Grammar Reinforcement Learning: path and cycle counting in graphs with a Context- Free Grammar and Transformer approach

Jason Piquenot

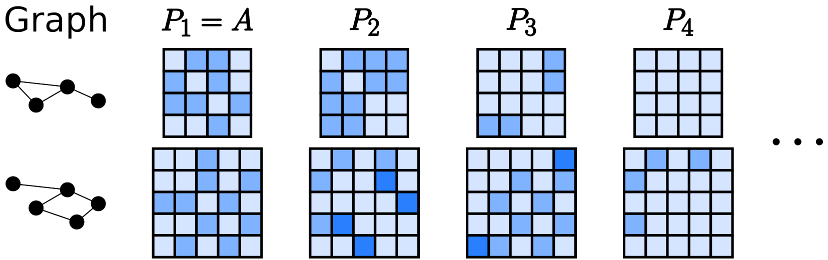
Supervisors : Sébastien Adam Romain Raveaux
Co-supervisors : Pierre Héroux Jean-Yves Ramel Maxime Bérar

(ANR-21-CE-0025)

Path matrix

l-path matrix

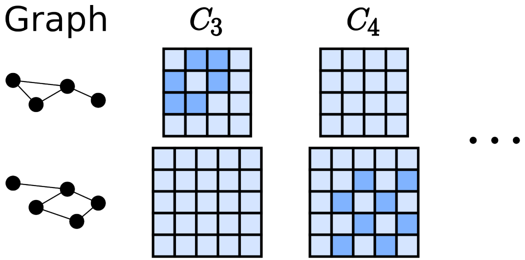
For any non-negative integer l, let P_l represent the l-path matrix where $(P_l)_{i,j}$ is the number of l-length paths connecting vertex i to vertex j.



Cycle matrix

l-cycle matrix

For $l > 2$, let C_l represent the l-cycle matrix where $(C_l)_{i,j}$ denotes the number of l-cycles connecting vertex i to its adjacent vertex j.

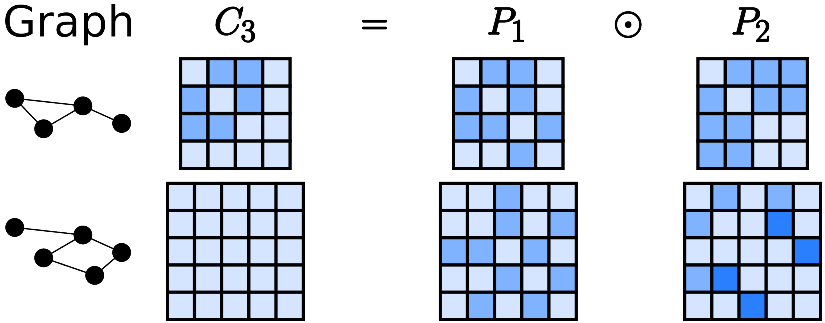


Link between cycle and path matrices

Path-cycle formula

For $l > 2$, we have from [Voropaev and Perepechko, 2012]:

$$C_{l+1} = P_1 \odot P_l$$

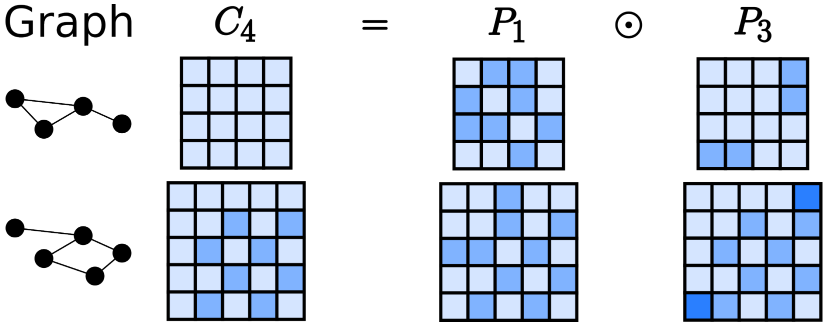


Link between cycle and path matrices

Path-cycle formula

For $l > 2$, we have from [Voropaev and Perepechko, 2012]:

$$C_{l+1} = P_1 \odot P_l$$



State of the art formulae

[Voropaev and Perepechko, 2012] introduced the most efficient formulae to date for calculating P_2 to P_6.

