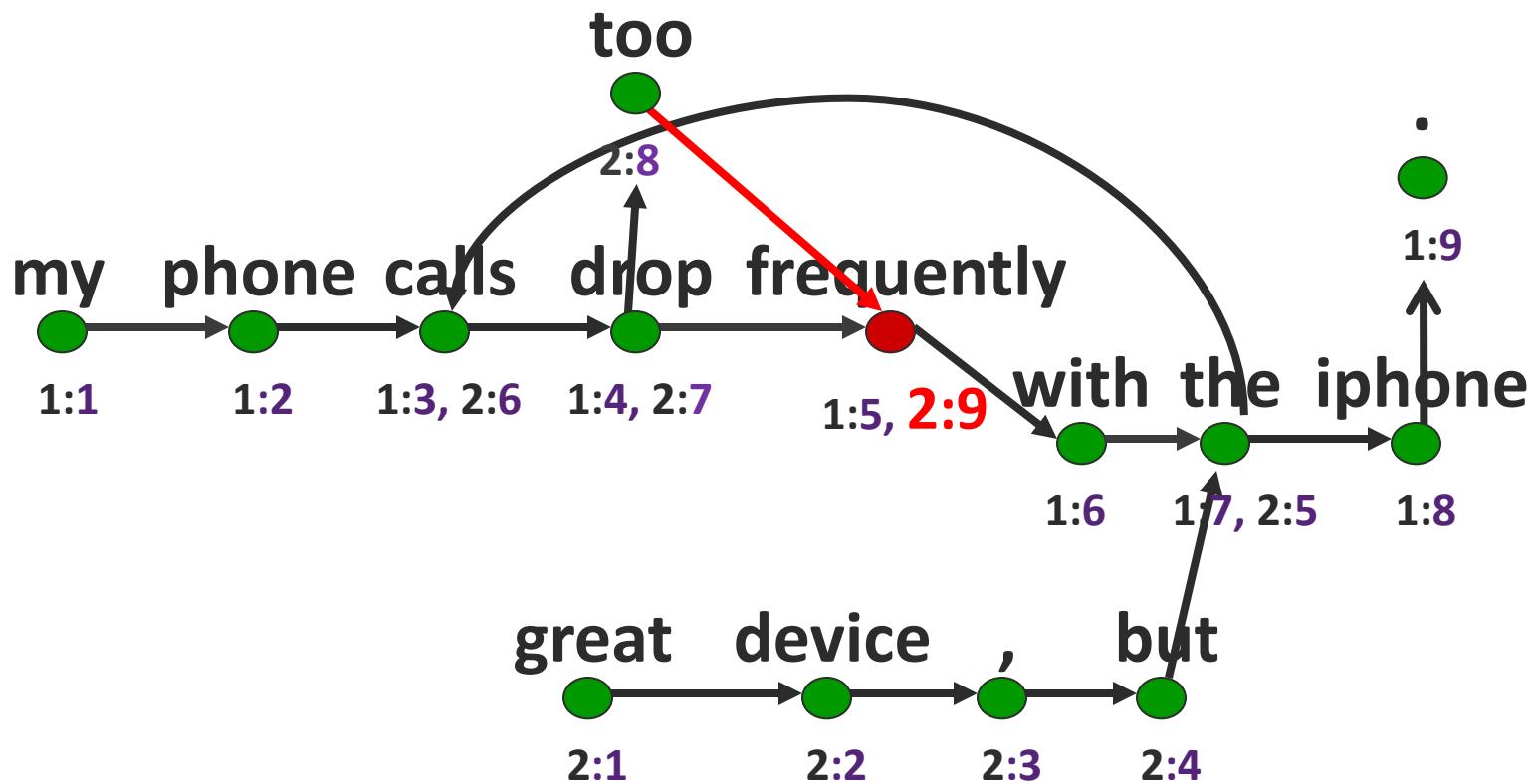
Building Opinosis-Graph

2. *Great device, but the calls drop too frequently.*



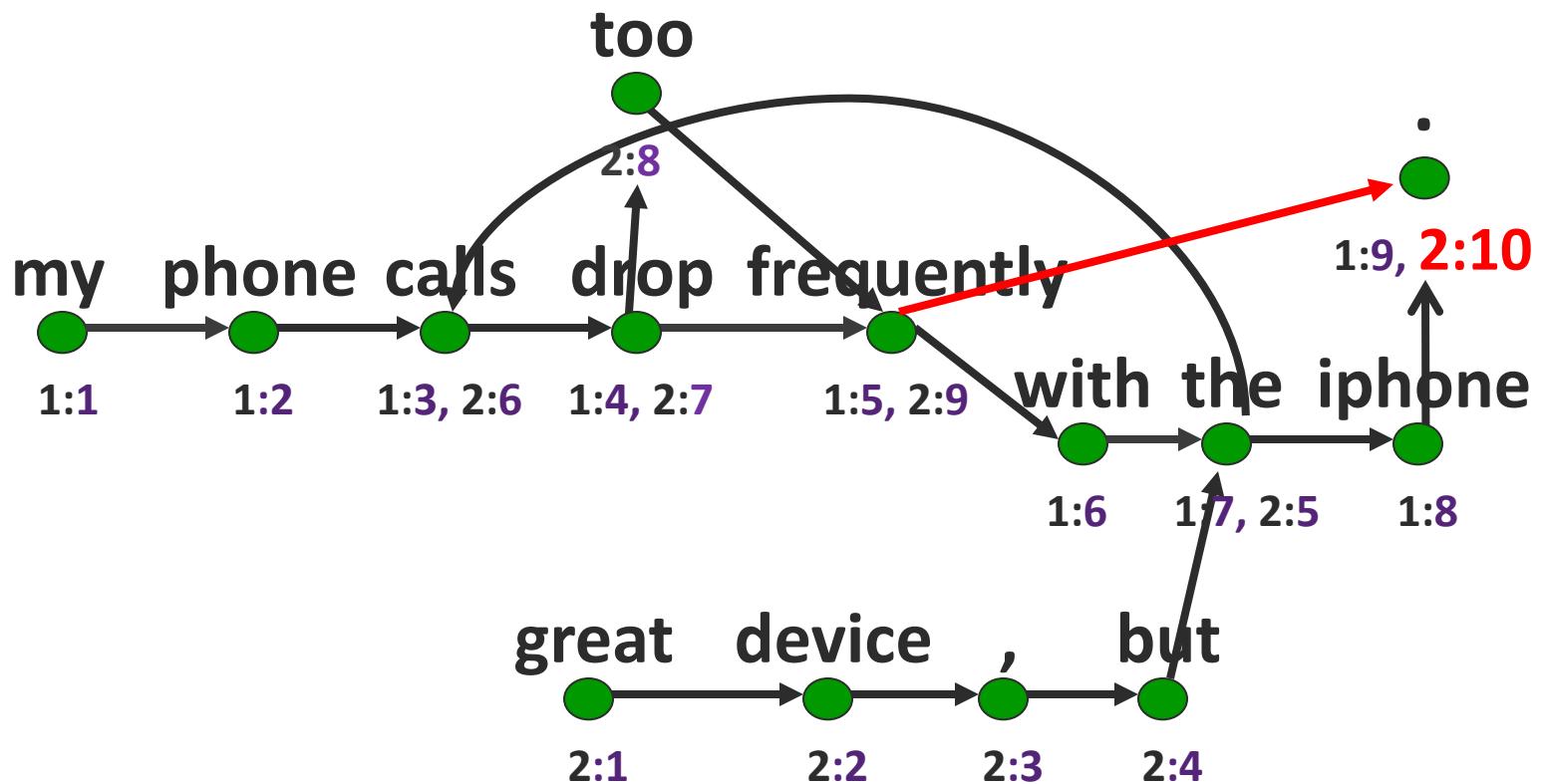
Building Opinosis-Graph

2. *Great device, but the calls drop too frequently.*



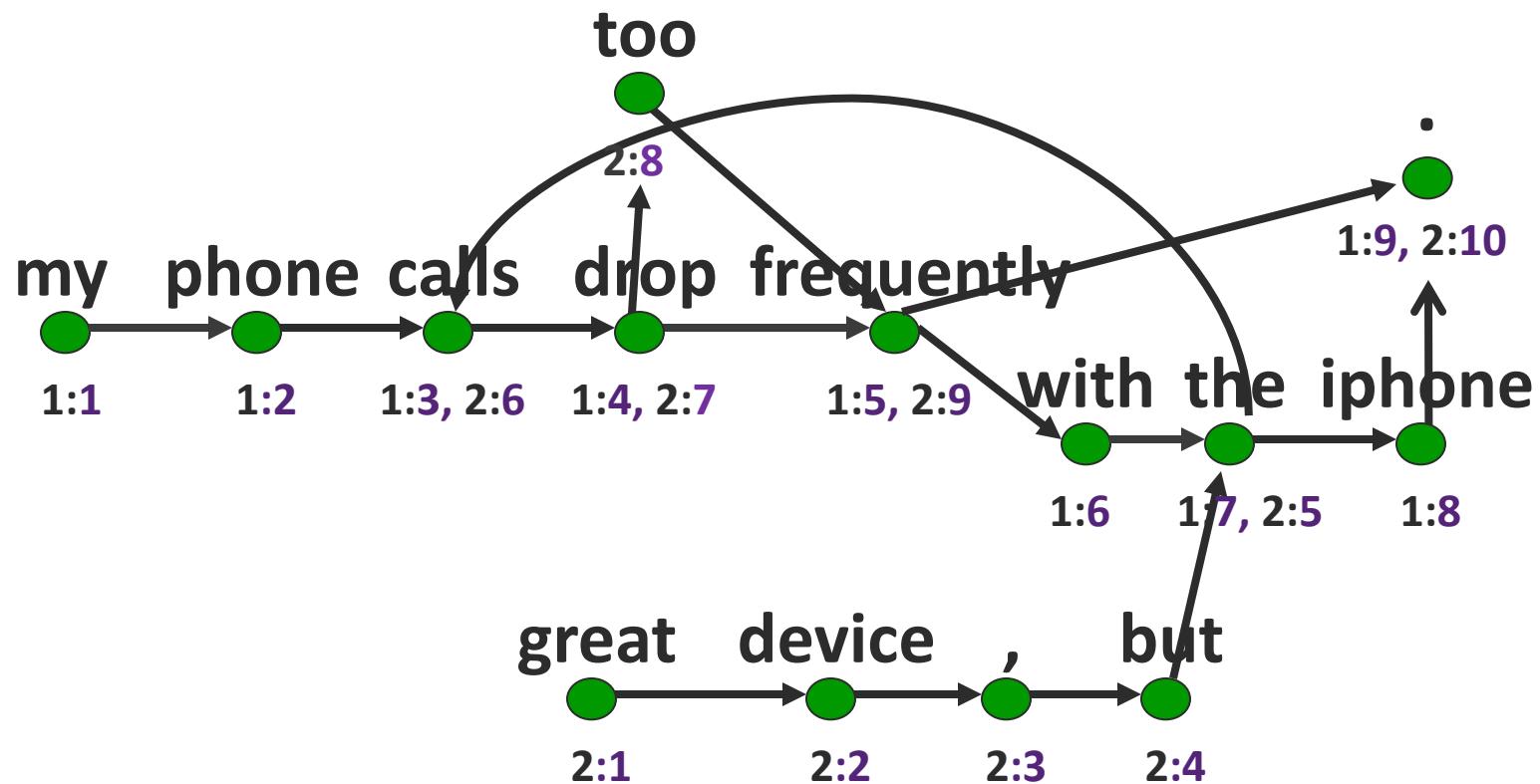
Building Opinosis-Graph

2. *Great device, but the calls drop too frequently.*



Building Opinosis-Graph

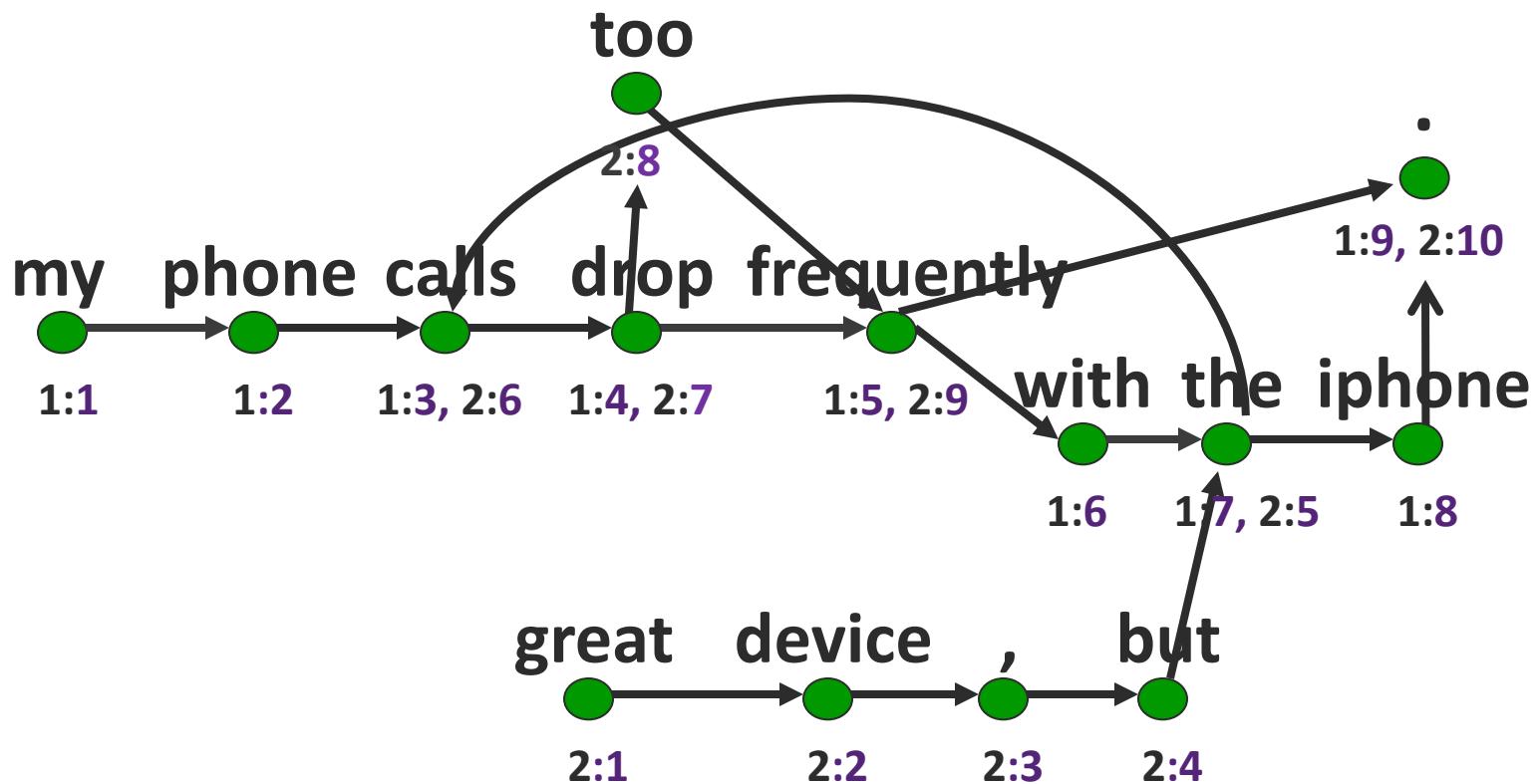
Graph is now ready for Step 2!



