



HAL
open science

Intertwining NLP and CP-AI-OR Reasoning for Constrained Text Generation

Alexandre Bonlarron, Aurelie Calabrese, Pierre Kornprobst, Jean-Charles
Régis

► **To cite this version:**

Alexandre Bonlarron, Aurelie Calabrese, Pierre Kornprobst, Jean-Charles Régis. Intertwining NLP and CP-AI-OR Reasoning for Constrained Text Generation. Journée Natural Language Argumentation – GDR TAL – GDR RADIA, Nov 2023, Sophia Antipolis (06), France. . hal-04296814

HAL Id: hal-04296814

<https://hal.science/hal-04296814>

Submitted on 21 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Intertwining NLP and CP-AI-OR Reasoning for Constrained Text Generation

A. Bonlarron^(1,3), A. Calabrèse⁽²⁾, P. Kornprobst⁽¹⁾, J.-C. Régim⁽³⁾

alexandre.bonlarron@inria.fr



IJCAI/2023 MACAO



(1) Université Côte d'Azur, Inria, France
 (2) Aix Marseille Univ, CNRS, LPC, Marseille, France
 (3) Université Côte d'Azur, I3S, France

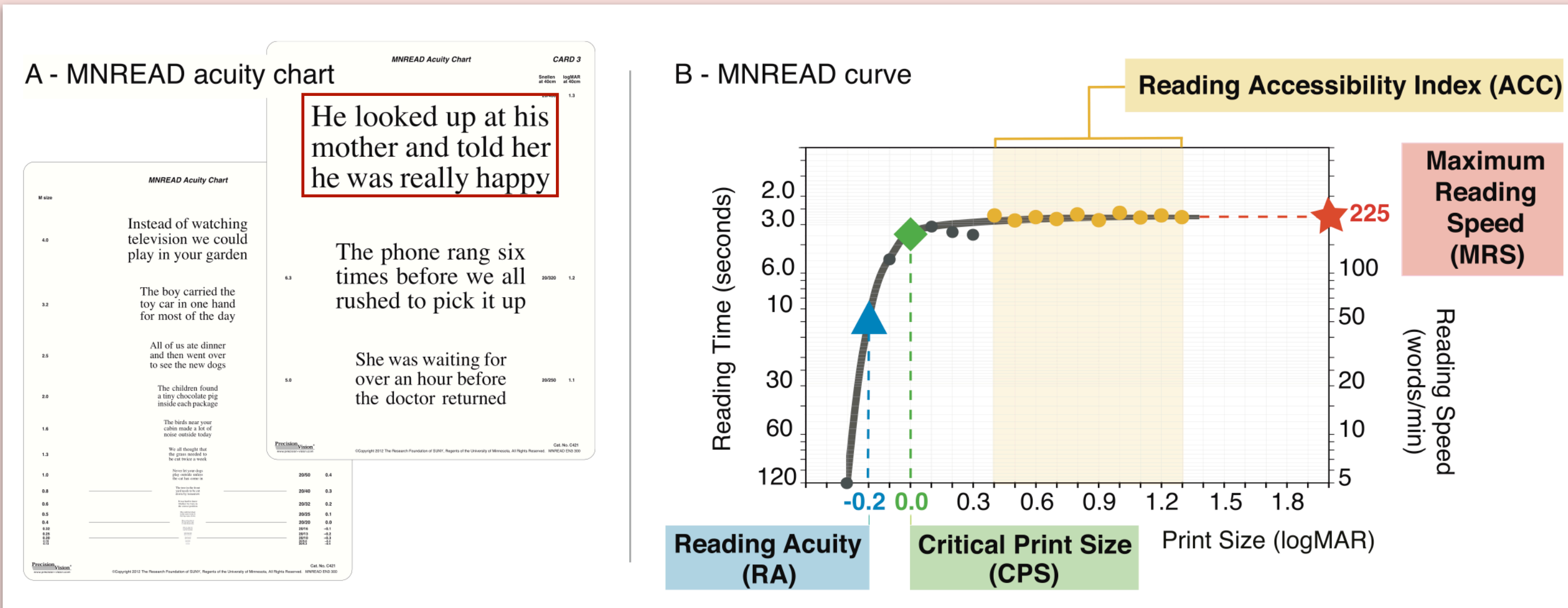


General Idea

Beam Search fails to generate constrained text when the solutions space is **strongly constrained** (e.g., length, display). When constraints are defined on the **whole text**. Our **CP-based** approach solves it.