State of the art formulae

[Voropaev and Perepechko, 2012] introduced the most efficient formulae to date for calculating P_2 to P_6.

$P_2 = \mathrm{J} \odot A^2$

State of the art formulae

[Voropaev and Perepechko, 2012] introduced the most efficient formulae to date for calculating P_2 to P_6.

$$P_2 = \mathrm{J} \odot A^2$$

$$P_3 = \mathrm{J} \odot A^3 - (\mathrm{I} \odot A^2)A - A(\mathrm{I} \odot A^2) + A$$

State of the art formulae

[Voropaev and Perepechko, 2012] introduced the most efficient formulae to date for calculating P_2 to P_6.

$$P_2 = \mathrm{J} \odot A^2$$

$$P_3 = \mathrm{J} \odot A^3 - (\mathrm{I} \odot A^2)A - A(\mathrm{I} \odot A^2) + A$$

$$\begin{aligned} P_4 = {} & \mathrm{J} \odot A^4 - \mathrm{J} \odot (A(\mathrm{I} \odot A^2)A) + 2(\mathrm{J} \odot A^2) \\ & - (\mathrm{I} \odot A^2)(\mathrm{J} \odot A^2) - (\mathrm{J} \odot A^2)(\mathrm{I} \odot A^2) \\ & - A(\mathrm{I} \odot A^3) - (\mathrm{I} \odot A^3)A + 3A \odot A^2 \end{aligned}$$

State of the art formulae

[Voropaev and Perepechko, 2012] introduced the most efficient formulae to date for calculating P_2 to P_6.

$$
\begin{aligned}
P_2 &= \mathrm{J} \odot A^2 \\
P_3 &= \mathrm{J} \odot A^3 - (\mathrm{I} \odot A^2)A - A(\mathrm{I} \odot A^2) + A \\
P_4 &= \mathrm{J} \odot A^4 - \mathrm{J} \odot (A(\mathrm{I} \odot A^2)A) + 2(\mathrm{J} \odot A^2) \\
&\quad - (\mathrm{I} \odot A^2)(\mathrm{J} \odot A^2) - (\mathrm{J} \odot A^2)(\mathrm{I} \odot A^2) \\
&\quad - A(\mathrm{I} \odot A^3) - (\mathrm{I} \odot A^3)A + 3A \odot A^2 \\
P_5 &= \mathrm{J} \odot A^5 - (\mathrm{I} \odot A^4)A - A(\mathrm{I} \odot A^4) - (\mathrm{I} \odot A^3)(\mathrm{J} \odot A^2) - (\mathrm{J} \odot A^2)(\mathrm{I} \odot A^3) \\
&\quad - (\mathrm{I} \odot A^2)(\mathrm{J} \odot A^3) - (\mathrm{J} \odot A^3)(\mathrm{I} \odot A^2) - \mathrm{J} \odot (A(\mathrm{I} \odot A^3)A) + 3A \odot A^3 \\
&\quad + 2A(\mathrm{I} \odot A^2)(\mathrm{I} \odot A^2) + 2(\mathrm{I} \odot A^2)(\mathrm{I} \odot A^2)A + 3A \odot A^2 \odot A^2 + (\mathrm{I} \odot A^2)A(\mathrm{I} \odot A^2) \\
&\quad - \mathrm{J} \odot (A(\mathrm{I} \odot A^2)A^2) - \mathrm{J} \odot (A^2(\mathrm{I} \odot A^2)A) + 3\mathrm{J} \odot ((A \odot A^2)A) + 3\mathrm{J} \odot (A(A \odot A^2)) \\
&\quad + (\mathrm{I} \odot (A(\mathrm{I} \odot A^2)A))A + A(\mathrm{I} \odot (A(\mathrm{I} \odot A^2)A)) - 6(\mathrm{I} \odot A^2)A - 6A(\mathrm{I} \odot A^2) \\
&\quad - 4A \odot A^2 + 3\mathrm{J} \odot A^3 + 4A.
\end{aligned}
$$