3 Important Properties of the Opinosis-Graph

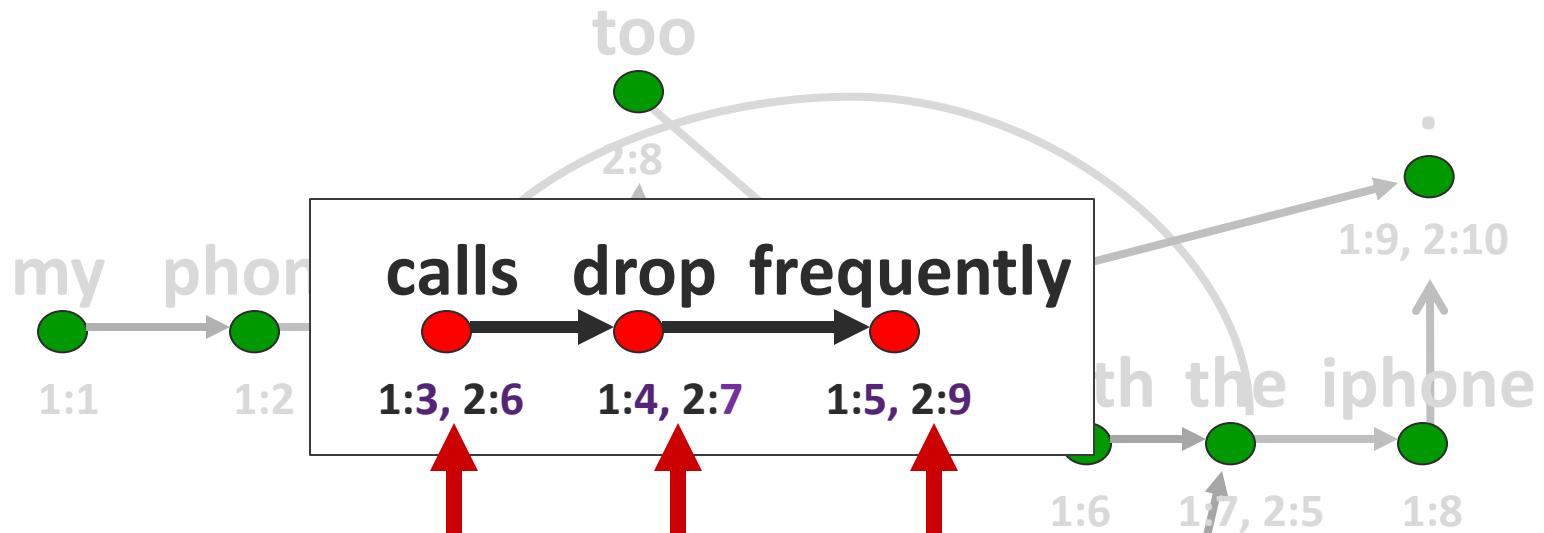
Property 1

Naturally captures redundancies



Property 1

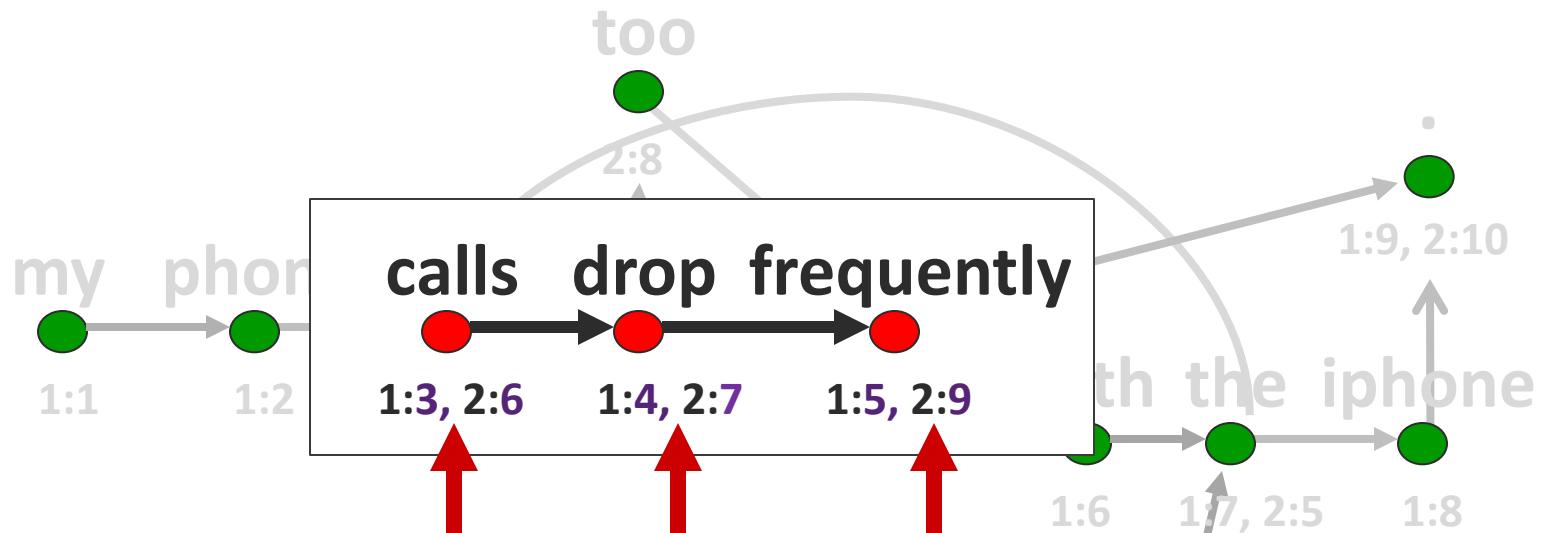
Naturally captures redundancies



Path shared by 2 sentences
naturally captured by nodes

Property 1

Naturally captures redundancies



Easily discover redundancies
for **high confidence** summaries

Property 2

Captures gapped subsequences

1. My phone ***calls drop frequently*** with the iPhone.
2. Great device, but the ***calls drop too frequently***.

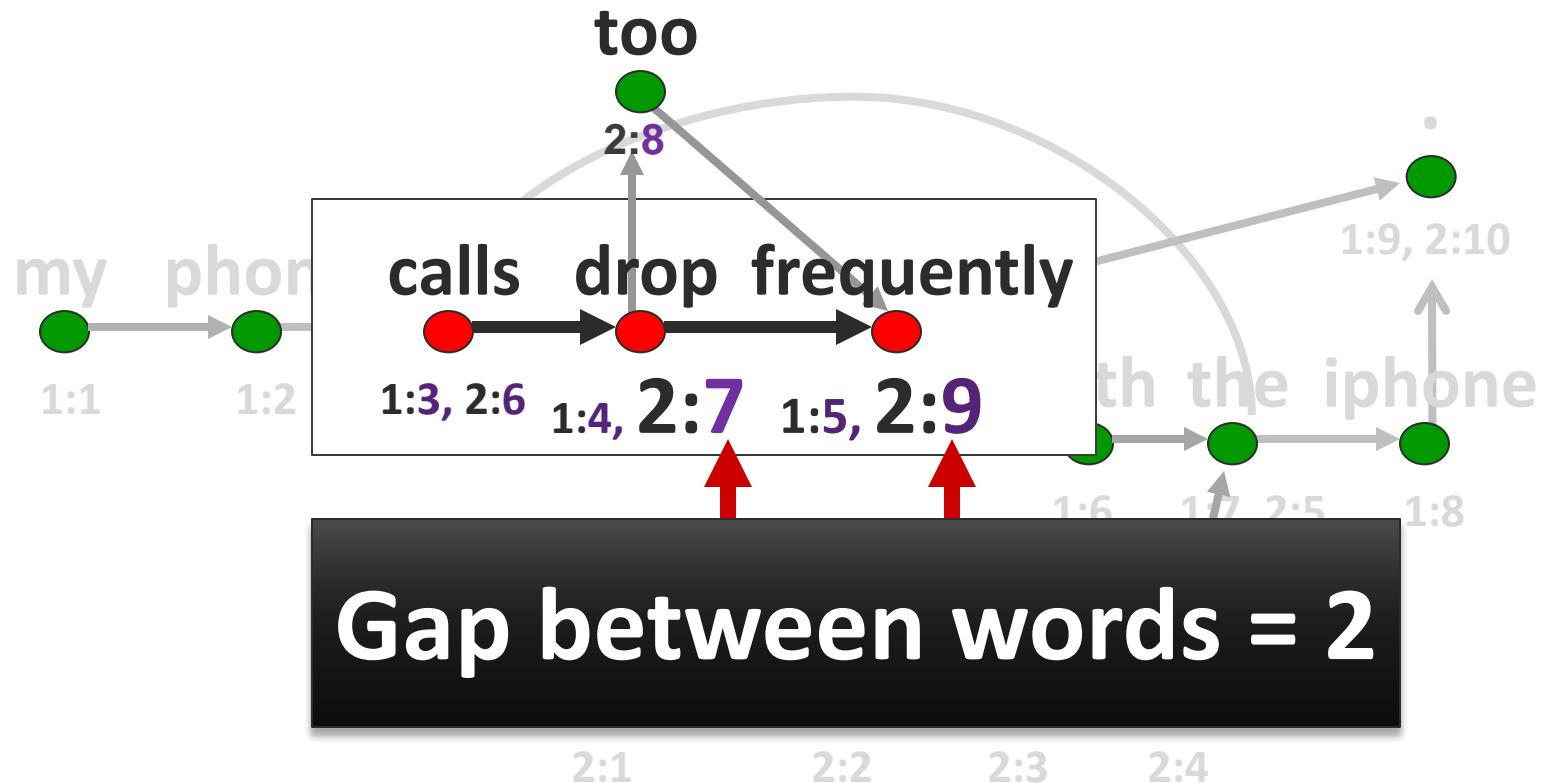
Property 2

Captures gapped subsequences

1. My phone ***calls drop frequently*** with the iPhone.
2. Great device, but the ***calls drop too frequently***.

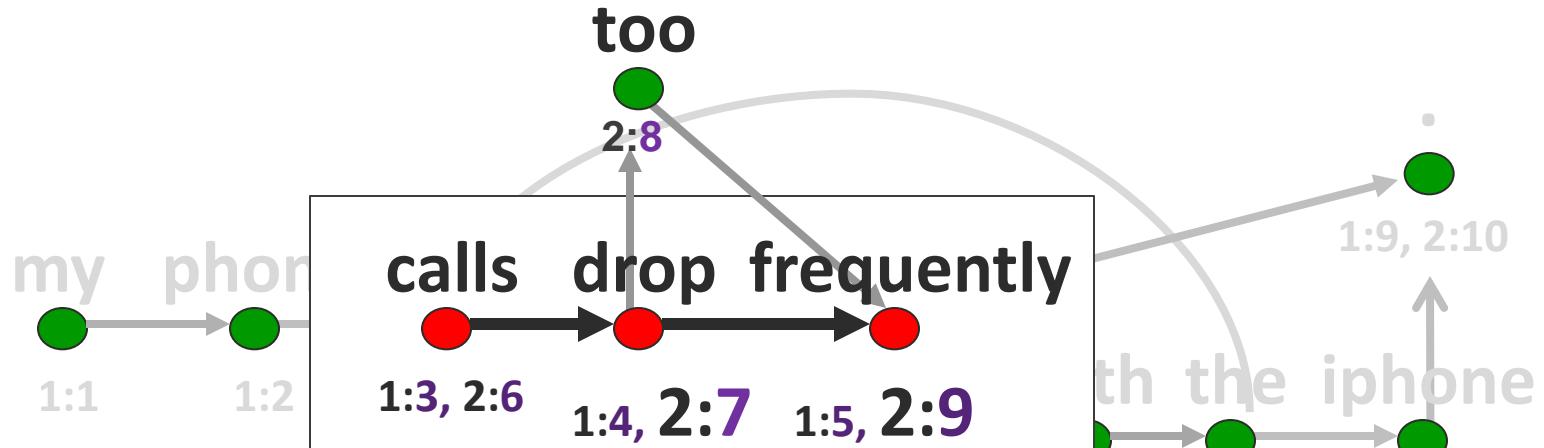
Property 2

Captures gapped subsequences



Property 2

Captures gapped subsequences



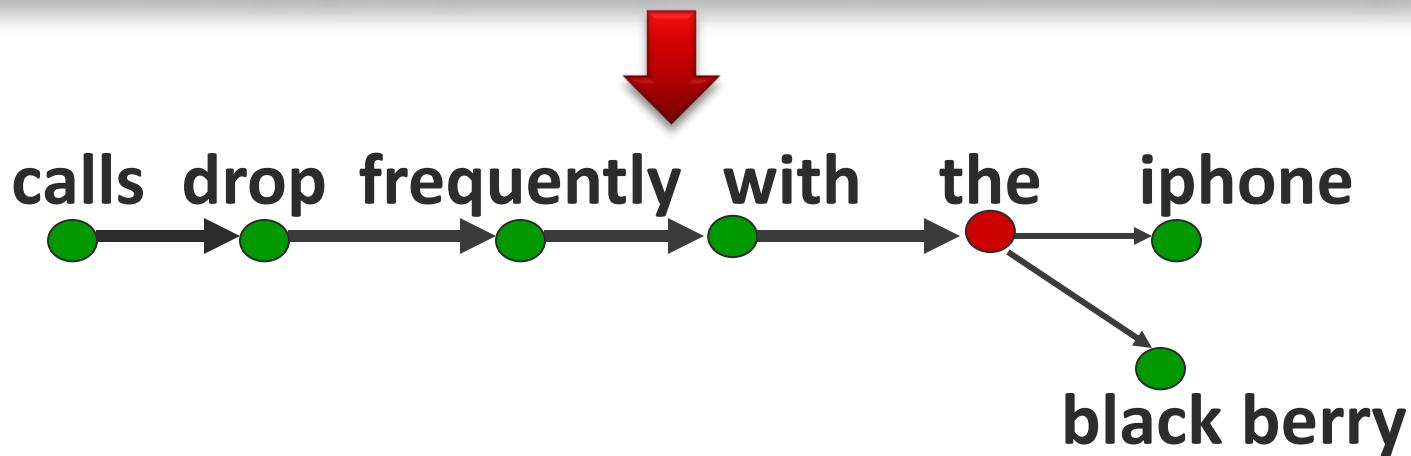
Gapped subsequences allow:

