

# Constraint-based Sentence Compression An Integer Programming Approach

James Clarke (jclarke@ed.ac.uk)  
School of Informatics, University of Edinburgh



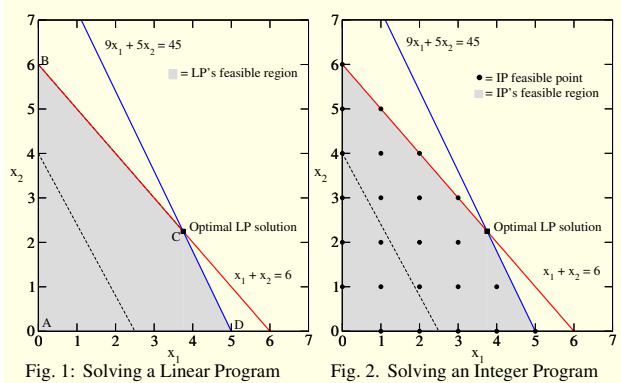
## What is Sentence Compression?

- Word deletion:** given a sentence of words  $W = w_1, w_2, \dots, w_n$ , form a compression by removing any subset of words (Knight and Marcu, 2000).
- Useful for applications:** text summarisation, subtitle generation from spoken transcripts, information retrieval.



## Why Integer Programming?

- Model sentence compression as an **optimisation problem**: search over  $2^n$  possible solutions.
- The decoding is **independent** of the underlying model.
- Linear programming (LP) finds the **global** minimum or maximum value of a linear **objective function** given some **constraints**.
- Integer Programming (IP) is an extension: all decision variables must be non-negative integers.



## Problem Formulation

- Use a **trigram language model** as objective function.
- Add a set of constraints to ensure the compressions are **structurally** and **semantically** valid.
- Add a **significance score** to help retain **important** content words.

### 1. Language Model

$$\max z = \sum_{i=1}^n p_i \cdot P(w_i | \text{start}) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n x_{ijk} \cdot P(w_k | w_i, w_j) + \sum_{i=0}^{n-1} \sum_{j=i+1}^n q_{ij} \cdot P(\text{end} | w_i, w_j) + \sum_{i=1}^n y_i \cdot I(w_i)$$

- Decision variables take **0** or **1** values:  

$$y_i = \begin{cases} 1 & \text{if } w_i \text{ is in the compression} \\ 0 & \text{otherwise} \end{cases} \quad (1 \leq i \leq n)$$
- $p_i$ ,  $q_{ij}$ , and  $x_{ijk}$  represent word sequences which start, end and appear in the sentence.
- Objective function is **sum of all possible trigrams** that can occur in **all compressions** of the original sentence.
- Constraints** ensure that trigrams are combined in a **valid manner**.
- Optional **significance score**  $I(w_i)$  retains **important** content words.

He became a power player in Greek Politics, when he founded the socialist Pasok Party.  
 He became a player in the Pasok.  
 Finally, AppleShare Printer Server, formerly a separate package, is now bundled with AppleShare File Server.  
 Finally, AppleShare, a separate, AppleShare.

Tab. 1: Compressions with language model (no significance score).

### 2. Modifier Constraints

- Ensure the **relationships** between **head** words and their **modifiers** remain grammatical.
- If a modifier is in the compression, its head word must be included:  

$$y_i - y_j \geq 0 \quad \forall i, j : w_j \in w_i \text{'s modifiers}$$
- Do not drop **not** if the head word is in the compression (same for words like *his*, *our* and genitives).

He became a power player in Greek Politics, when he founded the socialist Pasok Party.  
 He became a player in the Pasok **Party**.  
 Finally, AppleShare Printer Server, formerly a separate package, is now bundled with AppleShare File Server.  
 Finally, AppleShare **Server, is bundled**.

Tab 2: Compressions with modifier constraints (no significance score).

### 3. Sentential Constraints

- Take the **overall sentence** structure into account.
- If a verb is in the compression then so are its arguments, and vice-versa:  

$$y_i - y_j = 0 \quad \forall i, j : w_j \in \text{subject/object of verb } w_i$$

- The compression must contain **at least one verb**.

He became a power player in Greek Politics, when he founded the socialist Pasok Party.  
 He became a player in **politics**.  
 Finally, AppleShare Printer Server, formerly a separate package, is now bundled with AppleShare File Server.  
 Finally, AppleShare **Server, is bundled with Server**.

Tab. 3: Compressions with sentential constraints (no significance score).

### 4. Significance Score

- Language model does not know which **content words** to keep.
- Language model prefers words it has encountered **before**.
- Significance score** gives more weight to words **deeply embedded** in the syntactic tree (inspired by Hori and Furui, 2004).

He became a power player in Greek Politics, when he founded the socialist Pasok Party.  
 He became a player in politics **when he founded the Pasok Party**.  
 Finally, AppleShare Printer Server, formerly a separate package, is now bundled with AppleShare File Server.  
 AppleShare **Printer Server package is now bundled with AppleShare File Server**.

Tab. 4: Compressions with significance score.

## Results

- Compare language model with and without significance against state-of-the-art decision-tree model and human gold standard compressions.
- Evaluation on **40 sentences** (20 Ziff-Davis and 20 Broadcast News).
- Fifty-six** unpaid volunteers rate the compressions on a **five point** scale.

Model	CompR	Rating
Decision-tree	56.1%	2.22 <sup>*†</sup>
Language Model	49.0%	2.23 <sup>*†</sup>
Language Model+Significance	73.6%	2.83 <sup>*</sup>
Gold Standard	62.3%	3.68 <sup>†</sup>

Tab. 5: Compression results; compression rate (CompR) and average human judgements (Rating); \*: sig. diff. from gold standard; †: sig. diff. from Language Model+Significance.

## Summary

- We use integer programming to infer **globally optimal** compressions in the presence of **linguistic constraints**.
- Relatively simple and **knowledge-lean** compression model.
- Comparable compressions to state-of-the-art **without any supervision**.
- More details in** Clarke and Lapata (2006a, 2006b).