A CFG that encompasses them all

G_3

$$M \to (M \odot M) \mid (MM) \mid A \mid \mathrm{I} \mid \mathrm{J}.$$

A CFG that encompasses them all

G_3

$$M \to (M \odot M) \mid (MM) \mid A \mid \mathrm{I} \mid \mathrm{J}.$$

P_l and $L(G_3)$

[Voropaev and Perepechko, 2012] are linear combination of sentence in $L(G_3)$.

A CFG that encompasses them all

G_3

$$M \to (M \odot M) \mid (MM) \mid A \mid \mathrm{I} \mid \mathrm{J}.$$

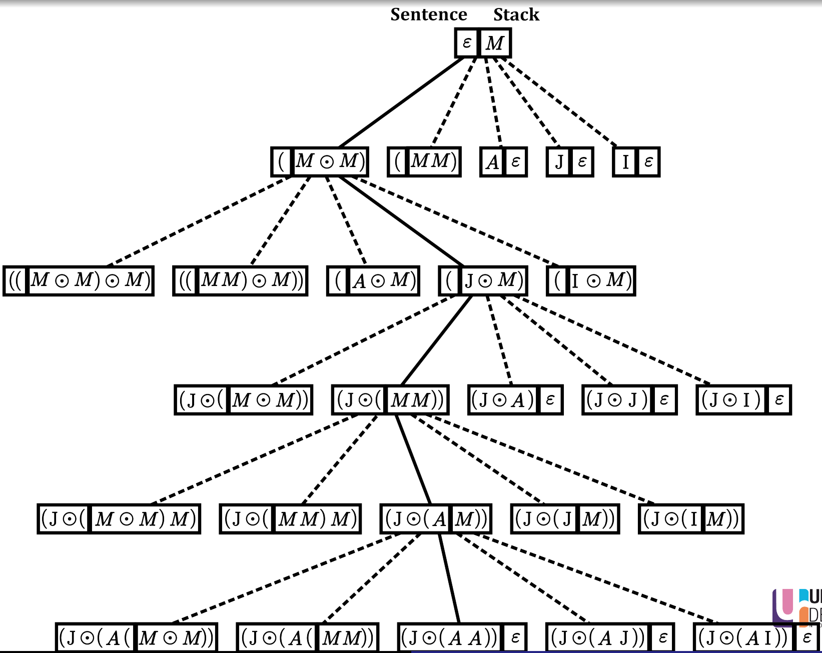
P_l and $L(G_3)$

[Voropaev and Perepechko, 2012] are linear combination of sentence in $L(G_3)$.

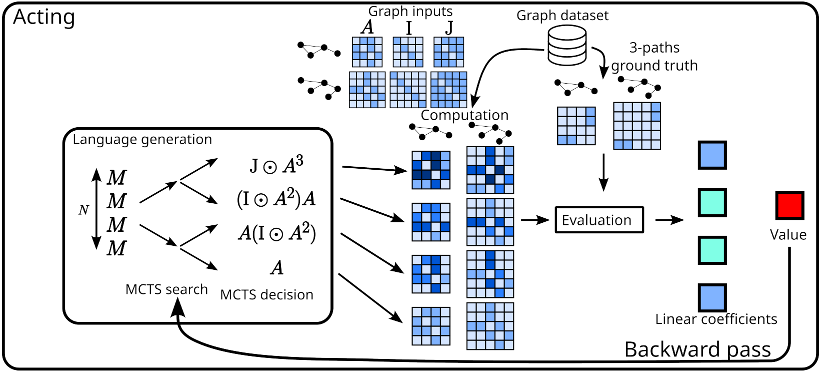
$\tilde{G}_3$

$$E \to (E \odot M) \mid (NE) \mid (EN) \mid A \mid \mathrm{J}$$
$$N \to (N \odot M) \mid (N \odot N) \mid \mathrm{I}$$
$$M \to (MM) \mid (EE)$$