- redundancy enforcements
- discovery of new sentences

Property 3

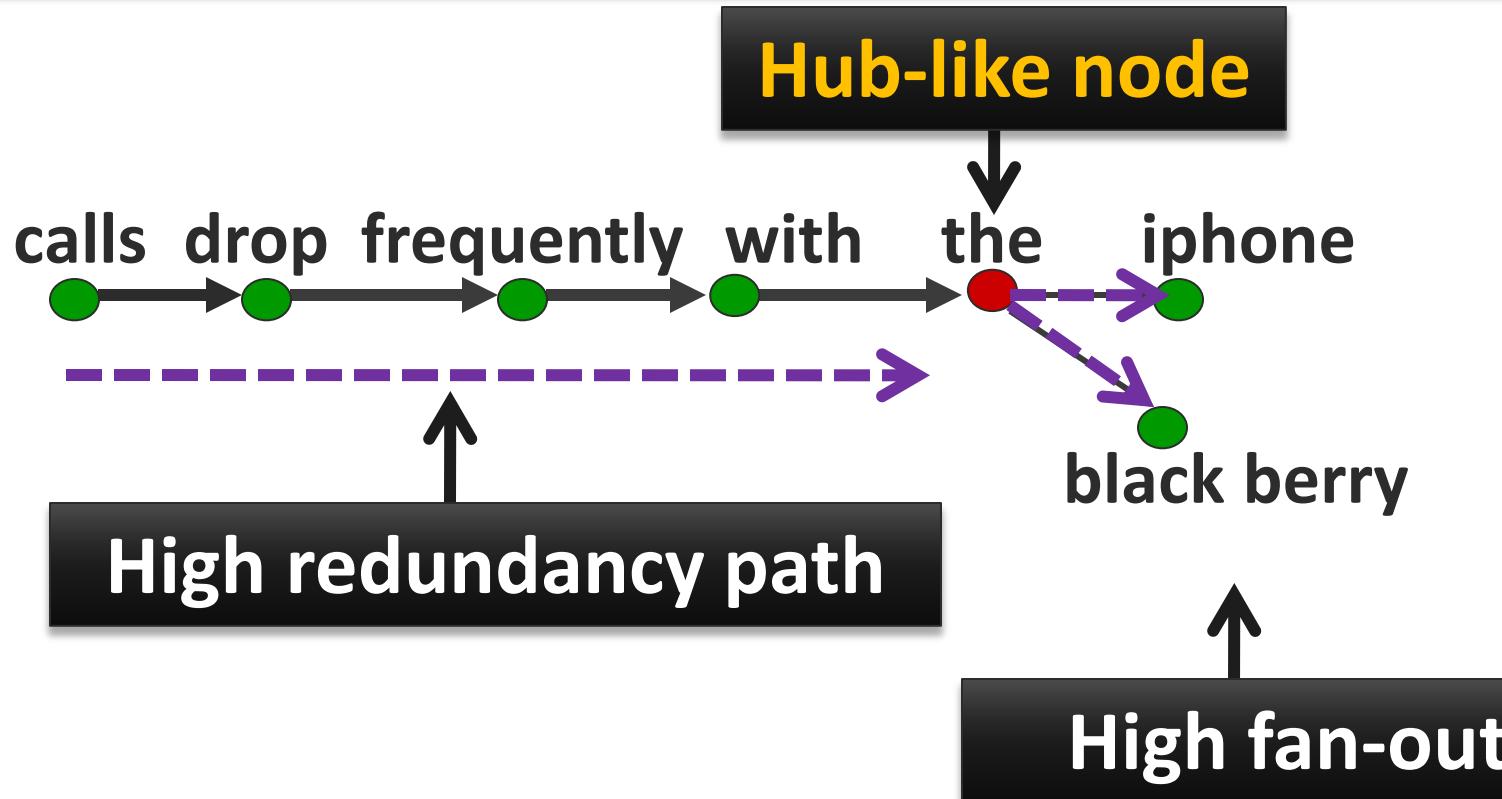
Captures collapsible structures

1. Calls drop frequently with the **iPhone**
2. Calls drop frequently with the **Black Berry**



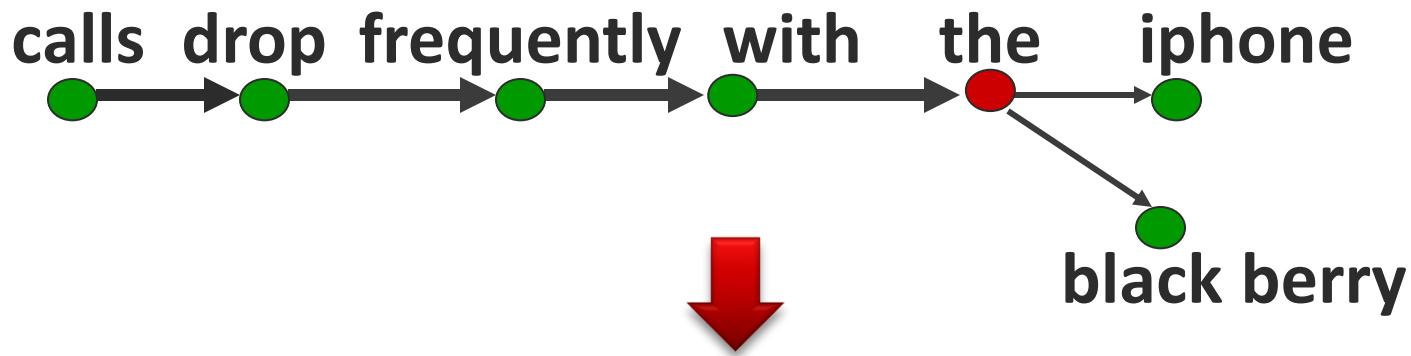
Property 3

Captures collapsible structures



Property 3

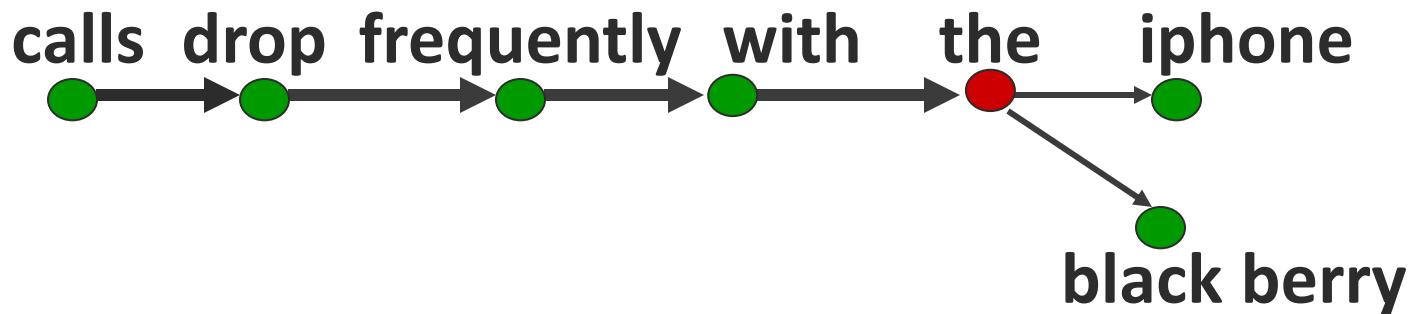
Captures collapsible structures



*"calls drop frequently with the iphone and
the black berry"*

Property 3

Captures collapsible structures



- Ideal for collapse & compression
- Can easily be discovered using OG

Step 2a: Generate Candidate Summaries

Generate Candidate Summaries

**Repeatedly search the Opinosis-
Graph for a *Valid Path***

Valid Path

- ▶ Set of connected nodes
- ▶ Has a Valid Start Node (**VSN**)
 - Natural starting point of a sentence
 - **Opinosis** uses average positional information
- ▶ Has a Valid End Node (**VEN**)
 - Point that completes a sentence
 - **Opinosis** uses punctuations & conjunctions

Finding Candidate Summaries

, calls drop frequently with the iphone .



VSN?

Finding Candidate Summaries

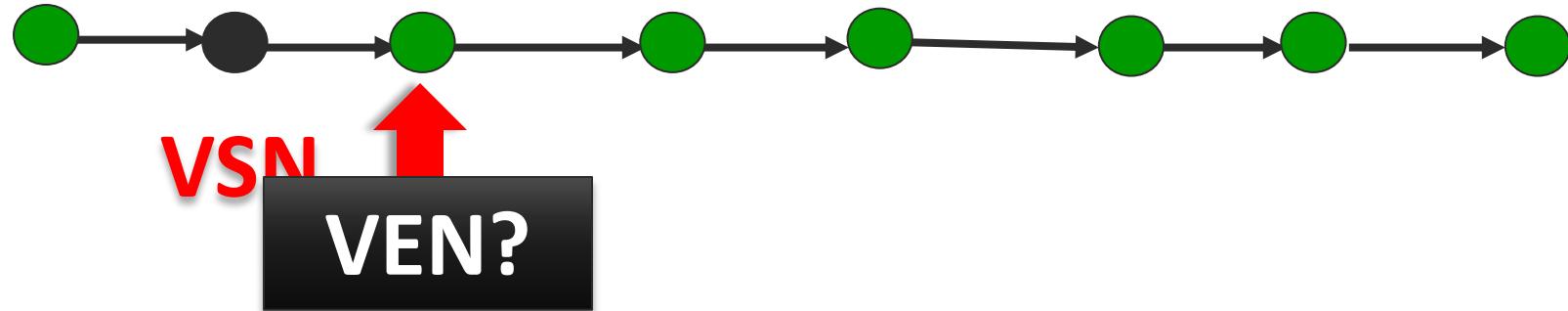
, calls drop frequently with the iphone .



Yes!

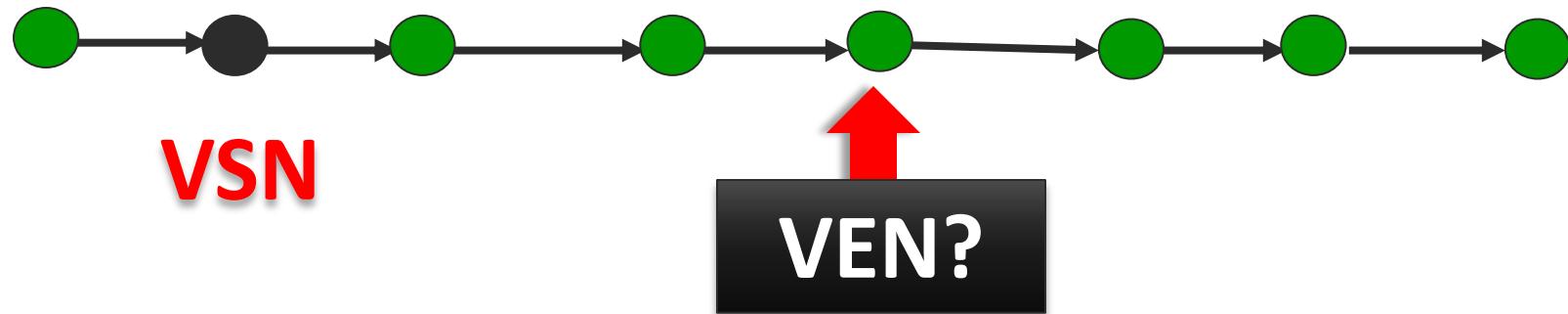
Finding Candidate Summaries

, calls drop frequently with the iphone .



Finding Candidate Summaries

, calls drop frequently with the iphone .



Finding Candidate Summaries

, calls drop frequently with the iphone .

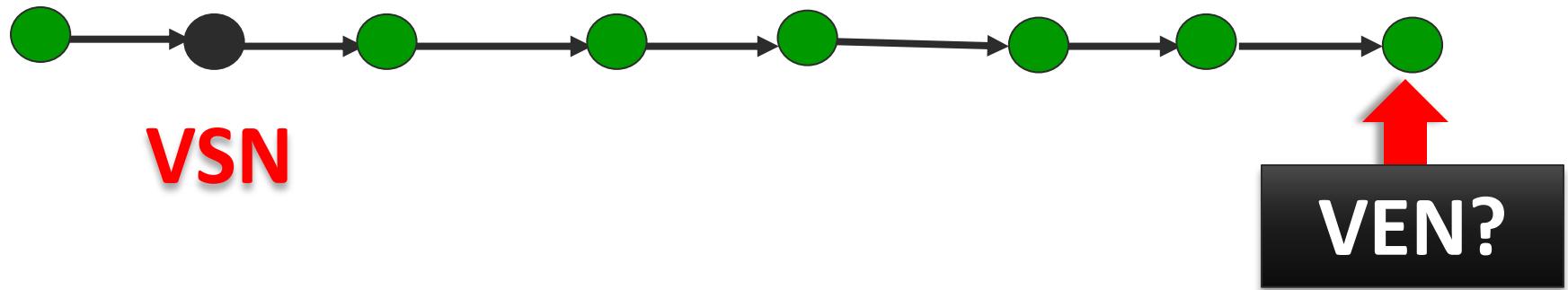


VSN



Finding Candidate Summaries

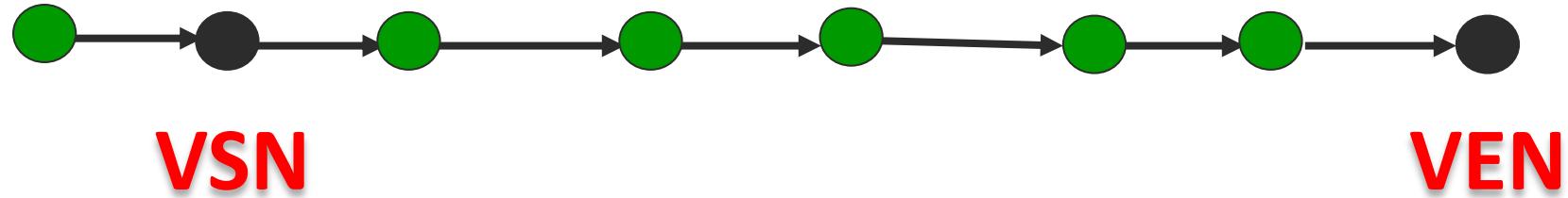
, calls drop frequently with the iphone .



Yes!

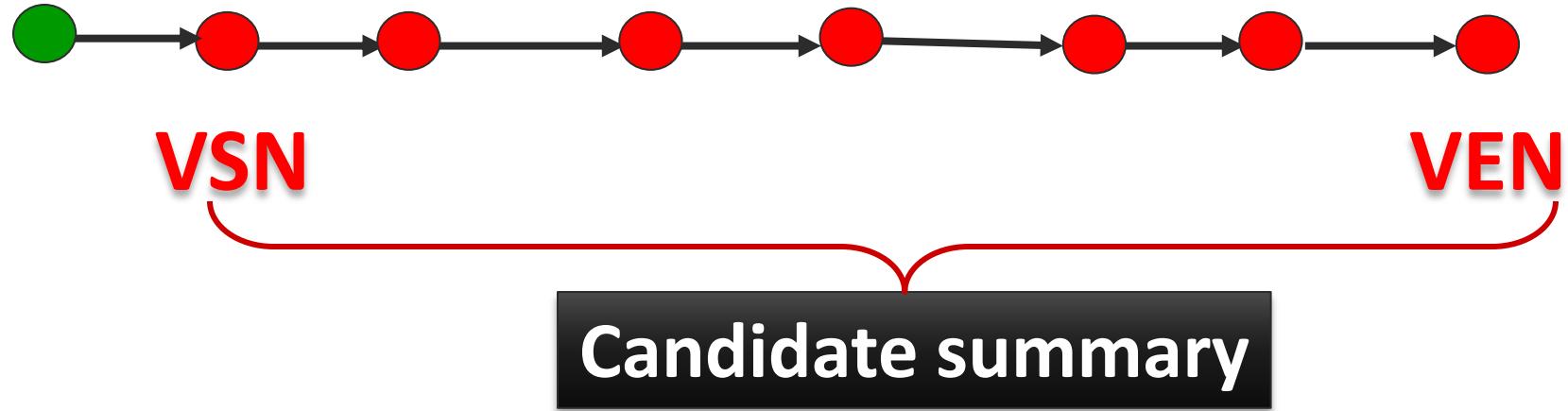
Finding Candidate Summaries

, calls drop frequently with the iphone .