Context

Case study: the MNREAD test



A psychophysical test based on standardised sentences, available in 19 languages

- Assessing the **reading speed** is crucial for evaluating patients with visual impairments
- The need for a collection of **standardized** sentences respecting grammar, lexicon, size, or display rules
- Limitation: **few sentences** available due to the rules being too restrictive (**only 38 MNREAD sentences in French...**)
- Problem: How can we **generate** sentences that comply with several rules?

Methods

Principle (n-gram)

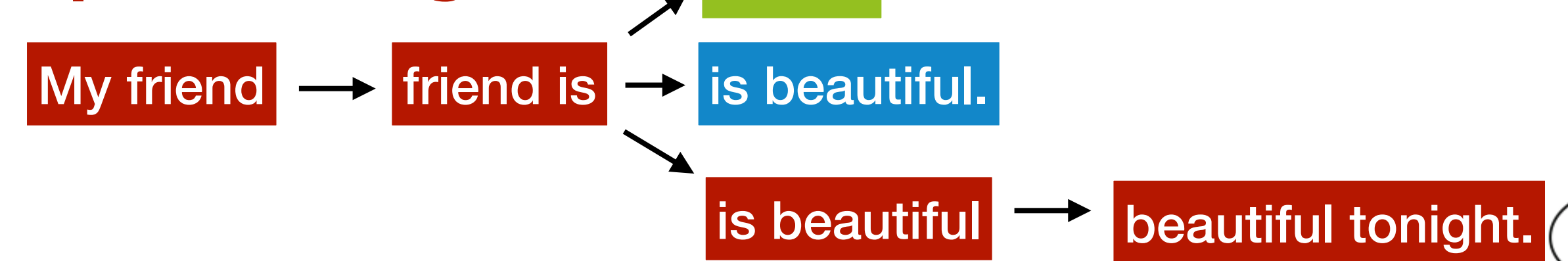
Step 1: Corpus

He is angry.
 My friend is sad.
 She is beautiful tonight.

Step 2: set of 2-grams

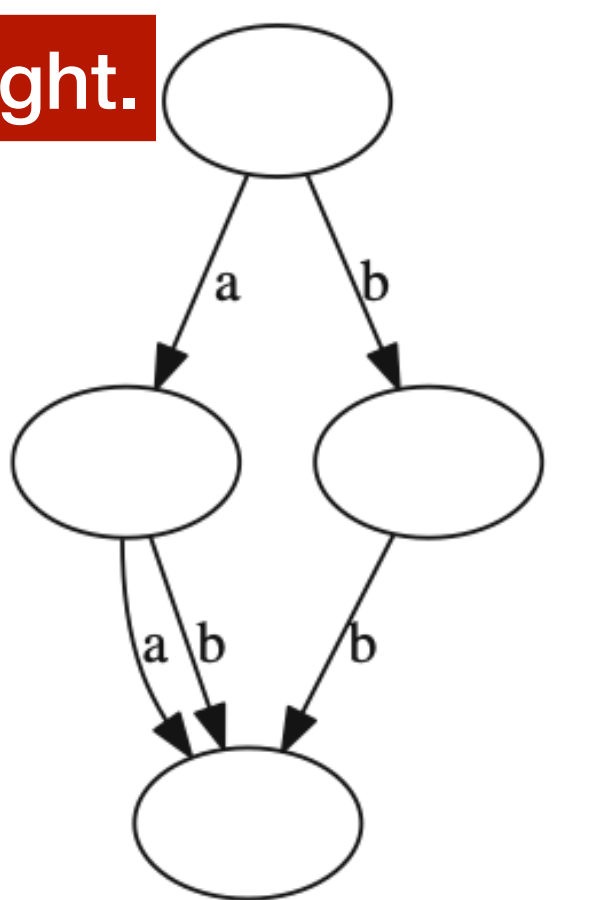
Is beautiful friend is Is beautiful He is
 My friend She is beautiful tonight Is sad.

Step 3: chaining



MDD Paradigm (Constraint Programming)

- Data structures for calculating and storing the solutions to a problem in a compressed form using an acyclic directed graph.
- Advantage: powerful modelling tool, computing solutions set without a search.