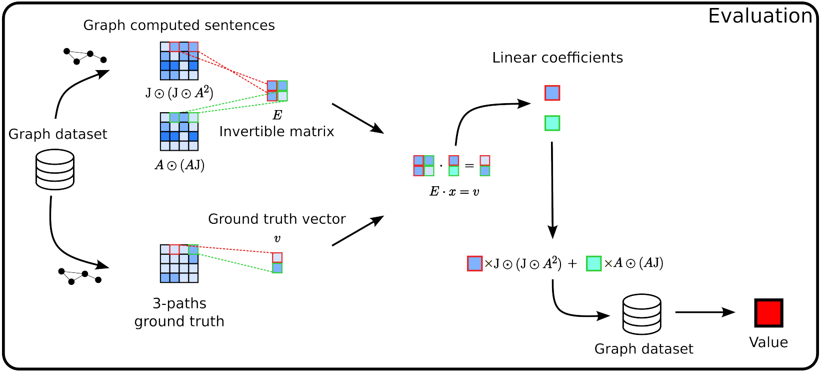
Tree structure of CFG



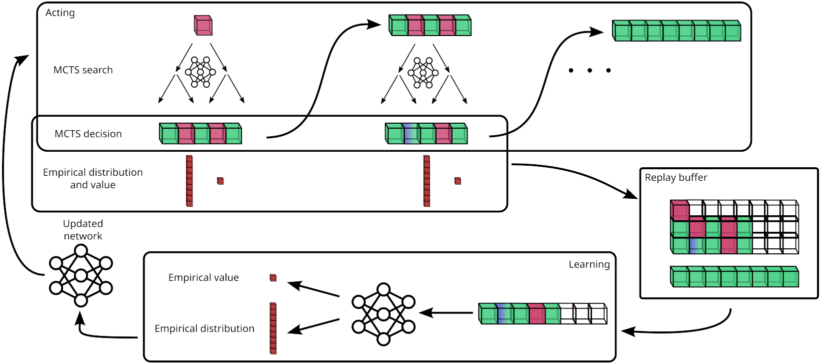
MCTS Search within grammar



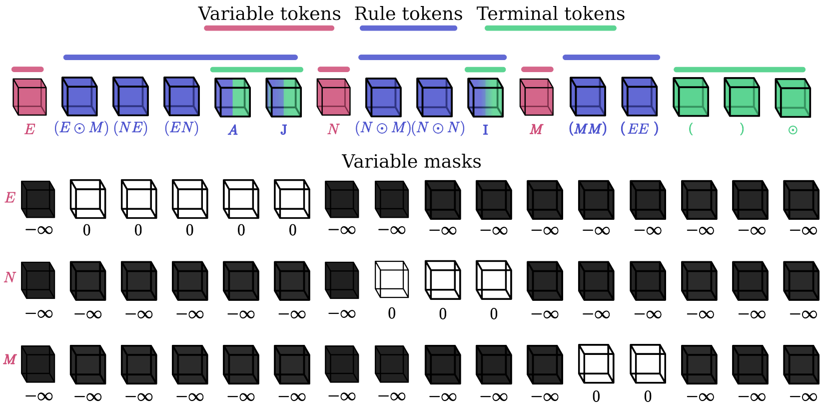
Focus on the evaluation



Grammar Reinforcement Learning

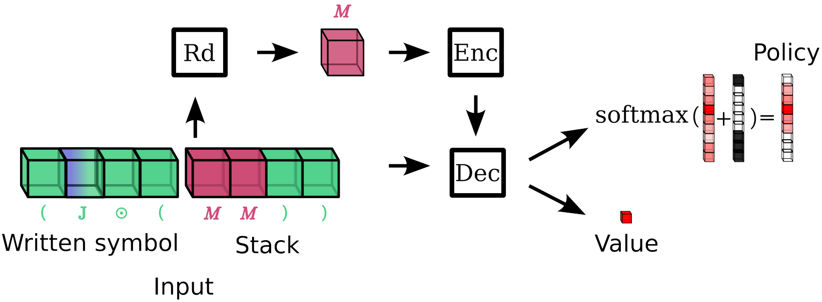


Tokenisation of grammar

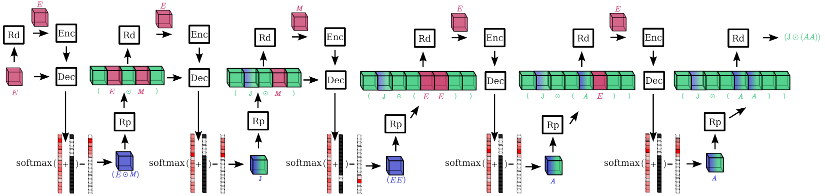


$$E \rightarrow (E \odot M) \mid (NE) \mid (EN) \mid A \mid \mathrm{J}$$
$$N \rightarrow (N \odot M) \mid (N \odot N) \mid \mathrm{I}$$
$$M \rightarrow (MM) \mid (EE)$$

Gramformer in GRL



Gramformer in autoregressive sentence generation



Formulae

GRL applied on $\tilde{G}_3$:

Formulae

GRL applied on $\tilde{G}_3$:

$P_2 = \mathrm{J} \odot A^2$

Formulae

GRL applied on $\tilde{G}_3$:

$P_2^* = \mathrm{J} \odot A^2$

Formulae

GRL applied on $\tilde{G}_3$:

$$P_2^* = \mathrm{J} \odot A^2$$

$$P_3 = \mathrm{J} \odot A^3 - (\mathrm{I} \odot A^2)A - A(\mathrm{I} \odot A^2) + A$$

Formulae

GRL applied on $\tilde{G}_3$:

$$P_2^* = \mathrm{J} \odot A^2$$
$$P_3^* = \mathrm{J} \odot (A(\mathrm{J} \odot A^2)) - A(\mathrm{I} \odot A^2) + A$$