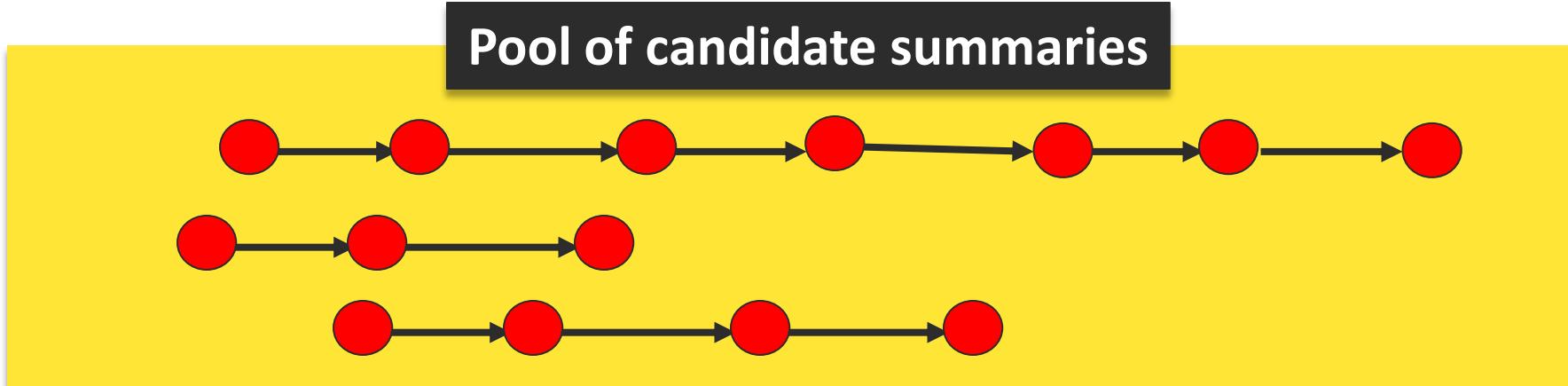
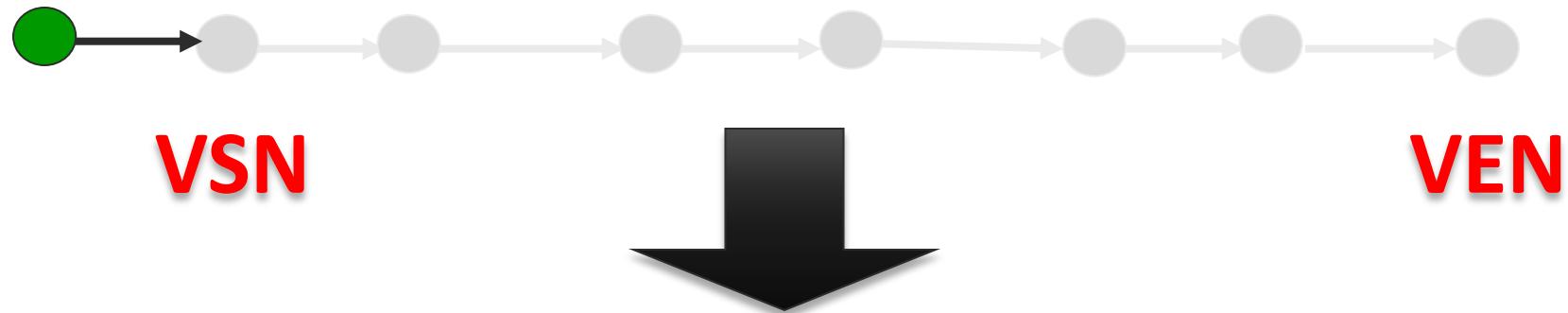
Finding Candidate Summaries

, calls drop frequently with the iphone .



Finding Candidate Summaries

, calls drop frequently with the iphone .



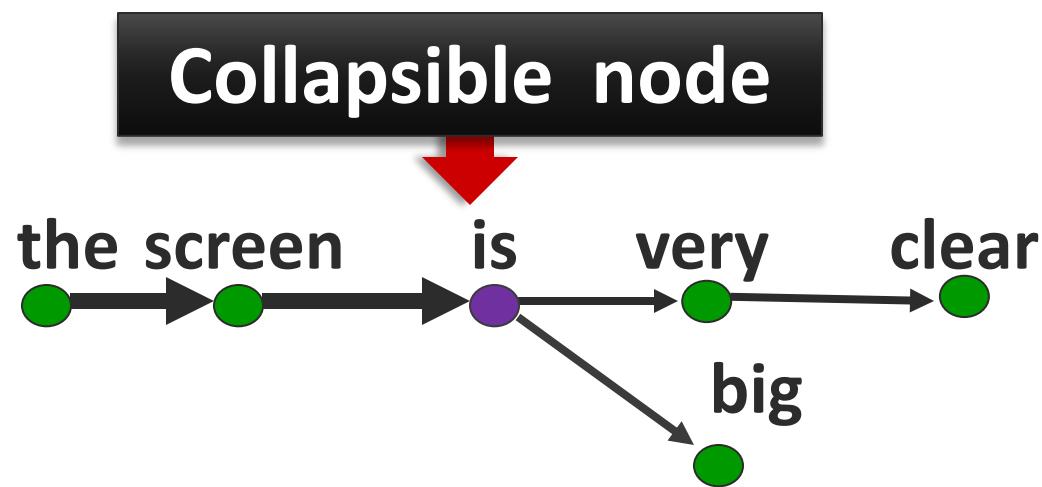
Collapsible Structures

- ▶ Some paths are collapsible
- ▶ We need to identify such paths

Identifying a Collapsible Structure

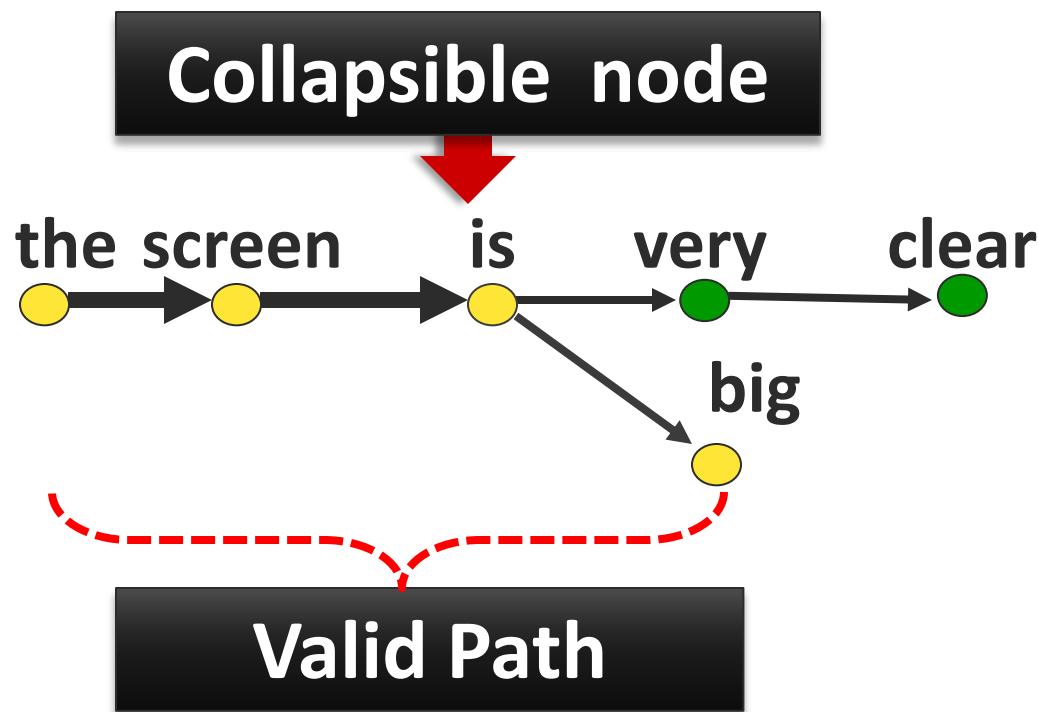
- ▶ Find **collapsible nodes** (hub-like nodes) during traversal
- ▶ Opinosis: Treat **linking verbs** (e.g. is, are) as **collapsible nodes**
 - Linking verbs have hub-like properties
 - Commonly used in opinion text

A Collapsible Structure



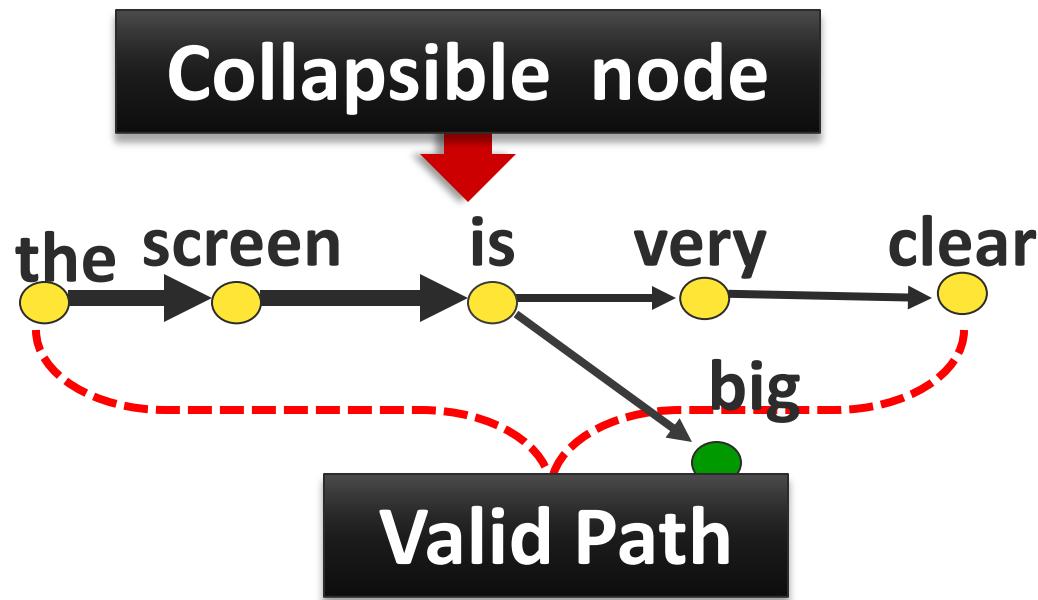
A Collapsible Structure

- ▶ Each path is a **Valid Path**

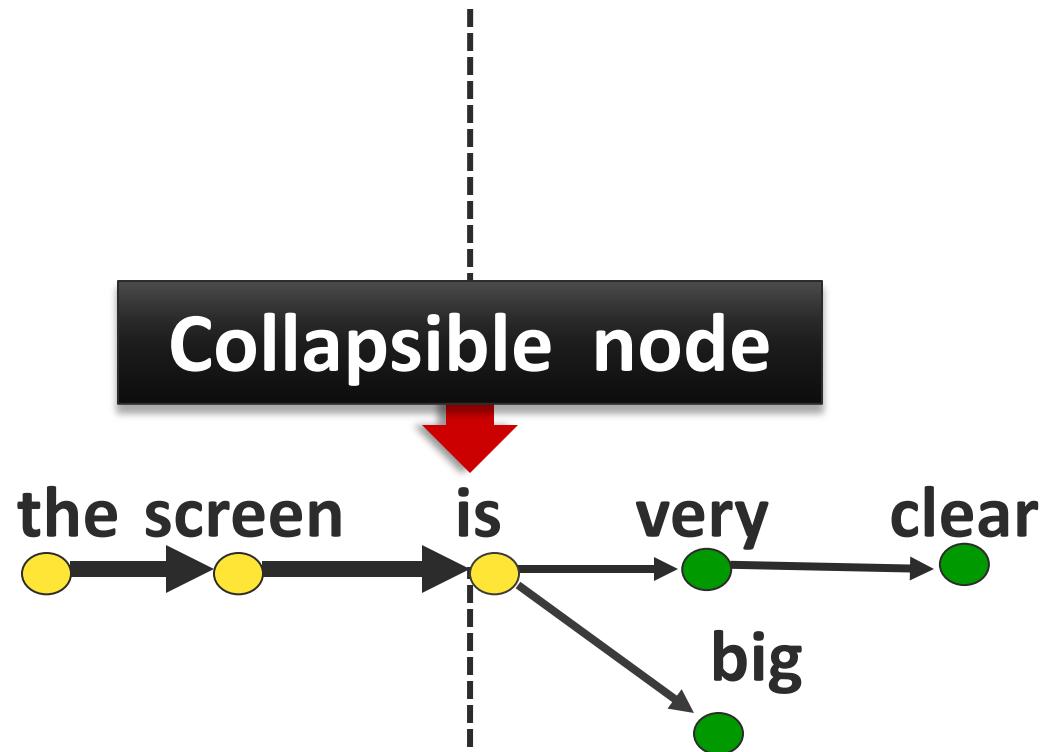


A Collapsible Structure

- ▶ Each path is a **Valid Path**



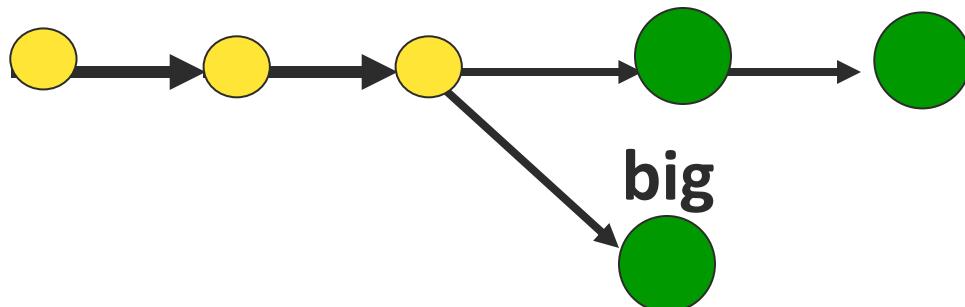
A Collapsible Structure



- *Common structure*
- *High redundancy path*

A Collapsible Structure

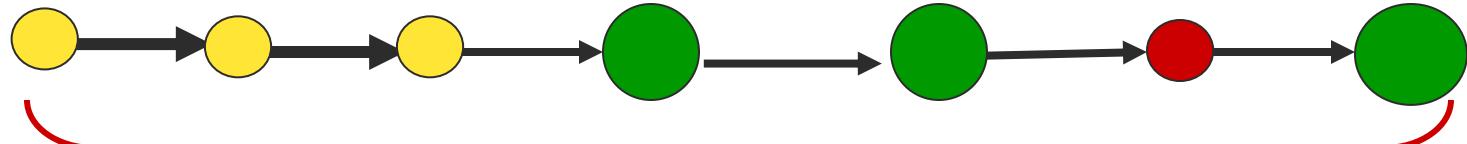
the screen is very clear



COLLAPSE



the screen is very clear and big



Candidate summary

How to collapse?