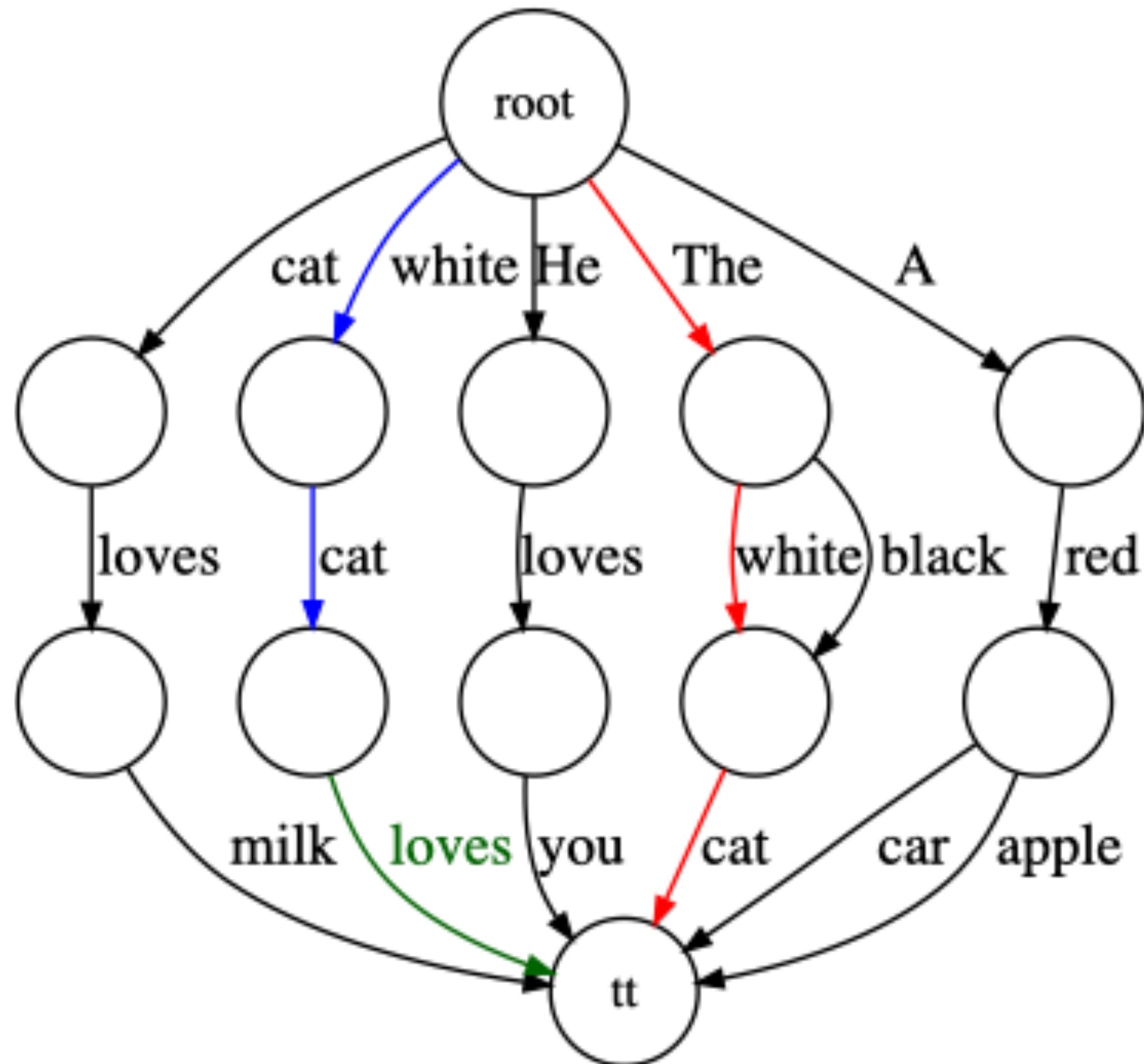


Large Language Model (LLM) can be used to score generated sentences. This score is viewed as a quality measurement of the generated sentences.