Formulae

GRL applied on $\tilde{G}_3$:

$$P_2^* = \mathrm{J} \odot A^2$$
$$P_3^* = \mathrm{J} \odot (A(\mathrm{J} \odot A^2)) - A \odot (A\mathrm{J})$$

Formulae

GRL applied on $\tilde{G}_3$:

$$
\begin{aligned}
P_2^* &= \mathrm{J} \odot A^2 \\
P_3^* &= \mathrm{J} \odot (A(\mathrm{J} \odot A^2)) - A \odot (A\mathrm{J}) \\
P_4 &= \mathrm{J} \odot A^4 - \mathrm{J} \odot (A(\mathrm{I} \odot A^2)A) + 2(\mathrm{J} \odot A^2) \\
&\quad - (\mathrm{I} \odot A^2)(\mathrm{J} \odot A^2) - (\mathrm{J} \odot A^2)(\mathrm{I} \odot A^2) \\
&\quad - A(\mathrm{I} \odot A^3) - (\mathrm{I} \odot A^3)A + 3A \odot A^2
\end{aligned}
$$

Formulae

GRL applied on $\tilde{G}_3$:

$$
\begin{aligned}
P_2^* &= \mathrm{J} \odot A^2 \\
P_3^* &= \mathrm{J} \odot (A(\mathrm{J} \odot A^2)) - A \odot (A\mathrm{J}) \\
P_4^* &= \mathrm{J} \odot (A(\mathrm{J} \odot (A(\mathrm{J} \odot A^2)))) - \mathrm{J} \odot (A(A \odot (A\mathrm{J}))) \\
&\quad - \mathrm{J} \odot ((A \odot (A\mathrm{J}))A) - A \odot ((A \odot A^2)\mathrm{J}) + 2A \odot A^2
\end{aligned}
$$