- ▶ Different ways to collapse
- ▶ Linking verbs: concatenate using commas
“The screen is very clear, bright, big”
- ▶ Better readability:
*“The screen is very clear, bright **and** big”*

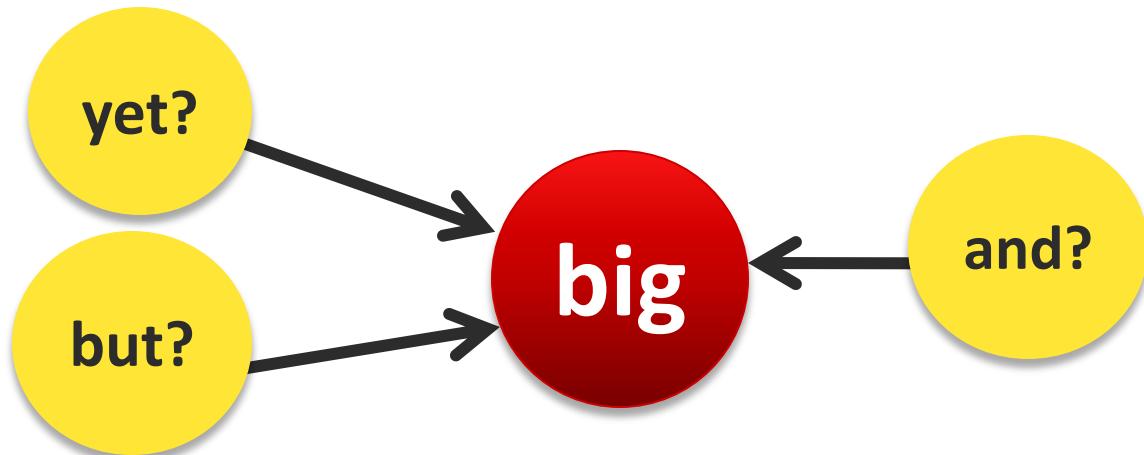


Find last connector

How to collapse?

- ▶ Use existing Opinosis-Graph:
 - Find conjunction that appears most frequently with last collapsed candidate

*“The screen is very clear, bright ? **big**”*



Step 2b: Score Candidate Summaries

Scoring Candidate Summaries

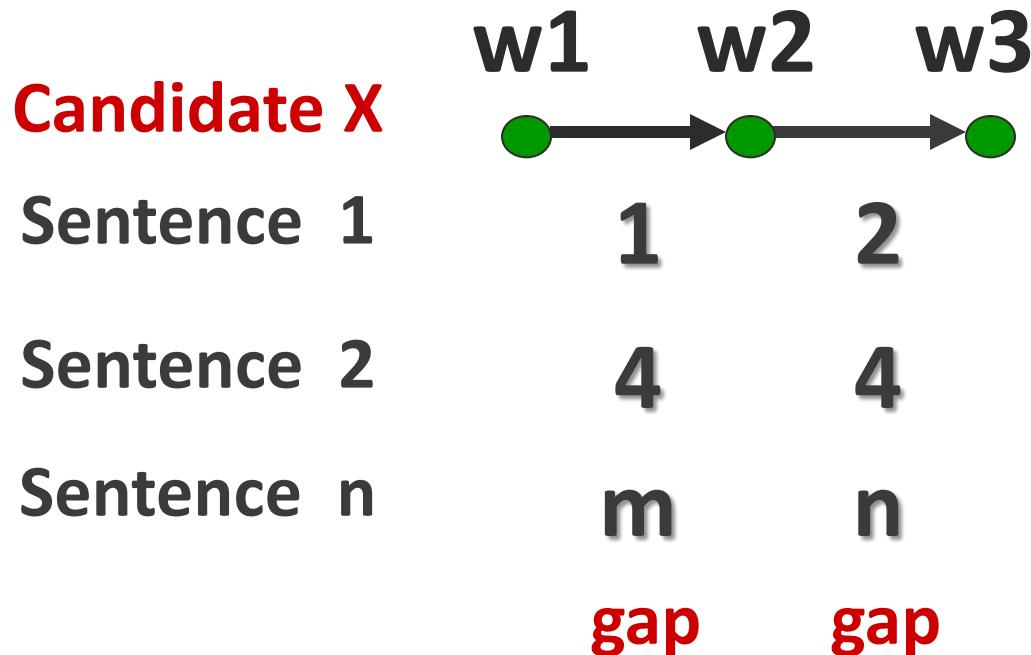
- ▶ Various properties can be used
- ▶ We introduce **2 types of scoring**

Scoring Candidate Summaries

- ▶ **We want high confidence summaries**
 - Select candidates with **high redundancy**
 - # of sentences sharing same path
 - controlled by gap threshold, σ_{gap}
- ▶ **+ Summaries with good coverage**
 - Select **longer sentences**
 - level of redundancy * length of candidate paths
 - Favor longer but redundant sentences

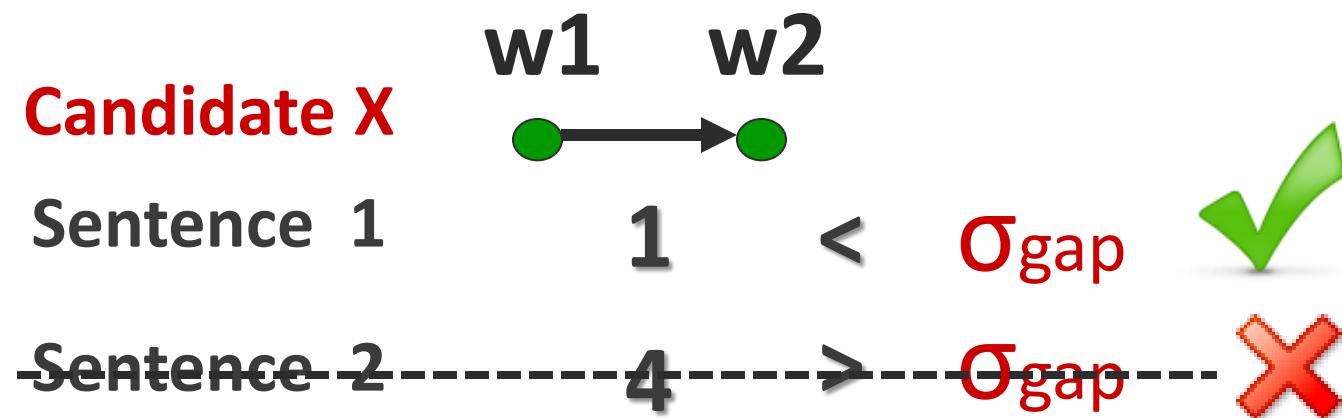
Gap Requirement (σ_{gap})

- Gaps vary between sentences sharing nodes



Gap Requirement (σ_{gap})

- ▶ σ_{gap} enforces maximum allowed gap between two adjacent nodes



- Lower risk of ill-formed sentences
- Avoids over-estimation of redundancy

Step 3: Final Opinosis Summaries

- ▶ Select **top 2** scoring candidates that are most **dissimilar**

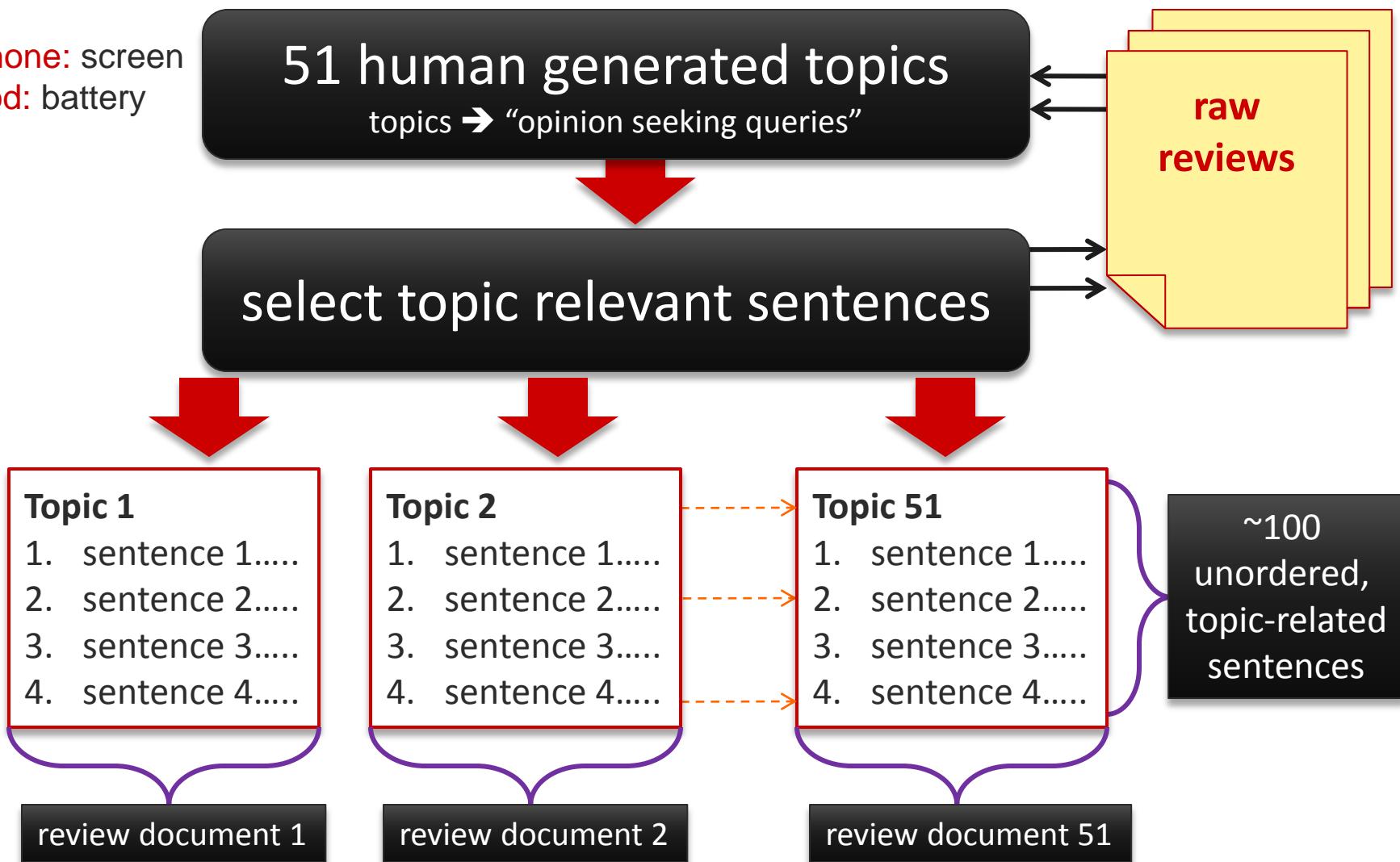
Evaluation

Raw Data: User Reviews

- ▶ Hotels: Tripadvisor.com 
- ▶ Products: Amazon.com 
- ▶ Cars: Edmunds.com 

Dataset Construction

Ex. iPhone: screen
iPod: battery



Summarization Task

► Input: A review document

Topic X

1. sentence 1.....
2. sentence 2.....
3. sentence 3.....
4. sentence 4.....

► Output: A concise abstractive opinion summary

Opinosis Summary:

The screen is big and clear.

Gold Standard

- ▶ **Human written summaries**
 - Concise (~25 words)
 - Focus on summarizing major opinions
- ▶ **Obtained using Amazon's Online Workforce (mturk.com)**
- ▶ **~4 human summaries per topic**

Baseline

- ▶ Hard to find ‘general’ abstractive summarizer
- ▶ Use **MEAD** - Extractive based method
[Radev et al.2000]
 - Ideal: selects sentences that capture most essential information
 - Select **2 sentences** as the summary

Evaluation Measures

► ROUGE (rouge-1, rouge-2, rouge-su4)

- Standard measure for summarization tasks
- Measures: agreement between system & human summaries

► Readability Test

- Measures: How different Opinosis summaries are compared to human composed summaries?

Results

Human-Human Agreement

- ▶ **Estimate:** How much one summary writer agrees with the rest

Human-Human Agreement

ROUGE Scores		
ROUGE-1	Precision	0.34
	Recall	<u>0.32</u>
	F-score	0.31

ROUGE-SU4	P	0.11
		<u>0.12</u>

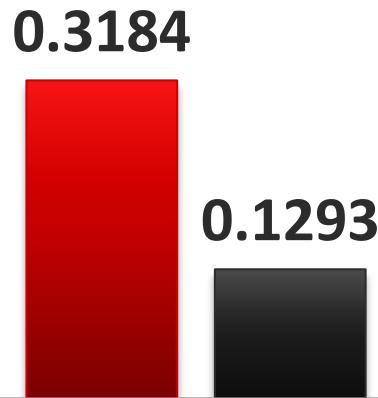
Manual inspection: Human summaries semantically similar. Difference in word usage.