Results

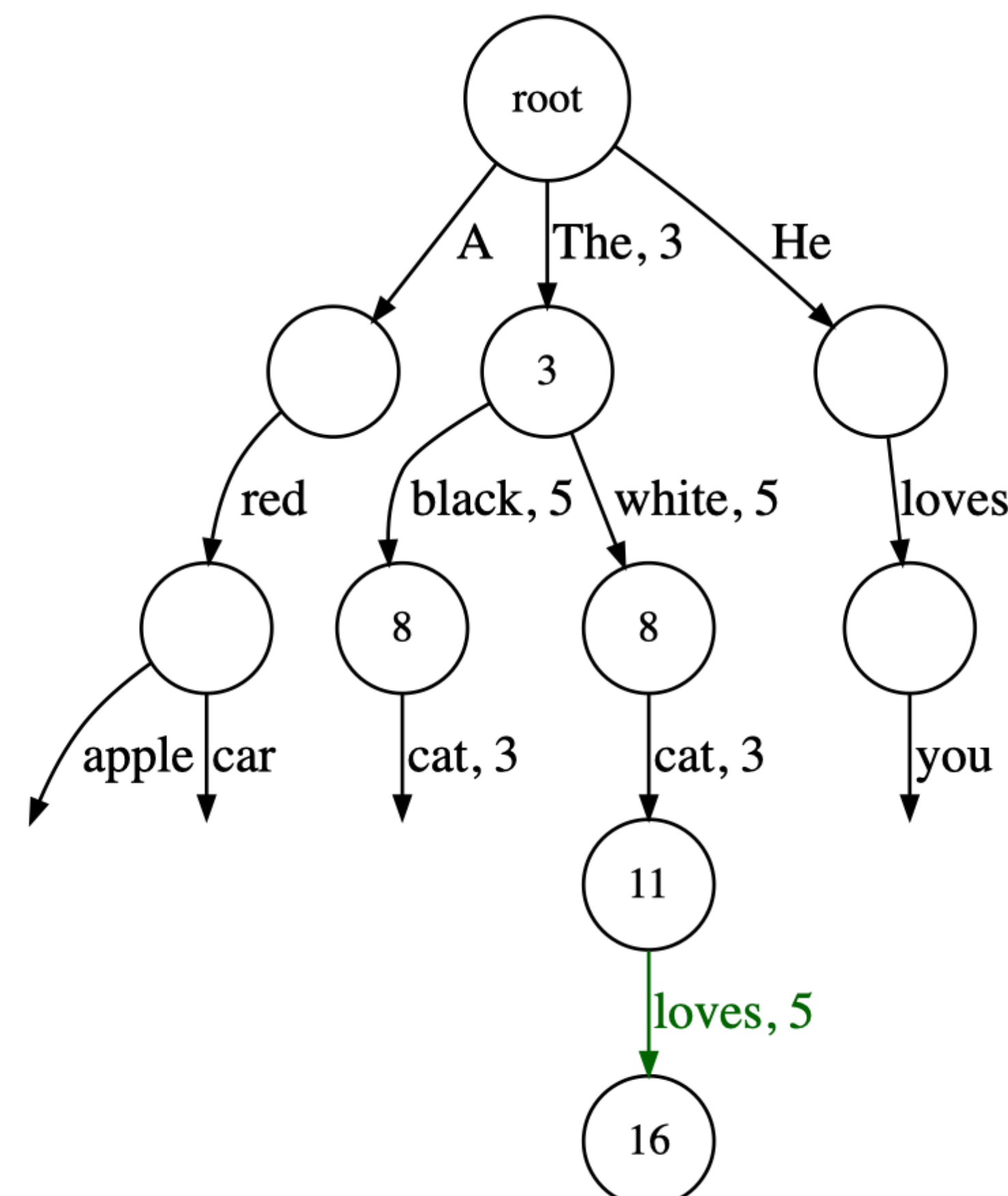
From MNREAD rules to Sentences Generation in 3 steps

1 - MDD as a Trie (prefix tree)



➔ What is the successor of **The white cat** ?

2 - MDD top-down compilation



3 - Sentences scoring with LLM

$$P(X_1 \dots X_n) = \prod_{i=1}^n P(X_i | X_1 \dots X_{i-1})$$

$$PPL(X_1 \dots X_n) = \sqrt[n]{\frac{1}{P(X_1 \dots X_n)}}$$

$$PPL(X_1 \dots X_n) = \sqrt[n]{\frac{1}{\prod_{i=1}^n P(X_i | X_1 \dots X_{i-1})}}$$

- 'Perplexity' is derived from Shannon's Entropy.
- It quantifies the uncertainty of a model with respect to a sample
- The lower, the better; range is [1 ; + inf]