Formulae

GRL applied on $\tilde{G}_3$:

$$
\begin{aligned}
P_2^* &= \mathrm{J} \odot A^2 \\
P_3^* &= \mathrm{J} \odot (A(\mathrm{J} \odot A^2)) - A \odot (A\mathrm{J}) \\
P_4^* &= \mathrm{J} \odot (A(\mathrm{J} \odot (A(\mathrm{J} \odot A^2)))) - \mathrm{J} \odot (A(A \odot (A\mathrm{J}))) \\
&\quad - \mathrm{J} \odot ((A \odot (A\mathrm{J}))A) - A \odot ((A \odot A^2)\mathrm{J}) + 2A \odot A^2 \\
P_5 &= \mathrm{J} \odot A^5 - (\mathrm{I} \odot A^4)A - A(\mathrm{I} \odot A^4) - (\mathrm{I} \odot A^3)(\mathrm{J} \odot A^2) - (\mathrm{J} \odot A^2)(\mathrm{I} \odot A^3) \\
&\quad - (\mathrm{I} \odot A^2)(\mathrm{J} \odot A^3) - (\mathrm{J} \odot A^3)(\mathrm{I} \odot A^2) - \mathrm{J} \odot (A(\mathrm{I} \odot A^3)A) + 3A \odot A^3 \\
&\quad + 2A(\mathrm{I} \odot A^2)(\mathrm{I} \odot A^2) + 2(\mathrm{I} \odot A^2)(\mathrm{I} \odot A^2)A + 3A \odot A^2 \odot A^2 + (\mathrm{I} \odot A^2)A(\mathrm{I} \odot A^2) \\
&\quad - \mathrm{J} \odot (A(\mathrm{I} \odot A^2)A^2) - \mathrm{J} \odot (A^2(\mathrm{I} \odot A^2)A) + 3\mathrm{J} \odot ((A \odot A^2)A) + 3\mathrm{J} \odot (A(A \odot A^2)) \\
&\quad + (\mathrm{I} \odot (A(\mathrm{I} \odot A^2)A))A + A(\mathrm{I} \odot (A(\mathrm{I} \odot A^2)A)) - 6(\mathrm{I} \odot A^2)A - 6A(\mathrm{I} \odot A^2) \\
&\quad - 4A \odot A^2 + 3\mathrm{J} \odot A^3 + 4A
\end{aligned}
$$

Formulae

GRL applied on $\tilde{G}_3$:

$$
\begin{aligned}
P_2^* &= \mathrm{J} \odot A^2 \\
P_3^* &= \mathrm{J} \odot (A(\mathrm{J} \odot A^2)) - A \odot (A\mathrm{J}) \\
P_4^* &= \mathrm{J} \odot (A(\mathrm{J} \odot (A(\mathrm{J} \odot A^2)))) - \mathrm{J} \odot (A(A \odot (A\mathrm{J}))) \\
&\quad - \mathrm{J} \odot ((A \odot (A\mathrm{J}))A) - A \odot ((A \odot A^2)\mathrm{J}) + 2A \odot A^2 \\
P_5^* &= \mathrm{J} \odot (A(\mathrm{J} \odot (A(\mathrm{J} \odot (A(\mathrm{J} \odot A^2)))))) - \mathrm{J} \odot (A(\mathrm{J} \odot (A(A \odot (A\mathrm{J}))))) \\
&\quad - \mathrm{J} \odot (A(\mathrm{J} \odot ((A \odot (A\mathrm{J}))A))) - \mathrm{J} \odot (A(A \odot ((A \odot A^2)\mathrm{J}))) + 2\mathrm{J} \odot (A(A \odot A^2)) \\
&\quad - (\mathrm{J} \odot A^2) \odot ((A \odot A^2)\mathrm{J}) - A \odot ((A \odot (A(\mathrm{J} \odot A^2)))\mathrm{J}) + A \odot ((A \odot (A\mathrm{J}))\mathrm{J}) \\
&\quad - (A\mathrm{J}) \odot (\mathrm{J} \odot (A(\mathrm{J} \odot A^2))) + (A\mathrm{J}) \odot A \odot (A\mathrm{J}) + \mathrm{J} \odot (A(\mathrm{J} \odot A^2)) \\
&\quad + A \odot (A(\mathrm{J} \odot A^2)) - 2A \odot (A\mathrm{J}) + 2A \odot A^2 \odot A^2 + 3\mathrm{J} \odot ((A \odot A^2)A) - 4A \odot A^2
\end{aligned}
$$