Human vs. Opinosis vs. MEAD

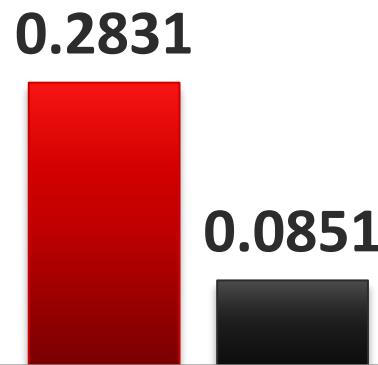
Human vs. Opinosis vs. MEAD

Rouge Recall Scores

- ROUGE-1
- ROUGE-SU4



HUMAN
(17 words)



OPINOSISbest
(15 words)

Highest recall

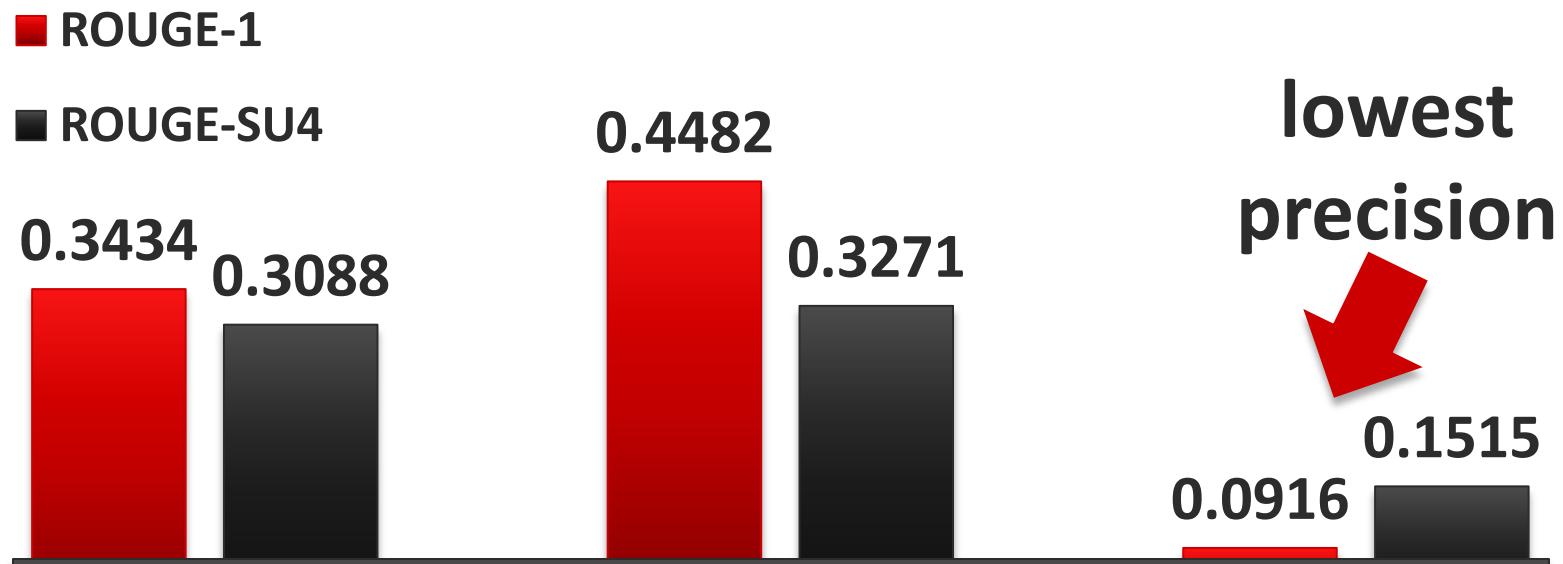


MEAD
(75 words)

Much longer sentences

Human vs. Opinosis vs. MEAD

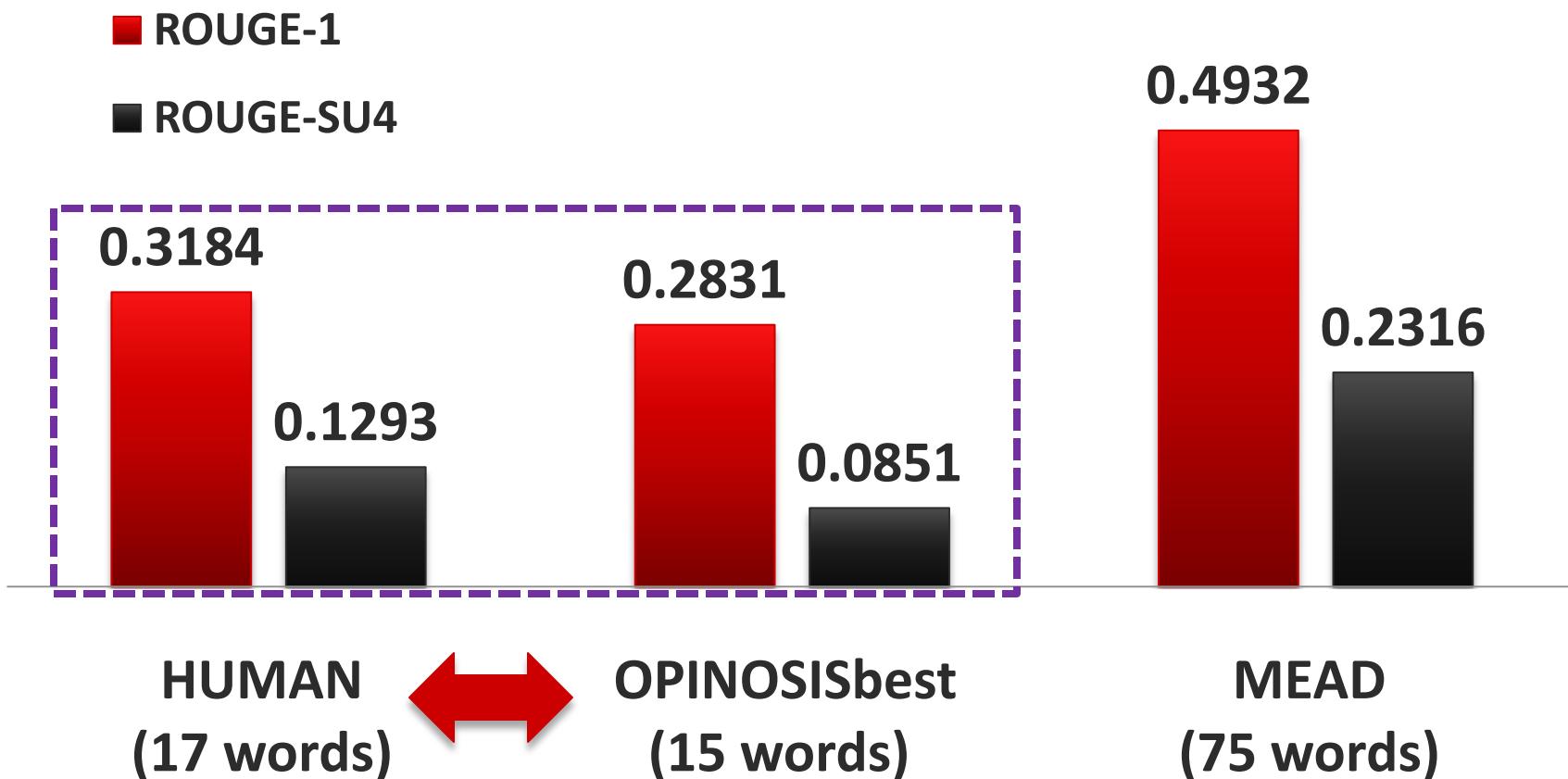
Rouge Precision Scores



Overall: Baseline does not do well in generating concise summaries.