General observation

- n of n-grams ↗
- Quality of sentences ↗
- Number of sentences ↘

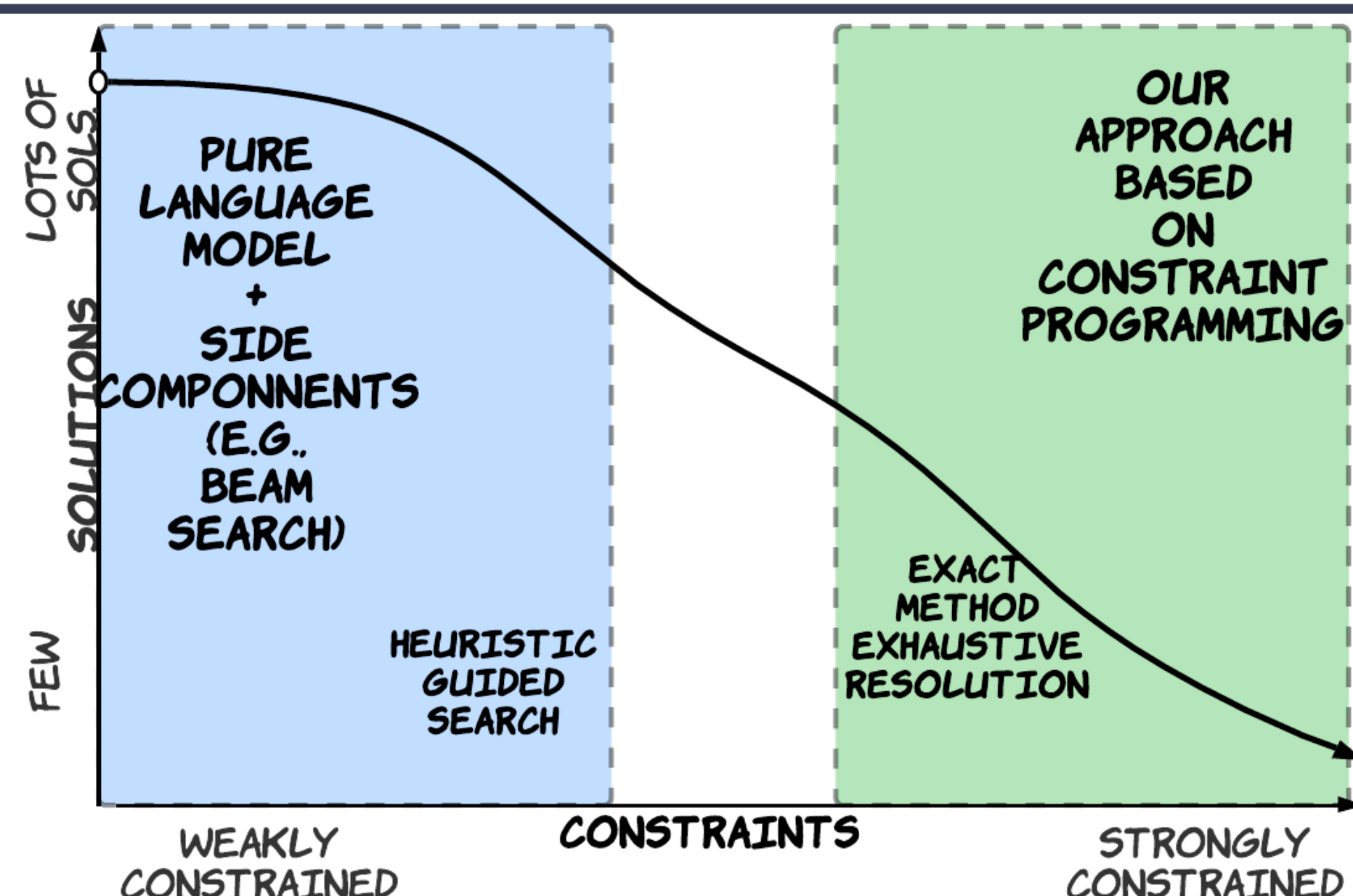
MNREAD sentence generation

	FR	EN	RAM	Time	Number of sentences
FR	443	75	3GB	72s	7028
EN	75	<<1GB	3s	204	

Examples of sentences

Quality: Good	Quality: Average	Quality: Poor
They are one of the most famous in the history of the sport	He stepped in front of the double doors and into the branch	So strange to be on the edge of my chin between his fingers
PPL = 12	PPL = 160	PPL = 224

CP-ML



Conclusion

- Promising method: more suitable than generic methods (e.g. GPT, Bert) for **managing constraints** and more **flexible** than the ad-hoc method of Mansfield et al.
- Advantages: **modularity** (easy to add and/or to remove rules), **constraints taken into account at the generation stage**, can be applied to **other languages**
- Perspectives: Bridging CP and ML

For more details, read the paper →