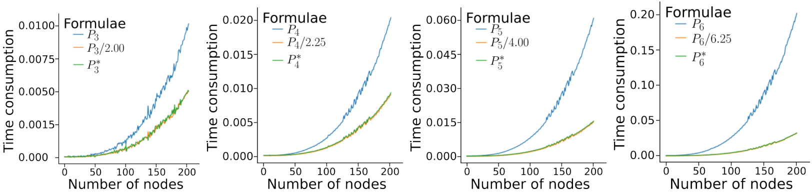
Formulae

$$
\begin{aligned}
P_6 = {} & \mathrm{J} \odot A^6 - (\mathrm{I} \odot A^5)A - A(\mathrm{I} \odot A^5) - (\mathrm{I} \odot A^2)(\mathrm{J} \odot A^4) - (\mathrm{J} \odot A^4)(\mathrm{I} \odot A^2) \\
& - (\mathrm{I} \odot A^4)(\mathrm{J} \odot A^2) - (\mathrm{J} \odot A^2)(\mathrm{I} \odot A^4) - \mathrm{J} \odot (A(\mathrm{I} \odot A^4)A) + 3A \odot A^4 \\
& - (\mathrm{J} \odot A^3)(\mathrm{I} \odot A^3) - (\mathrm{I} \odot A^3)(\mathrm{J} \odot A^3) - \mathrm{J} \odot (A(\mathrm{I} \odot A^2)A^3) - \mathrm{J} \odot (A^3(\mathrm{I} \odot A^2)A) \\
& - \mathrm{J} \odot (A(\mathrm{I} \odot A^3)A^2) - \mathrm{J} \odot (A^2(\mathrm{I} \odot A^3)A) + 4A(\mathrm{I} \odot A^2)(\mathrm{I} \odot A^3) + 4(\mathrm{I} \odot A^3)(\mathrm{I} \odot A^2)A \\
& + 6A \odot A^2 \odot A^3 + (\mathrm{I} \odot A^2)A(\mathrm{I} \odot A^3) + (\mathrm{I} \odot A^3)A(\mathrm{I} \odot A^2) + 3\mathrm{J} \odot ((A \odot A^3)A) \\
& + 3\mathrm{J} \odot (A(A \odot A^3)) + (\mathrm{I} \odot (A(\mathrm{I} \odot A^3)A))A + A(\mathrm{I} \odot (A(\mathrm{I} \odot A^3)A)) - \mathrm{J} \odot (A^2(\mathrm{I} \odot A^2)A \\
& + 2A(\mathrm{I} \odot A^2)(\mathrm{I} \odot A^2) + 2(\mathrm{I} \odot A^2)(\mathrm{I} \odot A^2)A + \mathrm{J} \odot A^2 \odot A^2 \odot A^2 + (\mathrm{I} \odot A^2)(\mathrm{J} \odot A^2)(\mathrm{I} \odot \\
& + 3\mathrm{J} \odot ((A \odot A^2)A^2) + 3\mathrm{J} \odot (A^2(A \odot A^2)) + (\mathrm{I} \odot (A(\mathrm{I} \odot A^2)A))(\mathrm{J} \odot A^2) + (\mathrm{J} \odot A^2)(\mathrm{I} \\
& + \mathrm{J} \odot ((\mathrm{I} \odot A^2)A(\mathrm{I} \odot A^2)A) + \mathrm{J} \odot (A(\mathrm{I} \odot A^2)A(\mathrm{I} \odot A^2)) + 2\mathrm{J} \odot (A(\mathrm{I} \odot A^2)(\mathrm{I} \odot A^2)A) \\
& + (\mathrm{I} \odot (A(\mathrm{I} \odot A^2)A^2))A + A(\mathrm{I} \odot (A(\mathrm{I} \odot A^2)A^2)) + (\mathrm{I} \odot (A^2(\mathrm{I} \odot A^2)A))A + A(\mathrm{I} \odot (A^2(\\
& + 3\mathrm{J} \odot ((A \odot A^2 \odot A^2)A) + 3\mathrm{J} \odot (A(A \odot A^2 \odot A^2)) - 12(\mathrm{I} \odot A^2)(A \odot A^2) - 12(A \odot A^2 \\
& - 4\mathrm{J} \odot A^2 \odot A^2 - 8A \odot (A(A \odot A^2)) - 8A \odot ((A \odot A^2)A) - 3A \odot (A(\mathrm{I} \odot A^2)A) \\
& + 3\mathrm{J} \odot (A(A \odot A^2)A) + \mathrm{J} \odot (A(\mathrm{I} \odot (A(\mathrm{I} \odot A^2)A))A) - 4\mathrm{J} \odot (A(A \odot A^2)) - 4\mathrm{J} \odot ((A \odot \\
& + 4\mathrm{J} \odot A^4 - 5A(\mathrm{I} \odot A^3) - 5(\mathrm{I} \odot A^3)A - 4(\mathrm{I} \odot (A(A \odot A^2)))A - 4A(\mathrm{I} \odot (A(A \odot A^2))) \\
& - 4(\mathrm{I} \odot ((A \odot A^2)A))A - 4A(\mathrm{I} \odot ((A \odot A^2)A)) - 7(\mathrm{I}A^2)(\mathrm{J} \odot A^2) - 7(\mathrm{J} \odot A^2)(\mathrm{I}A^2) \\
& - 10\mathrm{J} \odot (A(\mathrm{I} \odot A^2)A) + 44A \odot A^2 + 12\mathrm{J} \odot A^2
\end{aligned}
$$