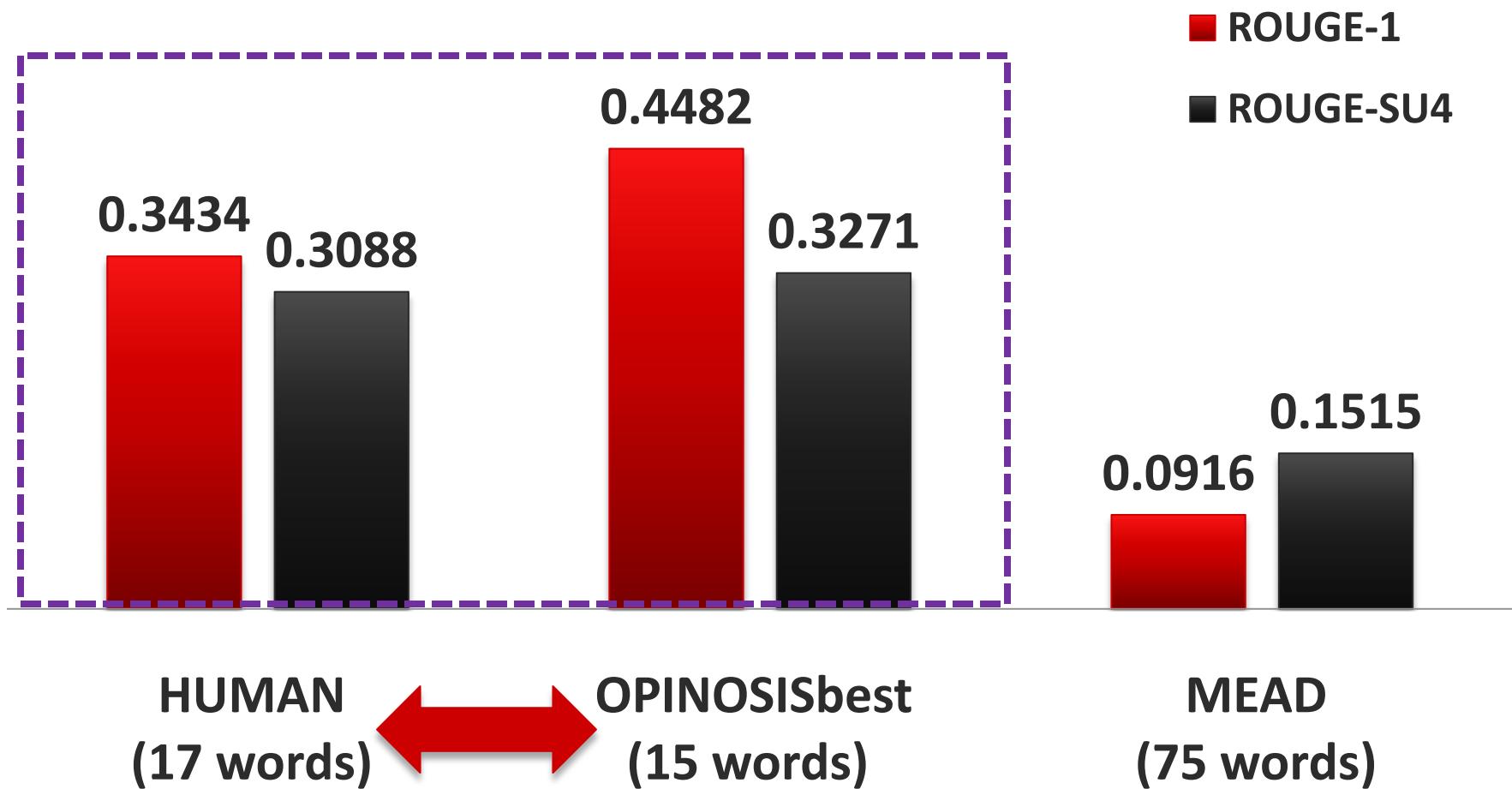
Human vs. Opinosis vs. MEAD

Rouge Recall Scores

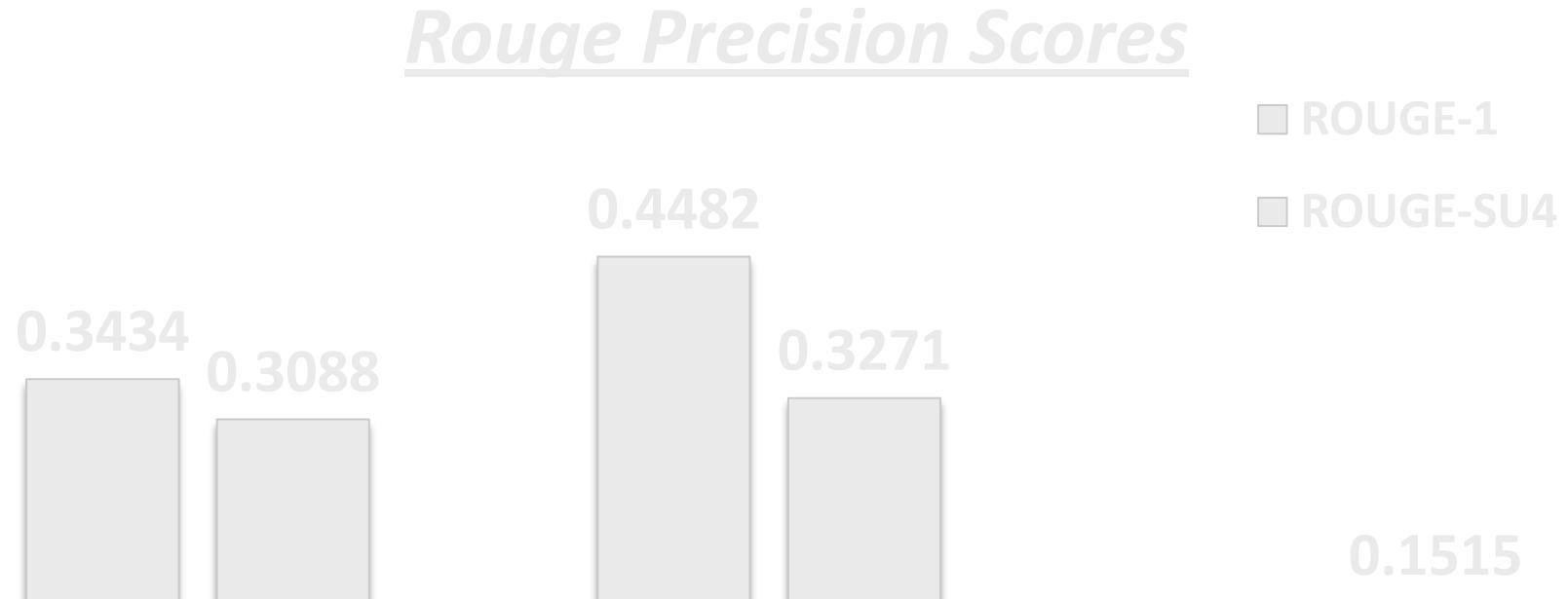


Human vs. Opinosis vs. MEAD

Rouge Precision Scores



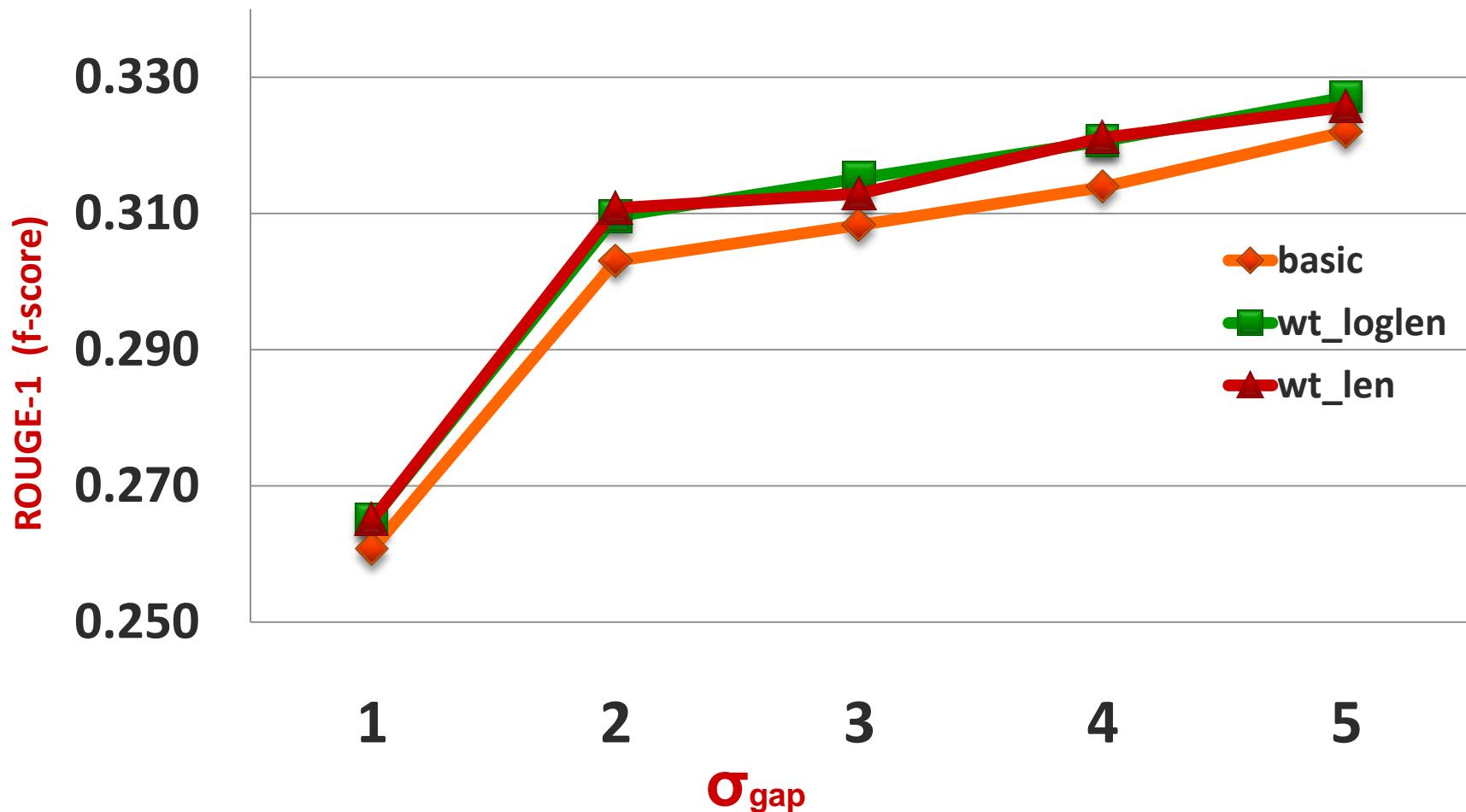
Human vs. Opinosis vs. MEAD



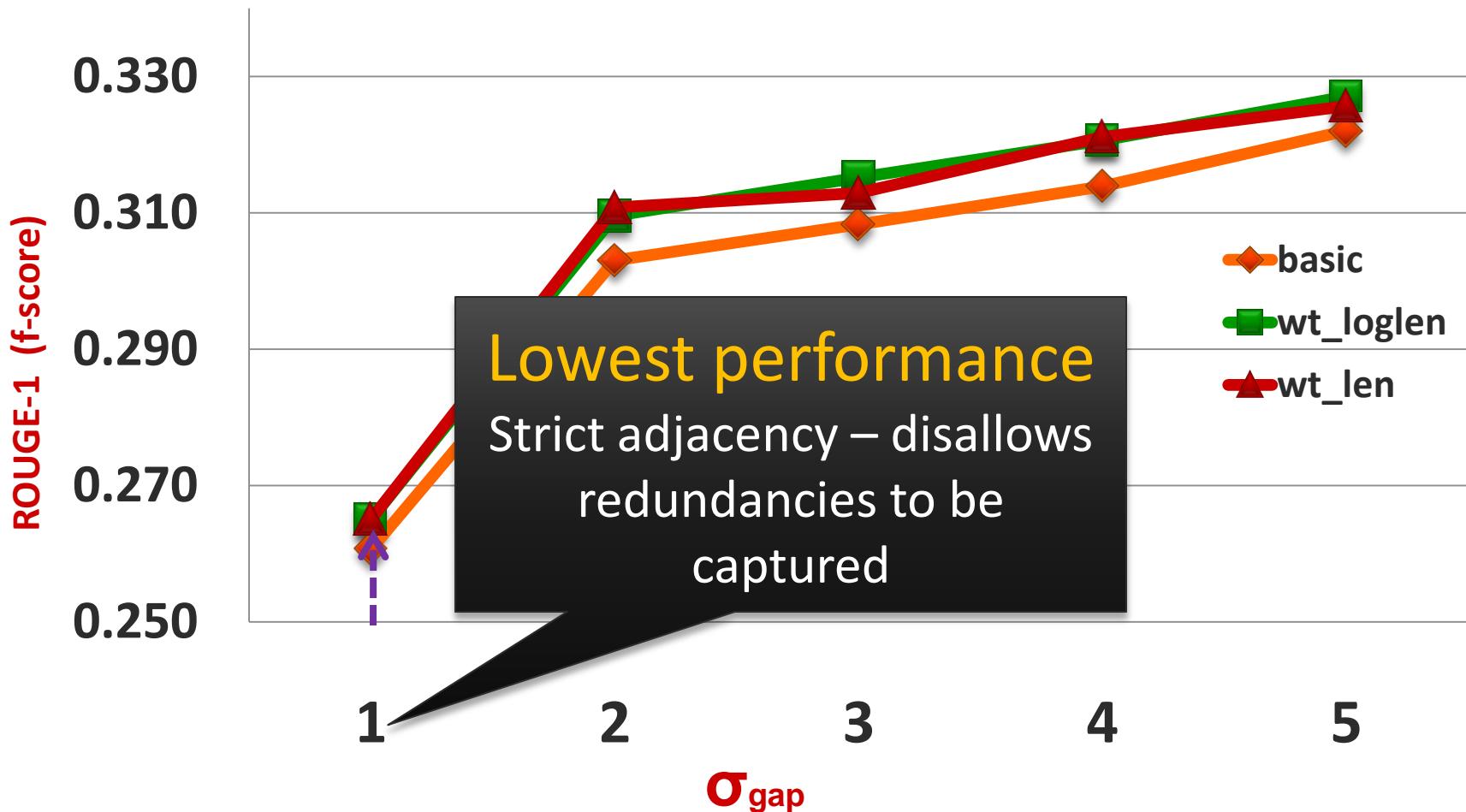
Performance of Opinosis **similar**
to Human performance

Effect of Gap Threshold (σ_{gap})

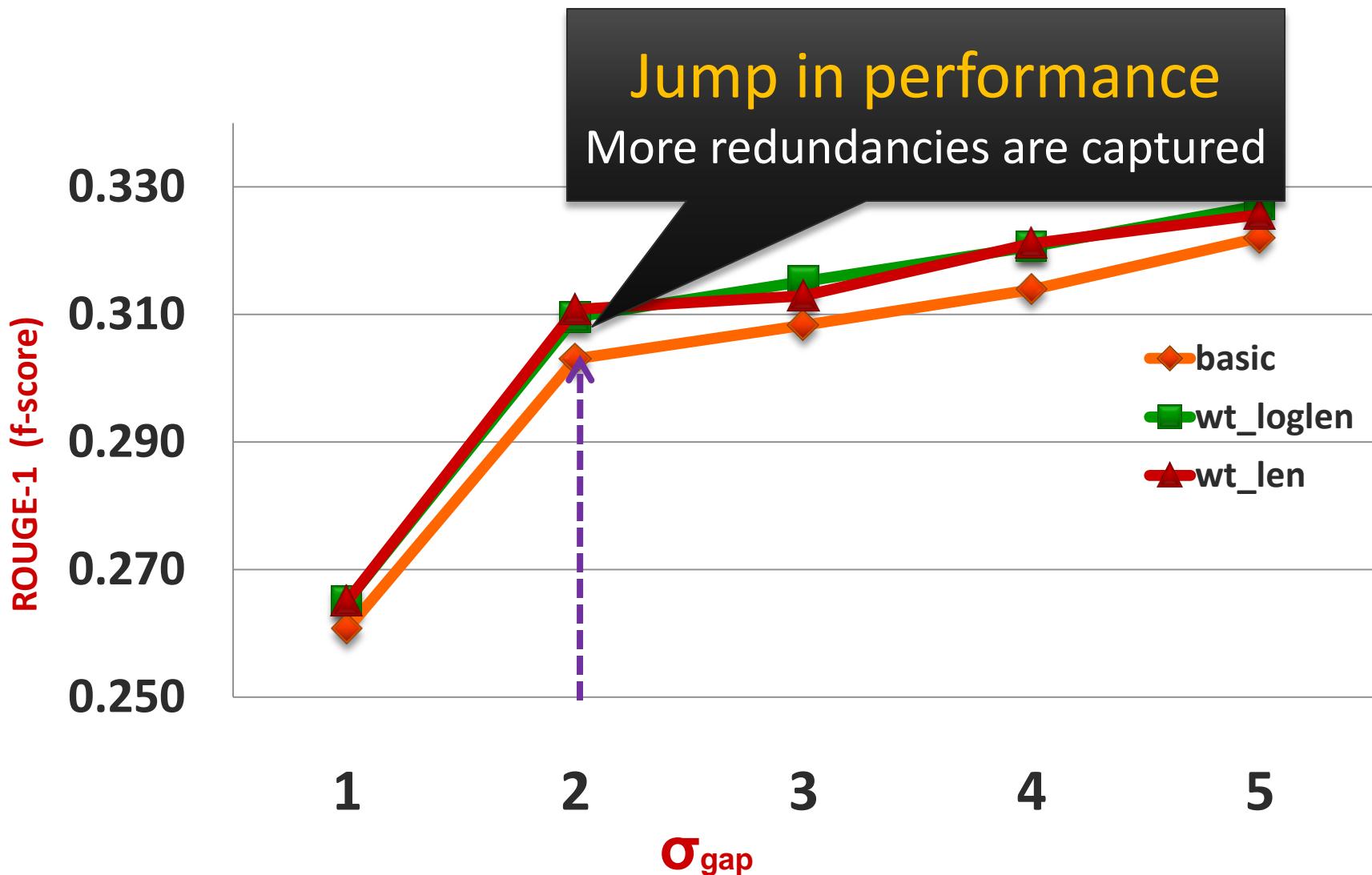
Effect of Gap Threshold (σ_{gap})



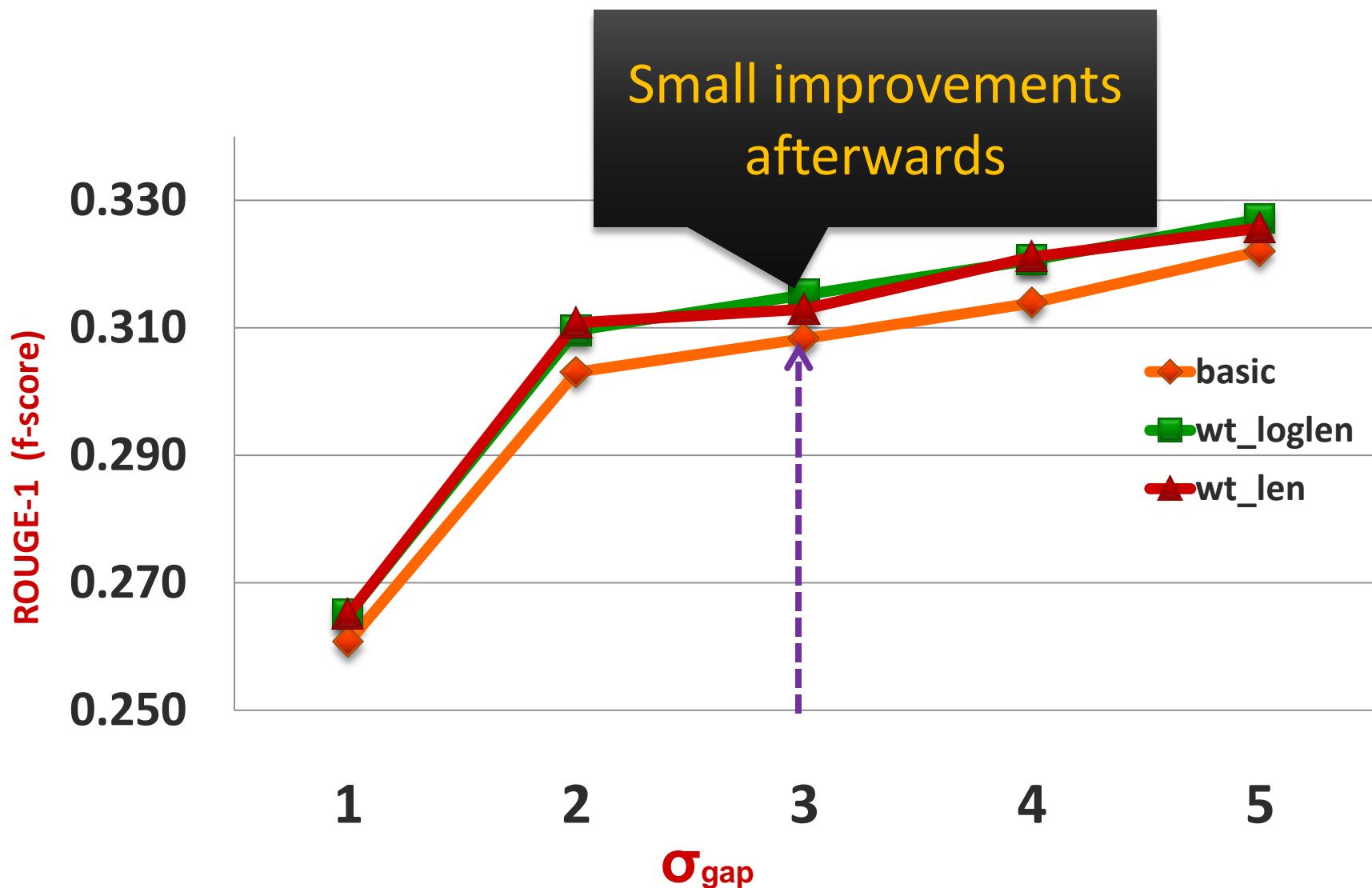
Effect of Gap Threshold (σ_{gap})



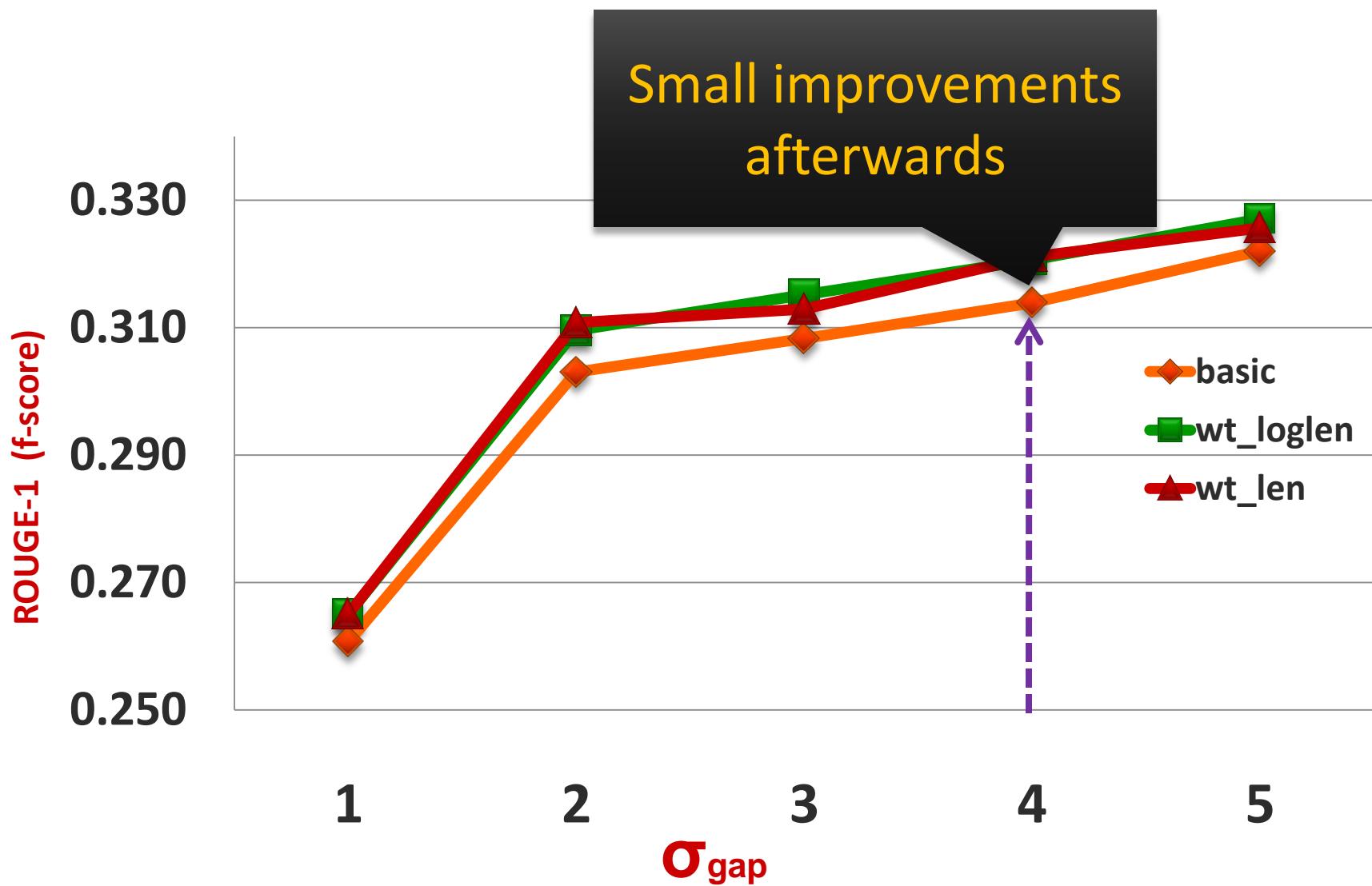
Effect of Gap Threshold (σ_{gap})



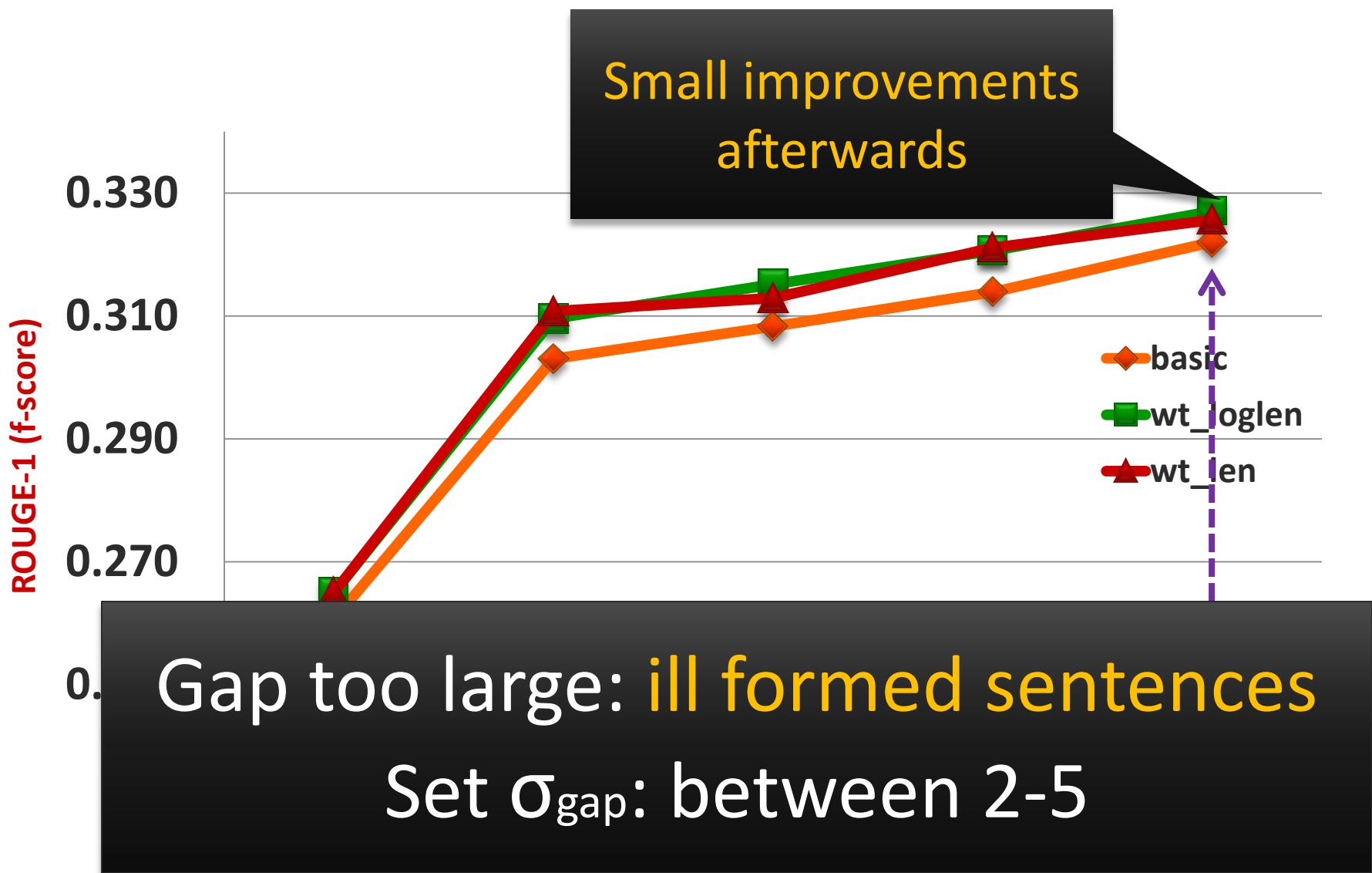
Effect of Gap Threshold (σ_{gap})



Effect of Gap Threshold (σ_{gap})

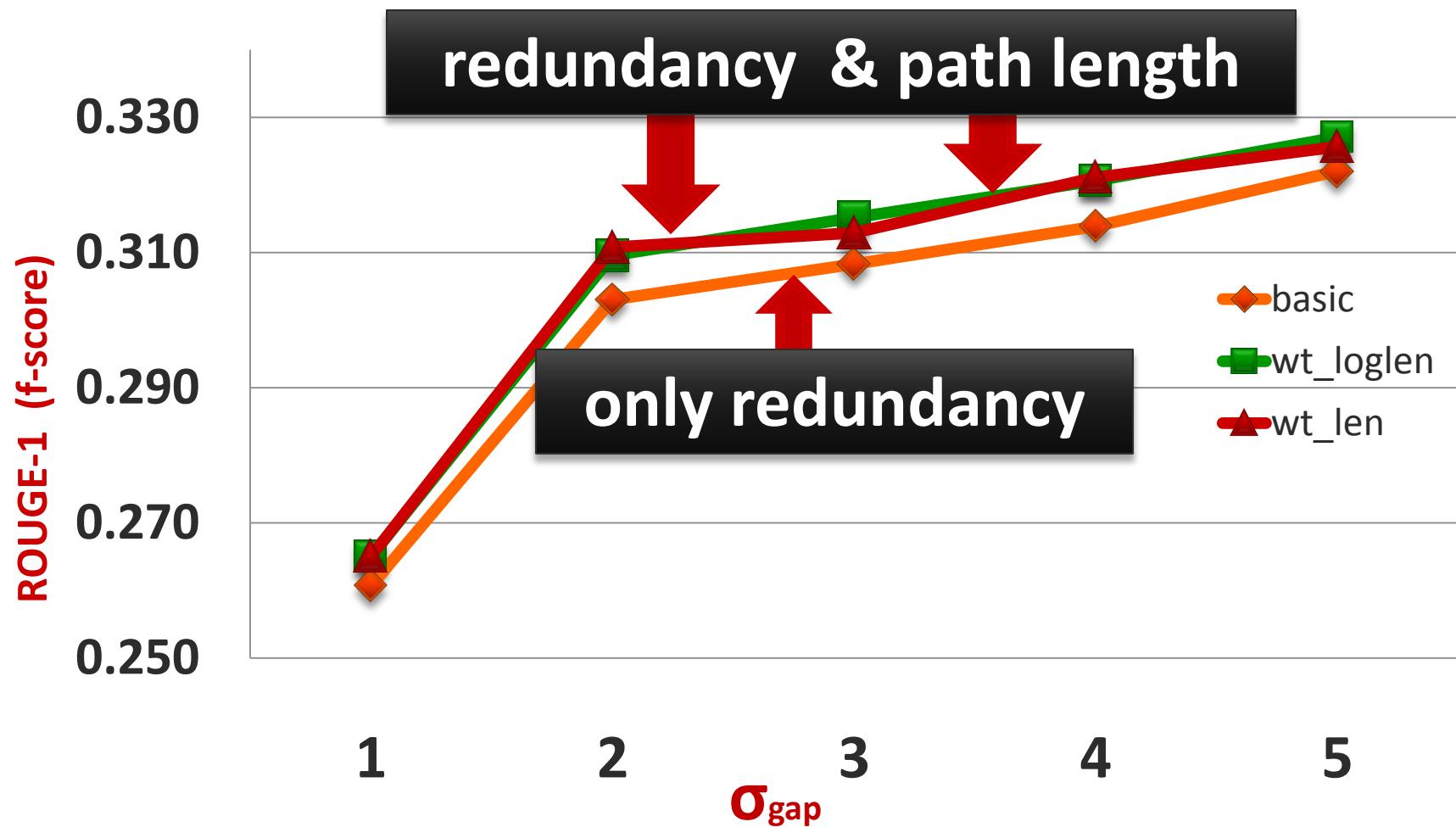


Effect of Gap Threshold (σ_{gap})

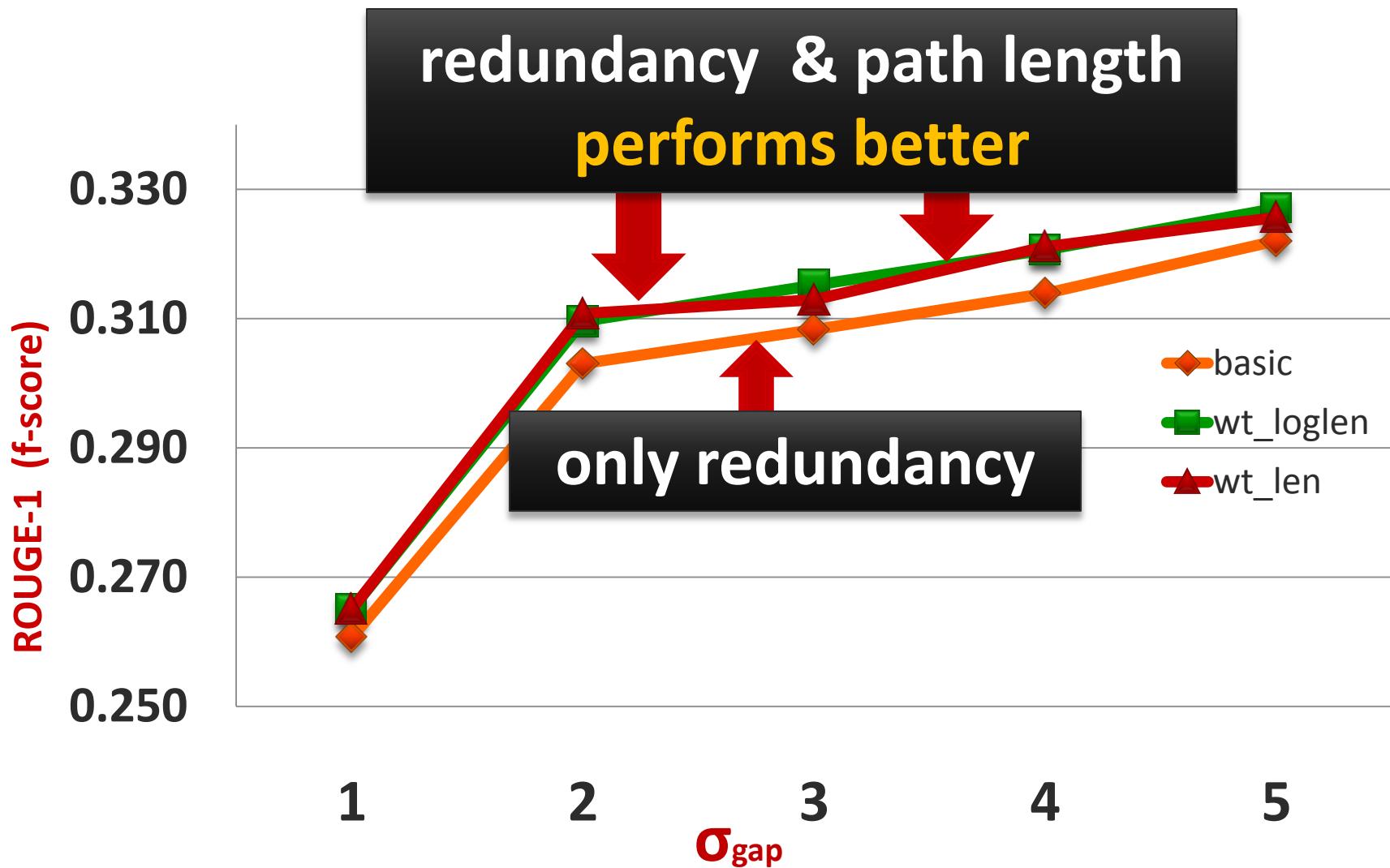


Compare: Scoring Functions

Compare: Scoring Functions



Compare: Scoring Functions



Readability Test

How Readability Test Works

Topic X

Opinosis Summary

1. sentence 1.....
2. sentence 2.....

Human Summary 1

1. sentence 1.....
2. sentence 2.....
3. sentence X..

Human Summary M

1. sentence 1.....
2. sentence 2.....
3. sentence X.....



Topic X

Mixed Sentences

1. sentence 1.....
 2. sentence 3.....
 3. sentence 2.....
 4. sentence 4.....
 5. sentence 8.....
 6. sentence 6.....
 7. sentence 7.....
 8. sentence 5.....
-
.....

**Pick at most 2 sentences
that are least readable**

Readability Test

- ▶ Assessor often picks **Opinosis** sentences:
 - Opinosis summaries have issues with readability
- ▶ Assessor often picks **wrong sentences or makes no picks**:
 - Opinosis summaries similar to human summaries

Our Readability Test

- ▶ Use summaries from Opinosis^{best}
- ▶ 565 sentences in total
 - 102 sentences from Opinosis summaries
 - 463 sentences from human summaries
- ▶ > 60% of Opinosis generated sentences are no different from human composed sentences

Summary

- ▶ A framework for summarizing **highly redundant opinions**
- ▶ **Use graph representation** to generate concise abstractive summaries
- ▶ **General & lightweight:** Can be used on any corpus with high redundancies (Twitter comments, Blog comments, etc)

Thank You...



ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

TIMan