Formulae

$$\begin{aligned}
P_6^* = {} & \mathrm{J} \odot (AP_5^*) - P_3^* \odot ((A \odot A^2)\mathrm{J}) - A \odot (C_{5f}\mathrm{J}) - P_4^* \odot (A\mathrm{J}) \\
& + P_4^* + C_{5f} - (\mathrm{J} \odot A^2) \odot (C_{4f}\mathrm{J}) + 4A \odot A^2 \odot P_3^* + 3\mathrm{J} \odot ((A \odot A^2)(\mathrm{J} \odot A^2)) \\
& + \mathrm{J} \odot A^2 \odot A^2 \odot A^2 + 3\mathrm{J} \odot ((A \odot A^2 \odot A^2)A) - 4\mathrm{J} \odot A^2 \odot A^2 - 8A \odot (A(A \odot A^2)) \\
& - 8A \odot ((A \odot A^2)A) - 4\mathrm{J} \odot ((A \odot A^2)A) - 3A \odot ((A \odot A^2)\mathrm{J}) + 17A \odot A^2 + 3\mathrm{J} \odot A^2
\end{aligned}$$

Gain of time complexity



Perspectives

Explore k-WL CFGs.

Try to find explicit formulas for real life datasets with RL.

Grammar in other types of data.

Questions ?

Thanks for your attention.

Bibliographie

Voropaev, A. and Perepechko, S. (2012).
The number of fixed length cycles in undirected graph explicit formula in case of small lengths.
Discrete and Continuous Models and Applied Computational Science, (2):6–12.